Fachgebiet Forstgenetik

Genetic characterisation of populations from the European natural range of Norway spruce (*Picea abies* (L.) Karst.) by means of EST markers

Maryna Valdivia Chevarria

Vollständiger Abdruck der vom Wissenschaftszentrum Weihenstephan für Ernährung, Landnutzung und Umwelt der Technischen Universität München zur Erlangung des akademischen Grades eines

Doktors der Naturwissenschaften (Dr. rer. nat.)

genehmigten Dissertation.

Vorsitzender: Univ.-Prof. Dr. Gerhard Wenzel

Prüfer der Dissertation:

1. Univ.-Prof. Dr. Gerhard Müller-Starck
2. Univ.-Prof. Dr. Gert Forkmann

Die Dissertation wurde am 08.11.2004 bei der Technischen Universität München eingereicht und durch die Fakultät Wissenschaftszentrum Weihenstephan für Ernährung, Landnutzung und Umwelt am 09.03.2005 angenommen.

Partial results of this dissertation were published as follows:

M. Bozhko, R. Riegel, R. Schubert and G. Müller-Starck

A cyclophilin gene marker confirming geographical differentiation of Norway spruce populations and indicating viability response on excess soil-born salinity

Molecular Ecology (2003) **12**, 3147-3155

Acknowledgements

I wish to thank Prof. Gerhard Müller-Starck who gave me possibility to work in the DFG project, commented, edited and added on the chapters of this thesis and the publication in Molecular Ecology.

I am grateful to Dr. Roland Schubert for his supervision, his help in writting of the article concerning cyclophilin marker (common with Dr. Ricardo Riegel). I would like to thank Dr. Roland Schubert and Dr. Ricardo Riegel for the methodological advice in the laboratory and common discussions.

I wish to say "Thank you very much!" to Eliane Röschter for her great help in the lab and supporting in the "fight with the complicated computer programs".

Beatrix Kain and Maria Thole solved all my "bureaucratic" problems. Thanks a lot!

I am thankful to Christiane Bittkau, Sabina La Scala, Rüdiger Baumann, Holger Paetsch, Cristina Valcu, Sheela Veerasenan for their "every day" help and friendly atmosphere in the group.

I thank Dr. Ch. Sperisen, Dr. A. Rigling, Dr. T. Skørppa, Dr. I. Shvadchak, Dr. P. Robakowski, Dr. M Konnert and Dr. F. Bergmann for providing spruce samples.

My mother, my family, my husband Jorge and my daughter Valeria supported me every day, I thank them.

Contents

Abbreviations

AFLP	Amplified fragment length polymorphism
APS	Ammoniumpersulfate
BSA	Bovine serum albumin
cDNA	Complementary (to an RNA) DNA
cpSSR	Chloroplast simple sequence repeats
CTAB	Cetyl-trimethyl-ammoniumbromide
DNA	Deoxyribonucleic acid
dNTP	Deoxyribonucleotide-triphosphate
EDTA	Ethylene Diamine TetraAcetate
EST	Expressed sequence tag
IPTG	Isopropylthiogalactoside
$MgCl_2$	Magnesium chloride
mt DNA	Mitochondrial DNA
PAG	Polyacrilamide gel
PCR	Polymerase chain reaction
RAPD	Random amplified polymorphic DNA
RFLP	Restriction fragment length polymorphism
RNA	Ribonucleic acid
SAMPLs	Selective amplification of microsatellites polymorphic loci
SCAR	Sequence-characterised amplified region
SSCP	Single stranded conformational polymorphism
SSR	Simple sequence repeat
STS	Sequence-tagged-site
TBE	Tris-borate EDTA
TE	Tris ethylenediaminetetraacetic acid
TEMED	Tetramethylenediamine
UV	Ultraviolet

Figures

Tables

1 Introduction

1.1 Genetic variation of forest trees

The majority of forest stands reveal high levels of genetic variation compared to other organisms (Hamrich and Godt, 1989; Müller-Starck, 1995). This realized amount of genetic variation is considered as a necessity for the maintenance of the long-lived tree populations under heterogeneous environmental stresses and climate changes (Gregorius, 1991; Müller-Starck, 1995).

Many traits of forest trees, in particular, the phenological traits, show great variation between populations but little intrapopulational variation. The most well-known among them follow latitudinal, altitudinal or longitudinal gradients (Skrøppa *et al*., 1995; Skrøppa and Johnsen, 2000). On the other hand, genetic traits such as nuclear DNA markers or biochemical genetic markers clearly indicate greater variation between trees of the same stand than their variation among populations (Bergmann, 1991; Lagercrantz and Ryman, 1990; Müller-Starck and Ziehe, 1991; Müller-Starck, 1995; Ouborg *et. al*., 1999; Sperisen *et al*., 2001; Sunnucks, 2000; Vendramin *et al*., 1999). This apparent contrast gives expression to the action of the different evolutionary forces effective on the molecular and phenotypical levels. While the phenotypical traits are often a result of adaptation to the prevailing environmental conditions, the vast majority of the molecular markers remain neutral. The diversity of the first group rather reflects the pressures of natural selection, while that of the second reflects the history of populations (effects of migration, genetic drift).

Today the genetic variation of European forest tree species are primarily dependent on the following main factors (Bergmann, 1991):

1. The **number and types of glacial refugia** generating genetic differentiation due to mutations, isolation, selection and genetic drift.
2. The **migration routes** during postglacial periods promoting genetic differentiation through environmental adaptation if refugial populations remained separated or diminishing genetic differentiation, if refugial populations have met and exchanged genetic material.
3. The **human activities** in the last centuries in many European countries, of which intensive exploitation of tree species, replacement of native stands by artificial

plantations and introduction of exotic populations, races and species have distributed the natural structures.

Since the efficacy of these three factors may considerably differ among the European tree species, their genetic variation patterns may differ accordingly (Bergmann, 1991; Petit *et al*., 2003). In addition to these factors, some life-time fecundity and the seed dispersal mechanisms, should also influence the level and distribution of genetic variation.

The knowledge of the species-specific genetic variation pattern is an essential prerequisite for managing and preserving forest gene resources, which were and are threatened by drastic man-made environmental changes, such as local pollution (Scholz *et al*., 1989) or the inevitable, global climatic shift (Ellis, 1990; Hattemer and Gregorius, 1990). Therefore, the level of intrapopulational diversity, the degree of interpopulational differentiation, the occurrence of local alleles and further information about the genetic system are subjects requiring analysis.

1.2 The role of late Quaternary climatic changes in present plant distribution in Europe

The period of strong climatic oscillations, with a dominant series of cold and dry glacial intervals (each lasting approximately 100 000 years), began 700 000 years ago (Petit, 2001). It was interrupted by shorter intervals (approximately 10-20 thousand years) of warmer and moister interglacial climates (Webb and Bartlein, 1992). Apart from these major climatic oscillations, climatic instabilities have been documented for the last full glacial period (approximately 115 000 years ago until 10 000 years ago). Vast areas in northern Europe were repeatedly covered by massive ice-sheets, although independent centres of glaciations also occurred at low-latitude Mountains, such as the Alps (Bond *et al*., 1993). Quaternary paleoecology in Europe has revealed a series of southward range contractions of plant species during the last glacial period followed by rapid northward range expansions in the wake of deglaciation (Hewitt, 1996; Riddle, 1996).

In Europe, geographic mapping of radiocarbonated pollen spectra has been invaluable for monitoring and changing distribution limits of plant species and shifts in their abundance. Detailed pollen map of European tree species (Huntley and Birks, 1983) shows that forest tree species spread individually in response to postglacial climatic change. Range of expansions occurred at different times, at varying rates, and in different directions. Fossil data alone, however, do not always allow a correct and complete interpretation of the

biogeographical history of species or entire plant biotas (Comes and Kadereit, 1998). Recently, different molecular markers, such as cpDNA (chloroplast DNA), mtDNA (mitochondrial DNA) and nuclear DNA markers were used for the analysis of intraspecific historical events with respect to angiosperms and gymnosperms. The studies of number of European trees using molecular markers allowed recognizing of possible glacial refugia and colonization routes of these species as well as the centres of gene diversity (Petit *et al.*, 2003).

In addition to the provision of evidence for or against the location of previously recognized (by fossil analysis) refugia and postglacial spread, molecular studies have also postulated formerly unknown refugial areas by pinpointing locations with high molecular diversity and a high proportion of unique alleles or haplotypes (Demesure *et al*., 1996; King and Ferris, 1998).

Cycles of contraction/expansion of geographical ranges following Quaternary climatic changes did not only infer the geographic distribution of intraspecific polymorphisms, but also affected patterns of genetic diversity (Hewitt, 1996). An increasing number of molecular studies provide substantial evidence that putative refugial plant populations harbour higher levels of genetic diversity relative to their likely descendant populations (Comes and Kadereit, 1998). Such a reduction in genetic diversity with increasing distance from a refugium is a general phenomenon to be expected from repeated population bottlenecks at an advancing edge of a range in any species during postglacial expansion (generally towards the north) (Hewitt, 1996). On the other hand, recent studies in angiosperms demonstrate that during re-immigration genetic diversity can increase as a consequence of fusion of routes (Petit *et al*., 2003). On a local scale, postglacial spread following such a mode of colonization is likely to result in high levels of interpopulational (and low levels of intrapopulational) diversity for maternally inherited genes (Petit *et al*., 1996). Higher levels of genetic diversity in refugia may simply reflect more stable population dynamics and larger population sizes. In cases where refugial areas coincide with mountainous regions, such as in southern Europe, the risk of population bottlenecks may be reduced by limited altitudinal migration in response to climatic warming and cooling (Comes and Kadereit, 1998). In any event, if interglacial range expansion of plant species results in extinction in situ at the end of the interglacial, it seems likely that any genetic diversity that arises during an interglacial range expansion will be lost at the end of the interglacial. By contrast, genetic diversity originating in refugia may accumulate over

several climatic oscillations. Following this line of reasoning, populations in long-term refugial localities (such as in southern Europe) are likely to have maintained most of the species extant genetic variation (Comes and Kadereit, 1998).

Intraspecific differentiation of *Picea abies* detected by allozymes was reported (Lagercrantz and Ryman, 1990). These differentiation profiles have been interpreted to be compatible with historical events related to approximately 40 000 years ago (Lagercrantz and Ryman, 1990; Comes and Kadereit, 1998).

1.3 *Picea abies* (L.) Karst.

Norway spruce is a coniferous species belonging to the family Pinacea, genus *Picea* and known also as Norway spruce. Norway and Siberian spruces are considered to be different but very closely related species with a wide zone of introgressive hybridization along both sides of the Ural Mountains (Pravdin, 1975).

1.3.1 History and distribution of *Picea abies* in Europe

Picea abies (L.) Karst. belongs to gymnosperms that represent an ancient group of plants first recorded as fossils in the Upper Devonian (350 million years ago) (Biswas and Jorhri, 1997).

Today, the natural range of Norway spruce is divided into three major domains, i.e. the Alpine, Hercyno-Carpathian and Baltic-Nordic domains, which are generally considered to be the result of postglacial recolonization from three putative refuges located in the Dinaric Alps, in the Carpathians and in the Moscow area (Huntley and Birks, 1983; Schmidt-Vogt, 1986).

The Alpine domain of Norway spruce distribution represents mainly mountainous and sub alpine locations. These are the Alps, the Alpine foreland, the French and Swiss Jura, The Black Forest, Bavarian Forest and several mountain-ranges of the Balkan Peninsula including the Rhodopes. The Italian distribution of natural stands of Norway spruce spans the entire Alpine range, and includes a single stand located in the northern Apennines, near Campolino (Scotti, 2000).

The Hercyno-Carpathian domain is area covering Carpathian Mountains, the Thuringian Forest, the Harz, the Bohemiam Forest as well as the adjoining ranges of distribution in lower altitudes.

The Baltic-Nordic area of spruce distribution is Northeast European. This area covers northern-Baltic, the Scandinavian-northern Sarmatian and adjoins to the north and east of the Ural regions. Norway spruce first occurred in Scandinavia approximately 2,500 years ago; its immigration from Europe is attributed to colder Scandinavian winters coupled with increased precipitation and storm events which allowed Norway spruce to colonize areas that were formerly too dry (Bradshaw, 2000).

The Central- and Southeast European spruce area is separated from the Northeast European spruce area by the Central Polish disjunction. It was for a long time a matter of discussion, whether this disjunction was devoid of spruce at all times.

Today, the most widely held opinion is, that a zone of contact where spruce migrations from the Russian refuge and from the refuges of the Carpathian area met during a warmer postglacial period, stretches across the spruce less strip (Schmidt-Vogt 1986).

Within its native range, the life span of Norway spruce is 200 years, and reaches 400 years at the northern limits of its range. Norway spruce occurs at elevations up to 6,560 feet (2,000 m) in the Bavarian Alps, up to 4,920 feet (1,500 m) in the Black Forest (Köstler, 1956).

Picea abies has been widely sown and planted in Central Europe, particularly during the last two centuries, so that large areas with artificial forests exist outside the natural habitat (Schmidt-Vogt, 1986).

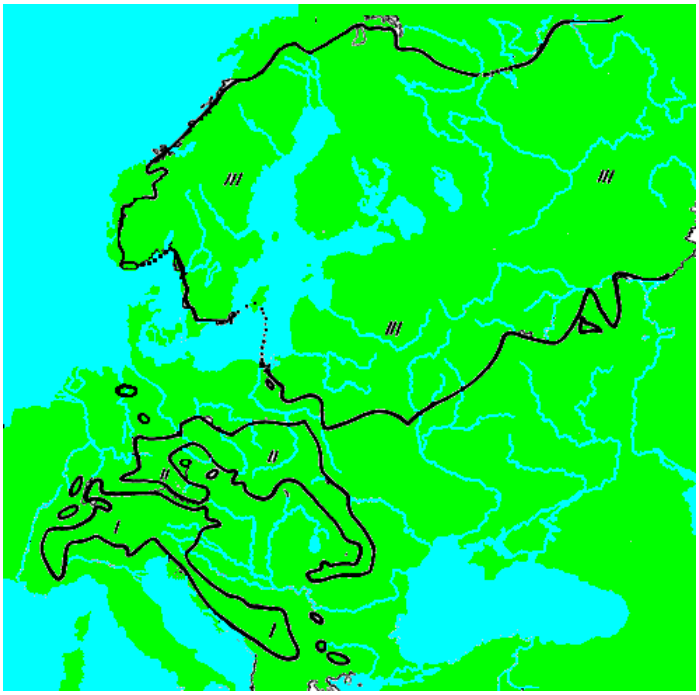


Figure 1-1. Distribution of *Picea abies* in Eurasia (Schmidt-Vogt, 1986)

I = **Alpine** - south-eastern European region
II = **Herzyno-Carpathian** region
III = **Nordic - Baltic** region

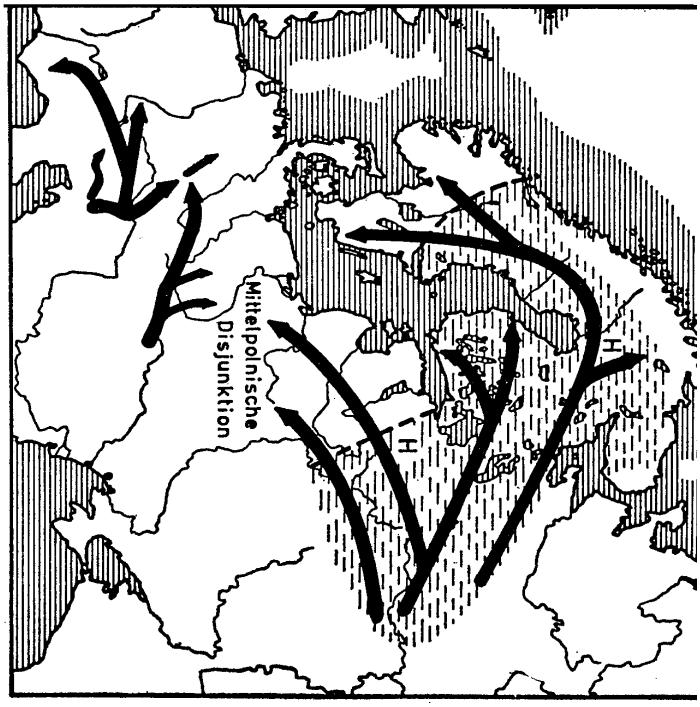


Figure 1-2. Postglacial migration of *Picea abies* in Europe (Schmidt-Vogt, 1986)

1.3.2 Ecological and economical significance of *Picea abies*

Norway spruce is an evergreen tree occurring within its native range in pure stands, transitional stands mixed with Scotch pine (*Pinus sylvestris*), or mixed stands with European beech (*Fagus sylvatica*) and European silver fir (*Abies alba*). Scattered Norway spruce occurs in sereal stands of European aspen (*Populus tremula*) or hairy birch (*Betula pubescens*). Common groundlayer species include bilberry (*Vaccinium myrtillus*), lingonberry (*V. vitis-idaea*), heather (*Calluna vulgaris*), and woodsorrel (*Oxalis spp.*). Mature Norway spruce forests typically have very little ground layer vegetation.

Norway spruce grows best in cool, humid climates on rich soils preferable including well-drained sandy loams. Soils under Norway spruce stands are often more acidic than soils under other species. Soil acidity appears to increase with stand age as soil buffering capacity decreases with age. Growth rates increase with increased soil organic material and are positively correlated to the nitrogen content of the soil.

In Central Europe, *Picea abies* reaches heights of up to 61 m; the range is usually between 30-35 m. Norway spruce seeds are wind dispersed, but do not usually travel much farther than the height of the parent tree. Movement after dispersal, however, can be considerable when seeds are dispersed onto crusted snow and are pushed along on the surface by wind.

Norway spruce cones are conspicuously large (10-18 cm long). They open from May to June. Seeds ripen in late autumn the same year. They are released on warm days in late autumn and winter, but are sometimes retained until spring. Due to the tendency of flat rooting of many sites, *Picea abies* is not wind firm and is also subject to snow break. It provides important winter cover for a number of species.

In Europe, *Picea abies* is the focus of increasing concern about forest decline. It is exhibiting a specific set of symptoms. These symptoms include needle chlorosis combined with magnesium deficiency and losses of needles in parts or all over. Explanations usually centre on air pollution (ozone, acid deposition, or toxic metals contamination) coupled with acidified, depleted soils that cause, among other problems, foliar magnesium deficiency.

Picea abies is one of the most common and economically important coniferous species in Europe and Scandinavia. Softwoods account for 80% of lumber consumed in Europe. Norway spruce wood is strong, soft, straight- and fine-grained, and easily worked (Safford 1974). It is widely used for construction, pulp, furniture, and musical instruments. Norway spruce is widely planted for windbreaks and shelterbelts, especially in humid, severe-winter regions, as well as Christmas trees and as an ornamental. Norway spruce resin has been used to make Burgundy pitch, and the twigs used to make Swiss turpentine, the needles and twigs used to make antiscorbutic and diuretic beverages (Safford 1974).

1.4 Genetic characterisation of *Picea abies*

1.4.1 Genome

The gymnosperms are characterized by large genomes and relatively low chromosome numbers (Murray, 1998). The high presence of repetitive DNA is also peculiar to most of gymnosperms (Murray *et al*., 2002). Generally, plants with large genomes were found to posses a much higher proportion of repeated DNA sequences than did those with smaller genomes (Dean and Schmidt, 1995). Reassociation kinetic studies reveal 75% of the gymnosperms genome to be repetitive (Dhillon, 1987).

The size of nuclear genome of Norway spruce is 3-4 x 10^{10} bp (Govindaraju and Cullis, 1991). Significant part of the nuclear genome is represented by no coding DNA (Schmidt *et al*., 2000). Chromosome number is 24 (2n) (Khoshoo, 1961). The structure and composition of Norway spruce genome as a whole is not known because of its large size and high presence of repetitive DNA. Several classes of repeated sequences are observed

within nuclear genome of *Picea abies*. Tandem repeats consisting from rDNA (ribosomal DNA), satellite DNAs representing short tandem repeats with a high degree of length variability, microsatellite repeats (mono-, di-, or trinucleotide repeats) (Scotti *et al*., 2002), dispersed repeated sequences of active and inactive derivatives of retrotransposable elements (Friesen et al., 2001; Murray et al., 2002) are found in Norway spruce.

Genetic map of Norway spruce (http://www.pierroton.inra.fr/genetics/Picea) has been constructed using a combination of marker types including RAPDs (Binelli and Bucci, 1994; Bucci et al., 1997), AFLPs, selective amplification of microsatellites polymorphic loci (SAMPLs) and SSRs (Paglia et al., 1998). Individual cDNA-based STS markers conserved across other species within the Pinaceae are incorporated into such maps.

CpDNA (chloroplast DNA) shows paternal inheritance in Norway spruce (Sutton *et al*., 1991). The chloroplast genome of *Picea abies* represents one master chromosome of the size about 120 kb that is peculiar for most of the conifers (Strauss *et al*. 1988; Raubeson and Jansen, 1992). Conifer cpDNA contains dispersed repetitive DNA that is associated with structural rearrangements (Hipkins *et al*., 1994). Chloroplast genome inversions and length mutations often occur in localized areas of the genome containing repeated sequences and tRNA genes (Hipkins *et al*. 1994). The presence of the genes required for dark synthesis of chlorophyll (a function that appears to be present in conifers and absent in other land plants) was identified in cpDNA of *Picea abies* (Lidholm and Gustafsson, 1991).

Maternal inheritance of mitochondrial DNA was found in *Picea* (Sutton *et al*., 1991). Mitochondrial DNA size and structure remain relatively unstudied in conifers and other gymnosperms. However, results to date indicate that the genome is large, contains dispersed repetitive DNA, exhibits a low nucleotide substitution rate (Hipkins *et al.*, 1994, Sperisen *et al*., 2001)). Intraspecific mtDNA variation in conifers appears to be generated by recombination among repeated sequences resulting in complex insertions/deletions or rearrangements (Hipkins *et al.*, 1994). Numerous editing events of RNA have been observed in mitochondria of Norway spruce (Hiesel et *al.*, 1994).

1.4.2 Development and application of the genetic markers in *Picea abies*

Genetic markers are defined as heritable polymorphic characters that simply reflect differences in DNA sequences directly at the nucleotide level or indirectly at the level of gene expression.

Morphological and biochemical markers are widely applied to study intraspecific genetic variation. Several population genetic studies concerning gymnosperms have demonstrated existence of great intrapopulational variation and presence of genetic differentiation among races, subspecies and geographical groups of populations within the same species (Müller-Starck *et al.*, 1992). During the last years, DNA markers get a leading position in the genetic characterisation of forest trees species because of their multiformity, affecting of different parts of both nuclear and organelle genomes and owing to development of variety of molecular biological methods. Most molecular markers measure apparently DNA variation and are being useful in the analysis of phylogenetic relationships, population structure, mating system, gene flow, parental assignment, introgressive hybridisation, marker-assisted selection and genetic linkage.

Large genome size and high proportion of repetitive DNA restrict identification and application of molecular markers to study conifer species (Pfeiffer *et al.*, 1997).

Isozymes

Isozymes (isoenzymes) represent biochemical class of genetic markers and appear structurally different species-specific molecular forms of an enzyme system with, qualitatively, the same catalytic function. Isoenzymes originate through amino acid alterations which cause changes in net charge, or the conformation of the enzyme molecules and therefore in their electrophoretic mobility. Isozymes mark allelic variation at single structural gene loci by difference of alleles in electrophoretic mobility. They are suitable as genetic markers and widely used to study biodiversity (Müller-Starck and Ziehe, 1991; Müller-Starck, 1998). Isoenzymes possess several advantages as genetic markers: they affect coding genes, are codominantly expressed and easy in performing. The restricted number of polymorphic functional gene loci and specifity of expression limit the application of these markers.

Using isozymes, genetic variation of *Picea abies* was examined along natural European range by Lagercrantz and Ryman (1989), in Germany by Bergmann (1991), in Italian

populations by Morgante and Vendramin (1991), on the high and low elevations of Switzerland by Müller-Starck (1995b). It has been shown that Norway spruce, like other conifers (Hamrick *et al*., 1983; Mitton, 1983; Loveless and Hamrick, 1984; Ledig, 1986), exibits a relatively large amount of genetic variability within populations and little differentiation among populations (see *Tab.1-1.*). A clear geographic pattern in the variation of allele frequencies was found by Lagercrantz and Ryman (1989), who analysed 70 populations from European natural range of Norway spruce using 22 isozyme loci. A striking correspondence between the linear pattern of genetic differentiation identified by the principal component analysis and a line connecting the two general presumed refugia was detected. The central European provenances were supposed to have originated from refugia in the Carpathians and the Dinaric Alps. The provenances from Eastern Poland, where spruce from the eastern and western refugia are supposed, to have met, were found genetically intermediate between those from locations closer to the both presumed refugia.

Molecular markers

Recent advances in molecular technology have greatly increased the number of genetic markers available for assessing geneticvariation. DNA markers based on the polymerase-chain reaction (PCR) require very small amounts of DNA and may be essentially unlimited in number. Also, unlike isozymes markers, they can provide direct access to the genomic regions that they mark. Molecular DNA markers address the nuclear genome as well as chloroplast and mitochondrial genomes. Chloroplast and mitochondrial DNA markers carry important information for phylogeographic investigations and studies of species history due to their uniparental inheritance, and sensitivity of haploid genomes to the effect of random drift because of the absence of recombination (Petit *et al*., 1993; Ennos, 1994). Nuclear markers represent interest for study of selection and adaptation as well as for creating of genetic map. Several kinds of molecular DNA markers were developed for *Picea abies* during last years.

RAPD

The random amplified polymorphic DNA (RAPD; Williams *et al*. 1990) is a technique by which individuals can be scored at a large number of loci that are randomly distributed within the genome. RAPD technique requires only the presence of a single „randomly chosen“ oligonucleotide which acts in PCR as both a forward and reverse primer. RAPD markers are characterised by simplicity of technique and high level of detected polymorphism. However, some authors report about difficulties in reproducibility of these

markers in different laboratories. The most part of RAPD markers represent anonymous regions of the nuclear genome.

In the study of Scheepers *et al*. (1997) it was showed that RAPD method can be used to check the clonally identity of *Picea abies* plants propagated from cuttings. In RAPD investigation of Collignon and Favre (2000), who studied marginal branch of spruce distribution area in France, Norway spruce appears to maintain a high level of diversity within populations and low level of inter-populational differentiation (see *Tab. 1-1*). The pattern obtained by the same authors on the basis of RAPDs results using cluster analysis revealed a clear longitudinal gradient across the large area sampled and supported the hypothesis of re-immigration of *Picea abies* in an East-West direction from Eastern Europe to France (Collignon and Favre, 2000).

SSRs

Simple sequence repeats (SSRs, microsatellites) consist of tandem of repeated units, each between 1 and 10 base-pairs in length, such as $(TG)_n$ or $(AAT)_n$. They are widely dispersed throughout eukaryotic genomes and are often highly polymorphic due to variation in the number of repeat units (Elsik and Williams, 2001; Schmidt *et al*., 2000).

Microsatellites are highly informative markers for genome analysis and mapping because of their codominant character, widespread occurence, hypervariability in repeat number and their interspersed multilocus distribution in eukaryotic genomes. As reported by Pfeiffer *et al*. (1997), microsatellites developed for Norway spruce reveal about five times more genetic types as isozymes and could provide valuable source of markers for genome mapping and genetic diversity studies. However the development of microsatellite markers can be a time-consuming process in conifers (Scotti *et al*., 2002). Microsatellite markers are reported for *Picea abies* (Pfeiffer *et al*., 1997; Scotti *et al*., 2002). The presence of dinucleotide microsatellites in expressed regions of the genome of Norway spruce was shown by Scotti *et al*. (2000).

Nuclear SSRs are multiallelic and highly variable within and among populations, and inherited in a codominant Mendelian manner (Morgante *et al*., 1993). Chloroplast SSRs have advantageous in the study of species history and geographical patterns of genetic variation because of their uniparental inheritance and the absence of recombination within haploid genome (Petit *et al*., 1993). The presence of high polymorphic microsatellite

regions has been documented in the chloroplast genome of *Picea abies* (Bucci *et al.*, 1998; Vendramin *et al*., 1996; Vendramin *et al*., 1999).

In the study of Vendramin *et al*. (1999) based on chloroplast microsatellites (cpSSRs) in Norway spruce, the distribution of some haplotypes showed a clear geographic structure and seems to be related to the existence of different refugia during the last glacial period. The overall picture was very similar to that obtained by Sperisen *et al.* (1998) who found specific haplotypes for the above-mentioned regions using specific mitochondrial DNA markers (see subsequent topic). On the basis of principal component analysis, good separation between two main groups of populations was recognized: Sarmathic-Baltic population group including stands from Scandinavian and Sarmathic regions (Fennoscandia, Baltic Republics, Russia, and eastern Poland) and Alpine-Centre European group, including populations from Italy, Switzerland, Austria, and Slovenia. Moreover, some stands from Western Alps and Northern Apennines showed to diverge to some extent from the Alpine-Central European group. This picture reveals similar trends than the one reported by Lagercrantz and Ryman (1990) using isozyme loci. Bucci and Vendramin (2000) found direct evidence of large-scale geographical structure of cpSSR haplotype frequency over the European natural range of Norway spruce using geostatistical methods.

Specific mitochondrial DNA (mtDNA) markers

mtDNA is maternally inherited in *Picea abies*, unlike cpDNA, which is inherited paternally (Birky, 1995). Variants of mtDNA are thus dispersed by seed, and not by pollen, so that the resulting geographical distributions mirror gene flow through migration.

Two polymorphic tandem repeats in the second intron of the mitochondrial NADH dehydrogenase subunit 1 (*nad1*) gene of Norway spruce were investigated by Sperisen *et al.* (2001). These authors show a clear separation of the tandem repeat marker into two phylogenetic lineages (North/North-East Europe and South-East/Central Europe). The survey of a mitochondrial tandem repeat in *Picea abies* also revealed a pattern of genetic variation which showed a genetic differentiation between Eastern and Western populations within the Alpine range due to completely monomorphic Western Alpine and slightly to highly polymorphic Eastern Alpine populations (Gugerli *et al*., 2001). Population differentiation based on frequencies of mtDNA size variants is high: G_{ST} is 0,676 and indicates that a large proportion of the total genetic variation is attributed to among-population variation. This value of differentiation of Norway spruce is lower than the one reported for paternally inherited cpDNA microsatellites (Vendramin *et al*., 1999), but

higher than the value observed analysing biparentally inherited isozyme loci (Müller-Starck, 1995b; Lagercrantz and Ryman, 1990) (see *Tab. 1-1*).

STS and SCAR markers

Sequence-tagged-site (STS) markers differ from so-called anonymous molecular markers such as RAPD by sequence characterized (and often functionally identified) region of PCR-amplification. The data obtained from genomic clones or cDNA libraries are used for STS-marker development. The STS polymorphism can be observed without manipulations with amplified products or using such techniques as PCR-RFLP, SSCP. Also a number of molecular markers (RAPD, SSRs) can be converted into sequence-characterised amplified regions (SCAR) using standard methods (Paran, Michelmore, 1993).

STS markers have been developed for many plants (Bradshaw *et al*., 1994; Talbert *et al.*, 1994; Tragoonburg *et al*.,1992) including the conifers *Cryptomerica japonica* (Tsumura *et al*.1997), *Picea mariana* (Perry, Bousquet 1998) and Picea abies (Perry *et al.* 1999).

Scotti *et al*., 2000 proved seven SCAR/STS markers to detect genetic variation in 8 of Italian populations of Norway spruce. A significant excess of heterozygotes in most populations and high overall level of genetic diversity due to population differentiation were detected.

Perry *et al.*(1999) examined the amount and nature of variation revealed by STS markers in Norway spruce. The nine markers that showed exclusively codominant polymorphisms in Norway spruce had an average observed heterozygosity of 0.30 which is similar or greater than those revealed by polymorphic isoenzyme markers in Norway spruce.

Table 1-1.. Examples for genetic variation observed within and among populations of *Picea abies* by different genetic markers. [1)]H_O : average of observed heterozygosity; [2)]H: Shannon´s diversity; [3)]S_W : within population variance; [4)]G_{ST}, [7)]F_{ST} : differentiation between populations; [5)] Φ_{ST} : differentiation among massifs; [6)]R_{ST} : overall divergence among populations

Marker system	Sampling	Examined populations	Analysed loci	Variation within populations	Differentiation among populations
Isoenzymes (Lagercrantz & Ryman, 1990)	European range; provenance test	70	22	H_O [1)] = 0.115;	G_{ST} [4)]= 0.052;
(Müller-Starck, 1995)	Alpine range; natural populations	20	18	H_O [1)] = 0.226;	G_{ST} [4)] = 0.043;
RAPD (Collignon & Favre, 2000)	French massifs (Alps, Jura, Vosges) and central Europe; clonal & provenance tests	77	31	H [2)] = 0.675-0.718;	Φ_{ST} [5)] = 0.063;
cpSSR (Vendramin *et al.*, 1999)	European range; clonal & provenance tests	97	3	S_W [3)] = 0.635;	R_{ST} [6)] = 0.0997;
mtDNA (Gugerli *et al.*, 2001)	Alpine range; natural stands	36	1	S_W [3)] = 0.592	F_{ST} [7)] = 0.41
SCAR (Scotti et al., 2000)	Italian range; natural range	8	7	H_O [1)] = 0.30	F_{ST} [7)] = 0.118

1.5 EST markers as a new tool in population genetics of forest tree species

These markers are alternative to complete genome sequencing in trees and many other organisms, which are based on identifying only DNA that refers to coding genes that are expressed. ESTs, are partial or complete sequences of complementary DNA (cDNA) obtained from mRNA which is isolated from different tissues and therefore represent genes expressed in these tissues with often known or suggested function. These DNA sequences and appropriate amino acid matches are compared with all other sequences in gene data bases in order to identify matches likely representing highly homologous genes. If there is a high similarity to some other gene sequence, whose identity has been determined, then the identity of the EST can be immediately inferred.

Only small number of the molecular markers are useful for measuring of adaptive genetic diversity. The marker ideal for estimating of adaptive variation should meet the following criteria: 1) be directly involved in the genetic control of adaptive traits; 2) have identified DNA sequence and known function; 3) have easily identifiable allelic variation. No marker fully satisfies all these criteria. However, the new EST marker (expressed sequence tag) seems to satisfy most of them.

Large libraries of partial or complete sequences of thousands of expressed genes have already been obtained. Global, multi-tissue EST projects have been reported for *Arabidopsis* (Delseny *et al*. 1997), rice (*Oriza sativa*) (Ewing *et al*. 1999), maize (*Zea mays*) (Fernandes *et al*. 2002), and soybean (*Glycine max* L. Merr.) (Shoemaker *et al*., 2002). Tissue-specific EST projects have been reported for root hair enriched *Medicago truncatula* tissue (Covitz et al. 1998), flower buds of Chinese Cabbage (*Brassica campestris* subsp. *pekinesis*) (Lim et al. 1996), and wood-forming tissues of poplar (Populus spp.) (Sterky *et al*. 1998).

ESTs provide an opportunity to study gene evolution, to make comparative analyses between genera, and when coupled with genetic mapping, to identify candidate genes for important biological processes and phenotypes (Hatey *et al*., 1998). Large libraries of partial or complete sequences of expressed genes have already been obtained, and data bases of EST are available for several tree species. ESTs can be used as a source for identifying candidate genes for QTL involved in genetic control of adaptive traits.

EST polymorphisms are derived from ESTs. PCR (Polymerase Chain Reaction) primer pairs designed using EST sequences provide amplification of unique sequence tag for the gene (Harry *et al*. 1998). Polymorphism is generally detected as a size difference of the amplified products or using RFLP procedure. Allelic polymorphism in the amplification product (ESTs) can be revealed using different modern methods for detection and visualisation of DNA alterations (Kristensen *et al*. 2001). Such PCR-based approach requires sequence information, but does not require cloning. The design and creation of good primers may involve a significant investment. ESTs mostly reveal genetic variation within genes, although variation can be found in both coding and non-coding regions of genes. Thus, ESTs are the most informative markers in terms of gene function among the most recently developed one and are the first genetic markers that offer real potential for detecting adaptive genetic diversity broadly.

For the present, ESTs databases are available for several conifer tree species, such as Monterey or radiata pine (*Pinus radiata*), loblolly pine (*P. taeda*) (Temesgen *et al*. 2001), Douglas-fir (Krutovskii *et al*. 2002) and Norway spruce (Schubert *et al*. 2001). Harry *et al*. (1998) outlined a foundation for developing codominant PCR-based genetic markers from loblolly pine ESTs.

1.6 Objectives

Objectives of this thesis were:

- EST marker development for *Picea abies* (L.) Karst.
- Monitoring of genetic variation in *Picea abies* along its natural European range by means of newly developed and already existing EST markers

2 Materials and methods

2.1 Population sampling

2.1.1 Control population of the forest of Kranzberg

The forest of Kranzberg is located near Freising (Bavaria) (for geographical location see *Tab. 2-1*). Bud samples of 18 spruce trees collected in the forest of Kranzberg were used for development of EST-PCR markers. The corresponding single tree seed probes were sampled in order to test the mode of inheritance of newly developed EST marker. Buds and seeds were stored at –20°C and +4°C, respectively, until use. Additional bud samples from 110 randomly selected spruce trees growing in the forest of Kranzberg were collected later (Schubert 2001). This material was utilised in order to examine genetic variation of the Norway spruce population of Kranzberg by newly developed EST-PCR markers.

2.1.2 Populations from European natural range

Material from 18 European populations was used in order to study the genetic variability of Norway spruce by means of EST-PCR markers.

Naturally regenerated Norway spruce populations were sampled from different geographical sites along the European natural range of this species. All samples were kindly provided by the people and institutions surveyed in Appendix A. Single tree samples of twigs with buds were stored at –20°C and later used for DNA isolation.

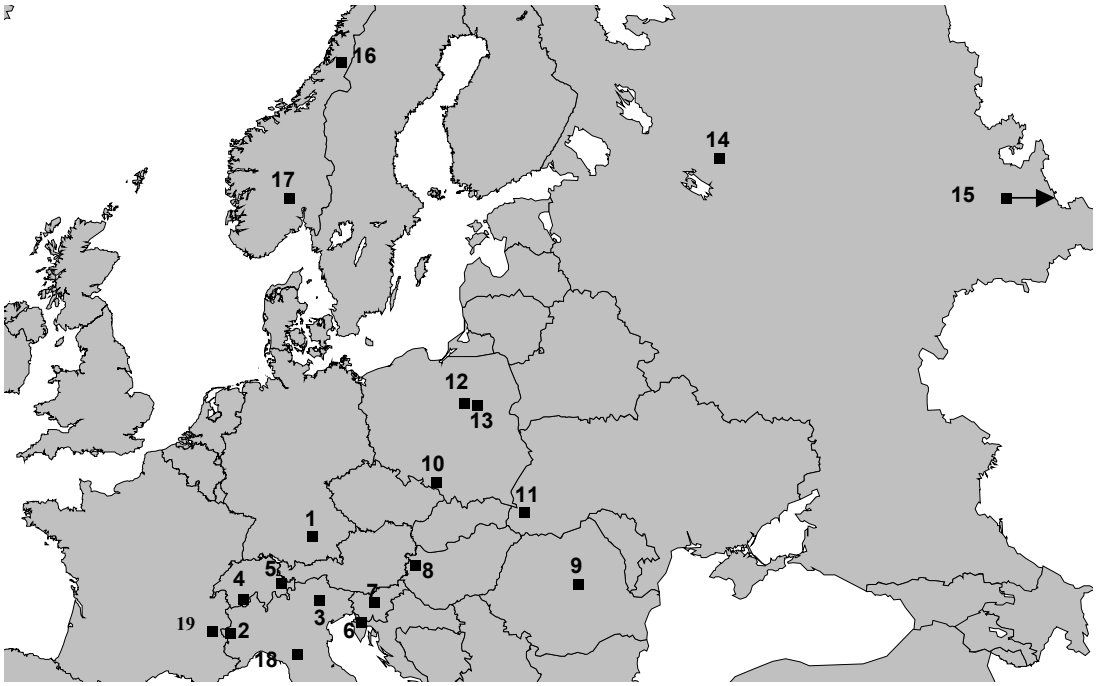
Populations from three general European domains of geographical distribution of *Picea abies* were analysed. South-European Mountains (Alpine domain) was represented by Italian, Swiss, French, Croatian, Slovenian as well as German populations. East-European natural area of Norway spruce (Hercyno-Carpathian domain) was characterized by populations from Ukraine, Poland, Hungary, and Romania. Population stands from Russia and Norway have constituted North-European part of *Picea* abies natural range (Nordic-Baltic domain). The geographical location of all analysed population stands is represented in *Tab. 2-1* and *Fig. 2-1*.

For each population 48 trees were tested.

Table 2-1. Local characterization and sample size of 19 populations samples of *Picea abies*

Population No.	Country	Longitude	Latitude	Altitude M a.s.l.	Sample size
1	Germany	11° 38' E	48° 24' N	500	128
2	Italy	6° 53' E	45° 4' N	1525-1700	48
3	Italy	11° 44' E	46° 17' N	1520-1550	48
4	Switzerland	9° 52' E	46° 47' N	1620-1720	48
5	Switzerland	8° 3' E	46° 23' N	1000-2000	48
6	Croatia	14° 33' E	45° 31' N	1100	48
7	Slovenia	15° 1' E	46° 3' N	1000-2000	48
8	Hungary	16° 31' E	47° 39' N	420	48
9	Romania	25° 46' E	46° 39' N	920-1150	48
10	Poland	18° 95' E	49° 57' N	600-640	48
11	Ukraine	24° 27' E	48° 19' N	1380-1400	48
12	Poland	23° 46' E	52° 40' N	110-120	48
13	Poland	23° 43' E	52° 42' N	110-120	48
14	Russia	40° E	59° 2' N	160	48
15	Russia	121° 25' E	56° 35' N	360	48
16	Norway	13° 21' E	65° 56' N	80	48
17	Norway	10° 43' E	60° 46' N	300	48
18	Italy	10° 42' E	44° 7` N	1450-1589	48
19	France	7° 2' E	45° 45' N	1525-1700	48

Figure 2-1. Geographical location of 19 studied Norway spruce populations (the arrow indicates that population number 15 is situated outside the map)



2.2 DNA extraction

Total genomic spruce DNA was isolated from single buds and corresponding megagametophytes according to the method of Doyle and Doyle (1987) as well as using DNeasy Plant Mini Kit, Qiagen, according to instructions of the manufacturer (see Appendix B). The method of Doyle and Doyle is based on the lysis of cells by CTAB (Cetyl-trimethyl –ammoniumbromide), a detergent which selectively precipitates DNA while maintaining the solubility of many polysaccharides (see *Note 1*).
Megagametophytes were removed from the embryos and the seed coats prior to extraction. Approximately 3 buds per individual, depending on their size, were used.

Note 1. Method of DNA extraction from the buds and megagametophytes (Doyle and Doyle, 1987)

-Buds (megagametophyte) tissue ground in 1.5 ml microfuge tube under 300µl of preheated 2% CTAB buffer (cell lysis buffer) that contains 0.2% 2-mercaptoethanol, 0.5 mg/ml proteinase-K and 1 mg/ml RNase A, and homogenize

- 400 µl of CTAB buffer add to the mixture; incubate the mixture in a 65°C water bath for 30 min

- after cool the solution to 25°C, and add 500µl of 24:1 chloroform:isoamyl alcohol (CIA)

- after centrifugation (12,000 x g ; 5 min ; 25°C), transfer the upper aqueous layer, and extract with 1/10 volume of 10% CTAB in 0.7 M NaCl

- repeat previous step

- after centrifugation, the DNA in the supernatant precipitate with an equal volume of isopropanol (500 µl)

- the nucleic acid pellet collect by centrifugation (12,000 x g, 15 min, 25°C)

- the pellet wash with 70% and than with 95% ethanol, dry and resuspend in 200 µl of TE buffer (10 mM Tris-HCl, pH 8 ; 1 mM EDTA) containing 10 µg/ml RNAase

Plasmid DNA (for sequencing procedure) was isolated using the Plasmid DNA Mini Kit and spin columns (Qiagen, Hilden, Germany) as well as Novagen`s SpinPrep^{TM} 20 DNA purification kit.
The concentration of isolated DNA was measured using spectrophotometer GeneQuan II (Pharmacia, Biotech,) in UV at A_{260} nm. The spruce DNA was diluted in water to a working concentration 5 ng/µl. It was measured by comparison of different dilutions with

the fluorescence of lambda DNA concentration standards on ethidium bromide stained agarose gel. After extraction and dilution DNA samples were stored at –20°C.

2.3 Designing and analysing polymorphic EST-PCR marker

2.3.1 Primer construction

Identification of gene functions for all nucleic acid sequences and deduced amino acid sequences was performed online (http://www.ncbi.nlm.nih.gov/BLAST/) by means of the *BLASTN* and *BLASTX* programmes respectively (Altschul *et al.,* 1990). Using the *HIBIO DNASIS* programme (version 2.5, *Hitachi Software Engineering Europe*, Olivet Cedex, France), cDNA sequences were employed in order to design primer pairs optimal for the clone-specific amplification of spruce DNA.

All primers vary in length from 20 to 30 bases; contain not less than 40% of G and C bases. The annealing temperatures of both primers belonging to the same pair differ for not more than 7°C. The minimal potential for creation of the second structure by oligonucleotides was allowed for primer construction. As is reported by Perry and Bousquet (1998), polymorphisms, and length polymorphisms in particular, are most likely to occur in no coding regions. Thus, all reverse primers were placed in the 3´non-coding region (UTR). Forward primers were selected upstream, within coding region. All reverse primers were originally labelled at the 5´site by the Cy5 dye for fluorescent detection of PCR fragments. Oligonucleotides and labelling reactions were provided by MWG Biotech, Ebersberg, Germany.

2.3.2 PCR

All PCR mix reagents were provided by Invitrogene GmbH, Karlsruhe, Germany. PCR reactions were performed using U Platinum *Tag* DNA polymerase in a total volume of 25 µl (see *Tab 2-2*). Different profiles of annealing temperature as well as concentration of $MgCl_2$ were tested for different primer pairs in order to optimise specific PCR conditions. Standard conditions for PCR amplification (optimised for EST-PCR marker PA005) are summarized in *Tab. 2-2*. Thermocycler instrument Uno, Biometra, Göttingen, Germany (lid temperature 94°C) was used. PCR programmes consisted of an initial denaturation of DNA and activation of the *Taq* polymerase at 94°C (5 min). Three different PCR

temperature profiles were used for analysis of six EST-PCR markers (see *Tab. 2-6*). All PCR products were directly analysed by electrophoresis on standard agarose gels (Manniatis, Sambrook, 1990).

2.3.3 Electrophoresis in PAG and agarose gels

5 µl of PCR mix in 6x loading buffer (0, 25% Bromphenolblue, 50% Glycerin in 1xTBE) was put on 1% agarose gel contained Ethidiumbromid. Electrophoresis was running in 0.5xTBE buffer (20 min, 100 V). "1kb plus" DNA size standard (Invitrogene GmbH, Karlsruhe, Germany) was used for comparable size analysis of PCR fragments. Results were documented using Digital camera (EDAS 120, Eastman Kodak).

For clear segregation between fluorescently labelled PCR products, ALF-express system (Amersham Pharmacia, Biotech) was used. Laser fluorescent technology represents a highly sensitive method which allows detecting of picograms of DNA fragments by the means of laser. Electrophoresis is carried out in a vertical gel where the fluorescently labelled fragments, migrating downwards through the gel, are excited by the laser beam and detected by photo detectors. The signals are transmitted directly into a computer database and processed.

Note 2. Major steps of preparation 6% polyacrilamide gel stock solution for ALF-express system

500 ml

-210g Urea dissolve in 100 ml of 30% acrylamide (AA) and 30 ml 10x TBE

-add water to reach final volume of 500ml

-perform vacuum filtration to avoid contamination

-add 26.5 ml TEMED (N,N,N`,N`-Tetramethylethylenediamine) and 105 µl APS 10% (Ammonium persulfate) to 30 ml of polyacrilamide gel just before pouting the gel into the short gel cassette (20 cm long and 0,5 mm thick) in order to polymerise the monomers

10xTBE (Tris-borate) stock solution: (1L)

Tris	121.14 g
Boric acid	51.3g
EDTA (0.5M)	3.72g

Each DNA probe (the products of DNA digestion or PCR) was diluted with 3 µl 1xTE buffer and mixed with 3 ml formamide loading dye. To each sample, an internal standard

(PCR-DNA fragments with definite length) was added and a Cy5 Sizer 50-500 (with increment of 50 bp) was used as external standard. Samples were then denatured at 85° C for 5 min, quench on ice and loaded onto gel. For procedure of gel preparation see *Note 2*. Following denaturation the amplificates were detected by electrophoresis on 7 *M* urea/6% (w/v) acrylamide gels (small plate, 60-240 min (depending on the fragment length), 1500 V, 38 mA, 34 W, 50° C).

The size of PCR products was calculated by the ALLELE LINKS version 1.0 software (Amersham Pharmacia, Biotech) using both "internal" and "external" DNA size standards.

2.3.4 Digestion

In the cases of EST-PCR markers PA0005, PA0038 and PA0055 the amplification products were digested by one of the restriction enzymes *Rsa I*, *Hinf I* and *Dra I* respectively (see *Tab. 2-6* and chapter 3.1.2)

Digestion procedures were performed according to the instructions of the manufacturer, MBI Fermentas GmbH, St. Leon-Rot, Germany (see *Tab. 2-3*). Cy-5 labelled restriction fragments were detected by an ALF-express instrument (Amersham Bioscines, Freiburg) during electrophoresis in acrylamide gels (small plate, 1500 V, 38mA, 34 W, 50° C). For each probe, 5 µl of digestion mix together with 3 µl of loading dye containing internal standards were used. Additionally, the digested PCR products of EST marker PA0055 were analysed on 2% agarose gels in 0.5x TBE buffer using standard electrophoresis conditions.

Table 2-2. PCR mix and conditions for amplification of spruce DNA using cDNA clone-specific primers

25µl mix		Temperature profile:		
$MgCl_2$	1,5 (0,75-2) mM	Initial denaturation	94 °C, 5 min	
10x PCR buffer	2,5 µl			
dNTP, each	0,1 mM	Denaturation	94 °C, 1min	40 cycles
Primer, each	10 pmol	Annealing	60 °C (+/-10°C), 1 min	
Enzyme	0,5 U	Elongation	72 °C , 1min	
DNA	30-100 ng			
dH_2O		Final elongation	72 °C , 10 min	

Table 2-3. Scheme of digestion of PCR products by restriction enzymes

15 µl mix		Temperature	
10x buffer (10% BSA)	1,5 µl		
Enzyme	5 U	37 ° C	6 h, over night
PCR product	5-10 µl		
dH_2O			

2.4 Analysing inheritance of EST-PCR marker

The mode of polymorphism inheritance was tested by studying the segregation among megagametophytes that are evident in coniferous species as haploid tissue surrounding the developing embryo, and that represent the female gametic contribution to the diploid embryo. Based on seed samples of heterozygous mother trees, conformity of each observed segregation with the expected 1:1 proportion was tested statistically by means of a chi-square test of goodness of fit (N = 24). In addition, segregation among the offspring of a full sib family (N = 81) was compared statistically with the expected proportions according to the Mendelian mode of inheritance. Genetic types which followed these proportions were considered as alleles at single gene loci. DNA extracts from the full-sib individuals and both parents were kindly provided by I. Scotti (University of Udine, Italy).

2.5 DNA sequencing of polymorphic PCR fragments

2.5.1 Cloning

Polymorphic PCR products separated by electrophoresis were extracted from 2% agarose gel using Gel extraction kit (Qiagen, Hilden, Germany) and diluted each in 5 µl of water. Approximately 30 ng of each samples was utilized for cloning in pETBlue™ Vector using Novagen`s Clonables ^{TM} Ligation/Transformation Kit (Novagen, VWR Deutschland GMBH, Darmstadt). This cloning system (*Note 3*) combines the visual identification of recombinants and high plasmid copy number of blue/white screening vectors. Insertion of target sequences into the multiple cloning site (MCS) disrupts expression of the lacZ a-peptide and produces a white colony phenotype in strain NovaBlue when plated in the presence of X-gal and IPTG. Colonies derived from the unmodified vector turn blue.

Note 3. Components pETBlue System

pETBlue DNA (uncut)

pETBlue UP primer

pETBlue Down primer

NovaBlue Competent cells

SOC Medium

Test Plasmid for Transformation

Note 4. Transformation protocol for Novagen`s Clonables™ Ligation/Transformation Kit

-Thaw the required number of tubes of cells on ice and mix gently to ensure that the cells are evenly suspended

-place the required number of 1.5 ml polypropylene microcentrifuge tubes on ice to pre-chill. Pipette 20 µl aliquots of cells into the pre-chilled tubes

-add 1µl of the DNA solution directly to the cells. Stir gently to mix

-place the tubes on ice for 5 min

-heat the tubes for exactly 30 sec in a 42°C water bath; do not shake

-place on ice for 2 min

-add 80 µl of room temperature SOC medium to each tube

-plate 5-50 µl cells directly on LB agar media containing 50 ng/ml ampicillin, 15 µg /ml tetracycline, 70 µg /ml X-gal, and 80 µM IPTG

-incubate inverted plates overnight and screen colonies

All components except competent cells were stored at –20°C. Competent cells were stored at –70°C. For procedure of transformation see *Note 4*.

Ligation procedures were performed according to the instructions, in a total volume of 10 µl, contained 50 ng of linearized, dephosphorylated vector, an equal molar to 5-fold excess of target insert, ligation buffer, ATP, DTT and T4 DNA ligase. Ligation was enhanced by a final concentration of 1mM ATP. 1µl of ligation reaction per 20 µl of competent cells was used for transformation procedure.

White colonies were picked up from the plates containing ampicillin. Prior to growing colonies for plasmid isolation, the presence of the appropriate insert and its orientation were checked by direct colony PCR. To verify the presence of an insert, 5 pmol (1µl) of each from vector-specific and insert-specific primers, 1, 25 U AmpliTag DNA polymerase 10 mM dNTP, 5 µl of 10x PCR buffer, in total volume 50 µl were used.

2.5.2 Sequencing

Recombinants were sequenced on both strands by an oligonucleotide walking strategy, employing the Cy5-AutoRead Sequencing Kit and the Cy5-dATP Labeling Mix (Amersham Pharmacia Biotech, Freiburg) (see Appendix E) long ranger gels (large plate, 800 min, 800 V, 60 mA, 25 W, 55°C, 0.5xTBE), an ALFexpress automated laser sequencer (Amersham Pharmacia Biotech, Freiburg). All sequences were analysed and compared using HIBIO DNASIS programme (version 2.5, *Hitachi Software Engineering Europe*, Olivet Cedex, France).

2.6 Quantifying of genetic variability of Norway spruce

2.6.1 EST markers

Five codominant nuclear EST markers were used to evaluate genetic variation along natural range of Norway spruce. EST markers *PA0034, PA0038, PA0043, PA0055, PA0066* (Schubert *et al.*, 2001; Riegel, 2001) were utilised in addition to the newly developed codominant marker *PA0005*. Three of them, *PA0034*, *PA0043* and *PA0066*, are described as markers detecting PCR fragments length polymorphism directly after PCR reaction. Two others, *PA0038* and *PA0055,* based on the point mutation polymorphism require digestion of the PCR products using specific restriction enzymes (Schubert *et al.*, 2001; Riegel, 2001).

Classifications of PCR primer pairs and observed alleles of five additional markers are adduced in *Tab. 2-4* and *Tab. 2-5* respectively. Three different PCR programs used for amplification of spruce genomic DNA are summarized in *Tab. 2-6***.** In the case of markers *PA0034*, *PA0043* and *PA0066* fluorescently labelled products of amplification were detected using ALF-express system immediately after PCR reaction. Digestion of the PCR fragments using specific restriction enzymes was carrying out for EST markers *PA0038* and *PA0055* according to the instructions of the manufacturer, MBI, Fermentas GmbH (see *Tab.2-5*). The fluorescently labelled products of digestion were detected by electrophoresis in acrylamide gels using ALF-express system (Amersham Pharmacia Biotech).

Table 2-4. PCR primer pairs for amplification of five EST markers and assessed gene function for the corresponding cDNA clones

Locus	**Primers** *A*: **Forward**; *B*: **Reverse**	**Gene function predicted for the corresponding cDNA clone**
PA0034	**A**: 5′ AGG TCT GCT AAT GGT TCT 3′ **B**: 5′ Cy5-ATA TCA AAC AAA CTG TGT TAG CTC 3′	no homology found, contains repeat $(CT)_6TT(CT)_5$
PA0038	**A**: 5′ AAC GGC ATT TGG ACT CTA TCT C 3′ **B** : 5′Cy5-TTA GAT GAT ATG CAA TGT AGA TTG A 3′	72% identity with halotoleranceprotein HAL3 of *A. thaliana*
PA0043	**A**: 5′ AAA ACT GGA GGA CCT TCT GG 3′ **B**: 5′ Cy5-GTG AAC CTC TAC AGA AAC ACA A 3′	86% identity with 78 kDa glucose regulated protein of *N. tobacco* (HSP 70-heatshock family)
PA0055	**A**: 5′ TTG GGT TTT AGG TCA TGA CTG C 3′ **B**: 5′ Cy5-TCC GAT TTA TTA TAT CAA AAC TGC CTC 3′	88% identity with ß-subunit of ATP-syntasecomplex of *H. brasiliensis*
PA0066	**A**: 5′ CAA GCG GTT GGT TGG AGT TCG GTT 3′ **B**: 5′ Cy5- GCA GCA AAA ACT GCA CCT CTC TTC TG 3′	78% identity with 60 S ribosomal protein L13-2 (cold-induced protein 24B) of *B. napus*

Table 2-5. Nomenclature of alleles, used amplification programs and digestion enzymes for five EST markers

EST Marker	**PCR program**	**Restriction enzyme**	**Type of gel**	**Classification and size of alleles**
PA0034	RR2	-	ALF-express	**A***:* 212***;*** **B**: 214***;*** **C**: 226; **D**:228; **E**:230
PA0038	JAN1	*HinfI*	ALF-express	**A**: 186; **B**: 192; **C**: 228
PA0043	RR3	-	ALF-express	**A**: 368; **B**: 372; **C**: 389; **D**: 420
PA0055	*RR3*	*DraI*	ALF-express	**A***:* 267+39 *(cut);* **B**: 306
PA0066	RR2	-	ALF-express	**A**: 154; **B**: 161; **C**: 171; **D**: 178

Tab 2-6. Amplification conditions for 3 PCR programs used for proceeding of EST markers

Program of	Denaturation	Primer annealing	Extension	Number cycles
JAN 1	94°C / 5 min	60°C / 5 min	72°C / 1 min	1
	94°C / 1 min	60°C / 1 min	72°C / 1 min	42
			72°C / 9 min	1
RR 2	94°C / 4 min			1
	94°C / 1 min	50°C / 1 min	72°C / 1 min	34
			72°C / 5 min	1
RR 3	94°C / 4 min			1
	94°C / 1 min	55° C / 1 min	72°C / 1,5 min	30
			72°C / 10 min	1

2.6.2 Estimation of genetic variation

All calculations of genetic parameters for single loci and gene pool (all loci) were realized by means of the *GSED* version 2.0 (E. Gillet, University of Goettingen, Germany) and *POPGENE* version 1.32 (F. C. Yeh, R. Yang, University of Alberta and T. Boyle, Center for International Forestry Research, USA) computer software.
Genetic variation within populations was quantified on the basis of allele and genotype frequencies observed for single gene loci.

Actual (observed) heterozygosity (H_a)

The proportion of heterozygotes among all studied individuals within a population is determined as $\mathbf{H_a}$ (average degree of heterozygosity for a single locus). For a set of loci, the arithmetic mean of the single locus heterozygosities coincides with the average individual heterozygosity.

Conditional heterozygosity (H_c, Gregorius *et al.* 1986)

The conditional heterozygosity at a single locus takes into account that the proportion of heterozygosity is conditional on the allele frequencies. In order to avoid dependencies on the underlying gene frequencies, the actual heterozygosity H_a is normalized by the maximum proportion of heterozygosity H_{max} obtainable for the underlying allele

frequencies, where H_{max} equals 1 if all allele frequencies are less than or equal to 0,5 and $H_{max}=2(1-p)$ if the most frequent allele has frequency p greater than 0,5.

$$H_C = \frac{H_a}{2(1-p)} \qquad \text{with } p \text{ as allele frequency } (p<0,5)$$

The multilocus mean of H_C is equal to the ratio of the summed H_a-values to summed maximum attainable values with

Diversity (v)

Gene diversity has been defined by Nei as the probability that two alleles taken at random within a population are different:

$$\mathbf{v}_N = 1 - \sum_i p_i^2$$

where p_i is the frequency of allele i. For a random mating population, gene diversity is equal to heterozygosity.

Genetic diversity defined by Gregorius 1978 quantifies the number of different types (alleles or genotypes) according to their frequencies in the population.

$$\mathbf{v}_{Gr} = \sum_i p_i^2)^{-1}$$

Genetic differentiation

The difference in the level of polymorphism between the different levels is defined as differentiation.

Genetic differentiation as stated by Nei (1987) is the extent of gene differences among populations that is measured by numerical quantity. It is formally identical to means of

Wright`s **F_{ST}** (1978). F_{ST} range from zero to one; if F_{ST}=1 the populations are fully differentiated, i.e. no common genetic types.

Genetic distance

Genetic distance between pair wise of populations was estimated using GSED program. Genetic distance, as defined by Gregorius (1974), is specified as the proportion of genetic elements (alleles, genotypes) which the two collections do not share.

$$d_0(p,p') = \frac{1}{2} \sum_{k=1}^{n} (p_k - p'_k)$$

On the basis of calculated genetic distance, UPGMA dendrogram was constructed using *PHYLIP* and *TREE VIEW*, version 6.0 (Page, R. D. M. 1996) programs.

Subpopulation differentiation Dj and δ (Gregorius, Roberds, 1986)

The amount of genetic differentiation of one subpopulation to the remainder of the population is specified as "the proportion of genetic elements (alleles, genotypes) by which a deme differs from the remainder of the population in type". The proportion is defined as

$$D_j = d_0(p_j, p_j^r)$$

where p_j and p_j^r are the frequency distributions of the types in deme j and in the remainder of the population, respectively, and d_0 is the genetic distance defined above.
The subpopulation differentiation is then defined by

$\delta = \sum_j c_j D_j$, where the weights c_j express the proportion of genetic elements present in the jth deme.

2.7 Statistical analysis

2.7.1 Testing for Hardy-Weinberg proportions

The data of single locus genotype frequencies in a population were standard cross-checked with Hardy-Weinberg expectations using Chi-square test and likelihood-ratio test statistic.

. Pearson´s **χ^2 goodness-of-fit test** statistic:

$$X^2 = \sum_{types} \frac{(N. - E(N.))^2}{E(N.)}$$

2.7.2 Isolation by distance test

Relationship between geographical remoteness among analysed populations and their genetic differentiation along natural range was statistically examined. For this purpose, Isolation-by-distance model (Slatkin and Maddison, 1990; Slatkin, 1993) based on Mantel's test statistics (Mantel, 1967) was applied.

"Isolation by distance" analysis involves plotting the genetic similarity (or distance) among population pairs as a function of the geographic distance between those pairs (Bohonak, 2002). Isolation by distance plots assess whether more distant population pairs are more different genetically.

Significance in the isolation by distance relationship can be tested statistically using a Mantel test. This test measures the association between the elements in two matrices by a suitable statistic, and determines the significance of this by comparison with the distribution of the statistic found by randomly reallocating the order of the elements in one of the matrices. A null distribution is generated by randomising rows and columns of one matrix while holding the other constant. Mantel test assess whether the pair-wise genetic distance matrix is correlated with the pair-wise geographic distance matrix.

The matrix of the Euclidean geographical distance was created online (http://www.indo.com/distance) using the coordinates of analysed populations.

The matrix of genetic distance (Gregorius) and the matrix of the Euclidean geographical distance were pair-wise proven by Mantel's test statistics using 1000 permutations and

employing the *ISOLDE* menu offered within the *GENEPOP* program (Version 3.1d, R. Raymond and F. Rousset, Laboratoire de Genetique et Environment, Montpellier, France) at the world web site (http://wbiomed.curtin.edu.au/genepop/genepoppop6.html). A null hypothesis of independence of both matrices was tested for possible rejection under the assumption that it is true. Significance of Mantel test was considered by next way: $P(\delta) \geq 5\%$ as "not significant", $1\% < P(\delta) < 5\%$, as "significant" (i.e., is not simply due to chance) (denoted *), and $P(\delta) < 1\%$ as "highly significant" (denoted **).

The populations were divided according to their sampling sites into the Alpine group (consisting of populations No. 2, 3, 4, 5, 6, 7,18 and 19), the Hercyno-Carpathian group (consisting of populations No. 8, 9, 10, 11) and the Northern European-Russian group (consisting of populations No. 12, 13, 14, 15, 16, and 17) (see *Tab. 2-1*). This grouping is considered to be related to the three different gene pools of Norway spruce that currently exist: Alpine domain, Hercyno-Carpathian domain, and Baltic-Nordic domain, respectively.

In order to test *Isolation by distance* between populations sampled along European range, the matrices of genetic and geographic distances were created for all populations together, for populations belonging to every group separately and for combined groups of populations such as "Alpine + Hercyno-Carpathian groups", "Alpine + Northern European-Russian groups", "Hercyno-Carpathian + Northern European-Russian groups".

2.7.3 Neutrality test

The **Ewens-Watterson Homozygosity test** is based on Ewens (1972) sampling theory of neutral alleles. It derives the expected number of alleles in a sample under the infinite-alleles model (infinite-allels mutation model: an allele can mutate into an non-existing allele and can never back-mutate). Ewens suggested that it is possible to calculate the probability of observing any configuration of allele frequencies and one way of summarising the frequency spectrum is to use the information

$$B = -\sum_{types} x_i \ln x_i \text{ , where } x_i \text{ is the frequency of allele } I$$

Equal frequency of alleles results in a large statistical value. In contrast, when there is a high frequency of a single mutation, and all others are low frequencies the statistical value

is small. The information in any observed data set can be compared to the distribution of information expected under the WF model. If the observed information is higher than the 97.5 percentile, or lower than the 2.5 percentile, the WF model can be rejected at the 5% level.

Watterson (1977) has shown that if the alternative hypothesis is that alleles are maintained by heterozygote advantage (with an equal fitness for all heterozygotes). The effect of such selection is to make allele frequencies more even than expected under neutrality. The likelihood tio test for balancing selection can be reduced to a function of the population homozygosity (the probability that two alleles picked at random from a population is identical).

$$H = \sum_{types} x_i^2$$

Homozygosity than is considered as a sufficient statistic for testing the hypothesis of symmetric heterozygote advantage (which decreases homozygosity) (McVean, 2002).

The Ewens-Watterson test for neutrality (Manly, 1985) was performed for all analysed EST markers, with respect to study of geographical differentiation, using *POPGENE* program, version 1.32.

The distribution of selectively neutral type (alleles) frequencies could be conveniently summarized by the sum of type (allele) frequencies (*F*), equivalent to the expected homozygosity for diploids. The null distribution of *F* is generated by simulating random neutral samples having the same number of genes and the same number of types using the algorithm of Stewart (1977). The probability of observing random samples with *F* values identical or smaller than the original sample is recorded. The "Obs. *F*" should be compared to "L*95" and "U*95", which are respectively, the lower and upper 95% confidence interval. If "Obs. *F*" is within this confidence interval, the locus is neutral; otherwise, it is not. Small "Obs. *F*" values indicate that allele frequencies are too even, suggesting that there is a tendency in favor of heterozygotes in the population. Large "Obs. *F*" values indicate selection against the heterozygotes in the population.

3 Results

3.1 Development of an EST-PCR marker for cyclophylin in Norway spruce

3.1.1 cDNA clone encoding cyclophylin, PCR amplification

A full-length cDNA clone PA0005 was selected from Norway spruce cDNA library (Bozhko *et al*., 2003; Schubert *et al*., 2001). For construction of this cDNA library, poly (A) +RNA was isolated from photomixotrophic suspension cells of *Picea abies* after treatment with a fungal elicitor (Galliano *et al*., 1993).

Sequencing of the recombinant plasmid pPA0005 revealed that the full-length cDNA clone, recorded under the EMBL accession number AJ271126, is 1026bp in length and contains a short 5´non –coding region of 101 bp as well as a longer 3´non-coding region of 406 bp including the poly(A) tail (Bozhko *et al*., 2003).

Its longest open reading frame encodes a polypeptide of 172 amino acid residues, and this query as well as nucleotide sequence were run on public molecular databases to search for similarities with already identified proteins of other organisms. Significant matching with the highest values reaching up to 84%-85% sequence identity were found amongst this protein and several cyclophilins that have been previously identified from a number of angiosperms including *Digitalis lanata*, *Ricinus communis*, *Phaseolus vulgaris, and Catharanthus roseus.*

Cyclophilins (CyPs) are described as ubiquitous proteins with an intrinsic enzymatic activity of peptidyl–prolyl cis-trans isomerase that catalyses the rotation of X-Pro peptide bonds and accelerates the folding of certain proteins (Godoy *et al*., 2000). CyPs play a role in signal transduction and in the plant response to environmental stress (Hare *et al*., 1999; Meza-Zepeda *et al*., 1998; Scholze *et al*., 1999).

As can be inspected from the EMBL library, 14 highly conserved amino acid residues (R, F, M, Q, G, A, N, A, Q, F, W, L, K, and H) were found in positions 62, 67, 68, 70, 79, 108, 109, 110, 118, 120, 128, 129, 132, and 133. These amino acid residues are known to be crucial for an intrinsic biochemical feature of all currently known cyclophilins: the cyclosporin A binding activity (Bozhko *et al*., 2003).

Based on significant homologies with known genes for cyclophilins of other organisms, the Norway spruce cDNA clone PA0005 was identified as putative stress-related gene. To

amplify expressed sequence tag (EST) site and to obtain a molecular marker for the newly identified nucleotide sequence, primer pair P1/P2 derived from *Picea abies* cDNA clone pPA0005 was used for PCR amplification of spruce genomic DNA. There is known that useful sequence variation for polymorphic DNA markers can often be found at non-translated parts of mRNA sequences (Gil *et al.*, 1997). Thus, the forward primer P1 5´GAACTTAAGATGTCGAACCCAA3` and reverse fluorescent primer P2 5`Cy5-CCCTCGAAACCTCTATAGTTGCCA3` were both constructed for annealing to the 5`non-translated and 3´non-translated cyclophilin gene regions, respectively. The amplification fragment with size 772 bp was expected.

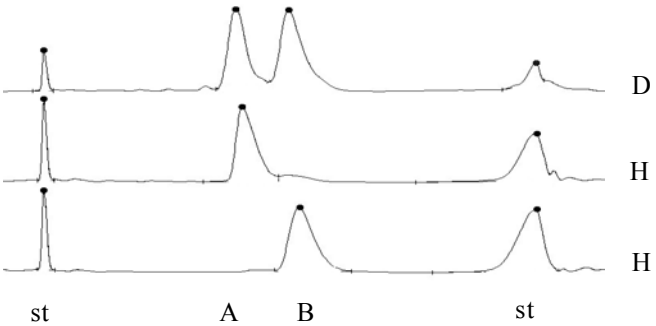
Utilizing 14 individual bud samples from test population number 1, the genomic PCR bands, generated under the direction of primer pair P1/P2 and analysed in agarose gel, were found to be polymorphic in size. To optimise the analysing the Cy5-labelled fragments by Polyacryl Amide Gel Electrophoresis (PAGE), the digestion of obtained PCR products by the restriction enzyme *Rsa*I was performed. Following digestion and PAGE, two polymorphic products, one with expected size (526 bp) and other 21 bp shorter (505 bp) were detected.

3.1.2 Co-dominant inheritance of an EST marker

In order to determine the mode of polymorphism inheritance, three maternal trees, each carrying two different *Rsa*I generated cyclophilin gene marker bands at the diploid DNA panel (bud sample), were selected and compared with the PCR products obtained from corresponding haploid megagametophytes (shown for one maternal tree in *Fig. 3-1*).

Figure 3-1. Example for co-dominant segregation pattern of the fluorescent PCR marker PA0005 as indicated by the bud sample of a heterozygous spruce tree [D] and corresponding megagametophytes samples [H] (alleles **A**:505 bp; **B**:526 bp; st=internal DNA size standards used for calibrating the Alfexpress instrument)

.

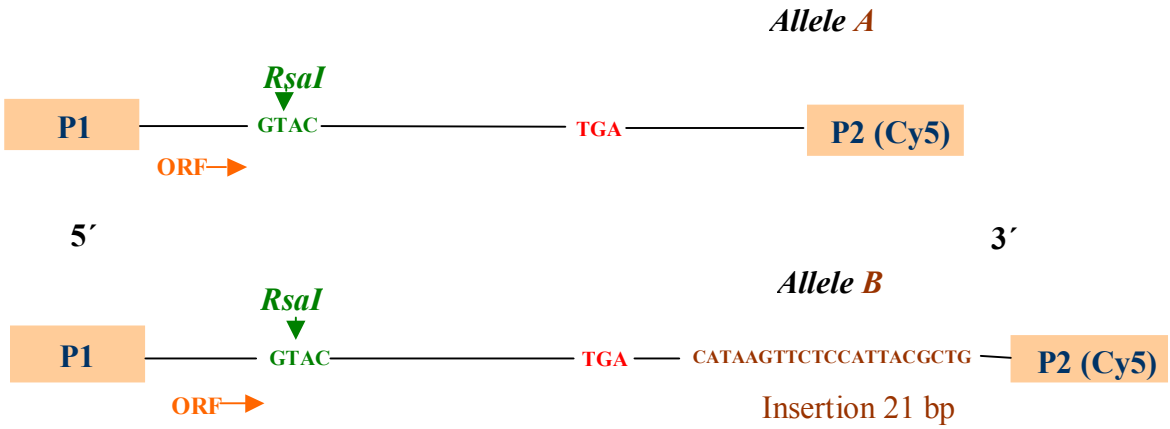


The marker bands exhibited regular gametic segregation because each observed segregation statistically conformed to the expected 1:1 segregation (chi-square values of 0, 0.2, and 0.4, respectively). Moreover, by analysing the F_1 progeny of one monomorphic and one polymorphic parent, the codominant mode of inheritance of the molecular cyclophilin gene marker was confirmed, since the observed segregation of 38:44 did not deviate statistically from the expected Mendelian 1:1 segregation (chi-square value of 0.44). Altogether, these tests clearly demonstrate that the EST marker, named PA0005 with respect to the cDNA clone, detects allelic polymorphism at a single Mendelian locus, proving that the observed RsaI patterns are not a result of artificial PCR or partial enzymatic digestion. The analysis of 110 trees (population No.1) implied a tendency towards minor polymorphism for the genetic variability detected by newly EST marker PA0005.

3.1.3 Sequence data analysis

To verify the observed patterns of digestion, the bud samples revealing PCR/*Rsa*I fragment A or B were applied for PCR with primer P1 and nonfluorescent primer P2. The amplification products were digested by the enzyme *RsaI* and separated by Polyacryl Amide Gel Electrophoresis. The fragments with sizes corresponding to alleles A (505 bp) and B (526 bp) were extracted from Polyacryl Amide Gel, purified, and used for the cloning procedure in pETBlue™ Vector. The presence of analysed sequences inside recombinant plasmids was detected by PCR under direction of P1/P2 primers. DNA sequence analysis of the recombinants containing both products amplified, confirms that *P.abies* cyclophilin genes are devoid of introns. Two genomic fragments corresponding to the alleles A and B had been shown to differ by the short inserted DNA motif 5´CATAAGTTCTCCATTACGCTG3´, lying in the 3´noncoding region, 155 bp downstream from the stop codon (*Fig. 3-2*).

Figure 3-2. Differences between two alleles of the cyclophilin gene in *Picea abies*. Two fragments amplified under direction of forward primer P1 and reverse fluorescent primer P2 differ in length by 21 bp. The insertion located in 3`non-translated region (ORF: opening reading frame) lies 155 bp downstream from the stop codon (TGA) and 440 bp downstream from the recognition site for the restriction enzyme *RsaI* (GTAC).



3.2. Monitoring of genetic variation of Norway spruce in Europe

The newly developed codominant EST marker PA0005 as well as five additional already existing nuclear EST markers PA0034, PA0038, PA0043, PA0055, and PA0066 was applied to verify genetic variation within and among 19 naturally regenerated populations of Norway spruce.

3.2.1. Variation within populations

3.2.1.1. Allele frequencies at single loci

Tab. 3-1 (a-f) and *Fig. 3-3(a-f)* survey allele frequencies at 6 polymorphic EST-markers for each of 19 analysed populations. Four markers were found to be polymorphic in all populations. The population from Siberia (No.15) was represented by only one allele in the case of the markers PA0005 and PA0055.

For the newly developed EST marker PA0005, a total of 7 alleles were identified among 19 spruce populations from a broad European distribution (*Tab. 3-1a*). The polymorphism at the locus PA0005 described in this study seems to be highly informative, since the uneven frequency distribution associated with allele B shows a clear geographic pattern. The marker band of 505 bp (allele A) represents the most frequent allele for all populations tested. Five rare alleles (C, D, E, F, and G) show low frequencies (≤ 3%) in some populations, or are completely absent in others. Allele B (526 bp) is, however, informative as indicated in *Fig. 3-3a* and *Tab. 3-1a*, since it shows a remarkable uneven frequency distribution in space. It ranges from 6% to 21% in the examined Croatian, Slovenian, Hungarian, Ukrainian, Southern German, Southern Polish and Italian, French and Swiss populations. Its maximum values were found in the Alpine region. In the populations from Russia, Norway, Northern Poland and Romania, however, allele B is missing or revealing a low frequency distribution (≤ 4%).

A comparable picture of uneven distribution of allele frequencies in space was detected by means of EST marker PA0055 (*Fig. 3-3b, Tab. 3-1b*). Considering allele frequencies at locus PA0055 within 19 analysed European spruce populations, definite geographical distribution of frequencies was found. Thus, smallest values of the allele B (303 bp) at the locus PA0055 were detected within populations from Russian-Northern–European group (Northern Polish populations No.12 and 13, Russian populations No.14 and 15, Norwegian

populations No.16 and 17), which varies from 0 to 4%. The populations from Alpine-Southern-European region (Italian populations No. 2 and 3, Swiss populations No. 4 and 5, Croatian and Slovenian populations No. 6 and 7, and Italian population No.18) constitute the opposite group with highest values of allele B within these populations varying from 26 to 40 %. The observed mean of indicative allele (B) at locus PA0055 within Romanian population No. 8 was comparable with Russian-Northern-European group as in the case of marker PA0005. But the allele A (267+39 bp, as a result of the digestion with *DraI*) is classified as predominant for all examined stands.

Outstanding position of Russian-Northern-European group in comparison with others analysed populations was evident in the case of distribution of allele frequencies at locus PA0043. The presence of an additional allele (M) with size 366 bp was found only in the populations from Russia and Norway (number 14, 15, and 16) with small frequencies between 2 and 7 % (*Tab. 3-1c, Fig. 3-3c*). The populations from Russian-Northern-European group (No. 14, 15, 16, 17) together with Romanian population (No. 11) were characterized at locus PA0043 by highest values of allele B (372 bp) (29-46%).

The same effect was observed for Russian-Northern-European domain at locus PA0066. The populations No. 12, 13, 14, 15, 16, 17 were represented by smallest frequencies of allele D (178 bp). It ranges between 4 and 12% in the group mentioned above and reaches 40% in others (*Tab. 3-1d, Fig. 3-3d*).

In the case of EST marker PA0034, the frequency of the allele C (226 b.p.) dominates in all populations. The variation in frequencies of all alleles does not show obvious geographical trends (*Tab. 3-1e, Fig. 3-3e*).

The allele A (186 bp) of marker PA0038 is rare in all populations excluding the Russian sample 15 where it is completely absent. The frequency distribution of the two other alleles of the marker PA0038 appears homogenous along the studied range of spruce distribution. (*Tab. 3-1f, Fig. 3-3f*).

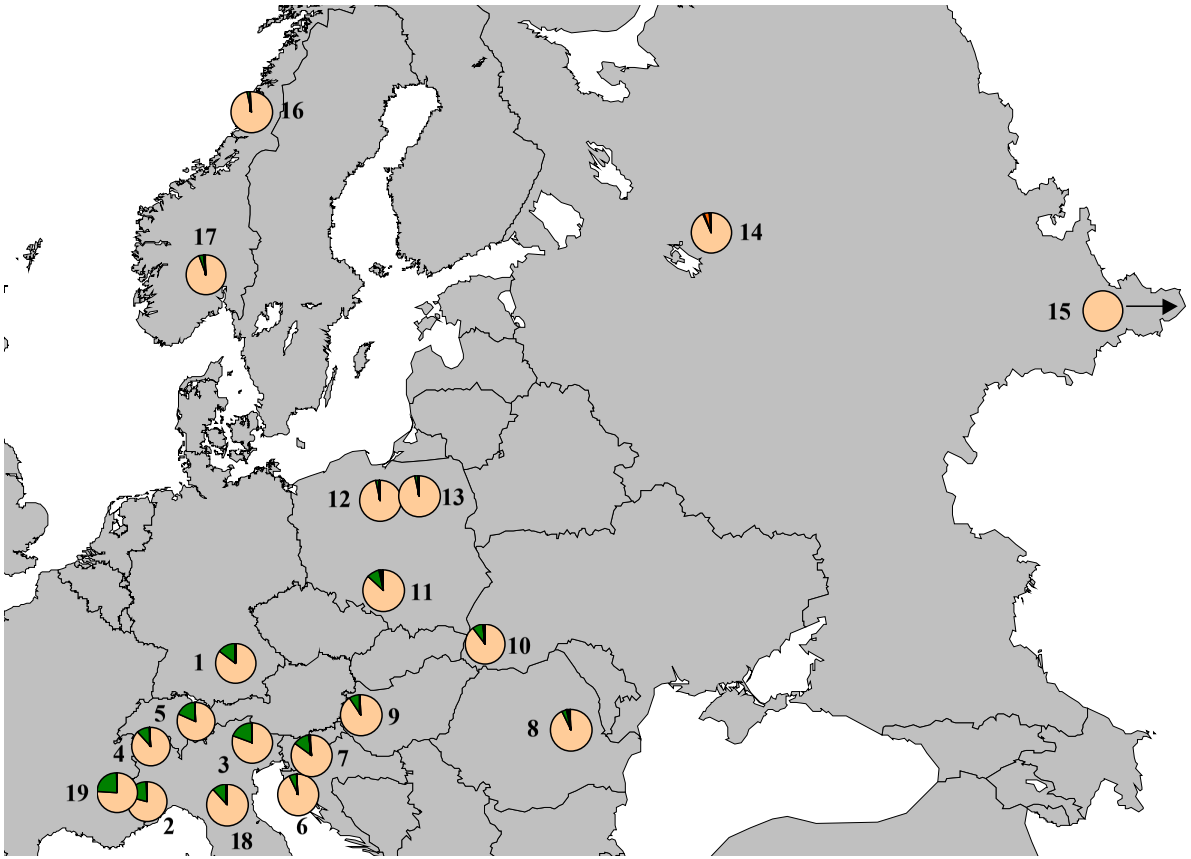


Figure 3-3a. Allele frequency distribution within European populations of Norway spruce at locus **PA0005.** The numbers and coordinates of populations as defined in Tab. 2-1, respectively; the arrow indicates that population no.15 is situated outside the map.

A
B
C
D
E
F
G

Table 3-1a. Frequency distribution for 7 alleles (A-G) of marker **PA0005**, measured in 19 European populations of *Picea abies* (populations are numbered according to *Tab. 2;* geographic positions are indicated in Fig. 3-3a; bp = size of the PCR fragment in base pairs).

Population		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
Allele	A (504bp)	90	80	79	80	88	94	85	89	90	86	90	97	97	94	100	95	98	88	75
	B (526bp)	9	20	21	20	12	6	14	9	8	11	9	3	3	2	0	4	2	12	25
	C (544bp)	0	0	0	0	0	0	0	1	1	2	1	0	0	0	0	0	0	0	0
	D (465bp)	1	0	0	0	0	0	1	1	1	1	0	0	0	0	0	0	0	0	0
	E (479bp)	0	0	0	0	0	0	0	0	0	0	0	0	0	3	0	1	0	0	0
	F (490bp)	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1	0	0
	G (375bp)	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0

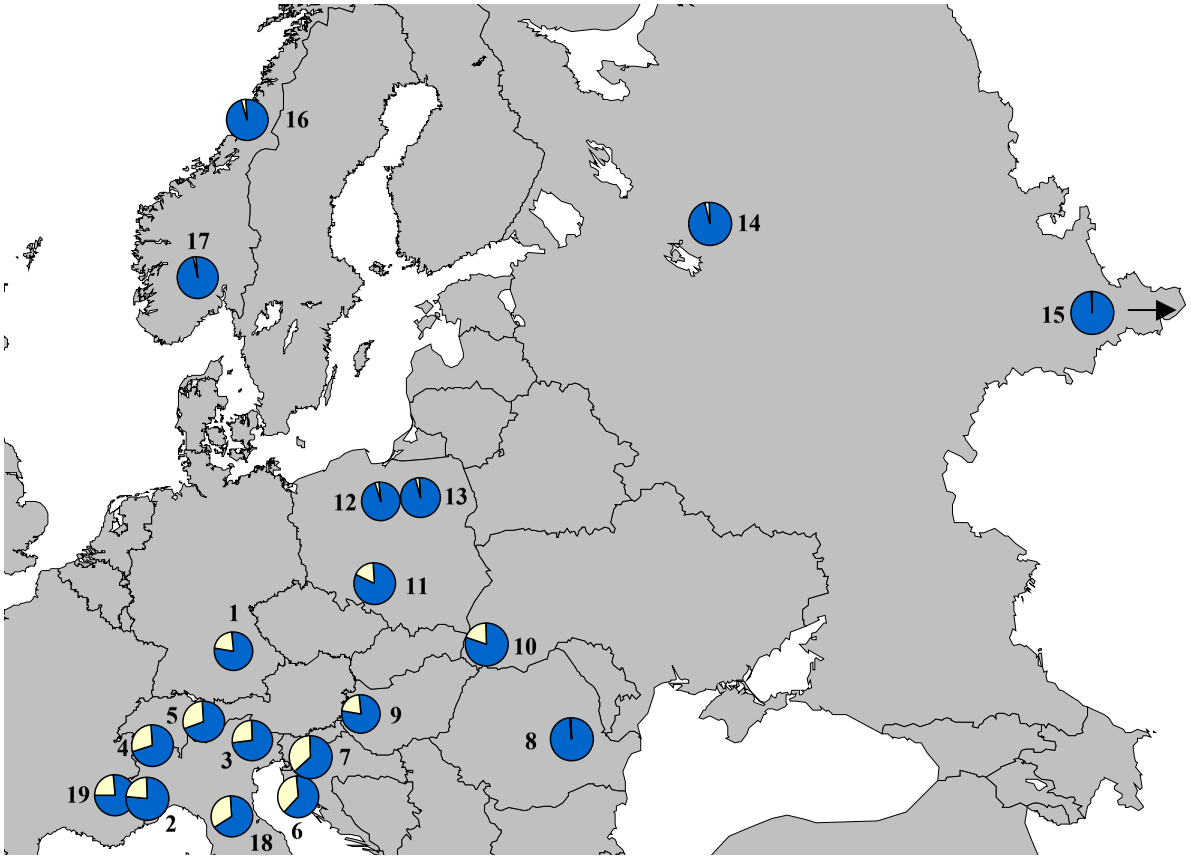


Figure 3-3b. Allele frequency distribution within European populations of Norway spruce at locus **PA0055.** The numbers and coordinates of populations as defined in Tab. 2, respectively; the arrow indicates that population no. 15 is situated outside the map.

Table 3-1b. Frequency distribution for 2 alleles (A,B) of marker **PA0055**, measured in 19 European populations of *Picea abies* (populations are numbered according to Tab. 2; geographic positions are indicated in Fig. 3-3b; bp = size of the PCR fragment in base pairs)

Population		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
Allele	A (267+39bp)	76	74	72	69	67	60	63	99	77	80	81	96	96	97	100	96	97	64	76
	B (306bp)	24	26	28	31	32	40	38	1	23	20	19	4	4	3	0	4	3	36	24

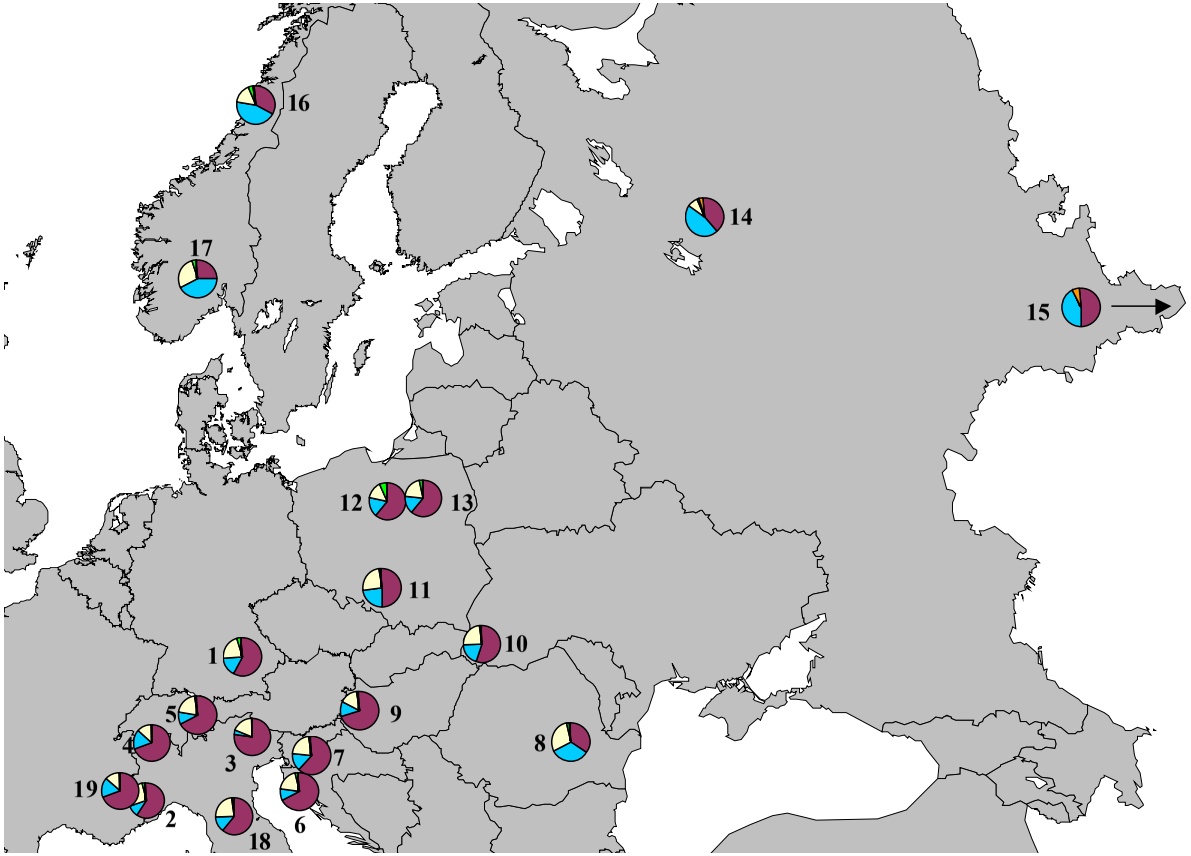


Figure 3-3c. Allele frequency distribution within European populations of Norway spruce at locus **PA0043.** The numbers and coordinates of populations as defined in Tab. 2, respectively; the arrow indicates that population no.15 is situated outside the map.

A
B
C
D
M

Table 3-1c. Frequency distribution for 5 alleles (A-D, M) of marker **PA0043**, measured in 19 European populations of *Picea abies* (populations are numbered according to Tab.2; geographic positions are indicated in Fig. 3-3c; bp = size of the PCR fragment in base pairs)

Population		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
Allele	A (370bp)	56	57	76	68	65	65	60	35	68	54	49	60	61	39	49	33	26	60	68
	B (374bp)	16	11	5	18	11	12	15	29	14	19	22	18	17	46	44	43	39	13	18
	C (391bp)	24	29	19	14	23	20	23	32	16	25	28	15	21	10	0	19	31	26	14
	D (422bp)	4	3	0	0	1	3	2	3	2	2	1	7	3	1	0	3	4	1	0
	M (366bp)	0	0	0	0	0	0	0	0	0	0	0	0	0	4	7	2	0	0	0

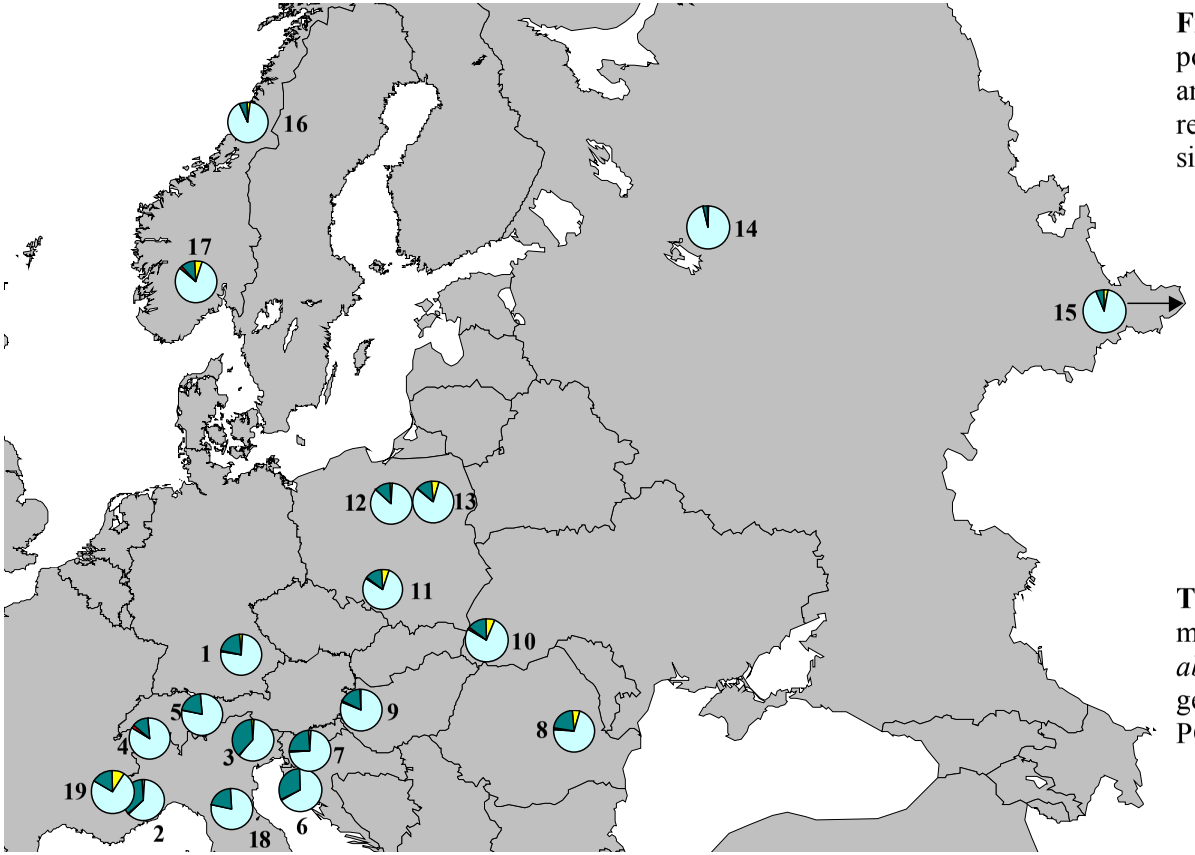


Figure 3-3d. Allele frequency distribution within European populations of Norway spruce at locus **PA0066.** The numbers and coordinates of populations as defined in Tab. 2, respectively; the arrow indicates that population no.15 is situated outside the map.

A
B
C
D

Table 3-1d. Frequency distribution for 4 alleles (A-D) of marker **PA0066**, measured in 19 European populations of *Picea abies* (populations are numbered according to Tab.2; geographic positions are indicated in Fig. 3-3d; bp = size of the PCR fragment in base pairs)

Population		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
Allele	A (154bp)	3	3	3	0	0	0	2	7	1	7	7	2	6	1	3	2	7	0	10
	B (161bp)	73	58	57	83	78	67	70	68	79	77	77	85	80	95	91	92	79	77	74
	C (171bp)	1	1	0	3	0	0	2	1	2	1	1	0	0	0	0	0	2	0	0
	D (178bp)	23	38	40	14	22	33	26	24	18	15	15	13	14	4	6	6	12	23	16

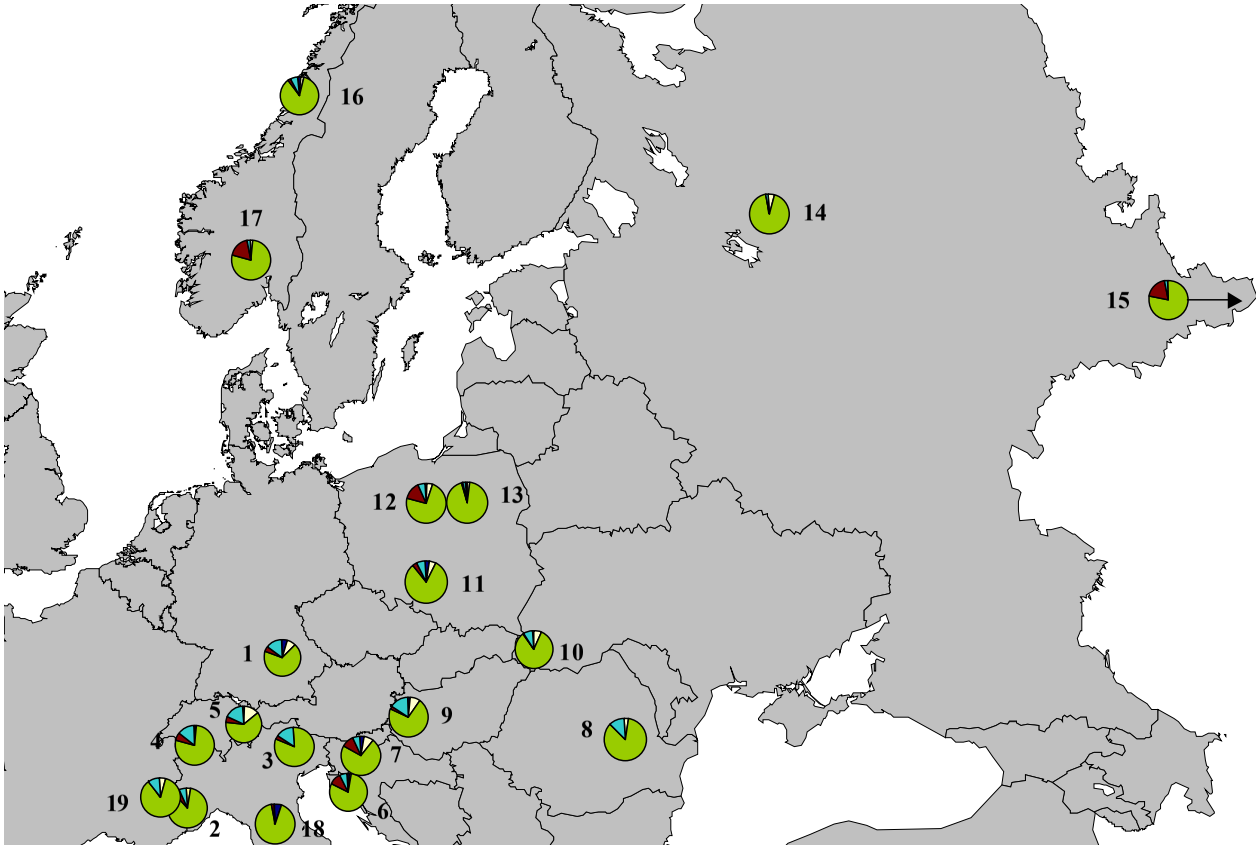


Figure 3-3e. Allele frequency distribution within European populations of Norway spruce at locus **PA0034.** The numbers and coordinates of populations as defined in Tab. 2, respectively; the arrow indicates that population no.15 is situated outside the map.

Table 3-1e. Frequency distribution for 5 alleles (A-E) of marker **PA0034**, measured in 19 European populations of *Picea abies* (populations are numbered according to Tab.2; geographic positions are indicated in Fig. 3-3e; bp = size of the PCR fragment in base pairs)

Population		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
Allele	A (212bp)	6	0	0	1	1	3	4	0	3	1	3	1	1	0	0	3	0	6	0
	B (214bp)	7	4	0	1	14	1	8	4	8	7	6	6	3	5	0	3	3	0	5
	C (226bp)	67	84	82	77	62	78	71	84	70	82	80	72	92	92	78	84	76	92	85
	D (228bp)	5	5	2	6	5	11	11	0	3	1	4	14	1	1	20	3	18	2	0
	E (230bp)	15	7	16	15	18	7	6	12	16	9	7	7	3	2	2	7	3	0	10

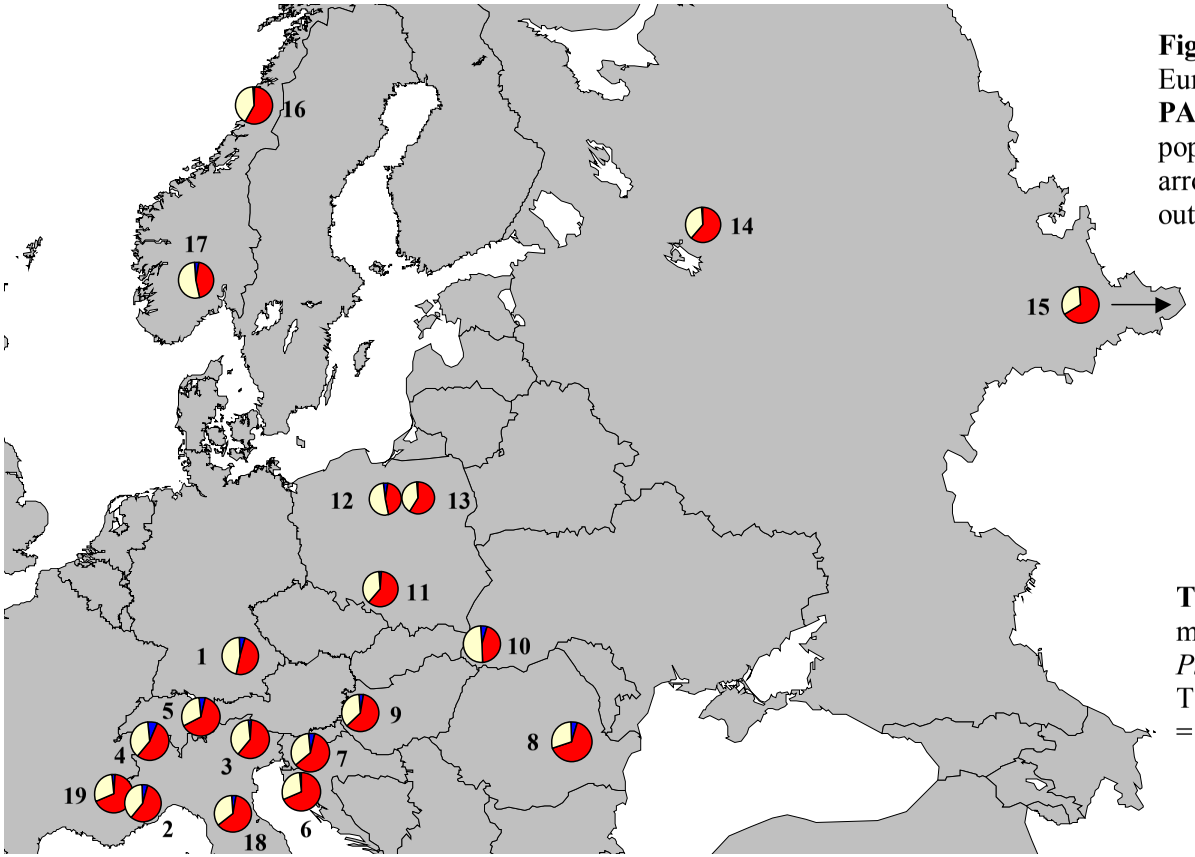


Figure 3-3f. Allele frequency distribution within European populations of Norway spruce at locus **PA0038.** The numbers and coordinates of populations as defined in Tab. 2, respectively; the arrow indicates that population no.15 is situated outside the map.

Table 3-1f. Frequency distribution for 3 alleles (A-C) of marker **PA0038**, measured in 19 European populations of *Picea abies* (populations are numbered according to Tab.2; geographic positions are indicated in Fig. 3-3f; bp = size of the PCR fragment in base pairs).

Population		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
Allele	**A (186 bp)**	6	3	3	9	5	1	5	5	4	5	3	5	2	2	0	1	4	4	5
	B (192 bp)	47	63	55	50	59	66	57	64	58	43	56	41	54	58	65	56	41	60	55
	C (228 bp)	47	34	41	41	35	33	38	31	38	52	41	55	43	40	35	43	55	36	40

3.2.1.2. Genotype frequencies at single loci

Genotypes frequencies of 19 Norway spruce populations at 6 analysed EST loci are represented in the *Tab.3-2 (a-f)* and *Fig.3-4 (a-f)*. The number of the found genotypes per loci varies between 13 (EST locus PA0034) and 3 (EST locus PA0055).
Dominating frequencies of one genotype (more than 35%) are observed in all analysed populations in cases of the markers PA0005, PA0034, PA0055.
For the marker PA0005, greatest values of genotype BA are detected in the populations from Italy (32-33%), Southern France (31,8%) and Switzerland (19,6-39,6 %) which represent the Alpine region. This genotype is absent or appears rare in the samples from Northern Poland (6,1-6,7%), Russia (0-2%) and Norway (4,2-8,7%) (*Tab.3-2a, Fig.3-4a*). All genotypes with low frequencies were found only in the populations from Russia or Central–Eastern Europe, but not within the Alpine region.

The comparable uneven distribution of the frequencies of genotype AB in space was found at the locus PA0055 (*Tab.3-2b, Fig.3-4b*). This genotype is present with highest values (29, 4-57 %) in the area of Alps (Italian, Swiss, French populations) and appears rare (0-8, 3 %) in the populations from northern Poland, Russia and Norway.
The Russian and Norwegian populations also reveal lowest frequencies of the genotype DB at the locus PA0066 (*Tab.3-2c, Fig.3-4c*). It amounts to 8, 3-17, 8 % in contrast to 19-56, 7 % found for the same genotype in the samples from Alpine, Central and Eastern Europe excluding the Romanian population (16, 7 %). The maximal values of the genotype BB were found at this locus in all populations from Russia and Norway. Two Russian and one Norwegian samples demonstrate also the highest frequencies of the genotype BA.
The highest frequencies of the genotype BA (populations No. 14, 15, 16) and the genotype BB (populations No. 14, 15, 16, 17) were detected within Russian-Northern-European group of analysed populations at the locus PA0043 (*Tab.3-2d, Fig.3-4d*)
Analysing genotypes frequencies at the loci PA0034 (*Tab.3-2e, Fig.3-4e*) and PA0038 (*Tab.3-2f, Fig.3-4f*), heterogeneous distributions without geographical trends were observed.

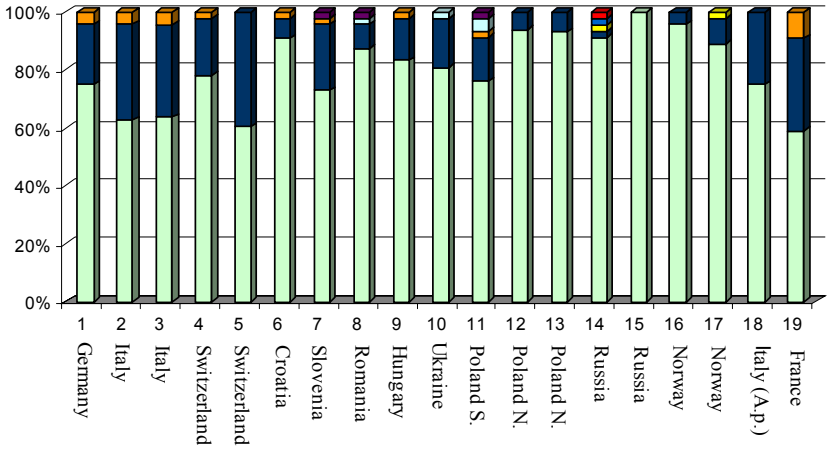


Figure 8 (a). Distribution of genotype frequencies at locus PA0005 found within 19 European populations.

Table 8 (a). Genotype frequencies (%) at locus PA0005 within 19 populations representing natural range of *Picea abies* in Europe.

Genotypes	*Populations*																		
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**
AA	75	63	64	78,3	60,4	91,1	72,9	87,5	83	80,9	76,3	93,9	93,3	91	100	95,8	89	75	59,1
BA	20,8	33	32	19,6	39,6	6,7	22,9	8,3	15	17	15,2	6,1	6,7	2,2	-	4,2	8,7	25	31,8
BB	4,2	4,2	4,3	2,2	-	2,2	2,1	-	2,1	-	2,2	-	-	-	-	-	-	-	9,1
CA	-	-	-	-	-	-	-	2,1	-	2,1	4,3	-	-	-	-	-	-	-	-
DA	-	-	-	-	-	-	2,1	2,1	-	-	2,2	-	-	-	-	-	-	-	-
EA	-	-	-	-	-	-	-	-	-	-	-	-	-	2,2	-	-	2,2	-	-
EE	-	-	-	-	-	-	-	-	-	-	-	-	-	2,2	-	-	-	-	-
FA	-	-	-	-	-	-	-	-	-	-	-	-	-	2,2	-	-	-	-	-

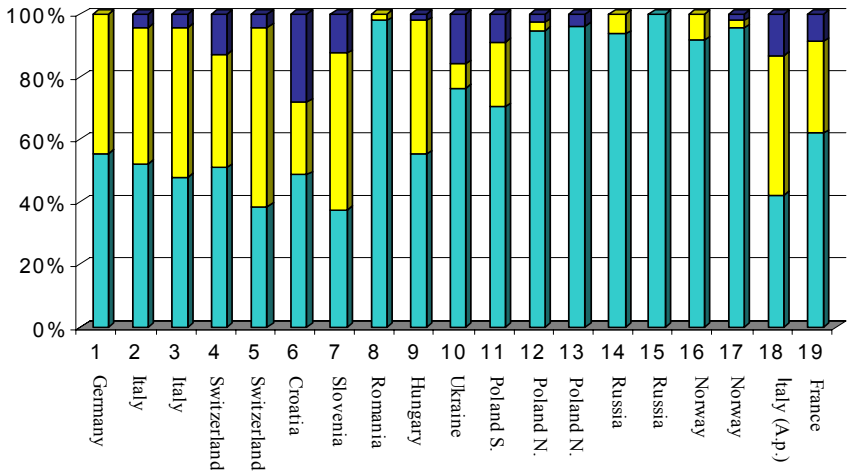


Figure 8 (b). Distribution of genotype frequencies at locus PA0055 found within 19 European populations.

Table 8 (b). Genotype frequencies (%) at locus PA0055 within 19 populations representing natural range of *Picea abies* in Europe.

Genotypes	*Populations*																		
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**
AA	54,2	52	48	51	38	49	38	98	55	76,3	70,5	94,9	96,3	96,3	100	91,7	95,7	42,2	61,8
BA	43,7	44	48	36	57	23	50	2,1	43	7,9	20,5	2,6	-	6,4	-	8,3	2,2	44,4	29,4
BB	-	4,2	4,2	13	4,3	28	13	-	2,1	15,8	9,1	2,6	3,7	-	-	-	2,2	13,3	8,8

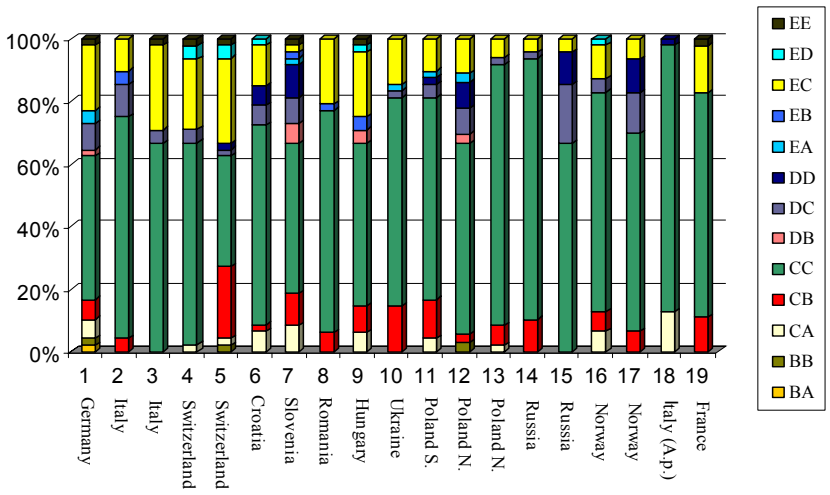


Figure 8 (c). Distribution of genotype frequencies at locus PA0034 found within 19 European populations of Norway spruce.

Table 8 (c). Genotype frequencies (%) at locus PA0034 within 19 populations representing natural range of *Picea abies* in Europe.

Genotypes	*Populations*																		
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**
BA	2,1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
BB	2,1	-	-	-	2,1	-	-	-	-	-	-	2,8	-	-	-	-	-	-	-
CA	6,2	-	-	2,2	2,1	6,4	8,3	-	6,2	-	4,2	-	2,1	-	-	6,4	-	13	-
CB	6,2	4,2	-	-	23	2,1	10	6	8,3	14,6	13	2,8	6,2	10	-	6,4	6,5	-	10,9
CC	45,8	71	67	63	35	64	48	71	52	66,7	65	61	83,3	83	66,7	70,2	63	85	71,7
DB	2,1	-	-	-	-	-	6,2	-	4,2	-	-	2,8	-	-	-	-	-	-	-
DC	8,3	10	4	4,3	2,1	6,4	8,3	-	-	2,1	4,2	8,3	2,1	2,1	18,7	4,3	13	-	-
DD	-	-	-	-	2,1	6,4	10	-	-	-	2,1	8,3	-	-	10,4	-	11	2,1	
EA	4,2	-	-	-	-	-	2,1	-	-	2,1	2,1	2,8	-	-	-	-	-	-	-
EB	-	4,2	-	-	-	-	2,1	2	4,2	-	-	-	-	-	-	-	-	-	-
EC	20,8	10	27	22	27	13	2,1	21	21	14,6	10	11	6,2	4,2	4,2	10,6	6,5	-	15,2
ED	-	-	-	4,3	4,2	2,1	-	-	2,1	-	-	-	-	-	-	2,1	-	-	-
EE	2,1	-	2	2,2	2,1	-	2,1	-	2,1	-	-	-	-	-	-	-	-	-	2,2

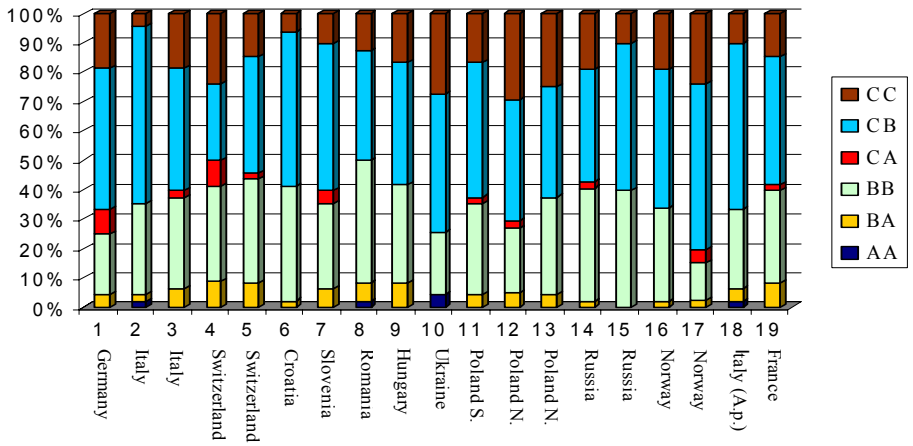


Figure 8(d). Distribution of genotype frequencies at locus PA0038 found within 19 European populations of Norway spruce.

Table 8 (d). Genotype frequencies (%) at locus PA0038 within 19 populations representing natural range of *Picea abies* in Europe.

Genotypes	*Populations*																		
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**
AA	-	2,1	-	-	-	-	-	2,1	-	4,3	-	-	-	-	-	-	-	2,1	-
BA	4,2	2,1	6,2	8,7	8,3	2,2	6,2	6,2	8,3	-	4,2	4,9	4,2	2,1	-	2,1	2,2	4,2	8,3
BB	20,8	31	31	33	35	39,1	29	42	33,3	21	31,2	22	33,3	38,3	39,6	31,9	13	27,1	31,2
CA	8,3	-	2,1	8,7	2,1	-	4,2	-	-	-	2,1	2,4	-	2,1	-	-	4,3	-	2,1
CB	47,9	60	42	26	40	52,2	50	38	41,7	47	45,8	41,5	37,5	38,3	50	46,8	56,5	56,3	43,7
CC	18,7	4,2	19	24	15	6,5	10	13	16,7	28	16,7	29,3	25	19,1	10,4	19,1	23,9	10,4	14,6

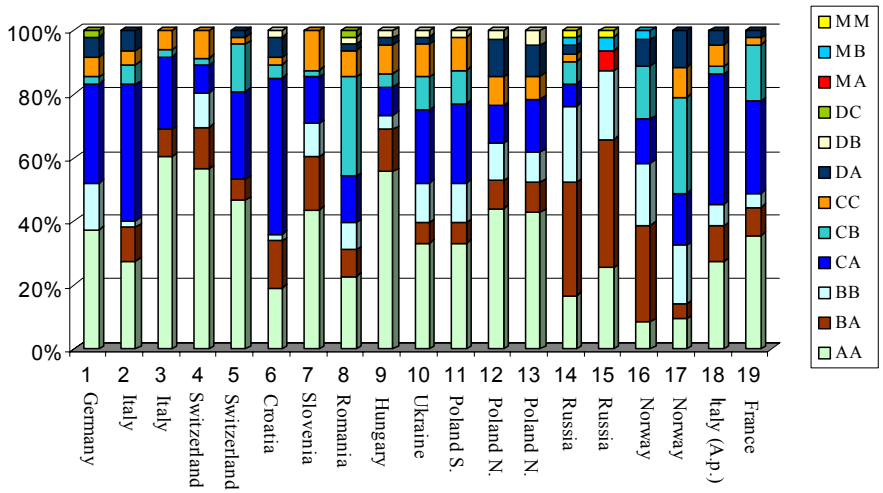


Figure 8 (e). Distribution of genotype frequencies at locus PA0043 found within 19 European populations of Norway spruce.

Table 8 (e). Genotype frequencies (%) at locus PA0043 within 19 populations representing natural range of *Picea abies* in Europe.

Genotypes	*Populations*																		
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**
AA	37,5	27,7	60,4	56,5	46,8	19,1	43,7	22,9	55,6	33,3	33,3	44,1	42,9	16,7	25,5	7,7	9,3	27,3	35,6
BA	-	10,6	8,3	13	6,4	14,9	16,7	8,3	13,3	6,2	6,2	8,8	9,5	35,7	40,4	28,2	4,7	11,4	8,9
BB	14,6	2,1	-	10,9	-	2,1	10,4	8,3	4,4	12,5	12,5	11,8	9,5	23,8	21,3	17,9	18,6	6,8	4,4
CA	31,2	42,6	22,9	8,7	27,7	48,9	14,6	14,6	8,9	22,9	25	11,8	16,7	7,1	-	12,8	16,3	40,9	28,9
CB	2,1	6,4	2,1	2,2	14,9	4,3	2,1	31,2	4,4	10,4	10,4	-	-	7,1	-	15,4	30,2	2,3	17,8
CC	6,2	4,3	6,2	8,7	2,1	2,1	12,5	8,3	8,9	10,4	10,4	8,8	7,1	2,4	-	-	9,3	6,8	2,2
DA	6,2	6,4	-	-	2,1	6,4	-	2,1	2,2	2,1	-	11,8	9,5	2,4	-	7,7	11,6	4,5	-
DB	-	-	-	-	-	2,1	-	2,1	2,2	2,1	2,1	2,9	4,8	-	-	-	-	-	-
DC	2,1	-	-	-	-	-	-	2,1	-	-	-	-	-	-	-	-	-	-	-
MA	-	-	-	-	-	-	-	-	-	-	-	-	-	-	6,4	-	-	-	-
MB	-	-	-	-	-	-	-	-	-	-	-	-	-	2,4	4,3	2,6	-	-	-
MM	-	-	-	-	-	-	-	-	-	-	-	-	-	2,4	2,1	-	-	-	-

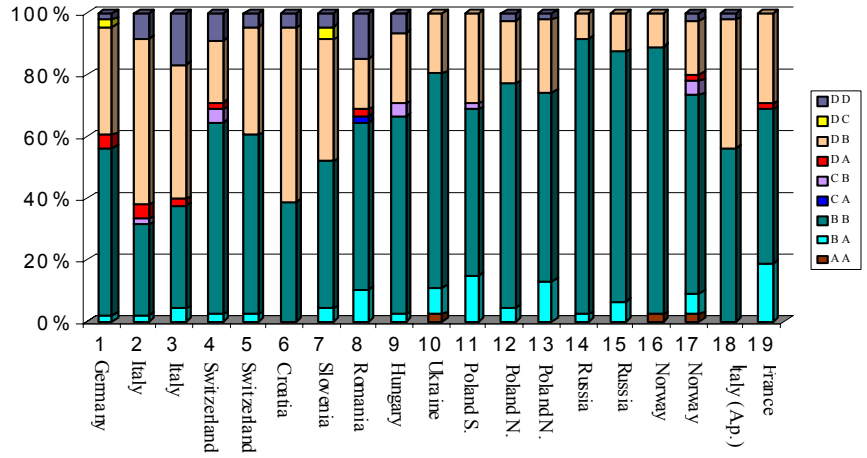


Figure 8 (f). Distribution of genotype frequencies at locus PA0066 found within 19 European populations of Norway spruce.

Table 8 (f). Genotype frequencies (%) at locus PA0066 within 19 populations representing natural range of *Picea abies* in Europe.

Genotypes	*Populations*																		
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**
AA	-	-	-	-	-	-	-	-	-	2,2	-	-	-	-	-	2,2	2,2	-	-
BA	2,1	2,1	4,2	2,2	2,1	-	4,2	10,4	2,1	8,7	14,6	4,5	12,8	2,1	6,2	-	6,7	-	18,7
BB	54,2	29	33	62,2	58,3	38,6	47,9	54,2	64,6	69,6	54,2	72,7	61,7	89,6	81,3	87	64,4	56,3	50
CA	-	-	-	-	-	-	-	2,1	-	-	-	-	-	-	-	-	-	-	-
CB	-	2,1	-	4,4	-	-	-	-	4,2	-	2,1	-	-	-	-	-	4,4	-	-
DA	4,2	4,2	2,1	2,2	-	-	-	2,1	-	-	-	-	-	-	-	-	2,2	-	2,1
DB	35,4	54	44	20	35,4	56,8	39,6	16,7	22,9	19,6	29,2	20,5	23,4	8,3	12,5	10,9	17,8	41,7	29,2
DC	2,1	-	-	-	-	-	4,2	-	-	-	-	-	-	-	-	-	-	-	-
DD	2,1	8,3	17	8,9	4,2	4,5	4,2	14,6	6,2	-	-	2,3	2,1	-	-	-	2,2	2,1	-

3.2.1.3 Heterozygosity

The heterozygosity observed within 19 populations of Norway spruce for six EST markers estimates range from 18% to 51% with a mean of 35% (*Tab. 3-4*). The marker PA0038 reveals greatest proportion of heterozygosity. In the case of this marker level of observed heterozygosity varies between 42% and 63%. Lowest values of heterozygosity were observed at loci PA0005. It ranges from 4% to 40% with highest proportion of heterozygosity in the populations from Italy (No. 2, No.3), Switzerland (No.5) and France (No.19) with range from 32% to 40%. The lowest heterozygosity values were found at the same locus in the populations from Croatia (No.6), northern Poland (No.12, No.13), Russia (No.14) and Norway (No.16, No.17) (6%-11%). Extreme difference in geographical distribution of observed heterozygosity reveals EST marker PA0055. High proportion of heterozygosity (36%-57%) was found in the German (No.1), Italian (No.2, No.3, No.18), Swiss (No.4, No.5) and Slovenian (No.7) populations, while in the populations from Romania (No.8), Ukraine (No.10), northern Poland (No.12, No.13), Russia (No.14) and Norway (No.16, No.17) it estimates range from 0 to 8%.

3.2.1.4 Diversity

Total means of genetic diversity at six analysed EST loci based on allelic frequencies is 1,7. Greatest average value of diversity estimated for alleles was found at locus PA0043 (*Tab. 3-3*). Smallest mean is observed at locus PA0005. Interesting that at this locus populations from Italy (No.2, No.3), Switzerland (No.5) and France (No.19) were found most diverse whereas populations from northern Poland (No.12, No.13), Russia (No.14) and Norway (No.16, No.17) demonstrate the lowest values of the same parameter. Russian (No.14), Norwegian (No.16, No.17), northern Polish (No.12, No.13) as well as Romanian (No.8) populations indicated the lowest values of diversity at the locus PA0055. The same effect was observed for the populations from Russia (No.14, No.15) and Norway (No.16) at the locus PA0066.

Table 3-3. The means of observed genetic diversity v_N based on allele frequencies and measured within 19 populations of Norway spruce at 6 EST-marker loci.

EST locus	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
PA0005	1,33	1,49	1,48	1,27	1,47	1,12	1,34	1,14	1,21	1,21	1,31	1,10	1,10	1,12	1	1,04	1,12	1,28	1,6
PA0034	2,09	1,42	1,42	1,61	2,26	1,61	2,29	1,38	1,92	1,45	1,53	1,82	1,19	1,19	1,54	1,40	1,64	1,19	1,37
PA0038	2	1,96	2,12	2,34	2,08	1,83	2,12	1,98	2,07	2,16	2,01	2,14	2,06	2,01	1,8	2	2,1	2,01	2,12
PA0043	3	2,35	1,62	1,97	2,05	2,53	2,30	3,17	1,98	2,77	2,73	2,37	2,35	2,63	2,30	3,12	3,29	2,44	2,47
PA0055	1,57	1,63	1,68	1,74	1,79	1,92	1,88	1,02	1,56	1,46	1,45	1,08	1	1,06	1	1,09	1,07	1,85	1,56
PA0066	1,71	2,08	2,06	1,63	1,56	1,79	1,80	1,91	1,50	1,4	1,61	1,35	1,52	1,11	1,21	1,17	1,56	1,55	1,72

Table 3-4: The means of observed H_O and conditional H_C heterozygosity measured within 19 tested populations of Norway spruce at 6 EST marker loci.

PA0005

Population No.	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**
$\mathbf{H_O}$	0,21	0,33	0,32	0,20	0,40	0,07	0,25	0,13	0,15	0,19	0,22	0,06	0,07	0,07	0	0,04	0,11	0,25	0,32
$\mathbf{H_C}$	0,72	0,8	0,79	0,82	1	0,6	0,86	1	0,79	1	0,83	1	1	0,6	*	1	1	1	0,64

PA0034

Population No.	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**
$\mathbf{H_O}$	0,5	0,29	0,31	0,35	0,58	0,30	0,40	0,29	0,46	0,33	0,33	0,23	0,17	0,17	0,23	0,30	0,26	0,13	0,26
$\mathbf{H_C}$	0,75	0,87	0,88	0,76	0,78	0,67	0,53	0,93	0,76	0,94	0,84	0,50	1	1	0,52	0,93	0,55	0,75	0,86

PA0038

Population No.	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**
$\mathbf{H_O}$	0,60	0,63	0,50	0,44	0,50	0,54	0,60	0,44	0,50	0,47	0,52	0,49	0,42	0,07	0,50	0,49	0,63	0,60	0,54
$\mathbf{H_C}$	0,60	0,83	0,56	0,44	0,62	0,81	0,71	0,60	0,60	0,47	0,60	0,50	0,46	0,51	0,71	0,56	0,69	0,71	0,63

PA0043

Population No.	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**
$\mathbf{H_O}$	0,44	0,63	0,50	0,29	0,37	0,57	0,48	0,31	0,29	0,28	0,46	0,25	0,36	0,10	0,19	0,11	0,31	0,42	0,50
$\mathbf{H_C}$	0,81	0,75	0,58	0,59	0,82	0,86	0,79	0,48	0,70	0,87	1	0,85	0,90	1	1	0,72	0,74	0,91	0,96

PA0055

Population No.	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**
$\mathbf{H_O}$	0,44	0,44	0,48	0,36	0,57	0,23	0,50	0,02	0,43	0,08	0,21	0,03	0	0,06	0	0,08	0,02	0,44	0,29
$\mathbf{H_C}$	0,91	0,84	0,85	0,59	0,87	0,29	0,67	1	0,91	0,20	0,53	0,33	0	1	*	1	0,33	0,63	0,63

PA0066

Population No.	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**
$\mathbf{H_O}$	0,44	0,6 3	0,50	0,29	0,38	0,57	0,48	0,31	0,29	0,28	0,46	0,06	0,36	0,10	0,19	0,11	0,31	0,42	0,5
$\mathbf{H_C}$	0,81	0,75	0,59	0,59	0,82	0,86	0,79	0,48	0,7	0,87	1	0,85	0,90	1	1	0,71	0,74	0,91	0,96

3.2.2 Interpopulational variation

In order to test genetic differentiation between populations, Gregorius genetic distance, Gregorius subpopulation differentiation (D_j, δ) and Wright`s F-statistic (F_{ST}) were measured.

3.2.2.1 Genetic distance

The genetic distance between two populations was specified as the proportion of alleles which the two populations do not share. This genetic parameter was calculated for all pairs of populations. The matrices of Gregorius genetic distance on the basis of the allele frequencies were used for creation of the loci-specific and summarized (gene pool) dendrograms (*Fig. 3-5*). The values of genetic distance calculated for the gene pool vary between 0.011 and 0.3. The dendrogram created for gene pool groups together in cluster I three branches represented by the very closely related Alpine populations No. 2, 3, 5, Central-European populations No. 4, 7, 18, 1 and Eastern-European populations No.11, 9, 10 respectively. Apart from the Croatian population No. 6 and the Romanian population No. 8, cluster II only comprises population samples from the Northern European-Russian gene pool (No.14, 13, 12, 17, 15, and 16). Surprisingly, the position of Southern French population No.19 is not corresponding with geographically structured clustering of the rest of remaining populations. This population is represented on the dendrogram of gene pool absolutely separated.

Each of the dendrograms created for EST loci PA0005 and PA0055 groups the geographically related populations into three clusters. In the case of the marker PA0005 first cluster is represented by the Alpine populations (No. 19, 2, 3, 5); second cluster comprises heterogeneous populations (No. 8, 10, 11) from the Hercyno-Carpathian gene pool in relationship with the Alpine populations No. 4 and 7, the population No.18 from Appenine peninsula and the Southern German population No. 1; the Croatian population No. 6 and the Romanian population No. 9 are included in the cluster III together with all populations from the Northern European-Russian gene pool (No.14, 13, 12, 17, 15, 16). Interestingly is clustering for the locus PA0055. Most separated cluster III the Romanian population No. 8 consists only from populations representing Northern European-Russian gene pool (No. 14, 13, 12, 17, 15, 16); cluster I and cluster II are grouped together. Cluster I comprises the populations from Hercyno-Carpathian gene pool (No. 9, 10, 11), the Southern German population No.1, Southern French population No. 19 and Alpine populations from Italy (No.

2, 3); cluster II comprises Swiss populations (No. 4, 5), the populations from Balkan peninsula (No. 6, 7) and population No.18 from Appenine peninsula.

A bit less evident but still related to the geographical location clustering of remaining populations is illustrated by the dendrograms for the loci PA0043 and PA0066.

In contrast, such a geographically structured genetic differentiation was not obvious when the populations mentioned above were assessed by the co-dominant EST-PCR markers PA0034 and PA0038.

3.2.2.2 Differentiation among populations

D_j, δ (Gregorius)

This measure of genetic differentiation is based on genetic distances between one sample (population) and the remaining ones which are pooled as the respective complement population. The *Fig. 3-6* shows graphically the genetic differentiation between 19 analysed populations, based on the allele frequencies, and measured for 6 single EST loci and the gene pool with respect to all loci together. The graphs illustrate the average level of differentiation (δ) as the radius of the circle and the proportion of genes in which one population differs from the remainder (D_j) as the radii of the population specific sectors.

The obvious deviations in the D_j-values observed among populations suggest the geographical character of detected genetic variation. Thus, the populations from Russia (No. 14 and 15) are shown like samples with great amount of differentiation between it and the remainder for all single loci and for the gene pool (excluding No.14 for the locus PA0038). The same trend is observed for Romanian and both Norwegian populations for the gene pool and 4 loci (No.16 for the loci PA0005, PA0043, PA0055, PA0066, No.17 for the loci PA0034, PA0038, PA0043, PA0055, No.9 for the loci PA0034, PA0038, PA0043, PA0055 respectively). In contrast, the populations No.8 from Hungary and No.10 from Poland reveal no differences to remainder for any locus. The average level of differentiation as well as its specific character vary among loci and reflect a general trend in the case of the gene pool, because all studied loci are addressed as a whole (Müller-Starck, 1995). The gene pool graph reveals that all analysed samples from Northern Europe (No. 12, 13, 14, 15, 16, 17) together with samples from Croatia (No. 6), Romania (No. 9), and two samples from Italy (No. 3, No. 18) tend to carry specific information in contrast to other populations. The same graph shows the greatest value of differentiation for Russian and Norwegian populations.

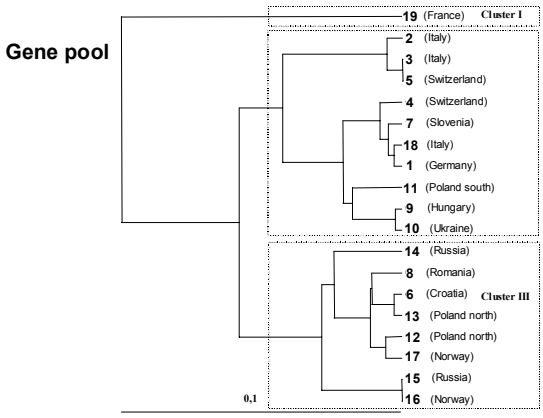
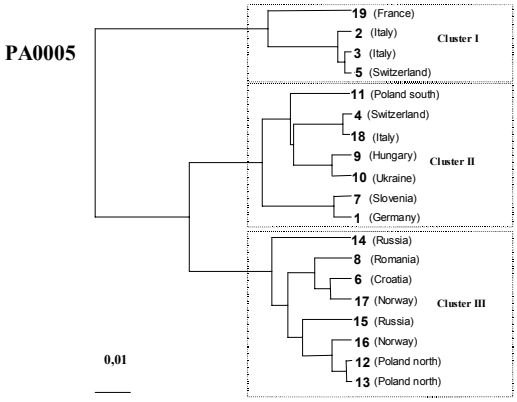
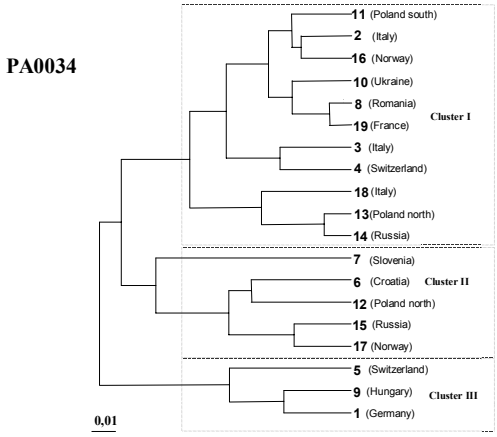
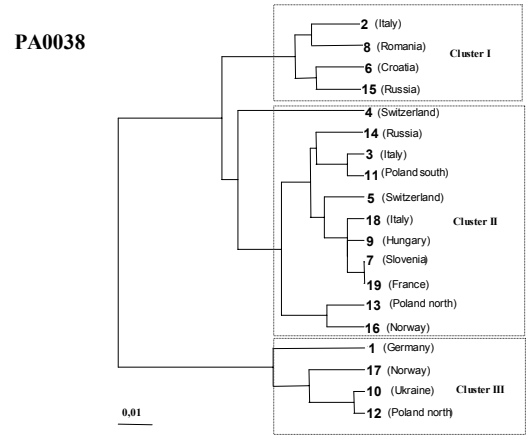
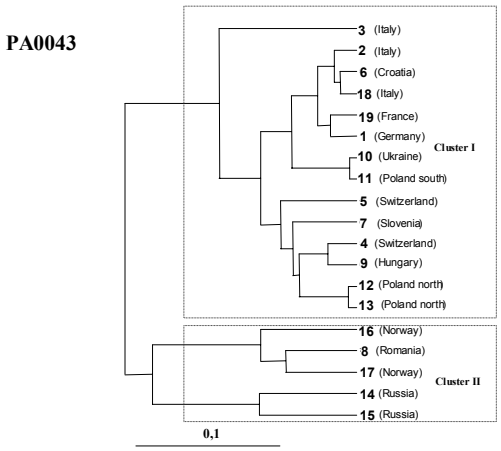


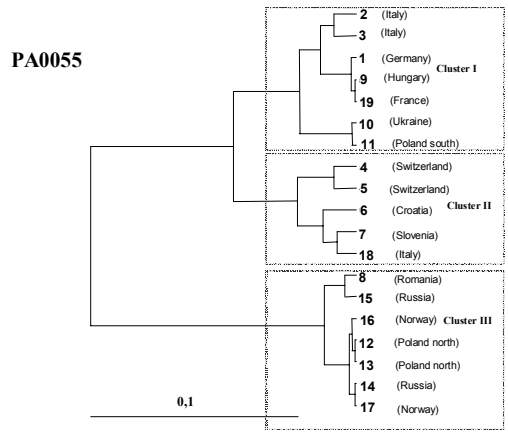
Figure 3-5. Dendrograms based on Gregorius genetic distancees for 19 European spruce populations, single EST markers and 6 analysed loci pooled together (gene pool) (population numbers according to Table 2-1).

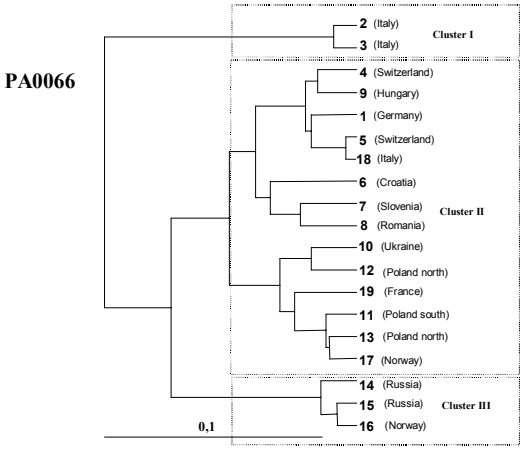












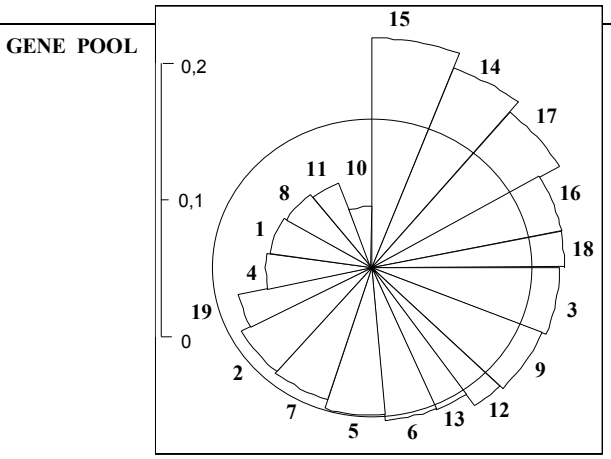
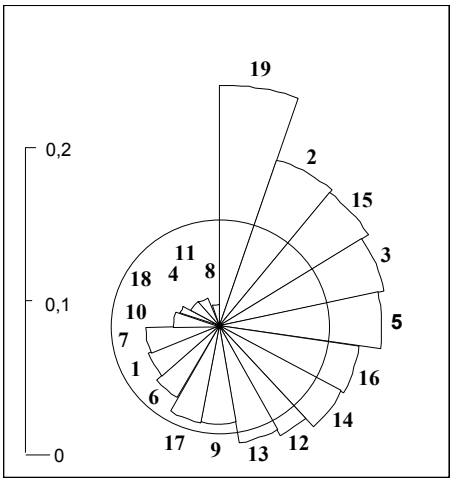
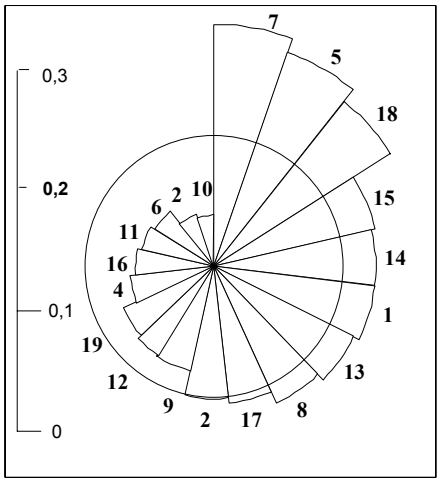


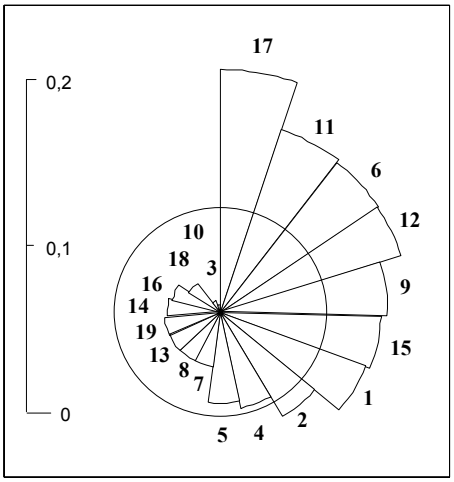
Figure 3-6. Genetic differentiation (D_j, δ) among 19 tested populations of *Picea abies* for 6 single EST loci and gene pool with respect to all loci together.



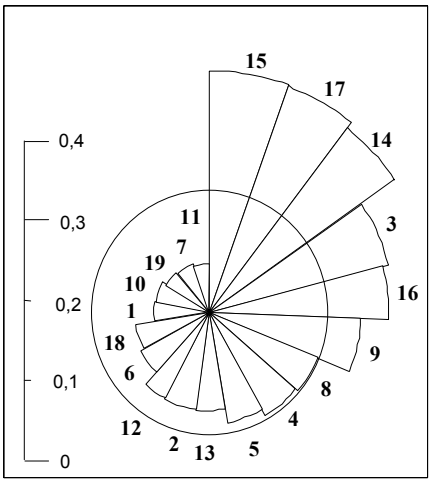
PA0034



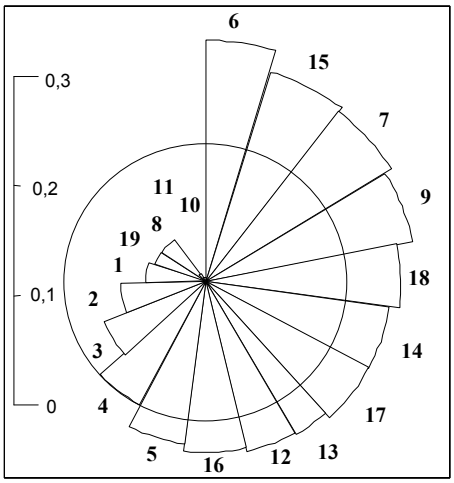
PA0038



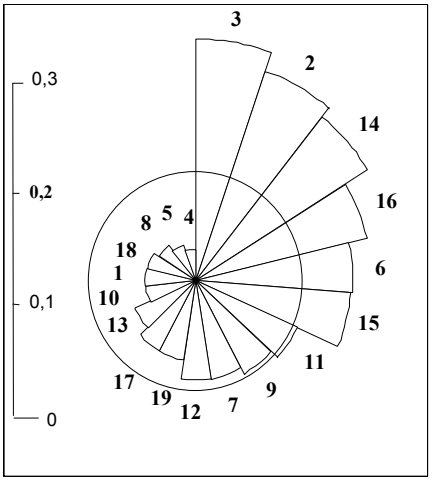
PA0043



PA0055



PA0066



F_{ST} (Wright)

Wright`s F_{ST} measures populational differentiation under the "island model". The values of F_{ST} observed for 6 EST loci are small and vary between 0,02 (PA0038) and 0,2 (PA0055) (Tab. 3-5). The mean value combines 0,0585.

Table 3-5. The F_{ST} values measured among 19 populations of Norway spruce at 6 EST loci

EST locus	**PA0005**	**PA0034**	**PA0038**	**PA0043**	**PA0055**	**PA0066**	**mean**
F_{ST}	0,0536	0,0393	0,0217	0,059	0,1181	0,0572	0,0585

3.2.3 Isolation by distance test

The matrix of Euclidean geographical distance between populations was compared with different matrices of genetic distance calculated for 6 EST loci and the gene pool (all 6 loci together) to test positive correlations between genetic differentiation and geographical separation of the pairs of populations. The permutation test resulted in high significant P-values ($P(\delta)$) characterizing correlations between genetic and geographic distances for the gene pool as well as for the EST markers PA0005, PA0043, PA0055, PA0066 (see *Tab. 3-6*). Statistically significant P-value rejects the null hypothesis of independence of two matrices and indicates that populations are genetically isolated by distance. No significant correlations were found when the loci PA0034 and PA0038 were tested.

Table 3-6. Correlations between matrices of genetic and geographic distances for the 19 Norway spruce populations tested under *Isolation-by-distance* model. **P(δ)** – P-value; ***R***-correlation coefficient
*- significant value; **- high significant value; n.s.- not significant value.

Locus tested	All loci	**PA0005**	**PA0034**	**PA0038**	**PA0043**	**PA0055**	**PA0066**
P(δ)	0.002**	0.006**	0.37200 n.s.	0.19200 n.s.	0.0030**	0.0034**	0.0110*
R	0.0799	0.0777	0.1414	0.0831	0.1418	0.1266	0.1192

Following *Isolation-by-distance* test was performed for the gene pool and for the loci PA0005, PA0043, PA0055, and PA0066 were positive significant correlations were indicated. All populations were separated in accordance to their latitudes on the three

different geographical domains (see *Materials and Methods*). Positive correlations between matrices of genetic and geographic distances were tested for the populations belonging to the same geographical domain and two domains in different combinations (*Tab. 3-7*).

Not significant P-values were detected in all tests of single Northern-European domain that indicates not found correlations of geographic and genetic separation at analysed EST loci between populations from Russia, Norway and Northern Poland. No significant correlations were found between the populations from Italy, Switzerland, Southern France, Croatia and Slovenia belonging to the Southern-European domain. However, significant mean of P (δ) was detected within Eastern-Central-European domain (populations from Southern Germany, Southern Poland, Hungary, Ukraine and Romania) analysing all loci (pooled) and locus PA0005. All tests of the combination Northern-European and Southern-European domains indicate significant or high significant correlations of genetic and geographic distances between populations from these two domains. No significant P-values are found at locus PA0055 testing the populations in combination Northern-European and Eastern-Central-European domains. But tests of the same combination of domains indicate significant and high significant correlations at loci PA0005, PA0066 and in the cases of the locus PA0043 and the gene pool respectively.

Table 3-7. Correlations detected between matrices of geographical and genetic distances among populations located within one out of three geographical domains and three combinations of these domains.

Loci	**Single domains**			**Combinations of domains**		
	(NE) Northern-European domain	(***SE***) Southern-European domain	(***CE***) Eastern-Central European domain	***NE*** and ***SE***	***NE*** and ***CE***	***SE*** and ***CE***
Pool (4 loci)	0,207 n.s.	0,281 n.s.	0,043 *	0,040 *	0,007 **	0,047 *
PA0005	0,295 n.s.	0,237 n.s.	0,035 *	0,013 *	0,053 *	0,018 *
PA0055	0,108 n.s.	0,076 n.s.	0,096 n.s.	0,001 **	0,141 n.s.	0,004 **
PA0043	0,116 n.s.	0,942 n.s.	0,335 n.s.	0,002 **	0,005 **	0,038 *
PA0066	0,536 n.s.	0,742 n.s.	0,528 n.s.	0,015 *	0,035 *	0,385 n.s.

3.2.4 Neutrality test

The Ewens-Watterson homozygosity test was performed for each from 6 analysed EST loci to test the hypothesis of selective neutrality. The number of 1000 simulations for computing 95% lower and upper confidence limits (see POPGENE, version 1.32) was selected. The values of "Obs. F" were found within 95% confidence interval indicating neutrality for all populations analysing 6 EST loci (*see Tab. 3-8*). However, large "Obs. F" values (> 0,80) where detected within populations No. 6, 8, 9, 10, 12, 13, 14, 16, 17 at the locus PA0005, within populations No. 13, 14, 18 at the locus PA0034, within populations No. 8, 12, 13, 14, 16, 17 at the locus PA0055, and within populations No. 14, 15, 16 at the locus PA0066. It suggests still not significant but stronger tendency in favour of homozygotes within these populations in comparison with others.

Table 3-8. Results of the Ewens-Watterson homozygosity test provided for 19 populations of Norway spruce and 6 EST loci. L95* and U95* are 95% lower and upper confidence limits respectively. Blue colour indicates large values of "Obs. F" (>0, 80)

Locus		Population No.																		
		1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**
PA0005	**Obs. F**	0,75	0,67	0,68	0,79	0,68	0,90	0,75	0,88	0,83	0,83	0,77	0,94	0,94	0,89	1	0,96	0,90	0,78	0,63
	L95*	0,50	0,50	0,50	0,50	0,50	0,50	0,37	0,30	0,50	0,38	0,30	0,50	0,50	0,31	-	0,50	0,36	0,50	0,50
	U95*	0,97	0,97	0,98	0,98	0,98	0,98	0,96	0,92	0,98	0,96	0,90	0,97	0,98	0,91	-	0,98	0,96	0,96	0,98
PA0034	**Obs. F**	0,48	0,70	0,70	0,62	0,44	0,62	0,53	0,73	0,52	0,69	0,65	0,55	0,81	0,84	0,65	0,71	0,61	0,84	0,73
	L95*	0,27	0,31	0,37	0,26	0,26	0,25	0,25	0,36	0,26	0,27	0,26	0,26	0,26	0,30	0,38	0,26	0,32	0,37	0,37
	U95*	0,84	0,90	0,96	0,82	0,84	0,84	0,77	0,96	0,84	0,84	0,84	0,82	0,83	0,90	0,96	0,84	0,90	0,96	0,96
PA0038	**Obs. F**	0,44	0,51	0,47	0,43	0,48	0,55	0,47	0,50	0,48	0,46	0,48	0,47	0,48	0,50	0,54	0,47	0,45	0,52	0,48
	L95*	0,36	0,37	0,36	0,36	0,37	0,36	0,37	0,38	0,37	0,36	0,36	0,36	0,36	0,36	0,50	0,36	0,36	0,39	0,37
	U95*	0,95	0,96	0,96	0,96	0,96	0,96	0,96	0,96	0,96	0,95	0,96	0,94	0,95	0,95	0,98	0,95	0,95	0,96	0,98
PA0043	**Obs. F**	0,40	0,43	0,62	0,51	0,49	0,46	0,51	0,32	0,51	0,38	0,37	0,42	0,41	0,38	0,44	0,33	0,32	0,47	0,54
	L95*	0,31	0,31	0,38	0,37	0,30	0,32	0,36	0,31	0,30	0,36	0,31	0,32	0,32	0,25	0,37	0,26	0,30	0,38	0,38
	U95*	0,90	0,90	0,96	0,96	0,92	0,90	0,94	0,90	0,89	0,89	0,90	0,89	0,92	0,82	0,96	0,77	0,84	0,94	0,96
PA0055	**Obs. F**	0,64	0,62	0,60	0,57	0,56	0,52	0,53	0,97	0,64	0,68	0,69	0,93	0,93	0,94	1	0,92	0,94	0,54	0,64
	L95*	0,50	0,50	0,50	0,50	0,50	0,50	0,50	0,50	0,50	0,50	0,50	0,50	0,50	0,50	-	0,50	0,50	0,50	0,50
	U95*	0,98	0,98	0,98	0,98	0,98	0,98	0,98	0,98	0,98	0,97	0,98	0,98	0,96	0,98	-	0,98	0,98	0,98	0,97
PA0066	**Obs. F**	0,59	0,48	0,49	0,71	0,65	0,56	0,56	0,52	0,66	0,72	0,62	0,74	0,66	0,90	0,83	0,86	0,64	0,65	0,58
	L95*	0,32	0,31	0,38	0,36	0,50	0,50	0,30	0,31	0,31	0,36	0,31	0,36	0,36	0,37	0,37	0,37	0,30	0,50	0,37
	U95*	0,90	0,90	0,96	0,94	0,98	0,98	0,92	0,92	0,90	0,96	0,90	0,96	0,96	0,96	0,96	0,96	0,91	0,98	0,96

Discussion

Main focus of the present study was the combination of EST marker development and its use in the quantification of genetic variation along natural European range of Norway spruce. This kind of molecular markers is addressing the wide spectrum of the functionally active genes. In the contrast, other DNA markers highlight polymorphism often within not transcribed genome regions, and biochemical markers such as isoenzymes correspond to the limited number of expressed genes. It makes application of EST markers very important for the study of genetic variability, understanding of the genome structure and the mechanisms of adaptability in Norway spruce.

Picea abies is a major species distributed in Central, Eastern and Northern Europe characterized by various elevations and climatic conditions, respectively. The patterns of genetic variation of Norway spruce in Europe were described by morphological, biochemical and chloroplast SSR markers previously. EST markers are applied to such study in *Picea abies* for the first time.

4.1 Cyclophilin EST marker

In the processing of markers, a new codominant EST marker (PA0005) was developed for the cDNA clone pPA0005 which encodes a protein with a molecular weight of 18 kDa, lying within the range found for the majority of cyclophilins. Its deduced amino acid sequence reveals an identity of up to 85% with already known cyclophilins from different plant sources. Further evidence of the cyclophilin character of the cDNA clone provided the discovery that the corresponding amino acid sequence of the cDNA clone pPA0005 contains the putative cyclosporin A binding sites unique for cyclophilins, including the conserved trypthophane residue (position 128) that is essential for the drug binding (detailed information on these conserved residues given by Ostoa-Saloma *et al.* 2000), as cyclophilins are characterized by both peptidyl-propyl cis-trans isomerase and cyclosporin A binding activity.

Cyclophilins, catalyzing cis-trans isomerization of proline imidic peptide bonds in oligopeptides, have been reported to be widely distributed in many organisms including bacteria, fungi, plants and a variety of invertebrates, metazoans, and vertebrates (reviewed by Pliyev and Gurvits 1999; Maruyama and Furutani 2000). They were shown to be involved in the folding of proteins as part of the cellular chaperone machinery, controlling

the assembly of protein complexes and membrane translocation. Twenty-three amino acid sequences of cyclophilins have been previously aligned for numerous angiosperms, indicating relatively high identity that rarely falls below 60% (Galat 1999).

Cyclophilins are known to be involved in different stress responses. Stimulated cyclophilin-encoding mRNA accumulation has been detected during different stress conditions. This fact makes the polymorphic cyclophilin EST locus potentially useful for the study of genetic variation, especially in the cases of the samples variable in stress response. Currently we have no detailed information on the expression of cyclophilin in Norway spruce.

After the amplification of the cyclophilin gene sequence, containing the whole coding and part of the 3`untranslated regions, by a fluorescent oligonucleotide primer pair and subsequent RsaI digestion, a co-dominant inherited polymorphic PCR marker was observed. This marker tends in population studies towards minor polymorphism as indicated by the presence of one high-frequency allele.

The sequence analysis of two most frequent alleles of locus PA0005 has shown that 21 bp insertion is localised within 3´ untranslated region and therefore does not influence directly on the structure of cyclophilin protein. Perry and Bousquet (1998) have developed the nine EST co-dominant markers of coding genes. Because these markers are essentially the result of polymorphisms located in transcribed but untranslated regions of arbitary genes, as in the case of newly developed cyclophilin marker, it was anticipated that their variation could be essentially neutral (Jaramillo-Correa et al., 2001).

4. 2 Genetic variation of *Picea abies* in Europe based on EST markers

Five additional EST markers (Schubert *et al*., 2001), together with newly developed marker for cyclophilin, were applied to study of genetic variation of *Picea abies* along its European natural range.

Picea abies being wind-pollinated, is characterised by high gene flow. In such organism, potentially neutral EST markers were expected to reflect a migration/drift equilibrium.

Despite for low average mean of genetic differentiation observed at analysed EST loci, an uneven distributions of allele and genotype frequencies as well as positive correlations between matrices of genetic and geographical distances were detected. The patterns of geographically structured differentiation across 19 test populations show a slight clinal

variation south-north through Europe. Four out of six analysed EST markers revealed different trends of differentiation.

The polymorphism at the newly developed EST locus PA0005 among the 19 European spruce populations described in this study seems to be highly informative, since the uneven frequency distribution associated with allele B shows a clear geographic pattern. This pattern was compared with different naturally existing gene pools of *P. abies,* which are known to result from different glacial refugee and post-glacial re-colonization routes. Italian and Swiss populations share the highest frequency in the case of allele B that can be a result of a random-drift mutation, which appears within the Alpine population group, in contrast to the putatively older, and therefore most frequent, allele A. This allele obviously dominates the whole European distribution range of *P. abies*. The populations of Alpine group (Italy, Switzerland, and France) were found at this locus as most diverse and showing highest level of observed heterozygosity in comparison with other populations. Once established, allele B might have been migrating from its origin of mutation to neighbouring populations by ongoing gene flow *via* pollen transfer and seed dispersal. Following the basic concepts of Wright (1943) and Slatkin (1993), gene flow, intensively studied in plant populations by different genetic markers as has been reviewed by Ouborg *et al.* (1999), will substantially decrease as a function of the geographic distance between populations. This fact leads to a higher genetic differentiation between distantly related populations in comparison to closely related populations. Results of Mantel's tests suggest that the Alpine population group and the Northern European-Russian population group are indeed genetically isolated by distance. The results of the Ewens-Watterson test, which has determined neutrality for marker PA0005, provide statistical evidence that the detected geographic differentiation results from the action of genetic drift and migration, and does not reflect adaptation processes following natural selection. However, large values of "Obs. F" detected in populations from Northern Europe suggest statistically still not significant but a tendency against heterozygosity in these populations in comparison with other. This fact indicates special dynamics of genetic variation for the populations mentioned above in contrast to other.

Uneven distribution of certain allele and genotype frequencies was detected also in the case of EST marker PA0055. This data are statistically supported using *Isolation-by-distance* test. The highest value of genetic differentiation in this study was detected for the marker PA0055 (12%) in comparison with a lower average of F_{ST} for all markers (6%).

This mean value is comparable with F_{ST} value described for the mtDNA markers of Norway spruce (Sperisen *et al*., 2001).

Furthermore, the markers PA0066 and PA0043 also reveal geographically structured differentiation among the European populations of Norway spruce. Outstanding positions of the Russian and Norwegian populations are evident from the analysis of allele and genotype frequencies distribution, UPGMA dendrogram clustering and the *Isolation-by-distance* test provided for the both loci.

However, UPGMA dendrograms based on the genetic distance between populations at loci PA0034 and PA0038 do not show definite cluster grouping of the samples originating from the same geographical region. Futhermore, there was not significant correlation detected between matrices of genetic and geographical distances for the both markers mentioned above.

It appears that the observed genetic structures reflect the effects of evolutionary factors imposed by historical events related to the last glaciation. There is an agreement between the current patterns of genetic differentiation and the recolonisation history suggested from pollen analyses (Schmidt-Vogt, 1977; Huntley and Birks, 1983). Spruce from the two central European refugia appear to have spread north and west, whereas the Russian relicts expanded westward to Scandinavia and toward the southwest to meet spruce from other refugia in the area of the present eastern Poland (Schmidt-Vogt, 1977). The results of present study also seem to indicate the migration processes from at least two glacial refugia, one located in western Russia and the other in the South-Eeastern European mountains. The observed genetic structures of Eastern European domain (Ukraine, Hungary, South Poland) indicate process of the gene flow between populations from this area and two neighbouring northern and southern domains. This data do not support the hypothesis about existence of additional refugia in the Carpathian mountains (Lang, 1994), and correspond to the suggestion of Huntley and Birkes (1983) that the Carpathian refugium was linked with the refugial area of Norway spruce in the Balkan.

The existence of the geographical trends of genetic variation over the European natural range of Norway spruce has already been described by several authors (Borgetti et al., 1988; Lagercrantz and Ryman, 1990; Vendramin *et al*., 2000; Bucci and Vendramin, 2001). The investigation on genetic variation of *Picea abies* in Europe based on cpSSR markers (Vendramin *et al*., 2000) identified the two major gene pools (Alpine-Center European and Sarmathic-Baltic), what is fully consistent with the evolutionary history of the species (Huntley and Birkes, 1983).

Two out of 19 analysed populations deviate outstandingly from detected genetic structure. Most east-ward located population No.15 from Siberia was found to be monomorphic at the loci PA0005 and PA0055. This fact can be explained by the highly probable hybridisation with *Picea obovata* in this geographical region, a species close relative to *Picea abies* and widely distributed in Eastern Siberia (Krutovskij *et al.*, 1989). Following study of the population No.15 using different genetic markers could be helpfull for the species determination of this sample.

The Romanian population No. 8 appears to be genetically similar on both levels of variation, within and among populations, with populations belonging to Russian-Northern-European pool. *Picea abies* was widely planted in Central Europe during last years (Bergmann *et al.,* 1989). This fact could explain such strong similarity of the Romanian population with northern populations detected on the genetic level.

The differences in the observed patterns of variation at six analysed EST loci are remarkable especially taking into account the addressing of these markers to the functionally active genes. Thus, the genetic structure observed using some EST markers could be explained not only by the occurrences of random-drift mutations and gene flow during postglacial recolonization of Norway spruce in Europe, but also as a possible result of selection. Genetic drift and migration should affect all loci similarly, whereas natural selection should affect some loci differently (Jaramillo-Correa *et al*., 2001). From other side, as predicted by the neutral theory (Kimura, 1983), the genome regions involving a change in a protein or a change in gene regulation are on average more conserved than non-functional sites (Ford, 2002). Despite the fact, that Ewens-Watterson test statistically quantified all analysed EST loci as neutral, the explanation for the observed differentiation patterns can be due to the “hitchhiking” of putatively neutral variation with sites that are under varying selection (Ford, 2002).

There is suggested that detected genetic structures largely reflect relatively recent historical events related to the last glaciations. But Norway spruce is still in a process of adaptation and differentiation. The patterns of detected geographical differentiation at several EST-marker loci are similar to that observed for isoenzymes and cpDNA markers. It suggests that the same evolutionary forces have acted upon different genetic markers.

4.3 Indicative potential of the newly developed cyclophilin gene marker for different environmental impacts on populations of Norway spruce

The codominant EST marker PA0005, reported in present study, address the functionally active gene of cyclophilin in Norway spruce. Evolutionary conserved cyclophilins (Galat, 1999) are known to be involved in different plant responses to stress, such as heat-shock treatment (Luan *et al*., 1994), low CO_2 conditions (Somanchi and Moroney, 1999), heavy metal pollution (Sturzenbaum *et al*., 1999), salt, cold, hormonal and osmotic stress (Marivet *et al*., 1994; Kullertz *et al*., 1999), wounding and fungal infection (Godoy *et al*, 2000). Also distinct plant cyclophilins were found to be developmentally regulated (Sholze *et al*., 1999). On the other hand, cyclophilin-encoding cDNA clone used for processing of codominant EST marker reported above, was isolated from elicitor induced (fungus *Rhizosphaera kalkhoffii*) cDNA library. All these facts together made a cyclophilin gene marker interesting for study of genetic variation in populations being under different climatic or stress conditions.

The present study of the current minor polymorphism at locus PA0005 revealed a geographically structured differentiation pattern across 19 test populations, showing a slight clinal variation south-north through Europe. Neutrality of this marker was indicated within all analysed populations by Ewens-Watterson test. Thus, this marker yields valuable information on the geographical origin of population samples, confirming the postulated re-imigration of Norway spruce.

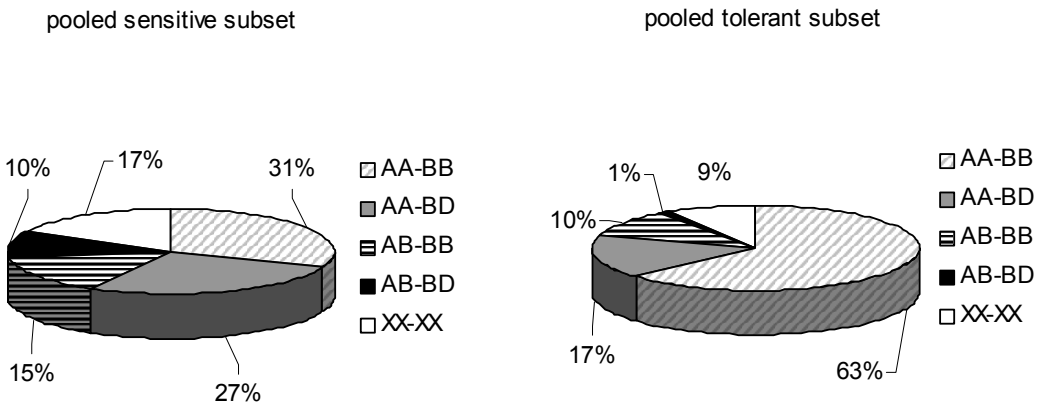
In parallel study (Riegel, 2001), the cyclophilin gene marker of *Picea abies* was found to be informative to face completely different objective (Bozhko *et al*., 2003). Analysis of tolerant and susceptible subsets of two adjacent Bavarian spruce populations affected by soil-born NaCl pollution, revealed that the cyclophilin marker locus also confirms biased genotype frequencies. Considering an unlinked PCR marker of a ribosomal protein-encoding EST clone (PA0066), deviations between pooled tolerant and pooled sensitive subsets were proven to be more significant for two-locus homozygous genotypes than for each locus alone (Bozhko *et al*., 2003) (*Fig. 4-1*). It seems likely that PA0005 marker gene indicate stress defence mechanisms in Norway spruce. Combination of both unlinked marker loci revealed a greater discriminative power than each locus alone, suggesting a complex multi-locus control for the avoidance mechanism studied as has been previously reported by Dvorack et al. (1992).

Considering, however, that not all the examined trees with the two-locus genotype AA-BB of the cyclophilin-ribosomal protein gene marker combination are exhibiting tolerance to excess salinity, it possible to conclude that not these marker loci themselves, but linked loci are candidate genomic regions for adaptation effects, and selection driven by the stress conditions tested. Once a genetic map becomes available for *P. abies,* which is offering molecular markers in close vicinity to PA0005 and PA0066, these unknown gene loci could be addressed in the future by fine-scale linkage analysis. The present study of recombination between both mentioned marker loci did not verify statistically significant deviations from random segregation, so that these loci are currently not considered to belong to the same linkage group. Moreover, the genetic control of salinity tolerance is known to be a complex and multi-gene trait (Dvorack *et al.,* 1992). For that reason, additional chromosomal units, which have not yet been detected by molecular markers, are expected to be involved in the genetic control of salt tolerance in Norway spruce (Bozhko et. al., 2003).

The highest values of both genotypes in the study of genetic variation along European natural range of Norway spruce were found within populations belonging to the Russian-Northern-European domain.

Taken into account the results of he Ewens-Watterson test and the sequencing data, confirming that the most frequent alleles A and B of locus PA0005 exclusively differ within th 3' non-coding regions, there is currently no evidence for a selection process on the cyclophilin gene locus surveyed. In general, selection is known to act on whole individuals rather than on genes, and acts on multiple traits simultaneously (reviewed by Namkoong et al. 2000). There was concluded that not these loci themselves, but tightly linked loci are candidate genomic regions for adaptation effects and selection driven by salt stress, taking into account, that not all the examined trees with informative two-locus genotype of the markers combination are exhibiting an advantageous phenotype.

Figure 4-1. Frequency distributions for two-locus genotypes of marker combination PA0005-PA0066 when analyzing a pooled sensitive subset and a pooled tolerant subset of two NaCl-affected spruce populations (symbol XX-XX indicates the remaining two-locus genotypes which are pooled because of their low individual frequencies <5%) (from Bozhko *et al*., 2003; modified from R. Riegel. 2001).



5 Concluding remarks

A codominantly inherited EST marker developed using cyclophilin-encoding cDNA clone of Norway spruce presents tends towards minor polymorphism as indicated by the presence of one high-frequency allele, in combination with six low-frequency alleles. Despite the fact that minor polymorphism are generally considered not to be very useful for studying the history of forest populations , the cyclophilin locus detected is suitable to verify simultaneously two completely different objectives. Firstly, the allele frequencies reveal slight geographical cline through Europe, resulting from a random–drift mutation that presumably appeared very recently within the Alpine population group. For that reason, this marker yields valuable information on the geographical origin of population samples, confirming the postulated re-immigration of Norway spruce in Europe. Secondly, the cyclophilin gene marker confirms together with the ribosomal protein gene marker PA0066 a genetic differentiation between tolerant and susceptible subsets of trees for two local salt-polluted Bavarian populations (see Discussion), therefore offering a novel tool to test the feasibility of marker-aided selection for improving NaCl-tolerance of *P.abies* in the future (Bozhko *et al*., 2003). Sequencing of the most frequent and informative alleles (A and B) of the locus PA0005 determines the location of observed polymorphism within 3` non-coding region. The detected mutation consequently does not affect the structure of the protein coded by cyclophilin gene. There is concluded that not EST locus PA0005 themselves but tightly linked locus is a candidate genomic region for adaptation effects. Integration of the cyclophilin marker into genetic map of Norway spruce in future will highlight this question.

Three out of five already existing co-dominant EST markers additionally used are shown to be indicative in order to verify geographically differentiated genetic structure between populations along the natural European range of Norway spruce. The uneven distribution of allele and genotype frequencies in space as well as significant correlations between genetic and geographical distances detected among Northern European-Russian and Southern European groups of populations at loci PA0055, PA0043 and PA0066, show at least two geographically isolated genetic zones of Norway spruce in Europe caused by postglacial migration from two supposed refugia. These markers together with geographically structured organelle-type DNA markers as well as isoenzymes provide valuable information on the geographic origin of population samples, which is important for the forest management to control planting of locally adapted provenances.

6 Summary

The sequence of the clone encoding for cyclophilin was picked up from elicitor induced cDNA library of *Picea abies* (L.) Karst. and used for primer construction in order to analyse EST polymorphism. The designed primer pair, targeting completely coding and partially 3` non-coding (potentially hypervariable) regions, amplified during genomic PCR polymorphic patterns. The PCR products following RsaI digestion were analysed using vertical gradient polyacrilamid gel electrophoresis (ALF-express system). In the processing the EST marker PA0005 was developed.

Based on segregation studies, codominant mode of inheritance was verified for polymorphism at the locus PA0005 (Mendelian segregation 1:1). Sequencing of two alleles observed during segregation analysis, identified the position of an insertion (21 bp) within 3´ non-coding region of the cyclophilin gene. The study of the current minor polymorphism at locus PA0005 revealed a geographically structured differentiation pattern across 19 natural European populations of Norway spruce with pronounced genetic distance existing among distantly related populations from the Alpine gene pool and the Russian-Northern European gene pool. Based on the frequency of alleles, isolation-by-distance analysis and the Ewens-Watterson test, it was concluded that a selectively neutral random-drift mutation, recently occurred within the Alpine population group, can be considered to be responsible for the genetic variation detected. In the study of 19 natural Norway spruce populations, five additional alleles represented by very low frequencies (not more than 3%) were detected at the locus PA0005. Additionally, the potential of newly developed cyclophilin marker to test the feasibility of marker-assisted selection of NaCl-tolerant spruce tree populations is discussed. There is suggested that PA0005 locus is linked to adaptive genomic region. Five additional already existing EST markers were applied to verify genetic variation within and among 19 populations representing European natural range of Norway spruce. All markers detected high levels of intrapopulational genetic variation. Three of them indicated statistically significant genetic differentiation between populations, showing a slight clinal variation south-north through Europe. Using Ewens–Watterson test, all five markers were shown not to deviate from selective neutrality in all populations tested. The observed patterns of genetic variation are suggested to be related with two gene pools (Russian-Northern and Alpine), considered to be the result of postglacial recolonization from putative refuges. Heterogeneous effects of genetic variation detected by different EST markers are discussed in connection with the neutrality theory and mechanisms of natural selection.

References

Altschul SF, Gish W, Miller W et al. (1990): Basic local alignment search tool. Journal of Molecular Biology 215: 403-410.

Bergmann F (1978): The allelic distribution at an acid phosphatase locus in Norway spruce (Picea abies) along similar climatic gradients. Theoretical and Applied Genetics 52: 57-64.

Bergmann F (1991): Causes and consequences of species-specific genetic variation patterns in European forest tree species: examples with Norway spruce and Silver Fir. In: Genetic *variation in European populations of forest trees* (eds. Müller-Starck G., Ziehe M): 192-204. J.D. Sauerländer`s Verlag, Frankfurt am Main.

Biswas C, Johri BM (1997): The gymnosperms. *New Delhi: Narosa Publishing House*.

Bond G (1993): Correlations between climate records from North Atlantic sediments and Greenland ice. *Nature*365: 143-147.

Bozhko M, Riegel R, Schubert R, Müller-Starck G (2003): A cyclophilin gene marker confirming geographic differentiation of Norway spruce populations and indicating viability selection on excess soil-born salinity. *Molecular Ecology* 12:3147-3155.

Binelli G, Bucci G (1994): A genetic linkage map of *Picea abies* Karst., based on RAPD markers, as a tool in population genetics. *Theoretical and Applied Genetics* 88: 283-288.

Bucci G, Kubisiak TL, Nance WL *et al.* (1997): A population 'consensus', partial linkage map of *Picea abies* Karst., based on RAPD markers, as a tool in population genetics. *Theoretical and Applied Genetics 95*: 643-654.

Bucci G, Vendramin GG (2000): Delineation of genetic zones in the European Norway sdruce natural range: preliminary evidence. *Molecular Ecology* 9: 923-934.

Bradshaw RHW, Holmqvist BH, Cowling SA, Sykes MT (2000): The effects of climate change on the distribution and management of *Picea abies* in southern Scandinavia. *Canadian Journal of Forest Research* 30: 1992-1998.

Bucci G, Vendramin GG (2000): Delineation of genetic zones in the European Norway spruce natural range: preliminary evidence. *Molecular Ecology* 9: 923-934.

Burga CA (1988): Swiss vegetation history during the last 18000 years. *New Phytology* 110: 581-602.

Collignon AM, Favre JM (2000): Contribution to the postglacial history at the western margin of *Picea abies'* natural area using RAPD markers. *Annals of Botany* 85: 713-722.

Comes HP, Kadereit JW (1998): The effect of Quaternary climatic changes on plant distribution and evolution. *Trends in plant science* Vol. 3, No.11: 432-438.

Dean C, Schmidt R (1995): Plant genomes: a current molecular description. *Annual Reviews: Plant Physiology and Plant Molecular Biology* 46: 395-418

Dvorack J, Epstein E, Galvez A *et al.* (1992): Genetic basis of plant tolerance of soil toxicity. In: *Proceedings of a symposium on plant breeding in the 1990s* (eds. Stalker HT, Murphy JP): 201-217. CAB International, Wallingford, England.

Ellis R (1990): Climatic change and plant genetic resources. *Biologist* 37: 4-6.

Elsik CG, Williams CG (2001): Low-copy microsatellite recovery from a conifer genome. *Theoretical and Applied Genetics* 103: 1189-1195.

Finkeldey R, Mátyás G (2000): Assessment of population history and adaptive potential by means of gene markers. In: *Forest Genetics and Sustainability* (eds. Mátyás C): 49-58. Kluwer Academic Publishers, Netherlands.

Friesen N, Brandes A, Heslop-Harrison JS (2001): Diversity, origin, and distribution of retrotransposons (*gypsy* and *copia*) in conifers. *Molecular Biology and Evolution* 18(7): 1176-1188.

Ford MJ Applications of selective neutrality tests to molecular ecology. *Molecular Ecology* 11, 1245-1262.

Galat A (1999): Variations of sequences and amino acid compositions of proteins that sustain their biological functions: An analysis of the cyclophilin family of proteins. *Archives of Biochemistry and Biophysics* 137: 149-162.

Galliano H, Cabané M, Eckerskorn C *et al.* (1993): Molecular cloning, sequence analysis and elicitor-/ozone-induced accumulation of cinnamyl alcohol dehydrogenase from Norway spruce (*Picea abies* L.). *Plant Molecular Biology* 23: 145-156.

Godoy AV, Lazzaro AS, Casalongué CA *et al.* (2000): Expression of a *Solanum tuberosum* cyclophilin gene is regulated by fungal infection and abiotic stress conditions. *Plant Science* 152: 123-134.

Gregorius HR (1991): Gene conservation and the preservation of adaptability. S. 31-47 In: Seitz A, Loeschke V (eds.) 1991.

Gregorius HR (1974): Genetischer Abstand zwischen Populationen. Konzeption der genetischen Abstandsmessung. *Silvae Genetica* 23: 22-27.

Gregorius HR (1989): The attribution of phenotypic variation to genetic or environmental variation in ecological studies. In: *Genetic aspects of air pollutants to forest tree populations* (eds. Scholz F, Gregorius HR, Rudin D): 3-15. Springer Verlag, Berlin - Heidelberg.

Gugerli F, Senn J, Anzidei M *et al.* (2001): Chloroplast microsatellites and mitochondrial nad1 intron 2 sequences indicate congruent phylogenetic relationships among Swiss stone pine (*Pinus cembra*), Siberian stone pine (*Pinus sibirica*), and Siberian dwarf pine (*Pinus pumila*). *Molecular Ecology* 10: 1489-1497.

Hamrick JL, Godt MJW (1989): Allozyme diversity in plant species. In: Brown AHD, Clegg MT, Kahler AL and Weir BS (eds.) Plant population genetic, breeding and germoplasm resources. Sinauer Associates, Sunderland: 43-63.

Hansen JK, Saxe H, Ræbild A *et al.* (1998): Decline and physiological response to foliar-deposited salt in Norway spruce genotypes: a comparative analysis. *Canadian Journal of Forest Research* 28: 1879-1889.

Hare PD, Cress WA, van Staden J (1999): Proline synthesis and degradation: a model system for elucidating stress-related signal transduction. Journal of experimental Botany 50(333): 413-434.

Hattemer HH, Gregorius HR (1990): Is gene conservation under global climatic change meaningful? In: *Climatic change and genetic resources* (eds. Jackson MT, Ford-Lloyd BV, Parry MJ): 155-166. Belhaven press, London.

Hewitt GM (1996): Some genetic consequences of ice ages, and their tole in divergence and speciation. *Biol. J. linn. Soc*. 56: 247-276.

Hiesel R, Combettes B, Brennicke A (1994): Evidence for RNA editing in mitochondria of all major groups of land plants exept the Bryophyta. Proceedings of the National Academy of Sciences of the United States of America 91: 629-633.

Hipkins VD, Krutovski KV, Strauss SH (1994): Organelle genomes in conifers: structure, evolution, and diversity. *Forest Genetics* 1(4):179-189

Huntley B, Birks HJB (1983): *An atlas of past and present pollen maps for Europe: 0-13000 years ago.* Cambrige University press.

Jaramillo-Correa JP, Beaulieu J, Bousquet J (2001): Contrasting evolutionary forces driving population structure at expressed sequence tag polymorphisms, allozymes and quantitative traits in white spruce. *Molecular Ecology* 10: 2729-2740.

Köstler J.N. (1956): Waldbauliche Beobachtungen an Wurzelstöcken sturmgeworfener Nadelbäume. *Forstwissenschaftliches Centralblatt*: 63, S. 65-91.

Kral F (1979): Spät-und postglaziale waldgeschichte der Alpen auf grund der Bisherigen pollenanalysen .*Veröffentlichung des Institutes für Waldbau an der Universität für Bodenkultur in Wien*.

Krutovski KV, Bergmann F (1995): Introgressive hybridization and phylogenetic relationships between Norway, Picea abies (L.) Karst., and Siberian, P. obovata Ledeb., spruce species studied by isozyme loci. *Heredity* 74: 464-480.

Kullertz G, Liebau A, Rucknagel P *et al.* (1999): Stress-induced expression of cyclophilins in proembryonic masses of *Digitalis lanata* does not protect against freezing/thawing stress. *Planta* 208: 599-605.

Lagercrantz U, Ryman N (1990): Genetic structure of Norway spruce (*Picea abies*): concordance of morphological and allozymic variation. *Evolution* 44: 38-53.

Langebartels C, Anegg S, Hahn K *et al.* (2001): Ozone and UV-B responses of trees and the question of forest sustainability. In: *Trends in European Forest Tree Physiology Research* (eds. Huttunen S, Heikkilä H, Bucher J *et al.*): 157-166. Kluwer Academic Publishers, Dordrecht - Boston - London.

Ligholm J, Gustafsson P (1991): Homologues of the green algal *gid*A gene and the liverwort *frx* C gene are present on the chloroplast genomes of conifers. *Plant Molecular Biology* 17: 787-798.

Longauer R, Gomory D, Paule L *et al.* (2001): Selection effects of air pollution on gene pools of Norway spruce, European silver fir and European beech. *Environmental pollution* 115: 405-411.

Luan S, Lane WS, Schreiber SL (1994): pCyP B: a chloroplast-localised, heat-shock-responsive cyclophilin from faba bean. *Plant Cell* 6: 885-892.

Manly BFJ (1985): *The Statistics of Natural Selection.* Chapman & Hall, London.

Mantel JB (1967): The detection of disease clustering and generalised regression approach. *Cancer Research* 27: 209-220.

Marivet J, Margis-Pinheiro, Frendo P *et al.* (1994): Bean cyclophilin gene expression during plant development and stress conditions. *Plant Molecular Biology* 26: 1181-1189.

Markussen T, Tusch A, Stephan R *et al.* (1999): QTL mapping of selected wood properties in Norway spruce *Picea abies (*L.) Karst. In: Book of abstracts, International Congress: Application of biotechnology to forest genetics, Vitoria-Gasteiz, Spain, Sept. 22-25: 157-159.

Maruyama T, Furutani M (2000): Archaeal peptidyl prolyl cis-trans isomerases (PPIases). *Frontiers in Bioscience* 5: 821-836.

McVean G (2002): Natural selection. http://www.stats.ox.ac.uk/~mcvean/L4notes.pdf

Meza-Zepeda L, Baudo MM, Palva ET, Heino P (1998): Isolation and characterization of a cDNA corresponding to a stress-activated cyclophilin gene in *Solanum commersonii*. Journal of Experimental Botany 49 (325): 1451-1452.

Murray BG (1998): Nuclear DNA amounts in gymnosperms. *Annals of Botany* 82: 3-15.

Murray BG, Friesen N, Heslop-Harrison JS (2002): Molecular Cytogenetic Analysis of *Podocarpus* and comparison with other gymnosperm species. *Annals of Botany* 89:483-489.

Müller-Starck G, Ziehe M (1991): *Genetic variation in European populations of forest trees*. JD Sauerlaender's Verlag, Frankfurt am Main.

Müller-Starck G (1995): Genetic variation in high elevated populations of Norway spruce (Pica abies (L.) Karst.) in Switzerland. *Silvae Genetica* 44: 5-6.

Nei M (1972): Genetic distance between populations. *American Naturalist* 106: 283-292.

Ostoa-Saloma P, Carrero JC, Petrossian P *et al.* (2000): Cloning, characterisation and functional expression of a cyclophilin of *Entamoeba histolytica*. *Molecular and Biochemical Parasitology* 107: 219-225.

Ouborg NJ, Piquot Y, Van Groenendael JM (1999): Population genetics, molecular markers and the study of dispersal in plants. *Journal of Ecology* 87: 551-568.

Paglia GP, Olivieri AM, Morgante M (1998): Towards second-generation STS (sequence-tagged sites) linkage maps in conifers: a genetic map of Norway spruce (*Picea abies* K.). *Molecular and General Genetics* 258: 466-478.

Perry DJ, Bousquet J (1998): Sequence-Tagged-Site (STS) Markers of Arbitrary Genes: Development, Characterization and Analysis of Linkage in Black Spruce. *Genome* 149: 1089-1098.

Perry DJ, Isabel N, Bousquet J (1999): Sequence-tagged-site (STS) markers of arbitrary genes: the amount and nature of variation revealed in Norway spruce. *Heredity* 83: 239-248.

Petit RJ, Demesure B, Pineau E (1996): Chloroplast DNA differentiation at a local and continental scale in European oak species: the importance of historical factors. In: *Inter- and Intra-specific Variation in European Oaks: Evolutionary Implications and Practical Consequences* (eds. Kremer A, Muhs HJ):145-164. Official Publications of the European Communities, Brussels.

Petit RJ, Bialozyt R, Brewer S, Cheddadi R, Comps B (2001): From spatial patterns of genetic diversity to postglacial processes in forest trees. In: *Integrating Ecology and Evolution in a spatial context* (eds. Silvertown J, Antonovics J):295-318. Blackwell.

Petit RJ, Itziar Aguinagalde, Jacques-Louis de Beaulieu, Christiane Bittkau, Simon Brewer, Rachid Cheddadi, Richard Ennos, Silvia Fineschi, Delphine Grivet, Martin Lascoux, Aparajita Mohanty, Gerhard Müller-Starck, Brigitte Demesure-Musch, Anna Palmé, Juan Pedro Martín, Sarah Rendell, and Giovanni G. Vendramin (2003): Glacial Refugia: Hotspots But Not Melting Pots of Genetic Diversity. *Sience* Jun 6: 1563-1565.

Pfeiffer A, Oliveri AM, Morgante M (1997): Identification and characterization of microsatellites in Norway spruce (*Picea abies* K.). *Genome* 40: 411-419.

Pilyev BK, Gurvits BY (1999): Peptidyl-prolyl cis-trans isomerases: structure and functions. *Biochemistry (Moscow)* 64: 738-751.

Purugganan MD (2000): The molecular population genetics of regulatory genes. *Molecular Ecology* 9: 1451-1461.

Pravdin (1975): *Picea abies* (L.) Karst and *Picea obovata* Ledeb. In the USSR. *Nauka Publ.*. Moscow.

Raubeson LA, Jansen RK (1992): A rare chloroplast DNA structural mutation is shared by all conifers. *Biochemical and Systematic Ecology* 20: 17-24.

Riddle BR (1996): The molecular phylogeographic bridge between deep and shallow history in continental biotas. *Trends Ecol.Evol.* 11: 207-211.

Riegel R (2001): *Entwicklung molekulargenetischer Marker bei der Fichte [Picea abies (L.) Karst.] für genetische Erhebungen in umweltbelasteten Populationen*. PhD thesis, Technical University of Munich.

Riegel R, Schubert R, Müller-Starck G (2001): Genetic variation in two heavily polluted stands of Norway spruce (*Picea abies* [L.] Karst.) as indicated by nuclear and organelle DNA markers. In: *Genetic Response of Forest Systems to Changing Environmental Conditions* (eds. Müller-Starck G, Schubert R.): 21-34. Kluwer Academic Publishers, Dordrecht -Boston - London.

Safford LO (1974): *Picea*, spruce. In: *Seeds of woody plants of the United States* (Schopmeyer CS, tech. Coord) Forest Service: 587B597. Washington.

Saez-Vasquez J, Raynal M, Meza-Basso L (1993): Two related, low-temperature-induced genes from *Brassica napus* are homologous to the human tumour bbc1 (Breast basic conserved) gene. *Plant Molecular Biology* 23: 1211-1221.

Schantz v M , Juvonen S (1966): Chemotaxonomische Untersuchungen in der Gattung *Picea*. *Annales Botanici Fennici* 73: 3-51.

Scheepers D, Eloy MC, Briquet M (1997): Use of RAPD patterns for clone verification and in studying provenance relationships in Norway spruce (*Picea abies*). *Theoretical and Applied Genetics* 94: 480-485.

Schmidt A, Doudrick RL, Heslop-Harrison JS, Schmidt T (2000): The contribution of short repeats of low sequence complexity to large conifer genomes. *Theoretical and Applied Genetics* 101: 7-14

Schmidt-Vogt H (1986): *Die Fichte, Band II/1, Wachstum-Züchtung-Boden-Umwelt-Holz.* Paul Paray Verlag Hamburg – Berlin.

Schmidt-Vogt H (1987): *Die Fichte, Band I, Taxonomie-Verbreitung-Morphologie-Ökologie-Waldgesellschaften, 2nd edn.* Paul Paray Verlag, Hamburg - Berlin.

Schmidt-Vogt H (1989): Die Fichte, Band II/2, *Krankheiten-Schäden-Fichtensterben*. Paul Paray Verlag Hamburg – Berlin.

Scholz F, Gregorius HR, Rudin D (1989): *Genetic effects of air pollutiants in forest tree populations*. Springer, Berlin, Heidelberg.

Scholze C, Peterson A, Diettrich B *et al.* (1999): Cyclophilin isoforms from *Digitalis lanata*. Sequences and expression during embryogenesis and stress. *Journal of Plant Physiology* 155: 212-219.

Schömig M (1988): *Das Wachstum von Fichten-Kiefern-Beständen unter dem Einfluss von im Grundwasser angereicherten Strassenauftausalzen im Forstamt Gunzenhausen.* Master thesis, Ludwig Maximilians University of Munich.

Schubert R, Müller-Starck G, Riegel R (2001): Development of EST-PCR markers and monitoring their intrapopulational genetic variation in *Picea abies* (L.) Karst. *Theoretical and Applied Genetics* 103: 1223-1231.

Schütt P, Cowling EB (1985): Waldsterben, a general decline of forest in central Europe: symptoms, development and possible causes. *Plant Disease* 69: 548-558.

Scotti I, Troggio M, Soranzo N, Vendramin GG, Bucci G (1998): A new set of PCR-based, locus-specific markers for *Picea abies* (L.) Karst.. *Molecular Ecology 7*: 783-792.

Scotti I, Vendramin GG, Matteotti LS, Scarponi C, Sari-Gorla M, Binelli G (2000): Postglacial recolonization routes for *Picea abies* K. in Italy as suggested by the analysis of sequence-characterized amplified (SCAR) markers. *Molecular Ecology* 9: 699-708.

Scotti I, Magni F, Fink R *et al.* (2000) Microsatellite repeates are not randomly distributed within Norway spruce (*Picea abies* K.) expressed sequences. *Genome* 43: 41-46.

Scotti I, Paglia GP, Magni F, Morgante M (2002): Efficient development of dinucleotide microsatellite markers in Norway sdruce (*Picea abies* Karst.) through dot-blot selection. *Theoretical and Applied Genetics* 104: 1035-1041

Skrøppa T, Johnsen Ø, Dietrichson J (1995): Effects of natural and artificial selection in Norway spruce. In: *Population genetics and genetic conservation of forest trees* (eds. Baradat Ph., Adams WT, Müller-Starck G): 453-458.SPB Academic Publishing, Amsterdam, The Netherlands.

Skrøppa T, Johnsen Ø (2000): Patterns of adaptive genetic variation in forest tree species; the reproductive enviroment as an evolutionary force in *Picea abies*. In: *Forest Genetics and Sustainability* (eds. Mátyás C): 49-58. Kluwer Academic Publishers, Netherlands.

Slatkin M (1993): Isolation by distance in equilibrium and non-equilibrium populations. *Evolution* 47: 264-279.

Slatkin M, Maddison W (1990): Detecting isolation by distance using phylogenies of genes. *Genetics* 126: 249-260.

Somanchi A, Moroney JV (1999): As *Chlamydomonas reinhardtii* acclimates to low-CO_2 conditions there is an increase in cyclophilin expression. *Plant Molecular Biology* 40: 1055-1062.

Sperisen C, Büchler U, Gugerli F, Mátyás G, Geburek T (2001): Tandem repeats in plant mitochondrial genomes: application to the analysis of population differentiation in the conifer Norway spruce. *Molecular Ecology* **10**: 257-263.

Strauss SH, Palmer JD, Howe GT, Doerksen AH (1988): Chloroplast genomes of two conifers lack a large inverted repeat and are extensively rearranged. *Proceedings of the National Academy of Sciences of the United States of America* 85: 3898-3909.

Storz JF Contrasting patterns of divergence in quantitative traits and neutral DNA markers: analysis of clinal variation. *Molecular Ecolgy* 11: 2537-2551.

Sturzenbaum SR, Morgan AJ, Kille P (1999): Characterisation and quantification of earthworm cyclophilins: identification of invariant and heavy metal responsive isoforms. *Biochimica et Biophysica Acta,* 1489: 467-473.

Sunnucks P (2000): Efficient genetic markers for population biology. *Tree,* 15: 199-203.

Sutton BCS, Flanagan DJ, Gawley JR, Newton CH, Leister DT, El-Kassaby YA (1991): Inheritance of chloroplast and mitochondrial DNA in *Picea* and composition of hybrids from introgression zones. *Theoretical and Applied genetics* 82: 242-248.

Taberlet P, Fumagalli L, Wust-Saucy AG *et al.* (1998): Comparative phylogeography and postglacial colonization routes in Europe. *Molecular Ecology* 7: 453-464.

Vendramin GG, Anzidei M, Madaghiele A, Sperisen C, Bucci G (1999): Chloroplast microsatellite analysis reveals the presence of population subdivision in Norway spruce (*Picea abies* K.). *Genome* 42: 1-11.

Webb T, Bartlein PJ (1992): Global changes during the last 3 million years: climatic controls and biotic response. *Annu. Rev. Ecol. Syst.*, 23: 141-173.

Wentzel KF (1974): Salz-Spritzwaserschäden von den Autobahnen in die Tiefe der Waldbestände. *Journal of Forest Pathology* 4: 45-46.

Wright S (1943): Isolation by distance. *Genetics* 28: 114-138.

Wright S (1978): *Evolution and the genetics of populations*, Vol. 4: Variability within and among natural populations. University of Chicago press.

Ye XY, Ng TB (2000): Mungin, a novel cyclophilin-like antifungal protein from the mung bean. *Biochemical and Biophysical Research communications* 273: 111-1115.

Appendix

Persons and institutions provided spruce material

Population No.	Country (origin of population)	Provider
1	Germany	**R. Riegel,** Technical University Munich
2	Italy	**C. Sperisen**, Swiss Federal Institute for Forest, Snow and Landscape Research, Birmensdorf
3	Italy	**C. Sperisen,** Swiss Federal Institute for Forest, Snow and Landscape Research, Birmensdorf
4	Switzerland	**C. Sperisen**, Swiss Federal Institute for Forest, Snow and Landscape Research, Birmensdorf
5	Switzerland	**C. Sperisen**, Swiss Federal Institute for Forest, Snow and Landscape Research, Birmensdorf
6	Croatia	**M. Konnert**, Bayerisches Amt für Forstliche Saat- und Pflanzenzucht, Teisendorf
7	Slovenia	**M. Konnert**, Bayerisches Amt für Forstliche Saat- und Pflanzenzucht, Teisendorf
8	Hungary	**C. Sperisen**, Swiss Federal Institute for Forest, Snow and Landscape Research, Birmensdorf
9	Romania	**C. Sperisen**, Swiss Federal Institute for Forest, Snow and Landscape Research, Birmensdorf
10	Poland	**C. Sperisen**, Swiss Federal Institute for Forest, Snow and Landscape Research, Birmensdorf
11	Ukraine	**I. Shvadchak**, Ukrainian State University of Forestry, Lviv
12	Poland	**C. Sperisen**, Swiss Federal Institute for Forest Snow and Landscape Research, Birmensdorf
13	Poland	**P. Robakowski**, Polish Agricultural University, Poznan
14	Russia	**C. Sperisen and A. Rigling,** Swiss Federal Institute for Forest, Snow and Landscape Research, Birmensdorf
15	Russia	**C. Sperisen and A. Rigling**, Swiss Federal Institute for Forest, Snow and Landscape Research, Birmensdorf

Population No.	Country	Provider
16	Norway	**T. Skrøppa**, Norwegian Forest Research Institute, Ås
17	Norway	**T. Skrøppa**, Norwegian Forest Research Institute, Ås
18	Italy	**F. Bergmann** Department of Forest genetics, University Göttingen
19	France	**F. Bergmann** Department of Forest genetics, University Göttingen