Genetics of *Pinus pinaster* Aiton with Cytoplasmic and Nuclear Markers

Maria Margarida Ribeiro

Department of Forest Genetics and Plant Physiology

Umeå

Doctoral thesis

Swedish University of Agricultural Sciences

Umeå 2001

Acta Universitatis Agriculturae Sueciae

Silvestria 177

ISSN 1401-6230

ISBN 91-576-6061-1

Printed by: SLU, Grafiska Enheten, SLU, Umeå, Sweden, 2001

Abstract

Ribeiro, M. M., 2001. Genetics of *Pinus pinaster* Aiton with Cytoplasmic and Nuclear Markers. Doctoral dissertation.
ISSN: 1401-6230
ISBN: 91-576-6061-1

This thesis summarizes and discusses results of three studies in which biochemical and molecular techniques were used to study the genetic variation in *Pinus pinaster*. In particular, the investigation focused on: (i) the within- and among-population genetic diversity in the region hypothesised as a putative refugium for the species during the last glaciation; (ii) the comparison of nuclear and cytoplasmic estimates of diversity within and between two regions of the species; and (iiii) the design of a test for provenance identification using knowledge about the levels of genetic variation between the two regions.

The distribution of the genetic variation of *P. pinaster* in Portugal, as revealed by chloroplast microsatellites (cpSSR), indicated that there are low levels of differentiation among populations and that the diversity is found mainly within populations. No discernable geographic pattern was found. Evidences of strong anthropogenic influence associated with extensive gene flow could explain these findings. Fossil, charcoal and palynological records supported the presence of the species in Portugal before and during the last glaciation; therefore, the hypothesis of a putative refugium in this country cannot be excluded.

The genetic variation of 24 populations from France and Portugal was investigated with amplified fragments length polymorphisms (AFLPs) and cpSSRs. Both types of markers could discriminate between the two provenances and the diversity of the French provenance was higher compared with that from Portugal. Similar differentiation estimates were found with nuclear and cytoplasmic markers. Extensive gene flow could account for this result, but higher mutation rates and homoplasy at cpSSR loci are not to be excluded. Despite the different modes of inheritance, a high correlation was found between the genetic distances matrices with both types of markers, which suggests that migration surpassed genetic drift in moulding the genetic structure of this species in the regions studied.

A provenance diagnostic test was designed, based on cpSSRs, to screen the putative origin of stands of *P. pinaster* in southwestern France and compared with the currently used terpene-based test. Five stands of unknown origin were diagnosed with both tests. The cpSSR-based test proved to be faster and more accurate to determine if stands were of French or northwest Iberian (Portugal and Galicia) origin. The result obtained was probably due to the higher capability of the DNA-based markers to discriminate between both provenances, compared to that of the terpene markers.

Key words: cpSSR, AFLP, terpenes, genetic variation, provenance identification, *Pinus pinaster*.

Author's address: Maria Margarida Ribeiro, Escola Superior Agrária de Castelo Branco, 6000 Castelo Branco, Portugal. E-mail: Maria.Ribeiro@genfys.slu.se

In memory of my father
to my daughter Filipa
and my mother

SEXTO / *D. DINIS*

Na noite escreve um seu Cantar de Amigo
O plantador de naus a haver,
E ouve um silêncio múrmuro consigo:
É o rumor dos pinhais que, como um trigo
De Império, ondulam sem se poder ver.

Arroio, esse cantar, jovem e puro,
Busca o oceano por achar;
É a fala dos pinhais, marulho obscuro,
É o som presente desse mar futuro,
É a voz da terra ansiando pelo mar.

Fernando Pessoa
in *Menssagem*

Contents

Appendix

This thesis is based on studies reported in the following papers, which will be referred to in the text by the corresponding Roman numerals:

I. Ribeiro M.M., Plomion C, Petit R, Vendramin G.G. & Szmidt A.E. (2001) Variation of chloroplast simple-sequence repeats in Portuguese maritime pine (*Pinus pinaster* Ait.). *Theoretical and Applied Genetics*. 102(1):97-103.

II. Ribeiro M.M.[*], Mariette S.[*], Vendramin G.G., Szmidt A.E., Plomion C. & Kremer A. (2001) Comparison of maritime pine (*Pinus pinaster* Ait.) diversity estimates using SSRcp and AFLP data. (Manuscript)

III. Ribeiro M.M.[*], LeProvost G.[*], Gerber S., Vendramin G.G., Anzidei, M., Decroocq S., Marpeau A., Mariette S., & Plomion C. (2001) Origin identification of maritime pine stands in France using chloroplast simple-sequence repeats. *Annals of Forest Science*. (Accepted)

Paper **I**, figures 3 and 4 are reproduced with permission from the publishers.

[*] To be considered as joint first authors

Introduction

Pithys, a Greek nymph had two lovers, Boreas and Pan. Boreas became jealous and threw Pitys against a rock ledge. She turned instantly into a pine, and resin drops are her tears… "Greek Myth"

An insight into population genetics

Population genetics, which studies the consequences of genetic information transmission, requires information about changes in the frequencies of genes within and among populations, and aims at understanding the influence and interaction of different factors of evolution in shaping genetic variation. Among the candidate explanations are a balance between natural selection, mutation, gene flow and drift (Stearns and Hoekstra 2000).

Population refers to a group of organisms of the same species living within a restricted geographical area where random mating is potentially possible. A precise definition of the term is difficult and varies from species to species due to the usual non-random pattern in the spatial distribution of the individuals, i.e., due to *genetic geographical structure*. Population subdivision can be caused by environmental patchiness. Therefore, local interbreeding units are defined as *local populations*, *sub-populations* or *demes*, and they constitute the units of population genetics. The level of *genetic differentiation* existing among local populations measures the genetic structure (or population subdivision) of a species, i.e., the frequencies of the alleles may differ from one local population to the next (Hartl and Clark 1997).

In any population, the *genotype frequencies* are determined largely by the patterns in which genotypes of the previous generation come together to form mating pairs. In *random mating*, genotypes form mating pairs in the proportions expected by chance alone. The expectations of genotype frequencies under random mating are described by the *Hardy-Weinberg equilibrium* (HWE). This equilibrium can be disturbed by a number of factors such as selection, mutation, migration and genetic drift, then allele frequencies will change over time (see e.g. Weir (1996) for details).

Several processes create new types of genetic variation or allow the reorganization of the previously existing variation either within genomes or among populations. The ultimate source of genetic variation is *mutation*, that is any heritable change in the genetic material. It includes a change in the nucleotide sequence or a chromosome rearrangement, such as an inversion or translocation. This process provides, by itself, a very weak possibility for changing allele frequencies.

Mutation is a *rare event* (10^{-4} to 10^{-6} per gene and per generation for wild-type genes), but in a large population there are many genes at risk of mutation. *Recombination* brings mutations of different parts of the genome together or conversely, splits portions of the chromosome that were previously together. *Gene flow* is the introduction of genes from one population to another by mating or migration. Like mutation, this process introduces new genetic variants into local populations. Gene flow enables mutations to spread among populations and sets the limit of genetic divergence among populations; moreover, it can prevent local differentiation if selection is weak and the distance that genes move in each generation is large (Stearns and Hoekstra 2000). Gene flow levels vary greatly among species, populations and seasons, but the isolation by distance is frequently not enough to prevent its homogenizing effect against genetic drift and diversifying selection (Ellstrand 1992). Nevertheless, a small amount of gene flow is usually not sufficient to disperse rare alleles among populations; therefore, rare alleles are often unique to one or few populations (Hartl and Clark 1997).

The random fluctuation of allelic frequencies in finite populations due to non-representative sampling of genes from one generation to another is called *random genetic drift*. This phenomenon changes the distribution of genetic variation in two ways: (i) the decrease of variation within populations (loss of heterozigosity and eventual fixation of alleles), and (ii) the increase of differentiation among populations. The effect of this process increases with decreasing population size and depends on allele frequencies of the parental populations (Ellstrand and Elam 1993). Wright (1931) predicted that genetic drift would substantially alter the organization of genetic variation of populations when it is much greater then mutation rate and selection. Populations founded by few individuals do not contain a representative sample of genes in the parental population (*founder effect*). Some alleles that may be completely absent or rare in the parental population can reach high frequencies in the new population simply because they were present in the founders. Similarly, if the composition of the population changes dramatically as it passes through a *genetic bottleneck*, many alleles are lost and others rise to high frequency (reviewed by Barrett and Kohn 1991).

According to the *neutral theory* (Kimura 1968) many mutations have so little effect on the organism that their influence in survival and reproduction is negligible. The frequencies of neutral alleles are not influenced by natural selection, but by genetic drift. Therefore, they have no particular role in the adaptation to new environments, which makes them particularly suitable to trace the geographical genetic structure of populations among other purposes.

The major questions of population genetics are: how to measure the genetic variation in populations and how much it affects individual reproductive success. In the middle sixties, molecular methods started to be used and they increased the possibility to obtain estimates of genetic variation at previously "invisible" loci. The amount of genetic variation is measured by *genetic diversity* or "heterozygosity", defined as the probability that two alleles chosen at random at a

locus in the population are different. *Molecular methods* do not solve the problem of deducting the genetic variation from observed phenotypic variability: they circumvent the problem. They tell us how much genetic variation is present in a particular part of the genome, but they do not tell us how this genetic variation affects phenotypic variation (Stearns and Hoekstra 2000).

Pinus pinaster Ainton: origin, taxonomy and biology

Pines belong to an old genus – the oldest in the whole family of *Pinaceae* – which, originated in the middle Mesozoic (190-136 My BP, million years before present) in middle latitudes. During the Cretaceous (136-65 My BP), the genus was already differentiated into two subgenera, and pines were widely distributed throughout the Northern Hemisphere. In general, pine populations were fragmented and displaced during the Eocene due to major climate changes (54-34 My BP). In the terminal Eocene, temperature and humidity decreased drastically leading to angiosperm taxa extirpation and expansion of pines and other cool and dry adapted taxa in middle-latitude locations. Fossil records indicate that pines rose in abundance throughout middle latitudes in North America, Europe and Asia in the Miocene (26-7 My BP); the direct ancestors of many modern pines can be traced to Miocene pines (Millar 1998 and references therein). There are indications that Mediterranean pines migrated to the region from eastern Asia along the mountain ranges that once extended north of and parallel to the Himalayas (Mirov 1967).

Pinus contains more species than any other genus of conifers (Little and Critchfield 1969), and they all share uniformity in chromosome number (2n = 24) (Sax and Sax 1933). More than 40 different classification systems have been proposed for this genus (Millar 1993) depending on the criteria used (e.g., Shaw 1914; Mirov 1967; Little and Critchfield 1969; Farjon 1984). This genus is usually divided into two subgenera *Strobus* (= *Haploxylon*, soft pines) and *Pinus* (= *Diploxylon*, hard pines), which are further divided into sections and sub-sections (Little and Critchfield 1969; Farjon 1984). According to Little and Critchfield (1969), *Pinus pinaster* belongs to the *Pinus* subgenus, subsection *Sylvestris*.

The evolution of the genus *Pinus* has been studied recently using different approaches, in particular molecular markers (e.g., Strauss and Doerksen 1990; Wang and Szmidt 1993; Farjon 1996; Wang *et al*. 1999; Wang *et al*. 2000). According to a study based on sequence divergence of chloroplast regions made with Eurasian pines (Wang *et al*. 1999), the Mediterranean pines formed one strongly supported clade within the subgenus *Pinus*, and within that clade, *P. pinaster* grouped with *P. canariensis* and *P. pinea*, but with a weak bootstrap support.

Pinus pinaster reaches a height of about 35 m. The needles are thick, rigid and shiny grey-green, with conspicuous rows of stomata on all sides. They are grouped in pairs, 15-20 cm long, with a smooth surface, and a half-circular cross-section.

Male strobili develop on the lower part of the new shoots, whereas female strobili form in the whorl around terminal buds. The cones are 10-20 cm long and remain in the branches for several years. Young trees 6-7 years old can start to produce cones on upper shoots. The seeds are about 1 cm long, with a 2-3 cm long wing (Farjon 1984).

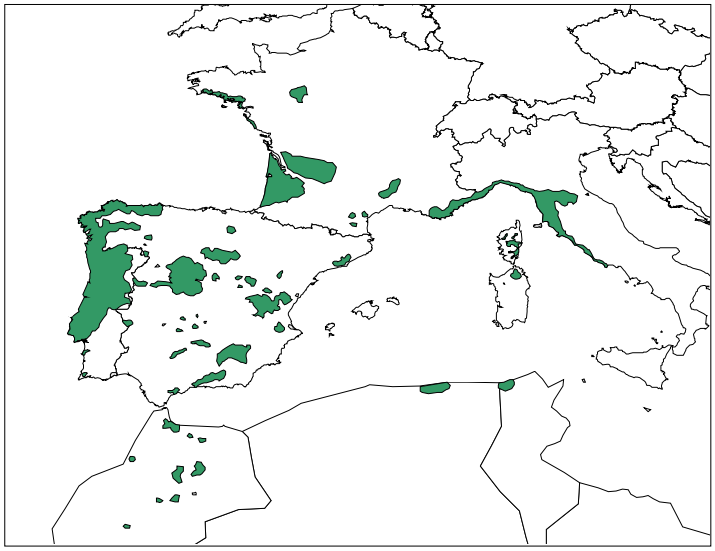


Figure 1. Natural distribution of *P. pinaster*. (After Baradat and Marpeau-Bezard 1988)

Distribution area

Some studies have shown that climatic changes during the Quaternary (2.4 My to present) in Europe have played a major role in shaping the phylogeography of European plant and tree species by contracting and expanding their natural ranges (Bennett *et al*. 1991; Hewitt 1996; Hewitt 2000). The successive glacial-interglacial oscillations of the Pleistocene (2.4-0.01 My BP) have moulded the range of plant species, by isolating refugia, which subsequently provided the source for new colonization (Comes and Kadereit 1998; Taberlet *et al*. 1998). By compiling several data sets available for various European tree species, Taberlet *et al*. (1998) were able to identify some general trends according to the location of their refugia in southern Europe, but they concluded that each species would exhibit a particular phylogeographic pattern. In the case of *P. pinaster*, it has been hypothesised that the actual distribution of this species is the result of events that occurred during the last glaciation (0.7-0.01 My BP) (Baradat and Marpeau-Bezard 1988). Nevertheless, the distribution of *P. pinaster* has also been greatly modified by human activities during historic times throughout the Mediterranean Basin until the recent expansion of its cultivated range (Barbero *et al*. 1998; Le Maitre 1998).

According to Scott (1962) and Farjon (1984) *Pinus pinaster* occurs naturally in southwest Europe and northwest Africa between latitude 31° and 46° N and longitude 9° and 13° E; from southwest Morocco to the mouth of the Gironde in France, and from the west coast of Portugal to the west coast of Italy. In France, Algeria, Tunisia and Italy the distribution is mainly coastal, but in Portugal, Spain, Morocco and Corsica this species grows from near the coast to far inland and high into the mountains This species is distributed throughout its range area in a discontinuous way, due to geographic isolation of populations and to the ancient human impact in the Mediterranean Basin (Fig. 1). The areas occupied by the species in different countries are shown in Fig. 2. *Pinus pinaster* has been planted in many other places in Europe and it has often become well established outside its natural range, especially in coastal regions. *Pinus pinaster* is also found in Australia, South Africa and New Zealand as an exotic species, where it constitutes extensive and successful stands.

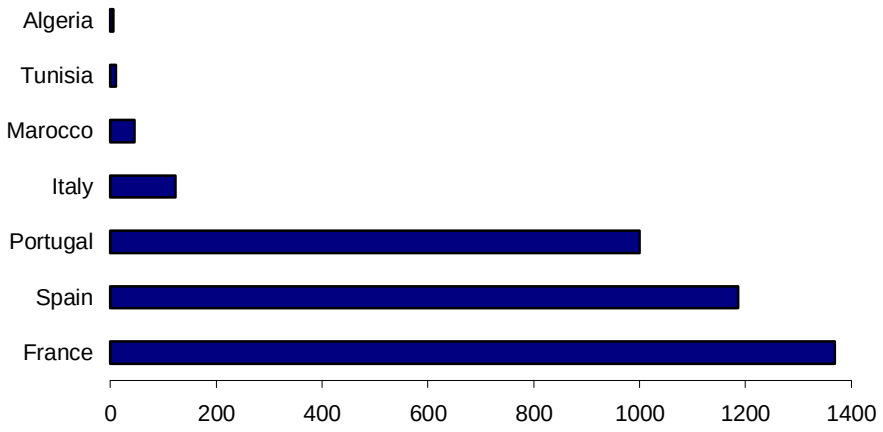


Figure 2. Areas occupied with *P. pinaster* in the countries of its natural range (x 1000 ha). (http://pinus.dgf.min-agricultura.pt/estatistica/invent.htm; Alía *et al*. 1996; Deroy 2000)

Migration pathways of *Pinus pinaster* in Europe

Tree species show different patterns in the distribution of genetic diversity within and among populations (Hamrick *et al*. 1981). As discussed earlier, several factors are responsible for moulding the genetic variation patterns: gene flow, selection, genetic drift, human activities and climatic changes. However, the relative significance of these factors is likely to vary greatly among populations and species (Stearns and Hoekstra 2000).

The genetic variation of *P. pinaster* has been studied using various ways. Intraspecific variation in the species has been investigated in numerous provenance

trials established in different countries (Harfouche and Kremer 2000 and references therein), and these experiments showed that morphological and adaptive traits vary significantly among provenances. Several range-wide diversity studies have focused on terpenes, isozymes, denatured proteins and chloroplast microsatellites (Baradat and Marpeau-Bezard 1988; Bahrman *et al*. 1994; Petit *et al*. 1995; Vendramin *et al*. 1998). Some other studies have been undertaken at a regional level using isozymes, AFLPs (amplified fragment length polymorphisms) and nuclear microsatellite markers (Castro 1989; Salvador *et al*. 2000; González-Martínez *et al*. 2001; Mariette *et al*. 2001b).

Based on terpene markers, palynological and paleoclimatological records, Baradat and Marpeau-Bezard (1988) discriminated three major groups of *P. pinaster* populations: the Atlantic group, comprising populations from southwestern France, Portugal, and Galicia in Spain; the Mediterranean group, extending from central Spain to the Ligurian coast in Italy; and finally the North African group that includes stands from Morocco, Algeria and Tunisia. In another study, Bahrman *et al*. (1994) included eastern Spain in the Atlantic group. Using mitochondrial DNA markers, C. Burban (unpublished manuscript) discriminated three groups: the Moroccan (Rif and Atlas), the Occidental (Iberian Peninsula and southwestern France) and the Oriental group (southeastern France, Corsica, Sardine, Italy, Pantelleria, Tunisia and Algeria).

The typically scattered distribution of this species may have prevented or limited gene flow among the different groups of populations, causing high genetic divergence among populations due to genetic drift (Baradat and Marpeau-Bezard 1988; Bahrman *et al*. 1994; Petit *et al*. 1995; Vendramin *et al*. 1998). Nevertheless, gene flow and human activity were probably responsible for the low differentiation found at a fine geographic scale (Castro 1989; Mariette *et al*. 2001b).

The presence of a centre of origin of *P. pinaster* in the southwest of the Iberian Peninsula at the end of the Pliocene (3 My BP) was hypothesised by Baradat and Marpeau-Bezard (1988) and supported by fossil findings (Teixeira 1945). The authors drew a different picture for the migration pathways of the species before and after the last glaciation. The preglacial hypothesis supposes the presence of three distinct pathways towards the north of Portugal, Spain and France, towards the south of Spain, France and Italy, and towards the north of Africa (Fig. 3). The postglacial hypothesis assumes that migration occurred only along the first two pathways. Moreover, the authors claim that the successive ice ages during the Pleistocene had several times stopped the northerly advance of *P. pinaster*, and eventually diminished its presence to scattered refugia in the south of the Iberian Peninsula, particularly during the Pleniglacial (0.35 My BP).

According to the study made by Vendramin *et al*. (1998) based on chloroplast microsatellites, a more complex picture of the possible migration pathways was drawn. Two main reservoirs of haplotypic diversity (Landes and Pantelleria) were identified (Fig. 3). The Pantelleria population (or North of Africa populations)

might have represented a starting point of the migration process. Pantelleria could represent an ancient population originating in the preglacial period from refugia located in the central part of North Africa, from which the migration took place towards West and East. The French area might have represented refugia from which the migration towards Italy began. Conversely, low levels of within-population diversity were observed in the Portuguese populations (Alcácer and Monção) with characteristics of recent establishment (possible founder effects). Such a reduction in genetic diversity with increasing distance from a refugium is a general phenomenon to be expected from repeated population bottlenecks at the advancing edge of a range in any species during postglacial expansion (Hewitt 1996).

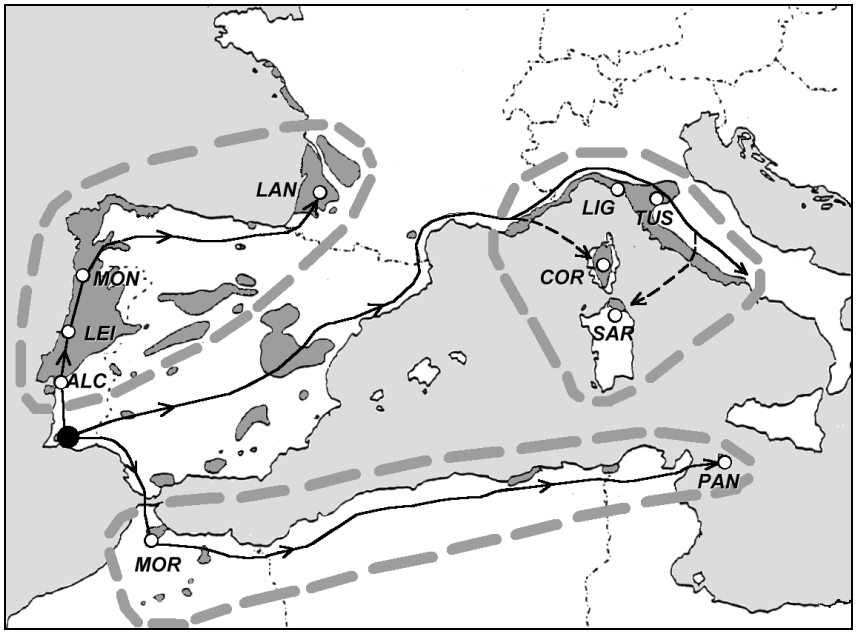


Figure 3. The three major groups of populations and preglacial migration pathways hypothesis considered for *P. pinaster*. ALC=Alcácer, LEI=Leiria, MON=Monção, LAN=Landes, PAN=Panteleria, LIG=Liguria, TUS=Tuscania, COR=Corsica, SAR=Sardinia, MOR=Morocco. (Courtesy of Vendramin *et al*. 1998)

The Iberian Peninsula has been reported as glacial refugium for numerous plant species, (Hewitt 1996; Comes and Kadereit 1998)) and it is one of the most important native areas of *P. pinaster*. Salvador *et al*. (2000) observed a generally higher level of allozyme diversity in the Spanish populations than the Portuguese population included in their study, and hypothesised a migration pathway from Spain to Portugal. The authors used six other populations from Portugal from another study (Castro 1989) and recomputed the data by using the same loci. Afterwards, Salvador *et al*. (2000) concluded that the idea of refugia in Portugal

could not be supported because no special variants were detected. In conclusion, the authors suggested the disappearance of *P. pinaster* from that area during the last glaciation.

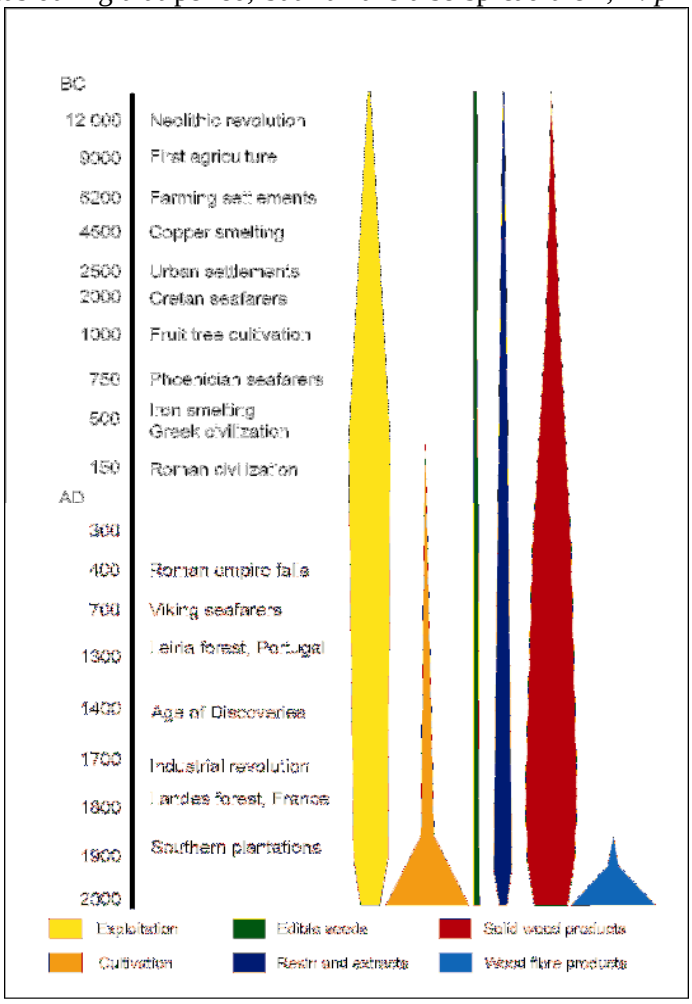
The phylogenetic analysis made with allozyme markers by González-Martínez *et al*. (2001) showed a high geographical structure in the Iberian Peninsula. The northwestern populations form a cluster and the southeastern populations another, in accordance with the study made by Baradat and Marpeau-Bezard (1988). The authors observed the highest levels of diversity in the eastern and the southern populations and an important reduction of gene diversity in the northwestern range of the species in the Peninsula. Nevertheless, a putative refugium in Portugal was not excluded, because *P. pinaster* could have survived during the last glaciation in sheltered areas at low altitude close to the Atlantic Ocean in Portugal, as suggested by Figueiral (1995) based on charcoal records.

Human impact in the Mediterranean region forests

The Mediterranean Basin is characterized by traditional human impacts on the forest (Thirgood 1981) and the modification of genetic diversity of species by human activity (Ledig 1992). The land use by humans over millennia has played a dramatic role in shaping the region's vegetation. In no other part of the natural range of pines has there been such a complex interplay between this genus and humans. The effects, however, have not been the same throughout the entire Mediterranean range, being more important in coastal areas and fertile soils. Various ecological factors, such as forest fires, drought, soils, etc., also played a major role in the adaptation of *P. pinaster*, in the isolation between stands, and, to some extent, in the genetic variation of the species (Barbero *et al*. 1998).

People and pines have had a long association. The first evidence of hominid habitation within the natural range of *Pinus* has been dated to about 1.4 My BP on the Eastern Shore of the Mediterranean Sea (Wood and Turner 1995). According to Le Maitre (1998) and references therein, by the time hominids encountered pines, the different species occupied most of their current range. Mediterranean-type climates, with their dry summers and recurrent fires, were also well established. Key events between humans and pines took place: the marked climatic fluctuations during the Pleistocene and Holocene (10 000 BP until present); the increasing use of fire that is essential for maintaining pine populations and reducing competition from hardwood species; the population growth that followed the domestication of agricultural crops; the copper and iron smelting that requires large amounts of wood; and the forest clearing necessary to produce open fields for agriculture and pasture (Fig. 4). Although population densities were low, human impacts during the prehistoric period were substantial and could have altered the abundance of pines, particularly in the Mediterranean Basin (Barbero *et al*. 1998).

The use of timber for the sea-borne trade and military protection further increased during the Graeco-Roman period. Also, the rapid population growth, particularly around cities, caused higher wood demands for housing, domestic use, bridges and roadways (Fig. 4). The emergence of agriculture and biological sciences with higher control over plant cultivation accelerated forest depletion, but regulation of forest cutting was common both for wood conservation and religious purposes (Thirgood 1981; Le Maitre 1998). The earliest evidence of humans altering the distributions of pine species through cultivation dates from the Graeco-Roman period. Palynological studies suggest that pine populations expanded locally in degraded areas during that period, but humans also spread them, *P. pinaster* being



a paradigmatic case (Willis 1992; Klaus 1989).

Figure 4. A diagrammatic representation of events in the history of cultivation of pines and the relative importance of the products or categories used by Man. (Adapted from Le Maitre 1998)

Growing concerns about forest depletion led to the first recorded large-scale reforestation. Natural forests in Portugal had suffered intense use before the 14th century due to fire, agriculture, grazing, mineral exploitation, wood consumption and shipbuilding; therefore, in 1310, King Don Dinis issued a law to protect pine forests (Scott 1962; Buting and Rego 1988). In the 13th century Cistercian monks established the "Leiria's pineyard" to stabilise the sand dunes, but many others were established elsewhere for soil protection and to fight desertification of the country (Mattoso and Sousa 1993). From the 15th century onwards, the naval and merchant fleets of Portugal, Spain, France and Great Britain also consumed large quantities of wood (Thirgood 1981). According to Le Maitre (1998), global timber trade increased significantly in the 18th century and sawmill industry expanded, encouraging forest harvesting. Forest nurseries were established for large-scale planting in Portugal and by the end of the 18th century and the beginning of the 19th century, *P. pinaster* was clearly expanding its presence and supplanting other species or invading non-cultivated areas (N. Devy-Vareta, personal communication). In France, attempts at reforestation were made as early as 1500, and *P. pinaster* was used to stabilise the dunes near Bordeaux in 1713. Reforestation of the sandy coastal plains of Landes in southwestern France continued until a significant area of the region was covered (Scott 1962).

In Europe, the two World Wars were critical and devastating periods for the forests, but the development of petroleum-based synthetics has, to a great extent, replaced wood-based products. Nevertheless, paper consumption increased notably with the growing information demands of the 20th century. In the Mediterranean Basin, the primary concern has been the protection of soil and reforestation of degraded areas (Thirgood 1981). For example, in Portugal, about 400 000 ha of *P. pinaster* had been planted between 1900 and the 1960s (Devy-Vareta 1993).

A particular feature of *P. pinaster* is its ability to grow well on very poor sandy soils, low in nutrients, while tolerating summer drought, winter flooding, and sea winds. Under favourable climate, such soils give a good yield of wood for lumber or pulp, and also resin (Scott 1962; Farjon 1984). Those characteristics played a central role in its preferential use for reforestation. Therefore, since at least the 20th century, reforestation programmes in the *P. pinaster* range area have notably spread this species. The present situation, therefore, is a complex of indigenous populations and planted forests of unknown origin (Devy-Vareta 1993; Alía *et al*. 1996; Barbero *et al*. 1998; Salvador *et al*. 2000).

Seed certification

Human settlement and action over the centuries have influenced forests in Europe, particularly in the Mediterranean area as discussed in the preceding paragraphs, and planted forests constitute a major part of the resource of some countries. The

concern about the deterioration of forests throughout Europe led to an increasing awareness of their economic, ecological, social and cultural value.

Human impact on forests alters the genetic structure of tree species in several ways; an important example concerns the introduction of seeds from other regions. Because introduced seeds produce, in general, less adapted stands than native ones, introductions may provoke economical losses and further affect the productivity of autochthonous stands. The *origin* of an *indigenous stand* is defined as the place in which the trees are growing and the origin of a *non-indigenous stand* is the place from which the reproductive material was originally introduced. *Reproductive material* may comprise fruits, seeds, pollen, scions or tissues for tissue culture. Planting non-indigenous reproductive material may alter local patterns of variation by influencing adjacent indigenous stands due to pollen and seed dissemination. The consequences are expected to be more important where intensive plantings were made by using reproductive material from regions with very different environmental characteristics from those at the site of material introduction (Jones and Burley 1973).

According to Zobel and Taberlet (1984), seed certification has been a concern that started as early as the beginning of the 20th century in Japan with *Cryptomeria japonica*. There are many meanings for seed certification, and different methods for the collection and handling of seeds from forest trees. Some authors regard certification as the correct *labelling* in which the seed size, purity, germination and other information about the seed is given, others consider that the information about where the seed was obtained, *source certification*, should also be included (Barber *et al*. 1962; Zobel and Taberlet 1984). According to Jones and Burley (1973), *seed certification* is an official statement that a seed lot conforms to certain standards, which may include specific identity, origin, genetic characters and seed purity. For forestry, seed certification systems have been developed largely to provide labels and records that give officially authenticated details of *identification* of the seeds, which involves the step of identifying its origin, and also its *quality*, meaning the genetic superiority, when that information is available. Therefore, reproductive material identification and certification in forestry has become a relevant issue (Barber *et al*. 1962; Matthews 1964; Jones and Burley 1973).

Morphological data and biochemical markers (terpenes, isozymes and denatured proteins) have all been used for provenance identification and seed certification in forest trees (e.g., Falkenhagen 1985, Boisseaux 1986, Bahrman *et al*. 1994, Espinel *et al*. 1995). Molecular markers based on nuclear and organelle DNA analysis have also been used lately for those purposes (e.g., Szmidt *et al*. 1988, Aragones *et al*. 1997, Sinclair *et al*. 1998, Bucci and Vendramin 2000). For example, in France, a law under the supervision of the Ministère d'Agriculture et Forêts regulates the collection of commercial *P. pinaster* seed-lots in the Aquitaine region (Réglement technique récolte pin maritime, Arrête du 8 février 1990). Candidate stands for seed collection in the Aquitaine region must have their origin certified using a diagnostic test developed by Baradat and Marpeau-Bezard (1988),

based on a discriminant analysis of terpene profiles. The term *provenance* has different meanings, reviewed by Jones and Burley (1973); in this study, it refers to the original geographical source of a given lot of plant reproductive material.

Recently, consensus on reproductive material certification in forestry is under discussion in Europe, allowing regional differences and specificities. The Third Pan-European Ministerial Conference on the Protection of Forests in Europe, held in Lisbon, 2-4 June 1998, adopted criteria for sustainable forest management, one of them being the maintenance, conservation and appropriate enhancement of biological diversity in forest ecosystems. This resolution involves the establishment of standards for forest reproductive material certification. For reforestation, indigenous species and local provenances that are well adapted to site conditions should be preferred. Introduced species or provenances should be used only after the evaluation of the impacts on the ecosystem and on the genetic integrity of indigenous species or provenances. For further details the following web page should be consulted: http://www.pefc.org/content.htm.

Objectives

The purpose of this thesis was to study the genetic structure of *P. pinaster* at a regional level. In particular, the investigation focused on a practical application to be used in the forest management of this species.

The main objectives of the three studies included in this thesis were: (1) to assess the distribution of genetic diversity within and among populations of *P. pinaster* in the region hypothesised as a putative refugium for the species during the last glaciation; (2) to compare nuclear and cytoplasmic estimates of diversity within and between two regions of the species; and (3) to design a test in order to identify the origin of the stands in one region.

In studies reported in Papers **I** and **II**, cpSSR (chloroplast microsatellites) and AFLP (amplified restriction fragment polymorphism) markers were used to provide information on the level and distribution of genetic variation among and within populations of *P. pinaster* at the regional level. Two different approaches were used: the first involved analysis of chloroplast repetitive simple-sequence repeats (cpSSR) (Paper **I**) and the second involved a comparative analysis between the cpSSR and AFLP markers (Paper **II**). In the analysis described in Paper **III**, a cpSSR-based test was developed in order to determine the putative origin of maritime pine stands in the Aquitaine region (southwest of France) and to compare the cpSSR-based test with the test based on terpene profile analysis.

Methods

Knowledge about the genetic structure of a species can be obtained from polymorphic markers that allow the determination of gene and/or genotypic frequencies. Genetic variation of forest trees can also be inferred using the traditional quantitative analysis of morphological traits, but due to the environmental influence, the polygenic character of some traits, and the time and cost to retrieve the information, other methods have been sought to obtain the same type of information (Wang and Szmidt 2000). The differences among individuals can be traced using secondary compounds, such as terpenes and flavonoids, but they also fail to be good candidates due to difficulties of inferring the genotype from phenotypes and to a possible environmental influence (Crawford 1983; Hanover 1992). In the last decades, methods that look at protein polymorphisms (Strauss *et al*. 1992) and the direct analysis of polymorphisms at the DNA level have provided a diverse array of molecular tools for genetic analysis in forest tree populations (reviewed by Morgante *et al*. 1996; Wang and Szmidt 2000).

A *molecular marker* can be defined as a sequence of DNA or a protein which can be readily detected and whose inheritance can be monitored. It is the *polymorphism* of molecular markers that can be used to study genetic diversity. Polymorphism in proteins has been studied, e.g., through *allozymes*, i.e., different molecular forms of an enzyme coded by different *alleles* at one gene locus. Polymorphism can be identified in different types of DNA: nuclear and cytoplasmic or organellar DNA (in the chloroplast, cpDNA and in the mitochondria, mtDNA) (Mitton 1994; Vekemans and Jacquemart 1997; Parker *et al*. 1998).

The desirable properties of a marker are: polymorphic expression; *codominant* inheritance (the different forms of a marker should be detectable in diploid organisms to allow discrimination of homozygotes and heterozygotes); even distribution throughout the genome; easy, fast and inexpensive detection; and reproducibility within and between laboratories. No single molecular marker meets all those criteria; the choice of a particular molecular marker will therefore depend on the objectives of the study (reviewed by Karp and Edwards 1997; Parker *et al*. 1998; Szmidt and Wang 2000).

DNA-based markers have a great advantage over terpenes or isozymes, for they provide pure genetic information, since they are not the products of transcription or translation. Other advantages of these markers are the large variety of scales on which evolutionary processes can be studied and their great potential for detecting variation in all kinds of organisms, in both living and dead tissues (Parker *et al*. 1998).

Markers mostly employed in the current studies are described in the following paragraphs.

Terpenes

Monoterpenes (C_{10} hydrocarbons) and sesquiterpenes (C_{15} hydrocarbons) are two classes of isoprenoid derivatives, which are elaborated from pyrophosphorylated precursors: respectively, gerandyl (C_{10}) and farnesyl (C_{15}) diphosphates. These compounds are found in conifers, in particular in the genus *Pinus*, where they accumulate in the resin ducts of different tissues (needles, primary cortex, and conducting tissues). Terpenes constitute 20% or more of the volatile fraction of oleoresin (Baradat *et al*. 1995). Terpenoid biosynthesis and metabolism were further reviewed by Chappell (1995).

In the genus *Pinus*, single-gene inheritance of monoterpenes has been demonstrated in several species (Baradat *et al*. 1995 and references therein), including *P. pinaster* (Baradat *et al*. 1972). Marpeau *et al*. (1975) also demonstrated single-gene inheritance of sesquiterpenes in *P. pinaster*. The use of terpenes as markers may present difficulties, as their expression can be affected by the age of the tree and the sample source (position of the sample in the tree and type of tissue used) (Bernard-Dagan *et al*. 1971). In *P. pinaster* the cortex of young but completely lignified shoots, the oleoresin composition remains unchanged (Baradat *et al*. 1972), therefore such plant material is appropriate for extraction of terpenic compounds for analysis. Moreover, the terpenic composition in the cortex of that plant material remains stable as soon as the tree achieves 7-10 years of age (Baradat *et al*. 1991).

Terpene markers have been used to study geographical genetic differentiation in conifers, as reviewed by Müller-Starck *et al*. (1992) and Strauss *et al*. (1992), but they can also be used to study mating patterns, to assess the geographic origin of stands and to study seed orchard pollen contamination (Baradat *et al*. 1991; Coppen *et al*. 1993; Schiller and Genizi 1993; Baradat *et al*. 1995). Nevertheless, terpeniod composition, while strongly inherited, probably plays a role in tree's disease and insect resistance, and is thus undoubtedly subject to natural selection (Hanover 1992). Some terpenes, in particular, are believed to be involved in the resistance to the caterpillar *Dioryctria spendidela* that attacks the bark of *P. pinaster* (Baradat and Marpeau-Bezard 1988). Besides, for terpenoids, only genetic changes that substantially alter gene expression or biosyntetic enzyme activity will be detected; thus, products of many different mutational events will be confounded or undetected (Strauss *et al*. 1992).

PCR-based markers

Dominant markers

In past years, a new generation of markers, based on the *polymerase chain reaction* (PCR), has been developed. PCR is a technique to amplify specific DNA sequences by primer extension of complementary strands of DNA with the action of the thermostable DNA polymerase (Mullis and Faloona 1987; Saiki *et al*. 1988). This technique is very powerful for amplifying tiny amounts of DNA sequences several million times over only in a few hours, involving several cycles of heating and cooling, each time the newly synthesised strand becoming a template for the subsequent replication. Theoretically, the cycling of temperatures increases in an exponential way the amount of the specified sequence (Fig. 5).

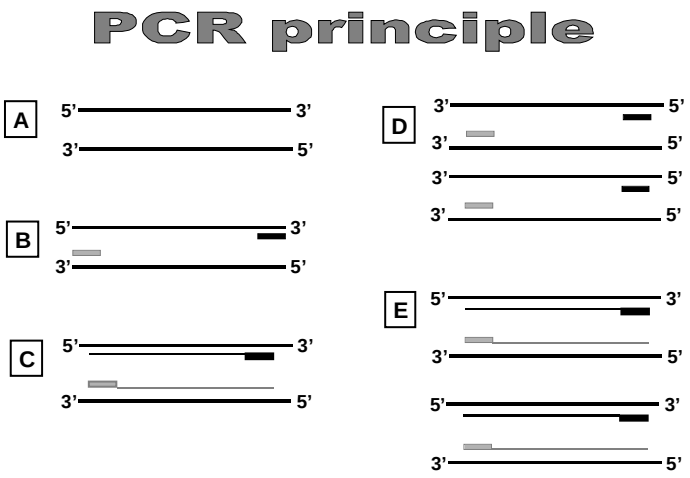


Figure 5. Diagrammatic representation of the PCR principle. A: In the denaturing step, the heat opens the double strand. B: In the annealing step, the cooling allows the primers (in grey and black) to bind the complementary regions. C: In the elongation step, the DNA polymerase synthesises complementary strands. D and E: The cycles of heating and cooling are repeated, each time the newly synthesised strand becoming a template for the subsequent replication.

In 1990, the use of short primers (usually 10 base-long) of arbitrary sequence was initiated to generate PCR amplification products, in low stringency conditions, from genomic DNA (Williams *et al*. 1990; Welsh and McClelland 1990). Depending on the specific conditions of the amplification or product separation and detection, different methods were termed: RAPD (*Random Amplified Polymorphic DNA*, Williams *et al*. 1990); AP-PCR (*Arbitrary Primer PCR*, Welsh and McClelland 1990); or, DAF (*DNA Amplification Fingerprinting*, Caetano-Annolés *et al*. 1991). RAPD analysis is performed at a low annealing temperature

(stringency), implying that the binding of the primer to the genomic DNA is partly non-specific. Therefore, in order to obtain reproducible results, the reaction conditions must be kept strictly constant and the analysis must be made in the same laboratory (Penner *et al*. 1993; Hallden *et al*. 1996; Jones *et al*. 1997; Rafalski 1997). Nevertheless, the method's speed, sensitivity and versatility make it suitable to survey large number of samples in population genetics of forest trees (e.g., Bucci and Menozzi 1995; Nesbitt *et al*. 1995; Schierenbeck *et al*. 1997; Gallois *et al*. 1998; Wu *et al*. 1999). RAPD markers usually show dominant Mendelian inheritance; amplification either occurs at a locus or not, leading to scores of band presence/absence. This means that the heterozygotes and homozygotes for the presence of the band cannot be distinguished (Isabel *et al*. 1995; Lu *et al*. 1995). Moreover, the amplified regions may represent both coding and non-coding sequences (Kazan *et al*. 1993; Lu *et al*. 1997).

Amplified Restriction Fragment Polymorphism (AFLP) is a powerful method for detecting polymorphism throughout the genome, based on a two-step amplification strategy that combines restriction enzymes and PCR (Zabeau and Vos 1993). This highly reproducible technique allows the simultaneous screening of a large number of molecular markers, randomly distributed throughout the genome (Vos *et al*. 1995; Zhu *et al*. 1998).

The genomic DNA is digested with two restriction enzymes, a frequent cutter and a rare cutter (in Fig. 6 *Mse*I and *Eco*RI). Afterwards, two adapters with 3' ends complementary to the sequences recognised by the enzymes are ligated to the DNA fragments, to provide known sequences for the first PCR amplification. In the pre-amplification step, primers with the 5' ends complementary to the adapters, but extended with one to several nucleotides at the 3' ends (two nucleotides in Fig. 6), are used to amplify a subset of the restricted fragments (PCR I). In the selective PCR amplification step, primers with the 5' ends, complementary to the adapters but extended with one to several nucleotides at the 3' ends (referred as selective nucleotides), are used to amplify the selected subset of restricted fragments (PCR II). Polymorphisms are detected by differences in the length of the amplified fragments by polyacrylamide electrophoresis. The amplified fragments are resolved in a sequencing gel and visualised by radioactivity, fluorescence or silver staining.

The AFLP method generates a large number of bands in a single reaction (Vos *et al*. 1995; Powell *et al*. 1996) and gives higher reproducibility compared with the RAPD method (Jones *et al*. 1997). The AFLP technique has been used for a wide range of purposes in tree species, among others, the investigation of genetic diversity (e.g., Winfield *et al*. 1998; Lerceteau and Szmidt 1999; Cervera *et al*. 2000; Mariette *et al*. 2001b), the generation of linkage maps, and the identification of molecular markers linked to phenotypic traits and/or genetic loci (Cervera *et al*. 1996; Marques *et al*. 1998; Travis *et al*. 1998; Cato *et al*. 1999; Marques *et al*. 1999; Remington *et al*. 1999; Arcade *et al*. 2000; Costa *et al*. 2000; Lerceteau *et al*. 2000; Sewell *et al*. 2000). Similarly to RAPD markers, AFLP markers show

predominantly dominant Mendelian inheritance (Paglia and Morgante 1998; Lerceteau and Szmidt 1999; Nikaido *et al*. 1999) and they detect variation in anonymous nuclear sequences.

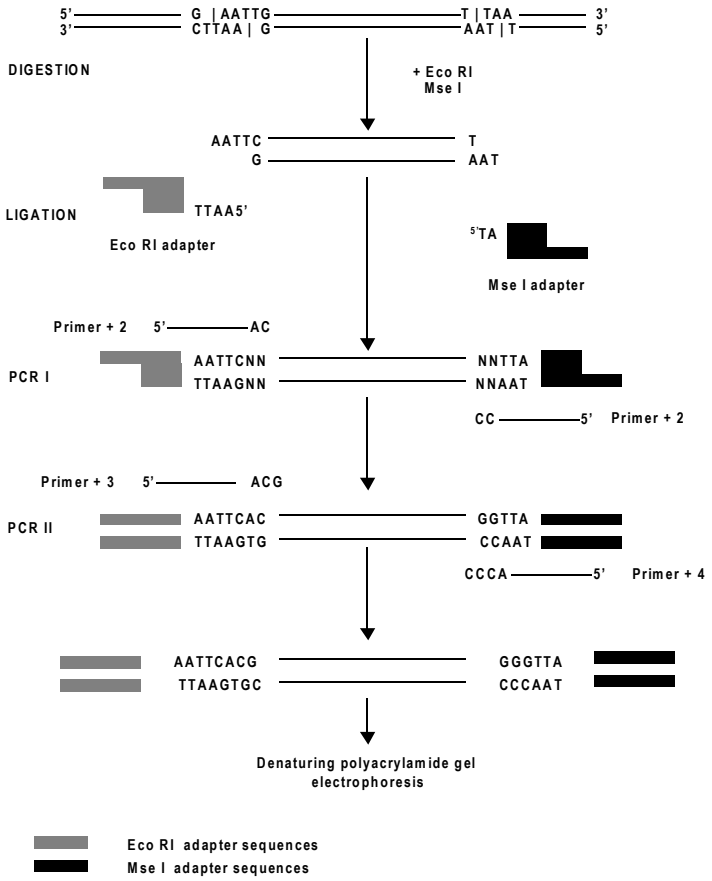


Figure 6. A diagrammatic representation of the AFLP procedure reported in this thesis. See text for explanation.

Only recently has this method been used in conifers because of the large size of their genome, which was responsible for the complicated band pattern, therefore the protocol had to be adapted in order to decrease the number of bands obtained per primer combination. The number of generated fragments can be restricted by changing the nucleotide extensions and/or the type of enzymes used in the digestion of the DNA (methylation sensitive or insensitive) (e.g., Paglia and Morgante 1998; Lerceteau and Szmidt 1999; Cervera *et al*. 2000; Costa *et al*. 2000). AFLPs are more technically demanding than RAPDs, but their automation and the availability of kits (e.g., Lerceteau and Szmidt 1999) made possible their use on a larger scale.

PCR-based methods such as RAPD and AFLP markers are more easily obtained than most non-PCR alternatives and their analysis does not require sequence information or laborious cloning. However, since most of these markers are dominant, the genotypes are not unambiguously traced, therefore biases are introduced in the estimation of population-genetic parameters (Lynch and Milligan 1994; Isabel *et al*. 1995; Szmidt *et al*. 1996; Isabel *et al*. 1999; Krutovskii *et al*. 1999; Zhivotovsky 1999). According to Lynch and Milligan (1994), dominant markers can be used to estimate unbiased population genetics parameters, provided that the loci with low frequency of the null allele are pruned from the analysis. Also Zhivotovsky (1999) developed a new approach to overcome the dominance situation for these markers. Isabel *et al*. (1999) refers that the reliability of RAPDs fingerprints in estimating population structure can be improved if prior knowledge exists of the matting system and levels of populations structuring, and if the fragments meet some polymorphism criterion (Lynch and Milligan 1994).

Microsatellite markers

A new type of marker, know as *simple-sequence repeat* (SSR) or *microsatellite* has been developed based on DNA sequence variation. This marker is based on tandem DNA repeats characterised by short motifs (1 to 6 bp), repeated from two to many thousands of times (Tautz 1989). A different allele occurs at a SSR locus as a result of changes in the number of times a core element is repeated, altering the length of the repeated region. Differences in length at a SSR locus are detected with DNA amplification by PCR using two oligonucleotide primers that complement unique sequence flanking at the SSR locus. Polymorphism is detected by electrophoretic separation of fragment sizes that can differ by as few as two base pairs.

Current research suggests that the length variation between alleles at a SSR locus are created by slippage of DNA polymerase during the replication of tandem repeats followed by failure of DNA mismatch repair to restore the original sequences (Strand *et al*. 1993). Microsatellites are very useful because they are codominant and highly-polymorphic markers, but their identification is a very expensive and time-consuming process, which generally requires the construction and screening of a genomic library. Known primers are not likely to amplify the same locus across related taxa, unless the flanking regions where priming sites are located are highly conserved (Ellegren 1992), which happens, usually in closely related species (Kijas *et al*. 1995). Therefore, the success of cross-amplification diminishes with increasing species divergence (Whitton *et al*. 1997).

Single-locus SSR markers have been identified in conifers (Lefort *et al*. 1999 and references therein), but due to their large genome size, this task has proven to be difficult. Moreover, only a small fraction of SSR clones selected from genomic libraries can be converted into informative SSR markers (e.g., Echt and Maymarquardt 1997; Pfeiffer *et al*. 1997; Mariette *et al*. 2001a). One strategy to

increase the efficiency of the identification of microsatellite regions is to transfer SSR markers across genera. Nevertheless, the SSR information generally does not transfer across *Pinus* species (Echt and Maymarquardt 1997; Echt *et al*. 1999), e.g., 47 SSRs primer pairs developed in three *Pinus* species were tested in *P. pinaster*, but only one amplified at a single polymorphic locus (Mariette *et al*. 2001a). The SSR primer pair that cross-amplified was identified in *P. halepensis*, which is from the same subgenus *Pinus* and belongs to a strongly supported clade with *P. pinaster*, according to (Wang *et al*. 1999).

The availability of the entire chloroplast sequence of *P. thunbergii* (Wakasugi *et al*. 1994) allowed the identification of *chloroplast simple-sequence repeats* (cpSSR). Primers flanking a mononucleotide repeat located in the intergenic region between the *trnK* and *pbsA* genes were used to detect variation in different pine species (Powell *et al*. 1995b; Vendramin *et al*. 1996).

The evolution rate of cpDNA genes is estimated to be several times slower than of the nuclear genes (Wolfe *et al*. 1987), with a low average level of sequence variation (Clegg and Zurawski 1992). The genetic information contained in the chloroplast genomes of plants, including the arrangements of genes and intergenic sequences, is very conservative compared with the nuclear and mitochondrial genomes, which contain vast amounts of DNA of no apparent function (Birky 1988). Those properties confer to the cpSSR markers primer binding sites a high degree of conservation (Powell *et al*. 1995a), therefore the designed primers ought to work across taxa.

The high degree of conservation of sequences in the chloroplast genome of conifers and the universality of the primers was confirmed by several recent studies in conifers (Powell *et al*. 1995b; Vendramin *et al*. 1996; Vendramin and Ziegenhagen 1997; Sperisen *et al*. 1998). Moreover, the primers have been used with success in 110 different conifer species belonging to different taxa, in particular, *Pinaceae*, *Cupressaceae* and *Taxodiaceae* (G.G. Vendramin, unpublished results). The universality of cpSSRs allows the transfer of primers across taxa alleviating the cost involved in their identification for each species. The use of automated DNA sequencing apparatus and adequate software can increase the efficiency and allows obtaining a large set of data in a relatively short period of time.

The use of cpSSRs has allowed the investigation of the distribution of chloroplast haplotypes and haplotypic diversity in different conifers at the range level of the species (Lefort *et al*. 1999 and references therein), including *P. pinaster* (Vendramin *et al*. 1998), and that evidence appears to be associated with the migration processes from glacial refugia that occurred in the most recent postglacial period.

In contrast to nuclear genomes, plant organelle genomes are haploid and uniparentally inherited. For most angiosperms the chloroplast genome is maternally transmitted, but in conifers it is generally paternally inherited (Neale *et al*. 1986; Szmidt *et al*. 1987; Neale and Sederoff 1989; Wagner *et al*. 1989; Stine and

Keathley 1990; Dong *et al*. 1992; Ziegenhagen *et al*. 1995; Cato and Richardson 1996; Vendramin and Ziegenhagen 1997; Stoehr *et al*. 1998) including in *P. pinaster* (Plomion *et al*. 2001). Due to the uniparental mode of inheritance, the chloroplast genome behaves as a haploid single-locus and does not undergo recombination (Chiu and Sears 1985). Except for occasional mutations, this molecule is inherited unaltered with linked associated loci, therefore the sequences are a source of evolutionary history information.

Due to the uniparental inheritance and to the smaller effective population size, organelle genomes are more sensitive than nuclear DNA to severe reductions in the number of individuals in a population; they are expected to show different population dynamics and may be more sensitive to population subdivision (Birky *et al*. 1989). Studies in conifers have shown that cytoplasmic markers generally display higher values of differentiation compared with nuclear makers (e.g., Hong *et al*. 1993; Petit *et al*. 1993; Strauss *et al*. 1993; Ennos 1994; Wang and Szmidt 1994; Hong *et al*. 1995).

Patterns of population subdivision in conifers may be influenced by the contrasting mode of cpDNA and mtDNA inheritance (Hu and Ennos 1999). Gene flow of organelle genes distributed only through seed (e.g., maternal inheritance for mtDNA in pines) can be significantly less among wind-pollinated tree species compared to organelle genes distributed by pollen and by seed (e.g., paternal inheritance for cpDNA in pines). As a consequence, in wind-pollinated outcrossers such as pines, population subdivision can be weaker when cpDNA markers are used compared with mtDNA, because the wind-dispersed pollen is the main agent of gene flow (Dong and Wagner 1994; Ennos 1994; Latta and Mitton 1997).

Mutation rates are generally ignored because they are considered to be much lower than the migration rates, but this might not always be valid. Recent data show that mutation rates are higher at the cpSSR loci than substitution rates elsewhere in the chloroplast genome, and generally higher than in the nuclear genome sequences, except for the nuclear SSRs (Provan *et al*. 1999) and references therein), and in this case higher mutation rates could reduce population subdivision. Another possibility is size homoplasy, which has been observed at chloroplast microsatellites (Doyle *et al*. 1998). This could also lead to underestimates of differentiation when cpSSRs are used, by erasing some of the differences in haploytypes that have arisen in the past. Comparisons to other markers are needed to test whether or not those novel markers detect reliable estimates of genetic diversity and differentiation.

Summary of the results

Genetic variation of *Pinus pinaster*

Studies presented in Papers **I** and **II** provide information on the level and distribution of genetic variation among and within populations of *P. pinaster* at chloroplast and nuclear loci.

Variation within populations

In Paper **I,** some Portuguese populations of *P. pinaster* showed high within-population haplotypic diversity. The scattered distribution of *P. pinaster* in Portugal in ancient times, with different populations evolving separately and later mixed by human activity, could have been a source of haplotypic diversity. An additional hypothesis is the possible migration pathway originating from putative refugia in southeastern Spain, as defended by Salvador *et al*. (2000). When allozyme markers were used, Salvador *et al*. (2000) found higher values of within-population diversity of *P. pinaster* in the eastern region of the Iberian Peninsula than in the Atlantic region (which includes Portugal). Vendramin *et al*. (1998) also detected relatively low values of diversity in two of the populations from Portugal in their range-wide study of cpSSR variation in *P. pinaster*. The earlier hypothesis of Baradat and Marpeau-Bezard (1988), that Portugal was a refuge for the species during the last glaciation, was rejected by Vendramin *et al*. (1998) due to the low values for within-population diversity found in some of the Portuguese populations of *P. pinaster*. Those populations included in Vendramin *et al*. (1998) study were probably of very recent origin and experienced a severe reduction in population size, with consequent founder effect. Moreover, further comparisons (Paper **I**) based on a subset of common loci proved one population of Portuguese origin (Leiria) to be the most polymorphic of those used by Vendramin *et al*. (1998) and in Paper **I**.

Fossil, charcoal and palynological records indicate that *P. pinaster* was present in Portugal during the Middle Würm glaciation (55 000-25 000 BP), the late Plenigacial (25 000-15 000 BP) and terminal Pleistocene (12 000-11 000 BP) (Mateus and Queiroz 1993; Figueiral 1995 and references therein). In addition, pollen analyses have also indicated the presence of *P. pinaster* forests in the coastal area south of Lisbon, during the Atlantic period (7 580-6 550 BP). This species was probably able to survive the latest glaciation in Portugal in sheltered areas close to the Atlantic Ocean (Figueiral 1995), in contrast to what is suggested by Salvador *et al*. (2000). In their study, the latter also suggested the disappearance of *P. pinaster* in Portugal during the last glaciation. According to Bennett *et al*. (1991), the influence of the Atlantic Ocean causes western Europe to be anomalously warm for its latitude, and various evidence of southern European

climate and topography from 18 000 BP suggest that sites within this region should have been suitable for tree survival throughout the last glaciation, perhaps even at densities too low to disperse a detectable pollen rain.

In Paper **II,** the diversity of the French populations computed with cpSSRs and AFLPs (pruning those loci that showed less than 4 null homozygotes, as recommended by Lynch and Milligan 1994) was higher than the diversity of the Portuguese populations. This supports the trend found by Vendramin *et al*. (1998), i.e., a higher diversity in the French compared with the Portuguese populations, and the authors concluded that French populations represent one reservoir of haplotypic diversity in the range of the species. This finding does not exclude the possibility of a putative refugium in Portugal, and conclusions about the migration history of the species should be drawn in the context of historical, fossil and palynological information, due to the known human impacts experienced by this species in its range, as discussed in the foregoing and in Paper **I**. In particular, due to the intensive reforestation of this species in the southwest of France (Scott 1962), the high diversity found in the *P. pinaster* populations of that region could be a result of the mix of plant material coming from different places, representing a "melting pot" rather than a natural centre of diversity. It should be recognized that anthropogenic influences could have erased the fingerprints of migration pathways for *P. pinaster*, particularly in those regions where these effects were more intense, which was the case for Portugal and southwestern France.

Variation among populations

According to El-Kassaby (1991) and references therein, the majority of genetic variation for pine species is found within populations, with a small but significant component among populations. In contrast, in *P. pinaster* a high intraspecific variation and genetic differentiation has been found among populations across the range using different types of markers (phenotypic traits, terpenes, denatured proteins, allozymes and chloroplast microsatelites) (Baradat and Marpeau-Bezard 1988; Bahrman *et al*. 1994; Alía *et al*. 1995; Petit *et al*. 1995; Vendramin *et al*. 1998). The geographical isolation and the possibility of scattered refugia during the last glaciation have been offered as possible explanations for this pattern of differentiation (Baradat and Marpeau-Bezard 1988; Vendramin *et al*. 1998).

In studies **I** and **II**, both the Portuguese and French populations of *P. pinaster* showed a low level of among-population diversity. Nevertheless, those studies were made at a fine geographical scale, and both natural gene flow and human activities could have erased the differences among populations at a regional level. Castro (1989) using allozymes and Mariette *et al*. (2001b) using AFLP and nuclear microsatellites, both obtained a similar pattern of differentiation in groups of Portuguese and French populations. Results obtained with allozymes by Salvador *et al*. (2000) and González-Martínez *et al*. (2001) showed clinal trends of genetic variation and fine-scale spatial structure in the Iberian Peninsula, but orientation of

the main mountain ranges and the scattered distribution of *P. pinaster* in Spain constituted natural barriers preventing the genetic homogenizing effect of gene flow. Interestingly, one of the few stands that showed to be divergent from the others in Paper **I** is situated in a very isolated part of Portugal and is surrounded by a high range of mountains.

In Portugal, it was common in *P. pinaster* planting programmes since at least the beginning of the 20th century to use seeds from different parts of the country and even from abroad. The fact that no discernible geographical genetic pattern was found in study **I** reflects this practice and strong anthropogenic influences are confirmed by historical records (Devy-Vareta 1993). Good agreement was found between the results presented in Paper **I** and those obtained by Salvador *et al*. (2000), González-Martínez *et al*. (2001) and Burban *et al*. (1999). In the former two studies, the authors observed a lack of relationship between geographic and genetic distances among the northwestern populations of the Iberian Peninsula, in a set of populations from the Iberian Peninsula, using isozyme markers. In the latter study, the authors found a blurred geographic pattern for a specific pest of maritime pine, the bast scale *Matsucocus feytaudi*, in Portugal.

In southwestern France, the human impact through reforestation with *P. pinaster* in that region, which started as early as the 18th century and continued until a significant area was covered (Scott 1962), should also be taken in consideration. The low differentiation exhibited by cpSSRs and AFLPs of the French populations detected in Paper **II** is strong evidence of the probable influence of this factor, together with gene flow, in obscuring the divergence among populations within this region.

In study **II,** the differentiation obtained with all populations considered together (Portuguese and French) was higher than within each provenance separately, and both provenances could clearly be distinguished with both types of markers (AFLP and cpSSR). Several earlier studies at the range level of the species and using different type of markers (terpenes and denatured proteins) showed that populations from southwestern France and Portugal clustered together (Baradat and Marpeau-Bezard 1988; Bahrman *et al*. 1994). In a more recent study the two provenances were discriminated with the results obtained using cpSSR (Vendramin *et al*. 1998).

Nuclear versus cytoplasmic genetic variation

In Paper **II,** the AFLPs revealed much lower within-population diversity than the cpSSRs. This could be explained by the nature of microsatellites that are usually highly polymorphic markers (Lefort *et al*. 1999 and references therein), and the cpSSR used in the study **II** were already known to show polymorphism in *P. pinaster* (Vendramin *et al*. 1998 and Paper **I**). Nevertheless, the same trend was found with both markers, i.e., a lower diversity in the Portuguese compared with the French provenance, after the AFLP loci with less than 4 null homozygotes were

pruned from the analysis, as suggested by Lynch and Milligan (1994) (LM restriction).

The AFLP markers used in study **II** exhibited similar levels of among-population diversity compared with the cpSSR markers, within both the French and Portuguese provenances, when the LM restriction was used. In the absence of extensive gene flow, uniparentally inherited markers show, in general, less variation within populations and more among populations, than nuclear biparentally inherited markers, due to a lower effective number of genes and to differences in seed and pollen migration (Birky *et al*. 1983; Petit *et al*. 1993; Wade *et al*. 1994).

In pines, population subdivision can be weaker when cpDNA markers are used compared with mtDNA, because the wind-dispersed pollen is the main agent of gene flow (Dong and Wagner 1994; Mitton *et al*. 2000). Allozymes and cpDNA revealed little population structure in *P. flexilis*, whereas mtDNA showed high population differentiation (Latta and Mitton 1997). Additionally, use of the Mantel test in Paper **II** showed that the genetic distance matrix calculated with AFLP loci was highly correlated with that calculated with cpSSRs. Despite the fact that both types of markers have different modes of inheritance, it appears that the effect of gene flow through pollen surpasses the effect of genetic drift in shaping the genetic variation of the species, at least at the geographical scale studied in Paper **II**.

In Paper **I,** using cpSSR the results indicated that there is little or no geographic genetic pattern in Portuguese populations, due not only to the effect of human activity, but also to extensive gene flow among populations, as described in the previous paragraph. Extensive gene flow could also explain similar differentiation values found for both nuclear and cytoplasmic markers within the Portuguese provenance, by smoothing differences due to differences in effective population sizes and genetic drift. Moreover, in study **II** the genetic differentiation among the Portuguese populations with AFLP (after LM restriction) and cpSSR data that are similar to that reported by Castro (1989) using allozyme markers (G_{ST} = 0.020) in six populations of *P. pinaster* widely spaced across Portugal.

The differentiation estimates obtained with the cpSSR and LM-restricted AFLP data in Paper **II** for the French provenance are also similar to the estimate obtained with nuclear microsatellite markers in a study made by Mariette *et al*. (2001b), in the same group of populations. Again, the low differentiation showed with different type of markers is a strong evidence of the genetic homogenising effect among populations due to extensive gene flow.

Studies in conifers that report lower population subdivision with nuclear makers compared with cytoplasmic markers have not included chloroplast microsatellites (e.g., Birky 1988, Petit *et al*. 1993; Dong and Wagner 1994; Ennos 1994), with the exception of a study in *P. leucodermis* with cpSSR and allozymes reported by Powell *et al*. (1995b). Recent data showed that mutation rates are higher at the cpSSR loci than substitution rates elsewhere in the chloroplast and nuclear genome, except for the nuclear SSRs (Provan *et al*. 1999 and references therein), and in

this case higher mutation rates could decrease population subdivision. Another argument to be considered is the size homoplasy observable at cpSSR loci (Doyle *et al*. 1998 and references therein). Since cpSSRs are generated by mutations at a limited number of hotspots, they are prone to suffer identical mutations occurring independently in different populations, which, in turn, biases the estimates of differentiation downward by erasing the differences in haplotypes the populations would otherwise possess.

Reproductive material identification and certification

Following World War II, during which large areas of forest were burned, reforestation programs were undertaken in the southwest of France (the Aquitaine region) with *P. pinaster* seeds of northwestern Iberian origin (Portugal and Galicia), and the stands they formed suffered considerable frost damage (Boisseaux 1986). To overcome this problem, and to avoid further damage, a terpene-based test was developed to test the putative origin of adult stands in Aquitaine, before seeds could be distributed for commercial purposes in France (Baradat and Marpeau-Bezard 1988). In Paper **III**, we describe a new test employing cpSSRs, based on randomisation tests to facilitate identification of stand origin of stands and seed-lots of P. *pinaster* in the Aquitaine region in France, and thus providing an alternative to the former terpene-based test for reproductive material identification and certification.

The origin of five stands of unknown origin was determined with both the cpSSR and biochemical (terpene profile analysis) tests. The results obtained with terpenes proved to be less efficient than those obtained with the cpSSRs. The terpene test was initially inconclusive for two out of the five tested stands and it had to be repeated before a reliable answer was obtained. Thus, use of the terpene test risks the need for repetitions, increasing the amount of plant material, costs and time needed to get the same information as obtained from one cpSSR test.

The results obtained in Paper **III** can be partially explained by the fact the terpene markers discriminate less the Portuguese from the French provenances as indicated by the range-wide *P. pinaster* study that employed terpene analysis (Baradat and Marpeau-Bezard 1988). Chloroplast microsatellites revealed high levels of genetic differentiation among populations across the range due to differences in allele size (Vendramin *et al*. 1998), compared with the other markers (Bahrman *et al*. 1994; Petit *et al*. 1995). The cpSSRs data were used to design the test, due to the fact that they showed a homogeneous distribution of the polymorphism within groups and clear differentiation between the two groups of populations (French and Portuguese). Therefore, the cpSSR markers used in the Paper **III** were suitable for the principal purpose of that study, the design of a test to determine the putative origin of *P. pinaster* stands in the Aquitaine region of France.

Closing remarks

Studies **I** and **II** presented in this thesis illustrate the patterns of population genetic structure of the investigated populations of *P. pinaster*. Paper **I** presents strong evidence of Man's influence in shaping the genetic structure of a species and how the mixing of genetic material of unknown origin can affect the genetic resources. Little or no geographic genetic pattern was found in Portuguese populations of *P. pinaster* due to human activities during the last century, and to the extensive associated gene flow among populations. As a consequence of the blurred pattern found in the Portuguese populations of *P. pinaster*, the interpretation of the history of this species in this area may prove difficult if not impossible. In Paper **II** a similar population structure was observed in the southwestern region of France. The low levels of differentiation found among populations of that region was probably due to intense gene flow, but the reported human influence in the region's forest building is not to be excluded.

In future research, it is recommended to perform a range-wide study with a denser sampling of *P. pinaster* populations, to confirm the hypothesis drawn by Vendramin *et al*. (1998) to explain the history and pathways of the species since the last glaciation. The use of other markers is suggested, such as mtDNA markers, due to the lower scale dispersal of the seeds compared with the pollen in pine species. Gene flow through pollen could have blurred the information retrieved by paternally inherited chloroplast markers. Unfortunately, the polymorphism found so far with mtDNA markers (C. Burban, unpublished results) is very low, and further research needs to be pursued. Additionally, conclusions about the migration history of the species should also be presented in the context of historical, fossil and palynological information, due to the known anthropogenic influence in the forests of the Mediterranean Basin.

Similar levels of differentiation were found with AFLPs and cpSSRs markers, when the LM restriction was applied, in the two regions studied in Paper **II**. Moreover, the differentiation they exhibited was similar to that computed with codominant markers (nuclear microsatellites and allozymes) for the same regions. Despite the fact that the two kinds of markers have different modes of inheritance, the trends found with the genetic distances computed with both types of markers were similar. This observation supports the hypothesis that the effect of gene flow through pollen surpassed the effect of genetic drift in shaping the genetic variation of the species at the geographical scale studied.

In the genus *Pinus*, pollen flow among adjacent populations is generally high and lowers population subdivision, which may have smoothed the differences found here between estimates of differentiation by nuclear and chloroplast markers. However, it is likely that species that show low levels of differentiation on a

regional scale may reveal very high differentiation on a macro-scale. *Pinus pinaster* seems to be a good example of this kind of behaviour.

The test for provenance identification designed in Paper **III** can easily be applied to other commercial species, provided that there is a homogeneous distribution of the polymorphism within groups and clear differentiation among groups of populations. In particular, the cpSSR primers have already been shown to cross-amplify sequences from several species, which could be very advantageous given the long time and high costs involved in identifying markers (Powell *et al*. 1995b, Vendramin *et al*. 1996). The availability of reliable tests for identifying the origins of reproductive material will be very valuable for providing solutions to seed certification problems and also in the context of gene conservation.

References

Alía, R., Gil, L., & Pardos, J.A. (1995) Performance of 43 *Pinus pinaster* Ait. provenances on 5 locations in central Spain. *Silvae Genetica,* 44, 75-81.

Alía, R., Martín, S., De Miguel, J., Galera, R., Agúndez, D., Gordo, J., Salvador, L., Catalán, G., & Gil, L. (1996) *Regiones de procedencia de Pinus pinaster Aiton*. Organismo Autónomo Parques Nacionales, Madrid.

Aragones, A., Barrena, I., Espinel, S., Herran, A., & Ritter, E. (1997) Origin of Basque populations of radiata pine inferred from RAPD data. *Annales des Sciences Forestières,* 54, 697-703.

Arcade, A., Anselin, F., Rampant, P.F., Lesage, M.C., Paques, L.E., & Prat, D. (2000) Application of AFLP, RAPD and ISSR markers to genetic mapping of European and Japanese larch. *Theoretical and Applied Genetics,* 100, 299-307.

Bahrman, N., Zivy, M., Baradat, P., & Damerval, C. (1994) Organization of the variability of abundant proteins in seven geographical origins of maritime pine (*Pinus pinaster* Ait.). *Theoretical and Applied Genetics,* 88, 407-411.

Baradat, P.H., Bernard-Dagan, C., Fillon, A., Marpeau, A., & Pauly, G. (1972) Les terpènes du Pin maritime: aspect biologiques et génétiques. II. Hérédité de la teneur en monoterpènes. *Annales des Sciences Forestières,* 29, 307-334.

Baradat, P.H., Maillart, A., Marpeau, A., Slak, M.F., Yani, A., & Pastuszka, P. (1995) Utility of terpenes to acess population structure and mating patterns in conifers. In: *Population genetics and genetic conservation of forest trees,* (ed Baradat, P., Adams, W.T., & Müller-Starck, G.), 5-27. Academic Publishing, Amsterdam.

Baradat, P.H., Marpeau, A., & Walter, J. (1991) Terpene Markers. In: *Variation in European Populations of Forest Trees,* (ed Müller-Stark, G., & Ziehe, M.), 40-65. Sauerländer's Verlag, Frankfurt am Main.

Baradat, P.H., & Marpeau-Bezard, A. (1988) *Le pin maritime, Pinus pinaster Ait.. Biologie et génétique des terpènes pour la connaissance et l'amélioration de l'espèce*. Ph.D. Thesis, University of Bordeaux I.

Barber, J.D., Callahan, R.Z., Wakeney, P.C., & Rudolf, P.O. (1962) More on tree seed certification and legislation. *Journal of Forestry,* 60, 349-352.

Barbero, M., Lisel, R., Quezel, P., Richardson, D.M., & Romaine, F. (1998) Pines of the Mediterranean basin. In: *Ecology and biogeography of Pinus,* (ed Richardson, D.M.), 153-170. Cambridge University Press, Cambridge.

Barrett, S.C.H., & Kohn, J.R. (1991) *Genetic and Evolutionary Consequences of Small Population Size in Plants - Implications for Conservation.* Oxford Univ Press, 200 Madison Avenue/New York/NY 10016.

Bennett, K.D., Tzedakis, P.C., & Willis, K.J. (1991) Quaternary refuge of north Europenan trees. *Journal of Biogeography,* 18, 103-115.

Bernard-Dagan, C., Fillon, A., Pauly, G., & Baradat, P.H. (1971) Les terpènes du Pin maritime: aspect biologiques et génétiques. I. Variabilité de la composition monoterpénique dans un individu, entre individus et entre provenances. *Annales des Sciences Forestières,* 28, 223-258.

Birky, C.W. (1988) Evolution and variation in plant chloroplast and mitochondrial genomes. In: *Plant evolutionary biology*, (ed Gottlieb, L., & Jain, S.), 23-53. Chapman & Hall, London.

Birky, J., C W, Fuerst, P., & Maruyama, T. (1989) Organelle gene diversity under migration, mutation, and drift: equilibrium expectations, Approach to equilibrium, effects of heteroplasmic cells, and comparison to nuclear genes. *Genetics*, 121, 613-627.

Birky, W., Maruyama, T., & Fuerst, P. (1983) An approach to population and evolutionary genetic theory for genes in mitochondria and chloroplasts, and some results. *Genetics*, 103, 513-527.

Boisseaux, T. (1986) *Influence de l'origine génétique (landaise ou ibérique) des peuplements de Pin maritime sur les dégâts causés par le froid de janvier 1985 au massif forestier aquitain. Mise au point d'un test variétal précoce utilisable pour le contrôle de lots de graines*. Mémoire de l'ENITEF Thesis, Université de Bordeaux.

Bucci, G., & Menozzi, P. (1995) Genetic variation of RAPD markers in a *Picea abies* Karst. population. *Heredity*, 75, 188-197.

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

Burban, C., Petit, R.J., Carcreff, E., & Jactel, H. (1999) Rangewide variation of the maritime pine bast scale *Matsucoccus feytaudi* Duc. (Homoptera : *Matsucoccidae*) in relation to the genetic structure of its host. *Molecular Ecology*, 8, 1593-1602.

Buting, S.C., & Rego, F.C. (1988) Human impact on Portugal's vegetation. *Rangelands*, 10, 251-255.

Caetano-Annolés, G., Bassam, B.J., & Gresshoff, P.M. (1991) High resolution DNA amplification fingerprinting using very short arbitrary oligonucleotide primers. *Biotechnology*, 9, 553-557.

Castro, L.F.T. (1989) *Isoenzimas do Pinus pinaster Ait. numa prespectiva de aplicação ao melhoramento genético da espécie*. Ph.D. Thesis, Universidade de Trás-os-Montes e Alto Douro.

Cato, S.A., Corbett, G.E., & Richardson, T.E. (1999) Evaluation of AFLP for genetic mapping in *Pinus radiata* D. Don. *Molecular Breeding*, 5, 275-281.

Cato, S.A., & Richardson, T.E. (1996) Inter- and intraspecific polymorphism at chloroplast SSR loci and the inheritance of plastids in *Pinus radiata* D. Don. *Theoretical and Applied Genetics*, 93, 587-592.

Cervera, M.T., Gusmão, J., Steenackers, M., Vangysel, A., Vanmontagu, M., & Boerjan, W. (1996) Application of AFLP(TM)-based molecular markers to breeding of *Populus* spp. *Plant Growth Regulation*, 20, 47-52.

Cervera, M.T., Remington, D., Frigerio, J.-M., Storme, V., Ivens, B., Boerjan, W., & Plomion, C. (2000) Improved AFLP analysis of tree species. *Canadian Journal of Forest Research*, 30, 1608-1616.

Chappell, J. (1995) The biochemistry and molecular biology of isoprenoid methbolism. *Plant Physiology*, 107, 1-6.

Chiu, W.-L., & Sears, B.B. (1985) Recombination between chloroplast DNAs does not occur in sexual crosses of *Oenothera*. *Molecular General Genetics*, 198, 525-528.

Clegg, M.T., & Zurawski, G. (1992) Chloroplast DNA and the study of plant phylogeny: present status and future prospects. In: *Molecular Systematics of Plants*, (ed Soltis, P.S., Soltis, D.E., & Doyle, J.J.), 1-13. Chapman and Hall, New York.

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

Coppen, J.J.W., Gay, C., James, D.J., Robinson, J.M., & Supriana, N. (1993) Variability in xylem resin composition amongst natural populations of Indonesian *Pinus merkusii*. *Phytochemistry*, 33, 129-136.

Costa, P., Pot, D., Dubos, C., Frigerio, J.M., Pionneau, C., Bodenes, C., Bertocchi, E., Cervera, M.-T., Remington, D.L., & Plomion, C. (2000) A genetic map of Maritime pine based on AFLP, RAPD and protein markers. *Theoretical and Applied Genetics*, 100, 39-48.

Crawford, D.J. (1983) Phylogenetic and systematic inferences from electrophoretic studies. In: *Isozymes in Plant Genetics and Breeding, Part A*, (ed Tanksley, S.D., & Orton, T.J.), pp. 257-287. Elsevier Science Publishers B. V., Amsterdam.

Deroy, J. (2000) *Etude de la diversité génétique du pin maritime en France en utilisant des marquers microsatellites*. Mémoire de l'ENITEF Thesis, Université de Bordeaux.

Devy-Vareta, N.L. (1993) *A floresta no espaço e no tempo em Portugal. A arborização da Serra da Cabreira (1919-1975).* Ph.D. Thesis, Faculdade de Letras da Universidade do Porto.

Dong, J., Wagner, D.B., Yanchuk, A.D., Carlson, M.R., Magnussen, S., Wang, X.R., & Szmidt, A.E. (1992) Paternal chloroplast DNA inheritance in *Pinus contorta* and *Pinus banksiana* - independence of parental species or cross direction. *Journal of Heredity*, 83, 419-422.

Dong, J.S., & Wagner, D.B. (1994) Paternally inherited chloroplast polymorphism in *Pinus*: Estimation of diversity and population subdivision, and tests of disequilibrium with a maternally inherited mitochondrial polymorphism. *Genetics*, 136, 1187-1194.

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

Echt, C.S., & Maymarquardt, P. (1997) Survey of microsatellite DNA in pine. *Genome*, 40, 9-17.

Echt, C.S., Vendramin, G.G., Nelson, C.D., & Marquardt, P. (1999) Microsatellite DNA as shared genetic markers among conifer species. *Canadian Journal of Forest Research*, 29, 365-371.

El-Kassaby, Y.A. (1991) Genetic variation within and among conifer populations: review and evaluation of methods. In: *Biochemical markers in the population genetics of forest trees*, (ed Fineschi, S., Malvolti, M.E., Cannata, F., & Hattemer, H.H.), 61-76. SPB Academic Publishing bv, The Hague.

Ellegren, H. (1992) Polymerase-Chain-Reaction (PCR) analysis of microsatellites - a new approach to studies of genetic relationships in birds. *Auk*, 109, 886-895.

Ellstrand, N.C. (1992) Gene flow by pollen - implications for plant conservation genetics. *Oikos*, 63, 77-86.

Ellstrand, N.C., & Elam, D.R. (1993) Population genetic consequences of small population size - Implications for plant conservation. *Annual Review of Ecology and Systematics*, 24, 217-242.

Ennos, R. (1994) Estimating the relative rates of pollen and seed migration among plant populations. *Heredity*, 72, 250-259.

Espinel, S., Aragones, A., & Ritter, E. (1995) Performance of different provenances and of the local population of the Monterey pine (*Pinus radiata* D. Don) in northern Spain. *Annales des Sciences Forestières*, 52, 515-519.

Falkenhagen, E.R. (1985) Isozyme studies in provenance research of forest trees. *Theoretical and Applied Genetics*, 69, 335-347.

Farjon, A. (1984) *Pines: Drawings and descriptions of the genus*. E. Brill/DR. Backhuys, Leiden, The Netherlands.

Farjon, A. (1996) Biodiversity of *Pinus* (*Pinaceae*) in Mexico: speciation and palaeo-endemism. *Botanical Journal of the Linnean Society*, 121, 365-384.

Figueiral, I. (1995) Charcoal analysis and the history of *Pinus pinaster* (cluster pine) in Portugal. *Review of Palaeobotany and Palynology*, 89, 441-454.

Gallois, A., Audran, J.C., & Burrus, M. (1998) Assessment of genetic relationships and population discrimination among *Fagus sylvatica* L. by RAPD. *Theoretical and Applied Genetics*, 97, 211-219.

González-Martínez, S.C., Agúndez, D., Alía, R., Salvador, L., & Gil, L. (2001) Geographical variation of gene diversity of *Pinus pinaster* Ait. in the Iberian Peninsula. In: *Genetic Response of Forest Systems to Changing Environmental Conditions*, (ed Müller-Starck, G.), in press. Kluwer Academic Press, Dordrecht.

Hallden, C., Hansen, M., Nilsson, N.O., Hjerdin, A., & Sall, T. (1996) Competition as a source of errors in RAPD analysis. *Theoretical and Applied Genetics*, 93, 1185-1192.

Hamrick, J.L., Mitton, J.B., & Linhart, Y.B. (1981) Levels of genetic variation in trees: influence of life history characteristics. In: *Proceedings of the symposium on isozymes of North American forest trees and forest insects*, (ed Conkle, M.T.), 35-41. PSW-48, USDA, Forest Service, Berkeley, Ca.

Hanover, J.W. (1992) Tree fitness as a function of terpene composition. *New Forests*, 6, 159-178.

Harfouche, A., & Kremer, A. (2000) Provenance hybridization in a diallel mating scheme of maritime pine (*Pinus pinaster*). I. Means and variance components. *Canadian Journal of Forest Research*, 30, 1-9.

Hartl, D.L., & Clark, A.G. (1997) *Principles of population genetics.* Sinauer Associates, Sunderland, MA.

Hewitt, G. (2000) The genetic legacy of the Quaternary ice ages. *Nature*, 405, 907-913.

Hewitt, G.M. (1996) Some genetic consequence of ice ages, and their role in divergence and speciation. *Biological Journal of the Linnean Society*, 58, 247-276.

Hong, Y.P., Bonoy, B., & Carlson, J.E. (1995) Genetic diversity and phylogeny in douglas-fir based on RFLP and RAPD (DAF) analysis of nuclear, chloroplast, and mitochondrial genomes. In: *Population genetics and genetic conservation of forest trees*, (ed Baradat, P., Adams, W.T., & Müller-Starck, G.), 247-266. Academic Publishing, Amsterdam.

Hong, Y.P., Hipkins, V.D., & Strauss, S.H. (1993) Chloroplast DNA diversity among trees, populations and species in the California closed-cone pines (*Pinus radiata*, *Pinus muricata* and *Pinus attenuata*). *Genetics*, 135, 1187-1196.

Hu, X.S., & Ennos, R.A. (1999) Impacts of seed and pollen flow on population genetic structure for plant genomes with three contrasting modes of inheritance. *Genetics*, 152, 441-450.

Isabel, N., Beaulieu, J., & Bousquet, J. (1995) Complete congruence between gene diversity estimates derived from genotypic data at enzyme and random amplified polymorphic DNA loci in black spruce. *Proceedings of the National Academy of Sciences, USA*, 92, 6369-6373.

Isabel, N., Beaulieu, J., Theriault, P., & Bousquet, J. (1999) Direct evidence for biased gene diversity estimates from dominant random amplified polymorphic DNA (RAPD) fingerprints. *Molecular Ecology*, 8, 477-483.

Jones, C.J., Edwards, K.J., Castaglione, S., Winfield, M.O., Sala, F., Vandewiel, C., Bredemeijer, G., Vosman, B., Matthes, M., Daly, A., Brettschneider, R., Bettini, P., Buiatti, M., Maestri, E., Malcevschi, A., Marmiroli, N., Aert, R., Volckaert, G., Rueda, J., Linacero, R., Vazquez, A., & Karp, A. (1997) Reproducibility testing of RAPD, AFLP and SSR markers in plants by a network of European laboratories. *Molecular Breeding*, 3, 381-390.

Jones, N., & Burley, J. (1973) Seed certification, provenance nomenclature and genetic history in forestry. *Silvae Genetica*, 22, 53-58.

Karp, A., & Edwards, K.J. (1997) DNA markers: a global overview. In: *DNA markers: protocols, applications and overviews*, (ed Caetano-Anollés, G., & Gresshoff, P.M.), 1-13. Wiley-VCH, New York.

Kazan, K., Manners, J.M., & Cameron, D.F. (1993) Inheritance of random amplified polymorphic DNA markers in an interspecific cross in the genus *Stylosanthes*. *Genome*, 36, 50-56.

Kijas, J.M.H., Fowler, J.C.S., & Thomas, M.R. (1995) An evaluation of sequence tagged microsatellite site markers for genetic analysis within *Citrus* and related species. *Genome*, 38, 349-355.

Kimura, M. (1968) Evolutionary rate at the molecular level. *Nature*, 217, 624-626.

Klaus, W. (1989) Mediterranean pines and their history. *Plant Systematic and Evolution*, 162, 133-163.

Krutovskii, K.V., Erofeeva, S.Y., Aagaard, J.E., & Strauss, S.H. (1999) Simulation of effects of dominance on estimates of population genetic diversity and differentiation. *The Journal of Heredity*, 90, 499-503.

Latta, R.G., & Mitton, J.B. (1997) A comparison of population differentiation across four classes of gene marker in limber pine (*Pinus flexilis* James). *Genetics*, 146, 1153-1163.

Le Maitre, D.C. (1998) Pines in cultivation: a global view. In: *Ecology and biogeography of Pinus*, (ed Richardson, D.M.), 407-431. Cambridge University Press, Cambridge.

Ledig, F.T. (1992) Human impacts on genetic diversity in forest ecosystems. *Oikos*, 63, 87-108.

Lefort, F., Echt, C., Streiff, R., & Vendramin, G.G. (1999) Microsatellite sequences: a new generation of molecular markers for forest genetics. *Forest Genetics,* 6, 15-20.

Lerceteau, E., Plomion, C., & Andersson, B. (2000) AFLP mapping and detection of quantitative trait loci (QTLs) for economically important traits in *Pinus sylvestris*: a preliminary study. *Molecular Breeding,* 6, 451-458.

Lerceteau, E., & Szmidt, A.E. (1999) Properties of AFLP markers in inheritance and genetic diversity studies of *Pinus sylvestris* L. *Heredity,* 82, 252-260.

Little, J.E.L., & Critchfield, W.B. (1969) Subdivisions of the genus *Pinus* (pines). USDA Forest Service, Washington D.C., Miscellaneous Publication No. 1144, Report, December 1969.

Lu, M.-Z., Szmidt, A.E., & Wang, X.-R. (1995) Inheritance of RAPD fragments in haploid and diploid tissues of *Pinus sylvestris* (L.). *Heredity*, 74, 582-589.

Lu, M.-Z., Wang, X.-R., & Szmidt, A.E. (1997) Molecular characterisation of RAPDs in *Pinus sylvestris* L. *Forest Genetics,* 4, 227-234.

Lynch, M., & Milligan, B.G. (1994) Analysis of population genetic structure with RAPD markers. *Molecular Ecology,* 3, 91-99.

Mariette, S., Chagne, D., Decroocq, S., Vendramin, G.G., Lalanne, C., Madur, D., & Plomion, C. (2001a) Microsatellite markers for *Pinus pinaster* Ait. *Annals of Forest Science,* in press.

Mariette, S., Chagne, D., Lezier, C., Pastuszka, P., Raffin, A., Plomion, C., & Kremer, A. (2001b) Genetic diversity within and among *Pinus pinaster* populations: comparison between AFLP and microsatellite markers. *Heredity,* in press.

Marpeau, A., Baradat, P.H., & Bernard-Dagan, C. (1975) Les terpènes du Pin maritime: aspect biologiques et génétiques. IV. Hérédité de la teneur en deux sesquiterpènes: le longifolène et le caryophyllène. *Annales des Sciences Forestières,* 32, 185-203.

Marques, C.M., Araujo, J.A., Ferreira, J.G., Whetten, R., O'Malley, D.M., Liu, B.H., & Sederoff, R. (1998) AFLP genetic maps of *Eucalyptus globulus* and *E. tereticornis. Theoretical and Applied Genetics,* 96, 727-737.

Marques, C.M., Vasquezkool, J., Carocha, V.J., Ferreira, J.G., O'Malley, D.M., Liu, B.H., & Sederoff, R. (1999) Genetic dissection of vegetative propagation traits in *Eucalyptus tereticornis* and *E. globulus. Theoretical and Applied Genetics,* 99, 936-946.

Mateus, J.E., & Queiroz, P.F. (1993) Os estudos da vegetação quaternária em Portugal: Contextos, balanço de resultados, prespectivas. O quaternário em Portugal, balanço e prespectivas. *Colibri,* 105-131.

Matthews, J.D. (1964) Seed production and seed certification. *Unasylva,* 18, 104-118.

Mattoso, J., & Sousa, A. (1993) *História de Portugal*. Círculo de Leitores, Lisboa.

Millar, C.I. (1993) Impact of the Eocene on the Evolution of *Pinus* L. *Annals of the Missouri Botanical Garden,* **80**, 471-498.

Millar, C.I. (1998) Early evolution of pines. In: *Ecology and biogeography of Pinus,* (ed Richardson, D.M.), 69-91. Cambridge University Press, Cambridge.

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

Mitton, J.B. (1994) Molecular approaches to population biology. *Annual Revue of Ecology and Systematics,* 25, 45-69.

Mitton, J.B., Kreiser, B.R., & Rehfeldt, G.E. (2000) Primers designed to amplify a mitochondrial nad1 intron in ponderosa pine, *Pinus ponderosa*, limber pine, *P. flexilis*, and Scots pine, *P. sylvestris*. *Theoretical and Applied Genetics*, 101, 1269-1272.

Morgante, M., Pfeiffer, A., Costacurta, A., & Olivieri, A.M. (1996) Molecular tools for population and ecological genetics in coniferous trees. *Phyton - Annales Rei Botanicae*, 36, 129-138.

Müller-Starck, G., Baradat, P., & Bergmann, F. (1992) Genetic variation within European tree species. *New Forests*, 6, 23-47.

Mullis, K.B., & Faloona, F.A. (1987) Specific synthesis of DNA *in vitro via* a polymerase-catalyzed chain reaction. In: *Methods in Enzymology*, (ed Wu, R.), 335-350. 155, Academic Press, New York and London.

Neale, D.B., & Sederoff, R.R. (1989) Paternal inheritance of chloroplast DNA and maternal inheritance of mitochondrial DNA in loblolly pine. *Theoretical and Applied Genetics*, 77, 212-216.

Neale, D.B., Wheeler, N.C., & Allard, R.W. (1986) Paternal inheritance of chloroplast DNA in Douglas-fir. *Canadian Journal of Forest Research*, 16, 1152-1154.

Nesbitt, K.A., Potts, B.M., Vaillancourt, R.E., West, A.K., & Reid, J.B. (1995) Partitioning and distribution of RAPD variation in a forest tree species, *Eucalyptus globulus* (*Myrtaceae*). *Heredity*, 74, 628-637.

Nikaido, A., Yoshimaru, H., Tsumura, Y., Suyama, Y., Murai, M., & Nagasaka, K. (1999) Segregation distortion for AFLP markers in *Cryptomeria japonica*. *Genes & Genetic Systems*, 74, 55-59.

Paglia, G., & Morgante, M. (1998) PCR-based multiplex DNA fingerprinting techniques for the analysis of conifer genomes. *Molecular Breeding*, 4, 173-177.

Parker, P.G., Snow, A.A., Schug, M.D., Booton, G.C., & Fuerst, P.A. (1998) What molecules can tell us about populations: choosing and using a molecular marker. *Ecology*, 79, 361-382.

Penner, G.A., Bush, A., Wise, R., Kim, W., Domier, L., Kasha, K., Laroche, A., Scoles, G., S.J., M., & Fedak, G. (1993) Reproducibility of Random Amplified Polymorphic DNA (RAPD) analysis among laboratories. *PCR methods and application*, 2, 341-345.

Petit, R.J., Bahrman, N., & Baradat, P. (1995) Comparison of genetic differentiation in maritime pine (*Pinus pinaster* Ait.) estimated using isozyme, total protein and terpenic loci. *Heredity*, 75, 382-389.

Petit, R.J., Kremer, A., & Wagner, D.B. (1993) Finite island model for organelle and nuclear genes in plants. *Heredity*, 71, 630-641.

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

Plomion, C., LeProvost, G., Pot, D., Vendramin, G., Gerber, S., Decroocq, S., Brach, J., Raffin, A., & Pastuszka, P. (2001) Estimation of pollen contamination in a maritime pine polycross seed orchard and certification of improved seeds using cpSSR. *Canadian Journal of Forest Research*, submitted.

Powell, W., Morgante, M., Andre, C., Hanafey, M., Vogel, J., Tingey, S., & Rafalski, A. (1996) The comparison of RFLP, RAPD, AFLP and SSR (microsatellite) markers for germplasm analysis. *Molecular Breeding*, 2, 225-238.

Powell, W., Morgante, M., & McDevitt, R. (1995a) Hypervariable microsatellites provide a general source of polymorphic DNA markers for the chloroplast genome. *Current Biology*, 5, 1023-1029.

Powell, W., Morgante, M., McDevitt, R., Vendramin, G., & Rafalski, J. (1995b) Polymorphic simple sequence regions in chloroplast genomes: applications to the population genetics of pines. *Proceedings of the National Academy of Sciences, USA*, 99, 7759-7763.

Provan, J., Soranzo, N., Wilson, N.J., Goldstein, D.B., & Powell, W. (1999) A low mutation rate for chloroplast microsatellites. *Genetics*, 153, 943-947.

Rafalski, J.A. (1997) Randomly amplified polymorphic DNA (RAPD) analysis. In: *DNA Markers: protocols, applications and overviews*, (ed Caetano-Anollés, G., & Gresshoff, P.M.), 75-83. Wiley-VCH, New York.

Remington, D.L., Whetten, R.W., Liu, B.H., & O'Malley, D.M. (1999) Construction of an AFLP genetic map with nearly complete genome coverage in *Pinus taeda*. *Theoretical and Applied Genetics*, 98, 1279-1292.

Saiki, R.K., Gelfand, D.H., Stoffel, S., Scharf, S.J., Higuchi, R., Horn, G.T., Mullis, K.B., & Erlich, H.A. (1988) Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. *Science*, 239, 487-491.

Salvador, L., Alía, R., Agúndez, D., & Gil, L. (2000) Genetic variation and migration pathways of maritime pine (*Pinus pinaster* Ait) in the Iberian peninsula. *Theoretical and Applied Genetics*, 100, 89-95.

Sax, K., & Sax, H.J. (1933) Chromosome number and morphology in the conifers. *Journal of the Arnold Arboretum*, XIV, 356-375.

Schierenbeck, K.A., Skupski, M., Lieberman, D., & Lieberman, M. (1997) Population structure and genetic diversity in four tropical tree species in Costa Rica. *Molecular Ecology*, 6, 137-144.

Schiller, G., & Genizi, A. (1993) An attempt to identify the origin of *Pinus brutia* Ten. Plantations in Israel by needle resin composition. *Silvae Genetica*, 42, 63-68.

Scott, C.W. (1962) A summary of information on *Pinus pinaster*. *Forestry Abstracts*, 23, i-xviii.

Sewell, M.M., Bassoni, D.L., Megraw, R.A., Wheeler, N.C., & Neale, D.B. (2000) Identification of QTLs influencing wood property traits in loblolly pine (*Pinus taeda* L.). I. Physical wood properties. *Theoretical and Applied Genetics*, 101, 1273–1281.

Shaw, G.R. (1914) *The genus Pinus*. Riverside Press, Cambridge, MA,

Sinclair, W.T., Morman, J.D., & Ennos, R.A. (1998) Multiple origins for Scots pine (*Pinus sylvestris* L.) in Scotland: evidence from mitochondrial DNA variation. *Heredity*, 80, 233-240.

Sperisen, C., Büchler, U., Mátyás, G., Anzidei, M., Madaghiele, A., Skrøppa, T., & Vendramin, G.G. (1998) Polymorphic tandem repeats in the chloroplast and mitochondrial genomes of Norway spruce. In: *Genetics and Breeding of Norway Spruce.*, (ed Skrøppa, T., Paule, L., & Gömöry, D.), 15-24. Arbora Publisher, Zvolen.

Stearns, S.C., & Hoekstra, R.F. (2000) *Evolution: an introduction*. Oxford University Press, Oxford.

Stine, M., & Keathley, D.E. (1990) Paternal inheritance of plastids in Engelmann spruce x blue spruce hybrids. *Journal of Heredity*, 81, 443-446.

Stoehr, M.U., Orvar, B.L., Vo, T.M., Gawley, J.R., Webber, J.E., & Newton, C.H. (1998) Application of a chloroplast DNA marker in seed orchard management evaluations of Douglas-fir. *Canadian Journal of Forest Research*, 28, 187-195.

Strand, M., Prolla, T.A., Liskay, R.M., & Petes, T.D. (1993) Destabilzation of tracts of simple repetitive DNA in yeast by mutations affecting DNA mismatch repair. *Nature*, 365, 274-276.

Strauss, S., Bousquet, J., Hipkins, V., & Hong, Y.-P. (1992) Biochemical and molecular genetic markers in biosystematic studies of forest trees. *New Forests*, 6, 125-158.

Strauss, S.H., & Doerksen, A.H. (1990) Restriction fragment analysis of pine phylogeny. *Evolution*, 44, 1081-1096.

Strauss, S.H., Hong, Y.P., & Hipkins, V.D. (1993) High levels of population differentiation for mitochondrial DNA haplotypes in *Pinus radiata*, *muricata*, and *attenuata*. *Theoretical and Applied Genetics*, 86, 605-611.

Szmidt, A.E., Alden, T., & Hällgren, J.-E. (1987) Paternal inheritance of chloroplast DNA in *Larix*. *Plant Molecular Biology*, 9, 59-64.

Szmidt, A.E., El-Kassaby, Y.A., Sigurgeirsson, A., Alden, T., Lindgren, D., & Hällgren, J.-E. (1988) Classifying seedlots of *Picea sitchensis* and *P. glauca* in zones of introgression using restriction analysis of chloroplast DNA. *Theoretical and Applied Genetics*, 76, 841-845.

Szmidt, A.E., & Wang, X.-R. (2000) Genetic markers in forest genetics and breeding - the tunnel remains dark. In: *Forest Genetics and Sustainability*, (ed Matyas, C.), 31-48. Kluwer Academic Publishers, The Netherlands.

Szmidt, A.E., Wang, X.-R., & Lu, M.-Z. (1996) Empirical assessment of allozyme and RAPD variation in *Pinus sylvestris* L. using haploid tissue analysis. *Heredity*, 76, 412-420.

Taberlet, P., Fumagalli, L., Wustsaucy, A.G., & Cosson, J.F. (1998) Comparative phylogeography and postglacial colonization routes in Europe. *Molecular Ecology*, 7, 453-464.

Tautz, D. (1989) Hypervariability of simple sequences as a general source for polymorphic DNA markers. *Nuclec Acid Research*, 17, 6463-6471.

Teixeira, C. (1945) Subsídios para a história evolutiva do pinheiro dentro da flora portuguesa. *Boletim da Sociedade Broteriana*, 19, 209-221.

Thirgood, J.V. (1981) *Man and the mediterranean forest*. Academic Press, London.

Travis, S.E., Ritland, K., Whitham, T.G., & Keim, P. (1998) A genetic linkage map of Pinyon pine (*Pinus edulis*) based on amplified fragment length polymorphisms. *Theoretical and Applied Genetics*, 97, 871-880.

Vekemans, X., & Jacquemart, A.L. (1997) Perspectives on the use of molecular markers in plant population biology. *Belgian Journal of Botany*, 129, 91-100.

Vendramin, G.G., Anzidei, M., Madaghiele, A., & Bucci, G. (1998) Distribution of genetic diversity in *Pinus pinaster* Ait. as revealed by chloroplast microsatellites. *Theoretical and Applied Genetics*, 97, 456-463.

Vendramin, G.G., Lelli, L., Rossi, P., & Morgante, M. (1996) A set of primers for the amplification of 20 chloroplast microsatellites in *Pinaceae*. *Molecular Ecology*, 5, 595-598.

Vendramin, G.G., & Ziegenhagen, B. (1997) Characterisation and inheritance of polymorphic plastid microsatellites in *Abies*. *Genome*, 40, 857-864.

Vos, P., Hogers, R., Bleeker, M., Reijans, M., Vandelee, T., Hornes, M., Frijters, A., Pot, J., Peleman, J., Kuiper, M., & Zabeau, M. (1995) AFLP: A new technique for DNA fingerprinting. *Nucleic Acids Research*, 23, 4407-4414.

Wade, M.J., Mcknight, M.L., & Shaffer, H.B. (1994) The effects of kin-structured colonization on nuclear and cytoplasmic genetic diversity. *Evolution*, 48, 1114-1120.

Wagner, D.B., Govindaraju, D.R., Yeatman, C.W., & Pitel, J.A. (1989) Paternal chloroplast DNA inheritance in a diallel cross of Jack pine (*Pinus banksiana* Lamb.). *Journal of Heredity*, 80, 483-485.

Wakasugi, T., Tsudzuki, S.I., Shibata, M., & Sugiura, M. (1994) A physical map and clone bank of the black pine (*Pinus thunbergii*) chloroplast genome. *Plant Molecular Biology Reporter*, 12, 227-241.

Wang, X.R., Tsumura, Y., Yoshimaru, H., Nagasaka, K., & Szmidt, A.E. (1999) Phylogenetic relationships of Eurasian pines (*Pinus*, *Pinaceae*) based on chloroplast *rbc*L, *mat*K, *rpl*20-*rps*18 spacer, and *trn*V intron sequences. *American Journal of Botany*, 86, 1742-1753.

Wang, X.-R., & Szmidt, A.E. (1993) Chloroplast DNA-based phylogeny of Asian *Pinus* species (*Pinaceae*). *Plant Systematics and Evolution*, 188, 197-211.

Wang, X.-R., & Szmidt, A.E. (1994) Hybridization and chloroplast DNA variation in a *Pinus* species complex from Asia. *Evolution*, 48, 1020-1031.

Wang, X.-R., & Szmidt, A.E. (2000) Molecular markers in population genetics of forest trees. *Scandinavian Journal of Forest Research*, in press.

Wang, X.-R., Szmidt, A.E., & Nguyen, N.-H. (2000) The phylogenetic position of the endemic flat-needle pine *Pinus krempfii* (*Pinaceae*) from Vietnam, based on PCR-RFLP analysis of chloroplast DNA. *Plant Systematics and Evolution*, 220, 21-36.

Weir, B.S. (1996) *Genetic data analysis II*. Sinauer Associates Inc. Publ., Sunderland, Mass.

Welsh, J., & McClelland, M. (1990) Fingerprinting genomes using PCR with arbitrary primers. *Nucleic Acids Research*, 18, 7213-7218.

Whitton, J., Rieseberg, L.H., & Ungerer, M.C. (1997) Microsatellite loci are not conserved across the *Asteraceae*. *Molecular Biology and Evolution*, 14, 204-209.

Williams, J.G.K., Kubelik, A.R., Livak, K.J., Rafalski, J.A., & Tingey, S.V. (1990) DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. *Nucleic Acids Research*, 18, 6531-6535.

Willis, K.J. (1992) The late Quaternary vegetational history of northwest Greece. III. A comparative study of two contrasting sites. *New Phytologist*, 121, 139-155.

Winfield, M.O., Arnold, G.M., Cooper, F., Leray, M., White, J., Karp, A., & Edwards, K.J. (1998) A study of genetic diversity in *Populus nigra* subsp. *betulifolia* in the Upper Severn area of the UK using AFLP markers. *Molecular Ecology*, 7, 3-10.

Wolfe, K.H., Li, W.-H., & Sharp, P.M. (1987) Rates of nucleotide substitution vary greatly among plant mitochondrial, chloroplast, and nuclear DNAs. *Proceedings of the National Academy of Sciences of USA*, 84, 9054-9058.

Wood, B., & Turner, A. (1995) Out of Africa and into Asia. *Nature*, 378, 239-240.

Wright, S. (1931) Evolution in Mendelian populations. *Genetics*, 16, 97-159.

Wu, J., Krutovskii, K., & Strauss, S.H. (1999) Nuclear DNA diversity, population differentiation and phylogenetic relationships in the California closed-cone pines based on RAPD and allozyme markers. *Genome*, 42, 893-908.

Zabeau, M., & Vos, P. (1993) Selective restriction fragment amplification: a general method for DNA fingerprinting. In: *European Patent Application No. 92402629.7.*

Zhivotovsky, L.A. (1999) Estimating population structure in diploids with multilocus dominant DNA markers. *Molecular Ecology*, 8, 907-913.

Zhu, J., Gale, M.D., Quarrie, S., Jackson, M.T., & Bryan, G.J. (1998) AFLP markers for the study of rice biodiversity. *Theoretical and Applied Genetics*, 96, 602-611.

Ziegenhagen, B., Kormutak, A., Schauerte, M., & Scholz, F. (1995) Restriction site polymorphism in chloroplast DNA of Silver Fir (*Abies alba* Mill.). *Forest Genetics*, 2, 99-107.

Zobel, B.J., & Taberlet, J.T. (1984) *Applied Forest Tree Improvement*. Wiley, New York, NY.

Acknowledgments

I am in debt to Aquilino Ribeiro and to his sharp and courageous book "When the wolves howl" ("Quando os lobos uivam"), that called my attention, for the first time, to the large-scale impact of compulsory forestation after 1938 in the communitarian areas of the small villages in the Centre-North of my country, and the strong change produced onwards in the social tissue, a story yet to be told[1]…and the winner is: maritime pine! Over the range of the species, Portugal possesses the highest ratio of *P. pinaster* surface compared with the surface of the country! Men and forests do interact! I read that book for the first time during some of the (beautiful) train trips in between my hometown and Lisbon, where I started to study Forestry. I just had to come to that book once again, and to try to "find" what was so special with that tree that made people cross the borders and seek for a new place[2]. I am also an example of that, after all!

I am in debt to many, many people… I will mention my supervisor, Alfred, who challenged me to start this adventure 4 000 km from home (if home is anywhere!). Thanks to my co-supervisors: Christophe (aussi ta petite famille, sans oublier les moules, les frites et le merveilleux vin rouge des Graves!) and Nica (obrigado pelo apoio, amizade e encorajamento). To Nicole, for her precious help in framing historical and ecological events related with *P. pinaster* in Portugal. To God (Antoine), for replying so gently to all of my (too) frequent questions! To my collaborators and friends: Laura, Stéfi, Santi, Sophie (et Fred!), Christian, David and Beppe. To all the staff in Pierroton, who made my stay there so profitable and nice. A special memory to Paulo Costa. To Gertrude, for her friendship and for taking care, always! To Maud, for lightening some dark paths in the way. To Tim, for his close friendship and help in editing my Summary. To Zé, my brother in mind and love, for keeping in touch all the time. To John, for editing my papers. A special thanks to Jarmo, who was so helpful with the impossible-to-do figures of the thesis. To the Mediterranean wing of Umeå: Leo & Laurence, Estelle & Stephan (sorry, you became also a bit Med!), Andreas, Dimitris, Grazia, Angela, Vicent, Julien, Ana & Alberto. To the other wing: Robert & Carina, Emilie and her lovely family, Per & Nemo (Greta & Svea), Erik & Margareta, Ivone, Agneta, David, Adolfo, Kyu-Suk, Monika, Madoka, Maria & Magnus, Xiao-Ru, Vikram, Wan, Toming, Ewa and all the staff from the Department, for being so kind and helpful all the time!

[1] It is worthwhile to read J.A. Estevão (1983) A florestação fos baldios, Análise Social: 1157-1260.

[2] Acording to J.Serrão (1977, A Emigração Portuguesa, Lisboa) until the beginning of the 50's the emigration in Portugal was negligible.

To my mother, who gave me (at least half) of her genes and a lot of support; whenever my ideas turned out to be too odd (like coming to the North of Sweden) she always understood! To my sisters, Cris and Vica, for being there all the time on the other side of the line, and to Icha and Rui for being so patient with them! To Xani, André and Joana, the best grown up kids in the world. To my daughter Filipa for having to spend four years with a far-away mother! To my father, who called my attention to science by showing himself so enthusiastic when the first man stepped on the Moon, that was also my big first contact with a New World!

To all the colleagues of the ESACB and of my Department in particular, that made possible my coming to Sweden, and for believing that I would be back!

This thesis would never be possible without grants from the PRODEP II Acção 5.2, the KSLA/Carl Fredrik von Horns Foundation, the SLU (Swedish University of Agricultural Sciences), the INRA/SJFR 2000 Cooperation in Agricultural Research, the SJFR (Swedish Council for Forestry and Agricultural Research), the SIDA (Swedish International Development Cooperation Agency), the European Union (IC18970200 and INCO), the BRG (Bureau des Ressources Génétiques), the INRA (Institut National de la Recherche Agronomique), and GIS (Programme Coopératif Pin maritime du futur).