Genetic characterization of Portuguese allochthonous populations of *Pinus nigra* using ISSRs and SCoTs and extrapolation of their infraspecific taxonomy

Alexandra Dias¹ · Maria Lemos² · Ivo Pavia¹ · Maria João Gaspar^{3,4,5} ·
Maria Emília Silva¹ · José Luís Louzada¹ · José Lima-Brito^{1,3,5} · Ana Carvalho^{1,3,5}

Received: 16 November 2018 / Revised: 11 January 2019 / Accepted: 5 February 2019 / Published online: 23 February 2019
© Prof. H.S. Srivastava Foundation for Science and Society 2019

Abstract The *Pinus nigra* distribution in Portugal is restricted to six allochthonous populations with unknown origin and infraspecific taxonomy. This work intends to evaluate their genetic diversity, structure and relationships, and to infer about their infraspecific taxonomy by comparing molecular patterns produced by inter-simple sequence repeat and Start Codon Targeted markers among Portuguese and foreign samples with known taxonomy and provenance. 127 Portuguese *P. nigra* individuals were clustered per population. The genetic differentiation was higher within rather than among populations. The pooled molecular data indicated high genetic proximity among the Portuguese and foreign samples of subspecies *laricio*. However, the separate analysis per marker system demonstrated that two varieties of subspecies *laricio* (*corsicana* and *calabrica*) may have been used in the

plantations of the Portuguese *P. nigra* stands performed in the last century. The genetic characterization and extrapolation of the infraspecific taxonomy of these populations provide useful information for forest management, afforestation and germplasm use.

Keywords European black pine · Genetic structure · Infraspecific *taxa* · Molecular markers

Introduction

Pinus nigra Arnold has high ecological and economic importance and its distribution ranges from Southern Europe, Asia Minor, Corsica and Sicily islands, and NW Africa (Afzal-Rafii and Dodd 2007; Rubio Moraga et al. 2012; Dias et al. 2018). The discontinuous distribution resulted in populations with high genetically variability (Scaltsoyiannes et al. 1994; Afzal-Rafii and Dodd 2007) which raised taxonomic issues and various revisions of the *P. nigra* infraspecific taxonomy (Akkemik et al. 2010). The most recognized infraspecific taxonomic classification considers five subspecies with Eurasian distribution: *pallasiana* (Turkey and Crimea), *dalmatica* (Balkans), *salzmannii* (France and Spain), *laricio* (Corsica, Sicily and Calabria) and *nigra* (Alps) (Gaussen et al. 1964; Farjon 1998, 2010), each one including regional varieties (Christensen 1997; Barbéro et al. 1998; Price et al. 1998; Akkemik et al. 2010). A sixth subspecies, *mauretanica* (Maire and Peyerimh) Heywood, was assigned to the NW of Africa (Little and Critchfield 1969).

The Portuguese *P. nigra* populations were planted 50- to 90-years ago (Dias et al. 2018) with plant material of unknown origin and infraspecific taxonomy.

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s12298-019-00649-5>) contains supplementary material, which is available to authorized users.

✉ Ana Carvalho
anacar@utad.pt

¹ Center of Research and Technology of Agroenvironmental and Biological Sciences (CITAB), University of Tras-os-Montes and Alto Douro, 5000-801 Vila Real, Portugal

² University of Tras-os-Montes and Alto Douro, 5000-801 Vila Real, Portugal

³ BioISI - Biosystems & Integrative Sciences Institute, University of Tras-os-Montes and Alto Douro (BioISI/UTAD), 5000-801 Vila Real, Portugal

⁴ Forest Research Centre (CEF), Instituto Superior de Agronomia (ISA), University of Lisbon, Lisbon, Portugal

⁵ Department of Genetics and Biotechnology (DGB), Ed. Blocos Laboratoriais, A1.09, University of Tras-os-Montes and Alto Douro, 5000-801 Vila Real, Portugal

Different molecular markers, including the ISSRs, have been used to characterize *P. nigra* populations (Liber et al. 2003; Naydenov et al. 2006; Afzal-Rafii and Dodd 2007; Rubio Moraga et al. 2012; Lucas Borja et al. 2013; Candel Pérez 2014; Šarac et al. 2015; Bonavita et al. 2016; Giovannelli et al. 2017). The SCoT markers have been successfully used alone or combined with other markers for genetic diversity and fingerprinting in various species (Collard and Mackill 2009; Cabo et al. 2014; Alikhani et al. 2014; Zhang et al. 2015), but have never been used in *P. nigra*.

This work intended to characterize genetically the Portuguese *P. nigra* populations using ISSRs and SCoTs concerning to their genetic diversity, structure, relationships, and infraspecific taxonomy.

Materials and methods

Needles were sampled on a total of 127 *P. nigra* individuals randomly selected in six populations that were dendrometrically characterized by Dias et al. (2018) (Table 1).

Needles from foreign *P. nigra* individuals of known infraspecific *taxa* and provenances (Table 2) were also included in this study for further comparison with the Portuguese samples. Both Portuguese and foreign needles were used for DNA extraction using the Doyle and Doyle (1987) protocol.

ISSRs were amplified according to Carvalho et al. (2005) with the primers reported by Rubio Moraga et al. (2012) and Cipriano et al. (2016). For SCoTs amplification, the Collard and Mackill (2009) primers and the Cabo et al. (2014) procedure were used. Each band was considered a marker, and bands with the same molecular weight produced by the same primer were considered as the same *locus* in the presence (1)/absence (0) analyses.

The combined ISSR and SCoT data allowed the estimation of genetic diversity, structure and relationships of the Portuguese populations. An Unweighted Pair Group Method with Arithmetic means (UPGMA) dendrogram of genetic similarity using the “simple matching” coefficient,

the sequential agglomerative hierarchical nested (SAHN) module and the software NTSYS v.2.02 (Rohlf 1998) was constructed. The clustering robustness was estimated by bootstrap analysis using 2,000 replications, a confidence level of 99% (Hedges 1992) and WinBoot (Yap and Nelson 1996). The genetic structure was inferred by STRUCTURE Ver.2.3.4 (Pritchard et al. 2000; Falush et al. 2007) using the ‘no admixture’ parameter, 50,000 generations of burn-in period, 100,000 Markov chain Monte Carlo iterations and different *K*-values. The optimal *K*-value was assessed with STRUCTURE HARVESTER (Earl and vonHoldt 2012). The Principal Coordinates Analyses (PCoA) and Analysis of Molecular Variance (AMOVA) were performed with GenALEX 6.5 (Peakall and Smouse 2006, 2012).

The number of observed (*na*) and effective (*ne*) alleles, Nei’s gene diversity (*h*), Shannon’s Information index (*I*), total genetic diversity (H_T), genetic diversity within populations (H_S), relative magnitude of differentiation among populations (G_{ST}), and genetic diversity among populations ($D_{ST} = H_T - H_S$) were calculated with POPGENE 1.32 (Yeh et al. 1999).

Results and discussion

Sixteen primers, eight for ISSRs and eight for SCoTs amplification, were selected. Both ISSRs and SCoTs detected intra- and inter-population polymorphism (Online Resource 1; Table 3).

SCoTs produced a higher average of bands per primer (18.63) than the ISSRs (14.63) but the latter revealed the highest total percentage of polymorphism (100%) (Table 3). ISSRs have been highly discriminative in *P. nigra* (Rubio Moraga et al. 2012) and *P. sylvestris* (Cipriano et al. 2016). This work also demonstrated that SCoTs could be highly discriminative in *P. nigra* (Table 3).

The pooled ISSR and SCoT data revealed a total genetic diversity ($H_T = 0.343$) higher than the genetic diversity within populations ($H_S = 0.2036$), an inter-population

Table 1 Data of the sampled Portuguese *P. nigra* populations

Population	Geographic coordinates	Average age (years)	Number of sampled individuals
Manteigas	40°22'47.00"N; 7°33'18.00"W	93.3	20
Vale do Zêzere	40°19'19.00"N; 7°34'26.00"W	59.1	20
Paredes de Coura	41°52'0.00"N; 8°36'21.00"W	57.8	20
Caminha	41°50'15.00"N; 8°43'57.00"W	57.9	25
Campeã	41°19'9.12"N; 7°53'28.35"W	58.1	20
Vila Pouca de Aguiar	41°31'02.72"N; 7°35'31.36"W	74.7	22

Table 2 Data of the foreign *P. nigra* samples

Infraspecific taxonomy:		Provenance (sampling local)	Code
Subspecies	Variety		
<i>P. nigra</i> subsp. <i>laricio</i> (Poir.) Maire	<i>calabrica</i>	Cosenza, Calabria, Italy (provenance trial)	COSE
		Aspromonte Calabria, Italy (provenance trial)	ASPR
<i>P. nigra</i> subsp. <i>laricio</i> (Poir.) Maire	<i>corsicana</i>	Sologne-Vayrières-VG (seed orchard)	PLO-VG-01
		Corse-Haute Serre-VG (seed orchard)	PLO-VG-02
<i>P. nigra</i> subsp. <i>nigra</i> J.F. Arnold	<i>austriaca</i>	Doblhoff, Austria (provenance trial)	DOBL
		Banat, Romania (provenance trial)	BANA
<i>P. nigra</i> subsp. <i>salzmannii</i> (Dunal) Franco	–	Saint Guilhem le Désert, France (natural population)	PNS

Table 3 Number of total amplified (T) and polymorphic (P) bands, and percentage of polymorphism (%P) per primer

Marker	Primer	Sequence (5' → 3')	T	P	%P
ISSR	817	(CA) ₈ A	20	20	100
	827	(AC) ₈ G	18	18	100
	834	(AG) ₈ YT*	11	11	100
	835	(AG) ₈ YC*	18	18	100
	836	(AG) ₈ YA*	7	7	100
	841	(GA) ₈ YC*	13	13	100
	850	(GT) ₈ YC*	12	12	100
	866	(CTC) ₆	18	18	100
Total			117	117	100
Marker	Primer	Sequence (5' → 3')	T	P	%P
SCoT	SCoT3	CAACAATGGCTACCACCG	24	24	100
	SCoT7	CAACAATGGCTACCACGG	21	21	100
	SCoT17	ACCATGGCTACCACCGAG	11	11	100
	SCoT18	ACCATGGCTACCACCGCC	22	22	100
	SCoT25	ACCATGGCTACCACCGGG	22	22	100
	SCoT26	ACCATGGCTACCACCGTC	15	12	80
	SCoT27	ACCATGGCTACCACCGTG	14	13	92.85
	SCoT31	CCATGGCTACCACCGCCT	20	20	100
Total			151	147	97.14

*Y = C or T

genetic diversity (D_{ST}) of 0.1398 and a high differentiation among the Portuguese populations ($G_{ST} = 0.4072$). The total number of observed alleles (n_a) was 1.9851 and the highest value was registered in Caminha (Online Resource 2). The total number of effective alleles (n_e) was 1.5896, and its highest value was detected in Vila Pouca de Aguiar (Online Resource 2). The total values of Nei's gene diversity (h) and Shannon's Information index (I) were 0.3458 and 0.5186, respectively (Online Resource 2). The Portuguese *P. nigra* individuals showed 59% of genetic similarity and were clustered into five main groups (I to V) for a cutoff value of 0.68 (Fig. 1). However, considering the subgroups II-a and II-b, six differentiated clusters, matching the number of studied populations, are evidenced

(Fig. 1). The bootstrap analysis estimates the confidence level of the phylogenetic hypothesis proposed by a dendrogram (Felsenstein 1985). Only bootstrap values higher than 50% are presented in Fig. 1, indicating a reliable clustering. Despite presenting different average ages (Table 1; Dias et al. 2018), the populations of Manteigas, V. Zêzere, Campeã and P. Coura showed high genetic similarity (Figs. 1, 2a). Probably, the plant material used in the first stands was reused later in other plantations explaining such genetic relationships (Figs. 1, 2a).

The genetic structure analyses determined $K = 6$ and F_{ST} values surrounding 0.25 (Online Resource 3), meaning high genetic differentiation (Savolainen et al. 2007) and corroborating the UPGMA clustering (Fig. 1). The

Fig. 1 UPGMA dendrogram of genetic similarity among 127 *P. nigra* individuals based on the pooled ISSR and SCoT data, showing five main groups (I to V) and two subgroups (II-a and II-b). Bootstrap percentage values (99% of confidence) are indicated

Fig. 2 Principal Coordinates Analyses based on the pooled ISSR and/or SCoT data and Nei's genetic distance matrix. **(a)** Projection of the Portuguese *P. nigra* populations, four showing higher proximity (circle). **(b)** Higher genetic proximity of the Portuguese populations with the foreign samples PLO-VG-02 (subsp. *laricio* var. *corsicana*) and COSE (subsp. *laricio* var. *calabrica*). **(c)** The ISSRs revealed higher genetic proximity among the Portuguese and foreign samples of subsp. *laricio* var. *corsicana*. **(d)** The SCoTs showed higher similarity among the Portuguese and foreign samples of subsp. *laricio* var. *calabrica*

AMOVA (Excoffier et al. 1992) indicated higher molecular variation within (57%) rather than among (43%) populations. This is expected for long lifespan, highly fertile and allogamous species like gymnosperms (Hamrick et al. 1992). In fact, similar results were found in other *P. nigra* populations (Liber et al. 2003; Naydenov et al. 2006; Afzal-Rafii and Dodd 2007; Rubio Moraga et al. 2012; Šarac et al. 2015; Bonavita et al. 2016).

A botanical characterization of the Portuguese *P. nigra* stands performed in the 1980's classified them as belonging to subspecies *laricio*, *salzmannii* and *nigra* (Louro 1982). Therefore, we compared the Portuguese populations with foreign samples representative of these infraspecific taxa. High genetic proximity among the Portuguese and foreign samples of subsp. *laricio* var. *corsicana* (PLO-VG-02) and subsp. *laricio* var. *calabrica* (COSE) was found (Fig. 2b). By analyzing each marker system separately, the ISSRs revealed higher similarity among the Portuguese and foreign samples of var. *corsicana* (Fig. 2c), whereas the SCoTs suggested higher proximity with var. *calabrica* (Fig. 2d). The first three axes of each PCoA in Fig. 2 explained a cumulative percentage of total variation higher than 70%, being statistically relevant (Reis 1997).

Our work corroborated partially the assumptions of Louro (1982) who reported a predominance and best adaptation of subsp. *laricio* var. *corsicana* in Portugal. However, plant material of subsp. *laricio* var. *calabrica* were probably also used in the plantation of the Portuguese stands, justifying the observed genetic relationships (Figs. 1, 2a).

This work provided valuable information for future forest management, afforestation and germplasm use of *P. nigra* in Portugal and abroad.

Acknowledgements The authors acknowledge to Prof. Bruno Fady (INRA, Avignon, France) and Dr. Marc Calvignac (Vilmorin S.A., La Ménétré, France) that kindly provided dehydrated needles and certified seeds, respectively, representative of different *Pinus nigra* taxa.

Funding This study was funded by “Fundação para a Ciência e a Tecnologia” (FCT) (grant SFRH/BD/91781/2012 co-financed by “Fundo Social Europeu” under the POPH-QREN program and projects UID/AGR/04033/2019 (to CITAB/UTAD), UID/MULTI/04046/2019 (to BioISI), and POCI-01-0145-FEDER-006958 - FEDER/COMPETE/POCI – Operational Competitiveness and Internationalization Program (to CITAB/UTAD).

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflicts of interest.

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: phylogeography of European black pine. *Mol Ecol* 16:723–736
- Akkemik Ü, Yılmaz H, Oral D et al (2010) Some changes in taxonomy of pines (*Pinus* L.) native to Turkey. *J Fac For Istanbul U* 61(1):63–78
- Alikehani L, Rahmani M-S, Shabanian N et al (2014) Genetic variability and structure of *Quercus brantii* assessed by ISSR, IRAP and SCoT markers. *Gene* 552:176–183
- Barbéro M, Loisel R, Quézel P et al (1998) Pines of the Mediterranean basin. In: Richardson DM (ed) *Ecology and biogeography of Pinus*. Cambridge University Press, Cambridge, 527 pp
- Bonavita S, Vendramin GG, Bernardini V et al (2016) The first SSR-based assessment of genetic variation and structure among *Pinus laricio* Poiret populations within their native area. *Plant Biosyst* 150:1271–1281
- Cabo S, Ferreira L, Carvalho A et al (2014) Potential of Start Codon Targeted (SCoT) markers for DNA fingerprinting of newly synthesized tritordeums and their respective parents. *J Appl Genet* 55:307–312
- Candel Pérez D (2014) *Pinus nigra* Arn. ssp. *salzmannii* forest management in the context of climate change: ecological and genetic factors. Tesis Doctoral, Universidad de Castilla-La Mancha, Albacete
- Carvalho A, Matos M, Lima-Brito J et al (2005) DNA fingerprint of F1 interspecific hybrids from the Triticeae tribe using ISSRs. *Euphytica* 143:93–99
- Christensen K (1997) Pinaceae, Cupressaceae, Taxaceae, Ephedraceae, Salicaceae, Juglandaceae, Betulaceae, Fagaceae, Ulmaceae Moraceae. In: Strid A, Tan K (eds) *Flora Hellenica*. Koeltz Scientific Books, Königstein
- Cipriano J, Carvalho A, Fernandes C et al (2016) Evaluation of genetic diversity of Portuguese *Pinus sylvestris* L. populations based on molecular data and inferences about the future use of this germplasm. *J Genet* 93(2):41–48. <https://doi.org/10.1007/s12041-013-0241-3>
- Collard BCY, Mackill DJ (2009) Start Codon Targeted (SCoT) Polymorphism: a simple, novel DNA marker technique for generating gene-targeted markers in plants. *Plant Mol Biol Rep* 27:86–93
- Dias A, Gaspar MJ, Carvalho A et al (2018) Within- and between-tree variation of wood density components in *Pinus nigra* at six sites in Portugal. *Ann For Sci* 75:58. <https://doi.org/10.1007/s13595-018-0734-6>
- Doyle JJ, Doyle JL (1987) A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochem Bull* 19:11–15
- Earl DA, vonHoldt BM (2012) STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conserv Genet Resour* 4:359–361
- 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(2):479–491
- Falush D, Stephens M, Pritchard JK (2007) Inference of population structure using multilocus genotype data: dominant markers and null alleles: technical article. *Mol Ecol Notes* 7:574–578
- Farjon A (1998) World checklist and bibliography of conifers. Royal Botanic Gardens, Kew
- Farjon A (2010) A handbook of The World's Conifers, vol 1. Brill Leiden, Boston
- Felsenstein J (1985) Confidence limits on phylogenies: an approach using the bootstrap. *Evolution* 39:783–791
- Gaussen H, Heywood VH, Chater AO (1964) The genus *Pinus* L. In: Tutin TG, Heywood VH, Burgers NA, Valentine DH, Walters SM, Webb DA (eds) *Flora Europaea*, vol 1. Cambridge University Press, Cambridge, pp 32–35
- Giovannelli G, Roig A, Spanu I et al (2017) A new set of nuclear microsatellites for an ecologically and economically important conifer: the European Black Pine (*Pinus nigra* Arn.). *Plant Mol Biol Rep* 35(3):379–388
- Hamrick JL, Godt MJW, Sherman-Broyles SL (1992) Factors influencing levels of genetic diversity in woody plant species. *New For* 6:95–124
- Hedges SB (1992) The number of replications needed for accurate estimation of the bootstrap p value in phylogenetic studies. *Mol Biol Evol* 9:366–369
- Liber Z, Nikolic T, Mitic B et al (2003) RAPD markers and black pine (*Pinus nigra* Arnold) intraspecific taxonomy-evidence from the study of nine populations. *Acta Soc Bot Pol* 72:249–257
- Little E, Critchfield W (1969) Subdivisions of the genus *Pinus* (Pines). U.S. Department of Agriculture Forest Services, Miscellaneous Publications, Washington
- Louro V (1982) O pinheiro larício: *Pinus nigra* Arnold em Portugal. Direcção Geral do Ordenamento e Gestao Florestal, Lisboa
- Lucas Borja ME, Candel Pérez D, Molero Carrasco J et al (2013) Utilización de los marcadores ISSR para la determinación de diversidad genética en poblaciones de *Pinus nigra* Arn.. 6º Congreso Forestal Español, 10-14 junio 2013, Vitoria-Gasteiz, pp 1–10. ISBN: 978-84-937964-9-5
- Naydenov KD, Tremblay FM, Fenton NJ et al (2006) Structure of *Pinus nigra* Arn. populations in Bulgaria revealed by chloroplast microsatellites and terpenes analysis: provenance tests. *Biochem Syst Ecol* 34:562–574
- Peakall R, Smouse PE (2006) GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. *Mol Ecol Notes* 6:288–295
- Peakall R, Smouse PE (2012) GenAIEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research—an update. *Bioinformatics* 28:2537–2539
- Price RA, Liston A, Strauss SH (1998) Phylogeny and systematics of *Pinus*. In: Richardson DM (ed) *Ecology and biogeography of Pinus*. Cambridge University Press, London, pp 49–68
- Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics* 155:945–959
- Reis E (1997) Estatística multivariada aplicada. Edições Sílabo, Lda., Lisboa
- Rohlf FJ (1998) NTSYS-pc ver. 2.02. Numerical taxonomy and multivariate analysis system. Exeter Publishing, Setauket
- Rubio Moraga A, Candel Pérez D, Lucas Borja ME et al (2012) Genetic diversity of *Pinus nigra* Arn. populations in Southern Spain and Northern Morocco revealed by Inter-Simple Sequence Repeat profiles. *Int J Mol Sci* 13:5645–5658
- Šarac Z, Aleksic J, Dodos T et al (2015) Cross-species amplification of nuclear EST-microsatellites developed for other *Pinus* species in *Pinus nigra*. *Genetika* 47:205–217
- Savolainen O, Pyhäjärvi T, Knürr T (2007) Gene flow and local adaptation in trees. *Annu Rev Ecol Evol Syst* 38:595–619
- Scaltsoyiannes A, Rohr R, Panetsos KP, Tsaktsira M (1994) Allozyme frequency distributions in 5 European populations of black pine (*Pinus nigra* Arnold). 1. Estimation of genetic variation within and among populations. 2. Contribution of isozyme analysis to the taxonomic status of the species. *Silvae Genet* 43:20–30
- Yap I, Nelson R (1996) Winboot: a program for performing bootstrap analysis of binary data to determine the confidence limits of

- UPGMA-based dendrograms. International Rice Research Institute (IRRI), Manila
- Yeh FC, Yang RC, Boyle TB et al (1999) POPGENE version 1.32, the user-friendly shareware for population genetic analysis. Molecular Biology and Biotechnology Centre, University of Alberta, Edmonton
- Zhang J, Xie W, Wang Y et al (2015) Potential of Start Codon Targeted (SCoT) markers to estimate genetic diversity and relationships among Chinese *Elymus sibiricus* accessions. Molecules 20:5987–6001
- Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.