UNIVERSITÀ DELLA CALABRIA

CRA SAM

UNIVERSITÀ DELLA CALABRIA

FACOLTÀ DI SCIENZE MATEMATICHE, FISICHE E NATURALI

DOTTORATO DI RICERCA IN BIOLOGIA VEGETALE

SCUOLA DI DOTTORATO "LIFE SCIENCES"

XXV CICLO

SETTORE DISCIPLINARE: 05/E2

GENETIC DIVERSITY ASSESSMENT IN *PINUS LARICIO* POIRET POPULATIONS USING MICROSATELLITES ANALYSIS AND INFERENCES ON POPULATION HYSTORY

Coordinatore:

Prof. Aldo Musacchio

Dottorando:

Dott. Savino Bonavita

Supervisore UNICAL:

Dott.ssa Teresa M.R. Regina

Supervisore CRA-SAM

Dott. Vincenzo Bernardini

Anno Accademico 2011-2012

TO MY LOVE MARY AND MY FAMILY

INDEX

Summary

Forests are complex and dynamic ecosystems characterized by trees with remarkable longevity and reproductive forms, particularly, cross-pollination which tend to increase their degree of genetic diversity.

If on the one hand, cross-pollination between different individuals continuously shuffles the genetic material ensuring heterogeneity, on the other hand, an effective pollen and seeds dispersal ensures a reliable gene flow between individuals and, therefore, high levels of intra-specific variability. Therefore, more genetically different individuals are, greater is their ability to adapt to changed environmental conditions.

Pinus laricio Poiret, usually considered as the most divergent and genetically original subspecies of European black pine (*Pinus nigra* Arnold), is the most widespread conifer occurring in Calabria, Sicily (Etna Mount) and Corsica. In Calabria, it grows on the Aspromonte mountain and mainly on the Sila plateau, where laricio pine forests cover more than 40,000 ha and characterize the landscape from 900 m up to 1,700 m above sea level. Thermophilic, xerophilous and heliophilous species, *Pinus laricio* can reach large sizes and 350 years of age, as documented for the Fallistro's Giants Biogenetical Reserve, within the Sila National Park.

To the best of our knowledge, until now no studies have been conducted on the genetic diversity of *Pinus laricio* forests in their natural range of distribution. Furthermore, an in-depth investigation on the within- and among-population genetic differentiation is greatly needed to preserve *Pinus laricio* diversity, but also to establishing appropriate strategy of management and conservation of this specie.

In this thesis, genetic polymorphism among geographically distant laricio pine natural populations from Sila, Aspromonte, Etna, and Corsica National Parks was detected using chloroplast and nuclear SSR markers.

Both types of markers revealed that the higher diversity was found mainly within populations, while there were low levels of differentiation among populations, very likely associated with extensive gene flow and strong anthropogenic influence. However, a geographical discontinuity was identified, clearly indicating genetic subdivision of the investigated laricio pine populations at both inter- and intra-population level.

All populations within the Sila area were found differentiated from the rest, particularly the Fallistro population that appeared the most genetically distinct.

Results issued from this study shed light on the gene pool and evolutionary history of *Pinus laricio* populations providing a genetic perspective for exploitation and conservation of this not yet sufficiently explored resources in forestry.

1. Introduction

1.1 *Biodiversity and forests*

The diversity of life forms or biodiversity is the basis of all biological studies. The amplify in the rate of extinction and the need to intensify the biological monitoring make it increasingly urgent to identify and characterize all living beings.

To date, the most complete and well-known definition of biodiversity indicates the variability at genetic, species and community levels of biological organization.

Even though genetic diversity is at the lowest hierarchy, it has a remarkable impact on the higher levels of biodiversity. Genetic diversity analysis consists in the appreciation of variations and/or similarities found in the primary sequence of the nucleic acids (DNA and RNA) of individuals of the same species.

Individuals can be distinguished by a different assortment for allelic *locus* that, together with the allelic distribution at the group level, determines the degree of genetic diversity among and within populations.

At specific level, biodiversity refers to the genetic differences that exist between species: species with high levels of biodiversity will be made up of individuals with different genetic information in a more or less wide. Individuals belonging to populations and species with reduced levels of biodiversity, on the contrary, tend to be similar to each other and, therefore, to react in a substantially uniform manner to environmental stress.

The biodiversity, in this case, is to be closely connected with the potential for adaptation of populations (climate, land, pollution or other dangers related to human

activity) and more individuals are different and more likely that at least a part thereof being able to tolerate changes that occur in the environmental conditions.

The appearance of a new pest, or modification of the climate, could have disastrous effects if all individuals of populations prove homogenous and lack of genetic mechanisms that confer resistance or tolerance to adversity.

Finally, at ecosystematic level, biodiversity generally refers to the number of species present in a given site or habitat: more this is high, more stable is the ecosystem as a whole. In recent years, there has been increasing concern that modern agricultural production practices are contributing to the decline of both species diversity and genetic diversity. Furthermore, it is known by experience that altering even a seemingly small component of an ecosystem can result in very dramatic and undesirable results.

Forests are complex ecosystems that cover 30 percent of the global land area, providing habitat for countless terrestrial species.

They are vital for livelihoods as well as economic and social development, providing food, raw materials for shelter, energy and manufacturing. Forests are, also, critical for environmental protection and conservation of natural resources and contain more carbon than the atmosphere. However, with climate change, forests, with their dual roles as both producers and absorbers of carbon, take on a new importance.

The relationship between genetic variability and adaptability plays a particular importance in the case of forests and forest species. These, in fact, have very long life cycles (even higher than the century) and, in a period of time so extended, is much more likely to attend to environmental variations, where the population must be able to respond in an appropriate manner (Hamrick, 2004).

The immobility of the plants makes it possible, moreover, that these are not able to escape to any environmental stress, but rather are particularly exposed to all their effects.

Genetic diversity provides the fundamental basis for evolution of forest tree species enabling them to adapt to changing and adverse conditions for thousands of years.

Thus, forest genetic diversity has resulted in a unique and irreplaceable portfolio of tree genetic resources, the vast majority of which remains unknown.

Until recently, studies of forest tree genetic resources have concentrated on domesticating those few deemed most applicable for wood fiber and fuel production from plantations and agroforestry systems.

The progressive deterioration of the natural environment, particularly in forest ecosystems, constitutes a severe threat to their future survival. The indiscriminate action of man in the area (over-exploitation, inappropriate forestry practices, forest fires, indiscriminate urbanization, various pollution) is, in fact, not only reducing the number of individuals who survive, but also their genetic diversity.

The main objective of conserving a species is to allow survival in its natural area of growth (*in situ* conservation). To achieve this, it is necessary to safeguard most of the genetic heritage of a species, protecting primarily the autochthonous populations better adapted to their habitat of origin and, for this reservoir of a gene pool allows the species survival. Any loss of individuals lead to the irreversible disappearance of some key genes in the constitution of the genetic variability of a species.

The preservation of great potential of forest genetic resources also requires direct human intervention targeted the reconstruction of conditions suitable for the conservation of the species of interest. In particular, attention is paid to maintaining protected areas

(National and Regional Parks, Natural Reserves), as well as employing sustainable forestry practices, where the species can grow and reproduce naturally.

1.2 *Biodiversity and molecular markers*

To aim safeguarding biodiversity of forest ecosystems, it appears essential to evaluate components and analyze the processes that influence, and/or the consequences of its possible reduction.

Therefore, it become necessary to complete information on the distribution of genetic variability within and among populations of forest species, on ecological characteristics and on all the biological variables that influence the distribution of species object of study.

The changes that a population can undergo are numerous and may include its size and diversity of the individuals who compose it. The first case, purely quantitative, consists in demographic oscillations, while the second hear each of the components of biodiversity (genes, individuals, species, ecosystems).

The estimate of biodiversity within and between populations must be supported with appropriate markers (Agarwall et al., 2008)

As a first step, the markers used to detect and analyze genetic diversity were morpho-physiological (leaf shape, flower color, structure of the pollen) or phenology (vegetative growth, flowering, ripening of seeds, fallen leaves).

These markers are, however, some serious drawbacks: first, it is often not known the genetic basis of control (hereditary transmission), for which it is difficult to understand the exact correspondence between the observed variability and the actual genetic diversity.

In addition, the phenotypic manifestation, what we observe and we can measure, of these characters is strongly influenced by the environment, subjective interpretation of the observer, is also susceptible to human error, and this, again, prevents a reliable estimate of effective diversity between individuals.

Much more reliable are biochemical markers (Agarwall et al., 2008), which analyze products of the metabolism in plants, such as isoenzymes and terpenes.

Isoenzymes have several advantages that make them particularly suitable for studies of genetic variability: are ontogenetically stable and their phenotypic expression is not subject to environmental influences.

Furthermore, their genetic basis is usually very simple, type Mendelian monoallelica, while the expression codominant allowing an immediate distinction between heterozygotes and homozygotes.

However, the isoenzyme technique also presents some limits. Among these, the fact that the genes be studied represent a small and not always random sample of those present in the whole genome of the species: the extent of genetic variability present in the population can, therefore, be underestimated or overestimated.

Furthermore, this approach is not always capable of detecting all changes that occur at the DNA level.

The limitations of phenotype and isoenzyme-based genetic markers led to the development of more general and useful direct DNA-based markers that became known as molecular markers.

A molecular marker is defined as a particular segment of DNA that is representative of the differences at the genome level.

Actually, basic marker techniques can be classified into two categories:

- ✓ non-PCR-based techniques or hybridization based techniques
- ✓ PCR-based techniques.

Non-PCR based techniques are principally constituted by RFLP (Restriction Fragment Length Polymorphism) and VNTR (Variable Number of Tandem Repeats).

The principally PCR-based techniques are represented by RAPD (Random Amplified Polymorphic DNA), AFLP (Amplified Fragment Length Polymorphism) and SSR (Simple Sequences Repeats).

Techniques differ from each other with respect to important features such as genomic abundance, level of polymorphism detected, locus specificity, reproducibility, technical requirements and cost. Depending on the need, modifications in the techniques have been made, leading to a second generation of advanced molecular markers (Tab. 1).

Particularly, the SSR markers are constituted by repetitions of a short basic motif of length generally between 1-6 base pairs (bp) which occur as interspersed repetitive elements in all eukaryotic genomes (Tautz and Renz, 1984). The technique is based on the observation that the sequences flanking a given microsatellite locus in the genome are conserved within species, between species within a genus and, more rarely, even among related genera.

Variation in the number of tandemly repeated units is mainly due to strand slippage during DNA replication where the repeats allow matching via excision or addition of

repeats (Schlotterer and Tautz, 1992). As slippage in replication is more likely than point mutations, microsatellite loci tend to be hypervariable.

Today, this technique is the most useful in the study of population genetics because its use in the identification of the presence of genetic variability allows to highlight a high polymorphism, ensures a high repeatability of results and provides the opportunity to perform Multiplexing PCR experiments, that is the ability to use two or more primers simultaneously when the amplification products differ in size so as not to overlap with one another in the same path electrophoretic (Agarwall et al., 2008).

In order to estimate the genetic variability of a plant population, it is necessary to investigate the genetic polymorphisms using a pool of molecular markers selected through precise experimental criteria. The higher will, in fact, the number of markers investigated, the greater the reliability of the information obtained.

In this context, it is noteworthy that all the three cellular genomes, nuclear, plastidial and mitochondrial, characterizing the plant systems, can be used for population genetic studies or intraspecific studies in general.

The organellar genomes are frequently used because chloroplasts and mitochondria are mostly uniparentally inherited in seed plants and, thus, have some great advantages over biparentally inherited nuclear markers.

The main advantage is that there is typically only one allele per cell and organism, and consequently no recombination between two alleles occurs. Due to different dispersal distances, biparentally, maternally, and paternally inherited genomes also exhibit strong differences in genetic differentiation between populations. Especially, maternally inherited markers mostly show a much higher population subdivision (Petit et al., 2005).

The mitochondrial genome of plants is considerably larger and more complex than that of animals, and pronounced differences in size and organization also exist among plant taxa (Kubo and Newton, 2008).

Intramolecular recombination, leading to genome rearrangements and variable gene order even within single individuals, as well as duplications and deletions, are rather common. Furthermore, base substitution rates in plant mitochondria are frequently low (Wolfe et al., 1987), resulting in only minute differences within specific loci among individuals or even species.

Chloroplast genomes, on the other hand, exhibit a much more stable structure and also higher substitution rates than those of mitochondria (Wolfe et al., 1987). Intraspecific or even intrapopulational chloroplast variation was found high enough for population studies regarding gene flow (Wagner et al., 1987; Milligan, 1991; Heuertz et al., 2010).

An interesting approach is the contrast of paternally or biparentally inherited markers with maternally inherited markers. Using this combination, the ratio and the distances of pollen vs seed-based gene flow can be measured (Dong and Wagner, 1994; Latta et al., 1998; Fénart et al., 2007).

In gymnosperms, the situation is somewhat different. Here, mitochondria are mainly maternally inherited and thus dispersed via seeds only (Wagner, 1992), while chloroplasts are inherited mainly paternally and are, therefore, dispersed through pollen and seed.

Since pollen is normally distributed over far longer distances than seeds (Liepelt et al., 2002), mitochondrial markers exhibit a much stronger population differentiation than chloroplast markers and are important characters used for population genetic studies in

gymnosperms (Johansen and Latta, 2003), sometimes also used in connection with chloroplast or nuclear markers (Chiang et al., 2006).

DNA polymorphisms also in non-coding regions are widely used for phylogenetic inferences of species relationships (Bachmann, 2001). Indeed, some non-coding regions exhibit enough intraspecific variability and, thus, are used to analyze phylogenetic relationships of subspecies, varieties, and domesticated forms or populations structure (Pleines et al, 2009).

1.3 *The Mediterranean Basin: a hot spot of biodiversity for plants*

The Mediterranean Basin is considered as one of the most complex regions on Earth in terms of geological history, geography, morphology and natural history and the interplay between complex historical processes and heterogeneous environmental conditions has given rise to considerable plant biodiversity and endemism in this region (Thompson, 2005).

Since the late Tertiary, many paleogeographical events, such as the Messinian Salinity Crisis or the Milankovitch climate oscillations, could explain the heterogeneous evolutionary history of Mediterranean plant lineages but mainly the Pleistocene climatic cycles have changed profoundly the phylogeographical imprint of Mediterranean species (Weiss and Ferrand, 2007).

During cold periods (glacials), the decrease in temperature led to formation of large icecaps and glaciers, which were subsequently undergone partial melting during warm periods (interglacials).

During major glaciations, the icecap was expanded considerably limiting the sea, the temperate zones and the arboreal vegetation in a latitudinal strips or in small refuge areas (Williams et al., 1998).

The glaciers that covered mountain ranges such as the Alps, the Andes and the Rockies mountains stored large volumes of water, leading to a lowering of the sea level of approximately 120 m (Rholing et al., 1998) and subsequent formation of territorial linkages between regions that before were separated by the sea, thus promoting the dispersion of all living species.

These climate changes appear to have different effects depending on the latitude, the ocean currents and regional geographic characteristics for which the species have modified their distribution according the climatic and local geographical features.

In Europe, the last glacial maximum (LGM), about 18,000 years ago, led to the formation of a huge glacier that covered Britain and northern Europe and an icecap on top of the highest mountain ranges such as the Pyrenees, the Alps and the Caucasus (Frenzel, 1973; Nilsson, 1983). At the edge of the glaciers was the tundra, which covering Europe (Adams and Faure, 1997; Tzedakis et al., 2002) (Fig. 1).

The glaciations caused great changes in the distribution of species, with alternating periods of expansion and contraction. The advance of glaciers and of permafrost led to the loss of different habitats, with the consequent extinction of local populations and/or their confinement in areas called “glacial refuges”.

These represent areas where the temperate fauna and flora founded conditions (habitats) suitable for their survival during adverse climate periods.

During the interglacial periods, i.e. after the Holocene climatic warming, populations expanded towards north away from their southern refuges (Taberlet et al.,

1998), and rapidly colonizing repeatedly distant areas (Hewitt, 2000) This is especially true for taxa today having both northern and southern distributions and that have thus benefited from newly available sites to colonize (Afzal-Rafii and Dodd, 2007).

However, in northernmost regions, expansion and founder effects are expected to undermine allelic richness and heterozygosity of colonizing populations (Petit et al., 2003), while southern regions since free from icecap and permafrost soils, would have permitted much more stable population dynamics for many species with resulting higher genetic diversity (Hewitt, 1996, 2000).

The Mediterranean region, in particular, has constituted a global refuge for relict plants where floristic exchange and active speciation were favorite and at present the Mediterranean Basin is believed one of the world's major biodiversity hotspots.

Recently, Médail and Diadema (2009) identified 10 Mediterranean regional hotspots, representing main areas of plant biodiversity and including 52 putative refuges (Fig. 2). About 50% of them are present in Iberian, Italian and Balkan peninsulas, supporting the key role played during both glacial and interglacial periods by the three major peninsulas and demonstrating as from these refuges started each colonization phase that have marked the expansion periods in interglacial stages.

Conifers and, particularly, some species belonging to the *Pinus* genus, were the most important components of the above peninsular refuges from which they plan the subsequent migrations, during postglacial stages.

1.4 *Pinus laricio* Poiret*: a complex and interesting Mediterranean forest tree species*

Among pine species largely present in the Mediterranean Basin, European black pine (*Pinus nigra* Arnold), belonging to the Pinaceae family, is considered a relict species pioneer that has differentiated itself in a very large and fragmented distribution area, extending from Asia Minor in the east through the Balkans to western habitats on the Iberian Peninsula and in North-Western Africa (Gaussen et al., 1993; Vidakovic, 1991).

Indeed, the variability in morphological, physiological and ecological traits of *Pinus nigra* have led to consider this species as a complex of six allopatric subspecies: *Pinus nigra* ssp. *nigra*, ssp. *dalmatica*, ssp. *pallasiana*, ssp. *mauritanica*, ssp. *salzmannii* and ssp. *laricio* Poiret (Bogunić et al., 2011; Quézel and Médail, 2003) (Fig. 3).

The first black-pine-type fossils date to the Miocene, about 20 million years ago. The ice cycles that shaped the Quaternary period in Europe are believed to have been responsible for the currently very discontinuous range of this species. However, studies using morphological and genetic markers have confirmed the common phylogenetic origin of all black pines.

However, *Pinus laricio* Poiret is treated as the most divergent subspecies of *Pinus nigra* and, some scientists, according to fossil records (Studt, 1926; Stojanoff and Stefanoff, 1929), morphological (Gaussen et al., 1993), chemotaxonomical-anatomical (Fineschi, 1984) and karyological (Cesca and Peruzzi, 2002) studies, regard laricio pine as a taxonomically independent species.

The species grows up to 50 m tall, with a trunk that is usually straight and gradually tapered toward to top of the plant. Moreover, the trunk can reach dimensions of diameter

up to 180 cm, in oldest plants. It has very high longevity mainly compared with *Pinus nigra*: 4-5 centuries in isolated plants (Fig. 4).

Pinus laricio is a monoecious wind-pollinated species, and its seeds are wind dispersed. The megasporangiate strobili occur predominantly high in the crown and near the ends of branches in the mid- or lower crown, and the microsporangiate strobili are borne predominantly on the interior and lower portions of the crown.

The microsporangiate strobili contains pollen grains while the macrosporangiate strobili contains ovules. Each ovule has multiple archegonia and may be pollinated by a different number of pollen grains. Multiple archegonia and multiple pollination events also provide an opportunity for competition and selection among embryos within the ovule, since only one usually survives to germinate.

Nevertheless, in *Pinus laricio*, such as in other *Pinus* species, the lack of self-incompatibility mechanisms provides a major flexibility in mating system that could be affected by elevate levels of inbreeding than outcrossing and events of self-sterility (Richardson, 2000).

Pinus laricio trees reach sexual maturity at 20 - 30 years in woods while at 15 - 18 years in marginal or isolated plants (Avolio, 2003).

In addition, it is a termo-xerophile species, though needs of bioclimatic conditions of humid (800-1,500 annual rainfall) to counter the summer aridity. The values of the mean annual temperature range from -2°C to 25°C, registered in January and August, respectively. Moreover, it is clearly a heliophile and pioneer species able to colonize open ground (Quézel and Médail, 2003). For reproduction it needs a high degree of lighting and the presence of poorly differentiated soil, not too rich in humus; when *Pinus laricio* finds these conditions form pure and unpeers woods.

Pinus laricio naturally occurs in restricted and discontinued areas in Calabria, Sicily and Corsica. In Calabria it grows on the Aspromonte mountain and mainly on the Sila plateau where laricio pine forests cover 4,000 ha and 40,000 ha, respectively, characterizing the landscape from 900 m up to 1,700 m above sea level, but with an environmental optimum in lower mountain plan, slopes hot, arid and siliceous soils by granitic origin.

The current distribution of populations configure how much of them remains of the largest coating forest of southern Italy, the so-called "Silva brutia" of the Romans (Avolio and Ciancio, 1985). The remnants forest of *Pinus laricio* in Sicily and Corsica cover 4,000 ha and 21,000 ha, respectively.

In Sicily, the tree forest of *Pinus laricio* grows on lower differentiated soil of volcanic origin while in Corsica almost exclusively on granitic acid soils or sandy soils.

In addition, in Sicily, on the eastern slopes of Etna, pure populations of the *Pinus laricio* are located between 1,200 m to 2,000 m above the sea level while in Corsica the of altitude levels ranged from 1,000 m up to 1,800 m above sea level.

Pinus laricio in Calabria is slightly more tolerant to limestone soils compared to those present in Corsica. Moreover, *Pinus laricio* in Corsica avoids the dry soils, those are too acids and with high humidity, but it can grow on clay soils. The Corsican pine trees also have a certain resistance to the action of the sea winds.

Finally, laricio pine is widely used in commercial production of hybrids in forest regeneration due to its high survival rate. Wood is durable and rich in resin, easy to process, and is appreciated for building and roofing because of its straightness and thin branches. If properly thinned, its low amount of duramen makes it a fine carpentry and cabinetry wood (Isajev et al, 2004). More recently, collaborative efforts between

government and private sector are focusing on the use of *Pinus laricio* as a source of biofuel.

2. Aim of research

In-depth genetic studies about the *Pinus laricio* Poiret natural populations present in Calabria, Sicily and Corsica were never carried out, though the high forestry, phytogeographical, landscape and environmental interest on this species.

The key objective of this study is to estimate genetic polymorphism level and distribution, and characterize population structure of *Pinus laricio* Poiret, using microsatellite loci, and relate the detected patterns to available information about the life history of the species and landscape characteristics.

The accurate characterization of the *Pinus laricio* population structure within its native range is further needed to develop appropriate management or silvicultural strategies and, thus, to preserve genetic diversity.

3. Methods

3.1 *Study site and plant sampling*

Green tissues (needles) used in this study were collected from seven uneven-aged populations of *Pinus laricio*, four of them were from the Sila's National Park, namely Fallistro (FAL), Roncino (RON), San Salvatore (SAN) and Trenta Coste (TRE); one from the Maesano area (ASP), within the Aspromonte's National Park and one from the Linguaglossa territory (ETN), within the Etna's Regional Park. Finally, Corsican pine individuals (COR) were collected from Restonica Valley in the vicinity of Corte's city, within Regional Natural Park of Corsica (Fig. 5, Tab. 2).

A total number of 459 individuals were sampled for all populations, following a "random and non-contiguous" method for each plant.

Based on dendrometric information, obtained by operators of CRA-SAM (Consiglio per la Ricerca e la Sperimentazione in Agricoltura - Unità di ricerca per la Selvicoltura in Ambiente Mediterraneo), sampled individuals from Sila, Sicily and Aspromonte were subdivided into four age classes, considering the tree-trunk diameter at breast height (DBH): 1) *Fustaia Stramatura* (FS), including trees with diameter > 300 cm and more than 200 years old; 2) *Fustaia Matura* (FM), comprising individuals 120 years old with a circumference range of 180-300 cm, 3) *Fustaia Adulta* (FA), with trees 80 years old and 90-150 cm of circumference range and 4) *Novelleto* (NV), with juvenile individuals (10-15 years old) and a circumference range of 30-50 cm.

An additional age class, *Fustaia Giovane* (FG), including trees with diameter 60-80 cm and a mean age of 40 years, was sampled only for FAL population. All the 11 Corsican samples came from NV individuals.

Furthermore, geographic coordinates were obtained for each sampled individual with a GPS (Global Positioning System) receiver (Leica-Magellan); flat coordinates were recorded in Gauss–Boaga coordinate system and transformed into UTM geographic coordinate system.

Individually sampled needles were conserved at -80°C until processing.

3.2 *DNA extraction and microsatellite analysis*

Total genomic DNA was isolated from needles of individual trees using the DNeasy Plant Mini Kit (Qiagen). Screening for polymorphism was conducted with 5 chloroplast (cp) (Pt30204, Pt36480, Pt45002, Pt71936 and Pt87268), 1 mitochondrial (mt) (Nad3-1) and 3 nuclear (nu) (PtTX4001, PtTX3107 and SPAG7.14) microsatellite primer pairs (Tab. 3), originally designed for *Pinus taeda* and *Pinus sylvestris* population genetic studies (Vendramin et al., 1996; Soranzo et al., 1998; Soranzo et al., 1999; Auckland et al., 2002).

PCR was performed using both a Perkin Elmer GeneAmp PCR System 9600 thermal cycler and an Eppendorf Mastercycler Pro with conditions varied for different primers and for the presence of the forward ones VIC, 6-FAM, NED and PET labeled.

Each amplification reaction contained approximately 10 ng genomic DNA, 1X reaction buffer [Tris-HCl (100 mM) pH 8.3, KCl (500 mM), gelatin (1 mg/mL)], 2.5-3.5

mM $MgCl_2$, 0.2 mM dNTP mix, 1 pmol/µL of each primer, 1 U Red Taq DNA polymerase (Euroclone) and deionised water to a total reaction volume of 20 µL.

PCR conditions using mt- or cpSSR primers were as follows:

- Initial denaturing step of 3 minutes at 94°C
- 30 cycles of:
    - 94°C (30 seconds) denaturation
    - 55°C (30 seconds) annealing
    - 72°C (30 seconds) extension
- Final extension period of 5 minutes at 72°C.

The touchdown PCR conditions with nuSSR primers were set as follows:

- Initial denaturing period of 3 minutes at 94°C
- 10 touchdown cycles:
    - 94°C (30 seconds) denaturation
    - 60°C (30 seconds) annealing, declining 1°C every cycle until 50°C
    - 72°C (30 seconds) extension
- 30 cycles
    - 94°C (30 seconds) denaturation.
    - 50°C (30 seconds) annealing
    - 72°C (30 seconds) extension
- Final extension step of 5 minutes at 72°C

Particularly, for SPAG 7.14 primer pair, the touchdown cycles were 5 and annealing temperature started with 55°C instead of 60°C.

Amplicons were visualized through 3% agarose gel electrophoresis in order to determine presence or absence of specific products and, then, purified with QUIAquick PCR Purification Kit (Qiagen).

Next sequencing analysis on an ABI Prism 310 Sequencer (Applied Biosystems) was carried out to confirm both the presence of the microsatellites in the amplified fragments and the fact that variation in length was due to different numbers of repeats within the microsatellite regions.

3.3 *Microsatellite genotyping*

Aliquots of each fluorescent dye labeled PCR product (0,5 µL) were additioned at Formamide (11,5 µL) and internal size standard LIZ 500 (0,5 µL) (Applied Biosystems).

After denaturation for 5 min at 95°C, labeled amplicons were separated by capillary electrophoresis on an Applied Biosystems Prism 3730*XL* Genetic Analyser using filter set G5.

The data were analyzed using Peak Scanner v. 1.0 software (Applied Biosystem). Each resulting peak was considered as an allele at a codominant locus and the genotype of each individual at each locus was recorded.

3.4 *Data analyses: within and among population diversity and genetic structure of populations*

Main values of chloroplast and nuclear genetic diversity [allelic frequencies, number of alleles (Na), number of effective alleles (Ne), haplotype diversity (H), observed heterozygosity (Ho), expected heterozygosity (He), fixation index (F) and inbreeding coefficient (F_{IS})] were computed with GenALEx version 6.4 software (Peakall and Smouse, 2006).

Allelic richness was calculated with the FSTAT version 2.9.3.2 program (Goudet, 2001), using the rarefaction approach proposed by El Mousadik and Petit (1996) to correct for differences in sample size. Standard sample size consisted of 11 individuals, which corresponded to the smallest *Pinus laricio* population sampled (COR) (Tab. 2).

A test for Hardy-Weinberg (H-W) Equilibrium was performed using the program GENEPOP version 4.1 (Rousset, 2008) and according to Fisher (1935), to determine whether the observed genotypes are consistent with the expectations under random mating.

When significant deficiencies of heterozygotes from H-W expectations were found, the presence of a relatively high frequency of null alleles was suspected.

Therefore, loci with high frequencies of null alleles were identified by estimating null allele frequencies for each locus and each population, using the software FREENA (Chapuis and Estoup, 2007).

Values > 0.19 of null allele frequency have been considered as a threshold over which significant underestimate of He due to null alleles can be found (Chapuis and Estoup, 2007).

An Analysis of Molecular Variance (AMOVA) (Excoffier et al., 1992; Michalakis and Excoffier, 1996) was performed to estimate levels of genetic differentiation assigned

by computing Φ_{PT} and F_{ST} estimators, assuming the infinite allele model (IAM), and the R_{ST} estimator, assuming the stepwise-mutation model (SMM) (Slatkin, 1995). Statistical significance of all the Φ_{PT}, F_{ST} and R_{ST} estimators were tested using 10,000 permutations.

GENETIX version 4.04 software (Belkhir et al., 2001), also, was used to estimate Gst value.

In addition, a Principal Coordinate Analysis (PCA), also available in GenALEx version 6.4 program, was conducted using Nei's unbiased genetic distance pairwise population matrix to determine whether observed patterns in the molecular data support the partitioning of the *Pinus laricio* samples into specific groupings.

The degree of isolation-by-distance was assessed by testing the association between geographic distances (log-transformed) and genetic distances (expressed as linearized Φ_{PT} and $F_{ST}/1\text{-}F_{ST}$ for cpSSR and nuSSR, respectively) for all pairs of populations. The statistical significance of the associations was tested based on a Mantel test (Mantel, 1967).

Spatial Autocorrelation Analysis was conducted for each *Pinus laricio* population and for both cpSSR and nuSSR to test the existence (H1) or not (H0) of a non-random distribution of genotypes in space. The analysis, performed comparing genetic distance and geographic genetic distance matrices, was based on the parameter of spatial autocorrelation (*r*), that should be > or < than 0, estimated with random permutations and bootstraps, selecting two different options of "Even Distance Classes" and "Even Sample Sizes".

Finally, Bayesian clustering methods were applied to infer the number of genetic units and their spatial delimitation.

Bayesian clustering methods use genetic information to ascertain population membership of individuals without assuming predefined populations. They can assign either the individuals or a fraction of their genome to a number of clusters based on

multilocus genotypes. The methods operate by minimizing H-W and linkage disequilibria, and the assignment of each individual genotype to its population of origin is carried out probabilistically (Chen et al., 2007).

Among Bayesian methods, STRUCTURE 2.3.3 program (Pritchard et al., 2000; Pritchard et al., 2010) was used to identify clusters (K) of genetically similar individuals. The analysis was conducted under the admixture model and the option of correlated allele frequencies between populations.

Five independent runs (iterations) were performed for each K, that varied from K = 1 to 10. The maximum number of clusters used was greater than the number of populations, for detection of potential substructuring within samples. All runs were performed with burn-in length of 10,000 and repetition number of 100,000 iterations.

The number of clusters was determined using the ad-hoc statistic ΔK according with Evanno et al. (2005) and based on the rate of change of log likelihood of data [L(K)] between consecutive K values used to select the optimal K.

TESS software (Durand et al., 2009) uses a Hidden Markov Random Field model that assumes that the log-probability that an individual belongs to a particular cluster given the cluster membership of its closest neighbors is equal to the number of neighbors belonging to this cluster.

The probability that two neighboring individuals belong to the same spatial cluster is controlled by a parameter known as the "interaction parameter" (Ψ). Any non-zero value introduces spatial dependence, with default value set to 0.6.

Three quantities, that is, a graph specifying the set of neighbors of each individual, the interaction parameter and the maximum number of clusters (K), were involved in the TESS algorithm. Models with and without admixture are also available. The maximum

number of clusters that best fit the data is, thus, chosen by the user from the statistical Deviance Information Criterion (DIC).

In our TESS analysis using both cpSSr and nuSSR data set, a Ψ value = 0.6, without admixture and with admixture, was tested. In addition, 9 runs were performed for a total number of sweeps of 50,000, a burn-in number of sweeps equal to 10,000, with the number of clusters K from 2 to 10, and 3 repetitions (for each data set) for a total of 27 runs. Then, the DIC values of the 27 runs were averaged.

GENELAND program implements a method developed by Guillot et al. (2005) that quantifies the amount of spatial dependence in a data set, estimates the number of populations, assigns individuals to their population of origin, and detects individual migrants between populations, while taking into account uncertainty on the location of sampled individuals.

The spatial domain of the sample is partitioned into a union of a random number of polygons by Voronoi tessellation that is randomly assigned to one of potential K spatial clusters. K is considered unknown, with maximal value (Kmax) input by users, and estimated by the algorithm. When polygons are assigned to different spatial clusters, the joint probability that any two polygons belong to the same spatial cluster decreases with geographical distance between them.

Estimates of K (the number of spatial clusters) and individual assignment probabilities are obtained using a Markov Chain Monte Carlo (MCMC) algorithm, an iterative procedure which starts from arbitrary values for all unknown parameters and modifies them so that after many iterations they approximate true values.

In our analysis, 10 indipendent runs were performed for K from 1 to 50. Each run consisted of 10^5 MCMC iterations with a thinning interval of 100, using correlated allele frequencies and spatial information.

In order to obtain the posterior probabilities of population membership of each individual and each spatial area, 10 runs were then post processed with a burn-in of 10^4 MCMC and 100 pixels along the X-axis and 200 pixels along the Y-axis. Then, the consistency of the results was visually checked by comparing the outputs across the 10 runs.

BARRIER program (Manni et al., 2004) was used to verify the presence of genetic barriers among populations. This software implements the Monmonier's algorithm based on a neighbor graph (such as a Delaunay triangulation) between sampled populations or individuals, and calculates the genetic distances associated with each edge of the graph. The algorithm builds growing barriers from the edge with the largest genetic distance, and extends it to the adjacent edges associated with the next largest genetic distance.

DIYABC (Cornuet et al., 2010) is a inference program based on Approximate Bayesian Computation (ABC), in which scenarios can be customized by the user to fit many complex situations involving any number of populations and samples.

Such scenarios involve any combination of historical population events (divergences, admixtures, population size fluctuations). DIYABC can be used to compare competing scenarios, estimate parameters for one or more scenarios and compute bias and precision measures for a given scenario and known values of parameters.

4. Results

4.1 *Chloroplast microsatellites analysis*

4.1.1. *Genetic diversity estimates*

Three out the five used cp markers (Pt30204, Pt71936 and Pt87268) (Tab. 3) were found polymorphic in all populations studied, although the degree of variability was different at each locus. Indeed, the number of amplified size variants (or alleles) for each polymorphic site ranged from 6 to 11, the maximum of which was showed by the locus Pt30204 with 11 size variants followed by Pt87268 and Pt71936 (Tab. 4).

On a population/ per locus basis the mean number of alleles (Na) was 5.91, a higher value compared with that previously reported for *Pinus nigra* Arnold (Naydenov et al., 2006).

The size of all the detected alleles ranged from 138 to 173 bp, reflecting a large difference in the number of repeats between the different alleles (Tabs. 3, 4).

4 out of 27 size variants identified at the 3 loci occurred at low frequencies (<1%) and, thus, represent a group of "rare size variants" (Tab. 4).

Two of this rare alleles (alleles 138 at locus Pt30204 and allele 145 bp at locus Pt71936) (Tab. 4), are also "private", that is, exclusively present in individuals of TRE and ASP populations, respectively.

Instead, the other rare size variants such as alleles 167 bp and 173 bp at locus Pt87268 in Table 4, are common in ETN and many Sila populations.

The most common size variants were fragments 143 bp at locus Pt30204 with a frequency of 31.3%, 147 bp at locus Pt71936 with a frequency of 57.5% and 170 bp at locus Pt87268 with a frequency of 51.9% (Tab. 4).

The mean number per locus and populations of effective alleles (Ne) was 3.05, a genetic variability value comparable with the allelic richness parameter (Ar = 4.52) (Tab. 5), calculated on minimum sample size of 11 aploid individuals.

Haploid diversity (h) was on average 0.65 with the largest value observed in the Maesano population (ASP = 0.74), while Fallistro stand showed the lowest variability (FAL = 0.60) (Tab. 5).

The 27 size variants were combined in 84 different haplotypes, 25 of these were found in the FAL population versus 34 present in the ETN stand. Only 8 haplotypes was detected in the Restonica population (COR), very likely due to small size sampling (Tab. 5).

39 out of 84 haplotypes were unique or private (frequency <1%) and are due to the presence of rare size fragments and/or rare combinations of size fragments. These unique haplotypes were found in all areas of the *Pinus laricio* natural distribution, with exception for the population located in the Corsican territory. In particular, three and eight unique haplotypes occurred in the SAL and ASP/ETN stands, respectively (Tab. 5).

The effective number of haplotypes in each stand ranged from 6.37 in Corsica to 20.5 in Sicily (Tab. 5). Furthermore, within the Calabrian area, Maesano (ASP) exhibits the largest percentage of individuals with private haplotypes (21%) versus only the 10% detected in the FAL stand (Tab. 5).

Three haplotypes (h-18, h-62 and h-68 in Table 6) are common to all *Pinus laricio* populations analyzed, even though they occur with different frequency. The most frequent

haplotype h-41 (11.6%) was found in the Sila, Etna and Corsica stands but not in the Maesano (ASP) population; conversely, the haplotype h-72 (frequency 3.5%) was detected in the populations from Sila, Aspromonte and Etna but it was absent in the Corsica stand (Tab. 6).

Overall, the predominant haplotypes which occur in at least ½ of the analyzed populations correspond to 45.3% of the laricio pine genetic structure.

Intra-population gene diversity values (He) based on haplotype frequency varied from 0.917 in FAL to 0.966 in ETN (Tab. 5) while the total haplotype diversity calculated across all populations was 0.942 on average.

4.1.2. *Genetic variability among populations*

Considering the data generated with 3 cpSSRs, in the above cited laricio pine stands, an analysis of molecular variance (AMOVA) was conducted to partition the cpSSR variance into within- and among populations (Tab.7). The above analysis revealed that 3.4% of the variation was found among populations with 96.6% of the diversity being expressed within populations.

The ϕ_{PT} value is slightly lower compared with that previously reported for *Pinus nigra* Arn. and other *Pinus* species (Naydenov et al., 2006; Bucci et al., 2007; Dzialuk et al., 2009). Indeed, the calculated F_{ST} and G_{ST} that estimate the degree of population differentiation were 3.4 and 4.1%, respectively, confirming the AMOVA results. The slight discrepancy is probably due to the different principles of G-statistics and AMOVA-population statistics.

The AMOVA analysis was also performed using pairwise comparison of populations between two (or more) different areas, revealing an higher molecular variance value (5.3%) between Sila and Aspromonte populations followed by lower differentiation values for Sila-Etna (2.7%) and Aspromonte-Etna (1.9%) comparisons (Tab. 7).

Genetic distance between populations, calculated according Nei (1978), varied from 0.023 between SAL and TRE to 0.227 between SAL and ASP (Tab. 8). Particularly, the smallest values, ranging from 0.023 (TRE-SAL) to 0.064 (RON-TRE), were observed within *Pinus laricio* stands located in the Sila National Park (Tab. 8).

According to AMOVA analysis, the highest dissimilarity values were, thus, found between Sila and Aspromonte populations (Tab. 8), while the Calabrian populations displayed high genetic similarity with those of the Linguaglossa (ETN) and Restonica (COR).

The average gene flow (N_m) is estimated to be equal to 14.21 migrants per generation.

A Mantel (1967) test was performed to verify the hypothesis about the spatial pattern in the distribution of chloroplast gene diversity at the stand level. To this end, the log matrix of geographic genetic distance (*x* data) and the linearized Φ_{PT} matrix (*y* data) were associated. A negative and not significant relationship (rxy = -0.265, $P < 0.05$) was found (Fig. 6), indicating the absence of a pattern of isolation by distance among *Pinus laricio* populations within their natural area of distribution.

In addition, to verify the hypothesis of the presence (H1) of a non-random distribution of genotypes in space, an analysis of spatial autocorrelation was carried out for each *Pinus laricio* populations (Figs. 7, 8).

The results of this analysis showed, in each population and for both the "Even Distance Classes" (Fig. 7) and "Even Sample Sizes" (Fig. 8) options used, a value of the parameter of spatial autocorrelation included between Upper (U) and Lower (L) confidence limits bound the 95% confidence interval about the null hypothesis (H0) (r = 0). Thus the non-random distribution of genotypes was discarded.

Next, for better interpretation of the genetic distances found among laricio pine natural populations, the Principal Coordinates Analysis (PCA) was conducted. This analysis, performed on Nei's unbiased genetic distance matrix and based on 27 different size variants, clearly distinguished the seven laricio pine stands into three main groups, one including the FAL, TRE, SAL and RON populations, the second group comprising the Linguaglossa (ETN) and Restonica (COR) populations and the third cluster including the Maesano (ASP) stand alone (Fig. 9).

In the PCA analysis, the first coordinate explains 73.24% of unbiased genetic distance while the second coordinate explains 13.82%, with a total value on the two main coordinate equal 87.06% of the same parameter.

Also, the population structure analysis of natural populations of *Pinus laricio* was carried out using a Bayesian clustering algorithm, implemented in the TESS software (Durand et al., 2009), that incorporates spatial information when identifying clusters of individuals.

The TESS software allows the testing of different values of the spatial dependent parameter Ψ that weights the relative importance given to spatial connectivities. A value of Ψ = 0.6 was tested in this analysis.

For Ψ value considered in this analysis, 9 runs for *K* from 2 to 10 were performed, with a total number of sweeps of 50,000, a burn-in number of sweeps equal to 10,000 and 3 repetitions for a total of 27 runs.

Then, the DIC values of the 27 runs with the smallest DIC were averaged. In plotting the average DIC values *versus* K (from 2 to 10) an inflection point at K = 5 was obtained. Sometimes the effective number of clusters in the data may be a smaller value of K and, in this case, the best K seems to be 2 because for this value, in the model without admixture, was observed the best barplot stabilizes, thus indicating the presence of a moderate spatial correlation between individuals (Fig. 10).

Therefore, the seven natural populations of *Pinus laricio* are grouped into two main clusters: the first cluster contained all the four Sila populations of laricio pine while the second cluster included the remaining stands of Maesano, Linguaglossa and Restonica.

4.1.3. *Evolutionary inferences of Pinus laricio populations based on cpSSRs.*

A DIYABC analysis, a variant of ABC analysis (EXcoffier et al., 2005), was performed considering two main populations of *Pinus laricio*: the first including all the Sila National Park populations (Pop 1 in Fig. 11) while the second comprises the populations of Restonica, Linguaglossa and Maesano (Pop 2 in Fig. 11). In the DIYABC analysis were tested three possible scenarios to describe the demographic evolution of natural populations of *Pinus laricio* through time.

In the first scenario, it was assumed that Pop 2 was separated from Pop 1 at time T1. In the second scenario, it was assumed that Pop 2 was separated from an ancestral

population at time T3; at time T2 this population was subject to bottleneck that originated the Pop 1 at time T1. Finally, in the third scenario, it was assumed that Pop 2 was separated from an ancestral population at time T2 and at time T1 Pop 1 was separated from Pop 2 (Fig. 11a).

These scenarios were tested on 3,000,000 simulations, computing the main parameters as PCA (Principal Component Analysis), logistic regression, the values of F_{ST}, mean genic diversity and the posterior distribution of parameters. The results of this analysis, and for each parameter analyzed, showed that the third scenario was better supported than the other two scenarios because the most observed summary statistics are well in the range of those simulated (Fig. 10 and data not shown).

Moreover, it was possible to plot the values of T1 and T2 on the time scale not in generations but in number of mutational events, based on the posterior distributions, with ABC analysis. So, for the third scenario the values of T1 and T2 was respectively 2.40E+002 and 5.47E+003 (data not shown).

Knowing the estimated time values of mutational events and the years for a reproductive generation in *Pinus laricio* (20 years), the time T2 of separation of Pop 2 from ancestral population (109,000 years ago) and the time T1 of separation of Pop 1 from Pop 2 (4,800 years ago) was detected.

4.2 *Nuclear microsatellites analysis*

4.2.1 *Allelic diversity of microsatellite loci and genetic variation within populations*

Three published primer pairs (Auckland et al., 2002; Soranzo et al., 1998) flanking nuclear microsatellites were employed to investigate the level of genetic variation among the 7 *Pinus laricio* populations sampled in this study.

The three markers (SPAG7.14, PtTX4001 and PtTX3107) were found to be polymorphic in all analyzed populations, producing a variable number of alleles per locus (Tab. 9).

Indeed, between 6 (at locus PtTX3107) and 38 (at locus SPAG7.14) size variants were identified with an average (Na) for all loci and populations of 11.48 (Tab. 10). The effective number of allele (Ar = allelic richness), calculated on minimum sample size of 11 diploid individuals, ranged from 5.33 in the Restonica population (COR) to 7.05 in the RON stand with an average of 6.56 (Table 10).

A total of 52 size variants at the 3 loci were identified, of which 15 with frequencies <1% and, thus, considered as "rare" alleles. Four of them are also "private" since they are present in ETN (alleles 222 and 244 bp at locus SPAG7.14), RON (allele 213 bp at locus PtTX4001) and SAL (allele 232 bp at locus SPAG7.14) trees (Tab. 9).

The remaining "rare" alleles (8, 2 and 1 alleles at the SPAG7.14, PtTX4001 and PtTX3107 loci, respectively) were found common to two or more *Pinus laricio* populations (Tab. 9).

The most common size variants, namely the 202 bp (at locus SPAG7.14), 211 bp (at locus PtTX4001) and 160 bp (at locus PtTX3107) alleles, were found with a frequency of 17.9%, 50% and 71.6%, respectively (Tab. 9).

Examination of intra-population genetic diversity revealed the lowest value of Nei's unbiased genetic diversity (He = 0.628) among individuals from FAL population while samples from ASP stand showed a highest variability (He = 0.701) (Tab. 10).

In all sampled populations, the observed heterozigosity (mean Ho = 0.528) was lower than expected (mean He = 0.674). The difference determines a significant positive value for mean Fixation Index (F = 0.204) (Table 10) and inbreeding coefficient (F_{IS} = 0.197) (Table 11), that could very well be attributed to non-random mating and null alleles.

Therefore, FREENA software (Chapuis and Estoup, 2007) was used to recompute the null allele frequencies and, thus, adjust F estimates. Frequencies lower than 0.19 were obtained for all loci for each sampled population, except for COR stand, where null allele frequencies higher than 0.2 were only found at the loci SPAG7.14 and PtTX4001 (Tab. 12).

A recalculation of all the F-statistics were also made and no significant differences were observed. For example, the global F_{ST} (Weir, 1996) is generally calculated in order to estimate the proportion of the total genetic variation due to the differentiation among populations. Taking or not into account the null allele frequencies, the global F_{ST} values were 0.018 and 0.019, respectively (Tab. 11).

Moreover, a χ^2 test using Fisher's method as implemented in GENEPOP 4.1 software (Rousset, 2008) was performed to evaluate if deviations from the H-W equilibrium were probabilistically significant or no. We found very highly significant tests for all locus and populations (Tab. 11), thus, suggesting that the excess of homozigosity

observed in all the *Pinus laricio* populations is due to inbreeding between closely related individuals.

Furthermore, to investigate the potential effects of inbreeding over long periods of time, individuals from each *Pinus laricio* populations were partitioned in diameter or age classes, and F values computed and compared (Tab. 13). A considerable excess of heterozygotes was only found for trees 10-15 years olds (NV in Tab. 13) belonging to the FAL and ETN stands, as indicated by the negative Fixation Index obtained (-0.024 and -0.104, respectively).

Instead, from 80 years old individuals (FA) to those 120 years olds (FM), the heterozygosity levels (radically) declined and an excess of homozygous is recorded (Tab. 13).

4.2.2. *Genetic diversity among populations*

Global and hierarchical population genetic structure were evaluated by analysis of molecular variance (AMOVA) (Excoffier et al.,1992) under (IAM) and (SMM) models (Tab. 14).

The AMOVA-IAM model analysis revealed that 2,1% of the variation was found among populations with 97,9% of the diversity being expressed within populations. The F_{ST} value was comparable to the Nei's coefficient G_{ST} (3.1%) but lowest than R_{ST} value (9.2%) obtained when the genetic differentiation among the seven laricio pine populations was calculated under the SMM model (Slatkin, 1995).

The average gene flow (N_m), based on the F_{ST} value via AMOVA, was equivalent to 11.65 migrants per generation, with RON stand showing the greatest value (data not shown).

AMOVA analysis was also performed on pairwise populations, evaluating the differences in the F_{ST} parameter. The comparison between ETN, ASP and Sila populations and the COR stand showed values of molecular variance among populations ranging from 1.9% to 3.3% while no genetic difference seems to emerge between ASP-ETN populations (Tab. 14).

The greatest values of genetic distance, estimated according to Nei (1978), were observed between COR and each other laricio pine stand, sampled in this study (Tab. 15). Particularly, the topmost dissimilarity was found with FAL (0.246), the latter also genetically different with respect to ASP stand (0.097). Moderately high distance values were recorded among the four populations within Sila National Park while ASP and ETN populations were found closely related genetically (0.021) (Tab. 15).

To ascertain whether there is a statistically significant relationship between genetic and geographic distance, a Mantel (1967) test was performed. By plotting the genetic distance among population pairs as a function of the geographic distance between those pairs, no significant correlation was found (rxy = 0.266, $P < 0.05$) (Fig. 12), thus, rejecting the assumption of an isolation by distance among *Pinus laricio* natural populations here sampled.

Each single population of *Pinus laricio* was further tested by a spatial autocorrelation analysis to verify the hypothesis (H1) of a non-random distribution of genotypes in the space. The estimated parameter of spatial autocorrelation (r), if in agreement with this hypothesis, should be > or < than 0.

Values of r = 0 ($P < 0.05$) were obtained for each laricio pine population, thus, indicating that the individual genotypes in space are randomly distributed in space (Figs. 13, 14).

Stand genetic structure was tested using three different approaches. First, the Principal Coordinate Analysis (PCA), performed on Nei's unbiased genetic distance matrix and based on 52 different size variants, showed that the 7 sampled populations of *Pinus laricio* were separated in three main groups (Fig. 15).

Group I contains three out the four populations from the Sila National Park (RON, SAL, TRE) as well Maesano (ASP) and Linguaglossa (ETN) stands. Group II and III include FAL and COR populations, respectively. In the PCA analysis the first two principal axes explain in total 86.41% of unbiased genetic distance (Fig. 14).

Next, Bayesian clustering algorithms such those implemented in TESS 2.3 and GENELAND programs were used for inferring laricio pine population structure (Durand et al., 2009; Guillot et al., 2005). Both programs include spatial coordinates into classical analyses based on multilocus genotypes.

In TESS analysis, a Ψ value of 0.6, without admixture and with admixture, was tested. For the above Ψ value, number of runs and clusters were performed as reported in Methods.

Graphical outputs by plotting the average DIC values *versus* K (from 2 to 10) showed an inflection point at K = 5, without admixture model, and an inflection point at *K* = 6 with admixture model.

However, when the assignment probabilities of individuals to different clusters were displayed in a bar chart, the barplot stabilizes at K = 2 for both models, thus, showing a moderate spatial correlation between individuals (Fig. 16).

Therefore, the natural populations of *Pinus laricio* were assigned to 2 main clusters: cluster 1 includes all Sila populations while cluster 2 contains the remaining stands of Maesano (ASP), Linguaglossa (ETN) and Restonica (COR) (Fig. 16).

The ability of the GENELAND program (Guillot et al., 2005) to infer population structure is well recognized compared to other spatially implicit clustering methods, particularly, when genetic differentiation is weak, a typical situation for trees collected over a large geographic area.

As result of my GENELAND analysis, the best number of clusters along the chain (after burn-in) for the seven populations of *Pinus laricio* analyzed in this study was K = 2 (Fig. 17a), such as TESS outputs. Two main clusters were evident, the first including the four populations of Sila plateau, the second one ASP, ETN and COR stands (Fig. 17b).

Furthermore, to test the effect of the landscape characteristics on the genetic structure of *Pinus laricio* populations, an additional method implemented in BARRIER software, based on the Monmonier's maximum difference algorithm, was used (Manni et al., 2004).

Nei's genetic distances were used to generate 10 bootstrapped matrices allowing to identify one most supported genetic barrier (Fig. 18) that separates the FAL population from the other *Pinus laricio* populations located in the Sila National Park.

A further Bayesian clustering approach, implemented in STRUCTURE software (Pritchard et al., 2000; Pritchard et al., 2010), was used as final approach for population structure analysis.

This program performs a multilocus analysis on the genotypes of individuals sampled in different geographical areas, but it does not require geographic information. STRUCTURE estimates the population number (K) in a sample assigning individuals to

one or more of these populations. Analyses were conducted under the admixture model and the option of correlated allele frequencies between populations.

My STRUCTURE analysis based on the information from three nuSSR loci demonstrated that the model with K = 3 had the highest ΔK value explaining adequately the data (Fig. 19a).

This finding suggests that the most probable number of populations for 459 sampled *Pinus laricio* individuals was 3. The first group includes only the FAL population while the remaining stands of Sila National Park were assigned to group 2. Finally, the third cluster comprises the populations of Maesano (ASP), Linguaglossa (ETN) and Restonica (COR) (Fig. 19b).

4.2.3. *Inferences on Pinus laricio population history based on nuSSRs*

Laricio pine population history was analyzed with the DIYABC method (Cornuet et al., 2010) considering two main populations: the first (Pop 1 in Fig. 20a)grouping all the four stands of Sila National Park, the second one (Pop 2 in Fig. 20a) includes the populations of Restonica (COR), Linguaglossa (ETN) and Maesano (ASP).

Three possible scenarios describing demographic evolution over time of *Pinus laricio* populations were tested in the DIYABC analysis.

In the first scenario (Fig. 20a) it was assumed that, at time T3, Pop 2 diverges from an ancestral population and, at time T2, it separates from Pop 1. In the second scenario, it was assumed that, at time T3, Pop 2 diverges from an ancestral population; at time T2, this

population splits originating Pop 1 that was subject to bottleneck at time T1. Finally, in the third scenario, it was assumed that Pop 2 diverges from Pop 1 at time T1.

These scenarios were tested on 3,000,000 simulations, computing the main parameters as PCA (Principal Component Analysis), logistic regression, F_{ST} values, mean genic diversity and posterior distribution of parameters.

Based on all computed parameters, the first scenario was found highly supported with most of the observed summary statistics in the range of those simulated (Figs 20a, 20b, 20c and data not shown).

Moreover, with DIYABC analysis,it was possible to plot the values of T2 and T3 on the time scale not in generations but in number of mutational events, based on the posterior distributions.

Thus, for the first scenario the values of T2 and T3 were 1.35E+002 and 5.16E+003 respectively (data not shown).

Knowing the estimated time values of mutational events and the years for a reproductive generation in *Pinus laricio* (20 years), 103,000 years ago the time T3 of separation of Pop 2 from ancestral population and 2,700 years ago the time T2 of separation of Pop 1 from Pop 2 was assigned.

5. Discussion

Microsatellites are considered the best molecular markers for population genetic studies(Agarwall et al., 2008). However, development of SSR markers for a particular species is a costly and time-consuming affair. The way to solve this is cross species amplification by using existing universal SSR primers.

Until now no studies have surveyed on the polymorphism level and distribution of natural forests of *Pinus laricio*. Indeed, previous genetic analyses were carried out under assumption of allopatric subspecies of *Pinus nigra* Arn. and, therefore, a reduced sampling has involved black pine representatives of southern Italian Apennines as well insular landscapes (Afzal-Rafii et al., 1996; Afzal-Rafii and Dodd, 2007).

In this study, genetic variability within and among seven laricio pine natural populations from Sila, Aspromonte, Etna and Corsica National and Regional Parks was tested and compared using both cpSSR and nuSSR markers, originally developed for other *Pinus* species (Vendramin et al.,1996; Soranzo et al.,1998; Auckland et al., 2002).

Pinus laricio is an anemophilous species that produces large amounts of pollen, with high dispersal potential. By using cpSSR markers, total haploid diversity (h = 0.65) and total haplotypic diversity (He = 0.94) was found high within populations. The above values were comparable to those, based on the same set of cpSSR loci, reported for several western European populations of *Pinus nigra* (Afzal-Rafii and Dodd, 2007) but also higher with respect to values found in other *Pinus* species (Dzialuk et al., 2009).

However, the high value of chloroplast haplotype diversity is a classical result for several gymnosperms, because of their long lifespans, high outcrossing rates and fecundity (Hamrick and Godt, 1989).

Substantial variations in the number of haplotypes per population as well haplotypic diversity were observed among sampled laricio pine populations, with the latter being generally higher (0.966) in the ASP stand than in the FAL population (0.917).

Conversely, AMOVA analysis provided clear evidence of a lower genetic differentiation (Φ_{PT} = 3.4%) among *Pinus laricio* stands and revealed that most of the total variation (96.6%) was expressed within populations.

Moreover, the global genetic differentiation among populations, calculated through other methods such as F- and G-statistics, retrieved values (F_{ST} = 3.4%, G_{ST} = 4.1%), very similar to those previously reported for *Pinus nigra* populations from Bulgaria (Naydenov et al., 2006) but lower than laricio pine stands from the western Europe (Afzal-Rafii and Dodd, 2007).

The low population differentiation found in my study could be explained by the human, also illegal, activities (deforestation and/or forest fires) that have decreased the father trees number (Young et al., 1996) or, alternatively, consistent with the hypothesis of extensive pollen exchange that homogenizes populations, over a long period of time, as also confirmed by the value of the number of migrants per generation (Nm = 14.21).

Interestingly, the highest Φ_{PT} value (5.3%) among populations was found when the *Pinus laricio* population of Sila's plateau was compared with that of Maesano (ASP). This result suggests a clear genetic differentiation between the above cited populations very likely due to their geographic distance (about 120 Km) and that, without a continuous ecological corridor, would have reduced the dispersal potential of pollen.

The genetic differentiation observed between the population of Sila and Aspromonte was also supported by the mean value of Nei's genetic distance (0.203).

A moderate level of genetic diversity appears to be maintained among the 7 *Pinus laricio* natural populations analyzed by nuSSRs, with a mean unbiased expected eterozigosity He of 0.674, also supported by the high values (estimated over all loci and populations) of both size variants (52 alleles) number and allelic richness (Ar = 6.56).

The He value was comparable with that reported in literature for *Pinus nigra* (González-Martínez et al., 2004) and other *Pinus* species analyzed with the same nuSSR markers (Echt et al., 1999; Mariette et al., 2001). The high level of diversity detected is rather common to woody plants, mainly for some species belonging to *Pinus* genus with marked habitat preference and fragmented distribution (Hamrick et al., 1992).

In addition, the estimates of allelic richness as well size variants number appear to be highest in comparison with the homologous values obtained from my previous cpSSR analysis, thus, demonstrating that SSR markers with biparentally inheritance are much more efficient with respect to those uniparentally inherited for allelic diversity detection.

Differences between observed (mean Ho = 0.528**)** and expected (mean He = 0.674) heterozigosity levels modify the H-W equilibrium, producing positive F indices in all *Pinus laricio* populations researched in this study.

Particularly, the higher homozygous levels were found in COR stand followed by two of the four populations sampled in the Sila National Park (RON and FAL).

The deficit of heterozigosity (mean F value = 0.204, mean F_{IS} value = 0.197) could be due to undetected null alleles or natural inbreeding events. The presence of null alleles in my data set was excluded since after recalculation of their frequency, according to Chapuis and Estoup (2007), the Weir's F_{ST} index was found altered only of 0.01%.

Therefore, the observed F positive values and, thus, H-W equilibrium deviations were very likely attributed to non-random mating effects or inbreeding.

The lack of self-incompatibility mechanisms, frequently observed in several pine species where the affect of the flexible mating system, could further contribute to the significant degree of inbreeding (about 20%) found within all the seven *Pinus laricio* populations analyzed in this study.

Wind conditions and pollen release timing largely influence gene flow direction and distance and, thus, levels of outcrossing. Indeed, very turbulent air disperses pollen to notable distances promoting outcrossing (Richardson, 2000). At the other extreme, reduction or total absence of wind would cause each tree to be enveloped in selfing pollen and consequently facilitate events of inbreeding within populations, even if tiny competition occurs among embryos.

Generally, the primary effect of selfing is heterozygosity reduction. This finding is particularly evident at least in two *Pinus laricio* populations (i.e. FAL and ETN) which have thus experienced severe changes in heterozygosity over time.

Indeed, the high heterozygosity levels (negative F values -0.024 and -0.104 in FAL and ETN stands, respectively) found in juvenile trees (NV = 10-15 years olds) radically declined in adult and mature (FA and FM) tree classes, thus, indicating an improved proportion of homozygous individuals within population.

The F negative values observed in the juvenile tree class from FAL and ETN populations were according to the hypothesis of strong selection against homozygous at the early stages of life, as reported in many conifers (Eriksson et al., 1973).

Conversely, the excess of heterozygotes surprisingly found in the FA tree class from Maesano stand (ASP) (Tab. 13) could be the result of anthropic interventions.

Values of F_{ST} (0.021) and R_{ST} (0.092) estimated by AMOVA analysis under the IAM and SMM models, respectively, indicated a non significant differentiation among all the seven natural populations of *Pinus laricio*.

This means, that the populations could not be genetically distinct, though the higher intra-population differentiation levels observed.

The lowest values of genetic distance were observed between RON and TRE stands (F_{ST} = 0.003), within Sila's National Park, as well between ASP and ETN populations (0.004), on the whole demonstrating their strict genetic relationship.

This result, also supported by the high value of the number of migrants per generation Nm (11.65), could be explained by an unexpected and high dispersal potential of seeds in this species resulting in high levels of gene flow between populations over a long period.

Although the low inter-population genetic differentiation was expected, mainly considering the relatively small geographic distances among the sampled laricio pine populations, however, a slight population structuring was evident.

The topmost dissimilarities were found in the comparison between the Corsican population with those from Etna, Aspromonte and Sila (F_{ST} values ranging from 0.024 to 0.033).

Particularly, the highest value of genetic distance (0.246), estimated according to Nei (1978), was obtained when *Pinus laricio* populations from COR and FAL were compared, very likely linked to the ecological barrier presence, that is, the Mediterranean Sea, that reduces gene exchanges.

In *Pinus laricio*, observed F_{ST} and R_{ST} values as well G_{ST} estimator (0.031) were found more o less in agreement with those previously reported for other pine species

surveyed both with allozyme and analogous SSR methods (Hamrick et al., 1992; Afzal-Rafii and Dodd, 2007; Eveno et al., 2008; Bucci et al., 2007: Provan et al., 1998; Robledo-Arnuncio et al., 2005).

As revealed by the Mantel results, the different pattern of genetic differentiation observed among *Pinus laricio* populations was independent of their geographical distribution, as also confirmed by the same test carried out using cpSSR data. The lack of a clear correlation between genetic and geographic distances among all sampled *Pinus laricio* stands very likely was due to the limited natural area of distribution of this species.

The above result was further confirmed by spatial autocorrelation analysis (carried out also with cpSSR markers), showing a random distribution of individuals in each population of laricio pine investigated in this study.

Although the "isolation by distance" hypothesis was rejected, however, PCA results, using both the cpSSR and nuSSR markers, showed interesting (even if slightly different) grouping of *Pinus laricio* stands, thus, highlighting the astonishing genetic peculiarity of the three relatively small geographic distanced Sila, Aspromonte and Corsica areas, here investigated.

Bayesian analyses (TESS and/or GENELAND), performed on the cpSSR and nuSSR data sets, further separated the seven natural populations of *Pinus laricio* in two main clusters: in the first Sila's populations were included, while the second cluster grouped the remaining populations of Maesano (ASP), Linguaglossa (ETN) and Restonica (COR).

STRUCTURE analysis supports and enhances both TESS and GENELAND results, since it allows to include ASP, ETN and COR populations once more in a same cluster and to further subdivide the *Pinus laricio* populations located on the Sila's plateau.

Particularly, the FAL stand appeared to be genetically distinct from all others, although a certain level of admixture was evident within the TRE, RON and SAL stands.

In addition, the most significant genetic boundary, identified by Monmonier's algorithm, unambiguously confirmed the FAL separation from the other laricio pine Sila's populations (Fig. 17).

It is noteworthy that among endemic populations of the Sila massif**,** constituting priority habitats under the EU Natura 2000 directive, the **"**Giant Pines of Sila" Natural Reserve**,** also known as the "*Fallistro Giants"*, is a very truly ancient pine forest.

The Fallistro Reserve, though covers only 5 ha, exceptionally includes about uneven-aged 1500 trees, with more than 40 individuals up to 350 years old (Avolio and Ciancio, 1985). The reserve is left to natural evolution as the local conditions favor the permanence of *Pinus laricio*, which ages and regenerates abundantly without human interventions.

The results obtained by several Bayesian approaches, for the first time, support the genetic distinctiveness of Fallistro Natural Reserve with respect to all other *Pinus laricio* populations.

The overall population structuring observed in this study could be explained considering the geological origin of natural area of laricio pine distribution.

It is well known that during the late Pliocene and Pleistocene, a region between the Balearic islands and Provence, belonging to the Alpine mountain range, began to fracture and sink.

This region, that geologists call the *Sardinian-Corsican block* and that was made up of Corsica, Sardinia and small parts of Sicily and southern-central Calabria, began to move

away towards the East, and so rotating anti-clock wise, a new sea was opened (the Tirrenia sea) (Fig. 20).

This finding seems to clarify the genetic affinities that I observed between Corsica, Sicily and southern Calabria (ASP) populations, further suggesting their common origin from a ancestral population (maybe *Pinus nigra* s.l.) which, later, may have rapidly expanded its range, thus, originating the Sila's populations.

As derived from stratigraphical and microtectonic investigations, in late Pliocene-Pleistocene times, the so-called Calabrian arc was block-segmented by several basins (i.e. lower Crati, Catanzaro and Siderno basins) (Tansi et al., 2007) (Fig. 21). Climate warming has produced sea retreat and subsequent formation of territorial linkages between southern (Aspromonte) and central (Sila) Calabrian regions, thus, improving pine species distribution.

Simply local climatic features, ecological barriers (i.e. mountains) as well anthropic interventions over time might have contributed to genetic differentiation of Sila's stands and, also, to genetic uniqueness of Fallistro Natural Reserve through gene flow limitation among genetically identical laricio pine groups.

DIYABC results also reject the hypothesis regarding the possible derivation of Sila's populations from different glacial refugia (Médail and Diadema, 2009), indeed, showing that COR, ETN and ASP populations evolve, before the last glaciation, from a single ancient population and that Sila's population separation occurs only later (after Pleistocene glaciation), very likely due to migration and genetic drift.

6. Conclusion

This work represents the first detailed genetic characterization of *Pinus laricio* populations within their native range, using the SSR methodology.

Three nuclear and three chloroplast microsatellite markers provided sufficient resolution to estimate genetic diversity, structure and relationships among laricio pine populations from Sila, Aspromonte, Etna and Corsica areas. As expected, within-population variation was high compared to among-population variation.

The relatively weak differentiation level among populations, measured with different efficacy by cpSSR and nuSSR markers (3.4% and 2.1% respectively), confirmed the important role of pollen dispersal in mitigating habitat fragmentation effects.

High levels of seed and pollen flow over long distances was the possible cause for mixed populations, while a low admixture proportion was observed only in the FAL populations, which is likely to exchange pollen only with neighboring populations.

Furthermore, by using both the two SSR markers, a clear genetic discontinuity was identified subdividing the investigated laricio pine populations into two major spatially and genetically distinct groups, which expansion/recolonisation pattern, after Pleistocene glaciations, can not be predicted certainly.

The data reported in this study, thus, not only improve the level of knowledge about the gene pool of *Pinus laricio* populations, but also provide valuable information for planning appropriate management or silvicultural strategies for the conservation and safeguard of this remarkable and, so far, neglected genetic resource particularly of Calabrian territory.

FIGURES AND TABLES

Fig. 1. The Last Glacial Maximum in Europe (from Adams and Faure, 1997).

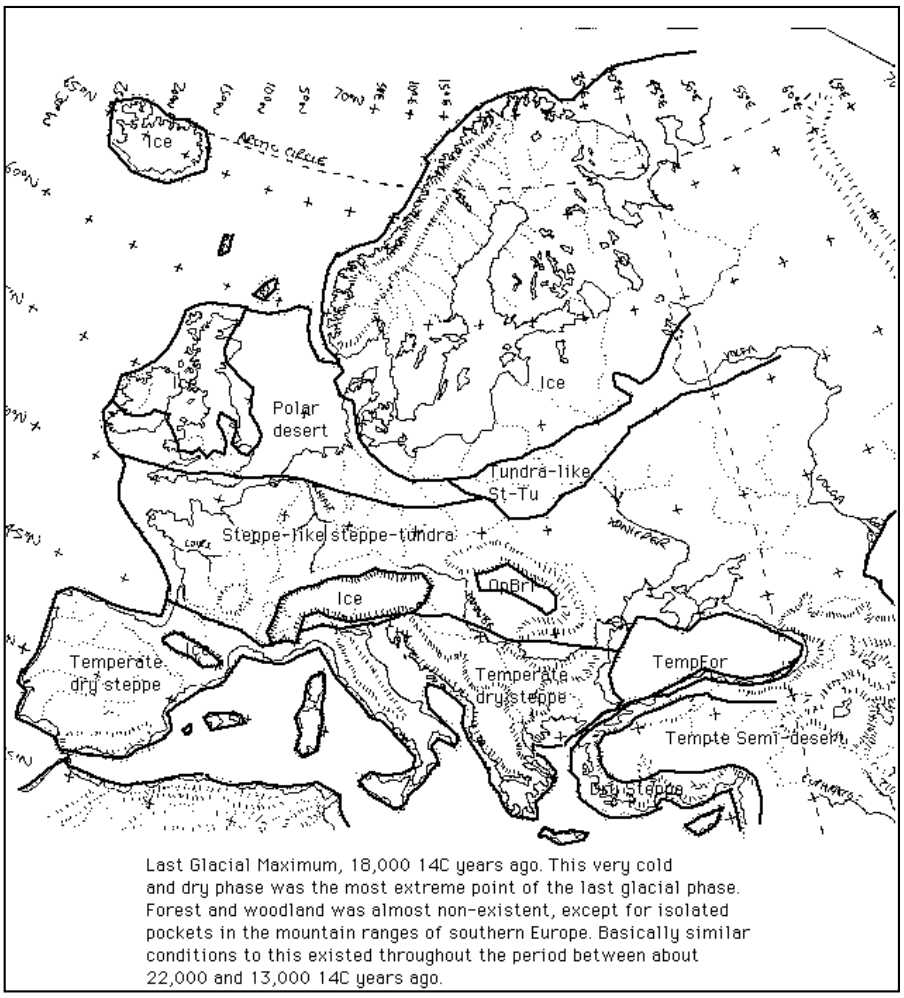


Fig. 2. Geographical distribution of the 52 putative refugia within the Mediterranean Basin, according to Médail and Diadema (2009).

Fig. 3. Geographic distribution of *Pinus nigra* subspecies according to Quézel and Médail (2003): 1 = ssp.mauritanica; 2 = ssp. salzmannii; 3 = ssp.laricio; 4 = ssp. nigra; 5 = ssp. pallasiana; 6 = ssp. dalmatica.

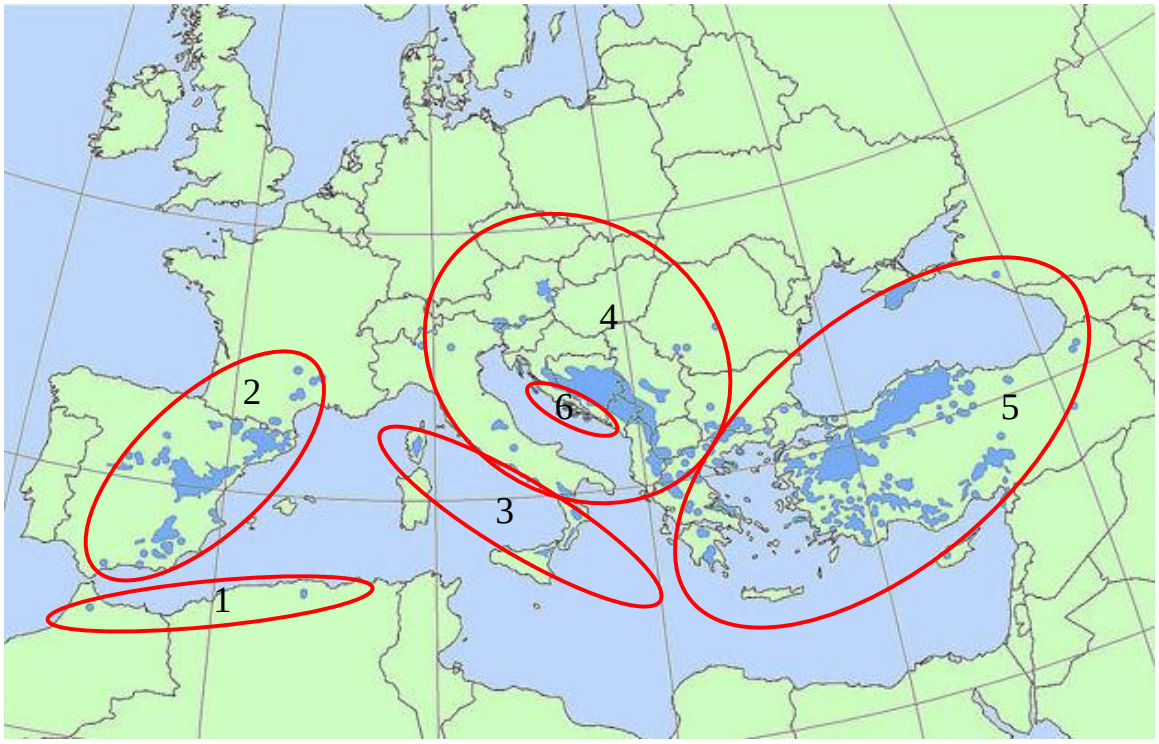


Fig. 4. *Pinus laricio* Poiret of Fallistro's Giants Biogenetical Reserve.

Fig. 5. Sampling locations of *Pinus laricio* natural populations analyzed in this study.

Fig. 6. Relationship between genetic and geographic distances among the populations of *Pinus laricio* included in this study. The line is the regression fitted to the data. *X*-axis geographic distance (logarithmic scale); *Y*-axis pairwise population linearized PhiPT values.

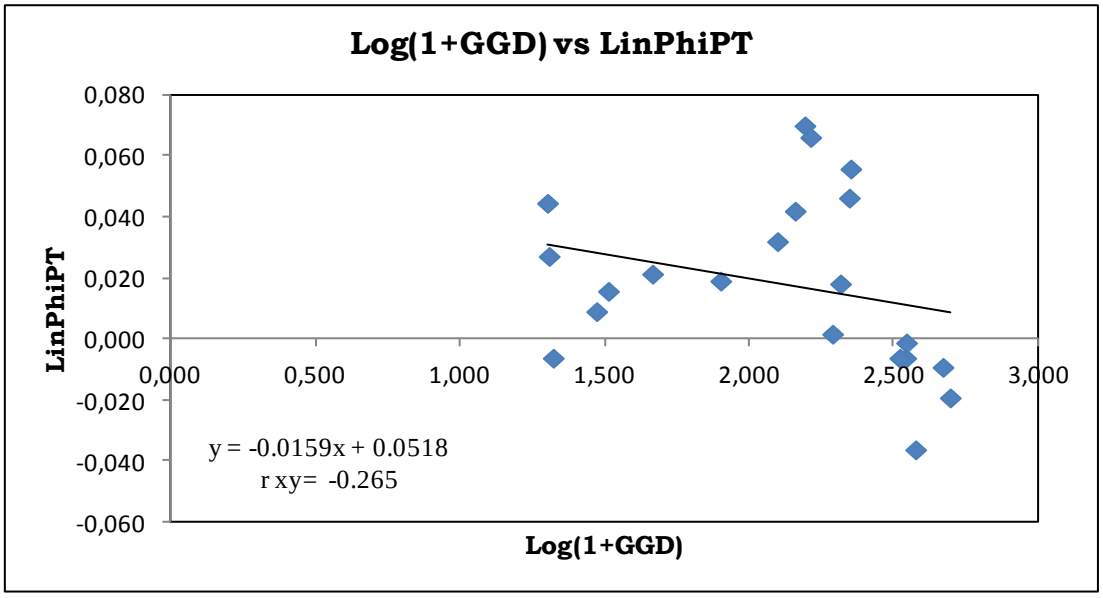


Fig. 7. Results of Spatial Autocorrelation Analysis with Even Distance Classes option. In picture r = autocorrelation coefficient (blue line). U and L were Upper and Lower limits of 95% confidence interval (red dashed line).

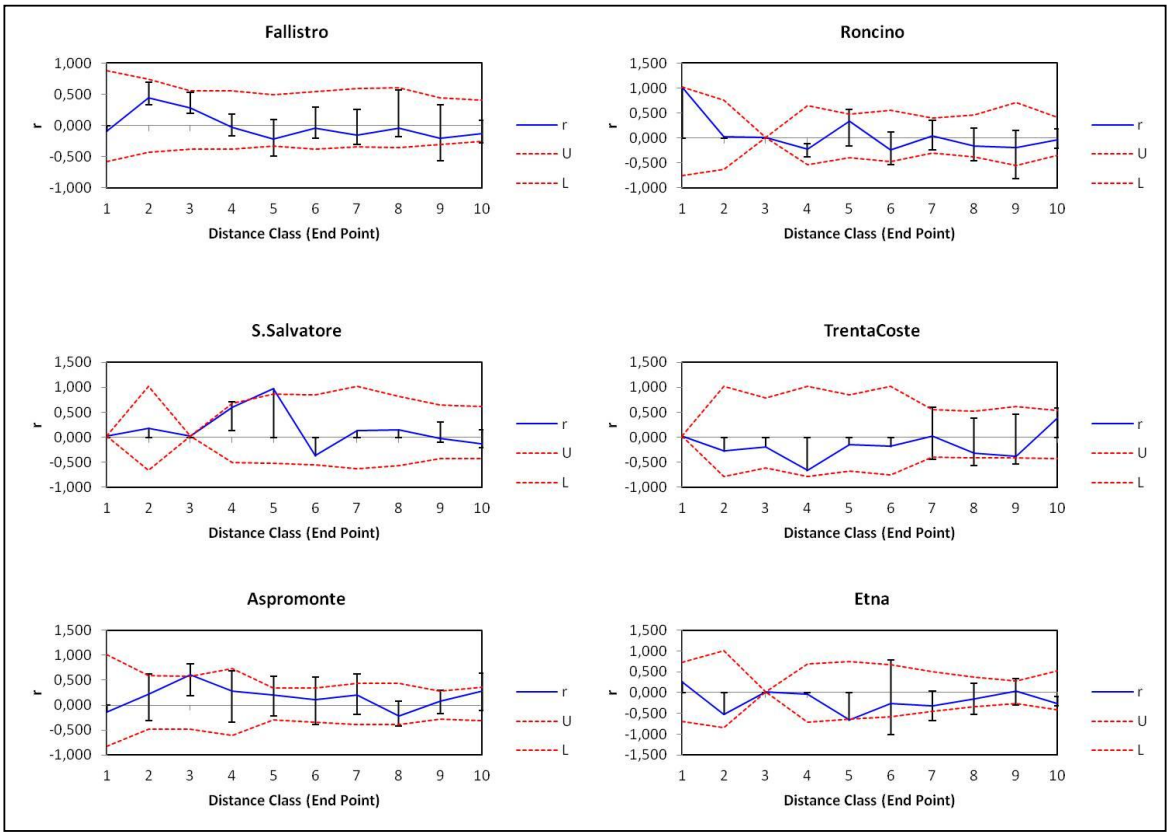


Fig. 8. Results of Spatial Autocorrelation Analysis with Even Sample Sizes option. In picture r = autocorrelation coefficient (blue line). U and L were Upper and Lower limits of 95% confidence interval (red dashed line).

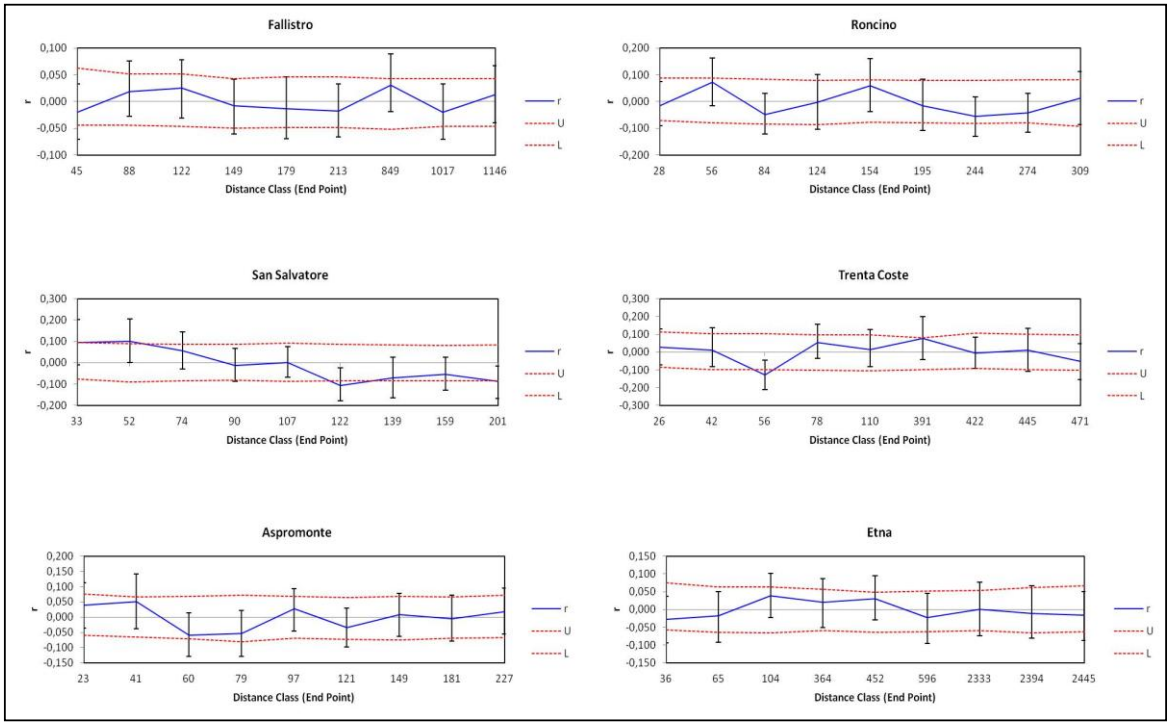


Fig 9. Principal Coordinates Analysis (PCA) plot of the 7 *Pinus laricio* populations based on the first two principal coordinates (coord. 1 = 73.24 % and coord. 2 = 13.82%). Population names are abbreviated as in Table 2.

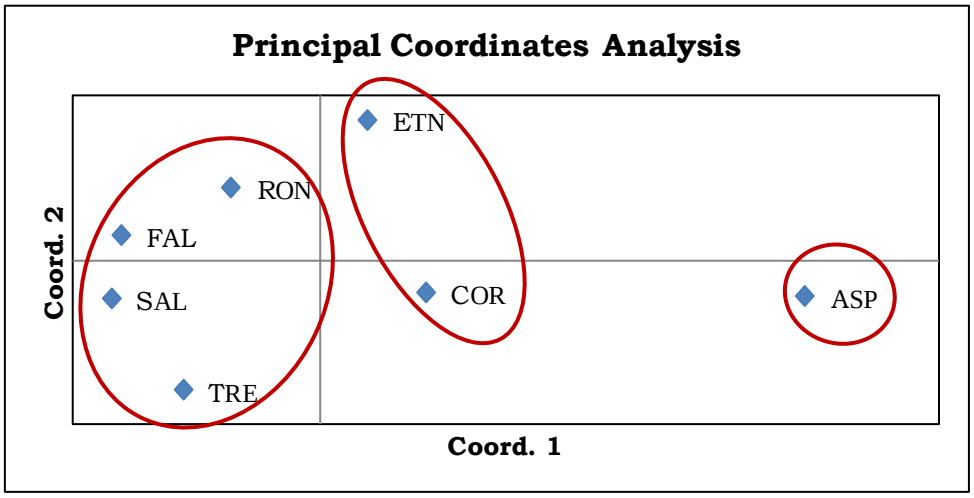


Fig. 10. The results of the TESS analysis for cpSSR markers at Kmax = 2 are shown as a barplot of the individual ancestry fraction for each tree. Each individual is represented by a thin vertical line divided into two coloured segments that represent the proportion of membership in each cluster. Black lines separate individuals from different populations

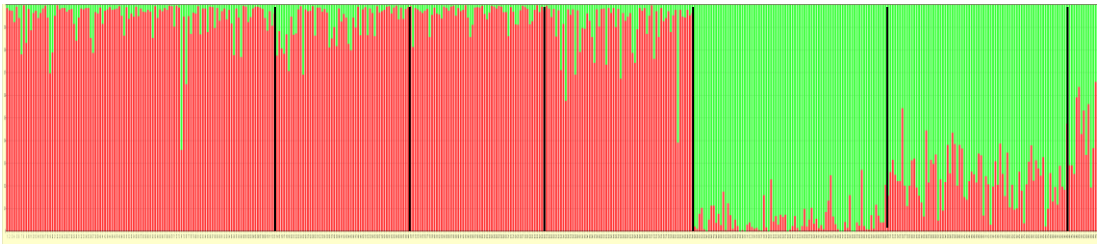
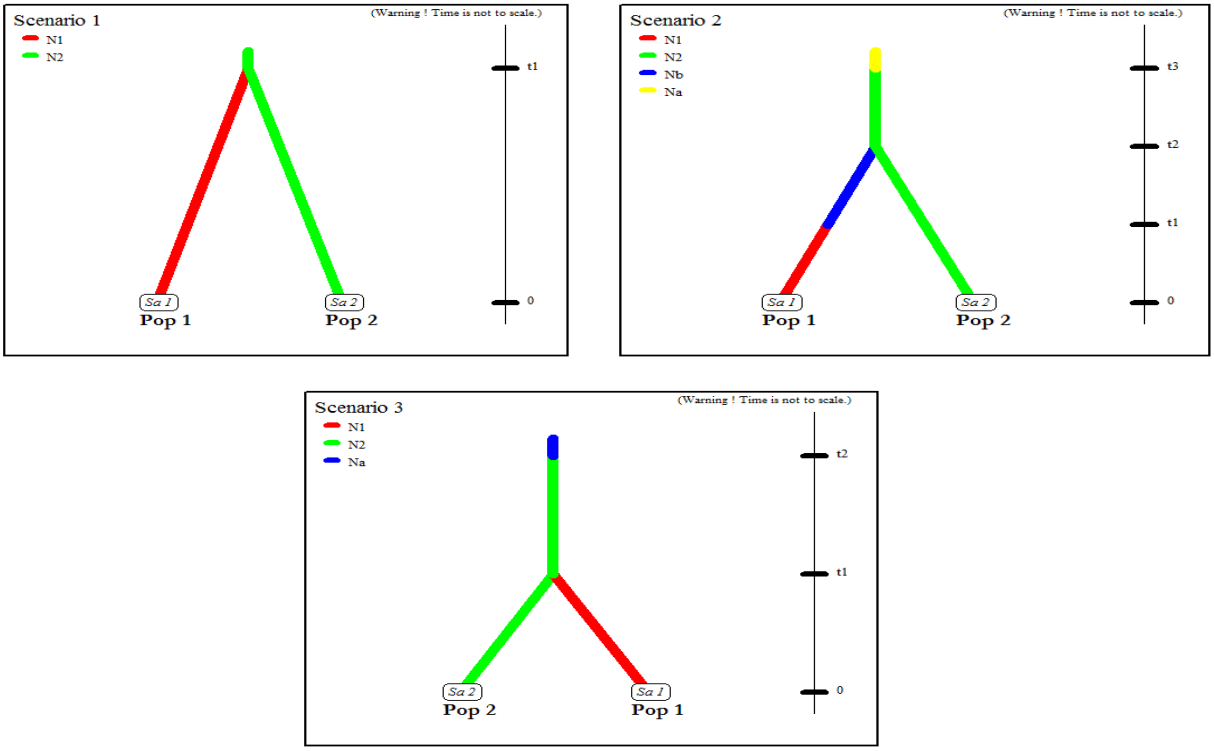
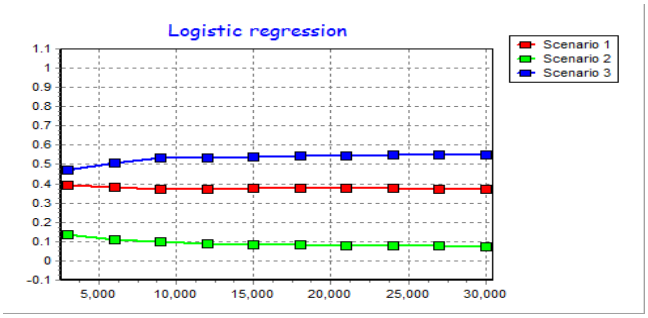


Fig. 11. Results of DIYABC analysis with cpSSR. Picture a) represented the scenarios hypothesized b) the logistic regression and c) the PCA analysis.

a)



b)



c)

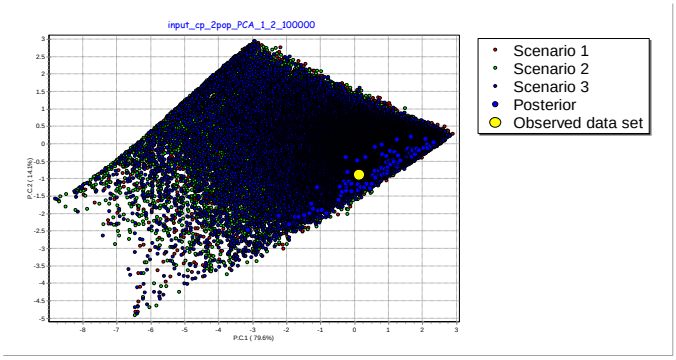


Fig. 12. Relationship between genetic and geographic distances among the populations of *Pinus laricio* included in this study. The line is the regression fitted to the data. *X*-axis geographic distance (logarithmic scale); *Y*-axis pairwise matrix [Fst/(1-Fst)].

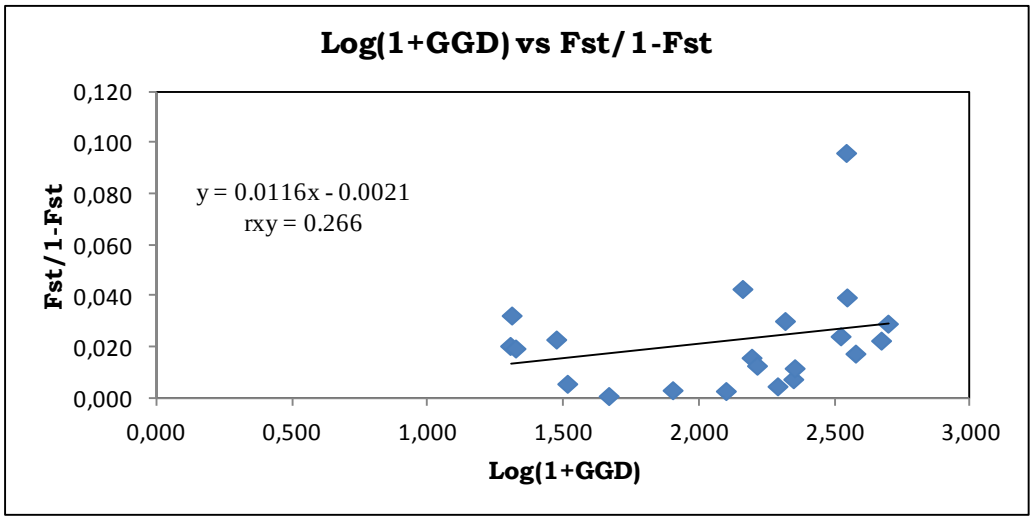


Fig. 13. Results of Spatial Autocorrelation Analysis with Even Distance Classes option. In picture r = autocorrelation coefficient (blue line). U and L were Upper and Lower limits of 95% confidence interval (red dashed line).

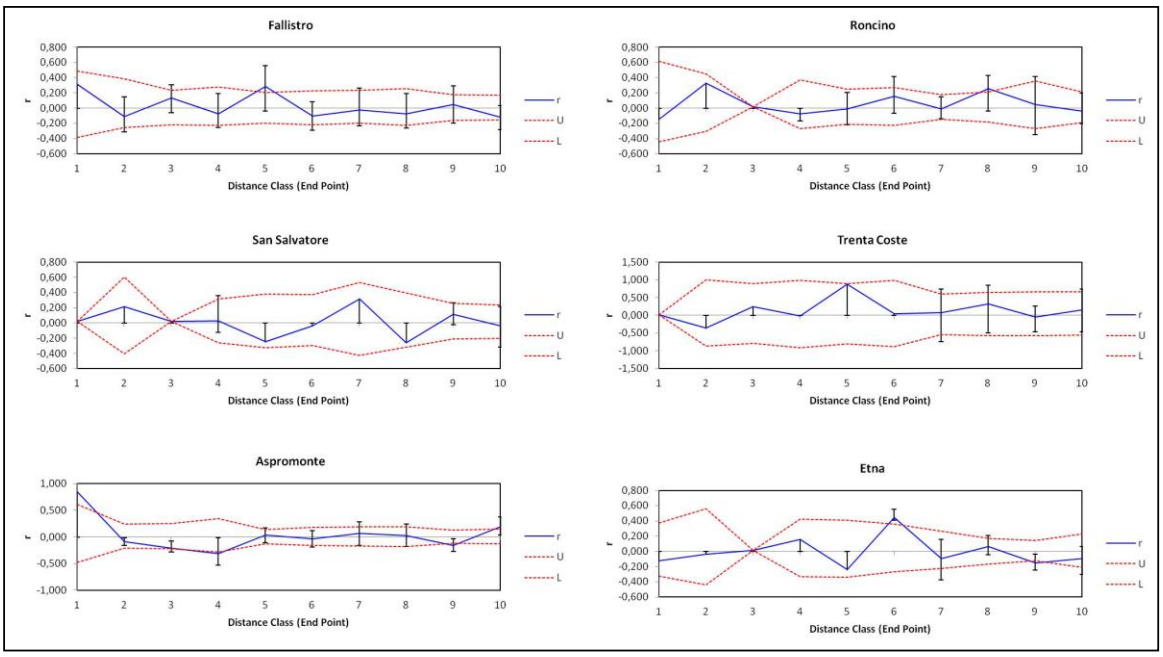


Fig. 14. Results of Spatial Autocorrelation Analysis with Even Sample Sizes option. In picture r = autocorrelation coefficient (blue line). U and L were Upper and Lower limits of 95% confidence interval (red dashed line)

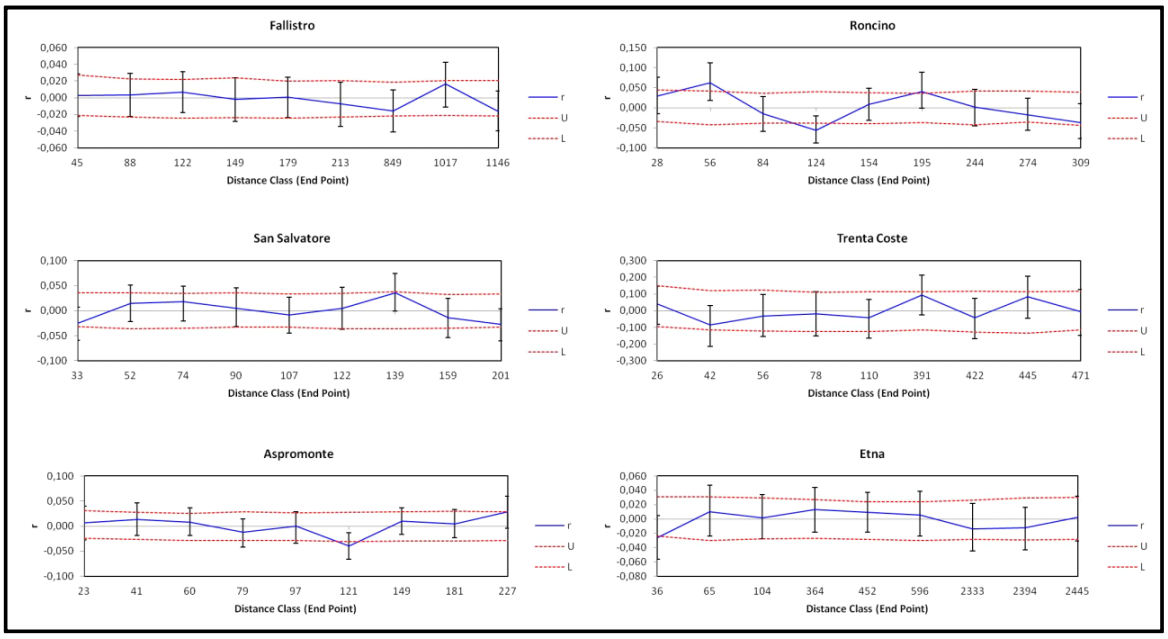


Fig. 15. Principal Coordinates Analysis (PCA) plot of the 7 *Pinus laricio* populations based on the first twoprincipal coordinates (coord. 1 = 65.92 % and coord. 2 = 20.48%).

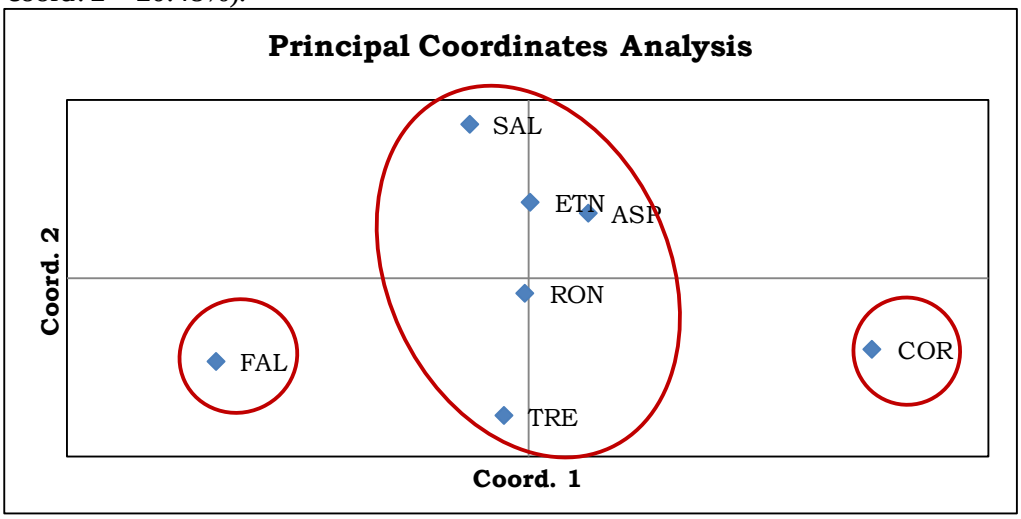


Fig. 16. The results of the TESS analysis for nuSSR markers at Kmax = 2 are shown as a barplot of the individual ancestry fraction for each individual. Each individual is represented by a thin vertical line divided into two coloured segments that represent the proportion of membership in each cluster. Black lines separate individuals from different populations.

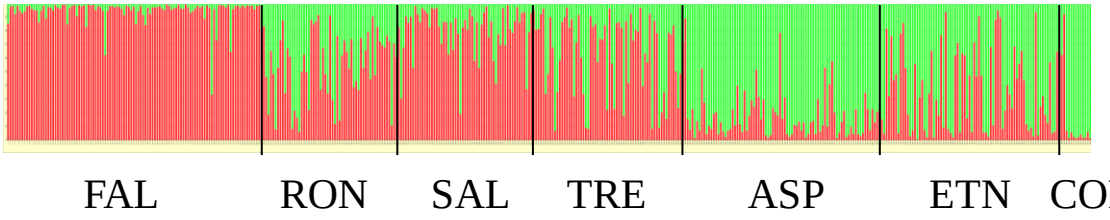
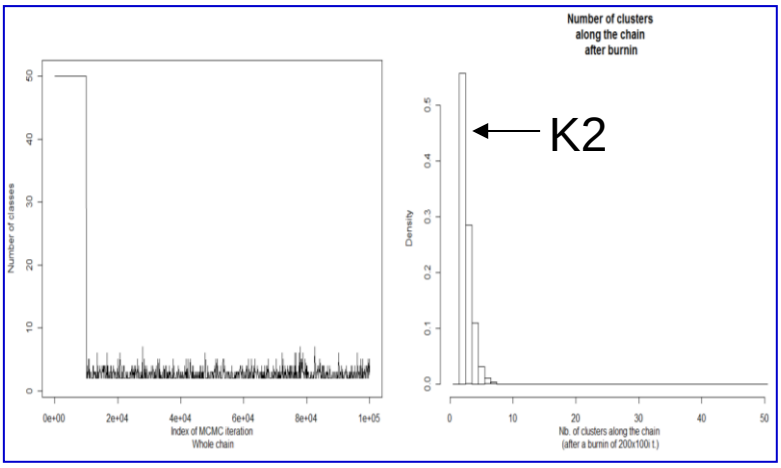


Fig 17. a) GENELAND result showing the number of clusters along the chain (after burn-in) for the seven populations of *Pinus laricio* analyzed in this study. **b**) The Geneland result was reworked graphically with program Mapinfo to verify the geographical organization of the groups identified in this analysis.

a)



b)

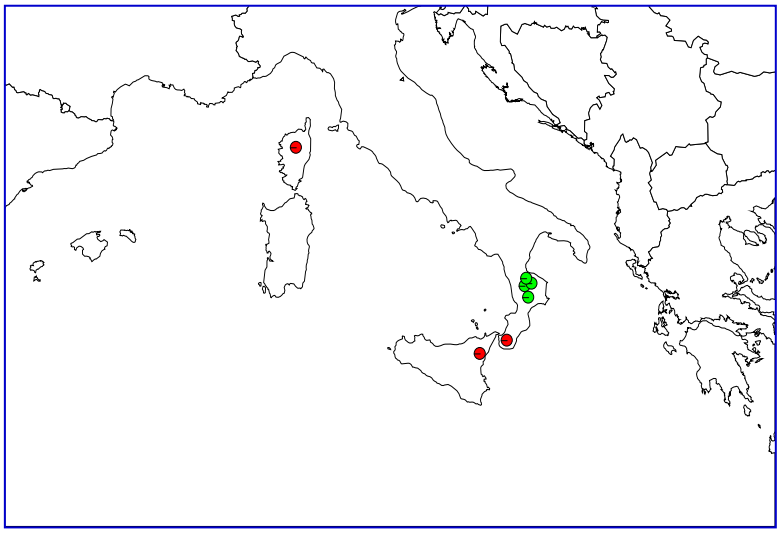


Fig. 18. Primary genetic discontinuity identified in the FAL population of *Pinus laricio* and using Monmonier's algorithm implemented in BARRIER program (Manni et al., 2004) .

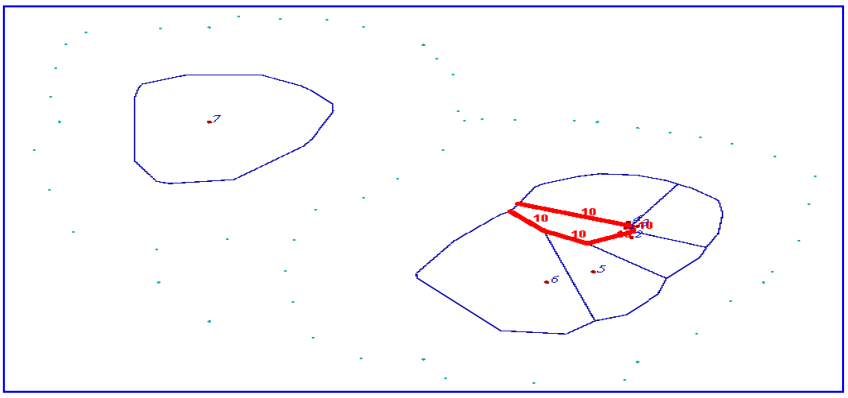


Fig. 19. STRUCTURE analysis results: **a**) the relationship between K (number of inferred clusters) and DeltaK. **b**) the probability of belonging to each of the three inferred clusters, according to the method by Evanno et al. (2005) for each of the individuals.

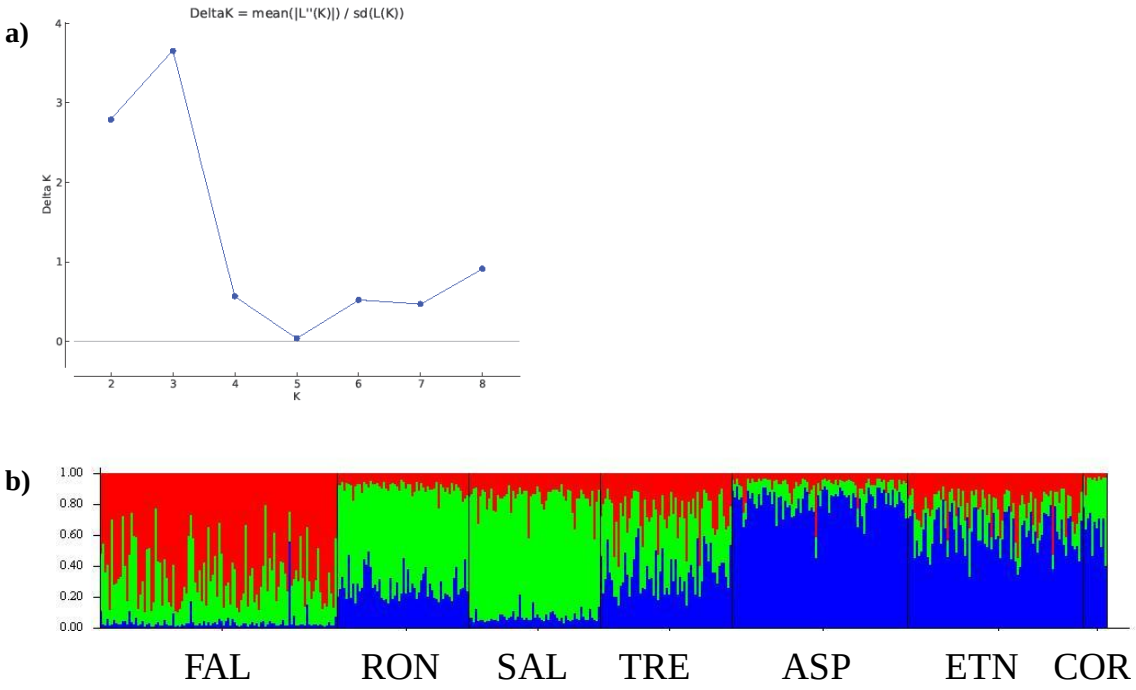
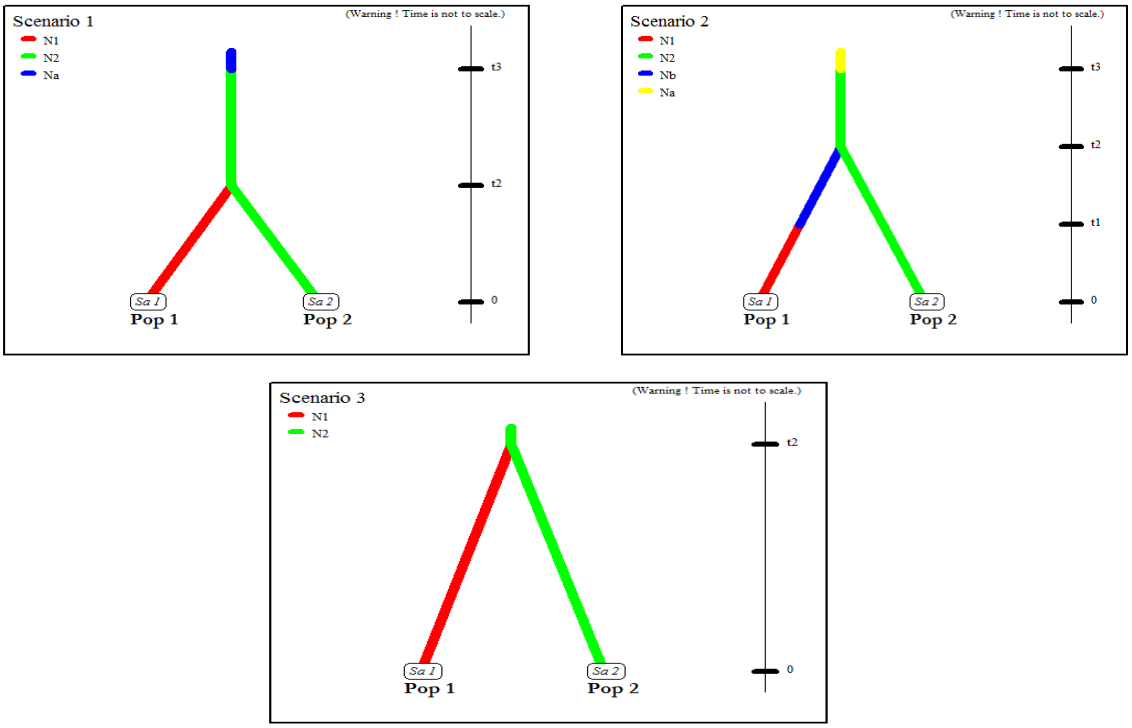
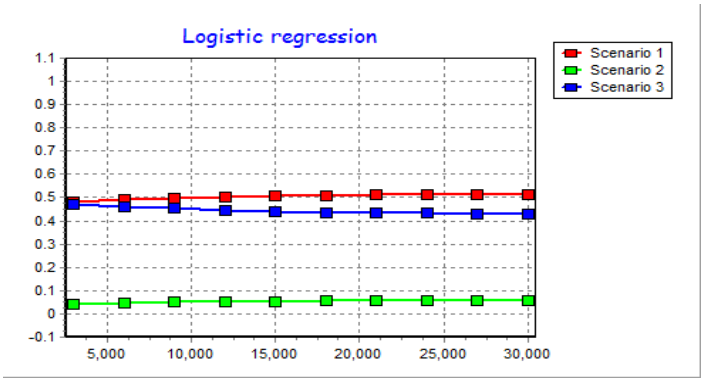


Fig. 20. Results of DIYABC analysis with nuSSR. Picture a) represented the scenarios hypothesized b) the logistic regression and c) the PCA analysis.

a)



b)



c)

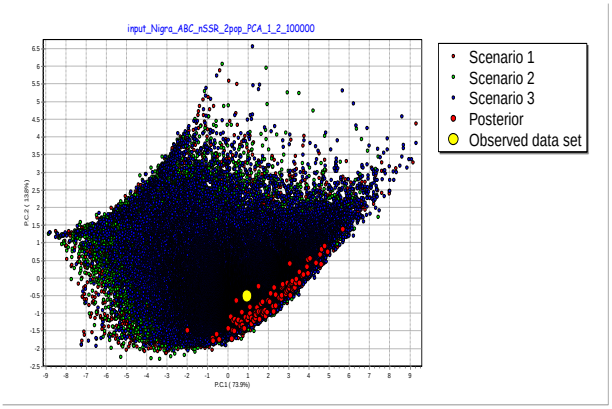


Fig. 21. Geological map of the Central Calabria area with geological section on bottom (according Van Dijk and Scheepers, 1995 and Van Dijk et al., 2000).

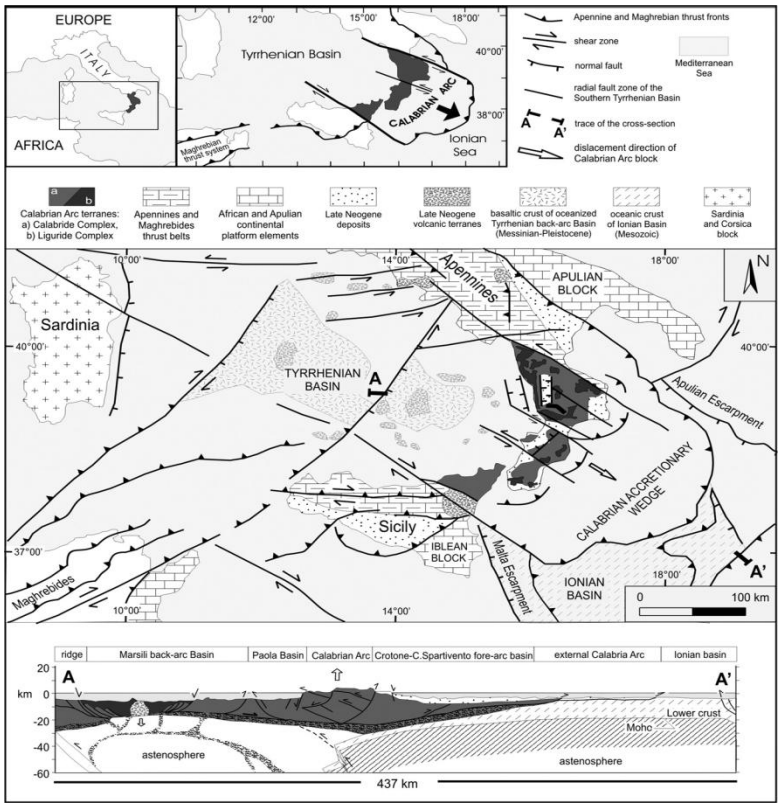
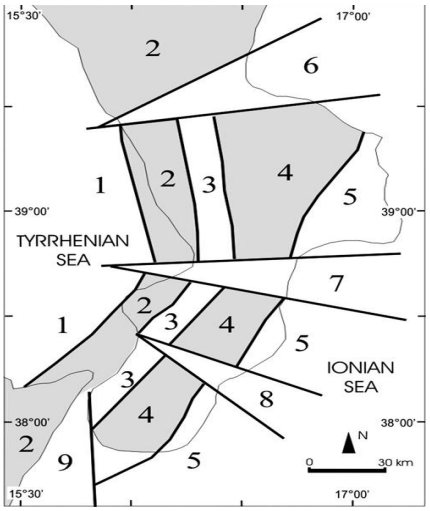


Fig. 22. Plio-Pleistocene schematic block-segmentation map of the Calabrian Arc. Black lines indicate the main faults. (1) Paola and Gioia Tauro peri-Tyrrhenian basins; (2) Pollino Massif. Coastal Chain. Capo Vaticano and Mt. Peloritani highs; (3) Crati and Mesima basins; (4) Sila. Serre and Aspromonte highs; (5) Crotone-Capo Spartivento peri-Ionian basins; (6) Sibari basin; (7) Catanzaro basin; (8) Siderno basin; (9) Messina basin.



Tab. 1. Comparison of various aspects of frequently used molecular marker techniques (Agarwall et al. 2008).

Molecular marker	Abundance	Reproducibility	Degree of polymorphism	Locus specificity	Quantity of DNA required	Major application
RFLP	High	High	Medium	Yes	High	Physical mapping
RAPD	High	Low	Medium	No	Low	Gene tagging
SSR	Medium	Medium	Medium	No	Low	Genetic diversity
SSCP	Low	Medium	Low	Yes	Low	SNP mapping
CAPS	Low	High	Low	Yes	Low	Allelic diversity
SCAR	Low	High	Medium	Yes	Low	Gene tagging and physical mapping
AFLP	High	High	Medium	No	Medium	Gene tagging
IRAP/REMAP	High	High	Medium	Yes	Low	Genetic diversity
RAMPO	Medium	Medium	Medium	Yes	Low	Genetic diversity

Tab. 2. Details of site characteristics of *Pinus laricio* populations sampled in this study.

Population/Code	N	Mountain ranges	Latitude (N)	Longitude (E)	Altitude (m a.s.l.)
Fallistro/FAL	108	Sila	39°19’	16°28’	1,410
Roncino/RON	60	Sila	39°04’	16°34’	1,200
San Salvatore/SAL	60	Sila	39°22’	16°41’	1,300
Trenta Coste/TRE	60	Sila	39°29’	16°31’	1,350
Maesano/ASP	80	Aspromonte	38°06’	15°54’	1,410
Linguaglossa/ETN	80	Etna	37°48’	15°04’	1,430
Restonica/COR	11	Corsica	42°18’	9°07’	1,200

N: number of samples in each population.

Tab. 3. SSR primers sequence and PCR conditions used to amplify microsatellite loci in seven *Pinus laricio* populations.

Locus	Primer Sequences (5'-3')	Tm (°C)	Repeat motif	Range of allele sizes (bp)
SPAG7.14 [1]	TTCGTAGGACTAAAAATGTGTG	54.7	$(TG)_n(AG)_n$	180-256
	CAAAGTGGATTTTGACCG	51.4		
PtTX4001 [1]	CTATTTGAGTTAAGAAGGGAGTC	57.1	$(CA)_n$	201-217
	CTGTGGGTAGCATCATC	52.8		
PtTX3107 [1]	AAACAAGCCCACATCGTCAATC	58.4	$(CAT)_n$	151-169
	TCCCCTGGATCTGAGGA	55.2		
Pt30204 [2]	TCATAGCGGAAGATCCTCTTT	50.0	$(A)_n(G)_n$	138-149
	CGGATTGATCCTAACCATACC	58.3		
Pt36480 [2]	TTTTGGCTTACAAAATAAAAGAGG	58.1	$(T)_n$	
	AAATTCCTAAAGAAGGAAGAGCA	57.8		
Pt45002 [2]	AAGTTGGATTTTACCCAGGTG	58.0	$(T)_n$	
	GAACAAGAGGATTTTTTCTCATACA	58.0		
Pt71936 [2]	TTCATTGGAAATACACTAGCCC	58.1	$(T)_n$	145-150
	AAAACCGTACATGAGATTCCC	57.9		
Pt87268 [2]	GCCAGGGAAAATCGTAGG	58.1	$(T)_n$	164-173
	AGACGATTAGACATCCAACCC	58.0		
Nad3-1 [3]	TTCCCCATGAATGGAAGAAG	55.3	$(G)_n$	
	ATTGATTCGATGTAGGCATCG	55.9		

1: Auckland et al., 2002 and Soranzo et al., 1998; 2: Vendramin et al., 1996; 3:Soranzo et al., 1999.

Tab. 4. Chloroplast size variant frequencies in seven populations of *Pinus laricio* considered in this study.

Loci	Size variants	Populations							Average
		FAL	RON	SAL	TRE	ASP	ETN	COR	
Pt30204	**138**	0.000	0.000	0.033	0.000	0.000	0.000	0.000	0.005
	140	0.009	0.000	0.000	0.068	0.026	0.014	0.000	0.017
	141	0.000	0.083	0.133	0.051	0.000	0.000	0.000	0.038
	142	0.236	0.183	0.333	0.407	0.260	0.247	0.364	0.290
	143	0.415	0.417	0.317	0.339	0.143	0.288	0.273	**0.313**
	144	0.170	0.200	0.067	0.068	0.403	0.219	0.364	0.213
	145	0.038	0.050	0.033	0.034	0.078	0.192	0.000	0.061
	146	0.028	0.067	0.017	0.034	0.026	0.027	0.000	0.028
	147	0.019	0.000	0.000	0.000	0.039	0.014	0.000	0.010
	148	0.000	0.000	0.067	0.000	0.000	0.000	0.000	0.010
	149	0.085	0.000	0.000	0.000	0.026	0.000	0.000	0.016
Pt71936	**145**	0.000	0.000	0.000	0.000	0.013	0.000	0.000	0.002
	146	0.252	0.169	0.119	0.136	0.234	0.108	0.182	0.171
	147	0.579	0.559	0.627	0.542	0.455	0.716	0.545	**0.575**
	148	0.122	0.203	0.220	0.254	0.273	0.122	0.273	0.210
	149	0.047	0.034	0.034	0.051	0.013	0.027	0.000	0.029
	150	0.000	0.034	0.000	0.017	0.013	0.027	0.000	0.013
Pt87268	**164**	0.065	0.000	0.000	0.000	0.000	0.000	0.091	0.022
	165	0.028	0.073	0.100	0.136	0.117	0.050	0.000	0.072
	166	0.009	0.000	0.000	0.000	0.091	0.025	0.000	0.018
	167	0.000	0.000	0.017	0.000	0.000	0.013	0.000	0.004
	168	0.019	0.036	0.017	0.034	0.000	0.013	0.000	0.017
	169	0.000	0.018	0.033	0.000	0.091	0.075	0.000	0.031
	170	0.692	0.527	0.633	0.593	0.273	0.463	0.455	**0.519**
	171	0.140	0.309	0.200	0.136	0.195	0.238	0.364	0.226
	172	0.019	0.018	0.000	0.102	0.234	0.113	0.091	0.082
	173	0.028	0.018	0.000	0.000	0.000	0.013	0.000	0.008

In blue were underlined the most frequency size variants; in yellow were both the private and rare size variants; in green were the rare size variants.

Tab. 5. *Pinus laricio* genetic diversity statistics for cpSSR

.

Pop	Na	Ne	Ar	h	Nh	Neh	He	UHe	Nhp	Ihp
FAL	6.67	2.71	4.58	0.60	25	10.86	0.908	0.917	6	0.096
RON	6.00	3.02	4.68	0.66	28	12.15	0.918	0.935	7	0.166
SAL	6.00	2.85	4.57	0.62	23	12.57	0.920	0.936	3	0.118
TRE	5.67	2.84	4.65	0.64	29	14.57	0.931	0.948	7	0.140
ASP	6.67	3.97	4.99	0.74	33	20.41	0.951	0.964	8	0.208
ETN	7.00	3.21	4.86	0.65	34	20.50	0.951	0.966	8	0.119
COR	3.33	2.75	3.33	0.63	8	6.37	0.843	0.927	-	-
Total	**5.91**	**3.05**	**4.52**	**0.65**	**25.71**	**13.92**	**0.918**	**0.942**	**39**	**0.134**
SE		0.19	0.19	0.02			0.014	0.007		

Na: numberof DifferentAlleles; Ne: numberof EffectiveAlleles; Ar: Allelic Richnessh: Haploid Diversity; Nh: number of Different Haplotypes; Neh: number of Effective Haplotypes; He: HaplotypesDiversity; UH: unbiasedHaplotypesDiversity; Nhp: numberof Private Haplotypes; Ihp: percentageof individualswith PrivateHaplotypes

Tab. 6. The most common haplotypes frequencies in seven populations of *Pinus laricio*. In green was underlined the most frequent haplotype.

Haplotypes	Code	Populations							Average
		FAL	RON	SAL	TRE	ASP	ETN	COR	
142/146/170	**h-11**	0.029	-	0.017	0.035	0.014	0.030	-	0.018
142/147/170	**h-17**	0.038	0.019	0.119	0.105	-	0.075	-	0.051
142/147/171	**h-18**	0.077	0.056	0.051	0.088	0.083	0.045	0.182	0.083
142/147/172	**h-19**	0.019	-	-	0.018	0.014	0.015	0.091	0.022
143/147/170	**h-41**	0.106	0.222	0.169	0.175	-	0.045	0.091	**0.116**
144/147/170	**h-62**	0.077	0.111	0.017	0.018	0.042	0.090	0.091	0.064
144/148/170	**h-68**	0.048	0.019	0.017	0.035	0.028	0.030	0.273	0.064
145/147/170	**h-72**	0.019	0.019	0.017	0.018	0.056	0.119	-	0.035

Tab. 7. Analysis of Molecular Variance (AMOVA) based on ϕ_{PT} among sampled *Pinus laricio* populations.

Source of variance	*df*	Variance component	Variation (%)	P
All populations				
Among populations	6	0.072	3.4	<0.001
Within populations	458	2.055	96.6	
Sila - Aspromonte				
Among populations	4	0.109	5.3	<0.001
Within populations	363	2.112	94.7	
Sila - Etna				
Among populations	4	0.054	2.7	<0.001
Within populations	363	1.955	97.3	
Sila - Corsica				
Among populations	4	0.043	2.1	<0.001
Within populations	294	1.985	97.9	
Aspromonte - Etna				
Among populations	1	0.042	1.9	<0.001
Within populations	159	2.185	98.1	
Aspromonte - Corsica				
Among populations	1	0.000	0	>0.05
Within populations	89	2.460	100	
Etna - Corsica				
Among populations	1	0.000	0	>0.05
Within populations	89	1.823	100	

Tab. 8. Pairwise population matrix of genetic distance calculated by Nei's method. The values in bold represented the maximum and minimum genetic distances.

FAL	RON	SAL	TRE	ASP	ETN	COR	Population
-							**FAL**
0.043	-						**RON**
0.048	0.043	-					**SAL**
0.056	0.064	**0.023**	-				**TRE**
0.223	0.169	**0.227**	0.191	-			**ASP**
0.078	0.053	0.068	0.084	0.127	-		**ETN**
0.105	0.063	0.101	0.100	0.081	0.080	-	**COR**

Tab. 9. Nuclear size variant frequencies in seven populations of *Pinus laricio* analyzed in this study.

Loci	Size variants	Populations							Average
		FAL	RON	SAL	TRE	ASP	ETN	COR	
SPAG7.14	**180**	-	0.008	-	-	0.019	0.019	-	0.007
	182	-	-	-	-	0.013	0.006	0.091	0.016
	184	-	0.008	-	-	-	0.006	-	0.002
	186	0.032	0.067	0.033	0.167	0.038	0.025	0.091	0.065
	188	0.005	0.025	-	0.033	0.056	0.013	-	0.019
	190	0.009	0.075	0.008	0.108	0.100	0.038	0.091	0.061
	192	0.005	0.008	0.067	0.008	0.056	0.063	-	0.030
	194	-	-	0.017	-	0.025	-	-	0.006
	196	0.046	-	0.058	0.017	0.044	0.044	-	0.030
	198	0.005	0.008	-	0.042	0.019	0.031	0.136	0.034
	200	-	-	-	0.017	0.006	0.006	-	0.004
	202	0.199	0.217	0.167	0.133	0.219	0.181	0.136	**0.179**
	204	0.028	0.050	0.025	0.025	0.019	0.013	0.091	0.036
	206	0.005	0.067	0.017	0.008	0.056	0.069	0.091	0.045
	208	0.046	0.058	0.100	0.075	0.063	0.056	-	0.057
	210	0.023	0.017	0.050	0.017	-	0.013	-	0.017
	212	0.005	0.050	0.008	0.017	-	-	-	0.011
	214	0.005	0.008	-	0.017	0.013	0.025	0.091	0.023
	216	0.009	-	0.017	-	0.013	0.025	-	0.009
	218	0.005	0.042	0.017	0.008	0.006	0.019	-	0.014
	220	0.139	0.025	0.042	0.058	0.031	0.056	-	0.050
	222	-	-	-	-	-	0.006	-	0.001
	224	-	-	-	-	-	0.006	0.091	0.014
	226	0.046	0.017	0.008	0.067	0.006	0.075	-	0.031
	228	0.014	0.008	0.000	0.008	0.013	0.038	-	0.012
	230	0.005	-	0.017	-	0.006	-	0.091	0.017
	232	-	-	0.008	-	-	-	-	0.001
	234	0.023	0.017	-	0.033	-	0.006	-	0.011
	236	0.005	0.017	-	-	0.006	-	-	0.004
	238	0.009	0.050	0.033	0.017	-	0.006	-	0.017
	240	0.079	0.100	0.142	0.058	0.044	0.031	-	0.065
	242	0.037	0.008	0.008	0.025	0.006	0.006	-	0.013
	244	-	-	-	-	-	0.006	-	0.001
	246	0.009	-	0.042	0.008	0.006	0.019	-	0.012
	248	0.014	-	0.017	0.008	0.019	-	-	0.008
	252	0.005	-	-	-	0.088	0.031	-	0.018
	254	0.181	0.033	0.100	0.025	0.013	0.063	-	0.059
	256	0.009	0.017	-	-	-	-	-	0.004
PtTX4001	**201**	0.375	0.224	0.117	0.310	0.203	0.200	0.182	0.230
	203	0.009	-	-	0.009	-	-	-	0.003
	207	0.125	0.121	0.117	0.103	0.082	0.044	0.091	0.098
	209	0.148	0.138	0.233	0.121	0.203	0.225	-	0.153
	211	0.329	0.474	0.483	0.440	0.513	0.531	0.727	**0.500**
	213	-	0.009	-	-	-	-	-	0.001
	215	-	0.009	0.008	-	-	-	-	0.002
	217	0.014	0.026	0.042	0.017	-	-	-	0.014
PtTX3107	**151**	0.060	0.058	0.050	0.017	0.056	-	-	0.034
	157	0.019	0.067	0.125	0.092	0.019	0.050	0.227	0.085
	160	0.843	0.717	0.733	0.725	0.650	0.750	0.591	**0.716**
	163	0.069	0.133	0.050	0.133	0.200	0.131	0.182	0.128
	166	-	-	-	-	0.006	0.006	-	0.002
	169	0.009	0.025	0.042	0.033	0.069	0.063	-	0.034

In blue were underlined the most frequency size variants; in yellow were both the private and rare size variants; in green were the rare size variants.

Tab. 10. *Pinus laricio* genetic diversity based on nuSSRs.

Pop	Na	Ar	Ho	He	UHe	F
FAL	13.33	6.13	0.500	0.628	0.631	0.209
RON	12.33	7.05	0.522	0.687	0.693	0.249
SAL	11.33	6.88	0.550	0.679	0.684	0.162
TRE	12.00	6.84	0.590	0.683	0.689	0.144
ASP	12.67	6.77	0.586	0.696	0.701	0.171
ETN	13.33	6.93	0.583	0.657	0.661	0.114
COR	5.33	5.33	0.364	0.631	0.661	0.383
Total	**11.48**	**6.56**	**0.528**	**0.666**	**0.674**	**0.204**
SE		0.216	0.044	0.045	0.046	0.039

Na: number of Different Alleles; Ar: Allelic Richness Ho: observed eterozigosity; He: expected eterozigosity; UHe: unbiased expected eterozigosity; F: Fixation Index.

Tab. 11. Value of F_{ST} without null allele correction. F_{ST}' with null allele correction and F_{IS} calculated per loci over *Pinus laricio* populations.

Locus	F_{ST}	F_{ST}'	F_{IS}
SPAG 7.14	0.016	0.014	0.225***
PtTX4001	0.022	0.022	0.240***
PtTX3107	0.019	0.020	0.124***
Mean	0.019	0.018	0.197***

P values: *** $P < 0.001$.

Tab. 12. Values of null allele frequencies per loci over all populations of *Pinus laricio* analyzed in this study.

Pop	Loci		
	SPAG 7.14	**PtTX4001**	**PtTX3107**
FAL	0.03864	0.15242	0.06219
RON	0.09047	0.10644	0.09750
SAL	0.11878	0.10286	0.01974
TRE	0.06091	0.02001	0.08005
ASP	0.06120	0.05496	0.11805
ETN	0.04784	0.05911	0.03071
COR	0.33654	0.20156	0.00001

Tab. 13. Values of Fixation Index (F) calculated on 7 populations of *Pinus laricio* subdivided in age classes.

Pop		FAL	RON	SAL	TRE	ASP	ETN
FS (>200 years)	**Mean**	0.197	-	-	-	0.237	0.149
	SE	0.095	-	-	-	0.025	0.056
FM (120 years)	**Mean**	0.259	0.214	0.176	0.112	0.259	0.234
	SE	0.141	0.127	0.022	0.071	0.116	0.060
FA (80 years)	**Mean**	0.271	0.273	0.174	0.238	-0.017	0.104
	SE	0.196	0.131	0.029	0.185	0.124	0.051
FG (40 years)	**Mean**	0.182	-	-	-	-	-
	SE	0.141	-	-	-	-	-
NV (10-15 years)	**Mean**	-0.024	0.248	0.108	0.063	0.137	-0.101
	SE	0.040	0.096	0.181	0.064	0.120	0.050

Tab. 14. Analysis of Molecular Variance (AMOVA) based on F_{ST} among *Pinus laricio* populations.

Source of variance	*df*	Variance component	Variation (%)	P
All populations				
Among populations	6	0.022	2.1	<0.001
Within populations	911	1.008	97.9	
Sila - Aspromonte				
Among populations	4	0.022	2.2	<0.001
Within populations	731	1.012	97.8	
Sila - Etna				
Among populations	4	0.019	1.9	<0.001
Within populations	731	0.999	98.1	
Sila - Corsica				
Among populations	4	0.024	2.4	<0.001
Within populations	593	1.001	97.6	
Aspromonte - Etna				
Among populations	1	0.004	0	0.068
Within populations	318	1.021	100	
Aspromonte - Corsica				
Among populations	1	0.000	0	0.018
Within populations	89	2.460	100	
Etna - Corsica				
Among populations	1	0.033	3.3	<0.001
Within populations	180	0.991	96.7	

Tab. 15. Pairwise population matrix of genetic distance calculated by Nei's method. The values in bold represented the maximum and minimum genetic distances.

FAL	RON	SAL	TRE	ASP	ETN	COR	Population
-							**FAL**
0.058	-						**RON**
0.076	0.036	-					**SAL**
0.053	0.025	0.066	-				**TRE**
0.097	0.027	0.056	0.049	-			**ASP**
0.068	0.027	0.032	0.041	**0.021**	-		**ETN**
0.246	0.108	0.154	0.117	0.114	0.116	-	**COR**

7. References

- **Adams** J.M., Faure H. (1997) (eds.), QEN members *Review and Atlas of Palaeovegetation: Preliminary land ecosystem maps of the world since the Last Glacial Maximum*. Oak Ridge National Laboratory, TN USA. Website: http://www.esd.ornl.gov/em/qen/adams1.ht
- **Afzal-Rafii** Z., Dodd R.S. (2007) *Chloroplast DNA supports a hypothesis of glacial refugia over postglacial recolonization in disjunct populations of black pine (Pinus nigra) in western Europe*. Molecular Ecology 16: 723-736.
- **Afzal-Rafii** Z., Dodd R.S., Zavarin E. (1996) *Genetic Diversity in foliar terpenoids among natural populations of european black pine*. Biochemical Systematics and Ecology 24: 325-339.
- **Agarwal** M., Shrivastava N., Padh H. (2008) *Advances in molecular marker techniques and their applications in plant sciences*. Plant Cell Reports 27: 617–631.
- **Auckland** L.D., Bui T., Zhou Y., Shepherd M., Williams C.G. (2002) *Conifer microsatellite handbook*. Corporate Press, Raleigh, N.C.
- **Avolio** S. (2003) *Schede monografiche delle specie forestali calabresi peculiari: Pino laricio*. In: Foreste di Calabria, Grafiche Ghiani (Cagliari): 252-254.
- **Avolio** S., Ciancio O. (1985) *I giganti della Sila*. Annali ISSEL XVI: 373-421.
- **Bachmann** K. (2001) *Evolution and the genetic analysis of populations:1950-2000*. Taxon 50: 7-45.
- **Belkhir** K., Borsa P., Chikhi L., Raufaste N., Bonhomme F. (2001) *GENETIX vs. 4.04, logiciel sous Windows ™ pour la génétique des populations*. Laboratoire Génome, Populations, Interactions, CNRS UMR 5000, Université de Montpellier II, Montpellier, France.
- **Bogunić** F., Siljak-Yakovlev S., Muratović E., Ballian D. (2011) *Different karyotipe patterns among allopatric Pinus nigra (Pinaceae) populations revealed by molecular cytogenetics*. Plant Biology 13: 194-200.
- **Bucci** G., González-Martínez S.C., Le Provost G., Plomion C., Ribeiro M.M., Sebastiani F., Alía R., Vendramin G.G. (2007) *Range-wide phylogeography and gene zones in Pinus pinaster Ait. revealed by chloroplast microsatellite markers*. Molecular Ecology 16: 2137-2153.

- **Cesca** G., Peruzzi L. (2002) *Pinus laricio Poir. and P. leucodermis Antoine: karyotipe analysis in Calabrian populations (Southern Italy)*. Caryologia 55: 21-25.
- **Chapuis** M.P., Estoup A. (2007) *Microsatellite null alleles and estimation of population differentiation*. Molecular Biology and Evolution 24: 621-631.
- **Chen** C., Durand E., Forbes F., François O. (2007) *Bayesian clustering algorithms ascertaining spatial population structure: a new computer program and a comparison study*. Molecular Ecology Notes 7: 747-756.
- **Chiang** Y.C., Hung K.H., Schaal B.A., Ge X.J., Hsu T.W., Chiang T.Y. (2006) *Contrasting phylogeographical patterns between mainland and island taxa of the Pinus luchuensis complex*. Molecular Ecology 15: 765-779.
- **Cornuet** J.M., Ravigné V. Estoup A. (2010) *Inference on population history and model checking using DNA sequence and microsatellite data with the sofware DIYABC (v1.0)*. BMC Bioinformatics 11: 401-411.
- **Dong** J., Wagner D.B. (1994) *Paternally inherited chloroplast polymorphism in Pinus: estimation of diversity and population subdivision, and tests of disequilibrium with a maternally inherited mitochondrial polymorphism*. Genetics 136: 1187-1194.
- **Durand** E., Jay F., Gaggiotti O.E., Francois O. (2009) *Spatial inference of admixture proportions and secondary contact zones*. Molecular Biology and Evolution 26: 1963-1973.
- **Dzialuk** A., Muchewicz E., Boratyński A., Montserrat J.M., Boratyńska K., Burczyk J. (2009) *Genetic variation of Pinus uncinata (Pinaceae) in the Pyrenees determined with cpSSR markers*. Plant Systematics and Evolution 277: 197-205.
- **Echt** C.S., Vendramin G.G., Nelson C.D., Marquardt P. (1999) *Microsatellite DNA as shared genetic markers among conifer species*. Canadian Journal of Forest Research 29: 365-371.
- **El Mousadik** A., Petit R.J. (1996) *High level of genetic differentiation for allelic richness among populations of the argan tree [Argania spinosa (L.) Skeels] endemic to Morocco*. Theoretical and Applied Genetics 92: 832-839 .
- **Eriksson** G., Schelander B., Akebrand V. (1973) *Inbreeding depression in and old experimental plantation of Picea abies*. Hereditas 73: 160-166.
- **Evanno** G., Regnaut S., Goudet J. (2005) *Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study*. Molecular Ecology 14: 2611-2620.
- **Eveno** E., Collada C., Guevara M.A., Léger V., Soto A., Díaz L., Léger P., González-Martínez S.C., Cervera M.T., Plomion C., Garnier-Gére P.H. (2008) *Contrasting patterns*

of selection at Pinus pinaster Ait. drought stress candidate genes as revealed by genetic differentiation analyses. Molecular Biology and Evolution 25: 417-43.

- **Excoffier** L., Estoup A., Cornuet J.M. (2005) *Bayesian analysis of an admixture model with mutations and arbitrarily linked markers*. Genetics 169: 1727-1738.
- **Excoffier** L., Smouse P.E., Quattro J.M. (1992) *Analysis of molecular variance inferred from metric distances among DNA haplotypes: Application to human mitocondrial DNA restriction sites*. Genetics 131: 479-491.
- **Fénart** S., Austerlitz F., Cuguen J., Arnaud J.F. (2007) *Long distance pollen-mediated gene flow at a landscape level: the weed beet as a case study*. Molecular Ecology 16: 3801-3813.
- **Fineschi** S. (1984) *Determination of the origin of an isolated group of trees of Pinus nigra through enzyme gene markers*. Silvae Genetica 33: 169-172.
- **Fisher** R.A. (1935) *The logic of inductive inference (with discussion)*. Journal of the Royal Statistical Society 98: 39-82.
- **Frenzel** B. (1973) *Climatic Fluctuations of the Ice Age*. The press of Case western Reserve University, Cleveland and London.
- **Gaussen** H., Webb D.A., Heywood H.V. (1993) *Pinus*. In: Tutin G.H., Heywood H.V., Burges V.A., Moore D.M., Valentine D.H., Walters S.M., Webb D.A. (Eds.), Flora Europaea, vol. 1. Cambridge University Press, Cambridge, UK: pp 40-44.
- **González-Martínez** S.C., Robledo-Arnuncio J.J., Collada C., Díaz A., Williams C.G., Alía R., Cervera M.T. (2004) *Cross-amplification and sequence variation of microsatellite loci in Eurasian hard pines*. Theoretical and Applied Genetics 109: 103-111.
- **Goudet** J. (2001) *FSTAT, a program to estimate and test gene diversities and fixation indices, version 2.9.3.2*. Website: http://www2.unil.ch/popgen/softwares/fstat.htm.
- **Guillot** G., Mortier F., Estoup A. (2005) *Geneland: a computer package for landscape genetics*. Molecular Ecology Notes 5: 712-715.
- **Hamrick** J.L. (2004) *Response of forest trees to global environmental changes*. Forest Ecology and Management 197: 323-325.
- **Hamrick** J.L., Godt M.J.W., (1989) *Allozyme diversity in plant species*. pp. 43-63 In Brown A.H.D., Clegg M.T., Kahler A.L., Weir B.S. (eds.) *Plant Population Genetics. Breeding and Genetic Resources*. Sinauer Press, Sunderland, MA.
- **Hamrick** J.L., Godt M.J.W., Sherman-Broyles S.L. (1992) *Factors influencing levels of genetic diversity in woody plant species*. New Forests 6: 95-124.

- **Heuertz** M., Teufel J., González-Martínez S.C., Soto A., Fady B., Alía R., Vendramin G.G. (2010) *Geography determines genetic relationships between species of mountain pine (Pinus mugo complex) in western Europe*. Journal of Biogeography 37: 541-556.
- **Hewitt** G.M. (1996) *Some genetic consequences of ice ages, and their role in divergence and speciation*. Biological Journal of the Linnean Society 58: 247-276.
- **Hewitt** G.M. (2000) *The genetic legacy of the Quaternary ice ages*. Nature 405: 907-913.
- **Isajev** V., Fady B. , Semerci H., Andonovski V. (2004) *EUFORGEN Technical guidelines for genetic conservation and use for European black pine (Pinus nigra)*. International Plant Genetic Resources Institute, Rome.
- **Johansen** A.D., Latta R.G. (2003) *Mitochondrial haplotype distribution, seed dispersal and patterns of postglacial expansion of ponderosa pine*. Molecular Ecology 12: 293-298.
- **Kubo** T., Newton K.J. (2008) *Angiosperm mitochondrial genomes and mutations*. Mitochondrion 8: 5-14.
- **Latta** R.G., Linhart Y.B., Fleck D., Elliot M. (1998) *Direct and indirect estimates of seed versus pollen movement within a population of ponderosa pine*. Evolution 52: 61-67.
- **Liepelt** S., Bialozyt R., Ziegenhagen B. (2002) *Wind-dispersed pollen mediates postglacial gene flow among refugia*. Proceedings of the Nationall Academy of Sciences USA 99: 14590-14594.
- **Manni** F., Guérard E., Heyer E. (2004) *Geographic patterns of genetic, morphologic, linguistic variation: how barriers can be detected by "Monmonier's algorithm"*. Human Biology 76: 173-190.
- **Mantel** N. (1967) *The detection of disease clustering and a generalized regression approach*. Cancer Research 27: 209-220.
- **Mariette** S., Chagne D., Decroocq S., Vendramin G.G., Lalanne C., Madur D., Plomion C. (2001) *Microsatellite markers for Pinus pinaster Ait.* Annals of Forest Science 58: 203-206.
- **Médail** F., Diadema K. (2009) *Glacial refugia influence plant diversity patterns in the Mediterranean Basin*. Journal of Biogeography 36: 1333-1345.
- **Michalakis** Y., Excoffier L., (1996) *A generic estimation of population subdivision using distances between alleles with special reference for microsatellite loci*. Genetics 142: 1061-1064.
- **Milligan** B.G. (1991) *Chloroplast DNA diversity within and among populations of Trifolium pratense*. Current Genetics 19: 411-416.

- **Naydenov** K.D., Tremblay F.M., Fenton N.J., Alexandrov A. (2006) *Structure of Pinus nigra Arn. populations in Bulgaria revealed by chloroplast microsatellites and terpenes analysis: Provenance tests*. Biochemical Systematics and Ecology 34: 562-574.
- **Nei** M. (1978) *Estimation of average heterozygosity and genetic distance from a small number of individuals*. Genetics 89: 583-590.
- **Nilsson** T. (1983) *The Pleistocene: Geology and Life in the Quaternary Ice Age*. Ferdinand Enke Verlag, Stuttgart.
- **Peakall** R., Smouse P.E. (2006) *GenAlEx 6: genetic analysis in Excel. Population genetic software for teaching and research*. Molecular Ecology Notes 6: 288-295.
- **Petit** R.J., Aguinagalde I., de Beaulieu J.-L., Bittkau C., Brewer S., Cheddadi R., Ennos R., Fineschi S., Grivet D., Lascoux M., Mohanty A., Müller-Starck G., Demesure-Musch B., Palme´ A., Martín J.P., Rendell S., Vendramin G.G. (2003) Glacial *refugia: hotspots but not melting pots of genetic diversity*. Science 300: 1563-1565.
- **Petit** R.J., Duminil J., Fineschi S., Hampe A., Salvini D., Vendramin G.G. (2005) *Comparative organization of chloroplast, mitochondrial and nuclear diversity in plant populations*. Molecular Ecology 14: 689-701.
- **Pleines** T., Jakob S.S., Blattner F.R. (2009) *Applications of non-coding DNA regions in intraspecific analyses*. Plant Systematics and Evolution 282: 281-294.
- **Pritchard** J.K. ,Wen W., Falush D. (2010) *Documentation for STRUCTURE software version 2.3*. Department of Human Genetics, University of Chicago, Chicago, Illinois, USA. Website: http://pritch.bsd.uchicago.edu.
- **Pritchard** J.K., Stephens M., Donnelly P. (2000) *Inference of population structure using multilocus genotype data*. Genetics 155: 945-959.
- **Provan** J., Soranzo N., Wilson N.J., McNicol J.W., Forrest G.I., Cottrell J., Powell W., (1998) *Gene-pool variation in Caledonian and European Scots pine (Pinus sylvestris L.) revealed by chloroplast simple-sequence repeats*. Proceedings of the Royal Society of London series B-Biological Sciences 265: 1697-1705.
- **Quézel** P., Médail F. (2003) *Ecologie et biogéographie du bassin méditerranéen*. Elsevier Collection Enviroments, Paris, France, 573.
- **Rholing** E.J., Fenton M., Jorissen F.J., Bertrand P., Ganssen G., Caulet J.P. (1998) *Magnitudes of sea-level lowstands of the past 500,000 years*. Nature 394: 162-165.
- **Richardson** D.M. (2000) *Ecology and Biogeography of Pinu*s. Cambridge University Press, Cambridge, UK.

- **Robledo-Arnuncio** J.J., Collada C., Alía R., Gil L. (2005*) Genetic structure of montane isolates of Pinus sylvestris L. in a Mediterranean refugial area*. Journal of Biogeography 32: 595-605.
- **Rousset** F. (2008) *GENEPOP 007: A complete re-implementation of the GENEPOP software for Windows and Linux*. Molecular Ecology Resources 8: 103-106.
- **Schlotterer** C, Tautz D (1992) *Slippage synthesis of simple sequence DNA*. Nucleic Acids Research 20: 2211-2215.
- **Slatkin** M. (1995*) A measure of population subdivision based on microsatellite allele frequencies*. Genetics 139: 457-462.
- **Soranzo** N., Provan J., Powell W. (1998) *Characterization of microsatellite loci in Pinus sylvestris L*. Molecular Ecology 7: 1260-1261.
- **Soranzo** N., Provan J., Powell W. (1999) *An example of microsatellite length variation in the mitochondrial genome of conifers*. Genome 42: 158-161.
- **Stojanoff** N., Stefanoff B. (1929). *Beitrag zur Kenntnis der Pliozäinflora der Ebene von Sofia*. Zeitschrift der Bulgarischen Geologischen Gesellschaft 1: 3-115.
- **Studt** W. (1926) *Die heutige und frühere Varbreitung der Koniferen und die Geschichte ihren Arealgestaitung*. Mitteilungen aus dem Institut für allgemeine Botanik in Hamburg 6: 167-308.
- **Taberlet** P., Fumagalli L., Wust-Saucy A.G., Cosson J.E. (1998) *Comparative phylogeography and postglacial colonization routes in Europe*. Molecular Ecology 7: 453-464
- **Tansi** C., Muto F., Critelli S., Iovine G. (2007) *Neogene-Quaternary strike-slip tectonics in the central Calabrian Arc (southern Italy)*. Journal of Geodynamics 43: 393-414.
- **Tautz** D., Renz M. (1984) *Simple sequences are ubiquitous repetitive components of eukaryotic genomes*. Nucleic Acids Research 12: 4127–4138.
- **Thompson** J.D. (2005) *Plant evolution in the Mediterranean*. Oxford University Press, New York, USA.
- **Tzedakis** P.C., Lawson I.T., Frogley M.R., Hewitt G.M., Preece R.C. (2002) *Buffered tree population changes in a Quaternary refugium: evolutionary implications*. Science 297: 2044-2047.
- **Van Dijk** J.P., Bello M., Brancaleoni G.P., Cantarella G., Costa V., Frixa A., Golfetto F., Merlini S., Riva M., Torricelli S., Toscano C., Zerilli A. (2000) *A regional structural model for the northern sector of the Calabrian Arc (southern Italy)*. Tectonophysics 324: 267-320.

- **Van Dijk** J.P., Scheepers P.J.J. (1995) *Neogene rotations in the Calabrian Arc; implications for a Pliocene-Recent geodynamic scenario for the Central Mediterranean*. Earth-Science Reviews 39: 207-246.
- **Vendramin** G.G., Lelli L., Rossi P., Morgante R. (1996) *A set of primers for the amplification of 20 chloroplast microsatellite in Pinaceae*. Molecular Ecology 5: 595-598.
- **Vidakovic** M. (1991) *Conifer morphology and variation*. Grafički zavod Hrvatske, Zagreb, Hrvatska.
- **Wagner** D.B., Furnier G.R., Saghai-Maroof M.A., Willians S.M., Dancik B.P., Allard R.W. (1987) *Chloroplast DNA polymorphisms in lodgepole and jack pines and their hybrids*. Proceedings of the National Academy of Sciences USA 84: 2097-2100.
- **Wagner** DB (1992). *Nuclear, chloroplast, and mitochondrial DNA polymorphisms as biochemical markers in population genetic analyses of forest trees*. New Forests 6: 373-390.
- **Weir** B.S. (1996) *Genetic data analysis II*. Sunderland (MA): Sinauer Associates.
- **Weiss** S., Ferrand N. (2007) *Phylogeography of southern European refugia. Evolutionary perspectives on the origins and conservation of European biodiversity*. Springer, Berlin.
- **Williams** D., Dunkerley D., DeDecker P., Kershaw P., Chappell M. (1998) *Quaternary Environnements*. Arnold, London.
- **Wolfe** K.H., Li W.H., Sharp P.M. (1987) *Rates of nucleotide substitution vary greatly among plant mitochondrial, chloroplast, and nuclear DNAs*. Proceedings of the National Academy of Sciences USA 84: 9054-9058.
- **Young** A., Boyle T., Brown T. (1996) *The population genetic consequences of habitat fragmentation for plants*. Trends in Ecology and Evolution 11: 413-418.

ACKNOWLEDGMENTS

A sincere thanks to Dr. Vincenzo Bernardini and Dr. Silvano Avolio from CRA-SAM (Rende, CS) for tecnical advice and scientific support during sampling plant materials.

I am profoundly grateful to Dr. Giovanni Giuseppe Vendramin from IGV-CNR (Sesto Fiorentino, FI), for precious and skillful assistance in genotyping and statisticals analyses.

A special thanks to my advisor, Dott.ssa Teresa M.R. Regina from University of Calabria (Arcavacata di Rende, CS). This thesis would not have been possible without her guidance, patience, and support.

I would like to thank the heads of Sila, Aspromonte and Etna National and Regional Parks as well all operators of the Corpo Forestale dello Stato Italiano for allowing me to collect plant material.

This study was financially supported by CRA-SAM [Consiglio per la Ricerca e la Sperimentazione in Agricoltura - Unità di ricerca per la Selvicoltura in Ambiente Mediterraneo, Rende (CS)] and University of Calabria.

Study results have been presented at the following scientific meetings:

- **S. BONAVITA**, T.M.R REGINA, " GENETIC DIVERSITY EVALUATION OF CALABRIAN PINE (*PINUS LARICIO* POIRET) REVEALED BY MICROSATELLITE ANALYSES: A PRELIMINARY STUDY". 106° CONGRESSO SOCIETA' BOTANICA ITALIANA, GENOVA (GE), 2011.
- **S. BONAVITA**, V. BERNARDINI, S. AVOLIO, T.M.R. REGINA, "INDAGINE PRELIMINARE SULLA DIVERSITA' GENETICA DI *PINUS LARICIO* POIRET NELL ' AREALE NATURALE CALABRESE DI DIFFUSIONE". 8° CONGRESSO SISEF, RENDE (CS), 2011.
- **S. BONAVITA**, V. BERNARDINI, S. AVOLIO, T.M.R. REGINA, " APPLICAZIONE DEI MARCATORI MICROSATELLITI ALLO STUDIO DELLA VARIABILITA' GENETICA DELLA POPOLAZIONE DI *PINUS LARICIO* POIRET". 9° CONGRESSO NAZIONALE SULLA BIODIVERSITA', VALENZANO (BA), 2012.