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

DOCTORAL (PH.D.) DISSERTATION

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

JURY MEMBERS

CHAIRMAN:

DR. GYÖRGY DÉNES BISZTRAY (PH.D.)

SCIENTIFIC COMMITTEE:

DR. JÚLIA HALÁSZ (PH.D.)
DR. LÁSZLÓ NAGY (PH.D.)
DR. ATTILA BOROVICS (PH.D.)

OPPONENTS:

DR. SÁNDOR BORDÁCS (C.SC.)
DR. TAMÁS DEÁK (PH.D.)

SECRETARY:

DR. GABRIELLA SOMOGYI (PH.D)

CONTENT

1. ABBREVIATIONS

AMOVA	Analysis of Molecular Variance
ANOVA	One-way Analysis of Variance
ANS	Anthocyanin Synthase
APX	Ascorbate Peroxidase
ATP	Adenosine Triphosphate
BLAST	Basic Local Alignment Search Tool
BLASTN	Program for search nucleotide databases using a nucleotide query
BLASTX	Program for search protein databases using a translated nucleotide query
bp	Base pairs
BP	Before Present
cal BP	calibrated years before the present
cDNA	Complementary DNA
CEGMA	Estimation of the completeness and contiguity of the assembly
CHI	Chalcone Isomerase
CHS	Chalcone Synthase
CoA	Coenzyme with role in the synthesis and oxidation of fatty acids
Contig N50	Ordering every contig by length from longest to shortest
Contig	Set of overlapping DNA segments, a consensus region of DNA
cpDNA	Chloroplast DNA
cpSSR	Chloroplast Simple Sequence Repeats
DFR	Dihydroflavonol 4-Reductase
DNA	Deoxyribonucleic Acid
dNTP	DeoxyriboNucleotide TriPhosphate
EST	Expressed Sequence Tags
EUFORGEN	European Forest Genetic Resources Programme
F3’5’H	Flavonoid 3’5’ Hydroxylase
F3’H	Flavonid 3’ Hydroxylase
F3H	Flavanone Hydroxylase
GC content	Guanine-cytosine content
GST	Glutathione S-Transferase
LD	Linkage Disequilibrium
LDR	Leucoanthocyanidin Reductase
LGM	Last Glacial Maximum (LGM: 21 ± 2 ka cal BP)
LINE	Long Interspersed Nuclear Element
LPG	Late Pleniglacial (LPG: 26.5–15 ka cal BP)
MADS	MCM1, AGAMOUS, DEFICIENS, SRF
MDH	Malate Dehydrogenase
mtDNA	Mitochondrial DNA
mya	Million years ago
MYB	Myeloblastosis family of transcription factors
NADPH	NAD Malate Dehydrogenase
NCBI	National Center for Biotechnology Information
ncRNA	Noncoding transcript
NGS tech	Next-Generation Sequencing technology
nSSR	Nuclear Simple Sequence Repeats
PAL	Phenylalanine Ammonia-Lyase
PCR	Polymerase chain reaction
QTL	Quantitative Trait Locus
RNA	Ribonucleic Acid

SINE	Short Interspersed Nuclear Element
SNP	Single-nucleotide polymorphism
SSR	Simple Sequence Repeats
TE	Transposable element
TF	Transcription factor
TPS	Terpene Synthase
UTR	Untranslated Region
UV	Ultraviolet light

2. INTRODUCTION AND OBJECTIVES

Scots pine *(Pinus sylvestris* L.) on its extremely large Eurasian distribution range has adapted to a wide variety of climates. As a light-demanding pioneer species it is able to colonize recently disturbed sites (Durrant *et al.* 2016), and can grow in areas with annual precipitation exceeding 1,780 mm but also where annual precipitation is less than 200 mm (Steinbeck 1966). Moreover populations are able to sustain where the winter temperatures have been recorded as low as -64° C, but can also tolerate high temperatures growing at the middle altitudes of the Mediterranean region.

In Europe, the species grows on a wide variety of soil types: peatbogs, rocky substrates and also young glacial deposits, siliceous and acidic soils, frequently with deep litter and raw humus layer (Steinbeck 1966). Its northern range limit is mostly determined by climatic conditions, where high temperature amplitudes are considered to be the most important restrictive factors (Boratyński 1991). The 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). 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 interpopulation differentiation 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 the 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.

The resolution and usability of these markers has been improved considerably with the advancement in the techniques of nucleic acid hybridization, polymerase chain reaction and DNA sequencing. Furthermore, the rapid development of next-generation sequencing technology provided an opportunity to develop novel genomic tools. Appreciated for their advantages such as accuracy, high throughput and relatively low cost, these technologies are widely applied also to the qualitative and quantitative analysis of transcriptomes (Bräutigam and Gowik 2010). Massively parallel sequencing of cDNA now is an efficient route for generating enormous sequence collections that represent expressed genes. The lower cost and greater sequence yield has allowed the identification of candidate genes, functional genomic level data with genome characteristics. These data have high functional information content, often correspond to genes with known or predicted functions, have proven to be unavoidable for comparative genomics (Vera *et al.* 2008) and for 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.

To describe signs of adaptive variation in selected peripheral populations of *Pinus sylvestris* L. in the Pannonian Basin and the Carpathian Mountains overall objectives of this research were as follows:

- Detect the level of phenotypic differentiation based on cone morphology and needle anatomy in response to the geographical distribution.
- Based on morphoanatomical differentiation, 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.

3. LITERATURE REVIEW

3.1. Taxonomic classification and morphology of Scots pine (*Pinus sylvestris* L.)

3.1.1. Taxonomy

Pinus (*Pinaceae*), with over 100 species, is the largest extant genus of conifers (Price *et al.* 1998) and a dominant component of the boreal, subalpine, temperate, and tropical forests, as well as arid woodlands (Richardson and Rundel 1998). Price *et al.* (1998) proposed a classification for the genus, recognizing 111 species in two subgenera, four sections, and 17 subsections. The species circumscriptions showed over 90 % correspondence to a separate compilation by Farjon, who recognised 109 species (Gernandt *et al.* 2005). Among all the pine species, *Pinus sylvestris* L. has the largest geographic distribution, which ranges from northern Scandinavia to southern Spain and from western Scotland to the Okhotsk Sea in eastern Siberia. The species' chromosome number is: 2n = 24 (Mirov and Stanley 1959)

3.1.2. Morphology

Scots pine is a medium-sized conifer, that reaches 23 - 27 m in height on average but can attain over 40 m and lives for 400 years or more (Praciak 2013). The bark on the upper part of the stem develops a distinct reddish-orange colour while the lower part is brown to grey-brown and becomes deeply fissured. Its blue-green or grey-green needles are in pairs, generally slightly twisted and are around 5 - 7 cm long. They stay on the tree for at least two, and in some cases up to six years (Durrant *et al.* 2016).

The needles are adapted to deal with cold and drought, having embedded stomata and a waxy layer on the thick walled epidermis to protect the needles from water loss (Skilling 1990). It is a wind-pollinated species and is normally monoecious but mature trees may vary, occasionally bearing only male or only female flowers (Durrant *et al.* 2016). The male flowers are situated at the base of new shoots, coloured yellow or pink; the female flowers occur at the tips of the new shoots and develop a purple colour.

The cones develop the year following pollination and are conic-oblong, 5 - 8 cm in size. The seeds are winged and about 3 - 4 mm in size, the wing is about three times longer than the size of the seed (Gencsi 1992) (**Figure 1**). They require alternating periods of dry and wet weather to open and shed the winged seeds, which are dispersed from the parent tree (Durrant *et al.* 2016).

Figure 1. *Pinus sylvestris* L. (Scots pine)
http://www.conifers.org/pi/Pinus_sylvestris.php

3.1.3. Leaf anatomy

In mature plants of Scots pine the leaves of the brachyblast are the only assimilating leaves. Typical short shoots or bracyblasts have two needle leaves, which are inserted on a strongly reduced short shoot axis (Dörken *et al.* 2010). In the most basal parts the two needle leaves are connate. In cross sections (**Figure 2**) leaves on the short shoot have a semicircular shape. A prominent cuticula covers the epidermis. A hypodermis with strongly thickened cell walls is developed. The mesophyll is homogenous. Five resin ducts are distributed more or less symmetrically and equidistant near the entire abaxial surface of the needle. In the middle part of the leaf the vascular bundle gets strongly subdivided by a parenchymatic band, which is several cell rows broad. Its cells become later lignified and can be distinguished from the tracheids only by a larger size. In this region two distinct vascular strands can be seen, which again are surrounded by a common bundle sheath with Casparian strips. In the most basal and the most distant parts this parenchyma is absent. The vascular bundle is not subdivided in these regions (Dörken and Stützel 2012).

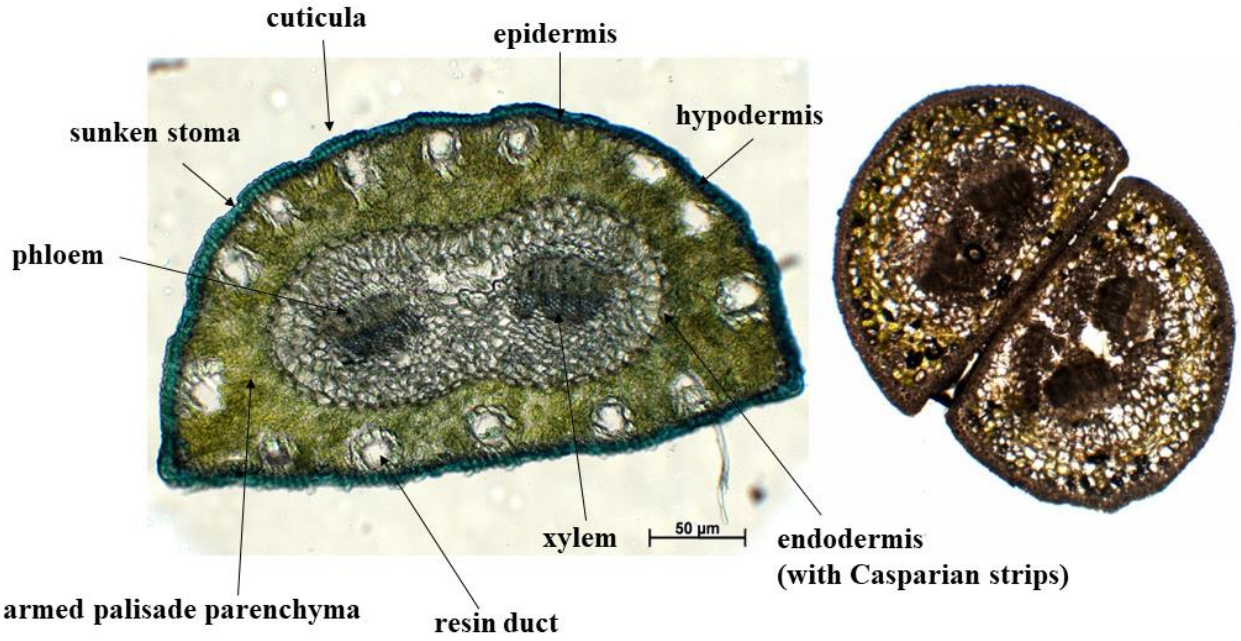

Figure 2. *Pinus sylvestris* needle cross-section (photo credit: M. Höhn)

Seedlings of *Pinus sylvestris* have several cotyledons and the same number of primary leaves. The cotyledons are about 2 - 3 cm long and have a pointed leaftip. In crosssections they have a triangular shape. A cuticula covers the epidermis. The hypodermis is lacking. The mesophyll is monomorphic. There are two vascular bundles which are surrounded by a bundle sheath. Xylem is placed to the adaxial and phloem to the abaxial side. Xylem and phloem are not subdivided by parenchyma. The primary leaves are about 2 cm long. In crosssections they are oval. In all other features they are similar to the cotyledons (Dörken and Stützel 2012).

3.1.4. Seed morphology, production and dissemination

Under favorable growth conditions Scots pine starts to produce male and female flowers at from five to eight years, although the average is between 10 and 15 years (Wright *et al.* 1966). Viable seeds are produced up to at age 200, with seed quality and size greatly reduced at this age. Good seed crops are produced at intervals at three to six years with lightcrops in most intervening years. If properly stored, the seeds remain viable for 15 years. One kilogram of average size cones produces approximately 3,300 seeds (Heit 1969). Cones begin to open in late October, and seed dispersal continues into December. At times, large quantities of seeds are dispersed on to snow cover. Seed dispersal for natural areas is normally limited to 50 and 100 m from the parent tree (Heit 1969).

3.2. Natural distribution

Scots pine can be characterised by the most extensive range of all species of the *Pinaceae* family (Giertych and Mátyás 2017). It can be found all the way across Eurasia. The huge pine forests of Siberia are the largest stands in the world. Even the name "sylvestris" comes from the Latin "of forests" (Durrant *et al.* 2016). The range of the species covers longitudinally over a distance of about 14,000 km, with the westernmost occurences in Spain and the easternmost in the Russian Far East (Bennett 1995). In north-south direction, species' range extends to northern Scandinavia and in the southern part extends to the Sierra Nevada mountains, in Spain (Labra *et al.* 2006, Pyhäjärvi *et al.* 2008, Scalfi *et al.* 2009). Regarding to the altitudinal distribution, the highest altitudes are reached at the southern limits, in the Caucasus, where the species forms stands up to 2,600 m above sea level, with single individuals even at 2,700 m. In Europe, the altitudinal maximum is reached in the Alps at 2,400 m.

In its southern limits, Scots pine has the lower altitudinal limit of occurrence in the Sierra Nevada in Spain, at 1,600 m (Soranzo *et al.* 2000). In the last decades an increased reproduction and growth has been detected at the northern limit of the species as a response to increased temperature, whereas at its southern limit decreased growth and massive mortality can be observed, caused by increased drought stress (Cipriano *et al.* 2013). The gene pools of the species have also been affected by isolation, by mutations and selection due to ecological conditions as well as by genetic drift. Recently growing populations to a considerable extent are also affected by human activity connected with the management of forest resources and silviculture. Pollution linked to the development of industry has also influenced the populations through the negative selection of susceptible genotypes (Bergmann and Hosius 1996, Oleksyn *et al.* 1994, Prus-Głowacki *et al.* 2012).

In the future, an increasing level of ozone in the troposphere and predicted global warming are expected to have substantial effects on processes of selection and genetic variation of Scots pine populations (Fowler *et al.* 1999, Staszak *et al.* 2004) The actual range of *Pinus sylvestris* is characterised by many refugees and colonization routes throughout Europe, which strongly influenced its genetic diversity and local adaptation capability (Matías and Jump 2012).

Habitat characteristics of Scots pine

Scots pine is adapted to a wide variety of climates due to its extremely large natural range. As a consequence of its vast distribution, it has adapted to a large variety of soils and climates – from

the arid mountains of Spain and Asia Minor to the subarctic forests of Northern Scandinavia and Siberia (**Figure 3**).

Figure 3. *Pinus sylvestris* populations from different habitats
Left: raised bog (Mohoș-RO) (photo credit: M. Höhn); Middle: mixed forest on sandy soil (Fenyőfő-HU) (photo credit: L. Udvardy); Right: rocky substrate (Kvacianska dolina-SK) (photo credit: M. Höhn).

Temperature and light are considered to be the most important environmental factors in its survival (Notivol *et al.* 2007). It is a light-demanding pioneer species and can colonize recently disturbed sites if competition and grazing pressure are low (Durrant *et al.* 2016). It grows in areas with an annual precipitation exceeding 1,780 mm and in areas with an annual precipitation as little as 200 mm (Steinbeck 1966). Scots pine survives in eastern Siberia where winter temperatures have been recorded as low as -64° C. In some areas it grows where the subsoil is permanently frozen. It can also survive high temperatures, and it is found also at the middle altitudes in the Mediterranean region. Its primary distribution, however, indicates that it is a tree of the continental climates (Burns and Honkala 1990).

In Europe, Scots pine grows on a wide variety of soil types: bogs, rocky substrates and also on the most recent glacial deposits, siliceous and acidic soils, frequently with deep litter and raw humus layer (Steinbeck 1966). The soils exhibit various degrees of podzolization. Studies of the mineral nutrient content of the foliage of several Scots pine provenances showed that the species has evolved an efficient mechanism to extract nutrients from the infertile sites to which it is relegated in its native range. Significant differences were found among seed sources in their ability to accumulate nitrogen, phosphorus, sodium, magnesium, and boron. Magnesium was one of the key minerals in Scots pine nutrition. The faster-growing seed sources accumulates higher levels of foliar magnesium (Heit 1969). Although Scots pine can grow on soils with pH from 4.0 to 7.0, it grows best on soils in between 4.5 to 6.0 (Dickson and Winch 1970). Although the seedlings grow very well on fertile soils, they can be found also on sandy dry soils because of the lack of competition of other trees and herbs. The best regeneration is found in stands with the following characteristics: large seed supply, open or light tree canopy, light understory ground cover, and

exposed mineral soil or no continuous layer of raw humus (Steinbeck 1966). It cannot cope with atmospheric pollution or salty sea wind. It requires a period of winter chilling to break autumn dormancy, and starts to grow in the spring when temperature reaches about 5° C (Steinbeck 1966).

Scots pine frequently grows in large single-species stands, but across its huge range forms mixed stands with most of the boreal species of Europe and Asia. In Europe it can be found growing with broad leaved trees such as oaks (*Quercus petraea* L., *Quercus robur* L.), beech (*Fagus sylvatica* L.), birch (*Betula pendula* L.), and other conifers including spruce (*Picea abies* L.), larch (*Larix decidua* L.), fir (*Abies alba* L.) and other pines (*Pinus nigra* L., *Pinus uncinata* L.), but no single species or species group is associated with it over its entire range (Burns and Honkala 1990).

3.3. The species' phylogeography in Europe

Population genetics and phylogenetics of gymnosperms with use of the initial data source of chloroplast DNA evolved enormously in the last 20 years, revealing not only the genetic heterogeneity throughout the range of the species but also historical rangeshifts and recolonization routes (Taberlet *et al.* 1998). Networks of fossilized pollen and plant micro- and macrofossils show that in response to the Quaternary climatic fluctuations, the biosphere has experienced dramatic changes, with large-scale species rangeshifts, population contractions, expansions and extinctions, as well as aggregation and disassociation of forest communities (Petit *et al.* 2008).

According to fossil data of closely related species, the ancestral gene pool of *Pinus* was located along the middle latitudes of North America and Western Europe (Millar 1993, Mirov 1967). Scots pine was present in the glacial (full glaciation) and interglacial periods with documented population fluctuations (Bennett 1995, Cohen *et al.* 2013, Kelly and Connolly 2000; Richardson and Rundel 1998). Macrofossil evidences show the presence of *P. sylvestris* in the last full-glacial period (Richardson and Rundel 1998) in central and southern Moravia and the territory of the Pannonian Basin (Jankovska and Pokorný 2008). Macro-charcoal remains dating between ca. 70,000 BP and 15,000 years BP, mostly consists *P. sylvestris* from the south of the Carpathian basin (Rudner and Sümegi 2001). Another paleobotanical survey in the Hungarian plain showed pollen fossils dating back to 30,000–20,000 BP, suggesting, that before the last glacial maximum, the area was covered by *P. sylvestris*, on lower and higher flood plain zones (Magyari 2011). Pollen and macrofossil evidences from the Northern Carpathians (Slovakia and Czech Republic) show, that Weichselian full-glacial montane forests were dominated by *Larix spp.*, *P. cembra*, *P. sylvestris*, between 50,000 BP and 16,000 BP (Jankovska and Pokorný 2008). Late Pleniglacial

(LPG: 26,500–15,000 BP) pollen records showed forest patches dominated by *P. sylvestris, P. mugo, P. cembra* (Magyari *et al.* 2014a).

By the Last Glacial Maximum (LGM: 21,000 ± 2,000 BP), the species survived only in patchy small and discontinuous glacial refugia mostly on ice-free areas (Matías and Jump 2012, Willis *et al.* 1998). The largest refugial territories were localized in southern Europe, Italy, the Iberian Peninsula and in the Balkans (Cheddadi *et al.* 2006, Labra *et al.* 2006). Morphological and molecular genetic studies confirm that southern refugia were located in Asia Minor (Jasińska *et al.* 2014, Pyhäjärvi *et al.* 2008, Savolainen *et al.* 2007) the northern refugia being found in Scandinavia (Parducci *et al.* 2012) and near Moscow (Buchovska *et al.* 2013). Between 16,000 and 12,000 BP, started an expansion of refugial populations towards the European continent, in response to climate warming (Pyhäjärvi *et al.* 2008, Richardson and Rundel 1998). In the eastern Carpathians the expansion started at 1,630 BP (Magyari *et al.* 2014b) mainly on low-elevation sites (Pèrez-Obiol and Julià 1994).

Following the late-glacial period, at the beginning of the Holocene, in the Alps and in Central Europe a decline of pine density was inferred (Huttunen *et al.* 1992, Lowe 1992) in the Hungarian basin and the Iberian Peninsula. Pines sustained their populations with outdisturbance and no change has been detected (Pèrez-Obiol and Julià 1994, Willis *et al.* 1998). In this period an elevational change was documented in the treeline elevation in the Southern Carpathians (Magyari *et al.* 2011). After the gradual warmup of the exposed tundra pines were forced to migrate northward, towards the European lowlands (Huntley and Birks 1983).

In the work of Tóth *et al.* (2017), 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. These populations preserved high genetic diversity, with historical divergence from the ancestral population in mid-Pleistocene (178,000 BP – 213,600 BP). Later, the diverged populations have expanded, which led to an admixture event (between 11,800 BP – 14,100 BP). This expansion was followed by a recent fragmentation in the early Holocene period. At the same time in the Mediterranean region of Southern Europe, the previously steppe-dominated vegetation began to experience invasion by deciduous woodlands, causing the decline of *Pinus* species (Adams and Faure 1997). During their northward advance, pines reached much of Scandinavia by 10,000 BP and started to increase their presence in Britain. They became dominant in Finland by 9,000 BP (Pyhäjärvi *et al.* 2008). In ca. 9,900 BP, Scots pine became locally dominant in Scotland (Richardson and Rundel 1998).

The species reached its maximum extent of distribution across Europe around 8,000 years ago (Kullman and Kjällgren 2006, Matías and Jump 2012). By 6,800 years ago, a range retraction was inferred for the southern parts of Fennoscandia, the distribution extending eastward to the

Siberian steppes (Kremenetski *et al.* 1997). Beginning from between 4,800 BP and 4,200 BP the species declined in the northern and western parts of the British Isles (Matías and Jump 2012). The decline of populations continued at about 4,000 BP (Birks and Birks 1980, Birks and Williams 1983, Watts 1985). In the Iberian Peninsula and Italy between 5,700 BP and 3,200 BP, populations retreated from their maximum distribution to the present distribution area *(*Matías and Jump 2012, Willis *et al.* 1998). Between 5,000 BP and 3,000 BP, *Pinus sylvestris* gradually withdrew from northern Scandinavia (Eronen 1979, Eronen and Hyvärinen 1982). The species survived on marginal territories, in extreme conditions on poor soils, far from the refugia of deciduous species (Bennett 1984, Matías and Jump 2012, Richardson and Rundel 1998, Willis *et al.* 1998).

From 4,000 BP to the present, natural withdrawal of Scots pine continues as a consequence of current climate warming and land-use change, the conversion of forests to croplands, livestock grazing, soil depletion by pasturing and timber production having the effect of shrinkage and disappearance of pine populations (Tipping *et al.* 2008).

In recent history, populations of Scots pine are also affected by secondary recolonization most likely caused by forest management and agricultural practices with the result of spread on abandoned agricultural landscapes. *Pinus sylvestris* has become an established woodland tree of the natural and human-disturbed areas in many geographical sites (Richardson and Rundel 1998).

3.4. Peripheral populations, their role and importance in species' conservation and during the ongoing climate change

As we dicussed also in a previous chapter regarding to the species' distribution, Scots pine is characterised by a very extensive range. In mountanious regions, the species does not reach as far south as some other members of the *Pinaceae* family, but the southernmost occurences are situated on higher altitudes. These southern populations represents the highly fragmented range of the species' distribution. They are isolated, and can be found in Europe in the Grampian Mountains (Scotland) (Pears 1968), the Cordillera Iberica (Prus-Głowacki *et al.* 2012), in the Pyrenees (Galiano *et al.* 2010), in the Northern Appenines (Labra *et al.* 2006) , the Dinarian Alps (Boratyński 1991), Rodopi (Argyropoulou *et al.* 2005), Stara Planina (Stankova and Shibuya 2007), Pindus Mountains (Papanastasis *et al.* 2009), the Northern, Southern and Eastern Carpathians, Central Island Mts. (Köbölkuti *et al.* 2017), Balkan Peninsula (Petit 2003), Crimea (Marcysiak 2006) and the Caucasus (Belletti *et al.* 2012) and also in the mountains of Northern Anatolia (Durkaya *et al.* 2009). In Europe the altitudinal maximum is reached in the Alps (Boratyński 1991), in other mountains the upper limit being much lower. The altitudinal

occurrence is determined by environmental and climatic conditions, which deteriorate as latitude increases.

Peripheral occurrences cannot become established beyond the up mentioned ranges because they would have negative growth rates in the new, restrictive, unfavourable habitats (Bridle and Vines 2007) (**Figure 4**). There are two contrasting explanations for the failure of their continuous expansion: if the range edge is highly fragmented, genetic drift and the low rate of mutations into marginal populations might limit the availability of locally beneficial alleles, preventing adaptation and, therefore, the continous range expansion (Lennon *et al.* 1997). The other explanation is, that if populations at the margins remain connected to large central populations, the continual flow of deleterious alleles could swamp the establishment of locally adaptive alleles, maintaining negative population growth, and preventing expansion (Garcia-Ramos and Kirkpatrick 1997, Kirkpatrick and Barton 1997). Much discussion of why expansion fails at range margins hinges on determining how much gene flow is necessary to maintain the adaptive potential of these margins without swamping local adaptation (Alleaume-Benharira *et al.* 2006). According to several authors, populations situated near the core of a species' geographic distribution exhibit greater abundance than populations near the periphery (Curnutt *et al.* 1996). When a species colonizes a geographical gradient of environmental conditions, it should become most abundant where the individual survival and reproduction is highest, and less abundant as conditions depart from its optimum (Hengeveld and Haeck 1982). The differences between the core and periphery are further exacerbated if peripheral populations experience rapid cycles of extinction, recolonization, or population bottlenecks.

As we discussed in the previous chapter, in Central and Eastern Europe the gene pools of *Pinus sylvestris* were formed in postglacial times during the migration of the species from different glacial refugias, having intensive gene exchange between populations. The abundance and variability of core populations is thought to be driven by spatial patterns in habitat quality (Curnutt *et al.* 1996). May arise because core populations, inhabiting productive habitats, are the sources of the species', and peripheral populations inhabiting lower quality habitat, are sinks (Pulliam 2017). However, not all populations are expected to exhibit these patterns. Available surveys report, that the peripheral populations commonly exist at lower densities but exhibit higher variability (Lesica and Allendorf 1995, Scudder 1989). According to Vucetich and Waite (2003), compared with core populations, edge populations may exhibit levels of genetic diversity that fall from much lower to much higher, explaining the findings of peripheral populations with less genetic diversity (Schnabel and Hamrick 1990) or, in some cases, with higher genetic diversity on periphery (Betancourt *et al.* 1991, Lesica and Allendorf 1995).

Accordingly, the importance and role of the peripheral Scots pine populations in biodiversity conservation is still very controversial (Eckert *et al.* 2008, Hampe and Petit 2005). Nevertheless, according to several authors, the dynamics of latitudinal margins inhabitant populations are critically important in the response of the species to expected climate change (Iverson *et al.* 2004, Thomas *et al.* 2001, Travis and Dytham 2004). Populations residing at the current low-latitude margins of species distribution ranges are extremely important for the long-term conservation of species and their investigation and conservation deserve high priority (Hampe and Petit 2005). Peripheral populations maintain substantial interpopulation differentiation, by the fact that they may diverge because of adaptation from central populations owing to different selective pressures and reduced gene flow (Lenormand 2002). With proper demographic and genetic properties, populations from the edge may also facilitate shifts in the geographical distributions of the species, in response to rapid climate change (Etterson 2001, Parmesan 2006). Such populations should occur mostly in regions, that have provided suitable conditions for species persistence under both cold stage and warm stage conditions (Tzedakis *et al.* 2004). Their successful long-term persistence in spite of their small sizes at least during interglacial periods indicates, that extinction because of demographic stochasticity has played a relatively minor role.

Recent phylogeographic studies have shown that marginal populations constitute a reservoir of within-species genetic diversity (Hewitt 2004, Petit 2003). Population dynamics varied with life history and geography the present genetic constitution of the populations of some species carrying attenuated signals of past dynamics. In regions with a recent history of glaciation or some other manifestation of climate change, species are liable to have experienced large-scale fluctuations in the effective population size and the rate of gene flow that may not be reflected among contemporary populations (Pamilo and Savolainen 1999). Geographical ranges have shifted in the past 20,000 years, with lead to significant genetic consequences for nowaday populations. Partition of genetic diversity that compares central and peripheral populations of gymnosperms was reviewed by Channell (2004), revealing, that peripheral populations in case of several species experience demographic processes, either historical, contemporary or both, that can lead to lower genetic diversity, but higher genetic differentiation on the periphery (in case of *Pinus contorta*, *Pinus jeffreyi* Balf., *Pseudotsuga menziesii* (Mirb.) Franco) while other species like *Picea abies* (L.) H. Karst. , *Pinus edulis* Engelm. *Pinus rigida* Mill. and *Pinus sylvestris* showed that the manifestation of this hypothesis on the partition of genetic diversity is contrary.

It is generally accepted that the genetic structure of Scots pine populations in Europe is determined by the post-glacial migration of the species from its southwestern, southern and southeastern refugias (Langlet 1959). For that reason, populations from Central and Northern Europe are supposed to be ancient, with importance in the response of the species to expected

climate change. Scots pine peripheral populations, in addition to possessing imprints of historical events, are also distinctive due to adaptations to different habitat extremes. The hypothesis, that natural populations in the refugial peripheral areas from Central and Northern Europe, the Balkans, Iberia, and Anatolia possess imprints of historical events and are also distinctive due to adaptations to different habitat extremes is supported by studies performed on these peripheral populations (Alía *et al.* 2001, Bilgen and Kaya 2007, Dzialuk *et al.* 2009, Jasińska *et al.* 2014, Labra *et al.* 2006, Mejnartowicz 1979, Prus-Głowacki *et al.* 2003, Pyhäjärvi *et al.* 2007, Semiz *et al.* 2007, Staszkiewicz 1961, Tobolski and Hanover 1971, Turna 2003). If the genetic diversity of these peripheral populations are not inherently inferior to those of central ones, then central and peripheral populations both have important roles in conservation (Channell 2004). This also may represent valuable opportunities to conserve biodiversity that is threatened by global climate change, the loss of them will reduce the choices in developing conservation plans to deal with the threat posed by global warming.

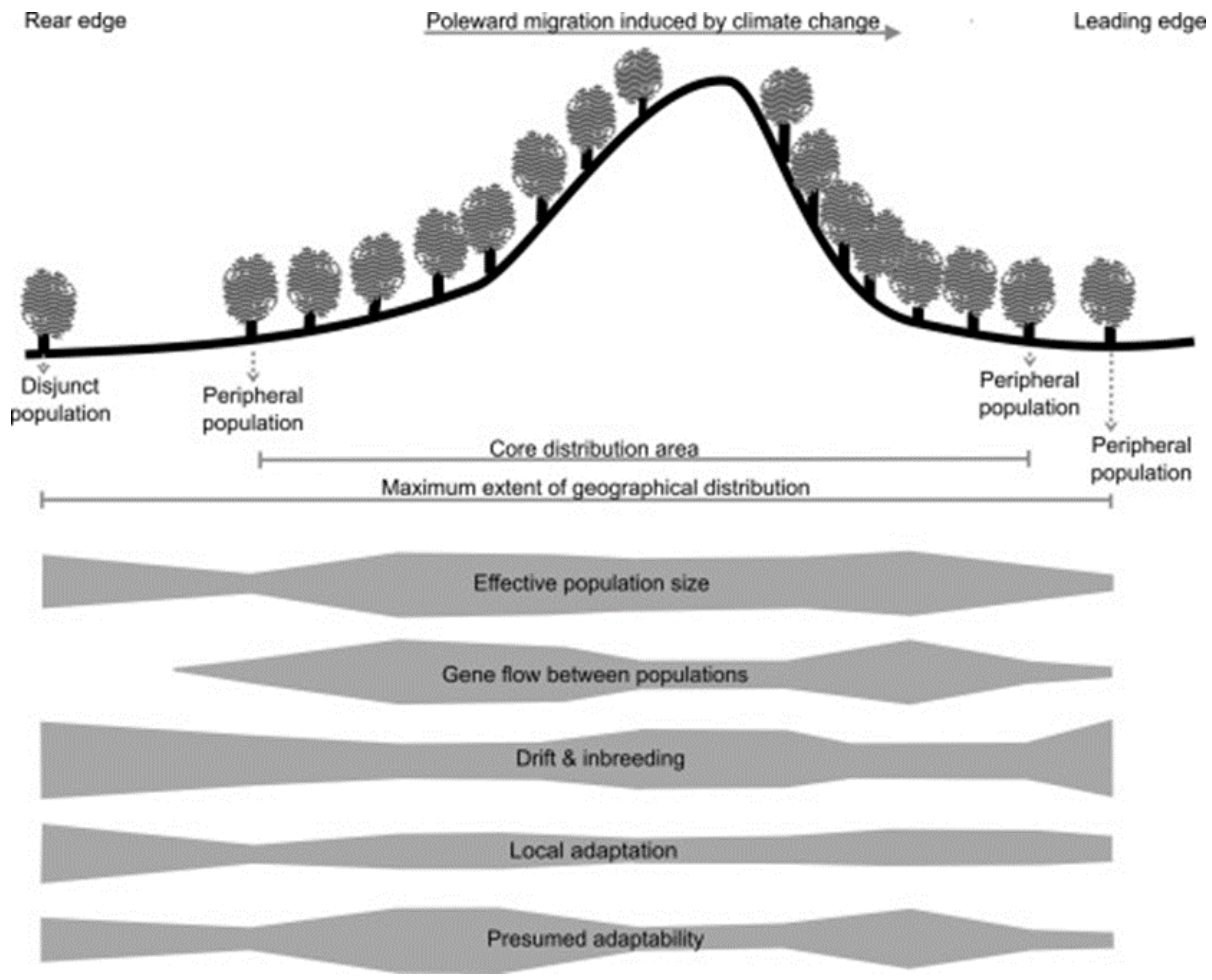

Figure 4. Demographic and genetic processes affecting forest tree populations across their distribution range. In the figure, the species range is fragmented and divided into two geographic entities, separated by a mountain. Grey shapes: Genetic and demographic processes in variable ways across the species distribution range, influenced by geography. Source: Fady *et al.* 2016

3.5. Scots pine communities and different habitat types in Central Europe

Natural selection in plant populations often have as result genotype – environment interactions. In the absence of other forces, selection should cause each local population to evolve traits that provide an advantage under its local environmental conditions. As a result, resident genotypes in each local population would have on average a higher relative fitness in their local habitat than genotypes originating from other habitats.

At first, the development of the seed is highly influenced by the environment, its characteristics being strongly determined by the genotype and the environmental factors acting on the mother tree. Secondly, the germination and growing of juvenile seedlings are also influenced through the environment, leading to local adaptation. In case of *Pinus sylvestris*, a species with a wide distribution, different populations are under the effect of different biotic and abiotic environmental conditions. The knowledge about the factors determining seedling establishment of Scots pine in different areas could thus aid the understanding of the patterns and processes involved in the recruitment of plant populations growing in contrasting ecological habitats. Since local adaptation in populations connected by gene flow must be due to recent natural selection, any information about the habitat may reveal specific genotype – environment interactions.

In Europe, Scots pine is distributed mainly in central and northern parts, and, in the south, is restricted to the high mountains of the Mediterranean basin (Boratyński 1991). In Central and Northern Europe, the main abiotic constraints for establishment are the low temperatures (Domisch *et al.* 2002, James *et al.* 1994), major biotic factors being the invertebrate herbivores (Nystrand and Granström 2000) pathogens (Burdon *et al.* 1994), or interference with the existing vegetation (Valkonen *et al.* 2002). The interactions between these factors exerts a powerful impact on the spatial pattern of recruitment, concentrating regeneration in some microhabitats and preventing seedling establishment in others (Burdon *et al.* 1994).

Current European distribution covers several climatic zones. These can be divided into oceanic, suboceanic, subcontinental, continental, alpine and Mediterranean (Ellenberg 1988). Scots pine also grows in a variety of soil conditions. The ecology is largely characterised by stress tolerance. On one hand this allows it to occupy a range of habitats that are unfavourable to other tree species, through tolerating various combinations of climatic and edaphic stress, including low temperatures, extremes of acidity and alcalinity, extremes of waterlogging and of drought (Kelly and Connolly 2000). On the other hand, the species is excluded from more favourable sites through competition. Thus in Central Europe its natural ecological range is restricted to dry acid, wet acid and dry alkaline soils through its inability to compete with species such *as Fagus sylvatica* and *Abies alba* (Ellenberg 1988).

The limits of *Pinus sylvestris* distribution correlate well with the -1º C isotherm for the mean temperature of the coldest month, and the +33º C isotherm for the mean temperature of the warmest month (Dahl 1988). Scots pine is thought to be kept at its northern limit by summer temperatures, while *Picea abies* is thought to be controlled more by early and late winter conditions (Engelmark 1999).

East-Central European communities

The East-Central European region is in the north temperate zone and is bissected by the 50th parallel. The climate is characterized by a change from a fairly warm, frost-free summer, to a more or less cold winter. The differences between the two seasons are more or less reduced by the position of being between the oceanic western and the increasingly continental eastern side. Therefore the air temperature in summer seldom exceeds 30° C, and in winter only exceptionally falls below -20° C. This situation causes the transition periods of spring and autumn to be significantly extended and the growing season for many plants to be correspondently longer. Furthermore the fact that cyclonic rain may fall at any time of the year is favourable for the vegetation of Central Europe. A very important character of the continental climate, which favours conifers over broadleaved deciduous trees is the shortness of the growing period (Ellenberg 1988). The region is dominated by broadleaved forest, in which conifer species only become significant under certain edaphic and environmental conditions or at higher elevations and latitudes. This is somewhat given that fir, spruce, pine and larch are able to grow in almost every habitat that supports broadleaved trees and Scots pine is the species, that has the broadest physiological amplitude of all native trees (Leuschner and Ellenberg 2017).

In East-Central Europe, the area of the Carpathian Basin is about 300,000 km^2. It is highly complex in terms of its geology, topography, climate and vegetation. It consists of the following main geographical units: peripheral mountains (the Carpathians, Alps, Dinaric Alps, and Transylvanian Mountain Range) and fertile alluvial plains and basins in the middle of the region (Little Hungarian Plain, Great Hungarian Plain, Transylvanian basin, Drava-Sava interfluve). The Alps and the Dinarides include calcareous and neutrophilous coniferous forests, predominantly in dry, warm climate areas, on rocky soils, especially on calcareous base rock. These stands are relatively competitive, where, despite their slower growth and relatively high light necessity, *Pinus sylvestris* and *Pinus mugo* can be favoured, and where their special adaptations can successfully exploit all these conditions. Their associations prefer the slowly and mostly physically decomposing rocks (dolomite, hard limestone and serpentine) (Borhidi 2003). These lime-grove Scots pine communities persists always in extreme habitats, in calcareous, dry conditions.

They are present in the Alps and in the northern part of the Dinaric Mountains, only in a small area of Central and Eastern Europe, always appearing as relic populations.

These hot, dry habitats are present on dewatered sites, on soils formed on limestone and with nutrient deficiency **(Figure 5**). Due to the lime base, the sparse canopy of the pine forest and to its light-rich forest interior, the poor water supply and the debris rich soil, the vegetation is characterized by calcareous and drought-tolerant species. Furthermore, due to the strong leaching, the tempering effect caused by the accumulation of spruce needle acidity, calcareous and mesophilic species are also characterizing these associations. The plant populations are small. Typical species: *Anthericum ramosum*, *Buphthalmum salicifolium*, *Calamagrostis varia*, *Cytisus supinus*, *Cytisus nigricans*, *Linum flavum*, *Senecio ovirensis*, *Thesium bavarum* (Borhidi 2003, Bölöni *et al.* 2011, Pócs 1960).

Figure 5. Scots pine stand on rock surface in the Eastern Carpathians, Cohárd mts. (RO). (photo credit: M. Höhn)

The 1,500 km long chain of the Carpathians consist of southwest to northeast belts of different ancient crystalline and Mesozoic blocks and Tertiary volcanic mountains (Inner Peripheral Volcanic Ranges with some small volcanic islands), and a Cretaceous-Tertiary folded mountain belt with Molasse Belt which can be found in the middle zone of this range and constituting the outermost part of the Carpathian region (Rudner and Sümegi 2001) (**Figure 6**).

Figure 6. Coniferous peat forest in the Eastern Carpathians, Fenyőkút (RO) peat bog. (photo: Z. A. Köbölkuti)

The climatic conditions in the Carpathian Basin are determined by its geographical position and topographical relief. Four climatic regions meet in this area (Rudner and Sümegi 2001). The western part of the Carpathian Basin is influenced by oceanic factors. The southern part of the Great Hungarian Plain and Transdanubia have a sub-Mediterranean influence. Continental climatic effects influence the eastern and central parts of the basin. A highland and submontane climate prevails in the mountains and lower mountain ranges. Thus, this climatic pattern further complicates the picture by causing large-scale altitudinal and latitudinal variations in both precipitation and temperature.

The vegetation of the Carpathian Basin is classified according to climate and topography (Rudner and Sümegi 2001) in several zones. The first temperate zone (0–600 m above sea level) contains Central and Eastern European forest types including *Quercus petraea*, *Q. pubescens*, *Carpinus betulus*, *Tilia cordata*, and *Acer platanoides*. Temperate zone 2 (0–600 m a.s.l.) is characterised by sub-Mediterranean thermophilous forest type with *Quercus pubescens*, *Q. dalechampii*, *Q. frainetto*, *Fraxinus ornus*, and *Carpinus orientalis*. Temperate zone 3 (700–1,700 m) contains west-central and southern forest types with *Fagus sylvatica*. Pontia-anatolian zone 1 is marked by forest-steppe and steppe zones with *Quercus robur*, *Q. pubescens*, *Tilia tomentosa*, and *Acer tataricum*. This vegetation zone developed in the central part of Great Hungarian Plain (Rudner and Sümegi 2001).

A specific Hungarian Scots pine habitat is Fenyőfő, located at the height of about 270 meters in the northern Bakony mountains of Trans-Danubian region. It is a loose forest-steppe like community where the herb layer preserves species of dry, pannonic sandy substrate.

The indigenous associations are on low humus sandy soil formed on lime consisting sand. The water management of the stands is poor, extremely dry or dry. Vegetation is mostly *Pinus* on sandy soil (*Festuco vaginatae – Pinetum sylvestris*), other types of *Pinus* (*Festuco rupicolae –Pinetum sylvestris*) and mixed sessile oak (*Quercetum petraeae – cerris pannonicum*) (Dövényi *et al.* 2008, Pócs 1960). The area is divided into several compartments with various sizes from two to 30 hectares. Unfavorable conditions of soil (low water capacity, sandy soils with calcium carbonate content in the topsoil) together with high temperature values in summer periods in the last years led the trees to die of or suffer from disease (Bölöni *et al.* 2011).

South-western Hungarian Transdanubian forests like Petőhenye are characterised by xerotherm forest association that is rich in continental elements. It is a relict-type association. The stands are on sand substrate, with shallow humus layer, or on carbonate residues, brown forest soils. The soils are alkaline, dry, with poor water management (Borhidi 2003).

3.6. Adaptation processes of Scots pine

A major focus of plant functional biology is the study of evolutionary adaptation and its basis in traits such as photosynthetis, morphology and development. On intraspecific level, the study of adaptation is to compare traits among some species' different populations that differ in or share a common ecology (Ackerly 2000). The comparative approach to the study of adaptation has a long and distinguished history (Darwin 1867, Harvey and Pagel 1991) and is motivated by observations of divergence in form and function and convergence among taxa from different environments or life forms and of convergence among taxa from similar environments or life forms (Ackerly and Reich 1999, Ehleringer and Monson 1993, Reich *et al.* 1999).

The process of adaptation to evolution by natural selection can be divided into two phases: (1) changes in the distribution of a phenotypic trait within a generation from differences in survivorship, growth, fertility, or mating success among parents with different phenotypes and (2) the evolutionary response to selection between generations, consisting of genetically based changes in trait distribution in offspring that result from fitness differences among parental phenotypes (Geber and Griffen 2003).

Trees with their long life span are more vulnerable than annual plants to rapid climate changes since they are not able to respond by migration or genetic selection in a short period of time. Many current models predict that rapid temperature increases as well as other climate changes will occur within the lifetime of one generation of trees (Schwartz 1991). Beside this, the overwhelming causes of population decline and extinctions world-wide are habitat destruction and the introduction of exotic species of parasites, predators, and competitors.

The restoration and maintenance of healthy habitats and ecosystems should be of great concern for conservation biology. The long-term preservation of biodiversity requires understanding the demography and genetics also of small populations. Although adaptive evolution can occur by mutations of large effect, the divergence in the quantitative traits that distinguish both different populations within a species and closely related species usually has a polygenic basis (Senner *et al.* 2017). In the short-term, genetic variability is often less critical than other determinants of population persistence (Lande and Shannon 1996), but in long-term, it can play decisive role in allowing a population to persist and adapt to a changing environment.

In order to examine intraspecific responses of Scots pine to changing environmental conditions, several experiments have been established. These experiments have been used to model climate change effects and to study genotype ×environment interactions (Fries 2017, Mátyás 1994, Oleksyn 1988, Oleksyn and Bialobok 1986, Semerci *et al.* 2017, Tarakanov *et al.* 2017). Only in the last two years several simulation models (Ermakova and Zuev 2017, González de Andrés *et al.* 2018), laboratory experiments (Domisch *et al.* 2017, Kulmala *et al.* 2017) and field studies (Ghimire *et al.* 2017; Rasheed *et al.* 2017, Repo *et al.* 2017) have been published to study the effects of different factors and to examine intraspecific responses. All these studies indicate that populations from different parts of the waste range of Scots pine have distinct eco-physiological differences. After all, all those analyses demonstrate that differentiation can comprise significant segments of the genome, because it influences many characteristics associated in some way with fitness.

Several characters show significant differentiation between Scots pine populations from different localities: winter dormancy (Pakharkova *et al.* 2016, van der Maaten *et al.* 2017) seed characteristics, germination (Geras'kin *et al.* 2016, Semerci *et al.* 2017), anthocyanin *(*Razzak *et al.* 2017) flowering time (Franke *et al.* 2017, Whittet *et al.* 2017), stem morphology (Kivimäenpää *et al.* 2017, González de Andrés *et al.* 2018, Kulmala *et al.* 2017), needle morphology (Jankowski *et al.* 2017, Schiestl-Aalto and Mäkelä 2017, Schönbeck *et al.* 2017) water uptake (González de Andrés *et al.* 2018, Martínez-Sancho *et al.* 2017), transpiration *(*Tor-Ngern *et al. 2017,* Wang *et al.* 2017) and photosynthesis (Feichtinger *et al.* 2017, Gao *et al.* 2017, Riikonen 2017). Both continuous and discontinuous character gradients have been documented for single-gene polymorphisms and for quantitatively varying characters. Nearly all of the studies reveal that the species is quite variable and extensively differentiated on a large geographic scale, for morphological (Köbölkuti *et al.* 2017, Zadworny *et al.* 2017) physiological *(*Aaltonen *et al.* 2016, Timofeeva *et al.* 2017) isozyme (Geras'kin *et al.* 2016, Volkova *et al.* 2018), or nucleic acid (Kujala *et al.* 2017, Stival *et al.* 2017, Tereba *et al.* 2017) traits.

In early studies of variation along geographical sites, one major focus of the analyses was to determine whether the variation showed a continuous clinal pattern or a discontinuous ecotypic one, or none. The same results were sometimes interpreted as clinal or ecotypic by different groups of researchers. Now, detailed analyses of variation show that, within the same species, some characters can vary gradually, others discontinuously, depending on, for example, gene flow, intensity of selection, or number of genes involved.

There is a preponderance of the association in all types of environments for all types of characteristics. These characteristics span many traits of plant anatomy, morphology, function, development, chromosomal makeup, life history, and biochemistry. At first, different environments generate different selection pressures, and these, in turn, lead to genetic heterogeneity. Secondly, environmental heterogeneity such as elevation, exposure, and water availability quite often generates significant barriers to gene flow, and thus, it enhances genetic differentiation among partially isolated or isolated populations. These two forces work synergistically, producing genetic heterogeneity in populations. Genetic differentiation can also be produced by genetic drift and isolation by distance (Scalfi *et al.* 2009), but these forces are considered to be less important in the connection between environmental and genetic heterogeneity across small spatial and temporal scales (Linhart 1988). Exceptions to this pattern can be describe in case of special historical events, or in case of low genetic variability (Linhart 1989). It is also true that very high levels of phenotypic plasticity often show low levels of genetic variability.

Selection can operate if specific traits are associated with fitness differences (Endler 1986), if trait frequencies or distributions vary between parent and offspring. The frequencies of heritable traits vary in the presence of selective factors.

Physical and biotic components of the environment often differ from each other in the patterns of differentiation they generate. Physical features (moisture, soil conditions, exposure) are typically contiguous in space and time, producing patterns of differentiation between neighbouring populations or generations (Linhart and Grant 1996). In contrast, biotic components vary much more dynamically, for the reason that competitors, herbivores and parasites can move within a given area. Plant competitors or parasites can move by seeds or spores from one generation to the next and they may undergo evolutionary change, having impact on plant population structure (Linhart 1989). Consequently, microscale differentiation patterns generated and maintained by biotic forces tend to be less sharply defined than those associated with changes in the physical environment. Biotic components can produce differentiation within populations but typically in a mosaic, unsegregated way. Thus, individuals with higher fitness, for example those that are better competitors or more parasite resistant, grow intermixed with individuals of lower fitness *(*Fritz *et al.* 1999).

3.6.1. Factors of differentiation in space

Soils – as long as edaphic conditions are extreme in terms of pH, mineral contents, or other features, they generate selection pressures. Soils generally characterized by high levels of magnesium and low levels of calcium, commonly produce strongly differentiated populations (Mayer *et al.* 2017). The forces producing this differentiation in plants of these regions produce effects including specialized morphology (Ostonen *et al.* 2007), physiology (Feichtinger *et al.* 2017), development (Hari *et al.* 2017). Kruckeberg (1992) reports intraspecific, genetically based differentiation in resistance to soils with a very low silica content across many plant taxa that include pines (Furnier and Adams 1986). Cyanogenic morphs occur with high frequencies in carbonate soils, but at much lower frequencies in acidic silicate soils (Miller and Cumming 2000). Several decades ago cyanogenesis already has been recognized as an important antiherbivore defense in several species (Dirzo and Harper 1982), but it may also be important in the nitrogen economy of plants (Cánovas *et al.* 2007). Specifically, cyanogenic glucosides may serve as nitrogen storage structures, accumulated in times favorable for rapid nitrogen cycling, to be used later under more stringent nitrogen conditions. For this reason, cyanogenic plants may be favored under the physiological conditions associated with carbonate soils (Urbanska 1984), which tend to be nitrogen deficient. Differentiation can be associated also with specific physiological features (Hatchell *et al.* 1970); other studies have detected differences in subpopulations on different soil types at the level of aconitase, alcohol dehydrogenase, diaphorase, fluorescent esterase, glutamate dehydrogenase, glutamate-oxaloacetate transaminase, glucose-6-phosphate dehydrogenase, hexoseaminidase, isocitrate dehydrogenase, leucine aminopeptidase, malate dehydrogenase, menadione reductase (Furnier and Adams 1986). Soil moisture, nutrients, and pH can also affect distribution of resident microorganisms such as mycorrhizal fungi (Balogh-Brunstad *et al.* 2017, Wang *et al.* 2017). As a result, plant populations spanning gradients of soil conditions may also be associated with different communities of soil microorganisms.

Moisture, temperature, and elevation – These factors often vary. Consequently, studies of differentiation that have been done along such habitat gradients consider them as acting together. For example, genetic differentiation has been demonstrated in frequencies of allozyme alleles between spire-shaped trees within closed- canopy forests and individuals growing nearby in both *Abies lasiocarpa* and *Picea engelmanii* (Grant and Mitton 1977). Differences between these habitats include temperature extremes, snow accumulation, insolation, wind, herbivory and competition on seedlings. Authors cannot determine which of these factors are important in producing the genetic differences observed. Also can be concluded that there are significant genetic differences associated with growth morphology, and detectable at the scale of 100 m or

less (Di Pierro *et al.* 2017, Lind *et al.* 2017). Small-scale gradients also occur on the sides of small depressions that fill with water. As the water evaporates and recedes, microhabitat zones with different characteristics including moisture availability, soil pH, temperatures, soil aeration, and vegetation composition appear. Genetic differentiation was demonstrated in response to water-logging of phenology and reproductive output (Cendán *et al.* 2011, Delpierre *et al.* 2016).

Light intensity – Shade is often provided by adjacent plants, so it is, in part, abiotic effect with which plants must cope, but its primary impact is the physical reduction of available light. Evidence suggests that if adjacent habitats are characterized by predictable conditions that create different selective pressures, populations will show evolutionary divergence *(*Merry *et al.* 2017, Riodrío *et al.* 2017, Zhu *et al.* 2014).

Competition – Plants compete each-other for light, water, nutrients, space, pollinators, and other necessities. Therefore, competition entails many different kinds of competitive interactions (González de Andrés *et al.* 2018, Uria-Diez and Pommerening 2017, Wang *et al.* 2017). Because of interspecific differences in morphology, development, physiological requirements, and tolerances, intraspecific and interspecific competition are likely to differ (Linhart and Grant 1996).

Herbivory, predation, and parasitism – Interactions between Scots pine and their herbivores and parasites can lead to two broad classes of selection scenarios. In some cases, adjacent populations are either exposed or not to specific herbivores or parasites, and under these conditions there can be spatially identifiable patterns of genetic differentiation in response to fungi (Kwaśna *et al.* 2017), mollusc herbivory (O'Reilly-Wapstra *et al.* 2014), or insect parasitism (Kovalchuk *et al.* 2015). The important roles of fungi and arthropods in shaping evolutionary change in plants have also been demonstrated, usually in studies involving interactions between a specific plant species and a single fungus or insect species (Burdon 1987). Most herbivores and parasites seldom have spatial distributions with sharp boundaries. For this reason, differentiation also seldom involves contiguous plant populations, but rather it forms a mosaic of differentiated individuals within populations. In the presence of a disease organism or parasite, some individuals are susceptible while others can be either tolerant or resistant. Patterns of species-specific host selection are not always associated with a single compound. In *Pinus ponderosa*, relative amounts of a whole suite of compounds in xylem and phloem determine whether a tree will be attacked by the beetle *Dendroctonus ponderosae*, the dwarf-mistletoe *Arceuthobium vaginatum*, or the squirrel *Sciurus aberti* within stands where all these species coexist. Trees are seldom attacked by all three (Linhart 1989). One potential consequence of multispecies challenges to a host plant is diversifying selection (Gillespie and Turelli 1989). In the case of traits controlled by several loci, diversifying selection can operate when individual hosts with different phenotypes are favored

under different conditions (Hay and Fenical 1988). In contrast to diversifying selection, stabilizing selection can also occur if different dependent species differentially attack individuals at the extremes of a continuously distributed trait. Directional selection can occur if two dependent species favor similar or positively correlated phenotypes, or no selection may be detected (Linhart and Grant 1996).

3.6.2. Population differentiation in time

A temporal component is relevant in differentiation and adaptation of populations in two contexts. One is to determine rates of evolutionary change. The other is that rapid changes in environmental conditions can expose different groups within a population to different selection pressures, thereby generating temporal differentiation in genetic constitution. Because ecological conditions can change quickly over time, different groups can be exposed to different selection regimes at different times. This means that such age-structured populations can consist of groups occupying the same site but having very different genetic constitutions (Linhart 1989).

Temporal differentiation may also occur in the context of succession; early arriving colonisers experience one set of environmental conditions, and their progeny, acclimatized through establishing in the same geographic area, may face different conditions of light, moisture, competition, herbivory, and other factors, so different genotypes are selected (Beckman *et al.* 2016). Temporal differences with changes in the mating system can also be stored in seed banks. Such seed banks, in turn, can be stored either in soil or in cones kept on plants for several decades (Gibson and Hamrick 1991).

3.6.3. Characteristics of Scots pine affected by differentiation

Local differentiation in *Pinus sylvestris* has been documented for most important features of its structure and function. Examples can be found for seed size (Sevik and Topacoglu 2015), needle pubescence (Repo *et al.* 2014), seed germination (Semerci *et al.* 2017), seed colour (Tikhonova *et al.* 2014), root morphology (Zadworny *et al.* 2017), plant stature (Woodruff 2016), flowering phenology (Duputié *et al.* 2015), photosynthetic abilities (Kolari *et al.* 2014), heavy metal tolerance (De Beeck *et al.* 2015), herbivory resistance (Zadworny *et al.* 2017), pathogen resistance (Giertych and Mátyás 2017), response to competitors (Wang *et al.* 2017), and organelle traits (García Gil *et al.* 2015).

3.7. Morphological traits variation of Scots spine in Central – Eastern Europe

Under extreme conditions, peripheral populations with specific structures are exposed to dramatic environmental changes which will impose novel selection pressures and may therefore cause adaptive responses (Bone and Farres 2001). Although trees are usually well adapted to local environment with high phenotypic plasticity, the composition and structure of populations is influenced or determined by a combination of historical events, geological changes in the growing sites, and *in situ* adaptation to ecological factors (Losos 1996). The adaptedness and the adaptability of edge populations to varying environmental conditions are dependent on the diversity accumulated in the gene stock and long potential exposure to divergent selection pressures (Gregorius 1989).

Although each peripheral population possesses a specific migration history, natural selection over extended periods of time leaves traces on gene based morphological and anatomical traits. Natural populations in the refugial areas, in addition to possess imprints of historical events, they are also distinctive due to adaptations to different habitat extremes on the periphery. This is supported by studies performed on populations from Central and Northern Europe, the Balkans, Iberia, and Anatolia (Alía *et al.* 2001, Bilgen and Kaya 2007, Dzialuk *et al.* 2009, Jasińska *et al.* 2014, Labra *et al.* 2006, Mejnartowicz 1979, Prus-Glowacki and Stephan 1994, Prus-Glowacki *et al.* 2003, Pyhäjärvi *et al.* 2007, Semiz *et al.* 2007, Staszkiewicz 1961, Tobolski and Hanover 1971, Turna 2003). Based on these works, it can be concluded that populations in refugial areas display not only “geographic marks”, but also genetic differences and differentiation at the phenotypic level.

As plants are integrated systems, the study of traits that influence their survival and reproduction can reveal diversification in contrasting environments (Reich *et al.* 2003). However, populations or genotypes can be preadapted to a given selection factor or environmental condition. When they colonize new habitats or geographic areas, their survival can depend on their functional traits, which are or are not suited to the environment (Reich *et al.* 2003). Thus, current species distributions may reflect ecological pre-sorting processes, in addition to *in situ* adaptive evolution (Losos 1996), indicating their origin and their historical colonization routes. In this aspect, morphological patterns leading to local adaptation are fundamental in plant life histories, and have profound consequences on many aspects of plant ecology and evolution. By this way, morphological studies offer a feasible alternative to compare local populations, which have evolved under different conditions. The small populations in Central - Eastern Europe persisted in extreme habitat types with elevated groundwater tables (such as peat bogs) or grow on sunny, rocky surfaces (Urbaniak 1994). They continue to survive under diverse environmental conditions within a scattered geographic area.

3.7.1. Variations in cone morphology and needle anatomy

As phenological marks, needles are directly exposed to the environment's physical factors, such as altitude, air temperature, atmospheric pressure, precipitation, and wind (Aaltonen *et al.* 2016, Friend and Woodward 1990, Kivimäenpää *et al.* 2016, Körner 2007, Mõttus *et al.* 2017, Schiestl-Aalto and Mäkelä 2017, Tiwari *et al.* 2013). They are also responsible for photosynthesis, carbon assimilation, and exchange of gas and water, and they may vary in overall dimensions, as well as in details of key anatomical characteristics important in processes of adaptation (Donnelly *et al.* 2016).

Cone formation is also known to be variable, depending greatly upon climatic factors (Ovington 1957). Differentiation of populations from the Iberian Peninsula has been described using needle (Boratynska and Hinca 2003, Jasińska *et al.* 2010) 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 based on cones and needles are also listed as distinguishing characteristics among populations in the work of Bobowicz and Korczyk (2000). Cone characteristics are traits of the highest discriminating power in inter-populational comparisons among regions in the work of Bobowicz and Korczyk (2000) and

Jasińska *et al.* (2014). The principal variables which proved to be indicative of discriminating populations were also found to be needle characteristics by Androsiuk *et al.* (2011). Several studies which describe differentiation based on morphology among populations of other species from the genus *Pinus* can also be found in the literature of Marcysiak (2004), Bączkiewicz *et al.* (2005) and in the work of Sobierajska and Marcysiak (2010). Morphological traits of other conifer species are also known to vary adaptively with geographic, climatic, and edaphic variables (Ji *et al.* 2011). In several works in the existing literature, significant differences between isolated populations have been observed, mostly with regard to their cone and needle morphological features (Szweykowski and Urbaniak 1982). The information due to structural characteristics of conifer needles are often strongly related to gradients in long-term light availability within canopies and across stands also can be found in several studies (Lhotáková *et al.* 2007, Niinemets *et al.* 2007, Richardson *et al.* 2001, Richardson *et al.* 2000). Scots pine, typically occuring on different types of well-drained mineral soils, representing a broad range of variation in pH, nutrient availability, and vegetation (Persson 1980) has not only the geographical range, but also the ecological tolerance very wide. For that reason, variation in cone and needle morphological traits is an important characteristic of ecological processes that are driving forces for biogeochemical cycles in ecosystems (Reich *et al.* 1992). These studies allow us to gain important insights and predict ecosystem responses to changes in the environment.

3.7.2. Differentiation by the morphology and germination of seeds

Among morphological surveys, reproductive phenological observations are some of the most sensitive data in identifying how plant species respond to regional climate conditions and to climatic changes. As reproductive organs, seeds are in highly specific interaction with the environment, influence not only the germination of individuals, but also the distribution of the species. It is important for the storage of life in the context of protection of the embryo, and its shape, size and weight are strongly influenced by the genotype and the environment where the mother tree grows (Castro 1999).

The morphology of the seed is in its turn dependent upon the genetic constitution and the modifying effect of the environmental factors acting on seed formation and seed maturity (Andersson 1965). At the first instance, seeds have specific morphological adaptations that influence seed movement into suitable germination microsites (Chambers and Macmahon 1994). These variation in morphological characters could be due to the fact that the species grows over a wide range of rainfall, temperature and soil type. Soil properties, climate, and disturbance

characteristics determine the physical attributes and micro- topography of exposed soils. In turn, these soil attributes influence both the horizontal and vertical movement of seeds.

After seed germination, the new individuals would react to soil, climate and management with a type of growth and development that is optimal under the offered conditions, emphasizing the genetic component of performance (Andersson 1965), but also the existence of ecological factors which influence individual and stand development. Variations in seed morphology in relation to habitat have been reported in a number of tree species (Kaushik *et al.* 2007). In *Pinus*, variations can mainly be attributed to the influence of the mother tree on the genetic composition of the seed coat and gametophyte, and to environmental conditions during seed development (Surles *et al.*1993). The size of the seed easily changes, not only with the climatic conditions of the year but even with the difference in cone size, the number of seeds per cone and the position within the cone (Ehrenberg *et al.* 1955). As a broad generalisation, it appears that *Pinus* species associated with stressful environments have smaller cones (Richardson and Rundel 1998) and it is intuitive to expect the largest seeds in the largest cones, but Keeley and Zedler (2000) showed, that correlation between cone size and seed size is week.

The approximately 110 species of the genus *Pinus* (Richardson and Rundel 1998) exhibit one or two seed-dispersal systems. At some haploxylon *Pinus* species i.e. *Pinus cembra*, *Pinus pinea* the seeds are enveloped in stone-hard shell, are wingless and disseminated by animals, some are large-seeded pines like Jeffrey pine (*Pinus jeffreyi*) and sugar pine (*Pinus lambertiana*) with wind-dispersed winged seeds, but of which the dispersal can be also animal-mediated (Vander Wall 2002). The seeds of several diploxylon species, among them also *Pinus sylvestris* L. are dispersed mostly by the wind. Their seeds are typically small with relatively large wings that have the potential to carry them well beyond the canopy of the parent tree. The cones open at seed maturity and shed seeds in the fall (Castro 1999).

As it was discussed more in detail in the previous chapter, the peripheral populations of Scots pine occur on different types of well drained mineral soils, representing a broad range of variation in pH, nutrient availability and sustain on different vegetation types. They can be found both in habitats with high ground water table such as peat bogs, or in other extremeties like dry rocky substrates or forming dry coniferous forests. During adaptation, certainly there are differences in genetic characters with regard to juvenile growth. To a large extent, however, the differences can be referred to casual characteristics of the seeds, embryo development, endosperm state, seed weight and in maturity in relation to the climate (Ehrenberg *et al.* 1955). The distance over which the tree disperse seeds depends on the plant traits as well as environmental conditions and varies strongly in time and space. The observation regarding the distance of seed deposition from the source is very difficult methodologically, but according to Castro (1999) an area of 100

m^2 might be required to catch a single seed. Studies of seed dispersal do not seem to have been done in significant number. As the phenotypic variation could be a result of different parental environments, the production of polymorphic seeds (differing in size, shape, colour, germinability or dispersability) is thought to broaden the range of conditions under which Scots pine can germinate and thus, increases the chances of reproducing in an unpredictable environment.

Understanding the nature, extent and pattern of variation existing in some peripheral populations in respect to seed characters and the degree of correlation between the morphology of the seeds and their germination can reveal significant differentiation among populations growing in different habitats. The results obtained can be evaluated as signs of local adaptation with detectable phenotypic patterns. Taking into consideration that the production of seeds with different sizes and shape may be a more effective and evolutionary stable strategy than the production of uniform crop (Haig 1996), can be presumed, that local adaptation to extreme ecological sites could have consequences not only on cone and needle morphology but also on the morphological characters and germination of seeds.

3.8. The importance of the provenance researches in the detection of local adaptation

Many species occur over large geographical regions, sustaining populations in a wide range of ecologically distinct conditions. As a result, individuals in different parts of the species' range may experience substantially different environmental challenges which can result in different phenotipic traits and subsequently in different local adaptation (Kawecki and Ebert 2004). Marginal populations of a distribution area are useful study systems, as they raise questions about the nature of evolutionary adaptation to local environments and about the limits to this adaptation (Vergeer and Kunin 2013). Common garden experiments provide a powerful tool in differentiating adaptive, non-adaptive and plastic changes in plant traits across a species' range (Kawecki and Ebert 2004).

Provenance research is the study of common garden plantations of tree populations originating from geographically different locations or different type of habitat. These studies are generally done to examine traits of significant economic interest, such as viability, growth rate, wood quality, frost hardiness, resistance to drought and diseases (Krutovsky 2006). The provenances have their origin seeds collected from identified stands or regions. The early juvenile stage, early growth, and different developmental and physiological traits of forest trees can be studied with success on level of provenances and correlate well with the climatic data at the location of origin (Matyas 1996). In common garden experiments, a number of genotypes are

grown in a common environment in order to quantify the genetic component of phenotypic variation (Moloney *et al.* 2009). Such experiments are ideally suited for disentangling how genetic and environmental factors contribute to the success of the species in their new non-native range. The original principal goal of provenance tests was to identify stands, populations or areas which provide the most desired traits and commercially best results at the test location. They often demonstrate geographic differences for the studied traits, suggesting the specific genetic basis of the given adaptive variation, which often has a latitudinal or altitudinal clinal form, caused by gradient natural selection and local adaptation. It is a classical approach to quantify genetic based phenotypic differentiation among populations, differences due to evolutionary responses within the species' range (Colautti *et al.* 2009).

Beyond the morphological variation study, these practice-oriented intentions also offer the opportunity to analyse intraspecific allozymic, chemical, physiological diversity (Matyas 1996). With the increase in understanding of the significance of genetic diversity, provenance tests have become valuable approaches for breeding and conservation, representing a powerful tool in testing climatic adaptation in trees, the observed geographic variation being interpreted as an adaptive response to climatic conditions. Provenance tests can be utilized also for biological site indication, the growth and vigor of populations characterising the environment. To demonstrate the plant-habitat interaction for fitness requires an experiment in which samples of the local genotypes are directly compared under the same environmental conditions. Another alternative is recreating the properties of different habitats in greenhouses, experimental plots (Kawecki and Ebert 2004). The tradition to investigate genetic variation among forest tree provenances dates back to 1823 (Matyas 1996).

Provenance tests with different conifer species were established over time, the distinct evolutionary and ecological implications making these tests important and interseting objects to study beyond silvicultural applications. Scots pine (*Pinus sylvestris*) is the species with most seed orchards have been established in Europe (Burczyk 1991). From the last three decades' literature can be mentioned several common garden studies with conifers. Reich *et al.* (1996), studied needle retention and longevity of spruce and pine populations; Hosius *et al.* (2000) established seed orchards of silver fir (*Abies alba* Mill.), Alía *et al.*(2001), studied seeds of Spanish and German provenances of Scots pine (*Pinus sylvestris*) with common garden establishment; Gömöry (2003) measured fertility variation and flowering asynchrony in *Pinus sylvestris*, Oleksyn *et al.* (2003) studied nutrient conservation with latitude of origin in European *Pinus sylvestris* populations and needle nutrients in geographically diverse populations. Four species (*Cedrus atlantica* Manetti, *Pinus nigra* Arn. ssp. *nigricans* Host. var. austriaca, *Pinus halepensis* Mill. and *Cupressus sempervirens* L.) were planted in a common trial by Froux *et al.* (2004). Icgen (2006) made a

common garden trial with Turkish red pine (*Pinus brutia* Ten.); Strimbeck *et al.* (2007) assessed the midwinter low-temperature tolerance of foliage from eight temperate and boreal species in each of the genera *Abies*, *Picea*, and *Pinus*. The effects of seed transfer and climate change on the width and basal area of tree rings were studied in 21 provenances of jack pine (*Pinus banksiana* Lamb.) by Savva *et al.* (2007). Voltas *et al.* (2008) evaluated the growth and survival for several Aleppo pine populations covering its geographic range and grown in two common-garden tests; Mallik *et al.* (2008) worked with replicated plots of black spruce (*Picea mariana*), white spruce (*Picea glauca*), and red pine (*Pinus resinosa*). Paternity analyses using microsatellites were conducted in two conifer clonal seed orchards, *Abies alba* and *Larix kaempferi* by Hansen *et al.* (2008); Voltas *et al.* (2008) measured the climate-related variability in carbon and oxygen stable isotopes among populations of Aleppo pine. Vitasse *et al.* (2009) evaluated the altitudinal differentiation in growth and phenology among populations in tree species of temperate zones, among them *Abies alba* growing in a common garden. Soil faunal activity was measured by Rożen *et al.* (2010). Sevik *et al.* (2010) estimated the genetic diversity among populations in Scots pine (*Pinus silvestris* L.) seed stands of Western Black Sea Region in Turkey. Mueller *et al.* (2012) investigated effects of temperate tree species (among them conifer species) on interactions among carbon, nitrogen, and acidity in mineral soils from an experiment with replicated monocultures. Salmela *et al.* (2013) evaluated spring phenology's genetic variation among and within populations in seedlings of Scots pine. Study about ectomycorrhizal fungal communities of native and non-native *Pinus* and *Quercus* species in a common garden of 35-year-old trees was performed by Trocha *et al.* (2012).

At the last years also several work can be mentioned: Whittet *et al.* (2017) studied variation in the timing of pollen production between distant populations of *Pinus sylvestris* in Scotland, Rudawska *et al.* (2017) were performing a study with forest litter amendment during nursery stage with influence of field performance and ectomycorrhizal community of Scots pine (*Pinus sylvestris*) seedlings. Semerci *et al.* (2017) followed morphological and physiological responses to drought stress of European provenances of Scots pine.

Provenance tests of trees are time consuming, require large areas and are very costly. Nevertheless, they became an important procedure in maintaining vitality and growth vigor of forest stands, being applied to several conifer species. In time, they have been used to study quantitative genetic traits, but beyond this, they also serve as an excellent opportunity for the analysis of intraspecific diversity and testing geographic variations and adaptive responses in different climatic conditions (Matyas 1996).

3.9. Functional genomic markers with role in adaptation research

3.9.1. Development and application

Genetic markers have many different uses in plant breeding, evolutionary, and conservation studies. One of the goals of population genetic studies based on markers is inferring the amount and distribution of variation in adaptively important traits. During the last four decades, the advancement in the field of molecular biology has facilitated development of a range of molecular genetic markers. Resolution and usability of DNA markers improved considerably with the advancement in the techniques of nucleic acid hybridization, polymerase chain reaction and lately DNA sequencing.

The rapid development of next-generation sequencing (NGS) technologies provides an opportunity to develop novel genomic tools. These technologies are appreciated for their advantages such as accuracy, high throughput and relatively low cost, and are widely applied also in qualitative and quantitative analysis of transcriptomes (Zhang *et al.* 2012). Development of gene based molecular markers led to the popularity of RNA-based (or cDNA or transcriptomic or functional) markers preceding development of alternative techniques for genome and transcriptome profiling. Massively parallel sequencing of cDNA now is an efficient route for generating enormous sequence collections that represent expressed genes. This approach provides a valuable starting point for characterizing functional genetic variation in non-model organisms, especially where whole genome sequencing efforts are currently cost and time prohibitive.

RNA (cDNA) sequencing has accelerated gene discovery when genome sequences are not available, facilitating gene identification and development of new molecular markers. The lower cost and greater sequence yield has allowed the identification of genes, even when they are expressed at low levels. Genotype- and developmental stage-specific transcriptomes (Garg *et al.* 2016, Miguel *et al.* 2015) generated from multiple individuals within a species, from different geographical origins, at different growth stages and subjected to environmental stresses under natural field condition, will ultimately lead to a comprehensive catalogue of gene expression.

Furthermore, transcriptome, or Expressed Sequence Tag (EST) sequencing is an efficient means to generate functional genomic level data for non-model organisms or those with genome characteristics prohibitive to sequence the whole genome. The method is an attractive alternative to whole genome sequencing because the majority of most eukaryotic genomes is non-coding DNA, and EST sequences lack introns and intragenic regions for that the analysis and interpretation of data is more difficult. ESTs thus have a high functional information content, and often correspond to genes with known or predicted functions (Andersen and Lübberstedt 2003). Large collections of EST sequences have proven to be unavoidable for gene annotation and

discovery (Emrich *et al.* 2007), comparative genomics (Vera *et al.* 2008), development of molecular markers (Barbazuk *et al.* 2007), and for population genomic studies of genetic variation associated with adaptive traits (Parchman *et al.* 2010).

The analysis of gene content and measurement of gene expression level in non-model organisms, are also very helpful for developing molecular markers, such as single nucleotide polymorphisms (SNPs) (Ashrafi *et al.* 2012). Single nucleotide polymorphisms (SNPs) identified in coding regions represent a useful tool for understanding the adaptive response against pathogens and stressful environmental conditions (Núñez-Acuña and Gallardo-Escárate 2013). Using the RNA-seq technologies transcriptomes have been sequenced and characterized in many plants, e.g. *Rubus, Cucurbita, Isatis* (Hyun *et al.* 2014, Tang *et al.* 2014, T. Wu *et al.* 2014), or willow (Rao *et al.* 2014).

The large and complex genomes of pines have hindered the fast development of genomic resources, despite the ecological and economical importance of the group. While most genomic studies have focused on just a few species (*Pinus taeda, Pinus lambertiana, Pinus cembra, Pinus contorta*), genomic level resources for other pines are insufficiently developed to facilitate ecological genomic research. After sequencing the transcriptome, *de novo* assembly is facilitated by the possibility of increased coverage depth for the much smaller number of nucleotides in the transcriptome than in the whole genome (Emrich *et al.* 2007). In addition, the reduced amount of repetitive DNA found in genes compared to non-coding regions ameliorates one of the principal obstacles to de novo assembly of short reads (Pop and Salzberg 2008). The construction of large EST collections is thus the most promising approach for providing functional genomic level information in pines (Neale 2007). EST collections will contribute to the development of molecular markers for members of the *Pinus* and are going to facilitate comparative genomics and the study of adaptive variation across the genus.

3.9.1.1. Genome annotation

With the completion of sequencing of the human and other mammalian genomes (Human Genome Sequencing Consortium 2004) scientists have turned their attention to the annotation of genomes for functional content, including gene-coding transcription units and cis-acting regulatory and epigenetic elements that modulate gene expression (ENCODE Project Consortium 2004). The advances in the fields of high-throughput genomic technologies and bioinformatics have paved the road for biologists to approach a comprehensive catalog of functional elements in the genome and

discover unrecognized patterns in the genomic context. The process of identifying genomic elements and their functions is referred to as genome annotation.

Several definitions of the process can be found in the literature, here I guote the MedicineNet.com: „*Genome annotation: The process of identifying the locations of genes and all of the coding regions in a genome and determining what those genes do. An annotation (irrespective of the context) is a note added by way of explanation or commentary. Once a genome is sequenced, it needs to be annotated to make sense of it."* The "unit" of genome annotation is the description of an individual gene and its protein (or RNA) product, and the focal point of each such record is the function assigned to the gene product. It is not only essential for interpretation of the genome sequences but also an active research area in the field of genomics. It will improve the process to select molecular markers and to approach underlying systems (Abugessaisa *et al.* 2014). Genome sequences reflect the history of life, and are shaped by negative selection against disadvantageous genomic variations (alleles), positive selection of favored ones, and genetic drift on neutral ones. Many of the protein coding sequences (CDSs) are conserved. Promoter regions of noncoding transcripts (ncRNA) are conserved as well as promoters of protein-coding transcripts (Carninci 2005). These ultra-conserved regions, where an almost complete identity is observed between orthologous regions of different species are relevant components in genome annotation.

The first step towards the successful annotation of any genome is determining whether its assembly is ready for annotation. Several summary statistics are used to describe the completeness and contiguity of a genome assembly, and by far the most important is N50. A contig N50 is calculated by first ordering every contig by length from longest to shortest. Next, starting from the longest contig, the lengths of each contig are summed, until this running sum equals one-half of the total length of all contigs in the assembly. The contig N50 of the assembly is the length of the shortest contig in this list (Earl *et al.* 2011). Other useful assembly statistics are the average gap size of a scaffold and the average number of gaps per scaffold (Li *et al.* 2010). Two assemblies can have identical scaffold N50s but can still differ in their percent gaps. Estimates of gap lengths are often made based on library insert sizes and read lengths; when these are available, the number of 'N's in these gaps usually represents the most likely estimate of that gap's size. Although there are no strict rules, an assembly with an N50 scaffold length that is gene-sized is suitable for annotation (Yandell and Ence 2012). Although genome annotation pipelines differ in their details, they share a core set of features. Some of the main procedures of annotation such as predicting gene functions, statistical gene prediction and search of general-purpose databases for sequence similarity are central in the sense that this is done as part of any genome project.

Generally, genome-wide annotation of gene structures is divided into two distinct phases. The first phase is the computation phase, during which EST-s and protein sequences are aligned

to the genome and gene predictions are generated. In the second phase, all these results are synthesised into genome annotations. Not considering the annotation of transcriptomes, during gene prediction the single most likely coding sequence (CDS) of a gene is going to be found, and the result do not report untranslated regions (UTRs) or alternatively spliced variants (Wang *et al.* 2010). Gene annotation is thus a more complex task than gene prediction.

A genome annotation protocol must not only deal with heterogeneous types of evidence in the form of the expressed sequence tags (ESTs), RNA-seq data, protein homologies and gene predictions, but it must also synthesize all of these data into coherent gene models and produce an output that describes its results in sufficient detail to become suitable inputs to genome browsers and annotation databases (Kent *et al.* 2002). Because this process is complicated and involves many different tools, the programs that assemble all the data and use it to create genome annotations are generally referred to as annotation pipelines (Brent 2005).

3.9.1.1.1. The computation phase of the annotation

The first step in the computation phase of genome annotation is to find low-complexity sequences as well as transposable elements, such as viruses, long interspersed nuclear elements (LINEs) and short interspersed nuclear elements (SINEs) (Kapitonov and Jurka 2003, 2008). Repeats complicate genome annotation, thus they need to be identified and annotated, but the tools used to identify repeats are distinct from those used to identify the genes of the host genome. Their identification is complicated by the fact that repeats are often poorly conserved (Alba and Guigo 2004). Available tools for finding repeats can be grouped into two classes: homology based and *de novo* tools. Homology-based repeat identification is powerful at detecting transposable elements that share sequence similarity with known elements, but it is inadequate at identifying full length or novel ones. Methods using *de novo* approaches can discover all TEs as long as they have multiple copies. However, the disadvantage of this approach is that its output is a mixture of TEs from all superfamilies and non-TE repeats. For that reason, the manual identification and classification of TEs from the output of *de novo* methods is often very time consuming (Han and Wessler 2010). Usable programs that have been developed exclusively to find repeats: TRANSPO (Santiago *et al.* 2002), FINDMITE (Tu 2001) and MUST (Chen *et al.* 2009). TRANSPO is a homology-based program, FINDMITE and MUST are structure-based transposable element discovery programs that can be used because they search for common structural features rather than similar sequences. However, *de novo* tools identify repeated sequences — not just mobile elements — so their outputs can include highly conserved protein-coding genes, such as histones

and tubulins, as well as transposon sequences. After it has been created, the repeat library can be used in conjunction with RepeatMasker 3.3.0 (Chalopin and Volff 2017), which uses BLAST (Altschul *et al.* 1990) and Crossmatch (Green *et al.* 1998) to identify stretches of sequence in a target genome that are homologous to known repeats. The masking step signals to downstream sequence alignment and gene prediction tools that these regions are repeats. Left unmasked, repeats can produce millions of false BLAST alignments, with the result of several incorrect evidence for gene annotations.

After this step, proteins, ESTs and RNA-seq data have to be aligned to the genome assembly. These sequences include previously identified transcripts and proteins from the organism whose genome is being annotated. Sequences from other organisms are also included. Generally, these are restricted to proteins, as these retain substantial sequence similarity over much greater spans of evolutionary time than nucleotide sequences do. In principle, TBLASTX (Altschul *et al.* 1990) can be used to align ESTs and RNA-seq data from phylogenetically distant organisms. UniProtKB/SwissProt (Barrell *et al.* 2011) is an excellent tool for protein sequences. One easy way to assess protein and EST data sets is to download from related organisms using the NCBI taxonomy browser (Wheeler *et al.* 2008). EST and protein sequence data sets have to be aligned to the genome in two steps: at first, BLAST is used to identify rapidly approximate regions of homology, then alignments are usually filtered to identify and to remove marginal alignments on the basis of percent similarity. After filtering, the remaining data are clustered to identify overlapping alignments. By clustering, results are grouped into a single cluster of data, all supporting the same gene and are identified redundant evidences like highly expressed genes. After clustering, highly similar sequences identified by BLAST are realigned to the target genome to obtain greater precision at exon boundaries. BLAST, although rapid, has no model for splice sites, and so the edges of its sequence alignments are only rough approximations of exon boundaries (Slater *et al.* 2005). For this reason, splice-site-aware alignment algorithms, such as Splign (Kapustin *et al.* 2008), Spidey (Wheelan *et al.* 2001), and Exonerate (Slater *et al.* 2005), are often used to realign highly similar ESTs, mRNAs and proteins to the genomic input sequence.

Of all, RNA-seq data have the greatest potential to improve the accuracy of gene annotations, as they provide copious evidence for better delimitation of exons, splice sites and alternatively spliced exons. RNA-seq reads can be assembled *de novo*, independently of the genome, using tools such as ABySS 2.0 (Jackman *et al.* 2017), SOAPdenovo (Xie *et al.* 2014) and Trinity (Grabherr *et al.* 2011), the resulting transcripts being realigned to the genome in the same way as ESTs. The tools for a fast and easy identification of genes in assembled DNA sequences are called *ab initio* gene predictors. They use mathematical models rather than external evidence (such as EST and protein alignments) to identify genes and to determine their intron–exon

structures. These tools have practical limitations from an annotation perspective. For instance, most of them find the single most likely coding sequence and do not report untranslated regions or alternatively spliced transcripts. Gene predictors use organism-specific genomic traits, such as codon frequencies and distributions of intron–exon lengths, to distinguish genes from intergenic regions and to determine intron–exon structures. As they come with precalculated parameter files that contain information for a few classic genomes, they need to be trained on the genome that is under study, as even closely related organisms can differ with respect to intron lengths, codon usage and GC content (Korf 2004).

In principle, alignments of ESTs, RNA-seq and protein sequences to a genome can be used to train gene predictors even in the absence of pre-existing reference gene models. The most important *ab initio* tools are: TwinScan, FGENESH, Augustus, GAZE and SNAP (Korf 2004), Geneid (Grenville-Briggs *et al.* 2017), Genemark (Min *et al.* 2017), GenomeScan (Yeh *et al.* 2001), Conrad (Decaprio *et al.* 2007), Contrast (Gross *et al.* 2007), Gnomon (Souvorov *et al.* 2010), GeneSeqer (Schlueter *et al.* 2003).

3.9.1.1.2. The annotation phase

The ultimate goal of annotation efforts is to obtain a synthesis of the alignment-based evidences with *ab initio* gene predictions to obtain a final set of annotated genes. This process, if it's done manually, results in high-quality annotation, but it is extremely labour-intensive. For this reason, smaller genome projects are increasingly being done by automated annotations. The simplest form of automated annotation is to run different gene finders on the genome and then to use an algoritm, which can select the best prediction whose intron–exon structure represents the best the consensus of the models with the overlapping predictions. This process can be done by JIGSAW (Allen and Salzberg 2005) and EVidenceModeler (Altimiras *et al.* 2017). Like *ab initio* gene predictors, JIGSAW must be trained for each new genome, EVM allows a set of expected evidence error rates manually. GLEAN and Evigan require no additional training. Usually, JIGSAW, EVM or Evigan are performed similarly. Another approach is to give the alignment evidence to the gene predictors at run time. This is the process used by PASA (Haas *et al.* 2011), Gnomon (Souvorov *et al.* 2010) and MAKER (Cantarel *et al.* 2008). NCBI uses BLAST alignments together with predictions from Gnomon and GenomeScan to produce gene models (Pruitt *et al.* 2012).

Doing the annotation process, from several pipelines available, should use the best approach in terms of effort versus accuracy. Simply running a single *ab initio* gene finder over even a very large genome can be done in a few hours. By contrast, a full run by an annotation

pipeline such as MAKER or PASA can take weeks, but because these pipelines align evidence to the genome, their outputs provide starting points for annotation curation and downstream analyses, such as differential expression analyses using RNA-seq data. Another factor to consider is the phylogenetic relationship of the studied genome to other annotated genomes. If it is the first of its taxonomic order or family to be annotated, it would definitely be preferable to use a pipeline that can use the full repertory of external evidence, especially RNAseq data (Elsik *et al.* 2007).

As long as tools and sequencing technologies continue to develop, periodic updates to every genome's annotations will remain necessary. Incorrect and incomplete annotations can undermine every experiment that makes use of them. For that reason, providing accurate and up-to-date annotations is extremely important. But doing it properly, it can lead to the discovery and application of new molecular markers to high-throughput and ultrahigh-throughput levels. These newly developed markers will lead onto high-density genetic map construction, identification of QTLs, breeding and conservation strategies.

3.9.1.2. Primer design

The next-generation sequencing methods have revolutionized genomic and transcriptomic approaches in biology. These new sequencing tools are also valuable to discover, and assess genetic markers in forest tree populations. Many biological questions can now be answered with high accuracy (Helyar *et al.* 2011). But the study of these discrete molecular fragments of the naturally occurring DNA needs the multiplication of the template molecule, and this is not possible without molecular cloning.

The *de novo* synthesis of oligonucleotides and their alignment into a new DNA molecule is called polymerase chain reaction (PCR). The method consists in repetitive cycles of denaturation, hybridization and polymerase extension, increasing the amount of the target molecule, defined by the positions of the two primers on the template DNA. So before performing a polymerase chain reaction experiment, a pair of primers to clip the target DNA subsequence is required. Optimal primer sequence and appropriate primer concentration are essential for maximal specificity and efficiency of PCR. While the outcome of PCR can be influenced by many other conditions such as the template DNA preparation and reaction conditions, designing a good pair of primers is a critical factor.

The ability for an oligonucleotide to serve as a primer is dependent on the kinetics of association and dissociation of primer-template duplexes at the annealing and extension temperatures, the duplex stability and location of mismatched nucleotides, and on the efficiency

with which the polymerase can recognize and extend the mismatched duplex (Abd-Elsalam 2003). Due to the properties of oligonucleotides that influence the efficiency of the PCR amplification, the optimal primer design includes criteria (Clackson *et al.* 1991) such as melting temperatures, length, base composition, termini, repeated and self-complementary sequences and complementarity between members of a primer pair. To ensure that the primers will be efficiently annealed during each cycle of the PCR, the calculated values for the melting temperature of the forward and reverse part of a primer pair should not differ by more than 5 °C. The primers lengths should be between 16 – 28 nucleotide. The lengths of the members of a primer pair should not differ by more than 3 bp. The GC content of the members of a primer pair should be between 40 and 60 %. The nature of the 3' end of the primers is also crucial, if possible, the 3' end of each primer should be G or C. Actual differences for these criteria are aggregated by weighting sums (Abd-Elsalam 2003). Additional requirements may also apply in certain cases. In reverse transcription PCR, (RT-PCR), to avoid unwanted amplification, it is recommended that a primer pair span an intron, or that one of the primers to be located at an exon-exon junction. Another concern is the possible impact of SNPs in the primer regions. Since a SNP may act as a mismatch, one should consider picking primers outside of such regions (Ye *et al.* 2012). One critical primer property is the target specificity. Ideal for a primer pair is to amplify only the intended target.

The primer design is the process of construction candidates and to select the best one. Various kinds of approaches in designing primers have been proposed in the last few decades. The manual primer design method can find a primer that fits the primer design constraints. However, it is too time consuming, besides, accuracy can easily be lost through human errors. Since experiments are expensive, and a minor mistake may cause the experiment to fail, the manual primer design method is considered to be potentially unstable (Wu *et al.* 2004). The design of primers for PCR has been intensely studied, especially in application-specific contexts. In recent years, several software programs have been developed to aid in primer design. Examples include OligoArray (Rouillard *et al.* 2003), FindGDPs (Blick *et al.* 2003), VPCR (Lexa *et al.* 2001), UniFrag/GenomePrimer (van Hijum *et al.* 2003), Primer3 (Rozen and Skaletsky 1999), CODEHOP (Rose *et al.* 2003), Gene Fisher (Giegerich *et al.* 1996) and several others. Almost all these packages focus on physical primer design: ensuring that an oligonucleotide has appropriate annealing and melting characteristics. PUNS (Primer UniGene Selectivity) is a system which complements physical primer design by controlling primer specificity within the transcriptome and the genome (Boutros and Okey 2004).

The multitude of these programs for PCR primer design reflects the central role of PCR in modern molecular biology. Nevertheless, all these programs were not developed to replace the human experience. When planning a PCR experiment, it is worthwhile to evaluate the predictions

of numerous different programs and to use the laboratory experience of the researcher to evaluate the suggested primers before any resort to their synthesis use.

3.9.2. Overview of the newly developed molecular markers with possible role in adaptation. Enzymes, transcription factors and their coding genes

Detection and analysis of genetic variation can help us to understand the molecular basis of the biological phenomena of adaptation. Since the entire plant kingdom cannot be covered under sequencing projects, molecular markers and their correlation to phenotypes provide us with requisite landmarks for elucidation of genetic variation. Genetic markers are ultimately based on the variation of DNA sequences, however, mostly particular aspects of the variation are investigated of the currently most commonly used genetic markers in *Pinus*. The amplified genomic regions are usually either unknown or are located in non-coding regions of the DNA. Accordingly, most of the variation at molecular DNA-based markers is assumed to be selectively neutral. Furthermore, the enzymes and transcription factor coding sequences can be used to assess important functions in the metabolism of the plant, with possible role in adaptation.

The most promising markers for the study of this biological phenomena at the moment are SNPs (Brookes 1999). The analysis is not restricted to special enzymes, SNP markers can also be used to analyse regions with role in the transcription of genes , for example, transcription factor coding sites. Unfortunately, SNP analyses in human populations revealed, that only few SNPs can be associated with phenotypic traits (Yoshiura *et al.* 2006). Some of these SNPs with a direct impact on phenotypes are likely to be under selection, while the wast majority of SNPs are likely to behave selectively neutral. Candidate genes can be chosen based on literature surveys suggesting an impact of the genes on the adaptation process. **Table 1** presents the candidate genes selected on base of literature, annotated and PCR cloned during the study.

Table 1. Selected candidate genes from the Pinaceae family related to adaptation based on literature survey.

Name (abbreviation)	Gene	Role in the adaptation process	Reference with investigated species
APX	Ascorbate peroxidase	Role in freezing tolerance during cold acclimatization; against high production of toxic oxygen; low light; pollution	Tao *et al.* 1998 (*Pinus sylvestris*); Anderson *et al.* 1992 (*Pinus strobus*); Tang 2000 (*Pinus taeda*); Polle *et al.* 1992; Whittet *et al*. 2017); Yang *et al.* 2008 (*Picea asperata*); Yamazaki *et al*. 2003 (*Abies mariesii*); Polle *et al*. 1996 (*Pinus sylvestris*); Pukacka and Pukacki 2000 (*Pinus sylvestris*);
CHS; CHI, F3H, F3'H, F3'5'H, LDR; ANS	Chalcone synthase; chalcone isomerase; flavanone hydroxylase; flavonid 3'hydroxylase; flavonoid 3'5' hydroxylase; leucoanthocyanidin reductase; anthocyanidin synthase	Role in responses to abiotic signals, including UV damage, temperature fluctuations and low availability of nutrients and water; fungal infection;	Anderson *et al.* 1992 (*Larix decidua, Picea abies, Pinus silvestris P. parviflora, Abies procera, A. koreana, and A. concolor*); Schröder and Rennenberg 1992 (*Pinus sylvestris*); Messner *et al.* 1991 (*Picea abies*); Teklemariani and Blake 2004 (*Pinus banksiana*)
MDH	NAD malate dehydrogenase	Role in responses to adverse environmental conditions, invading organisms and ultraviolet irradiation	El-Kassaby 1981 (*Abies concolor, Abies grandis; Picea abies; Picea glauca; Picea mariana; Picea sitchensis; Pinus banksiana; Pinus contorta; Pinus ponderosa; Pinus rigida; Pinus syJvestris; Pinus taeda; Pseudolsuga menziesii; Thuja plicata*)
SOD	Superoxide dismutase	Role in response to toxic oxygen species	Polle and Rennenberg 1991 (*Pinus sylvestris, Picea abies*); Wingsle and Hällgren 1993 (*Pinus sylvestris*)
diTPS	Diterpene synthase	Role in response to biotic and abiotic environmental factors	Miller *et al.* 2005 (*Pinus strobus*); Johnson and Croteau 1987 (several conifer species)
MADS	MADS box proteins	Role in floral development, organ identity	Chen *et al.* 2017; Di Stilio *et al.* 2017; Mouradov *et al.* 1996; Mouradov *et al.* 1997; Rutledge *et al.* 1998; Slater *et al.* 2005; Sundström *et al.* 1999; Tandre *et al.* 1995 (*Picea abies, Picea mariana, Pinus radiata, Pinus resinosa*)
MYB	MYB transcription factor superfamily	Role in tolerance to freezing, drought and salt stress; plant hormone and pathogen-mediated stress responses; regulation of flavonoid accumulation; in regulation of lignin synthesis enzymes	Bedon *et al.* 2007, 2010 (*Picea glauca, Pinus taeda*); Xue *et al.* 2003 (*Picea mariana*); Osakabe *et al.* 2009 (gymnosperms)
WRKY	WRKY transcription factor proteins	Role in reprogramming of plant immune responses; regulation of disease resistance	Ersoz *et al.* 2010 (*Pinus elliottii, Pinus sylvestris, Pinus radiata, Pinus taeda*); Donini *et al.* 2009 (*Pinus monticola*)

4. MATERIAL AND METHODS

4.1. Study sites, sampling design for cone morphological, needle anatomical and molecular study of Scots pine

For the cone morphological and needle anatomical study 16 natural populations of *Pinus sylvestris* from the Carpathians and the Pannonian Basin were sampled between 2011 - 2015 (**Table 2**, **Figure 7**).

Table 2. Location and population parameters of the studied *Pinus sylvestris* populations from the Central and Eastern European peripheral distribution of the species.

No.	Pop. Abbrev.	Country	Residential area	Latitude (N)	Longitude (E)	Altitude (m)	Est. size (km^2)	Habitat
1	HKO	Hungary	Kőszeg	47.22	16.33	630	0,04	Mixed forest
2	HFE	Hungary	Fenyőfő	47.35	17.77	252	4.49	Mixed forest
3	HVE	Hungary	Pethőhenye	46.87	16.92	306	0.04	Mixed forest
4	HZA	Hungary	Szalafő	46.87	16.30	231	0.08	Mixed forest
5	HOR	Hungary	Csörötnek	46.93	16.35	296	0.10	Mixed forest
6	SKV	Slovakia	Kvacany	49.18	19.54	799	0.48	Rock surface
7	STU	Slovakia	Svarin	49.02	19.91	1107	0.70	Rock surface
8	SLI	Slovakia	Liptovszky Hrádok	49.04	19.74	729	0.02	Rock surface
9	RFE	Romania	Fantana Brazilor	46.50	25.26	953	0.32	Peat bog
10	RPO	Romania	Poiana Stampei	47.30	25.12	878	1.43	Peat bog
11	RMO	Romania	Baile Tusnad	46.13	25.91	1052	0.58	Peat bog
12	RPA	Romania	Voineasa	45.38	23.91	753	3.42	Rock surface
13	RBI	Romania	Roșia	46.84	22.37	393	0.13	Rock surface
14	RBE	Romania	Posaga de sus	46.49	23.36	524	0.84	Rock surface
15	RML	Romania	Ponor	46.33	23.34	925	0.10	Peat bog
16	RMH	Romania	Calatele	46.73	23.02	913	0.58	Peat bog

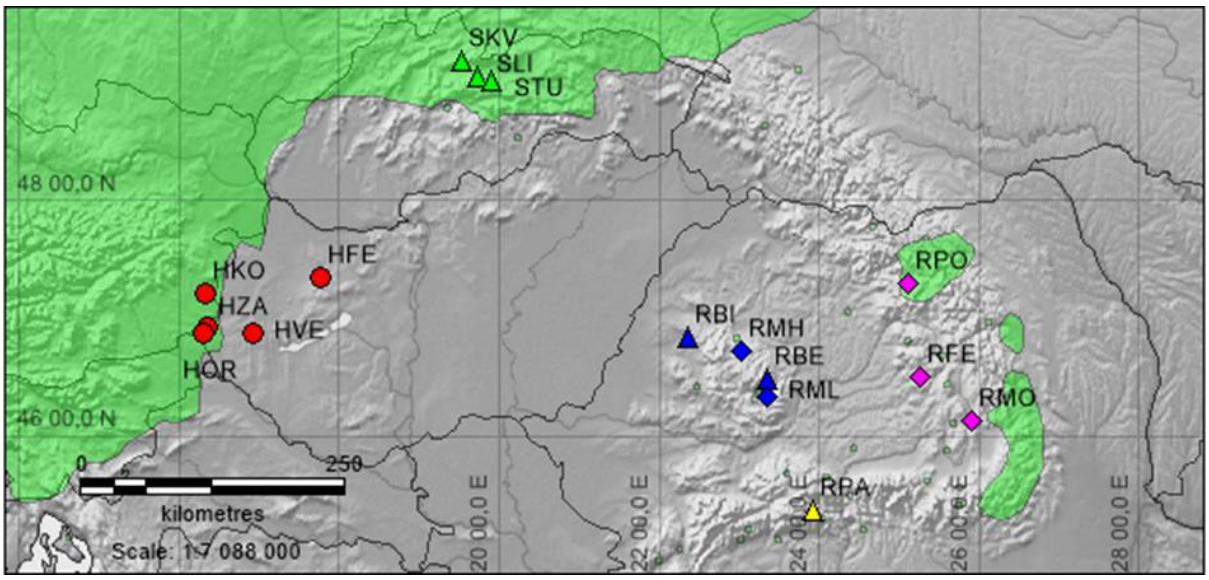

Figure 7. Sampled populations of *Pinus sylvestris* for cone morphological and needle anatomical study. The colors indicate the geographical affiliation (red: Pannonian basin (PB), green: Northern Carpathians (NC), blue: Central-Island Mountains/ Apuseni Mt. (CIM), pink: Eastern Carpathians (EC) and 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 Table 2.

All of these marginal populations within the natural range of the species occupy specific habitat types, such as raised bogs, dry rocky surfaces, or mixed forests on specific substrates with low nutrient content (**Figure 7**). These mixed forests have developed on less extreme substrates, on pebble, loam and clay, or directly on a solid silicate base. This, in addition to the typical Dicrano-Pinion species combination and acidophilic nature, allows a higher proportion of broad-leaved deciduous species in the composition of the association (Borhidi 2003, Pócs 1960). In this vegetation mixture of conifers and broad-leaved deciduous trees, Scots pine seems to be less competitive.

Trees were chosen using randomized sampling design at each stand. 25 to 40 fully ripened, brown, two years-old cones from each population and 10 - 20 two years-old brachyblasts with healthy needles were collected from each tree, from four tree per population. Needles were taken from 30 - 40 year-old adult trees at 2 - 3 m above ground level from a well-illuminated part of the crown. Leaf collections were stored in sealed plastic bags at -80° C in a freezer. Before measurements a cross section was made from the middle of each needle. Cross sectioning was carried out with an Ernst Leitz GMBH Wetzlar tissue microtome. Sections were stained with Toluidin–blue, washed with 10 % hydrochlorid acid, and placed on glass slides in glycerinated water.

4.2. Sampling design for the *Pinus sylvestris* seed and germination study; seed collection and storage

Since the cones were lacking from seeds in case of several populations sampled, the study of seed morphology and germination was completed on less populations than the previous cone morphological survey. The materials were collected from natural stands in Central - Eastern Europe, in the fall of 2015. Within the natural range of the species, we sampled populations with peripheral distribution as in case of the morphological study, from habitat types like raised bogs, dry rocky surfaces or mixed forests on lower elevation, characterized by specific competition features. The geographic locations and habitat conditions are presented in **Table 3** and **Figure 8**.

Table 3. Location and population parameters of the studied *Pinus sylvestris* populations from the Central and Eastern European peripheral distribution of the species, tested in the seed morphological and germination study.

No.	Pop. abbrev.	Country	Residential area	Latitude (N)	Longitude (E)	Altitude (m)	Est. size (km^2)	Habitat
1	RFE	Romania	Fântâna Brazilor	46.50	16.33	953	0.32	Peat bog
2	SME	Slovakia	Medzi bormi	49.16	19.37	813	0.06	Peat bog
3	RML	Romania	Ponor	46.33	23.34	925	0.10	Peat bog
4	RMO	Romania	Băile Tușnad -Mohos	46.13	25.91	1052	0.58	Peat bog
5	RPS	Romania	Poiana stampei	47.30	25.12	878	1.43	Peat bog
6	RBE	Romania	Poșaga de sus	46.49	23.36	524	0.84	Rock surface
7	RCO	Romania	Lacu roșu	46.47	25.47	1507	0.04	Rock surface
8	CHR	Czech Rep.	Hřensko	50.53	14.23	162	79	Rock surface
9	HZA	Hungary	Szalafő	46.87	22.37	231	0.08	Mixed forest
10	RVR	Romania	Putna-Vrancea	45.55	26.30	906	0.3	Mixed forest

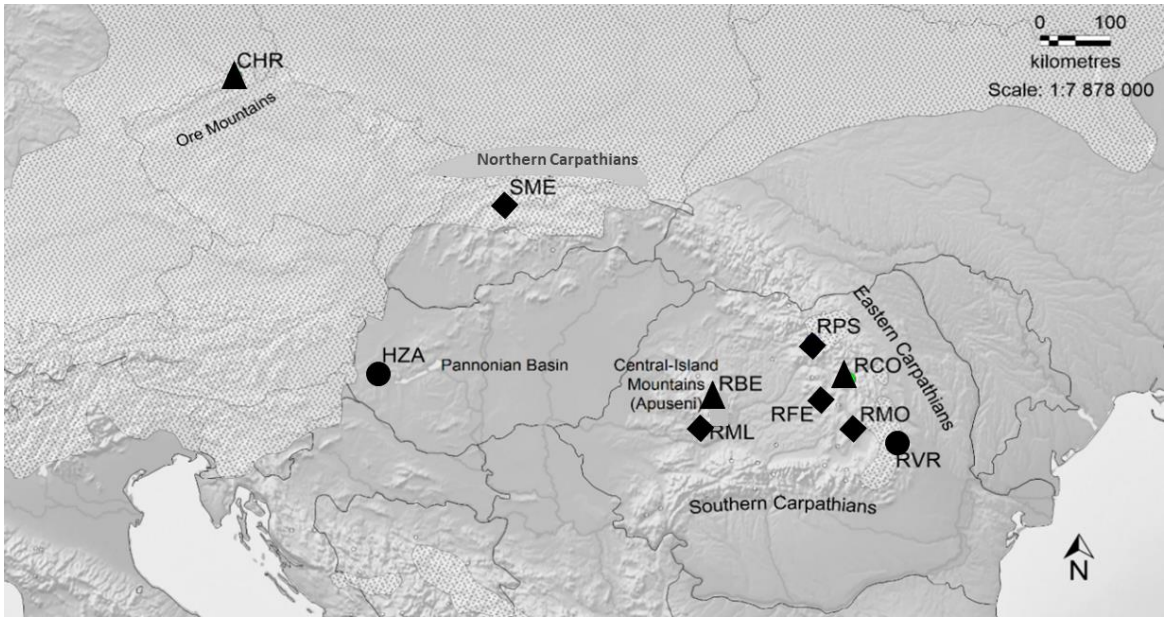

Figure 8. The populations of *Pinus sylvestris* analysed in the seed morphological and germination study. The shape of icons indicate the type of habitat (rectangulars: peat bogs, triangles: rock surfaces, circles: mixed forests) the distribution of the species (based on Euforgen map) is shaded. For population abbreviation see Table 3.

Four healthy trees in each population were randomly selected. 25 to 40, fully ripened, brown colored, two years old cones were collected from each tree. Cones were collected randomly from the crown of the sample tree. Only non-diseased cones were included in the study. Following the collection, cones were kept on room temperature, so that they opened after three weeks, then all seeds were extracted manually with a laboratory lancet and stored in paper bags, from each sample tree being wrapped separately.

4.3. The progeny trial

After the morphological measurements, the extracted seeds were used to establish a common-garden trial located in the Botanical Garden of Soroksár (47°24' - 19°09') in spring 2016. 15 seeds/mother tree were sampled from four mother trees per population and sown in pots of size 0.5 l (40 : 40 : 20 garden soil for conifers : peat : perlite), 0.5 cm depth in 5th of March 2016 under common-garden greenhouse conditions. During germination, seedlings were kept under natural light conditions (in a periodically shaded greenhouse) with watering applied one or two times per week (**Figure 9**).

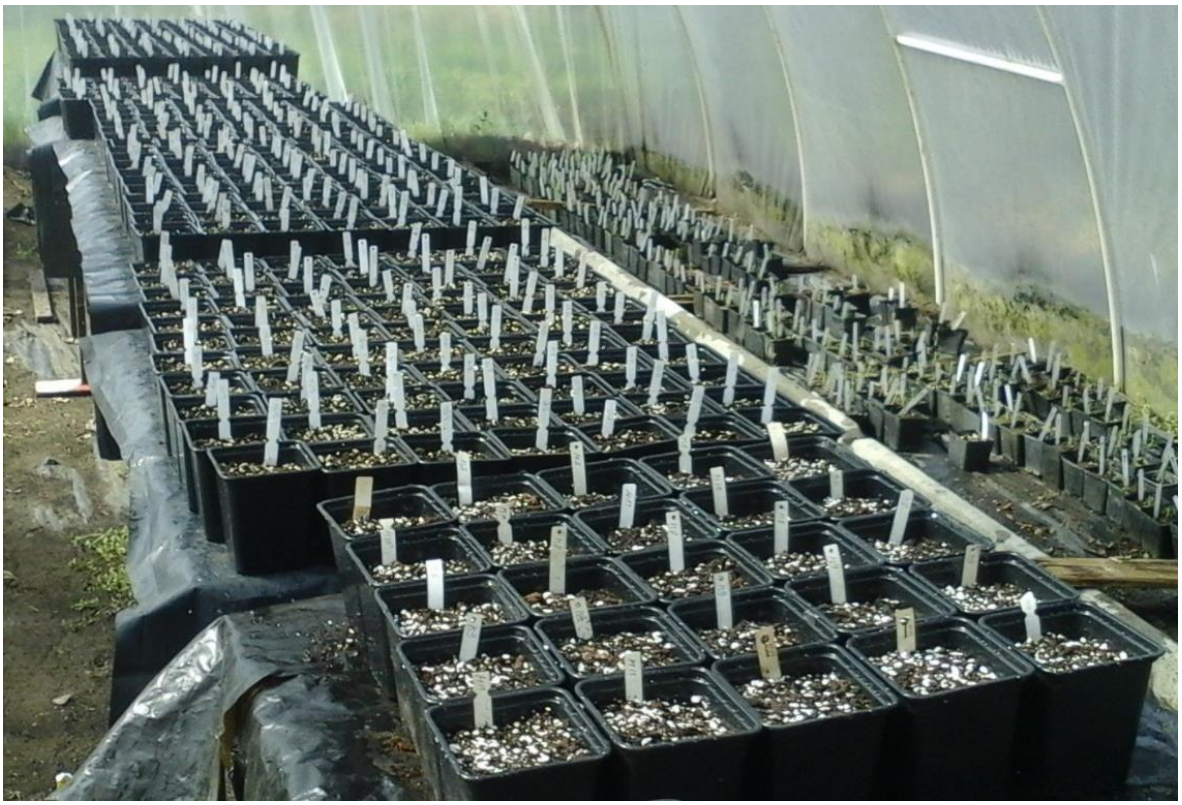

Figure 9. The *Pinus sylvestris* common garden trial located in the Botanical Garden of Soroksár in 5th of March 2016. (photo: Z.A. Köbölkuti)

4.4. Characters and their measurements

4.4.1. Cone morphological and needle anatomical characters

In reference to all twenty studied traits, six cone morphological and eight needle anatomical characters were measured and four cone morphological and four needle anatomical ratios were calculated. Six morphological characters were measured on each cone: cone width (CW), cone length (CL), number of scales (NBS), width (AW), length (AL), and thickness (AT) of apophysis (**Figure 10**). Four ratios were calculated: cone length/width (CL/CW), cone length/number of scales (CL/NBS), apophysis length/apophysis thickness (AL/AT), and apophysis length/width (AL/AW). These parameters were measured with an electronic caliper with 0.1 mm accuracy.

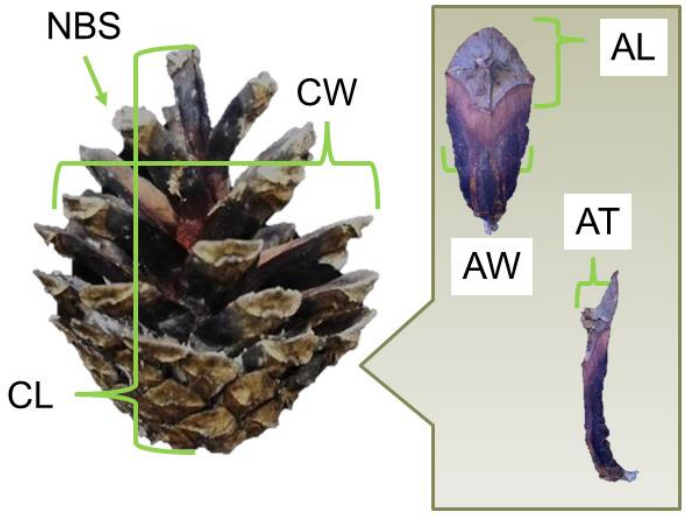

Figure 10. Measured cone parameters in Scots pine populations: number of cone scales (NBS), cone length (CL), cone width (CW); apophysis length (AL), apophysis width (AW) and thickness (AT). (photo: E.G. Tóth)

Needle anatomical data were obtained from 256 needles: four needles per tree on four trees per population (a total of 16 needles per population). The anatomical parameters (eight in total) measured for each needle were: number of resin ducts (NRD) and number of rows in the armed palisade parenchyma on the concave (NBRCC), convex (NBRCV) and crosswise mesophyll (NBRCW) (**Figure 11**). Additional characteristics (measured in µm) were assessed, including needle height (NH) and width (NW), the height (CCH) and width (CCW) of the central cylinder, and four ratios: central cylinder width/height (CCW/CCH), needle width/height (NW/NH), central cylinder height/needle height (CCH/NH), and central cylinder width/needle width (CCW/NW). The number of resin ducts and parenchyma cells were examined under an Olympus XC21 microscope, with the ocular and objective magnification of 10x and 40x, respectively.

The height and width of the needle and central cylinder were measured in micrometers (µm) using a Zeiss Axio Imager A2 microscope with 5x / 0.13 magnification, and they were photographed using a HRc Axiocam with AxioVision Microscopy software (Carl Zeiss, Germany).

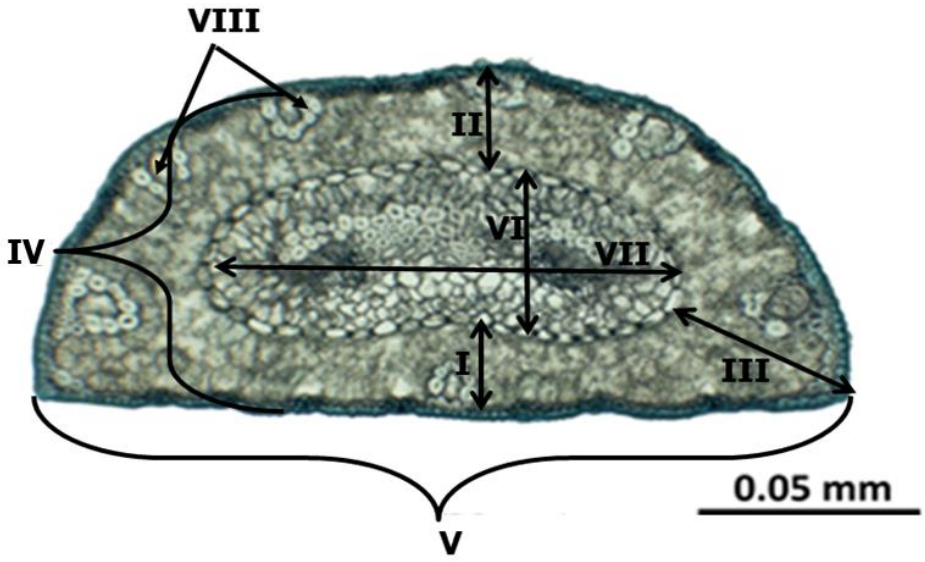

Figure 11. Measured needle anatomical parameters in Scots pine: I, II, III – number of rows in the armed palisade parenchyma on the concave (NBRCC), convex (NBRCV) and crosswise mesophyll (NBRCW); IV, V – needle height (NH) and width (NW); VI, VII – height (CCH) and width (CCW) of the central cylinder; VIII – number of resin ducts (NRD). (photo: Z.A. Köbölkuti)

4.4.2. Seed data collection

In reference to all studied traits (40 seed/population), five seed morphological traits were measured, and four ratios were calculated. On each winged seed four morphological characters were measured: seed length (SL); seed width (SW); wing length (WL); wing width (WW) (**Figure 12**). These parameters were measured with an electronic caliper with 0.1 mm accuracy. Ten seeds from every tree were taken to measure seed weight (g/ten seed) (SWE). Additional characteristics were assessed by calculating four ratios: seed length/seed width (SL/SW); wing length/wing width (WL/WW); seed length/wing length (SL/WL) and seed width/wing width (SW/WW).

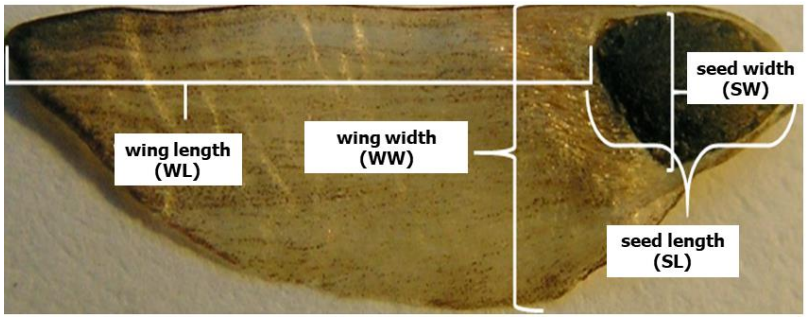

Figure 12. Measured seed parameters in Scots pine populations: seed length (SL); seed width (SW); wing length (WL); wing width (WW). (photo: Z.A. Köbölkuti)

4.4.3. Collection of germination data

Germination was scored once weekly between 5th of April to 19th of June 2016. We calculated Seed germination percentage using the following formula (International Seed Testing Association-ISTA 1985):

Germination % = Number of germinated seeds/Total number of seeds × 100.

Germination associate parameters were calculated by using:

a. Speed of germination (Ginwal *et al.* 2005):

$$GS = \frac{n1}{d1} + \frac{n2}{d2} + \frac{n3}{d3} + \ldots$$

Where, $n1$, $n2$, $n3$ = number of germinated seeds in day 1, 2, 3…; $d1$, $d2$, $d3$ … = day 1, 2, 3…

b. Mean Germination Time (MGT) (Roberts and Ellis 1989):

$$MGT = \frac{n1 \times d1 + n2 \times d2 + n3 \times d3 + \ldots}{\text{Total number of days}}$$

Where, $n1$, $n2$, $n3$ = number of germinated seed in day 1, 2, 3; $d1$, $d2$, $d3$ = day 1, 2, 3

c. Mean daily germination (MDG) (Ginwal *et al.* 2005):

$$MDG = \frac{\text{Total number of germinated seeds}}{\text{Total number of days}}$$

d. Peak Value (PV) (Ginwal *et al.* 2005):

$$PV = \frac{\text{Accumulated number of germinated seeds}}{\text{Corresponding number of days}}$$

e. Germination Value (GV) (Ginwal *et al.* 2005):

$$GV = PV \times MDG$$

4.5. Marker development on the *Pinus cembra* transcriptome

4.5.1. Screening the transcriptome data for sequence homologs

A previus pre-analysis was done in Swiss stone pine by Daniela Jahn (PhD student) at the Genetic Department, BFW, Vienna, Austria supervised by Dr. Berthold Heinze. For sampling the material, four sites were selected, which are scattered through the Austrian Alps. RNA was extracted by using MasterPure Plant RNA purification Kit (Invitrogen, Epicentre, USA), using the manufacturer's protocol from two cones in different developmental stages and a needle sample from the same mother tree. Messenger RNA was isolated with Dynabeads mRNA DIRECT Micro Kit (ThermoFisher Scientific, Carlsbad, CA, USA) by protocol, followed by an evaluation according to their RIN value (Bioanalyzer). The mRNA has been used to amplify cDNA libraries, according to the Ion Torrent RNA-Seq protocol (ThermoFisher Scientific, Carlsbad, CA, USA). The analysis was done using an Ion Torrent platform. For each sample a single library was prepared using all reagents and protocols for the Ion Torrent™ Personal Genome Machine™ (PGM) System. The quality and quantity of each preparation step was proofed using a Bioanalyzer Instrument and each library was loaded on a 316 Chip. The reads were quality checked and further analysis was done using the CLC Bio Genomics Workbench. The libraries were analysed separately by doing a *DeNovo* assembly.

A BLAST database has been constructed, downloaded from NCBI database (http://www.ncbi.nlm.nih.gov, 07.12.2016), after the following criteria: at first the search was conducted after the enzymes' and transcription factors' sequences that were annotated in the *Pinus* genus . In case of those enymes' and transcription factors' sequences that couldn't be found as *Pinus* sequences, „Pinaceae" was the following search term and in case of those that were lacking using „Pinaceae", „land plants" was the following term.

The *de novo* assembled contigs were searched using BLASTN and BLASTX toolkits against the database. After the search based on sequence homology, the putative function of the

sequences have been estimated according to the highest BLAST hits by the following thresholds: lowest E-value < 0.001; greatest identity > 98 %; greatest HSP length > 100. From all libraries, only those sequences have been selected, which could be found in both BLASTN and BLASTX MultiBLAST results. For the reason, that the starting material was RNA, the sequence list has been BLAST searched again with *Pinus lambertiana* and *Pinus taeda* genome scaffold, downloaded from theTreeGenes database, to find out how many from the selected sequences that could be identified in the two conifer genomes have been spliced or not in the transcriptome. Following this step, sequences have been selected for primer design (**Figure 13**).

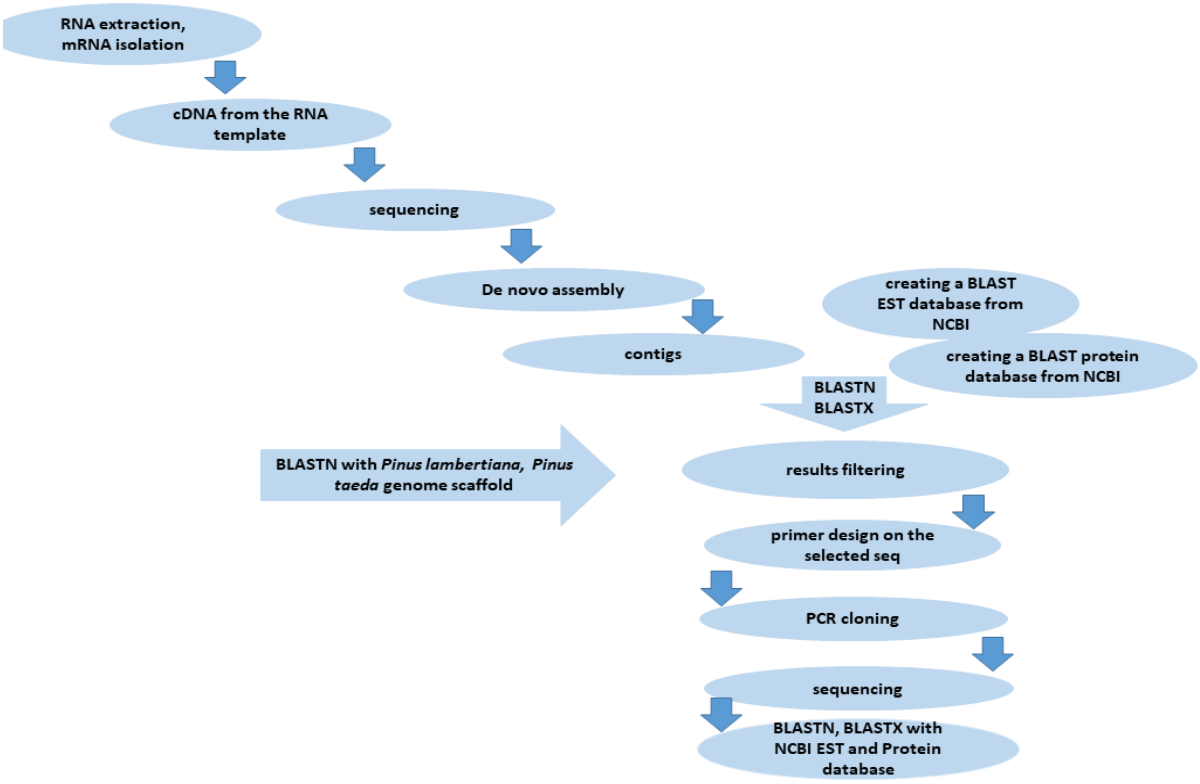

Figure 13. Overview of the annotation process of the *Pinus cembra* transcriptome.

4.5.2. Primer design

Primers have been designed with use of CLC Bio Genomics Workbench's Primer design tool, with lengths between 16 - 28 nt. The lengths of the members of a primer pair were not differed by more than 3 bp, the GC content of the primer pairs were between 40 and 60 %. Primers have been named by the coded protein and the sequence number from each contig, where the annotated sequence has been identified.

The primer-test in the lab was made on 84 Swiss stone pine DNA samples, selected also by random, and as it follows: in a 14 µl volume containing 1 µl of genomic DNA (about 10 ng), 0.14 µl Polim Phite Taq, GeneAmp PCR bufferII (Applied Biosysytems/Roche, Branchburg, N.J.), 3 mM $MgCl_2$, 200 µM of dNTPs, 0.2 mM of each primers, by the following PCR (SPECPCR) protocol: denaturation on 94° C for 3 min (1); 60° C for 1 min (2), 70° C for 1 min (3); 9x up to step 3 (4); 94° C for 30 sec (5); low stringency annealing 55° C for 50 sec (6); 70° C for 2 min (7); 34 cycles up to step 7 (8).

PCR products were analyzed by 2 % agarose gel electrophoresis with 3:1 Biogyn Sieve agarose: normal agarose and 1xTAE as electrophoresis buffer. PCR products have been cleaned by enzymatic PCR cleanup, with Exonuclease 1, Shrimp alkaline Phosphatase (New England Biolabs, USA) and sequenced after preparation by GenomeLab DTCS - Quick Start Kit (Beckman Coulter, Inc., USA) protocol. The results of the sequencing was searched against the nucleotide database of NCBI (default discontiguous megaBLAST settings).

4.6. Test of the developed markers on Scots pine samples

Total DNA was extracted from 20 - 25 mg of plant material of one-year old Scots pine needles by using DNeasy Plant Mini Kit (QIAGEN, Valencia, CA, USA) according to the manufacturer's protocol. For testing the newly developed molecular markers, a preliminary PCR test was carried out to evaluate the functionability of the designed primers.

The primer-test in the lab has been made on one Scots pine DNA sample (Mohos 26), with all primers tested on Swiss stone pine samples, PCR amplification was conducted in a 15 µl volume containing 1 µl of genomic DNA (about 10 ng), 0.14 µl Polim Phite Taq and GeneAmp PCR bufferII (Applied Biosysytems/Roche, Branchburg, N.J.), 3 mM $MgCl_2$, 200 µM of dNTPs, 0.2 mM of each primers, by the following PCR protocol: denaturation (1) on 94° C for 3 min; followed by (2) 94° C for 30 sec; annealing (3) on 55° C for 0.45 sec and extension (4) on 72° C for 1.20 min, (5) 72° C for 10 min; the first 4 cycles being repeated 34 times. PCR products were

analyzed by 1 % agarose gel electrophoresis with 1xTAE as electrophoresis buffer. Those which occured as a single band, were used to detect potential SNP polimorphism in three *Pinus sylvestris* DNA samples, providing from different types of habitat (Mohos – peat bog; Kvacianska dolina – rocky substrate; Őrség – mixed forest.

The post PCR products' purification was effectuated by hydrolyzing the excess primers and dephosphorylated unincorporated dNTP-s, in one step, with CleanSweep PCR purification reagent (ThermoFisher Scientific, Carlsbad, CA, USA), according to the manufacturer's protocol. Purified products were sequenced using the forward primers, in one direction, at Biomi Ltd., Hungary and at the Institute of Genetics, Biological Research Centre of the Hungarian Academy of Sciences, Szeged, Hungary.

4.7. Data analysis

4.7.1. Analysis of the cone morphological and needle anatomical dataset

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. The normality of the residuals was accepted on the basis of their skewness and kurtosis (Tabachnick 2007). Since the assumption of homogeneity of variances was moderately violated according to Levene's test ($p < 0.05$), we separated the significantly different groups by using Games-Howell's post hoc test. We applied discriminant function analysis to predict a categorical dependent variable and determine whether a set of variables was effective in predicting category membership.

We considered previous molecular studies as well as macrofossil and pollen data analyses, about the existence of refugial locations in East - Central Europe, e.g. the Hungarian plain (Bernhardsson *et al.* 2016, Cheddadi *et al.* 2006, Naydenov *et al.* 2005) as well as studies and data published by Magyari *et al.* (2014b), who have contended that the species persisted in Eastern Carpathian refugia. Therefore, we decided to apply discriminant function analysis separately for both morphological and anatomical datasets to detect samples grouping after a previous sorting of

populations according to geographical distribution (the Pannonian Basin, the Northern Carpathians, Central Island Mts., the Eastern Carpathians, and the Southern Carpathians). We also sorted the populations under examination according to the type of their habitat (peatbogs, rock surfaces, and mixed forests) to detect any grouping by traits, which are (or are not) suited to the specific environment. The analysis was performed according to the stepwise method by computing the group sizes within groups. Morphological variation was analyzed with IBM SPSS 20.0 (IBM Corp.) and Microsoft Excel. Mantel test (Mantel 1967) was performed to test the relationship between the geographical and morphological multi-character differences among the populations. Euclidean distances and geographical distances among populations were used for the evaluation using GenAlEx 6.5 (Peakall and Smouse 2012) software.

4.7.2. Statistical analysis of the seed morphological and germination data

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 was 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 (parameter) 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 with classify by computing from group sizes within groups with combined groups plots. We sorted the studied populations according to the type of their habitat (peatbog, rocky surface, and mixed forest) to detect any grouping by traits, which are or not suited to the specific environment. Under mixed forests we mention all forest types grow in tree comunity including beside the pines also broad-leaf species. 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.

4.7.3. Analysis of coding candidate gene dataset

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).

5. RESULTS

5.1. Morphological and anatomical differentiation of peripheral populations from the Carpathian region

5.1.1. Differentiation by geographical distribution

Table 4. Average values with standard deviations of the analyzed Scots pine cone characteristics by geographical distribution.

Abbrev.	Characters	Pannonian basin (HFE, HVE, HZA, HOR)	Northern Carpathians (SKV, SME, STU, SLI)	Central Island Mts. (RBI, RBE, RML, RMH)	Eastern Carpathians (RFE, RPO, RMO)	Southern Carpathians (RPA)
CW	Cone width (cm)	4.14 ± 0.68	3.97 ± 0.48	4.84 ± 0.67	3.97 ± 0.86	4.36 ± 0.81
CL	Cone length (cm)	3.94 ± 0.74	3.38 ± 0.56	4.31 ± 0.75	3.78 ± 1.02	3.84 ± 0.75
NBS	Number of scales (pc.)	72.33 ± 11.65	58.97 ± 8.36	81.50 ± 11.76	67.58 ± 11.15	63.75 ± 9.44
AW	Width of apophysis (mm)	6.71 ± 1.31	6.40 ± 0.95	6.26 ± 0.90	7.23 ± 1.02	6.36 ± 1.24
AL	Length of apophysis (mm)	6.88 ± 1.03	6.77 ± 0.83	6.76 ± 1.00	6.70 ± 1.57	7.19 ± 0.64
AT	Thickness of apophysis (mm)	2.56 ± 0.74	2.54 ± 0.61	3.31 ± 0.85	3.27 ± 1.06	2.68 ± 0.79
CL/CW	Cone length /cone width	0.95 ± 0.11	0.85 ± 0.10	0.89 ± 0.16	0.94 ± 0.13	088 ± 0.06
CL/NBS	Cone length /number of scales	0.05 ± 0.01	0.05 ± 0.00	0.05 ± 0.00	0.05 ± 0.01	0.06 ± 0.01
AL/AW	Apophysis length /apophysis width	1.05 ± 0.20	1.07 ± 0.15	1.09 ± 0.19	0.93 ± 0.21	1.16 ± 0.22
AL/AT	Apophysis length /apophysis thickness	2,.2 ± 0.96	2.72 ± 0.54	2.21 ± 0.87	2.18 ± 0.81	2.87 ± 0.92

One-Way MANOVA test on 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 (AW, AT, CW, and AL/AW, AL/AT), populations from the Northern Carpathians and the Pannonian Basin form one group, while cones from the Eastern Carpathians are separated by four (AW, AT and AL/AT and AL/AW ratios) variables, being in one group with Central Island Mountains by two (AT, AL/AT) variables (**Figure 14**).

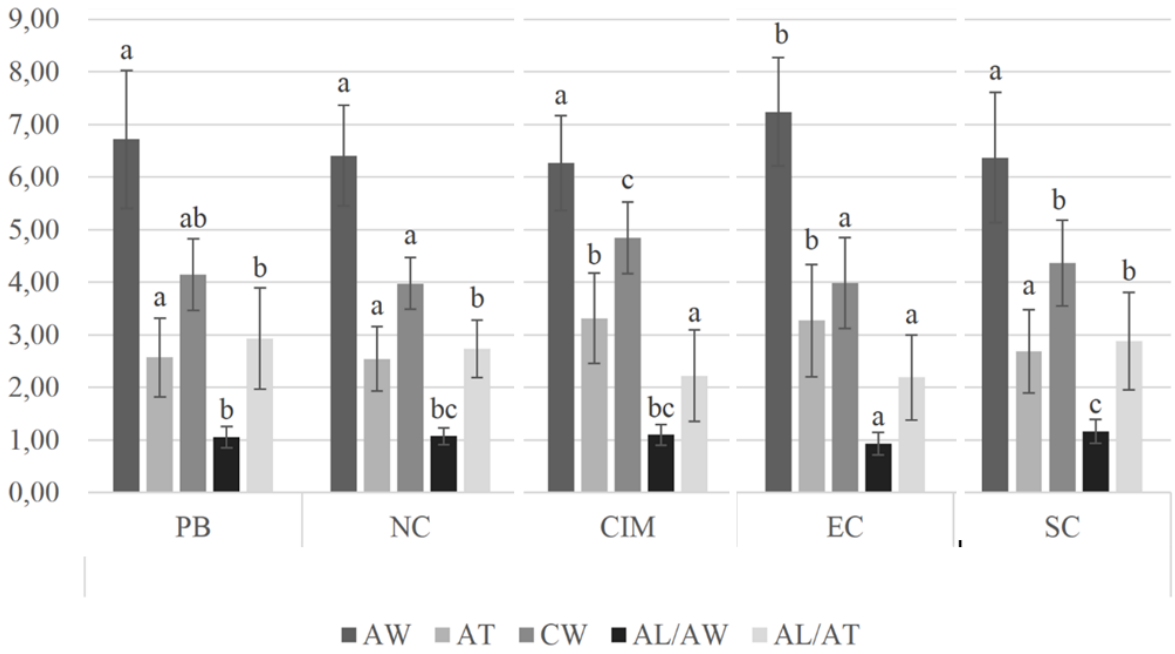

Figure 14. Statistically significant morphometric variables of Scots pine between the studied geographical regions. The PB abbreviation means Pannonian basin, NC: Northern Carpathians, CIM: Central-Island Mountains (Apuseni), EC: Eastern Carpathians and SC: Southern Carpathians respectively. For cone morfological abbreviation see Table 4.

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 15A**). NBS and CW were discriminating variables of Function 1.

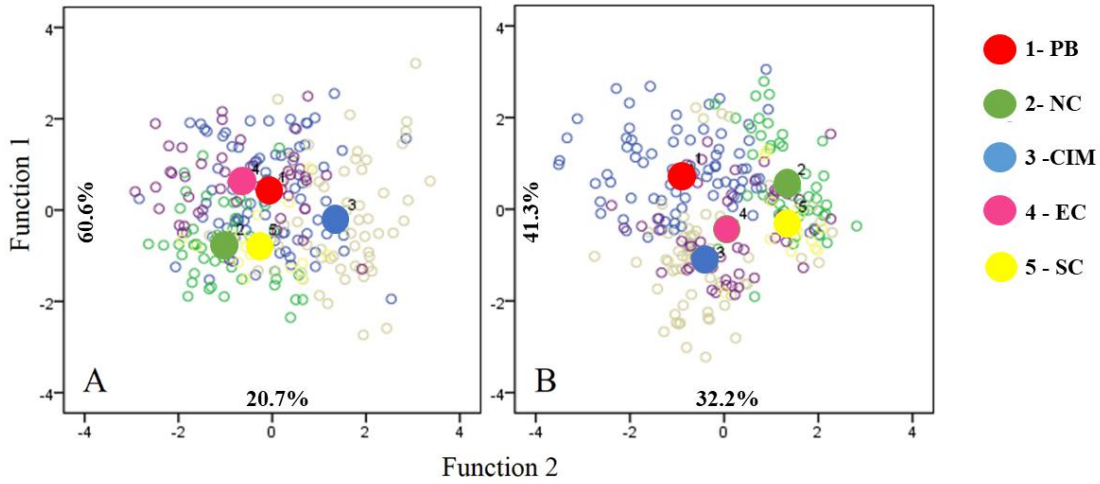

Figure 15. Canonical discriminant function analysis result of *Pinus sylvestris* populations by geographical distribution based on cone morphological characters (A) and needle anatomical variables (B). The PB abbreviation means Pannonian basin, NC: Northern Carpathians, CIM: Central-Island Mountains (Apuseni), EC: Eastern Carpathians and SC: Southern Carpathians respectively.

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) by CL/CW, AW and AL/AW, but the Northern (2) and Southern Carpathians (5) still remained in one group. Furthermore, by considering Function 3 (17 %) it could discriminated the Pannonian Basin (1) from the Eastern Carpathians (4) on the basis of AL and AT variables.

Table 5. Average values of the analyzed needle characteristics of Scots pine populations with standard deviations by geographical distribution.

Code	Characters	Pannonian basin	Northern Carpathians	Central Island Mts.	Eastern Carpathians	Southern Carpathians
		(HFE, HVE, HZA, HOR)	(SKV, SME, STU, SLI)	(RBI, RBE, RML, RMH)	(RFE, RPO, RMO)	(RPA)
NRD	Number of resin ducts (pc.)	12.96 ± 3.13	12.06 ± 2.93	8.98 ± 3.22	10.72 ± 1.44	9.62 ± 2.02
NBRCC	Number of rows on the concave (adaxial) mesophyll (pc.)	2.42 ± 0.49	2.95 ± 0.20	2.15 ± 0.44	2.5 ± 0.54	2.93 ± 0.25
NBRCV	Number of rows on the convex (abaxial) mesophyll (pc.)	2.95 ± 0.61	3.02 ± 0.14	2.78 ± 0.48	3.14 ± 0.68	3.00 ± 0.00
NBRCW	Number of rows on the crosswise mesophyll (pc.)	4.95 ± 0.80	5.14 ± 0.41	4.65 ± 0.64	5.85 ± 0.87	5.00 ± 0.00
NH	Needle height (μm)	703.66 ± 97.70	620.03 ± 86.59	606.15 ± 95.77	613.11 ± 86.63	585.36 ± 59.38
NW	Needle width (μm)	1379.77 ± 220.10	1206.30 ± 169.82	1149.43 ± 193.16	1238.09 ± 184.16	1158.19 ± 110.33
CCH	Central cylinder height (μm)	313.49 ± 48.20	269.69 ± 34.75	282.18 ± 43.75	294.94 ± 37.11	274.60 ± 26.79
CCW	Central cylinder width (μm)	911.11 ± 166.21	724.35 ± 12413	688.90 ± 143.26	764.06 ± 141.10	676.05 ± 95.11
CCW/CCH	Central cylinder width/central cylinder height	2.92 ± 0.44	2.69 ± 0.33	2.46 ± 0.46	2.60 ± 0.46	2.47± 0.36
NW/NH	Needle width/needle height	1.96 ± 0.15	1.95 ± 0.16	1.90 ± 0.25	2.02 ± 0.20	1.98 ± 0.13
CCH/NH	Central cylinder height/needle height	0.44 ± 0.04	0.43 ± 0.03	0.46 ± 0.04	0.48 ± 0.04	0.47 ± 0.03
CCW/NW	Central cylinder width/needle width	0.65 ± 0.03	0.59 ± 0.04	0.59 ± 0.04	0.61 ± 0.05	0.58 ± 0.03

Performing the MANOVA test on needle anatomical traits, significant differences based on four variables could be found: CCH, CCW, NH, NW and three ratios: CCW/CCH, CCH/NH and CCW/NW ($F(16;758) = 10.35$; $p < 0.001$; $F(126;659) = 12.62$; $p < 0.001$) 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: CCH, CCW, NH, and NW and three ratios: CCW/CCH, CCH/NH, and CCW/NW (**Figure 16**).

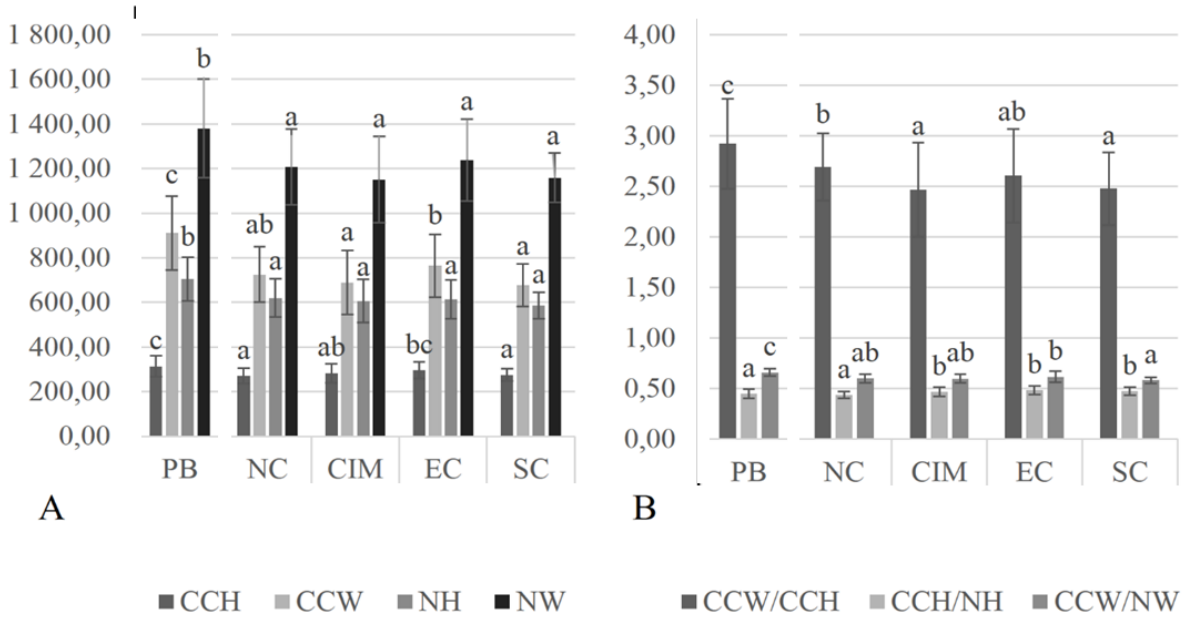

Figure 16. Statistically significant anatomical variables of Scots pine needles between the studied geographical regions. The PB abbreviation means Pannonian basin, NC: Northern Carpathians, CIM: Central-Island Mountains (Apuseni), EC: Eastern Carpathians and SC: Southern Carpathians respectively. For needle anatomical abbreviation see Table 5.

By carrying out discriminant function analysis by the first variable (Function 1), which was responsible for 41.3 % of the variation, populations formed two distinctive groups (**Figure 15B**): the Pannonian Basin (1), the Central Island Mts. (3), the Eastern Carpathian (4) vs. Northern (4) and Southern Carpathians (5). NBRCC and CCH were discriminating variables. 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 (4). Discriminating variables were NRD, CCW, NW, and NH. Additionally, Function 3 (24 %) indicated a difference between the populations in the Central Island Mts. (3) and the populations in the Eastern Carpathians (4) according to NBRCW and NBRCV. The Mantel correlation test was not significant ($R^2 = 0.017$, $p < 0.05$) (**Figure 17**).

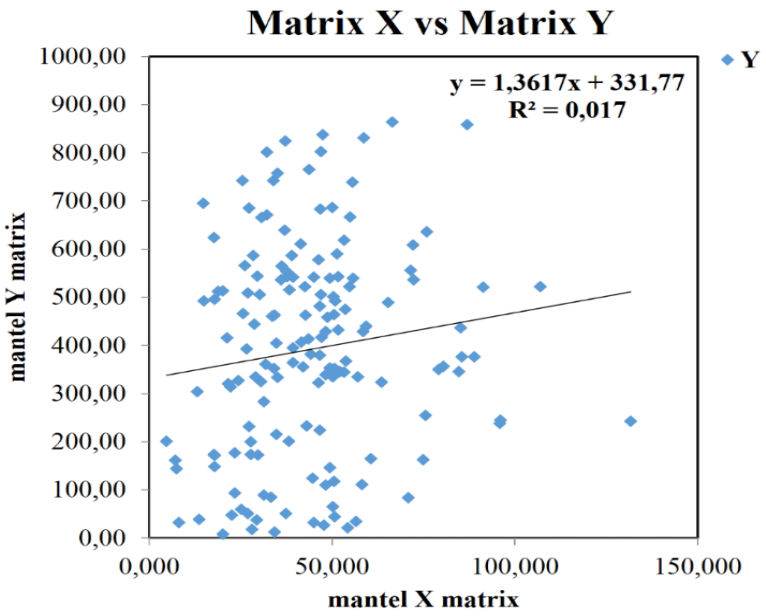

Figure 17. Relationship between pairwise Euclidean distance and geographic distances (r_{xy} = 0.130, p = 0.130, R^2 = 0.017) for the 16 *Pinus sylvestris* populations.

5.1.2. Differentiation by habitat type

Table 6. Average values of the analyzed needle characteristics in Scots pine samples with standard deviations by the habitat type.

Abbrev.	Characters	Peat bogs (RMO, RFE, RPO, RML, RMH)	Rock surfaces (HKO, RBE, RBI, RPA, SLI, SKV, STU)	Mixed forests (HFE, HZA, HVE, HOR)
NRD	Number of resin ducts (pc.)	10.43 ± 2.15	10.29 ± 3.20	13.62 ± 3.13
NBRCC	Number of rows on the concave (adaxial) mesophyll (pc.)	2.40 ± 0.51	2.67 ± 0.50	2.32 ± 0.47
NBRCV	Number of rows on the convex (abaxial) mesophyll (pc.)	3.02 ± 0.63	2.93 ± 0.30	2.92 ± 0.67
NBRCW	Number of rows on the crosswise mesophyll (pc.)	5.45 ± 0.91	4.97 ± 0.59	4.82 ± 0.82
NH	Needle height (µm)	630.06 ± 92.73	592.39 ± 75.79	732.66 ± 85.53
NW	Needle width (µm)	1220.30 ± 196.57	1154.40 ± 155.55	1451.40 ± 182.98
CCH	Central cylinder height (µm)	293.80 ± 39.48	272.41 ± 34.47	322.21 ± 49.11
CCW	Central cylinder width (µm)	745.94 ± 151.61	698.25 ± 113.10	958.74 ± 148.50
CCW/CCH	Central cylinder width/central cylinder height	2.54 ± 0.42	2.58 ± 0.42	3.00 ± 0.43
NW/NH	Needle width/needle height	1.94 ± 0.21	1.95 ± 0.20	1.98 ± 0.13
CCH/NH	Central cylinder height/needle height	0,. 6 ± 0.04	0.46 ± 0.04	0.44 ± 0.04
CCW/NW	Central cylinder width/needle width	0.60 ± 0.05	0.60 ± 0.04	0.65 ± 0.03

Table 7. Average values of the analyzed cone characteristics in Scots pine samples with standard deviations by the habitat type.

Abbrev.	Characters	Peat bogs	Rock surfaces	Mixed forests
		(RMO, RFE, RPO, RML, RMH)	(HKO, RBE, RBI, RPA, SLI, SKV, STU)	(HFE, HZA, HVE, HOR)
CW	Cone width (cm)	4.31 ± 0.85	4.31 ± 0.74	4.14 ± 0.71
CL	Cone length (cm)	4.06 ± 0.95	3.73 ± 0.74	3.95 ± 0.77
NBS	Number of scales (pc.)	74.90 ± 13.83	66.72 ± 12.98	72.39 ± 11.43
AW	Width of apophysis (mm)	6.90 ± 1.10	6.25 ± 0.91	6.90 ± 1.36
AL	Length of apophysis (mm)	6.56 ± 1.30	6.90 ± 0.92	7.00 ± 1.04
AT	Thickness of apophysis (mm)	3.46 ± 1.01	2.69 ± 0.62	2.50 ± 0.79
CL/CW	Cone length/cone width	0.94 ± 0.16	0.86 ± 0.10	0.95 ± 0.12
CL/NBS	Cone length/number of scales	0.05 ± 0.01	0.05 ± 0.00	0.05 ± 0.01
AL/AW	Apophysis length/apophysis width	0.96 ± 0.21	1.12 ± 0.17	1.04 ± 0.21
AL/AT	Apophysis length/apophysis thickness	2.01 ± 0.69	2.67 ± 0.75	3.08 ± 0.99

A statistical analysis with three pre-formed groups according to habitat type was performed. By carrying out a one-way MANOVA test using the morphological cone dataset (AL, AT, AW, CL/CW, NBS), 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 peat bog provenience differed significantly from those of rocky outcrops and mixed forests in the case of two variables: the length and thickness of apophysis (AL, AT). Samples originating from dry rocky outcrops were significantly separated according to three variables: AW, NBS, and CL/CW (**Figure 18A**).

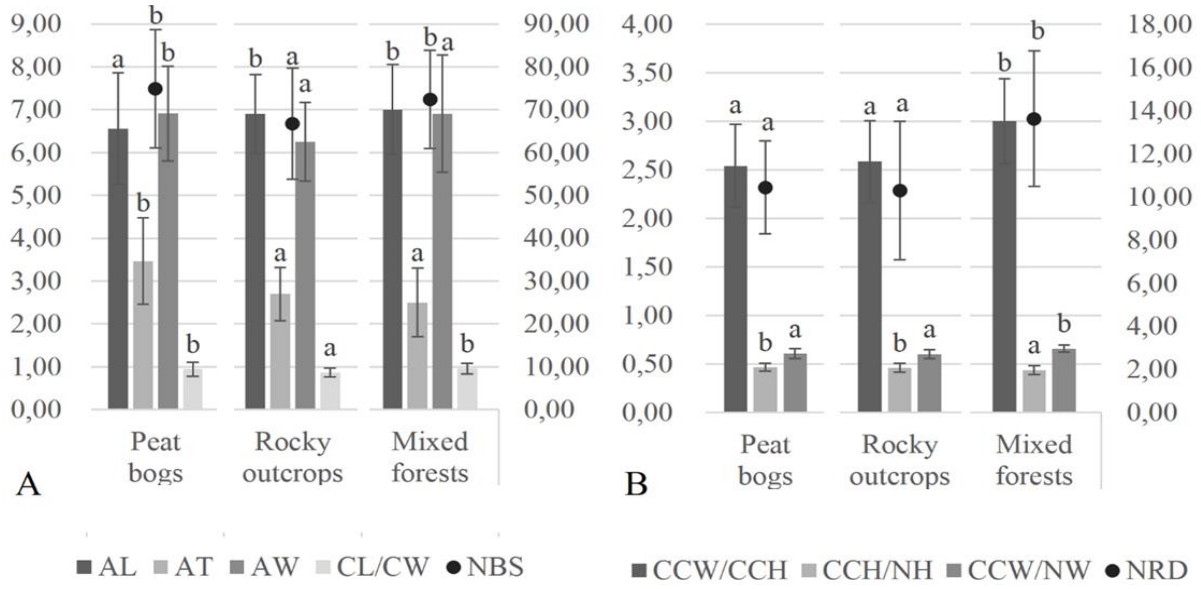

Figure 18. Statistically significant morphometric (A) and anatomical (B) variables of Scots pine individuals within the studied habitat types (peat bogs, rocky outcrops and mixed forests). For morphoanatomical abbreviation see Table 6,7.

By performing one-way MANOVA test with ratios (CCW/CCH, CCH/NH, CCW/NW) and with NRD, significant differences were detected again ($F(8;500) = 12.64$; $p < 0.001$) with significant between-subjects effects ($F(2;253) > 8.00$; $p < 0.001$). Significantly less resin ducts (NRD) among populations from peat bogs and on rocky surfaces have been found. In the case of three proportions (CCW/CCH, CCH/NH and CCW/NW), populations with mixed forest origin differed significantly by high values at CCW/NW and CCW/CCH and lower values at CCH/NH ratios (**Figure 18B**). Therefore, these anatomical traits seem to be very useful in describing differentiation among habitats.

Discrimination analysis based on the measured 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 the separation into two groups: mixed forests (3) with rocky surfaces (2) vs. peat bogs (1). The highest level of differences between the populations was defined by the following variables; for function 1: AT, AL/AT, AL/AW, NBS, AW, and for function 2, which separated mixed forests (3) from rocky surfaces (2): CL/CW (**Figure 19A**).

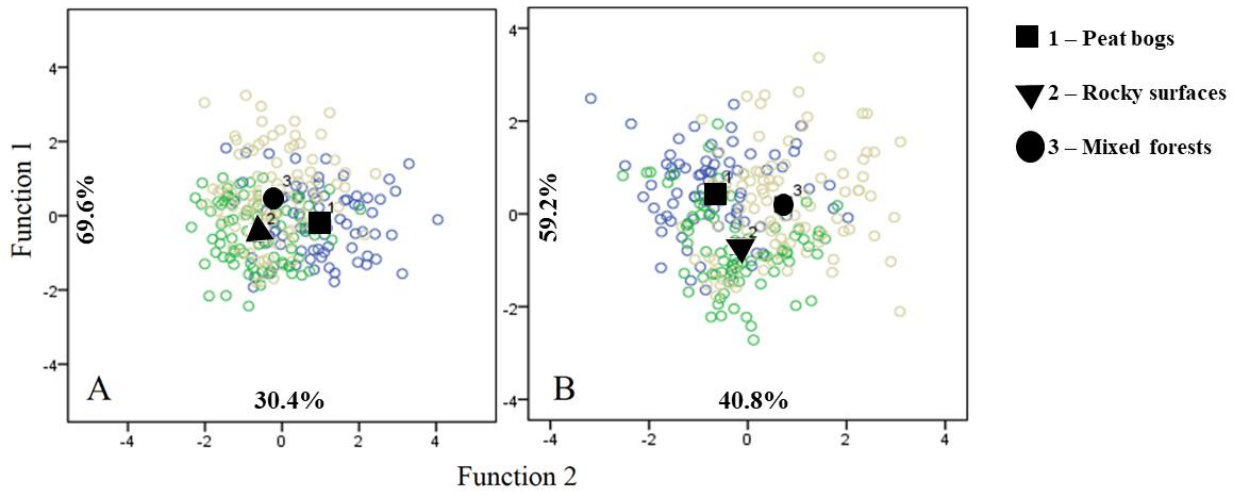

Figure 19. Canonical discriminant function analysis result of *Pinus sylvestris* populations by habitat type detected on morphological cone characters (A) and on needle anatomic variables (B): 1- peat bogs (RMO, RFE, RPO, RML, RMH), 2- rocky surfaces (HKO, RBE, RBI, RPA, SLI, SKV, STU), 3- mixed forests (HFE, HZA, HVE, HOR). For population abbreviation see Table 2.

Discriminant function 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 19B**). Differences among populations were defined by three variables represented by Function 1: central cylinder width (CCW), number of resin ducts (NRD), and needle width (NW). Populations from mixed forests were discriminated by Function 2 with five variables: central cylinder height (CCH), needle height (NH), number of rows on the crosswise, and concave and convex mesophyll (NBRCW, NBRCV, NBRCC) (**Figure 19B**).

5.2. Analysis of the habitat type differentiation in peripheral populations based on seed traits and germination data

5.2.1. Seed morphological characters

Nine morphological traits and ratios with the average values and standard deviations are summarized in **Table 8**.

Table 8. Average values with standard deviations of the analyzed Scots pine seed characteristics.

Code	Characters	RFE (peat bog)	SME (peat bog)	RML (peat bog)	RMO (peat bog)	RPS (peat bog)	RBE (rock surface)	RCO (rock surface)	CHR (rock surface)	HZA (mixed forest)	RVR (mixed forest)
SL	Seed length (mm)	3.80 ± 0.60	3.66 ± 0.50	3.90 ± 0.42	4.23 ± 0.48	4.06 ± 0.60	4.82 ± 0.52	4.20 ± 0.56	4.18 ± 0.78	4.30 ± 0.52	4.40 ± 0.50
SW	Seed width (mm)	2.10 ± 0.26	2.23 ± 0.27	2.40 ± 0.32	2.50 ± 0.29	2.44 ± 0.34	2.84 ± 0.45	2.46 ± 0.45	2.45 ± 0.35	2.53 ± 0.38	2.80 ± 0.60
WL	Wing length (mm)	13.23± 2.26	10.00± 1.18	15.07± 1.61	14.13± 1.61	13.14± 2.09	15.56± 1.59	13.72± 1.61	12.38± 1.66	14.51 ± 1.22	13.04± 1.83
WW	Wing width (mm)	4.65 ± 0.74	4.35 ± 0.49	5.10 ± 0.66	4.51 ± 0.65	4.57 ± 0.80	5.89 ± 0.78	4.78 ± 0.53	4.81 ± 0.87	5.31 ± 0.66	5.17 ± 0.85
SL/SW	Seed length/ seedwidth	1.83 ± 0.33	1.65 ± 0.28	1.64 ± 0.23	1.70 ± 0.20	1.67 ± 0.25	1.72 ± 0.21	1.74 ± 0.28	1.71 ± 0.26	1.72 ± 0.22	1.61 ± 0.27
WL/WW	Wing length/ wing width	2.89 ± 0.54	2.30 ± 0.23	2.99 ± 0.41	3.16 ± 0.36	2.94 ± 0.57	2.68 ± 0.38	2.89 ± 0.37	2.63 ± 0.49	2.76 ± 0.35	2.60 ± 0.59
SL/WL	Seed length/ wing length	1.84 ± 0.33	0.66 ± 0.28	1.65 ± 0.23	1.70 ± 0.20	1.67 ± 0.25	1.72 ± 0.21	1.74 ± 0.28	1.71 ± 0.27	1.72 ± 0.22	1.61 ± 0.27
SW/WW	Seed width/ wing width	0.46 ± 0.10	0.52 ± 0.08	0.48 ± 0.08	0.56 ± 0.10	0.55 ± 0.11	0.49 ± 0.10	0.52 ± 0.08	0.52 ± 0.08	0.49 ± 0.11	0.55 ± 0.11
SWE	Seed weight (g/ten seed)	0.008	0.010	0.011	0.009	0.008	0.007	0.010	0.010	0.010	0.008

By performing One-Way ANOVA test, seed length (SL), seed width (SW), wing length (WL), wing width (WW) and seed weight (SWE) (**Figure 20**) showed significantly higher values in case of RBE population, samples from SME had significantly lower values considering wing length (WL), wing length/wing width (WL/WW) (**Figure 20B,D**) and significantly higher in case

of seed length/wing length (SL/WL) ratios (**Figure 20E**). A distinct separation could be also observed in one group of SME, RMO, RPS populations, all from peat bogs, defined by low values of wing width (WW) variable (**Figure 20B**). Samples from RML showed significantly lower values at seed weight (SWE), seed length/seed width (SL/SW), seed length/wing length (SL/WL), seed width/wing width (SW/WW) traits (**Figure 20C-E**).

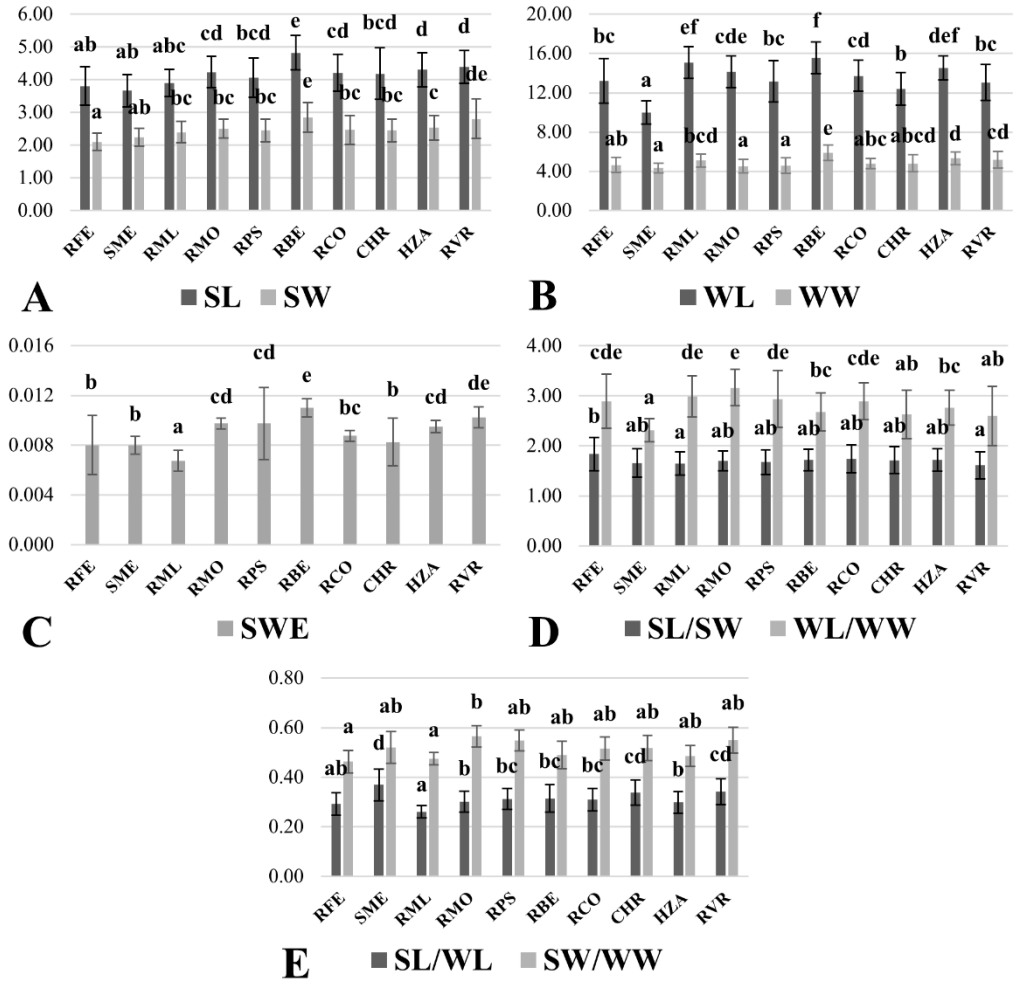

Figure 20. Statistically significant seed morphological variables of Scots pine among the sampled populations. Unit of the measurement is mm (SL, SW, WL, WW) and g (SWE), except SL/SW, WL / WW, SL/WL and SW/WW. For seed morphological abbreviation see Table 8.

Before carrying out the discriminant function analysis using Ward Linkage, three groups were formed according to habitat type: peat bog, rocky substrate and mixed forest. Based on the measured morphological datasets, 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 the separation into two groups: mixed forest (3) with rocky substrate (2) vs. peat bog (1). The highest level of differences between the populations was defined for function 1 by seed length (SL), wing width (WW) and seed weight (SWE) (**Figure 21**).

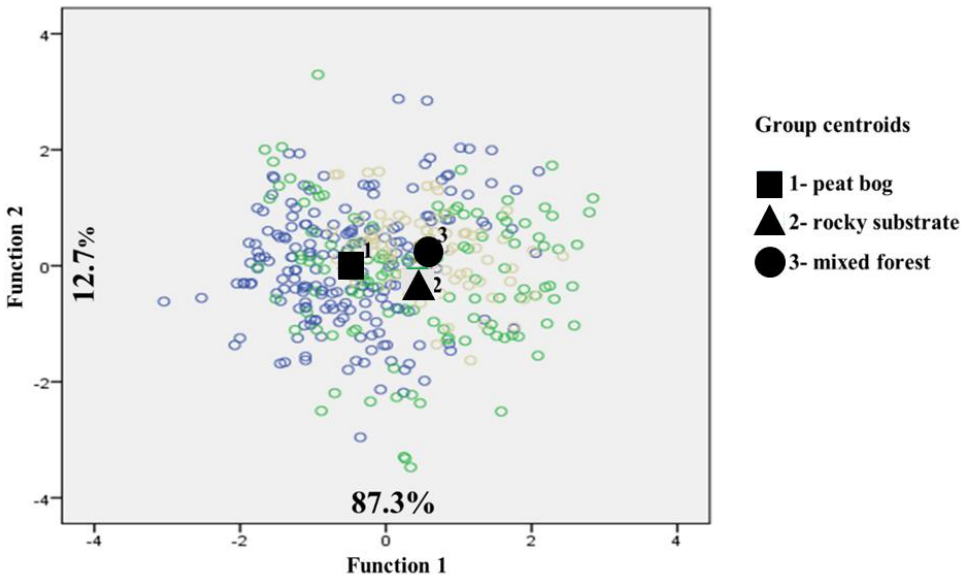

Figure 21. Differentiation of *Pinus sylvestris* L. populations by habitat type detected by discriminant function analysis, on the basis of morphological seed characteristics: 1 – peat bogs (RFE, SME, RML, RMO, RPO), 2 – rocky substrates (RBE, RCO, CHR), 3 – mixed forests (HZA, RVR). For population abbreviation see Table 3.

The hierarchical cluster tree using Ward's linkage is presented in **Figure 22**. Within the dendrogram, samples from RMO and RPS with similar type of habitat (oligotrophic peat bog) were forming one subcluster, also the RCO and CHR samples (rocky substrate), with a relatively well supported relationship with HZA (mixed forest) samples. Another subcluster comprises RBE and RVR (rocky substrate and mixed forest) samples and a special highlight is needed on SME population, characterised by a completely distant position.

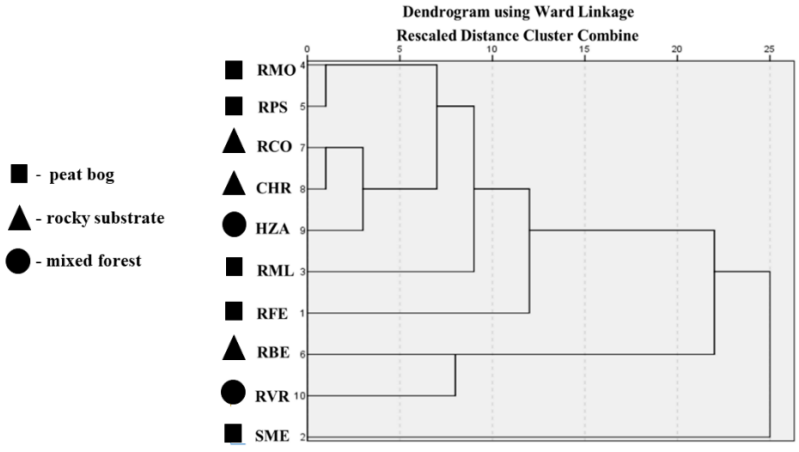

Figure 22. Dendrogram generated with IBM SPSS 20.0 using Ward Linkage among 10 *Pinus sylvestris* L. populations from different habitats (RMO, RPS, RML, RFE, SME – peat bog; RCO, RBE, CHR – rocky substrate; HZA, RVR – mixed forest). For population abbreviation see Table 3.

By performing Bivariate Correlation-analysis to detect relationship between each variable and defining correlation as significant at the 0.01 and 0.05 level, the following results have been obtained, presented in **Table 9**.

Table 9. Correlation between each morphological seed variable in the studied Scots pine populations (**significant at level 0,01; *significant at level 0.05).

	SL	SW	WL	WW	SL/SW	WL/WW	SL/WL	SW/WW
SL		**0.496****	**0.467****	**0.304****	**0.428****	**0.171****	**0.415****	**0.188****
SW	**0.496****		**0.245****	**0.360****	**-0.532****	-0.102*	**0.172****	**0.547****
WL	**0.467****	**0.245****		**0.470****	**0.176****	**0.494****	**-0.589****	**-0.190****
WW	**0.304****	**0.360****	**0.470****		-0.096	**-0.512****	**0.208****	**0.556****
SL/SW	**0.428****	**-0.532****	**0.176****	-0.096		**0.273****	**0.232****	**-0.365****
WL/WW	**0.171****	-0.102*	**0.494****	**-0.512****	**0.273****		**-0.347****	**0.394****
SL/WL	**0.415****	**0.172****	**-0.589****	**-0.208****	**0.232****	**-0.347****		**0.346****
SW/WW	**0.188****	**0.547****	**-0.190****	**-0.556****	**-0.365****	**0.394****	**0.346****	

The outcome of Correlation analysis showed strong correlation between most of the variables, with two exceptions: between wing length/wing width - seed width (WL/WW – SW) and seed length/seed width – wing width (SL/SW – WW) variables. Mantel test was carried out for the study of relationship between seed morphological divergence and geographic distance.

The Mantel correlation was equal to r_{xy} = -0.100. p = 0.410 (R^2 = 0.01). Our results are statistically significant at an alpha (p) of 0.05. The scatterplot between elements in the two matrices showed no linear relationship (**Figure 23**).

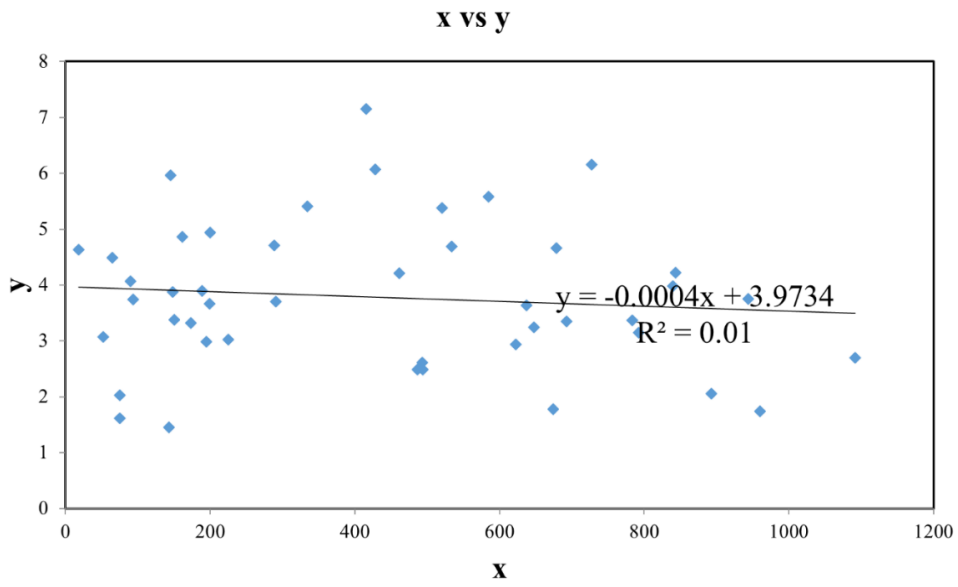

Figure 23. Relationship between pairwise Euclidean distance and geographic distances (r_{xy} = 0.100. p = 0.410. R^2 = 0.01) for the 10 *Pinus sylvestris* populations.

5.2.2. Germination associated parameters

The germination associated parameters for the studied Scots pine populations are presented in **Table 10** and **Figure 24**.

Table 10. Germination associated parameters for the 10 *Pinus sylvestris* populations. For population abbreviation see Table 3.

Pop abbrev.	RFE	SME	RML	RMO	RPS	RBE	RCO	CHR	HZA	RVR
GV (Germination value)	0.05	0.02	0.02	0.00	0.02	0.02	0.01	0.04	0.02	0.01
PG (% of germination)	76.67	41.67	50.00	43.33	36.67	28.33	31.67	75.00	51.67	23.33
GS (Germination speed)	13.50	2.62	6.29	5.96	3.18	5.11	3.83	12.74	8.35	3.78
MGT (Mean germination time)	138.41	119.77	139.88	149.71	377.94	220.35	197.65	381.29	134.71	242.35
MDG (Mean daily germination)	0.19	0.12	0.14	0.16	0.38	0.25	0.03	0.39	0.21	0.26
PV (Peak Value)	0.10	0.06	0.14	0.06	0.12	0.08	0.10	0.12	0.08	0.08

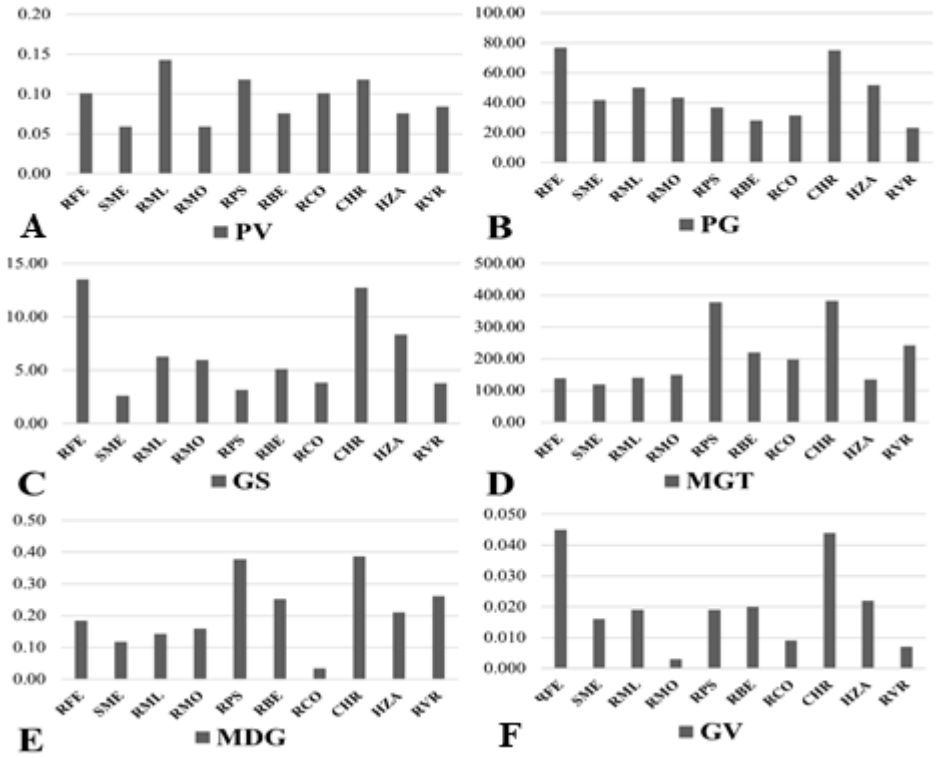

Figure 24. Parameters associated with germination for the studied Scots pine populations (A – Peak Value (PV), B – Germination % (PG), C – Speed of germination (GS), D – Mean germination time (MGT), E – Mean daily germination (MDG), F – Germination Value (GV)). For germination parameters' abbreviation see Table 10, for population abbreviation see Table 3.

Maximum Germination percentage (PG) (**Figure 24B**) has been recorded at RFE (76.67) and CHR (75.00) populations. The other values declined up to the lowest value of 23.33, the Germination percentage (PG) in case of RVR. Similar to seed germination percent, at the Speed of germination (GS) (**Figure 24C**) the highest values were also recorded in case of the RFE and CHR populations (13.50 and 12.74), the minimum value being recorded at SME samples (2.62). The Mean Germination Time (MGT) (**Figure 24D**) also reached its maximum levels in case of CHR and RPS (381.29 and 377.94), declining up to the lowest level in case of SME (119.76). Mean daily germination (MDG) (**Figure 24E**) has its maximum values (0.39 and 0.38) at the same populations (CHR and RPS), but with the lowest registered data in case of RCO (0.03). In case of Peak Value (PV) (**Figure 24A**), parameters varied from the minimum of 0.06 (SME and RMO) to the maximum of 0.14 in case of RML population. Germination Value (GV) (**Figure 24F**) was recorded maximum in case of RFE (0.05) and minimum at RCO and RVR (0.01) samples.

By carrying out the discriminant function analysis using Ward Linkage, with pre-formed three groups according to habitat type (peat bog, rocky substrate and mixed forest), but with germination associate 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 forest (3) with rocky substrate (2) vs. peat bog (1), but with

a more stronger pattern of differentiation between these two groups. The highest level of differences between the populations was defined by the following variables; for function 1: PG (Germination %), SL (seed length) and WL (wing length) and for function 2 by GS (Speed of germination), SWE (seed weight), SW (seed width), WW (wing width) and GV (Germination Value) (**Figure 25**).

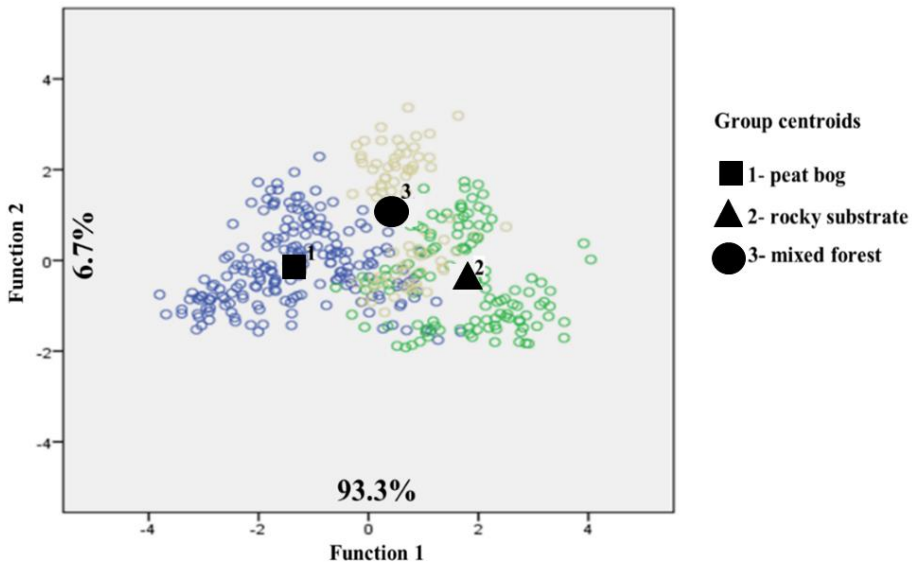

Figure 25. Differentiation of *Pinus sylvestris* populations by habitat type detected by discriminant function analysis, on the basis of morphological seed characteristics and germination associated parameters:1 – peat bogs (RFE, SME, RML, RMO, RPS), 2 – rocky substrate (RBE, RCO, CHR), 3 – mixed forest (HZA, RVR). For population abbreviation see Table 3.

By performing Bivariate Correlation-analysis to detect relationship between seed morphology and germination associated data and defining correlation as significant at the 0.01 and 0.05 level, significant correlation at 0.05 level between GS with SL/WL and SW/WW likewise between PG with SL, SW and SWE variables have been recorded (**Table 11**).

Table 11. Correlation between seed morphology and germination associated variables in the studied 10 Scots pine populations (**significant at level 0,01; *significant at level 0.05).

Variable	SL	SW	WL	WW	SWE	SL/SW	WL/WW	SL/WL	SW/WW	GV	GS	PG
SL		**0.496****	**0.467****	**0.304****	**0.473****	**0.428****	**0.172****	**0.413****	**0.188****	0.035	0.086	**0.178****
SW	**0.496****		**0.245****	**0.360****	**0.353****	**-0.532****	-0.101*	**0.173****	**0.548****	0.037	0.116*	**0.151****
WL	**0.467****	**0.245****		**0.470****	**0.303****	**0.177****	**0.495****	**-0.590****	**-0.191****	-0.027	-0.109*	0.001
WW	**0.304****	**0.360****	**0.470****		**0.275****	-0.097	**-0.512****	**-0.208****	**-0.555****	0.011	-0.028	0.059
SWE	**0.473****	**0.353****	**0.303****	**0.275****		0.068	0.039	0.106*	0.090	0.026	0.094	**0.224****
SL/SW	**0.428****	**-0.532****	**0.177****	-0.097	0.068		**0.274****	**0.228****	**-0.365****	-0.023	-0.059	-0.004
WL/WW	**0.172****	-0.101*	**0.495****	**-0.512****	0.039	**0.274****		**-0.348****	**0.391****	0.000	-0.044	-0.015
SL/WL	**0.413****	**0.173****	**-0.590****	**-0.208****	0.106*	**0.228****	**-0.348****		**0.348****	0.026	**0.152****	0.108*
SW/WW	**0.188****	**0.548****	**-0.191****	**-0.555****	0.090	**-0.365****	**0.391****	**0.348****		0.043	**0.135****	0.096
GV	0.035	0.037	-0.027	0.011	0.026	-0.023	0.000	0.026	0.043		**0.866****	**0.867****
GS	0.086	0.116*	-0.109*	-0.028	0.094	-0.059	-0.044	**0.152****	**0.135****	**0.866****		**0.917****
PG	**0.178****	**0.151****	0.001	0.059	**0.224****	-0.004	-0.015	0.108*	0.096	**0.867****	**0.917****	

5.3. Results of the marker development and the cross species transfer test

Candidate genes were chosen based on literature survey suggesting an impact of the genes on the adaptation process. **Table 1** presents the results of the data mining. The *de novo* assembled *Pinus cembra* contigs were searched using BLASTN and BLASTX toolkits against a database based on this data mining, constructed with aim to identify to date unknown genes involved in the conifers adaptation process to biotic and abiotic factors The constructed BLAST database is presented in **Table 12**.

Table 12. BLAST database with the downloaded protein and EST sequences resulted by using the search terms „Pinus", „Pinaceae" and in some cases „land plants" involved in plant's adaptive processes.

ABBREV.	NAME	PROTEIN SEQ	EST SEQ
PAL	Phenylalanine ammonia lyase	790	9
CHS	Chalcone synthase	277	25
CHI	Chalcone isomerase	14	260
F3H	Flavanone 3-hydroxylase	43	110
F3'H	Flavanoid 3'-hydroxylase	56	2
F3'5'H	Flavonoid 3'5'-hydroxylase	235	6
DFR	Dihydroflavonol reductase	26	7
LDOX	Leucoanthocyanidin dioxygenase	15	201
LAR	Leucoanthocyanidin reductase	41	34
ANS	Anthocyanidin synthase	363	2
ANR	Anthocyanidin reductase	26	1
3GT	Flavonoid 3-glucosyltransferase	3	41
AAT	Anthocyanin acyltransferase	96	48
5GT	Anthocyanin 5-o-glucosyltransferase	31	8
AMT	Anthocyanin methyltransferase	32	19
3RT	Rhamnosyltransferase	26	2
IGSTP	Glutathione S-transferase	365	31
UFGT	Flavonoid-3-O-glucosyltransferase	3	2
ALT	Alanine aminotransferase	8	17
AST	Aspartate aminotransferase	36	8
NAD+	NAD malate dehydrogenase	24	29
AroE	Shikimate dehydrogenase	22	65
GPI	Glucose 6 phosphate isomerase	24	18
6PGD	6 phosphogluconate dehydrog.	15	29
RuBisCo	Ribulose-1,5-bisphosph. carboxyl./oxyg.	8	65
diTP	Diterpene synthase	15	58
SUS	Sucrose synthase	29	34
SOD	Superoxide dismutase	26	32
APX	Ascorbate peroxidase	50	67
PFK	Phosphofructokinase	49	23
AcI	Acid invertase	36	76
GR	Glutathione reductase	8	73
MADS	MADS-box	11	3
WRKY	WRKY-transcription factor	555	163
Myb	Myb-transcription factor	9	252
Total		**5187**	

Based on these data, after BLASTN and BLASTX search with all samples' libraries, from the MultiBLAST results 399 sequences could be selected, annotated and renamed by the coded protein. Primers could be designed on 164 sequences (**Supplemental Materials Table S2**). By testing the markers on *Pinus sylvestris* DNA samples, specific PCR products have been obtained in case of 53 markers. Those with PCR products that occurred as a single band in the electrophoresis gel, have been obtained in case of 25 (**Figure 26**).

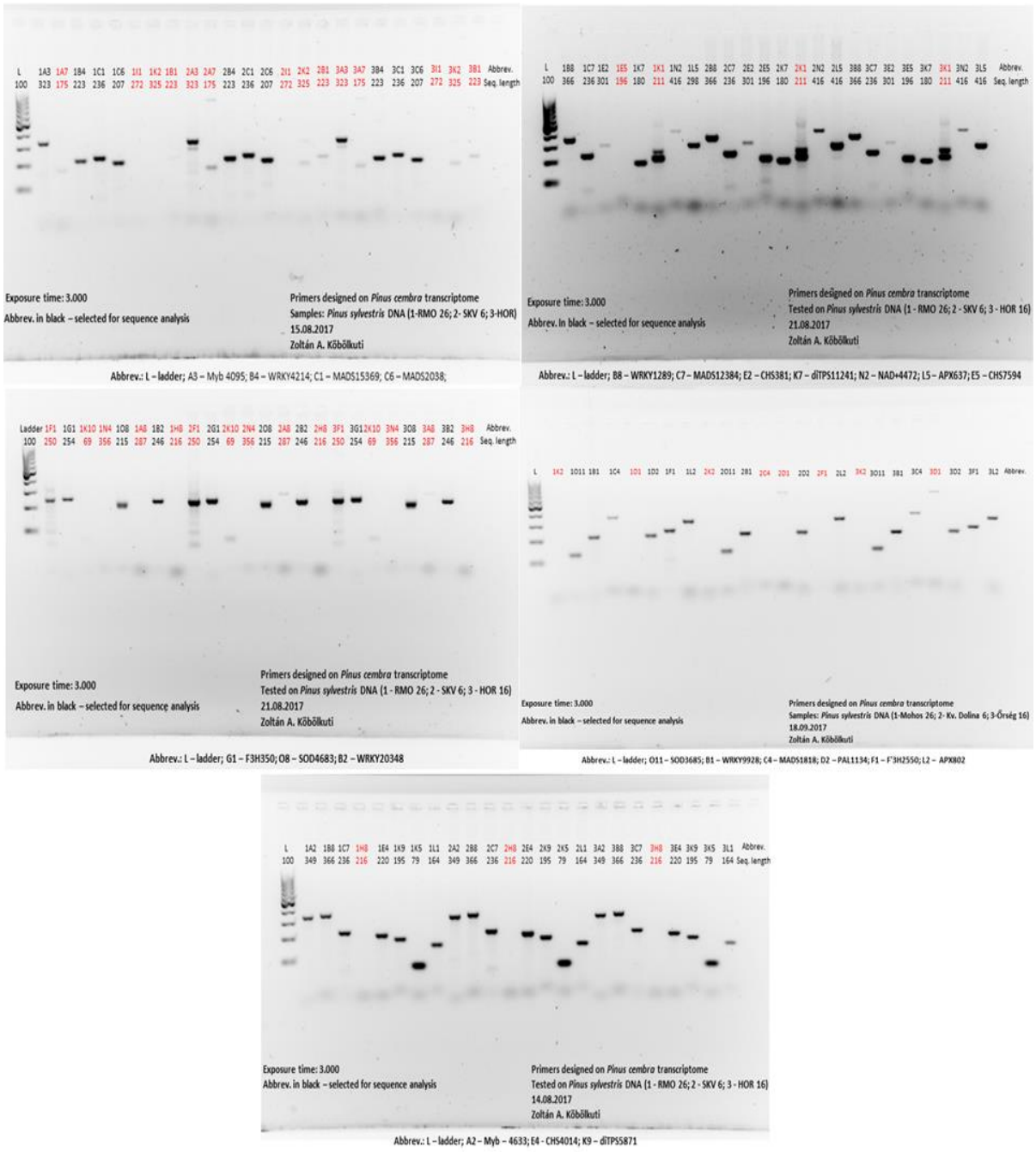

Figure 26. The PCR products of the selected primers analyzed on 1 % agarose gels. Sequences abbreviated in black were selected for sequence analysis. The number under each abbreviation indicates the length of the sequence. (photos: Z.A. Köbölkuti)

5.4. Analysis of newly developed molecular markers

Fragments from 25 different genes were successfully amplified, each of them on three DNA samples from three types of habitat: peat bog (RMO), mixed forest (HZA) and rocky substrate (SKV). After sequencing, conclusive results on all three samples' amplified products have been obtained in case of 20 sequences. The rest of five (NAD+4472, F'3H2550, CHS7594, CHS381, MADS1818) had convincing chromatograms – each of them – only on two samples' PCR product.

All fragments were verified at first using BLASTN search. The putative similarity of the sequences was estimated according to the best BLAST hit. In total, 5223 bp were analysed. After the homology based search, in case of three previously annotated sequences (APX8272, APX637, SOD4683) have been found no significant similarity, these sequences being excluded from further analyses (**Table 13**).

Table 13. BLASTN search results against the NCBI database of the 25 candidate genes. As result are presented only the best BLAST hits' accession numbers of the sequences in the NCBI database.

Query sequence	Max score	Total score	Query cover	E-value	Ident	Accession number in NCBI
Myb 4633	459	459	99 %	3e-125	95 %	BT070763.1
Myb 4095	337	337	72 %	1e-88	96 %	FJ125922.1
WRKY9928	278	278	85 %	4e-71	100 %	EU393816.1
WRKY20348	305	305	100 %	2e-79	94 %	EF676996.1
WRKY4214	281	281	85 %	3e-72	100 %	EU393816.1
WRKY1289	510	510	100 %	8e-141	97 %	BT070825.1
MADS15369	241	228	100 %	4e-56	92 %	BT119355.1
MADS1818	459	459	99 %	3e-125	93 %	BT104866.1
MADS2038	215	215	100 %	3e-52	94 %	BT117271.1
MADS12384	251	261	97 %	5e-66	92 %	BT119355.1
PAL1134	294	294	100 %	4e-76	97 %	AY321089.1
CHS381	298	298	100 %	4e-77	91 %	JN400054.1
CHS4014	237	237	100 %	6e-59	97 %	KJ796482.1
CHS7594	206	206	99 %	2e-49	94%	AJ002156.1
F'3H2550	350	350	100 %	1e-92	97 %	KF704818.1
F3H350	363	363	100 %	1e-96	99 %	KF704818.1
IGSTP657	363	363	98 %	2e-96	93 %	JX962799.1
diTPS11241	235	235	100%	2e-58	99 %	KJ158966.1
diTPS5871	237	237	100 %	5e-59	98 %	GU045757.1
APX802	217	217	57 %	2e-52	93 %	EF677796.1
APX8272			no significant similarity found			
APX637			no significant similarity found			
NAD+4472	525	525	100 %	3e-145	93 %	BT111057.1
SOD4683			no significant similarity found			
SOD3685	111		111	1e-21	98 %	FN564373.1

In case of the rest of 22 candidate genes, at first instance the 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 have been calculated, using Dna Sequence Polimorphism v6.10.01 (Librado and Rozas 2009) with the ClustalW results (BioEdit Sequence Alignment Editor version 7.0.9.0) (Hall 1999). as inputs. The results are presented in **Table 14**.

Table 14. Outputs of the sequence analysis in Scots pine samples with DnaSP v.5.0 effectuated on 22 candidate genes' 70 PCR products.

Abbrev.	Nr. of sequences	Nr. of sites	Nr. of polymorphic sites	Nr. of Haplotypes	Haplotype (gene) diversity	Variance of Haplotype diversity	Standard Deviation of Haplotype diversity	Nucleotide diversity
Myb 4633	3	296	0	1	0	0	0	0
Myb 4095	3	291	9	3	1	0.07407	0.272	0.02083
WRKY9928	3	176	0	1	0	0	0	0
WRKY20348	3	202	0	1	0	0	0	0
WRKY4214	3	306	0	1	0	0	0	0
WRKY1289	3	306	0	1	0	0	0	0
MADS15369	3	166	1	2	0.667	0.09877	0.314	0.00402
MADS1818	2	321	0	1	0	0	0	0
MADS2038	3	143	0	1	0	0	0	0
MADS12384	3	188	1	2	0.667	0.09877	0.314	0.00355
PAL1134	3	175	4	2	0.667	0.09877	0.314	0.01533
CHS381	2	215	2	2	1	0.25	0.5	0.0093
CHS4014	3	143	0	1	0	0	0	0
CHS7594	2	136	14	2	1	0.25	0.5	0.10294
F'3H2550	2	215	6	2	1	0.25	0.5	0.02871
F3H350	3	199	1	2	0.667	0.09877	0.314	0.00335
IGSTP657	3	255	5	3	1	0.07407	0.272	0.01307
diTPS11241	3	133	0	1	0	0	0	0
diTPS5871	3	137	0	1	0	0	0	0
APX802	3	278	2	3	1	0.07407	0.272	0.00481
NAD+4472	2	359	14	2	1	0.25	0.5	0.039
SOD3685	3	75	2	3	1	0.07407	0.272	0.01778

In total, 61 SNPs were found differently distributed over the analysed gene fragments. Because the low number of investigated trees made not possible a clearly definement of any tag of SNPs, an investigation of these SNPs have not been done. Excluding from the results the once appeared single nucleotid polimorphisms, nine SNPs were found in case of transcription factor coding genes, but these could be observed only in case of one sequence (Myb 4095). At enzyme coding genes, the number of SNPs was 49 and ranged from two to 14 (CHS7594, NAD+4472). The number of haplotypes ranged from one to three, higher in case of the enzyme coding sequences. The average number of nucleotide differences was the highest (14) in case of CHS7594 and NAD+4472, these two genes and also Myb 4095 have been found to be of special interest. Furthermore, the nucleotid diversity in enzyme coding sites was at more cases higher than at transcription coding sites. In a second phase of the analysis, the number of indels, and the character

of SNPs (synonymuous or non-synonymous) was determined, the results being summed in **Table 15**.

Table 15. Indels, synonymous and non-synonymous SNPs of the amplified candidate genes in the studied Scots pine samples.

Abbrev.	Number of sequences	Average InDel length event	Average InDel length	Number of InDel Haplotypes	InDel Haplotype Diversity	InDel Diversity, k(i)	InDel Diversity per site, Pi(i)	Theta (per sequence) from I, Theta(i) -W	No of non-synonymous SNPs	No of synonymous SNPs
Myb 4633	3	0	0	0	0	0	0	0	0	0
Myb 4095	3	1	1	2	0.667	1.333	0.0046	1.333	7	2
WRKY9928	3	1	1	2	0.667	0.667	0.00379	0.667	0	0
WRKY20348	3	0	0	0	0	0	0	0	0	0
WRKY4214	3	1	1	2	0.667	0.667	0.00372	0.667	0	0
WRKY1289	3	0	0	0	0	0	0	0	0	0
MADS15369	3	0	0	0	0	0	0	0	1	0
MADS1818	2	0	0	0	0	0	0	0	0	0
MADS2038	3	0	0	0	0	0	0	0	0	0
MADS12384	3	0	0	0	0	0	0	0	1	0
PAL1134	3	1	1	2	0.667	0.667	0.00381	0.667	3	1
CHS381	2	0	0	0	0	0	0	0	2	0
CHS4014	3	0	0	0	0	0	0	0	0	0
CHS7594	2	0	0	0	0	0	0	0	14	0
F'3H2550	2	1.2	1.2	2	1	5	0.02326	5	3	2
F3H350	3	0	0	0	0	0	0	0	1	0
IGSTP657	3	0	0	0	0	0	0	0	4	1
diTPS11241	3	1	1	2	0.667	0.667	0.00498	0.667	0	0
diTPS5871	3	0	0	0	0	0	0	0	0	0
APX802	3	0	0	0	0	0	0	0	0	2
NAD+4472	2	0	0	0	0	0	0	0	6	8
SOD3685	3	1	1	2	0.667	0.667	0.00877	0.667	2	0

As can be seen, indels were found in case of seven from 22 candidate genes. However, the detected values were higher in case of F'3H2550, but it was the gene on which the analysis was done only on two sequences. The highest values of InDel Diversity per site were found in case of SOD3685. Synonymous single nucleotid polimorphisms were found in case of six sequences, ranging from one to eight, in case of Myb 4095 (2), PAL1134 (1), F'3H2550 (1), IGSTP657 (1), APX802 (2), NAD+4472 (8), this partial sequence of NAD malate dehydrogenase being found to be of special interest. Non-synonymous single base mutations could be detected ranging as number

from one to 14, in case of 11 candidate genes, NAD+4472 (6), Myb 4095 (7) and especially CHS7594 with 14 SNPs being considered important for further analyses. To determine whether the 22 previously annotated candidate genes can indeed be associated to the putative enzyme or transcription factor, BLASTX search against the NCBI database was completed with every coding marker's each haplotype (amplified on samples from different type of habitat). At sequences, where only one haplotype was found, the homology based search was made only once, with one sample's PCR product. In several cases the highest BLAST results' definition was ambiguous (e.g. unknown [*Picea sitchensis*]). In that cases, the determination of the region where from that particular protein originates can carry great information content. For that reason efforts have been made also for characterising the region of each protein form the NCBI metatable by literature data mining (**Table 16**).

Table 16. BLASTX search results against the NCBI database of the 22 candidate genes detected in the studied Scots pine samples. Table contains beside the definition also the proteins' region. In case of markers with several haplotypes the search was made with use of all variants.

No of haplotypes	Query sequence/sample abbrev. (type of habitat)	Best BLASTX hit results	
		Protein definition	**Region**
2	Myb 4633/RMO (peat bog)	unknown [*Picea sitchensis*]	Helix-loop-helix DNA-binding domain
	Myb 4633/HZA (mixed forest)	unknown [*Picea sitchensis*]	Helix-loop-helix DNA-binding domain
3	Myb 4095/RMO (peat bog)	unknown [*Picea sitchensis*]	Myb-like DNA-binding domain
	Myb 4095/HZA (mixed forest)	unknown [*Picea sitchensis*]	Myb-like DNA-binding domain
	Myb 4095/SKV (rocky substrate)	unknown [*Picea sitchensis*]	Myb-like DNA-binding domain
1	WRKY9928/RMO (peat bog)	unknown [*Picea sitchensis*]	DNA binding domain
1	WRKY20348/RMO (peat bog)	hypothetical protein [*Oryza sativa* Indica Group]	box helicase
1	WRKY4214/RMO (peat bog)	unknown [*Picea sitchensis*]	WRKY DNA binding domain
1	WRKY1289/RMO (peat bog)	EF-hand domain [*Macleaya cordata*]	calcium ion binding
1	MADS15369/RMO (peat bog)	unknown [*Picea sitchensis*]	GUN4-like; pfam05419
2	MADS15369/ HZA (mixed forest)	unknown [*Picea sitchensis*]	GUN4-like; pfam05419
	MADS15369/SKV (rocky substrate)	unknown [*Picea sitchensis*]	GUN4-like; pfam05419
1	MADS1818/ RMO (peat bog)	xyloglucan galactosyltransferase KATAMARI1 [*Populus euphratica*)	Exostosin family
1	MADS2038/RMO (peat bog)	RNA polymerase II degradation factor 1 [*Nelumbo nucifera*)	Apolipophorin-III
	MADS12384/RMO (peat bog)	unknown *[Picea sitchensis*]	GUN4-like; pfam05419
3	MADS12384/HZA (mixed forest)	unknown [*Picea sitchensis*]	GUN4-like; pfam05419
	MADS12384/ SKV (rocky substrate)	unknown [*Picea sitchensis*]	GUN4-like; pfam05419
	PAL1134/ RMO (peat bog)	phenylalanine ammonia-lyase [*Pinus pinaster*]	phenylalanine ammonia-lyase
3	PAL1134/ HZA (mixed forest)	phenylalanine ammonia-lyase [*Pinus pinaster*]	phenylalanine ammonia-lyase
	PAL1134/SKV (rocky substrate)	phenylalanine ammonia-lyase, partial [*Pseudotsuga menziesii*]	phenylalanine ammonia-lyase
2	CHS381/RMO (peat bog)	unknown [*Picea sitchensis*]	synthesis and degradation of fatty acids
	CHS381/HZA (mixed forest)	unknown [*Picea sitchensis*]	synthesis and degradation of fatty acids
1	CHS4014/RMO (peat bog)	chalcone synthase-like protein, partial [*Picea sitchensis*]	synthesis and degradation of fatty acids
2	CHS7594/HZA (mixed forest)	chalcone synthase, partial [*Taxus baccata*]	synthesis and degradation of fatty acids
	CHS7594/SKV (rocky substrate)	chalcone synthase, partial [*Taxus baccata*]	synthesis and degradation of fatty acids
2	F'3H2550/HZA (mixed forest)	flavanone 3-hydroxylase *[Pinus radiata*]	2OG-Fe(II) oxygenase superfamily
	F'3H2550/SKV (rocky substrate)	flavanone 3-hydroxylase [*Pinus radiata*]	2OG-Fe(II) oxygenase superfamily

	F3H350/SKV (rocky substrate)	flavanone 3-hydroxylase [*Pinus radiata*]	2OG-Fe(II) oxygenase superfamily
3	F3H350/HZA (mixed forest)	flavanone 3-hydroxylase [*Pinus radiata*]	2OG-Fe(II) oxygenase superfamily
	F3H350/ RMO (peat bog))	flavanone 3-hydroxylase [*Pinus radiata*]	2OG-Fe(II) oxygenase superfamily
	diTPS11241/ RMO (peat bog	CPS1 [*Pinus tabuliformis*]	Isoprenoid Biosynthesis enzymes
3	diTPS11241/HZA (mixed forest)	CPS1 [*Pinus tabuliformis*]	Isoprenoid Biosynthesis enzymes
	diTPS11241/SKV (rocky substrate)	CPS1 [*Pinus tabuliformis*]	Isoprenoid Biosynthesis enzymes
1	diTPS5871/ SKV (rocky substrate)	ent-copalyl diphosphate synthase [*Picea sitchensis*]	Isoprenoid Biosynthesis enzymes
	APX802/RMO (peat bog)	unknown [*Picea sitchensis*]	Ascorbate peroxidases and cytochrome C peroxidases
3	APX802/ HZA (mixed forest)	unknown [*Picea sitchensis*]	Ascorbate peroxidases and cytochrome C peroxidases
	APX802/ SKV (rocky substrate)	unknown [*Picea sitchensis*]	Ascorbate peroxidases and cytochrome C peroxidases
2	NAD+4472/ HZA (mixed forest)	malate dehydrogenase [*Larix kaempferi*]	Glyoxysomal and mitochondrial malate dehydrogenases
	NAD+4472/SKV (rocky substrate)	malate dehydrogenase [*Larix kaempferi*]	Glyoxysomal and mitochondrial malate dehydrogenases
	SOD3685/RMO (peat bog)	unknown [*Picea sitchensis*]	Heme-dependent peroxidases similar to plant peroxidases
3	SOD3685/HZA (mixed forest)	unknown [*Picea sitchensis*]	Heme-dependent peroxidases similar to plant peroxidases
	SOD3685/SKV (rocky substrate)	unknown [*Picea sitchensis*]	Heme-dependent peroxidases similar to plant peroxidases

As the data from **Table 16** presents, it is evident, that all haplotypes of a certain gene could be assigned to the same type of protein. Secondly, the BLASTX search revealed, that at superficial sight, the annotation process was made with less accuracy in case of transcription factor coding genes than at the enzyme coding ones. The definition of the search-resulted proteins in most cases are not consistent with the putative function of the gene. In case of Myb 4633, its putative coded protein is a Myb transcription factor, and according to the search results, it's function is unknown, with a *Picea sitchensis* origin, from a Helix-loop-helix DNA-binding domain region. However, a basic helix-loop-helix (bHLH) is a protein structural motif that characterizes one of the largest families of dimerizing transcription factors. In case of WRKY9928, the putative protein is a WRKY transcription factor, and the homology search has as result a hypothetical protein from the *Oryza sativa* Indica group, box helicase region. However, box RNA helicases are involved in almost every aspect of RNA metabolism, associated with diverse cellular functions including plant growth and development, and their importance in response to biotic and abiotic stresses. According to Shamimuzzaman and Vodkin (2013), the WRKY family of transcription factors also may contain a repeated box helicase-binding domain. The WRKY1289 sequence's putative protein was also a WRKY transcription factor, and due to the BLASTX search, it is an EF-hand domain

(*Macleaya cordata*), from a calcium ion binding region. As in the previous case, looking in more detail, there is evidence in the literature, that CaM binds specifically to the Ca^{2+}-dependent CaM-binding domain (CaMBD) of AtWRKY7 (Park *et al.* 2005). GUN4 seems to be involved in sensing elevated levels of photoreactive tetrapyrrole intermediates originating from chlorophyll biosynthesis (Brzezowski *et al.* 2014). The MADS1818 sequence, having as BLASTX result to encode xyloglucan galactosyltransferase KATAMARI1 (*Populus euphratica*) has been searched with UniGene and Conserved Domain Arhitecture Retrieval Tool, to find any similarity with MADS conserved domains. MADS box proteins as well as cofactors interacts of type-II membrane protein metabolism, catalyzed by xyloglucan galactosyltransferase (Messenguy and Dubois 2003).

6. DISCUSSION

6.1. Morphological and anatomical trait differentiation in peripheral populations from the Carpathian region

Peripheral populations of Scots pine involved in our study are considered to be of natural origin and most probably have persisted in refugial territories with specific local ecological conditions. Plants are integrated systems, and the study of traits that influence their survival and reproduction can reveal diversification in contrasting environments (Reich *et al.* 2003). However, populations or genotypes can be preadapted to a given selection factor or environmental condition. When they colonize new habitats or geographic areas, their survival can depend on their functional traits, which are or are not suited to the environment (Reich *et al.* 2003). Thus, current species distributions may reflect ecological pre-sorting processes, in addition to *in situ* adaptive evolution (Losos 1996), indicating their origin and their historical colonization routes.

Differences among peripheral populations of Scots pine found in our study are in line with the earlier described variation of the species. Differentiation of populations from the Iberian Peninsula has been described using 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). The findings of Oleksyn *et al.* (1998) suggests that in order to develop new models to predict the responses of Scots pine to changing environmental conditions, it is necessary to consider intraspecific differentiation in the acclimatization and adaptation to environmental factors. Cone characteristics are traits of the highest discriminating power in inter-populational comparisons among regions in the work of Bobowicz and Korczyk (2000) and Jasińska *et al.* (2014). The principal variables which proved to be indicative of discriminating populations were also found to be needle characteristics by Androsiuk *et al.* (2011). Several studies which describe differentiation among populations of other species from the genus *Pinus* can also be found in the literature of Marcysiak (2004), Bączkiewicz *et al.* (2005) and Sobierajska and Marcysiak (2010).

6.1.1. Phenotypic differentiation of populations by geographical distribution

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 (Figure 15) and two needle characteristics (Figure 17). 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). Accordingly one main recolonization route in Europe presumably originated from around the Eastern Alps and the surroundings of the Danube plain, and there is evidence 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).

Literature data are also consistent with our findings based on needle anatomical characteristics, according to which 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. Pollen based vegetation reconstructions of LGM vegetation in the Eastern Carpathians between ~22,870 and 19,150 cal yr BP, supported the persistence of *Pinus sylvestris* in this geographical region (Magyari *et al.* 2014b).

The application of the Mantel test to correlate the morfo-anatomical and geographic distances and seek for spatial patterns, yielded no significant correlation. We assume that Scots pine from the Carpathians represents only a small geographic range of the species' large distribution. Evidence from various fossil proxies, palaeoclimatic modelling, and genetic research suggests that conifers and some broadleaf trees were continuously present throughout LGM in refugial territories around the Pannonian Basin (Mitka *et al.* 2014, Ronikier 2011, Willis and Van Andel 2004).

6.1.2. Differentiation by the habitat type

Under extreme conditions, peripheral populations with specific structures are exposed to dramatic environmental changes which will impose novel selection pressures and may therefore cause adaptive responses (Bone and Farres 2001). Morphological traits of conifer species are known to vary adaptively with geographic, climatic, and edaphic variables (Ji *et al.* 2011). In several works in the existing literature, significant differences between isolated populations have been reported, mostly with regard to the morphological features (Szweykowski and Urbaniak 1982). On the other hand, abiotic factors, such as temperature, light, soil type, available nutrients, and other derived factors (Pawlaczyk *et al.* 2010) lead to the development of local phenotypes, i.e. populations with distinguishing characteristics. These populations are present in a given area within the range of the species (Remlein *et al.* 2015). The samples of Scots pine which we examined here were mainly from specific habitat types, such as raised bogs, dry rocky outcrops, young glacial deposits, siliceous and acidic or mesic sandy substrates, and they were from different altitudinal gradients which varied from 252 m (HFE) to 1,107 m (STU).

Discriminant function analysis based on the eight measured cone characteristics revealed clearly discernible groups of populations. 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 also can be due to the lack of competition with other pines (Gil *et al.* 2002).

However, the results of one-way MANOVA test on the cone datasets revealed significant differentiation among populations growing in peat bogs and on rocky surfaces in the case of the following traits: length, width, and thickness of apophysis (AL, AW, AT), number of scales (NBS), and the cone length/cone width (CL/CW) ratio. If we take 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). Beaulieu and Simon (1995) have shown that no general geographical grouping can be detected in the observed variation in *P. strobus* collected from ten natural populations in the Canadian province of Quebec. Furthermore, cones from individuals sustaining on dry rocky outcrops have been separated as being significantly smaller and, on the basis of apophysis width (AW), number of scales (NBS), and cone length/width ratio (CL/CW), well-proportioned but less dense in structure. Significant

differences in needle anatomy were found among populations with different habitat types. One-way MANOVA test applied to needle anatomical data showed significantly less resin ducts (NRD) among populations from peat bogs and rocky surfaces or significantly more NRD among populations with 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 collected from 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 adaptation variations within and among populations. The calculated CCH/NH and CCW/NW proportions actually estimate the dimensional relationship of central cylinder with the photosyntetically active mesophyll. Lower values indicate a thicker mesophyll and, consequently, increased photosynthetic activity. We have found significantly high values of needle width (NW), cell rows in the armed palisade parenchyma on the concave and convex mesophyll (NBRCC, NBRCW), and significantly lower values of CCH/NH proportion among populations from mixed forests. Considering that our mixed forest samples were collected from areas in the Pannonian Basin with the lowest altitudinal gradient, our NW 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 higher NBRCC and NBRCW values and lower CCH/NH proportion indicates a thicker mesophyll and, therefore, increased photosynthetic activity. Plants growing under strong light have well developed palisade parenchyma, thicker leaves, relatively large leaf area, higher biomass, increased photosynthesis, and lower contents of chlorophyll, carotenoid, and nitrogen (Huang *et al.* 2008, Je *et al.* 2006, Volkova *et al.* 2010, Yang *et al.* 2007). Structural characteristics of conifer needles are often strongly related to gradients in long-term light availability within canopies and across stands (Niinemets *et al.* 2007, Richardson *et al.* 2001, Richardson *et al.* 2000). The lower NBRCC and NBRCW values and higher CCH/NH proportion in peat bog populations and populations growing on rock surfaces can be explained as a response to increasing environmental stress, which is accompanied by a decline in photosynthetic and growth rate, since higher leaf age compensates for low photosynthesis (Reich *et al.* 1995, Schoettle *et al.* 1994). According to some authors, Scots pine needles from the sites with limited phosphorus availability (such as peat bogs or rock surfaces) are narrower and thinner than needles from nutrient-rich site (Niinemets *et al.* 2001).

Our findings are congruent with these data, the width and height of needles in populations from peat bogs and on rock surfaces being significantly low. Scots pine, however, typically occurs on different types of well-drained mineral soils, representing a broad range of variation in pH, and

nutrient availability (Persson 1980). Consequently, not only the geographical range, but also the ecological tolerance of the species is wide. Variation in the leaf traits is an important characteristic of ecological processes that are driving forces for biogeochemical cycles in ecosystems (Reich *et al.* 1992). Studies of needle traits allow us to gain important insights into these processes and predict ecosystem responses to changes in the environment.

6.2. Habitat type differentiation in peripheral populations based on seed traits and germination data

In natural habitats, the resources such as mineral nutrients, water and light necessary for plant growth are heterogenously distributed in space and time (Zhang *et al.* 2017). In these habitats the quality of seeds depends on the genetic constitution of the seeds (Andersson 1965) and on the modifying effect of the habitat, the environmental properties of the growing sites of the mother tree (Castro 1999). Latter interactions of seeds with their environment also can be highly specific, influencing the germination of individuals and the distribution of species (Chambers and Macmahon 1994).

Different morphological characters, such as various wings, seed shapes and sizes, lead to further differences in seed dispersal patterns that are related to various ecological strategies. Species produces seeds with highly divergent morphology. The level of differences in 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). Morphological differences of seeds with regard to different environmental conditions are also listed in works of several authors. Seed size (mass) is essential to many aspects of plant ecology and evolution (Leubner-metzger *et al.* 2010, Moles *et al.* 2005). Korona (1996) suggested that the seed size of a species represents the amount of maternal investment in individual offspring. Navie *et al.* (1998) argued that the less dispersible heavier seeds produced in natural areas would likely form a persistent seed bank. Larger seeds have better survival in dry conditions, and also larger seeds have higher percentage of emergence and survival (Leishman and Westoby 1994).

6.2.1. Seed morphological characters

Seed size often increases with increasing dryness (Baker 1972), presumably because of the need on dry sites for vigorous early seedling development. Latter findings supports our observation regarding to the significantly higher values after performing One-Way ANOVA test in case of RBE (dry rocky substrate) population at SL, SW, WL, WW, SWE traits (seeds with significantly higher size and mass). Our samples from RML showed significantly lower values at SWE, SL/SW, SL/WL, SW/WW traits. This peat bog population has seeds with lower mass, longer shape associated with shorter and slimmer wings. In literature, the size and shape of the seed is described as a feature which can easily change, not only with the climatic conditions of the year but even with the difference in cone size, the number of seeds per cone and the position within the cone (Ehrenberg *et al.* 1955). The number of seeds in each cone sampled was not determined in our study, but our previous results suggested significant differences rather in structure (by higher values of length, width, and thickness of apophysis) than in the size of cones from peat bog populations (Köbölkuti *et al.* 2017). However, due to Keeley and Zedler (2000), correlation between cone size and seed size is poor. Although seed size is one of the least variable trait in plants (Marshall 1986), seeds do show a considerable degree of phenotypic plasticity in response to the environmental conditions under which they develop (Fenner 1992). The lower seed weight, seed length/seed width, seed length/wing length and seed width/wing width (SWE, SL/SW, SL/WL, SW/WW) traits of seeds from RML 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) and a peat bog can be characterized by specific edaphic conditions. Our significantly low values for wing width (WW) variable in case of SME, RMO and RPS peat bog populations are supported by the findings of (McGinley *et al.* 1990) in lodgepole pine.

Based on the measured morphological datasets, discriminant function analysis revealed a slight separation of populations from peat bogs defined by lower values of seed length, wing width and seed weight (SL, WW, SWE) variables. Nevertheless, these morphological traits are probably influenced by the cone structure (McGinley *et al.* 1990), tree age, general health of the trees, and the specific macro- and micro- habitat of the parent trees (Dangasuk and Panetsos 2004).

Within the hierarchical clustering, some samples (from RMO and RPS both from peat bog) were included in one subcluster, also the RCO and CHR samples (rocky substrate), with a relatively well supported relationship with HZA (mixed forest) samples. Nevertheless, these mixed forest stands grow on silicious substrate with low nutrient content. Despite of this, our dendrogram showed that populations are not homogeneous in grouping by habitat type regarding seed and wing morphology. 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 within hybrid swarm populations of *Pinus sylvestris* and *P. mugo* formerly mentioned by Christensen and Dar (1997) and Wachowiak and Prus-Glowacki (2008).

The results of our coefficient correlation analysis concerning 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 betwen the morfological 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. We emphasize also that based on our former genetic research, as well as paleoclimatic modelling data and fossil evidences conifer species and some broadleaf trees were continuously present throughout the LGM period in refugial territories around the Pannonian Basin (Mitka *et al.* 2014, Ronikier 2011, Willis and Van Andel 2004).

6.2.2. Germination associated parameters

Morphology and seed size are usually correlated, but how morphology affects germination and seedling growth is less understood. The chance that a seed will develop into an established seedling is dependent upon the site, which provides the specific conditions for its germination and development (Sheldon 1974).

By carrying out the discriminant function analysis with pre-formed three groups according to habitat type (peat bog, rock surface and mixed forest), with both seed morphological and germination associated parameters, our results revealed a strong pattern of differentiation. Peat bog populations were defined by lower values of Germination %, seed length and wing length (PG, SL and WL) for Function 1 and for Function 2 by values of Germination Speed, seed weight, seed width, wing width and Germination Value (GS, SWE, SW, WW and GV) situated between the values of populations from mixed forests and rocky substrate. Our lower seed length (SL) associated with intermediate seed width (SW) values means smaller seeds. 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 size of the seeds from peat bogs could be explained by these specific conditions characterizing this type of habitat.

The results of our Bivariate Correlation analysis was significant at 0.05 level between of Germination % (PG) with seed length, seed width and seed weight (SL, SW and SWE) likewise between Germination Speed (GS) with seed length/wing length (SL/WL) and seed width/wing width (SW/WW) variables. Early studies have shown that seedlings grown from large seeds have higher seedling establishment, growth and survival. Our lower PG values also correlates and commensurates with smaller seed size. The GS variable is one of the oldest concept of seedling vigor. The interest in germination speed is based on the theory that only those seeds which germinate rapidly and vigorously under favourable conditions are likely to be capable of producing vigorous seedling in field conditions. The SL/WL and SW/WW variables defines the relation between the size of the seeds in relation to the wings. As these two variables have higher values, seeds have increased size compared to their wings. The presence of endosperm in mature seeds provides the size of the seed and stored proteins in endosperm are an important source of amino acids and energy production during early germination (Angelovici *et al.* 2011). Higher seed size values in comparison to the wings means increased amount of endosperm and consequently vigorous germination.

6.3. Analysis of newly developed molecular markers

As results of the analysis of the newly developed molecular markers, partial sequences of 25 different candidate genes have been investigated. The cDNA fragments previously annotated on *Pinus cembra* were successfully amplified on Scots pine DNA samples originating from three types of habitat: peat bog (RMO), mixed forest (HZA) and rocky surface (SKV). The first aim was to test these molecular markers, developed on one pine species, and test the applicability on the other species, Scots pine in this case. The second drafted specific aim consisted in detection of polimorphisms.

For testing the corectness of annotations, all fragments were searched against the NCBI database by BLASTN search. In case of three previously annotated sequences no significant similarity was found. There are several possible reasons of these previous annotation mistakes. It may be the low length of the genomic query, or the database is too exiguous, being insufficient to detect homology in a reliable way. Or could represent an accurate inference of function from homology. Typical database searching methods are valuable for finding evolutionarily related proteins. Nevertheless, there are only about 1,000 major superfamilies in nature (Brenner *et al.*

1997), and most homologs must have different molecular and cellular functions. Another annotation problem could be in case of genes with incorrect functions uploaded into public databases (Bork and Bairoch 1996).

In case of 22 candidate genes, at first instance the number of polymorphic sites have been determined. The values of polymorphic sites ranged from one to 14, Myb4095 (9), CHS7594 (14) and NAD+4472 (14) being considered of special interest. These candidate genes, by determining biochemical or physiological functions under their control could possibly relate a QTL to some phenotypic variation. Myb genes seems to be transcriptional activators, with ability to regulate lignin synthesis enzymes with an important role in rigidity and impermeability of wood (Bedon *et al.* 2007), in the isoprenoid and flavonoid biosynthesis (Bedon *et al.* 2010) and in the control of the PAL gene transcription (Osakabe *et al.* 2009). The chalcone synthase (CHS) gene expression is influenced by many kind of stress and environmental factors (Dixon and Paiva 1995). High levels of NADP+ occur in conditions such as high light, high CO2, NH4+ nutrition, cold stress (Scheibe 2004). This sites' polimorphism represents non-conserved positions. Mutations have as result possible changes in above mentioned transcription factors and enzymes' constitution and consequently, their function (Fu 1995).

Until haplotypes are specific alleles in a cluster of tightly linked genes or a set of SNPs on one chromosome that tend to occur always together, haplotype diversity is a measure of the uniqueness of a particular haplotype in a given population (Nei and Tajima 1981). By using environment-fitting analyses, shifts in haplotype frequency can be used for association between these transitions in haplotype frequency and environmental attributes. Elective forces acting on allelic variants of genes have a profound effect on local levels of the genetic diversity and linkage disequilibria. Positive directional selection leads to reduced variability and increased linkage disequilibrum in the respective region (Kohn *et al.* 2000), and the so-called selective sweep regions provide clues to genes that have been subjects of evolutionary forces as well as selection (Clark *et al.* 2004). In this aspect, the number of haplotypes of the newly developed markers (ranging from 1 to 3) and haplotype diversity (between 0 and 1) – limited of course by the low number of the studied samples/gene – affords the opportunity to examine the connection between selection and diversity in *Pinus sylvestris*.

The nucleotide diversity found ranged from 0 to 0.1. Nucleotide diversity, as a measure of genetic variability is defined as the number of nucleotide differences per site between two or more randomly chosen sequences from a population (Li and Sadler 1991). It is influenced by several factors including mutation, migration, selection and random genetic drift (Wachowiak *et al.* 2011). The observed nucleotide differences were not distributed randomly among the genes studied. From 22, ten sequences analysed on three (or in some cases two) samples were identical whereas a

number of 12 showed multiple differences. This distribution could be partly due to variation in mutation rate among regions (Torre *et al.* 2017), partly due to sequencing errors (case of NAD+4472, F'3H2550, CHS7594, CHS381, MADS1818, without convincing chromatograms), where only two samples' PCR products were analysed, or due to differences in sequence lengths. It should be noted that in terms of polimorphism, our candidate genes were chosen due to their biological importance in adaptation, based on literature and not for any known mutations or protein variants. It should be also noted, that the constructed BLAST database during the annotation was far from being complete and this could be also a source of systematic bias, the sequences for the database being chosen randomly. However, the relative levels of diversity in different DNA regions shown in Table 14, can be explained by the relative stringencies of selective constraints in different regions (Kohn *et al.* 2000). Table 14 suggests that the nucleotide diversity in *Pinus sylvestris* is at most of the order of 0.1 (case of CHS7594). A few possible explanations for this low value might be that the mutation rate in the species is low, or the gene flow is high, at least in case of the sampled populations. The effective population size has been relatively small in the past and the species has gone through a severe bottleneck in the past. Another explanation is that the studied genes are under positive directional selection. Scots pine biology provides for the long-term maintenance high within-population and low among-population genetic diversity (Robledo-Arnuncio *et al.* 2005). Nucleotide diversity in central and northern European populations is compatible with an ancient bottleneck (Pyhäjärvi *et al.* 2007). Also, the selected sequences have as products proteins with importance in adaptation. The conclusion that can be traced is, that excepting the low mutation rate, the relatively low nucleotide diversity is a result of combination of their colonization history and the influence of *in situ* ecological factors. It is interesting to note that sequences from the same type of gene (e.g. Myb 4633 and Myb 4095); all the MADS related sequences or the chalcone synthase (CHS) candidates may differ from each other regarding nucleotide diversity. One simple explanation for the differences between these sequences is that they have been maintained in the peripheral populations for a long time by overdominant selection (Gillespie and Turelli 1989). It should be noted, that even under overdominant selection many of the polymorphic alleles would have been lost if severe bottleneck occurred in the past. Of course, if we think, that for example allozyme heterosis can be poorly explained at some conifers (Strauss and Libby 1987), or that overdominant lethals are parts of the conifer embryo lethal system (Williams *et al.* 2003), the deducible conclusion is, that the reason may be the specific colonization history of the species. Olson *et al.* (2010) also found in *Populus balsamifera* that nucleotide diversity in gene fragments is affected by the type of the mutation, being higher in gene fragments with indels, than in fragments that did not contain indels.

The results of Table 14 can be used to estimate the heterozygosity at the nucleotide level. In the present study, among the 22 sequences on three samples examined, 12 differ by at least one nucleotide and so the average heterozygosity at the nucleotide level is 12/22 = 54.5 %. Obviously, the heterozygosity for a sequence is dependent on its size.

The study was able to detect numerous SNPs and indels, potentially involved in adaptation processes to different environmental conditions. Indels in the exon regions are important due to their influence on protein structures and by this, on phenotypic trait differences. However, the sequences modified by indels were able to be associated to the putative enzyme or transcription factor by BLASTX search. Some of the non-synonymous SNPs detected are of special interest because they might have an influence on the protein structure and function. For example, Myb transcription factors (Myb 4095, with seven non-synonymous SNPs) are transcriptional activators, with ability to regulate lignin synthesis enzymes. and central regulators of anthocyanin biosynthesis. Conformation changes of the protein could have effect on the function of this transcription factor family, on the isoprenoid and flavonoid biosynthesis.

The variation found can be used to develop SNP markers and to apply them additionally or instead of neutral SSR or AFLP markers. For many applications, SNP markers are more optimal, since they are suitable for high throughput analysis, low-cost, highly reproductible, easy to score, comparable between laboratories and some of them can show high differentiation values. The virtually unlimited number of SNP markers in different parts of the genome of pines create opportunities for the investigation of genetic variation within the species with numerous applications in population genomics. The SNPs found have important applications in the study of adaptation. The Myb gene family members are transcriptional activators (Abe *et al.* 2003), lignin synthesis (Douglas 1996), isoprenoid (Bedon *et al.* 2010), flavonoid biosynthesis enzyme regulators (Winkel-Shirley 1999); the MADS box family members have roles on the male and female strobili (Mouradov *et al.* 1998), in needle primordia (Carlsbecker *et al.* 2004), vegetative shoots (Michaels 1999), microsporophyll primordia (Sundström and Engström 2002), cone axis (Carlsbecker *et al.* 2003) or the apical meristem formation (Egea-Cortines *et al.* 1999). Many WRKY proteins are involved in the defense against attack from pathogenic bacteria and fungi, viruses and oomycetes (Eulgem and Somssich 2007). Further, WRKY genes are involved in the response to the abiotic stress of wounding (Hara *et al.* 2000), oxidative stress (Liu *et al.* 2014), the combination of drought and heat, cold and salinity (Chen *et al.* 2012). It is also evident that some members of the family may play important regulatory roles in morphogenesis of trichomes (Zhang and Wang 2005), embryos (Ueda *et al.* 2011), senescence, dormancy (Robatzek and Somssich 2001), plant growth (Singh *et al.* 2002) and metabolic pathways nutrient stress, UV radiation, and dark treatment and tolerance to phosphate (Chen *et al.* 2012). Furthermore, enzymes like ascorbate

peroxidase are hydrogen peroxide-scavenging (Asada 1992), terpene synthases are implicated in several ecological and physiological functions in response to biotic and abiotic environmental factors (Martin *et al.* 2004), NAD malate dehydrogenases have defense functions to conditions such as light intensity, high CO_2, NH_4^+ nutrition, cold stress (Edwards and Nakamoto 1985); superoxide dismutases are indicators of enhanced O_2 production and inductors of antioxidant activity in winter-photooxidation injury (McKersie *et al.* 1999). Flavonoid biosynthesis enzymes have important role in providing floral pigments, antibiotics, UV protectants and insect repellents (Ferreyra *et al.* 2012). All these proteins coding loci can be directly linked with adaptive genetic variation, these genetic differences having direct effect on fitness. As genetic diversity and fitness are the basis for the ability of species to adapt to changes of the environment (Krutovsky and Neale 2005), adaptive genetic variation in relevant genes are essential for the long term adaptation to stressful conditions. For that reason, the knowledge of adaptive genetic variation is a basis for future management and conservation strategies.

The results presented here could be considered as „tools" for further association mapping studies in order to identify genomic regions with role in phenotypic variation. The study revealed differences in diversity among the investigated candidate genes. In this view, the application of these molecular markers developed by the identification of genes, the characterisation of SNPs in these markers may contribute to investigation of the genetic basis of adaptive variation. Of course, always will be the possibility to use the whole genome for these applications. But considering the cost and technical possibilities to study the whole genome, the sequencing and analysis of particular genes still will be the best available method.

7. NEW SCIENTIFIC ACHIEVEMENTS

I. The study based on cone morphological and needle anatomical marks provided evidence for population differentiation on geographical scale in the Carpato-Pannonian region. Cone morphometry and needle anatomy analyses revealed common origin of the Scots pine populations from the Northern Carpathians and the Pannonian Basin formerly highlighted also by SSR marker study. These findings together support phylogeographical evidence of refugial locations existed most probably in the Eastern Alps and East-Central Europe, and suggest one main recolonization route in Central Europe originated from around the Eastern Alps and the surroundings of the Danube plain.

II. Significant differentiation was revealed among populations growing on different extreme ecological sites, as peat bogs and rocky surfaces. This was based on cone morphology and needle anatomy. Statistical analysis applied to needle anatomical data showed significantly less resin ducts among populations from peat bogs and rock surfaces or significantly more resin ducts among populations with mixed forest provenience. As variation in needle traits is an important characteristic of ecological processes that are driving forces for biogeochemical cycles in ecosystems, these findings allow us to gain important insights into these 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. Seeds collected from a dry, exposed habitat have evidenced to be increased in size and weight, most probably because of the need of early seedling development. Higher seed size values means increased amount of endosperm. The correlation between the seed morphology and germination associated data may define a successful multi-trait selection. Seeds collected from peat bogs were revealed as smaller, with longer shape. The correlation between these smaller seeds and lower germination percent could be helpful in the early evaluation for seed selection.

IV. Development and test of new molecular markers on candidate genes previously annotated with possible role in adaptation. These markers were developed on *Pinus cembra* transcriptome, then tested and investigated on *Pinus sylvestris* DNA samples from different

habitat types. In case of several analysed genes polymorphic sites, haplotype diversity, nucleotide diversity, insertions/deletions as well as synonymous and non-synonymous SNPs were detected. The variation found is considered to be important in the study of adaptation, as the relevant genes are essential for the long term adaptation to stressful conditions.

8. SUMMARY

Carpathian Scots pine populations having peripheral distribution within the species' range are often sustained in specific types of habitats, such as peat bogs and rocky surfaces or lime consisting sandy substrates. On specific habitats or geographic areas, their survival can depend on functional traits, which are or are not suited to the environment. Due to the long time adaptive processes, which involve genotypes that can be preadapted to a given selection pressure, historically isolated populations always have been subjects of particular interest in studies of in situ adaptation. Earlier studies on the basis of morphological character expression concluded, that differentiation of populations are influenced by both the environmental conditions and the distinct genetic structure resulted from historical heritage. Based on the compiled literature, cone and needle characteristics are traits of the highest discriminating power in inter-populational comparisons among regions. These findings also suggest that in order to develop new models to predict the responses of Scots pine to changing environmental conditions, it is necessary to include intraspecific differentiation in acclimatization and adaptation to environmental factors.

Seed morphology and seed germination power can be also useful to characterize genotypes in natural populations. Measures of size and shape, their correlation and relationship with the germination capacity may be either the result of developmental programs or the response to a specific environmental condition. The seed size and weight are strongly determined by the genotype and the environmental factors acting on the mother tree. The plants' survival strongly depends on the dispersal potential of the seeds, and specific morphological adaptations, influencing their movement towards a suitable germination microsites. Germination and characteristics of juvenile seedlings are also determined through seed morphology.

Being able to distinguish between genotypes that are relevant to a trait of interest, molecular markers have totally changed our view of nature. The latest techniques promise to provide cheap, high-throughput methods, SNP genomic markers becoming indispensable tools for understanding adaptive responses against stressful environmental conditions. The genetic differences on nucleotide level have direct effect on fitness. As genetic diversity and fitness provide the basis for species' ability to adapt to the changes of the environment, the knowledge of adaptive genetic variation has high priority in future management and conservation strategies.

In our study, on one hand 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. Six cone morphological and eight needle anatomical characters were measured and four cone morphological and four needle anatomical ratios were calculated. 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 annotated with possible role in adaptation were tested and investigated on *Pinus sylvestris* samples. In case of three annotated sequences no significant similarity were found with the putative encoded protein. The other 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 number of haplotypes of the newly developed markers ranged from one to three and haplotype diversity was between zero and one. Considering the limitation of the low number of the studied samples, this result still affords the opportunity to examine the connection between selection and diversity. Nucleotid diversity was up to 0.1, was lower compared to the nucleotide diversities analyzed in other species, but this fact could be due to the low number of investigated samples. 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.

9. ACKNOWLEDGEMENT

First of all, I would like to thank the members of my thesis committee for their insightful comments and encouragement, but also for the questions which incented me to widen my research from more various perspectives.

Most of all, I would like to express my sincere gratitude to my supervisor Dr. Mária Höhn. Not only for the continuous support of my PhD study and related research, for her patience, motivation, and knowledge. Also because she treated me almost as a family member. At least I felt that way, especially in my homeless moments.

My sincere thanks goes to Dr. Berthold Heinze and the Genome Research Unit staff of the Federal Research and Training Centre for Forests, Natural Hazards and Landscape (BFW), Vienna, Austria, who provided me an opportunity to join their team, and who gave access to the laboratory and research facilities. Without their precious support it would not be possible to conduct this research.

I am grateful to the Department of Botany staff, Szent István University, for their cooperation and understanding that they have outlined all these years as such for somebody with a totally different professional background.

With a special mention to Dr. Attila Borovics, Dr. Klára Cseke and Department of Breeding staff, Forest Research Institute (ERTI) - NAIK for providing the understanding, time and professional support to work on my thesis in their institute.

I am also grateful to my colleague and friend, Dr. Endre György Tóth for the exhaustive labwork and assistance in sample collection. Many thanks for our soul preserving discussions, hard working hours, for all the fun we have had in the last three years. For the stimulating coffee-breaks.

I would also like to acknowledge Dr. Márta Ladányi for the professional support in dataset analysis, Krisztina Tóth for the assistance in data collection and Gábor Bottlik for all the work he offered maintaining the common-garden trial.

I am grateful to the National Research, Development and Innovation Office (Hungary) for the grant of the Hungarian Scientific Research Fund [NKFI K119208].

Last but not the least, I would like to thank my family. My brother, who always will be a life-coach for me in the scientific world, and especially my wife, for the limitless spiritual and financial support throughout working on this thesis and my life in general. I apologize for all these irretrievable years passed with so few common memories.

As ending, a few words to my son: please, consider this work as a faint proof of a great effort for at least to approach the person, who you think I am.

10. REFERENCES

Aaltonen, H., Lindén, A., Heinonsalo, J., Biasi, C., & Pumpanen, J. (2016). Effects of prolonged drought stress on Scots pine seedling carbon allocation. *Tree Physiology*, 37(4), 418–427.

Abd-Elsalam, K. A. (2003). Bioinformatic tools and guideline for PCR primer design. *African Journal of Biotechnology*, 2(5), 91–95.

Abe, H., Urao, T., Ito, T., Seki, M., & Shinozaki, K. (2003). Transcriptional activators in Abscisic acid signaling. *Society*, 15(1), 63–78.

Abugessaisa, I., Saevarsdottir, S., Tsipras, G., Lindblad, S., Sandin, C., Nikamo, P., ... & Tegnér, J. (2014). Accelerating translational research by clinically driven development of an informatics platform–a case study. *PLoS One*, 9(9), e104382.

Ackerly, D. D. (2000). Taxon sampling, correlated evolution, and independent contrasts. *Evolution*, 54(5), 1480–1492.

Ackerly, D. D., & Reich, P. B. (1999). Convergence and correlations among leaf size and function in seed plants: comparative tests using independent contrast. *American Journal of Botany*, 86(9), 1272–1281.

Adams, J. M., & Faure, H. (1997). Global land environments since the last interglacial. *Oak Ridge National Laboratory, TN, USA*, pp. 1-20. Available at: <http://www.esd.ornl.gov/ern/qen/nerc.html> [Accessed: 10.11.2017].

Alba, M. M., & Guigo, R. (2004). Comparative analysis of amino acid repeats in rodents and humans. *Genome Research*, 14(4), 549-554.

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.

Alleaume-Benharira, M., Pen, I. R., & Ronce, O. (2006). Geographical patterns of adaptation within a species' range: Interactions between drift and gene flow. *Journal of Evolutionary Biology*, 19(1), 203–215.

Allen, J. E., & Salzberg, S. L. (2005). JIGSAW: Integration of multiple sources of evidence for gene prediction. *Bioinformatics*, 21(18), 3596–3603.

Altimiras, F., Uszczynska-Ratajczak, B., Camara, F., Vlasova, A., Palumbo, E., Newhouse, S., … Cogram, P. (2017). Brain transcriptome sequencing of a natural model of alzheimer's disease. *Frontiers in Aging Neuroscience*, 9, 64.

Altschul, S. F., Gish, W., Miller, W., Myers, E. W., & Lipman, D. J. (1990). Basic local alignment search tool. *Journal of Molecular Biology*, 215(3), 403–10.

Andersen, J. R., & Lübberstedt, T. (2003). Functional markers in plants. *Trends in Plant Science*, 8(11), 554–560.

Anderson, J. V, Chevone, B., & Hess, J. L. (1992). Seasonal variation in the antioxidant system of Eastern white pine needles. *Plant Physiology*, 98(2), 501–508.

Andersson, E. (1965). *Cone and seed studies in Norway spruce (Picea abies (L.) Karst.)* (No. 23), pp. 1-278. Available at: < https://pub.epsilon.slu.se/8011/1/SFS023.pdf > [Accessed: 21.11.2017].

Androsiuk, P., Kaczmarek, Z., & Urbaniak, L. (2011). The morphological traits of needles as markers of geographical differentiation in European *Pinus sylvestris* populations, *Dendrobiology*, 65, 3–16.

Argyropoulou, S. (2005). The avifauna of the western Rodopi forests (N. Greece). *Belgian journal of zoology*, 135(2), 261.

Asada, K. (1992). Ascorbate peroxidase – a hydrogen peroxide-scavenging enzyme in plants. *Physiologia Plantarum*. 85(2), 235-24.

Ashrafi, H., Hill, T., Stoffel, K., Kozik, A., Yao, J., Chin-Wo, S., & Van Deynze, A. (2012). *De novo* assembly of the pepper transcriptome (*Capsicum annuum*): a benchmark for *in silico* discovery of SNPs, SSRs and candidate genes. *BMC Genomics*, 13(1), 571.

Bączkiewicz, A, Buczkowska, K, Wachowiak, W. (2005). Anatomical and morphological variability of needles of *Pinus mugo* Turra on different substrata in the Tatra Mountains. *Biological Letters*, 42(1), 21–32.

Baker, H. G. (1972). Seed weight in relation to environmental conditions in California. *Ecology*, 53(6), 997–1010.

Balogh-Brunstad, Z., Keller, C. K., Shi, Z., Wallander, H., & Stipp, S. L. (2017). Ectomycorrhizal fungi and mineral interactions in the rhizosphere of Scots and Red pine seedlings. *Soils*, 1(1), 5.

Barbazuk, W. B., Emrich, S. J., Chen, H. D., Li, L., & Schnable, P. S. (2007). SNP discovery via 454 transcriptome sequencing. *Plant Journal*, 51(5), 910–918.

Barrell, D., Bely, B., Bingley, M., Binns, D., Bower, L., Browne, P., … Zhang, J. (2011). Ongoing and future developments at the Universal Protein Resource. *Nucleic Acids Research*, 39(SUPPL. 1), 214–219.

Beaulieu, J., & Simon, J. (1995). Variation in cone morphology and seed characters in *Pinus strobus* in Quebec. *Canadian Journal of Botany*, 271(Caldwell 1978), 262–271.

Beckman, J. S., Mitton, J. B., The, S., Midland, A., & Jul, N. (2016). Peroxidase allozyme differentiation among successional stands of Ponderosa pine. *American Midland Naturalist*, 112(1), 43–49.

Bedon, F., Bomal, C., Caron, S., Levasseur, C., Boyle, B., Mansfield, S. D., … MacKay, J. (2010). Subgroup 4 R2R3-MYBs in conifer trees: Gene family expansion and contribution to the isoprenoid-and flavonoid-oriented responses. *Journal of Experimental Botany*, 61(14), 3847-3864.

Bedon, F., Grima-Pettenati, J., & Mackay, J. (2007). Conifer R2R3-MYB transcription factors: sequence analyses and gene expression in wood-forming tissues of white spruce (*Picea glauca*). *BMC Plant Biology*, 7(1), 17.

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.

Belletti, P., Ferrazzini, D., Piotti, A., Monteleone, I., & Ducci, F. (2012). Genetic variation and divergence in Scots pine (*Pinus sylvestris* L.) within its natural range in Italy. *European Journal of Forest Research*, 131(4), 1127–1138.

Bennett, K. D. (1984). The post-glacial history of *Pinus sylvestris* in the British Isles. *Quaternary Science Reviews*, 3(2–3), 133–155.

Bennett, K. D. (1995). Post-glacial dynamics of pine (*Pinus sylvestris* L.) and pinewoods in Scotland. *Our Pinewood Heritage*, 23–39.

Bennett, K. D. (1997). *Evolution and ecology: the pace of life*. Cambridge, Cambridge University Press. pp. 1-87.

Bergmann, F., & Hosius, B. (1996). Effects of heavy metal polluted soils on the genetic structure of norway spruce seedling populations. *Water, Air, and Soil Pollution*, 89(3–4), 363–373.

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.

Betancourt, J. L., Schuster, W. S., Mitton, J. B., & Anderson, R. S. (1991). Fossil and genetic history of a pinyon pine (*Pinus edulis*) isolate. *Ecology*, 72(5), 1685-1697.

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

Birks, H. J. B., & Birks, H. H. (1980). *Quaternary palaeoecology*. London, Edward Arnold, pp.1-289.

Birks, H. J. B., & Williams, W. (1983). Late-Quaternary vegetational history of the Inner Hebrides. *Proceedings of the Royal Society of Edinburgh*, 83B, 269–292.

Blick, R. J., Revel, A. T., & Hansen, E. J. (2003). FindGDPs: Identification of primers for labeling microbial transcriptomes for DNA microarray analysis. *Bioinformatics*, 19(13), 1718–1

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.

Bölöni, J., Molnár, Z., & Kun, A. (2011). *Magyarország élőhelyei. Vegetációtípusok leírása és határozója. ÁNÉR.* Vácrátót, MTA Ökológiai és Botanikai Kutatóintézete. pp. 1–442.

Bone, E., & Farres, A. (2001). Trends and rates of micro evolution in plants. In *Microevolution Rate, Pattern, Process*. Dordrecht, Springer. pp. 165-182.

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.

Boratyński, A. (1991). Range of natural distribution. In *Genetics of Scots pine*, Elsevier. pp. 19–30.

Borhidi, A. (2003). *Magyarország növénytársulásai*. Budapest, Akadémiai Kiadó. pp. 1–516.

Bork, P., & Bairoch, A. (1996). Go hunting in sequence databases but watch out for the traps. *Trends in Genetics*, 12(10), 425–427.

Boutros, P. C., & Okey, A. B. (2004). PUNS: Transcriptomic- and genomic- *in silico* PCR for enhanced primer design. *Bioinformatics*, 20(15), 2399–2400.

Bräutigam, A., & Gowik, U. (2010). What can next generation sequencing do for you? Next generation sequencing as a valuable tool in plant research. *Plant Biology*, 12(6), 831–841.

Brenner, S. E., Chothia, C., & Hubbard, T. J. (1997). Population statistics of protein structures: Lessons from structural classifications. *Current Opinion in Structural Biology*, 7(3), 369–376.

Brent, M. R. (2005). Genome annotation past , present , and future : How to define an ORF at each locus Genome annotation past , present , and future : How to define an ORF at each locus. *Genome Research*, 15(12), 1777–1786.

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.

Brookes, A. J. (1999). The essence of SNPs. *Gene*, 234(2), 177–186.

Brzezowski, P., Schlicke, H., Richter, A., Dent, R. M., Niyogi, K. K., & Grimm, B. (2014). The GUN4 protein plays a regulatory role in tetrapyrrole biosynthesis and chloroplast-to-nucleus signalling in *Chlamydomonas reinhardtii*. *Plant Journal*, 79(2), 285–298.

Buchovska, J., Danusevičius, D., Baniulis, D., Stanys, V., Šikšnianienė, J.-B., & Kavaliauskas, D. (2013). The location of the Northern post-glacial refugium of Scots pine based on the mitochondrial DNA markers. *Baltic Forestry*, 19(1), 2–12.

Burczyk, J. (1991). The mating system in a Scots pine clonal seed orchard in Poland. *Annales Des Sciences Forestières*, 48(4), 443–451.

Burdon, J. J. (1987). *Diseases and plant population biology*. Cambridge, Cambridge University Press. pp. 1–204.

Burdon, J. J., Wennstrom, A., Muller, W. J., & Ericson, L. (1994). Spatial patterning in young stands of *Pinus sylvestris* in relation to mortality caused by the snow blight pathogen *Phacidium infestans*. *Oikos*, 71(1), 130–136.

Burns, R. M., & Honkala, B. H. (1990). *Silvics of North America: 2 Hardwoods*, Agriculture Handbook. pp. 148–152.

Cánovas, F. M., Avila, C., Cantón, F. R., Cañas, R. A., & De La Torre, F. (2007). Ammonium assimilation and amino acid metabolism in conifers. *Journal of Experimental Botany*, 58(9),

Cantarel, B. L., Korf, I., Robb, S. M. C., Parra, G., Ross, E., Moore, B., … Yandell, M. (2008). MAKER: An easy-to-use annotation pipeline designed for emerging model organism genomes. *Genome Research*, 18(1), 188–196.

Carlsbecker, A., Sundström, J., Tandre, K., Englund, M., Kvarnheden, A., Johanson, U., & Engström, P. (2003). The DAL10 gene from Norway spruce (*Picea abies*) belongs to a potentially gymnosperm-specific subclass of MADS-box genes and is specifically active in seed cones and pollen cones. *Evolution and Development*, 5(6), 551–561.

Carlsbecker, A., Tandre, K., Johanson, U., Englund, M., & Engström, P. (2004). The MADS-box gene DAL1 is a potential mediator of the juvenile-to-adult transition in Norway spruce (*Picea abies*). *Plant Journal*, 40(4), 546–557.

Carninci, P. (2005). The transcriptional landscape of the mammalian genome. *Science*, 309(5740), 1559–1563.

Castro, J. (1999). Seed mass versus seedling performance in Scots pine: A maternally dependent trait. *New Phytologist*, 144(1), 153–161.

Cendán, C., Sampedro, L., & Zas, R. (2011). The maternal environment determines the timing of germination in *Pinus pinaster*. *Environmental and Experimental Botany*, 94, 66–72.

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

Chalopin, D., & Volff, J.-N. (2017). Analysis of the spotted gar genome suggests absence of causative link between ancestral genome duplication and transposable element diversification in teleost fish. *Journal of Experimental Zoology Part B: Molecular and Developmental Evolution*, (July), 1–9.

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.

Channell, R. (2004). The conservation value of peripheral populations: the supporting science. *Proceedings of the Species at Risk*, 1-17.

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.

Chen, F., Zhang, X., Liu, X., & Zhang, L. (2017). Evolutionary analysis of MIKCc-type MADS-box genes in gymnosperms and angiosperms. *Frontiers in Plant Science*, 8(5), 1–11.

Chen, L., Song, Y., Li, S., Zhang, L., Zou, C., & Yu, D. (2012). The role of WRKY transcription factors in plant abiotic stresses. *Biochimica et Biophysica Acta - Gene Regulatory Mechanisms*, 1819(2), 120–128.

Chen, Y., Zhou, F., Li, G., & Xu, Y. (2009). MUST: A system for identification of miniature inverted-repeat transposable elements and applications to *Anabaena variabilis* and *Haloquadratum walsbyi*. *Gene*, 436(1–2), 1–7.

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.

Cipriano, J., Carvalho, A., Fernandes, C., Gaspar, M. J., Pires, J., & Bento, J. (2013). Evaluation of genetic diversity of Portuguese *Pinus sylvestris* L . populations based on molecular data and inferences about the future use of this germplasm, *Journal of genetics, 92*, 41–48.

Clackson, T., Gussow, D., & Jones, P. T. (1991). PCR: a practical approach. In *PCR, A Practical Approach*, p. 202.

Clark, R. M., Linton, E., Messing, J., & Doebley, J. F. (2004). Pattern of diversity in the genomic region near the maize domestication gene tb1. *Proceedings of the National Academy of Sciences of the United States of America*, 101(3), 700–707.

Cohen, K. M., Finney S. C., Gibbard F. J. (2013). ICS international chronostratigraphic chart. *Episodes*, 36, 199–204.

Colautti, R. I., Maron, J. L., & Barrett, S. C. H. (2009). Common garden comparisons of native and introduced plant populations: Latitudinal clines can obscure evolutionary inferences. *Evolutionary Applications*, 2(2), 187–199.

Curnutt, J. L., Pimm, S. L., & Maurer, B. a. (1996). Population variability of sparrows in space and time. *Oikos*, 76(1), 131–144.

Dahl, E. (1988). *The phytogeography of northern Europe: British Isles, Fennoscandia, and adjacent areas. Vol. 297*. Cambridge, Cambridge university press. p. 297.

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.

Darwin, C. (1867). *The effects of cross and self fertilisation in the vegetable kingdom*. (J. Murray, Ed.). Available at: <https://books.google.hu > [Accessed: 11.12.2017].

De Beeck, M. O., Ruytinx, J., Smits, M. M., Vangronsveld, J., Colpaert, J. V., & Rineau, F. (2015). Belowground fungal communities in pioneer Scots pine stands growing on heavy metal polluted and non-polluted soils. *Soil Biology and Biochemistry*, 86, 58–66.

Decaprio, D., Vinson, J. P., Pearson, M. D., Montgomery, P., Doherty, M., & Galagan, J. E. (2007). Conrad : Gene prediction using conditional random fields. *Genome Research*, 17(9), 1389–1398.

Delpierre, N., Vitasse, Y., Chuine, I., Guillemot, J., Bazot, S., Rutishauser, T., & Rathgeber, C. B. K. (2016). Temperate and boreal forest tree phenology: from organ-scale processes to terrestrial ecosystem models. *Annals of Forest Science*, 73(1), 5–25.

Di Pierro, E. A., Mosca, E., González-Martínez, S. C., Binelli, G., Neale, D. B., & La Porta, N. (2017). Adaptive variation in natural Alpine populations of Norway spruce (*Picea abies* [L.] Karst) at regional scale: Landscape features and altitudinal gradient effects. *Forest Ecology and Management*, 405, 350–359.

Di Stilio, V. S., Melzer, R., & Hall, J. C. (2017). Editorial: A broader view for plant EvoDevo: novel approaches for diverse model systems. *Frontiers in Plant Science*, 8(1), 1–2.

Dickson, A., & Winch, F. E. (1970). *Plantation production of Christmas trees in New York*. New York, New York State College of Agriculture, Cornell University. p. 28.

Dirzo, R., & Harper, J. L. (1982). Experimental studies on slug-plant interactions: IV. The performance of cyanogenic and acyanogenic morphs of *Trifolium repens* in the field. *The Journal of Ecology*, 119–138.

Dixon, R. A., & Paiva, N. L. (1995). Stress-induced phenylpropanoid metabolism. *The Plant Cell*, 7(7), 1085.

Domisch, T., Finer, L., Lehto, T., & Smolander, A. (2002). Effect of soil temperature on nutrient allocation and mycorrhizas in Scots pine seedlings. *Plant and Soil*, 239, 173–185.

Domisch, T., Martz, F., Repo, T., & Rautio, P. (2017). Winter survival of Scots pine seedlings under different snow conditions. *Tree Physiology*, 1–15.

Donini, P., Liu, J.-J., & Ekramoddoullah, A. K. M. (2009). Identification and characterization of the WRKY transcription factor family in *Pinus monticola*. *Genome*, 52(1), 77–88.

Donnelly, K., Cavers, S., Cottrell, J. E., & Ennos, R. A. (2016). Genetic variation for needle traits in Scots pine (*Pinus sylvestris* L). *Tree Genetics & Genomes*, 12(3), 40.

Dörken, V. M., Stephan, G., & Stützel, T. (2010). Morphology and anatomy of anomalous short shoots in *Pinus* (Pinaceae) and their evolutionary meaning. *Feddes Repertorium*, 121(3–4), 133–155.

Dörken, V. M., & Stützel, T. (2012). Morphology, anatomy and vasculature of leaves in *Pinus* (Pinaceae) and its evolutionary meaning. *Flora: Morphology, Distribution, Functional Ecology of Plants*, 207(1), 57–62.

Douglas, C. J. (1996). Phenylpropanoid metabolism and lignin biosynthesis: From weeds to trees. *Trends in Plant Science*, 1(6), 171–178.

Dövényi, Z., Ambrózy, P., Juhász, Á., Marosi, S., Mezősi, G., Michalkó, G., ... & Tiner, T. (2008). Magyarország kistájainak katasztere. Inventory of microregions in Hungary. *OTKA Research Reports.* Available at: <http://real.mtak.hu/1416/1/46401_ZJ1.pdf > [Accessed: 17.12.2017].

Duputié, A., Rutschmann, A., Ronce, O., & Chuine, I. (2015). Phenological plasticity will not help all species adapt to climate change. *Global Change Biology*, 21(8), 3062–3073.

Durkaya, A., Durkaya, B., & Atmaca, S. (2009). Predicting the above-ground biomass of Scots pine (*Pinus sylvestris* L.) Stands in Turkey. *Energy Sources, Part A: Recovery, Utilization, and Environmental Effects*, 32(5), 485–493.

Durrant, H. T., de Rigo, D., & Caudullo, G. (2016). *Pinus sylvestris* in Europe: Distribution, habitat, usage and treats. *European Atlas of Forest Tree Species*, 132–133.

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.

Earl, D., Bradnam, K., St. John, J., Darling, A., Lin, D., Fass, J., … Paten, B. (2011). Assemblathon 1: A competitive assessment of *de novo* short read assembly methods. *Genome Research*, 21(12), 2224–2241.

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.

Edwards, G. E., Nakamoto H., Burnell, J., N., Hatch, M., D. (1985). Pyruvate, pi dikinase and NADP-malate dehydrogenase in C4 photosynthesis : Properties and mechanism of light / dark regulation. *Annual Review of Plant Physiology*, 36, 255–86.

Egea-Cortines, M., Saedler, H., & Sommer, H. (1999). Ternary complex formation between the MADS-box proteins SQUAMOSA, DEFICIENS and GLOBOSA is involved in the control of floral architecture in *Antirrhinum majus*. *EMBO Journal*, 18(19), 5370–5379.

Eguiluz, P. (1984). Geographic variation in needles, cones and seeds of *Pinus tecunumanii* in Guatemala. *Silvae Genetica*, 33(2/3), 72-79.

Ehleringer, J. R., & Monson, R. (1993). Evolutionary and ecologial aspects of photosynthetic pathway variation. *Annual Review of Ecology and Systematics*, 24(1), 411-439.

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

El-Kassaby, Y. A. (1981). Genetic interpretation of malate dehydrogenase isozymes in some conifer species. *The Journal of Heredity*, 72(12), 451–452.

Ellenberg, H. (1988). *Vegetation ecology of central Europe*. Cambridge, Cambridge University Press. p. 710. Available at: < https://books.google.hu> [Accessed: 23.11.2017].

Elsik, C. G., Mackey, A. J., Reese, J. T., Milshina, N. V, Roos, D. S., & Weinstock, G. M. (2007). Creating a honey bee consensus gene set. *Genome Biology*, 8(1), R13.

Emrich, S. J., Barbazuk, W. B., Li, L., & Schnable, P. S. (2007). Gene discovery and annotation using LCM-454 transcriptome sequencing. *Genome Research*, 17(1), 69–73.

ENCODE Project Consortium (2004). The ENCODE (ENCyclopedia of DNA elements) project. *Science*, 306(5696), 636–640.

Endler, J. A. (1986). *Natural selection in the wild (No. 21)*. Princeton, Princeton University Press. pp. 1339-40.

Engelmark, O. (1999). Boreal forest disturbances. In *Ecosystems of the world* , Elsevier, pp. 161–186. Available at: < https://books.google.hu > [Accessed: 19.11.2017].

Ermakova, E., & Zuev, Y. (2017). Interaction of Scots Pine defensin with model membrane by coarse-grained molecular dynamics. *Journal of Membrane Biology*, 250(2), 205–216.

Eronen, M. (1979). The retreat of pineforest in Finnish Lapland since the Holocene climatic optimum: a general discussion with radiocarbon evidence from subfossil pines. *Fennia*, 157, 93–114.

Eronen, M., & Hyvärinen, H. (1982). Subfossil pine dates and pollen diagrams from Northern fennoscandia. *Geologiska Föreningen i Stockholm Förhandlingar*, 103(4), 437–445.

Ersoz, E. S., Wright, M. H., González-Martínez, S. C., Langley, C. H., & Neale, D. B. (2010). Evolution of disease response genes in loblolly pine: Insights from candidate genes. *PLoS ONE*, 5(12).

Etterson, J. R. (2001). Constraint to adaptive evolution in response to global warming. *Science*, 294(5540), 151–154.

Eulgem, T., & Somssich, I. E. (2007). Networks of WRKY transcription factors in defense signaling. *Current Opinion in Plant Biology*, 10(4), 366–371.

Fady, B., Aravanopoulos, F. A., Alizoti, P., Mátyás, C., von Wühlisch, G., Westergren, M., … Zlatanov, T. (2016). Evolution-based approach needed for the conservation and silviculture of peripheral forest tree populations. *Forest Ecology and Management*, 375, 66–75.

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.

Feichtinger, L. M., Siegwolf, R. T., Gessler, A., Buchmann, N., Lévesque, M., & Rigling, A. (2017). Plasticity in gas-exchange physiology of mature Scots pine and European larch drive short-and long-term adjustments to changes in water availability. *Plant, Cell & Environment*. 40(9), 1972-1983.

Fenner, M. (1992). Environmental influences on seed size and composition. *Horticultural Reviews*, 13, 183–213.

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

Fowler, D., Cape, J. N., Coyle, M., Flechard, C., Kuylenstierna, J., Hicks, K., … Stevenson, D. (1999). The global exposure of forests to air pollutants. *Water, Air, and Soil Pollution*, 116(1–2), 5–32.

Franke, A. K., Bräuning, A., Timonen, M., & Rautio, P. (2017). Growth response of Scots pines in polar-alpine tree-line to a warming climate. *Forest Ecology and Management*, 399, 94–107.

Friend, A. D. & Woodward, F. (1990). Evolutionary and ecophysiological responses of mountain plants to the growing season environment. *Advances in Ecological Research*, 20, 59–124.

Fries, A. (2017). Damage by pathogens and insects to Scots pine and lodgepole pine 25 years after reciprocal plantings in Canada and Sweden. *Scandinavian Journal of Forest Research*, 32(6), 459–472.

Fritz, R. S., Moulia, C., & Newcombe, G. (1999). Resistance of hybrid plants and animals to herbivores, pathogens, and parasites. *Annual Review of Ecology and Systematics*, 30(1), 565–591.

Froux, F., Ducrey, M., Epron, D., & Dreyer, E. (2004). Seasonal variations and acclimation potential of the thermostability of photochemistry in four Mediterranean conifers. *Annals of Forest Science*, 61(3), 235–241.

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

Furnier, G. R., & Adams, W. T. (1986). Geographic patterns of allozyme variation in Jeffrey pine. *American Journal of Botany*, 1009–1015.

Galiano, L., Martínez-Vilalta, J., & Lloret, F. (2010). Drought-induced multifactor decline of Scots pine in the Pyrenees and potential vegetation change by the expansion of co-occurring oak species. *Ecosystems*, 13(7), 978–991.

Gao, Y., Markkanen, T., Aurela, M., Mammarella, I., Thum, T., Tsuruta, A., … Aalto, T. (2017). Response of water use efficiency to summer drought in a boreal Scots pine forest in Finland. *Biogeosciences*, 14(18), 4409–4422.

Garcia-Ramos, G., & Kirkpatrick, M. (1997). Genetic models of adaptation and gene flow in peripheral populations. *Evolution*, 51(1), 21.

García Gil, M. R., Floran, V., Östlund, L., Mullin, T. J. T., & Andersson Gull, B. (2015). Genetic diversity and inbreeding in natural and managed populations of Scots pine. *Tree Genetics and Genomes*, 11(2). 28.

Garg, R., Shankar, R., Thakkar, B., Kudapa, H., Krishnamurthy, L., Mantri, N., … Jain, M. (2016). Transcriptome analyses reveal genotype- and developmental stage-specific molecular responses to drought and salinity stresses in chickpea. *Scientific Reports*, 6(1), 19228.

Geber, M. A., & Griffen, L. R. (2003). Inheritance and natural selection on functional traits. *International Journal of Plant Sciences*, 164(S3), S21–S42.

Gencsi, L. V. R. (1992). *Dendrológia. (Dendrology)*. Budapest, Mezőgazda Kiadó. pp. 1-143.

Geras'kin, S. A., Dikareva, N. S., Oudalova, A. A., Vasil'ev, D. V, & Volkova, P. Y. (2016). The consequences of chronic radiation exposure of Scots pine in the remote period after the chernobyl accident. *Russian Journal of Ecology*, 47(1), 26–38.

Gernandt, D. S., López, G. G., García, S. O., & Liston, A. (2005). Phylogeny and classification of *Pinus*. *Taxon*, 54(1), 29–42.

Ghimire, R. P., Kivimäenpää, M., Kasurinen, A., Häikiö, E., Holopainen, T., & Holopainen, J. K. (2017). Herbivore-induced BVOC emissions of Scots pine under warming, elevated ozone and increased nitrogen availability in an open-field exposure. *Agricultural and Forest Meteorology*, 242(December 2016), 21–32.

Gibson, J. P., & Hamrick, J. L. (1991). Heterogeneity in pollen allele frequencies among cones, whorls, and trees of table mountain pine (*Pinus pungens*). *American Journal of Botany*, 1244–1251.

Giegerich, R., Meyer, F., & Schleiermacher, C. (1996). GeneFisher-software support for the detection of postulated genes. *ISMB*, 68-77.

Giertych, M., & Mátyás, C. (Eds:). (2017). *Genetics of Scots pine (Vol.3)*. Amsterdam, Elsevier. pp. 5-276.

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.

Gillespie, J. H., & Turelli, M. (1989). Genotype-environment interactions and the maintenance of polygenic variation. *Genetics*, 121(1), 129–138.

Ginwal, H. S., Phartyal, S. S., Rawat, P. S., & Srivastava, R. L. (2005). Seed source variation in morphology, germination and seedling growth of *Jatropha curcas* Linn. in Central India. *Silvae Genetica*, 54(2), 76–80.

Gömöry, D., Bruchánik, R., & Longauer, R. (2003). Fertility variation and flowering asynchrony in *Pinus sylvestris*: Consequences for the genetic structure of progeny in seed orchards. *Forest Ecology and Management*, 174(1), 117–126.

González de Andrés, E., Camarero, J. J., Blanco, J. A., Imbert, J. B., Lo, Y. H., Sangüesa-Barreda, G., & Castillo, F. J. (2018). Tree-to-tree competition in mixed European beech-Scots pine forests has different impacts on growth and water-use efficiency depending on site conditions. *Journal of Ecology*, 106(1), 59-75.

Grabherr, M. G., Haas, B. J., Yassour, M., Levin, J. Z., Thompson, D. A., Amit, I., … Regev, A. (2011). Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nature Biotechnology*, 29(7), 644–652.

Grant, M. C., & Mitton, J. B. (1977). Genetic differentiation among growth forms of Engelmann spruce and subalpine fir at tree line. *Arctic and Alpine Research*, 259–263.

Green, P., Ewing, B., Hillier, L., Wendl, M. C. (1998). Base-calling of automated sequencer traces using Phred. I. Accuracy assessment. *Genome research*, 8(3), 175-185.

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

Gregorius, H. (1989). The importance of genetic multiplicity for tolerance of atmospheric pollution. *Genetic Effects of Air Pollutants in Forest Tree Populations (Scholz F., Gregorius HR, Rudin D., Eds). Springer*, 163–172.

Grenville-Briggs, L. J., Kushwaha, S. K., Cleary, M. R., Witzell, J., Savenkov, E. I., Whisson, S. C., … Vetukuri, R. R. (2017). Draft genome of the oomycete pathogen *Phytophthora cactorum* strain LV007 isolated from European beech (*Fagus sylvatica*). *Genomics Data*, 12(5), 155–156.

Gross, S. S., Do, C. B., Sirota, M., & Batzoglou, S. (2007). CONTRAST: a discriminative, phylogeny-free approach to multiple informant *de novo* gene prediction. *Genome Biology*, 8(12), R269.

Haas, B. J., Zeng, Q., Pearson, M. D., Cuomo, C. A., & Wortman, J. R. (2011). Approaches to fungal genome annotation. *Mycology*, 2(3), 118–141.

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.

Haig, D. (1996). The pea and the coconut: seed size in safe sites. *Trends in Ecology & Evolution*, 11(1), 1–2.

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.

Han, Y., & Wessler, S. R. (2010). MITE-Hunter: A program for discovering miniature inverted-repeat transposable elements from genomic sequences. *Nucleic Acids Research*, 38(22), 1–8.

Hansen, O. K. (2008). Mating patterns, genetic composition and diversity levels in two seed orchards with few clones-Impact on planting crop. *Forest Ecology and Management*. 256(5), 1167-1177.

Hara, K., Yagi, M., Kusano, T., & Sano, H. (2000). Rapid systemic accumulation of transcripts encoding a tobacco WRKY transcription factor upon wounding. *Molecular & General Genetics : MGG*, 263(1), 30–37.

Hari, P., Aakala, T., Aalto, J., Bäck, J., Hollmén, J., Jõgiste, K., ... & Nikinmaa, E. (2017). Newtonian boreal forest ecology: The Scots pine ecosystem as an example. *PloS One*, 12(6), e0177927.

Harvey, P. H., & Pagel, M. (1991). *The comparative method in evolutionary biology* (Vol. 239). Oxford, Oxford university press. pp. 1-35. Available at: <http://www.sbs.utexas.edu/bio386/Readings/Text/Harvey.Ch.1.pdf> [Accessed: 21.01.2017].

Hatchell, G. E., Ralston, C. W., & Foil, R. R. (1970). Soil disturbances in logging: effects on soil characteristics and growth of loblolly pine in the Atlantic Coastal Plain. *Journal of Forestry*, 68(12), 772–5.

Hay, M. E., & Fenical, W. (1988). Marine plant-herbivore interactions: the ecology of chemical defense. *Annual Review of Ecology and Systematics*, 19(1), 111–145.

Heit, C. E. (1969). Propagation from seed: 19. Testing and growing Scotch pine seeds from different sources. *American Nurseryman*, 129(7), 10–15.

Helyar, S. J., Hemmer-Hansen, J., Bekkevold, D., Taylor, M. I., Ogden, R., Limborg, M. T., … Nielsen, E. E. (2011). Application of SNPs for population genetics of nonmodel organisms: New opportunities and challenges. *Molecular Ecology Resources*, 11(SUPPL. 1), 123–136.

Hengeveld, R. O. B., & Haeck, J. (1982). The distribution of abundance. I. Measurements. *Journal of Biogeography*, 303–316.

Hewitt, G. M. (2004). Genetic consequences of climatic oscillations in the Quaternary. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 359(1442), 183–195.

Hosius, B., Bergmann, F., Konnert, M., & Henkel, W. (2000). A concept for seed orchards based on isoenzyme gene markers. *Forest Ecology and Management*, 131(1-3), 143-152.

Huang, X., Yin, C., Duan, B., & Li, C. (2008). Interactions between drought and shade on growth and physiological traits in two *Populus cathayana* populations. *Canadian Journal of Forest Research*, 38(7), 1877–1887.

Human Genome Sequencing Consortium, I. (2004). Finishing the euchromatic sequence of the human genome. *Nature*, 431(7011), 931–945.

Huntley, B., & Birks, H. J. B. (1983). *An atlas of past and present pollen maps for Europe, 0-13,000 years ago.* Cambridge, Cambridge University Press. p. 667.

Huttunen, A., Huttunen, R. L., Vasari, Y., Panovska, H., & Bozilova, E. (1992). Late-glacial and Holocene history of flora and vegetation in the Western Rhodopes Mountains, Bulgaria. *Acta Botanica Fennica*, 144, 63–80.

Hyun, T. K., Lee, S., Kumar, D., Rim, Y., Kumar, R., Lee, S. Y., … Kim, J. Y. (2014). RNA-seq analysis of *Rubus idaeus* cv. Nova: transcriptome sequencing and *de novo* assembly for subsequent functional genomics approaches. *Plant Cell Reports*, 33(10), 1617–1628.

Icgen, Y., Kaya, Z., Cengel, B., Velioğlu, E., Öztürk, H., & Önde, S. (2006). Potential impact of forest management and tree improvement on genetic diversity of Turkish red pine (*Pinus brutia* Ten.) plantations in Turkey. *Forest Ecology and Management*, 225(1), 328–336.

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.

Jackman, S. D., Yeo, S., Coombe, L., & Warren, R. L. (2017). Supplemental material of ABySS 2 . 0 : Resource-efficient assembly of large genomes using a bloom filter effect of bloom filter false positive rate. *Genome research*, 27(5), 768–777.

James, J. C., Grace, J., & Hoad, S. P. (1994). Growth and photosynthesis of *Pinus sylvestris* at its altitudinal limit in Scotland. *Journal of Ecology*, 82(2), 297–306.

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.

Jankowski, A., Wyka, T. P., Żytkowiak, R., Nihlgård, B., Reich, P. B., & Oleksyn, J. (2017). [In Press] Cold adaptation drives variability in needle structure and anatomy in *Pinus sylvestris* L. along a 1900 km temperate-boreal transect. *Functional Ecology*. 31(12), 2212-2223.

Jasińska, A. K., Wachowiak, W., Muchewicz, E., Boratyńska, K., Montserrat, J. M., & Boratyński, A. (2010). Cryptic hybrids between *Pinus uncinata* and *P. sylvestris*. *Botanical Journal of the Linnean Society*, 163(4), 473–485.

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.

Je, S. M., Son, S. G., Woo, S. Y., Byun, K. O., & Kim, C. S. (2006). Photosynthesis and chlorophyll contents of *Chloranthus glaber* under different shading treatments. *Korean Journal of Agricultural and Forest Meteorology*, 8(2), 54–60.

Ji, M, Zhang, Q, Deng, J, Zhang, X, Wang, Z. (2011). Intra-versus inter-population variation of cone and seed morphological traits of *Pinus tabulaeformis* Carr. in northern China: impact of climate-related conditions. *Polish Journal of Ecology*, 59(4), 717–727.

Johnson, M. A., & Croteau, R. (1987). Biochemistry of conifer resistance to bark beetles and their fungal symbionts. *Ecology and Metabolism of Plant Lipids. ACS Symposium Series*, 325, 76–91.

Kapitonov, V. V., & Jurka, J. (2003). A novel class of SINE elements derived from 5S rRNA. *Molecular Biology and Evolution*, 20(5), 694–702.

Kapitonov, V. V., & Jurka, J. (2008). A universal classification of eukaryotic transposable elements implemented in Repbase. *Nature Reviews Genetics*, 9(5), 411–412.

Kapustin, Y., Souvorov, A., Tatusova, T., & Lipman, D. (2008). Splign: algorithms for computing spliced alignments with identification of paralogs. *Biology Direct*, 3(1), 20.

Kaushik, N., Kumar, K., Kumar, S., Kaushik, N., & Roy, S. (2007). Genetic variability and divergence studies in seed traits and oil content of Jatropha (*Jatropha curcas* L.) accessions. *Biomass and Bioenergy*, 31(7), 497–502.

Kawecki, T. J., & Ebert, D. (2004). Conceptual issues in local adaptation. *Ecology Letters*, 7, 1225–1241.

Keeley, J. E., & Zedler, P. H. (2000). Evolution of life histories in *Pinus*. In *Ecology and biogeography of Pinus*, Cambridge, Cambridge University Press. pp. 219–252.

Kelly, D. L., & Connolly, A. (2000). A review of the plant communities associated with Scots pine (*Pinus sylvestris* L .) in Europe , and an evaluation of putative indicator / specialist species. *Forest Systems*, 9(S1), 15–39.

Kent, W. J., Sugnet, C. W., Furey, T. S., & Roskin, K. M. (2002). The Human Genome Browser at UCSC W. *Journal of Medicinal Chemistry*, 19(10), 1228–31.

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

Kirkpatrick, M., & Barton, N. H. (1997). Evolution of a species' range. *American Naturalist*, 150(1), 1–23.

Kivimäenpää, M, Sutinen, S, Valolahti, H, Häikiö, E, Riikonen, J, Kasurinen, A, Holopainen, T. (2016). Warming and elevated ozone differently modify needle anatomy of Norway spruce (*Picea abies*) and Scots pine (*Pinus sylvestris*). *Canadian Journal of Forest Research*, 47(4), 488–499.

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.

Kohn, M. H., Pelz, H.-J., & Wayne, R. K. (2000). Natural selection mapping of the warfarin-resistance gene. *Proceedings of the National Academy of Sciences*, 97(14), 7911–7915.

Kolari, P., Chan, T., Porcar-Castell, A., BÃ¤ck, J., Nikinmaa, E., & Juurola, E. (2014). Field and controlled environment measurements show strong seasonal acclimation in photosynthesis and respiration potential in boreal Scots pine. *Frontiers in Plant Science*, 5(12), 1–12.

Korf, I. (2004). Gene finding in novel genomes. *BMC Bioinformatics*, 5, 59.

Körner, C. (2007). The use of " altitude " in ecological research. *Trends in Ecology & Evolution*, 22(11), 569–574.

Korona, R. (1996). Adaptation to structurally different environments. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 263(1377), 1665 LP-1669.

Kovalchuk, A., Raffaello, T., Jaber, E., Keriö, S., Ghimire, R., Lorenz, W. W., ... & Asiegbu, F. O. (2015). Activation of defence pathways in Scots pine bark after feeding by pine weevil (*Hylobius abietis*). *BMC Genomics*, 16(1), 352.

Kremenetski, C. V, Tarasov, P. E., & Cherkinsky, A. E. (1997). Postglacial development of Kazakhstan pine forests. *Geographie Physique et Quaternaire*, 51(3), 391–404.

Kruckeberg, A. R. (1992). Plant life of western North American ultramafics. In *The ecology of areas with serpentinized rocks*. Amsterdam, Springer Netherlands. pp. 31–73

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.

Krutovsky, K. V. (2006). From population genetics to population genomics of forest trees: integrated population genomics approach. *Russian journal of genetics*, 42(10), 1022–7954.

Kujala, S. T., Knürr, T., Kärkkäinen, K., Neale, D. B., Sillanpää, M. J., & Savolainen, O. (2017). Genetic heterogeneity underlying variation in a locally adaptive clinal trait in *Pinus sylvestris* revealed by a Bayesian multipopulation analysis. *Heredity*, 118(5), 413–423.

Kullman, L., & Kjällgren, L. (2006). Holocene pine tree-line evolution in the Swedish Scandes: Recent tree-line rise and climate change in a long-term perspective. *Boreas*, 35(1), 159–168.

Kulmala, L., del Rosario Dominguez Carrasco, M., Heinonsalo, J. (2017). The allocation of carbon in most common boreal dwarf shrubs compared to Scots pine seedlings. *EGU General Assembly Conference Abstracts*, 19, 12019.

Kwaśna, H., Mazur, A., Kuźmiński, R., Jaszczak, R., Turski, M., Behnke-Borowczyk, J., ... & Łakomy, P. (2017). Abundance and diversity of wood-decay fungi in managed and unmanaged stands in a Scots pine forest in western Poland. *Forest Ecology and Management*, 400, 438–446.

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.

Lande, R. & Shannon, S. (1996). The role of genetic variation in adaptation and population persistence in a changing environment. *Evolution*, 50(1), 434–437.

Langlet, O. (1959). A cline or not a cline–a question of Scots pine. *Silvae Genetica*, 13–22.

Leishman, M. R., & Westoby, M. (1994). The role of seed size in seedling establishment in dry soil conditions - Experimental evidence from semi-arid species. *Journal of Ecology*, 82(2), 249–258.

Lennon, J. J., Turner, J. R., & Connell, D. (1997). A metapopulation model of species boundaries. *Oikos*, 486-502.

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

Lesica, P., & Allendorf, F. W. (1995). When are peripheral populations valuable for conservation? *Conservation Biology*, 9(4), 753–760.

Leuschner, C., and Ellenberg, H. (2017). Pure and mixed coniferous forests. In *Ecology of Central European Forests*, pp. 521–606.

Lexa, M., Horak, J., & Brzobohaty, B. (2001). Virtual PCR. *Bioinformatics*, 17(2), 192–193.

Lhotáková, Z., Albrechtová, J., Malenovský, Z., Rock, B. N., Polák, T., & Cudlín, P. (2007). Does the azimuth orientation of Norway spruce (*Picea abies* / L ./ Karst.) branches within sunlit crown part influence the heterogeneity of biochemical , structural and spectral characteristics of needles ? *Environmental and Experimental Botany*, 59(3), 283–292.

Li, R., Zhu, H., Ruan, J., Qian, W., Fang, X., Shi, Z., … Wang, J. J. (2010). *De novo* assembly of human genomes with massively parallel short read sequencing. *Genome Research*, 20(2), 265–272.

Li, W. H., & Sadler, L. A. (1991). Low nucleotide diversity in man. *Genetics*, 129(2), 513–523.

Lind, B. M., Friedline, C. J., Wegrzyn, J. L., Maloney, P. E., Vogler, D. R., Neale, D. B., & Eckert, A. J. (2017). Water availability drives signatures of local adaptation in whitebark pine (*Pinus albicaulis* Engelm.) across fine spatial scales of the Lake Tahoe Basin, USA. *Molecular Ecology*, 26(12), 3168–3185.

Linhart, Y. B. (1988). Intrapopulation differentiation in annual plants. III. The contrasting effects of intra-and interspecific competition. *Evolution*, 1047-1064.

Linhart, Y. B. (1989). Interactions between genetic and ecological patchiness in forest trees and their dependent species. *Evolutionary Ecology of Plants*, 393–430.

Linhart, Y. B., & Grant, M. C. (1996). Evolutionary significance of local genetic differentiation in plants. *Annual Review of Ecology and Systematics*, 27(1), 237–277.

Liu, B., Hong, Y. B., Zhang, Y. F., Li, X. H., Huang, L., Zhang, H. J., … Song, F. M. (2014). Tomato WRKY transcriptional factor SlDRW1 is required for disease resistance against *Botrytis cinerea* and tolerance to oxidative stress. *Plant Science*, 227, 145–156.

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

Lowe, J. J. (1992). Late glacial and early holocene lake-sediments from the Northern Apennines, Italy - Pollen stratigraphy and radiocarbon dating. *Boreas*, 21(3), 193–208.

Magyari, E. K., Kuneš, P., Jakab, G., Sümegi, P., Pelánková, B., Schäbitz, F., ... & Chytrý, M. (2014a). Late Pleniglacial vegetation in eastern-central Europe : are there modern analogues in Siberia ? *Quaternary Science Reviews*, 95, 60–79.

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.

Magyari, E. K. (2011). Late quaternary vegetation history in the Hortobágy steppe and middle Tisza floodplain , NE Hungary. *Studia Botanica Hungarica*, 42(185), e203.

Mallik, A. U., Hossain, M. K., & Lamb, E. G. (2008). Species and spacing effects of northern conifers on forest productivity and soil chemistry in a 50-year-old common garden experiment. *Journal of Forestry*, 106(2), 83–90.

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

Marcysiak, K. (2004). Interpopulational variability of *Pinus uncinata* Ramond ex DC . in Lam . & DC .(Pinaceae) cone characters. *Dendrobiology*, 51, 43–51.

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.

Marshall, D. L. (1986). Effect of seed size on seedling success in three species of *Sesbania* (Fabaceae). *American Journal of Botany*, 73(4), 457–464.

Martin, D. M., Fäldt, J., & Bohlmann, J. (2004). Functional characterization of nine Norway spruceTPS genes and evolution of gymnosperm terpene synthases of the TPS-d subfamily. *Plant Physiology*, 135(8), 1908–1927.

Martínez-Sancho, E., Dorado-Liñán, I., Gutiérrez-Merino, E., Matiu, M., Helle, G., Heinrich, I., & Menzel, A. (2017). Increased water-use efficiency translates into contrasting growth patterns of Scots pine and sessile oak at their southern distribution limits. *Global Change Biology*. 24(3), 1012-1028.

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.

Matyas, C. (1996). Climatic adaptation of trees: rediscovering provenance tests. *Euphytica*, 92(1), 45–54.

Mátyás, C. (1994). Modelling climate change effects with provenance test data. *Tree Physiology*, 14(7-8–9), 797–804.

Mayer, M. S., Soltis, P. S., & Soltis, D. E. (2017). The evolution of the *Streptanthus glandulosus* complex (Cruciferae): genetic divergence and gene flow in serpentine endemics. *Botanical Society of America* , 81(10), 1288–1299.

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

McKersie, B. D., Bowley, S. R., & Jones, K. S. (1999). Winter survival of transgenic alfalfa overexpressing superoxide dismutase1. *Plant Physiology*, 119(3), 839–848.

Mejnartowicz, L. (1979). Genetic variation in some isoenzyme loci in Scots pine (*Pinus sylvestris* L.) populations. *Arboretum Kornickie*, 24(9), 1–1.

Merry, R., Jerrard, J., Frebault, J., & Verhoeven, A. (2017). A comparison of pine and spruce in recovery from winter stress; changes in recovery kinetics, and the abundance and phosphorylation status of photosynthetic proteins during winter. *Tree Physiology*, 1–12.

Messenguy, F., & Dubois, E. (2003). Role of MADS box proteins and their cofactors in combinatorial control of gene expression and cell development. *Gene*, 316(1–2), 1–21.

Messner, B., Boll, M., & Berndt, J. (1991). L-Phenylalanine ammonia-lyase in suspension culture cells of spruce (*Picea abies*), *Plant cell, tissue and organ culture*, 27(3), 267–274.

Michaels, S. D. (1999). Flowering locus C encodes a novel MADS domain protein that acts as a repressor of flowering. *The Plant Cell Online*, 11(5), 949–956.

Miguel, A., de Vega-Bartol, J., Marum, L., Chaves, I., Santo, T., Leitão, J., … Miguel, C. M. (2015). Characterization of the cork oak transcriptome dynamics during acorn development. *BMC Plant Biology*, 15(1), 158.

Millar, C. I. (1993). Impact of the Eocene on the evolution of *Pinus* L. *Annals of the Missouri Botanical Garden*, 471-498.

Miller, B., Madilao, L. L., Ralph, S., & Bohlmann, J. (2005). Insect-induced conifer defense. White pine weevil and methyl jasmonate induce traumatic resinosis, *de novo* formed volatile emissions, and accumulation of terpenoid synthase and putative octadecanoid pathway transcripts in Sitka spruce. *Plant Physiology*, 137(1), 369–382.

Miller, S. P., & Cumming, J. R. (2000). Effects of serpentine soil factors on Virginia pine (*Pinus virginiana*) seedlings. *Tree Physiology*, 20(16), 1129–1135.

Min, B., Grigoriev, I. V., & Choi, I. G. (2017). FunGAP: fungal genome annotation pipeline using evidence-based gene model evaluation. *Bioinformatics*, 33(18), 2936–2937.

Mirov, N. T., & Stanley, R. G. (1959). The pine tree. *Annual Review of Plant Physiology*, 10(1), 223–238.

Mitka, J., Bąba, W., & Szczepanek, K. (2014). Putative forest glacial refugia in the western and eastern Carpathians. *Modern Phytomorphology*, 5(8), 6–9.

Moles, A. T., Ackerly, D. D., Webb, C. O., Tweddle, J. C., Dickie, J. B., Pitman, A. J., & Westoby, M. (2005). Factors that shape seed mass evolution. *Proceedings of the National Academy of Sciences of the United States of America*, 102(30), 10540–10544.

Moloney, K. A., Holzapfel, C., Tielbörger, K., Jeltsch, F., & Schurr, F. M. (2009). Rethinking the common garden in invasion research. *Perspectives in Plant Ecology, Evolution and Systematics*, 11(4), 311–320.

Mõttus, M., Clemente, R. H., Perheentupa, V., & Markiet, V. (2017). *In situ* measurement of Scots pine needle PRI. *Plant Methods*, 13(1), 35.

Mouradov, A., Glassick, T., & Teasdale, R. (1997). Isolation and characterization of a new MADS-box cDNA from *Pinus radiata*. *Plant Physiology*, 113(664), 96-002.

Mouradov, A., Glassick, T., Vivian-Smith, A., & Teasdale, R. (1996). Isolation of a MADS box gene family from *Pinus radiata*. *Plant Physiology*, 110, 1047.

Mouradov, A., Glassick, T. V, Hamdorf, B. A., Murphy, L. C., Marla, S. S., Yang, Y., & Teasdale, R. D. (1998). Family of MADS-box genes expressed early in male and female reproductive structures of montrey pine. *Plant Physiology*, 117, 55–61.

Mueller, K. E., Eissenstat, D. M., Hobbie, S. E., Oleksyn, J., Jagodzinski, A. M., Reich, P. B., … Chorover, J. (2012). Tree species effects on coupled cycles of carbon, nitrogen, and acidity in mineral soils at a common garden experiment. *Biogeochemistry*, 111(1–3), 601–614.

Navie, S. C., Panetta, F. D., McFadyen, R. E., & Adkins, S. W. (1998). Behaviour of buried and surface-sown seeds of *Parthenium hysterophorus*. *Weed Research*, 38(5), 335–341.

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.

Neale, D. B. (2007). Genomics to tree breeding and forest health. *Current Opinion in Genetics and Development*, 17(6), 539–544.

Nei, M., & Tajima, F. (1981). DNA polymorphism detectable by restriction endonucleases. *Genetics*, 97(1), 145–163.

Niinemets, Ü, Lukjanova, A, Turnbull, M.H., Sparrow, A. (2007). Plasticity in mesophyll volume fraction modulates light-acclimation in needle photosynthesis in two pines. *Tree Physiology*, 27(8), 1137–1151.

Niinemets, U., Ellsworth, D. S., Lukjanova, a, & Tobias, M. (2001). Site fertility and the morphological and photosynthetic acclimation of *Pinus sylvestris* needles to light. *Tree Physiology*, 21(17), 1231–1244.

Notivol, E., García-Gil, M. R., Alía, R., & Savolainen, O. (2007). Genetic variation of growth rhythm traits in the limits of a latitudinal cline in Scots pine. *Canadian Journal of Forest Research*, 37(3), 540–551.

Mirov, N. T. (1967). *The genus Pinus*. New York ,The Ronald Press Company. pp. 1-602.

Núñez-Acuña, G., & Gallardo-Escárate, C. (2013). Identification of immune-related SNPs in the transcriptome of *Mytilus chilensis* through high-throughput sequencing. *Fish and Shellfish Immunology*, 35(6), 1899–1905.

Nystrand, O., & Granström, A. (2000). Predation on *Pinus sylvestris* seeds and juvenile seedlings in Swedish boreal forest in relation to stand disturbance by logging. *Journal of Applied Ecology*, 37(3), 449–463.

O'Reilly-Wapstra, J. M., Moore, B. D., Brewer, M., Beaton, J., Sim, D., Wiggins, N. L., & Iason, G. R. (2014). *Pinus sylvestris* sapling growth and recovery from mammalian browsing. *Forest Ecology and Management*, 325, 18–25.

Oleksyn, J. (1988). Height growth of different European Scots pine *Pinus sylvestris* L. provenances in a heavily polluted and a control environment. *Environmental Pollution (Barking, Essex : 1987)*, 55(4), 289–99.

Oleksyn, J., & Bialobok, S. (1986). Net photosynthesis, dark respiration and susceptibility to air pollution of 20 European provenances of scots pine *Pinus sylvestris* L. *Environmental Pollution. Series A, Ecological and Biological*, 40(4), 287–302.

Oleksyn, J., Prusglowacki, W., Giertych, M., & Reich, P. B. (1994). Relation between genetic diversity and pollution impact in a 1912 Experiment with East European *Pinus sylvestris* provenances. *Canadian Journal of Forest Research-Revue Canadienne De Recherche Forestiere*, 24(12), 2390–2394.

Oleksyn, J., Reich, P. B., Zytkowiak, R., Karolewski, P., & Tjoelker, M. G. (2003). Nutrient conservation increases with latitude of origin in European *Pinus sylvestris* populations. *Oecologia*, 136(2), 220-235.

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

Olson, M. S., Robertson, A. L., Takebayashi, N., Silim, S., Schroeder, W. R., & Tiffin, P. (2010). Nucleotide diversity and linkage disequilibrium in balsam poplar (*Populus balsamifera*), 186(2), 526–536.

Osakabe, Y., Osakabe, K., & Chiang, V. L. (2009). Characterization of the tissue-specific expression of phenylalanine ammonia-lyase gene promoter from loblolly pine (*Pinus taeda*) in *Nicotiana tabacum*. *Plant Cell Reports*, 28(9), 1309–1317.

Ostonen, I., Lohmus, K., Helmisaari, H.-S., Truu, J., & Meel, S. (2007). Fine root morphological adaptations in Scots pine, Norway spruce and silver birch along a latitudinal gradient in boreal forests. *Tree Physiology*, 27(11), 1627–1634.

Ovington, J. (1957). Dry-matter production by *Pinus sylvestris* L. *Annals of Botany*, 21(2), 287–314.

Pakharkova, N. V., Heilmeier, H., Gette, I. G., Andreeva, E. B., Grachev, A. M., Gaevskiy, N. A., & Grigoriev, Y. S. (2016). Quantitative characteristics of the phases of winter dormancy of conifer species at a site in Central Siberia. *Revista Brasileira de Botanica*, 39(4), 1005–1014.

Pamilo, P., & Savolainen, O. (1999). Post-glacial colonization, drift, local selection and conservation value of populations: A northern perspective. *Hereditas*. 130(3), 229-238.

Papanastasis, V. P., Mantzanas, K., Dini-Papanastasi, O., & Ispikoudis, I. (2009). Traditional agroforestry systems and their evolution in Greece. In: *Agroforestry in Europe: Current Status and Future Prospects*, 89–109.

Parchman, T. L., Geist, K. S., Grahnen, J. A., Benkman, C. W., & Buerkle, C. A. (2010). Transcriptome sequencing in an ecologically important tree species: assembly, annotation, and marker discovery. *BMC Genomics*, 11(1), 180.

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.

Parducci, L., Jorgensen, T., Tollefsrud, M. M., Elverland, E., Alm, T., Fontana, S. L., … Willerslev, E. (2012). Glacial survival of boreal trees in Northern Scandinavia. *Science*, 335(6072), 1083–1086.

Park, C. Y., Lee, J. H., Yoo, J. H., Moon, B. C., Choi, M. S., Kang, Y. H., … Cho, M. J. (2005). WRKY group IId transcription factors interact with calmodulin. *FEBS Letters*, 579(6), 1545–1550.

Parmesan, C. (2006). Ecological and evolutionary responses to recent climate change. *Annual Review of Ecology, Evolution, and Systematics*, 37(1), 637–669.

Pawlaczyk, E. M., Bobowicz, M. A., & Korczyk, A. F. (2010). Variability of three natural populations of *Pinus sylvestris* L. from different sites of Białowieża Primeval Forest estimated using needle traits. *Forest Research Papers*, 71(1), 83–92.

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.

Pears, N. V. (1968). The natural altitudinal limit of forest in the Scottish Grampians. *Oikos*, 71-80. 19(1), 71–80.

Pèrez-Obiol, R., & Julià, R. (1994). Climatic change on the Iberian peninsula recorded in a 30,000-Yr Pollen record from lake Banyoles. *Quaternary Research*, 41(1), 91-98.

Persson, H. (1980). Fine-root dynamics in a Scots pine stand with and without near-optimum nutrient and water regimes. *Acta Phytogeographica Suecica*, 68, 101–110.

Petit, R. J. (2003). Glacial refugia: hotspots but not melting pots of genetic diversity. *Science*, 300(5625), 1563–1565.

Petit, R. J., Hu, F. S., & Dick, C. W. (2008). Forests of the past: a window to future changes. *Science*, 320(5882), 1450–1452.

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

Polle, A., Mössnang, M., Von Schönborn, A., Sladkovic, R., & Rennenberg, H. (1992). Field studies on Norway spruce trees at high altitudes. *New Phytologist*, 121(1), 89–99.

Polle, A., & Rennenberg, H. (1991). Superoxide dismutase activity in needles of Scots pine and Norway spruce under field and chamber conditions: lack of ozone effects. *New Phytologist*, 117(2), 335–343.

Polle, A., Kroniger, W., & Rennenberg, H. (1996). Seasonal fluctuations of ascorbate-related enzymes: Acute and delayed effects of late frost in spring on antioxidative systems in needles of Norway spruce (*Picea abies* L). *Plant and Cell Physiology*, 37(6), 717–725.

Pop, M., & Salzberg, S. L. (2008). Bioinformatics challenges of new sequencing technology. *Trends in Genetics*, 24(3), 142–149.

Praciak, A. (2013). *CABI Encyclopedia of Forest Trees*. Wallingford,Wallingford Ltd. p. 490.

Price, R. A., Liston, A., & Strauss, S. H. (1998). Phylogeny and systematics of *Pinus*. In *Ecology and biogeography of Pinus*. pp. 49–68.

Pruitt, K. D., Tatusova, T., Brown, G. R., & Maglott, D. R. (2012). NCBI Reference Sequences (RefSeq): Current status, new features and genome annotation policy. *Nucleic Acids Research*, 40(D1), 130–135.

Prus-Glowacki, W, & Stephan, B. R. (1994). Genetic variation of *Pinus sylvestris* from Spain in relation to other European populations. *Silvae Genetica*, 43(1), 7–14.

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.

Prus-Głowacki, W., Urbaniak, L., Bujas, E., & Curtu, A. L. (2012). Genetic variation of isolated and peripheral populations of *Pinus sylvestris* (L.) from glacial refugia. *Flora: Morphology, Distribution, Functional Ecology of Plants*, 207(2), 150–158.

Pukacka, S. & Pukacki, M. (2000). Seasonal changes in antioxidant level of scots pine (*Pinus sylvestris* L.) needles exposed to industrial pollutin. I. Ascorbate and thiol content. *Acta Phsiologiae Plantarum*, 22(4), 451–456.

Pulliam, H. (2017). Sources, sinks, and population regulation. *The American Naturalist*, 132(5), 65.

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.

Pyhäjärvi, T., Salmela, M. J., & Savolainen, O. (2008). Colonization routes of *Pinus sylvestris* inferred from distribution of mitochondrial DNA variation. *Tree Genetics and Genomes*, 4*(2)*, 247–254.

Rao, G., Sui, J., Zeng, Y., He, C., Duan, A., & Zhang, J. (2014). *De Novo* transcriptome and small RNA analysis of two Chinese willow cultivars reveals stress response genes in *Salix matsudana*. *PLoS ONE*, 9(10), e109122.

Rasheed, M. U., Kasurinen, A., Kivimäenpää, M., Ghimire, R., Häikiö, E., Mpmah, P., … Holopainen, T. (2017). The responses of shoot-root-rhizosphere continuum to simultaneous fertilizer addition, warming, ozone and herbivory in young Scots pine seedlings in a high latitude field experiment. *Soil Biology and Biochemistry*, 114, 279–294.

Razzak, A., Ranade, S. S., Strand, Å., & García-Gil, M. R. (2017). Differential response of Scots pine seedlings to variable intensity and ratio of red and far-red light. *Plant Cell and Environment*, 40(8), 1332–1340.

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.

Reich, P. B., Koike, T., Gower, S. T., & Schoettle, A. W. (1995). Causes and consequences of variation in conifer leaf life-span. In *Ecophysiology of coniferous forests*, San Diego, Academic Press. pp. 225-254.

Reich, P. B., Walters, M. B., Ellsworth, D. (1992). Leaf life-span in relation to leaf, plant, and stand characteristics among diverse ecosystems. *Ecological Monographs*, 62(3), 365–392.

Reich, P. B., Ellsworth, D. S., Walters, M. B., Vose, J. M., Gresham, C., Volin, J. C., & Bowman, W. D. (1999). Generality of leaf trait relationships: a test across six biomes. *Ecology*, 80(6), 1955–1969.

Reich, P. B., Oleksyn, J., Modrzynski, J., & Tjoelker, M. G. (1996). Evidence that longer needle retention of spruce and pine populations at high elevations and high latitudes is largely a phenotypic response. *Tree Physiology*, 16(7), 643–7.

Reich, P. B., Wright, I. J., Craine, J. M., Oleksyn, J., Westoby, M., & Walters, M. B. (2003). The evolution of plant functional variation : traits , spectra , and strategies, *International Journal of Plant Sciences*, 164(S3), S143-S164.

Remlein, A., Jelonek, T., Tomczak, A., Jakubowski, M., & Grzywiński, W. (2015). Morphological architecture of different ecotypes of Scots pine (*Pinus sylvestris* L.) in Poland. *Annals of Warsaw University of Life Sciences*, 92, 355–363.

Repo, T., Launiainen, S., Lehto, T., Sutinen, S., Ruhanen, H., Heiskanen, J., ... & Finér, L. (2014). The responses of Scots pine seedlings to waterlogging during the growing season. *Canadian Journal of Forest Research*, 46(12), 1439–1450.

Repo, T., Heiskanen, J., Sutinen, M. L., Sutinen, R., & Lehto, T. (2017). The responses of Scots pine seedlings to waterlogging in a fine-textured till soil. *New Forests*, 48(1), 51–65.

Richardson, A. D., Berlyn, G. P., Gregoire, T. (2001). Spectral reflectance of *Picea rubens* (Pinaceae) and *Abies balsamea* (Pinaceae) needles along an elevational gradient, Mt. Moosilauke, New Hampshire, USA. *American Journal of Botany*, 88(4), 667–676.

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.

Richardson, A. D., Berlyn, G. P., Ashton, P. M. S., Thadani, R., & Cameron, I. R. (2000). Foliar plasticity of hybrid spruce in relation to crown position and stand age. *Canadian Journal of Botany*, 317(3), 305–317.

Riikonen, J. (2017). Content of far-red light in supplemental light modified characteristics related to drought tolerance and post-planting growth in Scots pine seedlings. *Forest Ecology and Management*, 390, 8–14.

Riodrío, J., del Río, M., Pretzsch, H., & Bravo, F. (2017). Changes in structural heterogeneity and stand productivity by mixing Scots pine and Maritime pine. *Forest Ecology and Management*, 405, 219-228.

Robatzek, S., & Somssich, I. E. (2001). A new member of the *Arabidopsis* WRKY transcription factor family , At WRKY6 , is associated with both senescence- and defence-related processes. *The Plant Journal*, 28(2), 123-133.

Roberts, E. H., & Ellis, R. H. (1989). Water and Seed Survival. *Annals of Botany*, 63(1), 39–52.

Robledo-Arnuncio, J. J., Collada, C., Alia, R., & Gil, L. (. (2005). Genetic structure of montane isolates of *Pinus sylvestris* L . in a Mediterranean refugial area. *Journal of Biogeography*, 32, 4.

Ronikier, M. (2011). Biogeography of high-mountain plants in the Carpathians: An emerging phylogeographical perspective. *Taxon*, 60(2), 373–389.

Rose, T. M., Henikoff, J. G., & Henikoff, S. (2003). CODEHOP (COnsensus-DEgenerate Hybrid Oligonucleotide Primer) PCR primer design. *Nucleic Acids Research*, 31(13), 3763–3766.

Rouillard, J. M., Zuker, M., & Gulari, E. (2003). OligoArray 2.0: Design of oligonucleotide probes for DNA microarrays using a thermodynamic approach. *Nucleic Acids Research*, 31(12), 3057–3062.

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.

Rozen, S. & Skaletsky, H. (1999). Primer3 on the WWW for general users and for biologist programmers. *Bioinformatics Methods and Protocols*, 365–386.

Rudawska, M., Leski, T., Aučina, A., Karliński, L., Skridaila, A., & Ryliškis, D. (2017). Forest litter amendment during nursery stage influence field performance and ectomycorrhizal community of Scots pine (*Pinus sylvestris* L.) seedlings outplanted on four different sites. *Forest Ecology and Management*, 395, 104–114.

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.

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

Rutledge, R., Regan, S., Nicolas, O., Fobert, P., Côté, C., Bosnich, W., … Stewart, D. (1998). Characterization of an AGAMOUS homologue from the conifer black spruce (*Picea mariana*) that produces floral homeotic conversions when expressed in Arabidopsis. *Plant Journal*, 15(5), 625–634.

Salmela, M. J., Cavers, S., Cottrell, J. E., Iason, G. R., & Ennos, R. A. (2013). Spring phenology shows genetic variation among and within populations in seedlings of Scots pine (*Pinus sylvestris* L.) in the Scottish Highlands. *Plant Ecology & Diversity*, 6(3-4), 523-536.

Santiago, N., Herráiz, C., Goñi, J. R., Messeguer, X., & Casacuberta, J. M. (2002). Genome-wide analysis of the Emigrant family of MITEs of *Arabidopsis thaliana*. *Molecular Biology and Evolution*, 19(12), 2285–2293.

Savolainen, O., Pyhäjärvi, T., & Knürr, T. (2007). Gene flow and local adaptation in trees. *Source: Annual Review of Ecology, Evolution, and Systematics*, 38, 595–619.

Savva, Y., Denneler, B., Koubaa, A., Tremblay, F., Bergeron, Y., & Tjoelker, M. G. (2007). Seed transfer and climate change effects on radial growth of jack pine populations in a common garden in Petawawa , Ontario , Canada. *Forest Ecology and Management*, 242(2), 636–647.

Scalfi, M., Piotti, A., Rossi, M., & Piovani, P. (2009). Genetic variability of Italian southern Scots pine (*Pinus sylvestris* L.) populations: The rear edge of the range. *European Journal of Forest Research*, 128(4), 377–386.

Scheibe, R. (2004). Malate valves to balance cellular energy supply. *Physiologia Plantarum*, 120(1), 21–26.

Schiestl-Aalto, P., & Mäkelä, A. (2017). Temperature dependence of needle and shoot elongation before bud break in Scots pine. *Tree Physiology*, 37(3), 316–325.

Schlueter, S. D., Dong, Q., & Brendel, V. (2003). GeneSeqer@PlantGDB: Gene structure prediction in plant genomes. *Nucleic Acids Research*, 31(13), 3597–3600.

Schnabel, A., & Hamrick, J. L. (1990). Organization of genetic diversity within and among populations of *Gleditsia triacanthos* Leguminosae. *American Journal of Botany*, 77(8), 1060–1069.

Schoettle, A. W., Fahey, T. J., & Shoettle, A. W. (1994). Foliage and fine root longevity of pines. *Ecological Bulletins*, 136–153.

Schönbeck, L., Gessler, A., Rigling, A., Schaub, M., & Li, M. (2017). Drought stress release increased growth rate but did not affect levels of storage carbohydrates in Scots pine trees, *EGU General Assembly Conference Abstracts*, 19, 6572.

Schröder, P., & Rennenberg, H. (1992). Characterization of glutathione S-transferase from dwarf pine needles (*Pinus mugo* Turra). *Tree Physiology*, 11(2), 151–60.

Schwartz, M. W. (1991). Potential effects of global climate change on the biodiversity of plants. *The Forestry Chronicle*, 68(4), 462–471.

Scudder, G. (1989). The adaptive significance of marginal populations: a general perspective. In: *Canadian Special Publication of Fisheries and Aquatic sciences/Publication Speciale Canadienne Des Sciences Halieutiques et Aquatiques.*Ottawa. pp.180-185.

Semerci, A., Semerci, H., Çalişkan, B., Çiçek, N., Ekmekçi, Y., & Mencuccini, M. (2017). Morphological and physiological responses to drought stress of European provenances of Scots pine. *European Journal of Forest Research*, 136(1), 91–104.

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.

Senner, N. R., Stager, M, Cheviron, Z. (2017). Spatial and temporal heterogeneity in climate change limits species' dispersal capabilities and adaptive potential. *Ecography*, (7491), 0–2.

Sevik, H., & Topacoglu, O. (2015). Variation and inheritance pattern in cone and seed characteristics of Scots pine (*Pinus sylvestris* L.) for evaluation of genetic diversity. *Journal of Environmental Biology*, 36(5), 1125.

Sevık, H., Ayan, S., Turna, I., & Yahyaoglu, Z. (2010). Genetic diversity among populations in Scotch pine (*Pinus silvestris* L.) seed stands of Western Black Sea Region in Turkey. *African Journal of Biotechnology*, 9(43), 7266–7272.

Shamimuzzaman, M., & Vodkin, L. (2013). Genome-wide identification of binding sites for NAC and YABBY transcription factors and co-regulated genes during soybean seedling development by ChIP-Seq and RNA-Seq. *BMC Genomics*, 14(1), 477.

Singh, K. B., Foley, R. C., & Oñate-Sánchez, L. (2002). Transcription factors in plant defense and stress responses. *Current Opinion in Plant Biology*, 5(5), 430–436.

Skilling, D. L. (1990). Scotch pine. In *Silvics of North America (ed. Note 86, Ottawa. by R.M. Burns and B.H. Honkala).* pp. 489–496.

Slater, G., Birney, E., Box, G., Smith, T., Waterman, M., Altschul, S., … Salzberg, S. (2005). Automated generation of heuristics for biological sequence comparison. *BMC Bioinformatics*, 6(1), 31.

Sobierajska, K., & Marcysiak, K. (2010). Variation of cone characters in *Pinus mugo* (Pinaceae) populations in the Giant Mountains (Karkonosze , Sudetes). *Dendrobiology*, 63, 33–41.

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

Souvorov, A., Kapustin, Y., Kiryutin, B., Chetvernin, V., Tatusova, T., & Lipman, D. (2010). Gnomon–NCBI eukaryotic gene prediction tool. *National Center for Biotechnology Information*, 1–24.

Stankova, T. V., & Shibuya, M. (2007). Stand density control diagrams for Scots pine and Austrian black pine plantations in Bulgaria. *New Forests*, 34(2), 123–141.

Staszak, J., Grulke, N. E., & Prus-Glowacki, W. (2004). Genetic differences of *Pinus ponderosa* [Dougl. ex Laws.] trees tolerant and sensitive to ozone. *Water, Air, and Soil Pollution*, 153(1–4), 53–14.

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.

Strauss, S. H., & Libby, W. J. (1987). Allozyme heterosis in radiata pine is poorly explained by overdominance. *The American Naturalist*, 130(6), 879–890.

Strimbeck, G. R., Kjellsen, T. D., Schaberg, P. G., & Murakami, P. F. (2007). Cold in the common garden : comparative low-temperature tolerance of boreal and temperate conifer foliage. *Trees*, 21(5), 557–567.

Sundström, J., Carlsbecker, A., Svensson, M. E., Svenson, M., Johanson, U., Theißen, G., & Engström, P. (1999). MADS-box genes active in developing pollen cones of Norway spruce (*Picea abies)* are homologous to the B-class floral homeotic genes in angiosperms. *Developmental Genetics*, 25(3), 253–266.

Sundström, J., & Engström, P. (2002). Conifer reproductive development involves B-type MADS-box genes with distinct and different activities in male organ primordia. *Plant Journal*, 31(2), 161–169.

Surles, S. E., White, T. L., Hodge, G. R., & Duryea, M. L. (1993). Relationships among seed weight components, seedling growth traits, and predicted field breeding values in slash pine. *Canadian Journal of Forest Research*, 23(8), 1550-1556.

Szweykowski, J. and Urbaniak, L. (1982). An interesting chemical polymorphism in *Pinus sylvestris* L. *Acta Societatis Botanicorum Poloniae*, 51(3–4), 441–452.

Tabachnick, F. L. (2007). Structural equation modeling. In *Using multivariate statistics* (5th ed.). Boston, MA. pp. 76–780.

Taberlet, P., Fumagalli, L., Wust-Saucy, A.-G., & Cosson, J.-F. (1998). Comparative phylogeography and postglacial colonization. *Molecular Ecology*, 7(4), 453–464.

Tandre, K., Albert, V. A., Sundås, A., & Engström, P. (1995). Conifer homologues to genes that control floral development in angiosperms. *Plant Molecular Biology*, 27(1), 69–78.

Tang, W. (2000). Peroxidase activity of desiccation-tolerant loblolly pine somatic embryos. *In Vitro Cellular and Developmental Biology - Plant*, 36(6), 488–491.

Tang, X., Xiao, Y., Lv, T., Wang, F., Zhu, Q. H., Zheng, T., & Yang, J. (2014). High-throughput sequencing and *de novo* assembly of the *Isatis indigotica* transcriptome. *PLoS ONE*, 9(9), e102963.

Tao, D. L., Oquist, G., & Wingsle, G. (1998). Active oxygen scavengers during cold acclimation of Scots pine seedlings in relation to freezing tolerance. *Cryobiology*, 37(1), 38–45.

Tarakanov, V. V., Chindyaeva, L. N., Tsybulya, N. V., Tikhonova, I. (2017). Variability of needle antimicrobic activity in clone plantation of Scots pine *Pinus sylvestris* L. *Sibirskij Lesnoj Zurnal\Siberian Journal of Forest Science*, 1, 95–104.

Teklemariani, T. A., & Blake, T. J. (2004). Phenylalanine ammonia-lyase-induced freezing tolerance in jack pine (*Pinus banksiana*) seedlings treated with low, ambient levels of ultraviolet-B radiation. *Physiologia Plantarum*, 122(2), 244–253.

Tereba, A. M., Konecka, A., & Nowakowska, J. A. (2017). Application of selected molecular markers in studies on forest trees. *Folia Forestalia Polonica, Series A*, 59(2), 146–151.

Thomas, C. D., Bodsworth, E. J., Wilson, R. J., Simmons, A. D., Davies, Z. G., Musche, M., & Conradt, L. (2001). Ecological and evolutionary processes at expanding range margins. *Nature*, 411(4837), 577–581.

Tikhonova, I. V., Tarakanov, V. V., Tikhonova, N. A., Barchenkov, A. P., & Ekart, A. K. (2014). Population variability of cones and seeds of scots pine by phenes of color and traits-indices in the south of Siberia. *Contemporary Problems of Ecology*, 7(1), 60–66.

Timofeeva, G., Treydte, K., Bugmann, H., Rigling, A., Schaub, M., Siegwolf, R., & Saurer, M. (2017). Long-term effects of drought on tree-ring growth and carbon isotope variability in Scots pine in a dry environment. *Tree Physiology*, 1(14), 1028-1041.

Tipping, R., Ashmore, P., Davies, A. L., Haggart, B. A., Moir, A., Newton, A., … Tisdall, E. (2008). Prehistoric *Pinus* woodland dynamics in an upland landscape in northern Scotland: The roles of climate change and human impact. *Vegetation History and Archaeobotany*, 17(3), 251–267.

Tiwari, S. P., Kumar, P., Yadav, D., & Chauhan, D. K. (2013). Comparative morphological , epidermal , and anatomical studies of *Pinus roxburghii* needles at different altitudes in the North-West Indian Himalayas. *Turkish Journal of Botany*, 37(1), 65–73.

Tobolski, J. J. and Hanover, J. (1971). Genetic variation in the monoterpenes of Scotch pine. *Forest Science*, 17(3), 293–299.

Tor-Ngern, P., Oren, R., Oishi, A. C., Uebelherr, J. M., Palmroth, S., Tarvainen, L., … Näsholm, T. (2017). Ecophysiological variation of transpiration of pine forests: Synthesis of new and published results: Synthesis. *Ecological Applications*, 27(1), 118–133.

Torre, A. R. D., Li, Z., Peer, Y. Van De, & Ingvarsson, K. (2017). Contrasting rates of molecular evolution and patterns of selection among gymnosperms and flowering plants, *Molecular biology and evolution*, 34(6), 1363-1377.

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.

Travis, J. M. J., & Dytham, C. (2004). A method for simulating patterns of habitat availability at static and dynamic range margins, *Oikos*, 104(2), 410–416.

Trocha, L. K., Kałucka, I., Stasińska, M., Nowak, W., Dabert, M., Leski, T., … Oleksyn, J. (2012). Ectomycorrhizal fungal communities of native and non-native *Pinus* and *Quercus species* in a common garden of 35-year-old trees. *Mycorrhiza*, 22(2), 121–134.

Tu, Z. (2001). Eight novel families of miniature inverted repeat transposable elements in the African malaria mosquito, *Anopheles gambia*e. *Proceedings of the National Academy of Sciences*, 98(4), 1699–704.

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.

Tzedakis, P. C., Frogley, M. R., Lawson, I. T., Preece, R. C., Cacho, I., & de Abreu, L. (2004). Ecological thresholds and patterns of millennial-scale climate variability: The response of vegetation in Greece during the last glacial period. *Geology*, 32(2), 109–112.

Ueda, M., Zhang, Z., & Laux, T. (2011). Transcriptional activation of *Arabidopsis* axis patterning genes WOX8/9 links zygote polarity to embryo development. *Developmental Cell*, 20(2), 264–270.

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.

Urbaniak, L. (1994). Morphometric differentiation of *Corex ligerica* Gay in Poland. *Acta Societatis Botanicorum Poloniae*, 67(4), 263–268.

Urbanska, K. M. (1984). Polymorphism of cyanogenesis in *Lotus alpinus* from Switzerland. II. Phenotypic and allelic frequencies upon acidic silicate and carbonate. *Bulletin of the Geobotanical Institute of Swiss Federal Institute of Technology, Zurich*, 51, 132–63.

Uria-Diez, J., & Pommerening, A. (2017). Crown plasticity in Scots pine (*Pinus sylvestris* L.) as a strategy of adaptation to competition and environmental factors. *Ecological Modelling*, 356, 117–126.

Valkonen, S., Ruuska, J., & Siipilehto, J. (2002). Effect of retained trees on the development of young Scots pine stands in Southern Finland. *Forest Ecology and Management*, 166(1–3), 227–243.

van der Maaten, E., Mehl, A., Wilmking, M., & van der Maaten-Theunissen, M. (2017). Tapping the tree-ring archive for studying effects of resin extraction on the growth and climate sensitivity of Scots pine. *Forest Ecosystems*, 4(1), 7.

van Hijum, S. A., de Jong, A., Buist, G., Kok, J., & Kuipers, O. P. (2003). UniFrag and GenomePrimer: Selection of primers for genome-wide production of unique amplicons. *Bioinformatics*, 19(12), 1580–1582.

Vander Wall, S. B. (2002). Masting in animal-dispersed pines facilitates seed dispersal. *Ecology*, 83(12), 3508–3516.

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.

Vergeer, P., & Kunin, W. E. (2013). Adaptation at range margins: Common garden trials and the performance of *Arabidopsis lyrata* across its northwestern European range. *New Phytologist*, 197(3), 989–1001.

Vitasse, Y., Delzon, S., Bresson, C. C., Michalet, R., & Kremer, A. (2009). Altitudinal differentiation in growth and phenology among populations of temperate-zone tree species growing in a common garden. *Canadian Journal of Forest Research*, 39(7), 1259–1269.

Volkova, P. Y., Geras' kin, S. A., Horemans, N., Makarenko, E. S., Saenen, E., Duarte, G. T., ... & Kudin, M. (2018). Chronic radiation exposure as an ecological factor: Hypermethylation and genetic differentiation in irradiated Scots pine populations. *Environmental Pollution*, 232, 105–112.

Volkova, L., Bennett, L. T., Merchant, A., & Tausz, M. (2010). Shade does not ameliorate drought effects on the tree fern species *Dicksonia antarctica* and *Cyathea australis*. *Trees - Structure and Function*, 24(2), 351–362.

Voltas, J., Chambel, M. R., Prada, M. A., & Ferrio, J. P. (2008). Climate-related variability in carbon and oxygen stable isotopes among populations of Aleppo pine grown in common-garden tests. *Trees - Structure and Function*, 22(6), 759-769.

Vucetich, J. A., & Waite, T. A. (2003). Spatial patterns of demography and genetic processes across the species' range: Null hypotheses for landscape conservation genetics. *Conservation Genetics*, 4(5), 639-645.

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., & 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.

Wang, H., Wan, P., Wang, Q., Liu, L., Zhang, G., & Hui, G. (2017). Prevalence of inter-tree competition and its role in shaping the community structure of a natural mongolian Scots pine (*Pinus sylvestris* var. mongolica) Forest. *Forests*, 8(3), 84.

Wang, M., Xue, J., Horswell, J., Kimberley, M. O., & Huang, Z. (2017). Long-term biosolids application alters the composition of soil microbial groups and nutrient status in a pine plantation. *Biology and Fertility of Soils*, 53(7), 799–809.

Wang, H., Tetzlaff, D., Dick, J. J., & Soulsby, C. (2017). Assessing the environmental controls on Scots pine transpiration and the implications for water partitioning in a boreal headwater catchment. *Agricultural and Forest Meteorology*, 240–241(November 2016), 58–66.

Wang, K., Li, M., & Hakonarson, H. (2010). ANNOVAR: Functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Research*, 38(16), 1–7.

Watts, W. A. (1985). Quaternary vegetation cycles. *The Quaternary History of Ireland*, London, Academic Press. pp.155–185.

Wheelan, S. J., Church, D. M., & Ostell, J. M. (2001). Spidey: A tool for mRNA-to-genomic alignments. *Genome Research*, 11(11), 1952–1957.

Wheeler, D. L., Barrett, T., Benson, D. A., Bryant, S. H., Canese, K., Chetvernin, V., … Yaschenko, E. (2008). Database resources of the National Center for Biotechnology Information. *Nucleic Acids Research*, 36(SUPPL. 1), 13–21.

Whittet, R., Cavers, S., Cottrell, J., Rosique-Esplugas, C., & Ennos, R. (2017). Substantial variation in the timing of pollen production reduces reproductive synchrony between distant populations of *Pinus sylvestris* L. in Scotland. *Ecology and Evolution*, 7(15), 5754–5765.

Williams, C. G., Auckland, L. D., Reynolds, M. M., & Leach, K. A. (2003). Overdominant lethals as part of the conifer embryo lethal system. *Heredity*, 91(6), 584–592.

Willis, K. J., Bennett, K. D., & Birks, H. J. B. (1998). *The late Quaternary dynamics of pines in Europe*. Cambridge, Cambridge University Press. pp.107-121.

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.

Wingsle, G., & Hällgren, J. (1993). Influence of SO2 and NO2 exposure on glutathione, superoxide dismutase and glutathione reductase activities in scots pine needles. *Journal of Experimental Botany*, 44(2), 463–470.

Winkel-Shirley, B. (1999). Evidence for enzyme complexes in the phenylpropanoid and flavonoid pathways enzyme complexes and subcellular localization in metabolic control. *Physiologia Plantarum*, 107, 142–149.

Woodruff, D. R., Meinzer, F. C., & McCulloh, K. A. (2016). Forest canopy hydraulics. In *Canopy photosynthesis: from basics to applications*. Dordrecht, Springer Netherlands. pp. 187–217.

Wright, J. W., Lemmien, W. A., & Bright, J. (1966). Early flowering patterns in Scotch pine. *Michigan Agricultural Experiment Station Bulletin*, 49, 189–199.

Wu, J. S., Lee, C., Wu, C. C., & Shiue, Y. L. (2004). Primer design using genetic algorithm. *Bioinformatics*, 20(11), 1710–1717.

Wu, T., Luo, S., Wang, R., Zhong, Y., Xu, X., Lin, Y., … Huang, H. (2014). The first Illumina-based *de novo* transcriptome sequencing and analysis of pumpkin (*Cucurbita moschata* Duch.) and SSR marker development. *Molecular Breeding*, 34(3), 1437–1447.

Xie, Y., Wu, G., Tang, J., Luo, R., Patterson, J., Liu, S., … Wang, J. (2014). SOAPdenovo-Trans: *De novo* transcriptome assembly with short RNA-Seq reads. *Bioinformatics*, 30(12), 1660–1666.

Xue, B., Charest, P. J., Devantier, Y., & Rutledge, R. G. (2003). Characterization of a MYBR2R3 gene from black spruce (*Picea mariana*) that shares functional conservation with maize C1. *Molecular Genetics and Genomics*, 270(1), 78–86.

Yamazaki, J. Y., Ohashi, A., Hashimoto, Y., Negishi, E., Kumagai, S., Kubo, T., … Kamimura, Y. (2003). Effects of high light and low temperature during harsh winter on needle photodamage of *Abies mariesii* growing at the forest limit on Mt. Norikura in Central Japan. *Plant Science*, 165(1), 257–264.

Yandell, M., & Ence, D. (2012). A beginner's guide to eukaryotic genome annotation. *Nature Reviews Genetics*, 13(5), 329–342.

Yang, Y., Han, C., Liu, Q., Lin, B., & Wang, J. (2008). Effect of drought and low light on growth and enzymatic antioxidant system of *Picea asperata* seedlings. *Acta Physiologiae Plantarum*, 30(4), 433–440.

Yang, Y., Liu, Q., Han, C., Qiao, Y. Z., Yao, X. Q., & Yin, H. J. (2007). Influence of water stress and low irradiance on morphological and physiological characteristics of *Picea asperata* seedlings. *Photosynthetica*, 45(4), 613–619.

Ye, J., Coulouris, G., Zaretskaya, I., Cutcutache, I., Rozen, S., & Madden, T. L. (2012). Primer-BLAST: a tool to design target-specific primers for polymerase chain reaction. *BMC Bioinformatics*, 13, 134.

Yeh, R., Lim, L. P., & Burge, C. B. (2001). Computational inference of homologous gene structures in the human genome computational inference of homologous gene structures in the human genome. *Genome Research*, 803–816.

Yoshiura, K. I., Kinoshita, A., Ishida, T., Ninokata, A., Ishikawa, T., Kaname, T., … Niikawa, N. (2006). A SNP in the ABCC11 gene is the determinant of human earwax type. *Nature Genetics*, 38(3), 324–330.

Zadworny, M., McCormack, M. L., Żytkowiak, R., Karolewski, P., Mucha, J., & Oleksyn, J. (2017). Patterns of structural and defense investments in fine roots of Scots pine (*Pinus sylvestris* L.) across a strong temperature and latitudinal gradient in Europe. *Global Change Biology*, 23(3), 1218–1231.

Zhang, K., Baskin, J. M., Baskin, C. C., Yang, X., & Huang, Z. (2017). Effect of seed morph and light level on growth and reproduction of the amphicarpic plant *Amphicarpaea edgeworthii* (Fabaceae). *Nature Publishing Group*, (November 2016), 1–8.

Zhang, Q., Li, J., Zhao, Y., Korban, S. S., & Han, Y. (2012). Evaluation of genetic diversity in chinese wild apple species along with apple cultivars using SSR Markers. *Plant Molecular Biology Reporter*, 30(3), 539–546.

Zhang, Y., & Wang, L. (2005). The WRKY transcription factor superfamily: its origin in eukaryotes and expansion in plants. *BMC Evolutionary Biology*, 5(1), 1.

Zhu, J., Wang, K., Sun, Y., & Yan, Q. (2014). Response of *Pinus koraiensis* seedling growth to different light conditions based on the assessment of photosynthesis in current and one-year-old needles. *Journal of Forestry Research*, 25(1), 53–62.

http://www.euforgen.org/species/pinus-sylvestris (downloaded 11.01.2018)
http://www.conifers.org/pi/Pinus_sylvestris.php (downloaded 11.01.2018)
http://www.ncbi.nlm.nih.gov (downloaded 11.01.2018)

11. SUPPLEMENTARY MATERIAL

Table S1. List of the designed 164 primers (with each pairs' nucleotide succession, melting temperature and length of the amplificated sequence).

Code	Sequence Fwd	Sequence Rev	Melt. temp. Fwd	Melt. temp. Rev	Fragment length (Fwd,Rev)
Myb1372	CTGATTGCAGGGCGTATT	CTTCAACATCATTGCGTCC	57.39	56.17	172
Myb4633	AAGCAGCAGCCAACAAAA	GCAAAGCTCCAGTGTAAT	55.80	55.54	349
Myb139	ACTTCGTGATGAAAAGCTG	ATAGTTTTGTTGGCTGCTG	56.11	56.29	159
Myb13629	GAGATTGAAAGCAGAAAAGG	CTGGAACAGGCATAGTTT	54.66	54.29	152
Myb10181	CTGATTGCAGGGCGTATT	CTTCAACATCATTGCGTCC	57.39	56.17	175
Myb4344	ATCCTGGAATGGCAATGT	GCAAAGCTCCAGTGTAAT	56.51	55.54	313
Myb4095	GAGAAAACTTATCAGCATGG	AATATCCCTTTCCCTCAC	53.44	53.10	323
Myb19497	GATTCTCTGTTACTTGCC	TGCTTCTCTCTATAGACC	52.36	51.43	287
Myb11378	CTGATTGCAGGGCGTATT	CTTCAACATCATTGCGTCC	57.39	56.17	172
Myb4661	GAAGACAGAGAAAAATGAGC	ACCAGCGAATCCATTTCA	54.41	56.63	468
Myb2894	GAAGGCAGAGAAAAATGAAC	CTGGAACAGGCATAGTTT	54.93	54.18	188
Myb5330	CTGATTGCAGGGCGTATT	CTTCAACATCATTGCGTCC	57.39	56.17	172
Myb4667	GAAGGCAGAGAAAAATGAAC	GGAACAGGCATAGTTTTG	54.93	54.32	186
Myb2882	GAGCTAAGCATCTTGTTGG	GTCTTGTTATTTGCTGCTTG	55.48	54.96	321
WRK9928	GCAAGAAAGCACGTAGAA	CATGGAGATTGCGCTAAA	53.13	53.95	223
WRY8624	ATGGAGGTGGCAGAAGAT	CCTCCATGAACACTTTTTTG	57.84	54.88	365
WRKY20368	GAAGAAACAGACGAAGGA	CACCAAACTATCTTAAGCC	53.44	52.49	246
WRKY4666	GCAGTATAACAACCAAAGAG	GATCCAAAAACTCCCCAAA	53.65	53.70	138
WRKY4214	GCAAGAAAGCACGTAGAA	CATGGAGATTGCGCTAAA	53.13	53.95	223
WRKY2370	ATGGAGGTGGCAGAAGAT	CCTCCATGAACACTTTTTG	57.84	54.17	364
WRKY16373	TGCACCCATCCAAATTGT	ATTCCCAGTGTCATGTAG	55.75	53.61	306
WRKY10472	TCACAACAAAGGAGCTTG	CCATTCTGGTCCTTGTCG	53.21	56.73	214
WRKY10455	GATGTCAAAAAGAAGGATGC	TGTACAAATCCCACCTCC	54.49	55.03	257
WRKY5091	AATAAAGGTGCAGAAGGC	CATCTCTTTCCTCCACCA	54.93	54.83	294
WRKY3145	GGAGCAGTTGACCGAAGA	CCGCAGATATGAAACCATT	57.50	55.11	305

WRKY2211	GGATAAGCTTACAGAGGA	AGATTTAGGAACTCGGCA	51.69	54.54	210
WRKY1289	CACAACAAAAGAACTGGG	CACTTGGCCATCATAACT	55.11	55.67	366
MADS15369	GGATATGGAACAAGGTGAA	CTCCTCAACAAAATCAAAGG	53.85	54.29	236
MADS6542	TGGCACTATTCTGGGGAA	CTGATTGCCAACTCCTTT	55.03	54.82	210
MADS3510	AGACACAGTGCTGGAAAA	CGGTCTGCAATCTCTTGT	56.66	56.73	357
MADS1818	CCATCGCTGTTTTCCTTT	CTAACATTTCTCTTCCGGAT	55.29	54.05	361
MADS3172	GGGAGCCTAAAAGTATGA	CAGCCCGCATTTCTTCTA	52.29	55.46	207
MADS20909	CCCAGATGGAAAGACAATA	CCTCCATCAAATCTTTCTC	53.45	53.47	269
MADS2190	ACCTCTGTATTTCCTCTG	CCATGAGCTACCTTCTTT	53.01	53.24	332
MADS2038	GGGAGCCTAAAAGTATGA	CAGCCCGCATTTCTTCTA	52.29	55.46	207
MADS12384	GGATATGGAACAAGGTGAA	CTCCTCAACAAAATCAAAGG	53.85	54.29	236
MADS3334	GACAGAACAATGGAAAGG	CATCCATAAGTTGAACACC	53.33	52.95	382
MADS1317	GTCATTGAAGAAGATAGTGG	GAGAGCAAAGCAACCAAA	52.97	52.81	197
MADS832	CAGAGTGGCGGATGATTT	AGCCACAAGCACAAAAAG	56.75	57.40	323
PAL2688	TTCGAATCTAAGTGGTGG	CCCTTTGTCATAACTTTCTC	52.06	53.18	612
CHS282	GACAAGTCGGCAATAAAG	AGAATCTCCTCCGTCAAG	52.33	56.25	50
PAL1134	TTTCAACAAGATTCCCGC	CAATATATTCGCCCGGAG	54.86	55.84	217
CHS1459	GGACATGGTTGTGGTGGA	GATCTTCGACTTGGGCTG	57.44	57.66	88
F3'H2550	GAAAGACGAGGTGAGCAA	GGTGAAGGGGATAAGAGAAA	55.82	55.77	250
F3H350	GCGTGAGAGACGAAGTAA	GTGAAGGGGATAAGAGAAG	53.85	54.07	254
F3H483	CAATGGCAAGTTCAAGAC	GGCAGACACCCATAAAAA	53.38	54.60	332
IGSTP346	AGAAAATTCCGGTGCTCAT	CTTTCACTCCATCTGCCT	57.67	57.69	462
PAL1050	CAACCAGGACGTCAATTC	AAAATCCGGTGTCGAGAA	54.92	54.94	227
IGSTP3726	GAGACAACCTGATGCAAA	CAATCTCCCTCTCAAACT	52.66	52.52	142
IGSTP17883	GTTCGCATAGCACTTTCT	CTCGGTCGTATGGATCTT	55.19	55.38	226
CHS381	GACAAGTCGGCAATAAAG	TTGGTACATCATCACTCTC	52.33	52.52	301
CHS1053	GCTAAGGCCATCAAGGAA	CACTGCAGACGACCAGAA	55.22	57.60	243
CHS2058	GCGACAAGTCAGCAATAA	CAGCCGCTTCTTTTCCAA	54.35	56.42	153
CHS4014	CGACAAGTCAGCAATAAAG	ACTGGTAGTGCAGAAAAC	53.68	53.88	220
CHS7594	CATGGTTGTTGTGGAGGT	TTGGTACATCATCACTCTCT	56.62	54.44	196

IGSTP18952	AGCTCATCACTTCTCGAAA	GTACTGGCCATCGTATTTT	54.94	55.04	185
IGSTP3638	GAATATTGTGCAAGGAGAG	GTTCTGAAGAGGTTGTATG	53.61	53.15	232
IGSTP657	TCTATGCGAGGAGTTTTC	TCGGGCATCACTTCTTTC	52.12	55.24	302
3RT9058	TGCAGAAGCCATAACGAA	AATTAGCATATCCCTCGC	53.63	54.24	272
IGSTP9956	GAATCAGAGATAAGTGTGG	TAAGGATTGGTTTTGGGC	53.07	52.97	312
IGSTP9874	GAGTACCTTGTGAATGCT	GTACTGGCCATCGTATTT	54.07	54.22	216
IGSTP985	TTTTGATAATGGAGGTGCC	CAGTCAGACCCAACAAAT	54.32	55.84	369
IGSTP554	CAGGTGGAAGAATGGGTG	GATCCGGCATAACTTCTTT	56.43	54.79	259
ALT14589	GTTTAGTTCTCCTCGCTG	TTGCTTTTTGTGGTAGGTG	54.42	54.32	474
ALT14413	TTGTGCTGGTAATTGTGG	TAATGCTTGGCTCTTGCT	53.91	54.70	331
ALT341930	TATTTGGCGGCACTCTTG	GCCGGATTTGGTGGATTA	56.92	55.91	514
AST11138	GAAGATTGGTTGGGCTAT	TTTGGGTCCATGTAGAAG	53.76	51.89	351
AST6108	TCTCAGAAAGCAGGGGTT	CACTACCAGCATATCCCA	56.47	55.51	49
AST3968	GACAGTTGCCTATAACAAAG	GGGCAAATACTCCTTATATC	54.13	52.12	157
AST3795	TATGGCAGGTGTTACTAC	ATGAGAAACCAACTGGAC	51.80	54.03	280
AST1776	GCTGTCCAAGCTCTATCT	GTCTCTGGATGGTAATAATGG	55.64	54.76	172
AST4	GTCTACATGACAAAGGATG	GATGTCAGGTGCTTCAAA	52.75	54.28	181
NAD+3631	CTCCAAATTGCAAGGTCT	CTTCCGTGCCTTGATAAT	54.67	54.23	356
AroE5840	GATTGGGCGTTTGAACAG	ATTATCAACAAGAAGCGGCA	56.50	55.48	142
GPI1664	TGTGGCCATTACTCAGGA	GTCCAACAGCTGACATTT	55.07	55.35	122
6PGD3462	GAGGCAATTGGTCGATGAT	GGCGGTAAGTGTCAAAATAGG	57.40	57.55	355
6PGD2129	GAATCTGAACTTTGGGGAA	TGTATGGAAAGAACCAGG	51.68	51.86	346
6PGD1285	GGGCTCTGTATGCATCTAA	TGGGCTTATTCTCTCTGG	53.62	53.63	500
RuBisCo4941	CCAATTATGATGAGCGCC	CACATCTTCCCTTTCTTGA	54.98	53.71	113
RuBisCo1424	TTGCTATAATGACAGGTGC	TCTAAGTTTTCGGTCCTC	52.69	52.62	311
RuBisCo459	GCCATTGAGGAAGGTATT	TTCTGTAGAGCACACCTTG	53.86	55.04	308
diTP9764	TTCGAAGGTGGTGGAAAT	TCGATGAACAGCTTGATG	55.07	54.89	211
diTP8780	CCTTGTTAGAGGCAATTTTC	TTGCAAAGACTCGAAGGG	55.11	57.27	325
diTP977	TCGAAGGTGGTGGAAATC	GTAAAAATGGCTCTACACG	55.92	53.97	129
AcI1998	TACAAGAATGCAGTTCCC	GTAGCACCCTTAGCATTG	52.95	55.32	158

SUS4909	TTGCAACGTGTAATGGAG	CCTTTCATAGATCCGCTG	54.46	54.66	182
SOD8804	ATGTCCGACCAGATTATC	GTCAACAACCAAGGAGCA	53.42	57.60	226
SOD7730	CATTGAGTGGACCTGATT	GCAACTCACTAATGCCAA	53.90	54.68	303
SOD6983	GGGTAATTCACAAGTCGAG	TCCATCAGAACCAGCAAC	54.74	55.54	250
SOD2665	CAGATTTTACAGAGCTCAAG	CTGTTGTATCCTTTTCGC	53.01	52.94	621
APX593	TTGACAAAGCCAGGAGGA	AAACTAGGATCAGCAAGCA	55.41	55.24	601
ALT2898	GGTCTGGATGTTGGGGAA	CAAAGTATCCTCCACGCT	57.95	57.61	343
AST3181	CCAACTGGACCCCTAAAA	GTGTTACTACAGGCAATG	52.75	53.05	264
AST2487	TCCCTTGACACGCTTAGA	TCTTTATATAGTCTGGCCAC	54.48	52.97	197
AST626	TTAAAATTGGCACCGCACT	TTGGCACTATGATGTGGTT	56.71	55.78	242
NAD+4462	CCTCTCTCACTGTTGATT	GCTCTAACAACATCCAAG	52.57	52.25	416
NAD+2599	GTGCCTATGATTGCTAGA	ATCAACTCCATCTTCACC	51.88	53.49	113
AroE22855	TGGGGAAGAGCATTAAGT	TAAGTGACAATTGCAGGG	53.99	53.55	268
6PGD5962	GTGACATACATTGGTAAAGG	TCCACAAGATACCCATCC	53.99	53.37	248
PFK6908	GCGATGCTTTTGGACAGA	CCACTGACGAGCTTAAAT	56.70	55.58	132
RuBisCo2097	GGACCGAAAACTTAGAATTG	TTCTGTAGAGCACACCTT	53.99	53.50	354
RuBisCo219	TGAACTGGAGAGTGGAGAT	AATTTTGCTGGCTCTCCT	56.06	56.73	39
diTPS18523	CAATGGAAGACGGTGAAAT	TCCAGGTTTGAAGTGCGA	56.74	56.83	219
diTPS11241	CCTTGTTAGAGGCAATTTTC	AGTCAGCCATTGTCCATA	55.11	54.48	79
diTPS10990	AATATACCCCCCATTGGA	GATGAGAGAGCATGAGTT	54.22	52.33	115
diTPS8769	CTTGGTCCGGGCAGTATA	GAGGCGTTCAATTCAATCA	54.74	54.79	180
diTPS8495	CGCCAAGGTGAAATAGTAA	CCATAAGTACCGCAGATT	53.51	53.58	106
diTPS5871	TCTTCTGTTTCTTCGGTC	CCTCTTCCTTTAAACCTTTG	52.80	52.80	195
diTPS4679	CCAAAAATGAACGATGAGC	CATCAGAGCAGGAAATACA	55.88	53.02	69
AcI15594	TACAAGAATGCAGTTCCC	AATTGCACAGAAGACACC	52.95	55.20	102
SUS2581	CTTTTACCCTTCCTGGTCT	AAAATCACCAGCAACCAC	55.72	55.80	338
SOD694	GGAAAGTCGTGAATTGGAAG	GTCAACAACCAAGGAGCA	57.19	57.60	196
GR1894	GTAGTTGAGGAGGGTTTG	TATCGGATTGGAAGACCT	54.09	53.46	223
APX858	TTGACAAAGCCAGGAGGA	AAACTAGGATCAGCAAGCA	55.41	55.24	600
APX637	GAAGCTTTCTGAATTGGG	GATTACAATATAGCGTGGG	53.24	53.21	298

ALT10459	TTAGTTCTCCTCGCTGATG	ACAAAGATTCACAGATGCC	55.66	55.62	246
AST18254	TTAAAATTGGCACCGCACT	TTGGCACTATGATGTGGTT	56.71	55.78	242
AST10567	TCCCTTGACACGCTTAGA	CCACTTTTCAGGCCATTT	54.48	55.35	181
AST4142	GTTTCTGCCTTCCAACAT	ACTTCCATCACCAAGAATCA	55.27	55.21	270
AST4000	TGACAAAGGATGGGCGTA	GATGTCAGGTGCTTCAAAA	56.46	55.84	174
AST2456	GGATAAGTATGGCAGGTC	TCAGTTCTTTGCCTCTTG	53.68	53.61	179
6PGD3004	GGGCTCTGTATGCATCTAA	TGGGCTTATTCTCTCTGG	53.62	53.63	500
6PGD930	GAATCTGAACTTTGGGGAA	TGTATGGAAAGAACCAGG	51.68	51.86	346
PFK7865	GGTGAGGGATATGGTGAAA	CTCTGTTCTATATGCTGGT	53.27	52.81	118
PFK3336	AATCCATCGCTCAACAATC	CCACTGACGAGCTTAAAT	55.76	55.58	245
RuBisCo14279	TGAACTGGAGAGTGGAGAT	AATTTTGCTGGCTCTCCT	56.06	56.73	39
RuBisCo944	TGAGGAAGGCATTGTTGT	AATAGAGCACACCCACCA	56.16	55.15	39
RuBisCo753	GGACCGAAAACTTAGAATTG	TTCTGTAGAGCACACCTT	53.99	53.50	354
diTPS24918	GTGAACAAGCTGCAAAGG	TGTTTCATGCAGAGATGGT	57.16	56.19	112
diTPS21613	TTAGAGATGAGAATGGCGA	TCTCCCCGAATGTTCTTT	53.31	54.37	216
diTPS17530	GGATTCGCAGAGCTCAAA	CATGCGCGTCATATAAGT	56.89	55.11	157
diTPS9193	CACCCTCACTTTCCTCAA	GTAGCGAGTAGTCTGTCA	54.82	54.86	107
SUS11741	GGAAATGTTTCTAGGAGG	GGCCAAAATAACCATGAG	52.05	53.41	71
SUS6064	TTTCACCCTTCCTGGTCT	AATATCACCAGCAACCACA	56.21	55.77	337
SOD3240	GAAAGTCGTGAATTGGAAGT	GGAGCGCATGTTCAAATTTA	55.18	55.59	243
SOD1764	GGTAATTCACAAGTCGAG	TTCCATATTTTCCAGGGC	51.83	52.40	454
SOD1790	CTCTTTCTGGGACACATT	CCAGCCTATAAATAACCAAC	53.83	52.84	177
GR3069	AGCTAATGGTGATATCGG	GCAGTACCTTATCTTGGTT	54.16	54.21	433
APX8272	GCCTCTCCAGTCTAATTAC	CAACTCATAAAAGGAACCTC	53.84	53.01	164
APX802	TTTGGGGTTTGCAGTAGG	GTTAGAAGTGGATTACAGCG	56.32	55.56	210
APX345	TTGACAAAGCCAGGAGGA	AAACTAGGATCAGCAAGCA	55.41	55.24	601
ALT4579	CTGGATGTTGGGGAATTA	GTTGATAAACCTCGTCTG	53.62	53.11	187
ALT10393	CTCGATTGACAACATTTCAC	CACTACCAACCTGCAAAC	55.56	55.49	265
AST6060	CAGAAGAAGTTGGATGGT	AGAGACATCCGTTGCAATA	55.22	55.00	117
AST8832	TCCCTTGACACGCTTAGA	CCACTTTTCAGGCCATTT	54.48	55.35	181

AST9313	TTAAAATTGGCACCGCACT	TTGGCACTATGATGTGGTT	56.71	55.78	242
AST10035	GTGTTACTACAGGCAATG	CCAACTGGACCCCTAAAA	53.05	52.75	264
AST12962	GGATATGCTGGTAGTGAA	CTAAAAAGTCTCTCCTCTC	51.25	50.68	67
NAD+2715	ACTGAAGTTGTGGAAGCAAA	GTGCACGGAGAGAAGATT	57.19	56.81	94
AroE16632	ATTGGGCGTTTGAACAGA	ATTATCAACAAGAAGCGGCA	56.05	55.48	142
AroE1082	GTTGATGCGGGAAATAGT	GAATGAGCACAAAAAAGTCC	54.84	54.88	225
6PGD5950	TATGGAGACATGCAGTTG	GCCATAATCATCCTTCACA	53.00	52.02	168
6PGD4229	GGAATCTGAACTTTGGGG	CGGCGGTAAGTATCAAAA	55.20	54.31	253
6PGD2942	GATATGTAGTTATGCGCAAG	AAACCAGGCATATCAGTC	53.48	53.98	396
6PGD823	GGAATCTGAACCTTGGGG	AATTCACCACTCTCCTCC	57.13	55.73	182
RuBisCo3648	GAGGACCGAAAACTTAGA	CATAGCATTATACCCCAC	52.65	52.44	282
diTPS1535	GGAGGTGGTGGAATTCAT	CTGGCTCAAACAAAAAAGAC	55.17	55.73	99
SOD9664	GGAAAGTCGTGAATTGGAAG	GTCAACAACCAAGGAGCA	57.19	57.60	196
SOD5066	TTGGTAGAGCACTTGTTG	CATATTTTCCAGGGCCTT	53.56	53.72	144
SOD4683	GGCATCTCCTCTAACATT	TTGTGGGACCGTTATCTT	53.48	54.80	215
SOD551	GATTTTACAGAGCTCAAGG	TCGAATGTGTCCCAGAAA	52.59	52.81	390
SOD4780	GCTGGAACATATGACAAGAA	TTATCTGTCAACCCCATTC	53.16	52.84	355
SOD3685	AAGGTCGAGATGTGTAAG	CGAGACGGAGCATGATTG	53.86	57.17	127
SOD74	ACAGTTCCCCATAATCAC	AACTAGGATCAGCAAGCA	53.60	54.44	389

Table S2. List of 53 primers with specific PCR product, tested each of them on the same 8 *Pinus cembra* DNA samples (with each pairs' nucleotide succession, melting temperature and length of the amplificated sequence).

Code	Sequence Fwd	Sequence Rev	Melt. temp. Fwd	Melt. temp. Rev	Fragment length (Fwd,Rev)
Myb1372	CTGATTGCAGGGCGTATT	CTTCAACATCATTGCGTCC	57.39	56.17	
Myb4633	AAGCAGCAGCCAACAAAA	GCAAAGCTCCAGTGTAAT	55.80	55.54	349
Myb139	ACTTCGTGATGAAAAGCTG	ATAGTTTTGTTGGCTGCTG	56.11	56.29	159
Myb13629	GAGATTGAAAGCAGAAAAGG	CTGGAACAGGCATAGTTT	54.66	54.29	152
Myb10181	CTGATTGCAGGGCGTATT	CTTCAACATCATTGCGTCC	57.39	56.17	175
Myb4344	ATCCTGGAATGGCAATGT	GCAAAGCTCCAGTGTAAT	56.51	55.54	313
Myb4095	GAGAAAACTTATCAGCATGG	AATATCCCTTTCCCTCAC	53.44	53.10	323
Myb19497	GATTCTCTGTTACTTGCC	TGCTTCTCTCTATAGACC	52.36	51.43	287
Myb11378	CTGATTGCAGGGCGTATT	CTTCAACATCATTGCGTCC	57.39	56.17	172
Myb4661	GAAGACAGAGAAAAATGAGC	ACCAGCGAATCCATTTCA	54.41	56.63	468
Myb2894	GAAGGCAGAGAAAAATGAAC	CTGGAACAGGCATAGTTT	54.93	54.18	188
Myb4667	GAAGGCAGAGAAAAATGAAC	GGAACAGGCATAGTTTTG	54.93	54.32	186
Myb2882	GAGCTAAGCATCTTGTTGG	GTCTTGTTATTTGCTGCTTG	55.48	54.96	321
WRKY20368	GAAGAAACAGACGAAGGA	CACCAAACTATCTTAAGCC	53.44	52.49	246
WRKY4214	GCAAGAAAGCACGTAGAA	CATGGAGATTGCGCTAAA	53.13	53.95	223
WRKY2370	ATGGAGGTGGCAGAAGAT	CCTCCATGAACACTTTTTG	57.84	54.17	364
WRKY16373	TGCACCCATCCAAATTGT	ATTCCCAGTGTCATGTAG	55.75	53.61	306
WRKY10455	GATGTCAAAAAGAAGGATGC	TGTACAAATCCCACCTCC	54.49	55.03	257
WRY5091	AATAAAGGTGCAGAAGGC	CATCTCTTTCCTCCACCA	54.93	54.83	294
WRKY3145	GGAGCAGTTGACCGAAGA	CCGCAGATATGAAACCATT	57.50	55.11	305
WRKY2211	GGATAAGCTTACAGAGGA	AGATTTAGGAACTCGGCA	51.69	54.54	210
WRKY1289	CACAACAAAAGAACTGGG	CACTTGGCCATCATAACT	55.11	55.67	366
MADS15369	GGATATGGAACAAGGTGAA	CTCCTCAACAAAATCAAAGG	53.85	54.29	236
MADS6542	TGGCACTATTCTGGGGAA	CTGATTGCCAACTCCTTT	55.03	54.82	210
MADS1818	CCATCGCTGTTTTCCTTT	CTAACATTTCTCTTCCGGAT	55.29	54.05	361
MADS3172	GGGAGCCTAAAAGTATGA	CAGCCCGCATTTCTTCTA	52.29	55.46	207

MADS20909	CCCAGATGGAAAGACAATA	CCTCCATCAAATCTTTCTC	53.45	53.47	269
MADS2190	ACCTCTGTATTTCCTCTG	CCATGAGCTACCTTCTTT	53.01	53.24	332
MADS2038	GGGAGCCTAAAAGTATGA	CAGCCCGCATTTCTTCTA	52.29	55.46	207
MADS832	CAGAGTGGCGGATGATTT	AGCCACAAGCACAAAAAG	56.75	57.40	323
PAL2688	TTCGAATCTAAGTGGTGG	CCCTTTGTCATAACTTTCTC	52.06	53.18	612
CHS282	GACAAGTCGGCAATAAAG	AGAATCTCCTCCGTCAAG	52.33	56.25	50
PAL1134	TTTCAACAAGATTCCCGC	CAATATATTCGCCCGGAG	54.86	55.84	217
CHS1459	GGACATGGTTGTGGTGGA	GATCTTCGACTTGGGCTG	57.44	57.66	88
F3'H2550	GAAAGACGAGGTGAGCAA	GGTGAAGGGGATAAGAGAAA	55.82	55.77	250
F3H350	GCGTGAGAGACGAAGTAA	GTGAAGGGGATAAGAGAAG	53.85	54.07	254
F3H483	CAATGGCAAGTTCAAGAC	GGCAGACACCCATAAAAA	53.38	54.60	332
IGSTP346	AGAAAATTCCGGTGCTCAT	CTTTCACTCCATCTGCCT	57.67	57.69	462
IGSTP3726	GAGACAACCTGATGCAAA	CAATCTCCCTCTCAAACT	52.66	52.52	142
IGSTP17883	GTTCGCATAGCACTTTCT	CTCGGTCGTATGGATCTT	55.19	55.38	226
CHS381	GACAAGTCGGCAATAAAG	TTGGTACATCATCACTCTC	52.33	52.52	301
CHS1053	GCTAAGGCCATCAAGGAA	CACTGCAGACGACCAGAA	55.22	57.60	243
CHS2058	GCGACAAGTCAGCAATAA	CAGCCGCTTCTTTTCCAA	54.35	56.42	153
CHS4014	CGACAAGTCAGCAATAAAG	ACTGGTAGTGCAGAAAAC	53.68	53.88	220
CHS7594	CATGGTTGTTGTGGAGGT	TTGGTACATCATCACTCTCT	56.62	54.44	196
IGSTP18952	AGCTCATCACTTCTCGAAA	GTACTGGCCATCGTATTTT	54.94	55.04	185
IGSTP3638	GAATATTGTGCAAGGAGAG	GTTCTGAAGAGGTTGTATG	53.61	53.15	232
IGSTP657	TCTATGCGAGGAGTTTTC	TCGGGCATCACTTCTTTC	52.12	55.24	302
3RT9058	TGCAGAAGCCATAACGAA	AATTAGCATATCCCTCGC	53.63	54.24	272
IGSTP9956	GAATCAGAGATAAGTGTGG	TAAGGATTGGTTTTGGGC	53.07	52.97	312
IGSTP9874	GAGTACCTTGTGAATGCT	GTACTGGCCATCGTATTT	54.07	54.22	216
IGSTP985	TTTTGATAATGGAGGTGCC	CAGTCAGACCCAACAAAT	54.32	55.84	369
IGSTP554	CAGGTGGAAGAATGGGTG	GATCCGGCATAACTTCTTT	56.43	54.79	259