Universidade de Trás-os-Montes e Alto Douro

Genetic diversity, provenances extrapolation and cytogenetics analysis of Portuguese *Pinus nigra* Arn. populations

Dissertação de Mestrado em
Genética Molecular Comparativa e Tecnológica

Maria Franco Rosa Costa de Lemos

Orientador: Professor Doutor José Eduardo Lima-Brito
Co-Orientador: Professora Doutora Maria João Magalhães Gaspar

Vila Real, 2016

Universidade de Trás-os-Montes e Alto Douro

Genetic diversity, provenances extrapolation and cytogenetics analysis of Portuguese *Pinus nigra* Arn. populations

Dissertação de Mestrado em
Genética Molecular Comparativa e Tecnológica

Maria Franco Rosa Costa de Lemos

Orientador: Professor Doutor José Eduardo Lima-Brito
Co-Orientador: Professora Doutora Maria João Magalhães Gaspar

Composição do Júri:

_____

_____

_____

Vila Real, 2016

This original research was developed to achieve the Master Degree in Molecular Comparative and Technological Genetics (Decreto-Lei nº 74/2006 with the modifications introduced by Decreto-Lei nº 107/2008 and Decreto-Lei nº 230/2009)

This dissertation is dedicated
to those who never stopped
believing in me.

Acknowledgments

This work was developed in the Plant Cytogenomics Lab of Centre of Genomics and Biotechnology (IBB/CGB) at University of Tras-os-Montes and Alto Douro (UTAD). This would not be possible without the help and support of several people and institutions, which I want to sincerely thank:

To the Rector of the UTAD, Professor António Fontainhas Fernandes, Ph.D., the possibility of the realization of my dissertation at this University.

To the Direction of the Master's degree in Technologic, Comparative and Molecular Genetics, for the accepting of this dissertation.

To the Director of the Master's degree in Technologic, Comparative and Molecular Genetics, Professor Paula Filomena Martins Lopes, Ph.D, for the availability to any clarification.

To the Professor José Eduardo Lima-Brito, Ph.D, head of the Plant Cytogenomics Lab (IBB-CGB/UTAD), for given me the opportunity to conduct the research in this Lab.

To Dr. Marc Calvignac from Vilmorin S.A., and to Professor Bruno Fady, Ph. D, for kindly given certified seed material and dehydrated needles material of *P. nigra*, respectively.

To all my teachers, for the transmitted knowledge which gave me bases for the performance and interpretation of this work.

To Professor José Eduardo Lima-Brito, Ph.D, my supervisor and to Professor Maria João Magalhães Gaspar, Ph.D, my co-supervisor, for the orientation of this dissertation, the teachings; the availability for any clarification and for the critical review of the work.

To Dr. Ana Isabel Ferreira de Carvalho, Ph.D, for all the help and teachings, for all time spent in corrections and clarifications, and for the friendship and moments of distraction inside and outside the laboratory.

To Alexandra Dias, MSc., Jani Pires, Eng., Luis Rocha, MSc., João Coutinho, Eng., for the help and the co-work in laboratory.

To Ivo Pavia, MSc., for the help, all advices and friendship.

To João Freitas, for the friendship, patience and comprehension.

To Andreia Delgado, Dr., for the help in this dissertation, for being always present and available to me, even when there is no time for anything; for the friendship in good and bad

moments during all the years of this journey that has been the university; and for believing always in me and in my capacities.

To my family, for all the support, love and affection.

To all that contribute to this dissertation, but that weren't referred, *thank you so much.*

Diversidade genética, extrapolação de proveniências e análise citogenética de populações Portuguesas de *Pinus nigra* Arn.

Resumo

Durante o século XX instalaram-se povoamentos de *Pinus nigra* em Portugal. Actualmente, o número de populações Portuguesas de *P. nigra* alóctones é restrito estando distribuídas por uma área reduzida. A sua origem e taxonomia são desconhecidas. A caracterização botânica destes povoamentos em 1980 revelou a presença das subespécies *nigra*, *salzmannii* e *laricio*.

Neste estudo pretendeu-se: i) caracterizar molecularmente estas populações alóctones; ii) extrapolar as suas proveniências; iii) determinar a sua taxonomia; iv) caracterizar amostras estrangeiras de diferentes *taxa* infraespecíficos e proveniências para comparar com as Portuguesas, utilizando marcadores ISSR e SCoT. Portugal constitui o limite mais ocidental de distribuição de *P.nigra* e *Pinus sylvestris* L.. Neste estudo objectivou-se ainda a caracterização citogenética de uma população de *P.nigra* (Campeã) e sua comparação com pinheiro-silvestre (população autóctone da Biduiça) usando as técnicas citogenéticas de FISH com uma sonda de rDNA, de ND-FISH com sondas SSR e de coloração com nitrato de prata.

O estudo de diversidade genética requereu a prévia extração de DNA de *P.nigra.* Devido ao elevado conteúdo de polissacáridos e polifenóis, o protocolo de extração foi optimizado para DNA de agulhas congeladas e desidratadas, sendo extensível a outros tecidos e espécies.

A caracterização molecular e extrapolação das proveniências das populações de *P.nigra* (Manteigas, Caminha, Vila Pouca de Aguiar, Vale do Zêzere, Paredes de Coura e Campeã) foram efectuadas em 127 indivíduos com oito *primers* para amplificação de ISSRs e oito *primers* para produção de SCoTs. Foram detetadas uma elevada diversidade genética intra-populacional (expectável para uma espécie *Pinus*) e uma estrutura genética definida (agrupamento dos indivíduos por população).

Os perfis genéticos das amostras estrangeiras foram determinados em *bulks* correspondentes a 120 indivíduos de 12 proveniências, representativos das subespécies (*nigra*, *salzmannii* e *laricio*) e das variedades (*corsicana*, *calabrica* e *austriaca*) usando ISSRs e SCoTs.

A subespécie *salzmannii* revelou-se geneticamente distante das amostras Portuguesas e restantes estrangeiras. Uma elevada identidade genética foi obtida entre amostras Portuguesas e estrangeiras da subespécie *laricio*. Os ISSRs evidenciaram maior similaridade genética entre os indivíduos Portugueses e amostras da subsp. *laricio* var. *calabrica* enquanto os SCoTs detetaram maior similaridade com a subespécie *laricio* var. *corsicana*. Segundo a caracterização botânica de 1980, esta subespécie foi apontada como a melhor adaptada e predominante em Portugal. Como a amplificação de SCoTs se baseia na ligação do *primer* a uma curta sequência conservada nos genes, consideraram-se estes resultados mais fidedignos. Globalmente, os dois sistemas marcadores demonstraram especificidade para determinar estrutura genética, extrapolar proveniências e discriminar *taxa* infraespecíficos em *P.nigra*.

Após ND-FISH observaram-se diferenças na hibridação de SSRs entre *P.nigra* e *P.sylvestris*. A sonda rDNA detetou 14 e 16 *loci* em *P.sylvestris* e *P.nigra*, respetivamente. O número de nucléolos por núcleo (1 a 12) em ambas as espécies revela que nem todos os *loci* rDNA estão transcripcionalmente activos. A presença de micronucléolos associada à ocorrência de irregularidades cromossómicas (e.g. cromossomas em anel, policêntricos e fragmentados) sugere que estas populações periféricas ainda estão em adaptação.

Os resultados deste trabalho contribuirão para o delineamento de estratégias de plantação/reflorestação de *P. nigra* em Portugal.

Palavras-chave: Citogenética; diversidade e estrutura genética, FISH, marcadores moleculares; *Pinus nigra* Arn.; taxonomia.

Abstract

During the 20th century, *P.nigra* stands were installed in Portugal. Currently, the number of Portuguese *P. nigra* allochthonous populations is restricted being distributed to a small area. Origin and taxonomy of material used are unknown. The botanic characterization of these stands in 1980 revealed the presence of the subspecies *nigra*, *salzmannii* e *laricio*.

In this study, we intend to: i) characterize molecularly these allochthonous populations; ii) extrapolate their provenances; iii) determine their taxonomy; and iv) characterize foreign samples from different infraspecifics *taxa* and provenances to compare with Portuguese, using ISSR and SCoT markers. Portugal constitutes the limit more occidental of *P.nigra* and *P.sylvestris* distribution. In this study, it was also aimed the cytogenetic characterization of a *P.nigra* population ('Campeã') and their comparison with Scots pine (autochthonous population of 'Biduiça') using the cytogenetic techniques FISH with a ribosomal rDNA probe, ND-FISH with SSR probes and silver nitrate staining.

The genetic diversity study required a previous extraction of *P.nigra* DNA. Due to the high content of polysaccharides and polyphenols, the extraction protocol was optimized for DNA from frozen and dehydrated needles, and was suitable to other tissues and species.

The molecular characterization and provenances extrapolation of *P.nigra* populations ('Manteigas', 'Caminha', 'Vila Pouca de Aguiar', 'Vale do Zêzere', 'Paredes de Coura', 'Campeã') were performed in 127 individuals with eight primers to ISSRs amplification and eight for SCoTs production. A high genetic diversity within populations (expectable for a *Pinus* species) and a defined genetic structure (clustering of individuals per population) were detected.

The genetic profiles of foreign samples were determined in bulks corresponding to 120 individuals of 12 provenances representatives of subspecies (*nigra*, *salzmannii* and *laricio*) and varieties (*corsicana*, *calabrica* and *austriaca*) using ISSRs e SCoTs.

The subspecies *salzmannii* revealed to be genetically distant from the Portuguese and other foreign samples. A high genetic identity was obtained among Portuguese and foreign samples from *laricio* subspecies. ISSRs evidenced higher genetic identity among the Portuguese and foreign samples from subsp. *laricio* var. *calabrica* whereas SCoTs estimated a higher genetic identity with subsp. *laricio* var. *corsicana*. According to botanic characterization of 1980, this subspecies was considered the best adapted and prevalent in Portugal. As the SCoTs amplification is based on the binding of primer to a short conserved

sequence in the genes, was considered these results more reliable. Globally, the two marker systems showed specificity for genetic structure determination, provenances extrapolation and infraspecific *taxa* discrimination of *P.nigra*.

After ND-FISH, differences in SSRs hybridization between *P.nigra* and *P.sylvestris* were observed. rDNA probe detected 14 and 16 *loci* in *P.sylvestris* and *P.nigra*, respectively. The number of nucleoli per interphase nucleus (1 to 12) in both species revealed that not all rDNA *loci* were transcriptionally active. The presence of micronucleoli associated to chromosomic irregularities (e.g. ring, polycentric and fragmented chromosomes) in both species, suggested that these peripheral populations are still in adaptation.

The results of this work will contribute to design strategies for planting/reforestation of *P.nigra* in Portugal.

Keywords: Cytogenetics; FISH; genetic diversity and structure; molecular makers; *Pinus nigra* Arn.; taxonomy.

Index

Index of Figures

Chapter V:

Index of Tables

Chapter V:

List of Abbreviations

4xT – Detection buffer composed by 4xSSC and 0.05% Tween-20
A – Absorbance
AADs – Arbitrary Amplified DNA markers
AFLP – Amplified Fragment Length Polymorphism
Ait. – Aiton
AMOVA – Analysis of Molecular Variance
Arn. – Arnold
bp – base pair
BSA – Bovine Serum Albumin
C – Nuclear DNA content (C-value)
CDDP – Conserved DNA-Derived Polymorphism
CDMs – Conserved DNA and Gene Family Based markers
cDNA – complementary DNA
cDNA-AFLP – complementary DNA - Amplified Fragment Length Polymorphism
cDNA-RFLP – complementary DNA - Restriction Fragment Length Polymorphism
cDNA-SCoT – complementary DNA - Start Codon Targeted Polymorphism
CGB – Centre of Genomics and Biotechnology
CIA – Chloroform: Isoamyl alcohol 24:1
CoRAP – Conserved Region Amplification Polymorphism
cpDNA – chloroplastidial DNA
cpSSRs – chloroplastidial Simple Sequence Repeats
CTAB – Cetyltrimethylammonium bromide
DALP – Direct Amplification of Length Polymorphisms
DAPI – 4',6-Diamidino-2-Phenylindole, Dihydrochloride
DAMD-PCR – Directed Amplification of Minisatellite-region DNA polymerase chain reaction
DBH – Diameter at the Breast Height
DNA – Deoxyribonucleic acid
dNTP – deoxynucleotide triphosphate solution mix
dUTP – Deoxyuridine triphosphate
EDTA – Ethylenediaminetetraacetic acid
EST-SSR – Expressed Sequence Tag-Simple Sequence Repeat
E.U. – European Union
FCT – *Fundação para a Ciência e a Tecnologia*
FEDER – European Fund of Regional Development
FISH – Fluorescence *In Situ* Hybridization
FRM – Forest Reproductive Material
FM – Functional Markers
Gbp – Giga base pair
GTM – Gene-Targeted Markers
GRIN – Germplasm Resources Information Network
I – Shannon's information index
iPBS – inter-Primer Binding Site
IRAP – Inter-Retrotransposon Amplified Polymorphism
ISAP – Inter-SINE Amplified Polymorphism
ISH – *in situ* hybridization

iSNAP – inter Small RNA Polymorphism
ISSR – Inter-Simple Sequence Repeat
ITP – Intron-Targeting Polymorphism
ITS – Internal Transcribed Spacer
kb – kilobase
MAS – Markers Assisted Selection
MCMC – Markov Chain Monte Carlo
Mya – Million years ago
NBS – Nucleotide Binding Site profiling
ND-FISH – Nondenaturating FISH
NTSYS – Numerical Taxonomy System
nSSRs – nuclear microsatellites
ON – overnight
PAAP – Promotor Anchored Amplified Polymorphism
PBA – cytochrome P450 Based Analogues
PCoA – Principal Coordinates Analysis
PCR – Polymerase Chain Reaction
PIC – Polymorphic Information Content
PVP-40 – Polyvinylpyrrollidone; molecular weight 40,000
QTL – Quantitative Trait *Loci*
r – correlation coefficient
RAPD – Random Amplified Polymorphic DNA
RBMs – RNA-Based Markers
rDNA – ribosomal DNA
REMAP – Retrotransposon-Microsatellite Amplified Polymorphism
RFLP – Restriction Fragment Length Polymorphism
RFLP cpDNA – Restriction Fragment Length Polymorphism of chloroplast DNA
RGAP – Resistance-Gene Analog Polymorphism
RGMs – Resistance-Gene based Markers
RT – Room Temperature
SAHN – Sequential Agglomerative Hierarchical Nested
SCoT – Start Codon Targeted
SDS – Sodium Dodecyl Sulfate
SM – Simple Matching
SNP – Single Nucleotide Polymorphism
S-RAP – Sequence-Related Amplified Polymorphism
S-SAP – Sequence-Specific Amplification Polymorphism
SSC – Saline Sodium Citrate buffer
SSR – Simple Sequence Repeat
Subsp. – Subspecies
Taq* –Thermus aquaticus* (DNA polymerase)
TBE – Tris-Borate-EDTA
TBP – Tubulin Based Polymorphism
TE – Tris-EDTA buffer solution
TFM – Targeted Fingerprinting Markers
TMEs – Transposable Element Based markers
TRAP – Targeted Region Amplified Polymorphism
UBC – University of British Columbia
UPGMA – Unweighted Pair Group Method with Arithmetic means

USDA – United States Department of Agriculture
UTAD – University of Tras-os-Montes and Alto Douro

Layout of this dissertation

This dissertation is constituted by six chapters: Chapter I – General introduction and Objectives; Chapter II to Chapter V corresponding to scientific articles (see below) submitted to international journals with peer review and belonging to the Scientific International Index; and Chapter VI to the Conclusions and perspectives of the developed research.

Lemos M, Dias A, Carvalho A, Delgado A, Pavia I, Coutinho J, Gaspar MJ, Coutinho J, Louzada JL, Lima-Brito J (2015) Single protocol for genomic DNA extraction from foliar, dry and vascular tissues of gymnosperms suitable for genetic studies. Paper Submitted to *Wood Science and Technology*. (Chapter II)

Dias A, Lemos M, Carvalho A, Fady Bruno, Louzada JL, Gaspar MJ, Lima-Brito J (2015) Genetic profiles of European allopatric subspecies and varieties of Pinus nigra Arn. based on ISSR and SCoT markers. Paper Submitted to *Plant Systematics and Evolution*. (Chapter III)

Lemos M, Dias A, Carvalho A, Gaspar MJ, Coutinho J, Louzada JL, Lima-Brito J (2015) First molecular characterization of the Portuguese populations of *Pinus nigra* Arnold and extrapolation of their provenances. Paper Submitted to *Plant Molecular Biology Reporter*. (Chapter IV)

Lemos M, Carvalho A, Pavia I, Delgado A, Gaspar MJ, Lima-Brito J (2015) Nucleolar activity, SSRs hybridization and chromosome irregularities in peripheral populations of *Pinus nigra* Arn. and *Pinus sylvestris* L. Paper Submitted to *Plant Cell Reports*. (Chapter V)

Chapter I:

General Introduction and objectives

I.1 – The genus *Pinus*

Pines are the most important group of conifers (Vidaković 1991; Vargas-Mendoza *et al.* 2011) and belong to the genus *Pinus* (family Pinaceae).

Pinus is one of the largest existent gymnosperm genus, including more than 100 species distributed across the Northern Hemisphere (Mirov 1967; Vidaković 1991). Pine species typically inhabit the boreal, temperate and mountainous tropical regions (Vidaković 1991). According to Gernandt *et al*. (2005), the genus *Pinus* can be divided into two subgenera *Pinus* and *Strobus*, each one with two sections and several subsections. The subgenus *Pinus* includes the sections *Pinus* and *Trifoliae* while the subgenus *Strobus* is composed by the *Parrya* and *Quinquefoliae* sections (Gernandt *et al*. 2005) (Table I.1). Nonetheless, the taxonomy of the genus *Pinus* is still incomplete for many *taxa* (Nkongolo *et al.* 2002).

Table I.1 – Infrageneric taxonomical classification of *Pinus* accordingly Gernandt *et al.* (2005). (Adapted from Wang *et al.* 2013).

Subgenus	*Section*	*Subsection*
Pinus	*Pinus*	*Pinus*
		Pinaster
	Trifoliae	*Contortae*
		Australes
		Ponderosae
Strobus	*Quinquefoliae*	*Gerardianae*
		Krempfianae
		Strobus
	Parrya	*Nelsoniae*
		Balfourianae
		Cembroides

Pinus species have been widely planted over the world (Richardson 1998a,b; Zlatanov *et al.* 2010), and since the 19th century they have been widely used for reforestation and land restoration (Zlatanov *et al.* 2010). Furthermore, their ecological and forest roles have been

widely studied in their native areas but also over the distribution range of their introduction (Richardson 1998a,b).

Some authors pointed out the existence of different pine species in Portugal, namely as mountain pines, *Pinus sylvestris* L., *Pinus nigra* Arn., *Pinus uncinata* Ramond ex DC and most predominantly, *Pinus pinaster* Ait. (Figueiral 1995; Figueiral and Carcaillet 2005).

In this dissertation was studied Portuguese *P. nigra* Arn. populations using both molecular and cytogenetic approaches.

I.2 – *Pinus nigra* species

I.2.1 – General characteristics and distribution

Pinus nigra Arnold (European Black Pine) is a conifer of the family Pinaceae (Mirov 1967; Price *et al.* 1998; Nkongolo and Mehes-Smith 2012) that is considered as a relict species (Mirov 1967; Vidaković 1991).

P. nigra populations have been subjected to different glacial and interglacial episodes of climatic change and geological events (Thompson 2005) which coupled in part with anthropogenic disturbance contributed for the actual discontinuous and fragmented distribution of the European Black Pine (Ozenda 1975, 1985; Dobrinov 1983; Isajev *et al.* 2004; Naydenov *et al.* 2006; Esteban *et al.* 2012).

Despite being native of Europe and Asia (Rezzi *et al.* 2001), *P. nigra* is widespread through an area of 230,000 ha (Tolun *et al.* 2000; Gerber *et al.* 2005; Afzal-Raffi and Dodd 2007) that extends from North Africa through the Northern Mediterranean, from eastwards to the Black Sea and also can be found on the islands of Corsica and Sicily (Lee 1968; Afzal-Raffi and Dodd 2007) (Fig. I.1). *P. nigra* can be found in a wide range of environmental conditions (Naydenov *et al.* 2006), being tolerant to poor soils and supporting a wide range of climates across its geographical range (Lee 1968; Arbez and Miller 1971; Debazac 1971). Moreover, *P. nigra* is a light-demanding species (Vidaković 1991; Isajev *et al.* 2004), intolerant to shade but resistant to wind and drought (Isajev *et al.* 2004).

This species is usually found at altitudes ranging from 250 to 1,400m (Vidaković 1991) but its optimal range is between 800 and 1500m (Isajev *et al.* 2004). Generally, it can be found at high mountains zones (Specht *et al.* 1988), on low elevation mountains around the Mediterranean basin (Naydenov *et al.* 2006) or at the sea level along the shores of the

Adriatic Sea (Lee 1968). Unfortunately, its habitat in low elevation forests has been highly affected by human habitation for millennia (Ozenda 1975, 1985; Dobrinov 1983).

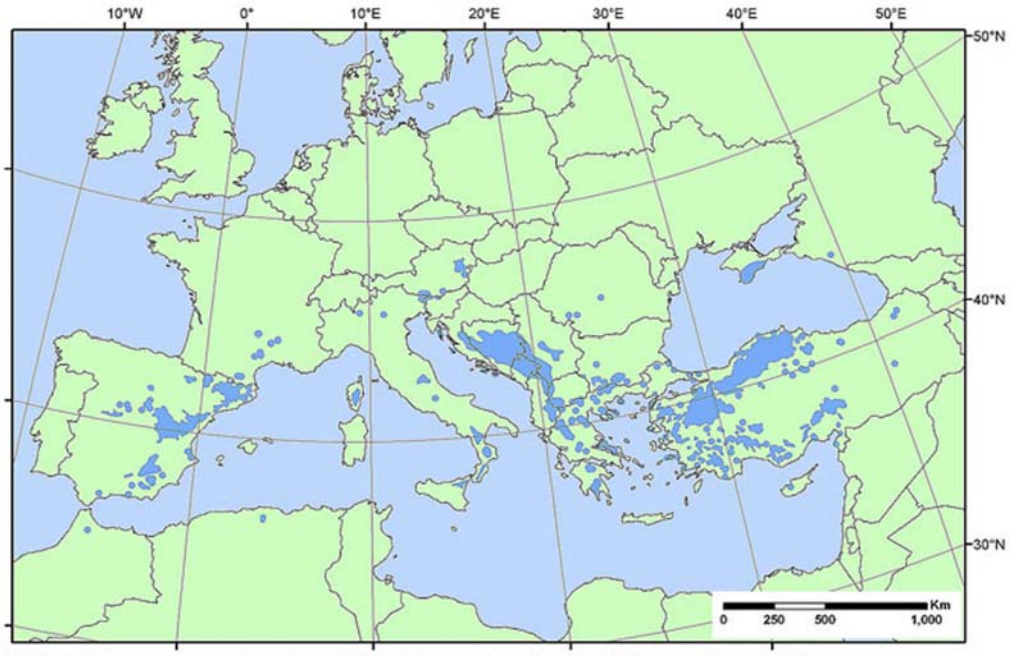

Figure I.1 – Global distribution of *P. nigra*. (Adapted from EUFORGEN 2009).

European Black Pine has a fast early growth, reaching up to 30m (rarely 40-50m) of height and its trunk is usually straight (Isajev *et al.* 2004). The bark color ranges from light grey to dark grey-brown (Pedro 1993; Isajev *et al.* 2004). The sexual maturity of European Black Pine is achieved after 15-20 years in natural habitat.

The natural and fragmented distribution of *P. nigra* contributed for the high variability of morphological, anatomical and physiological traits (Scaltsoyannes *et al.* 1994) among different regions. This high variability, also at genetic level, allowed the attribution of different taxonomic names which resulted on an infraspecific taxonomic classification that is still debated after several years. Thus, several authors proposed the occurrence of allopatric subspecies, varieties and/or races of *P. nigra* through the range of its natural distribution. Consequently, three groups could be assigned: i) the Eastern group, including the subspecies *pallasiana* (Lamb.) Holmboe; ii) the Central group composed by the subspecies *nigra*, subspecies *laricio* (Poiret) Maire and subspecies *dalmatica* (Vis.) Franco; iii) the Western group only represented by the subspecies *salzmannii* (Dunal) Franco (Esteban *et al.* 2012).

I.2.2 – Infraspecific taxonomic classification

Due to its widespread distribution, *P. nigra* is also named as Austrian Pine, Australian Pine, Corsican Pine, Crimean Pine or Pyrenees Pine (Isajev *et al*. 2004; Afzal-Raffi and Dodd 2007). More than one hundred Latin synonymous names were given to this species (Wright and Bull 1962; Akkemik *et al.* 2010) and 81 races and geographic varieties were ascribed to this species (Wheeler *et al.* 1976). The excessive number of described scientific names for *P. nigra* is due to the very narrow rank attribution for each *taxon* (species, subspecies or variety), being often issue of discussion. Besides, there was never concordance about the limits and names of varieties or geographical races (Wheeler *et al*. 1976) and some authors reported the difficulty in distinguish between subspecies and varieties (Bonnet-Masimbert and Bikay-Bikay 1978). Additionally, not all of the different scientific names are accepted by the International Code of Nomenclature for *Algae*, *Fungi* and Plants.

P. nigra has a complex botanic classification and some authors considered it as a collective species or an aggregate of microgeographic species due to the high morphological, physiological and ecological variability found among different geographical locations (Villar 1947; Gaussen 1960; Arbez and Millier 1971; Pajares and Escudero 1989; Blanco 1998; Richardson 1998b). However, other authors have proposed its botanic division (Schwarz 1938; Delevoy 1949; Mirov 1967). Since then, several revisions of its infraspecific taxonomic classification have been proposed resulting on a division of *P. nigra* into two to nine subspecies and respective allopatric varieties (e.g. Delevoy 1949; Debazac 1963; 1964; 1965; 1971; Gaussen *et al*. 1964; Mirov 1967; Vidaković 1974; 1991; Farjon 1984; Gaussen and Heywood 1993; Christensen 1997; Barbéro *et al.* 1998; Price *et al.* 1998; Tolun *et al.* 2000). Nonetheless, the infraspecific genetic structure of *P. nigra* still is not entirely known over its whole area of natural distribution (Naydenov *et al.* 2006).

This taxonomic problem increased with the existence of transitional forms resultant from cross hybridization between subdivisions of *P. nigra* and from efficient gene flow (Vidaković 1974; 1991). In addition to the incongruency about the number and designation of the *P. nigra* subspecies and/ or varieties, there is not also a complete concordance about their geographical distribution (Delevoy 1949; Wright and Bull 1962; Gellini 1968; Lee 1968; Arbez and Millier 1971, Afzal-Rafii and Dodd 2007). Additionally, the geographic or allopatric groups are inter-fertile (Vidaković 1991; Isajev *et al.* 2004), probably due to their common phylogenetic origin, increasing the classification issues (Delevoy 1949; Wright and

Bull 1962; Gellini 1968; Lee 1968; Arbez and Millier 1971; Isajev *et al.* 2004; Afzal-Rafii and Dodd 2007).

The infraspecific taxonomic classification of *P. nigra* has been debated since long time ago and remains unclear probably due to the: i) absence of a study covering the whole distribution range of the species; ii) impossibility of integrating results of certain areas; iii) high variability of traits which have been used so far (Liber *et al*. 2003). Presently, its infraspecific taxonomic classification is still not unanimously accepted (Scaltsoyiannes *et al.* 1994; Naydenov *et al.* 2006; Afzal-Raffi and Dodd 2007; Bogunić *et al.* 2007; del Cerro Barja *et al.* 2009). Nevertheless, most of the infraspecific taxonomic classifications proposed so far, considered five *P. nigra* subspecies, namely, *dalmatica*, *laricio*, *nigra*, *pallasiana* and *salzmannii* (Gaussen *et al.* 1964; 1993; Vidakovic 1974; Farjon 1998; 2010; 2013; Price *et al.* 1998). Vidakovic (1974) agreed with the classification of Gaussen *et al*. (1964) published in the Flora Europaea, but reported the existence of the subspecies *salzmannii* in the North Africa. This modification as well as the consideration of five *P. nigra* subspecies were maintained by Gaussen *et al.* (1993) and Farjon (2013). Although some studies were in agreement with the classification used (Scaltsoyiannes *et al.* 1994; Gerber *et al.* 1995), others presented some inconsistencies (Delevoy 1949; Arbez and Miller 1971; Liber *et al.* 2002; 2003; Giovannelli *et al.* 2014). Globally, these studies revealed high variation among the studied *P. nigra* populations that could be favored by the widespread and fragmented distribution of the species, mostly in Europe. The high number of common and scientific names attributed to *P. nigra*, corroborate the idea that this species is be extremely variable. However, such degree of variation is not uncommon if compared with other conifers widely distributed or with the degree of genetic diversity of other *Pinus* species (Scaltsoyiannes *et al.* 1994).

Considering all the reasons presented above, the taxonomic classification of *P. nigra* needs to be revised, as proposed by Giovannelli *et al.* (2014). These authors tried to study molecularly the maximum possible number of *P. nigra* subspecies considered taxonomically diverse in order to solve their phylogeny. After amplification and alignment of the sequences of the barcoding genes *Cox1*, *MatK*, *Rbcl*, *Trnh* and *Nad 5-4*, no differences were detected among subspecies (Giovannelli *et al*. 2014). Otherwise, the use of nuclear microsatellites (nSSRs), chloroplastidial SSRs (cpSSRs) and detection of Single Nucleotide Polymorphisms (SNPs) allowed the grouping of the subspecies *pallasiana* and *nigra* into the Eastern group, and the subspecies *mauretanica* and *salzmannii* in the Western group (Giovanelli *et al*. 2014).

The subspecies *laricio* included two varieties *corsicana* (related with the Western group) and *calabrica* (related to the Eastern group). However, the classification of the subspecies into three main groups also differed among authors. According to Lucas Borja *et al*. (2013), three core populations could be considered in the actuality: i) the Western group enclosing the subspecies *pallasiana* (Lamb.) Holmboe; ii) the central group *nigra* containing the subspecies *nigra*, *laricio* (Poiret) Maire and *dalmatica* (Vis) Franco; and iii) the Eastern Group composed by the subspecies *salzmannii*.

Over the years, the taxonomic classification of *P. nigra*, namely at infraspecific level, has been considered an unsolved question which still needs to be carefully analyzed based on other tools since its definition could be highly useful for the designing of conservation and reforestation strategies with economic and ecological importance.

In the present study, the infraspecific taxonomic classification of *P. nigra* proposed by Gaussen *et al.* (1964; 1993) in the Flora Europaea, which divided the Mediterranean *P. nigra* into five subspecies: *laricio*, *nigra*, *pallasiana*, *salzmannii* and *dalmatica* will be considered. In addition to these five Mediterranean subspecies, a sixth one, designed as *mauretanica,* was locally ascribed to the Northern Africa (Greuter *et al.* 1984; Barbéro *et al.* 1998).

I.2.3 – Economic and ecological importance

European Black Pine is one of the most economically important native conifers in Southern Europe (Isajev *et al.* 2004; Naydenov *et al.* 2006) due to their wood properties namely easy to process, durable and rich in resin (Espelta 1999; Isajev *et al.* 2004). Because of the variability among the *P. nigra* subspecies, their wood and sub-products have different applications. For example, subspecies *laricio* has straightness and thin branches, being appreciated for building and roofing, while the subspecies *nigra* has lower quality, being used to lower-grade building wood and the making of crates (Isajev *et al.* 2004). *P. nigra* have been used as Christmas tree, for fuel wood and poles (Isajev 2004). In many areas, it is one of the most important trees for timber production (Lee 1968). In the United States of America, *P. nigra* is mainly used for shelterbelts (Van Haverbeke 1990) and used as ornamental (Isajev *et al.* 2004). Moreover, *P. nigra* is also used with medicinal purposes (Tzulac and Erol 1999). The turpentine extracted from resin of *P. nigra* has been used due to their antiseptic effects on respiratory system and urinary diseases, for back pain, dermatologic, analgesic drug and as natural antioxidant (Gülçin *et al.* 2003). Additionally, the oil of *P. nigra* subspecies *laricio* revealed to be an efficient herbicidal (Amri *et al.* 2014).

This species is also ecologically important as carbon storage, protection against erosion and recreational activities, being able to colonize several ecological niches and evidencing its high adaptive capacity to different environments (Espelta 1999). Moreover, *P. nigra* plays a leading role in different stages of forestry ecological successions (Zaghi 2008).

Forest ecosystems of *P. nigra* are usually affected by crown fires. Furthermore, the regeneration after severe crown fire is reduced or almost null in this species; due to the fact of *P. nigra* neither produces serotinous cones nor maintain a seed bank. For that reason, *P. nigra* species represent a European priority on conservation (Christopoulou *et al.* 2013). The Mediterranean forests of *P. nigra* are included in the E.U. list of endangered natural habitats that require specific conservation measures (Resolution 4/1996 by the Convention on the Conservation of European Wildlife and Natural Habitats) (del Cerro Barja *et al.* 2009).

I.2.4 – Reforestation programs

In the last century, *P. nigra* was target of a high demand for plantation throughout its range of natural distribution. Due to its enormous adaptive plasticity, European Black Pine was the favorite species for reforestation projects over a wide range of environments, once this species has the ability to develop well on open lands and in ecologically demanding situations (Isajev *et al*. 2004).

Currently, *P. nigra* has been used in reforestation programs throughout its natural distribution (Naydenov *et al.* 2006) and represents an important pine species in forest management of Mediterranean areas (Bogunić *et al.* 2007). This is a result of its commercial interest, characteristics and adaptability to a wide range of environmental conditions. However, these planting strategies did not care about the genetic structure of the existent plantations and that of the new specimens. These facts could threaten the genetic diversity of *P. nigra* populations (Naydenov *et al.* 2006).

Some issues are very important to take into account during *P. nigra* plantations (both for conservation and reforestation programs), such as: global warning, site, altitude, slope, soil, forestry treatments (Sevgi and Akkemik 2007) and genetic structure of populations (Naydenov *et al.* 2006). This species is one of the most used for afforestation of arid and rocky terrains in the sub-Mediterranean region (Vidaković 1991) and it is widely planted in Central and Southern Europe in areas with poor soils (Luchi *et al.* 2005). *P. nigra* has been widely planted outside its natural range and has been showing promising results in New Zealand, Great Britain, France, Belgium, Argentina, United States of America and Portugal

(Lee 1968; Louro 1982). However, during the reforestation programs, it should be present that intraspecific hybridization can easily occur among different subspecies of European Black Pine, being important to avoid the plantation of *P. nigra* of unknown origin in the proximity of autochthonous pinewoods to reduce the risk of genetic pollution. Most of the plantations of *P. nigra* realized in Europe, were performed with material from unknown origin, being highly important to identify autochthonous populations in order to develop correct afforestation programs (Isajev *et al*. 2004; Zaghi 2008).

I.2.5 – Portuguese populations of *P. nigra*

During the middle of the 20th century, *P. nigra* was planted in Portugal (Louro 1982). The origin of the plant material used in these plantations is unknown in terms of provenance or infraspecific taxonomic classification. Presently, a few *P. nigra* allochthonous populations, with reduced size, are restricted to the North and Center of the country. Louro (1982) attempted to characterize these populations based on multidisciplinary data: dendrometric; botanical; natural regeneration ability; technological qualities; morphological and anatomical traits; evaluation of the phytosanitary status and edaphoclimatic conditions; and seed production/yield. The author tried to extrapolate about the origin of the seed material used in the plantation of these allochthonous pinewoods by characterizing the existent *P. nigra* stands in that decade. According to Louro (1982), three different subspecies of *P. nigra* existed in Portugal: subspecies *nigra*, subspecies *salzmannii*, and subspecies *laricio* (var. *calabrica* and var. *corsicana*) as the most predominant. Fabião and Oliveira (2006) indicated the presence of the same subspecies and also presumed *laricio* as the most frequent in Portugal. Louro (1982) focused his work on *P. nigra* subsp. *laricio* and highlighted that, regarding all aspects under study, *P. nigra* subsp. *laricio* var. *corsicana* seemed to had a more tendentiously favorable behavior comparatively to the other two subspecies (*nigra* and *salzmannii*).

Louro (1982) related that in the periods 1971-1975 and 1980-1981, about half the area of *P. nigra* subsp. *laricio* at North of the 'Douro' region had fired. In 1982, *P. nigra* subsp. *laricio* was considered an expansion species in Portugal, with an area of 9,000ha of which 6,000ha corresponded to young specimens (10 years or less). Moreover, the oldest stands had little more than 60 years at that time. This subspecies was found in several nuclei from 'Trás-os-Montes' to the North of 'Tejo' in pure or mixed stands, being almost all adults at the North of Portugal. The P. *nigra* subsp. *laricio* was found in soils generally composed by granites and schist at altitudes ranging from 400 to 1,600m. The evaluation of the plant development

and dendrometric data obtained in the 1980's, revealed that *P. nigra* was a well-adapted species. Despite the high adaptive potential, this species was always pointed out as problematic in terms of ability of natural regeneration. Louro (1982) only found evidences of natural regeneration in stands of 'Serra da Padrela' (subsp. *laricio* var. *corsicana*); 'Serra da Estrela (Covais)' (subsp. *laricio* var. *calabrica* and var. *corsicana*); 'Serra da Estrela (Nave)' (subsp. *laricio* var. *calabrica* and var. *corsicana*), 'Serra da Garraia' (subsp. *laricio* var. *calabrica* and var. *corsicana*) and 'Serra da Nogueira (Vilar d'Ouro)' (mixed stand composed by subsp. *laricio* var. *corsicana* and *P. sylvestris*).

Marques *et al.* (2012) only reported the existence of *P. nigra* subsp. *laricio* in Portugal, but with a smaller distribution area comparatively to the earlier study of Louro (1982).

Presently, the provenance of the seeds used in the plantation of the Portuguese *P. nigra* stands is still unknown. Additionally, no molecular studies of these allochtonous populations were developed in that sense and no information about their genetic diversity level is available. This study could help to answer to these molecular issues, taking into account the reduced distribution area and number of *P. nigra* stands in Portugal, which put this species in risk due to the occurrence of ongoing climatic change and forest fires, for further development of conservation and reforestation strategies.

I.3 – DNA extraction from tree species: an obstacle to genetic studies

Despite the need of genetic diversity characterization of *P. nigra* populations for reforestation and conservation strategies, other goals of the genetic studies performed in such pine species are related with its ambiguous taxonomic classification at the infraspecific level and with the search or identification of molecular markers suitable for population genetics studies.

However, before the development of the genetic diversity study in the Portuguese populations of *P. nigra* it was need to optimize the procedure of DNA extraction because pines constitute hard species for achieving genomic DNA of high quality due to presence of high amounts of endogenous tannins, phenolics and polysaccharides (Tibbits *et al.* 2006). These compounds are DNA contaminants and difficult the isolation of sufficient yield of high quality DNA (Khanuja *et al.* 1999; Tibbits *et al.* 2006; Ivanova *et al.* 2008). Polysaccharides

and tannins are difficult to separate from DNA (Murray and Thompson 1980) and polyphenols are a powerful oxidizing agent, reducing the yield and purity of DNA (Loomis 1974; Porebski *et al.* 1997). Moreover, the conifers have large size genomes that require isolation methods that minimize DNA degradation during the cells disrupting (Keller and Manak 1993; Muralitharan *et al.* 1994).

The extraction of total genomic DNA from a given sample constitutes a crucial and limiting step to any plant molecular study (Semagn *et al.* 2006; Ivanova *et al.* 2008). DNA can be extracted from fresh, lyophilized, preserved or dried samples, but fresh material is ideal for obtaining good quality DNA (Semagn *et al.* 2006). In order to obtain reproducible results, DNA must have high quality, high concentration and free of contaminants. The fidelity of results increase with DNA purity ratio, since contaminated nucleic acid often fails to give precise, reliable and reproducible results (Varma 2007). Furthermore, degraded and/ or unpurified DNA may affect subsequent molecular reactions, such as amplification, DNA digestion and cloning (Tibbits *et al.* 2006; Kumar *et al.* 2009) and can cause DNA degradation after long-term storage (Porebski *et al.* 1997; Schlink and Reski 2002; Sharma *et al.* 2002). The difficulty of separate DNA from co-extraction contaminants is a problem in the most cases, requiring many cleaning steps. Usually, repeated extractions are realized with organic solvents such as phenol, phenol-chloroform-isoamyl alcohol (25:24:1) and chloroform, followed by DNA precipitation with salt solutions (Tibbits *et al.* 2006). Cleaning steps of isolated DNA require long time to be effective, being limitative in molecular genetic studies involving hundreds of samples (Tibbits *et al.* 2006), and also decrease the DNA yield.

Many protocols for DNA extraction from various plant species and tissues have been published (Tibbits *et al.* 2006). However, they are typically time consuming and the quality of DNA obtained can be unpredictable, due to incomplete removal of PCR inhibitors (Ivanova *et al.* 2008). Frequently in tree species, it has been used extraction buffers containing cetyltrimethylammonium bromide (CTAB) (Murray and Thompson 1980; Doyle and Doyle 1987; Wagner *et al.* 1987) suggesting the utility of this reagent in DNA extraction. Other components such as soluble polyvynilpirrolidone (PVP-40) (Stewart and Via 1993; Devey *et al.*1996; Kim *et al.* 1997), SDS (sodium dodecyl sulfate) (Nelson *et al.* 1994; Jobes *et al.* 1995) and guanidine detergent (Lin and Kuo 1998) have also been suggested; despite the latter two revealed to be less effective in trees species, especially in pines (Tibbits *et al.* 2006).

The biochemical composition differs considerably among plant tissues and species (Khanuja *et al.* 1999). However, a universal procedure for DNA extraction does not exists justifying the high number of optimized protocols for DNA isolation from several plant species that are presently available.

I.4 – Molecular studies in *P. nigra*

Genetic markers have been used for studies in plant evolution, taxonomy, phylogenetic, ecology and genetics (Agarwal *et al.* 2008). A genetic marker consists in biological features (phenotype, protein or DNA fragment) that correspond to a genotype (O'Brien 1990). These characteristics are determined by allelic forms of genes or genetic *loci*, which may be transmitted from one generation to another and can be used as experimental probes or tags allowing the study of an individual, a tissue, a cell, a nucleus, a chromosome or a gene (Jiang 2013). Genetic markers should evidence polymorphism and can be used in combination with other markers as characteristic of this genotype (O'Brien 1990). Polymorphism is the identification of distinct alleles in the same *locus* and is the principal characteristic of a genetic marker (Ferreira and Grattaplaglia 1996; Andersen 2003). The concept of genetic marker has over than two centuries. It remits to the 19th century, when Gregor Mendel used phenotype-based genetic markers in his experiments with *Pisum sativum* (Agarwal *et al.* 2008). Genetic markers are inherited according to the Mendelian laws of inheritance, allowing differentiating genotypes and facilitating the detection of differences in the genetic information between individuals (Lefebvre and Chèvre 1995).

Over the years, different classifications for genetic markers were proposed (see Solanille 1994; Lefebvre and Chèvre 1995). More recent classification of genetic markers were published by Xu (2010) and Poczai *et al.* (2013), and are summarized in Figure I.2.

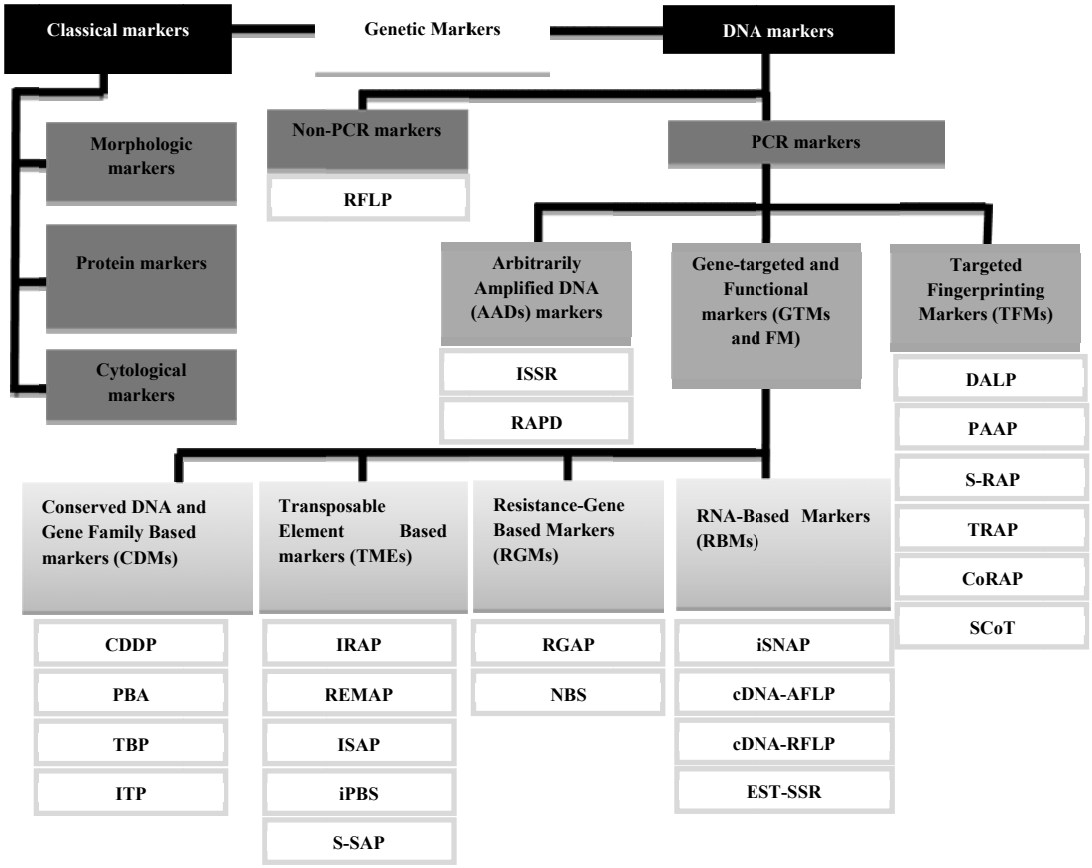

Figure I.2 – Schematic representation of two recent classifications of the genetic markers (Adapted from Xu 2010 and Poczai *et al.* 2013). The acronyms of the DNA markers are described in the abbreviation's list.

A useful molecular marker should be: i) highly polymorphic; ii) co-dominant allowing the distinction between heterozygous and homozygous; iii) able to allow the easy distinction between alleles; iv) ubiquitous and abundant through the genome (not clustered in certain regions); v) selective neutral; vi) easily detected; vii) cost-effective; viii) reproducible; xix) genome-specific in nature, especially with polyploids; and x) developed with small amounts of DNA or tissue (Agarwal *et al.* 2008; Xu 2010; Jiang 2013).

In spite of the large number of molecular markers available today, there is not an ideal or universal marker system to be used in all studies (Arif *et al.* 2010). The choice of a marker system should be made accordingly to the goal of each study.

Studies in *P. nigra* with molecular markers performed until now, used Restriction Fragment Length Polymorphism of chloroplast DNA (RFLP cpDNA) and Random Amplified Polymorphic DNA (RAPD) in Austrian and Dalmatian Black Pine (Liber *et al.* 1999); RAPD in Croatian, Austrian and Turkish populations (Liber *et al.* 2003) and in Siberian populations (Lučić *et al.* 2010); cpSSR in populations of Bulgaria (Naydenov *et al.* 2006) and in populations of Western Europe (Afzai-Raffi and Dodd 2007); barcoding genes *Cox1*, *MatK*,

Rbcl, *Trnh* and *Nad 5-4,* nSSRs, cpSSRs and detection of SNPs (Giovannelli *et al*. 2014); noncoding trn regions of cpDNA in subsp. *pallasiana* and their varieties (Gülsoy *et al.* 2014); Simple Sequence Repeat (SSR) in subsp. *laricio* (Bonavita *et al.* 2015) and Expressed Sequence Tag - Simple Sequence Repeat (EST-SSR) in the populations of Balkans (Šarac *et al.* 2015).

Since in the present study we used the ISSR and SCoT markers, only these marker systems will be described with more detail in the following subsections.

I.4.1 – Inter-Simple Sequence Repeat (ISSR) markers

ISSRs were first described by Meyer *et al.* (1993) but were designated as inter-microsatellites by Zietkiewicz *et al.* (1994). According to these authors, this technique allows the simultaneous analysis of a variety of genomic *loci* (Zietkiewicz *et al.* 1994) being widely reported as a multilocus marker (Tomar *et al.* 2010). This happens due to the presence of abundant repeated sequences all over the genome, which allows that SSR primers anneal in numerous regions and create a complex amplification pattern (Blanco *et al.* 2000).

This method consists in the amplification of a genomic DNA region flanked by two adjacent and inversely oriented microsatellite regions (Zietkiewicz *et al.* 1994; Ammiraju *et al.* 2001; Tomar *et al.* 2010) using a single primer (Blanco *et al.* 2000, Ammiraju *et al.* 2001; Prince 2015). The primers are based on di-, tri-, tetra- or pentanucleotide repeats, have 16-18bp long (Blanco *et al.* 2000; Chadha and Gopaloakrishna 2007; Tomar *et al.* 2010) that can be either unanchored (Gupta *et al*. 1994; Meyer *et al.* 1993; Wu *et al.* 1994) or more usually anchored at the 3'- or 5'-end by 2 to 4 arbitrary nucleotides (Ziekiewicz *et al.* 1994; Blanco *et al.* 2000) (Fig. I.3).

According to Blanco *et al.* (2000), Ammiraju *et al.* (2001) and Prince (2015), some of the advantages of ISSR marker system are: i) only a single step is need after DNA isolation, decreasing relatively the technique cost; ii) the amount of DNA required is reduced, so this can be used to study rare organisms; iii) does not require the previous knowledge of the sequence of target DNA region; iv) highly informative and polymorphic; and v) reproducible.

ISSR are dominant markers (Tomar *et al.* 2010; Prince 2015), and their analysis is based on a binary code which consists in the detection of presence or absence of band or marker among the individuals under study (Prince 2015).

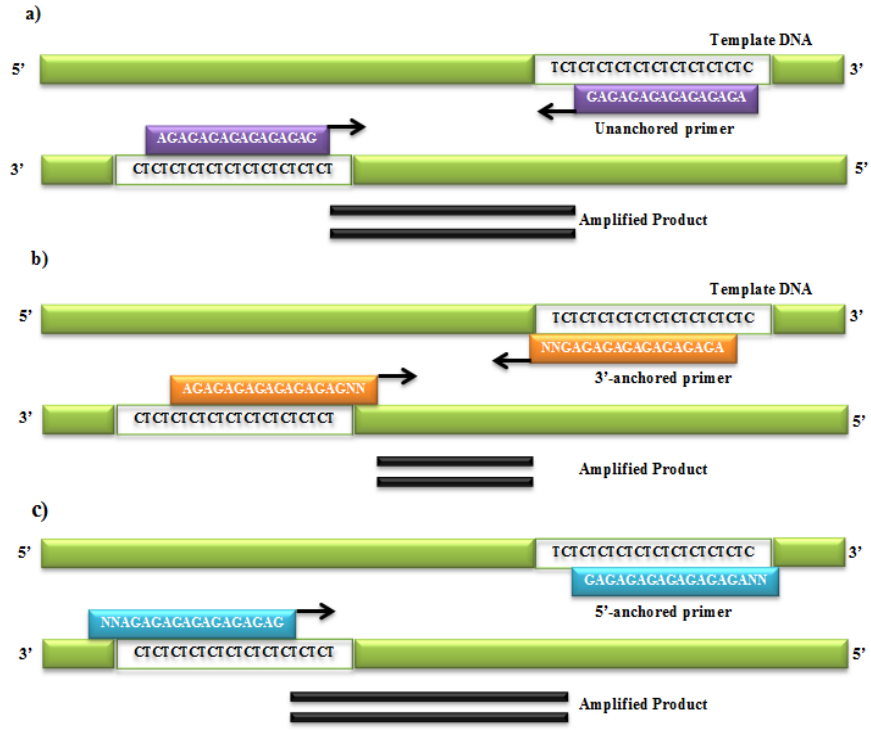

Figure I.3 – Schematic representation of ISSR technique with a single primer $(AG)_8$, unanchored (a), 3'-end anchored (b) or 5'-end anchored (c) targeting a (TC)*n* repeat, to amplify a ISSR flanked by two inversely oriented (TC)*n* sequences. (a) Unanchored (AG)*n* primer can anneal anywhere in the (TC)*n* repeat region on the template DNA leading to slippage and ultimately smear formation; (b) (AG)*n* primer anchored with two nucleotides (NN) at the 3'-end anneals at specific regions of the template DNA and produces clear bands; (c) (AG)*n* primer anchored with two nucleotides (NN) at the 5'-end anneals at specific regions of the template DNA and amplifies part of the repeat region also, leading to larger bands (Adapted from Reedy *et al.* 2002).

Over the years, ISSR have been used with success in several plant species for the study of: i) genetic diversity (Souframanien and Gopalakrishna 2004; Basha and Sujatha 2007; Cipriano *et al.* 2013; Coutinho *et al.* 2014a); ii) gene-tagging (Ammiraju *et al.* 2001); iii) cultivars identification and estimation of phylogenetic relationships (Blanco *et al.* 2000; Ammiraju *et al.* 2001; Seyedimoradi and Talebi 2014), iv) genetic mapping (Arcade *et al.* 2000); v) DNA fingerprinting (Carvalho *et al.* 2005; Cabo *et al.* 2014a; Coutinho *et al.* 2014b). ISSR markers were also used to study the genetic diversity in *P. nigra* populations from Southern Spain and Northern Morocco (Rubio-Muraga *et al.* 2012; Lucas Borja *et al.* 2013).

I.4.2 – Start Codon Targeted (SCoT) markers

In 2009, Collard and Mackill developed a new molecular marker system named Start Codon Targeted (SCoT) suitable for the amplification of plant genes since the primers were designed for a short conserved region flanking their start codon (ATG). SCoTs could be

amplified by using a single 18-mer primer and an annealing temperature of 50 ºC (Collard and Mackill 2009).

Due to the basis of SCoT primer design, SCoT markers are expected to be distributed within gene regions that contain genes on both plus and minus DNA strands. It is also possible that pseudogenes and transposable elements may be used as primer binding sites by SCoT polymorphism technique (Collard and Mackill 2009)

The amplification product can be visualized with a regular electrophoresis on agarose gel and staining. All these characteristics made the SCoTs a suitable technique for the majority of plant research labs with standard equipment (Collard and Mackill 2009). These authors verified that SCoT markers are generally reproducible and suggested that primer length and annealing temperature are not the sole factors determining reproducibility. PCR amplification profiles of SCoT marker indicated that they are dominant markers like RAPDs, which means that the fragments detected on gels are scored as absent (0) or present (1) (Fig. I.4).

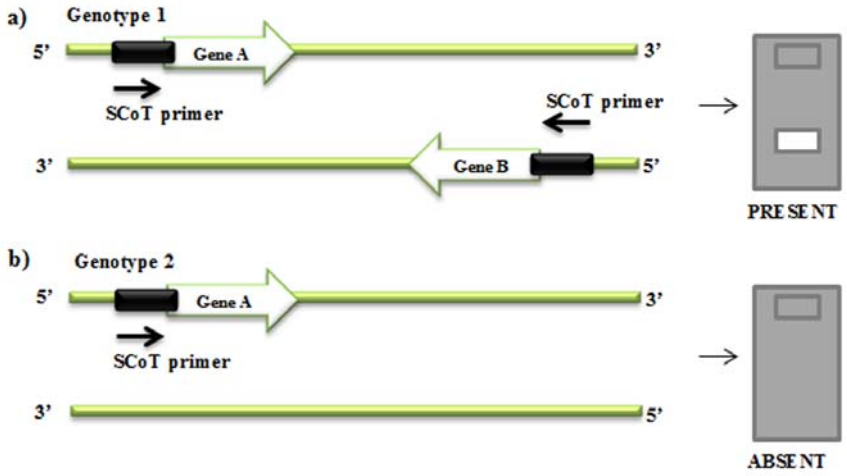

Figure I.4 – Schematic representation of SCoT marker amplification. (a) The SCoT marker is generated after binding of the primer to a complementary sequence in both DNA strands at a distance suitable for PCR amplification; (b) no amplified product is observed due to high distance between the complementary sequences of the primer, avoiding the PCR amplification (Adapted from Collard and Mackill 2009).

Collard and Mackill (2009) suggested that SCoTs can be used alone or in combination with other markers.

SCoTs have been an useful tool to estimate genetic relationships (Luo *et al.* 2010; Xiong *et al.* 2011; Amimoradi *et al.* 2012; Guo *et al.* 2012; Mulpuri *et al.* 2013; Jiang *et al.* 2014), germplasm management and clonal identification at field breeding stations (Luo *et al.*

2010), diagnostic fingerprinting of cultivars and genotypes (Gorji *et al.* 2011; Cabo *et al.* 2014b), estimate genetic diversity (Chen *et al.* 2010; Luo *et al.* 2010, 2011; Xiong *et al.* 2011; Amimoradi *et al.* 2012; Guo *et al.* 2012; Bhattacharyya *et al.* 2013; Alikhani *et al.* 2014; Chen and Liu 2014; Jiang *et al.* 2014; Que *et al.* 2014; Shahlaei *et al.* 2014; Zeng *et al.* 2014) and evaluate genetic homogeneity of micropropagated plantlet (Agarwal *et al.* 2015).

Thus, SCoT markers are a simple and novel marker system with several advantages, such as: simple, cost-effective; highly polymorphic and informative; versatile to different plant species (Collard and Mackill 2009; Luo *et al.* 2010, Guo *et al.* 2012). The SCoT polymorphism might be straightly associated to gene function (Poczai *et al.* 2013).

It is important to highlight that only one work used SCoT markers in the forestry species, *Quercus brantii* (Alikhani *et al*. 2014). So, the present study constitutes the first use of SCoTs in a gymnosperm species.

I.5 – Cytogenetics analysis in *P. nigra*

At beginning of 20th century, chromosome banding and *in situ* hybridization (ISH) techniques allowed to construct and analyze the karyotypes of a number of species (Jacobs *et al.* 2000). ISH turns a key to link molecular results and chromosomes (Heslop-Harrison *et al.* 1991), so it was important to develop reproducible high-resolution cytogenetic techniques (Doudrick *et al.* 1995). The development of fluorescence *in situ* hybridization (FISH) allowed a clear advance in chromosome studies specifically to physical mapping, genome analyses and evolutionary studies (Ribeiro *et al.* 2008). Actually, FISH technique is frequently used in *Pinus* species, mainly to localize DNA probes and to identify chromosomes or chromosomal regions (Hizume *et al.* 2002). Some examples of probes that have been used in FISH experiments performed in pine species are: i) the 5S and 45S rDNA which allowed the detection of rDNA *loci* in *Pinus densiflora*, *Pinus thunbergii*, *Pinus elliottii* var. *elliottii*, *Pinus sylvestris*, *Pinus radiata* or *Pinus taeda* (Hizume *et al.* 1992; Doudrick *et al.* 1995; Lubaretz *et al.* 1996; Hizume and Kondo 2000; Jacobs *et al.* 2000; Pavia *et al.* 2014); ii) a telomeric sequence that presented different chromosomal locations in *P. sylvestris* (Fuchs *et al.* 1995) and *P. elliottii* (Schmidt *et al.* 2000); and iii) SSR probes that allowed the construction of an ideogram for the probe $(AG)_{10}$ in *P. sylvestris* using ND-FISH (Pavia *et al.* 2014). However, the advances in this field has been limited in gymnosperms, particularly in

conifers, due to the constancy of chromosome number (average of 12 chromosomes per haploid set) and similar chromosome morphology (Muratova 1997; Nkongolo and Mehes-Smith 2012). These facts explain the difficulty in the identification of individual chromosomes (Pederick 1967), as well as, in the studies of genome organization and in the detection of chromosomal rearrangements (Cai *et al.* 2006). Nevertheless, interspecific differences among the Eurasian pines of section *Pinus*, including *P. nigra*, were found (Karvonen *et al*. 1993; Lubaretz *et al*. 1996; Hizume *et al*. 2002; Liu *et al.* 2003; Bogunić *et al*. 2011a). All Eurasian pines presented two 5S rDNA *loci* except for *Pinus massoniana* (only one *locus*) and *Pinus densata* (with three *loci*) (Bogunić *et al.* 2011b). The number of 18S rDNA *loci* ranged from 8 to 24 (Hizume *et al*. 2002; Liu *et al*. 2003). Additionally, different Mediterranean *P. nigra* populations have showed considerable variation in the number and position of the 18S rDNA *loci* at the intraspecific level (Bogunić *et al*. 2011a,b). Recently, Lemos *et al.* (2015) detected sixteen 45S rDNA *loci* after FISH experiments performed in mitotic chromosome spreads of Portuguese *P. nigra* individuals.

Geographic variation in the karyotypes of the Pinaceae species remains largely unexplored (Nkongolo and Mehes-Smith 2012). Kaya *et al.* (1985) detected differences among karyotypes of *P. nigra* from different sources. Bogunić *et al.* (2011a) found variations in number of hybridization signals of 5S rDNA, 18S rDNA sequences, number of chromomycin A (CMA) and DAPI bands in *P.nigra* species from different seed sources and regions. On the contrary, molecular cytogenetics and flow cytometry techniques revealed a conserved genome organization in *Pinus mugo* and *Pinus uncinata* (Bogunić *et al.* 2011b).

According to Nkongolo and Mehes-Smith (2012), advances in cytogenetic allowed the understanding of speciation mechanisms. In genus *Pinus*, the species share major karyotype characteristics and the mechanisms of speciation involving chromosome rearrangements may not represent large structural mutations but small or cryptic changes (Nkongolo and Mehes-Smith 2012). Alternatively, if the speciation has occurred as a consequence of large chromosome modification, changes such as paracentric inversions or reciprocal translocations with segments of equal size that did not modify the karyotype morphology have probably occurred (Nkongolo and Mehes-Smith 2012).

According to Siljak-Yakovlev *et al.* (2014), molecular cytogenetic analyses can unequivocally reveal the subtle chromosomal changes even among lower grade taxonomic categories. As a result, it could turn possible the determination of processes of species

differentiation, evolution and phylogenetic relationships among *taxa*, by combining cytogenetic with phytogeographical and ecological data.

I.6 – Objectives

With this study is intended to:

- Optimize a CTAB-based protocol of genomic DNA extraction from frozen and dehydrated needles of European Black Pine, and to evaluate the suitability of the isolated DNA samples for amplification of molecular markers in *P. nigra* and other forestry species;
- Characterize *P. nigra* samples collected in different geographical regions and belonging to different infraspecific *taxa* (allopatric subspecies and varieties) with molecular markers, for further estimation of genetic relationships and structure, and to infer about the proposed taxonomic classification based on the molecular data;
- Evaluate for the first time the genetic diversity, structure, relationships among 127 individuals from the six most representative Portuguese populations of *P. nigra*, in order to extrapolate their provenances by comparing with molecular patterns achieved in foreign European Black Pine samples of different allopatric subspecies and varieties, and to ascribe which *P. nigra* subspecies and/or variety effectively exists in Portugal;
- Physically locate SSR and the 45S rDNA sequences in two Portuguese populations of *P. nigra* and *P. sylvestris* using ND-FISH and FISH techniques, respectively; to detect possible cytogenetic instabilities; to study nucleolar activity using the salt-nylon silver staining technique and to compare with the FISH results within and between the two species.

I.7 – References

Afzal-Rafii Z, Dodd RS (2007) Chloroplast DNA supports a hypothesis of glacial refugia over postglacial recolonization in disjunct populations of black pine (*Pinus nigra*) in Western Europe. Molecular Ecology 16:723-736.

Agarwal M, Shrivastava N, Padh H (2008) Advances in molecular marker techniques and their applications in plant sciences. Plant Cell Reports 27:617-631.

Agarwal T, Gupta AK, Patel AK, Shekhawat NS (2015) Micropropagation and validation of genetic homogeneity of *Alhagi maurorum* using SCoT, ISSR and RAPD markers. Plant Cell Tissue and Organ Culture. doi:10.1007/s11240-014-0608-z.

Akkemik Ü, Yilmaz H, Oral D, Kaya A (2010) Some changes in taxonomy of pines (*Pinus* L.) native to Turkey. Journal of the Faculty of Forestry 61:63-78.

Alikhani L, Rahmani M-S, Shabanian N, Badakhshan H, Khadivi-Khub A (2014) Genetic variability and structure of *Quercus brantii* assessed by ISSR, IRAP and SCoT markers. Gene 552:176-183.

Amirmoradi B, Talebi R, Karami E (2012) Comparison of genetic variation and differentiation among annual Cicer species using Start Codon Targeted (SCoT) polymorphism, DAMD-PCR, and ISSR markers. Plant Systematics and Evolution 298:1679-1688.

Ammiraju JSS, Dholakia BB, Santra DK, Singh H, Lagu MD, Tamhankar, Dhaliwal HS, Rao VS, Gupta VS, Ranjekar PK (2001) Identification of inter simple sequence repeat (ISSR) markers associated with seed size in wheat. Theoretical and Applied Genetics 102:726–732.

Amri I, Hanana M, Jamoussi B, Hamrouni L (2014) Essential oils of *Pinus nigra* J.F. Arnold subsp. *laricio* Maire: Chemical composition and study of their herbicidal potential. Arabian Journal of Chemestry. doi:10.1016/j.arabjc.2014.05.026.

Andersen SB (2003) Genetic markers in plant breeding. Types of genetic markers in plant breeding. [http://www.kursus.kvl.dk].

Arbez M, Miller C (1971) Contribution a l'étude de la variabilité géographique de *Pinus nigra* Arn.. Añnales des Sciences Forestiéres 28:23-49.

Arcade A, Anselin F, Rampant PF, Lesage MC, Pâques LE, Prat D (2000) Application of AFLP, RAPD and ISSR markers to genetic mapping of European and Japonese larch. Theoretical and Applied Genetics 100:299-307.

Arif IA, Bakir MA, Khan HA, Al Farhan AH, Al Homaidan AA, Bahkali AH, Sadoon MA, Shobrak M (2010) A brief review of molecular techniques to assess plant diversity. International Journal of Molecular Sciences 11:2079-2096.

Barbéro M, Loisel R, Quézel P, Richardson DM, Romane F (1998) Pines of mediterranean basin. In: Ecology and Biogeography of Pinus (edited by DM Richardson), pp. 153-170, Cambridge University Press.

Barzegari A, Vahed SZ, Atashpaz S, Khani S, Omidi Y (2010) Rapid and simple methodology for isolation of high quality genomic DNA from coniferous tissues (*Taxus baccata*). Molecular Biology Reports 37:833-837.

Basha SD, Sujatha M (2007) Inter and intra-population variability of *Jatropha curcas* (L.) characterized by RAPD and ISSR markers and development of population-specific SCAR markers. Euphytica 156:375-386.

Bhattacharyya P, Kumaria S, Kumar S, Tandom P (2013) Start Codon Targeted (SCoT) marker reveals genetic diversity of *Dendrobium nobile* Lindl., an endangered medicinal orchid species. Gene 529:21-26.

Blanco E (1998) Los bosques Ibéricos. Una interpretación geobotánica. Planeta (Barcelona).

Blanco A, Bruno A, De Vita P, Di Fonzo N, Lotti C, Pasqualone A (2000) Use of ISSR markers for cultivar identification in durum wheat. In: Durum wheat improvement in the Mediterranean region: New challenges. Zaragoza: CIHEAM (edited by C Royo, M Nachit M, N Di Fonzo, JL Araus JL), pp. 157-161 (Options Méditerranéens: Série A. Séminaires Méditerranéens; n. 40).

Bogunić F, Muratovi E, Ballian D, Siljak-Yakolev S, Brown S (2007) Genome size stability of five subspecies of *Pinus nigra* Arnold s.l. Environmental and Experimental Botany 59:354-360.

Bogunić F, Siljak-Yakovlev S, Muratović E, Ballian D (2011a) Different karyotype patterns among allopatric *Pinus nigra* (Pinaceae) populations revealed by molecular cytogenetics. Plant Biology 13(1):194–200.

Bogunić F, Siljak-Yakovlev S, Muratović E, Pustahija, Medjedović (2011b) Molecular cytogenetics and flow cytometry reveal conserved genome organization in *Pinus mugo* and *P. uncinata*. Annals of Forest Science 68(1):179-187.

Bonavita S, Vendramin GG, Bernardini V, Avolio S, Regina TMR (2015) The first SSR-based assessment of genetic variation and structure among *Pinus laricio* Poiret populations within their native area. Plant Biosystems - An International Journal Dealing with all Aspects of Plant Biology: Official Journal of the Societa Botanica Italiana. doi:10.1080/11263504.2015.1027316.

Bonnet-Masimbert M, Bikay-Bikay V (1978) Variabilité intraspécifique des isozymes de la glumate-oxaloacetate-transaminase chez *Pinus nigra* Arnold Intérêt pour la taxonomie des sous espèces. Silvae Genetica 27:71-79.

Cabo S, Carvalho A, Rocha L, Martín A, Lima-Brito J (2014a) IRAP, REMAP and ISSR fingerprinting in newly formed hexaploid tritordeum (X *Tritordeum* Ascherson et Graebner) and respective parental species. Plant Molecular Biology Reporter 32:761-770.

Cabo S, Ferreira L, Carvalho A, Martins-Lopes P, Martín A, Lima-Brito JE (2014b) Potential of Start Codon Targeted (SCoT) markers for DNA fingerprinting of newly synthesized tritordeums and their respective parents. Journal of Applied Genetics 55:307-312.

Cai Q, Zhang D, Liu Z-L, Wang X-R (2006) Chromosomal localization of 5S and 18S rDNA in five species of subgenus Strobus and their implications for genome evolution of Pinus. Annals of Botany 97(5):715-722. doi:10.1093/aob/mcl030. PMID:16481361.

Carvalho A, Matos M, Lima-Brito J, Guedes-Pinto H, Benito C (2005) DNA fingerprint of F1 interspecific hybrids from the Triticeae tribe using ISSRs. Euphytica 143:93-99.

Chadha S, Gopalakrishna T (2007) Comparative assessment of REMAP and ISSR marker assays for genetic polymorphism studies in *Magnaporthe grisea*. Current Science 93:688-692.

Chen F-Y, Liu J-H (2014) Germplasm genetic diversity of *Myrica rubra* in Zhejiang Province studied using inter-primer binding site and start codon-targeted polymorphism markers. Scientia Horticulturae 170:169-175.

Chen H, He X, Luo C, Zhu J, Li F (2010) Analysis on the genetic diversity of 24 Longan (*Dimocarpus longan*) accessions by SCoT Markers. Acta Horticulturae Sinica 37:1651-1654.

Christensen KI (1997) Pinaceae, Cupressaceae, Taxaceae, Ephedraceae, Salicaceae, Juglandaceae, Betulaceae, Fagaceae, Ulmaceae Moraceae. In: Flora Hellenica, vol 1 (edited by A Strid, K Tan), pp. 1–17, Koeltz Scientific Books, Königstein, Germany.

Christopoulou A, Fule PZ, Andriopoulos P, Sarris D, Arianoutsou M (2013) Dendrochronology-based fire history of *Pinus nigra* forests in Mount Taygetos, Southern Greece. Forest Ecology and Management 293:132–139.

Cipriano J, Carvalho A, Fernandes C, Gaspar MJ, Pires J, Bento J, Roxo L, Louzada J, Lima-Brito J (2013) Evaluation of genetic diversity of Portuguese *Pinus sylvestris* L. populations based on molecular data and inferences about the future use of this germplasm. Journal of Genetics 92:e41-e48.

Collard BCY, Mackill DJ (2009) Conserved DNA-derived polymorphism (CDDP): a simple and novel method for generating DNA markers in plants. Plant Molecular Biology Reporter 27:558-562.

Coutinho JP, Carvalho A, Lima-Brito J (2014a) Genetic diversity assessment and estimation of phylogenetic relationships among 26 Fagaceae species using ISSRs. Biochemical Systematics and Ecology 54:247-256.

Coutinho JP, Carvalho A, Lima-Brito J (2014b) Fingerprinting of Fagaceae individuals using intermicrosatellite markers. Journal of Genetics 93:e132-e140. Online only: http://www.ias.ac.in/jgenet/OnlineResources/93/e132.pdf.

Debazac EF (1963) L'aire spontanée du pin de salzman en France. Revue Forestière Française.

Debazac EF (1964) Manuel des conifers. Ecole Nationale des Eaux et Forêts, Nancy.

Debazac EF (1965) Les pinerais de Calabre et de Sicile. Revue Forestière Française.

Debazac EF (1971) Contribution à la connaissance de la répartition et de l'écologie de *Pinus nigra* Arn. Dans le sud-est de l'Europe. Añnales des Sciences Forestiéres 28:91-139.

del Cerro Barja A, Lucas-Borja ME, Garcia EM, López Serrano FR, Abellán MA, Morote FAG, López RN (2009) Influence of stand density and soil treatment on the Spanish Black Pine (*Pinus nigra* Arn. ssp. *Salzmannii*) regeneration in Spain. Investigación Agraria: Sistemas y Recursos Forestales 18:167-180.

Delevoy G (1949) A propos de la systematique de *Pinus nigra* Arnold. Travaux Statation Recherches, Série B 12. Groenendaal, Belgique.

Devey M, Bell JC, Smith DN, Neale DB, Moran GF (1996) A genetic linkage map for *Pinus radiata* based on RFLP, RAPD, and microsatellite markers. Theoretical and Applied Genetics 92:673-679.

Dobrinov I (1983) Genetic and selection of forest tree. Zemizdat, Sofia, Bulgaria.

Doyle JJ, Doyle JH (1987) Isolation of plant DNA from fresh tissue. Focus 12:13-15.

Doudrick RL, Heslop-Harrison JS, Nelson CD, Schmidt T, Nance WL, Schwarzacher T (1995) Karyotype of Slash Pine (*Pinus elliottii* var. *elliottii*) using patterns of fluorescence *in situ* Hybridization and Fluorochrome banding. The Journal of Heredity 86:289-296.

Espelta JM (1999) La reconstrucció de paisatges forestals afectats per grans incendis. Projecte pilot a l' incendi del Bages-Berguedà del 1994- Silvicultura 27:6-9.

Esteban LG, Martín JA, de Palacios P, Fernández FG (2012) Influence of region of provenance and climate factors on wood anatomical traits of *Pinus nigra* Arn. subsp. *salzmannii*. European Journal of Forest Research 131:633-645.

EUFORGEN (2009) Distribution map of Black pine (*Pinus nigra*). www.euforgen.org Accessed on 6 August 2015.

Fabião AMD, Oliveira ÂC (2006) A Floresta em Portugal. Lisboa.

Farjon A (1984) Pines: drawings and Descriptions of the Genus Pinus. Leiden: E. J. Brill & Dr W. Baxkhuys.

Farjon A (1998) World checklist and bibliography of conifers. Royal Botanic Gardens, Kew. 300 p. ISBN 1-900347-54-7.

Farjon A (2010) A handbook of The World's Conifers Volume 1. Brill Leiden-Boston.

Farjon A (2013) *Pinus nigra*. The IUCN Red List of Threatened Species. Version 2015.1. <www.iucnredlist.org>. Downloaded on 9 June 2015.

Ferreira ME, Grattapaglia D (1996) Introdução ao uso de marcadores moleculares em análise genética - 2ª Edição, EMBRAPA-CENARGEN, Brasília, 218 p.

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

Figueiral I, Carcaillet C (2005) A review of Late Pleistocene and Holocene biogeography of highland Mediterranean pines (*Pinus* type *sylvertris*) in Portugal, based on wood charcoal. Quaternary Science Reviews 24:2466-2476.

Fuchs J, Brandes A, Schubert I (1995) Telomere sequence localization and karyotype evolution in higher plants. Plant Systematics and Evolution 196(3–4):227–241. doi:10.1007/BF00982962.

Gaussen H, Haywood VH, Chater AO (1964) *Pinus*. In: Flora Europaea, Vol. 1 (edited by TG Tutin, VH Heywood, NA Burges, DM Moore, DH Valentine, SM Walters, DA Webb), pp. 32-35, Cambridge University Press, Cambridge, UK.

Gaussen H, Webb DA, Heywood VH (1993) Genus *Pinus*. In: Flora Europaea, Vol 1 (edited by TG Tutin, VH Heywood, NA Burges, DM Moore, DH Valentine, SM Walters, DA Webb), pp. 40-44, Cambridge University Press, Cambridge, UK.

Gaussen H (1960) Les gymnospermes actuelles et fossiles. Fasc. VI. Les Coniférales. Faculté des Sciences, Toulouse, France.

Gellini R (1968) Posizione sistematica del pino nero di Villetta Barrea in base ai caratteri anatomici degli aghi. Annali Accademia Italiana de Scienze Forestali XVII:101–122.

Gerber S, Baradat P, Marpeau A, Arbez M (1995) Geographic variation in terpene composition of *Pinus nigra* Arn.. Forest Genetics 2:1-10.

Gernandt DS, Lopez GG, Garcia SO, Liston A (2005) Phylogeny and classification of Pinus. Taxon 54:29-42.

Gernandt DS, Magallon S, Lopez GG, Flores OZ, Willyard A, Liston A (2008) Use of simultaneous analysis to guide fossil-based calibrations of Pinaceae phylogeny. International Journal of Plant Sciences 169:1086–1099. ISSN 1058-5893.

Giovannelli G, Roig A, Spanu R, Vendramin GG, Fady B (2014) Using different molecular approaches to resolve the pylogeny and taxonomic complexity of the European black pine, *Pinus nigra* Arn.. MedPINE5 - 5th International Conference on Mediterranean Pines Solsona, Spain, September 22-26, 2014.

Gorji AM, Poczai P, Polgar Z, Taller J (2011) Efficiency of arbitrarily amplified dominant markers (SCoT, ISSR and RAPD) for diagnostic fingerprinting in tetraploid potato. American Journal of Potato Research 88:226–237.

Greuter W, Burdet HM, Long G (Eds) (1984) Med-checklist, Pterydophyta Gymnospermae, Dycotiledones (Acanthaceae-Cneoraceae). Editions des Conservatoire et Jardin botaniques de la Ville de Genève, Genève, 1:33-34.

Gülçin I, Büyükokuroğlu ME, Oktay M, Küfrevioğlu (2003) Antioxidant and analgesic activities of turpentine of *Pinus nigra* Arn. subsp. *pallasiana* (Lamb.) Holmboe. Journal of Ethnopharmacology 86:51-58.

Gülsoy AD, Gülsoy AM, Çengel B, Kaya Z (2014) The evolutionary divergence of *Pinus nigra* subsp. *pallasiana* and its varieties based on noncoding trn regions of chloroplast genome. Turkish Journal of Botany 38:627-636.

Guo D-L, Zhang J-Y, Liu C-H (2012) Genetic diversity in some grape varieties revealed by SCoT analyses. Molecular Biology Reports 39:5307-5313.

Gupta PK, Varshney RK (2000) The development and use of microsatellite markers for genetic analysis and plant breeding with emphasis on bread wheat. Euphytica 113:163–185.

Heslop-Harrison JS, Schwarzacher T, Anamthawat-Jonsson K, Leitch AR, Shi M, Leitch IJ (1991) *In situ* hybridization with automated chromosome denaturation. Technique 3:109-116.

Hizume M, Ishida F, Murata M (1992) Multiple locations of ribosomal RNA genes in chromosomes of pines, *Pinus densiflora* and *P. thunbergii*. The Japonese Journal of Genetics 67:389–396.

Hizume M, Kondo T (2000) Relationships between fluorescent chromosome bands and DNA sequences in conifers. In: Cytogenetic studies of forest trees and shrubs (edited by H Guttenberger, Z Borzan, SC Schlarbaum, TPV Hartman), pp 71-79, Abora Publ. Zvolen, Slovakia.

Hizume M, Shibata F, Matsusaki Y, Garajova Z (2002) Chromosome identification and comparative karyotypic analyses of four Pinus species. Theoretical and Applied Genetics 105:491–497.

Isajev V, Fady B, Semerci H Andonovski V (2004) EUFORGEN Technical Guidelines for genetic conservation and use for European black pine (*Pinus nigra*). International Plant Genetic Resources Institute, Rome.

Ivanova NV, Fazekas AJ, Hebert PDN (2008) Semi-automated, membrane-based protocol for DNA isolation from Plants. Plant Molecular Biology Reporter 26:186-198.

Jacobs MD, Gardner RC, Murray BG (2000) Cytological characterization of heterochromatin and rDNA in *Pinus radiata* and *P. taeda*. Plant Systematics and Evolution 223:71-79.

Jiang G-L (2013) Molecular markers and marker-assisted breeding in plants. InTech. http://dx.doi.org.10.5772/52583.

Jiang LF, Qi X, Zhang XQ, Huang LK, Ma X, Xie WG (2014) Analysis of diversity and relationships among orchardgrass (*Dactylis glomerata* L.) accessions using start codon-targeted markers. Genetics and Molecular Research 13:4406-4418.

Jobes DV, Hurley DL, and Thien LB (1995) Plant DNA isolation: method to efficiently remove polyphenolics, polysaccharides and RNA. Taxon 44:379-386.

Kaya Z, Ching KK, Stafford SG (1985) A statistical analysis of karyotypes of European black pine (*Pinus nigra* Arnold) from different sources. Silvae Genetica 34(3):148–156.

Karvonen P, Karkalainen M, Savolainen O (1993) Ribosomal RNA genes in Scots pine (*Pinus sylvestris* L.): chromosomal organization and structure. Genetica 88:59-68.

Keller GH, Manak MM (1993) DNA probes. Macmillan Publishers, New York, NY; 1993:70-150.

Khanuja SPS, Shasany AK, Darokar MP, Kumar S (1999) Rapid isolation of DNA from dry and fresh samples of plants producing large amounts of secondary metabolites and essential oils. Plant Molecular Biology Reporter 17:1-7.

Kim CS, Lee CH, Shin JS, Chung YS, and Hyung NI (1997) A simple and rapid method for isolation of high quality genomic DNA from fruit trees and conifers using PVP. Nucleic Acids Research 25:1085-1086.

Kumar P, Grupta VK, Misra AK, Modi DR, Pandey BK (2009) Potential of Molecular Markers in Plant Biotechnology. Plant Omics Journal 2:1836-3644.

Lee CH (1968) Geographic Variation in European Black Pine. Silvae Genetica 17:165-74.

Lefebvre V, Chèvre AM (1995) Tools for making plant disease and pest resistance genes: a review. Agronomic 15:3-19.

Lemos M, Pavia I, Azevedo C, Gaspar MJ, Carvalho A, Lima-Brito J (2015) Chromosomal irregularities and nucleolar activity studies in *Pinus nigra* Arn. and *Pinus sylvestris* L.

VII Jornadas de Genética e Biotecnologia, 26 a 28 de Março de 2015, Universidade de Trás-os-Montes e Alto Douro, Vila Real, Portugal, p. 58.

Liber Z, Nikolić T, Mitić B, Šatović Z (2002) Intra- and interpopulation relationships and taxonomic status of *Pinus nigra* Arnold in Croatia according to morphology and anatomy of leaves. Acta Societatis Botanicorum Poloniae 72:141-147.

Liber Z, Pavletió Z, Nikolió T, Fülgosi H (1999) DNA polymorphisms in Austrian and Dalmatian black pine. Wiss Mitt Niederösterr. Landesmuseum 12:53-62.

Liber Z, Nikolić T, Mitić B, Šatović Z (2003) RAPD markers and Black pine (*Pinus nigra* Arnold) intraspecies taxonomy- Evidence from the study of nine populations. Acta Societatis Botanicorum Poloniae 72:249-257.

Lin J-J, Kuo J (1998) A new reagent for simple isolation of plant genomic DNA. Focus 20:46-48.

Liu ZL, Zhang D, Hong D-Y, Wang X-R (2003) Chromosomal localization of 5S and 18S–5.8S–25S ribosomal DNA sites in five Asian pines using fluorescence *in situ* hybridization. Theoretical and Applied Genetics 106:198-204.

Loomis MD (1974) Overcoming problems of phenolics and quinones in the isolation of plant enzymes and organelles. Methods in Enzymology 31:528-544.

Louro V (1982) O Pinheiro Laricio (*Pinus nigra* Arnold) em Portugal. Divulgação Técnica N.º 2. Direcção Geral das Florestas. MACP.

Lubaretz O, Fuchs J, Ahne R, Meister A, Schubert I (1996) Karyotyping of three Pinaceae species via fluorescence *in situ* hybridization and computer-aided chromosome analysis. Theoretical and Applied Genetics 92:411-416.

Lucas Borja ME, Candel Pérez D, Molero Carrasco J, Ahrazem O, Gomez Gomez ML, Rubio Moraga A, Monreal Montoya J, Botella Miralles O (2013) Utilización de los marcadores ISSR para la determinación de diversidad genética en poblaciones de *Pinus nigra* Arn.. 6º Congreso Forestal Español. Montes: Servicio y desarollo rural. 10-14 junio 2013, Vitoria-Gasteiz.

Luchi N, Capretti P, Surico G, Orlando C, Pazzagli M, Pinzani P (2005) A Real-Time Quantitative PCR Assay for the detection of *Sphaeropsis sapinea* from inoculated *Pinus nigra* shoots. Journal of Phytopathology 153:37-42.

Lučić A, Isajev V, Nevenić R, Ćirković-Mitrović T, Brašanac-Bosanac L (2012) Analysis of genetic variability of Austrian Pine (*Pinus nigra* Arnold) in Serbia using protein markers. South-East European Forestry 3:1-8.

Luo C, He X-H, Chen H, Ou S-J, Gao M-P (2010) Analysis of diversity and relationships among mango cultivars using Start Codon Targeted (SCoT) markers. Biochemical Systematics and Ecology 38:1176–1184.

Luo C, He X, Chen H, Ou S, Gao M, Brown JS, Tondo CT, Schnell RJ (2011) Genetic diversity of mango cultivars estimated using SCoT and ISSR markers. Biochemical Systematics and Ecology 39:676–684.

Marques H, Pinto G, Pinto P, Teixeira C (2012) Regiões de Proveniência de Portugal. PROJECTO – DEFOR INTERREG III B SUDOE DEFOR SO2/1.3/F64, "The contribution of research for the development and Competitiveness of Southwest European forest sector", Autoridade Florestal Nacional.

Meyer W, Mitchell TG, Freedman EZ, Vilgalys R (1993) Hybridization probes for conventional DNA fingerprinting used as single primers in the polymerase chain reaction to distinguish strains of Cryptococcus neoformans. Journal of Clinical Microbiology 31:2274–2280.

Mirov NT (1967) The genus Pinus. Ronald Press, New York.

Mulpuri S, Muddanuru T, Francis G (2013) Start codon targeted (SCoT) polymorphism in toxic and non-toxic accessions of *Jatropha curcas* L. and development of a codominant SCAR marker. Plant Science 207:117-127.

Muralitharan MS, Stuart S, Graham M (1994) Methods in plant molecular biology techniques. Proceedings of Molecular Biology Workshop, Launceston, Tasmania. Published by Applied Biology, University of Tasmania, Launceston.

Murray MG, Thompson WF (1980) Rapid isolation of high molecular weight plant DNA. Nucleic Acids Research 8:4321-4326.

Muratova EN (1997) Studies on nucleolar chromosomes in representatives of Pinaceae Lindl. In: Cytogenetics of forest trees and shrub species (edited by Ž Borzan, SE Schlarbaum), pp. 45-72, Croation Forests, Inc. Zagreb and Faculty of Forestry, University of Zagreb.

Naydenov K, Tremblay FM, Fenton NJ, Alexandrov A (2006) Struture of *Pinus nigra* Arn. populations in Bulgaria revealed by chloroplast microsatellites and terpenes analysis: Provenance tests. Biochemical Systematics and Ecology 34:562-574.

Nelson CD, Kubisiak TL, Stine M (1994) A genetic linkage map of longleaf pine (*Pinus palustris* Mill.) based on random amplified polymorphic DNAs. Journal of Heredity 85:433-439.

Nkongolo KK, Michael P, Gratton (2002) Identification and characterization of RAPD markers inferring genetic relationships among Pine species. Genome 45:51-58.

Nkongolo KK, Mehes-Smith M (2012) Karyotype evolution in the Pinaceae: implication with molecular phylogeny. Genome 55:1–19.

O'Brien SL (1990) Genetics maps. Book 6. Plants Cold Spring Harbor Laboratory Press, New York.

Ozenda P (1975) Sur les étages de végétation dans les montagnes du bassin mèditerranéem. Documents Cartographique Ecologiques, Grenoble (France) 16:1e32.

Ozenda P (1985) La végétation de la chaîne alpine dans l'espace montagnard européen. Masson, Paris.

Pajares PR, Escudero A (1989) Caracterización fitocológica de las comunidades de *Pinus nigra* subsp. *salzmannii* en los afloramentos rocosos del sistema ibérico meridional. Botanica Complutensis 15:149-161.

Pavia I, Carvalho A, Rocha L, Gaspar MJ, Lima-Brito JL (2014) Physical location of SSR regions and cytogenetic instabilities in *Pinus sylvestris* chromosomes revealed by ND-FISH. Journal of Genetics 93:567-571.

Pederick LA (1967) The structure and identification of the chromosomes of *Pinus radiate* D. Don. Silvae Genetica 16:69-77.

Pedro JG (1993) Carta da distribuição do pinheiro manso, pinheiro de Alepo e outros. Atlas do ambiente, Lisboa.

Poczai P, Varga I, Laos M, Cseh A, Bell N, Valkonen JPT, Hyvönen J (2013) Advances in plant gene-targeted and functional markers: a review. Plant Methods 9:6.

Porebski S, Bailey LG and Baum BR (1997) Modification of a CTAB DNA extraction protocol for plants containing high polysaccharide and polyphenol components. Plant Molecular Biology Reporter 15:8-15.

Price RA, Liston A, Strauss SH (1998) Phylogeny and systematics of *Pinus*. In: Ecology and Biogeography of Pinus (edited by DM Richardson), pp. 49-68, Cambridge University Press, Cambridge.

Prince LM (2015) Plant genotyping using fluorescently tagged Inter-Simple Sequence Repeats (ISSRs): basic principles and methodology. In: Plant Genotyping: Methods and Protocols (edited by J Batley), pp. 63-75, Springer New York, New York.

Que Y, Pan Y, Lu Y, Yang C, Yang Y, Huang N, Xu L (2014) Genetic analysis of diversity within a Chinese Local Sugarcane Germplasm Based on Start Codon Targeted polymorphism. BioMed Research International. http://dx.doi.org/10.1155/2014/468375.

Reddy MP, Sarla N, Siddiq EA (2002) Inter simple sequence repeat (ISSR) polymorphism and its application in plant breeding. Euphytica 128:9–17.

Rezzi S, Bighelli A, Mouillot D, Casanova J (2001) Composition and chemical variability of the needle essential oil of *Pinus nigra* subsp. *laricio* from Corsica. Flavour and Fragance Journal 16:379-383.

Ribeiro T, Barão A, Viegas W, Morais-Cecílio L (2008) Molecular cytogenetics of forest trees. Cytogenetic Genome Research 120:220–227.

Richardson D (1998a) Forestry trees as invasive aliens. Conservation Biology 12:18–26.

Richardson DM (1998b) Ecology and biogeography of Pinus, Cambridge University Press, Cambridge.

Rubio-Moraga A, Candel-Perez D, Lucas-Borja ME, Tiscar PA, Viñegla, Linares JC, Gómez-Gómez L, Ahrazem O (2012) Genetic diversity of *Pinus nigra* Arn. populations in Southern Spain and Northern Morocco revealed by inter-simple sequence repeat profiles. International Journal of Molecular Sciences 13:5645-5658.

Šarac Z, Matejić JS, Stojanović-Radić ZZ, Veselinović JB, Džamić AM, Bojović S, Marin PD (2014) Biological activity of *Pinus nigra* terpenes-Evaluation of FtsZ inhibition by selected compounds as contribution to their antimicrobial activity. Computers in Biology and Medicine 54:72–78.

Scaltsoyiannes A, Rohr R, Panetos KP, Tsaktsira M (1994) Allozyme frequency distributions in five European populations of black pine (*Pinus nigra* Arnold). Silvae Genetica 43:20-30.

Schlink K, Reski R (2002) Preparing high-quality DNA from moss (Physcomitrella patens). Plant Molecular Biology Reporter 20:423a–f.

Schmidt A, Doudrick RL, Heslop-Harrison JS, Schmidt T (2000) The contribution of short repeats of low sequence complexity to large conifer genomes. Theoretical and Applied Genetics 101(1–2):7–14. doi:10.1007/s001220051442.

Schwarz O (1938) Über die Systematik und Nomenklatur der Europäischen Schwarzkiefern. Notizblatt des Bot. Garten zu Berlin Dahlem XIII 117:226–243.

Semagn K, Bjørnstad Å, Ndjiondjop MN (2006) An overview of molecular marker methods for plants. African Journal of Biotechnology 5:2540-2568.

Sevgi O, Akkemik U (2007) A dendroecological study on *Pinus nigra* Arn. at different altitudes of Northern slopes of Kazdaglari, Turkey. Journal of Environmental Biology 28:73-75.

Seyedimoradi H, Talebi R (2014) Detecting DNA polymorphism and genetic diversity in Lentil (*Lens culinaris* Medik) germplasm: comparison of ISSR and DAMD marker. Physiology and Molecular Biology of Plants 20:495-500.

Shahlaei A, Torabi S, Khhosroshahli M (2014) Efficacy of SCoT and ISSR markers in assessment of tomato (*Lycopersicum esculentum* Mill.) genetic diversity. International Journal of Biosciences 5:14-22.

Sharma AD, Gill PK, Singh P (2002) DNA isolation from dry and fresh samples of polysaccharide-rich plants. Plant Molecular Biology Reporter 20:415a–f.

Siljak-Yakovlev S, Pustahija F, Vicic V, Robin O (2014) Molecular Cytogenetics (FISH and Fluorochrome Banding): Resolving Species Relationships and Genome Organization. Methods in Molecular Biology 1115:309-23. doi: 10.1007/978-1-62703-767-9_15.

Solanille E L (1994) Localizacion cromosómica y control genético de marcadores moleculares obtenidos mediante la técnica de la reacción en cadena de ela polimerase

en centeno (Secale cereale), Tesina, Universidad Complutense de Madrid, Faculdade de Biología, Madrid.

Souframanien J, Gopalakrishna T (2004) A comparative analysis of genetic diversity in blackgram genotypes using RAPD and ISSR markers. Theorical and Applied Genetics 109:1687-1693.

Specht PL, Ashton DH, Clifford HT, Olsvig-Whittaker L, Robertson D, Scarlett N, Atkins L, Dixon, Hobbs RJ, Hopkins AJM, Lamont BB, Struthers R, Scheid GA, Zammit CA, Zedler PH, Gajardo RA, Montenegro G, Amandier L, Martinex C, Floret C, Romane F, Arianoutsou LTM, Kokkini S, Orshan G, Rabinovitz Vyn A, Quezel P, Catarino FM, Correia AIVD, Moll EJ, Merino GJ (1988) Natural vegetation ecomorphological characters. In: Mediterranean type ecosystems: A data source book. (edited by RL Specht.), pp.14-62, Kluwer Academic Publishers.

Stewart NC, Via LE (1993) A rapid CTAB DNA isolation technique useful for RAPD fingerprinting and other PCR applications. Biotechniques 14:748-749.

Thompson JN (2005) The geographic mosaic of coevolution. University of Chicago Press, Chicago, Illinois, USA.

Tibbits JFG, McManus LJ, Spokevicius AV, Bossinger G (2006) A rapid method for tissue collection and high-throughput isolation of genomic DNA from mature trees. Plant Molecular Biology Reporter 24:81-91.

Tolun AA, Velioğlu E, Çencel BN, Kaya Z (2000) Genetic Structure of Black Pine (*Pinus nigra* ARNOLD subspecies *pallasiana*) Populations Sampled from the Bolkar Mountains. Silvae Genetica 49:113-119.

Tomar RS, Parakhia MV, Patel SV, Golakiya BA (2010) Molecular markers and plant biotechnology. New India Publishing.

Tuzlacı E, Erol MK (1999) Turkish folk medicinal plants. Part II: Eğridir (Isparta). Fitoterapia 70:593–61.

Van Haverbeke, David F (1990) *Pinus nigra* Arnold European black pine. In: Burns, Russell M.; Honkala, Barbara H., technical coordinators. Silvics of North America. Vol. 1. Conifers. Agric. Handb. 654. Washington, DC: U.S. Department of Agriculture, Forest Service:395-404.

Vargas-Mendoza CF, Medina-Jaritz NB, Ibarra-Sánchez CL, Romero-Salas EA, Alcalde-Vázquez R, Rodríguez-Banderas A (2011) Phylogenetic analysis of mexican Pine species based on three *loci* from different genomes (nuclear, mitochondrial and chloroplast), Relevant Perspectives in Global Environmental Change, (edited by J Agboola), ISBN: 978-953-307-709-3, InTech, doi:10.5772/26988. Available from:http://www.intechopen.com/books/relevant-perspectives-in-global-environmental-change/phylogenetic-analysis-of-mexican-pine-species-based-on-three-*loci*-from-different-genomes-nuclear-mit.

Varma A, Padh H, Shrivastava N (2007) Plant genomic DNA isolation: An art or a science. Biotechnology Journal 2:386-392.

Vidaković M (1974) Genetics of European Black Pine (*Pinus nigra* Arnold). Annales Forestales 6/3:57-86.

Vidaković M (1991) Conifers morphology and variation, Translated from Croatian by Maja Soljan. Croatia, Graficki Zavod Hrvastske.

Villar EH (1947) Quel est le nom valable de *Pinus laricio* Poir. Bericht der Schweiz. Botan. Gesellsch, LVII 152.

Wagner DB, Furnier GR, Saghai-Maroof MA, Williams SM, Dancik BP, and Allard RW (1987) Chloroplast DNA polymorphisms in lodgepole and jack pines and their hybrids. Proceedings of the National Academy Sciences USA 84:2097-2100.

Wang B (2013) Hybridization and Evolution in the Genus Pinus. Department of Ecology and Environmental Science. Umeå University, Umeå, Sweden.

Wheeler NC, Kriebel HB, Lee CH, Read RA, Wright JW (1976) 15-year performance of European Black Pine in provenance tests North Central United States. Silvae Genetica 25:1-6.

Wright JW, Bull WI (1962) Geographic variation in European black pine – two year results. Forest Science 8:32-42.

Wu K, Jones R, Dannaeberger L, Scolnik PA (1994) Detection of microsatellite polymorphisms without cloning. Nucleic Acids Research 22:3257–3258.

Xiong F, Zhong R, Han Z, Jiang J, He L, Zhuang W, Tang R (2011) Start codon targeted polymorphism for evaluation of functional genetic variation and relationships in cultivated peanut (*Arachis hypogaea* L.) genotypes. Molecular Biology Reporter 38:3487-3494.

Xu Y (2010) Molecular Plant Breeding. CABI Publisher, Wallingford, UK.

Zaghi D (2008) Management of Natura 2000 habitats. 9530 *(Sub)-Mediterranean pine forests with endemic black pines. European Commission.

Zeng B, Zhang Y, Huang L, Jiang X, Luo D, Yin Guohua (2014) Genetic diversity of orchardgrass (*Dactylis glomerata* L.) germplasms with resistance to rust diseases revealed by Start Codon Targeted (SCoT) markers. Biochemical Systematics and Ecology 54:96-102.

Zietkiewicz E, Rafalski A, Labuda D (1994) Genome fingerprinting by simple sequence repeat (SSR)-anchored polymerase chain reaction amplification. Genomics 20:176-183.

Zlatanov T, Velichkov I, Lexer JM, Dubravec T (2010) Regeneration dynamics in aging black pine (*Pinus nigra* Arn.) plantations on the south slopes of the Middle Balkan Range in Bulgaria. New Forests 40:289-303.

Chapter II:

Single protocol for genomic DNA extraction from foliar, dry and vascular tissues of gymnosperms

---

Lemos M, Dias A, Carvalho A, Delgado A, Pavia I, Coutinho J, Gaspar MJ, Louzada JL, Lima-Brito J (2015) Single protocol for genomic DNA extraction from foliar, dry and vascular tissues of gymnosperms suitable for genetic studies. Paper Submitted to *Wood Science and Technology*

II.1 – Abstract

Most protocols of genomic DNA extraction were optimized for leaf tissue. However, harvesting leaves in trees may be very difficult, namely in gymnosperms, and the sampling of vascular tissues such as differentiating xylem and phloem could constitute a valuable alternative. Here we present a modified CTAB-based protocol that proved to be successful in the isolation of genomic DNA from foliar tissue of *Pinus pinaster* Ait., *Pinus sylvestris* L., *Pinus nigra* Arn. (dehydrated and frozen needles) and *Taxus baccata* L., and from frozen samples of differentiating xylem and phloem from *P. pinaster* and *P. sylvestris*. Regarding only the type of tissue, an initial amount of 200mg of grounded sample produced the following average DNA yields: i) 60.95µg in dehydrated needles; ii) 45.36µg in frozen needles; iii) 27.16µg in phloem; iv) 12.62µg in differentiating xylem; and v) 54.82µg in mature leaves of *T. baccata*. Despite the suboptimal values of A260/A230 ratio, the DNA proved to have enough quality to be successfully amplified, enabling the production of ISSR and SCoT markers as well as the specific amplification of the entire internal transcribed spacer (ITS) region (ca. 3,000bp) of the ribosomal DNA (rDNA) in *Pinus* sp.. Globally, this protocol could be useful for isolation of quality DNA in these four and other gymnosperms rich in polyphenols and polysaccharides; and for tree species where the foliar tissue is inaccessible even using extensible scissors; as well as for dry specimens, and presenting ability to be amplifiable. Besides, the isolation of high yields enable the production of DNA samples suitable for other downstream reactions such as enzymatic digestion and Southern Blotting, contributing for the widening of molecular genetics, taxonomic and/or phylogenetic studies in forestry species.

Keywords: Amplifiable DNA; Internal Transcribed Spacer (ITS); Inter-Simple Sequence Repeat (ISSR); *Pinus* sp.; Start Codon Targeted (SCoT) markers; *Taxus baccata* L..

II.2 – Introduction

Gymnosperms have large amounts of polyphenols, resins and terpenoids that render the isolation of genomic DNA difficult (Ziegenhagen *et al.* 1993; Telfer *et al.* 2013). Good or highquality DNA is required for several downstream molecular techniques. Certainly, none of the extraction methods published so far could be universally applicable across plant species, tissues or analytical approaches (Varma *et al.* 2007), explaining the need of improvement of available protocols to overcome specific difficulties. PCR-based molecular markers have been widely used for assessment of genetic diversity within and among populations; extrapolation of provenances; and/or estimative of gene flow, contributing to the definition of afforestation and conservation strategies (Cipriano *et al.* 2013; Sinha *et al.* 2013; Pavia *et al.* 2014 among many others). However, most of the available methods for DNA isolation and commercial extraction kits for plants were developed for leaf tissue. Sometimes due to the height of the trees, is impossible to access the foliar tissue, even using extensible scissors, and the only viable option is to sample the vascular cambium. By other hand, for estimation of phylogenies it could be need the inclusion of herbarium material (desiccated samples) which constitutes a hard material for genomic DNA isolation. In this sense, the use of a single protocol suitable for genomic DNA extraction from foliar, dry and vascular tissues of forestry species, namely gymnosperms, will turn easier the lab work, fostering the genetic and phylogenetic studies.
With this manuscript we intend to divulgate and describe an improved protocol for total genomic DNA isolation from foliar, dry and vascular tissues of four gymnosperm species (*Pinus pinaster* Ait.; *Pinus sylvestris* L.; *Pinus nigra* Arn. and *Taxus baccata* L.), based on the CTAB methods of Doyle and Doyle (1997) and Barzegari *et al*. (2010).

II.3 – Materials and methods

II.3.1 – Plant material

Adult trees of four gymnosperm species were sampled for needles and vascular tissues (Table II.1) to test the modifications performed relatively to two CTAB-based protocols previously published (Doyle and Doyle 1987; Barzegari *et al.* 2010).

Table II.1 – Gymnosperm species and respective tissue(s) tested in the optimized DNA extraction protocol

Species	Tissue
Pinus pinaster Ait.	Frozen needles
	Frozen differentiating xylem (DX)
	Frozen phloem (Ph)
Pinus sylvestris L.	Frozen needles
	Frozen differentiating xylem (DX)
	Frozen phloem (Ph)
Pinus nigra Arn.	Frozen needles
	Dehydrated needles
Taxus baccata L.	Frozen leaves

The differentiating xylem (DX) and phloem (Ph) (vascular or secondary tissues) were extracted from stems of adult *P. pinaster* and *P. sylvestris* trees during the spring and summer, according to Paiva *et al*. (2008). Paiva *et al.* (2008) suggested that outer bark may be removed from the trunk and the inner bark should be cut and removed using a knife. This sampled layer obtained will contain secondary phloem cells with attached cambium and immature secondary xylem cells in the early stages of differentiation (division and beginning of cell expansion). Then the exposed tissue on the trunk side should be scraped down to the hard lignified layer beneath (Paiva *et al*. 2008). We protected the exposed wound from fungal diseases with 'Calda Bordalesa' as described by Novaes *et al*. (2009). These tissues were immediately frozen in liquid nitrogen and conserved at -80 ºC till genomic DNA extraction.

II.3.2 – Improved protocol for genomic DNA isolation

All plant tissues (frozen and dehydrated) were ground with liquid nitrogen to achieve a very fine powder using mortar and pestle. About 200mg of pulverized sample were immediately transferred to a frozen microtube of 2mL, and the next steps were followed:

1. Add 900µL of CTAB-Lysis buffer[1] (previously warmed at 65 ºC) to each grounded sample and mix thoroughly in the vortex.
2. Place the samples in the bath at 65 ºC during 4 hours and mix occasionally every half hour.
3. Place the samples during 5 min. on ice.
4. Add 900µl of CIA solution (24:1) composed by chloroform (A0642, Applichem GmbH, Darmstadt, Germany) and isoamylalcohol (K40462879, Merck KGaA,

Darmstadt, Germany) and mix thoroughly in the vortex during 30 sec. till the visualization of a white emulsion.

5. Centrifuge at 12,662 *g* (13,200 rpm)[2] during 20 min. at 4ºC, and transfer the aqueous phase to a new microtube. Repeat steps 4 and 5.
6. Transfer the final aqueous phase to a new microtube and measure its volume.
7. Precipitate the genomic DNA by adding 1/10 of volume of 3M sodium acetate (pH 5.2) and 2.5 to 3 volumes of absolute ethanol (-20 ºC). The DNA should precipitate overnight at -20 ºC to maximize DNA yield. Alternatively, a shorter precipitation step of 2h at -20 ºC could also be performed.
8. Centrifuge for 30 min. at 4 ºC and 12,662 *g*.
9. Discard the supernatant and add 250 µL of cold (-20 ºC) washing buffer (10mM ammonium acetate (M = 77.08g/mol) and ethanol 76% in the proportion of 1:1). Mix gently. Place the samples on a -20 ºC freezer during 8 min..
10. Centrifuge for 5 min. at 4 ºC and 12,662 *g*, and air dry the DNA pellet.
11. Dissolve the DNA pellet in 50 µL of elution buffer (1x TE)[3] and 8 µL of Ribonuclease A (10 mg/ml; R5500, Sigma-Aldrich, Co. St. Louis, MO, USA) and place during 2 hours at 37ºC and then at 4 ºC till DNA quantification.

Notes:

[1]Commercial CTAB-Lysis buffer (Biochemica A4150, Applichem GmbH, Darmstadt, Germany): 20.00 g/L (2% w/v) CTAB; 7.44g/L (20mM) EDTA.Na_2.$2H_2O$ pH 8; 81.82 g/L (1.4M) NaCl and 12.11 g/L (100mM) Tris ultrapure. To this commercial CTAB-Lysis buffer, was added 2% PVP 40 (Sigma-Aldrich, Co. St. Louis, USA) and 1% of 2-mercaptoethanol (M6250, Sigma-Aldrich, Co. St. Louis, MO, USA).

[2]Centrifugations were carried on a centrifuge 5415 (Eppendorf) containing a rotor F45-24-11 with 5.5cm of radius.
[3]The elution buffer 1xTE was diluted from the commercial 100xTE buffer (Biochemica A0973, Applichem GmbH, Darmstadt, Germany).

II.3.3 – Evaluation of DNA integrity and quantification

The integrity of DNA samples was evaluated after electrophoresis at a constant voltage of 100 V on 0.8% agarose gels prepared with the 1x Tris-Borate-EDTA (TBE) buffer. In each lane 3 µL of genomic DNA and 3 µL of loading buffer (0.007% W/v of bromophenol blue, 30% V/v glycerol and 10x TBE) were loaded. The images were captured with the Molecular Imager® Gel DocTM XR+ System equipment using the Image LabTM Software - Bio-Rad® (Bio-Rad Laboratories, Inc., Hercules, California). The DNA samples were quantified in the NanoDrop™ 1000 Spectrophotometer (Thermo Scientific, Burlington, ON, Canada) and the values of the A260/A280 and A260/A230 ratios were registered to detect protein/RNA and polyphenolic/polysaccharide contaminations, respectively.

II.3.4 – Amplification and visualization of ISSR and SCoT markers

The isolated DNA samples were diluted with DNase/ RNase free distilled water (Gibco, Paisley, U.K.) to a concentration of 30ng/μL, and used as templates in PCR experiments for ISSRs and SCoTs amplification. ISSRs were amplified and visualized according to Carvalho *et al*. (2005). The reaction mixture for SCoTs amplification (final volume of 20 μL) contained: 60 ng of genomic DNA; 1x KCl reaction buffer (Thermo Scientific, Burlington, ON, Canada); 1.5mM MgCl2; 0.25mM of each dNTP; 0.8μM of SCoT primer and 0.1 units of *Taq* DNA polymerase (Thermo Scientific, Burlington, ON, Canada). SCoTs amplification and their visualization followed the procedures described by Collard and Mackill (2009). In each gel, the molecular weight marker Gene RulerTM 100 bp Plus DNA Ladder or Gene RulerTM DNA Ladder Mix, both from Thermo Scientific (Burlington, ON, Canada) was loaded.

II.3.5 – Amplification of the ITS rDNA region in *Pinus* sp.

The complete ITS region (ITS1-5.8S-ITS2) of the ribosomal DNA (rDNA) of *P. pinaster*, *P. sylvestris* and *P. nigra*, was amplified using the forward primer (5'-GTCCACTGAACCTTATCATTTAG-3') of Urbatsch *et al*. (2000) and the reverse primer ITS-4 (5'-TCCTCCGCTTATTGATATGC-3') described by White *et al*. (1990). The reaction mixture, amplification conditions and visualization of the ITS rDNA region were performed as described elsewhere (Carvalho *et al.* 2009).

II.4 – Results and discussion

The biochemical composition of plant tissues varies considerably among species (Choudhary *et al.* 2008), explaining the need of optimizing available protocols of DNA isolation for each specific plant tissue or species.

The extraction of nucleic acids from gymnosperms is a hard task. Thus, the existence of several optimized protocols for particular species and/or tissues (Cheng *et al.* 1997; Barzegari *et al.* 2010; Kejani *et al.* 2010; Lorenz *et al.* 2010; Telfer *et al.* 2013; Carvalho *et al.* 2015) is an understandable issue. Here, is presented a micro-scale extraction protocol tested in different tissues of four gymnosperms after modifications to protocols previously reported by Doyle and Doyle (1987) and/or Barzegari *et al*. (2010). Globally, the following

modifications were performed: i) use of frozen and dehydrated tissues instead of fresh leaves; ii) micro-scale procedure allowing the extraction of large number of samples per day; iii) higher volume of CTAB-Lysis buffer; iv) use of a commercial CTAB-Lysis buffer with EDTA.Na2.2H2O and higher concentrations of 2- mercaptoethanol and PVP-40; v) longer step of cell lysis and at higher temperature (maximizing DNA yield); vi) ice incubation before CIA addition; vii) higher centrifugation speed and time to ensure a cleaner/pure aqueous phase; viii) DNA precipitation with sodium acetate and absolute ethanol instead of isopropanol; ix) wash of the precipitated DNA ('pellet') at -20 ºC to improve its purification; x) higher amount of RNAse A per sample; and xi) longer incubation at 37 ºC for simultaneous RNA digestion and DNA elution.

According to Allen *et al.* (2006), usually from 200mg of fresh weight powder is possible to attain 5 to 30µg of genomic DNA. Starting with the same initial amount of grounded tissue, we achieved average values of DNA yield within this range but also higher than 30µg (Table II.2).

Table II.2 – Average concentration, purity (A_{260}/A_{280} and A_{260}/A_{230}) and yield values of the DNA samples isolated with the optimized extraction protocol

Species	Tissue	Average values of:			
		DNA concentration (ng/µL) (mean ± SD[a]) (min. and max. values)	**A_{260}/A_{280} (mean ± SD[a])**	**A_{260}/A_{230} (mean ± SD[a])**	**Yield[b] (µg per 200 mg)**
P. pinaster	Needles	599.18 ± 457.13 (min. 89.88; max. 1198.66)	1.80 ± 0.14	1.00 ± 0.39	29.96
	DX	231.25 ± 17.11 (min. 132.33; max. 320.47)	1.68 ± 0.09	0.62 ± 0.11	11.56
	Ph	565.10 ± 226.20 (min. 397.30; max. 884.86)	1.79 ± 0.04	0.56 ± 0.05	28.26
P. sylvestris	Needles	921.46 ± 335.68 (min. 454.88; max. 1230.56)	1.86 ± 0.08	1.07± 0.11	46.07
	DX	273.60 ± 23.33 (min. 242.33; max. 298.37)	1.67 ± 0.01	0.68 ± 0.03	13.68
	Ph	521.28 ± 372.59 (min. 152.48; max. 1031.60)	1.64 ± 0.11	0.63 ± 0.17	26.06
P. nigra	Needles	1200.96 ± 378.24 (min. 512.20; max. 2179.85)	1.87 ± 0.05	1.12 ± 0.16	60.05
	Dehydrated needles	1219.02 ± 228.02 (min. 947.41; max. 1565.56)	1.83 ± 0.03	0.80 ± 0.08	60.95
T. baccata	Leaves	1096.49 ± 270.15 (min. 752.80; max. 1610.35)	1.79 ±0.03	0.80 ±0.10	54.82

Notes: [a]SD – standard deviation; [b]Yield (µg) = DNA concentration (µg/µL) x 50µL.

Ostrowska *et al*. (1998) and Telfer *et al*. (2013) extracted genomic DNA from fresh needles of *Pinus radiata* using the CTAB extraction buffer proposed by Doyle and Doyle (1987; 1990). These authors reported average DNA yields of 59.66 ± 8.27µg obtained from 1g of needle tissue, and 5.85µg from 150 mg of initial amount of tissue, respectively. Considering that we used a similar CTAB-Lysis buffer and an initial amount of 200mg of sample, we may say that we achieved higher average DNA yields than these authors (Table II.2). Although the DNA yield and purity could be influenced by the type, age and quality of the plant tissue, in the present study, we obtained the higher average DNA yield in samples isolated from dehydrated needles of *P. nigra*, followed by frozen needles of the same species, frozen *T. baccata* leaves and *P. sylvestris* needles (Table II.2). The DNA samples extracted from foliar tissues also presented ideal A260/A280 ratios (Table II.2). Furthermore, Khanuja *et al.* (1999) obtained higher DNA yield in samples isolated from dry tissues comparatively to fresh ones. Although the statement that fresh or frozen plant tissues are ideal for obtaining high-quality DNA, sometimes, taxonomic, phylogenetic and evolutionary studies require the use of herbarium specimens (Drábková *et al.* 2002; DeCastro *et al.* 2004; Jankowiak *et al.* 2005). The effective isolation of DNA from dried specimens often requires CTAB extraction buffers and longer precipitation steps (Drábková *et al.* 2002; Soltis and Soltis 1993; Rogers 1994; Taylor and Swann 1994; Golenberg 1999; Cota-Sánchez *et al.* 2006). Our CTAB-based protocol also included a long precipitation step. Additionally, our protocol proved to be efficient for extraction of highly concentrated and pure genomic DNA not only from dehydrated needles, but also from frozen foliar and vascular tissues (Fig. II.1), without the need of additional modifications per type or condition of the plant tissue. We are confident that the suitability of this method for DNA isolation from DX and Ph of *Pinus* sp. constitutes a very important advantage. A similar approach was described for tissues sampling and isolation of DNA from stem bark of Leguminosae trees (Novaes *et al*. 2009). These authors deeply referred the inconvenient of sampling leaves in tall trees and of isolating genomic DNA from leaves with inappropriate phytosanitary status, and reinforced that contrastingly to leaves, bark can be easily sampled from the ground level using simple tools and shortly, being available during all the year. Although presenting the lowest average values of concentration, purity and yield (Table II.2), the DNA samples isolated from the vascular tissues were successfully used as templates for the amplification of microRNA precursors (data not shown). Moreover, despite the suboptimal values of the A260/A230 ratio (Table II.2), the isolated DNA samples proved to be suitable for the production of ISSR and SCoT markers

(Figs. II.2a-c) as well as for the specific amplification of the entire ITS rDNA region in pines (Fig. II.2d).

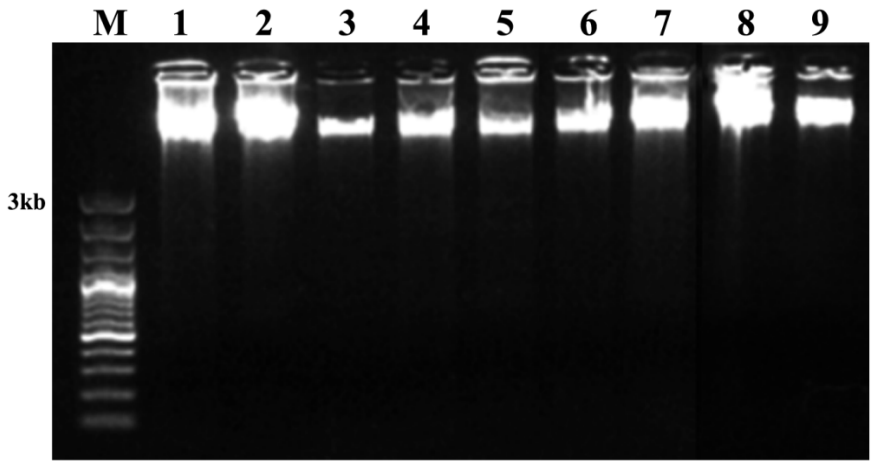

Figure II.1 – Genomic DNA samples isolated from: 1) needles; 2) Ph; and 3) DX of *P. pinaster*; 4) needles; 5) DX; 6) Ph of *P. sylvestris*; 7) frozen and 8) dehydrated needles of *P. nigra*; and 9) mature leaves of *T. baccata*, using the optimized protocol. M – Gene Ruler DNA ladder 100 bp Plus.

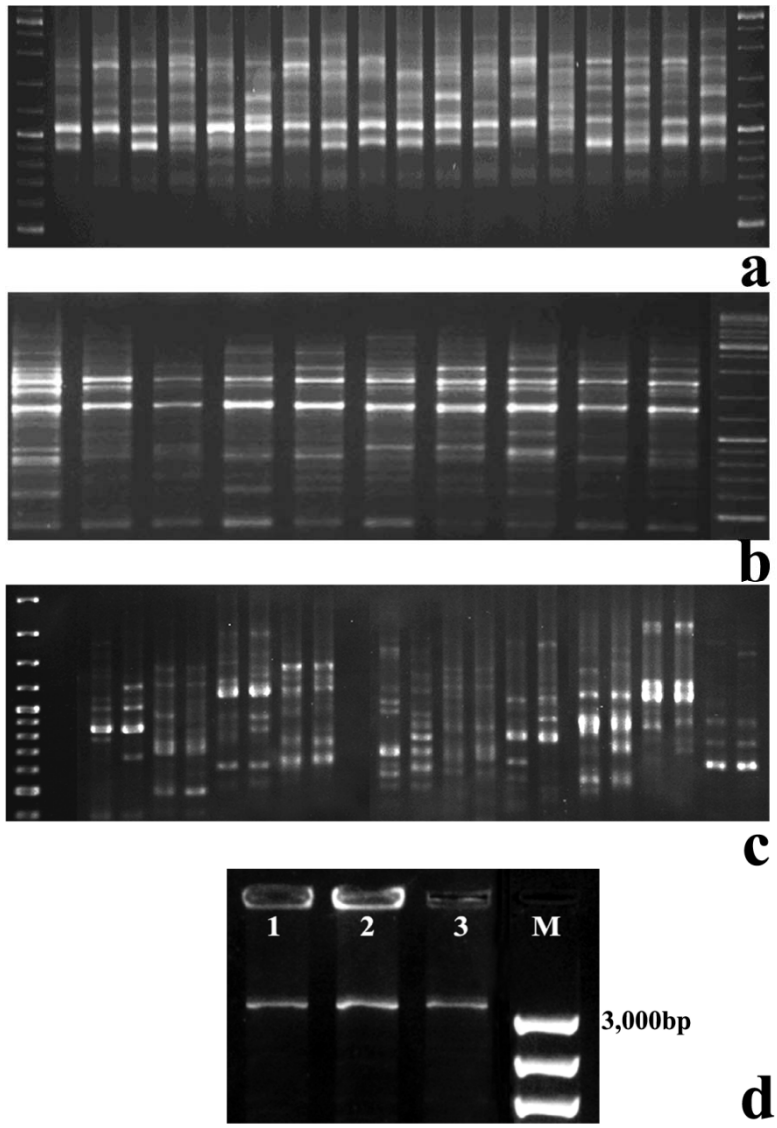

Figure II.2 – Agarose gels stained with ethidium bromide showing: a) ISSRs amplified with primer 836 (5'-AGAGAGAGAGAGAGAGYA-3' from the set9/100 of University of British Columbia) and b) SCoTs produced by primer SCoT3 (Collard and Mackill 2009), both in *P. nigra* individuals; c) Test of SCoT primers (Collard and Mackill 2009) in 20 *T. baccata* individuals; d) ITS rDNA amplicons with ca. 3,000bp, obtained in *P. pinaster* (1), *P. sylvestris* (2) and *P. nigra* (3). In the first and/or last well of each gel was loaded the molecular weight marker (M).

Polyphenolic compounds, polysaccharides, proteins and/or RNA are the main contaminants that co-precipitate with genomic DNA, conditioning the success of PCR experiments (Arif *et al.* 2010). However, the DNA samples extracted with our improved protocol proved to be amplifiable (Fig. II.2). This cost-effective protocol is highly versatile in terms of tissues and species and has been routinely used in our lab. It was also successful for the extraction of genomic DNA from cereals (e.g. tritordeum, wheat and barley), even performing a shorter incubation time (2h at 65 ºC), and high average values of DNA

concentration (2300ng/μL) and purity ratios (A260/A280=1.88 and A260/A230=1.50) were achieved. Thus, we are confident that this protocol could be also successful in other plant species and hard tissues, such as herbarium material.

II.5 – Acknowledgements

This work was partially supported by the project PTDC/AGR-CFL/110988/2009, attributed by the Portuguese Foundation for Science and Technology (Fundação para a Ciência e a Tecnologia, FCT) and co-financed by the European Fund of Regional Development (FEDER) under the scope of the COMPETE-QREN program.

Ana Carvalho and Alexandra Dias thank the SFRH/BPD/68932/2010 and SFRH/BD/91781/2012 Grants, respectively, both attributed by FCT and co-financed by the FSE/POPH-QREN program.

The authors acknowledge to Prof. Maria Margarida Ribeiro (Department of Natural Resources and Sustainable Development, Instituto Politécnico de Castelo Branco, Portugal) and to Prof. Bruno Fady (Unit of Ecology of Mediterranean Forests, INRA, Avignon, France) for kindly provided *T. baccata* leaves and *P. nigra* dehydrated needles, respectively.

II.6– References

Allen GC, Flores-Vergara MA, Krasnyanski S, Kumar S, Thompson WF (2006) A modified protocol for rapid DNA isolation from plant tissues using cetyltrimethylammonium bromide. Nature Protocols 1:2320-2325.

Arif IA, Bakir MA, Khan HA, Ahamed A, Al Farhan AH, Al Homaidan AA, Al Sadoon M, Bahkali AH, Shobrak M (2010) A simple method for DNA extraction from mature date palm leaves: impact of sand grinding and composition of lysis buffer. International Journal of Molecular Sciences 11:3149-3157.

Barzegari A, Vahed SZ, Atashpaz S, Khani S, Omidi Y (2010) Rapid and simple methodology for isolation of high quality genomic DNA from coniferous tissues (*Taxus baccata*). Molecular Biology Reporter 37:833-837.

Carvalho A, Matos M, Lima-Brito J, Guedes-Pinto H, Benito C (2005) DNA fingerprint of F_1 interspecific hybrids from the *Triticeae* tribe using ISSRs. Euphytica 143:93-99.

Carvalho A, Guedes-Pinto H, Lima-Brito J (2009) Genetic diversity among old Portuguese bread wheat cultivars and botanical varieties evaluated by ITS rDNA PCR-RFLP markers. Journal of Genetics 88(3):363-367.

Carvalho A, Graça C, Carocha V, Pêra S, Lousada JL, Lima-Brito J, Paiva JAP (2015) An improved total RNA isolation from secondary tissues of woody species for coding and non-coding gene expression analyses. Wood Science and Technology 49:647-658.

Cheng FS, Brown SK, Weeden NF (1997) A DNA extraction protocol from various tissues in woody species. Horticultural Science 32:921-922.

Choudhary K, Mathur N, Choudhary OP, Pillai U (2008) Protocol for isolation of genomic DNA from dry and fresh leaves of *Vigna* species suitable for RAPD and restriction digestion. Advanced Biomedical Research 2:83-89.

Cipriano J, Carvalho A, Fernandes C, Gaspar MJ, Pires J, Bento J, Roxo L, Louzada J, Lima-Brito J (2013) Evaluation of genetic diversity of Portuguese *Pinus sylvestris* L. populations based on molecular data and inferences about the future use of this germplasm. Journal of Genetics 92:e41-e48.

Collard BCY, Mackill DJ (2009) Start Codon Targeted (SCoT) polymorphism: A simple novel DNA marker technique for generating gene-targeted markers in plants. Plant Molecular Biology Reporter 27:86–93.

Cota-Sánchez JH, Remarchuk K, Ubayaena K (2006) Ready-to-use DNA extracted with a CTAB method adapted for herbarium specimens and mucilaginous plant tissue. Plant Molecular Biology Reporter 24:161–167.

DeCastro O, Mendle B (2004) PCR amplification of Michele Teniere's historical specimens and facility to utilize an alternative approach to resolve taxonomic problems. Taxon 53:147-151.

Doyle JJ, Doyle JL (1987) A rapid DNA isolation procedure for small quantities of fresh leaf tissue. Phytochemical Bulletin 19:11-15.

Doyle JJ, Doyle JL (1990) Isolation of plant DNA from fresh tissue. Focus 12:12-15.

Drábková L, Kirschner J, Vlcek C (2002) Comparison of seven DNA extractions in amplification protocols in historical herbarium specimens of *Juncaceae*. Plant Molecular Biology Reporter 20:161-175.

Golenberg EM (1999) Isolation, identification, and authentication of DNA sequences derived from fossil material. In: Fossil Plants and Spores: Modern Techniques (edited by TP Jones, NP Rowe), pp. 156-160, The Geological Society, London.

Jankowiak K, Buczkowska K, Szweykowka-Kulinska S (2005) Successful extraction of DNA from 100-year-old herbarium specimens of the liverwort *Bazzania trilobata*. Taxon 54:335-336.

Kejani AA, Tafreshi SAH, Nekouei SMK, Modif MR (2010) Efficient isolation of high quality nucleic acids from different tissues of *Taxus baccata* L. Molecular Biology Reporter 37:797–800.

Khanuja SPS, Shasany AK, Darokar MP, Kumar S (1999) Rapid isolation of DNA from dry and fresh samples of plants producing large amounts of secondary metabolites and essential oils. Plant Molecular Biology Reporter 17:1–7.

Lorenz WW, Yu Y-S, Dean JFD (2010) An improved method of RNA isolation from Loblolly Pine (*P. taeda* L.) and other conifer species. Journal of Visualized Experiments. doi: 10.3791/1751.

Novaes RML, Rodrigues JG, Lovato MB (2009) An efficient protocol for tissue sampling and DNA isolation from the stem bark of Leguminosae trees. Genetics and Molecular Research 8:86-96.

Ostrowska E, Muralitharan M, Chandler S, Volker P, Hetherington S, Dunshea F (1998) Technical review optimizing conditions for DNA isolation from *Pinus radiata*. In Vitro Cellular and Developmental Biology - Plant 34:108-111.

Paiva JAP, Garcés M, Alves A, Garnier-Géré P, Rodrigues JC, Lalanne C, Porcon S, Le Provost G, Da Silva Perez D, Brach J, Frigerio J-M, Claverol S, Barré A, Fevereiro P, Plomion C (2008) Molecular and phenotypic profiling from the base to the crown in maritime pine wood-forming tissue. New Phytologist 178:283–301.

Pavia I, Mengl M, Gaspar MJ, Carvalho A, Heinze B, Lima-Brito J (2014) Preliminary evidences of two potential native populations of *Pinus sylvestris* L. in Portugal based on nuclear and chloroplast SSR markers. Austrian Journal of Forest Science 131:1-22.

Rogers SO (1994) Phylogenetic and taxonomic information from herbarium and mummified DNA. In: Conservation of plant genes II.: Utilization of ancient and modern DNA. Miss Bot Gard, Monogr Vol. 48 (edited by RP Adams, J Miller, E Golenberg, JE Adams).

Sinha D, Singh J, Tandon PK, Kakkar P (2013) Genetic diversity of *Pinus roxburghii* Sarg. collected from different Himalayan regions of India assessed by Random Amplified Polymorphic DNA analysis. Toxicology International 20:208-213.

Soltis PS, Soltis PE (1993) Ancient DNA: Prospects and limitations. New Zealand Journal of Botany 31:203-209.

Taylor JW, Swann EC (1994) DNA from herbarium specimens. In: Ancient DNA (edited by B Hermann, S Hummel), pp 166-181, Springer-Verlag, New York.

Telfer E, Graham N, Stanbra L, Manley T, Wilcox P (2013) Extraction of high purity genomic DNA from pine for use in a high-throughput Genotyping Platform. New Zealand Journal of Forest Science 43:3.

Urbatsch LE, Baldwin B, Donoghue MJ (2000) Phylogeny of coneflowers and relatives (*Heliantheae: Asteraceae*) based on nuclear rDNA internal transcribed spacer (ITS) sequences and chloroplast DNA restriction site data. Systematic Botany 25:539–565.

White TJ, Bruns T, Lee S, Taylor J (1990) Amplification and direct sequencing of fungal ribosomal genes for phylogenetics. In: PCR protocols: a guide to methods and applications (edited by MA Innis, DH Gelfand, JJ Sninsky and TJ White), pp. 315-322, Academic Press, Orlando, USA.

Varma A, Padh H, Shrivastava N (2007) Plant genomic DNA isolation: An art or a science. Biotechnology Journal 2:386-392.

Ziegenhagen B, Guillemaut P, Scholz F (1993) A procedure for mini-preparations of genomic DNA from needles of silver fir (*Abies alba* Mill.). Plant Molecular Biology Reporter 11:117-121.

Chapter III:

Genetic profiles of European allopatric subspecies and varieties of *Pinus nigra* Arn. based on ISSR and SCoT markers

---

Dias A, Lemos M, Carvalho A, Fady Bruno, Louzada JL, Gaspar MJ, Lima-Brito J (2015) Genetic profiles of European allopatric subspecies and varieties of Pinus nigra Arn. based on ISSR and SCoT markers. Paper Submitted to *Plant Systematics and Evolution*

III.1 – Abstract

Pinus nigra Arnold is a Mediterranean pine with a large distribution area, presenting high morphological, physiological and ecological variation, which has led to identification and taxonomic classification uncertainties. The present work concerns the molecular characterization of *P. nigra* populations from distinct geographical locations that were previously classified as different infraspecific *taxa*, namely, subspecies and varieties. Populations from different geographic origins (France, Italy, Austria and Romania) were chosen and are representative of three allopatric subspecies and three regional varieties. The comparison of their molecular patterns was performed for the DNA fingerprinting and estimation of genetic relationships based on inter-simple sequence repeat (ISSR) and start codon targeted (SCoT) markers, to infer about the adopted taxonomic classification. According to the ISSR and SCoT data, the studied *P. nigra* provenances presented 76% of genetic similarity and were clustered into four groups and one branch considering a cut off value of 0.876 in the UPGMA dendrogram. These groups discriminated the bulked individuals per subspecies and variety except for the samples of subspecies *nigra* which were clustered in two different groups, suggesting that the studied individuals could belong to different varieties. The subspecies *salzmannii* was projected as a branch in the dendrogram, indicating its higher genetic dissimilarity relatively to the remaining bulked samples. In general, ISSRs and SCoTs revealed to be reliable marker systems for DNA fingerprinting, estimation of genetic relationships and discrimination of infraspecific *taxa* of *P. nigra*.

Keywords: DNA fingerprinting; infraspecific *taxa*; inter-simple sequence repeat (ISSR); *P. nigra*; start codon targeted (SCoT) polymorphism.

III.2 – Introduction

Pinus nigra Arnold is a coniferous species belonging to the group of Mediterranean pines, and is an economically and ecologically important pine species in Europe (Afzal-Raffi and Dodd 2007). This species has a fragmented distribution through the Northern Africa, Northern Mediterranean and eastwards to the Black Sea, Corsica and Sicily islands (Afzal-Raffi and Dodd 2007). *P. nigra* populations are distributed through a wide range of habitats and those studied so far, have presented high variation at morphological, physiological, biochemical, genetic, ecological and heterochromatin levels (Scaltsoyiannes *et al.* 1994; Raffi *et al.* 1996; Bogunić *et al.* 2003; 2011; Afzal-Raffi and Dodd 2007; del Cerro Barja *et al.* 2009; Rubio-Moraga *et al*. 2012). This high variation might be explained by their exposure to different glacial episodes and geological events over the years (Thompson 2005). Besides, such variation has raised some issues in terms of their identification and taxonomic classification (Scaltsoyiannes *et al*. 1994; Afzal-Raffi and Dodd 2007; Bogunić *et al.* 2007; del Cerro Barja *et al*. 2009; Akkemik *et al.* 2010).

Some authors considered *P. nigra* as an aggregate of microgeographical species (Pajares and Escudero 1989; Blanco *et al*. 1997; Richardson 1998). Others regarded it as a single species subdivided into allopatric subspecies, varieties or races (Vidaković 1974; Gaussen *et al.* 1993; Barbéro *et al.* 1998;). The infraspecific taxonomic classification of *P. nigra* has been revised by several authors over the years, and involved more than 85 different Latin names, due to the consideration of regional varieties for each subspecies (Christensen 1997).

The most commonly accepted infraspecific taxonomic classification of *P. nigra* is that proposed by Gaussen *et al.* (1993) which divided the species into five subspecies based on their distribution through the Mediterranean basin, length and stiffness of the needles. This author proposed the following five subspecies i) *salzmannii* (Dunal) Franco; ii) *nigra* Arnold; iii) *dalmatica* (Visiani) Franco; iv) *pallasiana* (Lambert) Holmboe; and v) *laricio* (Poiret) Maire. Four of these subspecies were also subdivided into regional varieties: i) *P. nigra* subsp. *salzmannii* var. *mauritanica*, *hispanica*, *salzmannii* or *cebennensis*; ii) *P. nigra* subsp. *nigra* var. *austriaca*, *illyrica*, *pindica* or *italica*; iii) *P. nigra* subsp. *pallasiana* var. *banatica*, *tatarica*, *caramanica* and *fenzlii*; and iv) *P. nigra* subsp. *laricio* var. *calabrica* or *corsicana* (see Christensen 1997; Barbéro *et al.* 1998; Price *et al.* 1998; Bussotti 2002).

In 1998 the USDA Agricultural Research Service recognized a sixth subspecies, *P. nigra* subsp. *mauretanica* (Maire and Peyerimh) Heywood, as being locally assigned to the populations of North Africa, and it still be accepted by the USDA Germplasm Resources Information Network (GRIN) (Bussotti 2002).

Taxonomic classifications based on some cone and/or needle characteristics overlap extensively throughout its distribution (Boydak 2001).

P. nigra is widely used for reforestation of difficult soils in the mountains of the Mediterranean region. However, the genetic structure of this species has not been taken into consideration in some areas, and due to the easy hybridization among subspecies, the plantation of European Black Pines of unknown origin in the proximity of autochthonous pinewoods should be avoided in order to reduce the risk of genetic pollution (Naydenov *et al.* 2006; Zaghi 2008).

Genetic variation is fundamental for the maintenance and long term stability of forest ecosystems, since it determines the adaptive potential of the forestry species to the environmental conditions and climatic changes (Muller-stark *et al.* 1992; Çengel *et al.* 2012). It is also an important factor to develop effective conservation strategies and germplasm management (Lucić *et al.* 2010) which could be achieved with the contribution of genetic studies that could allow at some extent the discrimination among subspecies or varieties.

When applied to large geographic areas, the DNA fingerprinting could allow the differentiation among populations and extrapolation of their origins, contributing for designing strategies of germplasm conservation and management and definition of plant breeding programs based on molecular profiles (Xie *et al*. 2015). Regarding the usefulness of molecular markers for phylogenetic studies and overcoming of taxonomic and systematic issues (Poczai *et al*. 2013), the DNA fingerprinting and molecular characterization of allopatric *P. nigra* populations could contribute for the identification or discrimination of infraspecific *taxa*. Some studies were already developed in that sense, using random amplified polymorphic DNA (RAPD) markers in Croatian *P. nigra* populations (Liber *et al*. 2003) and chloroplast DNA microsatellites (Naydenov *et al.* 2006; Afzal-Rafii and Dodd 2007). Croatian populations of *P. nigra* studied by Liber *et al.* (2003) were found to be undoubtedly different from the subspecies *nigra* and *pallasiana* (Austria and Turkey populations, respectively). Among the studied Croatian populations, Liber *et al.* (2003) proposed the existence of subsp. *illyrica*, subsp. *dalmatica* and a third population that may resulted of gene flow between the former two. Naydenov *et al.* (2006) used chloroplast microsatellites and

terpene analysis to study genetic structure and diversity of European Black Pine populations in Bulgaria. Also, Afzal-Rafii and Dodd (2007), used chloroplast DNA microsatellite markers to discriminated populations by sampling regions evidencing a defined genetic structure. This regional structure was supported by biogeographical analysis detecting five barriers, mainly between Alps, Corsica and Southern Italy, and Southern Spain from the Pyrenees (Afzal-Rafii and Dodd 2007). Assuming a high mutation rate, the authors inferred that European Black Pine populations existed in Western Europe through the Last Glacial Maximum (Afzal-Rafii and Dodd 2007). Rubio-Moraga *et al.* (2012) studied the genetic variation within and among seven populations of *P. nigra* from Southern Spain and Northern Morocco using inter-simple sequence repeat (ISSR) markers, in order to infer about strategies of conservation and restoration.

ISSRs are highly reproducible and polymorphic markers, allowing the detection of DNA variability at different levels, from single base changes to deletions and insertions (López-Aljorna *et al*. 2007; Naik *et al*. 2009; Chezhian *et al*. 2010). These markers present the advantage of not requiring any previous knowledge of the target DNA sequence for polymorphism detection (Gupta *et al.* 1994; Zietkiewicz *et al*. 1994; Godwin *et al.* 1997).

ISSRs have been widely used in several plant species, alone or in combination with other markers, e.g. the start codon targeted (SCoT) markers, for DNA fingerprinting; assessment of genetic diversity; estimation of phylogenies; extrapolation of provenances; gene-tagging, among other applications (Ammiraju *et al*. 2001; Culley and Wolfe 2001; Cabo *et al*. 2014a,b; Carvalho *et al*. 2005; 2009a,b; 2012; Cipriano *et al*. 2013; Coutinho *et al.* 2014a,b; Gorji *et al*. 2011; Xie *et al*. 2015).

The SCoT polymorphism technique was developed by Collard and Mackill (2009) and each SCoT primer was designed for a short conserved sequence of the start codon (ATG) of the plant genes. Thus, each SCoT amplified product may correspond to coding regions. Since the development of SCoTs by Collard and Mackill (2009), these markers have been widely used in different plant species with different purposes, such as genetic diversity assessment (Chen *et al*. 2010; Luo *et al.* 2010; 2011; Gorji *et al.* 2011; Xiong *et al.* 2011; Amirmoradi *et al.* 2012; Guo *et al.* 2012; Bhattacharyya *et al.* 2013; Mulpuri *et al.* 2013; Pakseresht *et al.* 2013; Alikhani *et al.* 2014; Chen and Liu 2014; Jiang *et al.* 2014; Que *et al.* 2014; Shahlaei *et al.* 2014; Zeng *et al.* 2014; Zhang *et al.* 2015); clonal identification (Luo *et al.* 2010); DNA fingerprinting (Collard and Mackill 2009; Gorji *et al.* 2011; Cabo *et al.* 2014b); estimation of genetic relationships (Luo *et al.* 2010; Gorji *et al.* 2011; Xiong *et al.* 2011; Amimoradi *et al.*

2012; Mulpuri *et al.* 2013; Zhang *et al.* 2015); and validation of genetic homogeneity (Agarwal *et al.* 2015).

In order to achieve these purposes, SCoTs markers could be used alone (Collard and Mackill 2009; Chen *et al*. 2010; Luo *et al.* 2010; Xiong *et al.* 2011; Guo *et al.* 2012; Bhattacharyya *et al.* 2013; Mulpuri *et al.* 2013; Cabo *et al.* 2014b; Jiang *et al.* 2014; Zeng *et al.* 2014; Zhang *et al.* 2015) or in combination with other marker systems, such as RAPDs (Gorji *et al.* 2011; Agarwal *et al.* 2015), ISSRs (Gorji *et al.* 2011; Luo *et al.* 2011; Amimoradi *et al.* 2012; Pakseresht *et al.* 2013; Alikhani *et al.* 2014; Shahlaei *et al.* 2014; Agarwal *et al.* 2015), directed amplification of minisatellite-region DNA-polymerase chain reaction (DAMD-PCR) (Amimoradi *et al.* 2012; Pakseresht *et al.* 2013), inter-retrotransposon amplified polymorphism (IRAP) (Alikhani *et al.* 2014) or inter primer binding site (iPBS) (Chen and Liu 2014).

Most of the studies performed with SCoT markers relied on non-forestry species, except for that of Alikhani *et al*. (2014) who assessed the genetic variation and differentiation in some populations of the oak *Quercus brantii*, and our ongoing molecular characterization of Portuguese allochthonous populations of *P. nigra* (unpublished data).

With this work, we intended to molecularly characterize bulked DNA samples of *P. nigra* from distinct geographical regions and belonging to different infraspecific *taxa* (allopatric subspecies and varieties); to estimate their genetic relationships and to evaluate the potential of both ISSR and SCoT markers for the discrimination of infraspecific *taxa*; and to confirm or decline the actual adopted taxonomic classification.

III.3 – Materials and methods

III.3.1 – Plant material

A total of 120 *P. nigra* individuals from 12 different geographical origins were used in the present study (Table III.1). Seeds from seven provenances were allowed to germinate in the dark at 25 ºC for three days in moistened filter paper. After three weeks, young needle tissue was harvested, immediately frozen in liquid nitrogen and maintained at -80 ºC till genomic DNA extraction. From the remaining five geographical origins, dehydrated needle tissue was maintained at room temperature and used for the isolation of genomic DNA.

Regarding the *P. nigra* taxonomical classification proposed by Gaussen *et al*. (1964; 1993), the plant material from 12 provenances used in this study was representative of four allopatric subspecies and three regional varieties (Table III.1).

Table III.1 – Characterization of the plant material used in this study: infraspecific taxonomic classification (according Gaussen *et al*. 1964; 1993), geographic origin and coordinates, provenance code, type of population where the samples were harvested, and code of the DNA bulked samples

Infraspecific taxonomic classification of *P. nigra*		Geographic origin	Geographic coordinates	Provenance code	Type of population	Code of DNA bulked samples
Subspecies	Variety					
laricio	*corsicana*	Ponteils et Brésis (France)	44°24'18''N; 3°58'39''E	PL0902	Mas de L'Aire National Forest (indigenous stand)	LCor1
		Saint Denis de Catus (France)	-	PL0-VG-02	Seed orchard	LCor2
		Yvoy le Marron (France)	47°38'01''N; 1°51'12''E	PL0901	(Stand)	LCor3
		Lavercantière (France)	-	PL0-VG-01	Seed orchard	LCor4
		Lisle sur Tarn (France)	-	PLA-VG-02	Seed orchard	LCal5
	calabrica	Cosenza (Calabria, Italy)	39º15'00''N; 16º17'00''E	COSE	Provenance test	LCal6
		Aspromonte (Calabria, Italy)	38º05'00''N; 16º00'00''E	ASPR	Provenance test	LCal7
nigra	*austriaca*	Valbelle (France)	44°09'10''N; 5°52'30''E	PNI902	Jabron National Forest (indigenous stand)	NAus8
		Villers la Faye (France)	47°06'47''N; 4°51'59''E	PNI901	Villers la Faye Community Forest (Stand)	NAus9
	----	Doblhoff (Austria)	48º00'22,83''N; 16º13'49,77''E	DOBL	Provenance test	Nig10
	----	Banat (Romania)	44º52'00''N; 22º25'00''E	BANA	Provenance test	Nig11
salzmannii	----	Saint Guilhem le Désert (France)	43º46'00''N; 3º34'00''E	PNS	Natural population	Salz12

III.3.2 – Genomic DNA extraction

The genomic DNA was extracted with a CTAB protocol that was based on that of Doyle and Doyle (1987) with some modifications.

After evaluation of the integrity and quantification of the genomic DNA samples, we diluted them with distilled ultra-pure water to the concentration of 30ng/µL.

The diluted samples of genomic DNA of ten individuals per *P. nigra* provenance were pooled, using the same volume from each individual (Furman *et al.* 1997; Liber *et al.* 2003), achieving a final DNA concentration of 30ng/µL per provenance.

III.3.3 – Amplification of ISSR and SCoT markers

For the ISSRs amplification a total of 15 primers were tested: seven from the set #9 of the University of British Columbia (UBC) that were previously used in *P. sylvestris* (Cipriano *et al.* 2013) and eight primers reported by Rubio-Moraga *et al.* (2012) for *P. nigra*. Each reaction mixture (final volume of 20 µL) contained 2 µL of total genomic DNA (30ng/µL), 1 µL of primer (5 µM), 10 µL of Taq-PCR master mix (Qiagen) and 7 µL of ultra-pure distilled water (Qiagen). The ISSR amplifications were carried on a T-Professional (Biometra) thermal cycler using: an initial denaturation step of 94 ºC for 5 min, followed by 45 cycles of denaturation of 94ºC for 30 sec, primer annealing at 47 ºC for 45 sec and extension at 72 ºC for 2 min, and a final extension of 72 ºC for 5 min.

For the amplification of SCoTs we tested the 36 primers developed by Collard and Mackill (2009). The reaction mixture (final volume of 20µL) for amplification of these markers was constituted by: 60ng/µL of genomic DNA; 3.2µL of SCoT primer (5µM); 0.25mM of dNTPs mixture; 1x reaction buffer with KCl (Thermo Scientific, Burlington, USA); 1.5mM of $MgCl_2$ (Thermo Scientific, Burlington, USA); and 2 units of *Taq* DNA polymerase (5U/µL) (Thermo Scientific, Burlington, USA). The conditions for the amplification of SCoTs were those reported by Collard and Mackill (2009).

Both the ISSR and SCoT amplified products were visualized after electrophoresis performed on 2% agarose gels stained with ethidium bromide. The electrophoresis occurred at a constant voltage of 100V. In each gel was loaded the molecular weight marker GeneRuler DNA Ladder Mix (Thermo Scientific, Burlington, USA). The gel images were captured on the Gel Doc™ XR+ equipment (BIO-RAD, Hercules, USA) using the Image Lab Software (BIO-RAD, Hercules, USA). Only the reproducible bands were considered for the presence (1)/absence (0) analyses in order to construct a binary matrix. Each ISSR or SCoT band was

considered a marker, and bands with the same molecular weight produced by the same primer were considered as being the same *locus*.

III.3.4 – Statistical analyses of the molecular data

The pool of the ISSR and SCoT data was used for the elaboration of a single binary matrix. This matrix was used for the construction of an UPGMA (Unweighted Pair Group Method with Arithmetic means) dendrogram of genetic similarity using the software NTSYS pc version 2.02 (Rohlf 1998), the DICE coefficient and the Sequential Agglomerative Hierarchical Nested (SAHN) cluster analysis module. This approach allowed the estimation of the genetic relationships among the bulked DNA samples corresponding to *P. nigra* individuals from 12 distinct provenances.

A Principal Coordinates Analysis (PCoA) was also performed with the software GenALEX 6.501 (Peakall and Smouse 2012).

The software POPGENE 1.32 (Yeh *et al.* 1999) enabled the calculation of the following indexes: i) the Shannon's information index (I) which measures the gene diversity (Shannon and Weaver 1949); ii) the Nei's gene diversity index ($h = 1 - \sum pi^2$, where p_i is the frequency of the i^{th} allele at the *locus*; Nei 1973); iii) the total genetic diversity (H_T); iv) the genetic diversity within subspecies (H_S); v) the relative magnitude of differentiation among subspecies (G_{ST}) (Nei 1987); and vi) the inter-subspecies genetic diversity ($D_{ST} = H_T - H_S$).

The Tree View software (Page 1996) was used to obtain an unrooted tree based on the Nei's unbiased measures of genetic identity and genetic distance (Nei 1978) among the 12 *P. nigra* provenances.

III.4 – Results and discussion

III.4.1 – ISSR and SCoT polymorphism

Among the primers tested for the amplification of ISSRs, we selected eight: one previously used in *P. nigra* by Rubio-Moraga *et al*. (2012) and which corresponds to the primer 866 from the UBC set 9/100, and seven previously used in *P. sylvestris* by Cipriano *et al*. (2013) (Table III.2).

After testing the 36 SCoT primers developed by Collard and Mackill (2009), we selected eight based on the production of discriminative patterns among the *P. nigra* DNA bulks (Table III.2).

The total percentage of polymorphism among the 12 bulked DNA samples was higher with the ISSR markers (Table III.2). However, the SCoTs amplified the highest number of bands (Table III.2).

Table III.2 – Number of total amplified bands (T), monomorphic bands (M), polymorphic bands (P), exclusive bands (E), percentage of polymorphism (%P) and range size (R) of ISSR and SCoT markers amplified per primer. Note:*Y=C or T

Marker	Primer [sequence 5'→3']	T	M	P	E	%P	R (bp)
ISSR	ISSR 817 [$(CA)_8$A]	14	2 (1,800; 1,700bp)	12	0	85.71	1,200-5,000
	ISSR 827 [$(AC)_8$G]	15	1 (2,400bp)	14	0	93.33	1,000-5,000
	ISSR 834 [$(AG)_8$YT*]	16	3 (1,200; 750; 700bp)	13	0	81.25	700-2,700
	ISSR 835 [$(AG)_8$YC*]	8	2 (2,000; 1,100bp)	6	0	84.62	900-2,500
	ISSR 836 [$(AG)_8$YA*]	6	4	2 (1,500; 1,200bp)	0	33.33	900-1,800
	ISSR 841 [$(GA)_8$YC*]	9	3 (1,700; 900; 800bp)	6	0	66.67	450-2,000
	ISSR 850 [$(GT)_8$YC*]	9	7	2 (2,200; 800 bp)	0	22.22	800-3,500
	ISSR 866 [$(CTC)_6$]	13	8	5	0	38.46	500-2,500
	Total	**90**	**30**	**60**	**0**	**66.67**	**450-5,000**
SCoT	SCoT3 [CAACAATGGCTACCACCG]	24	10	14	1 (550bp)	58.33	400-4,000
	SCoT7 [CAACAATGGCTACCACGG]	16	6	10	0	62.50	500-2,500
	SCoT17 [ACCATGGCTACCACCGAG]	11	4	7	0	63.64	900-3,000
	SCoT18 [ACCATGGCTACCACCGCC]	19	4	15	1 (2,400bp)	78.95	350-2,400
	SCoT25 [ACCATGGCTACCACCGGG]	20	0	20	2 (400; 350bp)	100	350-4,000
	SCoT26 [ACCATGGCTACCACCGTC]	13	5	8	0	61.54	700-2,500
	SCoT27 [ACCATGGCTACCACCGTG]	19	10	9	0	47.37	300-3,000
	SCoT31 [CCATGGCTACCACCGCCT]	22	10	12	2 (1,900; 400bp)	54.55	400-3,000
	Total	**144**	**49**	**95**	**6**	**65.60**	**300-4,000**

The percentage of ISSR polymorphism achieved here was lower than that reported by Cipriano *et al.* (2013) who used almost the same primers in *P. sylvestris* as well as RAPDs. In addition, the study developed by those authors involved a higher number of individuals. Contrastingly, the percentage of ISSR polymorphism obtained in the present study was higher than that reported by Li *et al.* (2005) in *P. sylvestris.* The ISSR markers have been proved to be highly discriminative among *Pinus* individuals from different species, populations or provenances (Li *et al.* 2005; Zhang *et al.* 2005; Feng *et al.* 2006; Labra *et al.* 2006; Rubio-Moraga *et al.* 2012; Cipriano *et al.* 2013). Despite this study was based on bulked DNA samples, most of the primers used proved to be highly discriminative in the molecular characterization of Portuguese *P. sylvestris* individuals, suitable for their clustering by population and for provenances extrapolation (Cipriano *et al*. 2013).

All primers except for the SCoT25 produced monomorphic bands (Table 2). These monomorphic bands were amplified in the 12 DNA bulked samples representative of 12 distinct provenances. By other hand, four SCoT primers amplified polymorphic exclusive bands (Tables III.2 and III.3). We adopted the term exclusive instead of unique bands since our study was performed in bulked DNA samples, and we could not ascertain if the amplified marker corresponds to a single or to all pooled individuals. Further studies will be developed in separated individuals to determine if those exclusive bands are diagnostic of a particular infraspecific *taxon* or only individual-specific. The SCoT markers produced six exclusive bands with the primers SCoT3, SCoT18, SCoT25 and SCoT31 in *P. nigra* subsp. *nigra* var. *austriaca* (NAust8), *P. nigra* subsp. *laricio* var. c*orsicana* (LCor1 and LCor2) and *P. nigra* subsp. *salzmannii* (Salz12) (Table III.3). In addition, some bands were exclusively present in the bulks representative of a single subspecies and/or variety (Table III.3).

Table III.3 – Infraspecific *taxa* and respective bulked DNA samples that presented exclusive polymorphic ISSR and SCoT markers with different molecular weight (bp)

Infraspecific *taxa*		**Code of DNA bulked samples**	**Molecular weight of exclusive polymorphic markers (bp)**	
Subspecies	**Variety**		**ISSR**	**SCoT**
laricio	*corsicana*	LCor1	-	$SCoT18_{2,400bp}$
		LCor2	-	$SCoT25_{400bp}$
nigra	*austriaca*	NAus8	$ISSR841_{1,300bp}$	$SCoT3_{550bp}$
				$SCoT31_{1,900bp}$
			$ISSR835_{2,500bp}$	$SCoT31_{400bp}$
				$SCoT26_{1,100bp}$
		NAus9	$ISSR841_{1,300bp}$	$SCoT26_{1,100bp}$
			$ISSR835_{2,500bp}$	
	-	Nig10	$ISSR841_{1,300bp}$	$SCoT27_{750bp}$
		Nig11	$ISSR841_{1,300bp}$	$SCoT27_{750bp}$
salzmannii	-	Salz12	-	$SCoT25_{350bp}$

Such markers might be useful for the future discrimination among the different European Black Pine allopatric subspecies and varieties. The amplification of unique ISSR bands in oak species was reported as being helpful, since they could represent putative species-diagnostic markers (Carvalho *et al.* 2009b).

The average of bands amplified per primer was 11.25 for the ISSR markers and 18 for SCoTs. Primers 834 and 827 showed the highest numbers of total amplified ISSR bands, while primer 835 showed the lowest one (Table III.2). According to Carvalho *et al.* (2009a,b), these data could allow a partial understanding about which sequences are more abundant in

the genomes under study. In that sense, regarding the repetition motifs of primers 827 and 834, $(AC)_8G$ and $(AG)_8YT$, respectively, we may say that the *P. nigra* genome seems to have a high abundance of the complementary SSR dinucleotide repeat motifs $(TG)_n$ and $(TC)_n$.

As far as we know, SCoT markers were never used in gymnosperms. The only tree species molecularly characterized by SCoTs was the angiosperm *Quercus brantii* (Alikhani *et al.* 2014). In this sense, we were unable to compare our SCoT results with previous studies developed in similar species.

Primer SCoT3 presented the highest number of bands while SCoT17 amplified the lowest one (Table III.2). The highest percentage of polymorphism was achieved with primer SCoT25 (100%) since it was the unique primer that did not amplified monomorphic bands (Table III.2). These primer sequences are similar except for their last two nucleotides at the 3'-end (see Table III.2). According to Collard and Mackill (2009), the primers that they designed for SCoT amplification have conserved nucleotide positions in the final three or four nucleotides positioned in the 3' end and this feature influences the primer-template specificity.

Both the ISSR and SCoT markers enabled the production of discriminative patterns among the DNA bulked samples (see codes in Table III.1) belonging to 12 distinct *P. nigra* provenances (Fig. III.1).

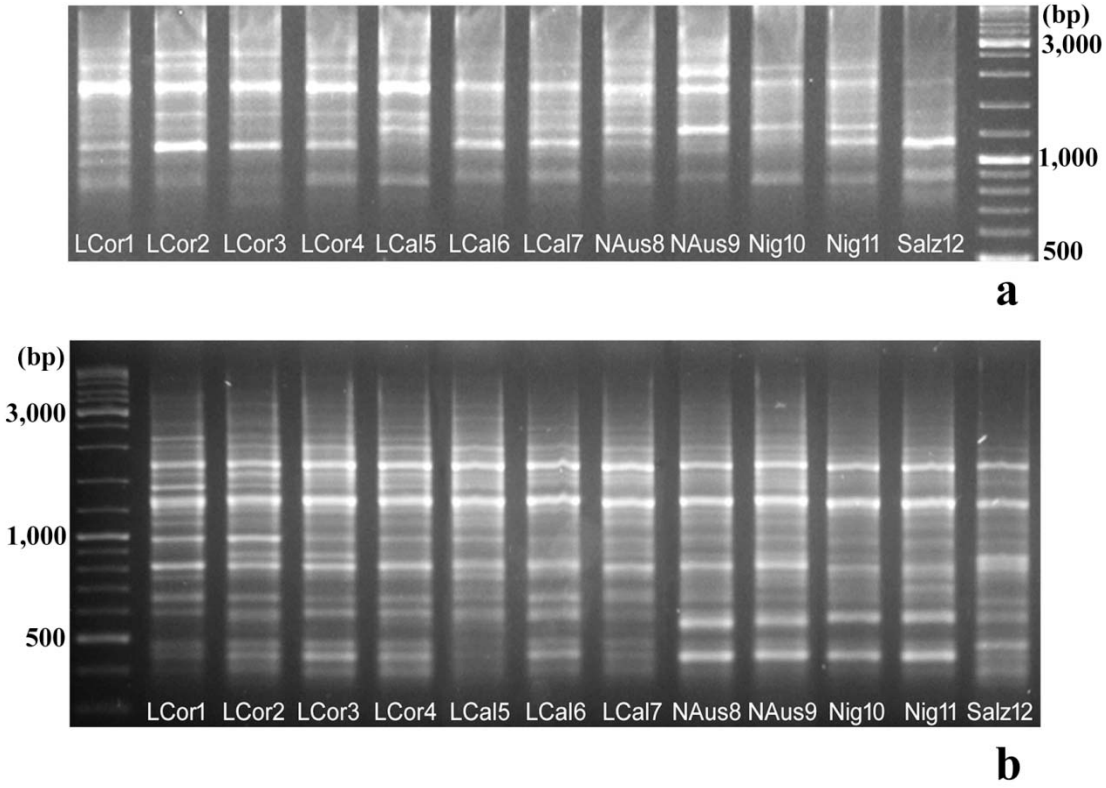

Figure III. 1 – (a) ISSR amplified products achieved with primer 835; and (b) SCoT markers produced by primer SCoT3, showing polymorphic patterns among the 12 *P. nigra* bulked DNA samples from different provenances

The molecular pattern obtained in the bulked samples belonging to *P. nigra* subsp. *salzmannii* with the ISSR markers is highly different from the remaining ones (Fig. III.1a). In addition, if we regard the bulked DNA samples of the same subspecies or variety we found similarities among them, as well as distinct patterns among subspecies or varieties (Fig. III.1).

DNA fingerprinting has been used for cultivars identification, becoming an important tool for genetic identification useful for plant breeding and germplasm conservation and management (Jondle 1992; Smith 1998). Several markers have been used for DNA fingerprinting and the choice of best system to be used with that purpose constitutes an important decision (Gorji *et al.* 2011). Both ISSR and SCoT markers have proved to be suitable for genetic diversity assessment and DNA fingerprinting in different plant species (Fang and Roose 1997; Wünsch and Hormaza 2002; Carvalho *et al*. 2005; 2009a,b; Balasaravanan *et al.* 2006; Gao *et al.* 2006; López-Aljorna *et al.* 2007; Collard and Mackill 2009; Gorji *et al.* 2011; Cabo *et al*. 2014; Coutinho *et al*. 2014a). The present study

demonstrated that both ISSRs and SCoTs could be combined and successfully used for genetic diversity assessment.

Since a high number of monomorphic bands were detected among the 12 bulked DNA samples and could constitute specific markers for lower grade *P. nigra taxa*, we present the number and molecular weight of those bands amplified per ISSR or SCoT primer in Table III.4.

Table III.4 Monomorphic ISSR and SCoT bands amplified in the 12 DNA bulked samples of *P. nigra*

Marker	Primer	Molecular weight of monomorphic bands (bp)
ISSR	817	1,800; 1,700
	827	2,400
	834	1,200; 750; 700
	835	1,100
	836	1,800; 1,300; 1,000; 900
	841	1,700; 900, 800
	850	3,500; 2,900; 2,500; 1,900; 1,800; 1,400; 1,100
	866	1400; 1300; 1100; 1000; 850; 700; 600; 500
SCoT	SCoT3	2,000; 1,900; 1,700; 1,300; 1,200; 1,100; 1,000; 800; 650; 600
	SCoT7	1,900; 1,600; 1,400; 1,100; 600; 500
	SCoT17	1,600; 1,300; 1,200; 1,000
	SCoT18	1,800; 1,600; 1,500; 1,000
	SCoT26	1,900; 1,300; 1,200; 800; 700
	SCoT27	2,500; 2,300; 2,000; 1,800; 1,500; 1,300; 900; 600; 500; 400
	SCoT31	2,500; 2,000; 1,800; 1,600; 1,500; 1,400; 1,200; 1,100; 750; 650

A total of 30 and 49 monomorphic bands were amplified with the ISSR and SCoT markers, respectively (Table III.4). Since the SCoTs could amplify coding regions due to the use of primers whose sequences are complementary to a short conserved region that flanks the initiation codon (ATG) of the plant genes (Collard and Mackill 2009), the common bands among the different provenances could correspond to *P. nigra* genes that have been conserved at the infraspecific level, explaining their presence among the studied plant material.

III.4.2 – Analyses of genetic relationships, variation and gene diversity

The pool of the ISSR and SCoT data was used to construct an UPGMA dendrogram of genetic similarity (Fig. III.2).

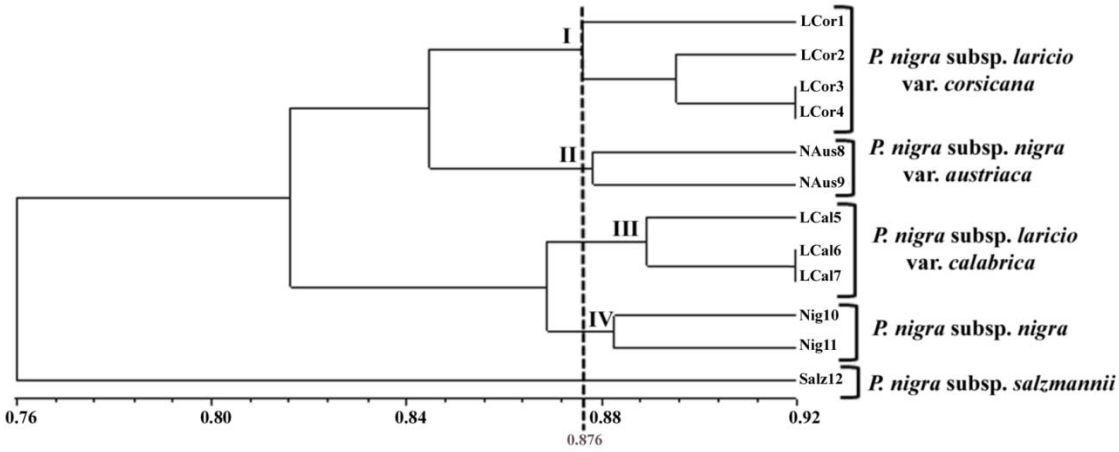

Figure III.2 – UPGMA dendrogram of genetic similarity among the 12 *P. nigra* bulked DNA samples from different provenances, calculated with the DICE coefficient and based on the pool of ISSR and SCoT data. For a cut off value of 0.876, four groups and one branch could be considered: I – enclosing bulks representative of subspecies *laricio* var. *corsicana* (LCor); II – composed by bulks of subspecies *nigra* var. *austriaca* (NAus); III – constituted by bulks of subspecies *laricio* var. *calabrica* (LCal); IV – represented by bulks of subspecies *nigra* but of unknown variety (Nig); and the branch was constituted by bulked samples of subspecies *salzmannii* (Salz12).

The dendrogram revealed 76% of genetic similarity among the 12 *P. nigra* bulked DNA samples corresponding to distinct provenances, and representative of three subspecies and three varieties (Fig. III.2). For a cut off value of 0.876, four groups and one branch could be considered (Fig. III.2). The branch was constituted by the DNA bulked sample of *P. nigra* subsp. *salzmannii* (Salz12) individuals (Fig. III.2), indicating that this subspecies is the most genetically distant and distinct from the others, as revealed by their molecular patterns (see Fig. III.1).

Group I enclosed all the samples representative of *P. nigra* subsp. *laricio* var. *corsicana* whereas Group III was solely composed by DNA bulks representative of *P. nigra* subsp. *laricio* var. *calabrica* (Fig. III.2). The clustering of the DNA bulks from *P. nigra* subsp. *laricio* in two separated clusters indicated the ability of both ISSR and SCoT markers to distinguish samples belonging to different regional varieties of the same subspecies (Fig. III.2). Besides, the bulked samples with codes LCal6 and LCal7 corresponded to individuals that were sampled in geographically closest regions, evidencing another advantage of these markers (Fig. III.2).

Group II was constituted solely by bulks representative of *P. nigra* subsp. *nigra* var. *austriaca* (NAus8 and NAus9). The bulks from *P. nigra* subsp. *nigra* (Nig10 and Nig11) of unknown regional variety were clustered apart in group IV (Fig. III.2). Since the varieties of

subspecies *laricio* were discriminated by the molecular markers used, we suggest that the DNA bulked samples with codes Nig10 and Nig11 might correspond to other regional variety than *austriaca*. By other hand, these samples were sampled in different geographical regions, corroborating the occurrence of high genetic variation among allopatric populations and varieties, and that has been raised identification and taxonomic issues, as previously reported by other authors (Scaltsoyiannes *et al.* 1994; Thompson 2005; Afzal-Raffi and Dodd 2007; Bogunić *et al.* 2007; del Cerro Barja *et al*. 2009; Akkemik *et al.* 2010; Rubio-Moraga *et al*. 2012). Due so, in order to minimize the variation introduced per individual in each studied population or provenance, and to achieve more accurate phylogenetic inferences among the studied infraspecific *taxa*, we adopted the use of DNA bulks in the present study. These features were accomplished with the present analyses since the values of gene diversity and genetic variation provided by the software POPGENE revealed a high value of total genetic diversity (H_T = 0.2382±0.0401); detection of a global percentage of 66.24% of polymorphic *loci* among the studied 12 DNA bulked samples; and higher inter-subspecies genetic diversity (D_{ST} = 0.1655), being $D_{ST} = H_T - H_S$, than intra-subspecies genetic diversity (H_S = 0.0727±0.0073). The ISSR and SCoT markers detected a high relative magnitude of differentiation (G_{ST} = 0.6949) among the *P. nigra* subspecies.

With the minimization of the intra-population variation, we were able to extrapolate about the genetic relationships among the studied infraspecific *taxa* of *P. nigra* based on statistical data provided by different software and coefficients (Figs. III.3 and III.4; Table III.5). The Principal Coordinates Analysis based on the Nei's genetic distance matrix constructed with the pool of the ISSR and ScoT data, projected the 12 DNA bulked samples into five groups (Fig. III.3), corroborating the UPGMA clustering, since four groups and one branch constituted by the subspecies *salzmannii* were considered in the dendrogram (Fig. III.2).

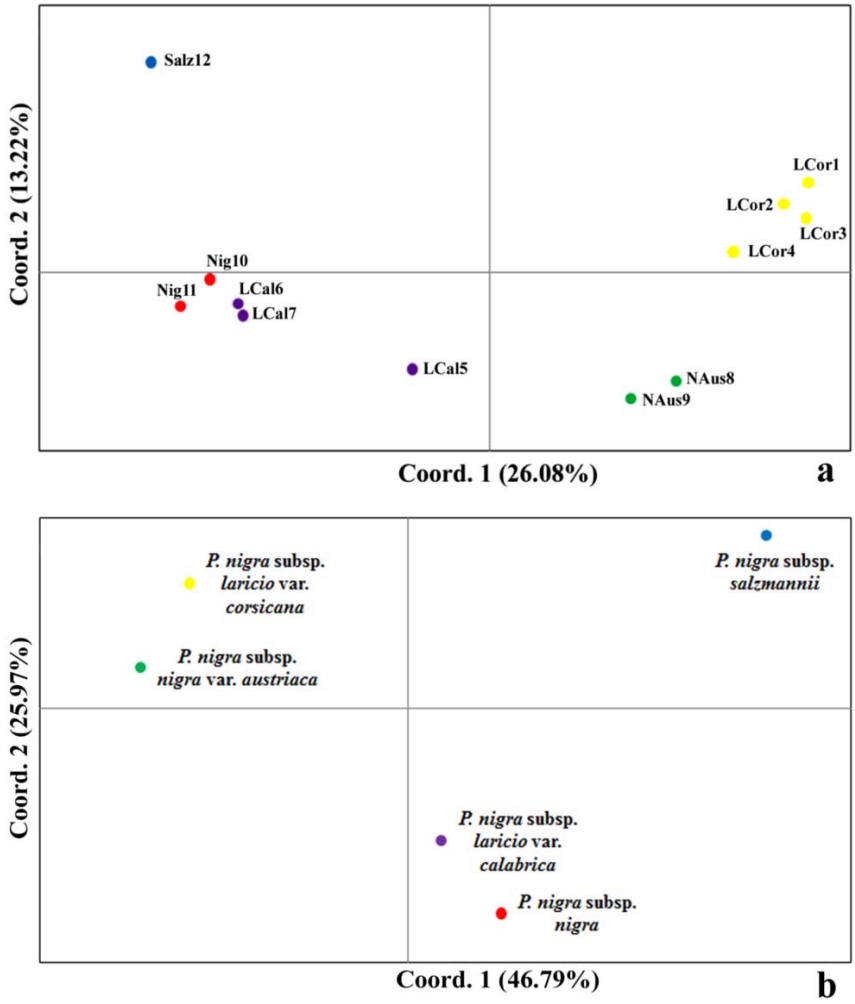

Figure III.3 – Principal Coordinates Analyses (PCoA) based on the Nei's genetic distance matrix constructed with the pooled ISSR and SCoT data: (a) projection of the 12 bulked DNA samples; and (b) their corresponding infraspecific taxonomic classification

We also detected a clear discrimination among the samples belonging to the two regional varieties of subspecies *laricio*, and among samples from the subspecies *nigra* (Fig. 3b), as verified in the dendrogram (Fig. III.2). Except for the clustering and closest projection in the PCoA of bulked DNA samples belonging to subspecies *laricio* var. *corsicana* and subspecies *nigra* var. *austriaca* (Figs. III.2 and III.3), the molecular data provided by the ISSR and SCoT markers, indicated their ability to discriminate *P. nigra* samples per subspecies and regional variety. Similar results were achieved with the Nei's method (Nei 1978) which determined high genetic identity between the samples belonging to these *taxa* (Table III.5), as graphically represented in Figure III.4. According to the Nei's method calculation, the subspecies *nigra* and subspecies *laricio* var. *calabrica* revealed the highest

genetic identity value (0.9051), followed by the subspecies *nigra* and *laricio* var. *corsicana* (0.8690) (Table III.5). The subspecies *laricio* var. *corsicana* and *salzmannii* showed the highest genetic distance value (0.3544) (Table III.5).

Table III.5 – Nei's unbiased measures of genetic distance (below diagonal) and identity (above diagonal) among the *P. nigra* infraspecific *taxa* calculated by the Nei's method (Nei 1978).

P. nigra* infraspecific *taxa	**subsp. *laricio* var. *corsicana***	**subsp. *laricio* var. *calabrica***	**subsp. *nigra* var. *austriaca***	**subsp. *nigra***	**subsp. *salzmannii***
subsp. *laricio* var. *corsicana*	-	0.8250	0.8690	0.7833	0.7016
subsp. *laricio* var. *calabrica*	0.1924	-	0.8258	0.9051	0.7818
subsp. *nigra* var. *austriaca*	0.1404	0.1914	-	0.7912	0.6745
subsp. *nigra*	0.2443	0.0998	0.2342	-	0.7772
subsp. *salzmannii*	0.3544	0.2461	0.3938	0.2520	-

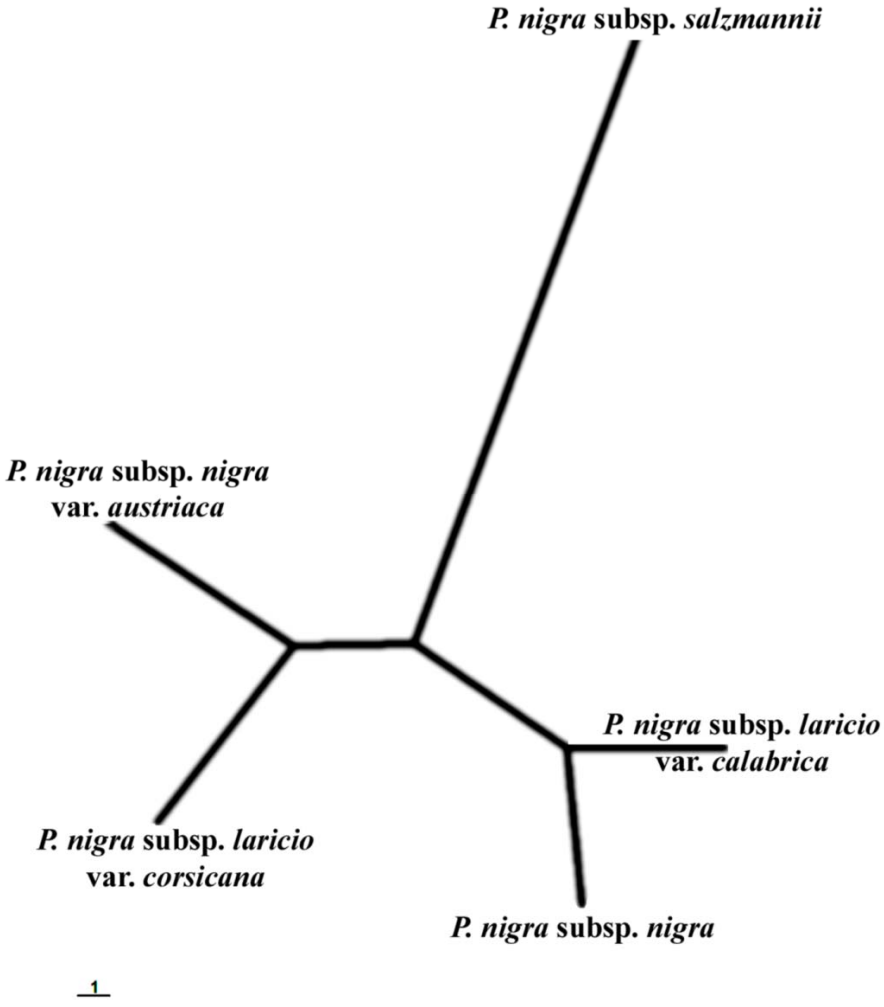

Figure III.4 – Unrooted tree based on the Nei's unbiased measures of genetic identity and distance (Nei 1978) constructed with the TreeView software

Over the years several attempts to revise the *P. nigra* taxonomy were made. Consequently, numerous classifications at the infraspecific level of *P. nigra* were proposed (Scaltsoyiannes *et al.* 1994; Raffi *et al.* 1996; Naydenov *et al.* 2006; Afzal-Raffi and Dodd 2007; Bogunić *et al.* 2007; del Cerro Barja *et al.* 2009). Louro (1982) compared taxonomical classifications made by several authors and indicated some coincidences among them.

According to Roiron *et al.* (2013), the *P. nigra* subspecies *nigra* and *salzmannii* can be easily differentiated by observation of their cone morphology. Contrastingly, the same was not true for the subspecies *salzmannii* and *laricio* (Roiron *et al.* 2013). The molecular results provided by this study evidenced the clear genetic distance among the bulked samples from subspecies *salzmannii* and the remaining ones, being useful for the discrimination of the infraspecific *taxa* of *P. nigra*. In addition, the different genetic analyses performed in this

study revealed genetic variation among bulked DNA samples belonging to the regional varieties *calabrica* and *corsicana* of subspecies *laricio.* In fact, several authors have been studied these two varieties and encountered differences that contested their taxonomic classification (Arbez and Miller 1971; Arbez *et al.* 1974; Bonnet-Masimbert and Bikay-Bikay 1978; Louro 1982; Gerber *et al.* 1995; Aguinagalde *et al.* 1997), allowing some authors to propose the elevation of subspecies *laricio* to an isolated species (Cesca and Peruzzi 2002).

III.5 – Conclusions

The present study comprised the molecular characterization of 12 DNA bulks representative of distinct *P. nigra* provenances representative of different infraspecific *taxa*. Globally, the two dominant marker systems, ISSRs and SCoTs, proved to be suitable for the discrimination of the bulked DNA samples per allopatric subspecies and regional varieties of *P. nigra*.

The molecular data provided by the ISSR and SCoT analyses corroborated the actual adopted taxonomic classification, clustering the bulked samples per allopatric subspecies and regional variety. Different regional varieties were suggested for the four DNA bulked samples belonging to subspecies *nigra* due to their separated clustering. The ISSR and SCoT markers confirmed the occurrence of genetic variation between the regional varieties *calabrica* and *corsicana* of subspecies *laricio*, and evidenced that subspecies *salzmannii* is highly genetically distant from the others analysed subspecies. Despite the preliminarily of this study, we verified that the cost-effective ISSR and SCoT dominant markers might provide an excellent tool for the discrimination of infraspecific *taxa* in *P. nigra*, contributing for the overcoming of issues related to morphological identification and taxonomic classification among the highly variable allopatric subspecies and varieties. Additional studies will be performed in separated individuals, in order to develop diagnostic markers for each infraspecific *taxon* that could be further used in DNA fingerprinting.

III.6 – Acknowledgements

This work was supported by the project UID/AGR/04033/2013, attributed by the Portuguese Foundation for Science and Technology ('Fundação para a Ciência e a Tecnologia', FCT)

Ana Carvalho and Alexandra Dias thank the postdoctoral SFRH/BPD/68932/2010 and doctoral SFRH/BD/91781/2012 Grants, respectively, both attributed by FCT and co-financed by Fundo Social Europeu/ Programa Operacional Potencial Humano - Quadro de Referência Estratégica Nacional (FSE/POPH-QREN).

The authors acknowledge to Prof. Bruno Fady (INRA, Unité des Recherches Forestières Méditerranéennes, Avignon, France) and to Dr. Marc Calvignac from Vilmorin S.A. (La Ménitré, France) for kindly provided dehydrated needles and certified seed material, respectively, of different subspecies and varieties of *P. nigra*.

III.7 – References

Afzal-Rafii Z, Dodd RS (2007) Chloroplast DNA supports a hypothesis of glacial refugia over postglacial recolonization in disjunct populations of black pine (*Pinus nigra*) in Western Europe. Molecular Ecology 16:723-736.

Agarwal T, Gupta AK, Patel AK, Shekhawat NS (2015) Micropropagation and validation of genetic homogeneity of *Alhagi maurorum* using SCoT, ISSR and RAPD markers. Plant Cell, Tissue and Organ Culture 120:313-323.

Aguinagalde I, Lorente F, Benito C (1997) Relationships among five populations of European black bine (*Pinus nigra* ARN.) using morphometric and isozyme markers. Silvae Genetica 46:1-5.

Akkemik Ü, Yilmaz H, Oral D, Kaya A (2010) Some changes in taxonomy of pines (*Pinus* L.) native to Turkey. Journal of the Faculty of Forestry 61:63-78.

Alikhani L, Rahmani M-S, Shabanian N, Badakhshan H, Khadivi-Khub A (2014) Genetic variability and structure of *Quercus brantii* assessed by ISSR, IRAP and SCoT markers. Gene 552:176-183.

Amirmoradi B, Talebi R, Karami E (2012) Comparison of genetic variation and differentiation among annual Cicer species using Start Codon Targeted (SCoT) polymorphism, DAMD-PCR, and ISSR markers. Plant Systematic and Evolution 298:1679-1688.

Ammiraju JSS, Dholakia BB, Santra DK, Singh H, Lagu MD, Tamhankar, Dhaliwal HS, Rao VS, Gupta VS, Ranjekar PK (2001) Identification of inter simple sequence repeat (ISSR) markers associated with seed size in wheat. Theoretical and Applied Genetics 102:726-732.

Arbez M, Bernard-Dagan C, Fillon C (1974) Variabilité intraspécifique des monoterpènes de *Pinus nigra* Arn. Annales des Sciences Forestiéres 31:57–70.

Arbez M, Miller C (1971) Contribution a l'étude de la variabilité géographique de *Pinus nigra* Arn.. Annales des Sciences Forestiéres 28: 23-49.

Balasaravanan T, Chezhian P, Kamalakannan R, Yasodha R, Varghese M, Gurumurthi K, *et al.* (2006) Identification of species-diagnostic ISSR markers for six *Eucalyptus* species. Silvae Genetica 55(3):119–122.

Barbéro M, Loisel R, Quézel P, Richardson DM, Romane F (1998) Pines of Mediterranean basin, pp. 153–170. In: Ecology and Biogeography of *Pinus*. (edited by DM Richardson), Institute for Plant Conservation, University of Cape Town, South Africa, Cambridge University Press, 2000, UK, ISBN: 0521789109527 p.

Bhattacharyya P, Kumaria S, Kumar S, Tandom P (2013) Start Codon Targeted (SCoT) marker reveals genetic diversity of *Dendrobium nobile* Lindl., an endangered medicinal orchid species. Gene 529:21-26.

Blanco Castro E, Casado Gonzalez MA, Costa Tenorio M *et al*. (1997) Los bosques ibéricos: Una interpretación geobotánica. Barcelona: Planeta. 572 p. ISBN 84-08-01924-4.

Bogunić F, Muratovi E, Ballian D, Siljak-Yakolev S, Brown S (2007) Genome size stability of five subspecies of *Pinus nigra* Arnold s.l.. Environmental and Experimental Botany 59:354-360.

Bogunić F, Muratovic E, Brown SC, Siljak-Yakovlev S (2003) Genome size of five *Pinus* from Balkan region. Plant Cell Reports 22:59–63.

Bogunić F, Siljak-Yakolev S, Muratovi E, Ballian D (2011) Diferent karyotype patterns among allopatric *Pinus nigra* (Pinaceae) populations revealed by molecular cytogenetics. Plant Biology 13:194-200.

Bonnet-Masimbert M, Bikay-Bikay V (1978) Variabilité intraspécifique des isoenzymes de la glutamate oxalacetate transaminase chez *Pinus nigra*. Silvae Geneta 27:71–79.

Boydak M (2001) A new variety of *Pinus nigra* J.F. Arnold subsp. *pallasiana* (Lamb.) Holmboe from Anatolia. Karaca Arbor Magazine 6:15–23.

Bussotti F (2002) *Pinus nigra* Arnold, pp. 266-282. In: Pines of silvicultural importance / compiled from the Forestry compendium, 1st Edition, Oxon, UK; New York: CAB International Publishers, c2002. ISBN-10: 085199539X, 531 p.

Cabo S, Carvalho A, Rocha L, Martín A, Lima-Brito J (2014a) IRAP, REMAP and ISSR fingerprinting in newly formed hexaploid tritordeum (X Tritordeum Ascherson et Graebner) and respective parental species. Plant Molecular Biology Reporter 32:761-770.

Cabo S, Ferreira L, Carvalho A, Martins-Lopes P, Martín A, Lima-Brito J (2014b) Potential of Start Codon Targeted (SCoT) markers for DNA fingerprinting of newly synthesized tritordeums and respective parents. Journal of Applied Genetics 55:307-312.

Carvalho A, Guedes-Pinto H, Lima-Brito J (2012) Genetic diversity in Old Portuguese durum wheat cultivars assessed by retrotransposon-based markers. Plant Molecular Biology Reporter 30:578-589.

Carvalho A, Lima-Brito J, Maçãs B, Guedes-Pinto H (2009a) Genetic diversity and variation among botanical varieties of Old Portuguese wheat cultivars revealed by ISSR assays. Biochemical Genetics 47:276-294.

Carvalho A, Matos M, Lima-Brito J, Guedes-Pinto H, Benito C (2005) DNA fingerprint of F1 interspecific hybrids from the Triticeae tribe using ISSRs. Euphytica 143:93-99.

Carvalho A, Paula A Guedes-Pinto H, Martins L, Carvalho J, Lima-Brito J (2009b) Preliminary genetic approach based on both cytogenetic and molecular characterizations of nine oak species. Plant Biosystems 1:25-33.

Çengel B, Tayanç Y, Kandemir G, Velioglu E, Alan M, Kaya Z (2012) Magnitude and efficiency of genetic diversity captured from seed stands of *Pinus nigra* (Arnold) subsp. *pallasiana* in established seed orchards and plantations. New Forests 43:303-317.

Cesca G, Peruzzi L (2002) *Pinus laricio* Poir. and *P. leucodermis* Antoine: Karyotype analysis in Calabrian populations (Southern Italy). Caryologia 55:21-25.

Chen F-Y, Liu J-H (2014) Germplasm genetic diversity of *Myrica rubra* in Zhejiang Province studied using inter-primer binding site and start codon-targeted polymorphism markers. Scientia Horticulturae 170:169-175.

Chen H, He X, Luo C, Zhu J, Li F (2010) Analysis on the genetic diversity of 24 Longan (*Dimocarpus longan*) accessions by SCoT Markers. Acta Horticulturae Sinica 37:1651-1654.

Chezhian P, Yasodha R, Ghosh M (2010) Genetic diversity analysis in a seed orchard of Eucalyptus tereticornis. New Forests 40:85-99.

Christensen KI (1997) Pinaceae, Cupressaceae, Taxaceae, Ephedraceae, Salicaceae, Juglandaceae, Betulaceae, Fagaceae, Ulmaceae Moraceae, pp 1–17. In: Flora Hellenica. Vol.1 (edited by A Strid, K Tan). Koeltz Scientific Books, Königstein, Germany.

Cipriano J, Carvalho A, Fernandes C, Gaspar MJ, Pires J, Bento J, Roxo L, Louzada J, Lima-Brito J (2013) Evaluation of genetic diversity of Portuguese *Pinus sylvestris* L. populations based on molecular data and inferences about the future use of this germplasm. Journal of Genetics 92: e41-e48.

Collard BCY, Mackill DJ (2009) Start Codon Targeted (SCoT) polymorphism: A simple novel DNA marker technique for generating gene-targeted markers in plants. Plant Molecular Biology Reporter 27:86-93.

Coutinho J, Carvalho A, Lima-Brito J (2014a) Fingerprinting of Fagaceae individuals using inter-microsatellite markers. Journal of Genetics 93:e132-e140. Online only: http://www.ias.ac.in/jgenet/OnlineResources/93/e132.pdf.

Coutinho J, Carvalho A, Lima-Brito J (2014b) Genetic diversity assessment and estimation of phylogenetic relationships among 26 Fagaceae species using ISSRs. Biochemical Systematics and Ecology 54:247–256.

Culley TM, Wolfe AD (2001) Population genetic structure of the cleistogamous plant species *Viola pubescens* Aiton (Violaceae), as indicated by allozyme and ISSR molecular markers. Heredity 86:545-556.

del Cerro Barja A, Lucas-Borja ME, Garcia EM, López Serrano FR, Abellán MA, Morote FAG, López RN (2009) Influence of stand density and soil treatment on the Spanish Black Pine (*Pinus nigra* Arn. ssp. *Salzmannii*) regeneration in Spain. Investigación Agraria: Sistemas y Recursos Forestales 18:167-180.

Fang DQ, Roose ML (1997) Identification of closely related citrus cultivars with inter-simple sequence repeat markers. Theoretical and Applied Genetics 95:408-417.

Feng FJ, Han SJ, Wang HM (2006) Genetic diversity and genetic differentiation of natural *Pinus koraiensis* populations. Journal of Forest Research 17:21-24.

Furman BJ, Grattapaglia D, Dvorak WS, O'Malley DM (1997) Analysis of genetic relationships of Central American and Mexical pines using RAPD markers that distinguish species. Molecular Ecology 6:321-331.

Gao J, Zhang S, Qi L, Zhang Y, *et al.* (2006) Application of ISSR markers to fingerprinting of elite cultivars (varieties/clones) from different sections of the genus *Populus* L. Silvae Genetica 55(1):1-6.

Gaussen H, Haywood VH, Chater AO (1964) *Pinus*. In: Flora Europaea. Vol. 1 (edited by TG Tutin, VH Heywood, NA Burges, DM Moore, DH Valentine, SM Walters, DA Webb), pp- 32-35, Cambridge University Press, UK.

Gaussen H, Webb DA, Heywood VH (1993). Genus *Pinus*. In: Flora Europaea. Vol 1 (edited by TG Tutin, VH Heywood, NA Burges, DM Moore, DH Valentine, SM Walters, DA Webb), pp 40–44, Cambridge University Press, UK.

Gerber S, Baradat P, Marpeau A, Arbez M (1995) Geographic variation in terpene composition of *Pinus nigra* Arn.. Forest Genetics 2:1-10.

Godwin ID, Aitken EAB, Smith LW (1997) Application of inter-simple sequence repeat (ISSR) markers to plant genetics. Electrophoresis 18:1524–1528.

Gorji A, Poczai P, Polgar Z, Taller J (2011) Efficiency of arbitrarily amplified dominant markers (SCoT, ISSR and RAPD) for diagnostic fingerprinting in tetraploid potato. American Journal of Potato Research 88:226-237.

Guo D-L, Zhang J-Y, Liu C-H. 2012. Genetic diversity in some grape varieties revealed by SCoT analyses. Molecular Biology Reports 39:5307-5313.

Gupta M, Chyi YS, Romero-Severson J, Owen JL (1994) Amplification of DNA markers from evolutionarily diverse genome using single primers of simple-sequence repeats. Theoretical and Applied Genetics 89:998-1006.

Jiang LF, Qi X, Zhang XQ, Huang LK, Ma X, Xie WG (2014) Analysis of diversity and relationships among orchardgrass (*Dactylis glomerata* L.) accessions using start codon-targeted markers. Genetics and Molecular Research 13:4406-4418.

Labra M, Grassib F, Sgorbatia S, Ferraric C (2006) Distribution of genetic variability in Southern populations of Scots pine (*Pinus sylvestris* L.) from the Alps to the Apennines. Flora 201:468-476.

Li HY, Jiang J, Liu GF, Ma XJ, Dong JX, Lin SJ (2005) Genetic variation and division of *Pinus sylvestris* provenances by ISSR markers. Journal of Forest Research 16:216-218.

Liber Z, Nikolić T, Mitić B, Šatović Z (2003) RAPD markers and Black pine (*Pinus nigra* Arnold) intraspecies taxonomy - Evidence from the study of nine populations. Acta Societatis Botanicorum Poloniae 72:249-257.

López-Aljorna A, Bueno MÁ, Aguinagalde I, Martín JP (2007) Fingerprinting and genetic variability in oak (*Quercus suber* L.) elite trees using ISSR and SSR markers. Annals of Forest Science 64:773-779.

Louro V (1982) O pinheiro larício (*Pinus nigra* Arnold) em Portugal. Divulgação Técnica, MACP, Direcção Geral das Florestas Nº 2 – Agosto 1982. 110 pp.

Lučić A, Mladenović-Drinić S, Stavretonić N, Isajev V, Lavadinivić V, Rakonkac L, Novaković (2010) Genetic diversity of Austrian Pine (*Pinus nigra* Arnold) populations in Serbia revealed by RAPD. Archives of Biological Science Belgrade 62:329-336.

Luo C, He X-H, Chen H, Ou S-J, Gao M-P (2010) Analysis of diversity and relationships among mango cultivars using Start Codon Targeted (SCoT) markers. Biochemical Systematics and Ecology 38:1176–1184.

Luo C, He X-H, Chen H, Ou S-J, Gao M-P, Brown JS, Tondo CT, Schnell RJ (2011) Genetic diversity of mango cultivars estimated using SCoT and ISSR markers. Biochemical Systematics and Ecology 39:676-684.

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

Mulpuri S, Muddanuru T, Francis G (2013) Start codon targeted (SCoT) polymorphism in toxic and non-toxic accessions of *Jatropha curcas* L. and development of a codominant SCAR marker. Plant Science 207:117-127.

Naik D, Singh D, Vartak V, Paranjpe S, Bhargava S (2009) Assessment of morphological and genetic diversity in *Gmelina arborea* Roxb. New Forests 38:99-115.

Naydenov K, Tremblay FM, Fenton NJ, Alexandrov A (2006) Structure of *Pinus nigra* Arn. populations in Bulgaria revealed by chloroplast microsatellites and terpenes analysis: provenance tests. Biochemical Systematics and Ecology 34:562-574.

Nei M (1973) Analysis of gene diversity in subdivided populations. Proceedings of the National Academy of Sciences 70:3321-3323.

Nei M (1978) Estimation of average heterozygosity and genetic distance from a small number of individuals. Genetics 89:583-590.

Nei M (1987) Molecular evolutionary genetics, pp. 176–187, Columbia University Press, New York, USA.

Page RD (1996) TreeView: An application to display phylogenetic trees on personal computers. Computer Applications in the Biosciences 12: 357–358.

Pajares PR, Escudero A (1989) Caracterización fitoecológica de las comunidades de *Pinus nigra* subsp. *salzmannii* en los afloramientos rochosos del sistema ibérico meridional. Botanica Complutensis 15:149-161.

Pakseresht F, Talebi R, Karami E (2013) Comparative assessment of ISSR, DAMD and SCoT markers for evaluation of genetic diversity and conservation of landrace chickpea (*Cicer arietinum* L.) genotypes collected from North-west of Iran. Physiology and Molecular Biology of Plants 19:563-574.

Peakall R, Smouse P (2012) GenAlEx 6.5: Genetic analysis in Excel. Population genetic software for teaching and research - an update. Bioinformatics 28:2537-2539.

Poczai P, Varga I, Laos M, Cseh A, Bell N, Valkonen JPT, Hyvönen J (2013) Advances in plant gene-targeted and functional markers: a review. Plant Methods 9:6.

Price RA, Liston A, Strauss SH (1998). Phylogeny and systematics of *Pinus*. In: Ecology and biogeography of *Pinus* (edited by DM Richardson), pp. 49–68, Cambridge University Press. London.

Que Y, Pan Y, Lu Y, Yang C, Yang Y, Huang N, Xu L (2014) Genetic analysis of diversity within a Chinese local sugarcane germplasm based on Start Codon Targeted Polymorphism. BioMed Research International Volume 2014, Article ID 468375, 10 p. http://dx.doi.org/10.1155/2014/468375.

Raffi ZA, Dodd RS, Zavarin E (1996) Genetic diversity in foliar terpenoids among natural populations of European Black Pine. Biochemical Systematics and Ecology 24:325-339.

Richardson DM (1998) Ecology and biogeography of *Pinus*. Cambridge University Press, London, 527p.

Rohlf FJ (1998) NTSYS-pc ver. 2.02. Numerical taxonomy and multivariate analysis system. Exeter Publishing, Setauket, USA.

Roiron P, Chabal L, Figueiral I, Terral J, Ali A (2013) Palaeobiogeography of *Pinus nigra* Arn. subsp. *salzmannii* (Dunal) Franco in the Northwestern Mediterranean Basin: A review based on macroremains. Review Palaeobototany and Palynology 194:1-11.

Rubio-Moraga A, Candel-Perez D, Lucas-Borja ME, Tiscar PA, Viñegla, Linares JC, Gómez-Gómez L, Ahrazem O (2012) Genetic diversity of *Pinus nigra* Arn. populations in Southern Spain and Northern Morocco revealed by inter-simple sequence repeat profiles. International Journal of Molecular Sciences 13:5645-5658.

Scaltsoyiannes A, Rohr R, Panetos KP, Tsaktsira M (1994) Allozyme frequency distributions in five European populations of black pine (*Pinus nigra* Arnold). Silvae Genetica 43:20-30.

Shahlaei A, Torabi S, Khhosroshahli M (2014) Efficacy of SCoT and ISSR markers in assessment of tomato (*Lycopersicum esculentum* Mill.) genetic diversity. International Journal of Biosciences 5:14-22.

Thompson JN (2005) The geographic mosaic of coevolution. University of Chicago Press, Chicago, Illinois, USA, 400p.

Vidaković M (1974) Genetics of European black pine (*Pinus nigra* Arn.). Ann Forest 6:57-86.

Wünsch A, Hormaza JI (2002) Cultivar identification and genetic fingerprinting of temperate fruit tree species using DNA markers. Euphytica 125:59–67.

Xie W, Zhang J, Zhao X, Zhang J, Wang Y (2015) Siberian wild rye (*Elymus sibiricus* L.): Genetic diversity of germplasm determined using DNA fingerprinting and SCoT markers. Biochemical Systematics and Ecology 60:186-192.

Xiong F, Zhong R, Han Z, Jiang J, He L, Zhuang W, Tang R (2011) Start codon targeted polymorphism for evaluation of functional genetic variation and relationships in cultivated peanut (*Arachis hypogaea* L.) genotypes. Plant Molecular Biology Reporter 38:3487-3494.

Yeh FC, Yang RC, Boyle TBJ, Ye ZH, Mao JX (1999) Popgene version 1.32, the user-friendly shareware for population genetic analysis. Molecular Biology and Biotechnology Centre, University of Alberta, Edmonton, USA.

Zaghi D (2008) Management of Natura 2000 habitats. 9530 *(Sub)-Mediterranean pine forests with endemic black pines. European Commission.

Zeng B, Zhang Y, Huang L, Jiang X, Luo D, Yin Guohua (2014) Genetic diversity of orchardgrass (*Dactylis glomerata* L.) germplasms with resistance to rust diseases revealed by Start Codon Targeted (SCoT) markers. Biochemical Systematics and Ecology 54:96-102.

Zhang J, Xie W, Wang Y, Zhao (2015) Genetic diversity and relationships among Chinese *Elymus sibiricus* accessions using Start Codon Targeted (SCoT) markers. Molecules 20:1-13.

Zhang Z-Y, Chen Y-Y, Li D-Z (2005) Detection of low genetic variation in a critically endangered Chinese pine, *Pinus squamata*, using RAPD and ISSR markers. Biochemical Genetics 43:239-249.

Zietkiewicz E, Rafalski A, Labuda D (1994) Genome fingerprinting by simple sequence repeat (SSR)-anchored polymerase chain reaction amplification. Genomics 20:176–183.

Chapter IV:

First molecular characterization of the Portuguese populations of *Pinus nigra* Arnold and extrapolation of their provenances

---

Lemos M, Dias A, Carvalho A, Gaspar MJ, Coutinho J, Louzada JL, Lima-Brito J (2015) First molecular characterization of the Portuguese populations of *Pinus nigra* Arnold and extrapolation of their provenances. Paper Submitted to *Plant Molecular Biology Reporter*

IV.1 – Abstract

European Black Pine (*Pinus nigra* Arnold) belongs to the Family Pinaceae and to the Mediterranean pines group. This study constitutes the first molecular characterization of six allochthonous Portuguese *P. nigra* populations ('Manteigas', 'Vale do Zêzere', 'Campeã', 'Vila Pouca de Aguiar', 'Paredes de Coura' and 'Caminha') using both inter-simple sequence repeat (ISSR) and Start Codon Targeted (SCoT) markers. These populations were installed 50 to 90 years ago in those areas but their provenances are unknown. Morphological and anatomical evaluation of the Portuguese stands of *P. nigra* performed during the 80s have characterized them as belonging to three subspecies, namely, *laricio* (var. *corsicana* and var. *calabrica*); *salzmannii*; and *nigra*. To extrapolate the provenances of the actual *P. nigra* populations, we compared the molecular patterns of the Portuguese samples with those achieved in foreign samples of *P. nigra* belonging to those subspecies and varieties. The pool of the ISSR and SCoT data revealed 65% of genetic similarity among the 127 Portuguese *P. nigra* individuals that were structured into six highly differentiated populations, and the polymorphism was higher within than among populations. A reduced correlation coefficient was detected between the genetic distance matrices achieved with ISSRs and SCoTs. Due so, for provenances extrapolation, we performed a separate molecular analysis for each marker system. ISSRs evidenced higher genetic identity among the Portuguese and foreign samples from *P. nigra* subsp. *laricio* var. *calabrica* whereas SCoTs estimated a higher genetic identity with *P. nigra* subsp. *laricio* var. *corsicana*. Both markers showed high genetic similarity with subspecies *laricio*, corroborating the morphological characterization performed in the *P. nigra* Portuguese stands in the 1980's.

Keywords: European Black Pine; Genetic diversity and structure; Inter-Simple Sequence Repeat (ISSR); Start Codon Targeted (SCoT).

IV.2 – Introduction

Pines are conifers with a vast ecological and economical importance. Six pine species can currently be found at the Iberian Peninsula: *Pinus pinaster* Aiton*, Pinus uncinata* Ramon ex DC, *Pinus sylvestris* L., *Pinus nigra* Arnold, *Pinus pinea* L .and *Pinus halepensis* Mill.

Based on macroremain records (wood, charcoals, fruits, seeds and leaves) and palaeoecological data, *P. nigra* is one of the three oldest pine species widely spread through the Iberian landscapes (García Antón *et al*. 1995; Carrión and Geel 1999; Rubiales *et al.* 2010). The species occupies about 1.4 million hectares in the central and Southern Iberian Mountains and presents specimens with over 650 years old (Génova and Fernández 1998; Martín-Benito *et al*. 2008; Rubiales *et al*. 2010).

The European Black Pine forests have an essential role in ecology, preservation of ecosystems, soil protection, as well as high economic importance for wood production (Espelta *et al.* 2003; Martín-Benito *et al.* 2008, Lucas-Borja *et al.* 2012).

P. nigra has a discontinuous distribution that extends from Southern Europe through Asia Minor and locally at Northwestern Africa (Little and Critchfield 1969; Rubio-Moraga *et al*. 2012). It is also found on the islands of Corsica and Sicily (Rubio-Moraga *et al*. 2012). European Black Pine populations experienced different glacial and interglacial episodes of climatic changes and geological events (Thompson 2005) that contributed for their fragmented distribution (Isajev *et al*. 2004), and to existence of highly variable populations at morphological, physiological, biochemical, genetic and ecological levels (Scaltsoyiannes *et al.* 1994; Raffi *et al.* 1996; Bogunić *et al.* 2003; Afzal-Raffi and Dodd 2007; del Cerro Barja *et al.* 2009; Rubio-Moraga *et al*. 2012). The high variation among allopatric populations hampered their interpretation and resulted on taxonomic classification issues. Despite being considered a collective species (Villar 1947; Svoboda 1953; Gaussen 1960; Debazac 1964; Arbez and Millier 1971; Pajares and Escudero 1989; Blanco 1998; Richardson 1998), other authors have proposed the botanic division of the *P. nigra* species by recognizing the existence of allopatric subspecies and varieties (Delevoy 1949; Gaussen *et al.* 1964; 1993; Mirov 1967; Vidaković 1974; Barbéro *et al.* 1998; Greuter *et al.* 1984; Christensen 1997).

In spite of physical separation, the different geographical subunits of the collective *P. nigra* species did not result in matting barriers under experimental conditions, being described as subspecies (Debazac 1964; Isajev *et al*. 2004; Fady *et al*. 2003).

The most widely recognized infraspecific classification was that proposed by Gaussen *et al.* (1964; 1993) in Flora Europaea which was based on the previous classification of Schwarz (1938). The authors considered the existence of five subspecies with Eurasian distribution: 1) *P. nigra* subsp. *pallasiana* (Lambert) Holmboe (Turkey and Crimea); 2) *P. nigra* subsp. *dalmatica* (Visiani) Franco (Balkans); 3) *P. nigra* subsp. *salzmannii* (Dunal) Franco (syn: *P. nigra clusiana*, *P. nigra pyrenaica*) (France and Spain); 4) *P. nigra* subsp. *laricio* (Poir.) Maire. (Corsica, Sicily and Calabria); and 5) *P. nigra* subsp. *nigra* (Visiani) Franco (syn: *P. nigra austriaca* Höss, *P.nigra nigricans* Host) (Alps). Additionally, a sixth allopatric subspecies should be added: *P. nigra* subsp. *mauretanica* from the Northwestern of Africa (Maire and Peyerimh.) Heywood (Greuter *et al.* 1984; Barbéro *et al.* 1998).

Molecular, cytogenetic and biochemical tools have been used to explore the intraspecific variation of *P. nigra* (Scaltsoyiannes *et al*. 1994; Bogunić *et al*. 2003; 2007; 2011; Afzal-Rafii-Dodd 2007; Rubio-Moraga *et al*. 2012). Moreover, morphological and molecular markers have demonstrated a common phylogenetic origin for all European Black Pines (Isajev *et al.* 2004). If the *taxa* have a common phylogenetic origin, then they are better described as members of a collective species rather than as members of separate species (see Avise 1994 for a review). Among the different subspecies of European Black Pine, the *P. nigra* subsp. *salzmannii* and *P. nigra* subsp. *laricio* are considered the most divergent and genetically distinct (Isajev *et al.* 2004). Surprisingly, the genetic structure of the species has not been taken in consideration for the definition of reforestation strategies (Naydenov *et al.* 2006). Thus, is important to plant European Black Pines of unknown origin apart from autochthonous pinewoods in order to reduce the risk of genetic pollution and to avoid hybridization among different subspecies (Zaghi 2008).

During the XX century, several provenance trials were established independently in Europe, USA and New Zealand. The Corsican and Calabrian Black Pine provenances were found to be the best adapted genotypes. Nonetheless, the extensive European plantations were done with material from unknown and/or very distant sources, and probably resulted on a mix of local and imported gene pools (Isajev *et al*. 2004).

Genetic diversity is fundamental for the maintenance and long term stability of forest ecosystems, since it determines the adaptive potential of the forestry species to the environmental conditions and climatic changes (Muller-stark *et al.* 1992; Çengel *et al.* 2012; Rubio-Moraga *et al*. 2012). The assessment of genetic variability among forestry populations could contribute for the definition of strategies for *in situ* conservation, exploitation and

restoration of genetic resources, and forest management (Lucić *et al.* 2010). In the last decade, a wide array of DNA markers has been used in genetic studies of forest populations to evaluate neutral variation, to analyze their structure, relationships and to estimate the gene flow. Some molecular studies were performed in *P. nigra* populations using Inter Simple Sequence Repeat (ISSR), Random Amplified Polymorphic DNA (RAPD), nuclear and chloroplast simple sequence repeats (nSSRs and cpSSRs) that revealed higher polymorphism at the intra-population level (Liber *et al.* 1999; 2003; Naydenov *et al.* 2006; Afzal-Rafii and Dodd 2007; Lucić *et al.* 2010; Rubio-Moraga *et al*. 2012; Šarac *et al.* 2014; Bonavita *et al.* 2015).

Start Codon Targeted (SCoT) markers developed recently by Collard and Mackill (2009) revealed to be useful for genetic diversity studies (Luo *et al.* 2010; Amimoradi *et al.* 2012; Gorji et al. 2012; Mulpuri *et al.* 2013; Cabo *et al.* 2014b). However, only one study with SCoTs was reported in the forestry species *Quercus brantii* (Alighani *et al*. 2014). The study presented here is the first that use SCoT markers in *Pinus* sp., namely, in *P. nigra.*

Although being one of the most common and oldest pine species in Iberian Peninsula (Rubiales *et al*. 2010), the actual Portuguese *P. nigra* populations exist in mountainous areas and result from afforestation performed about 50 to 90 years ago. However, they are reduced in size and restricted to the North and Centre of the country. Additionally, these allochthonous populations were never molecularly characterized and their provenances are unknown (Louro 1982).

With this study, we intend to evaluate, for the first time, the genetic diversity, relationships and the structure of the six most representative Portuguese *P. nigra* populations using ISSRs and SCoT markers. We also aim to extrapolate about their provenances by comparison of the molecular patterns achieved with those of foreign samples belonging to the following subspecies: i) *P. nigra* subsp. *laricio* (var. *corsicana* and var. *calabrica*); ii) *P. nigra* subsp. *salzmannii*; and iii) *P. nigra* subsp. *nigra* (var. *austriaca*).

IV.3 – Material and methods

IV.3.1 – Plant material

In Portugal, the distribution of *P. nigra* it is represented by six pure and adult stands with an average age ranging from 50 to 90 years old. These stands have a reduced size and are

located in mountainous areas (with altitudes higher than 800m) at the North and Centre of the country. These populations present a regular and clean spacing among trees, typical of planted stands and good forestry management for avoidance of fires.

A total of 127 adult individuals (20 to 25 individuals per population) were sampled for needle tissue (Table IV.1) to be further used for genomic DNA extraction.

Table IV.1 – Number and average dendrometric measurements of the *P. nigra* individuals sampled in each population. In all populations the individuals were casually selected for sampling. Notes: Age – average age per population; h – average height of the individuals per population; DBH – Average diameter at breast height per population.

Local of the sampled population	Number of sampled individuals	Age (years)	Height (m)	DBH (cm)
'Manteigas'	20	93.33	24.35	34.07
'Vale do Zêzere'	20	58.79	14.95	24.8
'Paredes de Coura'	20	53.47	14.76	21.09
'Caminha'	25	52.73	26.3	32.56
'Campeã'	20	51.75	23.12	37.11
'Vila Pouca de Aguiar'	22	75.60	26.82	40.08
Total	127			

These stands were planted in Portugal during the XX century but the origin of the plant material used remains unknown. Based on the characterization of morphological and anatomical traits, Louro (1982) identified three *P. nigra* subspecies in Portugal: subsp. *laricio* (var. *calabrica* and var. *corsicana*), subsp. *nigra* and subsp. *salzmannii*. In order to extrapolate the provenances of the allochthonous Portuguese populations of *P. nigra* based on molecular data, we used dehydrated needles and certified seeds of foreign *P. nigra* individuals belonging to the subspecies and varieties assigned by Louro (1982). The dehydrated needles were harvested in a provenance test (COSE, ASPR, DOBL and BANA) and in a natural population (PNS) at the Southern of France, and the seeds of *P. nigra* subsp. *laricio* var. *corsicana* (PL0-VG-01 and PL0-VG-02) were collected in seed orchards (Table IV.2).

Table IV.2 – Taxonomic classification (according to Gaussen *et al*. 1964; 1993) of the foreign samples of *P. nigra*, sampling local and respective code used in this study.

Taxonomic classification		Sampling local	Code
Subspecies	Variety		
P. nigra subsp. *laricio*	*calabrica*	Cosenza (Calabria, Italy)	COSE
		Aspromonte (Calabria, Italy)	ASPR
P. nigra subsp. *laricio*	*corsicana*	Sologne Vayrieres VG	PLO-VG-01
		Corse Haute-Serre VG	PLO-VG-02
P. nigra subsp. *nigra*	*austriaca*	Doblhoff (Austria)	DOBL
		Banat (Romania)	BANA
P. nigra subsp. *salzmannii*	----	Saint Guilhem le Désert (France)	PNS

Twenty seeds of the foreign samples of *P. nigra* subsp. *laricio* subsp. *corsicana*, namely, PL0-VG-01 and PL0-VG-02 (Table IV.2) were germinated in Petri dishes containing moistened filter paper in the dark at 25 ºC for three weeks. Young needle tissue from these samples was collected and frozen at -80 ºC till genomic DNA extraction.

IV.3.2 – Genomic DNA extraction

The frozen and dehydrated needles of the *P. nigra* samples were used for genomic DNA extraction using the CTAB-based protocol of Doyle and Doyle (1987) with some modifications (Lemos *et al.* 2015, unpublished data). The integrity of the genomic DNA was evaluated after electrophoresis on 0.8% agarose gels stained with ethidium bromide and quantified in the Nanodrop ND-1000 (Thermo Scientific) spectrophotometer. The DNA samples were diluted to a concentration of 30ng/µL using ultra-pure distilled water (Gibco).

IV.3.3 – Amplification of ISSR and SCoT markers

For the ISSRs amplification a total of 15 primers were tested: seven from the set #9 of the University of British Columbia (UBC) that were previously used in *P. sylvestris* (Cipriano *et al*. 2013) and eight primers reported by Rubio-Moraga *et al*. (2012) for *P. nigra*. Each reaction mixture (final volume of 20 µL) contained 2 µL of total genomic DNA (30ng/µL), 1 µL of primer (5 µM), 10 µL of Taq-PCR master mix (Qiagen) and 7 µL of ultra-pure distilled water (Qiagen). The ISSR amplifications were carried on a T-Professional (Biometra) thermal cycler using: an initial denaturation step of 94 ºC for 5 min, followed by 45 cycles of denaturation of 94ºC for 30 sec, primer annealing at 47 ºC for 45 sec and extension at 72 ºC for 2 min, and a final extension of 72 ºC for 5 min.

Thirty-six primers designed by Collard and Mackill (2009) were used to amplify the SCoT markers. Each reaction mixture (final volume of 20 µL) was constituted by: 2 µL of genomic DNA (30ng/µL), 3.2 µL of primer (5µM), 2 µL of dNTPs mixture (2.5mM), 2 µL of 10x reaction buffer with KCl (Thermo Scientific), 1.2 µL of $MgCl_2$ 25 mM (Thermo Scientific), 9.2 µl of ultra-pure distilled water (Gibco) and 0.4 µL of *Taq* DNA polymerase (5U/µL) (Thermo Scientific). The amplification conditions followed those described by Collard and Mackill (2009). All PCR reactions were repeated at least twice, on different days and using different thermal cyclers, as suggested by Collard and Mackill (2009). The SCoT amplified products were visualized after electrophoresis on 2 % agarose gels stained with ethidium bromide.

An electrophoresis was performed at a constant voltage of 100V to visualize both ISSR and SCoT amplified products. The molecular weight marker GeneRuler DNA Ladder Mix (Thermo Scientific) was loaded on each gel. The gel images were captured by Gel Doc™ XR+ equipment (Bio-Rad, Hercules, USA) using the Image Lab Software (Bio-Rad, Hercules, USA).

IV.3.4 – Statistical analyses of the molecular data

Only reproducible ISSR and SCoT markers were analyzed for their presence (1) or absence (0) among individuals for construction of binary matrices. Each ISSR or SCoT band was considered a marker and bands with the same molecular weight produced by the same primer, as being the same *locus*.

In order to characterize the ability of each primer to detect polymorphic *loci* among the genotypes, the mean Polymorphic Information Content (PIC) that measures the discriminative capacity of a marker *locus* was calculated. PIC is a modification of the heterozygosity, subtracting from the *h* value an additional probability that an individual in a linkage analysis does not contribute with information to the study: $PIC = 1-\sum_{i=1}^{k} p_i^2 - \sum_{j=i+1}^{k-1} 2p_i^2 p_j^2$, were p_i is the frequency of the i^{th} allele, p_j the frequency of the j^{th} allele and *k* is the number of alleles (Botstein *et al.* 1980).

A general binary matrix combining the pool of the ISSR and SCoT data was used to construct an Unweighted Pair Group Method with Arithmetic means (UPGMA) dendrogram of genetic similarity with the software NTSYS pc ver. 2.02 (Rohlf 1998). The simple matching (SM) coefficient and the module of sequential agglomerative hierarchical nested (SAHN) cluster analysis were also used.

For the study of genetic structure it was used the software STRUCTURE Ver.2.3.4 (Pritchard *et al.* 2000; Falush *et al*. 2003; 2007), the 'no admixture' parameter (suitable for dominant markers), 50,000 generations of burn-in period followed by 100,000 Markov chain Monte Carlo (MCMC) iterations and different values of *K* (number of populations). The software STRUCTURE HARVESTER (Earl and vonHoldt 2012) was also used to identify the optimal *K* value. This program used the Evanno method for a fast analysis of the STRUCTURE output data (Evanno *et al.* 2005).

The Principal Coordinates Analysis (PCoA) and the Analysis of Molecular Variance (AMOVA) and the Mantel test between the genetic distance matrices achieved with ISSRs and SCoTs were performed with the software GenALEX 6.5 (Peakall and Smouse 2006: 2012).

The ISSR and SCoT data were also analysed with the software POPGENE 1.32 (Yeh *et al.* 1999), for dominant character of the markers and diploid character of the plant material. Multiple populations were selected as input options in the initial menu of the software and enabled the calculation of the following parameters: i) Shannon's information index (*I)*, which measures gene diversity (Shannon and Weaver 1949); ii) Nei's gene diversity index (h = $1-\sum p^2 i$, where p_i is the frequency of the i^{th} allele at the *locus*; Nei 1973); iii) the total genetic diversity (H_T); iv) the genetic diversity within populations (H_S); v) the relative magnitude of differentiation among populations (G_{ST}) (Nei 1987); and vi) the genetic diversity among populations ($D_{ST} = H_T - H_S$). The POPGENE software also provided the estimation of the unbiased genetic identity (Nei 1973) and genetic distance (Nei 1978).

The genetic relationships among the Portuguese samples and among these and the foreign samples of *P. nigra* were estimated by the construction of unrooted trees using the TreeView software (Page 1996), based on the genetic distance (Nei 1987) calculated by the software POPGENE and using the neighbour-joining method (Saitou and Nei 1987).

IV.4 – Results and discussion

IV.4.1 – Genetic diversity, relationships and structure of the Portuguese populations

Molecular markers can be used to study natural, managed and cultivated tree stands, and can measure the interconnections extension between individuals and populations (Mahajan and Gupta 2012).

Since the Portuguese *P. nigra* populations were never molecularly studied, a separated analysis of their genetic diversity, relationships and structure was performed. For these analyses, we pooled the ISSR and SCoT data achieved in the 127 *P. nigra* individuals of the six Portuguese allochthonous populations.

For ISSRs amplification we tested a total of 15 primers previously used by Cipriano *et al*. (2013) and Rubio-Moraga *et al*. (2012) in *P. sylvestris* and *P. nigra*, respectively. Eight primers were selected (Table 3) based on their amplification potential, discriminative pattern and/or reproducibility.

Thirty-six SCoT primers described by Collard and Mackill (2009) were tested. Eight oligonucleotides (Table IV.3) were selected for the amplification of SCoTs in the Portuguese *P. nigra* individuals. This selection was based on the same criteria mentioned for ISSRs.

Table IV.3 – Primers used for the amplification of ISSRs and SCoTs, respective sequences, number of total amplified (T), polymorphic (P) and monomorphic (M) bands, percentage of polymorphism (%P) and average polymorphism information content (PIC) produced per primer. Note: *Y= C/T

Marker	Primer	Sequence (5'→3')	T	P	M	%P	Average PIC
ISSR	817	$(CA)_8A$	20	20	0	100	0.817
	827	$(AC)_8G$	18	18	0	100	0.768
	834	$(AG)_8YT$*	11	11	0	100	0.798
	835	$(AG)_8YC$*	18	18	0	100	0.805
	836	$(AG)_8YA$*	7	7	0	100	0.628
	841	$(GA)_8YC$*	13	13	0	100	0.760
	850	$(GT)_8YC$*	12	12	0	100	0.702
	866	$(CTC)_6$	18	18	0	100	0.750
		TOTAL	117	117	0	100	---
Marker	**Primer**	**Sequence (5'→3')**	**T**	**P**	**M**	**%P**	**Average PIC**
SCoT	SCoT3	CAACAATGGCTACCACCG	24	24	0	100	0.764
	SCoT7	CAACAATGGCTACCACGG	21	21	0	100	0.835
	SCoT17	ACCATGGCTACCACCGAG	11	11	0	100	0.944
	SCoT18	ACCATGGCTACCACCGCC	22	22	0	100	0.977
	SCoT25	ACCATGGCTACCACCGGG	22	22	0	100	0.798
	SCoT26	ACCATGGCTACCACCGTC	15	12	3	80	0.560
	SCoT27	ACCATGGCTACCACCGTG	14	13	1	92.85	0.489
	SCoT31	CCATGGCTACCACCGCCT	22	22	0	100	0.628
		TOTAL	151	147	4	97.14	---

Among the 127 *P. nigra* individuals, most of the primers revealed 100% of ISSR or SCoT polymorphism (Table IV.3).

A total of 117 ISSR markers were produced with the eight selected primers, corresponding to an average of 14.63 bands per primer. Primer 817 presented the highest number of amplified bands and of average PIC value (Table IV.3), suggesting a high discriminative resolution among the *P. nigra* individuals, as also revealed by their elevated PIC values (Table IV.3). With the exception of the 866, all primers selected here for ISSRs amplification had already proved to be discriminative and successful for genetic diversity analysis and extrapolation of provenances in the Portuguese populations of *P. sylvestris* (Cipriano *et al*. 2013). Although these authors have obtained a higher number of total amplified bands per primer, the total percentage of ISSR polymorphism was equal to that achieved here (100%, Table IV.3). In addition, some of these primers were also used by Li *et al*. (2005) and Labra *et al*. (2006) in *P. sylvestris*, showing their high discriminative potential in pine species.

The eight SCoT primers produced a total of 151 amplified bands, an average of 18.88 bands per primer, and a total percentage of 97.14% of polymorphism (Table IV.3). Most of the SCoT primers presented high average PIC values, particularly, SCoT17 and SCoT18

(Table IV.3). The primers SCoT26 and SCoT27 presented the lowest PIC values (Table IV.3) due to the amplification of three (1,220bp, 1,000bp and 650bp) and one (900bp) monomorphic bands, respectively, in the 127 *P. nigra* individuals.

Globally, the 16 primers selected for the assessment of genetic diversity among the *P. nigra* individuals and populations proved to be highly discriminative (Table IV.3), and each primer (used for ISSR or SCoT amplifications) allowed the detection of polymorphism at the intra- and inter-population levels (Fig. IV.1).

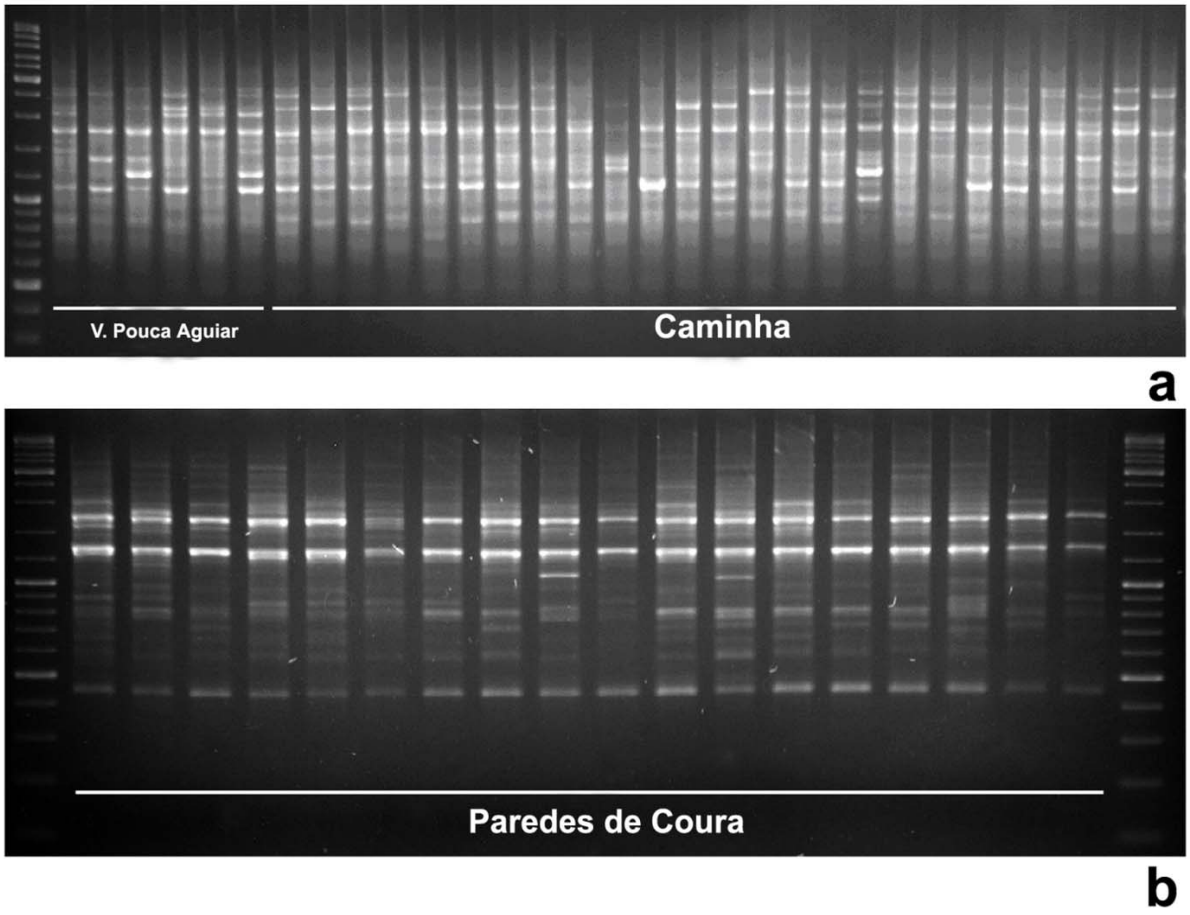

Figure IV.1 – a) Polymorphic ISSR patterns among 31 *P. nigra* individuals from the 'Vila Pouca de Aguiar' and 'Caminha' populations, produced with primer 835; and b) polymorphic SCoT patterns among 18 *P. nigra* individuals from the 'Paredes de Coura' population amplified with primer SCoT3.

ISSRs are dominant molecular markers with high reproducibility that can be used to detect DNA variability at different levels, from single base changes to deletions and insertions (López-Aljorna *et al*. 2007; Naik *et al*. 2009; Chezhian *et al*. 2010). Furthermore, polymorphisms can be detected without any previous knowledge of the target DNA sequence (Gupta *et al.* 1994; Godwin *et al.* 1997). ISSRs are highly polymorphic markers and have

been widely used in different plant species, alone or in combination with other markers, for the assessment of genetic diversity; estimation of phylogenies; DNA fingerprinting; extrapolation of provenances; gene-tagging, among other applications (Culley and Wolfe 2001; Carvalho *et al*. 2005; 2009; 2012; Ammiraju *et al*. 2011; Cipriano *et al*. 2013; Cabo *et al*. 2014a; Coutinho *et al.* 2014a,b; among many others). These markers have also been used in pine species for assessment of genetic diversity among populations or extrapolation of provenances (Li *et al.* 2005; Feng *et al.* 2006; Labra *et al.* 2006; Rubio-Moraga *et al.* 2012; Cipriano *et al.* 2013).

The oligonucleotide $(CTC)_6$ has the same sequence of the primer 866 from the set #9 of the University of British Columbia. This primer was used previously in *P. nigra* by Rubio-Moraga *et al.* (2012) who named it as ISCS41 [$(CTC)_6$] but since these authors did not report the number of amplified or polymorphic bands produced with this primer we could not compare with the present results. Rubio-Moraga *et al.* (2012) reported an average percentage of ISSR polymorphism of 51.04% among the populations of Southern Spain and Northern Morocco. The set of eight primers selected during our study evidenced 100% of ISSR polymorphism among the six most representative Portuguese populations of *P. nigra* (Table IV.3). A previous study performed by Lučić *et al*. (2010) with RAPDs in *P. nigra* revealed a total percentage of polymorphism of 88.5%. Thus, considering our results, ISSRs and SCoT seem to be highly polymorphic and more discriminative for genetic diversity assessment in such pine species.

The microsatellites are repetitive sequences randomly dispersed through the genomes. According to Morgante and Olivieri (1993), Wang *et al*. (1994) and Steinkellner *et al.* (1997), the $(GA)_n$ dinucleotide repeats are the most abundant in plants. Most of the primers selected in this work for ISSRs amplification are based on the repeat motifs $(AG)_n$ or $(GA)_n$. Those primers also produced the higher numbers of total amplified and polymorphic ISSR bands (Table IV.3). Moreover, the primer 866 with a trinucleotide repeat motif $(CTC)_6$, certainly hybridized with complementary microsatellite regions of motif $(GAG)_n$ to amplify the ISSR regions producing the highest number of total and polymorphic bands (Table IV.3). Based on these results we might suggest a high abundance of these microsatellite repeats also in the *P. nigra* genome. In this sense, these results allowed us to suggest a high abundance of SSRs with dinucleotide repeat motifs $(AG/TC)_n$ and $(CT/GA)_n$ in this pine species. This information could be useful for the future choice of SSR primers to be used in *P. nigra.* Due to the

dominant nature of the ISSRs, the use of codominant markers should be undertaken to complement the data provided by the former ones (Rubio-Moraga *et al*. 2012).

The dominant SCoT markers generate putative coding genomic regions in plants. Despite their potential for targeted fingerprinting or Quantitative Trait *Loci* (QTL) mapping purposes, they are also suitable for genetic diversity assessment (Poczai *et al*. 2013; Zhang *et al*. 2015). SCoTs could be used in combination with other markers, such as RAPDs, ISSRs, directed amplification of minisatellite-region DNA polymerase chain reaction (DAMD-PCR) or inter primer binding site (iPBS) for DNA fingerprinting to evaluate the genetic variation and relationships in different plant species (e.g. Luo *et al.* 2010; 2011; Amirmoradi *et al.* 2012; Guo *et al*. 2012; Mulpuri *et al.* 2013; Alikhani *et al.* 2014; Cabo *et al*. 2014b; Chen and Liu 2014; Agarwal *et al*. 2015). Only Alikhani *et al*. (2014) used SCoTs to assess the genetic diversity and structure of a forestry species, *Quercus brantii*. Nevertheless, our study constitutes the first use of SCoTs in a conifer species. The percentage of polymorphism and total number of amplified bands (Table IV.3), demonstrated their high discriminative potential for the genetic diversity evaluation of *P. nigra* within and among populations.

Collard and Mackill (2009) reported that the SCoT amplified products in rice cultivars ranged in size from 200 to 1,500bp. Recent studies of the application of SCoTs to other plant species showed wider ranges of molecular size for the amplified products: 250-4,000bp in mango cultivars (Luo *et al.* 2010); 220-2,250bp in *Cicer* species (Amirmoradi *et al.* 2012); 100-2,800bp in *Jatropha curcas* L. and *Quercus brantii* (Mulpuri *et al.* 2013; Alikhani *et al.* 2014); 500-2,250bp in newly synthesized tritordeums and their respective parents (Cabo *et al*. 2014b); 100-2,000bp in *Myrica rubra* (Chen and Liu 2014) and 100-1,100bp in *Alhagi maurorum* (Agarwal *et al.* 2015). In this study we visualized amplified SCoTs products with a molecular size ranging from 100 to 4,000bp and for ISSRs a molecular range from 450 to 4,000bp.

Usually, RAPDs and ISSRs present a molecular size ranging from 300 to 3,000bp (Carvalho *et al*. 2005; Lima-Brito *et al*. 2006; Carvalho *et al*. 2009; 2012; Coutinho *et al*. 2014a,b). The amplification of SCoT and ISSR fragments with large size (over than the usual 3,000bp) in the *P. nigra* individuals might be explained by the high complexity of the conifer genomes. Conifers are key representatives of gymnosperms and the complete size of their genomes represents a significant challenge for characterization, sequencing and assembling, since they are among the largest genomes of all known organisms (Murray 1998; Pavy *et al*. 2012).

Since genetic diversity assessment based on a high number of molecular markers is considered more statistically reliable, the ISSR and SCoT data produced in the 127 individuals with a total of 16 primers were pooled. The UPGMA dendrogram revealed 65% of genetic similarity among the 127 *P. nigra* individuals (Fig. IV.2).

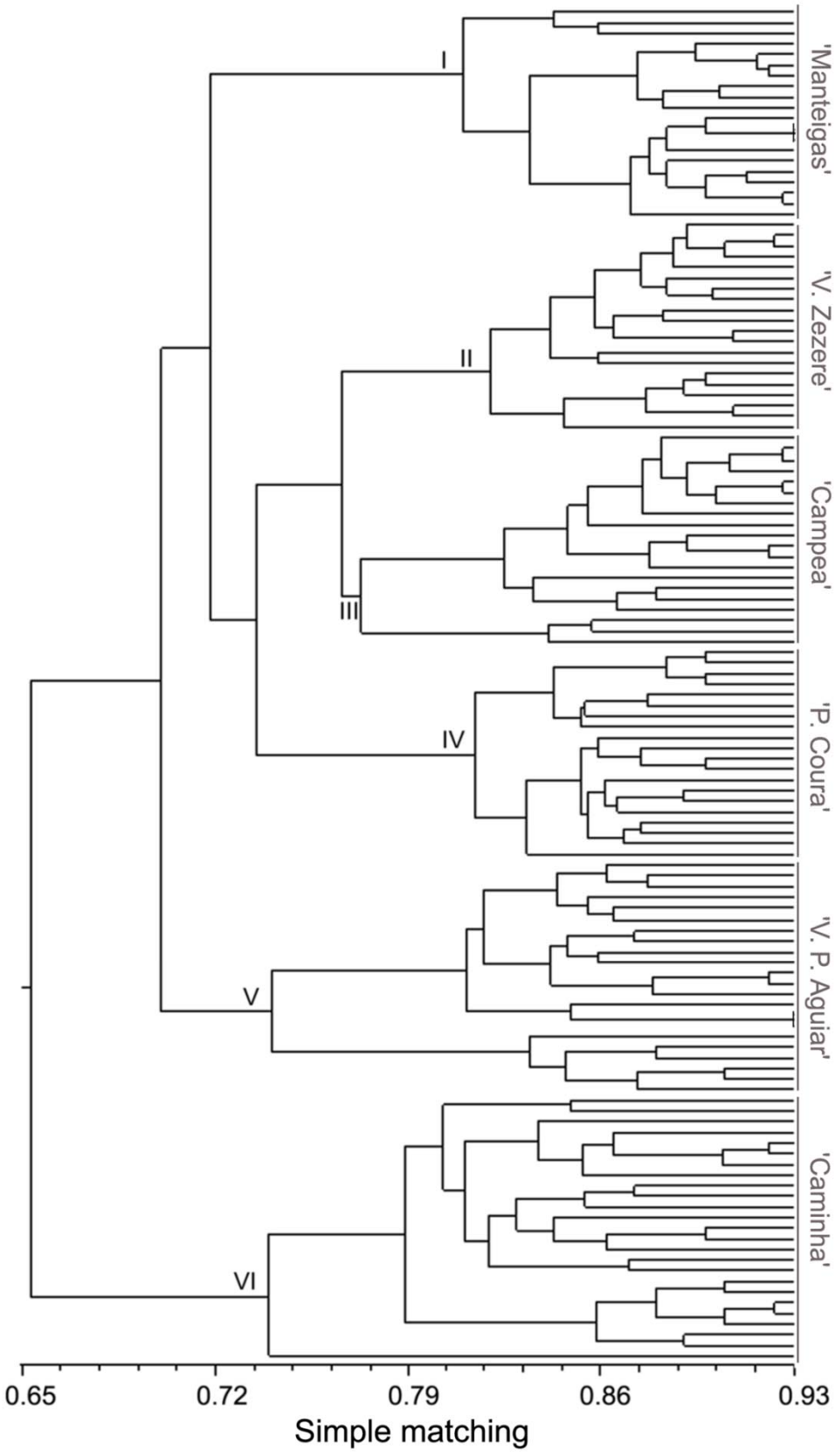

Figure IV.2. – UPGMA dendrogram of genetic similarity among the 127 *P. nigra* individuals based on the pool of the ISSR and SCoT data, with all individuals clustered into six groups (I to VI), corresponding to the number of studied populations.

Accordingly to previous dendrometric analyses (Dias *et al*. 2014), the population of 'Manteigas' was the oldest installed stand since it presents an average age of 90 years old (Table IV.1). 'Vila Pouca de Aguiar' has an average age of 70 years old, 'Vale do Zêzere' has an average age of 60 years old and the remaining three stands have an average age of 50 years old (Table IV.1). In the dendrogram, the populations were not clustered by the age criterion, probably due to the use of plant material with different provenances in the installation of these stands. After the morphological characterization of the Portuguese stands of *P. nigra* existent in the 1980's, Louro (1982) proposed the presence of three subspecies, namely, *laricio*, *nigra* and *salzmannii*, which might explain the previous result.

For a cut off value of 0.734 of genetic similarity, the UPGMA dendrogram showed the 127 individuals clustered into six groups (I to VI) that corresponded to the number of studied populations and suggested a highly defined structure (Fig. IV.2). The populations of 'Manteigas' and 'Vale do Zêzere', which are geographically close, showed the highest value of Nei's unbiased measures of genetic identity (0.9049) estimated by the Nei's method (Nei 1978) with the software POPGENE. By other hand, the highest Nei's genetic distance (0.1606) was detected between 'Manteigas' and 'Caminha', corroborating their separated UPGMA clustering (Fig. IV.2).

For the extrapolation of the genetic structure of the Portuguese *P. nigra* populations we used the 'no admixture' parameter and both the STRUCTURE and STRUCTURE HARVESTER software. After several runs of *K* (number of populations) based on the pool of the ISSR and SCoT data, the STRUCTURE software resolved $K = 6$ highly differentiated genetic clusters (Fig. IV.3; Table IV.4), confirming the UPGMA clustering (Fig. IV.2). The number of populations was further confirmed with the software STRUCTURE HARVESTER which defined $K=6$ as the optimal value. As reported by Pritchard *et al*. (2000), we may not always be able to know the true value of *K*, but we should aim for the smallest value of *K* that captures the major structure in the data. Besides, the mean F_{ST} values attributed to each group were near or above the value of 0.25, indicating a good genetic differentiation among the Portuguese *P. nigra* populations (Table IV.4), and being indicative of local adaptation as reported by Savolainen *et al*. (2007). The Portuguese *P. nigra* populations belong to pure stands with average values of height and diameter at the breast height (DBH) that also denote a highly adapted species to our climate and their ecological environments (Table IV.1; Dias *et al*. 2014).

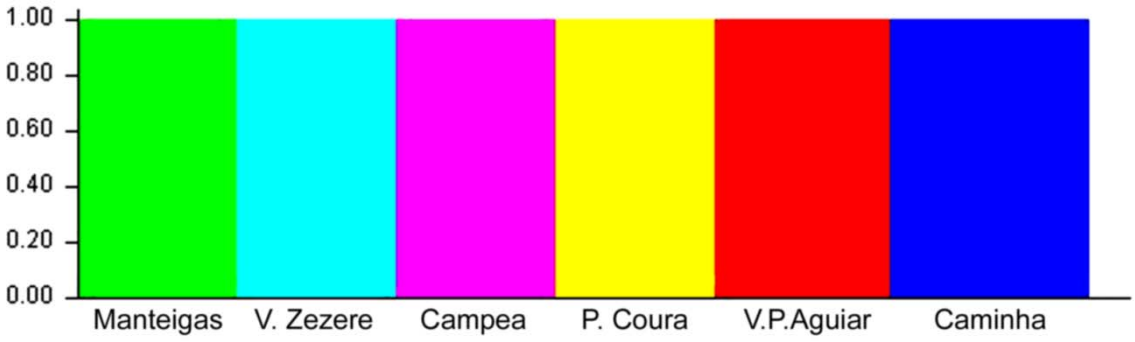

Figure IV.3 – Bar plot diagram provided by the STRUCTURE software estimating the genetic structure of the Portuguese populations of *P. nigra* into six populations (*K*=6) based on the pool of the ISSR and SCoT data.

Table IV.4 – Mean F_{ST} values per group achieved with $K = 6$ using the STRUCTURE software based on the pool of ISSR and SCoT data.

Group	Mean F_{ST} value for $K = 6$
I	$F_{ST}\_1 = 0.2758$
II	$F_{ST}\_2 = 0.2356$
III	$F_{ST}\_3 = 0.2552$
IV	$F_{ST}\_4 = 0.2487$
V	$F_{ST}\_5 = 0.2489$
VI	$F_{ST}\_6 = 0.2469$

AMOVA is recognized as an effective tool to define population structure and degree of genetic differentiation (Excoffier *et al.* 1992). This analysis allowed us to verify a higher molecular variation within (57%) rather than among populations (43%). Similar results were proposed in other genetic diversity studies of *P. nigra* and other pine species by Feng *et al.* (2006), Labra *et al.* (2006), Naydenov *et al. (*2006) and Rubio-Moraga *et al.* (2012). The POPGENE software also provided a higher value of intra-population genetic diversity (H_S = 0.1652 ± 0.0142) than the inter-population genetic diversity (D_{ST} = 0.0798), as well as a high total percentage of polymorphic *loci* (99.49%).

Some authors have reported that molecular variance could be related with geographical distance of the populations, with the natural selection of their environments, and with different ecological microenvironments, inducing significant differences in their genetic structure and ecological segregation (Taylor and Aarssen 1990; Feng *et al.* 2006). Moreover,

Hamrick and Godt (1989) suggested that gymnosperms with long life span, high outcrossing rates and fecundity, maintain high intra-population genetic diversity.

The present results are in concordance with other studies that indicated that *P. nigra* is a highly polymorphic species (Lucić *et al.* 2010; Šarac *et al.* 2014), generating problems with the definition of taxonomy in terms of species, subspecies and varieties (Afzal-Raffi and Dodd 2007).

The values of Shannon's Information index (I; Lewontin 1972), Nei's (1973) gene diversity (h) and number of polymorphic *loci* achieved per population are presented in Table 5. 'Caminha' presented the highest values of Shannon's Information index (I), Nei's gene diversity (h) and percentage of polymorphic *loci* (Table IV.5). In all populations, the values of Nei's gene diversity were lower than the values of Shannon's index (Table IV.5). Similar results were previously reported for Southern Spain and Northern Morocco populations of *P. nigra* studied by Rubio-Moraga *et al.* (2012), and for other pine species, namely, *Pinus koraiensis* (Feng *et al.* 2006) and *P. sylvestris* (Li *et al.* 2005, Cipriano *et al.* 2013). The maximum and minimum values of Shannon's information index (I) obtained in this work were higher than those reported by Rubio-Moraga *et al*. (2012), but these authors only used ISSRs. Comparatively to the work of Rubio-Moraga *et al*. (2012), we obtained higher values of total genetic diversity ($H_T = 0.2450 \pm 0.0283$), percentage of molecular variation and G_{ST} (0.3257), and these results could be explained by the use of a different set of primers for ISSRs amplification and inclusion of SCoTs. Higher H_T values were also achieved with ISSRs by other authors in different *Pinus* sp.: *P. sylvestris* (0.4965; Cipriano *et al.* 2013), *Pinus tabuleaformis* (0.4152; Wang and Hao 2010), *P. koraiensis* (0.3477; Feng *et al.* 2006) and *Pinus sibirica* (0.2699; Yang *et al.* 2005). Rubio-Moraga *et al.* (2012) argued that the differences in total genetic diversity detected among different pine species could be due to their geographic distribution, number of tested populations, population size and effect of climate changes during the last glacial period.

Table IV.5 – Summary of genetic variation and gene diversity statistical analyses achieved per population based on the pool of the ISSR and SCoT data.

Population	Mean ± standard deviation		% of Polymorphic
	I	h	*loci*
'Manteigas'	0.2196± 0.2702	0.1444± 0.1874	46.29
'Vale do Zêzere'	0.2200 ± 0.2743	0.1456 ± 0.1912	45.52
'Campeã'	0.2517± 0.2813	0.1672 ± 0.1963	50.38
'Paredes de Coura'	0.2304 ± 0.2711	0.1512 ± 0.1885	49.10
'Vila Pouca de Aguiar'	0.2702 ± 02736	0.1774 ± 0.1892	55.50
'Caminha'	0.3081 ± 0.2840	0.2053 ± 0.1997	61.38

The present results of genetic diversity, relationships and structure of the six populations representative of the European Black Pine distribution in Portugal, can be taken into account in future strategies of conservation of these genetic resources, forest management, definition of afforestation programs and/or use of germplasm.

IV.4.2 – Extrapolation of provenances

In this study we applied the approach of using bulks of populations (also called population DNA pooling strategy; Liber *et al.* 1993) as reported in other studies (Furman *et al.* 1997; Liber *et al.* 2003).

In order to extrapolate about the origin of the forest reproductive material (FRM) used in the installation of the Portuguese *P. nigra* stands during the last century, we compared the molecular patterns achieved in the Portuguese samples with those produced in seven DNA bulks representative of foreign individuals belonging to three allopatric subspecies and varieties: subsp. *nigra* var. *austriaca*; subsp. *laricio* var. *corsicana* and var. *calabrica*; and subsp. *salzmannii* (see Table IV.2). These three subspecies were previously assigned as present in Portugal in the 1980's by Louro (1982). The ISSR and SCoT patterns of the foreign samples were also analyzed by presence/ absence of band and included in the binary matrix of the Portuguese samples, resulting on a total of 134 individuals and 314 markers. After constructing an UPGMA dendrogram of genetic similarity among the Portuguese and foreign samples of *P. nigra* based on the pool of the ISSR and SCoT data, the foreign samples showed 75% of genetic similarity among them and 54% of genetic similarity with the Portuguese samples (Fig. IV.S1). Since the foreign samples were clustered into one single group apart from the Portuguese ones, this analysis was not very explicit about the extrapolation of provenances (Fig. IV.S1). Due so, we decided to perform a Mantel test between the genetic distance matrices achieved with ISSRs and SCoTs (calculated with the

Nei's coefficient and using 9999 permutations) using the GenAlEX software and we achieved a reduced correlation coefficient (r = 0.017). This reduced value could be explained by the principles underlying the two methodologies used. The ISSRs are abundant and disperse throughout the genome and could amplify coding and noncoding regions, whereas the SCoT markers use primers that bind to a short conserved sequence that flanks the start codon (ATG) of the plant genes (Collard and Mackill 2009). Regarding the conservation degree of plant genes even among different species, SCoT markers could provide more reliable molecular data than the ISSRs for the extrapolation of provenances of the Portuguese European Black Pine populations. Therefore, the following analyses were performed using ISSR and SCoT data separately.

The seven DNA bulks of foreign *P. nigra* samples were representative of four different *taxa*: i) *P. nigra* subsp. *laricio* var. *calabrica* (COSE and ASPR); ii) *P. nigra* subsp. *laricio* var. *corsicana* (PL0VG01 and PL0VG02); iii) *P. nigra* subsp. *nigra* var. *austriaca* (DOBL and BANA); and iv) *P. nigra* subsp. *salzmannii* (PNS).

Tables IV.6 and IV.7 present the results of Nei's unbiased measures of genetic identity among the Portuguese and foreign samples revealed by the ISSR and SCoT markers, respectively.

Table IV.6 – Pairwise matrix of Nei's genetic identity among the Portuguese *P. nigra* populations and foreign samples representative of four distinct infraspecific *taxa*, based on the pool of the ISSR data.

'Manteigas'	'V. Zêzere'	'P. Coura'	'Campeã'	'V. P. Aguiar'	'Caminha'	subsp. *laricio* var. *calabrica*	subsp. *nigra* var. *austriaca*	subsp.*laricio* var. *corsicana*	subsp. *salzmannii*	
1.000										'Manteigas'
0.953	1.000									'V. Zêzere'
0.957	0.950	1.000								'P. Coura'
0.950	0.950	0.951	1.000							'Campeã'
0.951	0.947	0.949	0.957	1.000						'V. P. Aguiar'
0.959	0.957	0.955	0.958	0.959	1.000					'Caminha'
0.826	**0.811**	**0.818**	**0.844**	**0.835**	**0.834**	1.000				**subsp. *laricio* var. *calabrica***
0.798	0.776	0.804	0.789	0.808	0.783	0.667	1.000			subsp. *nigra* var. *austriaca*
0.797	0.796	0.815	0.796	0.800	0.821	0.701	0.701	1.000		subsp. *laricio* var. *corsicana*
0.698	0.679	0.696	0.690	0.698	0.704	0.633	0.660	0.588	1.000	subsp. *salzmannii*

Table IV.7 – Pairwise matrix of Nei's genetic identity among the Portuguese *P. nigra* populations and foreign samples representative of four distinct infraspecific *taxa*, based on the pool of the SCoT data.

'Manteigas'	'V. Zêzere'	'P. Coura'	'Campeã'	'V. P. Aguiar'	'Caminha'	subsp. *laricio* var. *calabrica*	subsp. *nigra* var. *austriaca*	subsp.*laricio* var. *corsicana*	subsp. *salzmannii*	
1.000										'Manteigas'
0.945	1.000									'V. Zêzere'
0.942	0.957	1.000								'P. Coura'
0.946	0.957	0.955	1.000							'Campeã'
0.949	0.953	0.955	0.956	1.000						'V. P. Aguiar'
0.958	0.952	0.951	0.958	0.955	1.000					'Caminha'
0.782	0.786	0.786	0.791	**0.808**	0.780	1.000				subsp. *laricio* var. *calabrica*
0.785	0.813	0.766	0.767	0.779	0.773	0.674	1.000			subsp. *nigra* var. *austriaca*
0.799	**0.799**	**0.799**	**0.795**	0.789	**0.791**	0.699	0.673	1.000		**subsp. *laricio* var. *corsicana***
0.676	0.691	0.706	0.693	0.660	0.693	0.560	0.585	0.644	1.000	subsp. *salzmannii*

Regarding the ISSR data, the foreign DNA bulks of *P. nigra* subsp. *laricio* var. *calabrica* showed higher values of Nei's genetic identity with the six Portuguese *P. nigra* populations (Table IV.6).

The pairwise matrix of Nei's genetic identity calculated based on the pool of the SCoT data revealed higher genetic identity among five Portuguese populations and the foreign bulked samples of *P. nigra* subsp. *laricio* var. *corsicana* (Table IV.7). Based on the SCoT data, the Portuguese population of 'Vila Pouca de Aguiar' showed higher genetic identity with the foreign samples of *P. nigra* subsp. *laricio* var. *calabrica* (Table IV.7).

Both marker systems evidenced a higher genetic identity among the Portuguese *P. nigra* individuals and the foreign samples of subspecies *laricio* (Tables IV.6 and IV.7). These results are partially in accordance with the previous comparative analyses of phenotypic data, silviculture characterization, botanic classification and evaluation of the edaphoclimatic conditions performed in the European Black Pine stands existent in Portugal in the 1980's (Louro 1982). According to this author, *P. nigra* subsp. *laricio* var. *corsicana* was the most predominant and well adapted to Portugal.

Both marker systems revealed that the *P. nigra* subsp. *salzmannii* was the most genetically distant from the Portuguese populations (Tables IV.6 and IV.7) and this result was confirmed by the two PCoA performed (Fig. IV.4). The ISSRs explained a higher cumulative percentage of total variation (92.62%) than the SCoT markers (87.74%), based on the first three axes, among the Portuguese and foreign *P. nigra* samples (Fig. IV.4).

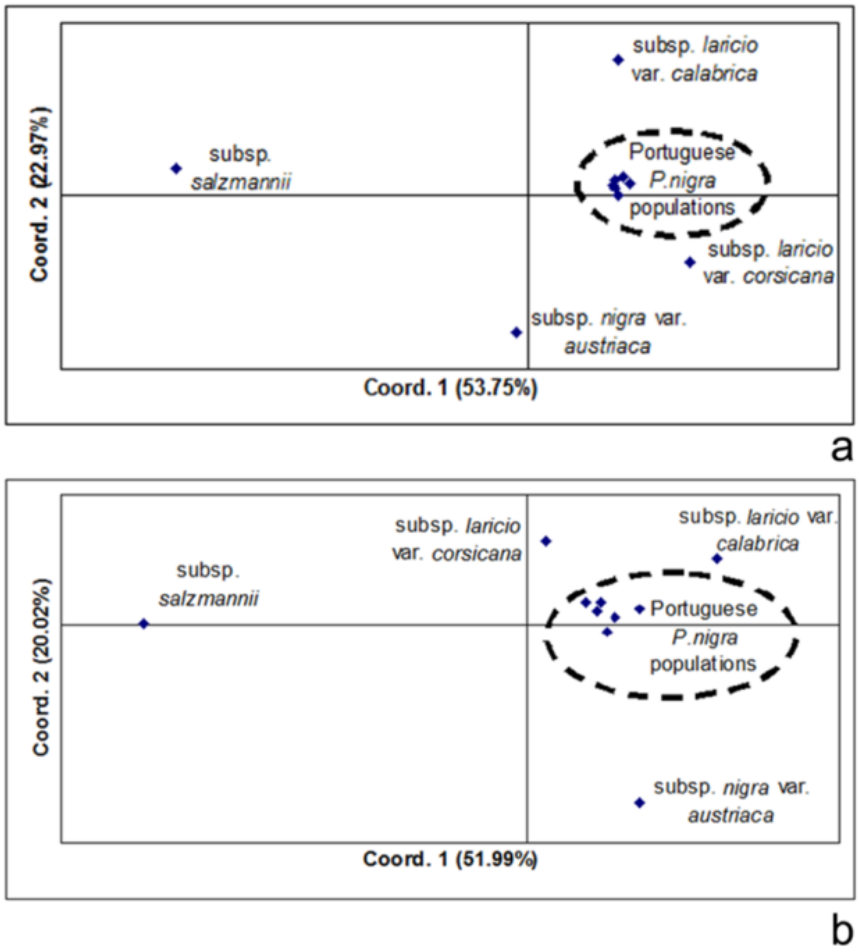

Figure IV.4 – Principal Coordinates Analysis (PCoA) based on the Nei's genetic distance matrix among the 127 Portuguese individuals and the seven DNA bulks of foreign *P. nigra* samples based on the pool of: a) ISSR; and b) SCoT data.

The two marker systems used in this work already proved to be suitable for DNA fingerprinting among individuals belonging to different *taxa* (Coutinho *et al*. 2014a,b); among interspecific hybrids and their parental species (Carvalho *et al*. 2005); among tritordeum and respective parents (Cabo *et al*. 2014a,b); and/or for extrapolation of provenances (Cipriano *et al*. 2013). However, the amplification of SCoTs requires the binding of the primer to a short conserved sequence that flanks the start codon, contributing for the comparison of highly conserved regions that could provide the estimation of more accurate genetic relationships and estimation of provenances.

IV.5 – Conclusions

This study constituted the first molecular characterization of the six *P. nigra* populations that are representative of the distribution of this pine species in Portugal. The understanding of the genetic diversity, structure and relationships of these populations will be highly important under the scope of forestry management, genetic improvement and/or for the definition of afforestation and conservation strategies. Based on the molecular data provided by both ISSRs and SCoTs we verified that these allochthonous populations have high genetic diversity, mostly at the intra-population level and a highly differentiated structure. The data achieved during this study also revealed a high adaptive potential of the species to our country, corroborating the previous dendrometric and ecological evaluations done during sampling (Dias *et al.* 2014).

In conclusion, the ISSR and SCoT markers proved to be reliable for the discrimination of individuals per population and also for the extrapolation of provenances of the Portuguese populations. Globally, both marker systems revealed that the Portuguese populations have higher genetic identity with the subspecies *laricio*, corroborating the assumption of Louro (1982) who suggested that this subspecies was the most predominant and well adapted to our country.

IV.6 – Acknowledgements

Ana Carvalho and Alexandra Dias thank the postdoctoral SFRH/BPD/68932/2010 and doctoral SFRH/BD/91781/2012 grants, respectively, both attributed by the Portuguese Foundation for Science and Technology (Fundação para a Ciência e a Tecnologia, FCT) and co-financed by the Social European Fund (FSE) under the POPH-QREN program.

The authors acknowledge to Prof. Bruno Fady for kindly given dehydrated needles of *P. nigra* and to Dr. Marc Calvignac from Vilmorin S.A. (La Ménitré, France) for kindly provided certified seed material of *P. nigra*, of different allopatric subspecies and varieties.

IV.6 – References

Afzal-Rafii Z, Dodd RS (2007) Chloroplast DNA supports a hypothesis of glacial refugia over postglacial recolonization in disjunct populations of black pine (*Pinus nigra*) in Western Europe. Molecular Ecology 16:723-736.

Agarwal T, Gupta AK, Patel AK, Shekhawat NS (2015) Micropropagation and validation of genetic homogeneity of *Alhagi maurorum* using SCoT, ISSR and RAPD markers. Plant Cell Tissue and Organ Culture 120(1):313-323.

Alikhani L, Rahmani M-S, Shabanian N, Badakhshan H, Khadivi-Khub A (2014) Genetic variability and structure of *Quercus brantii* assessed by ISSR, IRAP and SCoT markers. Gene 552:176-183.

Amirmoradi B, Talebi R, Karami E (2012) Comparison of genetic variation and differentiation among annual *Cicer* species using Start Codon Targeted (SCoT) polymorphism, DAMD-PCR, and ISSR markers. Plant Systematics and Evolution 298(9):1679-1688.

Ammiraju JSS, Dholakia BB, Santra DK, Singh H, Lagu MD, Tamhankar, Dhaliwal HS, Rao VS, Gupta VS, Ranjekar PK (2001) Identification of inter simple sequence repeat (ISSR) markers associated with seed size in wheat. Theoretical and Applied Genetics 102:726-732.

Arbez M, Millier C (1971) Contribution à l´étude de la variabilité géographique de *Pinus nigra*, Arn.. In: Annales des Sciences Forestiéres, Paris, vol. 28, no. 1.

Avise JC (1994) Molecular Markers, Natural History, and Evolution. Chapman and Hall, New York. 511 pp.

Barbéro M, Loisel R, Quézel P, Richardson DM, Romane F (1998) Pines of Mediterranean basin. In: Ecology and Biogeography of Pinus (edited by DM Richardson), pp 153–170, Cambridge University Press.

Blanco E (1998) Los bosques Ibéricos. Una interpretación geobotánica. Planeta Barcelona, Barcelona.

Bogunić F, Muratovi E, Ballian D, Siljak-Yakolev S, Brown S (2007) Genome size stability of five subspecies of *Pinus nigra* Arnold s.l. Environmental and Experimental Botany 59:354-360.

Bogunić F, Muratovic E, Brown SC, Siljak-Yakovlev S (2003) Genome size of five *Pinus* from Balkan region. Plant Cell Reports 22:59–63.

Bogunić F, Siljak-Yakovlev S, Muratović E, Ballian D (2011) Different karyotype patterns among allopatric *Pinus nigra* (Pinaceae) populations revealed by molecular cytogenetics. Plant Biology 13:194-200.

Bonavita S, Vendramin GG, Bernardini V, Avolio S, Regina TMR (2015) The first SSR-based assessment of genetic variation and structure among *Pinus laricio* Poiret populations within their native area, Plant Biosystems - An International Journal Dealing with all Aspects of Plant Biology: Official Journal of the Societa Botanica Italiana. doi: 10.1080/11263504.2015.1027316

Cabo S, Carvalho A, Rocha L, Martín A, Lima-Brito J (2014a) IRAP, REMAP and ISSR fingerprinting in newly formed hexaploid tritordeum (X Tritordeum Ascherson et Graebner) and respective parental species. Plant Molecular Biology Reporter 32(3):761-770.

Cabo S, Ferreira L, Carvalho A, Martins-Lopes P, Martín A, Lima-Brito JE (2014b) Potential of Start Codon Targeted (SCoT) markers for DNA fingerprinting of newly synthesized tritordeums and their respective parents. Journal of Applied Genetics 55:307-312.

Carrión JS, Van Geel B (1999) Fine-resolution upper Weichselian and Holocene palynological record from Navarrés (Valencia, Spain) and a discussion about factors of Mediterranean forest succession. Review of Palaeobotany and Palynology 106:209-236.

Carvalho A, Guedes-Pinto H, Lima-Brito J (2012) Genetic diversity in Old Portuguese durum wheat cultivars assessed by retrotransposon-based markers. Plant Molecular Biology Reporter 30:578-589.

Carvalho A, Lima-Brito J, Maçãs B, Guedes-Pinto H (2009) Genetic diversity and variation among botanical varieties of Old Portuguese wheat cultivars revealed by ISSR assays. Biochemical Genetics 47(3-4):276-294.

Carvalho A, Matos M, Lima-Brito J, Guedes-Pinto H, Benito C (2005) DNA fingerprint of F_1 interspecific hybrids from the Triticeae tribe using ISSRs. Euphytica 143:93-99.

Çengel B, Tayanç Y, Kandemir G, Velioglu E, Alan M, Kaya Z (2012) Magnitude and efficiency of genetic diversity captured from seed stands of *Pinus nigra* (Arnold) subsp. *pallasiana* in established seed orchards and plantations. New Forests 43:303-317.

Chen F-Y, Liu J-H (2014) Germplasm genetic diversity of *Myrica rubra* in Zhejiang Province studied using inter-primer binding site and Start Codon-Targeted polymorphism markers. Scientia Horticulturae 170:169-175.

Chezhian P, Yasodha R, Ghosh M (2010) Genetic diversity analysis in a seed orchard of Eucalyptus tereticornis. New Forests 40:85-99.

Christensen KI (1997) Pinaceae, Cupressaceae, Taxaceae, Ephedraceae, Salicaceae, Juglandaceae, Betulaceae, Fagaceae, Ulmaceae Moraceae. In: Flora Hellenica, vol 1 (edited by A Strid, K Tan), pp 1–17, Koeltz Scientific Books , Königstein, Germany.

Cipriano J, Carvalho A, Fernandes C, Gaspar MJ, Pires J, Bento J, Roxo L, Louzada J, Lima-Brito J (2013) Evaluation of genetic diversity of Portuguese *Pinus sylvestris* L. populations based on molecular data and inferences about the future use of this germplasm. Journal of Genetics 92:e41-e48.

Collard BCY, Mackill DJ (2009) Start Codon Targeted (SCoT) polymorphism: A simple novel DNA marker technique for generating gene-targeted markers in plants. Plant Molecular Biology Reporter 27:86–93.

Coutinho JP, Carvalho A, Lima-Brito J (2014a) Genetic diversity assessment and estimation of phylogenetic relationships among 26 Fagaceae species using ISSRs. Biochemical Systematics and Ecology 54:247–256.

Coutinho JP, Carvalho A, Lima-Brito J (2014b) Fingerprinting of Fagaceae individuals using inter-microsatellite markers. Journal of Genetics - Online Resources 93:e132-e140.

Culley TM, Wolfe AD (2001) Population genetic structure of the cleistogamous plant species *Viola pubescens* Aiton (Violaceae), as indicated by allozyme and ISSR molecular markers. Heredity 86:545-556.

Debazac EF (1964) Manuel des conifers. Ecole Nationale des Eaux et Forêts, Nancy.

del Cerro Barja A, Lucas-Borja ME, Garcia EM, López Serrano FR, Abellán MA, Morote FAG, López RN (2009) Influence of stand density and soil treatment on the Spanish Black Pine (*Pinus nigra* Arn. ssp. *salzmannii*) regeneration in Spain. Investigación Agraria: Sistemas y Recursos Forestales 18:167-180.

Delevoy G (1949) Contribution à l´etude de quelques variétés de *Pinus nigra*, Arn. en Bélgique. Travaux de la Station de Recherches de Groendaal, no. 8-12.

Dias AL, Fonseca TF, Silva ME, Gaspar MJ, Carvalho AI, Lousada JL (2014) *Pinus nigra* growth in Portugal. Book of Abstracts of the 5th International Conference on Mediterranean Pines (MedPine5), 22-26 September 2014, Solsona, Spain.

Doyle JJ, Doyle JL (1987) A rapid DNA isolation procedure for small quantities of fresh leaf tissue. Phytochem Bulletin 19:11-15.

Earl DA, vonHoldt BM (2012) STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. Conservation Genetics Resources 4:359-361.

Espelta JM, Retana J, Habrouk A (2003) An economic and ecological multi-criteria evaluation of reforestation methods to recover burned Pinus nigra forests in NE Spain. Forest Ecology and Management 180:185–198.

Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Molecular Ecology 14:2611-2620.

Excoffier L, Smouse PE, Quattro JM (1992) Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. Genetics 131:479–491.

Fady B, Lefèvre F, Reynaud M, Vendramin GG, Bou Dagher-Kharrat M, Anzidei M, Pastorelli R, Savoure´ A, Bariteau M (2003) Gene flow among different taxonomic units: evidence from nuclear and cytoplasmic markers in *Cedrus* plantation forests. Theoretical and Applied Genetics 107:1132–1138.

Falush D, Stephens M, Pritchard JK (2003) Inference of population structure using multilocus genotype data: linked *loci* and correlated allele frequencies. Genetics 164:1567-1587.

Falush D, Stephens M, Pritchard JK (2007) Inference of population structure using multilocus genotype data: dominant markers and null alleles. Molecular Ecology Notes 7:574-578.

Feng FJ, Han SJ, Wang HM (2006) Genetic diversity and genetic differentiation of natural *Pinus koraiensis* populations. Journal of Forest Research 17:21-24.

Furman BJ, Grattapala D, Dvorak WS, O'Malley DM (1997) Analysis of genetic relationships of Central American and Mexican pines using RAPD markers that distinguish species. Molecular Ecology 6:321-331.

García Antón M, Franco Múgica F, Ruiz MJ, Juaristi CM, Ollero HS (1995) Una secuencia polínica en Quintana Redonda (Soria). Evolución holocena del tapiz vegetal en el Sistema Ibérico septentrional. Anales del Jardín Botánico de Madrid 52:187-195.

Génova M, Fernández A (1998/1999) Tree rings and climate of *Pinus nigra* subsp. *salzmannii* in Central Spain, Dendrochronologia 16–17:75–86.

Gaussen H, Haywood VH, Chater AO (1964) *Pinus*. In: Flora Europaea, Vol. 1 (edited by TG Tutin, VH Heywood, NA Burges, DM Moore, DH Valentine, SM Walters, DA Webb), pp. 32-35, Cambridge University Press, Cambridge, UK.

Gaussen H, Webb DA, Heywood VH (1993) Genus *Pinus*. In: Flora Europaea, Vol 1 (edited by TG Tutin, VH Heywood, NA Burges, DM Moore, DH Valentine, SM Walters, DA Webb), pp. 40-44, Cambridge University Press, Cambridge, UK.

Gaussen H (1960) Les gymnospermes actuelles et fossiles. Fasc. VI. Les Coniférales. Faculté des Sciences, Toulouse, France.

Godwin ID, Aitken EAB, Smith LW (1997) Application of inter simple sequence repeat (ISSR) markers to plant genetics. Electrophoresis 18:1524–1528.

Gorji AM, Poczai P, Polgar Z, Taller J (2011) Efficiency of arbitrarily amplified dominant markers (SCoT, ISSR and RAPD) for diagnostic fingerprinting in tetraploid potato. American Journal of Potato Research 88:226–237.

Greuter W, Burdet HM, Long G (eds) (1984) Med-checklist, *Pterydophyta Gymnospermae*, *Dycotiledones* (*Acanthaceae-Cneoraceae*). Editions des Conservatoire et Jardin botaniques de la Ville de Genève, vol 1, Genève, pp. 33–34.

Guo D-L, Zhang J-Y, Liu C-H (2012) Genetic diversity in some grape varieties revealed by SCoT analyses. Molecular Biology Reports 39:5307–5313.

Gupta M, Chyi YS, Romero-Severson J, Owen JL (1994) Amplification of DNA markers from evolutionarily diverse genome using single primers of simple-sequence repeats. Theoretical and Applied Genetics 89:998–1006.

Hamrick J, Godt M (1989) Allozyme diversity in plant species. In: Differentiation patterns in higher plants (edited by K Urbanski), pp. 53–67, Academic Press, New York, USA.

Isajev V, Fady B, Semerci H Andonovski V (2004) EUFORGEN Technical Guidelines for genetic conservation and use for European black pine (*Pinus nigra*). International Plant Genetic Resources Institute, Rome.

Labra M, Grassi F, Sgorbati S, Ferrari C (2006) Distribution of genetic variability in southern populations of Scots pine (*Pinus sylvestris* L.) from the Alps to the Apennines. Flora 201:468-476.

Lemos M, Dias A, Carvalho A, Delgado A, Pavia I, Coutinho J, Gaspar MJ, Lima-Brito J (2015) Single protocol for genomic DNA extraction from foliar, dry and vascular tissues of gymnosperms suitable for genetic studies. Wood Science and Technology (*submitted*).

Lewontin RC (1972) Testing the theory of natural selection. Nature 236:181-182.

Liber Z, Pavletió Z, Nikolió T, Fülgosi H (1999) DNA polymorphisms in Austrian and Dalmatian black pine. Wiss Mitt Niederösterr. Landesmuseum 12:53-62.

Liber Z, Nikolić T, Mitić B, Šatović Z (2003) RAPD markers and Black pine (*Pinus nigra* Arnold) intraspecies taxonomy- Evidence from the study of nine populations. Acta Societatis Botanicorum Poloniae 72:249-257.

Li HY, Jiang J, Liu GF, Ma XJ, Dong JX, Lin SJ (2005) Genetic variation and division of *Pinus sylvestris* provenances by ISSR markers. Journal of Forest Research 16:216-218.

Lima-Brito J, Carvalho A, Martín A, Heslop-Harrison JS, Guedes-Pinto H (2006) Morphological, yield, cytological and molecular characterisation of a bread wheat x tritordeum F_1 hybrid. Journal of Genetics 85:123-131.

Little EL, Critchfield WB (1969) Subdivisions of the Genus Pinus (pines). U.S. Forest Services, Miscellaneous Publication No. 1144, U.S. Department of Agriculture, Washington, D.C., Forest Service, December 1969. 51 pp.

López-Aljorna A, Bueno MÃ, Aguinagalde I, Martín JP (2007) Fingerprinting and genetic variability in cork oak (*Quercus suber* L.) elite trees using ISSR and SSR markers. Annals of Forestry Science 64:773-779.

Louro V (1982) O pinheiro larício (*Pinus nigra* Arnold) em Portugal. Divulgação Técnica, MACP, Direcção Geral das Florestas Nº 2 – Agosto 1982. 110 pp.

Lucas-Borja ME, Fonseca TF, Lousada JL, Silva-Santos P, Garcia EM, Abellá MA (2012) Natural regeneration of Spanish black pine [*Pinus nigra* Arn. ssp. *salzmannii* (Dunal) Franco] at contrasting altitudes in a Mediterranean mountain area. Ecological Research 27:913–921.

Lučić A, Mladenović-Drinić S, Stavretonić N, Isajev V, Lavadinivić V, Rakonkac L, Novaković (2010) Genetic diversity of Austrian Pine (*Pinus nigra* Arnold) populations in Serbia revealed by RAPD. Archives Biological Sciences 62:329-336.

Luo C, He X-H, Chen H, Ou S-J, Gao M-P, Brown JS, Tondo CT, Schnell RJ (2011) Genetic diversity of mango cultivars estimated using SCoT and ISSR markers. Biochemical Systematics and Ecology 39:676–684.

Luo C, He X-H, Chen H, Ou S-J, Gao M-P (2010) Analysis of diversity and relationships among mango cultivars using Start Codon Targeted (SCoT) markers. Biochemical Systematics and Ecology 38:1176–1184.
Mahajan R, Gupta P (2012) Molecular markers: their use in tree improvement. Journal of Forest Science 58(3):137-144.
Martín-Benito D, Gea-Izquierdo G, Río M, Cañellas I (2008) Long-Term trends in dominant-height growth of black pine using dynamic models. Forest Ecology and Management 256:1230-1238.
Mirov NT (1967) The genus *Pinus*. Ronald Press, New York. 602 pp.
Morgante M, Olivieri AM (1993) PCR-amplified microsatellites as markers in plant genetics. Plant Journal 3:175-182.
Muller-Starck, Baradat Ph, Bergmann F (1992) Genetic variation within European tree species. New Forests 6:23-47.
Mulpuri S, Muddanuru T, Francis G (2013) Start codon targeted (SCoT) polymorphism in toxic and non-toxic accessions of *Jatropha curcas* L. and development of a codominant SCAR marker. Plant Science 207:117-127.
Murray BG (1998) Nuclear DNA amounts in gymnosperms. Annals of Botany 82:3-15.
Naik D, Singh D, Vartak V, Paranjpe S, Bhargava S (2009) Assessment of morphological and genetic diversity in *Gmelina arborea* Roxb. New Forests 38:99-115.
Naydenov K, Tremblay FM, Fenton NJ, Alexandrov A (2006) Structure of *Pinus nigra* Arn. populations in Bulgaria revealed by chloroplast microsatellites and terpenes analysis: Provenance tests. Biochemical Systematics and Ecology 34:562-574.
Nei M (1973) Analysis of gene diversity in subdivided populations. Proceedings of the National Academy of Sciences 70:3321-3323.
Nei M (1978) Estimation of average heterozygosity and genetic distance from a small number of individuals. Genetics 89:583-590.
Nei M (1987) Molecular evolutionary genetics, pp. 176–187, Columbia University Press, New York, USA.
Pajares PR, Escudero A (1989) Caracterización fitoecológica de las comunidades de *Pinus nigra* subsp. *salzmannii* en los afloramientos rochosos del sistema ibérico meridional. Botanica Complutensis 15:149-161.
Page RD (1996) TreeView: An application to display phylogenetic trees on personal computers. Computer Applications in the Biosciences 12: 357–358.
Pavy N, Pelgas B, Laroche J, Rigault P, Isabel N, Bousquet J (2012) A spruce gene map infers ancient plant genome reshuffling and subsequent slow evolution in the gymnosperm lineage leading to extant conifers. BMC Biology 10:84.
Peakall R, Smouse P (2012) GenAlEx 6.5: Genetic analysis in Excel. Population genetic software for teaching and research – an update. Bioinformatics 28:2537-2539.
Poczai P, Varga I, Laos M, Cseh A, Bell N, Valkonen JPT, Hyvönen J (2013) Advances in plant gene-targeted and functional markers: a review. Plant Methods 9:6.
Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. Genetics 155:945-959.
Raffi ZA, Dodd RS, Zavarin E (1996) Genetic Diversity in Foliar Terpenoids Among Natural Populations of European Black Pine. Biochemical Systematics and Ecology 24:325-339.
Richardson DM (1998) Ecology and biogeography of Pinus. Cambridge University Press. London.
Rohlf FJ (1998) NTSYS-pc ver. 2.02. Numerical taxonomy and multivariate analysis system. Exeter Publishing, Setauket, USA.

Rubiales J, García-Amorena I, Hérnandez L, Génova M, Martínez F, Manzaneque F, Morla C (2010) Late Quaternary dynamics of pinewoods in the Iberian Mountains. Review of Palaeobotany and Palynology 162:476-491.

Rubio-Moraga A, Candel-Perez D, Lucas-Borja ME, Tiscar PA, Viñegla, Linares JC, Gómez-Gómez L, Ahrazem O (2012) Genetic diversity of Pinus nigra Arn. populations in Southern Spain and Northern Morocco revealed by inter-simple sequence repeat profiles. International Journal of Molecular Sciences 13:5645-5658.

Saitou N, Nei M (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. Molecular Biology and Evolution 4:406-425.

Šarac Z, Matejić JS, Stojanović-Radić ZZ, Veselinović JB, Džamić AM, Bojović S, Marin PD (2014) Biological activity of *Pinus nigra* terpenes – Evaluation of FtsZ inhibition by selected compounds as contribution to their antimicrobial activity. Computers in Biology and Medicine 54:72-78.

Savolainen O, Pyhäjärvi T, Knürr T (2007) Gene flow and local adaptation in forest trees. Annual Review Ecology Evolution Systematics 37:595–619.

Scaltsoyiannes A, Rohr R, Panetos KP, Tsaktsira M (1994) Allozyme frequency distributions in five European populations of black pine (*Pinus nigra* Arnold). Silvae Genetica 43:20-30.

Schwarz O (1938) Über die Systematik und Nomenklatur der europäischen Schwarzkiefern. Notizblatt des Königlichen Botanischen Gartens und Museums zu Berlin-Dahlem. 17:226-243.

Shannon C, Weaver W (1949) The mathematical theory of communication. University of Illinois Press, Urbana, USA.

Steinkellner H, Lexer C, Turetscheck E, Glössl J (1997) Conservation of (GA)n microsatellite *loci* between *Quercus* species. Molecular Ecology 6:1189-1194.

Svoboda P (1953) Lesni drevini a jeijch porosity, I. Praha.

Taylor DR, Aarssen LW (1990) Complex competitive relationships among genotypes of three Perennial grasses: Implications for species coexistence. The American Naturalist 136:305–327.

Thompson JN (2005) The geographic mosaic of coevolution. University of Chicago Press, Chicago, Illinois, USA.

Vidaković M (1974) Genetics of European black pine (*Pinus nigra* Am.). Annales Forestales 6:57-86.

Villar EH (1947) Quel est le nom valable de Pinus larício Poir. Bericht der Schweiz. Botan Gesellsch, LVII 152.

Wang MB, Hao ZZ (2010) Range wide genetic diversity in natural populations of Chinese pine (*Pinus tabulaeformis*). Biochemical Genetics 48:590–602.

Wang Z, Weber JL, Zhong G, Tanksley SD (1994) Survey of plant short tandem DNA repeats. Theoretical and Applied Genetics 88:1–6.

Yang CP, Wei L, Jiang J, Liu GF, Zhao GY (2005) Analysis of genetic diversity for nineteen populations of Pinus sibirica Du Tour with technique of ISSR. Journal of Northeast Forestry University 33:1–3.

Yeh FC, Yang RC, Boyle TBJ, Ye ZH, Mao JX (1999) Popgene version 1.32, the user-friendly shareware for population genetic analysis. Molecular Biology and Biotechnology Centre, University of Alberta, Edmonton, USA.

Zaghi D (2008) Management of Natura 2000 habitats. 9530 (Sub)-Mediterranean pine forests with endemic black pines. Technical Report 2008 24/24. Directive 92/43/EEC on the conservation of natural habitats and of wild fauna and flora.

Zhang J, Xie W, Wang Y, Zhao X (2015) Genetic diversity and relationships among Chinese *Elymus sibiricus* accessions using Start Codon Targeted (SCoT) markers. Molecules 20:1-13.

IV.7 – Supplemental material

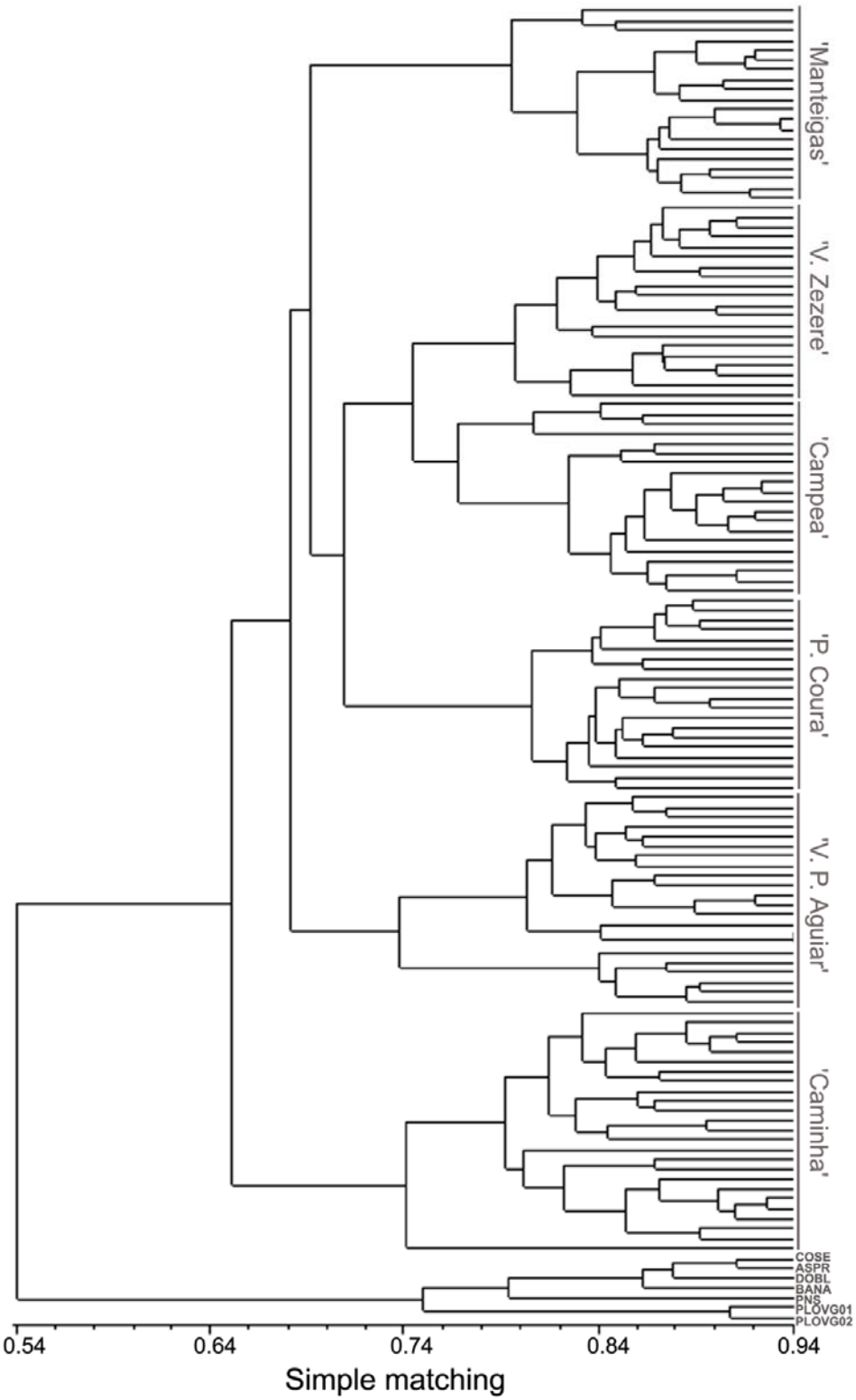

Figure IV.S1 – UPGMA dendrogram of genetic similarity among the Portuguese and foreign *P. nigra* individuals based on the pool of the ISSR and SCoT data. COSE and ASPR - *P. nigra* subsp. *laricio* var. *calabrica*; DOBL and BANA - *P. nigra* subsp. *nigra* var. *austriaca*; PNS - *P. nigra* subsp. *salzmannii*; PLOVG01 and PLOVG02 - *P. nigra* subsp. *laricio* var. *corsicana*.

Chapter V:

Nucleolar activity, SSRs hybridization and chromosome irregularities in peripheral populations of *Pinus nigra* Arn. and *Pinus sylvestris* L.

Lemos M, Carvalho A, Pavia I, Delgado A, Gaspar MJ, Lima-Brito J (2015) Nucleolar activity, SSRs hybridization and chromosome irregularities in peripheral populations of *Pinus nigra* Arn. and *Pinus sylvestris* L. Paper Submitted to *Plant Cell Reports*.

V.1 – Abstract

Portuguese populations of *Pinus nigra* and *Pinus sylvestris* are considered peripheral. These stands were installed in 20th century, but autochthonous *P. sylvestris* populations were ascribed to 'Serra do Gerês'. Only one *P. sylvestris* autochthonous population was cytogenetically characterized so far. We intended to characterize cytogenetically individuals from two Portuguese populations, one of *P. nigra* and other of *P. sylvestris*. We used nondenaturing-fluorescence in situ hybridization (ND-FISH) and FISH with simple sequence repeats (SSRs) and one 45S rDNA probes, respectively, to establish comparisons and detect potential chromosome irregularities in both species. We used silver nitrate staining to determine the number of nucleoli/interphase nucleus and compared with the rDNA *loci* number. A different number of SSR hybridization signals were detected between the species. After FISH, 14 and 16 rDNA *loci* were observed in *P. sylvestris* and *P. nigra*, respectively. A variable number of nucleoli/interphase nucleus (1 to 12) in both species, suggested that not all rDNA *loci* were transcriptionally active. A variable number of micronucleoli were also detected in both species and might be related with the observed irregularities (e.g. ring, polycentric and fragmented chromosomes). Such instabilities are common in Pinaceae peripheral populations growing in adverse environmental conditions, and should be regarded as microevolution elements that increase genetic diversity and adaptive ability.

Keywords: European Black Pine; Scots pine; ND-FISH; nucleoli; FISH; salt-nylon silver nitrate staining.

V.2 – Introduction

Pinus nigra Arnold (European Black Pine) is one of the most important pine species in Europe at the economic and ecological levels. This species belongs to the group of Mediterranean pines and has a fragmented distribution through the Northern Africa, Northern Mediterranean and eastwards to the Black Sea, Corsica and Sicily islands (Afzal-Raffi and Dodd 2007). European Black Pine populations experienced different glacial and interglacial episodes of climatic changes and geological events (Thompson 2005) that contributed for their scattered distribution (Isajev *et al*. 2004), and to high variation at the morphological, physiological, biochemical, genetic, ecological and heterochromatin levels (Scaltsoyiannes *et al.* 1994; Raffi *et al.* 1996; Bogunić *et al.* 2003; 2011a; Afzal-Raffi and Dodd 2007; del Cerro Barja *et al.* 2009; Rubio-Moraga *et al*. 2012). This variability has been responsible for some issues in terms of *P. nigra* identification and taxonomic classification (Scaltsoyiannes *et al*. 1994; Afzal-Raffi and Dodd 2007; Bogunić *et al.* 2007; del Cerro Barja *et al*. 2009; Akkemik *et al.* 2010). The most widely recognized infraspecific classification was that proposed by Gaussen *et al.* (1964; 1993). Molecular, cytogenetic and biochemical tools have been used to explore the intraspecific variation of *P. nigra* (Scaltsoyiannes *et al*. 1994; Bogunić *et al*. 2003; 2007; 2011a; Afzal-Rafii-Dodd 2007; Rubio-Moraga *et al*. 2012). However, much remain to be studied in these areas; in order to identify or discriminate the allopatric *P. nigra* subspecies and varieties.

The actual Portuguese *P. nigra* populations present reduced size and are restricted to the North and Centre of the country. Our research team recently started their molecular characterization in order to evaluate their genetic diversity and to extrapolate their provenances (Lemos *et al.* 2015, unpublished). Based on morphological and anatomic data, Louro (1982) proposed the existence of three subspecies in Portugal, namely, the subspecies *laricio*, *nigra* and *salzmannii*. Besides, the Portuguese *P. nigra* populations were never cytogenetically characterized.

Pinus sylvestris L. (Scots pine) is widely distributed through Eurasia and Portugal constitutes its most Southwestern end limit of distribution. This pine species is ecologically and economically important as timber source (Manson and Alia 2000). In contrast to *P. nigra* which was planted in Portugal, autochthonous populations of *P. sylvestris* were found at the NW of Portugal at 'Serra do Gerês'. Both autochthonous and nine planted Portuguese Scots

pine populations were recently characterized at the molecular level by our research group (Cipriano *et al*. 2013; Pavia *et al*. 2014a).

The Portuguese populations of *P. nigra* and *P. sylvestris* could be considered peripheral regarding their natural distribution area. Such populations are susceptible to marginal environments and could present phenotypic variations as well as cytogenetic instabilities, as previously reported for Scots pine (Muratova 1994; Sedel'nikova and Pimenov 2010; Pavia *et al*. 2014b). Several publications have reported the occurrence of structural rearrangements, presence of a giant chromosome, haploidy, aneuploidy, polyploidy and mixoploidy in *Pinaceae* species (revised by Muratova 1994; Sedel'nikova and Pimenov 2010). These cytogenetic instabilities have been detected in populations growing under extreme conditions, i.e., near and outside the borders of their distribution area or in response to adverse environmental factors or pollution (Sedel'nikova and Pimenov 2010).

The species from family *Pinaceae* present a stable karyotype with a constant chromosome number and morphology. *Pinus* species are diploid and present 24 chromosomes (10 metacentric and two smaller submetacentric chromosome pairs) (Hizume 1988; Hizume *et al*. 2002). The large metacentric chromosomes are hard to distinguish due to their similar size and morphology (Saylor 1964; 1972; 1983; Borzan 1977; Hizume 1988; Cai *et al.* 2006). Nonetheless, efforts have been made over the years for the cytogenetic characterization of different pine species, including *P. nigra* and *P. sylvestris*, regarding the: determination of number and location of small constrictions in the haploid complement (Pederick 1997); karyotyping or chromosomal identification by C-banding (Borzan and Papes 1978; Machperson and Filion 1981); G-banding (Drewry 1982); aceto-carmine staining (Kaya *et al.* 1984); fluorescence banding techniques based on the use of chromomycin A3 (CMA) and 4',6-diamidino-2-phenylindole (DAPI) (Hizume *et al.* 1983; 1989; 1990; Doudrick *et al*. 1995); or FISH (Hizume *et al*. 2002; Bogunić *et al*. 2006; 2011a,b; Pavia *et al*. 2014b). Sometimes the lack of resolution in the karyotypes constructed based on fluorescence banding techniques unable the identification of homologous chromosomes and comparison studies among species (Hizume *et al*. 2002), limiting the investigations of the genome organization and structural rearrangements in pines (Cai *et al.* 2006). Fluorochrome banding and silver nitrate staining were also sequentially used for the study of the patterns of heterochromatin regions distribution (AT-, GC-rich and DAPI-positive heterochromatin) and nucleolar activity, respectively, in the two tertiary relict pines, *Pinus heldreichii* and *P. nigra* (Bogunić *et al.* 2006). The silver nitrate staining technique was also used to study the nucleolar

structure along the cellular cycle of Scots pine (Kupila-Ahvenniemi and Hohtola 1979). Other classic cytogenetic techniques than silver nitrate staining have been used in studies of nucleolar activity, mitotic and meiotic disturbances in *P. sylvestris* growing under unfavourable environmental conditions (Butorina *et al*. 1999; Muratova and Sedel'Nikova 2000; Sedel'nikova and Pimenov 2010).

The use of fluorescence *in situ* hybridization (FISH) with ribosomal DNA (rDNA) probes proved to be an excellent tool for the physical mapping of the 18S-25 rDNA *loci* in different pine species (Hizume *et al.* 1992; 2002; Lubaretz *et al.* 1996; Jacobs *et al*. 2000; Liu *et al.* 2003; Cai *et al.* 2006; Pavia *et al*. 2014b).

The nondenaturing-FISH (ND-FISH) technique constitutes a variant of the conventional FISH procedure that does not involve the denaturing of chromosomal DNA (Cuadrado and Jouve 2010; 2011). This technique proved to be suitable for the hybridization of simple sequence repeats (SSRs) probes in different species (Cuadrado and Jouve 2010; 2011; Cabo *et al.* 2014; Pavia *et al.* 2014b). Recently, the use of ND-FISH with SSRs, as probes allowed the karyotyping and detection of chromosome irregularities in individuals from the autochthonous Portuguese population 'Ribeira das Negras' ('Serra do Gerês') of *P. sylvestris* (Pavia *et al.* 2014b). The use of SSRs as FISH probes have been widely used in plants, enabling chromosomes identification and the study of genome organization and evolution (Cuadrado *et al.* 2008a,b; Carvalho *et al.* 2013; Cabo *et al.* 2014; Pavia *et al.* 2014b).

As far as we know, except for the study of Kupila-Ahvenniemi and Hohtola (1979), this work constitutes the first analysis of nucleolar activity in both *P. nigra* and *P. sylvestris* using the salt-nylon silver staining technique, as well as the first use of ND-FISH performed with SSR probes in *P. nigra,* and FISH with the 45S rDNA probe in these two Portuguese populations. Another novelty of this study relies on the comparison between the number of active and inactive rDNA *loci* in these two pine species. Due so, with this study is intended to: i) characterize *P. nigra* and *P. sylvestris* individuals from the Portuguese populations of 'Campeã' and 'Biduiça', respectively, by physically mapping eight SSR and the 45S rDNA (pTa71) probes in mitotic chromosome spreads using ND-FISH and conventional FISH techniques, respectively; ii) to establish comparisons of hybridization patterns between the two species; iii) to analyze the number of nucleoli per interphase nucleus assayed by the salt-nylon silver staining technique, and to compare it with the number of rDNA *loci* detected by

FISH in each species; and iv) to search for possible cytogenetic instabilities in these peripheral populations.

V.3 – Material and methods

V.3.1 – Seeds germination, collection and fixation of root tips

Regarding the distribution area of *P. nigra* and *P. sylvestris*, Portugal constitutes their Southwestern end limit of distribution. Thus, the Portuguese populations of *P. nigra* and *P. sylvestris* under study could be considered peripheral.

In this study, we used seeds of *P. nigra* and *P. sylvestris* that were collected in the Portuguese populations of 'Campeã' (Vila Real, North of Portugal) and 'Biduiça' ('Serra do Gerês', NW Portugal), respectively.

The seeds were allowed to germinate in the dark at 25 ºC in Petri dishes containing filter paper moistened in distilled water. After two weeks, root tips with 1.5 cm length were harvested and treated in ice-cold water (0 ºC) for 24 h. The ice treated root tips were fixed in a solution of absolute ethanol and glacial acetic acid in the proportion of 3:1 (v/v). The fixed root tips were maintained at -20 ºC till the preparation of chromosome spreads by the squashing procedure.

V.3.2 – ND-FISH and FISH experiments

Eight SSR sequences: $(AC)_{10}$; $(AG)_{10}$; $(AG)_{12}$; $(AAG)_5$; $(AAC)_5$; $(GATA)_4$; $(GACA)_4$ and $(GGAT)_4$, were labeled with biotin-16-dUTP or digoxigenin-11-dUTP by the Random Primed Labeling Kit (Roche Applied Science, Mannheim, Germany) according to manufacturer's instructions, and used as probes in ND-FISH experiments performed on mitotic chromosome spreads of both *P. nigra* and *P. sylvestris*. ND-FISH experiments followed Pavia *et al.* (2014b) protocol. The hybridization mixture was only composed by the 2x SSC buffer (0.3M NaCl, 0.03 M sodium citrate) and 2 pmol/μL of SSR labeled probe. The post-hybridization washes followed the protocol described by Cuadrado and Jouve (2010). The slides were hybridized in the modified thermal block Omnislide (Hybaid) during 2 h at 24 ºC. After hybridization, the slides were washed twice for 10 min in the detection buffer 4xT (4x SSC, 0.05% Tween 20). To avoid the unspecific binding of the antibody, the chromosome spreads were incubated in 5% BSA (bovine serum albumin) diluted in 4xT for

10 min at room temperature (RT). Further washes in the 4xT detection buffer during 15 min at RT were performed. The hybridization signals were detected by the use of the antibodies fluoresceinavidin D (Vector Laboratories, Peterborough, UK) or anti-digoxigenin-rhodamine (Roche Applied Science, Mannheim, Germany), both diluted in 4xSSC, for probes labeled with biotin or digoxigenin, respectively. The slides were incubated with the antibodies for 1 h at 37 ºC in the modified thermal block Omnislide (Hybaid). After three washes in 4xT for 15 min at RT, the chromosomes were counterstained using the Vectashield Mounting Medium with DAPI (Vector Laboratories, Peterborough, UK).

The slides were observed in the epifluorescence microscope BX41 (Olympus, Hamburg, Germany). The images were captured by a color digital camera XC10 (Olympus, Hamburg, Germany) using the CellSens Entry v1.7.1. software (Olympus, Hamburg, Germany), and captured after double exposure with appropriate filters. The images were prepared for printing using the Adobe_Photoshop ver. 6.0 (Adobe Systems Incorporated, USA; http://www.adobe.com).

The conventional FISH experiments followed the Schwarzacher and Heslop-Harrison (2000) protocol and were performed with the 45S rDNA sequence pTa71 probe (Gerlach and Bedbrook 1979) labeled with digoxigenin-11-dUTP using the Nick Translation Kit (Roche Applied Science, Mannheim, Germany) according to manufacturer's instructions. The hybridization mixture (final volume of 40 μL) contained 50% of formamide; 10% of dextran sulphate; 2x SSC; 1 μL of salmon sperm DNA; 0.125% of SDS; 0.125% of EDTA and 7.5 ng/μL of pTa71 probe. Due to the high molecular weight (around 12 kb) of this cloned probe, the hybridization was performed overnight at 37◦C in the modified thermal block Omnislide (Hybaid). The post-hybridization washes, detection and visualization of hybridization signals were performed as described above for the ND-FISH experiments.

V.3.3 – Salt-nylon silver staining technique

The mitotic chromosome spreads of both species were aged for 2 h at 60 ºC on a dry heat oven. The salt-nylon silver staining technique followed the protocol of Stack *et al*. (1991) with the some modifications. The slides were incubated in 2x SSC at 55ºC during 5 min; thoroughly washed in distilled water for three to four times and air dried. For silver staining we used 100 μL of an aqueous solution of 100% (p/v) silver nitrate (AgNO3) that were dropped on the chromosome spread and covered by a nylon piece. The slides were incubated at 60 ºC for 30 minutes and then thoroughly rinsed in distilled water for several times. After

air dried, the preparations were mounted with mounting medium VectaShield (Vector Laboratories, Peterborough, UK) and a glass coverslip of 24x50mm. The number of nucleoli per interphase nucleus was scored using an optical microscope.

V.4 – Results

V.4.1 – SSR hybridization patterns

ND-FISH experiments were performed in mitotic chromosome spreads of both *P. nigra* and *P. sylvestris* individuals using eight SSR sequences as probes. The SSR $(AC)_{10}$ did not hybridize in both pine species investigated.

In the mitotic chromosome spreads of *P. nigra*, only the SSRs $(GACA)_4$ and $(AG)_{10}$ successfully hybridized (Table V.1; Fig. V.1). The number of hybridization signals of these two SSR probes was higher in *P. sylvestris* individuals (Table V.1).

Table V.1 – Number of hybridization signals of each SSR probe that successfully hybridized in mitotic metaphase cells of *P. nigra* and/or *P. sylvestris* individuals after ND-FISH.

SSR probe	*P. nigra*	*P. sylvestris*
$(AAC)_5$	-	4
$(AAG)_5$	-	4
$(AG)_{12}$	-	6
$(AG)_{10}$	12	34
$(GACA)_4$	2	4
$(GATA)_4$	-	2
$(GGAT)_4$	-	4

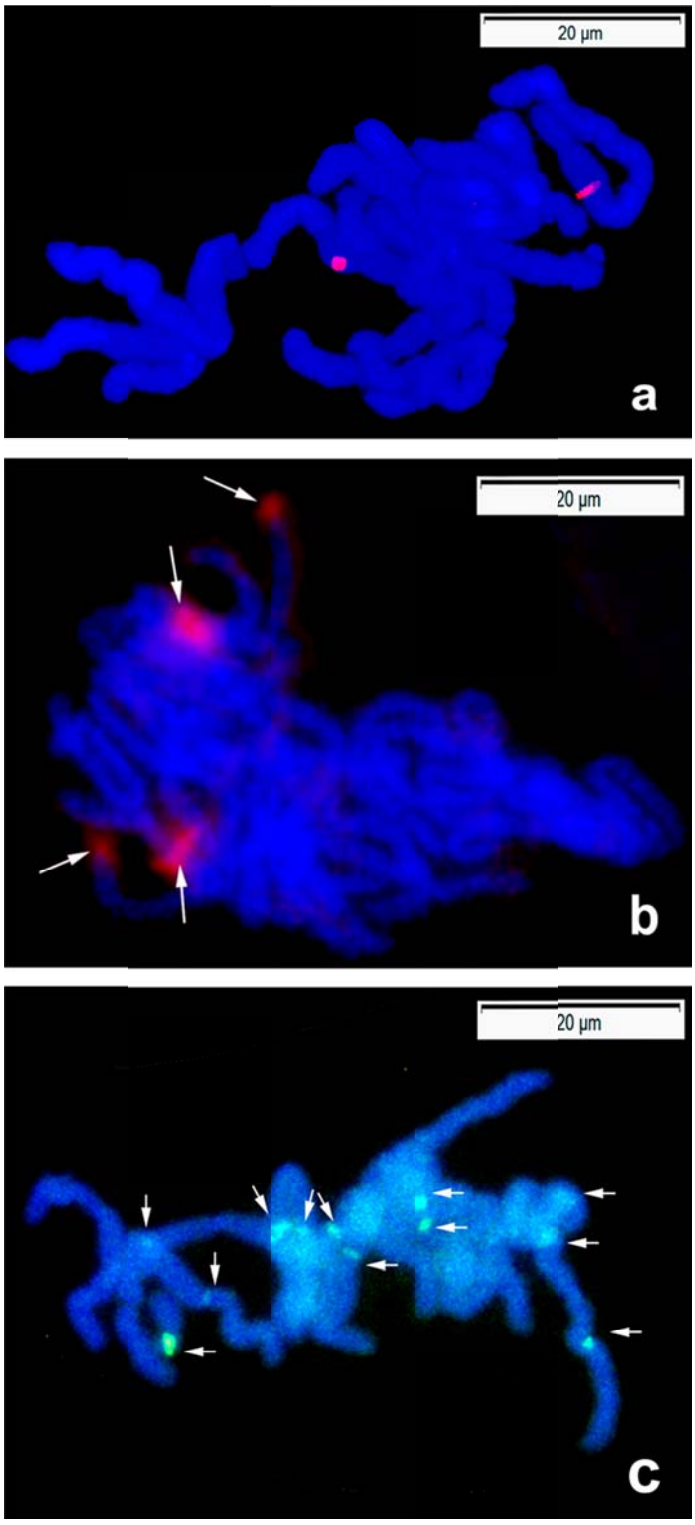

Figure V.1 – Mitotic chromosomes of *P. nigra* (a and c) and *P. sylvestris* (b) after ND-FISH performed with the SSR probes: (a and b) $(GACA)_4$ (red); and (c) $(AG)_{10}$ (green).

In *P. sylvestris* individuals from the 'Biduiça' population, seven of the eight tested SSR probes, namely, $(AG)_{10}$; $(AG)_{12}$; $(AAG)_5$; $(AAC)_5$; $(GATA)_4$; $(GACA)_4$ and $(GGAT)_4$, successfully hybridized (Table V.1; Figure V.1).

V.4.2 – rDNA physical localization and number of nucleoli per interphase nuclei

The mitotic chromosome spreads of *P. nigra and P. sylvestris* were hybridized with the 45S rDNA probe pTa71 using the conventional FISH procedure (two days protocol) due to the large size of this cloned sequence.

Fourteen and sixteen 45S rDNA *loci* per prometaphase cell were detected with the pTa71 probe in *P. sylvestris* and *P. nigra*, respectively (Fig. V.2).

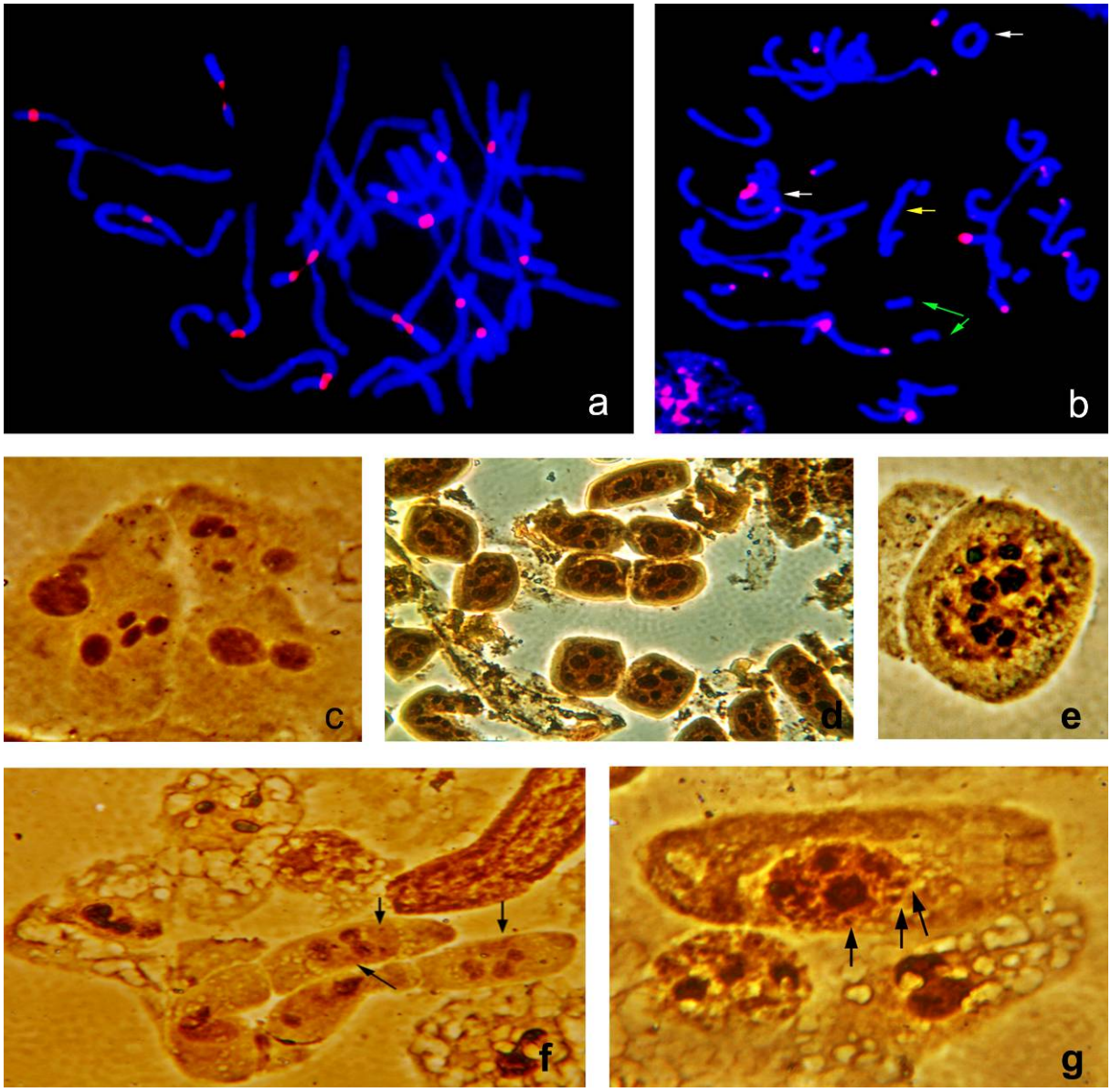

Figure V.2 – Mitotic cells of *P. sylvestris* (a, c, f) and *P. nigra* (b, d, e, g) after FISH (a, b) and silver staining (c, d, e, f, g): (a) the pTa71 probe (red) detected 14 rDNA *loci* in *P. sylvestris* and (b) 16 rDNA *loci* in *P. nigra*, in this cell is also observed two ring chromosomes (white arrows), a polycentric chromosome (yellow arrow) and various chromosome fragments (green arrows), some of them with rDNA *loci*; (c, d) variable number of nucleoli per interphase nucleus; (d) interphase nucleus showing nucleoli with unusual forms; and (f, g) presence of micronucleoli (black arrows) within interphase nuclei of both pine species.

The rDNA *loci* of *P. sylvestris* were detected in secondary constrictions located in subtelomeric regions of short or long chromosome arms (Fig. V.2a).

In the *P. nigra* chromosomes, most of the rDNA *loci* were detected in telomeric regions of complete or fragmented chromosomes (Fig. V.2b). This unexpected location of the rDNA sites was due to the occurrence of a high frequency of structural rearrangements

resulting on occurrence of ring chromosomes, polycentric and fragmented chromosomes (Fig. V.2b).

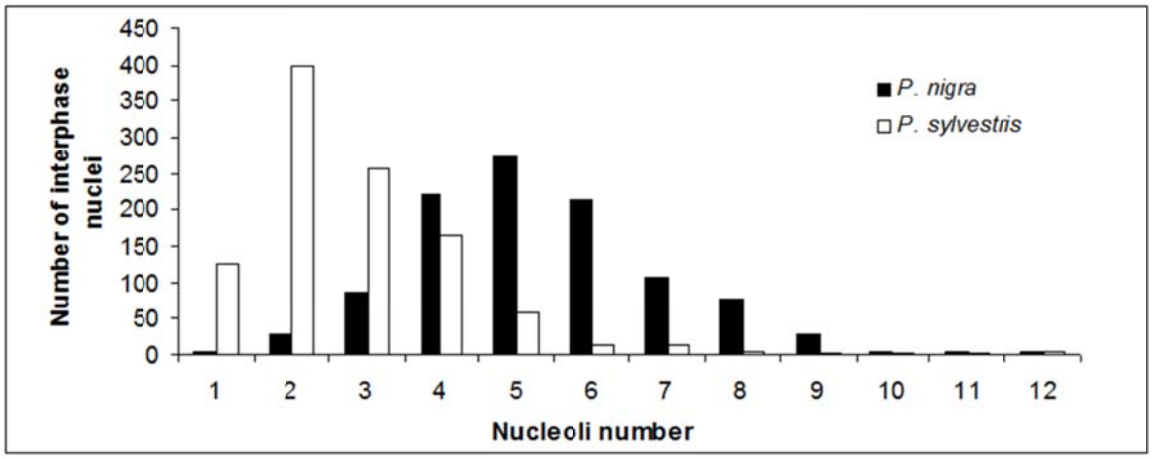

Figure V.3 – Variable number of nucleoli scored per interphase nucleus after silver nitrate staining in *P. nigra* (black bar) and *P. sylvestris* (white bar) individuals.

The nucleoli number per interphase nucleus ranged from 1 to 12 in both *P. nigra* and *P. sylvestris* (Fig. V.2c to g; Fig. V.3). The most frequent number of nucleoli per nucleus was 5 in *P. nigra* and 2 in *P. sylvestris* (Fig. V.3). A variable number of micronucleoli were also observed in the interphase nuclei of *P. sylvestris* and *P. nigra* (Fig. V.2f and V.2g).

Our data suggest that, in both *P. nigra* and *P. sylvestris*, not all rDNA *loci* detected by FISH were transcriptionally active, since we did not observe a coincident maximum number of nucleoli per interphase nucleus after silver staining (Fig. V.3).

V.4.3 – Chromosome irregularities

After DAPI counterstaining inherent to the FISH protocol, some chromosome irregularities such as ring chromosomes; centric and acentric chromosome fragments; polycentric chromosomes and chromatin bridges were observed in some mitotic cells of *P. sylvestris* and in all prometaphase cells of *P. nigra* (Figs. V.2b and V.4). Each prometaphase cell of *P. nigra* presented at least one chromosome irregularity.

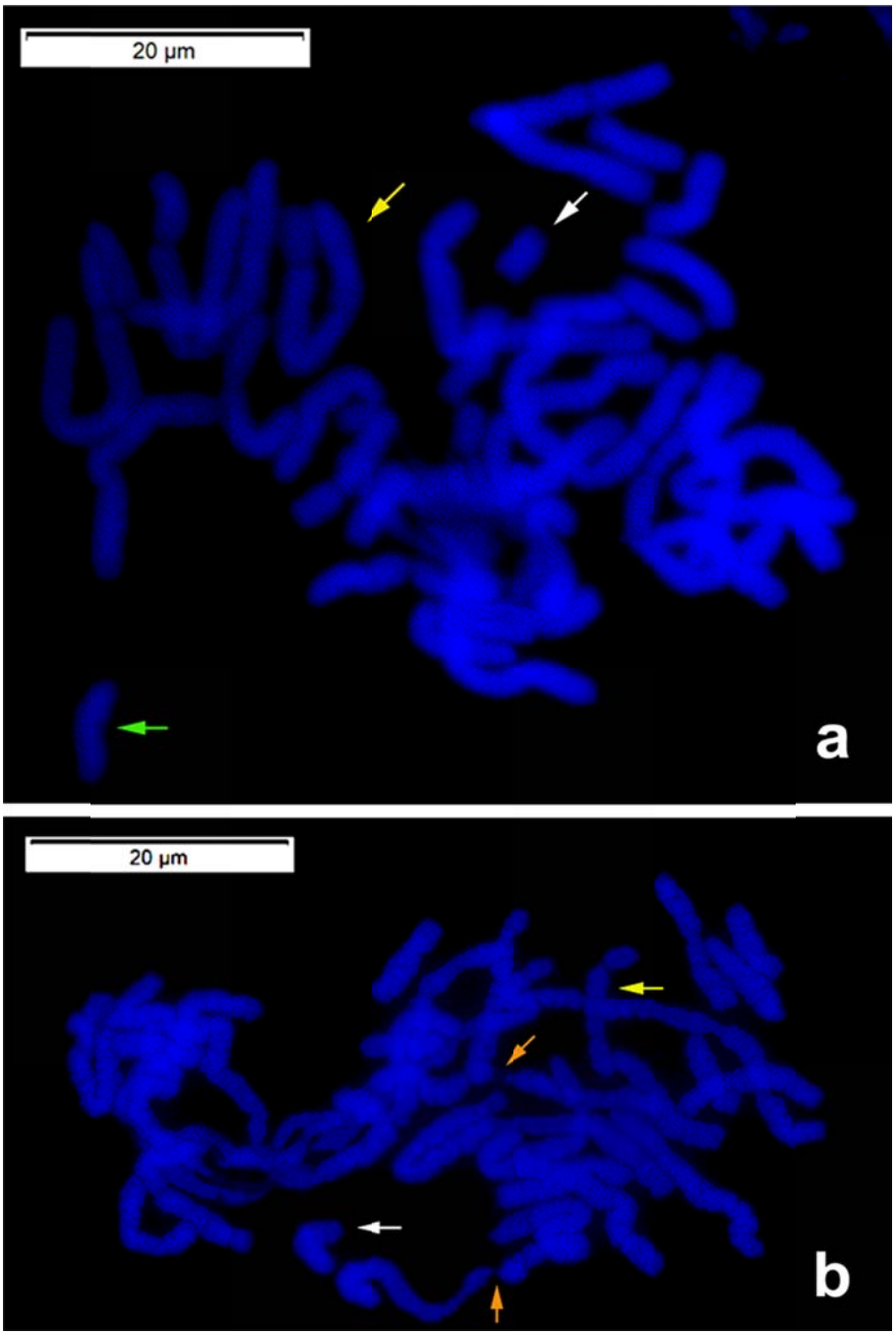

Figure V.4 – Prometaphase cells of *P. nigra* (a) and *P. sylvestris* (b) counterstained with DAPI presenting: (a) an acentric chromosome fragment (white arrow), a centric fragment (green arrow) and a polycentric chromosome (yellow arrow), among others in the same cell; (b) an acentric chromosome fragment (white arrow), an example of a polycentric chromosome (yellow arrow), and two chromatin bridges between different chromosomes (orange arrows).

Nucleoli with unusual forms and the occurrence of micronucleoli can also be considered cytogenetic anomalies and were detected after silver nitrate staining in nuclei of both pine species (Figs. V.2e to V.2g).

V.5 – Discussion

Previous cytogenetic studies were developed in multiple pine species using molecular and classic cytogenetic techniques. However, the ND-FISH technique performed with SSR probes was never been applied in in *P. nigra* chromosomes.

Among the eight SSR probes tested here, only the SSR $(AC)_{10}$ did not hybridize in both pine species investigated. This result is interesting since Schmidt *et al*. (2000) reported the successful hybridization of low complex SSRs with dinucleotide repeat motifs such as $(CA)_8$ and $(GA)_8$ in multiple gymnosperms, including pines.

The number of hybridization signals detected in *P. sylvestris* individuals from the autochthonous population of 'Biduiça' ('Serra do Gerês') with seven out of the eight tested SSR probes, were similar to those previously achieved by Pavia *et al*. (2014b) in another autochthonous of *P. sylvestris* from population ('Ribeira das Negras', located at 'Serra do Gerês'). Along with SSRs based on dinucleotide repeat motifs, other SSRs considered with low complexity, such as $(CAC)_5$, $(GGAT)_4$, $(GACA)_4$ and $(GATA)_4$, successfully hybridized in the following pine species: *Pinus echinata*, *Pinus elliottii*, *Pinus palustris*, *Pinus caribaea*, *Pinus oocarpa*, *Pinus banksiana*, *Pinus massoniana*, *Pinus resinosa* and *Pinus strobus* (Schmidt *et al.* 2000). Comparing with the results of these authors, we also verified hybridization of the SSR probe $(GACA)_4$ in both *P. nigra* and *P.* sylvestris, and of $(GGAT)_4$ in *P. sylvestris*. The SSR $(GACA)_4$ was reported as an infrequent repeat motif in conifer genomes (Schmidt *et al*. 2000), but it seems to be present in the pines studied so far, including *P. nigra* and *P. sylvestris*.

SSRs are an abundant class of repeats in plant genomes. However, is unknown the reason why some sequence motifs are more abundant or more polymorphic than others (Beyermann *et al.* 1992; Poulsen *et al.* 1993; Depeiges *et al.* 1995). Based on Southern blotting results, Schmidt *et al.* (2000) reported that the SSRs $(CAC)_5$ and $(GACA)_4$ are present in only low copy numbers whereas the dinucleotide repeat motifs $(CA)_8$ and $(GA)_8$ represented the major components of the gymnosperm genomes investigated by them. We also achieved successful hybridization with the dinucleotide repeat motifs $(AG)_{10}$ in both species and with $(AG)_{12}$ only in *P. sylvestris*.

Gymnosperms have large sized genomes (Ribeiro *et al.* 2008). In general, the pines genome size range from 22.10 to 36.89pg per 1C (Murray 1998), reaching ca. 23.1Gbp in *P. nigra* (Bogunić *et al.* 2003, 2007). Variation in genome size among plants is directly

correlated with the amount of repetitive DNA, such as transposable elements and tandemly repeated sequences (Ribeiro *et al.* 2008). The successful hybridization of repeats with low sequence complexity seems to indicate that they constitute a large fraction of the repetitive DNA of conifers (Schmidt *et al.* 2000; Ribeiro *et al.* 2008). The differences in terms of number of hybridization signals, abundance and organization of various repeat motifs among conifers could be explained by their different genome sizes and high levels of polymorphism, and suggest that these sequences evolved differently in gymnosperms (Schmidt *et al*. 2000; Ribeiro *et al*. 2008).

Another novelty of this study consisted on the use of salt-nylon silver staining technique for the study of the nucleolar activity in *P. nigra* and in *P. sylvestris*, and further comparison with the number of 45S rDNA *loci* detected by the conventional FISH procedure (two days protocol). The nucleoli of *P. sylvestris* were previously studied in terms of number, form, and dimension but with other techniques and goals (Butorina *et al*. 1999; Muratova and Sedel'Nikova 2000).

We detected 14 rDNA *loci* per mitotic cell in *P. sylvestris*, which agrees the previously results reported by Hizume *et al*. (2002) and Pavia *et al.* (2014). These hybridization signals were located on the secondary constrictions of short or long chromosome arms (Fig. V.2a), corroborating the karyotype previously published by Hizume *et al*. (2002).

In European Black Pine it was not found an unique normal (pro)metaphase cell, since each presented at least one chromosomal anomaly. Contrastingly to *P. sylvestris*, we were unable to confirm the chromosome complement of the studied *P. nigra* individuals, due to the presence of multiple centric and acentric chromosome fragments.

Despite the finding of the expected number of sixteen 45S rDNA *loci* in *P. nigra*, as previously reported by other authors (Hizume *et al*. 2002), some of the rDNA sites were located in the telomeric regions of fragmented chromosomes and also in the telomeres of longer chromosomes. These results allowed us to propose that: i) the telomeric hybridization signals of pTa71 in both longer and fragmented chromosomes might arose from a breakage within some rDNA *loci* with maintenance of hybridization signal in both resultant ends; or ii) the effective number of rDNA *loci* in the European Black Pine individuals studied here is smaller than 16, since karyotypic variability was previously reported among allopatric populations of this species (Bogunić *et al.* 2011a). This last hypothesis could partially explain why the number of nucleoli observed per interphase nucleus after silver nitrate staining was

smaller than that of rDNA *loci* detected after FISH (Fig. V.2 and V.3). However, the same occurred in *P. sylvestris* individuals (Fig.V.2 and V.3).

According to Hizume *et al.* (2001; 2002) and Cai *et al.* (2006), 45S rDNA *loci* could be observed in centromeric regions of some pine species, despite presenting weak hybridization signals. The quality of the chromosome spreads and morphology of *P. sylvestris* and *P. nigra* should be improved to confirm this assumption. Even though, a variation in number and position of rDNA *loci* among gymnosperms was previously reported, such as in *Podocarpus*, *Larix* and *Picea* which presented 45S rDNA *loci* in interstitial positions, near the centromere, or in the terminal regions of long and short chromosome arms (Doudrick *et al.* 1995; Lubaretz *et al.* 1996; Murray *et al.* 2002).

Comparing the number of rDNA *loci* detected after FISH with the observed maximum number of nucleoli per interphase nucleus after silver staining (12 in both *P. nigra* and *P. sylvestris*), we might say that not all rDNA *loci* are transcriptionally active.

Previous studies pointed out the occurrence of 3 to 10 (Muratova 1993), 1 to 11 (Muratova 1994) and 3 to 12 (Muratova 1997) nucleoli per interphase nucleus of *P. sylvestris*, and 1 to 11 nucleoli per interphase nucleus of *P. nigra* (Bogunic *et al*. 2006). The slight differences observed in our study concerning the nucleoli number per interphase nucleus can be explained by the fact that despite most of the conifer being diploid species with a constant number of chromosomes with similar morphology, they could present polymorphism in terms of number of nucleoli (Muratova 1997a,b). These authors pointed out several explanations to the apparent inconsistency between the

number of nucleoli and 45S rDNA *loci*, namely: (i) some chromosome pairs without secondary constrictions may have ability to organize the nucleolus; (ii) several chromosomes can take part in the organization of one nucleolus; and (iii) not all secondary constrictions are related to the nucleolus formation.

The salt-nylon silver staining technique also allowed the detection of micronucleoli in both pine species. The presence of micronucleoli was previously reported in other plant species: *Brassica alboglabra* (Cheng *et al.* 1995); tritordeum (Lima-Brito *et al.* 1998); and *Lolium multiflorum* (Bustamante *et al.* 2014). According to Cheng *et al.* (1995) and Lima-Brito *et al.* (1998), the occurrence of such structures could be related with chromosome pairs bearing *minor* nucleolar organizer regions (NORs). Otherwise, in the case of *L. multiflorum* (Bustamante *et al.* 2014) it was hypothesized that the micronucleoli formation may be associated with rearrangements caused by breaks or gaps in fragile sites such as translocations

(Schubert and Künzel 1990). The presence of micronucleoli may indicate reduced transcriptional activity of some NORs which underwent some type of rearrangement (Bustamante *et al.* 2014).

Chromosome irregularities have already been described in pines species, including in *P. sylvestris* individuals from peripheral populations or under unfavourable environmental conditions (Muratova 1997a,b; Pavia *et al.* 2014b).

Muratova (1997a,b) stated that the ring and polycentric chromosomes were not additions to the genomes, belonging to diploid set of chromosomes. Beyond pines, ring and polycentric chromosomes were also found in other conifers (see Muratova 1997a,b).

Chromosome and chromatic structural changes can led to formation of ring chromosomes due to the deletion of the telomere parts and conjugation of the sticky ends of the middle fragment (Muratova 1997a,b). Acentric fragments should be also originated as a consequence of this process (Muratova 1997a,b). In case of polycentric chromosomes, they can result from asymmetric translocations (Muratova 1997a,b).

Our results allowed us to suppose a probable correlation between the presence of micronucleoli and chromosome instabilities. Some authors previously argued the same and proposed that: (i) the micronucleoli formation may be associated with rearrangements such as translocations caused by breaks or gaps in fragile sites (Schubert and Künzel 1990); and (ii) the presence of micronucleoli may indicate reduced transcriptional activity of some NORs which underwent some type of rearrangement (Bustamante *et al.* 2014). In this way, the present results can support both hypotheses, since we observed the presence of micronucleoli (Fig. V.2), and cytogenetic irregularities derived from structural rearrangements in *P. nigra* and *P. sylvestris* (Fig. V.4).

According to Pavia *et al.* (2014a) peripheral *P. sylvestris* populations have high intraspecific genetic variability as revealed by chloroplastidial and nuclear SSR markers. Besides, cytogenetic instabilities could contribute to a reduced rate of concerted evolution emphasized by the high number of rDNA *loci* (Pavia *et al.* 2014b). Such irregularities were also found in the autochthonous *P. sylvestris* population of 'Biduiça', another autochthonous Portuguese population from 'Serra do Gerês', but in higher frequency in the *P. nigra* individuals. These chromosome anomalies may reflect that these populations have been exposed to different selection regimes due to their peripheral location relatively to their natural distribution areas. Thus, the cytogenetic results achieved so far by our group confirmed Portugal as an edge limit of the distribution area of these two pine species.

Despite the worldwide economical and ecological importance of forest trees, the progress in the understanding of their genomics and cytogenetics did not progress at the same rhythm of that of herbaceous species. According to Bogunić *et al*. (2011b), molecular cytogenetic studied based on FISH were employed only in 23 of 111 pine species, and this number should not have increased in the last four years.

V.6 – Conclusions

As far as we know, this study constitutes the first molecular cytogenetics characterization of the Portuguese *P. nigra* population of 'Campeã' and of the autochthonous *P. sylvestris* population of 'Biduiça' ('Serra do Gerês'). Despite the previous development of nucleolar activity studies performed in both species by other authors, here we present for the first time a simple and cost-effective silver staining protocol that was adapted from the Stack *et al*. (1991) method. We are confident that this technique could be extended to other forest trees.

Our study revealed that nucleolar activity results and their comparison with FISH might enable indications about the adaptive potential of peripheral populations to the surrounding environment and how the nucleolar activity could be involved in the adaptation process, even in geographic locations where the conditions seem to be favorable. Nonetheless, efforts will be made for the optimization of the cytogenetic protocols to be applied in pine for the detection of Ag-NORs, for the sequential use of FISH with 45S rDNA probes and for the preservation of morphology and spread of these large chromosomes that have been preventing the accurate identification of individual chromosomes. The improved morphology and identification of individual chromosomes will contribute for the detection of cytogenetic irregularities or for the discovering of new mutation types in individuals from peripheral populations.

V.7 – Acknowledgements

Ana Carvalho thank the postdoctoral SFRH/BPD/68932/2010 attributed by the Portuguese Foundation for Science and Technology (Fundação para a Ciência e a Tecnologia, FCT), co-financed by the Social European Fund (FSE) under the POPH-QREN program.

The authors acknowledge to the Eng. Goreti Pinto from Centro Nacional de Sementes Florestais (CENASEF, Portugal) for kindly provided certified seed material of *P. nigra*, and to Prof. João Bento and Eng. Luís Roxo for the harvesting of seeds in *P. sylvestris* individuals at 'Biduiça' ('Serra do Gerês').

V.8 – References

Afzal-Rafii Z, Dodd RS (2007) Chloroplast DNA supports a hypothesis of glacial refugia over postglacial recolonization in disjunct populations of black pine (*Pinus nigra*) in Western Europe. Molecular Ecology 16:723-736.

Beyermann B, Nürnberg P, Weihe A, Meixner M, Epplen JT, Börner T (1992) Fingerprinting plant genomes with oligonucleotide probes specific for simple repetitive DNA sequences. Theoretical and Applied Genetics 83:691-694.

Bogunić F, Muratovi E, Ballian D, Siljak-Yakolev S, Brown S (2007) Genome size stability of five subspecies of *Pinus nigra* Arnold s.l. Environmental and Experimental Botany 59:354-360.

Bogunić F, Muratovic E, Brown SC, Siljak-Yakovlev S (2003) Genome size of five *Pinus* from Balkan region. Plant Cell Reports 22:59–63.

Bogunić F, Muratovic E, Brown SC, Siljak-Yakovlev S (2006) Chromosomal differentiation between *Pinus heldreichii* and Pinus nigra. Annals of Forest Science 63:267-274.

Bogunić F, Siljak-Yakolev S, Muratovi E, Ballian D (2011a) Diferent karyotype patterns among allopatric *Pinus nigra* (Pinaceae) populations revealed by molecular cytogenetics. Plant Biology 13:194-200.

Bogunić F, Siljak-Yakolev S, Muratovi E, Pustahija F, Medjedović S (2011b) Molecular cytogenetics and flow cytometry reveal conserved genome organization in *Pinus mugo* and *P. uncinata*. Annals of Forest Science 68:179-187.

Borzan Z (1977) Contribution to the karyotype analysis of the European Black Pine (*Pinus nigra* Arn.). Annales Forestales 8:29-50.

Borzan Z, Papes D (1978) Karyotype analysis in *Pinus*: contribution to the standardization of the karyotype analysis and review of some applied techniques. Silvae Genetica 27:144-150.

Bustamante FO, Rocha LC, Torres GA, Davide LC, Mittelmann A, Techilo VH (2014) Distribution of rDNA in diploid and polyploid Lolium multiflorum Lam. and fragile sites in 45S rDNA regions. Crop Science 54:1-9.

Butorina AK, Kalaev VN, Najdenova OS, Myagkova OE, Vostricova TV, Mitroshina O, Dirdina OV, Polyakova NA (1999) Relationship between cytogenetic anomalies in forest trees subjected to radioactive contamination and industrial pollution with inherent defects in human infants in central Russia. In: Cytogenetic studies of forest trees and shrubs – review, present status, and outlook on the future. Proceedings of the Second IUFRO Cytogenetics Working Party S2.04.08 Symposium, September 6-12 (1998), Graz, Austria. (edited by H Guttenberger, Z Borzan, SE Schlarbaum, TPV Hartman), pp. 35-41, Special issue of the Forest Genetics (2000). Arbora Publishers, Zvolen, Slovakia, 206 p.

Cabo S, Carvalho A, Martín A, Lima-Brito J (2014) Structural rearrangements detected in newly-formed hexaploid tritordeum after three sequential FISH experiments with repetitive DNA sequences. Journal of Genetics 93:183-188.

Cai Q, Zhang D, Liu Z-L, Wang X-R (2006) Chromosomal localization of 5S and 18S rDNA in five species of subgenus Strobus and their implications for genome evolution of *Pinus*. Annals of Botany 97:715-722.

Carvalho A, Guedes-Pinto H, Lima-Brito J (2013) Polymorphism of the simple sequence repeat $(AAC)_5$ in the nucleolar chromosomes of Old Portuguese wheat cultivars. Journal of Genetics 92:583-586.

Castilho A, Heslop-Harrison J (1995) Physical mapping of 5S and 18S-25S rDNA and repetitive DNA sequences in *Aegilops umbellulata*. Genome 38:91-96.

Cheng BF, Heneen WK, Pedersen C (1995) Ribosomal RNA gene *loci* and their nucleolar activity in *Brassica alboglabra* Bailey. Hereditas 123:169-173.

Cuadrado A, Cardoso M, Jouve N (2008a) Increasing the physical markers of wheat chromosomes using SSRs as FISH probes. Genome 51:809-815.

Cuadrado A, Cardoso M, Jouve N (2008b) Physical organization of simple sequence repeats (SSRs) in Triticeae: structural, functional and evolutionary implications. Cytogenetic and Genome Research 120:210-219.

Cuadrado A, Jouve N (2010) Chromosomal detection of simple sequence repeats (SSRs) using nondenaturing FISH (ND-FISH). Chromosoma 119:495-503.

Cuadrado A, Jouve N (2011) Novel simple sequence repeats (SSRs) detected by ND-FISH in heterochromatin of *Drosophila melanogaster*. BMC Genomics 12:205.

del Cerro Barja A, Lucas-Borja ME, Garcia EM, López Serrano FR, Abellán MA, Morote FAG, López RN (2009) Influence of stand density and soil treatment on the Spanish Black Pine *(Pinus nigra* Arn. ssp. *salzmannii*) regeneration in Spain. Investigación Agraria: Sistemas y Recursos Forestales 18:167-180.

Depeiges A, Goubely C, Lenoir A, Cocherel S, Picard G, Raynal M, Grellet F, Delseny M (1995) Identification of the most-represented repeated motifs in *Arabidopsis thaliana* microsatellite *loci*. Theoretical and Applied Genetics 91:160-168.

Doudrick RL, Heslop-Harrison JS, Nelson CD, Schmidt T, Nance WL, Schwarzacher T (1995) Karyotype of slash pine (*Pinus elliottii* var *elliottii*) using patterns of fluorescence *in situ* hybridization and fluorescence banding. Journal of Heredity 86:289-296.

Drewry A (1982) G-banded chromosomes in *Pinus resinosa*. Journal of Heredity 73:305-306.

Gaussen H, Haywood VH, Chater AO (1964) *Pinus*. In: Flora Europaea, Vol. 1 (edited by TG Tutin, VH Heywood, NA Burges, DM Moore, DH Valentine, SM Walters, DA Webb), pp. 32-35, Cambridge University Press, Cambridge, UK.

Gaussen H, Webb DA, Heywood VH (1993) Genus *Pinus*. In: Flora Europaea, Vol. 1 (edited by TG Tutin, VH Heywood, NA Burges, DM Moore, DH Valentine, SM Walters, DA Webb), pp. 40-44, Cambridge University Press, Cambridge, UK.

Hizume M (1988) Karyomorphological studies in the family Pinaceae. Memoirs of the Facculty of Education Ehime University Natural Science 8:1-108.

Hizume M, Arai M, Tanaka A (1990) Chromosome banding in the genus *Pinus*. III. Fluorescent banding pattern of *P. luchuensis* and its relationships among the Japanese diploxylon pines. The Botanical Magazine, Tokyo 103:103-111.

Hizume M, Ishida F, Murata M (1992) Multiple locations of the rRNA genes in chromosomes of pines, *Pinus densiflora* and *P. thunbergii*. The Japanese Journal of Genetics 67:389-396.

Hizume M, Ohgiku A, Tanaka A (1983) Chromosome banding in the genus *Pinus*. I. Identification of chromosomes in *P. nigra* by fluorescent banding method. The Botanical Magazine, Tokyo 96:273-276.

Hizume M, Ohgiku A, Tanaka A (1989) Chromosome banding in the genus *Pinus*. II. Interspecific variation of fluorescent banding patterns in *P. densiflora* and *P. thunbergii*. The Botanical Magazine Tokyo 102:25-36.

Hizume M, Shibata F, Maruyama Y, Kondo T (2001) Cloning of DNA sequences localized on proximal fluorescent chromosome bands by microdissection in *Pinus densiflora* Sieb. & Zucc. Chromosoma 110:345-351.

Hizume M, Shibata F, Matsusaki Y, Garajova Z (2002) Chromosome identification and comparative karyotypic analyses of four *Pinus* species. Theoretical and Applied Genetics 105:491-497.

Isajev V, Fady B, Semerci H, Andonovski V (2004) EUFORGEN Technical Guidelines for genetic conservation and use for European black pine (*Pinus nigra*). International Plant Genetic Resources Institute, Rome, Italy, 6 p.

Jacobs MD, Gardner RC, Murray BG (2000) Cytological characterization of heterochromatin and rDNA in *Pinus radiata* and *P. taeda*. Plant Systematics and Evolution 223:71-79.

Kaya Z, Ching KK, Stafford SG (1984) A statistical analysis of karyotypes of European Black Pine (*Pinus nigra* Arnold) from different sources. Silvae Genetica 34:148-156.

Kupila-Ahvenniemi S, Hohtola A (1979) Structure of the nucleoli of developing microsporangiate strobili and root tips of Scotch pine. Protoplasma 100:289-301.

Lemos M, Dias A, Carvalho A, Gaspar MJ, Coutinho J, Louzada JL, Lima-Brito J (2015) First molecular characterization of the Portuguese populations of *Pinus nigra* Arnold and extrapolation of their provenances. Plant Molecular Biology Reporter (*submitted*).

Lima-Brito J, Guedes-Pinto H, Heslop-Harrison JS (1998) The activity of nucleolar organizing chromosomes in multigeneric F1 hybrids involving wheat, triticale, and tritordeum. Genome 41:763-768.

Liu Z-L, Zhang D, Hong D-Y, Wang X-R (2003) Chromosomal localization of 5S and 18S-5.8S-25S ribosomal DNA sites in five Asian *Pinus* species using fluorescence *in situ* hybridization. Theoretical and Applied Genetics 106:198-204.

Louro V (1982) O pinheiro larício (*Pinus nigra* Arnold) em Portugal. Divulgação Técnica, MACP, Direcção Geral das Florestas Nº 2 – Agosto 1982, 110 p.

Lubaretz O, Fuchs J, Ahne R, Meister A, Schubert I (1996) Karyotyping of three Pinaceae species via fluorescence *in situ* hybridization and computer-aided chromosome analysis. Theoretical and Applied Genetics 92:411-416.

Machperson P, Filion WG (1981) Karyotype analysis and distribution of constitutive heterochromatin in five species of *Pinus*. Journal of Heredity 72:193-198.

Moreno FJ, Rodrigo RM, Garcia-Herdugo G (1990) Ag-NOR proteins and rDNA transcriptional activity in plant cells. Journal of Histochemestry and Cytochemestry 38:1879-1887.

Müller M, Tausz M, Höfer B, Grill D (1999) Chromosomal changes in root tip meristems of spruce seedlings caused by jasmonic acid. In: Cytogenetic studies of forest trees and shrubs – review, present status, and outlook on the future. (edited by H Guttenberger, Z Borzan, SE Schlarbaum, TPV Hartman), pp. 113-119, Proceedings of the Second IUFRO Cytogenetics Working Party S2.04.08 Symposium, September 6-12 (1998), Graz, Austria. Special issue of the Forest Genetics (2000) Arbora Publishers, Zvolen, Slovakia, 206 p.

Muratova EN (1997a) Cytogenetical study on Scots pine (*Pinus sylvestris* L.) in the Central Yakutia, pp. 157-177. In: Cytogenetic studies of forest trees and shrub species. (edited by Z Borzan, SE Schlarbaum), Proceedings of the first IUFRO Cytogenetics Working Party S2.04-08 Symposium, September 8-11 (1993). Brijuni National Park, Croatia. Croatian Forests, Inc., Zagreb, Faculty of Forestry, University of Zagreb. ISBN 953-6253 15-1/ ISBN 953-6307-28-6.

Muratova EN (1997b) Studies on nucleolar chromosomes in representatives of Pinaceae Lindl, pp. 45-72. In: Cytogenetic studies of forest trees and shrubs, (edited by Z Borzan, Schlarbaum), Proceedings of the first IUFRO Cytogenetics Working Party S2.04-08 Symposium, September 8-11 (1993). Brijuni National Park, Croatia.

Croatian Forests, Inc., Zagreb, Faculty of Forestry, University of Zagreb. ISBN 953-6253 15-1/ ISBN 953-6307-28-6.

Muratova EN, Sedel'nikova TS (2000) Karyotypic variability and anomalies in populations of conifers from Siberia and Far East, pp. 129-141. In: Cytogenetic studies of forest trees and shrubs – review, present status, and outlook on the future. (edited by H Guttenberger, Z Borzan, SE Schlarbaum, TPV Hartman), Proceedings of the Second IUFRO Cytogenetics Working Party S2.04.08 Symposium, September 6-12 (1998), Graz, Austria. Special issue of the Forest Genetics, 2000. Arbora Publishers, Zvolen, Slovakia, 206 p.

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

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

Pavia I, Carvalho A, Rocha L, Gaspar MJ, Lima-Brito J (2014b) Physical location of SSR region and cytogenetic instabilities in *Pinus sylvestris* chromosomes revealed by ND-FISH. Journal of Genetics 93:567–571.

Pavia I, Mengl M, Gaspar MJ, Carvalho A, Heinze B, Lima-Brito J (2014a) Preliminary evidences of two potential native populations of *Pinus sylvestris* L. in Portugal based on nuclear and chloroplast SSR markers. Australian Journal Forest Science 131:Jahrgang Heft 1, S. 1-22.

Pederick LA (1967) The structure and identification of the chromosomes of *Pinus radiata* D. Don. Silvae Genetica 16:69-77.

Poulsen GB, Kahl G, Weising K (1993) Abundance and polymorphism of simple repetitive DNA sequences in *Brassica napus* L. Theoretical and Applied Genetics 85:994-1000.

Raffi ZA, Dodd RS, Zavarin E (1996) Genetic diversity in foliar terpenoids among natural populations of European Black Pine. Biochemical Systematics and Ecology 24:325-339.

Ribeiro T, Barão A, Viegas W, Morais-Cecílio L (2008) Molecular cytogenetics of forest trees. Cytogenetic and Genome Research 120:220-227.

Roa F, Guerra M (2012) Distribution of 45S rDNA sites in chromosomes of plants: Structural and evolutionary implications. BMC Evolutionary Biology 12:1-13.

Rubio-Moraga A, Candel-Perez D, Lucas-Borja ME, Tiscar PA, Viñegla B, Linares JC, Gómez-Gómez L, Ahrazem O (2012) Genetic diversity of *Pinus nigra* Arn. populations in Southern Spain and Northern Morocco revealed by inter-simple sequence repeat profiles. International Journal of Molecular Sciences 13: 5645-5658.

Saylor LC (1964) Karyotype analysis of *Pinus*-group *Lariciones*. Silvae Genetica 13:165-170.

Saylor LC (1972) Karyotype analysis of *Pinus* subgenus *Pinus*. Silvae Genetica 21:155–163.

Saylor LC (1983) Karyotype analysis of the genus *Pinus*-subgenus *Strobus*. Silvae Genetica 32:119-121.

Scaltsoyiannes A, Rohr R, Panetos KP, Tsaktsira M (1994) Allozyme frequency distributions in five European populations of black pine (*Pinus nigra* Arnold). Silvae Genetica 43:20-30.

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

Schubert I, Kunzel G (1990) Position-dependent NOR activity in barley. Chromosoma 99:352-359.

Schwarzacher T, Heslop-Harrison JS (2000) Practical *in situ* hybridization. BIOS Scientific Publishers Limited, Oxford, UK, ISBN 1 85996 138 X, 204 p.

Sedel'nikova TS, Pimenov AV (2010) Chromosomal mutations in Scots pine (*Pinus sylvestris* L.) Growing in a hydromorphic environment: first case of a giant chromosome detection. The Biological Bulletin 37:351–356.

Stack S, Herickhoff L, Sherman J, Anderson L (1991) Staining plant cells with silver. I. The salt-nylon technique. Biotechnic and Histochemistry 1:69-78.

Thompson JN (2005) The geographic mosaic of coevolution. University of Chicago Press, Chicago, Illinois, USA, 400 p.

Chapter VI:

Conclusions and perspectives

VI.1 – Conclusions and prespectives

P. nigra is an economically and ecologically important Mediterranean pine species. Due to an extensive plantation throughout its natural distribution area, most of the populations presently existent in Europe have unknown origin, including the Portuguese ones (Isajev *et al.* 2004). Efforts have been made to identify the provenances of the European *P. nigra* populations in order to develop both conscientious forest management and conservation strategies.

The infraspecific taxonomical classification of *P. nigra* has been intensely revised over the years based on multidisciplinary data and some authors advised the revision of its taxonomy.

Molecular studies based on markers such as nuclear and cpSSR markers, ISSRs or RAPDs have been developed to characterize allopatric populations of *P. nigra* belonging to different infraspecific *taxa* (Liber *et al.* 1999; 2003; Naydenov *et al.* 2006; Afzai-Raffi and Dodd 2007; Lučić *et al.* 2010; Rubio-Moraga *et al*. 2012; Lucas Borja *et al.* 2013; Gülsoy *et al.* 2014; among others).

The isolation of genomic DNA from conifers has been reported as a hard task. Before the development of the genetic diversity study in the Portuguese populations of *P. nigra*, the procedure of DNA extraction need to be optimized due to the high content of polysaccharides, endogenous tannins and polyphenols in pines (Tibbits *et al.* 2006), including *P. nigra*. Our optimized protocol proved to be useful for isolation of amplifiable DNA and was also versatile and suitable for DNA extraction from other tissues and species.

In order to determine specific genetic profiles for each *P. nigra* infraspecific *taxon*, we analyzed 120 foreign samples from 12 different provenances that were bulked in order to minimize the intra-population genetic diversity. The different molecular analyses, using ISSR and SCoT markers, corroborated the adopted taxonomic classification, clustering most of the bulked samples per allopatric subspecies and regional variety. With this study it was possible to verify that ISSRs and SCoTs are cost-effective markers with ability to discriminate allopatric infraspecific *taxa*, namely subspecies and varieties of *P. nigra*. Such result evidenced that these markers constitute valuable tools for discrimination of *P. nigra* infraspecific *taxa* that could be highly useful for the definition of forest management strategies, since allopatric subspecies of *P. nigra* are inter-fertile (Vidaković 1991; Isajev *et al.* 2004) and the plantation of stands of unknown origin in the proximity of autochthonous

pinewoods could induce genetic pollution (Zaghi 2008) or threaten the genetic diversity of European Black Pine populations (Naydenov *et al.* 2006).

The ISSR and SCoT markers were also reliable in the first assessment of genetic diversity within and among the six actual Portuguese populations of *P. nigra* that are representative of the species distribution in the country. Besides, these markers also allowed the extrapolation of their provenances and the first ascribing of their infraspecific taxonomy based on molecular data. Briefly, after the characterization of 127 *P. nigra* individuals from the populations of 'Manteigas', 'Caminha', Campeã', 'Vila Pouca de Aguiar', 'Paredes de Coura' and 'Alvão', a highly differentiated genetic structure was found, since all individuals were clustered per population. The genetic diversity was higher within rather than among populations, which is a characteristic feature of pine species.

The origin of the forest reproductive material used in the installation of *P. nigra* stands in Portugal during the last century was also extrapolated based on ISSR and SCoT data. For that purpose, the molecular patterns of the Portuguese samples were compared with those produced in seven DNA bulks representative of foreign individuals belonging to three allopatric subspecies and varieties: subsp. *nigra* var. *austriaca*; subsp. *laricio* var. *corsicana* and var. *calabrica*; and subsp. *salzmannii*, previously ascribed as existent in Portugal in the 1980's (Louro 1982). Separated genetic analyses revealed that the ISSR markers evidenced higher genetic identity among the Portuguese and foreign samples from *P. nigra* subsp. *laricio* var. *calabrica*, whereas SCoTs estimated a higher genetic identity with *P. nigra* subsp. *laricio* var. *corsicana*. Both markers showed high genetic similarity with subspecies *laricio*, corroborating the assumption of Louro (1982).

The two marker systems used in this work already proved to be suitable for DNA fingerprinting among individuals belonging to different *taxa* (Coutinho *et al*. 2014a,b); among interspecific hybrids and their parental species (Carvalho *et al*. 2005); among tritordeum and respective parents (Cabo *et al*. 2014a,b); and/or for extrapolation of provenances (Cipriano *et al*. 2013). The ISSRs are abundant and disperse throughout the genome and could amplify coding and noncoding regions, whereas the SCoT markers use primers that bind to a short conserved sequence that flanks the start codon (ATG) of the plant genes (Collard and Mackill 2009). Regarding the conservation degree of plant genes even among different species, SCoT markers could provide more reliable molecular data than the ISSRs for the extrapolation of provenances of the Portuguese Black Pine populations. The amplification of SCoTs requires the binding of the primer to a short conserved sequence that flanks the start codon,

contributing for the comparison of highly conserved regions that could provide the estimation of more accurate genetic relationships and estimation of provenances. In fact, the percentage of polymorphism and total number of amplified bands by SCoTs, demonstrated their high discriminative potential for the genetic diversity evaluation of *P. nigra* within and among populations.

Finally, our molecular studies in *P. nigra* were complemented by a cytogenetic characterization in one single population ('Campeã') and its comparison with an autochthonous population of *P. sylvestris*, 'Biduiça' ('Serra do Gerês'), regarding that both species have a peripheral distribution in Portugal, relatively to their natural distribution areas. As far as we know, this study constituted the first molecular cytogenetic characterization for both populations. Despite the previous development of nucleolar activity studies performed in both species by other authors, here it was present for the first time a simple and cost-effective silver staining protocol that was adapted from the Stack *et al*. (1991) method. This technique could be extended to other forest trees. In addition to the first molecular characterization of Portuguese individuals belonging to these two pine species, our study revealed that nucleolar activity results and their comparison with FISH might enable indications about the adaptive potential of peripheral populations to the surrounding environment and how the nucleolar activity could be involved in the adaptation process, even in geographic locations where the conditions seem to be favorable.

Hereafter, it will be useful the optimization of cytogenetic protocols to be applied in pine for the detection of Ag-NORs, for the sequential use of FISH with 45S rDNA probes and for the preservation of morphology and spread of these large chromosomes that have been preventing the accurate identification of individual chromosomes. Consequently, the improvement of the morphology and identification of individual chromosomes will contribute for the detection of cytogenetic irregularities or for the discovering of new mutation types in individuals from peripheral populations. Also, a more exhaustive cytogenetic approach involving different populations of *P. nigra* in Portugal and different subspecies/ varieties of this species, should be useful to increase the knowledge about these species, namely into SSR distribution and occurrence of chromosome aberrations in the remaining Portuguese *P. nigra* populations.

In spite of the results obtained in this dissertation, more studies should be performed in Portuguese *P. nigra* populations, as well as, with individuals of other subspecies and/or varieties. Additionally, it will be useful to use other molecular marker systems, such as

retrotransposons, nuclear and cpSSRs, among others, to confirm the taxonomy of the Portuguese *P. nigra* populations, preliminary determined with this study.

Results of silvicultural, ecological and genetic studies should be progressively integrated into the adaptive management, which is crucial for the sustainable supply of goods and services from forests in the face of climate change. In fact, forest species have evolved under several periods of climatic change and their genetic diversity gives them the ability to adapt to changing environments. Since, the Mediterranean forests of *P. nigra* are included in the E.U. list of endangered natural habitats that require specific conservation measures (Resolution 4/1996 by the Convention on the Conservation of European Wildlife and Natural Habitats) (del Cerro Barja *et al.* 2009), the results achieved in this work will provide valuable information for the future definition of conscientious plantation, reforestation and conservation strategies of *P. nigra* in Portugal and abroad.

VI.2 – References

Afzal-Rafii Z, Dodd RS (2007) Chloroplast DNA supports a hypothesis of glacial refugia over postglacial recolonization in disjunct populations of black pine (*Pinus nigra*) in Western Europe. Molecular Ecology 16:723-736.

Cabo S, Carvalho A, Rocha L, Martín A, Lima-Brito J (2014a) IRAP, REMAP and ISSR fingerprinting in newly formed hexaploid tritordeum (X Tritordeum Ascherson et Graebner) and respective parental species. Plant Molecular Biology Reporter 32(3):761-770.

Cabo S, Ferreira L, Carvalho A, Martins-Lopes P, Martín A, Lima-Brito JE (2014b) Potential of Start Codon Targeted (SCoT) markers for DNA fingerprinting of newly synthesized tritordeums and their respective parents. Journal of Applied Genetics 55:307-312.

Carvalho A, Matos M, Lima-Brito J, Guedes-Pinto H, Benito C (2005) DNA fingerprint of F_1 interspecific hybrids from the Triticeae tribe using ISSRs. Euphytica 143:93-99.

Cipriano J, Carvalho A, Fernandes C, Gaspar MJ, Pires J, Bento J, Roxo L, Louzada J, Lima-Brito J (2013) Evaluation of genetic diversity of Portuguese *Pinus sylvestris* L. populations based on molecular data and inferences about the future use of this germplasm. Journal of Genetics 92:e41-e48.

Collard BCY, Mackill DJ (2009) Start Codon Targeted (SCoT) polymorphism: A simple novel DNA marker technique for generating gene-targeted markers in plants. Plant Molecular Biology Reporter 27:86–93.

Coutinho JP, Carvalho A, Lima-Brito J (2014a) Genetic diversity assessment and estimation of phylogenetic relationships among 26 Fagaceae species using ISSRs. Biochemical Systematics and Ecology 54:247–256.

Coutinho JP, Carvalho A, Lima-Brito J (2014b) Fingerprinting of Fagaceae individuals using inter-microsatellite markers. Journal of Genetics - Online Resources 93:e132-e140.

del Cerro Barja A, Lucas-Borja ME, Garcia EM, López Serrano FR, Abellán MA, Morote FAG, López RN (2009) Influence of stand density and soil treatment on the Spanish Black Pine (*Pinus nigra* Arn. ssp. *Salzmannii*) regeneration in Spain. Investigación Agraria: Sistemas y Recursos Forestales 18:167-180.

Gülsoy AD, Gülsoy AM, Çengel B, Kaya Z (2014) The evolutionary divergence of *Pinus nigra* subsp. *pallasiana* and its varieties based on noncoding trn regions of chloroplast genome. Turkish Journal of Botany 38:627-636.

Liber Z, Pavletió Z, Nikolió T, Fülgosi H (1999) DNA polymorphisms in Austrian and Dalmatian black pine. Wiss Mitt Niederösterr. Landesmuseum 12:53-62.

Liber Z, Nikolić T, Mitić B, Šatović Z (2003) RAPD markers and Black pine (*Pinus nigra* Arnold) intraspecies taxonomy- Evidence from the study of nine populations. Acta Societatis Botanicorum Poloniae 72:249-257.

Louro V (1982) O Pinheiro Laricio (*Pinus nigra* Arnold) em Portugal. Divulgação Técnica N.º 2. Direcção Geral das Florestas. MACP.

Lucas Borja ME, Candel Pérez D, Molero Carrasco J, Ahrazem O, Gomez Gomez ML, Rubio Moraga A, Monreal Montoya J, Botella Miralles O (2013) Utilización de los marcadores ISSR para la determinación de diversidad genética en poblaciones de *Pinus nigra* Arn.. 6º Congreso Forestal Español. Montes: Servicio y desarollo rural. 10-14 junio 2013, Vitoria-Gasteiz.

Lučić A, Isajev V, Nevenić R, Ćirković-Mitrović T, Brašanac-Bosanac L (2012) Analysis of genetic variability of Austrian Pine (*Pinus nigra* Arnold) in Serbia using protein markers. South-East European Forestry 3:1-8.

Isajev V, Fady B, Semerci H Andonovski V (2004) EUFORGEN Technical Guidelines for genetic conservation and use for European black pine (*Pinus nigra*). International Plant Genetic Resources Institute, Rome.

Naydenov K, Tremblay FM, Fenton NJ, Alexandrov A (2006) Struture of *Pinus nigra* Arn. populations in Bulgaria revealed by chloroplast microsatellites and terpenes analysis: Provenance tests. Biochemical Systematics and Ecology 34:562-574.

Rubio-Moraga A, Candel-Perez D, Lucas-Borja ME, Tiscar PA, Viñegla, Linares JC, Gómez-Gómez L, Ahrazem O (2012) Genetic diversity of *Pinus nigra* Arn. populations in Southern Spain and Northern Morocco revealed By inter-simple sequence repeat profiles. International Journal of Molecular Sciences 13:5645-5658.

Stack S, Herickhoff L, Sherman J, Anderson L (1991) Staining plant cells with silver. I. The salt-nylon technique. Biotechnic and Histochemistry 1:69-78.

Tibbits JFG, McManus LJ, Spokevicius AV, Bossinger G (2006) A rapid method for tissue collection and high-throughput isolation of genomic DNA from mature trees. Plant Molecular Biology Reporter 24:81-91.

Vidaković M (1991) Conifers morphology and variation, Translated from Croatian by Maja Soljan. Croatia, Graficki Zavod Hrvastske.

Zaghi D (2008) Management of Natura 2000 habitats. 9530 *(Sub)-Mediterranean pine forests with endemic black pines. European Commission.

Supplemental Material

Supplement 1:

Wood Science and Technology

Single protocol for genomic DNA extraction from foliar, dry and vascular tissues of gymnosperms

--Manuscript Draft--

Manuscript Number:	
Full Title:	Single protocol for genomic DNA extraction from foliar, dry and vascular tissues of gymnosperms
Article Type:	Original Article
Corresponding Author:	Ana Carvalho, Ph.D. PORTUGAL
Corresponding Author Secondary Information:	
Corresponding Author's Institution:	
Corresponding Author's Secondary Institution:	
First Author:	Maria Lemos
First Author Secondary Information:	
Order of Authors:	Maria Lemos
	Alexandra Dias
	Ana Carvalho, Ph.D.
	Andreia Delgado
	Ivo Pavia
	João Coutinho
	Maria João Gaspar
	José Lima-Brito

Supplement Material

Supplement 2:

Plant Systematics and Evolution

Genetic profiles of European allopatric subspecies and varieties of Pinus nigra Arn. based on ISSR and SCoT markers

--Manuscript Draft--

Manuscript Number:	
Full Title:	Genetic profiles of European allopatric subspecies and varieties of Pinus nigra Arn. based on ISSR and SCoT markers
Short Title:	Genetic profiles of allopatric Pinus nigra subspecies and varieties
Article Type:	Original Article
Keywords:	Allopatric subspecies; infraspecific taxa; inter-simple sequence repeat (ISSR); Pinus nigra; start codon targeted (SCoT) polymorphism
Corresponding Author:	José Eduardo Lima-brito, PhD University of tras-os-montes and alto douro vila real, PORTUGAL
Corresponding Author Secondary Information:	
Corresponding Author's Institution:	University of tras-os-montes and alto douro
Corresponding Author's Secondary Institution:	
First Author:	Alexandra Dias
First Author Secondary Information:	
Order of Authors:	Alexandra Dias
	Maria Lemos
	Ana Carvalho
	José Luís Louzada
	Maria João Gaspar
	José Eduardo Lima-brito, PhD
Order of Authors Secondary Information:	
Funding Information:	Fundação para a Ciência e a Tecnologia (Project UID/AGR/04033/2013) — José Luís Louzada
	Fundação para a Ciência e a Tecnologia (Grant SFRH/BD/91781/2012) — Alexandra Dias
	Fundação para a Ciência e a Tecnologia (SFRH/BPD/68932/2010) — Ana Carvalho
Abstract:	Pinus nigra Arnold is a Mediterranean pine with a large distribution area, presenting high morphological, physiological and ecological variation, which has led to identification and taxonomic classification uncertainties. The present work concerns the molecular characterization of P. nigra populations from distinct geographical locations that were previously classified as different infraspecific taxa, namely, subspecies and varieties. Populations from different geographic origins (France, Italy, Austria and Romania) were chosen and are representative of three allopatric subspecies and three regional varieties. The comparison of their molecular patterns was performed for the DNA fingerprinting and estimation of genetic relationships based on inter-simple sequence repeat (ISSR) and start codon targeted (SCoT) markers, to infer about the adopted taxonomic classification. According to the ISSR and SCoT data, the studied P. nigra provenances presented 76% of genetic similarity and were clustered into four groups and one branch considering a cut off value of 0.876 in the UPGMA dendrogram. These groups discriminated the bulked individuals per subspecies and variety except for the samples of subspecies nigra which were clustered in two different groups, suggesting that the studied individuals could belong to different varieties. The subspecies salzmannii was projected as a branch in the dendrogram, indicating its higher genetic dissimilarity relatively to the remaining bulked samples. In general, ISSRs and SCoTs revealed to be reliable marker systems for DNA fingerprinting, estimation of genetic relationships and discrimination of infraspecific taxa of P. nigra.

Supplement 3:

Plant Molecular Biology Reporter

First molecular characterization of the Portuguese populations of Pinus nigra Arnold and extrapolation of their provenances

--Manuscript Draft--

Manuscript Number:	PMBR-D-15-00356
Full Title:	First molecular characterization of the Portuguese populations of Pinus nigra Arnold and extrapolation of their provenances
Article Type:	Original Article
Keywords:	European Black Pine; genetic diversity and structure; inter-simple sequence repeat (ISSR); Start Codon Targeted (SCoT).
Corresponding Author:	José Lima-Brito Institute for Biotechnology and Bioengeneering, Centre of Genetics and Biotechnology of the University of Tras-os-Montes and Alto Douro Vila Real, PORTUGAL
Corresponding Author Secondary Information:	
Corresponding Author's Institution:	Institute for Biotechnology and Bioengeneering, Centre of Genetics and Biotechnology of the University of Tras-os-Montes and Alto Douro
Corresponding Author's Secondary Institution:	
First Author:	Maria Lemos
First Author Secondary Information:	
Order of Authors:	Maria Lemos
	Alexandra Dias
	Ana Carvalho
	Maria João Gaspar
	João Coutinho
	José Luís Louzada
	José Lima-Brito
Order of Authors Secondary Information:	
Funding Information:	Fundação para a Ciência e a Tecnologia (SFRH/BPD/68932/2010) — Ana Carvalho
	Fundação para a Ciência e a Tecnologia (SFRH/BD/91781/2012) — Alexandra Dias
Abstract:	Pinus nigra Arnold (Family Pinaceae) belongs to the Mediterranean pines group. This study constitutes the first molecular characterization of six allochthonous Portuguese populations ('Manteigas', 'Vale do Zêzere', 'Campeã', 'Vila Pouca de Aguiar', 'Paredes de Coura' and 'Caminha') using both inter-simple sequence repeat (ISSR) and Start Codon Targeted (SCoT) markers. These populations were installed 50 to 90 years ago but their provenances are unknown. Morphological and anatomical evaluations of these Portuguese stands of P. nigra were performed during the 1980's. The existence of three subspecies was ascribed: laricio (var. corsicana and var. calabrica), salzmannii and nigra. Foreign samples belonging to these subspecies and varieties were included in this study in order to extrapolate the provenances of the actual Portuguese P. nigra populations based on molecular data. The pool of the ISSR and SCoT data revealed 65% of genetic similarity among the 127 Portuguese P. nigra individuals that were structured into six highly differentiated populations. The polymorphism was higher within than among populations. A reduced correlation coefficient was detected between the genetic distance matrices produced by ISSRs and SCoTs. Due so, for provenances extrapolation, we performed a separate

	molecular analysis for each marker system. ISSRs evidenced higher genetic identity among the Portuguese and foreign samples from P. nigra subsp. laricio var. calabrica, whereas SCoTs estimated a higher genetic identity with P. nigra subsp. laricio var. corsicana. Both markers showed high genetic similarity with subspecies laricio, corroborating the morphological and anatomical characterizations performed in the Portuguese P. nigra stands at the 1980's.

Supplement 4:

Plant Cell Reports

Nucleolar activity, SSRs hybridization and chromosome irregularities in peripheral populations of Pinus nigra Arn. and Pinus sylvestris L.

--Manuscript Draft--

Manuscript Number:	
Full Title:	Nucleolar activity, SSRs hybridization and chromosome irregularities in peripheral populations of Pinus nigra Arn. and Pinus sylvestris L.
Article Type:	Original Article
Funding Information:	Fundação para a Ciência e a Tecnologia (Grant SFRH/BPD/68932/2010) — Ph.D. Ana Carvalho
Abstract:	Key message Nucleolar activity was never studied in Pinus nigra. Not all rDNA loci in P.nigra and P.sylvestris are transcriptionally active. Chromosome irregularities in both revealed ongoing genomic adaptation of peripheral populations. Abstract Portuguese populations of Pinus nigra and Pinus sylvestris are considered peripheral. These stands were installed in 20th century, but autochthonous P. sylvestris populations were ascribed to 'Serra do Gerês'. Only one P. sylvestris autochthonous population was cytogenetically characterized so far. We intended to characterize cytogenetically individuals from two Portuguese populations, one of P. nigra and other of P. sylvestris. We used nondenaturing-fluorescence in situ hybridization (ND-FISH) and FISH with simple sequence repeats (SSRs) and one 45S rDNA probes, respectively, to establish comparisons and detect potential chromosome irregularities in both species. We used silver nitrate staining to determine the number of nucleoli/interphase nucleus and compared with the rDNA loci number. A different number of SSR hybridization signals were detected between the species. After FISH, 14 and 16 rDNA loci were observed in P. sylvestris and P. nigra, respectively. A variable number of nucleoli/interphase nucleus (1 to 12) in both species, suggested that not all rDNA loci were transcriptionally active. A variable number of micronucleoli were also detected in both species and might be related with the observed irregularities (e.g. ring, polycentric and fragmented chromosomes). Such instabilities are common in Pinaceae peripheral populations growing in adverse environmental conditions, and should be regarded as microevolution elements that increase genetic diversity and adaptive ability.
Corresponding Author:	José Lima-Brito, PhD Universidade de Tras-os-Montes e Alto Douro Vila Real, PORTUGAL
Corresponding Author Secondary Information:	
Corresponding Author's Institution:	Universidade de Tras-os-Montes e Alto Douro
Corresponding Author's Secondary Institution:	
First Author:	Maria Lemos
First Author Secondary Information:	
Order of Authors:	Maria Lemos Ana Carvalho Ivo Pavia Andreia Delgado Maria João Gaspar José Lima-Brito, PhD