Annals of Forest Science (2020) 77: 64
https://doi.org/10.1007/s13595-020-00967-9

RESEARCH PAPER

Portuguese *Pinus nigra* J.F. Arnold populations: genetic diversity, structure and relationships inferred by SSR markers

Alexandra Dias[1] · Guia Giovannelli[2] · Bruno Fady[2] · Ilaria Spanu[3] · Giovanni G Vendramin[3] · Francesca Bagnoli[3] · Ana Carvalho[1,4,5] · Maria Emília Silva[1,6] · José Lima-Brito[1,4,5] · José Luís Lousada[1,6] · Maria João Gaspar[4,5,7]

Received: 9 June 2019 / Accepted: 19 May 2020 / Published online: 19 June 2020

Abstract

• ***Key message*** ***Pinus nigra*** **J.F. Arnold has an ecological and economic interest in Europe, but many of the planted populations have an unknown origin and infraspecific taxonomy. Six Portuguese *P. nigra* populations characterised with microsatellites revealed high intra-population diversity structured into two clusters with low differentiation that might suggest two provenances or infraspecific taxa. Despite compared with foreign samples from different subspecies, we were not able to infer about the origin or infraspecific taxonomy of the Portuguese populations based on the pooled microsatellite data.**

• ***Context*** Many of the European *Pinus nigra* J.F. Arnold forests were planted with the material of unknown origin. Efforts have been made to determine their provenances and infraspecific taxonomy regarding their relevance for defining strategies of genetic resources conservation, germplasm use, forest management and genetic improvement. The Portuguese *P. nigra* populations are allochthonous, and their provenance and infraspecific taxonomy are unknown.

• ***Aims*** With this work, we intended to characterise the Portuguese *P. nigra* populations regarding its genetic diversity, structure and relationships using simple sequence repeat (SSR) markers and/or to infer about its origin and infraspecific taxonomy by comparing the SSR patterns of foreign *P. nigra* samples of known taxonomic classification.

• ***Methods*** A total of 224 Portuguese *P. nigra* individuals from six populations were characterised using 14 SSR markers developed specifically to different *Pinus* sp., including *P. nigra*, by other authors.

• ***Results*** Thirteen SSR loci were selected and showed 100% of polymorphism among individuals. The genetic diversity was higher within (95%) rather than among (5%) populations. The Portuguese individuals were structured into two main clusters (K = 2) with low genetic differentiation (F_{ST} = 0.04), and the foreign samples were genetically distant from the Portuguese individuals.

• ***Conclusion*** The six Portuguese *P. nigra* populations revealed high genetic diversity and seemed to be structured into two main clusters with low differentiation suggesting two provenances or infraspecific taxa. Nonetheless, the comparative analyses with foreign samples did not allow a clear inference about its origin and/or infraspecific taxonomy. Additional foreign samples and/or molecular markers will be tested to pursue these goals.

Keywords European Black Pine · Microsatellites · Polymorphism · Simple sequence repeats (SSRs)

Handling Editor: Ricardo Alia

✉ Maria João Gaspar
mjgaspar@utad.pt

Alexandra Dias
xana_dias@hotmail.com

Guia Giovannelli
guiagiovannelli@gmail.com

Bruno Fady
bruno.fady@inra.fr

Ilaria Spanu
ilaria.spanu@ibbr.cnr.it

Giovanni G Vendramin
gg.vendramin@gmail.com

Francesca Bagnoli
francesca.bagnoli@ibbr.cnr.it

Ana Carvalho
anacar@utad.pt

Maria Emília Silva
emil_ms@utad.pt

José Lima-Brito
jbrito@utad.pt

José Luís Lousada
jlousada@utad.pt

Extended author information available on the last page of the article

1 Introduction

Pinus nigra J.F. Arnold is one of the leading commercial important conifers of the Mediterranean region (Bonavita et al. 2016). *P. nigra* has been widely used for reforestation of difficult soils with hard climatic conditions (Portoghesi et al. 2013). This Mediterranean pine has an extensive and scattered distribution through Northern Africa, Northern Mediterranean, eastwards to the Black Sea, Corsica and Sicily islands (Vidaković 1991; Gaussen et al. 1993; Afzal-Rafii and Dodd 2007; Olsson et al. 2020). The populations of *P. nigra* studied by various authors presented high variation at different levels (Scaltsoyiannes et al. 1994; Rafii et al. 1996; Quézel and Médail 2003; Bogunić et al. 2003, 2007, 2009; Afzal-Rafii and Dodd 2007; del Cerro et al. 2009; Rubio-Moraga et al. 2012; Bonavita et al. 2016). Due to their geographically fragmented distribution areas, a reduced intra-population genetic diversity for these *P. nigra* populations is expected (Scotti-Saintagne et al. 2019). Nonetheless, recent molecular studies revealed a weak genetic spatial structure originated from events that occurred in the late Pleistocene or early Holocene (Giovannelli et al. 2019; Scotti-Saintagne et al. 2019). These events fragmented one ancestral *P. nigra* population into six to seven distinct genetic lineages with high gene flow among them (Giovannelli et al. 2019; Scotti-Saintagne et al. 2019). The same authors proposed the revision of the infraspecific taxonomy of *P. nigra* species on five subspecies with Mediterranean distribution based on the length and stiffness of needles: (i) *salzmannii* (Dunal) Franco; (ii) *nigra* Arnold; (iii) *dalmatica* (Visiani) Franco; (iv) *pallasiana* (Lambert) Holmboe and (v) *laricio* (Poiret) Maire (Gaussen et al. 1993), and a sixth subspecies, *P. nigra* subsp. *mauretanica* (Maire and Peyerimh) Heywood assigned to the North of Africa (Bussotti 2002). Scotti-Saintagne et al. (2019) proposed to name the recently discovered six lineages of *P. nigra* using regionally accepted subspecies-level names. Additionally to the six main inter-fertile *P. nigra* subspecies agreed so far (Quézel and Médail 2003), some of those are also divided into two to four regional varieties (Christensen 1997; Barbéro et al. 1998; Price et al. 1998; Bussotti 2002). The *P. nigra* subsp. *salzmannii* can present the varieties *mauretanica*, *hispanica*, *salzmannii* and *cebennensis*; the *P. nigra* subsp. *nigra* can be subdivided into the varieties *austriaca*, *illyrica*, *pindica* and *italica*; the *P. nigra* subsp. *pallasiana* contains the varieties *banatica*, *tatarica*, *caramanica* and *fenzlii*; and the *P. nigra* subsp. *laricio* includes the varieties *calabrica* or *corsicana* (Christensen 1997; Barbéro et al. 1998; Price et al. 1998; Bussotti 2002).

Beyond the taxonomic issues, most of the European *P. nigra* populations have an unknown origin and/or infraspecific taxonomy. This information is of utmost importance under the scope of afforestation/reforestation programs regarding the existence of multiple inter-fertile subspecies of *P. nigra* (Quézel and Médail 2003; Naydenov et al. 2006; Zaghi et al. 2008). Besides, the information about intra- and inter-population, genetic diversity is also relevant for the maintenance of the stability of forest ecosystems, the adaptive potential to biotic and abiotic stresses (Muller-Starck et al. 1992; Çengel et al. 2012), and to assign high priority populations that deserve to be preserved and managed (Bonavita et al. 2016). Knowledge of genetic diversity can give insights about the adaptive potential and/or the evolutionary history of a species, being particularly crucial in trees with a long life cycle (Hamrick 2004; Giovannelli et al. 2017). The gathering of genetic diversity and structure information of forestry populations is helpful for the development of suited and effective strategies of conservation, management, germplasm use and genetic breeding (Frankham et al. 2002; Hamrick 2004; Lučić et al. 2010). Different molecular marker systems have been performed in various pine species, including *P. nigra*, to assess the genetic diversity, provenances, genetic structure and relationships, and/or to discriminate closely related *taxa* (Vendramin et al. 1996; Liber et al. 2003; Jiménez et al. 2005; Naydenov et al. 2006; Afzal-Rafii and Dodd 2007; Soto et al. 2010; Rubio-Moraga et al. 2012; Liu et al. 2012; Šarac et al. 2015; Bonavita et al. 2016; Cipriano et al. 2016; Giovannelli et al. 2017; Dias et al. 2019).

The nuclear SSRs (nSSRs) are inherited biparentally, evolve faster than the chloroplast SSRs (cpSSRs) and are more polymorphic than the Expressed Sequence Tag-SSRs (EST-SSRs) (Wolfe et al. 1987; Mogensen 1996; Cho et al. 2000). The EST-SSRs generally consist of GC-rich trinucleotide repeats within transcribed genomic regions and their polymorphism is associated with gene diversity (Varshney et al. 2005). Since the EST-SSRs have few null alleles, they can be used to discriminate closely related taxa (Ellegren 2004; Varshney et al. 2005; Hayden et al. 2008).

The wide and fragmented distribution of *P. nigra*, gene flow, weak spatial genetic structure and admixed origin of the populations contributed to taxonomic classification issues (Giovannelli et al. 2019). Recent studies based on simple sequence repeat (SSR) markers were developed to resolve the complex taxonomy and evolutionary history of *P. nigra* in Europe (Giovannelli et al. 2017, 2019). These authors developed for the first time specific nSSRs to *P. nigra* but also revealed that nSSRs and EST-SSRs from other pine species are transferable to *P. nigra*.

With this work is intended to test a set of 14 nSSRs, which include nine nSSRs specifically developed in *P. nigra* by Giovannelli et al. (2017), three nSSRs specific to *Pinus sylvestris* and *Pinus taeda*, and two EST-SSRs specific to *Pinus halepensis* (Soranzo et al. 1998; Elsik and Williams 2001; Zhou et al. 2002; Leonarduzzi et al. 2016b) in 224 *P. nigra* individuals from six Portuguese populations, to evaluate the (i) intra- and inter-population genetic diversity; (ii) genetic structure and relationships and (iii) to infer about the

infraspecific taxonomy of the Portuguese populations by comparing with foreign samples with known taxonomy.

2 Materials and methods

2.1 Plant material and sampling

The distribution of *P. nigra* in Portugal is restricted to pure even-aged planted and managed stands located at the North and Centre of the country, mainly in mountain regions whose altitude ranges from 450 to 1600 m (Dias et al. 2019). The six *P. nigra* populations were chosen based on the species distribution in Portugal: Minho (NW), Trás-os-Montes (N) and Beira (Central). The previous dendrometric characterisation of these populations suggested an average age ranging from 57.8 to 93.3 years old (Dias et al. 2014, 2018; Table 1). The planted areas range from 5 to 40 ha. Two sample plots in each region were established.

In each population, individuals were randomly selected within plots with an average of 0.04 ha. A total of 224 individuals were sampled for needles or differentiating xylem (in the tallest trees) (Table 1).

The plant material was immediately frozen in liquid nitrogen and maintained at − 80 °C until the extraction of genomic DNA.

A total of 30 foreign samples of *P. nigra* (six samples per subspecies) with known infraspecific taxonomy and provenance, and previously genotyped with SSRs by Giovannelli et al. (2017), were included in this study (Table 2) for comparison with the Portuguese individuals.

2.2 DNA extraction and SSRs amplification

Frozen needles or differentiating xylem (250 mg) were grounded to a fine powder in the presence of liquid nitrogen. Genomic DNA extraction was performed with a CTAB-based protocol (Doyle and Doyle 1987). The genomic DNA integrity was verified after electrophoresis on 0.8% agarose gels and the DNA quantification was performed in the spectrophotometer Nanodrop ND-1000® (Thermo Scientific, Burlington, USA).

The SSR markers were amplified in the 224 *P. nigra* individuals using the primers and mixtures indicated in Table 3.

The same set of 14 SSR markers of Giovannelli et al. (2017) was used, but the procedure in terms of the final volume of the reaction mixture and amplification conditions was changed. Briefly, the primer mixes 1 and 2 were performed for a final volume of 50 μL. Each multiplex PCR contained 1–2 μL of genomic DNA (10 ng/μL), 0.6 μL of primer mix, 3 μL of QIAGEN Multiplex PCR Master Mix (Qiagen, Germany) and 1.4 μL of RNase-free water, for a total volume of 6 μL. For both multiplexes, the PCR thermal profile was the following: an initial step at 95 °C for 5 min, followed by 32 cycles at 95 °C for 30 s, 57 °C for 90 s and 72 °C for 30 s, and a final extension of 60 °C for 30 min. The amplified SSR loci labelled with fluorescent dyes were separated by capillary electrophoresis using the ABI 3500 automatic sequencer (Applied Biosystems, Foster City, CA, USA), using LIZ 500 as an internal size standard. The chromatograms were analysed with the GeneMarker software version 2.7.0 (SoftGenetics, USA).

2.3 Statistical analyses of the SSRs data

Alleles were sized manually to reduce the common allele calling and binning errors reported by Guichoux et al. (2011).

The software POPGENE 1.32 (Yeh et al. 1999) was used to calculate the (i) number of observed alleles (na); (ii) number of effective alleles (ne; Kimura and Crow 1964); (iii) Shannon's information index (*I*; Lewontin 1972) that measures gene diversity (Shannon and Weaver 1949); (iv) observed heterozygosity (Ho) and expected heterozygosity (He) (Levene 1949; Nei 1973); (v) Nei's gene diversity index (Nei 1973) (*h*); (vi) estimation of the unbiased genetic identity (Nei 1973) and genetic distance (Nei 1978) and (vii) inter-population genetic variation (D_{ST}).

The software STRUCTURE 2.3.4 (Pritchard et al. 2000, 2003, 2009; Falush et al. 2003, 2007) was used for genetic structure evaluation using the admixture model with correlated allele frequencies among clusters (suitable for codominant markers). For the Portuguese populations, ten independent runs for each *K* (from 1 to 10) were performed with 100,000 generations of a burn-in period followed by 500,000 Markov Chain Monte Carlo (MCMC) iterations. Additional genetic structure analyses were

Table 1 Portuguese populations of *P. nigra*: respective geographic coordinates, average age and number of individuals sampled per population

P. nigra population	Coordinates (latitude; longitude)	Age (years)	Number of sampled individuals
Manteigas	40° 22′ 47.00″ N; 7° 33′ 18.00″ W	93.3	37
Vale do Zêzere	40° 19′ 19.00″ N; 7° 34′ 26.00″ W	59.1	37
Paredes de Coura	41° 52′ 0.00″ N; 8° 36′ 21.00″ W	57.8	36
Caminha	41° 50′ 15.00″ N; 8° 43′ 57.00″ W	57.9	41
Campeã	41° 19′ 9.12″ N; 7° 53′ 28.35″ W	58.1	35
Vila Pouca de Aguiar	41° 31′ 02.72″ N; 7° 35′ 31.36″ W	74.7	38

Table 2 Foreign samples of *P. nigra* with known infraspecific taxonomy, provenance and geographic coordinates of the sampling sites

Taxonomic classification	Country	Latitude	Longitude
Pinus nigra subsp. *laricio* (Poir.) Maire	Italy	38° 14′ N	16° 00′ E
Pinus nigra subsp. *dalmatica* (Visiani) Franco	Croatia	42° 52′ 07″ N	17° 33′ 03″ E
Pinus nigra subsp. *salzmannii* (Dunal) Franco	Algeria	36° 27′ N	4° 06′ E
Pinus nigra subsp. *pallasiana* (Lamb.) Holmboe	Cyprus	34° 57′ N	32° 50′ E
Pinus nigra subsp. *salzmannii* (Dunal) Franco	France	43° 46′ N	3° 34′ E
Pinus nigra subsp. *nigra* J.F. Arnold	Romania	44° 52′ 57″ N	22°25′ 40″ E
Pinus nigra subsp. *pallasiana* (Lamb.) Holmboe	Ukraine	44° 25′ 39″ N	34° 03′ 05″ E

performed with the foreign samples and/or the Portuguese *P. nigra* populations, using ten independent runs for each *K* (from 1 to 20) and the same values of burn-in period and MCMC iterations. The STRUCTURE outputs of the previous analyses were summarised and tested with the software STRUCTURE Harvester (Earl and vonHoldt 2012) that quickly provided the results of the Evanno method concerning the determination of the optimal number of clusters (*K*) for the studied individuals without considering predefined populations (Evanno et al. 2005).

The software GenAlEx v6.5 (Peakall and Smouse 2012) was used to perform the analysis of molecular variance (AMOVA) and the principal coordinate analysis (PCoA) based on Nei's genetic distance pairwise population matrix (Nei 1978). The GenAlex also allowed the calculation of the (i) fixation index (*F*); (ii) Wright's *F* statistics inbreeding coefficient (FIS) (Holsinger and Weir 2009); (iii) overall inbreeding of an individual relative to the total populations (FIT); (iv) proportion of total genetic differentiation (F_{ST}); (v) Gst: coefficient of gene differentiation (analogue of FST) adjusted for bias; and (vi) differentiation index Jost's *D* that measures the relative degree of differentiation of allele frequencies (Jost 2009).

Table 3 nSSR and EST-SSR *loci*, respective primer sequences, repeat motif, reference and labelling

Multiplex	SSR *locus*	Type	Primer sequences (5′–3′)	Repeat motif	Reference	Labelling
Primer mix 1	SPAG7.14	nSSR	F:TTCGTAGGACTAAAAATGTGTG R:CAAAGTGGATTTTGACCG	$(TG)_{17}$ $(AG)_{21}$	Soranzo et al. (1998)	VIC
	pn6360	nSSR	F:CAACTTCCTTCACCTGGCAC R:GACCCTTCTCAGCATCAACAC	$(TC)_{20}$	Giovannelli et al. (2017)	6-FAM
	pn7754	nSSR	F:TCAAGCTAATGCTGGGAACTC R:GTTGGTCGTGGTGACAAAGG	$(TA)_{12}$		6-FAM
	pn2153	nSSR	F:TTGTGGGCTTAACTCGTCAG R:TGGCATTTGCTTCTCATTTG	$(TG)_{12}$		VIC
	pn4379	nSSR	F:ACAGGCACACAATTTGGTCC R:CATCATCTGCAGAAGGGAAGAC	$(AT)_{15}$		VIC
	pn2246	nSSR	F:TCATGTCCAGCTTTGGTTGC R:AAGGTTCATTTCCACTCCGG	$(TA)_{13}$		PET
	pn6175	nSSR	F:ATTTCCCGCCTACCATTACC R:GTTACCTGCAATTCGTGTGG	$(AT)_{12}$		NED
	pn1403	nSSR	F:AGGCAATGATCATGTGGGTC R:GGATAGCCTGCAACTCCAATG	$(GA)_{13}$		NED
Primer mix 2	PHA_6062	EST-SSRs	F:CTCTCTTTCTCTTCGGTACTCA R:AAGTTTCAGAGGATTCACACTC	$(TGC)_8$	Leonarduzzi et al. (2016b)	6-FAM
	PHA_4783	EST-SSRs	F:AATCTTCTCTTTCTTTGACTGTGT R:CATGGGAATGTTAGGAGTGC	$(TCC)_8$		VIC
	pn8747	nSSR	F:GCATTCTTTCCTTGCTCTCG R:CTCATCTGCTCCCTCAGTCC	$(AT)_{12}$	Giovannelli et al. (2017)	VIC NED
	pn6266	nSSR	F:TGCCTTGCTGATTTATTTCCG R:CACTTTCAGCTACGACCGAC	$(GT)_{14}$		NED PET
	PtTX4001	nSSR	F:CTATTTGAGTTAAGAAGGGA GTC R:CTGTGGGTAGCATCATC	$(GT)_{15}$	Zhou et al. (2002)	VIC
	PtTX3107	nSSR	F:AAACAAGCCCACATCGTCAATC R:TCCCCTGGATCTGAGGA	$(CAT)_{14}$	Elsik and Williams (2001)	PET

3 Results

3.1 SSR polymorphism

The 14 SSR loci were amplified in the 224 Portuguese *P. nigra* individuals. However, the SSR locus pn2153 was discarded from this study due to the production of low-quality results reflected by unspecific amplifications. Each SSR locus was polymorphic among the 224 *P. nigra* individuals studied (Dias et al. 2020).

The size of the amplified SSR fragments ranged from 99 to 476 bp, matching the expected size, except for the SSR loci PHA_6062, PtTX4001 and PtTX3107 (Table 4), without affecting the binning precision.

3.2 Genetic diversity and relationships

The average number of observed alleles (na) amplified per population with the 13 selected SSR markers ranged from 8.46 to 12.46, with a total mean of 11.32 for the six populations (Table 5).

The lowest na value was detected in locus SSR PHA_4783 in all populations (Table 7). On the other hand, the highest na value was verified in the SSR locus pn6175 in Vale do Zêzere (Table 7). The total mean number of na (11.32) was higher than the total mean of the effective number of alleles (ne), which was 6.00 (Table 5).

The number of alleles amplified per locus with the eight SSRs specifically developed to *P. nigra* ranged from 5 to 25 (Table 7). Besides, the number of alleles amplified per locus with the SSR primers that were transferable from other species ranged from 2 to 22 in the Portuguese *P. nigra* samples (Table 7).

Table 4 Expected and observed range of fragment size (bp) per SSR *locus*

SSR *locus*	Expected fragment size (bp)	Range of fragment size (bp)
pn6360	324	265–345
pn7754	114	99–145
SPAG7.14	209	189–257
pn4379	450	402–450
pn2246	359	348–410
pn6175	201	190–279
pn1403	307	294–340
PHA_6062	367	446–476
PHA_4783	472	470–473
PtTX4001	224	196–218
pn8747	278	273–295
PtTX3107	182	153–171
pn6266	317	301–329

The highest average values of na and ne were registered in the populations of Campeã and Paredes de Coura, respectively (Table 5).

The populations of Vale do Zêzere and Paredes de Coura presented the highest average values of Shannon's information index (*I*), Nei's gene diversity index (*h*) and expected heterozygosity (He). In contrast, Vila Pouca de Aguiar showed the lowest means for the same parameters (Table 5).

The highest value of observed heterozygosity (Ho) was shown by Paredes de Coura, while the lowest one was observed in Manteigas (Table 5).

The fixation index (*F*) revealed its maximum value in Manteigas, and its minimum in Vila Pouca de Aguiar (Table 5).

The pooled SSR data was used to calculate the Nei's unbiased measures of genetic identity and genetic distance among the six *P. nigra* populations (Table 6).

The population of Manteigas presented high values of Nei's unbiased measures of genetic identity with all populations except with Paredes de Coura (0.7902) (Table 6). These results were partially corroborated in the PCoA since higher proximity among Vale do Zêzere and the populations of Manteigas, Campeã and Caminha was expected (Fig. 1). The PCoA demonstrated that Campeã, Manteigas and Caminha are the most genetically proximal populations (Fig. 1).

The cumulative percentage of total variation explained by the first three coordinates of the PCoA represented in Fig. 1 was 96.46%.

3.3 Population differentiation and genetic structure

The statistical analysis of genetic variation and gene diversity estimated within and among the studied Portuguese *P. nigra* populations was performed with the calculation of the inter-population genetic diversity (D_{ST}), with a value of 0.122 and the proportion of total genetic differentiation (F_{ST}) with 0.040 ± 0.005.

The genetic differentiation among populations (*G*st) and the differentiation index Jost's *D* that measures the relative degree of differentiation of allele frequencies (Jost 2009) were estimated per SSR locus (Table 8). Overall, the 13 SSR markers estimated a reduced average genetic differentiation (*G*st = 0.025) among the six Portuguese populations of *P. nigra* (Table 8).

The bar plot generated after the STRUCTURE analysis performed to the six Portuguese *P. nigra* populations evidenced a different pattern in the populations of Paredes de Coura and Vila Pouca de Aguiar relatively to the four remaining ones (Fig. 2a). This analysis also indicated that the optimal number of genetic clusters was *K* = 2 (Fig. 2b).

In the bar plot achieved with the STRUCTURE analysis that combined the Portuguese and the foreign *P. nigra* samples, the former presented a common pattern that

Table 5 Genetic diversity statistics assayed per Portuguese *P. nigra* population. Mean (± standard deviation, s.d.) values of na, ne, *I*, *h*, H_o, H_e and *F* per SSR locus are presented. Notes: na, observed number of alleles; ne, effective number of alleles (Kimura and Crow 1964); *I*, Shannon's Information Index (Lewontin 1972); *h*, Nei's gene diversity index (Nei 1973); H_o, observed heterozygosity; H_e, expected heterozygosity (Levene 1949) and *F*, fixation index

Population	Mean values (± s.d.) of:						
	na	ne	*I*	*H*	H_o	H_e	*F*
Manteigas	11.54 ± 1.66	5.44 ± 0.78	1.84 ± 0.18	0.75 ± 0.05	0.54 ± 0.05	0.76 ± 0.05	0.270 ± 0.050
Vale do Zêzere	12.31 ± 1.78	6.86 ± 1.24	1.94 ± 0.20	0.76 ± 0.06	0.59 ± 0.06	0.77 ± 0.06	0.22 ± 0.06
Campeã	12.46 ± 1.81	6.55 ± 1.21	1.91 ± 0.21	0.75 ± 0.05	0.59 ± 0.05	0.76 ± 0.06	0.19 ± 0.07
Paredes de Coura	12 ± 1.70	6.80 ± 1.14	1.94 ± 0.21	0.76 ± 0.06	0.62 ± 0.07	0.77 ± 0.06	0.18 ± 0.05
Vila Pouca de Aguiar	8.46 ± 1.14	4.76 ± 0.71	1.61 ± 0.17	0.70 ± 0.05	0.61 ± 0.06	0.71 ± 0.05	0.13 ± 0.06
Caminha	11.15 ± 1.54	5.61 ± 0.91	1.81 ± 0.18	0.75 ± 0.05	0.60 ± 0.06	0.75 ± 0.05	0.18 ± 0.05
Total mean	11.32 ± 0.66	6.00 ± 0.41	1.84 ± 0.08	0.74 ± 0.02	0.74 ± 0.02	0.76 ± 0.02	0.20 ± 0.02

distinguished them from the foreign ones (Fig. 2c). Therefore, for this genetic structure analysis, the inferred number of clusters was also $K = 2$ (Fig. 2d).

The average values of total genetic differentiation proportion (F_{ST}) estimated for the two clusters of the Portuguese populations were reduced ($F_{ST}\_1 = 0.0867$ and $F_{ST}\_2 = 0.0711$). A reduced proportion of total genetic differentiation ($F_{ST} = 0.04$) was also estimated with the Genalex software. Hence, the F_{ST} values obtained with these two analyses suggested a low genetic differentiation ($F_{ST} < 0.25$) between the two clusters (*K*) (Fig. 2b) and among the six Portuguese populations.

3.4 Extrapolation about the infraspecific taxonomy

In order to extrapolate about the infraspecific taxonomy of the six Portuguese *P. nigra* populations, their SSR profiles were compared with others previously achieved by Giovannelli et al. (2017) in foreign *P. nigra* samples with known subspecies and provenances. After performing a PCoA analysis based on the Nei (1978) genetic distance matrix, the Portuguese *P. nigra* populations were clustered apart from the seven foreign samples (Fig. 3). The number of groups observed in the PCoA was corroborated by the STRUCTURE analysis of the same samples that also inferred two clusters ($K = 2$) (Fig. 2d). Nonetheless, these two clusters showed reduced genetic differentiation values, $F_{ST}\_1 = 0.099$ and $F_{ST}\_2 = 0.2018$. A reduced proportion of total genetic differentiation ($F_{ST} = 0.22$) among the Portuguese and the foreign samples was also inferred with the GenAlex software.

4 Discussion

Nowadays, the six populations of *P. nigra* located at the North and Centre of Portugal are representative of the species distribution in our country and show high adaptive potential, as demonstrated previously by their dendrometric and ecological evaluations (Dias et al. 2014, 2018).

The first molecular characterisation of the Portuguese *P. nigra* populations was performed with dominant molecular markers, namely, inter-simple sequence repeats (ISSRs) and Start Codon Targeted (SCoT) markers (Dias et al. 2019). The authors detected high intra-population polymorphism, good genetic differentiation and structure among populations as well as high genetic similarity with foreign samples belonging

Table 6 Nei's unbiased measures of genetic identity (above diagonal) and genetic distance (below diagonal) among the Portuguese populations of *P. nigra* based on the pooled SSR data

Population	Manteigas	Vale do Zêzere	Campeã	Paredes de Coura	Vila Pouca de Aguiar	Caminha
Manteigas	–	0.8859	0.9328	0.7902	0.8320	0.9503
Vale do Zêzere	0.1211	–	0.9098	0.9032	0.8107	0.8821
Campeã	0.0695	0.0945	–	0.8254	0.8305	0.9349
Paredes de Coura	0.2354	0.1018	0.1919	–	0.8033	0.7948
Vila Pouca de Aguiar	0.1839	0.2099	0.1858	0.2190	–	0.8299
Caminha	0.0510	0.1255	0.0673	0.2296	0.1864	–

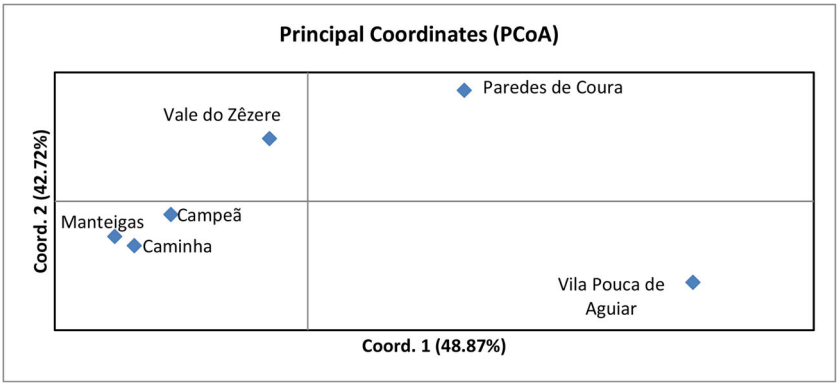

Fig. 1 PCoA showing the projection of the six Portuguese populations of *P. nigra* based on the Nei's genetic distance matrix calculated with Nei's coefficient (Nei 1978) and the pooled SSR data

to the subspecies *laricio* var. *calabrica* (according to SCoTs) and subsp. *laricio* var. *corsicana* (based only in ISSRs) (Dias et al. 2019).

In the present work, additional molecular characterisation of the same Portuguese populations using the codominant SSR markers was performed. A set of 14 SSR markers was tested including SSRs that were developed specifically to *P. nigra* by Giovannelli et al. (2017) as well as SSRs that were specific to other pine species (Soranzo et al. 1998; Elsik and Williams 2001; Zhou et al. 2002; Leonarduzzi et al. 2016b). Among the tested SSR markers, 13 revealed to be polymorphic and amplified successfully in the Portuguese *P. nigra* samples, confirming the transferability of SSRs markers developed specifically in *Pinus sylvestris* C. Linnaeus, *Pinus halepensis* P. Miller and *Pinus taeda* C. Linnaeus (Soranzo et al. 1998; Elsik and Williams 2001; Zhou et al. 2002; Leonarduzzi et al. 2016b) to *P. nigra*, corroborating the assumption of Giovannelli et al. (2017). Šarac et al. (2015) also studied the cross-transferability of SSRs markers developed in other species to *P. nigra*. These authors obtained successful

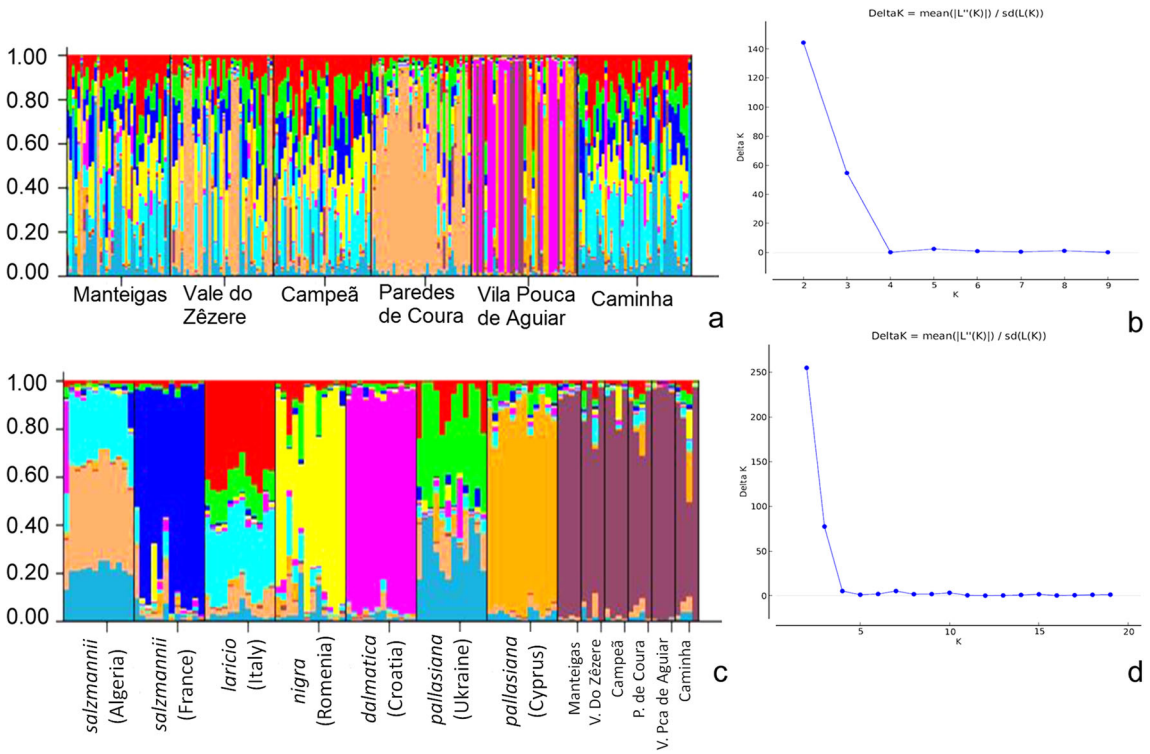

Fig. 2 Results of STRUCTURE analyses performed with the six Portuguese *P. nigra* populations (**a**, **b**) and/or the foreign *P. nigra* samples belonging to different subspecies (**c**, **d**). **a**, **c** Proportion of the membership coefficient for each individual for the inferred clusters when $K = 2$, according to STRUCTURE results. Each individual is represented by a thin vertical line, which is partitioned into coloured segments that indicate the individual's estimated membership fraction in $K = 2$ clusters. Different populations are separated by vertical lines. **b**, **d** Plot of the delta *K* value, calculated according to Evanno et al. (2005) and plotted against the number of the modelled clusters

amplification but also reduced the level of polymorphism or unspecific loci amplification. Here, the selected 13 SSR loci demonstrated to be polymorphic among the 224 *P. nigra* individuals similar to the results previously achieved by Giovannelli et al. (2017).

The SSR study revealed high genetic variability in the six *P. nigra* populations, with average diversity values for the six populations of 1.8416 ± 0.76 (*I*) and 0.74 ± 0.0209 (*h*). The maximum values of the Shannon index and Nei's gene diversity were found in Paredes de Coura and Vale do Zêzere populations. In contrast, the lowest ones were observed in Vila Pouca de Aguiar. Lower results were achieved by Jiménez et al. (2005) for the same species with Nei's gene diversity of 0.425 to 0.558, and by Naydenov et al. (2006) that reported values of 0.008 to 0.195, respectively. Concerning to the Shannon index presented by other species such as *Pinus pinaster* W. Aiton, Naydenov et al. (2014) also referred lower results (0.756) than those shown in this work.

In respect to the total average number of observed (na) and effective alleles (ne), values of 11.32 ± 5.78 and 6.00 ± 3.60, respectively, were achieved. In *P. nigra*, Bonavita et al. (2016) observed a lower value of na (5.9) but a higher value of ne (13.9).

Regarding the expected heterozygosity (H_e), a high number of alleles were detected, estimating high genetic diversity in this species (H_e = 0.75 ± 0.19). Reduced values of H_e (0.123 to 0.242) were found in Spanish populations of *P. nigra* using ISSR markers (Rubio-Moraga et al. 2012). The authors explained the reduced genetic diversity as the effect of previous forest management. Similarly, a reduced H_e value (0.183) was achieved with EST-SSRs in Balkan populations of *P. nigra* likely due to the use of primers previously developed for other pine species, which revealed to be less variable and/or monomorphic in European Black Pine (Šarac et al. 2015). The EST-SSRs used in this work were developed for *P. halepensis* (Leonarduzzi et al. 2016b) but showed high genetic diversity in *P. nigra*. Bonavita et al. (2016) achieved a H_e value (0.67) in *P. nigra* populations similar to the ones reported in this work. The same was found in other conifers such as *P. sylvestris* (0.850, Soranzo et al. 1998), P. *taeda* (0.679, Al-Rabab'ah and Williams 2002), P. *pinaster* (0.403, Naydenov et al. 2014) and *Taxus baccata* C. Linnaeus (0.621, Jaramillo-Correa et al. 2010). Slightly higher H_e values were found in *Pinus tabuliformis* E. A. Carrière (0.8739) and *Pinus henryi* M. T. Masters (0.8829, Liu et al. 2012). Contrarily, lower genetic diversity values were detected in some populations of *Pinus pinea* C. Linnaeus by different authors, possibly due to the occurrence of bottleneck effect (Fallour et al. 1997; Vendramin et al. 2008; Pinzauti et al. 2012), and also in *Abies alba* P. Miller (0.36 to 0.40, Leonarduzzi et al. 2016a). Overall, the H_e values were higher than those of H_o, suggesting low heterozygosity. The H_e and H_o are not expected to be under the Hardy Weinberg equilibrium due to the inexistence of descendants, being the sampled trees composed only by planted individuals. This fact does not allow a precise estimation of the null alleles, which may lead to an overestimation of the genetic differentiation between populations.

Previous works reported that the fragmented distribution of *P. nigra* has contributed to high intra-population variation generating taxonomic issues, particularly at the infraspecific level (Afzal-Rafii and Dodd 2007). According to Giovannelli et al. (2017), such SSR markers are highly informative and valuable for population genetic studies and can be interesting to resolve the complex taxonomy of the *P. nigra*. Giovannelli et al. (2019) recently reported the need for revising the infraspecific taxonomy of this pine species based on the integration of molecular data with its demographic history resulting from geological events. Regarding the wide and patchy distribution of *P. nigra*, it was expected a strong genetic differentiation among populations and a low within-population genetic diversity (Giovannelli et al. 2019; Scotti-Saintagne et al. 2019). However, these authors reported a weak genetic spatial structure probably resulting from the fragmentation of an ancestral population at the late Pleistocene or early Holocene into six to seven genetic lineages with high gene flow among them, actively contributing for admixture (Giovannelli et al. 2019; Scotti-Saintagne et al. 2019). Therefore, these authors recommended the revision of the infraspecific taxonomy of *P. nigra*. This task should be based on molecular markers data but also transcriptomics and focusing in major biogeographic regions where the species is growing naturally, as recently proposed by Olsson et al. (2020).

The six allochthonous Portuguese populations revealed a reduced proportion of total genetic differentiation among them. Their inferred genetic structure corroborated this result into two clusters with reduced F_{ST} values or genetic distinction between them. Similar results were observed in the PCoA that showed high genetic similarity among Manteigas, Caminha and Campeã, and the distance of this group relative to the populations of Vila Pouca de Aguiar and Paredes de Coura (Fig. 1).

The SSR data achieved in this work reinforced our previous hypothesis about the origin of the Portuguese populations in two provenances, based on their high

genetic similarity with two varieties of subsp. *laricio* (Dias et al. 2019). This hypothesis may also explain the low differentiation among the six Portuguese populations assayed by the SSRs since they had origin in one main genetic group, the same subspecies (*laricio*). Other authors reported a weak differentiation among populations of subsp. *laricio* based on biochemical (Fineschi 1984) and SSR data (Bonavita et al. 2016).

Giovannelli et al. (2017) genotyped *P. nigra* individuals of subspecies *laricio* from Corsica (France) and Italy. Dias et al. (2019) compared the ISSR and SCoT patterns of the Portuguese individuals with foreign samples of subsp. *laricio* from var. *calabrica* (with origin in the region of Cosenza, Italy) and var. *corsicana* (Corsica, France). The use of foreign samples from different geographic areas along with the high gene flow and admixture of *P. nigra* lineages (Giovannelli et al. 2019) may justify the genetic distance observed in this work among the Portuguese populations and the foreign samples (Fig. 3). Despite the achievement of reduced values of genetic differentiation among the Portuguese and foreign samples with the Genalex and STRUCTURE analyses, the PCoA based on the pooled SSR data projected these two groups separately and the structure analysis inferred two main groups (Figs. 2d and 3).

The hypothetic infraspecific taxonomic classification that we propose for the Portuguese planted *P. nigra* populations is partially supported by a phenotypic characterisation realised in the past century, by Louro (1982). This author considered that the Portuguese *P. nigra* populations existent at that time belong to the subspecies *laricio*, *salzmanni* and *nigra*. Louro (1982) pointed out that *laricio* was the best adapted and predominant species in Portugal. Nevertheless, and regarding the recent researches that reported high gene flow and admixture among the *P. nigra* lineages (Giovannelli et al. 2019; Scotti-Saintagne et al. 2019), we do not discard the realisation of further molecular studies that include subspecies and varieties of other provenances as foreign samples to compare and to pursue the extrapolation of the infraspecific taxonomy of the Portuguese populations of *P. nigra*.

5 Conclusion

This study constituted the first molecular characterisation of the six *P. nigra* populations representative of the species distribution in Portugal, using SSR markers.

The SSR molecular data demonstrated that these allochthonous *P. nigra* populations have high genetic diversity at the intra-population level but reduced genetic differentiation among them. Their genetic structure into two clusters with low genetic differentiation allowed us to suggest that in their plantation, plant material with two different origins was used. The reduced proportion of total genetic differentiation found among the Portuguese, and foreign samples indicated that the former ones had origin in plant material from other subspecies or provenances, not included in this work. Hence, additional molecular studies to pursue the extrapolation of the infraspecific taxonomy of the Portuguese allochthonous populations of *P. nigra* are required

The understanding of the genetic diversity, structure and relationships of these *P. nigra* populations will be highly important under the scope of forestry management, genetic improvement and/or for the definition of afforestation and conservation strategies.

Acknowledgements The author A.D. thanks Maria Celeste Labriola from the Institute of Biosciences and Bioresources, Sezione di Firenze, Consiglio Nazionale delle Ricerche, Sesto Fiorentino (Firenze), for the help and support during her laboratory work. The author A.C. thanks the FCT and UTAD for her contract as researcher under the scope of the D.L. no. 57/2016 of 29 August and Law no. 57/2017 of 19 July.

Author's contributions Conceptualization: MJG, AC, JLB and JLL; methodology: AC, IS, JP, MES, BF, GGV, JLL; formal analysis: FB, GG, AC, JLL and JLB; investigation: AC, JLL, JLB; resources: MJG, MES, JLL, JLB; writing—original draft: AD; writing—review and editing: all authors; supervision: AC, MJG, JLL, JLB; funding acquisition: JLB, GGV.

Funding information This work was funded by national funds provided by the FCT – Fundação para a Ciência e a Tecnologia) to CITAB under the project UIDB/04033/2020 and to the author A.D. by the attribution of the doctoral grant SFRH/BD/91781/2012, co-financed by the Social European Fund (FSE) under the POPH-QREN program; and also by the COST Action FP1202 "Strengthening conservation: a key issue for adaptation of marginal/ peripheral populations of forest trees to climate change in Europe (MaP-FGR)" (European Cooperation in Science and Technology) that financed a Short Term Scientific Mission to the author A.D. at Florence. The authors thank the Centre grants UIDB/04046/2020, UIDP/04046/2020 and UID/AGR/00239/2020 funded by the Portuguese Foundation for the Science and Technology (FCT – Fundação para a Ciência e a Tecnologia).

Data availability The datasets generated during and/or analysed during the current study are available in the Figshare repository (https://doi.org/10.6084/m9.figshare.11344271.v5).

Compliance with ethical standards

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

Annexes

Table 7 Genetic diversity statistics assayed per Portuguese *P. nigra* population and SSR *locus*. Notes: na, observed number of alleles; ne, effective number of alleles (Kimura and Crow 1964); *I*, Shannon's Information Index (Lewontin 1972); *h*, Nei's gene diversity index (Nei 1973); H_o, observed heterozygosity; H_e, expected heterozygosity (Levene 1949); *F*, fixation index; and s.d., standard deviation

Population	SSR *locus*	na	ne	*I*	*h*	H_o	H_e	*F*
Manteigas	pn6360	14	6.42	2.21	0.84	0.79	0.86	0.059
	pn7754	17	8.17	2.40	0.88	0.77	0.89	0.121
	SPAG_7.14	15	7.80	2.39	0.87	0.25	0.89	0.713
	pn4379	13	3.33	1.70	0.70	0.50	0.71	0.286
	pn2246	18	10.42	2.61	0.90	0.56	0.92	0.381
	pn6175	23	9.00	2.67	0.89	0.73	0.90	0.182
	pn1403	11	5.96	2.03	0.83	0.65	0.84	0.222
	PHA_6062	6	3.60	1.42	0.72	0.53	0.73	0.269
	PHA_4783	2	1.46	0.49	0.31	0.28	0.32	0.113
	PtTX4001	8	2.36	1.25	0.58	0.50	0.58	0.132
	pn8747	11	5.75	1.94	0.83	0.67	0.84	0.193
	PtTX3107	4	2.58	1.11	0.61	0.40	0.62	0.346
	pn6266	8	3.86	1.64	0.74	0.38	0.75	0.494
	Mean ± s.d.	11.54 ± 1.66	5.44 ± 0.78	1.84 ± 0.18	0.75 ± 0.05	0.54 ± 0.05	0.76 ± 0.05	0.270 ± 0.050
Vale do Zêzere	pn6360	15	7.54	2.30	0.87	0.86	0.88	0.003
	pn7754	16	9.92	2.50	0.90	0.78	0.91	0.128
	SPAG_7.14	17	11.76	2.63	0.92	0.24	0.93	0.736
	pn4379	13	5.05	1.96	0.80	0.76	0.81	0.055
	pn2246	21	14.92	2.86	0.93	0.56	0.95	0.401
	pn6175	25	13.83	2.90	0.93	0.68	0.94	0.272
	pn1403	12	6.83	2.17	0.85	0.81	0.87	0.050
	PHA_6062	8	4.60	1.75	0.78	0.75	0.79	0.042
	PHA_4783	2	1.35	0.43	0.26	0.25	0.26	0.034
	PtTX4001	6	3.99	1.51	0.75	0.59	0.76	0.207
	pn8747	10	4.66	1.79	0.79	0.68	0.80	0.140
	PtTX3107	7	1.88	1.05	0.47	0.27	0.48	0.424
	pn6266	8	2.90	1.42	0.66	0.42	0.66	0.364
	Mean ± s.d.	12.31 ± 1.78	6.86 ± 1.24	1.94 ± 0.20	0.76 ± 0.06	0.59 ± 0.06	0.77 ± 0.06	0.22 ± 0.06
Campeã	pn6360	13	5.50	2.04	0.82	0.76	0.83	0.074
	pn7754	12	7.23	2.18	0.86	0.91	0.87	-0.061
	SPAG_7.14	22	14.29	2.87	0.93	0.22	0.95	0.761
	pn4379	14	3.51	1.82	0.72	0.62	0.73	0.137
	pn2246	21	13.14	2.79	0.92	0.59	0.94	0.363
	pn6175	24	12.83	2.88	0.92	0.77	0.94	0.163
	pn1403	12	7.36	2.22	0.86	0.66	0.88	0.239
	PHA_6062	10	5.02	1.83	0.80	0.71	0.81	0.118
	PHA_4783	2	1.51	0.52	0.34	0.37	0.34	-0.103
	PtTX4001	7	2.87	1.31	0.65	0.69	0.66	-0.053
	pn8747	11	6.35	2.01	0.84	0.57	0.85	0.322
	PtTX3107	6	1.60	0.75	0.37	0.37	0.38	0.009
	pn6266	8	3.89	1.63	0.74	0.42	0.75	0.429
	Mean ± s.d.	12.46 ± 1.81	6.55 ± 1.21	1.91 ± 0.21	0.75 ± 0.05	0.59 ± 0.05	0.76 ± 0.06	0.19 ± 0.07
Paredes de Coura	pn6360	15	8.48	2.41	0.88	0.94	0.89	-0.069
	pn7754	17	7.78	2.43	0.87	0.83	0.88	0.044

Table 7 (continued)

Population	SSR *locus*	na	ne	*I*	*h*	H_o	H_e	*F*
	SPAG_7.14	14	8.94	2.40	0.89	0.37	0.91	0.583
	pn4379	17	6.68	2.29	0.85	0.77	0.86	0.093
	pn2246	16	10.18	2.51	0.90	0.52	0.92	0.429
	pn6175	23	16.01	2.93	0.94	0.63	0.95	0.330
	pn1403	16	10.12	2.50	0.90	0.77	0.91	0.144
	PHA_6062	8	5.47	1.88	0.82	0.86	0.83	-0.049
	PHA_4783	2	1.53	0.53	0.35	0.28	0.35	0.196
	PtTX4001	7	4.31	1.59	0.77	0.58	0.78	0.241
	pn8747	10	5.38	1.92	0.81	0.86	0.83	-0.053
	PtTX3107	4	1.76	0.85	0.43	0.31	0.44	0.293
	pn6266	7	1.76	0.96	0.43	0.34	0.44	0.205
	Mean ± s.d.	12 ± 1.70	6.80 ± 1.14	1.94 ± 0.21	0.76 ± 0.06	0.62 ± 0.07	0.77 ± 0.06	0.18 ± 0.05
Vila Pouca de Aguiar	pn6360	11	6.54	2.09	0.85	0.79	0.86	0.068
	pn7754	11	7.66	2.16	0.87	0.87	0.88	0.001
	SPAG_7.14	16	9.51	2.46	0.90	0.52	0.91	0.424
	pn4379	9	6.22	1.99	0.84	0.84	0.85	-0.003
	pn2246	9	5.05	1.83	0.80	0.54	0.81	0.326
	pn6175	14	6.04	2.13	0.83	0.55	0.85	0.338
	pn1403	7	4.23	1.66	0.76	0.66	0.77	0.139
	PHA_6062	6	3.15	1.38	0.68	0.84	0.69	-0.235
	PHA_4783	2	1.93	0.68	0.48	0.45	0.49	0.074
	PtTX4001	5	1.75	0.86	0.43	0.45	0.43	-0.043
	pn8747	11	6.43	2.04	0.84	0.84	0.86	0.003
	PtTX3107	4	1.69	0.79	0.41	0.29	0.41	0.291
	pn6266	5	1.71	0.84	0.41	0.26	0.42	0.365
	Mean ± s.d.	8.46 ± 1.14	4.76 ± 0.71	1.61 ± 0.17	0.70 ± 0.05	0.61 ± 0.06	0.71 ± 0.05	0.13 ± 0.06
Caminha	pn6360	11	7.31	2.15	0.86	0.85	0.87	0.011
	pn7754	13	6.50	2.15	0.85	0.85	0.86	-0.009
	SPAG_7.14	16	6.88	2.34	0.85	0.29	0.87	0.666
	pn4379	12	3.55	1.68	0.72	0.59	0.73	0.185
	pn2246	19	13.11	2.73	0.92	0.49	0.94	0.473
	pn6175	22	9.29	2.63	0.89	0.76	0.90	0.153
	pn1403	11	6.37	2.05	0.84	0.78	0.85	0.074
	PHA_6062	7	3.60	1.50	0.72	0.59	0.73	0.189
	PHA_4783	2	1.46	0.49	0.31	0.29	0.32	0.068
	PtTX4001	9	2.95	1.44	0.66	0.61	0.67	0.078
	pn8747	10	6.79	2.05	0.85	0.83	0.86	0.028
	PtTX3107	4	2.07	0.87	0.52	0.41	0.52	0.198
	pn6266	9	3.05	1.50	0.67	0.51	0.68	0.238
	Mean ± s.d.	11.15 ± 1.54	5.61 ± 0.91	1.81 ± 0.18	0.75 ± 0.05	0.60 ± 0.06	0.75 ± 0.05	0.18 ± 0.05
Total mean ± s.d.	11.32 ± 0.66	6.00 ± 0.41	1.84 ± 0.08	0.74 ± 0.02	0.74 ± 0.02	0.76 ± 0.02	0.20 ± 0.02	0.20 ± 0.02

Table 8 Genetic differentiation coefficient (Gst) and differentiation index of allele frequencies (Jost's *D*) estimated per SSR locus, and total mean value (± standard deviation, s.d.) determined with the 13 SSR markers

Multiplex	SSR *locus*	Gst	Jost's D
Primer mix 1	pn6360	0.008	0.065
	pn7754	0.023	0.217
	SPAG_7.14	0.020	0.280
	pn4379	0.073	0.343
	pn2246	0.021	0.281
	pn6175	0.020	0.269
	pn1403	0.015	0.111
Primer mix 2	PHA_6062	0.007	0.027
	PHA_4783	0.026	0.017
	PtTX4001	0.041	0.094
	pn8747	0.008	0.049
	PtTX3107	0.034	0.038
	pn6266	0.036	0.075
Mean ± s.d.		0.025 ± 0.005	0.095 ± 0.031

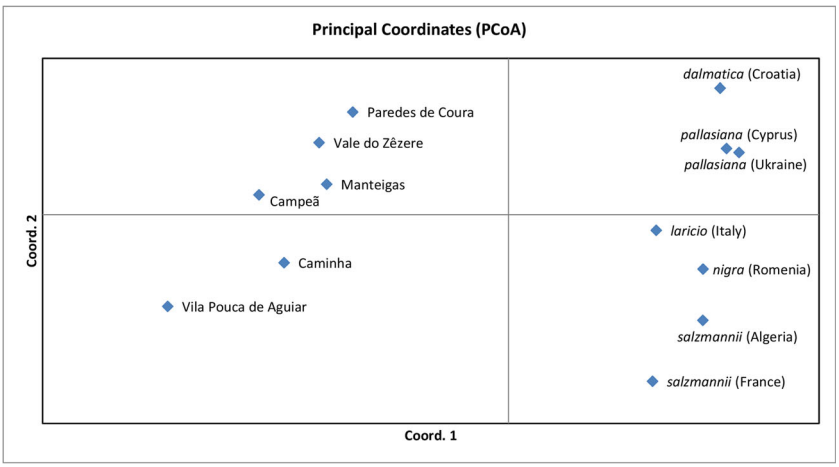

Fig. 3 PCoA based on the pooled SSR data using Nei's genetic distance coefficient (Nei 1978), showing the projection of the six Portuguese *P. nigra* populations apart from the seven foreign samples representative of the subspecies *dalmatica*, *pallasiana*, *laricio*, *nigra* and *salzmannii*. The cumulative percentage of total variation explained by the first three coordinates was 74.50%.

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. https://doi.org/10.1111/j.1365-294X.2006.03183.x

Al-Rabab'ah MA, Williams CG (2002) Population dynamics of *Pinus taeda* L. based on nuclear microsatellites. For Ecol Manag 163:263–271

Barbéro M, Loisel R, Quézel P et al (1998) Ecology and Biogeography of *Pinus*. Cambridge University Press, South Africa, David M. Richardson

Bogunić F, Muratovic E, Brown SC, Siljak-Yakovlev S (2003) Genome size and base composition of five *Pinus* species from the Balkan region. Plant Cell Rep 22:59–63. https://doi.org/10.1007/s00299-003-0653-2

Bogunić F, Muratovic E, Ballian D et al (2007) Genome size stability among five subspecies of *Pinus nigra* Arnold s.l. Environ Exp Bot 59:354–360. https://doi.org/10.1016/j.envexpbot.2006.04.006

Bogunić F, Siljak-Yakovlev S, Muratović E, Ballian D (2009) Different karyotype patterns among allopatric *Pinus nigra* (*Pinaceae*) populations revealed by molecular cytogenetics. Plant Biol 13:194–200. https://doi.org/10.1111/j.1438-8677.2010.00326.x

Bonavita S, Vendramin GG, Bernardini V, Avolio S, Regina TMR (2016) The first SSR-based assessment of genetic variation and structure among *Pinus laricio* Poiret populations within their native area. Plant Biosyst - Int J Deal Asp Plant Biol 150:1271–1281. https://doi.org/10.1080/11263504.2015.1027316

Bussotti F (2002) *Pinus nigra* Arnold. In: Pines of silvicultural importance / compiled from the Forestry compendium, 1[a]. CAB International Publishers, New York, pp 266–282

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 For 43:303–317. https://doi.org/10.1007/s11056-011-9282-8

Cho YG, Ishii T, Temnykh S, Chen X, Lipovich L, McCouch SR, Park WD, Ayres N, Cartinhour S (2000) Diversity of microsatellites derived from genomic libraries and GenBank sequences in rice (*Oryza sativa* L.). Theor Appl Genet 100:713–722. https://doi.org/10.1007/s001220051343

Christensen K (1997) *Pinaceae*, *Cupressaceae*, *Taxaceae*, *Ephedraceae*, *Salicaceae*, *Juglandaceae*, *Betulaceae*, *Fagaceae*, *Ulmaceae Moraceae*, Strid A, Tan K. Koeltz Scientific Books, Königstein, Germany, Flora Hellenica

Cipriano J, Carvalho A, Fernandes C, Gaspar MJ, Pires J, Bento J, Roxo L, Louzada J, Lima-Brito J (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 93:41–48. https://doi.org/10.1007/s12041-013-0241-3

del Cerro BA, Borja ML, Garcia EM et al (2009) Influence of stand density and soil treatment on the Spanish Black Pine (*Pinus nigra* Arn. ssp. *salzmannii*) regeneration in Spain. Investig Agrar Sist Recur For 18:167–180

Dias A, Fonseca T, Silva ME et al (2014) *Pinus nigra* growth in Portugal. In: Book of Abstracts of the 5th International Conference on Mediterranean Pines (MedPine 5). Solsona, Spain, p 123

Dias A, Gaspar MJ, Carvalho A, Pires J, Lima-Brito J, Silva ME, Louzada JL (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

Dias A, Lemos M, Pavia I, Gaspar MJ, Silva ME, Louzada JL, Lima-Brito J, Carvalho A (2019) Genetic characterization of Portuguese allochthonous populations of *Pinus nigra* using ISSRs and SCoTs and extrapolation of their infraspecific taxonomy. Physiol Mol Biol Plants 25:1–7. https://doi.org/10.1007/s12298-019-00649-5

Dias A, Giovanelli G, Fady B, Spanu I, Vendramin GG, Bagnoli F, Carvalho A, Silva ME, Lima-Brito J, Lousada JL, Gaspar MJ (2020) Portuguese Pinus nigra populations : genetic diversity, structure and relationships inferred by SSR markers. V5. Figshare. [Dataset]. https://doi.org/10.6084/m9.figshare.11344271.v5

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. https://doi.org/10.1007/s12686-011-9548-7

Ellegren H (2004) Microsatellites: simple sequences with complex evolution. Nat Rev Genet 5:435–445

Elsik CG, Williams CG (2001) Low-copy microsatellite recovery from a conifer genome. Theor Appl Genet 103:1189–1195. https://doi.org/10.1007/s001220100725

Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software structure: a simulation study. Mol Ecol 14:2611–2620. https://doi.org/10.1111/j.1365-294X.2005.02553.x

Fallour D, Fady B, Lefevre F (1997) Study on isozyme variation in *Pinus pinea* L.: Evidence for low polymorphism. Silvae Genet 46:201–207

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: technical article. Mol Ecol Notes 7:574–578. https://doi.org/10.1111/j.1471-8286.2007.01758.x

Fineschi S (1984) Determination of the origin of an isolated group of trees of *Pinus nigra* through enzyme gene markers. Silvae Genet 33:169–172

Frankham R, Briscoe DA, Ballou JD (2002) Introduction to conservation genetics. Cambridge University Press, Cambridge, UK ; New York

Gaussen H, Heywood V, Chater A (1993) Genus *Pinus*. In: Tutin T, Burges N, Chater, et al. (eds) Flora Europaea, Volume 1: Psilotaceae to Platanaceae. Cambridge University Press, Cambridge, pp 40–44

Giovannelli G, Roig A, Spanu I, Vendramin GG, Fady B (2017) A New Set of Nuclear Microsatellites for an Ecologically and Economically Important Conifer: the European Black Pine (*Pinus nigra* Arn.). Plant Mol Biol Report 35:379–388. https://doi.org/10.1007/s11105-017-1029-z

Giovannelli G, Scotti-Saintagne C, Scotti I, et al (2019) The genetic structure of the European black pine (*Pinus nigra* Arnold) is shaped by its recent Holocene demographic history.: Supplementary files. bioRxiv. https://doi.org/10.1101/535591

Guichoux E, Lagache L, Wagner S et al (2011) Current trends in microsatellite genotyping: trends in microsatellite genotyping. Mol Ecol Resour 11:591–611. https://doi.org/10.1111/j.1755-0998.2011.03014.x

Hamrick JL (2004) Response of forest trees to global environmental changes. For Ecol Manag 197:323–335. https://doi.org/10.1016/j.foreco.2004.05.023

Hayden MJ, Nguyen TM, Waterman A, McMichael GL, Chalmers KJ (2008) Application of multiplex-ready PCR for fluorescence-based SSR genotyping in barley and wheat. Mol Breed 21:271–281. https://doi.org/10.1007/s11032-007-9127-5

Holsinger KE, Weir BS (2009) Genetics in geographically structured populations: defining, estimating and interpreting FST. Nat Rev Genet 10:639–650

Jaramillo-Correa JP, Grivet D, Terrab A et al (2010) The Strait of Gibraltar as a major biogeographic barrier in Mediterranean

conifers: a comparative phylogeographic survey: genetic divergence across the strait of Gibraltar. Mol Ecol 19:5452–5468. https://doi.org/10.1111/j.1365-294X.2010.04912.x

Jiménez JF, Sánchez-Gómez P, Martínez JF, et al (2005) Variabilidad genética de *Pinus nigra* subsp. salzmannii en la Región de Murcia mediante microsatélites cloroplásticos. In: Anales de Biología. pp 105–112

Jost L (2009) D vs. GST: Response to Heller and Siegismund (2009) and Ryman and Leimar (2009). Mol Ecol 18:2088–2091. https://doi.org/10.1111/j.1365-294X.2009.04186.x

Kimura M, Crow JF (1964) The number of alleles that can be maintained in a finite population. Genetics 49:725–738

Leonarduzzi C, Piotti A, Spanu I, Vendramin GG (2016a) Effective gene flow in a historically fragmented area at the southern edge of silver fir (*Abies alba* Mill.) distribution. Tree Genet Genomes 12:. https://doi.org/10.1007/s11295-016-1053-4

Leonarduzzi C, Spanu I, Labriola M, González-Martínez SC, Piotti A, Vendramin GG (2016b) Development and characterization of three highly informative EST-SSR multiplexes for *Pinus halepensis* mill. and their transferability to other Mediterranean pines. Plant Mol Biol Report 34:993–1002. https://doi.org/10.1007/s11105-016-0980-4

Levene H (1949) On a matching problem arising in genetics. Ann Math Stat 20:91–94

Lewontin RC (1972) Testing the Theory of Natural Selection. Nature 236:181–182

Liber Z, Nikolic T, Mitic B, Satovic Z (2003) RAPD markers and black pine [*Pinus nigra* Arnold] intraspecies taxonomy-evidence from the study of nine populations. Acta Soc Bot Pol 72:249–257

Liu Z-L, Cheng C, Li J (2012) High genetic differentiation in natural populations of *Pinus henryi* and *Pinus tabuliformis* as revealed by nuclear microsatellites. Biochem Syst Ecol 42:1–9. https://doi.org/10.1016/j.bse.2011.07.005

Louro V (1982) O pinheiro larício : *Pinus nigra* Arnold em Portugal. Direcção geral do ordenamento e gestão florestal, Lisboa

Lučić A, Mladenovic-Drinic S, Stavretovic N et al (2010) Genetic diversity of Austrian pine (*Pinus nigra* Arnold) populations in Serbia revealed by RAPD. Arch Biol Sci 62:329–336. https://doi.org/10.2298/ABS1002329L

Mogensen H (1996) The hows and whys of cytoplasmic inheritance in seed plants. Am J Bot 83:383–404. https://doi.org/10.2307/2446172

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

Naydenov KD, Tremblay FM, Fenton NJ, Alexandrov A (2006) Structure of *Pinus nigra* Arn. populations in Bulgaria revealed by chloroplast microsatellites and terpenes analysis: Provenance tests. Biochem Syst Ecol 34:562–574. https://doi.org/10.1016/j.bse.2006.01.011

Naydenov KD, Alexandrov A, Matevski V, Vasilevski K, Naydenov MK, Gyuleva V, Carcaillet C, Wahid N, Kamary S (2014) Range-wide genetic structure of maritime pine predates the last glacial maximum: evidence from nuclear DNA: Genetics of maritime pine populations in the western Mediterranean Basin. Hereditas 151:1–13. https://doi.org/10.1111/j.1601-5223.2013.00027.x

Nei M (1973) Analysis of gene diversity in subdivided populations. Proc Natl Acad Sci 70:3321–3323

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

Olsson S, Grivet D, Cattonaro F, Vendramin V, Giovannelli G, Scotti-Saintagne C, Vendramin GG, Fady B (2020) Evolutionary relevance of lineages in the European black pine (*Pinus nigra*) in the transcriptomic era. Tree Genet Genomes 16:30. https://doi.org/10.1007/s11295-020-1424-8

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

Pinzauti F, Sebastiani F, Budde KB, Fady B, González-Martínez SC, Vendramin GG (2012) Nuclear microsatellites for *Pinus pinea* (Pinaceae), a genetically depauperate tree, and their transferability to *P. halepensis*. Am J Bot 99:362–365. https://doi.org/10.3732/ajb.1200064

Portoghesi L, Consalvo M, Angelini A, Ferrari B, Barbati A, Castaldi C, Corona P (2013) Multifunctional management of mountain reforestations: thoughts and perspectives from a case study in Central Italy. Ital For E Mont 68:305–315. https://doi.org/10.4129/ifm.2013.6.03

Price RA, Liston A, Strauss SH (1998) Phylogeny and systematics of Pinus. In: Ecology and biogeography of *Pinus*. DM Richardson. 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

Pritchard JK, Wen W, Falush D (2003) Documentation for structure software: version 2

Pritchard JK, Wen X, Falush D (2009) Documentation for structure software: Version 2.3

Quézel P, Médail F (2003) Écologie et biogéographie des forêts du bassin méditerranéen. Elsevier, Paris

Rafii ZA, Dodd RS, Zavarin E (1996) Genetic diversity in foliar terpenoids among natural populations of European black pine. Biochem Syst Ecol 24:325–339

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. Int J Mol Sci 13:5645–5658. https://doi.org/10.3390/ijms13055645

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. https://doi.org/10.2298/GENSR1501205S

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

Scotti-Saintagne C, Giovannelli G, Scotti I, Roig A, Spanu I, Vendramin GG, Guibal F, Fady B (2019) Recent, late Pleistocene fragmentation shaped the phylogeographic structure of the European black pine (*Pinus nigra Arnold*). Tree Genet Genomes 15:76. https://doi.org/10.1007/s11295-019-1381-2

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

Soranzo N, Provan J, Powell W (1998) Characterization of microsatellite loci in *Pinus sylvestris* L. Mol Ecol 7:1260–1261

Soto A, Robledo-Arnuncio JJ, González-Martínez SC et al (2010) Climatic niche and neutral genetic diversity of the six Iberian pine species: a retrospective and prospective view. Mol Ecol 19:1396–1409. https://doi.org/10.1111/j.1365-294X.2010.04571.x

Varshney R, Graner A, Sorrells M (2005) Genomics-assisted breeding for crop improvement. Trends Plant Sci 10:621–630. https://doi.org/10.1016/j.tplants.2005.10.004

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

Vendramin GG, Fady B, González-Martínez SC, Hu FS, Scotti I, Sebastiani F, Soto Á, Petit RJ (2008) Genetically depauperate but widespread: the case of an emblematic mediterranean pine. Evolution 62:680–688. https://doi.org/10.1111/j.1558-5646.2007.00294.x

Vidaković M (1991) Conifer morphology and variation. Grafički zavod Hrvatske, Zagreb, Hrvatska

Wolfe KH, Li W-H, Sharp PM (1987) Rates of nucleotide substitution vary greatly among plant mitochondrial, chloroplast, and nuclear DNAs. Proc Natl Acad Sci 84:9054–9058

Yeh FC, Yang RC, Boyle TB, et al (1999) POPGENE version 1.32, the user-friendly shareware for population genetic analysis. Molecular Biology and Biotecnology Centre, University of Alberta, Edmonton, USA

Zaghi D, European Commission, Directorate-General for the Environment (2008) Management of Natura 2000 habitats (sub-) Mediterranean pine forests with endemic black pines 9530: Directive 92/43/EEC on the conservation of natural habitats and of wild fauna and flora. European Commission, Brussels

Zhou Y, Bui T, Auckland L, Williams CG (2002) Undermethylated DNA as a source of microsatellites from a conifer genome. Genome 45: 91–99. https://doi.org/10.1139/g01-119

Affiliations

Alexandra Dias[1] · Guia Giovannelli[2] · Bruno Fady[2] · Ilaria Spanu[3] · Giovanni G Vendramin[3] · Francesca Bagnoli[3] · Ana Carvalho[1,4,5] · Maria Emília Silva[1,6] · José Lima-Brito[1,4,5] · José Luís Lousada[1,6] · Maria João Gaspar[4,5,7]

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

2 Écologie des Forêts Méditerranéennes, INRA, UR629, Avignon, France

3 Institute of Biosciences and Bioresources, IBBR-CNR, Sesto Fiorentino, FI, Italy

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

5 Department of Genetics and Biotechnology, University of Tras-os-Montes and Alto Douro, Quinta de Prados, 5000-801 Vila Real, Portugal

6 Department of Forestry Sciences and Landscape (CIFAP), University of Tras-os-Montes and Alto Douro, Quinta de Prados, 5000-801 Vila Real, Portugal

7 Forest Research Centre (CEF), School of Agronomy, University of Lisbon, Lisbon, Portugal