Università degli Studi di Sassari

SCUOLA DI DOTTORATO DI RICERCA

Scienze dei Sistemi Agrari e Forestali

e delle Produzioni Alimentari

Indirizzo Produttività delle Piante Coltivate

Ciclo XXIV

Molecular and phenoptypic characterization of *Quercus suber* L. and *Pinus uncinata* R. populations in the Mediterranean basin

Dr. Giovanni Michele Zucca

Direttore della Scuola: Prof. Giuseppe Pulina

Referente di Indirizzo: Prof. Antonino Spanu

Docente Guida: Prof.ssa Giovanna Attene

Triennio accademico 2009- 2011

Alla mia Famiglia,

che non ha mai smesso di credere in me.

INDEX

1. GENERAL INTRODUCTION

1.1 The diversity in the Mediterranean basin

The Mediterranean Basin is an hotspot for bio-geographical and evolutionary studies, showing an exceptional level of biodiversity (Greuter 1995; Quézel 1995; Blondel & Aronson 1999; Comes 2004; Thompson 2005; Magri *et al.*, 2007). A prominent feature of the Mediterranean flora is its high rate of regional endemism, due partly to a recent diversification (neo-endemism), and partly to the taxa conservation with relictual distribution, resulting from a long-term regional permanence without evolutionary changes (paleo-endemism) (Comes 2004; Thompson 2005).

Endemism represents an integral component of plant diversity in the Mediterranean region. About 60% of all native taxa in the Mediterranean region occur just in this basin, i.e. are endemic to the region as a whole (Thompson *et al.*, 2005, Greuter, 1991). In some regions, especially mountain and islands, rates of endemism often exceed 10% and sometimes 20% of the local flora (Thompson et al., 2005, Médail & Quézel, 1999).

Allopatric speciation associated with geographic isolation is no doubt an important process in long-lived species with efficient mechanisms of gene flow and high effective population sizes (Thompson, 1999). In fact, Allopatric speciation is a fenomena that occurs when biological populations of the same species become isolated due to geographical changes such as mountain building or social changes such as emigration. The isolated populations then undergo genotypic and/or phenotypic divergence as: (a) they become subjected to different selective pressures, (b) they independently undergo genetic drift, and (c) different mutations arise in the populations' gene pools. First, gene flow is often spatially limited; Local differentiation (and possible adaptation) is common in plants (Linhart & Grant, 1996), and Mediterranean species are no exception (Thompson, 1999). Second, in plants there is enormous potential for local speciation associated with several different types of evolutionary process: i.e., hybridization, polyploidisation and inbreeding. For example, hybridization have rise from some clear examples of the evolution of endemic species in the Mediterranean region (Thompson, 2005).

Hence, despite recent acknowledgement that small amounts of gene flow may facilitate the spread of advantageous genes and thus facilitate geographic speciation (Morjan & Rieseberg, 2004). This process would be extremely spread out in the Mediterranean mosaic landscape, where environmental gradients vary strongly in space, localized differentiation may also be an important evolution of endemism (Thompson *et al.*, 2005).

In addition to a very high level of species diversity, the Mediterranean region shows a considerable diversity among and within populations, documented by a large number of studies based on different genetic markers (Dumolin-Lapègue *et al.*, 1997b; Manos *et al.*, 1999; Fineschi *et al.*, 2000; Lumaret *et al.*, 2005; Magri *et al.*, 2007; Lumaret *et al.*, 2009).

Fady-Welterlen (2005) observed that the Mediterranean Basin accounts for more than 10% of the world's vascular plant biodiversity in an area less than 1.5% the size of continental Earth. Forest tree taxa are also exceptionally diverse: more than 100 species have been recorded around the Mediterranean, but less than 30 species can be found in temperate Europe. This amazing biodiversity was developed over millions of years due to the highly heterogeneous geology and climate of the Mediterranean (Fady-Welterlen 2005).

The main factor for the differentiation of European gene resources is generally found in Quaternary climate oscillations, causing repeated retreats and re-advances of plant populations into and from refuge areas (Comes & Kadereit 1998; Taberlet *et al.*, 1998; Hewitt, 1999, 2000). However a possible age for the genetic diversity of the tree populations remains a challenge and new investigations urges. The Mediterranean flora is especially appropriate for this purpose, as many species with a Tertiary origin have persisted through the Quaternary up to present, and so may bear witness of long-term changes of genetic characters (Magri *et al.*, 2007).

Different forces might act in the plants speciation process. In particular, the Mediterranean basin is largely characterized by "allopatric speciation". Geographic or allopatric speciation is an evolutionary process in which one species divides into two because the original homogenous population has become separated and both groups diverge from each other.

The isolated populations then undergo genotypic and/or phenotypic divergence as: a) they become subjected to different selective pressure; b) they independently undergo genetic drift; c) different mutations arise in the populations gene pools.

1.2 The Mediterranean forests

The Mediterranean forests have a various specific characteristic differentiating them from other forest sets. On the physiognomic, trees and evergreen conifers play a major role. These forests are also characterized by their richness and diversity of flora, which are evident both in species that make the procession partner (Médail & Quézel, 1999). These features are the results of the history of the Mediterranean region, and the establishment of its flora. Also frequent morpho-genetic instability, often linked to phenomena for separate area and hybridization or introgression, occurred during the post glacial re-colonization (Quézel *et al.*, 2003). The refuge areas, especially in regions of mountain and peninsulas, explains the diversity of forest phytocenoses and the originality of their genetic components. From the ecological point of view, two sets of factors are decisive. First, bioclimatic criteria are partly responsible for the development of specific functional types wherever there is a Mediterranean climate, which results in plants marked phenomena of water stress and summer drought adaptation (Quézel *et al.*, 2003). The second major criteria is the very strong impact of human activities on forest Mediterranean ecosystems, where, for over 10 millennia, have developed many successive civilization. The impact of human activity, based first on gathering and hunting, and finally on an intensive use of space by gradual clearing for primarily agricultural, tough played a key role in the tree selection process. The extreme variety of environmental conditions, but especially the maturation of these forest systems, sometimes causes a physiognomic convergence with forests not subject to a Mediterranean climate, but made by the same dominant species. In the Northern Mediterranean region, the spectacular growth of deciduous trees in areas previously dominated by sclerophyllae was accompanied by changes in flora and fauna processing to bio-geographical sets, dominated by species Medium-European of Eurasian (Quézel *et al.*, 2003).

1.3 The European forests seed management

The report of Forests Europe State shows that, in 2010, a total of 476.000 ha and 7.700 ha of forests were managed for *in situ* and *ex situ* genetic conservation, respectively, and 870.000 ha for seed production. These efforts included a total of 142 tree species, including subspecies and hybrids (Forest Europe Report, 2011).

There is considerable variation in these areas among different tree species. Large part of *in situ* areas and seeds production efforts have been targeted to widely occurring, stand-forming tree species which are important for forestry. A group of five economically important tree species, i.e. silver fir (*Abies alba*), beech (*Fagus sylvatica*), Norway spruce (*Picea abies*), Scots pine (*Pinus sylvestris*) and sessile oak (*Quercus petraea*), alone, accounts 74% and 66% of the total areas managed for *in situ* conservation and seed production, respectively (Forest Europe Report, 2011).

There are also several economically important tree species (e.g. *Castanea sativa*, *Junglans regia* and *Quercus suber*) for which only small areas are managed for *in situ* and/or *ex situ* conservation. Furthermore, very few genetic conservation areas are managed for scattered tree species (e.g. *Populus nigra*, *Sorbus domestica* and *Ulmus laevis*) which are often considered, incorrectly, or having low importance. In addition to specific uses, many scattered tree species have a high value in terms of maintaining forest biodiversity and/or providing ecosystem services (Forest Europe Report, 2011).

1.4 Evolution and causes of forest trees genetic diversity structure

Among European forest trees, most species have a cyclical history of natural colonization and retreats due to the succession of glacial and interglacial periods so that the timescale for populations divergence extends over a long period. The distribution of other forest tree species, however, has also been strongly modified by humans during the last 2000 years. The ability to distinguish between natural and introduced populations is extremely important in evolutionary studies. Non-native populations that have adapted to new environments can provide important insights into the ability of forest to cope with likely future climate change scenarios (Kremer *et al.*, 2007).

An important main question is: '*How fast will a tree population respond to climate change?*'. The first part will be devoted to the description of the genetic mechanisms that are likely to promote evolutionary changes in response to climate change. These mechanisms will be described at individual, population and species level. Secondly, examine the time scales, and the third part will be addresses to evolutionary rates in trees (*How fast do trees evolve?*). The final part tackles the potential responses of tree populations in terms of risks of extinction, local adaptation and migration.

Adaptive mechanisms of individual trees have been gathered under the generic term plasticity, but they may actually cover different processes (e.g. individual heterozygosity, acclimation, epigenetic response).

Older stands exhibit usually higher heterozygosity than younger stands, but interpretation of these results is still controversial (Mitton and Grant 1984, Ledig 1986, Mitton 1997). Whatever the evolutionary significance of heterozigosity might be, different cases supported correlation between individual heterozygosity and adaptation to strong environmental changes.

Acclimation is the phenotypic change of a single individuals to gradual environmental modification (it's a reversible process). The genetic basis of such phenotypic response has not been elucidated, and to date it is considered to probably be a physiological adaptation to environmental change. However, an investigation in quantitative trait loci (QTL) detection of ecophysiological traits in seedlings of forest trees has indicated that the number and contributions of QTLs might vary substantially according to the CO_2 concentration (Torti 2005).

Many epigenetic responses have been documented in plants as a response to temporary, severe environmental or biotic stresses (Madlung and Comai 2004). However, changes in gene expression may be generated by structural changes (e.g. quantitative modification of repetitive DNA, insertion or deletion of responsible elements), or by change in DNA methylation. In a series of repeated experiments on offspring originated from the same parents but with the mothers raised under different weather conditions, it has been shown that the climate change during sexual reproduction influences the development of seedling (Skrøppa *et al.*, 1994; Johnsen and Skrøppa 1997).

Natural selection has driven populations differentiation throughout the natural distribution of forest trees. The response is a continuous shift in gene frequencies or phenotypic values of traits (König 2005).

Other evidence that selection is the most likely evolutionary force responsible for provenance variation is given by the comparative analysis of 'historical' versus 'geographical' factors of variation (Kremer *et al.*, 2007).

In oaks, it was shown that extant populations stemming from the same source (refugial) of glacial origin but growing today in different ecological sites exhibit strong phenotypic differentiation for fitness-related traits, while the populations are not differentiated for neutral genetic markers (Le Corre *et al.*, 1997; Kremer *et al.*, 2002).

Magri *et al.* (2007), presents the fascinating case of a Mediterranean tree species whose population preserve the genetic imprints of plate tectonic events that took place between 25 million years and 15 million ago.

The study provide a unique insight into the place of evolution of trees, which, despite interspecific gene flow, can retain a cohesive species identity over timescales long enough to allow the diversification of entire plant genera (Magri *et al.*, 2007, Hampe & Petit 2007).

Plant domestication has been particularly intensive in the areas around the Mediterranean basin. Human-mediated domestication could have two contrasting effects on plant genetic diversity: first, it could increase in genetic diversity through admixing of formerly differentiated gene pools; second, it decrease in the genetic diversity through domestication bottleneck (Gepts *et al.*, 1988).

Plant domestication involves the selection of useful or desired traits, and thus can lead to a reduction in the variation present.

As a results, plants often experience a reduction in genetic variability and progressively divergence from wild congeners as they progress through the process of domestication. Such dramatic consequences of selection during domestication via bottleneck followed by a reduction of genetic diversity, have recently been observed in many crops (Tang & Knapp 2003; Liu & Burke, 2006; Cardwell *et al.*, 2006; Vigoroux *et al.*, 2005).

Also decrease genetic diversity was observed in woody plants such as *Olea europaea* (Lumaret *et al.*, 2004) and *Spondias purpurea* (Miller & Schaal, 2006).

Recent studies on *Malus x domestica* B. (Velasco *et al.*, 2010), Eurasian grapevine (Arroyo-Garcia *et al.*, 2006) and *Olea europaea* (Besnard *et al.*, 2001) propose a multilocal origin of the cultivars, which, together with a limited manipulation, might be common features of the domestication process of many woody species (Armelagos & Harper, 2005). It has been shown that species widely cultivated in the Mediterranean region during the last thousands of years may also retain genetic imprints of their natural colonization history. This had been observed, for example, in *Cupressus sempervirens*, an iconic Mediterranean tree, which is thought to have been a human-mediated introduction to Italy sometime in antiquity which has subsequently spread extensively so that is now a permanent component of the Italian landscape (Pignatti, 1982; Bagnoli *et al.*, 2009). This, also, may be for *Quercus suber*, whose genetic differentiation still reflects the footprints of its Tertiary history, despite relatively intense, albeit recent, human exploitation (Magri *et al.*, 2007). Possible explanations for the very low rates of change in the genetic structure of *Q. suber* include its longevity and long-term persistence in the same geographical area (Hampe & Petit, 2007).

1.5 Molecular tools to investigate genetic diversity

The advent of molecular markers has revealed a significant turning point in the world of plant genetics. Construction of genetic association maps, identification of genes responsible for some agronomic traits and varietal characterization (Barcaccia & Falcinelli, 2006).

A molecular marker could be defined as a genomic locus, detectable by specific probes or primers, that, for his presence, define a distinctive chromosomal trait (50-3000bp) (Barcaccia & Falcinelli, 2006).

Molecular markers can be classified according to the number of loci tested. Multi-locus markers are divide in: RAPD (Random Amplified Polymorphic DNA), I-SSR (Inter-Simple Sequence Repeats), AFLP (Amplified Fragment Length Polymorphism).

The molecular markers could be also divided in dominant or codominant markers. The first group (i.e. AFLP, RAPD...) discriminate individuals for the presence (1) or absence (0) of the locus studied. The second group instead, is able to evaluate also the heterozygose individuals (SSR, I-SSR…) (Barcaccia & Falcinelli, 2006).

A further classification can be formulated based on the technique used: molecular markers obtained by Southern blot hybridization, RFLP and VNTR such as (Variable Number of Tandem Repeat or Minisatellites), and molecular markers produced by PCR amplification (RAPD, SSR or SSR/AFLP-I).

The principal function of molecular markers in genetics and agronomic field, is to identification the potential differences in nucleotide genomic DNA sequence (Barcaccia & Falcinelli, 2006).

The differences (polymorphisms) could be due to insertions/deletions, translocations/duplications and mutations. A few future ambitious goal would be to develop association maps of interesting-agronomic traits (Resistance to pathogens, abiotic resistance stress, yield), and polymorphic marker.

In conclusion, the use of molecular markers seems to be usefully to explore the molecular bases of individuals phenotypically indistinguiscible. The limits of molecular markers use is related to the sequences information available (i.e. genomic sequences, flanking sequences…). In addition the cost of sequencing approach is extremely high and specific equipment and expertise are required (Barcaccia & Falcinelli, 2006).

In our study SSR markers and candidate genes loci were used to investigate a genetic diversity in two different Mediterranean species collection.

Microsatellites loci are particularly useful for intraspecific genetic comparison, but their high mutation rate makes them less useful for interspecific analyses. Size homoplasy, where alleles of identical size are not identical by descendent, is a particular problem when analyzing interspecific microsatellites data sets (e.g. Doyle *et al.*, 1998), however the lower mutation rate of chloroplast microsatellites compared to nuclear microsatellites (Provan *et al.*, 1999) suggest that size homoplasy may not be a problem for chloroplast microsatellites comparison between closely related species.

Bryan *et al.* (1999) suggested that cpSSR are great utility in population genetics, germplasm management, evolutionary and phylogenetic studies as well as in the analysis of material from introgression and somatic-fusion experiments.

Using a candidate genes loci we could be explore different metabolic pathways in plant. There are different genes which represented a keys for specific biological process. For example: a) Dehydrin genes, that are key components of dehydration tolerance (Close T.J. *et al.,* 1996, Wachowiak *et al.*, 2009); b) Chalcone synthase (*chcs*) is one of

the genes responsible for the production of polyketide synthase enzymes. This enzyme is associated to the production of chalcones (an organic compounds related to natural defense mechanism). c) Abscissic acid (ABA) plays a role in the abscission of plant leaves and ABA mediated signaling. It is also an important hormone involved in plant response to abiotic and biotic stress (Zhu JK. 2002).

1.6 Common garden experiment

In this study, we used a *Q. suber* seedling collection from Mediterranean basin country. The collection was transplanted in a "common garden" central Sardinia area by Department of Economy and Woody Plant (DESA) of Sassari University (1999).

A "common garden experiment", is an experiment where one or more organisms (in our cases plants) are moved from one environment to another environment. In a usual common garden experiment, more plants growing in their native environments would be transplanted in a common environment (Molles, 2002).

Transplant experiments are often performed to test if there is a genetic component to differ in populations. In recent decades the advances of molecular biology have provided researchers with the ability to study genetic variation more directly.

However, transplant experiments still have the advantages of being simple and requiring little technology. On the other hand, they may require considerable time and efforts, and the organisms test number is often limited (Molles, 2002).

Therefore, the presence of geographically related patterns in the distribution of intra-specific genetic differentiation (Gómez *et al.*, 2005) points to evidence differences among populations, which are likely related to the climatic characteristics at their geographic origin (Mátyás, 1996). Provenance research (i.e. the assessment of intra-specific variability under common-garden conditions) is the traditional approach to study the relevance of ecological factors, especially of climate, in shaping adaptive strategies of forest trees (Mátyás, 1996; Rehfeldt *et al.*, 2002). This information may be important to understand the risk of mal-adaptation of Mediterranean species.

2. GENERAL OBJECTIVES

The main aim of this research is to explore the genetic diversity in two different Mediterranean forest species: *Quercus suber* L. and *Pinus uncinata* R.

The cork oak has a typical areal Atlantic – Western Mediterranean. In Sardinia is abundant throughout the island, extending from the lowlands up to the average mountain areas, with sporadic presence on the limestone in the center of Eastern Orgosolo Supramonte and Oliena (Gellini & Grossoni, 1997). No information are known about the genetic structure diversity level for Sardinia *Q. suber* populations. This is a critical prerequisite for economical and environmental correct use of this specie. For this purpose, the first objective of this study was to investigate the genetic and phenotypic variability in a collection of cork oak populations.

In particular, in the first chapter we tested the level of genetic variability in 27 different gene pool *Quercus suber* populations using SSR microsatellites approach (Dumolin *et al.*, 1995; Besnard *et al.*, 2002; Jiménez *et al.*, 2004; Lumaret *et al.*, 2005; Magri *et al.*, 2007). These accessions are from different forests of Western-central Mediterranean (French, Spanish, Portuguese, Moroccan, Italian, Tunisia and Algerian). They were planted in a "common garden" central Sardinia area (Monte Grighini, Ente Foreste, Siamanna City council; 39°55'37" N, 8°48'45" E).

The chloroplast and nuclear microsatellite markers were used to evaluate the *Q. suber* population genetic and phenotipic structure across the Mediterranean origins.

In the second chapter we explore the genetic diversity of 5 *Pinus uncinata* populations from the Western Mediterranean sites (France, Spain and Andorra). We used 29 different nuclear candidate genes loci to investigate the genetic diversity level among *P. uncinata* populations (Heuertz *et al.*, 2006; Gonzalez-Martinez *et al.*, 2006; Päjärvi *et al.*, 2007; Eveno *et al.*, 2008, Wachowiack *et al.*, 2009).

Mountain pines, especially *Pinus uncinata*, are often used by homeowners and landscape architects for home security purposes. The aesthetic characteristics *P. uncinata*, is in conjugation with their home security qualities, makes them a considerable alternative to artificial fences and walls. They are widely used for ornamental purposes in few areas of North America (Christensen, 1987).

Finally for both species we evaluated some phenotypic qualitative and quantitative characters.

In fact, quantitative traits play a key role in adaptive evolution process (Lande and Barrowclough, 1987; Milligan *et al.*, 1994).

CHAPTER 1

Use of chloroplast and nuclear microsatellites

to study genetic diversity in a *Quercus suber* Mediterranean basin

collection.

1. INTRODUCTION

1.1 *Quercus* genus

Mediterranean evergreen *Quercus* species are a group with overlapping habitats, which often leads to their consideration as an homogeneous entity in botanical, bio-geographical or paleo-historical studies. In the Western Mediterranean Basin, *Q. suber* (cork oak), *Q. ilex* (holm oak) and *Q. coccifera* (kermes oak) are the dominant broadleaved species. The three species are sympatric in many areas, but they have some different evolutionary histories.

Previous studies have shown differences in their genetic variation patterns at both nuclear and cytoplasmic levels (Toumi and Lumaret, 1998, 2001; Manos *et al*., 1999; Belahbib *et al.*, 2001; Lumaret *et al.*, 2002).

The genus *Quercus* comprises approximately 500 species of trees and shrubs distributed throughout much of the Northern hemisphere (Nixon, 1993). Oaks are conspicuous members of the temperate deciduous forests of the Europe, North America and Asia, in addition to being important evergreen elements of Mediterranean woodlands and subtropical forests.

The largely temperate species of subgenus *Quercus* are distinguished from the strictly South-East Asian members of subgenus *Cyclobalanopsis* by several characters, notably the presence of expanded stigmatic surfaces on the pistillate flowers and small, inconspicuous bracts which subtend single-staminate flowers (Nixon, 1985, 1993).

The Middle and late Tertiary fossil record of North America and Asia clearly establishes the long-term abundance of subgenus *Quercus* and provides a framework considering the historical biogeography of the group (Axelrod, 1983; Crepet, 1989; Crepet and Nixon, 1989a; Zhou 1993).

Leaf impressions, some with clear affinities to modern species, establish a minimum age of 40 million years for several of the major oak groups recognized today (Elliot *et al.*, 1999). Oaks species are well known for their taxonomically perplexing patterns of intra-specific morphological variation which may be caused in part of

hybridization (e.g. Trelease, 1924; Palmer, 1948; Muller, 1952; Tucker, 1961; Hardin, 1975; Rushton, 1993; Spellenberg, 1995; Bacilieri *et al.*, 1996; Howard *et al.*, 1997).

The hypothesis of hybridization among oaks species is supported by several observations (Bacilieri *et al.*, 1996). DNA-based evidence of gene flow between species was first reported on the basis of shared patterns of chloroplast DNA (cpDNA) haplotypes in sympatric populations of white oak species (Whittemore and Schaal, 1991).

Additional studies on cpDNA from a broad sampling of populations of two Western European white oak species suggested that geographical patterning ad discontinuities among haplotypes also could be explained by geological migration via long-distance dispersal (Petit *et al.*, 1993a, 1997). Although it is well known that sterility barriers between oak species are poorly developed, hybridization appears to be limited to species that belong to the same major group or section within the genus (Stebbins, 1950; Grant, 1981; Cottam *et al.*, 1982).

Nonetheless, the broad boundary between apparently hierarchic relationships and interbreeding in the oaks presents an interesting opportunity to compare gene trees derived from both chloroplast and nuclear DNA at various taxonomic levels across the phylogenetic breadth of subgenus *Quercus*.

In fact, in this genus, phylogenetic relationships based on morphology are somewhat obscure, largely due to pronounced vegetative variation (Tucker *et al.*, 1974), which is sharply contrasted by a stabilized, seemingly constant set of floral characters.

Most studies on genetic variation of *Quercus* forest trees have been carried out using nuclear markers. The chloroplast genome typically shows low intraspecific variation, although in the last decade, the development of new markers has allowed the detection of enough variation to assess phylogeographic patterns. Many such studies in *Quercus* species have used the PCR-RFLP technique (Whittemore and Schaal, 1991; Ferris *et al.*, 1993; Petit *et al.*, 2002; Olalde *et al.*, 2002).

1.2 *Quercus suber* L.

Quercus suber L. is an emblematic Mediterranean evergreen sclerophyllous tree. Cork oak is a slow growing, extremely long-lived evergreen tree. It is a monoecious wind-pollinated species with a protandrous system to ensure cross-pollination.

Reproductive system (mostly sexual) and dissemination (gravity and zoochory) of cork oak are coherent with those of most oak species (Magri *et al.*, 2007; Gellini & Grossoni 1997). It may reach about 20m in height, with massive branches forming a round crown. Its thick a soft bark is the source of cork, which is stripped every 10-12 years from the outer layer of the bark along the lower portion of the trunk (Gellini & Grossoni 1997).

The modern distribution of the *Q. suber*, rather discontinuous, ranges from the Atlantic coasts of North Africa and Iberian Peninsula to the Southern regions of Italy (especially Sardinia and Sicily), and includes the main West Mediterranean island as well as the Morocco, coastal belts of Maghreb (Algeria and Tunisia), Provence (France) and Catalonia (Spain) (Fig. 1).

Quercus suber is widely cultivated within its natural range for the cork production. According to Carriòn *et al.* (2000), without human activities *Q. suber* would never develop pure stands in the Iberian Peninsula, and would form mixed forests with xerophyllous and deciduous oaks together with *Pinus pinaster* instead.

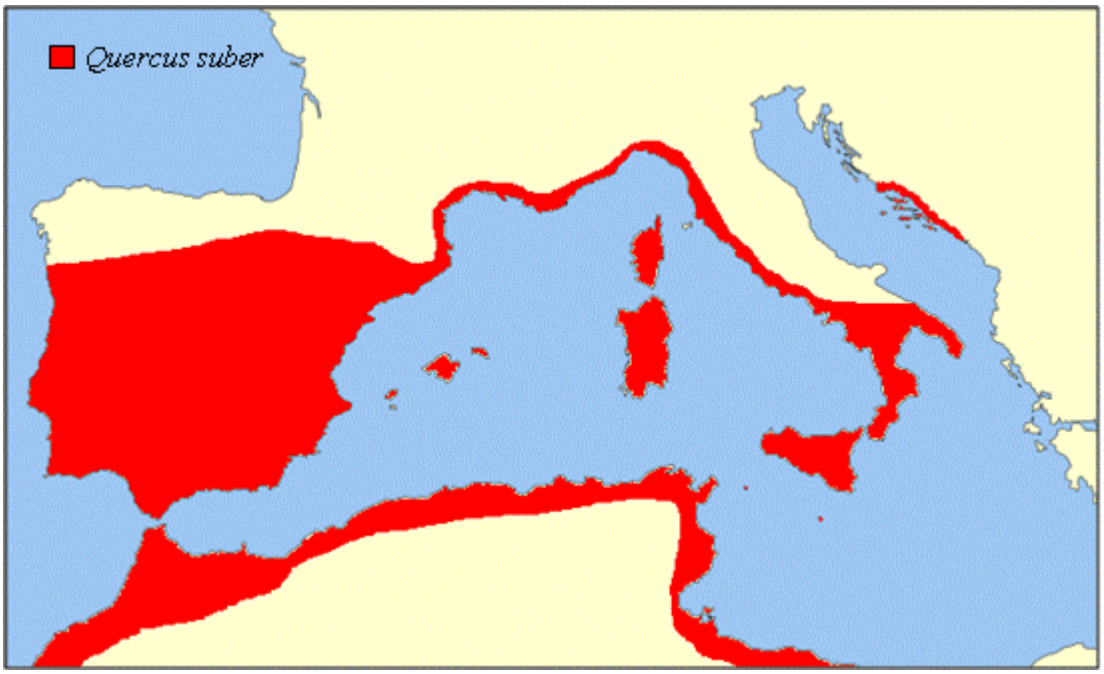
In recent years, cork oak has been the subject of intensive genetic studies by means of different markers, and a strong differentiation among populations has been detected around Mediterranean Basin by investigating both chloroplast and mitochondrial DNA (which are maternally inherited in oaks, as demonstrated by Dumolin-Lapègue *et al.*, 1995), as well as allozyme variations; moreover, evidence of cytoplasmic introgression of *Q. suber* by *Quercus ilex* genes has been reported (Toumi & Lumaret 1998, 2001; Belahbib *et al*. 2001; Lumaret *et al.* 2002, 2005; Lòpez de Heredia *et al.* 2005).

In fact, exceptionally, the allozyme diversity among populations of *Quercus suber* (11%) was relatively high when compared to other oak species (Toumi & Lumaret, 1998; Gandour *et al.*, 2007).

Although both *Quercus suber* and *Q. ilex* are evergreen oaks, they belong to different subgenera (*Q. suber*, subgenus *Cerris*; *Q. ilex*, subgenus *Sclerophyllodrys*) according to classical oak taxonomy (Schwarz 1993).

However, using a similar methodology, a more recent classification includes both species within the same Eurasian section *Cerris* (Manos *et al.*, 1999).

Figure 1. Cork oak (*Quercus suber* L.) Mediterranean areal distribution.



1.3 The importance to study the *Quercus suber* in Sardinia

In Sardinian ecosystem *Q. suber* has an important role for environmental, recreational and landscape purposes. Also, in Sardinia, the cork production had represented for a long time an important economically remunerative product, in the local industrial systems. The productions of cork stoppers, the most higher remunerative economical products from cork oak (*Q. suber* L.), stands, required raw materials with high elasticity to assure good bottle closure and with limited porosity (Pereira *et al.*, 1996, Corona *et al.*, 2005).

The quality of raw cork oak is determined by the interaction of genetic and environmental factors (Natividade 1934). Frequent allogamy in this species leads to the occurrence of very diversified phenol-genotypes: its total genetic diversity is among the

highest recorded in oak species (Toumi and Lumaret, 1998). Thus, it is possible determinate populations using morphological (Schirone & Bellarosa, 1996), biochemical (Toumi and Lumaret, 2001) and molecular (Bellarosa, 2003) descriptors.

However, so far, the few studies conducted have not been able to quantify a precise correlation between genotypes and cork quality. Ferreira *et al.* (2000) provides statistical grounds for a genetic selection program. As an example from the empirical side, in the year 1962 an important producer in Gallura (North Sardinia) planted 800 seedling obtained from a single mother tree selected for its excellent cork quality: the first stripping of the planted trees confirmed that the quality of the cork, even if virgin, was above average (Corona *et al.*, 2005).

In fact, according to all the observations and economical required, it is very important, to improve the genetic knowledge of Sardinian *Q. suber* accessions throughout different type of molecular markers. SSRs markers and candidate genes loci, could be useful to identify the genetic origin of *Quercus* in Sardinia and study the genetic structure of local *Q. suber* populations. Several recent common garden collection are made by *Quercus* from European germplasm; A complete genotypization of the common garden germplasm, and local germplasm is quite often missing. The molecular markers study could explore the state of biodiversity in Sardinia.

1.4 Microsatellite marker (SSR)

Simple sequence repeats (SSR), otherwise known as microsatellites, are tandem repeated nucleotide sequences 2-4bp long that vary in number of repeat motifs and flanked by conserved DNA sequences (Tautz, 1989). SSR containing fragments can be amplified via the polymerase chain reaction (PCR) using primers pairs complementary to the flanking regions. Polymorphism is due to different fragment lengths that can by discriminated by gel electrophoresis.

Due to the high mutation rates and resulting variability at SSR loci, these markers have been ideal for genetic mapping and characterizing genetic diversity in forestry and crop species at the inter-specific, inter-subspecific, inter-varietal and even intra varietal levels (Lee, 1995; Mitchell *et al.*, 1997; Matus and Hayes, 2000).

The SSRs markers, have high repeatability, transferability of results from one laboratory to other, the possibility of automation and multiplexing. These is an ability to use two or more primers simultaneously when the amplification products differ in weight so they do not overlap each other in the same line gel (Mitchell *et al.*, 1997). In addition, because of their codominant nature the SSR can be used for genetic maps construction.

However, they have same limitations: it can happen that the mechanism of mutational homoplasy is present among individuals carrying subversive. In fact, it is possible that two individuals show identity in the state but not identically by descent (Wright, 1951).

This phenomenon can have varying frequency depending on the species thus complicating the study of phylogenetic relationship between species within the same genus and between populations within species.

Chloroplast microsatellite marker (cpSSR)

The chloroplast genome generally has a highly conserved organization (Palmer, 1991; Raubeson and Jansen, 2005). It's composed of a single circular chromosome with a quadripartite structure that includes two copies of an inverted repeat (IR) that separate the large (LSC) and a small single copy (SSC) regions (Sasky *et al.*, 2005).

The analysis of chloroplast genome are emerging as a useful tool for population genetic and phylogenetic analysis because conserved gene order, the widespread availability of primers and a general lack of heteroplasmy and recombination (Provan *et al.*, 2001). Furthermore, its uniparental mode of inheritance (usually maternal in angiosperms and paternal in gymnosperms) makes it possible to elucidate the relative contributions of seeds and pollen gene flow to the genetic structure of natural populations by comparing nuclear and chloroplast markers (Ennos *et al.*, 1999).

CpSSR analysis technique was developed in the 1990's (Powell *et al.*, 1995; Vendramin *et al.*, 1996; Vendramin and Ziegenhagen, 1997). This technology is based on highly polymorphic sequences.

The simple repeat found in the chloroplast genome of higher plants is generally made up of short poly (A) or poly(T) tracts, with maximum sizes of about 20bp (Weising and Gardner, 1999).

Nevertheless, a number of studies employing flanking PCR primers have shown that such mononucleotide stretches are polymorphic among different species and

accessions of *Hordeum* (Provan *et al.*, 1999b), *Oryza* (Provan *et al.*, 1996; Ishii and McCouch, 2000), *Pinus* (Cuenca *et al.*, 2003), *Solanum* (Btyan *et al.*, 1999; Sukhotu *et al.*, 2006) and *Vitis* (Arroyo-Garcia *et al.*, 2002).

Microsatellite loci are particularly useful for intraspecific genetic comparisons, but their high mutation rate makes them less useful for interspecific analyses. Few studies focusing on the mutation rate of chloroplast microsatellites are available in literature; however estimates vary from 10^{-3} in *Pinus contorta* (Marshall *et al.*, 2002) to 3.2 x 10^{-5} - 7.9 x 10^{-5} number of mutations per generation (Provan *et al.*, 1999).

Studies of chloroplast microsatellites have often revealed much higher levels of diversity than have those chloroplast RFLP such as those on a set of wild and cultivated *Oryza* (Provan *et al.*, 1997) and on a set of 12 accessions of *Hordeum spontaneum* from Israel (Provan *et al.*, 1999b). Similarly, the analysis of natural populations of *Pinus sylvestris* using chloroplast microsatellite were resulted in within population genetic diversity (*H*) values always greater than 0.95, and with the 40% of the individual with a unique genotype. Moreover, Ribeiro *et al.* (2002) analyzing populations of *Pinus pinaster* observed that cpSSR revealed much higher genetic diversity than AFLPs.

This all highlighted the discriminating power and the usefulness of the cpSSR assay (Russell *et al.*, 2000).

Nuclaer misrosatellite marker (nuSSR)

Nuclear microsatellites are repeating units of DNA base pairs (A-T, G-C) which are used as markers in genetic studies of kinship, origins and migrations species. Microsatellites experience higher rates of genetic mutation compared to other parts of DNA and are therefore perfect for studying species ancestry (Barcaccia & Falcinelli, 2006).

In oaks, although nuclear microsatellites have been available for many years (Dow *et al.*, 1995; Steinkellner *et al.*, 1997; Kampfer *et al.*, 1998), multiplexing efforts were limited, with only two studies reporting multiplexing at no more than five loci (Dzialuk *et al.*, 2005; Lepais *et al.*, 2006). Thus, analysing large oak populations at multiple markers remains expensive and time consuming. Also, Guichoux *et al*. (2011) developed two multiplex kits (eSSRs and gSSRs) for *Q. petraea* and *Q. robur*.

2. OBJECTIVES

To have in the future a detailed knowledge on the genetic structure of *Q. suber* Sardinian populations, we decided to evaluated the distribution of genetic diversity across the whole Mediterranean area.

According to this, main objective of this study was to determine the level and the structure of the genetic diversity of *Q. suber* in the Mediterranean basin.

With this purpose, we first investigated a chloroplast and nuclear DNA of *Q. suber* grown in a *common garden* monospecific accessions Monte Grighini Sardinia collection (MGC) using microsatellites markers. We evaluated the genetic structure of the *Q. suber* and analyzed the level of polymorphism among and within populations (French, Spanish, Italian, Portuguese, Moroccan, Tunisian and Algerian accessions).

Secondly, we evaluated the genetic structure, of this species, investigated in a *Q. suber* area to others *Quercus* species (*i.e.*, *Q. coccifera*, *Q. ilex*, *Q. pubescens, Q. morisii, Q. congesta*) sampled in different Sardinian areas.

Finally, we analyzed phenotypic variation among populations of *Q. suber,* monitoring some phenological characters (*i.e.*, bud set and bud flush), and morphometric data (shape of trees).

3. MATERIALS AND METHODS

3.1 Site and Plant materials

Experimental plot of 3.3ha, of mono-specific forestation of Q. suber, were located at 440m above sea level (a.s.l.) in the perimeter of the Regional Monte Grighini Forestry (West-Central Sardinia, Siamanna City Council, 39°55'37.05"N; 8°48'45.68"E)(Figure 3.1).

The seedlings were from 27 populations from 7 Mediterranean contries (Spain, Portugal, Morocco, Algeria, Tunisia , France and different Italian regions, see Table3.1). Seedling were planted by DESA Department of Agriculture University of Sassari in 1999, after working with a harrowing year going and managed not to alter the genetic expression of the seedling were not subject to pruning.

The parcel was divided into equal parts by a geological rift between the complex and the Hercynian metamorphic Paleozoic, correspondent by granite and quartzite respectively (Figure 3.2). The soil texture is sandy or franco-sandy, sub acid and neutral. According to data compiled by the climate area is located in the plane with bioclimatic meso-Mediterranean ombrotype dry.

The experimental parcel, integrated in Mediterranean Oaks Network (FAIR I CT 95-0202; Valera, 2003), containing 3240 *Q. suber* trees distributed in 30 different blocks, with 108 individuals per block.

These seedlings were planted according to the following scheme: 3 x 3m distances among trees and4 different genotypes for any populations (Dettori *et al.*, 2006). We sampled 27 individuals per block,one for each population.

This sampling was repeated on 5 blocks for a total of 135 individuals. Two extremely different genotypes were also added to the collection in order to have representative genetic diversity dataset (from Sardinian and Moroccan). So, in Monte Grighini collection (MGC) parcel, we sampled 137 *Q. suber* individuals.

In the MGC collection, the Sardinian area was represented by populations from the north (province of Sassari), and the south (Province of Cagliari). As the Q. suber diffusion is all over the island, we decided to implemented this sampling with 14 individuals from

the Central Sardinia area (Province of Oristano). We sampled these 14 *Q. suber* individuals in June and observed for *L. dispar* response to the attack: 8 of them were not attacked by the phytophagous and the other six were attacked.

Furthermore 11 more Sardinian individuals, belonging to the *Fagaceae* family, were added as controls. Exactly were sampled: 3 *Quercus ilex*, 1 *Quercus coccifera*, 2 *Q.morisii*, 3 *Q.congesta*, 2 *Quercus pubescens.*

In total were sampled: 137 *Q. suber* individuals in MGC, 14 *Q. suber* accessions in central Sardinia area and more 11 samples from Sardinian place (of different *Quercus* species).

Figure 3.1 Distribution of *Q. suber* accessions grown in the common garden (Sardinia collection) located in MGC. The black circles indicate the different population analyzed.

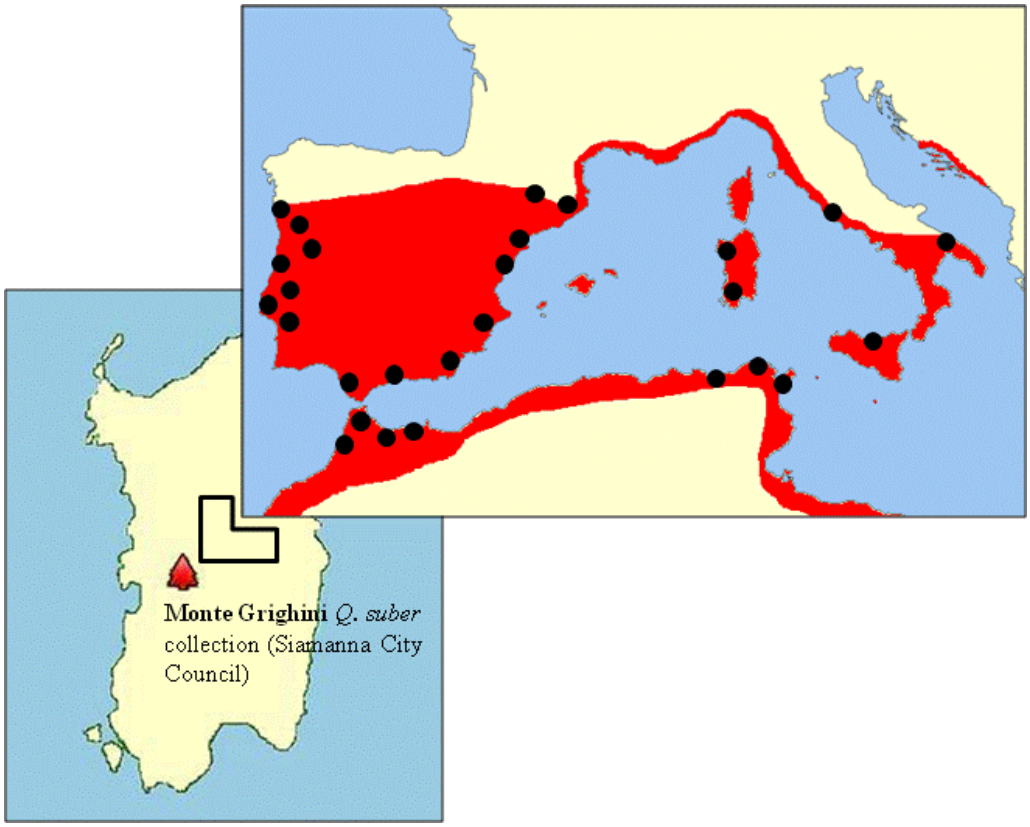


Figure 3.2 Experimental plots. Cork oak stand (white pints) is articulated in granite on the left (black) and quartzite on the right (grey).

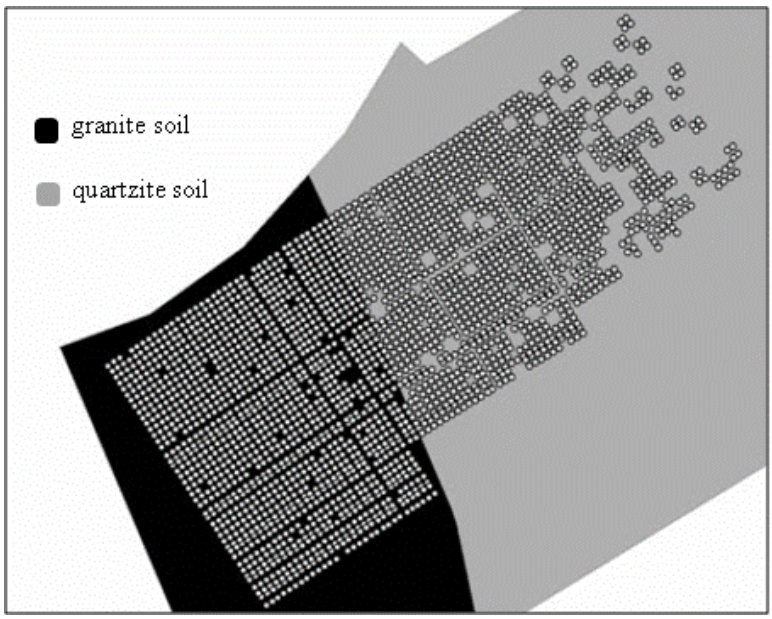


Table 3.1 Code, Country, Locality, Geographic, climatic and edaphic characteristic for the sampled populations of cork oak (*Q. suber* L.), grown in a Monte Grighini Sardinian collection. **N**, population number; $\mathbf{T_m}$, average annual Temperature (°C); **P**, average annual rainfall.

N	Code	Country	Lacality	Latitude	Altitude (m)	P (mm)	Tm (°C)	Substrate
1	AL I	Algeria	Guerbes	36° 53' N	42	-	-	-
2	ES 3V	Spain	Monte de Toledo	39° 22' N - 39° 25' N	600-800	1063	15,20	Siliceous
3	ES 4CR	Spain	Sierra Morena Oriental	38° 24' N - 38° 33' N	700-900	719	14,30	Siliceous
4	ES 5J	Spain	Sierra Morena Occidental	38° 13' N	400-500	666	16,00	Siliceous
5	ES 7 ALM	Spain	Parco de los Alcornocales	36° 16' N	20-120	993	16,60	Siliceous
6	ES 9 GE	Spain	Cataluna Litoral	41° 51' N - 41° 53' N	200-500	802	15,00	Siliceous
7	ES F PAR	Spain	Sierra de Guadarrama	40° 31' N	680-740	455	13,90	Siliceous
8	FR II	France	Le Rembaut Provence	42° 27' N	280	958	15,70	Schistose
9	FR III	France	Soustons Landes	43° 45' N	20	870	12,30	Siliceous
10	IT LA VT	Italy	Tuscania	42° 25' N	160	937	14,50	Eruptive
11	IT PU BR	Italy	Brindisi	40° 34' N	45	588	16,60	Clayey
12	IT SA CA	Italy	Cagliari (Santadi)	40° 27' N	200	883	17,00	Granite
13	IT SA SS	Italy	Sassari (Villanova Monteleone)	39° 05' N	300	910	14,20	Eruptive
14	IT SI CT	Italy	Catania	37° 07' N	250	448	17,70	Terra rossa

…continue…

…continue…

N	Code	Country	Lacality	Latitude	Altitude (m)	P (mm)	Tm (°C)	Substrate
15	MO I 1	Morocco	Boussafl	35° 11' N	150	574	-	Clayey
16	MO I 2	Morocco	Ain Rami Rif. Occidentale	35° 07' N	300	1280	-	Siliceous
17	MO III 1	Morocco	Ain Johra Maarnora	34° 07' N	160	536	-	Siliceous
18	MO III 2	Morocco	Oulmes-Plateau Central	34° 07' N	150	479	-	Clayey
19	PT IV 02	Portugal	Ponte de Sor	36° 35' N	70	710	16,00	Siliceous
20	PT IV 03	Portugal	Quinta da Serra	38° 28' N	120	681	14,30	-
21	PT IV 04	Portugal	Alcacer do Sal	38° 43' N	30	577	16,30	-
22	PT V 01	Portugal	Azaruja	36° 35' N	360	564	15,60	Siliceous
23	PT VI 01	Portugal	Santiago do Cacem	36° 35' N	140	736	15,60	Siliceous
24	PT VI 03	Portugal	S.Bras de Alportel	36° 35' N	440-485	874	15,90	Siliceous
25	PT ES	Port./Spain	Alent.B.Baixa/S.San Pedro	39° 12' N	450-515	778	15,40	Siliceous
26	TU I	Tunisia	B Fernanda	36° 35' N	12	948	17,90	Siliceous
27	TU II	Tunisia	Mekna	36° 57' N	270	1610	14,90	Siliceous

3.2 DNA extraction

For all 162 individuals (describe in the previous section), genomic DNA was extracted from fresh leaves. The DNA extraction was carried out on a single plant using the DNeasy® Plant kit (QIAGEN, Hilden, Germany).

3.3 Chloroplast SSR analysis

We screened the 162 individuals with 8 cpSSR primers, published in literature (Table 3.2). The first two primer pairs (defined *universal primer*) were designed by Weising and Gardner (1999) in *Tabacum* and used in different forestry species (Bruschi *et al.*, 2000; Gugerli *et al.*, 2001; Oddou-Muratorio *et al.*, 2001; Besnard *et al.*, 2002; Heuertz *et al.*, 2004, 2006; Magri *et al*., 2006, 2007). The others 6 primers pairs were designed by Sebastiani *et al.* (2004), where described a new set of chloroplast microsatellites, represented by both mono and dinucleotide stretches, in some species of Fagaceae and in *Fraxinus excelsior*, and their variation in *Castanea sativa*, *Fagus sylvatica* and *Quercus petraea*. The identification of this new scpSSR was possible because of the availability of the sequences chloroplast genome of *C. sativa* (Sebastiani, 2003).

PCR was carried out in a 25µl volume, containing 25ng template DNA, 10pm of each primer, 20µM dNTPs, buffer PCR 1X (200mM Tris HCL, pH 8.4, 500mM KCl), 50mM $MgCl_2$ and 1U Taq polymerase (Invitrogen).

The amplifications were conducted with a Perkin-Elmer 9700 thermocycler, with an initial 5min at 94°C that was followed by 35 cycles of: 1min at 94°C, 1min at 60°C, 1min at 72°C, and a final extension step at 72 °C for 10min (Ramping 75%) (Table 3.2). The amplification products were loaded onto 6% 8M denaturing polyacrylamide gels and separated for 1,50h at 70W. Gels were fixed in 7.5% acid acetic, washed and visualized by silver staining coloration system (Bassam *et al.*, 1991).

Table 3.2 Code, location, repeats, sequences **(5'→3')**, annealing temperature (°C) and size motif (*bp*) of 8 primers pairs for chloroplast microsatellites analysis designed by Weising & Gardner (1999), and Sebastiani *et al.* (2004).

Code	Location	Repeats motif	Primer Sequence (5'→3')[a-b]	T_a (°C)	PCR product size (bp)**	Source
ccmp4*	atpF intron	$(T)_{13}$	AATGCTGAATCGAYGACCTA CCAAAATATTBGGAGGACTCT	60	126	Weising and Gardner, 1999
ccmp6*	ORF 77-ORF 82 int.	$(T)_5\,C(T)_{17}$	CGATGCATATGTAGAAAGCC CATTACGTGCGACTATCTCC	60	103	Weising and Gardner, 1999
Cmcs1*	petD int.	$(AT)_7$	ATTCATTTCCTTTGCATTGA TTTACTTGTTACTAATAGGGTCTAGC	60	109	Sebastiani et al., 2004
Cmcs6*	ndhG-ndhI int.	$(T)_{10}$	GAAAAAGGACCCTTCCTAAT CTTATGATCGTCACGAATTG	60	199	Sebastiani et al., 2004
Cmcs7*	ndhH-rps15 int.	$(T)_{12}$	AAGCGAGATGAATGAGTTTT AAAATTGGATTGATTATTGACT	60	205	Sebastiani et al., 2004
Cmcs8*	ndhH-rps15 int.	$(A)_{10}$	GGTCTATTTTTCCACTCACAA AGAAATAAACACCCCCATTA	60	179	Sebastiani et al., 2004
Cmcs9*	ycf5-ndhD unt.	$(T)_9$	AAAAATACTTCTTTTTCGTTTTC CCTGAATAAAATTCAAAATCAA	60	101	Sebastiani et al., 2004
Cmcs14*	accD-psaI int.	$(AT)_7$	GGATTGTAACAAATTTTTCAGG GTGCAAGGAATGTCGAACTA	60	178	Sebastiani et al., 2004

T_a, annealing temperature; **PCR**, polymerase chain reaction. ***PCR conditions**: 5min at 94 °C; 35 cycles of 1min at 94 °C, 1min at 60 °C, 1min at 72 °C; 10min at 72 °C, 4 °C$_\infty$, Ramping 75%. **a)** the first sequences is forward and the second is reverse; **b)** Degenerate positions are Y (=C or T), B (=G, C or T). **expected size in tobacco for Weising & Gardner (1999) loci, and in *C. sativa* for Sebastiani *et al.* (2004) loci.

3.4 cpSSR data analysis

Gene diversity has been calculated as for a haplotype organism, using Nei's unbiased gene diversity (Nei, 1978):

$$H = (n/n-1)(1-\sum pi^2)$$

Where: ***pi*** is the frequency of the ***i*** variants and summation extends over ***n*** variants. This allowed direct comparison with other authors (e.g. Bryan *et al.*, 1999; Xu *et al.*, 2000; Ishi *et al.*, 2001; Chung *et al.*, 2003).

Gene diversity was calculated based on frequencies of cpSSR alleles and it has been estimated for the whole samples, the entire *Q. suber* Grighini accession and within *Q. suber* individuals for any country separately.

In the same time *H* was calculated considered all accessions, where admix *Q. suber* Grighini collection whit different trees collected in central area of Sardinia.

The cpSSR microsatellites data were analyzed following the infinite allele model (Otha and Kimura, 1973), with each band scored as present (1) or absent (0), and the matrix of similarity among genotypes was obtained using Nei and Li (1979) equation as follows:

$$S = 2Nij / (Ni + Nj)$$

Where ***Ni*** = number of alleles of the genotypes *i*, ***Nj*** = number of alleles of genotypes *j* and ***Nij***= number of alleles in common between genotypes *i* and *j*.

A dendrogram based on the Nei and Li (1979) pairwise distances matrix was then constructed with the unweighted pair group method of arithmetic clustering (UPGMA) method (Sneath and Sokal, 1973) (*e.g.* Powell *et al.*, 1996, Ishii and McCouch, 2000).

The level of support for the nodes was evaluated using by bootstrapping with 1000 replicates, which was carried out using the TREECON software (Van de Peer, 1994). Two different dendrograms were build considered separately *Quercus suber* MGC and all accessions sampled (162 total individuals).

The degree of genetic divergence among and within populations was investigated using analysis of variance framework (AMOVA). Genetic variation was analyzed for all chloroplast loci considered. The significance of genetic differentiation coefficient was evaluated by 1000 permutations of the samples between populations using ARLEQUIN ver. 3.5.1.2 (Excoffier *et al.,* 2005). Using the same software was calculated F_{ST}, pairwaise genetic distance within different accessions considered.

3.5 Sequence analysis

To confirm and support the polymorphism level observed in the previously molecular analysis (polyacrylamide gels), we sequenced one genotype for haplotype for locus for a total of 96 sequences.

At the main time we investigated whether the cpSSR length variations observed were due to SSR variation or other *INDEL* (insertion/deletion) events within the region between the annealing sites, all of the cpSSR loci were identified in the *Castanea sativa* chloroplast DNA (accession number: AY4973), which the most loci was originally used as reference genome (Sebastiani *et al.*, 2004). Only for 5 than 8 total loci used was possible analyzed the sequences, for a quality of fragments resulted.

Each putative amplification product, was blasted against the complete chloroplast genome of *Castanea mollissima* (accession number: HQ336406) and other *Fagaceae* spacies (*e.g. Quercus nugra* and *Fagus sylvatica*). DNA sequences were aligned using the program CLUSTALW ver. 1.74 (Thompson *et al.*, 1994). Putative insertion or deletion events where identified.

3.6 Nuclear SSR analysis

PCR was carried out in the same quantity and mix solutions described for cpSSR. For modification and detail see Table 3.3. The amplification products were loaded in the ABI Prism 310, mono-capillary sequencer using POP4 polymer solution (Applied Biosystems by Life Technologies).

We screened a total of 8 nuSSR primer pairs, published in literature (Table 3.3). The first one primer pairs were designed by Dow *et al.* (1995), on *Q. macrocarpa* and characterized for highly variable $(GA/CT)_n$ microsatellite. Four primer pairs (QpZAG9, QpZAG15, QpZAG36, QpZAG110), were designed by Steinkellner *et al.* (1997a), on *Q. petraea*.

In Steinkellner *et al.* (1997) was demonstrated a conservation of $(GA)_n$ microsatellite loci between a range of oak species (*Quercus* spp.), and other member of the *Fagaceae* family. They observed, for these loci, a highly ability of amplify PCR products and polymorphism range, especially for *Robur* and *Cerris* section in a *Quercus* genus.

The last three primer pairs (QrZAG7, QrZAG11, QrZAG31), were designed and characterized by Kampfer *et al.* (1998) in *Q. robur*. These authors revealed that 32 primers sets successfully amplified PCR products of expected size in both *Q. robur* and *Q. petraea*. 28 of these were informative in the intraspecific mapping cross of *Q. robur* (Kampfer *et al.*, 1998).

Table 3.3 Code, location, repeats, sequences **(5'→3')**, annealing temperature (°C) and size motif (*bp*) of 8 primers pairs for chloroplast microsatellites analysis designed by Weising & Gardner (1999), and Sebastiani *et al.* (2004).

Code	Repeats motif	Primer Sequence (5'→3')[a-b]	T_a (°C)	PCR[b] program	Product size bp[c]	Source
MSQ 4	$(GA/CT)_{17}$	TCTCCTCTCCCCATAAACAGG GTTCCTCTATCCAATCAGTAGTGAG	50	1	216	Dow et al., 1995
QpZAG9	$(AG)_{12}$	GCAATTACAGGCTAGGCTGG GTCTGGACCTAGCCCTCATG	55	1	197	Steinkellner et al., 1997
QpZAG15	$(AG)_{23}$	CGATTTGATAATGACACTATGG CATCGACTCATTGTTAAGCAC	51	1	134	Steinkellner et al., 1997
QpZAG36	$(AG)_{19}$	GATCAAAATTTGGAATATTAAGAGAG ACTGTGGTGGTGAGTCTAACATGTAG	50	1	220	Steinkellner et al., 1997
QpZAG110	$(AG)_{15}$	GGAGGCTTCCTTCAACCTACT GATCTCTTGTGTGCTGTATTT	48	1	242	Steinkellner et al., 1997
QrZAG7	$(TC)_{17}$	CAACTTGGTGTTCGGATCAA GTGCATTTCTTTTATAGCATTCAC	50	2	150	Kampfer et al., 1998
QrZAG11	$(TC)_{22}$	CCTTGAACTCGAAGGTGTCCTT GTAGGTCAAAACCATTGGTTGACT	50	2	273	Kampfer et al., 1998
QrZAG31	$(GA)_{31}$	CTTAGTTTGGTTGGGAAGAT GCAACCAAACAAATGAAAT	50	2	190	Kampfer et al., 1998

T_a, annealing temperature; **PCR**, polymerase chain reaction. **a)** the first sequences is forward and the second is reverse. **b)** PCR conditions: **1**: 5min at 94 °C; 35 cycles of 1min at 94 °C, 1min at T_a (°C), 1min at 72 °C; 10min at 72 °C, 4 °C$_\infty$; **2**: 3min at 95°C; 10 cycles of 15sec at 94 °C, 15sec at 50 °C; 25 cycles of 15sec at 89 °C, 15sec at 50 °C, 4 °C$_\infty$. **c)** expected size in *Q. macrocarpa* for Dow *et al.* (1995) loci, in *Q. petaea* for Steinkellner *et al.* (1997) and in *Q. robur* for Kampfer *et al.*(1998) loci.

3.7 nuSSR data analysis

To identify the number of genetic groups present within the set of the analyzed *Quercus* individuals the STRUCTURE software ver. 2.3.3 (Pritchard and Wen, 2003) has been used. This approach is articulated in two main steps:

a) To determine the most likely number of genetic groups (K) within the considered dataset.

b) To assign the individuals to the K genetic groups minimizing (as far as possible).

The most likely number of genetic groups within the set of *Quercus* individuals was determined following the procedure of Evanno *et al.* (2005). The model choice criterion implemented in structure to detect the K is an estimate of the posterior probability of the data for a given K, called 'Ln P(D)' (Pritchard *et al.*, 2000).

True number of populations (K) is often using the maximal value of L(K) returned by structure. However, Evanno *et al.*, (2005) reviewed that, in most cases, the distribution of L(K) is not appropriate to determine the true K. They proposed an *ad hoc* statistics based on the second order rate of change of likelihood function with respect to K (ΔK) that show a clear peak at true value of K. The rational for this ΔK was to make salient the break in slope of the distribution of L(K) at the true K, i.e. the value of K upon which gathering improved estimates becomes irrelevant. Most of parameters were set to their default values (Pritchard and Wen, 2003).

Specifically, initially, we choose the admixture model and the option of correlated allele frequencies among all populations (both *Q. suber* genotypes and the others *Quercus* species) which is the best choice when subtle population structure is present (Evanno *et al.*, 2005). In the second time, we checked separately the admixture model for a group A (only *Q. suber* genotypes) and a group B (more *Quercus* species, i.e., *Q. coccifera*, *Q. congesta, Q. ilex, Q. morisii* and *Q. pubescens*).

The degree of genetic divergence among populations was assessed by AMOVA (1000 permutations) using ARLEQUIN software version 3.5.1.2 (Excoffier *et al*. 2005).

The genetic differentiation among genetic groups inferred by STRUCTURE was estimated by hierarchical analysis of molecular variance. The degree of genetic divergence among population was evaluated as F_{ST} (Wright *et al.*, 1951).

3.8 Phenotypic assessment

Q. suber (common garden Monte Grighini collection) was also explored for some phenotypic characters.

The shape of trees of 27 *Q. suber* Mediterranean basin populations was evaluated according to Bellarosa *et al.* (2000) method. Five randomized block (137 MGC individuals) were monitored. We also observed the bud set and bud flush in two different years (2009 and 2010). The measures were performed from April to June every 3-4 days.

3.9 Phenotypic data analysis

The shape of trees differences were evaluated with contingence analysis by a Chi-Square test using JMP.7 software (SAS Institute, Cary, USA) (Heer *et al.*, 2006). ANOVA statistical analysis by JMP.7 was made to monitored bud set and bud flush differences among populations from different countries.

4. RESULTS

4.1 Chloroplast SSR diversity

Eight cpSSR primer sets were tested in 6 *Quercus* genera of *Fagaceae* family. All loci considered gave amplification product of all six genera described. In Table 4.1 are reported the size (bp) of the amplification products for each primer sets of all different species analyzed. All the loci considered were polymorphic in *Q. suber*. The primer pairs Ccmp4 detected polymorphism only within *Q. suber* individuals. In addition, for Ccmp6 primer pairs, it was observed that allele (102bp) was absent in *Q. suber* genotypes. Contrariwise, for a Cmcs9 locus, we detected 3 unique alleles (80, 118, 125bp), present only in a *Q. suber* samples collected (Table 4.1). In addition, for this locus, the allele 80bp correspond an insertion/deletion (≈30bp), were detected only for Italian genotypes (confirmed by Magri *et al.* 2007).

Table 4.1 Size (bp) of the amplification products for each of the 8 primers pairs and for each species considered.

Species	Samples size	Ccmp 4	Ccmp 6	Cmcs 1	Cmcs 6	Cmcs 7	Cmcs 8	Cmcs 9	Cmcs 14
Q. suber	151	117 119 122	104 106 108	106 112 115	197 202 207	208 210 212	180 181 182	80 114 118 120 125	172 174 176
Q. ilex	3	119	102 106	115	197	208	182	114	174
Q. coccifera	1	119	106	115	202	208	181	120	174
Q. congesta	3	119	102	112	202	210	182	120	176
Q. morisii	2	119	106	115	197 202	208	181	120	174
Q. pubescens	2	119	106	112	202	210	182	120	176
Total	**162**	**3**	**4**	**3**	**3**	**3**	**3**	**5**	**3**

In Table 4.2 were reported statistics diversity parameters considering separately MGC (without Sardinian accessions) and SQS (only *Q. suber* Sardinian accessions). Firstly, the highest degree of diversity was observed in Ccmp6 with 3 alleles detected and a gene diversity (*H*) of 0.519. The lowest level of diversity was shown in Cmcs8 with 3 alleles detected and an *H* of 0.216. The average number of alleles per locus was 3.25±0.707 and the average *H* was 0.370±0.100 (considering all polymorphic loci in the whole sample).

The lowest level of diversity was identified among *Q. suber* Sardinian accessions where only two primers (Ccmp6 and Cmcs6) were resulted polymorphic (*H*: 0.181, 0.173 respectively).

Table 4.2 Statistics diversity per locus: n_a = number of alleles; *H* = genetic diversity (Nei's, 1978) on the whole sample analyzed , on *Q. suber* accessions.

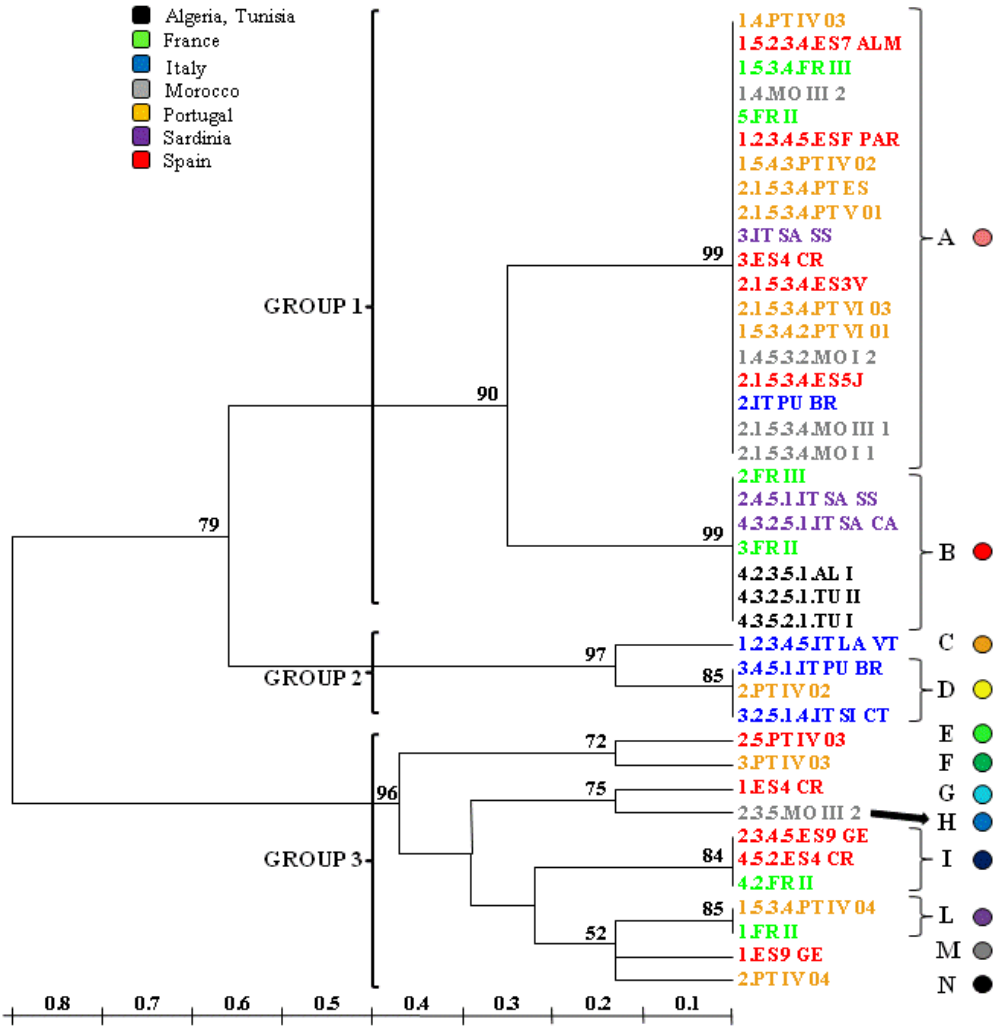
	MGC			**SQS**		
Locus	Sample Size	n_a	*H*	Sample Size	n_a	*H*
Ccmp4	126	3	0.391	23	1	0.00
Ccmp6	126	3	**0.519**	23	2	0.181
Cmcs1	126	3	0.328	23	1	0.00
Cmcs6	126	3	0.427	23	2	0.173
Cmcs7	126	3	0.288	23	1	0.00
Cmcs8	126	3	**0.216**	23	1	0.00
Cmcs9	126	5	0.439	23	1	0.00
Cmcs14	126	3	0.279	23	1	0.00
Mean	126	3.25	0.370	23	1.25	0.177
St. Dev		0.707	0.100			0.005

4.2 Relationship among genotypes analyzed

In Figure 4.2 are shown the relationship among genotypes analyzed. In particular, the dendrogram supports the relation among *Q. suber* accessions sampled in the Monte Grighini collection (MGC).

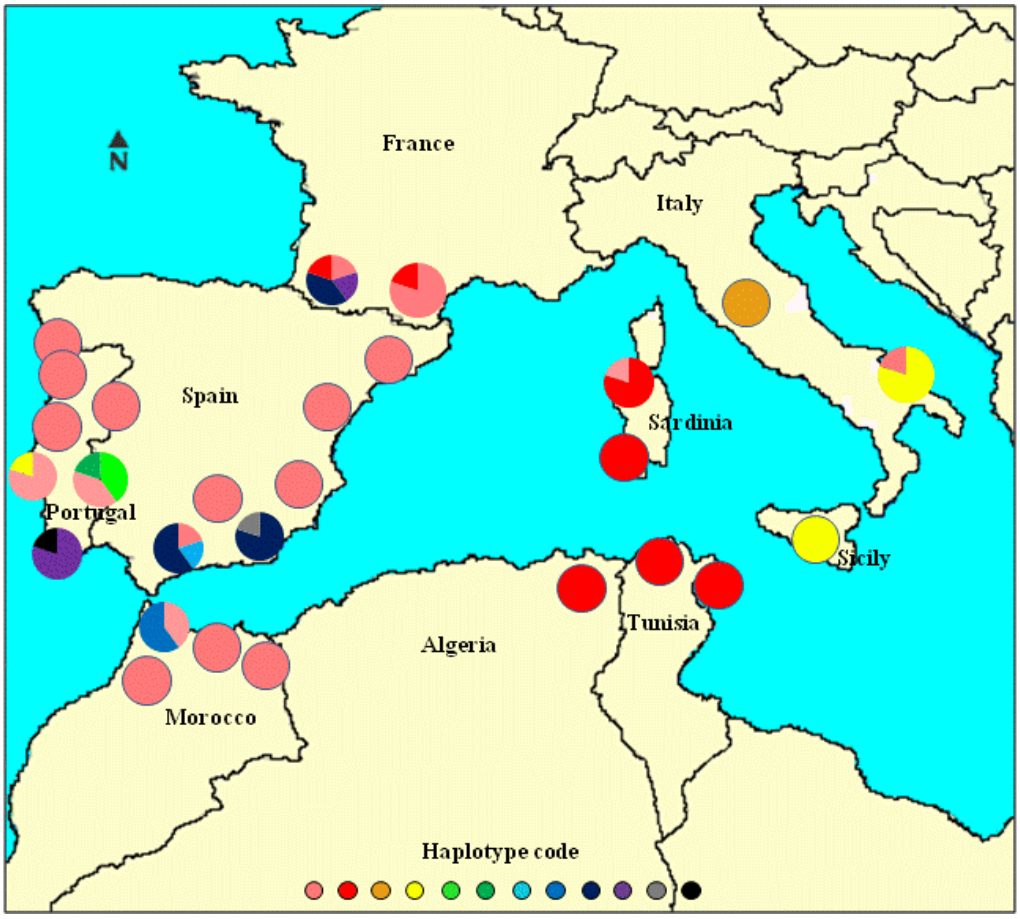
Figure 4.2 UPGMA cluster analysis on cpSSR MGC. In the noodles is written the bootstrap value (in percentage). The name (different for any population) and the color of each individuals indicate the different origin (region or country) (see Table 3.1). The number near to the name indicate the different block (5 total replications).

Genetic Similarity (Nei and Li, 1979).



Most populations are fixed for haplotype A (pink in the Figure 4.3), and few populations contain different haplotypes (see Figure 4.3). In fact, in the haplotype A found, more than half of accessions considered (≈52%). Sequence data confirmed (for the five loci analyzed) that the detected variation is due to differences in the number of repeats within the microsatellite stretches.

Figure 4.3 Distribution of the *Q. suber* populations in the Mediterranean basin. Different circles describe the 27 populations monitored. Different colors (pink to black), represent different haplotypes. For the correspondence haplotype code see circle on the right in Figure 4.2.
Note: the position of the circle does not exactly correspond with the site origin collection.



Population genetic structure was performed on 137 *Q. suber* individuals.

Considering the threshold at 0.4 genetic distance (Nei and Li, 1979), in dendrogram describe above, the genotypes were grouped in 3 principal groups.

Group 1 included haplotypes A and B with 73 and 26 individuals respectively. Group 2 is made by Italian haplotypes (C and D) for a total of 14 individuals. Group 3 contained 23 genotypes articulated on 8 haplotypes (E to N) (see Figure 4.2). Globally 83.1% were assigned to the Group 1 and 2 (haplotypes A to D) and 16.9% to the Group 3 (haplotypes E to N). The Western Mediterranean Basin (from France to Morocco) resulted the most variegate, containing ten different haplotypes, in particular the haplotype A was the most present. In the Italian peninsula and Sicily, two closely related haplotypes (C and D) were detected. Haplotype B is found only in the central Mediterranean range, from Sardinia to North-East Africa (Algeria and Tunisia), and in France (in this case, few accessions are present in the collection).

In Figure 4.4 is displayed the relationship among all species analyzed, for a total of 162 genotypes. In particular, this dendrogram support the relationship between *Q. suber* Monte Grighini collection (black name in the describe) and other different *Quercus* accessions sampled in a central Sardinia area. An identificative colour was assigned to each species (see caption of Figure 4.4).

Using the same threshold described in Figure 4.2, all genotypes considered were subdivided in five principal groups. Here (Fig. 4.4) are described the Sardinian genotypes (underline with colours), sampled outside MGC site.

In black are indicated the MGC individuals already described in figure 4.2. Group 1 included 14 *Q. suber* individuals sampled in the center Island of Sardinia. Group 2 is made by 3 *Q. ilex*, 1 *Q. coccifera* and 2 *Q. morisii*. Group 3 contained only MGC Italian *Q. suber* genotypes and a Portuguese. Group 4 have included 2 *Q. pubescens* and 3 *Q. congesta* genotypes, and finally, Group 5 is characterized by 22 MGC *Q. suber* genotypes.

The principal groups identified in Fig. 4.4 is compatible with the phylogeny differentiation of *Quercus* genus showed by Manos *et al.*, (1999).

In particular, molecular cpDNA data, supported the recognition of *Cerris* section (in Eurasian area), articulated in two sub-sections: "*Cerris* group" containing *Q. suber* and "*Ilex group*", containing *Q. ilex* and *Q. coccifera* and in our case *Q. morisii* (an hybrid *Q. suber* x *Q. ilex*).

The other two species *Q. pubescens* and *Q. congesta*, are collocated in a "*Robur* group", section *Quercus* (Eurasian distribution) (Tutin *et al.*, 1993; Steinkellner *et al.*, 1997).

Group 1, displayed 3 different haplotypes. The first two, contained *Q. suber* MGC individuals, with 6 SQS genotypes attacked by the phytophagous (names are indicated in red).The third haplotype contained *Q. suber* Algerian, Tunisian and Sardinian MGC plus, the last 8 SQS genotypes not attacked by the phytophagous (names are indicated in orange).

Genetic similarity (Nei and Li, 1979).

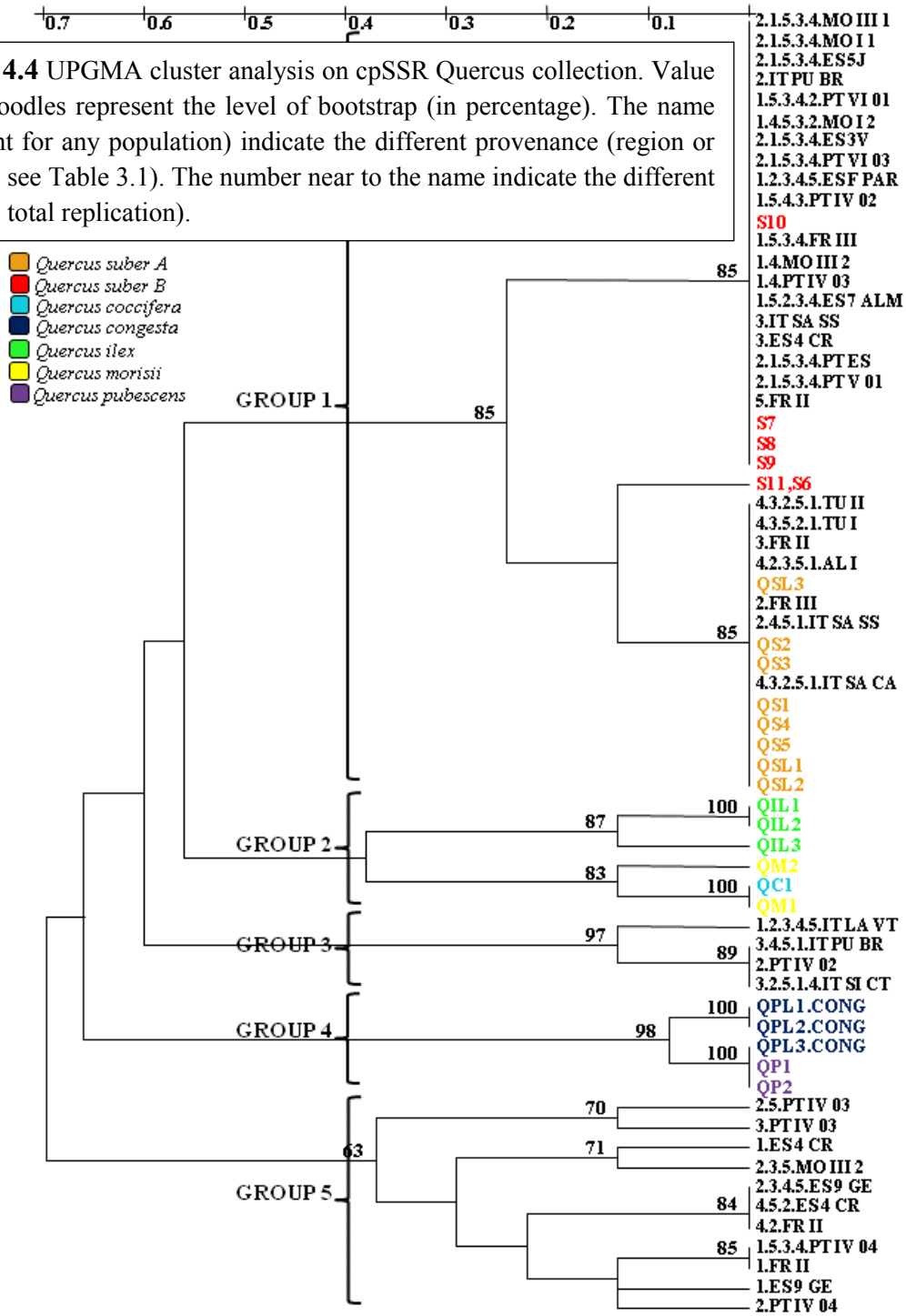


Figure 4.4 UPGMA cluster analysis on cpSSR Quercus collection. Value at the noodles represent the level of bootstrap (in percentage). The name (different for any population) indicate the different provenance (region or country, see Table 3.1). The number near to the name indicate the different block (5 total replication).

The portioning of the total genetic variance, calculated as F_{ST} value, of *Q. suber* MGC evaluated by AMOVA is shown in Table 4.3. The hierarchic AMOVA using all MGC accessions, articulated in 27 populations, showed a general high variance level (62.39%), especially among populations within Mediterranean basin countries. In the main time, significant variation, among country and within population (13.99 and 23.63 respectively) was found.

Table 4.3 AMOVA analysis on cpSSR markers, for MGC. Distance method based of number of different alleles (F_{ST}).

Source of variation	D.F.	S.S.	Variance components	Variation (%)	Fixation indices
Countries	6	56,31	0,220	13,99	FCT: 0.140
Among populations	20	106,08	0,980	62,39	FSC: 0.725
Within populations	111	40,47	0,371	23,63	
Total	137	202,85	1,571		**FST: 0.764**

4.3 Sequences data

Considering the limitate amount of *Q. suber* sequence information, in this study, we performed cpSSR sequence analysis based on *Castanea mollissima* (Figure 4.5), because these species belongue to the *Fagaceae* family.

The cpSSR loci positions were designed using the *Castanea* as reference genome. The Ccmp4 primer pairs was built in *Nicotiana tabacum,* but however, we observed an homology region in *Castanea* genome (red block in Figure 4.5).

The alignment between *Quercus suber* and the *Fagaceae* species (*Castanea mollissima*, *Castanea sativa*, *Fagus sylvatica* and *Quercus nigra*) sequences, showed a cpSSR polymorphism based on the SSR repeats motif (Table 4.4). At the main times few loci showed SNPs, and insertion-deletion (indel) events within microsatellites flanking region. Each locus explored, information about indel and SNPs positions , are shown in supplementary material A1.

Figure 4.5 Chloroplast DNA of *Castanea mollissima* genome. The black blocks represent the position of genes on the genome. The blue and red block define the investigated loci positions in that genome.

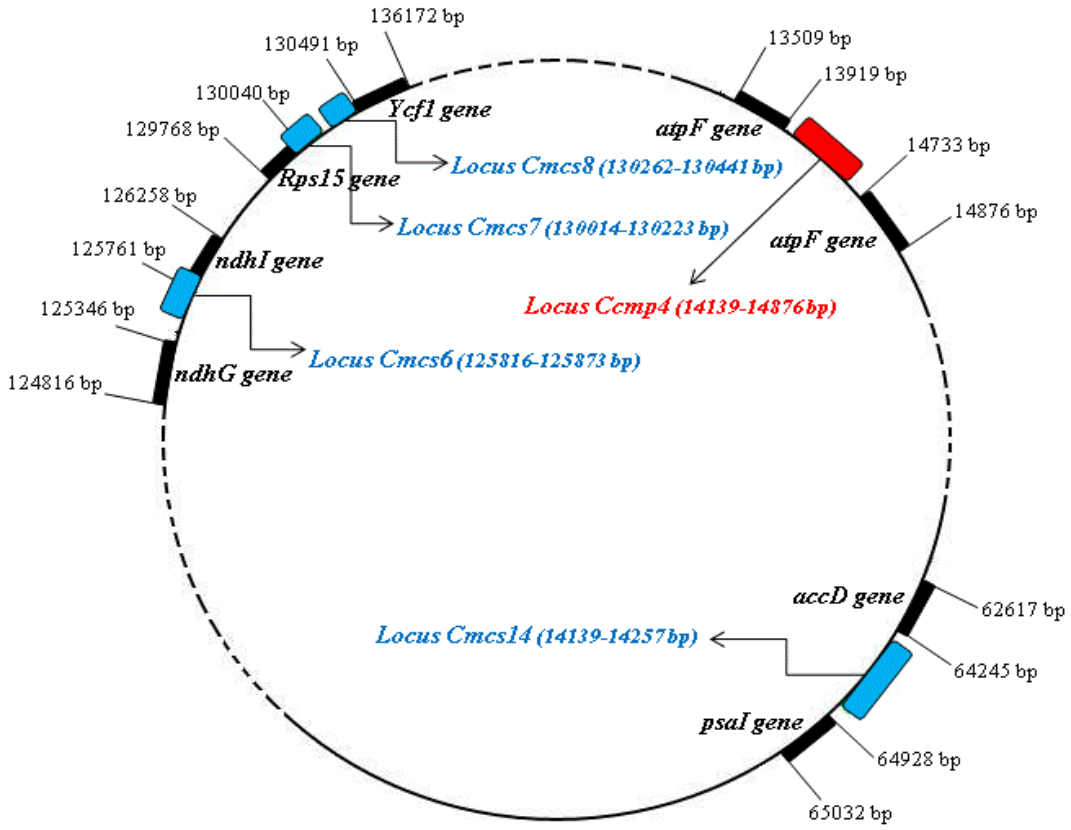


Table 4.4 SSR motif in *Quercus suber* and the other *Fagaceae* blasted. In blue deletions, in red insertion within the original microsatellites motif.

Loci	Repeats in *Q. suber* [a]	Identities (%)	*C. mollissima*	Identities (%)	*C. sativa*	Identities (%)	*Q. nigra*	Identities (%)	*F. sylvatica*	Identities (%)
Ccmp4[b]	$(T)_{13}$	90	$(T)_{11}$ $(T)_2$	96	-	-	$(T)_{12}$ C	95	$(T)_{11}$ $(T)_2$	96
Cmcs6[c]	$(A)_{13}$	67	$(A)_{12}$ A	95	$(A)_{10}$ $(A)_3$	94	-	-	-	-
Cmcs7[c]	$(A)_{14}$	98	$(A)_{12}$ $(A)_2$	99	$(A)_{12}$ $(A)_2$	96	-	-	-	-
Cmcs8[c]	$(T)_{11}$ $A(T)_8$	97	$(T)_{10}$ $TA(T)_6(T)_2$	97	$(T)_{10}$ $TA(T)_6(T)_2$	97	-	-	-	-
Cmcs14[c]	$(AT)_7$	73	$(AT)_7$	99	$(AT)_7$	99	-	-	-	-

a) In these case the repeats in the microsatellites motif were defined (within the twelve genotypes analyzed), for the biggest sequence (sample size in bp). **b)** this primer have been designed in *N. tabacum.* **c)** these primers pairs have been designed *in C. sativa.*

Two examples of *Q. suber* alignment with *Fagaceae* species are shown in figure 4.6 A and B. In Figure 4.6 A is displayed the alignment for the locus Ccmp4 (Weising and Gardner, 1999).

We observed that the microsatellites motif is always present (orange underlined). Few SNPs and indel events among species (i.e. *C. mollissima* and *Q. nigra* with *F. sylvatica* and *Q. suber*), and within different *Q. suber* genotypes were also detected. In fact, differences among *Quercus suber* genotypes in repeats number of microsatellites motif (T_{13} and T_9 respectively) were observed. The data is even more evident comparing Italian and Spanish-Portuguese genotypes.

Primer set Cmcs6 (Sebastiani *et al.*, 2004) alignment was reported in Figure 4.6 B. As shown for Ccmp4, the microsatellites motif (A_n) was always detected, (orange underlined). We observed a SNP between C*astanea* and *Quercus* species.

Among *Q. suber* genotypes an insertion event was noted, in correspondence of Italian haplotypes, and a big deletion for a Spanish (ES7ALM-IV) haplotype was also clearly evident in the alignment. In general we evaluated that Italian haplotypes have the longest microsatellites motif.

All the data about microsatellite and haplotypes were verified using two different methods: polyacrilamide gels and sequencing analysis. We found that the sequencing analysis discriminated more haplotypes.

Figure 4.6 A *Q. suber* L. haplotypes (12) alignment and Fagaceae species (3). On the left side is indicated the *Fagaceae* species names. The *Q. suber* genotypes analyzed were described with different color for any country and region (*i.e.* **red** for Spain, **purple** for Sardinia, etc.). On the right were described the sequence lenght (bp). In **blue**, primer pairs ccmp4 designed by Weising and Gardner (1999); in **green** the SNPs and insertion events; **orange underlined**: microsatellites motif.

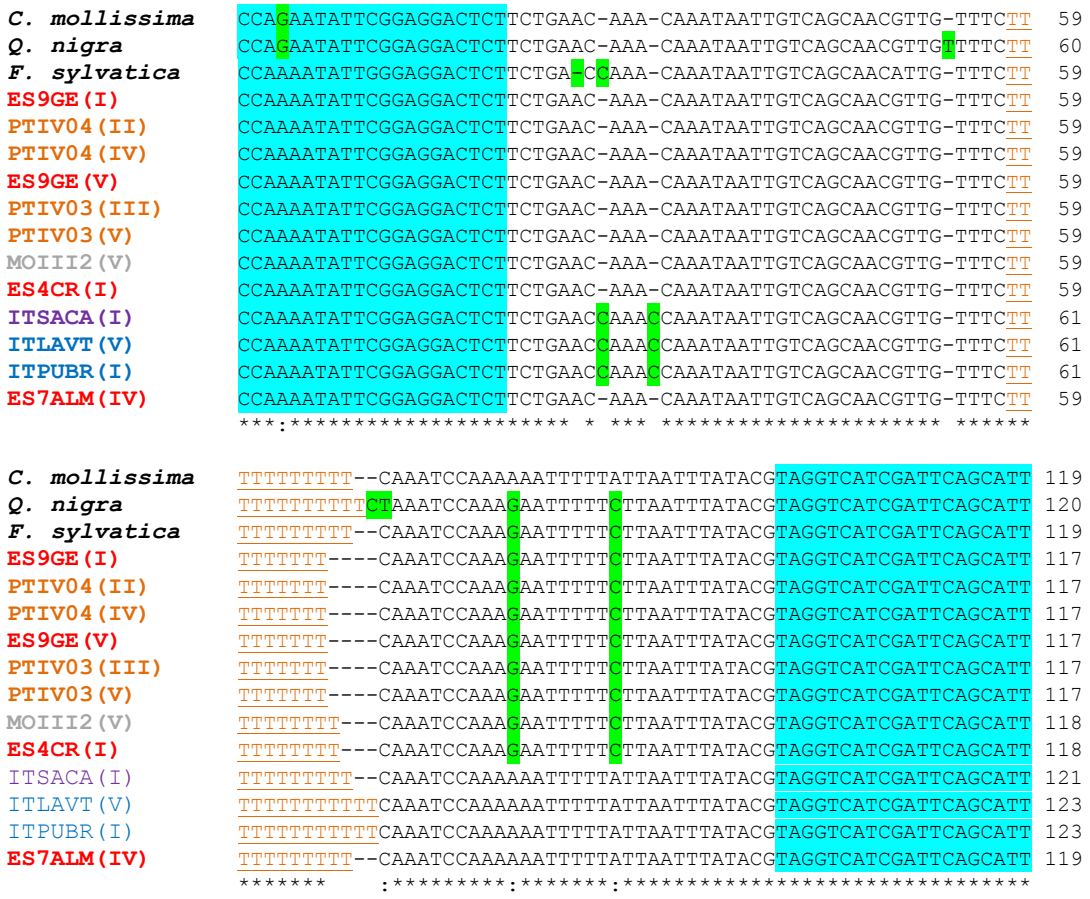


Figure 4.6 B *Q. suber* L. haplotypes (12) alignment and *Fagaceae* species (2). On the left side is indicated the *Fagaceae* species names. The *Q. suber* genotypes analyzed were described with different color for any country and region (*i.e.* **orange** for Portugal, grey for Morocco, etc.). On the right were described the sample sizes (bp). In **blue**, primer pairs Cmcs6 designed by Sebastiani *et al.* (2004); in **green** the SNPs and insertion events; **orange underlined**: microsatellites motif.

```
C. mollissima   CTTATGATCGTCACGAATTGAATTATAATCAAATTGCCTTGGGTCGTTTACCGATGTCAGTAATTGACGA  70
C. sativa       CTTATGATCGTCACGAATTGAATTATAATCAAATTGCCTTGGGTCGTTTACCGATGTCAGTAATTGACGA  70
ES9GE(I)        CTTATGATCGTCACGAATTGAATTATAATCAAATTGCCTTGGGTCGTTTACCGATGTCAGTAATTGACGA  70
PTIV04(II)      CTTATGATCGTCACGAATTGAATTATAATCAAATTGCCTTGGGTCGTTTACCGATGTCAGTAATTGACGA  70
PTIV04(IV)      CTTATGATCGTCACGAATTGAATTATAATCAAATTGCCTTGGGTCGTTTACCGATGTCAGTAATTGACGA  70
ES9GE(V)        CTTATGATCGTCACGAATTGAATTATAATCAAATTGCCTTGGGTCGTTTACCGATGTCAGTAATTGACGA  70
MOIII2(V)       CTTATGATCGTCACGAATTGAATTATAATCAAATTGCCTTGGGTCGTTTACCGATGTCAGTAATTGACGA  70
ES4CR(I)        CTTATGATCGTCACGAATTGAATTATAATCAAATTGCCTTGGGTCGTTTACCGATGTCAGTAATTGACGA  70
PTIV03(III)     CTTATGATCGTCACGAATTGAATTATAATCAAATTGCCTTGGGTCGTTTACCGATGTCAGTAATTGACGA  70
PTIV03(V)       CTTATGATCGTCACGAATTGAATTATAATCAAATTGCCTTGGGTCGTTTACCGATGTCAGTAATTGACGA  70
ITLAVT(V)       CTTATGATCGTCACGAATTGAATTATAATCAAATTGCCTTGGGTCGTTTACCGATGTCAGTAATTGACGA  70
ITPUBR(I)       CTTATGATCGTCACGAATTGAATTATAATCAAATTGCCTTGGGTCGTTTACCGATGTCAGTAATTGACGA  70
ES7ALM(IV)      CTTATGATCGTCACGAATTGAATTATAATCAAATTGCCTTGGGTCGTTTACCGATGTCAGTAATTGACGA  70
ITSACA(I)       CTTATGATCGTCACGAATTGAATTATAATCAAATTGCCTTGGGTCGTTTACCGATGTCAGTAATTGACGA  70
                ************************************************************************

C. mollissima   TTATACAATTCGAACAATTTTAAATTCACCTAAAAT------AAAATAAAAAAAAAAA--TAAGTAAATC  132
C. sativa       TTATACAATTCGAACAATTTTAAATTCACCTAAAAT------AAAATAAAAAAAAA---TAAGTAAATC  131
ES9GE(I)        TTATACAATTCGAACAATTTTAAATTCACCTAAAAT------AAAATAAAAAAA-----GAAGTAAATC  129
PTIV04(II)      TTATACAATTCGAACAATTTTAAATTCACCTAAAAT------AAAATAAAAAAA-----GAAGTAAATC  129
PTIV04(IV)      TTATACAATTCGAACAATTTTAAATTCACCTAAAAT------AAAATAAAAAAA-----GAAGTAAATC  129
ES9GE(V)        TTATACAATTCGAACAATTTTAAATTCACCTAAAAT------AAAATAAAAAAA-----GAAGTAAATC  129
MOIII2(V)       TTATACAATTCGAACAATTTTAAATTCACCTAAAAT------AAAATAAAAAAA-----GAAGTAAATC  129
ES4CR(I)        TTATACAATTCGAACAATTTTAAATTCACCTAAAAT------AAAATAAAAAAA-----GAAGTAAATC  129
PTIV03(III)     TTATACAATTCGAACAATTTTAAATTCACCTAAAAT------AAAATAAAAAAA-----GAAGTAAATC  129
PTIV03(V)       TTATACAATTCGAACAATTTTAAATTCACCTAAAAT------AAAATAAAAAAA-----GAAGTAAATC  129
ITLAVT(V)       TTATACAATTCGAACAATTTTAAATTCACCTAAAATTAAAATAAAATAAAAAAAAAAA-GAAGTAAATC  139
ITPUBR(I)       TTATACAATTCGAACAATTTTAAATTCACCTAAAATTAAAATAAAATAAAAAAAAAAA-GAAGTAAATC  139
ES7ALM(IV)      TTATACAATTCGAACAATTTTAAATTCACCTAAAA----------ATAAAAAAAAAAAAGAAGTAAATC  130
ITSACA(I)       TTATACAATTCGAACAATTTTAAATTCACCTAAAAT------AAAATAAAAAAAAAAGAAGTAAATC  134
                ********************************          **********     :*********

C. mollissima   CTTTGATTAAAGATATTAAAGATAAAGAAATGGTAATTTCAAATTCAAATTAGGAAGGGTCCTTTTTC  200
C. sativa       CTTTGATTAAAGATATTAAAGATAAAGAAATGGTAATTTCAAATTCAAATTAGGAAGGGTCCTTTTTC  199
ES9GE(I)        CTTTGATTAAAGATATTAAAGATAAAGAAATGGTAATTTCAAATTCAAATTAGGAAGGGTCCTTTTTC  197
PTIV04(II)      CTTTGATTAAAGATATTAAAGATAAAGAAATGGTAATTTCAAATTCAAATTAGGAAGGGTCCTTTTTC  197
PTIV04(IV)      CTTTGATTAAAGATATTAAAGATAAAGAAATGGTAATTTCAAATTCAAATTAGGAAGGGTCCTTTTTC  197
ES9GE(V)        CTTTGATTAAAGATATTAAAGATAAAGAAATGGTAATTTCAAATTCAAATTAGGAAGGGTCCTTTTTC  197
MOIII2(V)       CTTTGATTAAAGATATTAAAGATAAAGAAATGGTAATTTCAAATTCAAATTAGGAAGGGTCCTTTTTC  197
ES4CR(I)        CTTTGATTAAAGATATTAAAGATAAAGAAATGGTAATTTCAAATTCAAATTAGGAAGGGTCCTTTTTC  197
PTIV03(III)     CTTTGATTAAAGATATTAAAGATAAAGAAATGGTAATTTCAAATTCAAATTAGGAAGGGTCCTTTTTC  197
PTIV03(V)       CTTTGATTAAAGATATTAAAGATAAAGAAATGGTAATTTCAAATTCAAATTAGGAAGGGTCCTTTTTC  197
ITLAVT(V)       CTTTGATTAAAGATATTAAAGATAAAGAAATGGTAATTTCAAATTCAAATTAGGAAGGGTCCTTTTTC  207
ITPUBR(I)       CTTTGATTAAAGATATTAAAGATAAAGAAATGGTAATTTCAAATTCAAATTAGGAAGGGTCCTTTTTC  207
ES7ALM(IV)      CTTTGATTAAAGATATTAAAGATAAAGAAATGGTAATTTCAAATTCAAATTAGGAAGGGTCCTTTTTC  198
ITSACA(I)       CTTTGATTAAAGATATTAAAGATAAAGAAATGGTAATTTCAAATTCAAATTAGGAAGGGTCCTTTTTC  202
                ************************************************************************
```

4.4 Nuclear SSR diversity

Eight nuSSR primers were tested in the same collection describe above, for a total of 162 accessions. In particular 6 different *Quercus* genera were analyzed.

A successful amplification products were obtained for all of the 6 genera studied. Only 2 primers pairs (QrZAG11, QrZAG31) did not amplified the 11% of samples.

All loci considered were polymorphic in *Q. suber* (MGC), and in the other *Quercus* species considered. In Table 4.5 is reported the samples lenght (bp) of the amplification products for each locus in the different sampled species..

Primers QpZAG36 and QpZAG110 detected the highest level of polymorphism, with 22 alleles for each locus. Interestingly QpZAG110 primer amplification displayed a big insertion (≈60 bp).

All of the 24 *Q. suber* individuals displayed the "QpZAG110 insertion". In addition a significant samples length difference between *Q. ilex* and most of *Quercus*, was observed. Less polymorphism level was shown in QrZAG31, with a total of 9 alleles.

In Table 4.5 were described all the alleles scored in *Q. suber* accessions. In particular the first section of the table showed about 7.5 unique alleles for *Q. suber* genotypes (blue).

QpZAG110 SSR locus identified the largest number of specie-specific alleles for *Q. suber*. This locus is extremely informative and could be used for future *Quercus* genetic diversity analysis.

The second part of table 4.5 explored the *Q. ilex*, *Q. coccifera*, *Q. congesta*, *Q. morisii* and *Q. pubescens* SSR locus diversity (red). In particular *Q. coccifera* and *Q. morisii* contained a low number of unique alleles. Instead, *Q. congesta* and *Q. pubescens* showed to be extremely polymorphic with a high number of specie-specific alleles.

Table 4.5 Size (bp) of the amplification products for each of the 8 nuclear SSRs loci and for each species considered. In **blue**, specie-specific alleles detected for *Q. suber*. In **red**, *Quercus* species-specific alleles.

Genus	Species	Samples size	QpZAG9	QpZAG36	QpZAG15	MSQ4	QrZAG11	QrZAG7	QrZAG31	QpZAG110
Quercus	***Q. suber***	151	**219**, 220, **223**, **229**, 230, **233**, 234, **240**, **242**, 246	**207**, **208**, 209, **210**, **212**, **213** **214**, 215, **216**, 217, **218**, 219, 220, **221**, **222**	**101**, **102**, **115**, 117, **118**, **120**, **121**, **123**, 125, **132**, **133**	193, **194**, **197**, **198**, **199**, 201, 207, **208**, **209**, 211, **213**	258, 260, **262**, **264**, **270**, **272**, **274**, **276**	113, **115**, **117**, 119, 121, **123**, 125, **127**, 129, 131	134, **135**, **136**, **142**, 143	**210**, **215**, 217, 219, **221**, **223**, **225**, **227**, **229**, 230, **232**, **233**, **234**, **235**, **236**, **239**, **286**
	Q. ilex	3	230, 234, **239** 246	**202**, **205**	**111**, 117, **119**	193	**244**, **252**, **256**	113	**141**	**345**, **354**
	Q. coccifera	1	246	**200**, 209	**107**, 125	193	260	113	n.a.	n.a.
	Q. congesta	3	**190**, **196**, 220 **241**	**199**, **201**, 215, **223**	**104**, **106**	**203**, 207, 211, **212**	**250**, 258	**111**, 119, 121, 129, **139**	**145**, **149**	**203**, **213**
	Q. morisii	2	220	**200**, **206**, 215, 219	117	193, 207	**244**, 260	**111**, 125, 131	134, 143	217, 230
	Q. pubescens	2	**182**, **194**, **241**	217, 220	**104**, **106**, **112**	201, **214**	**246**, **253**	121, **137**, **146**	**145**, **152**	**203**, 217, 219, **226**
Total		162	17	22	18	14	14	14	9	22

In Table 4.6 is shown statistic diversity parameters considering MGC (without Sardinian accessions) and SQS (only *Q. suber* Sardinian accessions) separately.

In the first group, the highest degree of diversity was observed in QpZAG110 with 16 alleles and a gene diversity value (*H*) of 0.853. The lowest level of diversity was shown in QrZAG11 with 8 alleles and *H* of 0.506. The average number of alleles per locus was 10.88±3.56 and the average *H* was 0.663±0.126 (considering all polymorphic loci in the whole samples).

In the second group, the most polymorphic primer was QpZAG110 with 11 alleles and *H* of 0.496. The lowest level of diversity was found in QrZAG11 with 3 alleles and *H* of 0.188. The average number of alleles per locus was 5.88±3.27 and the average *H* was 0.318±0.129 (considering all polymorphic loci in the whole *Q. suber* Sardinian samples).

In order to evaluate the general genetic diversity difference between MGC and SQS we calculated Δ*H***.**

Despite the small n_a detected in both collections for the locus QrZAG31 (5; 3), we found that the QrZAG31 showed the highest Δ*H* value (0.420).

The opposite trend is found for QpZAG15, where the Δ*H* value (0.274) is the lowest score of genetic diversity difference between MGC and SQS. In this case we also found that the n_a value is extremely different between MGC (12) and SQS (4).

Table 4.6 Statistics diversity per locus: n_a = number of alleles; *H* = genetic diversity (Nei's, 1978) on the whole sample analyzed, on the *Q. suber* , separately for MGC and SQS.

	MGC			SQS			
Locus	**Sample Size**	**n_a**	***H***	**Sample Size**	**n_a**	***H***	***ΔH***
QpZAG9	127	10	0.597	24	4	0.253	*0.344*
QpZAG36	116	15	0.783	24	10	0.468	*0.315*
QpZAG15	127	12	0.537	24	4	0.263	***0.274***
MSQ4	123	11	0.637	22	4	0.233	*0.404*
QrZAG11	113	8	0.506	19	3	0.188	*0.318*
QrZAG7	127	10	0.770	24	8	0.377	*0.394*
QrZAG31	114	5	0.618	21	3	0.198	***0.420***
QpZAG110	124	16	0.852	24	11	0.497	*0.356*
Mean	121.38	10.88	0.663	22.75	5.88	0.318	*0.353*
St. Dev.		3.56	0.126		3.27	0.129	*0.050*

4.5 Population genetic structure

Population genetic structure was performed on 162 *Quercus* individuals. Based on 8 nuclear SSRs loci, STRUCTURE revealed a break in the slope of likelihood value at *K=2* (Figure 4.7), suggesting that the genetic structuring of *Quercus* Mediterranean genus monitored, should assume two main clusters-populations.

The obtained results suggest that the *Quercus* germplasm evaluated can be divided in two genetically distinct clusters and the spatial information did not influence the population clustering (Figure 4.8).

This results revealed the genetic difference between *Q. suber* and other *Quercus* species analyzed.

To investigate a genetic differentiation present within *Q. suber* populations (MGC) and among other *Quercus* analyzed (SQS), we decided to consider the STUCTURE *Qi* (coefficient of membership). Conventionally, we have grouped genotypes where a membership value (*Qi*) resulted ≥0.70. Globally 92.4% of individuals were assigned to the Cluster A (C.A.), and 7.5% to the Cluster B (C.B.) (see Figure 4.8).

Figure 4.7 ΔK as a function of the number of genetic groups, considering all collection.

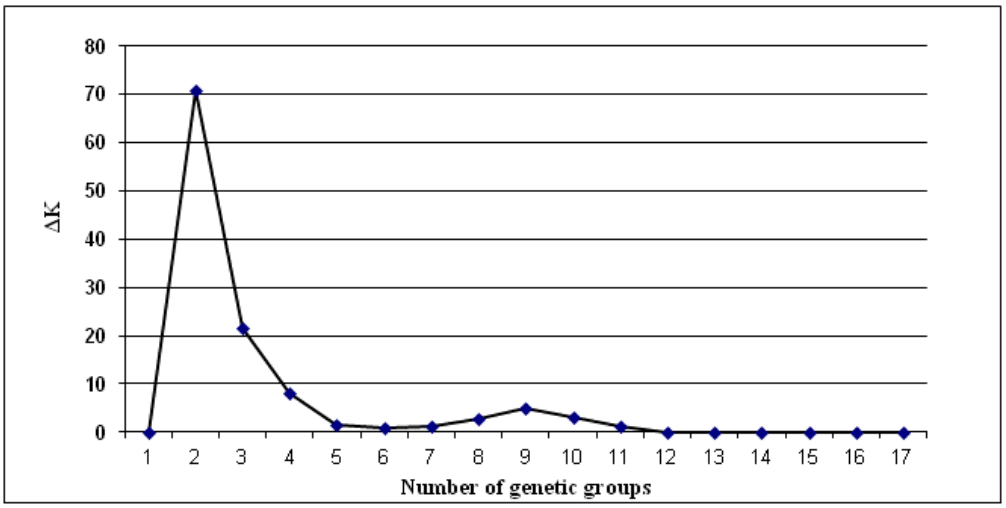
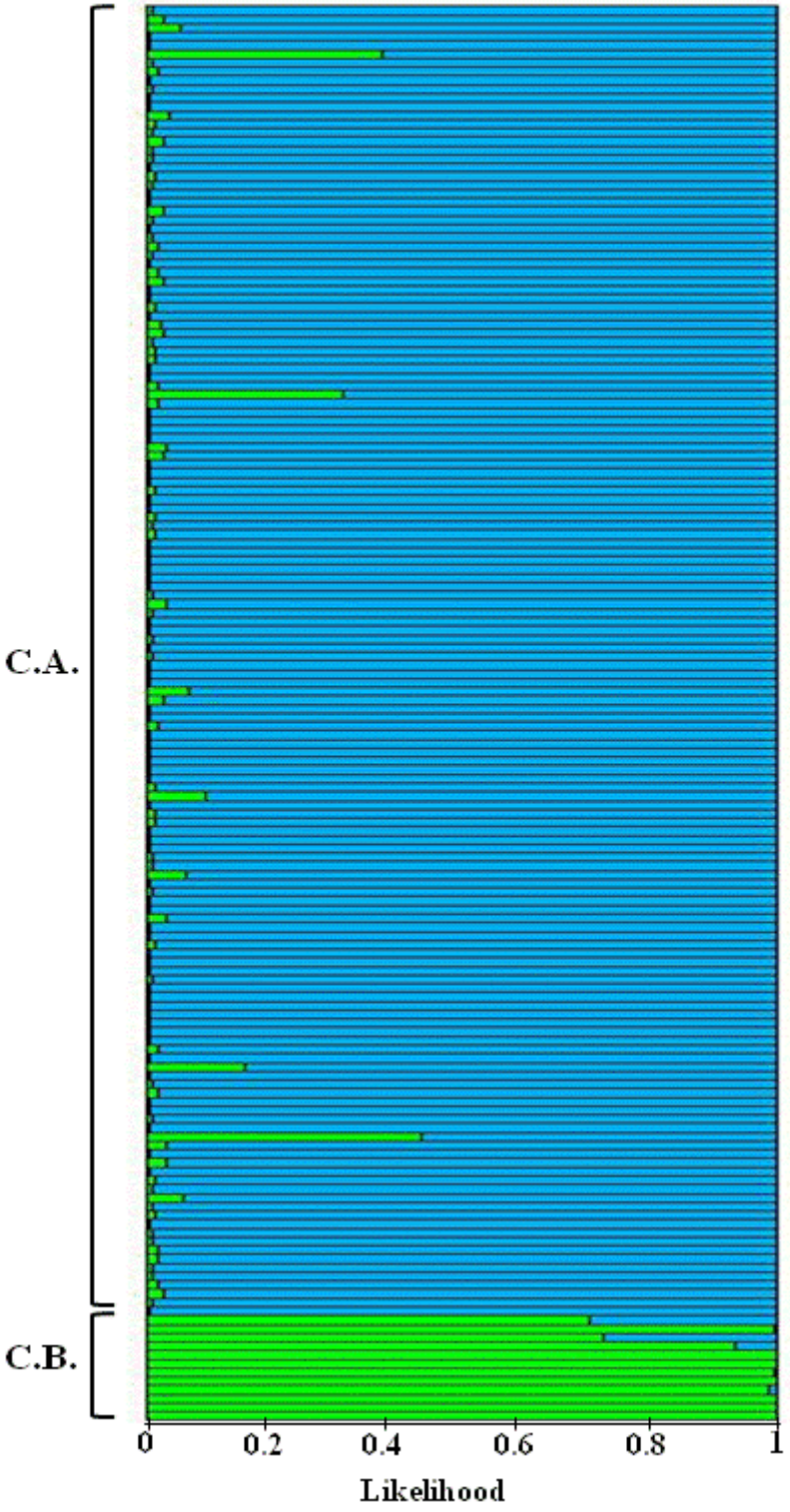


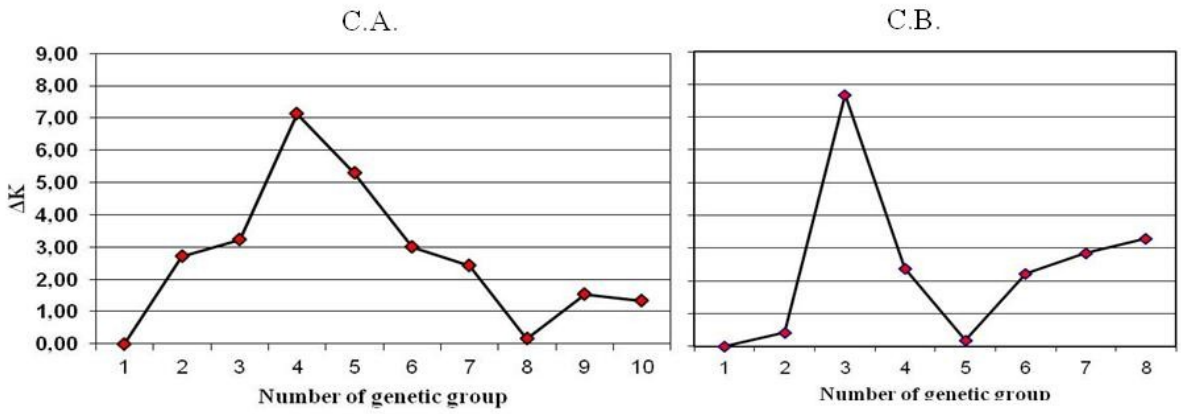
Figure 4.8 Genetic relationship among all 162 *Quercus* individuals analyzed. Different colors indicate the group identified by STRUCTURE software. First group (C.A.) represent by blue. Second group (C.B.) identified by green.



Most of the *Q. suber* genotypes (MGC and SQS) were preliminary assigned to C.A. The other *Quercus* species belonged to C.B. Only three *Q. suber* individuals, can't assign to any cluster, because the *Qi* was <0.7. As well, one sample (Sardinian genotypes), grown in MGC resulted to belonge to the C.B. cluster. With this grouping (based on the *Qi* value), we obtained two different situations: first group is made by *Q. suber* genotypes (C.A.); second group contained other *Quercus* species (grown in Sardinia areas) (C.B.).

To improve the robustness of results, we investigated the genetic population structure, of C.A. and C.B, treating the two clusters as separated entities. STRUCTURE analysis revealed a break in the slope of likelihood value at *K=4* for C.A. group, and at *K=3* for C.B. group (Figure 4.9).

Figure 4.9 ΔK as a function of the number of genetic groups, considering separately C.A. (*Q. suber* individuals), and C.B. (different *Quercus* species).



Among *Q. suber* accessions, we referred to C.A., to indicate four main clusters indentified in the fundamental genetic groups, and an further admixture group (see Figure 4.10 A). However, C.B. group (Figure 4.10 B), is made by three different clusters. The different genetic groups were mirror of *Quercus* genus division. In fact, group I containing *Q. ilex* and *Q. coccifera*, group II, *Q. morisii* and group III, *Q. pubescens* and *Q. congesta*.

Figure 4.10 A Genetic relationship among *Q. suber* species (C.A.). Different colors indicate the STRUCTURE groups. **Cluster C.A.:** sky=group I, Yellow= group II, Red=group III, Purple=group IV, and an Admixture group=group V. In right were reported sampled distribution per country (details in legend). Different pies represent the Mediterranean individuals distribution for each group.

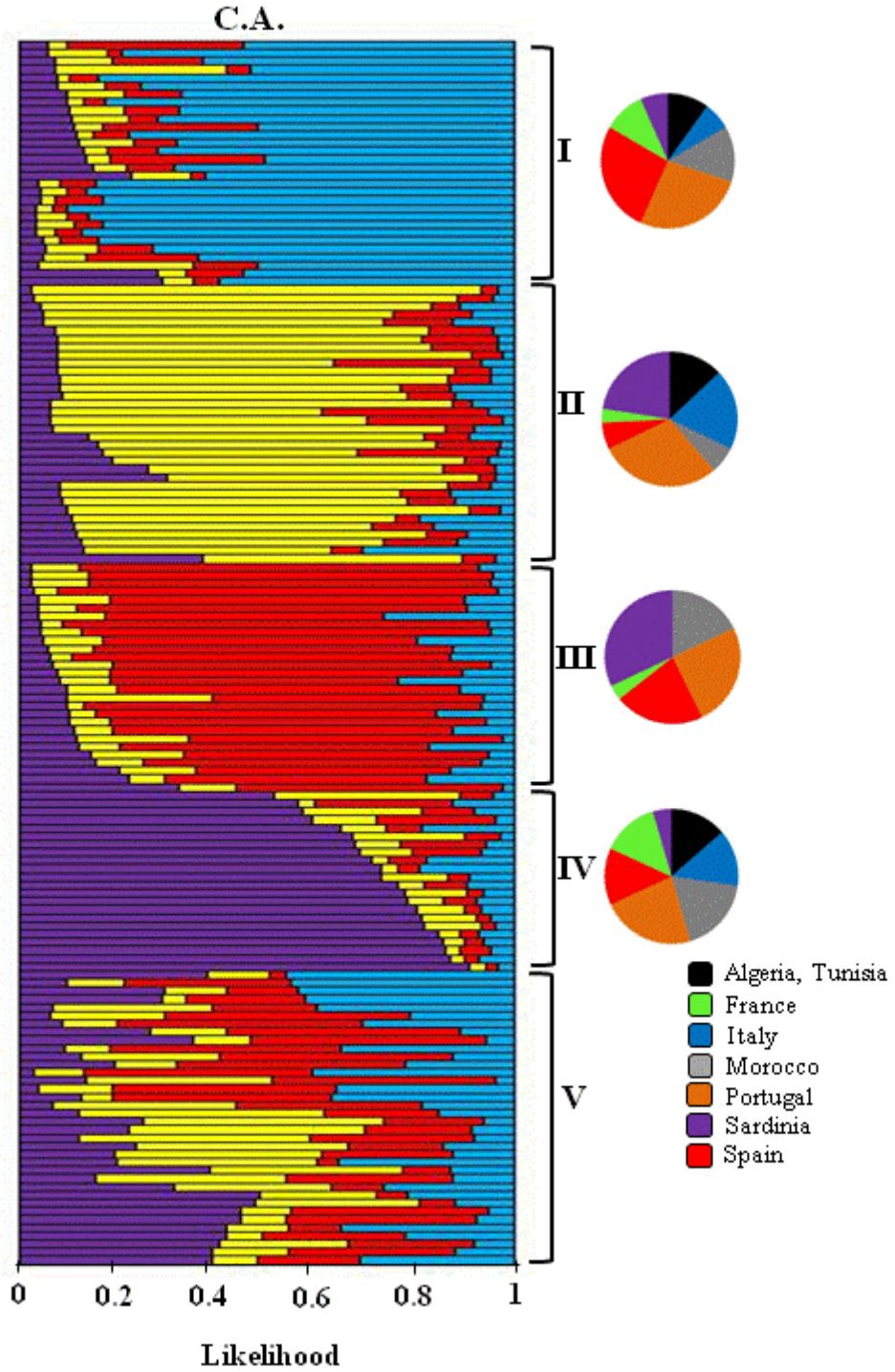
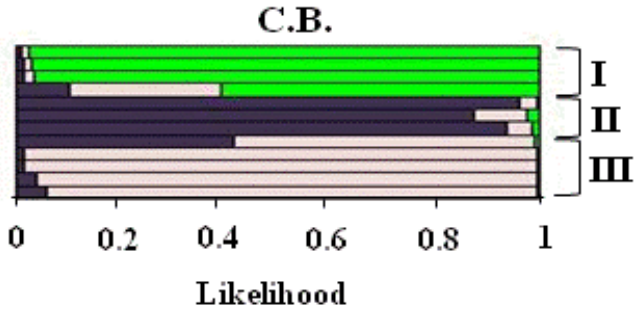


Figure 4.10 B Genetic relationship among different *Quercus* species (C.B). Different colors indicate the group identified by STRUCTURE software. **Cluster C.B.:** green=group I, violet= group II, gray=group III.



The *Q. suber* MGC distribution of genetic variance, calculated as F_{ST} value, was evaluated with AMOVA (Table 4.7). The hierarchic AMOVA using all MGC accessions, articulated in 27 populations, showed a higher difference than results observed to chloroplast situation (cp F_{ST} =0.764). According to our germplasm we observed that main component of variation was explained by "within population" (91.24%). However, 7.5% of variation was attributable to "among population ".

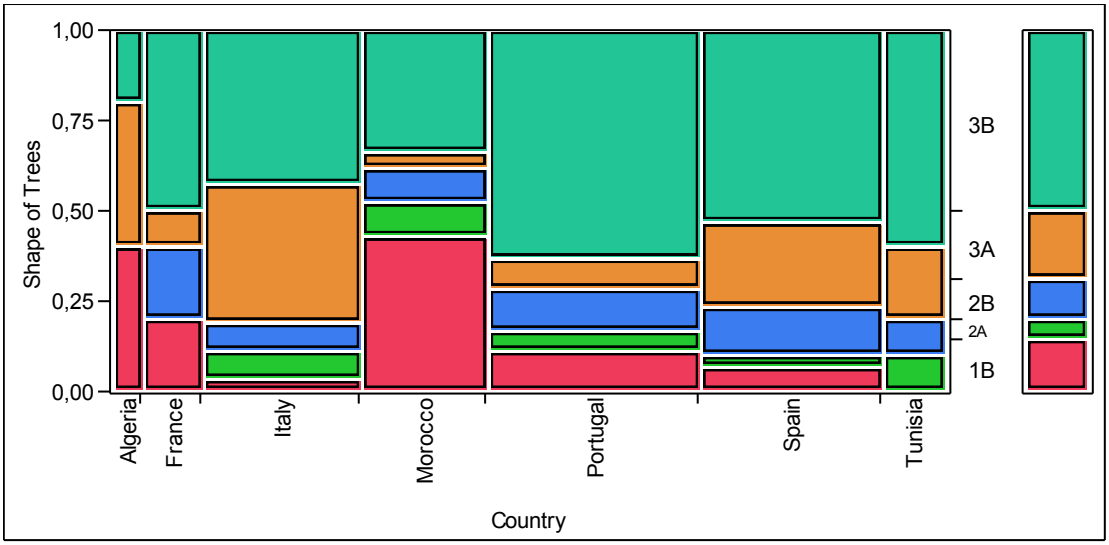
Table 4.7 AMOVA analysis on nuSSR markers, for MGC. Distance method based of number of different alleles (F_{ST}).

Source of variation	D.F.	S.S.	Variance components	Percentage of variation	Fixation indices
Country	6	25,85	0,025	1,23	FCT: 0.0122
Among populations	18	62,06	0,155	7,54	FSC: 0.0763
Within populations	227	427,03	1,881	91,24	
Total	251	514,94	2,062		**FST: 0.0877**

4.6 Phenotypic data

Some phenotypic data were recorded in MGC (137 individuals) in two different years (2009 and 2010). In Figure 4.11 were reported a contigent analysis among shape of trees in the Mediterranean Basin countries. The relationship between shape of tree and country showed to be significant correlated (p= 0.0347). In fact, was observed that the shape 3B is present in all countries with highest percentage in Portugal. Contrarily, there were a shape 1B more representative in Morocco and Algeria, and shape 3A in Italy.

Figure 4.11 Contingent analysis between shape of trees by Country (Mosaic Plot).



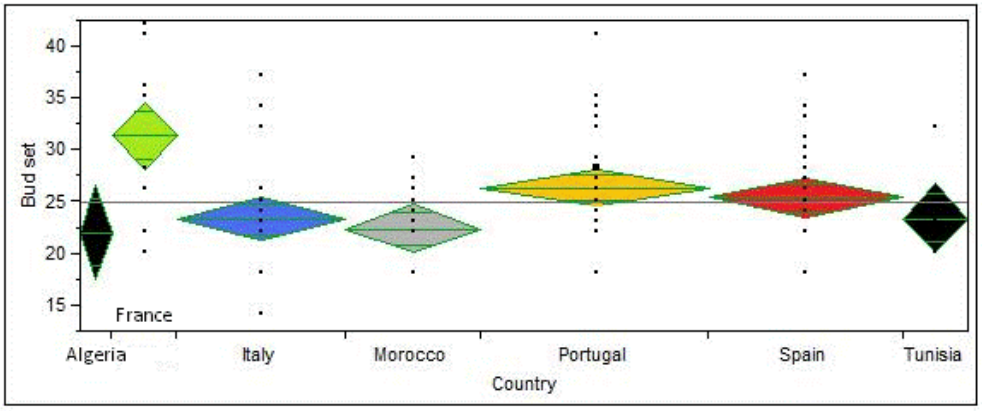
N	DF	Chi square	Prob>Chi square
137	24	37,1014	0,0347*

*p< 0.05 (Pearson index)

In figure 4.12 were shown the ANOVA results between bud set and bud flush stage of growing observed in *Q. suber* populations (MGC). In the statistical test was detected a significance difference among countries, and a *p-value* resulted very close among growing stages analyzed (0.0003 for bud set, 0.0004 for bud flush). The significance differences observed, pointed out a big differences between earlier Algerian-Tunisian and later French trees. The data shown in Figure 4.12 are related to the 2009 year, similar behavior was observed in 2010 (data not shown).

Figure 4.12 ANOVA analysis results for a bud set and bud flush (state grown in trees) differentiated by Country (2009 year). Method based of Chi-Square test.

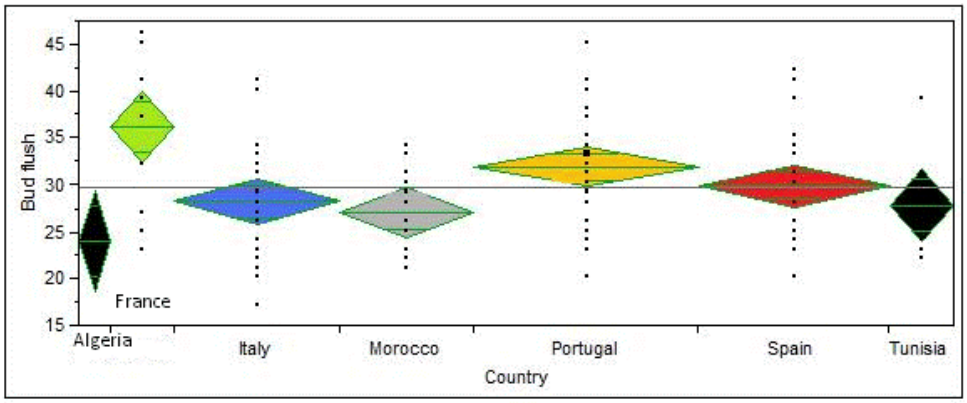
a) Bud set



Source	DF	S.S.	M.S.	F Ratio	Prob > F
Country	6	743.95	124.00	4.63	0.0003*
Error	130	3481.41	26.78		
C. Total	136	4225.36			

*p< 0.05 (Pearson index)

b)Bud flush



Source	DF	S.S.	M.S.	F Ratio	Prob > F
Country	6	981.73	163.62	4.48	0.0004*
Error	130	4750.59	36,54		
C. Total	136	5732.32			

*p< 0.05 (Pearson index)

5. DISCUSSION

In this study a Mediterranean collection of *Q. suber* accessions has been analyzed using molecular markers (chloroplastic and nuclear SSR) and phenotypic traits.

The sampling strategy was established with the aim to evaluate the level of genetic diversity within Sardinia *Q. suber* individuals. To test the method of analysis it was decided to sampled the Mediterranean *Q. suber* accession grown in a "common garden" central Sardinia site (MGC).

Past researches using the common garden approach, indicates that for tree species such as *P. abies* populations, originating from higher altitudes are characterized by lower above-ground growth rates, survival and greater resistance to snow and drought damage (Oleksyn *et al.*, 1998).

In recent study, Ramìrez-Valiente *et al.* (2009) used a common garden experiment to test the role of genetic drift and natural selection in cork oak differentiation regarding drought tolerance. They compared the molecular and quantitative genetic differentiation among *Q. suber* populations throughout rainfall and temperature gradient.

Chloroplast DNA (cpDNA) is maternally inherited in oaks (Dumolin-Lapègue *et al.*, 1997), it allows a direct study of seed-mediated dispersal and gene flow and its consequently particularly useful to infer re-colonization routes.

The reduced gene flow for this marker caused a cpDNA polymorphism much more structured than nuclear polymorphism (Petit *et al.*, 1993a, b). In fact, our data shown to be extremely well structured, with 63% of variance related to among populations component. This data is compatible with previous studies, conducted in the last 15 years by several authors (Dumolin-Lapègue *et al.*, 1997b; Petit *et al.*, 1993a, b).

Q. suber seems to be under a general trend characterized by highest level of genetic diversity (F_{ST}: 0.764) driven by chloroplastic markers. In the main time we have to consider that the uniparental inheritance evolves very slowly (four times slower than the plant nuclear genome; Wolfe *et al.*, 1987), a trend that is even more evidence in long lived tree species such as *Fagaceae* (Frascaria *et al.*, 1993).

The variability explicated by F_{ST} data, was also confirmed by Nei's (1978) gene diversity index (*H*). Considering our germplasm, in fact, we observed the major component of *H* is based on among populations coming from different countries (0.37) (Table 4.2).

Based on the cpSSR analysis, the Mediterranean collection showed a strong structure. Indeed, four haplotypes (represented by a different number of individuals), account for the 84.5% of the total number of individuals. One haplotype (A) characterize the most of accessions from Morocco, the Iberian Peninsula and south west France, while another two haplotypes occupy the Italian Peninsula (C and D). The fourth haplotypes (D) shows a striking distribution: it occurs in a narrow belt ranging from Tunisia and Algeria via Sardinia to south France (Fig. 4.2).

The remaining haplotypes (E to N), showed a discontinuous distribution. These last are collocated in the South Iberian Peninsula, near the Strait of Gibraltar. This genetic haplotypes structure could be interpreted as reflection of routes of post glacial re-colonization (Dumolin-Lapègue *et al.*, 1997a; Magri *et al*., 2007).

Magri *et al.* (2007), using a large number of chloroplast markers, found that the Italian haplotypes may be resulting of unidirectional cytoplasmic introgression by *Quercus Cerris*, an Eastern European deciduous oak that is widespread in the Italian peninsula. Similarly, Magri and coworkers observed, for a similar collection, a single haplotype (E to N, in our work), corresponds to a cpDNA lineage shared with *Quercus ilex*.

However, a similar relationship was observed in our study. Often, a genetic survey based on cpDNA (chloroplast DNA) showed that introduced populations harbour the same haplotypes as populations from the natural distribution, suggesting that the introduced gene pool is a representative sample of the original germplasm (Magni *et al.*, 2004; Kremer, 2007).

The multiplication of phylogeographic studies, evidence that there is an association between the geographic localization of the haplotypes and their position on geographic map (Dumolin-Lapègue *et al.*, 1997a).

Sequencing analysis revealed in our collection a general trend of longer microsatellites motive in Italian genotypes.

Most of the time the Italian haplotypes sequences followed the two outgroups (*C. mollissima* and *C. sativa*) sequences microsatellites repeats motif.

It might be speculated that Italian genotypes represent the original ancestral chloroplast *Quercus* sequences, and the rest of sequence data set under went to a deletion of the microsatellites motive.

This data was previously partially observed by Magri *et al.* (2007), where Cmcs9 cpSSR locus displayed a 30bp insertion/deletion stretch.

In general, the level of subdivision for cpDNA polymorphism in oak is much higher than for nuclear polymorphism. Previous study observed a mean coefficient of genetic differentiation of 0.003 in *Q. petraea* (Zanetto et al., 1994), as compared to 0.83 per cpDNA in white oak (Dumolin-Lapègue *et al.*, 1997b). The common garden MGC verified similar trend between nuclear and chloroplast SSR markers (Table 4.3 and 4.7).

In plants, such a contrast between the level of structuration for nuclear and maternally inherited genes has now been shown to be the roles than the exception.

Several factors contribute to increase genetic structure for organelle genes in comparison to nuclear genes: 1) effective gene flow will be limited to seeds for maternally inherited genome (Petit *et al.*, 1993b); 2) drift will be twice as strong for a haploid as compared to a diploid genome; 3) in hermaphrodite species such as oaks, the flowering and fruiting pattern result in an effective number of tree contributing to the next generation as female that is much reduced as compared to the effective number of tree acting as males (Demesure *et al.*, 1996; Dow and Ashley, 1996).

However no association between cpSSR and nuSSR has been detected. In fact the *Q. suber* diversity investigated using chloroplast SSR showed different scenario than nuclear SSR.

The low nuclear F_{ST} value (0.088) point out a limited differentiation between populations, showing that the 91.24% of the total variation is observed within populations. This results are supported by previous study conducted in natural selection in cork oak differentiation by Ramírez-Valiente *et al.* 2009.

When nuSSR variation was analyses in *Q. suber* using SRUCTURE, the model-based clustering revealed no geographical patterns.

Our data is in contrast with Mediterranean Eastern and Western main division reported by Lumaret *et al.* (2009). This authors used 10 nuSSR for 580 *Q. suber* individuals.

A possible justification could be related to the difference in the dataset samples size: in fact Lumaret *et al.* (2009) used more than 500 *Q. suber* individuals, while our germoplasm collection is made by about 150 individuals.

Using the STRUCTURE software, 4 major groups were occurred for *Q. suber* genotypes analyzed (C.A. group). In our samples the *Qi* is strikingly lower than in the other tree species of the Mediterranean cultivated taxa, such as *Olea europea* (Lumaret *et al.*, 2004; Baldoni *et al.*, 2006), *Castanea sativa* (Fineschi *et al.*, 2000) and *Cupressus sempervirens* (Bagnoli *et al.*, 2009). We may consider few observations related to the pie representation of structure data showed in figure 4.10, because the data set used is extremely not homogeneous.

Different number of individuals from different country was used (from the maximum data set of 35 Portuguese to 5 from Algerian). However few interesting data come out from the STRUCTURE analysis. The Italian, Algerian and Tunisian genepools showed the same trend: the amount of Italian-Algerian-Tunisian germplasm increase from the group I to group II, disappeared in group III and is again detected in group IV. Italian, Algerian and Tunisian germplasm seems to have the same behaviors and contribution on the 4 STRUCTURE groups identified. We may speculat that Italian-Algerian-Tunisian built a different peculiar genepool and genetic group, closely connected.

A strong genetic structure was observed in C.B. group, where different *Quercus* species were analyzed. The 6 *Quercus* species clustered in 3 structured genetic groups that partially followed the subgenus classification of *Quercus* explored by Manos et al. 1999 in cp and nrDNA. These authors defined in *Quercus* genus 4 different groups, based on ITS (nrDNA): Lobatae, Quercus s.s., Protobalanus and Cerris. Our data (Fig. 4.11) showed that *Q. ilex* and *Q. coccifera* clusterd in the same genepool (I) (green); while *Q. suber* and *Q. morisii* clustered in the genepool II (purple-violet).

The two genepools (I, II) are mirror respectively of "Ilex group and "Cerris group" reported by Manos *et al.* (1999).

The genepool III is made by *Q. pubescens* and *Q. congesta*. In Steinkeller *et al.* (1997), these last species were grouped with *Q. petraea* and *Q. robur.*

According to Tutin *et al.* 1993 in the *Quercus* genus different sections were shown: section Cerris, Robur and Rubrae.

In our case, the genepool III confirmed a differentiation among section into the *Quercus* genus.

The conservation of genetic diversity within species of economic value is important in order to insure the potentialities of populations for future adaptation and future breeding programs (Lynch, 1996; Eriksson *et al.*, 2001). The phenotypic characterization of few important traits is showing an increasing value for economical purposes.

For all phenotypic traits monitored, we found high correlation with trees provenance. Bud set and bud flush grown stages monitored showed to be significant correlated to the geographic origin.

The earlier time observed for Algerian, Tunisian, Italian and Moroccan individuals were probably occurred to search on the best of climatic and soil conditions. For France samples, the later bud set and bud flush time, could be hypothesized in a very different soil-climatic condition between MGC (implant site) and original provenance forest.

Different shape of trees was also evaluated in this study. In general we observed a non homogeneous geographic distribution of the morphologic data. In particular the form 3B was present in all the germplasm source analyzed, while form 2B resulted totally absent in Algerian individuals, but in our samples a few number of accessions are from Algerian (only one population).

It should be hypotize that shape of trees observed is a result of adaptation phenomena. Also, individuals originated by different Mediterranean country evidenced dominance for different type of shape.

At the same time, the shape of trees should be defined by anthropic selection. In fact, human activity, focused on economic income, looking for the most upright and linear grown, to have a large surface of cork planks.

The modern history of *Q. suber* is closely related to human activity for cork production.

For this reason, humans have been considered responsible for a reduction in genetic variation in some stands of cork oak, as well as for hybridization with congeners (Blondel & Aronson, 1999; Thompson, 2005; Magri *et al.*, 2007).

However, the geographical distribution of the cork oak haplotypes does not appears to be related to cultivation.

In fact, fossil pollen and wood records suggest that cork oak was distributed in approximately the same areas as today even before the Neolithic (Magri *et al.*, 2007).

The factors controlling hybridization in oaks are not well understood, but it is generally agreed that selection against intermediate phenotypes must be important (Stebbins *et al.*, 1947; Hardin, 1975).

However, it is difficult to determine whether hybrids are being eliminated before they can pass on their genes, or whether some of them are actually contribution to the variability of the parental species via backcrossing.

Oaks species are quite variable morphologically (Baranski, 1975), and they show agreed deal of convergence evolution of morphological characters (Tucker, 1974), so that it is difficult to find distinctive genetic markers or to be sure of the homologies of morphological studies whether there is significant interspecific gene flow, masked by strong selection for a limited number of genes controlling striking morphological and physiological features (Whittemore *et al.*, 1991).

6. CONCLUSION

In this study we explored the genetic diversity of two *Q. suber* collections using a Microsatellites markers. In particular the chloroplastic DNA diversity detected by CpSSR markers, was extremely usefully to evaluate genetic population structure in *Quercus* genus by several authors. Based on our results cpSSR markers have shown to be deeply informative for country-populations genetic diversity.

Chloroplast DNA is the results of maternal heritance, that is expressed mainly by among population from different Mediterranean basin (FST=0.764). The Cp markers has confirmed its key role discriminating and describing specifically populations from different states. In this case we were able to design a partial geographical fingerprinting of *Quercus* populations along Western Europe localities.

Considering the economical, recreational and landscape role of *Q. suber* in Sardinia we begin an evaluation of cpSSR reliability analysis for local populations.

However, we discovered low level of *H* among Sardinia accessions; this data is compatible with adaptation of endemisms to the local environmental, habitat and climatic conditions.

Abiotic and biotic forces would selected the germplasm that is most compatible with the habitat conditions; a decrease of genetic diversity among populations is a consequences of the lost of natural biodiversity.

Despite the high informative value of nuSSR markers analysis, our data showed to be not able to discriminate as much as the chloroplast DNA analysis.

In order to study the geographic specific *Q. suber* traits, we also evaluated few phenotypic characters such as bud set and bud flush stage of growing and shape of trees. Interestingly we detected high significance correlation between the morphological parameters and the countries of origin. Anthropic activity and natural selection could have influenced the geographical specificity of the morphological data considered.

Finally, the high variability found in this samples, showed the informanec of these trees in future breeding programs and in a changing climatic scenario.

In conclusion with this work a morphological and genotypic characterizations of two collections (MGC and SQS), was performed.

These results represent the beginning of future studies for the conservations and evaluations of Mediterranean *Quercus* biodiversity.

Sardinia, is an hot spot of ancestral plant biodiversity, and represents a natural storage of genetic diversity that is only partially known. Future studies for the natural biodiversity preservation and valorizations need to be done.

A1 Supplementary material

Alignment of *Q. suber* L. haplotypes (20) and other *Fagaceae* species considered. On the left was indicated the identity of the aligned species. The *Q. suber* genotypes analyzed were described with different colors according to the country and the region of origin (**red** for Spain, **orange** for Portugal, **blue** for Italy, gray for Morocco, **purple** for Sardinia). On the were described the sample sizes (bp). In **green** was signed the SNPs and insertion events ; **orange underlined**: microsatellites motif.

Locus **Ccmp4** (Weising and Gardner, 1999)

```
C. mollissima    CCAGAATATTCGGAGGACTCTTCTGAAC-AAA-CAAATAATTGTCAGCAACGTTG-TTTCTTTTTTTTTTT--CAAAT 73
Q. nigra         CCAGAATATTCGGAGGACTCTTCTGAAC-AAA-CAAATAATTGTCAGCAACGTTGTTTTCTTTTTTTTTTTTCTAAAT 74
F. sylvatica     CCAAAATATTGGGAGGACTCTTCTGA-CCAAA-CAAATAATTGTCAGCAACATTG-TTTCTTTTTTTTTT--CAAAT 73
ES9GE(I)         CCAAAATATTCGGAGGACTCTTCTGAAC-AAA-CAAATAATTGTCAGCAACGTTG-TTTCTTTTTTTTT----CAAAT 71
PTIV04(II)       CCAAAATATTCGGAGGACTCTTCTGAAC-AAA-CAAATAATTGTCAGCAACGTTG-TTTCTTTTTTTTT----CAAAT 71
PTIV04(IV)       CCAAAATATTCGGAGGACTCTTCTGAAC-AAA-CAAATAATTGTCAGCAACGTTG-TTTCTTTTTTTTT----CAAAT 71
ES9GE(V)         CCAAAATATTCGGAGGACTCTTCTGAAC-AAA-CAAATAATTGTCAGCAACGTTG-TTTCTTTTTTTTT----CAAAT 71
PTIV03(III)      CCAAAATATTCGGAGGACTCTTCTGAAC-AAA-CAAATAATTGTCAGCAACGTTG-TTTCTTTTTTTTT----CAAAT 71
PTIV03(V)        CCAAAATATTCGGAGGACTCTTCTGAAC-AAA-CAAATAATTGTCAGCAACGTTG-TTTCTTTTTTTTT----CAAAT 71
MOIII2(V)        CCAAAATATTCGGAGGACTCTTCTGAAC-AAA-CAAATAATTGTCAGCAACGTTG-TTTCTTTTTTTTTT---CAAAT 72
ES4CR(I)         CCAAAATATTCGGAGGACTCTTCTGAAC-AAA-CAAATAATTGTCAGCAACGTTG-TTTCTTTTTTTTTT---CAAAT 72
ITSACA(I)        CCAAAATATTCGGAGGACTCTTCTGAACCAAACCAAATAATTGTCAGCAACGTTG-TTTCTTTTTTTTTTT--CAAAT 75
ITLAVT(V)        CCAAAATATTCGGAGGACTCTTCTGAACCAAACCAAATAATTGTCAGCAACGTTG-TTTCTTTTTTTTTTTTTCAAAT 77
ITPUBR(I)        CCAAAATATTCGGAGGACTCTTCTGAACCAAACCAAATAATTGTCAGCAACGTTG-TTTCTTTTTTTTTTTTTCAAAT 77
ES7ALM(IV)       CCAAAATATTCGGAGGACTCTTCTGAAC-AAA-CAAATAATTGTCAGCAACGTTG-TTTCTTTTTTTTTT--CAAAT 73
                 *************************** ************************** ************    :****

C. mollissima    CCAAAAAATTTTTATTAATTTATACGTAGGTCATCGATTCAGCATT 119
Q. nigra         CCAAAGAATTTTTCTTAATTTATACGTAGGTCATCGATTCAGCATT 120
F. sylvatica     CCAAAGAATTTTTCTTAATTTATACGTAGGTCATCGATTCAGCATT 119
ES9GE(I)         CCAAAGAATTTTTCTTAATTTATACGTAGGTCATCGATTCAGCATT 117
PTIV04(II)       CCAAAGAATTTTTCTTAATTTATACGTAGGTCATCGATTCAGCATT 117
PTIV04(IV)       CCAAAGAATTTTTCTTAATTTATACGTAGGTCATCGATTCAGCATT 117
ES9GE(V)         CCAAAGAATTTTTCTTAATTTATACGTAGGTCATCGATTCAGCATT 117
PTIV03(III)      CCAAAGAATTTTTCTTAATTTATACGTAGGTCATCGATTCAGCATT 117
PTIV03(V)        CCAAAGAATTTTTCTTAATTTATACGTAGGTCATCGATTCAGCATT 117
MOIII2(V)        CCAAAGAATTTTTCTTAATTTATACGTAGGTCATCGATTCAGCATT 118
ES4CR(I)         CCAAAGAATTTTTCTTAATTTATACGTAGGTCATCGATTCAGCATT 118
ITSACA(I)        CCAAAAAATTTTTATTAATTTATACGTAGGTCATCGATTCAGCATT 121
ITLAVT(V)        CCAAAAAATTTTTATTAATTTATACGTAGGTCATCGATTCAGCATT 123
ITPUBR(I)        CCAAAAAATTTTTATTAATTTATACGTAGGTCATCGATTCAGCATT 123
ES7ALM(IV)       CCAAAAAATTTTTATTAATTTATACGTAGGTCATCGATTCAGCATT 119
                 *****:*******:********************************
```

Locus **Cmcs6** (Sebastiani *et al.*, 2004)

```
C. mollissima    CTTATGATCGTCACGAATTGAATTATAATCAAATTGCCTTGGGTCGTTTACCGATGTCAGTAATTGACGATTATACAATTCGAACAATTTTAAATTCACCTAAA 104
C. sativa        CTTATGATCGTCACGAATTGAATTATAATCAAATTGCCTTGGGTCGTTTACCGATGTCAGTAATTGACAATTATACAATTCGAACAATTTTAAATTCACCTAAA 104
ES9GE(I)         CTTATGATCGTCACGAATTGAATTATAATCAAATTGCTTGGGGTCGTTTACCGATGTCAGTAATTGACGATTATACAATTCGAACAATTTTAAATTCACCTAAA 104
PTIV04(II)       CTTATGATCGTCACGAATTGAATTATAATCAAATTGCTTGGGGTCGTTTACCGATGTCAGTAATTGACGATTATACAATTCGAACAATTTTAAATTCACCTAAA 104
PTIV04(IV)       CTTATGATCGTCACGAATTGAATTATAATCAAATTGCTTGGGGTCGTTTACCGATGTCAGTAATTGACGATTATACAATTCGAACAATTTTAAATTCACCTAAA 104
ES9GE(V)         CTTATGATCGTCACGAATTGAATTATAATCAAATTGCTTGGGGTCGTTTACCGATGTCAGTAATTGACGATTATACAATTCGAACAATTTTAAATTCACCTAAA 104
MOIII2(V)        CTTATGATCGTCACGAATTGAATTATAATCAAATTGCTTGGGGTCGTTTACCGATGTCAGTAATTGACGATTATACAATTCGAACAATTTTAAATTCACCTAAA 104
ES4CR(I)         CTTATGATCGTCACGAATTGAATTATAATCAAATTGCTTGGGGTCGTTTACCGATGTCAGTAATTGACGATTATACAATTCGAACAATTTTAAATTCACCTAAA 104
PTIV03(III)      CTTATGATCGTCACGAATTGAATTATAATCAAATTGCTTGGGGTCGTTTACCGATGTCAGTAATTGACGATTATACAATTCGAACAATTTTAAATTCACCTAAA 104
PTIV03(V)        CTTATGATCGTCACGAATTGAATTATAATCAAATTGCTTGGGGTCGTTTACCGATGTCAGTAATTGACGATTATACAATTCGAACAATTTTAAATTCACCTAAA 104
ITLAVT(V)        CTTATGATCGTCACGAATTGAATTATAATCAAATTGCTTGGGGTCGTTTACCGATGTCAGTAATTGACGATTATACAATTCGAACAATTTTAAATTCACCTAAA 104
ITPUBR(I)        CTTATGATCGTCACGAATTGAATTATAATCAAATTGCTTGGGGTCGTTTACCGATGTCAGTAATTGACGATTATACAATTCGAACAATTTTAAATTCACCTAAA 104
ES7ALM(IV)       CTTATGATCGTCACGAATTGAATTATAATCAAATTGCTTGGGGTCGTTTACCGATGTCAGTAATTGACGATTATACAATTCGAACAATTTTAAATTCACCTAAA 104
ITSACA(I)        CTTATGATCGTCACGAATTGAATTATAATCAAATTGCTTGGGGTCGTTTACCGATGTCAGTAATTGACGATTATACAATTCGAACAATTTTAAATTCACCTAAA 104
                 ********************************************************************************************************

C. mollissima    AT------AAAATAAAAAAAAAAA--TAAGTAAATCCTTTGATTAAAGATATTAAAGATAAAGAAATGGTAATTTCAAATTCAAATTAGGAAGGGTCCTTTTTC 200
C. sativa        AT------AAAATAAAAAAAAAAA---TAAGTAAATCCTTTGATTAAAGATATTAAAGATAAAGAAATGGTAATTTCAAATTCAAATTAGGAAGGGTCCTTTTTC 199
ES9GE(I)         AT------AAAATAAAAAAAA-----GAAGTAAATCCTTTGATTAAAGATATTAAAGATAAAGAAATGGTAATTTCAAATTCAAATTAGGAAGGGTCCTTTTTC 197
PTIV04(II)       AT------AAAATAAAAAAAA-----GAAGTAAATCCTTTGATTAAAGATATTAAAGATAAAGAAATGGTAATTTCAAATTCAAATTAGGAAGGGTCCTTTTTC 197
PTIV04(IV)       AT------AAAATAAAAAAAA-----GAAGTAAATCCTTTGATTAAAGATATTAAAGATAAAGAAATGGTAATTTCAAATTCAAATTAGGAAGGGTCCTTTTTC 197
E9GE(V)          AT------AAAATAAAAAAAA-----GAAGTAAATCCTTTGATTAAAGATATTAAAGATAAAGAAATGGTAATTTCAAATTCAAATTAGGAAGGGTCCTTTTTC 197
MOIII2(V)        AT------AAAATAAAAAAAA-----GAAGTAAATCCTTTGATTAAAGATATTAAAGATAAAGAAATGGTAATTTCAAATTCAAATTAGGAAGGGTCCTTTTTC 197
ES4CR(I)         AT------AAAATAAAAAAAA-----GAAGTAAATCCTTTGATTAAAGATATTAAAGATAAAGAAATGGTAATTTCAAATTCAAATTAGGAAGGGTCCTTTTTC 197
PTIV03(III)      AT------AAAATAAAAAAAA-----GAAGTAAATCCTTTGATTAAAGATATTAAAGATAAAGAAATGGTAATTTCAAATTCAAATTAGGAAGGGTCCTTTTTC 197
PTIV03(V)        AT------AAAATAAAAAAAA-----GAAGTAAATCCTTTGATTAAAGATATTAAAGATAAAGAAATGGTAATTTCAAATTCAAATTAGGAAGGGTCCTTTTTC 197
ITLAVT(V)        ATTAAAATAAAATAAAAAAAAAAAA-GAAGTAAATCCTTTGATTAAAGATATTAAAGATAAAGAAATGGTAATTTCAAATTCAAATTAGGAAGGGTCCTTTTTC 207
ITPUBR(I)        ATTAAAATAAAATAAAAAAAAAAAA-GAAGTAAATCCTTTGATTAAAGATATTAAAGATAAAGAAATGGTAATTTCAAATTCAAATTAGGAAGGGTCCTTTTTC 207
ES7ALM(IV)       A----------ATAAAAAAAAAAAAA-GAAGTAAATCCTTTGATTAAAGATATTAAAGATAAAGAAATGGTAATTTCAAATTCAAATTAGGAAGGGTCCTTTTTC 198
ITSACA(I)        AT------AAAATAAAAAAAAAAAA-GAAGTAAATCCTTTGATTAAAGATATTAAAGATAAAGAAATGGTAATTTCAAATTCAAATTAGGAAGGGTCCTTTTTC 202
                 *         **********     :************************************************************************
```

Locus **Cmcs7** (Sebastiani *et al.*, 2004)

```
C. mollissima     AAAATTGGAGTGATTATTGACTACTAAAATAAGTAAGTTTGTTTGACAAAATTAAGACTTCTTAACGAAATTAAGAATTATTGTGTATAGCAAATTAAAATG 102
C. sativa         AAAATTGGATTGATTATTGACTACTAAAATAAGTAAGTTTGTTTGACAAAATTAAGACTTCTTAACGAAATTAAGA-----TGTGTATAGAAAATTAAAATG  97
ES7ALM(IV)        AAAATTGGATTGATTATTGACTACTAAAATAAGTAAGTTTGTTTGACAAAATTAAGACTTCTTAACGAAATTAAGAATTATTGTGTATAGCAAATTAAAATG 102
ITSACA(I)         AAAATTGGATTGATTATTGACTACTAAAATAAGTAAGTTTGTTTGACAAAATTAAGACTTCTTAACGAAATTAAGAATTATTGTGTATAGCAAATTAAAATG 102
ITLAVT(V)         AAAATTGGATTGATTATTGACTACTAAAATAAGTAAGTTTGTTTGACAAAATTAAGACTTCTTAACGAAATTAAGAATTATTGTGTATAGCAAATTAAAATG 102
ITPUBR(I)         AAAATTGGATTGATTATTGACTACTAAAATAAGTAAGTTTGTTTGACAAAATTAAGACTTCTTAACGAAATTAAGAATTATTGTGTATAGCAAATTAAAATG 102
PTIV03(V)         AAAATTGGATTGATTATTGACTACTAAAATAAGTAAGTTTGTTTGACAAAATTAAGACTTCTTAACGAAATTAAGAATTATTGTGTATAGCAAATTAAAATG 102
PTIV03(III)       AAAATTGGATTGATTATTGACTACTAAAATAAGTAAGTTTGTTTGACAAAATTAAGACTTCTTAACGAAATTAAGAATTATTGTGTATAGCAAATTAAAATG 102
ES4CR(I)          AAAATTGGATTGATTATTGACTACTAAAATAAGTAAGTTTGTTTGACAAAATTAAGACTTCTTAACGAAATTAAGAATTATTGTGTATAGCAAATTAAAATG 102
MOIII2(V)         AAAATTGGATTGATTATTGACTACTAAAATAAGTAAGTTTGTTTGACAAAATTAAGACTTCTTAACGAAATTAAGAATTATTGTGTATAGCAAATTAAAATG 102
ES9GE(V)          AAAATTGGATTGATTATTGACTACTAAAATAAGTAAGTTTGTTTGACAAAATTAAGACTTCTTAACGAAATTAAGAATTATTGTGTATAGCAAATTAAAATG 102
PTIV04(IV)        AAAATTGGATTGATTATTGACTACTAAAATAAGTAAGTTTGTTTGACAAAATTAAGACTTCTTAACGAAATTAAGAATTATTGTGTATAGCAAATTAAAATG 102
PTIV04(II)        AAAATTGGATTGATTATTGACTACTAAAATAAGTAAGTTTGTTTGACAAAATTAAGACTTCTTAACGAAATTAAGAATTATTGTGTATAGCAAATTAAAATG 102
ES9GE(I)          AAAATTGGATTGATTATTGACTACTAAAATAAGTAAGTTTGTTTGACAAAATTAAGACTTCTTAACGAAATTAAGAATTATTGTGTATAGCAAATTAAAATG 102
                  ************************************************************************     ********************

C. mollissima     AGTGTAATTATTATTACTGATCAATAAATATATGGATAAAAATATCTAAAAAAAAAAAA--GAGAGAGGGGCGATTCATTTTTATGGTACAAAACTCATTCATCTCGCTT 210
C. sativa         AGTGTAATTATTATTACTGATCAATAAATATATGGATAAAAATATCTAAAAAAAAAAAA--GAGAGAGGGGCGATTCATTTTTATGGTACAAAACTCATTCATCTCGCTT 205
ES7ALM(IV)        AGTGTAATTATTATTACTGATCAATAAATATATGGATAAAAATATCTAAAAAAAAAAAAAAGAGAGAGGGGCGATTCATTTTTATGGTACAAAACTCATTCATCTCGCTT 212
ITSACA(I)         AGTGTAATTATTATTACTGATCAATAAATATATGGATAAAAATATCTAAAAAAAAAAAAAAGAGAGAGGGGCGATTCATTTTTATGGTACAAAACTCATTCATCTCGCTT 212
ITLAVT(V)         AGTGTAATTATTATTACTGATCAATAAATATATGGATAAAAATATCTAAAAAAAAAAAAAAGAGAGAGGGGCGATTCATTTTTATGGTACAAAACTCATTCATCTCGCTT 212
ITPUBR(I)         AGTGTAATTATTATTACTGATCAATAAATATATGGATAAAAATATCTAAAAAAAAAAAAAAGAGAGAGGGGCGATTCATTTTTATGGTACAAAACTCATTCATCTCGCTT 212
PTIV03(V)         AGTGTAATTATTATTACTGATCAATAAATATATGGATAAAAATATCTAAAAAAAAAAAA---GAGAGAGGGGCGATTCATTTTTATGGTACAAAACTCATTCATCTCGCTT 209
PTIV03(III)       AGTGTAATTATTATTACTGATCAATAAATATATGGATAAAAATATCTAAAAAAAAAAAA---GAGAGAGGGGCGATTCATTTTTATGGTACAAAACTCATTCATCTCGCTT 209
ES4CR(I)          AGTGTAATTATTATTACTGATCAATAAATATATGGATAAAAATATCTAAAAAAAAAAAA--GAGAGAGGGGCGATTCATTTTTATGGTACAAAACTCATTCATCTCGCTT 210
MOIII2(V)         AGTGTAATTATTATTACTGATCAATAAATATATGGATAAAAATATCTAAAAAAAAAAAA---GAGAGAGGGGCGATTCATTTTTATGGTACAAAACTCATTCATCTCGCTT 209
ES9GE(V)          AGTGTAATTATTATTACTGATCAATAAATATATGGATAAAAATATCTAAAAAAAAAAAA---GAGAGAGGGGCGATTCATTTTTATGGTACAAAACTCATTCATCTCGCTT 209
PTIV04(IV)        AGTGTAATTATTATTACTGATCAATAAATATATGGATAAAAATATCTAAAAAAAAAAA----GAGAGAGGGGCGATTCATTTTTATGGTACAAAACTCATTCATCTCGCTT 208
PTIV04(II)        AGTGTAATTATTATTACTGATCAATAAATATATGGATAAAAATATCTAAAAAAAAAAA----GAGAGAGGGGCGATTCATTTTTATGGTACAAAACTCATTCATCTCGCTT 208
ES9GE(I)          AGTGTAATTATTATTACTGATCAATAAATATATGGATAAAAATATCTAAAAAAAAAAA---GAGAGAGGGGCGATTCATTTTTATGGTACAAAACTCATTCATCTCGCTT 209
                  ************************************************************    ************************************************
```

Locus **Cmcs8** (Sebastiani *et al.*, 2004)

```
C. mollissima     AGAAATAAACACCCCCATTATCTTATTATCTTACACATCATTCTGATACAGGATACTTAATTTAATGACGCGTTGAATTATATTAATATTTATTCGATCTAGAAATGAAT 110
C. sativa         AGAAATAAACACCCCCATTATCTTATTATCTTACACATCATTCTGATACAGGATACTTAATTTAATGACGCGTTGAATTATATTAATATTTATTCGATCTAGAAATAAAT 110
PTIV03(V)         AGAAATAAACACCCCCATTATCTTATTATCTTACACATCATTCTGATACAGGATACTTAATTTAATGACGCGTTGAATTATATTAATATTTATTCGATCTAGAAATGAAT 110
PTIV03(III)       AGAAATAAACACCCCCATTATCTTATTATCTTACACATCATTCTGATACAGGATACTTAATTTAATGACGCGTTGAATTATATTAATATTTATTCGATCTAGAAATGAAT 110
PTIV04(II)        AGAAATAAACACCCCCATTATCTTATTATCTTACACATCATTCTGATACAGGATACTTAATTTAATGACGCGTTGAATTATATTAATATTTATTCGATCTAGAAATGAAT 110
ES9GE(I)          AGAAATAAACACCCCCATTATCTTATTATCTTACACATCATTCTGATACAGGATACTTAATTTAATGACGCGTTGAATTATATTAATATTTATTCGATCTAGAAATGAAT 110
PTIV04(IV)        AGAAATAAACACCCCCATTATCTTATTATCTTACACATCATTCTGATACAGGATACTTAATTTAATGACGCGTTGAATTATATTAATATTTATTCGATCTAGAAATGAAT 110
ES9GE(V)          AGAAATAAACACCCCCATTATCTTATTATCTTACACATCATTCTGATACAGGATACTTAATTTAATGACGCGTTGAATTATATTAATATTTATTCGATCTAGAAATGAAT 110
MOIII2(V)         AGAAATAAACACCCCCATTATCTTATTATCTTACACATCATTCTGATACAGGATACTTAATTTAATGACGCGTTGAATTATATTAATATTTATTCGATCTAGAAATGAAT 110
ES4CR(I)          AGAAATAAACACCCCCATTATCTTATTATCTTACACATCATTCTGATACAGGATACTTAATTTAATGACGCGTTGAATTATATTAATATTTATTCGATCTAGAAATGAAT 110
ITLAVT(V)         AGAAATAAACACCCCCATTATCTTATTATCTTACACATCATTCTGATACAGGATACTTAATTTAATGACGCGTTGAATTATATTAATATTTATTTGATCTAGAAATGAAT 110
ITPUBR(I)         AGAAATAAACACCCCCATTATCTTATTATCTTACACATCATTCTGATACAGGATACTTAATTTAATGACGCGTTGAATTATATTAATATTTATTTGATCTAGAAATGAAT 110
ES7ALM(IV)        AGAAATAAACACCCCCATTATCTTATTATCTTACACATCATTCTGATACAGGATACTTAATTTAATGACGCGTTGAATTATATTAATATTTATTTGATCTAGAAATGAAT 110
ITSACA(I)         AGAAATAAACACCCCCATTATCTTATTATCTTACACATCATTCTGATACAGGATACTTAATTTAATGACGCGTTGAATTATATTAATATTTATTTGATCTAGAAATGAAT 110
                  ********************************************************************************************************

C. mollissima     CAACAAAGTCGTCTTATTTATTTTCAGGTCTAATTTTTTTTT-TA--TTTTTTGTGAGTGGAAAAATAGACC 179
C. sativa         CAACAAAGTCGTCTTATTTATTTTCAGGTCTAATTTTTTTTT-TA--TTTTTTGTGAGTGGAAAAATAGACC 179
PTIV03(V)         CAACAAAGTCGTCTTATTTATTTTCAGGTCTAATTTTTTTTT--ATTTTTTTTGTGAGTGGAAAAATAGACC 180
PTIV03(III)       CAACAAAGTCGTCTTATTTATTTTCAGGTCTAATTTTTTTTT--ATTTTTTTTGTGAGTGGAAAAATAGACC 180
PTIV04(II)        CAACAAAGTCGTCTTATTTATTTTCAGGTCTAATTTTTTTTTT-ATTTTTTTTGTGAGTGGAAAAATAGACC 181
ES9GE(I)          CAACAAAGTCGTCTTATTTATTTTCAGGTCTAATTTTTTTTTT-ATTTTTTTTGTGAGTGGAAAAATAGACC 181
PTIV04(IV)        CAACAAAGTCGTCTTATTTATTTTCAGGTCTAATTTTTTTTTT-ATTTTTTTTGTGAGTGGAAAAATAGACC 181
ES9GE(V)          CAACAAAGTCGTCTTATTTATTTTCAGGTCTAATTTTTTTTTT-ATTTTTTTTGTGAGTGGAAAAATAGACC 181
MOIII2(V)         CAACAAAGTCGTCTTATTTATTTTCAGGTCTAATTTTTTTTTT-ATTTTTTTTGTGAGTGGAAAAATAGACC 181
ES4CR(I)          CAACAAAGTCGTCTTATTTATTTTCAGGTCTAATTTTTTTTTT-ATTTTTTTTGTGAGTGGAAAAATAGACC 181
ITLAVT(V)         CAACAAAGTCGTCTTATTTATTTTCAGGTCTAATTTTTTTTTTTATTTTTTTTGTGAGTGGAAAAATAGACC 182
ITPUBR(I)         CAACAAAGTCGTCTTATTTATTTTCAGGTCTAATTTTTTTTTTTATTTTTTTTGTGAGTGGAAAAATAGACC 182
ES7ALM(IV)        CAACAAAGTCGTCTTATTTATTTTCAGGTCTAATTTTTTTTTT-ATTTTTTTTGTGAGTGGAAAAATAGACC 181
ITSACA(I)         CAACAAAGTCGTCTTATTTATTTTCAGGTCTAATTTTTTTTTT-ATTTTTTTTGTGAGTGGAAAAATAGACC 181
                  ******************************************  *  ************************
```

Locus **Cmcs14** (Sebastiani *et al.*, 2004)

```
C. mollissima     GTGCAAGGAATGTCAAACTAAATAAATATAATTTTTCATCTTTCTACATGAAATATATATATATATATGATAATCGCCTTTT 80
C. sativa         GTGCAAGGAATGTCGAACTAAATAAATATAATTTTTCATCTTTCTACATGAAATATATATATATATATGATAATCGCCTTTT 80
ES7ALM(IV)        GTGCAAGGAATGTCGAACTAAATAAATATAATTTTTCATCTTTCTACATGAAATATATATATATATATGATAATCGCCTTTT 80
ITSACA(I)         GTGCAAGGAATGTCGAACTAAATAAATATAATTTTTCATCTTTCTACATGAAATATATATATATATATGATAATCGCCTTTT 80
ITLAVT(V)         GTGCAAGGAATGTCGAACTAAATAAATATAATTTTTCATCTTTCTACATGAAATATATATATATATATGATAATCGCCTTTT 80
ITPUBR(I)         GTGCAAGGAATGTCGAACTAAATAAATATAATTTTTCATCTTTCTACATGAAATATATATATATATATGATAATCGCCTTTT 80
PTIV03(V)         GTGCAAGGAATGTCGAACTAAATAAATATAATTTTTCATCTTTCTACATGAAATATATATATATAT---AATCGCCTTTT 77
PTIV03(III)       GTGCAAGGAATGTCGAACTAAATAAATATAATTTTTCATCTTTCTACATGAAATATATATATATAT---AATCGCCTTTT 77
ES4CR(I)          GTGCAAGGAATGTCGAACTAAATAAATATAATTTTTCATCTTTCTACATGAAATATATATATAT--GATAATCGCCTTTT 78
MOIII2(V)         GTGCAAGGAATGTCGAACTAAATAAATATAATTTTTCATCTTTCTACATGAAATATATATATAT--GATAATCGCCTTTT 78
ES9GE(V)          GTGCAAGGAATGTCGAACTAAATAAATATAATTTTTCATCTTTCTACATGAAATATATATATAT--GATAATCGCCTTTT 78
PTIV04(IV)        GTGCAAGGAATGTCGAACTAAATAAATATAATTTTTCATCTTTCTACATGAAATATATATATAT--GATAATCGCCTTTT 78
PTIV04(II)        GTGCAAGGAATGTCGAACTAAATAAATATAATTTTTCATCTTTCTACATGAAATATATATATAT--GATAATCGCCTTTT 78
ES9GE(I)          GTGCAAGGAATGTCGAACTAAATAAATATAATTTTTCATCTTTCTACATGAAATATATATATAT--GATAATCGCCTTTT 78
                  ************************************************************     ************

C. mollissima     TTATTATCTTGTTTTTGTCTTATAATTTGAAT-TTCATAAAATGGAATAAAATAAAGAAAGAGTTTCTTACAACTTCCTGAAAAATTTGTTACAATCC 177
C. sativa         TTATTATCTTGTTTTTGTCTTATAATTTGAATCTTCATAAAATGGAATAAAATAAAGAAAGAGTTTCTTACAACTTCCTGAAAAATTTGTTACAATCC 178
ES7ALM(IV)        TTATTATCTTGTTTTTGTCTTATAATTTGAAT-TTCATAAAATGGAATAAAATAAAGAAAGAGTTTCTTACAACTTCCTGAAAAATTTGTTACAATCC 177
ITSACA(I)         TTATTATCTTGTTTTTGTCTTATAATTTGAAT-TTCATAAAATGGAATAAAATAAAGAAAGAGTTTCTTACAACTTCCTGAAAAATTTGTTACAATCC 177
ITLAVT(V)         TTATTATCTTGTTTTTGTCTTATAATTTGAAT-TTCATAAAATGGAATAAAATAAAGAAAGAGTTTCTTACAACTTCCTGAAAAATTTGTTACAATCC 177
ITPUBR(I)         TTATTATCTTGTTTTTGTCTTATAATTTGAAT-TTCATAAAATGGAATAAAATAAAGAAAGAGTTTCTTACAACTTCCTGAAAAATTTGTTACAATCC 177
PTIV03(V)         TTATTATCTTGTTTTTGTCTTATAATTTGAAT-TTCATAAAATGGAATAAAATAAAGAAAGAGTTTCTTACAACTTCCTGAAAAATTTGTTACAATCC 174
PTIV03(III)       TTATTATCTTGTTTTTGTCTTATAATTTGAAT-TTCATAAAATGGAATAAAATAAAGAAAGAGTTTCTTACAACTTCCTGAAAAATTTGTTACAATCC 174
ES4CR(I)          TTATTATCTTGTTTTTGTCTTATAATTTGAAT-TTCATAAAATGGAATAAAATAAAGAAAGAGTTTCTTACAACTTCCTGAAAAATTTGTTACAATCC 175
MOIII2(V)         TTATTATCTTGTTTTTGTCTTATAATTTGAAT-TTCATAAAATGGAATAAAATAAAGAAAGAGTTTCTTACAACTTCCTGAAAAATTTGTTACAATCC 175
ES9GE(V)          TTATTATCTTGTTTTTGTCTTATAATTTGAAT-TTCATAAAATGGAATAAAATAAAGAAAGAGTTTCTTACAACTTCCTGAAAAATTTGTTACAATCC 175
PTIV04(IV)        TTATTATCTTGTTTTTGTCTTATAATTTGAAT-TTCATAAAATGGAATAAAATAAAGAAAGAGTTTCTTACAACTTCCTGAAAAATTTGTTACAATCC 175
PTIV04(II)        TTATTATCTTGTTTTTGTCTTATAATTTGAAT-TTCATAAAATGGAATAAAATAAAGAAAGAGTTTCTTACAACTTCCTGAAAAATTTGTTACAATCC 175
ES9GE(I)          TTATTATCTTGTTTTTGTCTTATAATTTGAAT-TTCATAAAATGGAATAAAATAAAGAAAGAGTTTCTTACAACTTCCTGAAAAATTTGTTACAATCC 175
                  * ***************************** ************************************************************
```

CHAPTER 2

Use of candidate gene loci
to investigate genetic diversity in a *Pinus uncinata* Western
Mediterranean basin collection.

1. INTRODUCTION

Across the environmental range occupied by plant species, local adaptation occurs as a balance between the divergent pressure of natural selection in different populations and gene flow among them (Savolainen *et al.*, 2007). Adaptation to local climatic conditions is a widespread feature of pines. For istance, reciprocal transplant experiments have demonstrated the superior performance of native provenance (Howe *et al.*, 2003; Shutyaev & Giertych 1997).

Local adaptation can occur in response to many selective forces, such as climate, photoperiod and edaphic factors (Hedrick, 2006; Kawecki *et al.*, 2004) and different traits in different species may respond. In general, steep clines may be common where natural selection acts on large populations whereas, in isolated populations, genetic drift and founder effects may have more influence (Savolainen *et al.*, 2007). Therefore controlled assessment of the genetic component of variation is needed to effectively link genomic and quantitative trait variation as species and populations from different environments show variable responses to selective factors (Rehfeldt *et al.*, 1999; Carter, 1996).

Traits related to adaptation are usually highly heritable, as shown in several quantitative genetic studies (Home *et al.*, 2003). However there is a dearth of information on the molecular genetics underlying adaptive variation. Some indication of the genetic architecture of complex traits in forest tree species comes from QTL (quantitative trait loci) studies that have identified genomic regions related to phenology, stress resistance (drought, cold, salinity) or wood quality (Howe *et al.*, 2003; Hurme *et al.*, 2000; Frewen *et al.*, 2000).

However, the use of QTL methodology in conifers is difficult because of the lack of pedigree lines, the size of the conifer genome, and consequent missing genomic resources such as dense genetic maps for precise location of QTLs and identification of causative loci (Neale and Savolainen, 2004).

Therefore for conifers species which have a large genome size and low correlation (linkage disequilibrium) between polymorphic sites at short genomic distance (a few hundred base pairs), the best current genomic approaches to test hypotheses of the genetic basis of local adaptation are based on the study of candidate genes i.e. genes known function from model organisms potentially related to the interest in the target species,

known position in the genome, potentially collocated with QTLs of genes of known expression patterns in response to various stress factors (Gonzàlez-Martìnez *et al.*, 2006a).

Nucleotide polymorphism is influenced by several factors including mutation, migration, selection and random genetic drift. In tree species, the current increase in sequence data gathered from nuclear gene loci has been driven mostly by the search for the molecular signature of natural selection (Savolainen & Pyhajarvi, 2007; Neale and Ingvarsson, 2008; Achaz, 2009). Selection can leave its traces as deviations from neutrality in the level of nucleotide diversity, allele frequency distribution or linkage disequilibrium (LD) (Achaz, 2009).

In pines, due to high recombination rates and a lack of correlation between polymorphic sites at very short genomic distances (rapid dacay of linkage disequilibrium), natural selection is expected to have very localized effects on nucleotide variation. Therefore, it should be possible distinguish between the effects of selection and demography by comparing genetic variation at many loci. Selection should affect specific loci, whilst demographic effects should be genome-wide (Howe *et al.*, 2003). This approach has been successfully applied in other species of rapid decay of linkage disequilibrium (e.g. poplar, Ingvarsson *et al.*, 2006).

Pinus uncinata Ramond, sensu stricto (s.s.), know as mountain pine, is usually a monocormic (single-stemmed) tree of up to 20 m height, has a straight and erect trunk and asymmetrical cones (Gaussen *et al.*, 1993).

It grows in the subalpine vegetation belt up to the upper forest limit, and even up to the upper tree limit (1000-2300 m elevation) in the mountainous regions of Western Europe (Critchfield & Little, 1969).

It is found in the mountains regions in Spain (Sierra de Gùdar and Sierra de Cebollera), as well as the Pyrenees, the Massif Central, Mont Ventoux and the western Alps (Cantegrel, 1983; Christensen, 1987; Hamernìk & Musil, 2007).

In the Iberian Peninsula, fossil pollen indicates glacial populations of the taxa from the *Pinus mugo complex* (that includes *P. uncinata*) species in the Cantabrian Atlantic Mountains, in the Sierra de Cebollera in Spain, where *P. uncinata* grows today (Franco Mùgica *et al*., 1998; Ramil-Rego *et al.*, 1998; Gil Garcià *et al*., 2002), in eastern Spain (Carriòn, 2002), and in the Betic System (Pons & Reille, 1998).

Glacial refugia existed very probably in the Alps, where travertine formations dated to 11000-10000 yr BP unambiguously reveal cone imprints of *P. uncinata* (in the southern

French Alps at 220 m elevation, Ali *et al.*, 2003) and of *P. uncinata* and *P. mugo s.s.* (in the Susa Valley, Italy at *c.* 1800 and 1900 m elevation, Ali *et al.*, 2006).

Given that today *P. uncinata* is a species of the subalpine vegetation, the interglacial retreat to high elevation have played an important role in shaping their genetic structure. Populations of *P. uncinata* that have already reached the mountain tops of low mountains in Spain (Camarero *et al.*, 2005) were also affected by the upward pressure by *P. sylvestris*. The sub-optimal growth conditions of the marginal *P. uncinata* populations and the competition and the potential hybridization with *P. sylvestris* may represent a significant threat to their conservation. Furthermore, if the marginal *P. uncinata* populations are genetically distinct from core populations, their loss may represent a significant diminution of diversity and adaptive potential for the species (Hampe & Petit, 2005; Eckert *et al.*, 2008). The geographical pattern of genetic diversity has recently been described for *P. uncinata* based on several cpDNA microsatellite markers (Dzialuk *et al.*, 2009), and a lack of genetic differentiation between *P. uncinata* and some species from the *P. mugo s.s.* complex has been observed in the Alps based on Monteleone *et al.,* 2006.

Pinus uncinata form an excellent system for analysis of the demographic and evolutionary forces driving divergence in natural populations because they form undomesticated, randomly mating natural populations of considerable ecological diversity (Neale & Savolainen, 2004).

2. OBJECTIVE

The main objective in this work was to study genetic variation, within and between multiple natural *P. uncinata* populations across Western Europe, for the purposes of detecting the molecular signature of natural selection. Therefore, populations were first assessed under controlled conditions (progeny trial) to determine the extent of genetically-controlled differentiation in phenotypes.

Phenotypic differentiation was assessed using several different measurements: morphometric data (height of seedling), phenological data (start time of bud flush) and, physiological data (chlorophyll fluorescence). Once phenotypic patterns had been established, patterns of nucleotide and haplotype diversity, and allele frequency spectra were analyzed in a multilocus nuclear gene data-set and tests were carried out to establish whether the observed nucleotide polymorphism had arisen due to demographic factors or to selection.

The specific objectives of the study were to:

- Test for signs of population structure due to geographical isolation or allopatric distribution range;
- Test for departures from a neutral model of evolution at the candidate genes and assess causes.
- Quantify phenotypic differentiation between populations due to local adaptation.

3. MATERIALS AND METHODS

3.1 Sampling and DNA extraction

P. uncinata seed samples from 5 locations in Western Europe were included in this study (Figure 3.1). The trees were sampled across an environmental gradient related to differences in altitude, length of growing season, annual rainfall and average mean temperature in winter. Cones were collected from mature trees in recognized old-growth pine forest (Table 3.1).

Seeds were germinated for a few days (4-7 days) in moist petri dishes. Genomic DNA was extracted from haploid megagametophyte (maternal tissue that surrounds the embryo following a standard Qiagen DNasy protocol. in the seed). As DNA samples were haploid, the haplotypes could be determined directly by sequencing. In total, 50 DNA extracts were prepared, representing 10 different trees from each of the 5 populations.

Figure 3.1 Distribution of *P. uncinata* populations included in this study. Red circle represent seeds site origin.

Table 3.1 Geographical location and environmental conditions for the different *Pinus uncinata* populations.

N	Country	Area	Location	Latitude	Longitude	Altitude (m)	At	M min	M max	P	PET	S	W	WVP
1	Andorra	Pyrenees 1	Vall de Ransol	42°34'N	01°29'E	2.00	12.36	8.13	16.87	57.17	82.50	51.90	3.12	10.45
2	Andorra	Pyrenees 2	San Miguel de Engolasters	42°31'N	01°34'E	2.00	12.65	8.47	17.16	56.83	82.67	52.00	3.06	10.58
3	Spain	Pyrenees 3	La Trapa	42°41'N	00°32'E	1.72	6.58	2.19	10.73	61.58	84.42	53.25	4.36	7.82
4	Spain	Gúdar	Sierra de Gúdar,Valderinares	40°23'N	00°36'E	2.00	16.28	11.79	20.80	44.50	90.00	58.42	2.65	12.76
5	France	Massif 1	Col de Croix de Morand	45°41'N	03°03'E	1.40	10.96	5.73	16.12	51.75	59.50	40.92	2.24	9.80

N Population number, ***At*** Annual of Temperature (deg C), ***M min*** Monthly average daily minimum (deg C), ***M max*** Monthly average daily maximum (deg C), ***P*** Precipitation (mm-Litres/sqm), ***PET*** Potential evapo-transpiration (mm- Litres/sqm), ***S*** Sunshine (%),***W*** Wind speed (m/s), ***WVP*** Water vapour pressure (Hpa).

3.2 Loci studied

In total, 29 nuclear loci were analyzed (Table 3.2). The loci were selected to represent different metabolic pathways. These included 3-ketoacyl-CoA synthase 3 (*KCS-3, Pr11*), ABI3-interacting protein (*a3ip*), ribosomal protein in S4 and in S10 (*rps4, rps10*), epoxide hydrolase (*eph*), abscissic acid-responsive protein (*abaR*), phytocyanin protein (*phy*), phitocrome P protein (*phytP*), chalcone synthase (*chcs*), and dehydrin dehydrative stress response (dhn2PP) and several additional putative proteins (Table 3.2).

Many dehydration induced proteins are referred to as late embryogenesis abundant (LEA) proteins (Ingram J. *et al.*, 1996). Dehydrins, also called Lea D11-family group II or RAB (responsive to ABA), accumulate in plant tissues in response to cellular dehydration resulting from developmental events or environmental stimuli such as osmotic stress and low temperature but also abscissic acid treatment (Carpita N. *et al.,* 1979, Close T.J. *et al.,* 1996, Ouvrard O. *et al.,* 1996).

Dehydrin genes are key components of dehydration tolerance (Close T.J. *et al.,* 1996). It is hypothesized that LEA proteins bind water in their random coil conformation and protect cellular structures from dehydtration stress (Ingram J. *et al.*, 1996).

The others cold candidate loci were selected based on either their function potentially related to response to dehydrative stress or the level of expression during cold acclimation correlated with freezing tolerance in Scots pine seedlings (Joosen *et al*., 2006).

Chalcone synthase (*chcs*) is one of a family of genes responsible for production of polyketide synthase enzymes associated with the production of chalcones, a class of organic compounds found mainly in plants as natural defense mechanisms and as synthases have been extensively studied, some information is available on the enzymes from bryophytes (primitive plants).

Cloning of *chcs* from the moss *Physcomitella patens* revealed an important transition from the chalcone synthases present in microorganisms to those present in higher plants (Jiang *et al.*, 2006).

Abscissic acid (ABA) plays a role in the abscission of plant leaves and ABA mediated signaling is also an important part in plant response to environmental stress and plant pathogens (Zhu, 2002).

The genes for ABA biosynthesis and sequence of the pathway have been elucidated. Abscissic acid is also produced in the roots in response to decreased soil water potential and other situations in which the plant may be under stress. ABA then translocates to the leaves, where it rapidly alters the osmotic potential of stomatal guard cells, causing them to shrink and stomata to close. The ABA induced closure reduces transpiration, thus preventing further water loss from the leaves in times of low water availability. In preparation for winter, ABA is produced in terminal buds.

This slows plant growth and directs leaf primordial to develop scales to protect the dormant buds during the cold season. ABA also inhibits the division of cells in the vascular cambium, adjusting to cold conditions in the winter by suspending primary and secondary growth (Milborrow 2001; Nambara *et al.*, 2005).

Table 3.2 Full list of candidate genes used in the study, their putative role and PCR amplification conditions.

Locus Name	Putative Role	Ta(°C)	[$MgCl_2$]
Pr3	ethylen-resp.transc.factor15(A.thaliana)	60	1.5
Pr4	f-box family protein (Populus)	60	1.5
Pr5	hypothetic protein	60	1.5
Pr7	hypothetic protein	60	1.5
Pr9	hypothetic protein	60	1.5
Pr11	3-ketoacyl-CoA-synthase AG	60	1.5
Pr12	recept.prot.kinase (O.sativa Japonica)	60	1.5
Pr16	glutamyl-tRNA reduct.precurs.(G.max)	60	1.5
Pr17	Put.polyol/monos.transp. (V.vinifera)	60	1.5
Pr20	possible oligpep.transp.system (B.sp.)	60	1.5
Pr23	hypothetic protein	60	1.5
Pr25	hypothetic protein	60	1.5

…continue…

…continue…

Locus Name	Putative Role	Ta(°C)	[$MgCl_2$]
Pr28	hypothetic protein	60	1.5
Pr29	hypothetic protein	60	1.5
Pr30	hypothetic protein	60	1.5
Pr32	hypothetic protein	60	1.5
Pr34	hypothetic protein	60	1.5
Pr35	hypothetic protein (Populus)	60	1.5
rps10	Put.ribosomal protein S10 (P. sylvestris)	59	2
rps4	hypothetic protein	59	1.5
chcs	chalcone synthase (P.pinaster)	61	1.5
abaR	abscisic acid-resp. protein (P.sylvestris)	61	1.5
dh2PP	Put.dehydrin-dehidr. stress resp.	61	1.5
eph	Put. Epoxide hydrolase (P.sylvestris)	59	1.5
phy	Put.phytocyanin (P.mugo)	59	1.5
a3ip	hip.ABI3-interacting protein	61	1.5
175_Anna	hypothetic protein (V.vinifera)	52	2
umn_927	hypothetic protein	60	1.5
phytP	phytochrome P (P.sylvestris)	61	2

3.3 PCR amplification and sequencing

PCR-amplification was performed with Thermo MBS thermal cyclers and, carried out in a total 25µl containing about 10ng of haploid template DNA, 50µM of each of dNTP, 2µM of each primer and 0.25U *Taq* DNA polymerase (Biolabs) with the respective 1 x PCR buffer. PCR followed standard amplification procedures with $MgCl_2$ concentration optimized for each primer pair as described in Table 3.2. Standard amplification procedures were used with initial denaturation at 94°C for 3min followed by 35 cycles with 30sec. denaturation at 94°C, 30sec. annealing and 1 min. extension at 72°C, and a final 5min. extension at 72°C. PCR fragments were purified using *Exo–Sap* (exonuclease, alckaline-phosphatase) enzymatic treatment. About 20ng of PCR product was used as a template in 10µl sequencing reactions with the Big Dye Terminator DNA Sequencing Kit (Applied Biosystems, Carlsbad, CA, USA) performed by the GenePool sequencing service, University of Edinburgh.

50 *P. uncinata* haploid DNA samples were sequenced in one direction for each locus. Codon-Code Aligner software ver. 3.7.1 (Codon Code Corporation, Dedham, MA, USA) was used for editing and assembly to produce alignments on the basis of the nucleotide sequence.

3.4 Sequence analysis

Good quality sequence was obtained for most of the samples at 29 loci. Nucleotide sequence alignment were constructed in Codon-Code Aligner and were further manually adjusted using GENEDOC. All polymorphisms were visually rechecked from chromatograms edited with Codon-Code Aligner. Coding and non-coding regions (introns, untranslated regions) were annotated on the basis of alignment with known sequences identified in the international databases (National Center for Biotechnology Information http://www.ncbi.nlm/nih.gov/gorf/gorf.html) and ExPASy (http://www.expasy.ch/tools/dna.html). The influence of demography on the multilocus pattern of variation and locus-specific affects was assessed by looking at the amount of nucleotide diversity, correlation between polymorphic sites and allelic frequency distribution between populations.

3.5 Nucleotide diversity

Multilocus theta, an estimate of the population mutation parameter, was used to assess nucleotide diversity. Estimates of theta (Θ_w, equal to $4N_e\mu$, were N_e is the effective population size and μ is the mutation rate per nucleotide site per generation), were computed on basis of the number of segregating sites and the length of each locus (Pyhäjärvi *et al.*, 2007).

The estimates of nucleotide diversity were conducted for all samples per locus (29 candidate genes considered), combined and separately for each *P. uncinata* populations.

Deviations of particular genes from the frequency distribution spectrum under the standard neutral model of evolution were assessed with Tajima's *D* (Tajima 1989).

Negative value of Tajima's *D* statistics indicate an excess of polymorphism with low frequency, with can be caused by selective sweeps (or population growth), whereas positive value indicate an excess of polymorphism with intermediate frequency, which

could be due, among other things, to balancing selection, population structure with sampling from many populations or bottlenecks. The distribution of Tajima's *D* was investigated for each locus, combined and separately for all *P. uncinata* populations.

Statistically significant deviation of this test from neutral model of evolution was evaluated by coalescence simulations as implemented in DNAsp5 software (Librado & Rozas 2009).

3.6 Linkage disequilibrium and haplotype diversity

The level of linkage disequilibrium was measured as the correlation coefficient r^2 (Hill and Robertson 1968) using informative sites. The decay of linkage disequilibrium with physical distance was estimated using non-linear regression of r^2 between polymorphic sites and the distance in *bp* (*base pairs*) between sites as detailed in Wachowiak *et al.* (2009).

The non-linear last-squares estimate of ρ ($\rho=4N_e$c, were N_e is effective population size, *c* is the recombination rate) was fitted by the *nls*-function implemented in R statistical package (http://www.r-project.org).

Differences in the last-squares ρ estimate within and between populations were measured separately for *Pr3, Pr17, Pr25,Pr29, Pr32, Pr34, rps10, abaR, dnh2PP* and *chcs* and for the combined dataset. Loci with low polymorphism or sample size in each population were excluded from the analysis.

The number of haplotypes (*Ne*) and haplotypes diversity (*Hd*) were computed for each gene using DNAsp5, insertions and deletions were included.

3.7 Population genetic structure

Population genetic structure was investigated using an analysis of variance framework (AMOVA). Genetic variation among individuals (groups) and among populations was studied locus by locus at both haplotype and single nucleotide polymorphic/indel level and also by averaging pairwise F_{ST} over all polymorphic sites across loci.

The significance of the genetic differentiation coefficient was evaluated by 1000 permutations of the samples between groups using Arlequin ver. 3.0 (Excoffier *et al.,* 2005).

Population structure from the haplotypic data was tested by *Snn* and K_{ST}* statistics (Hudson *et al.,* 1992, Lynch and Crease, 1990, Hudson 2000), and its significance evaluated by 1000 permutations of the samples between groups as implemented in DNAsp. *Snn* measure the average proportion of nearest-neighbour haplotypes that are present in the same locality and it is expected to be near 1.0 for strong population differentiation, while an estimate of 0.5 would indicate that two groups are part of the same panmictic populations. K_{ST}* compares distance between pairs of sequences from different populations to distances from pairs sampled at large from the total pool.

Genetic clustering of individuals on the basis of both full sequence data and all segregating sites and indels at 29 loci was conducted using BAPS 5.2 (Corander and Tang, 2007).

3.8 Comparison of polymorphism and divergence

To assess the correlation between the level of nucleotide polymorphism and divergence, we used sequences for *Pinus pinaster* at six loci (*phy, dhn2PP, abaR, chcs, phytP, 175*) and *Pinus taeda* for the others.

In case of loci *rps4, rsp10 Pinus sylvestris* was used as an outgroup, whilst *Pinus nigra* was used as outgroup for locus *eph*.

Nucleotide polymorphism based non-synonymous (*Ps*), and synonymous (*Pa*) sites and nucleotide divergence based on non-synonymous (*Ks*), synonymous (*Ka*) sites relative to the outgroup species was calculated using Jukes-Cantor correction (Hughes and Nei 1988) as calculated in DNAsp5. The program conducts maximum likelihood analysis of multilocus polymorphic and divergence data providing an estimate of the selection parameter *k*, which measures the degree to which diversity is increased or reduced by the action of selection at candidate loci.

3.9 Variation in quantitative traits-phenotypic differentiation

The progeny trial took place in a glasshouse at CEH Edinburgh (an indoor common garden trial), which is located at an altitude of 190 m. According to UK Met Office data (Perry and Hollis, 2005), the site has an average GLS (growing seasonal lenght) of 234 days. The mean temperature is 1.8 °C in February and 13.1 °C in July (see Table 3.1).

A total of 551 seedlings (5 populations, 5 families per population, about 21 seedlings per family) were grown in randomised block design with one seedling per family in each block (21 blocks). The populations were, La Trapa, Gúdar-Valederinares to Spain; Col de la Croix to France; Val de Ransol and San Miguel to Andorra (Figure 3.1). The heights of all of the seedlings (planted in spring 2010) were measured with a metric ruler before re-starting vegetative stage, and the date of bud flush for each seedling was recorded. Monitoring for bud flush measurements began 10 March and ended 13 May 2011 with a recording frequency of 7/8 days.

To evaluate variations in photochemical capacity, chlorophyll fluorescence was measured in the seedlings on 10 occasions between October 2010 and March 2011. Chlorophyll fluorescence has been extensively used in plant physiology and to study stress tolerance among relatively small numbers of genotypes, but has yet to be-come a popular tool in evolutionary biology for assessing differences in photosynthetic activity in response to environmental signals among multiple populations (Lopez *et al.*, 2009; Salmela *et al.*, 2011).

Measurement were taken between 9.30 am and noon, using one current or previous-year needle. Immediately after removal, the needle was dark-adapted for approximately 20 min., then a Handy PEA portable fluorescence (Hansatech Instrument Ltd., Norfolk, UK) was used to obtain fluorescence data. For logistical reasons, chlorophyll fluorescence measurements were monitored on a subset of the trial, in this case, 5 from three populations, respectively Val de Ransol for Andorra, Valderinares for Spain and Col de la Croix for France.

ANOVA was used to assess differentiation between populations and among families within populations for seedling heights and bud flush on the same set of individuals. ANOVA analysis of chlorophyll fluorescence data (based on the ratio F_v/F_m) was based on the subset of seedlings only.

4. RESULTS

4.1 Nucleotide Polymorphism

In total 11971 base pairs of sequence was obtained from 29 nuclear loci. Candidate genes used in this study were polymorphic with a total of 208 polymorphic sites plus 14 indels distributed in 8 different genes (*Pr11, Pr29, Pr35, abaR, chcs, eph, a3iP, phytP*). Multilocus silent theta was 0.0064 (means value, data not shown), but the peak value was observed for the population La Trapa (blue in Figure 4.1).

The most polymorphic genes were: *Pr20* with 17 polymorphic sites, *chcs* with 16, *dhn2PP* and *Pr5* with 13; the least polymorphic was *Pr4* with only one polymorphic site. The average nucleotide diversity total (π) over all loci was 0.0040 with 0.0071 silent sites and 0.0019 non synonymous sites (Table 4.1). An excess of singleton mutations across genes compared to neutral expectation was detected by negative multi locus Tajima's *D* in the total average dataset (-0.2367) (Table 4.1).

Figure 4.1. Multilocus silent theta: distribution for all *Pinus uncinata* populations.

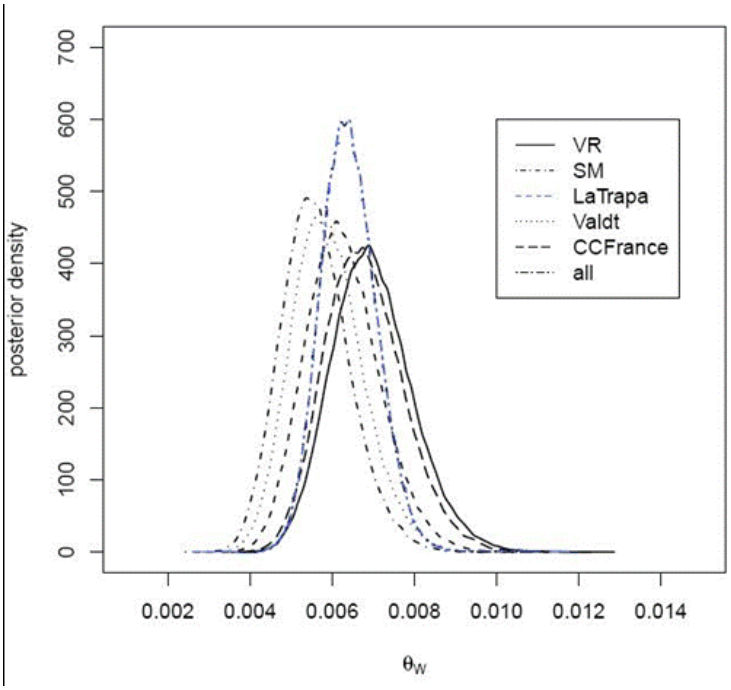


Table 4.1 Means values of summary statistics of nucleotide and haplotype variation, frequency distribution spectrum of polymorphism at the analyzed genes in all populations.

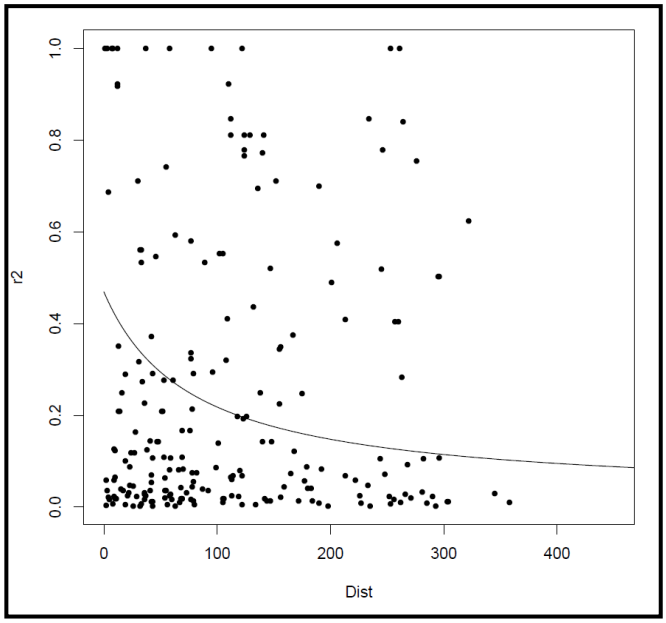
Population	L	N	π	SNP	Pa	Ps	PaS	LS	Sm	Ne	S	Hd	R	D
Vall de Ransol, Andorra	11971	9.7	0.004	3.93	0.007	0.002	0.007	181.5	2.8	3.4	1.62	0.582	1.00	-0.083
San Miguel de Engolasters. Andorra	11971	9.9	0.004	3.90	0.009	0.002	0.007	181.4	2.7	3.6	1.31	0.591	1.50	0.058
La Trapa. Spain	11971	10.0	0.004	4.62	0.008	0.002	0.008	181.5	3.2	3.9	1.76	0.638	1.00	-0.065
Sierra de Gúdar. Valderinares; Spain	11971	9.9	0.003	3.83	0.006	0.002	0.007	181.2	2.7	3.1	1.45	0.508	1.00	-0.062
Col de la Croix de Morand. France	11971	10.0	0.004	4.62	0.008	0.002	0.007	181.1	3.1	3.8	1.97	0.619	1.33	-0.216
Total	**11971**	**49.5**	**0.004**	**7.17**	**0.008**	**0.002**	**0.007**	**180.9**	**4.7**	**7.0**	**1.90**	**0.609**	**1.00**	**-0.237**

L, length of sequence in base pairs including indels; **N**, number of samples; **π**, nucleotide diversity; **SNP**, single nucleotide polymorphism; **Pa**, synonymous sites; **Ps**, non synonymous sites; **PaS**, synonymous sites and non coding position; **LS**, number of site; **Sm**, number of mutations; **Ne**, number of haplotypes; **S**, singleton variable sites; **Hd**, haplotypes diversity; **R**, recombination rate; **D**, test (Tajima, 1989).

4.2 Linkage disequilibrium and haplotype polymorphisms

Rapid decay of LD between pairs of parsimony informative sites at 9 loci was found across all populations (Figure 4.2). The average number of haplotypes (*Ne*) per genes was ~7. For *Pr 4* and *Pr 7* only two and three haplotypes were found respectively, whilst 13 were found for locus *abaR, dhn2PP, chcs* and 11 for *Pr5*. The haplotype diversity (*Hd*) means value was 0.6086, with and range of 0.892 (maximum at *dhn2PP*) to 0.230 (minimum at *eph*).

Figure 4.2 Scatter plot of the squared correlation coefficient of allele frequencies (r^2) as a function of distance in base pairs between pairs of polymorphic sites in all *P. uncinata* populations.



4.3 Population genetic structure

Population genetics structure was assessed by AMOVA (Table 4.2) based on pairwise genetic differentiation coefficients, F_{ST} (under the infinite allele model).

Variance component were estimated among populations, among groups and among populations within groups.

The vast majority of variation was partitioned within populations (96.76%). Significant pairwise F_{ST} values were found (P<0.05) between population at some loci (Table 4.3): *Pr5, Pr9, Pr11, Pr20, rps4, dhn2PP*, but the most substantial difference were at loci *Pr12, abaR* and *Umn-927*.

Table 4.2 Analysis of molecular variance (AMOVA) based on F_{ST} among and within *Pinus uncinata* populations in five different population from Western Europe.

Source of variation	**d.f.**	**S.S.**	**V. component**	**Variation(%)**
Among populations	4	99.52	0.62	3.24
Within populations	45	839.00	18.64	96.76
Total	49	938.52	19.27	
F_{ST} : 0.03236				

Significance test (1023 permutation)
Va and F_{ST}: P(rand. Value > obs. Value)= 0.0049
P(rand. Value= obs. Value)= 0.001

Table 4.3 Pairwise F_{ST} estimates among the five populations, with standard deviations. Significant value are shown in blue.

	VR_Andorra	SM_Andorra	La_Trapa	Valderinares	CC_France
VR.	*				
SM.	0.57658±0.0411	*			
LaTr.	0.16216±0.0445	0.25225±0.0326	*		
Vald.	**0.02703±0.0139***	0.06306±0.0237	**0.00901±0.0091***	*	
CC.	0.16216±0.0402	0.29730±0.0360	**0.00901±0.0091***	0.65766±0.0360	*

*P<0.05

Population differentiation based on number of differences between haplotypes was generally low across loci with mean *Snn* values of 0.504 and *Kst** of 0.021 (data not shown).

Sixteen out of 29 genes used showed significant differentiation based on *Snn*, and for four of these (*Pr12, abaR, chcs* and *Umn-927*), the significance level was high 0.001<P<0.01 (data not shown).

The same significant differentiation was based on *Kst** for thirteen different loci with a significant *p. value* between 0.001 and 0.01 (*Pr5, Pr12, Pr20, abaR, Umn-927*).

The most significance differentiation was observed among La Trapa and Gúdar-Valderinares from Spain (Table 4.4).

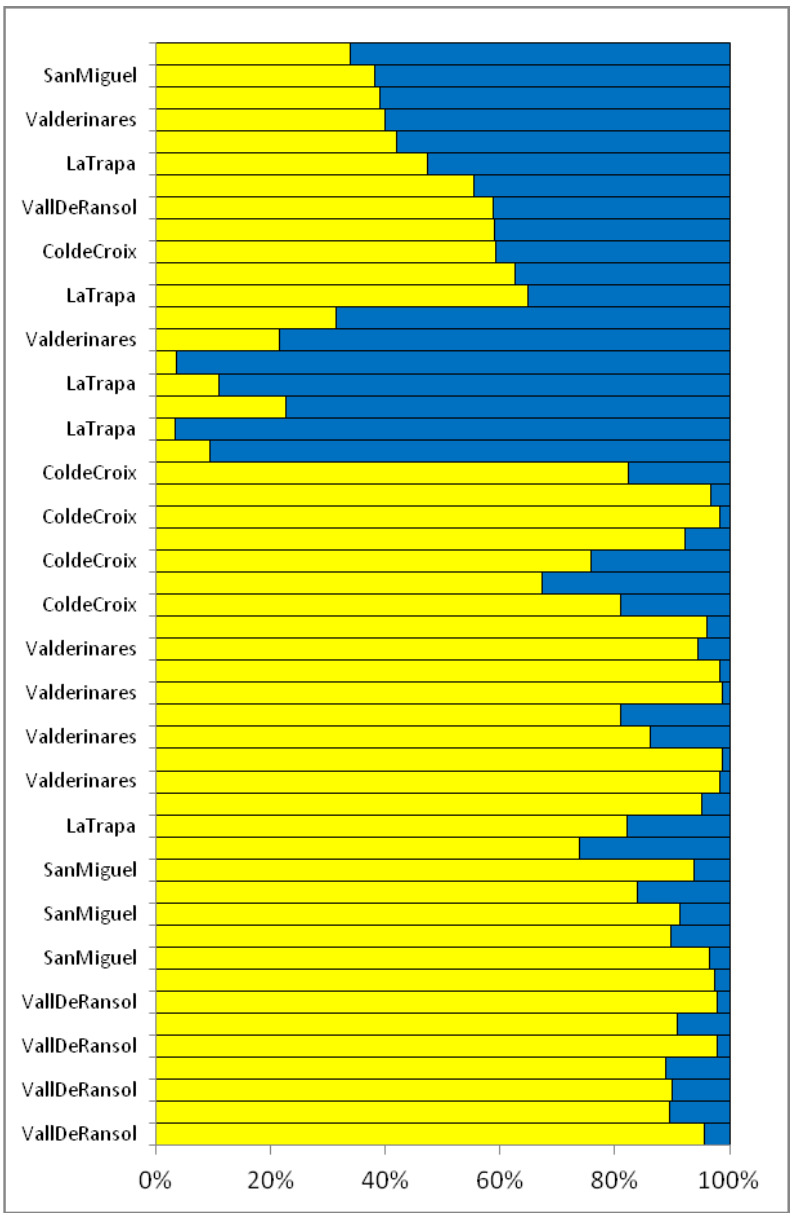
Table 4.4 Matrix of significance in pairwise comparisons between *P. uncinata* populations.

Populations	**Fst**					**Snn**					**Kst***				
	VR	SM	LaTR	Vald	CC	VR	SM	LaTR	Vald	CC	VR	SM	LaTR	Vald	CC
VR	*					*					*				
SM	0.0071	*				0.4850	*				0.0124	*			
LaTR	0.0274	0.0244	*			0.5064	0.4895	*			0.0078	0.0180	*		
Vald.	0.0578	0.0472	0.0866	*		0.5333	0.5059	0.5371	*		0.0361	0.0329	0.0576	*	
CC	0.0241	0.0060	0.0576	0.0076	*	0.5144	0.4941	0.5299	0.4779	*	0.0221	0.0071	0.0244	0.0036	*

VR. Vall de Ransol, Andorra; **SM.** San Miguel de Engolasters, Andorra; **LaTR.** La Trapa, Spain; **Vald.** Sierra de Gùdar, Valderinares, Spain; **CC.** Col de la Croix, France.

Genetic clustering of individuals was conducted using BAPS 5.2 (Corander and Tang, 2007) (Figure 4.3). This identified, two different groups across the *P. uncinata* populations considered (K=2 was optimal). However, although four of the individuals most clearly assigned to the second cluster were located in population La Trapa, the clusters were largely assigned across populations, and showed little geographic structure. The main geographic distinction was that individuals from the French population, Col de la Croix, were assigned a single cluster only.

Figure 4.3 BAPS results for K=2, for *P. uncinata* populations considered. Left: assignment of individuals to the two clusters; right: population of origin of individuals in cluster 1 and 2.



4.4 Phenotypic differentiations

Height measurement

ANOVA analysis of the seedling height showed statistically significant differences between populations and among families within populations with P <0.0001 and P<0.001 respectively (Table 4.5). This difference in the average heights of various populations (Table 4.6) comprised variation from maximum value in seedling from France to a minimum for a seedling from the Spanish population. (La Trapa).

Table 4.5 ANOVA tables of height measurement *Pinus uncinata* populations.

Change	**d.f.**	**s.s.**	**m.s.**	**v.r.**	**P**
Population	4	83.214	20.803	43.72	<0.0001***
Pop./Family	20	31.461	1.573	3.31	0.0065**
Block	24	12.845	0.535	1.12	0.311 ns
Residual	503	239.345	0.476		
Total	551	366.864	0.666		

**P<0.001. ***P<0.0001

Table 4.6 Variation in population height means (cm).

N	**Country**	**Population**	**Hm(cm)**	**s.d.**
1	France	Col de Croix de Morand	4.09	0.781
2	Spain	La Trapa	2.90	0.564
3	Andorra	San Miguel de Engolasters	3.32	0.619
4	Spain	Sierra de Gudar. Valderinares	3.32	0.815
5	Andorra	Vall de Ransol	3.65	0.804

N: populations number; Hm: Height mean (Cm)

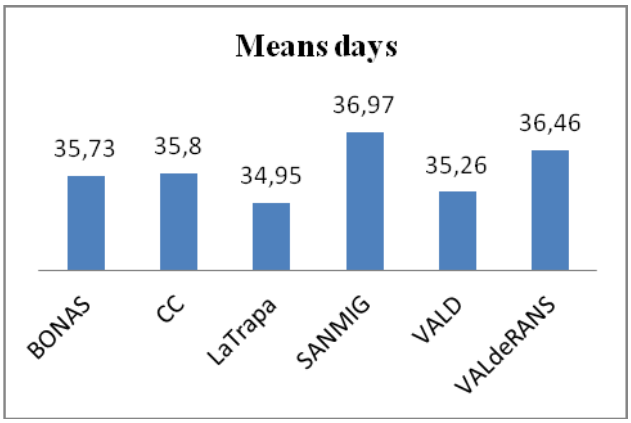
Phenology (bud flush)

Bud flush stage was monitored every 7-8 days on 551 *P. uncinata* individuals. Peak flushing date was between 8-15 April 2011, when 63.5% of individuals flushed. There were no statistically significant differences between populations or among families within populations, in fact the average value (bud flush time) were very similar for all groups (Table 4.7 and Figure 4.4).

Table 4.7 ANOVA analysis of phenolgy observations for six P. *uncinata* populations.

Change	d.f.	s.s.	m.s.	v.r.	P
Population	5	300.77	60.2	1.41	**0.3927**
Pop./Family	37	2068.98	55.9	1.31	**0.9752**
Block	24	2714.31	113.1	2.65	
Residual	523	22302.82	42.6		
Total	589	27386.87	46.5		

Figure 4.4 Phenology variation in six *P. uncinata* population of bud flush stage (means days).



Chlorophyll fluorescence

A seasonal pattern was observed in F_v/F_m with the peak level reached in November for all populations and the minimum in the second half of January for Andorra and Spanish populations, and in early February for the French population (Figure 4.5). The F_v/F_m means values on different dates varied from 0.76 in November to 0.59 in January. In fact, on dates when the minimum values of F_v/F_m were recorded (January 5 and January 20), significant differentiation between populations was observed, 0.038 and 0.040 respectively (Table 4.8). No statistically significant differences were observed among family within populations on any of these measurement dates.

Figure 4.5 Variation in population of F_v/F_m between October 27 2010 and March 18 2011 (distinct values for any population).

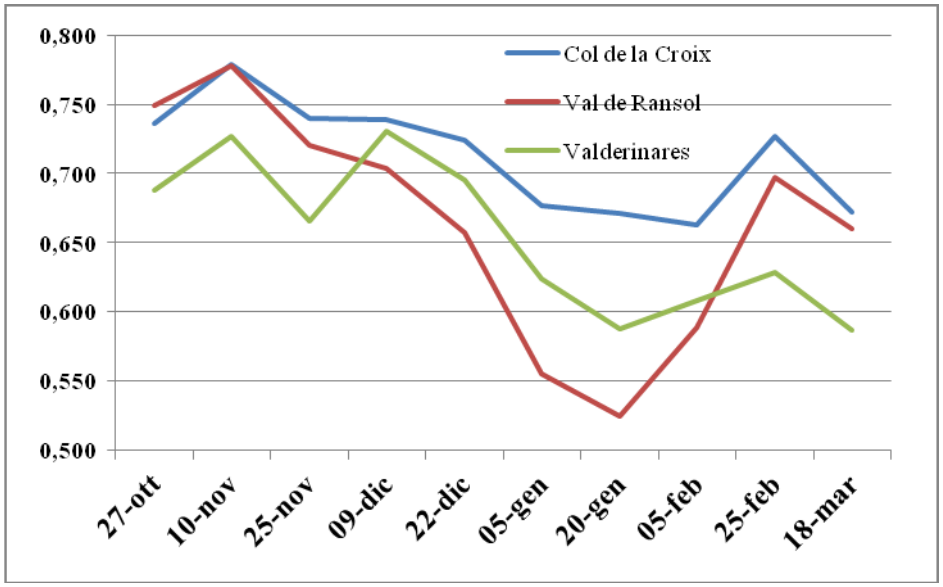


Table 4.8 ANOVA analysis of chlorophyll fluorescence for three different P. *uncinata* populations monitored.

Change	d.f.	S.S.	m.s.	v.r.	F
05_January 2011					
Populations	2	0.1823	0.0911	10.82	<0.001**
Pop./Family	12	0.2513	0.0209	2.49	0.012
Block	4	0.1024	0.0256	3.04	0.025
Residual	53	0.4467	0.0084		
Total	71	0.9826	0.0138		
20_January 2011					
Populations	2	0.2719	0.1360	9.16	<0.001**
Pop./Family	12	0.3856	0.0321	2.16	0.028
Block	4	0.1437	0.0359	2.42	0.06
Residual	53	0.7868	0.0149		
Total	71	1.5880	0.0224		

**P<0.001

5. DISCUSSION

As genetic variation in quantitative traits has been shown in *P. uncinata*. This suggests that some genes may be under differential selection among populations. In this situation, a locus-specific signature of selection could be potentially detectable in contrast to this background genetic variation (Pyhäjärvi *et al.*, 2007; Wachowiak *et al.*, 2009).

Variation in gene expression under different stress conditions and across spatial distributions in different plant tissues suggests functional specialization of the genes (Kontunen-Soppela *et al.*, 2000; Welling and Palva 2006; Wachowiak *et al.*, 2009). In *P. uncinata*, reduced polymorphisms at non-synonymous sites (0.0019) relative to silent sites (0.0071) suggests that the genes could be potentially under strong purifying selection.

The level of nucleotide polymorphism is determined by several factors such as effective population size, mutation rate, selection and demography (Wachowiak *et al.*, 2009). In general, nucleotide polymorphism in conifers is lower compared to angiosperms: e.g. poplar (Ingvarsson 2005), *Arabidopsis* (Wright *et al.*, 2003), and maize (Tiffin and Gaut 2001).

The multilocus level of silent nucleotide diversity detected in the 29 analyzed loci (Θ_W =0.0064), was lower than similar estimates for Scots pine for a set of 14 genes (Θ_W =0.0089, Wachowiak *et al.*, 2009). Compared with others conifers a similar level of silent nucleotide diversity (πs=0.0071) was found across several cold and drought stress candidate genes in *P. sylvestris* (πs=0.0077, Wachowiak *et al.*, 2009), in *Pinus taeda* (πs=0.0085, Gonzalez Martinez *et al.*, 2006a, b), in *P.pinaster* (πs=0.0085, Eveno *et al.*, 2008), and in *Pseudotsuga menziesii* (πs=0.0106, Krutovsky and Neale 2005).

An excess of low frequency polymorphisms marked by significant multilocus Tajima's *D* was also similar to findings for other conifer species (Brown *et al.*, 2004; Bouille and Bousquet 2005; Krutovsky and Neale, 2005; Gonzalez-Martinez *et al.*, 2006a; Heuertz *et al.*, 2006; Wachowiak *et al.*, 2009). Pooling differentiated samples can sometimes result in negative Tajima's *D*.

However, in this study was observed the same results as those indentified in Scots pine in Wachowiak *et al.* (2009), with negative value of Tajima's D at most cold candidate genes and others loci in *P. uncinata* populations.

Usually, average values of Tajima's *D* found in most plant species studied so far are likely the results of demographic processes (such as expanding populations, introgression, extinction and recolonisation events) and/or weak selection which allows deleterious variants rise to a low population frequency (Wachowiak *et al.*, 2009).

As conifers have many gene families, high polymorphism at some loci may potentially result from erroneous co-amplification of different family members (Wachowiak *et al.*, 2009).

We found very rapid decay of intragenic LD with physical distance (r^2=0.24) which is similar to estimates from some genes in Scots pine (Pyhäjärvi *et al.*, 2007; Wachowiak *et al.*, 2009).

The rapid decay of LD suggests that selective sweeps may potentially affect fairly short genomic regions. It seems that the current estimates of LD in conifers based on coding and in most cases relatively short regions are likely to be overestimates of the average genome-wide decay in the species as recombination rates very across the genome (Gaut *et al.*, 2007; Wachowiak *et al.*, 2009).

The 29 candidate genes analyzed had level of genetic diversity (*Hd* = 0.61), lower than those observed in *P. sylvestris* at similar loci (*Hd* = 0.68; Wachowiak *et al.*, 2009).

The possibility of survive of several tree species, including *Pinus uncinata*, was recently justified by predictive modeling on the Iberian Peninsula during the LGM (Last Glacial Maximum) and Holocene (Dzialuk *et al.*, 2009).

The larger than present area of distribution of this species during the LGM and mid-Holocene confirms the expansion of *P. uncinata*'s range in the lower mountain parts and hence a gradual retreat with climate warming to the uppermost parts (Dzialuk *et al.*, 2009). The population in de Sierra de Gùdar (Valderinares) is geographically isolated and separate from those in the Pyrenees.

The Ebro basin and coastal region of the Maestrazgo, where the Sierra de Gúdar is located, has been recognized as a potential Pleistocene refugium of *P. sylvestris* (González-Samperiz *et al.*, 2005; Cheddadi *et al.*, 2006; Dzialuk A. *et al.*, 2009).

It is very probable that *P. uncinata* has been able to remain and survive during the Holocene (van Andel 2002) in the Sierra de Gúdar, which are the most elevated mountains of the Cordillera Iberica peninsula (Dzialuk *et al.*, 2009).

Despite this, we were observed different results for the structure of populations analyzed. In fact our molecular analysis did not reveal fragmentation of populations in most groups, but identified only two groups for the different families of *P. uncinata* considered (AMOVA analyses by Baps software). The very weak geographic organization of these groups, allied to low F_{ST} values, suggests that gene flow is not particularly restricted among these populations, but that among certain localizations, such as between the Massif central in France and the Pyrenees, there is an isolation by distance effect and that gene exchange is rare.

In this study we aimed to establish whether or not populations showed phenotypic divergence by comparing growth under controlled environmental conditions. If such differentiation could be established, this would direct subsequent testing of genetic data for the presence of selection by identifying groups of plants among whom polymorphism might be segregated.

Statistically significant variation was identified in the height of seedlings, both among populations and among families within populations.

In fact, difference in eight was highly significant between populations ($P<0.0001$) with a means value of 2.9 cm for one Spanish population (La-Trapa), and ~4.1 m for France population (Col de la Croix de Morand).

One could hypothesize that the individuals from French location may have found the trial conditions more close to those of the source site, which has a low average altitude compared to other stations (1400 m), and hence may be pre-adapted to the trial environment.

It is likely that the common garden conditions (in the glasshouse), where temperature, humidity and substrate were the same for all seedling were more similar to those at low altitude than high altitude.

Interestingly, no statistically significant differences between populations or among families within populations were found in the timing of bud flush, despite the fact that variation in phenology with latitude has been observed and the populations tested covered

a range of 5 degrees. The latitudinal range is possibly not sufficiently wide among populations to cause differentiation. particularly in a species adapted to altitude.

To assess variation in the photosynthetic efficiency (estimated as F_v/F_m) among populations, chlorophyll fluorescence was measured for three different populations in the glasshouse. Significant differences among populations were detected when outside temperatures were at a minimum, in this case on two different dates (5 and 20 on January).

At these points, the photochemical capacity had minimum values (~0.62 and 0.59 respectively). It is interesting to compare these results with a similar study in Scots pine (Salmela M.J. *et al.*. 2011), when seasonal patterns of photochemical capacity were assessed (estimated with F_v/F_m).

In this case seedling were grown outdoors, and the mean temperature (~2.5°C) was substantially lower than the glasshouse temperature for this *P. uncinata* trial, with a mean value ~ 5°C. However, significant differences were also found when temperatures was a minimum and interestingly, the date at which this occurred (January 22 and March 21 2010. Salmela M.J. *et al.*. 2011), was approximately one month later than for the *P. uncinata* trial (January 5 and 20 2011).

Given the difference in seasonal temperatures between 2010 and 2011, where a cold spell in 2011 occurred approximately one month earlier than in 2010, this timing difference may be to the advance of severe cold conditions between years. Seasonality characterizes the natural habitats of temperate and boreal zone tree species, where environmental conditions are ideal for growth for only a limited period of time.

Trees have adapted to such environmental variation by switching between periods of active growth and dormancy in response to changes in temperature or photoperiod (Salmela M. *et al.*, 2011; Howe *et al.*., 2003). These data suggest that the onset of dormancy may be triggered by the arrival of very cold conditions in both species.

6. CONCLUSION

In this study we explored the genetic diversity of *P.uncinata* collection using a candidate genes loci. In particular the analysis showed a low differentiation in Western Mediterranean *P. uncinata* populations (FST=0.032). Most of the variation among species analyzed is expained within populations (Table 4.2).

One of the most expected hypothesis to explain the low differentiation is related to species history. In fact, for *P. uncinata*, the isolation took place from the Holocene, but this also occurred during each earlier warm interglacial period, similar to others *Pinaceae*. Also, if we suppose, that the Mediterranean *P. uncinata* had the same ancestor population, that we could speculated that there was no possibility to generate substantial differentiation among populations.

However, we found that some populations (e.g. La Trapa) are the most genetically distant from the rest, which suggest independent evolutionary process.

Not weth standing , we observed seasonal variations in photosynthetic activity in *Pinus*. These variations, are partly influenced by temperature and local climatic conditions.

These results represent the beginning of future studies for investigate the genetic variability in Mediterranean forest species, using the Candidate genes approach.

REFERENCES

Achaz G. Frequency Spectrum Neutrality Test: One for All and All for One. Genetics, **183:** 249-258, 2009.

Affre L., Thompson J.D., Debussche M. Genetic structure of continental and island populations of the Mediterranean endemic *Cyclamen balearicum* (Primulaceae). American Journal of Botany, **84(4):** 437-451, 1997.

Ali AA, Carcaillet C, Guendon JL, Quinif Y, Roiron P & Terral JF. The Early Holocene treeline in the southern French Alps: new evidence from travertine formations. Global Ecology and Biogeography, **12:** 411-419, 2003.

Ali AA, Martinez M, Fauvart N, Roiron P, Fioraso G, Guendon JL, Terral JF & Carcaillet C. Incendies at peuplements à *Pinus mugo* Turra dans les Alps occidentales (Val de Sue, Italie) durant la transtition Tardiglaciaire-Holocene: une zone refuge évidente. Comptes Rendus Biologies, **329:** 494-501, 2006.

Armelagos G.J., Harper K.N. Genomics and the origins of agriculture, part one. Evolutionary Anthropology, **14:** 68-77, 2005.

Arrigoni P.V. Le piante endemiche della Sardegna ed 2. Boll. Soc. Sarda Sci. Nat., **16:** 1-202, 1997.

Arroyo García R., Lefort F., de Andrés MT., Ibáñez J., Borrego J., Jouve N., Cabello F., Martínez-Zapater JM. Chloroplast microsatellites polymorphisms in *Vitis* species. Genome **45:** 1142-1149, 2002.

Arroyo García R., Ruiz García L., Bolling L., *et al.* Multiple origins of cultivate grapevine (*Vitis vinifera* L. ssp. *sativa*) based on chloroplast DNA polymorphism. Molecular Ecology, **15:** 3707-3714, 2006.

Axelrod D.I. Bopgeography of oaks in the Arcto-Tertiary province. Ann. Missouri Bot. Gard, **70:** 629-657, 1983.

Bacchetta G., Brullo S., Mossa L. Note tassonomiche sul genere Helichrysum Miller (Asteraceae) in Sardegna. Inform. Bot. Italiano, **35(1):** 217-225, 2003.

Bacilieri R., Ducousso A., Petit RJ, and Kremer A. Mating system and asymmetric hybridization in a mixed stand of European oaks. Evolution, **50:** 900-908, 1996.

Bagnoli F., Vendramin G.G., Buonamici A., Doulis A.G., González-Martínez S.C., La Porta N., Magri D., Raddi P., Sebastiani F., Fineschi S. Is *Cupressus sempervirens* native in Italy? An answer from genetic and paleobotanical data. Molecular Ecology, **18:** 2276-2286, 2009.

Baldoni JA., Tosti N., Riccioli C. *et al.* Genetic structure of wild and cultivated olives in the central Mediterranean basin. Annal of botany, **98:** 935-942, 2006.

Baranski MJ. An analysis of variation within white oak (*Quercus alba* L.). Diss. Abstr. Int., **35(10):** 4779, 1975.

Barcaccia G & Falcinelli M. Genetica e Genomica. Genomica e Biotecnologie genetiche. Vol. III: 783-827, 2006.

Barreneche T., Bodenes C., Frontin J.F., Fluch S., Steiff R., Plomion C., Roussel G., Steinkellner H. and Burg K. *et al.* A genetic linkage map of *Quercus robur* L. (pedunculate oak) based on RAPD, SCAR, microsatellite, minisatellite, isozyme and 5S rDNA markers. TAG **97:** 1090-1103, 1998.

Bassam B.J., Caetano-Anolles G., Gresshoff P.M. Fast and sensitive silver staining of DNA in polyacrylamide gels. Analysis Biochemical, **196(1):** 80-3, 1991.

Belahbib N., Pemonge MH., Ouassou A., Sbay H., Kremer A., Petit RJ. Frequent cytoplasmic exchanges between oak species that are not closely related: *Quercus suber* and *Q. ilex* in Morocco. Molecular Ecology, **10(8):** 2003-2012, 2001.

Bellarosa R. Introduction: brief synthesis of the current knowledge on cork oak. In Valera MC (eds.) Evaluation of genetic resources of cork oak for appropriate use in breeding and gene conservation strategies. Handbook of the concerted action on cork oak, reference FAIR 1 CT 95-0202, 2000. INIA-Estação Florestal Nacional, Portugal.

Besnard G., Baradat P., Bervillé A. Genetic relationships in the olive (*Olea europaea* L.) reflect multilocal selection of cultivars. Theoretical and Applied Genetics, **102:** 251-258, 2001.

Besnard G., Khadari B., Baradat P., Bervillé A. *Olea europaea* (Oleaceae) phylogeography based on chloroplast DNA polymorphism. Theoretical and Applied Genetics, **104:** 1353-1361, 2002.

Blondel J., Aronson J. *Biology and Wildlife of the Mediterranean Region*. Oxford University Press, Oxford. UK., 1999.

Boratyńska & Boratyński. Taxonomic differences among closely related pines *Pinus sylvestris*, *P. mugo*, *P. uncinata*, *P. rotundata* and *P. uliginosa* as revealed in needle sclerenchyma cells. Flora, **202:** 547-554. 2007.

Bouille M, Bousquet J. Trans-species shared polymorphisms at orthologous nuclear gene loci among distant species in the conifer *Picea* (Pinaceae): implications for the long-term maintenance of genetic diversity in trees. Am Journal Botany **92:** 63-73, 2005.

Brown GL, Gill GP, Kuntz RJ, Langley CH, Neale DB. Nucleotide diversity and linkage disequilibrium in loblolly pine. Proc Natl Acad Sci USA**, 101:**15255-15260, 2004.

Bruschi P., Vendramin GG., Bussotti F., Grossoni P. Morphological and Molecular Differentiation between *Quercus petraea* (Matt.) Liebl. and *Quercus pubescens* Willd. (Fagaceae) in Northern and Central Italy. Annals of Botany, **85(3):** 325-333, 2000.

Brush S.B. The issues of in situ conservation of crop genetic resources. In: Brush SB (eds.), Genes in the Field. IPGRI Rome, IDRC Ottawa, Lewis Boca Raton, pp. 3-26.

Bryan GJ., McNicoll J., Ramsay G., Meyer RC., De JWS. Polymorphic simple sequence repeats markers in chloroplast genomes of *Solanaceous* plants. Theoretical and Applied Genetics, **99:** 859-867, 1999.

Camarero JJ, Gutiérrez E, Fortin MJ & Ribbens E. Spatial patterns of tree recruitment in a relict population of *Pinus uncinata*: forest expansion trough stratified diffusion. Journal of Biogeography, **32:** 1979-1992, 2005.

Cantegrel R. Le Pin à crochets pyrénéen: biologie, biochimie, sylviculture. Acta Biologica Montana, **2-3:** 87-330, 1983.

Cardwell K.S., Russel J., Langridge P., Powell W. Extreme population-dependent linkage disequilibrium detected in an inbreeding plant species, *Hordeum vulgare*. Genetics, **172:** 557-567, 2006.

Carpita N, Sabularse D, Montezinos D, Delmer DP. Determination of pore size of cell walls of living plant cells. Science, **205:** 1144–1147, 1979.

Carrion JS. Patterns and processes of late Quaternary environmental change in a montane region of southwestern Europe. Quaternary Science Reviews, **21:** 2047-2066, 2002.

Carrión JS., Parra I., Navarro C., Munuera M. Past distribution and ecology of the cork oak (*Quercus suber* L.) in the Iberian Peninsula: a pollen-analytical approach. Diversity and Distributions, **6:** 29-44, 2000a.

Carter KK. Provenance tests as indicators of growth response to climate change in 10 north temperate tree species. Can J For Res, **26:** 1089-95, 1996.

Cheddadi R, Vendramin GG, Litt T, Francois L, Kageyama M, Lorentz S, Laurent JM, de Beaulieu JL, Sadori L, Jost A, Lunt D. Imprints of glacial refugia in the modern genetic diversity of *Pinus sylvestris*. Global Ecology Biogeography, **15:** 271-282, 2006.

Christensen KI. Taxonomy revision of *Pinus mugo* complex and *P.* x *rhaetica* (*P. mugo* x *sylvestris*) (Pinaceae). Nordic Journal of Botany, **7:** 383-408, 1987.

Chung S., Decker-Walters DS., Staub JE. Genetic relationships whitin the Cucurbitaceae as assessed by ccSSR marker and sequence analyses. Canadian Journal of Botany, **81:** 814-832, 2003.

Close TJ. Dehydrins: emergence of a biochemical role of a family of plant dehydration proteins, Physiol. Plant. **97:** 795–803, 1996.

Comes H.P., Kadereit J.W. The effect of Quaternary climatic changes on plant distribution and evolution. Trends in Plant Science **3:** 432-438, 1998.

Comes HP. The Mediterranean region a hotspot for plant biogeography research. New Phytologist **164:** 11-14, 2004.

Corander J, Tang J. Bayesian analysis of population structure based on linked molecular information. Math Biosci **205:** 19-31, 2007.

Corona PM., Dettori S., Filigheddu MR., Maetzke F., Scotti R. Site quality evaluation by classification tree: an application to cork quality in Sardinia. Eur. J. Forest Res., **124:** 37-46, 2005.

Cottam WP., Tucker JM and Santamour FS. "Oak Hybridisation at the University of Utah", State Arboretum of Utah Publ. No. 1. Salt Lake City, UT, 1982.

Crepet W.L. History and implications of the early North American fossil record of Fagaceae. In Evolution, Systematics, and Fossil History of the Hamamelidae, Vol. 2: 'Higher' Hamamelidae" (P.R. Crane and S. Blackmore, Eds.), pp. 45-66. Clarendon, Oxford, 1989.

Crepet WL., Nixon KC. Earlier megafossil evidence of Fagaceae: Phylogenetic and biogeographic implications. Am. J. Bot., **76:** 842-855, 1989.

Critchfield & Little. Geographic distribution of the pines of the World. USDA. Forest Service, 1969.

Cuenca A., Escalante E., Pinero D. Long-distance colonization, isolation by distance, and historical demography in a relictual Mexican pinyon pine (*Pinus nelsoni* Shaw) as revealed by paternally inherited genetic markers (cpSSRs). Molecular Ecology, **12:** 2087-2097, 2003.

Demesure B., Comps B., Petit RJ. Chloroplast DNA phylogeography of the common beech (*Fagus sylvatica* L.) in Europe. Evolution, **50:** 2515-2520, 1996.

Dettori S., Falqui A., Filigheddu MR., Sedda L. *Performances* di recenti rimboschimenti con quercia da sughero in ex-coltivi. Forest@, **3(3):** 327-338, 2006.

Dow B.D., Ashley M.V., Howe H.F. Characterisation of highly variable (GA/CT)n microsatellites in the bur oak, *Quercus macrocarpa.* Theor Appl Genet **91:** 137-141,1995.

Dow BD., and Ashley MV. Microsatellite analysis of seed dispersal and parentage of sapling in bur oak, *Quercus macrocarpa*. Mol. Ecology, **5:** 615-627, 1996.

Doyle JJ., Morgante M., Tingey SV., Powell W. Size homoplasy in chloroplast microsatellites of wild perennial relatives of soybean (Glycine subgenus Glycine). Molecular Biology and Evolution **15:** 215-218, 1998.

Dumolin-Lapègue S., Demesure B., Petit R.J. Inheritance chloroplast and mitochondrial genomes in pedunculate oak investigated with an efficient PCR-based method. Theoretical and Apllied Genetics **91:** 1253-1256, 1995.

Dumolin-Lapègue S., Demesure B., Fineschi S., Le Corre V., Petit RJ. Phylogeographic structure of white oaks throughout the European continent. Genetics, **146:** 1475-1487, 1997a.

Dumolin-Lapègue S., Pemonge MH., Petit RJ. An enlarged set of consensus primers for the study of organelle DNA in plants. Molecular Ecology, **6:** 393-397, 1997b.

Dzialuk A., Chybicki I., Burczyk J. PCR Multiplexing of Nuclear Microsatellites Loci in *Quercus* Species. Plant Mol. Bio. Rep., **23:** 121-128, 2005.

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

Eckert CG, Samis KE, Lougheed SC. Genetic variation across species geographical ranges: the central marginal hypothesis and beyond. Molecular Ecology, **17:** 1170-1188, 2008.

Ennos RA., Sinclair WT., Hu XS., Langdon A. Using organelle markers to elucidate the history, ecology and evolution of plant populations. In Mol. Syst. and Plant Evol. (Hollingsworth, P.M. *et al.*, eds.), pp.1-19, 1999.

Elliot KJ., Vose JM., Swank WT., and Bolstad PV. Long term patterns in vegetation-site relathionship in a southern Appalachian forest. Journal of Turr. Bot. Soc., **126(4):** 320-334, 1999.

Eriksson G. Conservation of noble hardwoods in Europe. Can. J. For. Res., **31:** 577-587, 2001.

Evanno G., Regnaut S., Goudet J. How to use the software STRUCTURE: a simulation study. Molecular Ecology, **14:** 2611-2620, 2005.

Eveno E, Collada C, Guevara MA, Léger V, Soto A, Díaz L, Léger P, González-Martínez SC, Cervera MT, Plomion C, Garnier Géré PH. Contrasting patterns of selection at Pinus pinaster Ait. Drought stress candidate genes as revealed by genetic differentiation analyses. Mol Bio Evol**, 25(2):** 417-437, 2008.

Excoffier L, Laval G, Schneider S. Arlequin: An integrated software package for population genetics data analysis. Evol Bioninform Online**, 1:** 47-50, 2005.

Fady-Welterlen B. Is there really more biodiversity in Mediterranean forest ecosystems?. Taxon **54:** 905-910, 2005.

Ferreira A., Lopes F., Pereira H. Caractérisation de la croissance at et de la qualitè du liège dans une région de production. Ann. For. Sci., **57:** 187-193, 2000.

Ferris C., Oliver RP., Davy AJ., and Hewitt GM. Native oak chloroplasts reveal an ancient divide across Europe. Molecular Ecology, **2:** 337-344, 1993.

Fineschi S., Taurchini D., Villani F., Vendramin GG. Chloroplast DNA polymorphism revealed little geographical structure in *Castanea sativa* Mill. (Fagaceae) throughout southern European countries. Molecular Ecology, **9:** 1495-1503, 2000.

Forest Europe Report. Status & Trends in Sustainable Forest Management in Europe. UNECE and FAO Congress, 2011.

Franco Múgica F, García Antón M & Sainz Ollero H. Vegetation dynamics and human impact in the Sierra de Guadarrama, Central System, Spain. The Holocene, **8**: 69-82, 1998.

Frascaria N., Maggie L., Michaud M., and Bousquet J. The *rbc*L gene sequence from chestnut indicates a slow rate of evolution in the Fagaceae. Genome, **36:** 667-671, 1993.

Fréville H., Colas B., Ronfort J., Riba M., Olivieri I. Predicting Endemism from Population Structure of a Widespread Species : Case Study in Centaurea maculosa Lam. (Asteraceae). Cons. Biol., **12(6):** 1269-1278, 1998.

Frewen BE, Chen THH, Howe GT, Davis J, Rohde A, Boerjan W and Bradshaw HD. Quantitative trait loci and candidate gene mapping of bud set and bud flush in *Populus*. Genetics, **154**: 837–45, 2000.

Gaussen H, Heywood VH & Chater AO. *Pinus. Flora Europaea*, Vol. **1**,2 edn (ed. by T.G. Tutin, N.A. Burges and A.O. Chaters), pp. 40-44. Cambridge University Press, Cambridge, 1993.

Gaut BS, Wright SI, Rizzon C, Dvorak J, Anderson LK. Opinion-Recombination: an underappreciated factor in the evolution of plant genomes. Nat Rev, Genetics, **8**: 77-84, 2007.

Gellini R, Grossoni P.. Botanica forestale, Vol. II Angiosperme. CEDAM, 1997.

Gepts P., Kmeieciz K., Pereira P., Bliss F.A. Dissemination patways of common bean (*Phaseolus vulgaris*, Fabaceae) deduced from phaseolin electroforetic variability. I The Americas. Econ Bot. **42:** 86-104, 1988.

Gil García MJ, Dorado-Valiño M, Valdeolmillos-Rodríguez A & Ruiz-Zapata MB. Late-glacial and Holocene paleoclimatic record from Sierra de Cebollera (northern Iberian Range, Spain). Quaternary International, **93-94**, 13-18, 2002

Gómez-Aparicio L., Valladares F., Zamora R., Quero JL. Response of tree seedlings to the abiotic heterogeneity generated by nurse shrubs: an experimental approach at different scales. Ecography, **28(6):** 757-767, 2005.

González -Sampériz P, Valero Garcés BL, Carrión JS, Peña-Monné JL, García Ruiz JM, Martí-Bono C. Glacial and Late glacial vegetation in northeastern Spain: new data and a review. Quat Int**,** **140-141**, 4-20.

González-Martínez SC, Ersoz E, Brown GR, Wheeler NC, Neale DB. DNA sequence variation and selection of tag single-nucleotide polymorphisms at candidate genes for drought-stress response in *Pinus taeda* L. Genetics **172**, 1915-1926, 2006a.

González-Martínez SC, Krutovsky KV, Neale DB. Forest-tree population genomics and adaptive evolution. New Phytologist **170**, 227-238, 2006b.

Grant V. “Plant Speciation”, Columbia University Press, New York, 1981.

Greuter W. Botanical diversity, endemism, rarity, and extinction in the Mediterranean area: an analysis based on the published volumes of the Med-Checklist. Bot. Chronika, **10:** 63-79, 1991.

Greuter W. Origin and peculiarities of Mediterranean island floras. Ecologia Mediterranea, **21:** 1-10, 1995.

Gugerli F., Sperisen C., Büchler U., Magni F., Geburek T., Jeandroz S., Senn J. Haplotype variation in a mitochondrial tandem repeat of Norway spruce (*Picea abies*) populations suggests a serious founder effect during postglacial re-colonization of the western Alps. Molecular Ecology, **10(5):** 1255-1263, 2001.

Guichoux E., Lagache L., Wagner S., Léger P., Petit RJ. Two highly validated multiplexes (12-plex and 8-plex) for species delimitation and parentage analysis in oaks (*Quercus spp*.). molecular Ecology Res., **11(3):** 578-585, 2011.

Hamerník J & Musil I. The *Pinus mugo* complex – its structuring and general overview of the used nomenclature. *Journal of Forest Science*, **53**, 253-266, 2007.

Hampe A & Petit RJ. Conserving biodiversity under climate change: the rear edge matters. *Ecology Letters*, **8**, 461-467, 2005.

Hampe A., Petit R.J. Ever deeper phylogeographies: trees retain the genetic imprint of Tertiary plate tectonics. Molecular Ecology, **16:** 5113-5114, 2007.

Hardin JW. Hybridization and introgression in *Quercus alba.* J. Arnold Arbor., **56(3):** 336-363, 1975.

Hedrick PW. Genetic polymorphism in heterogeneous environments: the age of genomics. *Annu Rev Ecol Evol Syst* **37**, 67-93, 2006.

Heer J., Agrawala M. Software design patterns for information visualization. Citeseer, **12(5):** 853-860, 2006.

Heuertz M., Fineschi S., Anzidei M., Pastorelli R., Salvini D., Paule L., Frascaria-Lacoste N., Hardy OJ., Vekemans X., Vendramin GG. Chloroplast DNA variation and postglacial recolonization of common ash (*Fraxinus excelsior* L.) in Europe. Molecular Ecology, **13(11):** 3437-3452, 2004.

Heuertz M, De Paoli E, Kallman T, Larsson H, Jurman I, Morgante M, Lascoux M, Gyllenstrand N. Multilocus patterns of nucleotide diversity, linkage disequilibrium and demographic history of Norway spruce [*Picea abies* (L.) Karst]. Genetics **174**, 2095-2105, 2006a.

Heuertz M., Carnevale S., Fineschi S., Sebastiani F., Hausman JF., Paule L., Vendramin GG. Chloroplast DNA phylogeography of European ashes, *Fraxinus* sp. (Oleaceae): roles of hybridization and life history traits. Molecular Ecology, **15(8):** 2131-2140, 2006b.

Hewitt G.M. Post-glacial re-colonisation of European biota. Biological Journal of the Linnean Society, **68:** 87-112, 1999.

Hewitt G.M. The genetic legacy of the Quaternary ice bags. Nature, **405:** 907-913, 2000.

Hill WG, Robertson A. Linkage disequilibrium in finite populations. Theor Appl Genet, **38:** 226-231, 1968.

Howard DJ., Preszler RW., Williams J., Fenchel and Boecklen S. How discrete are oaks species? Insights from a hybrid zone between *Quercus grisea* and *Quercus gambelii*. Evolution **51:** 747-755, 1997.

Howe GT, Aitken SN, Neale DB, Jermstad KD, Wheeler NC, Chen THH. From genotype to phenotype: unravelling the complexities of cold adaptation in forest trees. Can J Bot, **81**: 1247-1266, 2003.

Hudson RR, Boos DD, Kaplan NL. A statistical test for detecting geographic subdivision. Mol Biol Evol, **9:** 138-151, 1992.

Hudson RR. A new statistic for detecting genetic differentiation. Genetics, **155:** 2011-2014, 2000.

Hughes AL, Nei M. Pattern of nucleotide substitution at major histocompatibility complex class I loci reveals overdominant selection. Nature**, 335:** 167-170, 1988.

Hurme P, Sillanpää MJ, Repo T and Savolainen O. Genetic basis of climatic adaptation in Scots pine by Bayesian quantitative locus analysis. Genetics, **156:** 1309–22, 2000.

Hurtrez-Boussès S., Blondel J., Perret P., Fabreguettes J., Reanud F. Chick parasitism by bowflies affects feeding rates in a Mediterranean population of blue tits. Ecology Letters, **1(1):** 17-20, 1998.

Ingram J, Bartels D. The molecular basis of dehydration tolerance in plants. *Annual Rev.* Plant Physiology. Plant Mol. Biol., **47:** 377–403, 1996.

Ingvarsson PK. Nucleotide polymorphism and linkage disequilibrium within and among populations odf European aspen (*Populus tremula* L. Salicaceae). Genetics, **169:** 945-953, 2005.

Ingvarsson PK, García MV, Hall D, Luquez V and Jansson S. Clinal variation in *phy*B2, a candidate gene for day-length-induced growth cessation and bud set, across a latitudinal gradient in European aspen (*Populus tremula*). Genetics, **172:** 1845-1853, 2006.

Ishii T. and McCouch SR. Microsatellites and microsynteny in the chloroplast genomes of *Oryza* and eight other *Gramineae* species. Theoretical and Applied Genetics, **100:** 1257-1266, 2000.

Ishii T., Xu Y., McCouch SR. Nuclear and Chloroplast-microsatellite variation in A-genome species of rice. Genome, **44:** 658-666, 2001.

Jiang C, Schommer C, Kim SY and Suh DY. "Cloning and Characterization of Chalcone Synthase from the moss Physcomitrella patens". Phytochemistry, **67(23):** 2531–2540, 2006.

Johnsen O., & Skrøppa T. Parents of Norway spruce adjust the performance of their progeny according to changes in climate and weather conditions during female flowering, a review. Pp. 159-163, in G.M.J. Mohren, K. Kramer and S. Sabaté (editors). *Impact of global change on tree physiology and forest ecosystems*. Proceedings of the International Conference on Impacts of Global change on Tree Physiology and Forest Ecosystems, Wageningen, The Netherlands, 26-29 November 1996. Kluwer Academic Publishers, Dordecht, The Netherlands. Published as Forestry Science, Vol. **52.** 388 p. 1997.

Joosen RVL, Lammers M, Balk PA, Brommun P, Koning MCJM, Perks M, Stattin E, van der Geest AHM. Correlation gene expression to physiological parameters and environmental conditions during cold acclimation of *Pinus sylvestries*, identification of molecular markers using cDNA microarrays. *Tree Physiol*, **26:** 1297-1313, 2006.

Kampfer S., Lexer C., Glössl J. and Steinkellner H. Characterization of $(GA)_n$ microsatellite loci from *Quercus robur* . Hereditas **129:** 183-186, 1998.

Kawecki TJ, Ebert D. Conceptual issues in local adaptation. *Ecol Lett* **7:** 1225-41, 2004.

König AO. Provenance research: evaluating the spatial pattern of genetic variation. In: Geburek T. & Turok K. (eds.). Conservation and Management of Forest Genetic Resources in Europe, pp. 275-333. Arbora Pubblishers, Zvolen, Slovaka, 2005.

Kontunen-Soppela S, Taulavuori K, Taulavuori E, Lahdesmaki P, Laine K. Soluble proteins and dehydrins in nitrogen-fertilized Scots pine seedlings during deacclimation and the onset of growth. Physiology Plantarum **109:** 404-409, 2000.

Kremer A., Kleinschmit J., Cottrell J., Cundall EP., Deans JD., Ducousso A. et al. Is there a correlation between chloroplastic and nuclear divergence, or what are the roles of history and selection on genetic diversity in European oaks?. Forest Ecology and Management, **156(1-3):** 75-87, 2002.

Kremer A. How well can existing forests withstand climate change? In: *Climate Change and Forest Genetic Diversity: Implication for Sustainable Forest Management in Europe* (eds Koskela J., Buck A., Teissier du Cros E.), pp. 3-17. Biodiversity International, Rome Italy, 2007.

Krutovsky KV, Neale DB. Nucleotide diversity and linkage disequilibrium in cold-hardiness and wood quality-related candidate genes in Douglas fir. *Genetics* **171:** 2029-2041, 2005.

Lande R., and Barrowclough GF. Effective population size, genetic variation, and their use in population management, pp. 87-123, in Viable populations for conservation (M.E. Soulé, eds.). Cambridge University Press. Cambridge (189 pp), 1987.

Le Corre V., Dumolin-Lapègue S. & Kremer A. Genetic variation at allozyme and RAPD loci in sessile oak *Quercus petraea* (Matt.) Liebl.: the role of history and geography. Molecular Ecology, **6(6):** 519-529, 1997.

Ledig FT. Heterozygosity, heterosis, and fitness in outbreeding plants. Pp.77-104, in: M.E. Soulé (editor). *Conservation Biology. The science of scarcity and diversity*. Sinauer Associates, Sunderland, MA, USA.

Lee M. DNA markers and plant breeding programs. Advances in Agronomy Vol. **55:** 265-344, 1995.

Lepais O., Léger V., and Gerbes S. Short Note: High Troughput Microsatellites Genotyping in Oak Species. Silvae Genetics, **55:** 4-5, 2006.

Levin D.A. The Origin, Expansion, and Demise of Plant Species. Oxford University Press, Oxford, 2000.

Lewandowski A, Boratińsky A, Mejnartowicz L. Allozyme investigations on the genetic differentiation between *Pinus sylvestris*, *P.mugo*, *P.uncinata* and *P. uliginosa*. Plant Systematic and Evolution **221**, 15-24, 2000.

Linhart Y.B. & Grant M.C. Evolution significance of local genetic differentiation in plant. Annu. Rev. Ecol. Syst. **27:** 237-277, 1996.

Liu A., Burke I.M. Patterns of nucleotide diversity in wild and cultivated sunflower. Genetics, **173:** 321-330, 2006.

López R., Rodríguez-Calcerrada J., Gil L. Physiological and morphological response to water deficit in seedlings of five provenances of *Pinus canariensis*: potential to detect variation in drought-tolerance. Tree Structure and Function, **23:** 509-519, 2009.

López de Heredia U., Jiménez P., Díaz-Fernández P., Gil L. The Balearic islands: a reservoir of cpDNA genetic variation for evergreen oaks. Journal of Biogeography, **32:** 939-949, 2005.

Lumaret R., Mir C., Michaud H., Raynal V. Phylogeographical variation of chloroplast DNA in holm oak (*Quercus ilex* L.). Molecular Ecology, **11:** 2327-2336, 2002

Lumaret R., Ouazzani N., Michaud H. *et al.* Allozyme variation of oleaster populations (wild olive tree) (*Olea europaea* L.) in the Mediterranean basin. Heredity, **92:** 343-351, 2004.

Lumaret R., Tryphon-Dionnet M., Michaud H., Sanuy A., Ipotesi E., Born C., Mir C. Phylogeographical Variation of Chloroplast DNA in Cork oak (*Quercus suber*). Annals of Botany **96:** 853-861, 2005.

Lumaret R. and Jabbour-Zahab R. Ancient and current gene flow between two distantly related Mediterranean oak species, *Quercus suber* and *Q. ilex*. Annals of Botany, **104:** 725-736, 2009.

Lynch M, Crease TJ. The analyses of population survey data on DNA sequence variation. Mol Biol Evol **7**, 377-394, 1990.

Lynch M. A quantitative genetic perspective on conservation issues, in: Avise J., Hamrik J. (Eds.), Conservation genetics: case histories from nature Chapman and Hall, New York: pp. 471-501, 1996.

Madlung A. & Comai L. The effect of stress on genome regulation and structure. Annals of Botany, **94:** 481-495. 2004.

Magni Diaz CR. Reconstitution de l'introcution de *Quercus rubra* L. en Europe at consequences dans les populations allochtones. PhDF Thesis, ENGREF, Paris, France. 419 p., 2005.

Magri CR., Ducousso A., Caron H., Petit RJ and Kremer A. Chloroplast DNA variation of *Quercus rubra* L. in north America and comparison with other Fagaceae. Molecular Ecology, **14:** 513-524, 2005.

Magri D., Vendramin GG., Compos B., Dupanloup I., Geburek T., Gomory D., Latalowa M., Litt T, Paule L., Roure JM., Tantau I., van der Knapp WO., Petit RJ., and Beaulieur JL. A new scenario for the Quaternary history of European beech populations: palaeobotanical evidence and genetic consequences. New Phytologist, **171:** 199-221, 2006.

Magri D., Fineschi S., Bellarosa R., Buonamici A., Sebastiani F., Schirone B., Simeone M.C. and Vendramin G.G. The distribution of *Quercus suber* chloroplast haplotypes matches the paleogeographical history of the western Mediterranean. Molecular Ecology **16:** 5259-5266, 2007.

Manos PS., Doyle JJ., Nixon KC. Phylogeny, Biogeography, and Processes of Molecular Differnetitation in *Quercus* (Fagaceae). Molecular Phylogenetics and Evolution, **12(3):** 333-349, 1999.

Marshall HD., Newton C., Ritland K. Chloroplast phylogeography and evolution of highly polymorphic microsatellites in lodgepole pine (*Pinus contorta*). Theoretical and Appl. Genetics, **104:** 367-378, 2002.

Mátyás C. Climatic adaptation of trees: rediscovering provenance tests. Euphytica, **92:** 45-54, 1996.

Matus IA., and Hayes PM. Genetic diversity in three groups of barley germplasm assessed by simple sequence repeats. Genome, **45:** 1095-1106, 2000.

Médail F., Quézel P. Biodiversity Hotspots in the Mediterranean Basin: setting global conservation priorities. Conservation Biol. **13(6):** 1510-1513, 1999.

Milborrow BV. The patway of biosynthesis of abscisic acid in vascular plants: a review of the present state of knowledge of ABA biosynthesis. Journal of Experimental Botany **52**, 1145-1164, 2001.

Miller A., Schaal B.A. Domestication and the distribution of genetic variation in wild and cultivated populations of the Mesoamarican fruit tree *Spondias purpurea* L. (Anacardiaceae). Molecular Ecology, **15(6):** 1467-1480, 2006.

Milligan BG., Leebens-Mack J., Conservation of genetics: beyond the maintenance of marker diversity. Molecular Ecology, **3(4):** 423-435, 1994.

Mitchell SE., Kresovich S., Jester CA., Hernandez CJ. And Szewe-Fadden AK. Application of multiplex PCR and fluorescence-based, semi-automated alleles sizing technology for genotyping plant genetic resources. Crop. Sci., **37:** 617-624, 1997.

Mitton JB. & Grant MC. Association between protein heterozygosity, growth rate, and developmental homeostasis. Annual Review of Ecology and Systematics, **15:** 479-499, 1984.

Mitton JB. Selection in natural populations. Oxford University Press, Oxford, UK, 1997.

Molles M. C. Ecology: *Concepts and Applications* (International Edition). New York. The Mc Grown Hill Companies. Inc., pp.586. 2002.

Monteleone I, Ferrazzini D & Belletti P. Effectiveness of neutral RAPD markers to detect genetic divergence between the subspecies *uncinata* and *mugo* of *Pinus mugo* Turra. *Silva Fennica*, **40**, 391-406, 2006.

Morjan C.L. & Rieseberg L.H. How species evolve collectively: implications of gene flow and selection for the spread of advantageous alleles. Molecular Ecology **13:** 1341-1356, 2004.

Mugica Franco F., Garcia Anton M., Sainz Ollero H. Vegetation dynamics and human impact in the Sierra de Guadarrama, Central System, Spain. The Holocene **8:** 69.82, 1998.

Muller C.H. Ecological control of hybridization in *Quercus*: A factor in the mechanism of evolution. Evolution, **6:** 147-161, 1952.

Nambara E, Marion-Poll A. Abscisic acid biosynthesis and catabolism. *Ann. Rev. Plant Biology* **56**, 165-185, 2005.

Natividade JV. Cortiças. Contribução para o estudo do melhoramento da qualidade. Pubblicações da Direcção Geral dos Serviços Florestais e Aquícolas I (1), Lisboa, 1934.

Neale DB, Savolainen O. Association genetics of complex traits in conifers. *Tren Pl Sci* **9**:325-330, 2004.

Neale DB, Ingvarsson PK. Population, quantitative and comparative genomics of adaptation in forest trees. *Curr Opin Plant Biology*, **11**, 149-155, 2008.

Nei M. Estimation of average hetozygosity and genetic distance for a small number of individuals. Genetics, **89:** 583-590, 1978.

Nei M. and Li W. Mathematical Model for Studying Genetic Variation in Terms of Restriction Endonucleases. PNAS, **76(10):** 5269-5273, 1979.

Neophytou C.H., Dounavi A., Aravanopoulos F.A. Conservation of Nuclear Loci Reveals High Affinity of *Quercus infectoria* ssp. *veneris* A. Kern (Fagaceae) to section *Robur*. Plant Molecular Bilogy **26:** 133-141, 2008.

Nixon K.C. "A Biosystematic Study of *Quercus* series *Virentes* (the Live oaks) with Phylogenetic Analyses of Fagales, Fagaceae and *Quercus*", PhD dissertation, University of Texas, Austin, 1985.

Nixon K.C. Infrageneric classification of *Quercus* (Fagaceae) and typification of selection names. Ann. Sci. For. Suppl. 1 (Paris), **50:** 25-34, 1993.

Oddou-Muratorio S., Petit RJ., Le Guerroue B., Guesnet D., Demesure B. Pollen- versus seed-mediated gene flow in a scattered forest tree species. Evolution, 55(6): 1123-1135, 2001.

Olalde M., Herrán A., Espinel S., Goicoechea PG. White oaks phylogeography in the Iberian Peninsula. Forest Ecol. And Man., **156(1-3):** 89-102, 2002.

Oleksyn J., Modryzynski J., Tjoelker MG., Zytkowiak R., Reich PB., Karolewski P. Growth and Physiology of *Picea abies* populations from elevational transects: Common Garden evidence for Altitudinal Ecotypes and Cold Adaptation. Functional Ecology, **12 (4):** 573-590, 1998.

Otha T. and Kimura M. A model of mutation appropriate to estimate the number of electrophoretically detectable alleles in a genetic population. Genetic Resources, **22:** 201-204, 1973.

Ouvrard O, Cellier F, Ferrare K, Tousch D, Lamaze T, Dupuis JM, Casse-Delbard F. Identification and expression of water stressand abscisic acid-regulated genes in drought-tolerance sunflower genotype Plant. Mol. Biol., **31:** 819–829, 1996.

Palmer E.J. Hybrid oaks of North America. J. Arnold Arbor. Harv. University, **29:** 1-48, 1948.

Palmer E.J. Plastid chromosomes: structure and evolution. In: RG Hermann (Eds.), The Molecular Biology and Plastids. Cell Culture and Somatic Cell Genetics of Plants vil 7A, Springer-Verlag (Vienna), pp. 5-53, 1991.

Pereira H., Lopes F., Graça J. The evaluation of the quality of cork planks by image analysis. Holzforschung, **50(2):** 111-115, 1996.

Perry H, Hollis D. The generation of monthly gridded datasets for a range of climatic variables over the UK. International journal of Climatology, **25:** 1041-1054, 2005.

Petit RJ., Kremer A. and Wagner DB. Geographic structure of chloroplast DNA polymorphism in European oaks. Theor. Appl. Genet., **87:** 122-128, 1993a.

Petit RJ., Kremer A. and Wagner DB. Finite island model of organelle and nuclear genes in plants. Heredity, **71:** 630-641, 1993b.

Petit RJ, Csaikl UM., Bordács S., Burg K., Coart E., Cottrell J., van Dam B., Deans DJ., Sylvie Dumolin-Lapègue S., Fineschi S., Finkeldey R., Gillies A., Izabela Glaz I., Goicoechea PG., Jensen JS., König AO., Lowe AJ., Madsen SF., Mátyás G., Munro RC., Olalde M., Pemonge MH., Popescu F., Slade D., Tabbener H., D. Taurchini, de Vries SGM., Ziegenhagen B., Kremer A. Chloroplast DNA variation in European white oaks: Phylogeography and patterns of diversity based on data from over 2600 populations. Forest Ecol. And Man., **156 (1-3):** 1-26, 2002.

Pignatti S. Flora d'Italia, Vol. I, p.790. Edagricole, Bologna, Italy, 1982.

Pons A. & Reille M. 1988. The Holocene and upper Pleistocene pollen record from Padul (Granada, Spain): a new study. Palaeogeography, Palaeoclimatology, Palaeoecology, **66**, 243-263.

Powell W., Morgante M., McDevitt R., Vendramin GG., Rafalski J. Polymorphic simple-sequence repeat regions in chloroplast genomes: applications to the population genetics of pines. Proceeding of National Academy of Science USA **92:** 7759-7763, 1995.

Powell W., Morgante M., Doyle JJ., McNicol JW., Tingey SV., Rafalski AJ. Genepool variation in the genus *Glycine* subgenus *Soja* revealed by polymorphic nuclear and chloroplast microsatellites. Genetics, **144:** 792-803, 1996.

Pritchard JK., Stephen M., Rosenberg NA., Donnelly P. Association mapping in Structure Populations American Journal of Human Genetics **67:** 170-181, 2000.

Pritchard JK. and Wen W. Documentation for Structure Software, version 2. Department of Human Genetics, University of Chicago, Chicago, 2003.

Provan J., Corbett G., McNicol JW., Powell W. Chloroplast variability in wild and cultivated rice (*Oryza sativa*) revealed by polymorphic chloroplast simple sequence repeats. Genome, **40:** 104-110, 1997.

Provan J., Soranzo N., Wilson NJ., Goldstein DB., Powell W. A low mutation rate for Chloroplast Microsatellites. Genetics, **153:** 943-947, 1999.

Provan J., Powell W., Hollingsworth PM. Chloroplast microsatellites: new tools for studies in plant ecology and evolution. Trend in Ecol. & Evolution, **16(3)**, 2001.

Pyhäjärvi T, García-Gil MR, Knürr T, Mikkonen M, Wachowiak W, Savolainen O. Demographic history has influenced nucleotide diversity in European *Pinus sylvestris* populations. Genetics **177**, 1713-1724, 2007.

Quézel P. La flore du bassin Méditerranèen: origine, mise en place, endémisme. Ecologia Mediterranea **21:** 19-39, 1995.

Quézel P. & Médail F. Ecologie et biogéographie des forêts du bassin Méditerranéen. Elsevier (Collection Environnement), Paris, 2003.

Ramil-Rego P., Munoz-Sobrino C., Rodrìguez-Guitìan M., Gòmez-Orellana L.. Differences in the vegetation of the North Iberian Peninsula during the last 16.000 years. Plant Ecology, **138:** 41-62, 1998.

Ramírez-Valiente JA., Lorenzo Z., Soto A., Valladares F., Gil L., Arada I. Elucidating the role of genetic drift and natural selection in cork oak differentiation regarding drought tolerance. Molecular Ecology, **18:** 3803-3815, 2009.

Raubeson LA., Jansen RK. Method for obtaining and analyzing whole chloroplast genome sequences. Methods Enzymology, **395:** 348-384, 2005.

Rehfeldt GE, Ying CC, Spittlehouse DL, Hamilton DA. Genetic responses to climate in *Pinus contorta*: niche breadth, climate change and reforestation. Ecological Monographs **69**, 375-407, 1999.

Rehfeldt GE., Tchebakova NM., Parfenova YI., Wykoff WR., Kuzmina NA., Milyutin LI. Intraspecific responses to climate in *Pinus sylvestris*. Global Change Biology, **8(9):** 912-929, 2002.

Ribeiro MM., Mariette S., Vendramin GG., Szmidt AE., Plomion C., Kremer A. Comparison of genetic diversity estimates within and among populations of maritime pine using chloroplast simple-sequence repeat and amplified fragment length polymorphism data. Molecular Ecology, **11:** 869-877, 2002.

Roger & Barbee. Symetry in plants. World Scientific Publishing, 1998.

Rushton B.S. Natural hybridisation within the genus *Quercus* L. Ann. Sci. For. Suppl. 1 (Paris), **50:** 73-90, 1993.

Russell JR., Ellis RP., Thomas WTB, Waugh R., Provan J., Booth A., Fuller J., Lawrence P., Young G., Powell W. A retrospective analysis of spring barley germplasm development from "foundation genotypes" to currently successful cultivars. Molecular Breeding, **16:** 553-568, 2000.

Salmela MJ, Cavers S, Cottrell JE, Iason GR, Ennos RA. Seasonal patterns of photochemical capacity and spring phenology reveal genetic differentiation among native Scots pine (*Pinus sylvestris* L.) populations in Scotland. For Ecol and Menag, (doi: **10.1016**), 2011.

Sasky C., Lee S., Daniell H. Wood T., Tomkins J., Kim HG., Jansen RJ. Complete Chloroplast Genome of *Glycine max* and Comparative Analysis with other Legume Genomes. Plant Molecular Biology **59:** 309-322, 2005.

Savolainen & Pyhäjärvi. Genomic diversity in forest trees. Curr Opin Plant Biol, **10:** 1-6, 2007.

Savolainen O, Pyhäjärvi T, Knurr T. Gene Flow and local adaptation in trees. Annu Rev Ecol Evol Syst., **38:** 595-619, 2007.

Schirone B., Bellarosa R. Una proposta italiana per i descrittori della quercia da sughero (*Quercus suber* L.). Silvanova, CISDAM, De Rerum Natura 1 (Suppl): 5-13, 1996.

Schwarz O. *Quercus* L.: In : Flora Europaea (eds Tutin TG, Burges NA, Chater AO *et al.*), pp. 72-76. Cambridge University Press, Cambridge, UK, 1993.

Sebastiani F. Sequenziamento del genoma del cloroplasto del castagno, Castanea sativa *Mill.* PhD Thesis, University of Florence (Italy), 2003.

Sebastiani F., Carnevale S., Vendramin G.G. A new set of mono- and dinucleotide chloroplast microsatellites in *Fagaceae*. Molecular Ecology **4:** 259-261, 2004.

Shutyaev & Giertych. Height growth variation in a comprehensive Eurasian provenance experiment of *P. sylvestris* L. *Sil Genet* **46**, 332-49, 1997.

Skrøppa T., Nikkanen T., Ruotsalainen S. & Johnsen O. Effects of sexual reproduction at different latitudes on performance of the progeny of *Picea abies*. Silvae Genetica, **43:** 297-303, 1994.

Sneath, P.H.A. and Sokal R.R. Numerical Taxonomy. Freeman, San Francisco, 1973.

Soto A., Lorenzo Z., *et al.* Nuclear microsatellite markers for the identification of *Quercus ilex* L. and *Quercus suber* L. hybrids. Silvae Genetica **52:** 63-66, 2003.

Soto A., Lorenzo Z. and Gil L.. Differences in fine-scale genetic structure and dispersal in *Quercus ilex* L. and *Quercus suber* L.: consequences for regeneration of meditarranean open woods. Heredity **99:** 601-607, 2007.

Spellenberg R. On the hybrid nature of *Quercus basaseachicensis* (Fagaceae, sect. *Quercus*). Sida, **16:** 427-437, 1995.

Stebbins GL. Jr. "Variation and Evolution in Plants". Columbia University Press, New York, 1950.

Stebbins GL., Martze EB., Epling C. Types of Polyploids: Their classification and significance. Advantage Genetics, **1:** 403-429, 1947.

Steinkellner H., Fluch S., Turetschek E., Lexer C., Streiff R., Kremer A., Burg K., and Glössl J. Identification and characterization of (GA/CT)n – microsatellite loci from *Quercus petraea*. Plant Molecular Biology, **33:** 1096-1096 ,1997.

Sukhotu T., Kamijima O., Hosaka K. Chloroplast DNA variation in the most primitive cultivated diploid potato species *Solanum stenotomum* Juz. et Buk. and its putative wild ancestral species using high-resolution markers. Genetic Reources and Crop Evolution **53:** 53-63, 2006.

Taberlet P., Fumagalli L., Wust-Saucy A.G., Cosson J.F. Comparative phylogeography and postglacial colonization routes in Europe. Molecular Ecology **7:** 453-464, 1998.

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

Tang S., Knapp S.J. Microsatellites uncover extraordinary diversity in native American landraces and wild populations of cultivated sunflower. Theoretical and Applied Genetics, **160:** 990-1003, 2003.

Tarayre M. and Thompson J. D. Population genetic structure of the gynodioecious *Thymus vulgaris* L. (Labiatae) in southern France. Evolution Biol., **10**: 157-174, 1997.

Tautz D. Hypervariability of simple sequences as a general source of polymorphic DNA markers. Nucleic Acids Resources, **17:** 6463-6471, 1989.

Thompson J.D., Higgins DJ., Gibson TJ. CLUSTALW: improving the sensitivity of progressive multiple alignment trough sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Research, **22:** 4673-4680, 1994.

Thompson J.D. Population differentiation in Mediterranean plants: insights into colonization history and implications for species diversifications. Heredity **82:** 229-236, 1999.

Thompson J.D. Plant Evolution in the Mediterranean. Oxford University Press, Oxford, UK, 2005.

Thompson J.D., Lavergne S., Affre L., Gaudeul M., Debussche M. Ecological differentiation of Mediterranean endemic plants. Taxon **54(4)**: 967-976, 2005.

Tiffin P, Gaut BS. Sequence diversity in the tetraploid *Zea perennis* and the closely related diploid *Z. diploperennis*: insights from four nuclear loci. *Genetics* **158**, 401-412, 2001.

Torti X. Variabilité génétique de la réponse du métabolisme carboné du chêne pédonculé (*Quercus robur* L.) à des teneurs croissantes en CO2. PhD Thesis, Université Henri Poincaré, Nancy I, Franec. 130 p., 2005.

Toumi L., Lumaret R. Allozyme variation in cork oak (*Quercus suber* L.): the role of phylogeography and genetic introgression by other Mediterranean oaks species and human activity. TAG, **97(4):** 647-656, 1998.

Toumi L., Lumaret R. Allozyme characterisation of four evergreen Mediterranean oak species. Biochemical Systematics and Ecology, **29(8):** 799-817, 2001.

Trelease W. The Americans oaks. Mem. Natl. Acad. Sci., **20:** 1-255, 1924.

Tucker JM. Patterns of parallel evolution of leaf form in new world oaks. Taxon, **23(1):** 129-154, 1974.

Tucker JM. Studies in the Quercus undulate complex I. A preliminary statement. Am. J. Bot., **72:** 32-43, 1961.

Tutin TG., Burges NA., Chater AO. *et al.* Flora Europaea. Cambridge University Press, 1993.

Van Andel TH. The climate and landscape of the middle part of Weichselian Glaciation in Europe: The stage 3 project. *Quat Res* **57**, 2-8, 2002.

Van de Peer Y. and De Watcher R. TREECON for Windows: a software package for the construction and drawing of evolutionary trees for the Microsoft Windows environment. Computer Applications in the Bioscience, **10:** 569-570, 1994.

Varela MC. European network for the evaluation of genetic resources of cork oak for appropriate use in breeding and gene conservation strategies, Handbook FAIR (Varela eds.) MC 2003.

Velasco R., Zharkiku A., Affourtit J., Dhingra A., Cestaro A., Fontana P., *et al.* The genome of the domesticated apple (*Males x domestica* Borkh.). Nature genetics, **42(10):** 833-841, 2010.

Vendramin GG., Lelli L., Rossi P., Morgante M. A set of primers for the amplification of 20 chloroplast microsatellites in *Pinaceae*. Mol. Ecol., **5:** 595-598, 1996.

Vendramin GG., Ziegenhagen B. Characterization and inheritance of polymorphic plastid microsatellites in *Abies*. Genome, **40:** 857-864, 1997.

Vigouroux Y., Mitchell S., Matsuoka Y., Hamblin M., Kresovich S., Smith J.S.C., Jaqueth J., Smith O.S., Doebley J. An analysis of genetic diversity across the maize genome using microsatellites genetics. Genetics **169:** 1617-1630, 2005.

Wachowiak W, Lewandowski A, Prus-Głowacki W. Reciprocal controlled crosses between *Pinus sylvestris* and *P. mugo* verified by a species-specific *cp*DNA marker. J Appl Genetics **46** (1) 41 – 43, 2005.

Wachowiak W, Balk PA, Savolainen O. Search for nucleotide diversity patterns of local adaptation in dehydrins and other cold-related candidate genes in Scots pine (*Pinus sylvetsris* L.). Tree Gen & Genome **5**, 117-132, 2009.

Weir BS and Cockerham CC. Estimating F-statistics for the analysis of population structure. Evolution **38,** 1358-1370, 1984.

Weising K. and Gardner R.C. A set of conserved PCR primers for the analysis of simple sequence repeat polymorphisms in chloroplast genomes of dicotyledonous angiosperms. Genome **42:** 9-19, 1999.

Welling A, Palva ET. Molecular control of cold acclimation in trees. Physiol Plantarum **127:** 167-181, 2006.

Whittemore AT., Shaal BA. Interspecific gene flow in sympatric oaks. Proc. Natl. Acad. Sci. USA, **88:** 2540-2544, 1991.

Wolfe KH., Li WH., Sharp PM. Rates of nucleotide substitution vary greatly among plant mitochondrial, chloroplast and nuclear DNAs. Proc. Acad. Sci. USA, **84:** 9054-9058, 1987.

Wright S. The genetical structure of populations. Annal Eugenic **15:** 323-254, 1951.

Wright S., Lauga B, Charlesworth D. Subdivision and haplotype structure in natural populations of *Arabidopsis lyrata*. Mol Ecol **12**: 1247-1263, 2003.

Xu D.H., Abe J., Gai JY., Shimamoto Y. Diversity of Chloroplast DNA SSRs in wild and cultivated soybeans: evidence for multiple origins of cultivated soybean. Theoretical and Applied Genetics, **105:** 645-653, 2002.

Yamaguchi-Shinozaki K, Shinozaki K. Transcriptional regulatory networks in cellular responses and tolerance to dehydration and cold stresses. Ann Rev Plant Biol **57**: 781-803, 2006.

Zanetto A., Roussel G., Kremer A. Geographic variation of inter-specific differentiation between *Quercus robur* L. and *Quercus petraea* (Matt.). Liebl. Forest Genetics, **1:** 111-123, 1994.

Zhou Z. The fossil history of *Quercus*. Acta Bot. Yunnanica, **15:** 21-33, 1993.

Zhu JK, Xiong L, Shumaker KS. Cell signaling during cold, drought and salt stress. The Plant cell **14:** S165-S183, 2002.