CrossMark

Ancient split of major genetic lineages of European Black Pine: evidence from chloroplast DNA

Krassimir D. Naydenov[1] · Michel K. Naydenov[2] · Alexander Alexandrov[3] ·
Kole Vasilevski[4] · Veselka Gyuleva[3] · Vlado Matevski[5] · Biljana Nikolic[6] ·
Venceslas Goudiaby[7] · Faruk Bogunic[8] · Despina Paitaridou[9] · Andreas Christou[10] ·
Irina Goia[11] · Christopher Carcaillet[12] · Adrian Escudero Alcantara[13] · Cengiz Ture[14] ·
Suleyman Gulcu[15] · Lorenzo Peruzzi[16] · Salim Kamary[17] · Srdjan Bojovic[18] ·
Georgi Hinkov[3] · Anatoly Tsarev[19]

Abstract The European Black Pine (*Pinus nigra* Arn.) has a long and complex history. Genetic distance and frequency analyses identified three differentiated genetic groups, which corresponded to three wide geographical areas: Westerns Mediterranean, Balkan Peninsula and Asia Minor. These groups shared common ancestors (14.75 and 10.72 Ma). The most recent splits occurred after the Messinian Salinity Crisis (4.37 Ma) and the Early–Middle Pleistocene Transitions

This article is dedicated to the memory of Prof. Dr. Dimitar Velkov from Forest Research Institute, Bulgarian Academy of Science (1921–2001).

Communicated by Z. Kaya

Electronic supplementary material The online version of this article (doi:10.1007/s11295-016-1022-y) contains supplementary material, which is available to authorized users.

✉ Krassimir D. Naydenov
   KrassimirDN@gmail.com; KrassimirDN@hotmail.com

[1]   Ministry of National Defence—CAF, 25 Grand-Bernier, St-Jean-sur-Richelieu, QC J0J1R0, Canada

[2]   Faculty of Medicine, Ss. Cyril and Methodius University, 50 Divizija 6, 1000 Skopje, Republic of Macedonia

[3]   Forest Research Institute, Bulgarian Academy of Science, 132 Bd. Kliment Ochridskii, 1756 Sofia, Bulgaria

[4]   Faculty of Forestry, Ss. Cyril and Methodius University, Blvd. Edvard Kardelj bb, 1000 Skopje, Republic of Macedonia

[5]   Faculty of Natural Sciences and Mathematics, Institute of Biologie, 1000 Skopje, Republic of Macedonia

[6]   Institute of Forestry, Kneza Višeslava 3, 11000 Belgrade, Serbia

[7]   Forest Ecology and Forest Management, Wageningen University, P.O. Box 47, 6700 AA, Wageningen, The Netherlands

[8]   Faculty of forestry, University of Sarajevo, 20 Zagrebacka, Sarajevo, Bosnia and Herzegovina

[9]   Ministry of Reconstruction of Production, Environment, and Energy, 31 Xalkokondili, 10164 Athens, Greece

[10]   Departament of Forestry, 26 Loukis Akritas Avenue, 1414 Nicosia, Cyprus

[11]   Faculty of Biology and Geology, Babes-Bolyai University, Republicii Street 42, 400015 Cluj-Napoca, Romania

[12]   UMR5023 CNRS-Université Lyon 1, 43 Bd 11 novembre 1918, F-69622 Villeurbanne Cedex, France

[13]   Area de Biodiversidad and Conservacion, Universidad Rey Juan Carlos, Tulipan, 28933 Mostoles, Madrid, Spain

[14]   Science Faculty, Department of Biology, Ecology Program, Anadolu University, Biyoloji Bolumu, 26470 Eskisehir, Turkey

[15]   Forestry Faculty, Suleyman Demirel University, 32260 Isparta, Turkey

[16]   Dipartimento di Biologia, Università di Pisa, I-56126 Pisa, Italy

[17]   Faculté des Sciences, Université Mohamed Premier, Oujda, 63, rue Al maqdis Alqods, 60000 Oujda, Maroc

[18]   Institute for Biological Research "Siniša Stanković", University of Belgrade, Bul. Despota Stefana 142, 11000 Belgrade, Serbia

[19]   Petrozavodsk State University, Petrozavodsk, Lenin Avenue 33, Petrozavodsk, Republic of Karelia 185910, Russia

(0.93 Ma). The posterior ancestral population size (Na) is 260,000–265,000 individuals. Each pool is further fragmented, with evidence of a phylogeographic structure ($N_{st} > G_{st}$) typically observed in some natural populations from the Western Mediterranean region and the Balkan Peninsula. The laboratory analysis was performed by fragment analysis—i.e. electrophoretic sizing of polymerase chain reaction fragments, combined with the sequencing analysis of 33 % of all individuals as a control. Intense sampling of chloroplast DNA polymorphisms (3154 individuals and 13 markers: SNPs and SSRs) over the full area of the species' natural distribution indicated moderate among-population variability ($G_{st(nc)} \leq 0.177$) in various parts of its range. These results indicate that the natural populations have long migration histories that differ from one another and that they have been strongly phylogeographically affected by complex patterns of isolation, speciation and fragmentation. Long and varying climatic fluctuations in the region of the principal genetic group have been the probable cause of different forest community associations with different successional patterns resulting in *interglacial refugia* vs. *macro long-term refugia*.

Keywords Chloroplast DNA · Phylogeography · *Pinus nigra* · Population structure · SNP · SSR

Introduction

Forest trees are major components of the ecosystems in large geographical regions. Consequently, they are essential for the maintenance of both mountain and lowland biodiversity because many organisms depend on the habitats provided or shaped by woody plants. Past population dynamics may be used to improve the conservation and management of genetic resources, and numerous studies have focused on populations (past and present) of tree species in the southern European region where the *main glacial refugia* of biodiversity in Europe is believed to have occurred. The refugia are *hotspots* that appear to have allowed the survival of many species, with strong persisting effects on the distribution and diversity of current flora and fauna (Taberlet et al. 1998; Petit et al. 2003; Provan & Bennett 2008).

Conifers are the major components of many ecosystems, and many populations also have considerable economic importance (Farjon 2008; Eckenwalder 2009). Hence, the population demographics, genetic structures, migration patterns, and systematic relationships of conifers have been intensively studied, especially in the Northern Hemisphere (Naydenov et al. 2007; Godbout et al. 2008; Jaramillo-Correa et al. 2008).

In conifers, mitochondrial DNA (mtDNA) is maternally inherited and chloroplast DNA (cpDNA) is generally paternally inherited (Powell et al. 1995a, b; Vendramin et al. 1996). These differences can be highly valuable for tracing patterns of gene flow via seeds and pollen, respectively, and mtDNA and cpDNA have been widely used in analyses of the genetic structure of conifer populations (Naydenov et al. 2007; Bonavita et al. 2015). Because wind-dispersed plant seeds are thought to have a smaller effective radius of distribution than pollen (Wright 1979; Grant 1980), markers in these genomes have widely differing rates of gene flow. Additionally, the syntheses of multiple analyses of plants, birds and insects suggest that there is generally a negative correlation between the rates of intraspecific gene flow and introgression in the absence of selective pressures under a null neutral model (Petit & Excoffier 2009). This negative correlation indicates that the effective delimitation of species' (or subspecies') complexes will be more congruent with low mutation rates (i.e. long signal/memory of the population history) and high levels of variation in markers, such as cpDNA variations in coniferous species (Provan et al. 1999). This phenomenon is particularly important for analyses of the genetic structure of *collective species* with the following: (i) incomplete reproductive barriers; (ii) retained ancestral polymorphisms; (iii) fragmented distributions (e.g. by mountains); and (iv) a long history of anthropogenic pressure. Such is the case for the European Black Pine according to Fukarek (1958a, b), Dobrinov et al. (1982), and Vidakovic (1991). The meaning of collective species used by the authors is closer to the definition of a *species complex* and *cryptic species* of the present day.

The European Black Pine (*Pinus nigra* Arn., subgenus *Diploxilon*) is one of the most intensively studied conifers after the Scots Pine, Maritime Pine, and Norway Spruce. It is considered to be a Tertiary relict (i.e. 2.6–66.0 Ma; Mirov 1967) and has been one of the most economically important forest species in the Mediterranean region for at least 200 years. Its natural area of distribution is in southern Europe, northwestern Africa and Asia Minor. It is present in all of the main European glacial refugia, including the Iberian, Apennine, and Balkan Peninsulas, Turkey and some Mediterranean islands. Currently, its distribution is highly fragmented in its marginal regions and west of the Balkans (i.e. west of the Adriatic and Ionian Sea basins); however, it is more contiguously distributed in the Balkans and Asia Minor (Turkey). It occupies exposed sites with southern aspects and well-drained, poor soil at altitudes ranging from 100 to 1800 m above sea level according to Dobrinov et al. (1982).

The species was first taxonomically described by Miller (1768), and since then, several authors have identified more than 20 sub-species based on morphological and anatomical data (Fukarek 1958a, b; Vidakovic 1991). This delineation is of particular interest because the long-term fragmentation and persistence of individual populations has given rise to extreme adaptation to the local environment, which is expressed as high morphological variability among populations. Consequently, several authors consider it an example of a

collective species (Fukarek 1958a, b; Dobrinov et al. 1982; Vidakovic 1991). During the last 30–40 years, several attempts have also been made to analyse the genetic and chemical-phenotypical population and subspecies structures of the Black Pine in geographic regions of varying sizes using karyology, terpene and allozyme analyses (Fineschi 1984; Bojović 1995; Naydenov et al. 2003). The first molecular analysis of the species' genetic structure was based on cpDNA fragment SSR analyses of 324 individuals from nine Black Pine populations in Bulgaria (Naydenov et al. 2006). This analysis was followed by two studies using the same molecular markers in populations from the western Mediterranean in Europe (Raffi & Dodd 2007; Soto et al. 2010; Bonavita et al. 2015). The proposed structures have been generally linked with the species' geographic distribution, but not the full area of its natural distribution was analysed. Thus, the global genetic dynamics and structure of the Black Pine remain unclear despite multiple publications on its biogeography and population genetic variation (Dobrinov et al. 1982; Vidakovic 1991; Bojović 1995).

Therefore, in the present study, we analysed sequence variation in 13 regions of the cpDNA of *P. nigra* individuals from across its entire natural range. Then, we evaluated the acquired data using several statistical methods to more thoroughly elucidate the Black Pine diversity and genetic structure. We hypothesised that the acquired data would indicate a strong phylogeographic structure in all studied populations and a high level of among-population differentiation.

Material and methods

Sampling and laboratory analysis

A total of 3154 individuals from 106 natural populations (29.7 trees per population on average) were sampled from the entire area of the species' natural distribution from 5.1° N to 48.1° N latitude (from Morocco to Austria) and 4.7° W to 39.1° E longitude (from Morocco to Turkey, Fig. 1). Total DNA was extracted from megagametophytes or green needle tissues using the Plant Elute Genomic Mini Kit (Sigma-Aldrich, USA) following the manufacturer's instructions. cpSSRs primer pairs that were tailed, modified and fluorescently labelled with FAM-6, NED or VIC as described by Vendramin et al. (1996) (synthesized by Applied Biosystems) were initially used to determine size variations at six microsatellite loci following the PCR protocol of Naydenov et al. (2006). MapMarker 1000-ROX (Bio Ventures) size standards, an ABI-3130 Genetic Analyzer and the Run3130 DATA Collection v. 3.0 and GeneMAPPER v.3.7 software (Applied Biosystems) were used. The six-tailed modified cpSSR primers were Pt-30204, Pt-36480, Pt-45002, Pt-71936, Pt-79951 and Pt-87268. The length variations for each

chloroplast fragment were classified by size and population geographic origin. Then, for accuracy, both strands of the samples from a third of the individuals (i.e. 1051) were sequenced after purification using CleanSEQ and AMPure (Agencourt Bioscience) SPRI paramagnetic beads using the same six (non-labelled and non-modified) primers, a BigDye Terminator v. 1.1 Sequencing Kit (Applied Biosystems), and the same Genetic Analyzer. The obtained sequences were assembled by the SEQUENCHER v.4.8 software (Gene Code Corp.). Variations in 13 regions were investigated: seven SSRs (single sequence repeats) and six flanking regions with single nucleotide polymorphisms (SNPs). The seven SSRs were from six primer sets because the region sequenced with Pt-30204 had two independent repetitive sequences. The sequencing was performed to minimize homoplasy and false-positive signals in the basic data and to substantially increase the power of the statistical analysis.

Statistical analysis of the population structure

The six chloroplast region sequence analyses for the 1051 individuals were interpolated for fragments of the same size and from the same population (i.e. the same length fragments from sequenced and non-sequenced individuals were interpreted as identical if they were from same population); one third of the samples were sequenced as an additional control. We obtained the sequence for the six chloroplast regions from all 3154 analysed individuals with less possibility of homoplasy (i.e. electromorph size homoplasy) due to the intensive sequence analysis and interpolation within the limit of each population (Table S2 in the Appendix). The sequences obtained from the 3154 individuals by interpolation from the six chloroplast regions were used for all of the statistical analyses. Chloroplast DNA diversity was calculated for each natural population in terms of the number of haplotypes expressed as N_{hap}/N_{ind} (the number of haplotypes detected divided by the total number of analysed individuals from each population) and the nucleotide diversity. These results were compared to the Hcp chlorotype diversity (expected heterozygosity) estimated from 10^3 random permutations using ARLEQUIN v. 3.1 (Excoffier et al. 2002). We estimated an unbiased size variant number in the chloroplast SSR variable regions for each population with the HP-RARE v.1.0 program (Kalinowski 2005).

The population structure was studied using both distance-based (I) and model-based (II), i.e. frequency-based approaches. We used two distance-based methods with the obtained results interpreted as principals; the model-based method was used as a complementary approach. For the distance-based approach (I), we calculated pairwise genetic distances according to the method of Rogers (1972) based on the assumptions of no mutations and no selection. To avoid the risk of mutations in the models changing between populations and
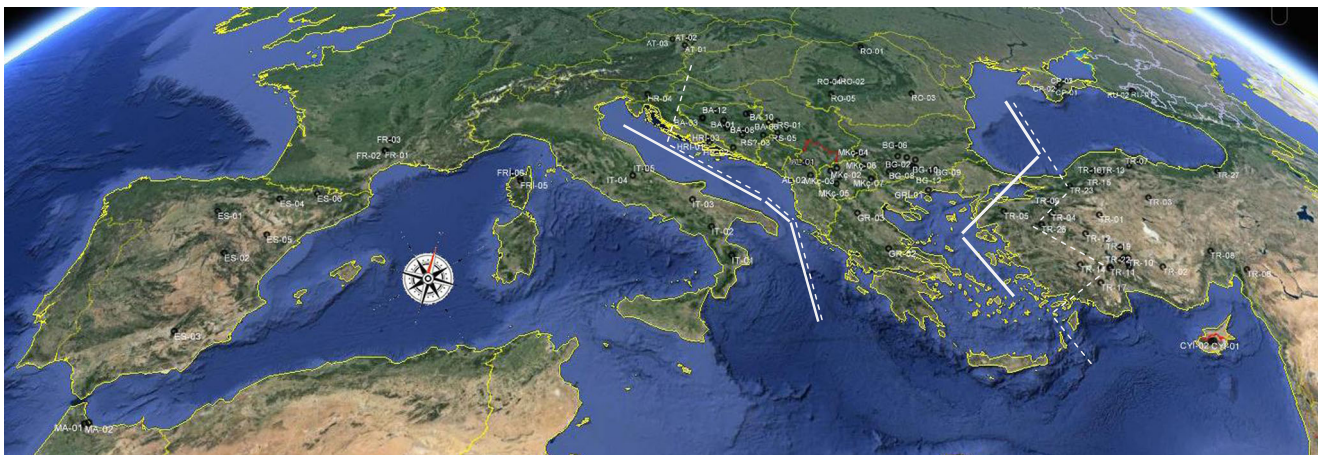
Fig. 1 The studied natural populations of *Pinus nigra* (Arn.) and the principal edges of genetic diversity determined by the BARRIER program (*dashed line*), and the model-based method of STRUCTURE/TESS programs (*solid line*)

group of populations in the large area of the Black Pine's natural distribution, we choose the non-mutation distance in the among-population analyses. The unweighted pair group method with arithmetic mean (UPGMA) algorithm for tree building was preferred to the neighbour joining (NJ) method after preliminary testing using the TreeFIT v.1.0 software (R^2_{av} = 0.94; Kalinowski 2009). The Kalinowski's R^2_{av} is the average proportion of variation in the matrix of genetic distances between populations that is explained by the trees at different scenarios of migration rates (*m*: 0.01; 0.001; 0.0001; 0.00001; 0.000001) and divergence times for the populations in the generations before sampling (*t*: 100; 500; 1000; 5000; 10,000; 50,000; 100,000), data not shown. The distances were calculated using PowerMarker v. 3.0 (Liu & Muse 2005) with 10^3 bootstrap replicates. The pairwise genetic distances, i.e. distance-based method have been used with success with cpSSRs for multiple species (Terrab et al. 2006; Wang et al. 2011; Vinceti et al. 2013). Monmonier's maximum difference algorithm implemented in BARRIER v. 2.2 (Manni et al. 2004) with 10^3 bootstrap replications and 1000 distance matrices from Rogers (1972) was used as a second strategy to study the natural population structure determined here using genetic distances combined with geographic data. The BARRIER program has often been successfully used for uniparental markers (Dzialuk et al., 2009; Hohn et al., 2009; Hodel & Gonzales 2013).

The geographic distribution of the haplotype frequencies was analysed using a Bayesian clustering algorithm (BYM) from two different programs for additional verification: STRUCTURE v.2.2 (Pritchard and Wen 2003; Falush et al. 2003a) and TESS v. 2.1 (Chen et al. 2007). The STRUCTURE program was successfully applied to linked loci to define bacterial populations (Falush et al. 2003b) and to plastid DNA (uniparental markers) to assess the population structure of plants (Pico et al. 2008; Delplancke et al. 2012; Perdereau et al. 2014). We used

the chloroplast haplotype frequencies to perform the model-based methods (i.e. a single locus with multiple alleles) as described in Coart et al. (2006) and Bonavita et al. (2015). The results of the model-based methods were used if they agreed with the results from the distance-based methods.

Initially, we expected that the STRUCTURE program would be better adapted for the cpDNA data; however, TESS was a better choice for the spatial population genetic analyses because it used the hierarchical MRF model from Voronoi and Dirichlet tessellation (François et al. 2006) that was suitable for species with fragmented and large natural distribution areas, such as the European Black Pine. We used two models (*without admixture* and *with admixture*) of natural populations and individuals. For the without admixture model, the maximal number of clusters (K_{max}) was initially set to range between 2 and 33 (i.e. approximately one third of the population numbers). At higher numbers of clusters ($K_{max} > 33$), the likelihood history was unstable and the obtained data were considered unusable. For the with admixture model, the K_{max} was initially set to range between 2 and 8 (i.e. close to the result of the BARRIER program). At higher numbers of clusters ($K_{max} > 6$), the likelihood profile was unstable and the obtained data were considered unusable. The true number of clusters (K_{true}) was determined by two methods: (1) an estimate of the posterior probability of the data for a given K by calculating Pr(KjX) and Pr(XjK) as described by Pritchard et al. (2000) and (2) K_{true} determined by $\Delta K = m(|L''(K)|)/s[L(K)]$, i.e. the mean likelihood L(K) values and a second-order rate of change of the likelihood (ΔK) calculated for three values of the interaction parameter ψ (0.6, 1.2 and 2.4) from 100 runs of 6×10^4 MCMC sweeps with 10^4 burn-in sweeps as recommended by Evanno et al. (2005). The K_{true} was compared to the results of the BARRIER program and the Rogers (1972) distance as described by Pritchard et al. (2000).

Additionally, the N_{st} and G_{st} indices were calculated from 10^4 permutations by the SPAGEDI v. 1.2 software (Hardy & Vekemans 2002) and GENETIX v. 4.05.2 (Belkhir 2000). The comparison between N_{st} vs. G_{st} was used to determine the presence of the phylogeographic structure (at $N_{st} > G_{st}$) conceived by the SSRs in the different groups of natural populations from the without admixture and with admixture models. The indices are calculated by 10^4 permutations by the SPAGEDI software and GENETIX. The population structure analyses were completed by hierarchical tests using AMOVA (with 10^5 permutations) in ARLEQUIN. To detect the spatial edges of differentiated groups of populations, we used BARRIER v. 2.2 (Manni et al. 2004) with 10^3 bootstrap replications.

Coalescence analysis

The coalescence process was used to model the ancestry of the samples and the variability-reducing events (Kingman 1982; Hudson 1990). Two sets of groups (AA-BB and BB-CC) were analysed with the BATWING v.1.03 software (Wilson et al. 2003); the principal Black Pine formations: AA (western Mediterranean), BB (the Balkan Peninsula) and CC (Asia Minor), were determined from the *admixture model* of the spatial population genetic analyses (BYM algorithm). BATWING's algorithm is Bayesian, which is adapted for haplotype data and assumes that the effect of selection is negligible. The population size, time of most recent common ancestor (TMRCA) and divergent time among the groups of populations were calculated by the Markov chain Monte Carlo (MCMC) algorithm to generate random samples from the posterior distribution of the population genetic parameter. The BATWING program has often been successfully used for uniparental markers (Kayser et al. 2006; Raffi & Dodd 2007; Balaresque et al. 2010; Duminil et al. 2010). The priors of the mutation per generation mu (μ) range varied from 3.2×10^{-5} to 8×10^{-5} (average 5.6×10^{-5}) according to Provan et al. (1999), and the priors of the effective population size was $v_1 = 4$ and $v_2 = 250{,}000$ in a pure linear growth population size model with a burn-in of the MCMC algorithm (10^5) and number of sample output equal to 10^4 interactions.

The number of years for one generation is very difficult to determine for forest trees and depends on many factors. Thus, reproductive activity, period of intensive body growth and longevity are more important. These factors have a strong relationship and determine the effective capacity of one individual to transmit its genome to the next generation (Pianka 1978; Lande et al. 1999, 2003; Spellman & Klicka 2006). For the European Black Pine, we calculated the TMRCA and more recent split time with three different generation ranges (20, 68 and 100 years). The first range (20 years per generation) corresponded to the average age of the starting reproductive activity. The second range (68 years per generation) corresponded to the average age of the limit of the period of intensive body growth from the timber harvesting growth table. The third range (100 years per generation) corresponded to the average age of the Black Pine's life limit. We expect that this question is broadly disputable. The number of years per generation depends on many ecological conditions and can vary at different times and different natural distribution areas. For this reason, the time calibration in the present publication is treated with caution.

The results of the coalescence analysis are presented by three different quantiles (5, 50 and 95 %), three different mutation rates per generation (3.2×10^{-5}, 8×10^{-5} and average) and three different numbers of years per generation (20, 68 and 100). This approach generated nine different scenarios for the posterior ancestral population size (N_a) and the relative sub-cluster size and 27 different scenarios for the TMRCA and time of the most recent split. To make the choice closer to reality, we calibrated the results, i.e. we compared the obtained time from different scenarios with climatic records timed from the same region as the Last Glacial Maximum (LGM; 0.01–0.02 Ma), the Early–Middle Pleistocene Transitions (0.892–1.2 Ma) and the Messinian Salinity Crisis (5.33–5.96 Ma). Mediterranean region tectonic history studies are used as a complementary method (i.e. from southern Europe, northern Africa and Asia Minor). The scenario obtained from the TMRCA and time of the most recent split calibration is used for the posterior ancestral population size (N_a) and relative sub-cluster size calculations. This scenario is based upon three guiding principles: (1) focus on the quantiles 95 %—i.e. high confidence scenario; (2) the period of most recent split must be from the same period (or after it) of important climatic records timed from the same region; (3) if there are no matches, the analysis must be performed with modification of the sample size and mutation rate (μ).

Results

A total of 272 distinct haplotypes were detected within the 13 examined cpDNA microregions (seven SSRs and six SNPs; Table S1 in the Appendix). We used the maximum number of chloroplast markers that were similar to other Black Pine studies. The region sequences with Pt-30204 have two independent repetitive sequences, whereas the remaining primer sets had one. The GenBank numbers of the sequences were doi:10.1007/s11295-016-1022-y (Data Accessibility-1). SNPs were detected by chance in the flanking regions of the SSRs. We integrated the SNPs to increase the power of the statistical analysis due to their low mutation rates, i.e. two haplotypes with the same SSR length, but different SNPs in the flanking regions were interpreted as two different haplotypes (Table S2 in the Appendix). The

method applied here has a missed interpolation error of 33–50 % less than that of published results without the sequencing analysis control.

A total of 48.9 % of the haplotypes were detected in less than three individuals. The number of distinct haplotypes specific for different geographic regions ranged from 4 to 136 (Fig. S1 in the Appendix). The gene diversity at these size-variant loci ranged from 0.123 (Pt-36480) to 0.716 (Pt-71936). We performed rarefactions to measure the size variant richness in the population and showed that 30 (at $p \leq 0.05$) and 50 (at $p \leq 0.01$) individuals per population (or groups of populations) were sufficient to accurately estimate the size variant frequencies (Data Accessibility-2). This parameters varied between natural populations and regions, but there was no evidence of a clear east-west divide. The average number of distinct haplotypes per population was 12.8, with a range from four to 27, a mean diversity index (H_{cp}) of 0.288 (0.192 to 0.442) and an average nucleotide diversity of 0.029 (Table 1). The average level of association between the number of haplotypes (N_{hap}) and the total number of analysed individuals per population (N_{ind}) is 0.42. Both parameters—N_{hap} and N_{ind}—are independent according to results of the T test ($t = 504.6$ with 208 degrees of freedom at $p = <0.001$), with weak correlation between them ($r^2 = 0.187$; at $p = 0.05$).

Present population structure

The present population structure was investigated using two approaches based on the following: (1) genetic distances irrespective of geographic distances using Rogers (1972) geometric distances (both genetic and geographic distances) in the BARRIER program and (2) combined distance and geographic parameters from the model-based methods (STRUCTURE and TESS programs).

The result of the first approach using UPGMA clustering with Rogers (1972) geometric distances indicated the presence of three large clusters (with ≥50 % bootstrap support) and an association of the haplotypes with the geographical regions (Figs. S1 and S2 in the Appendix). Each of the large UPGMA clusters contained at least four sizeable sub-clusters with bootstrap support. Finer geographical structuring was also apparent in the clusters or sub-clusters in several geographic regions. The first UPGMA cluster included most natural populations from west of the Adriatic and Ionian Sea basins France, Italy, Morocco and Spain (FR, IT, MA and ES) and three natural populations located in the western Balkan Peninsula (BA-11 and GR-2) and Asia Minor (TR-24). The second and third cluster generally represented natural populations from the Balkan Peninsula and Asia Minor, respectively. The misclassified

populations in the three principal clusters were most often observed in the Asia Minor formation, followed by the Balkan Peninsula and Western Europe. These misclassified populations are indicators of a limit of this algorithm for a large number of populations; thus, it was necessary to complete the analysis with different methods, and the presence of admixing (i.e. an admixture of a significant number of local and universal haplotypes; Figs. S1 and S2 in Appendix). The hierarchical AMOVA analyses also showed substantial genetic variation among groups (clusters) with $\Phi_{CT} = 0.081$ and among populations with $\Phi_{SC} = 0.066$ (Table S3 in the Appendix). The calculated F_{st} and G_{st} values suggested that migration and genetic drift contributed significantly to the genetic natural population structure.

The BARRIER results indicated the presence of six barriers with more than 50 % bootstrap support that were clearly consistent with the UPGMA, *no admixture* and *admixture* clusters (Fig. 2). The highest rate of change (0.80) was found at the barriers ii-iv and iii-iv between the Black Pine admixture clusters AA and BB, and (0.76) at the barrier v-v between admixture clusters BB and CC, followed by the barrier (i-i) coinciding with the Rhone valley (0.65). The other barrier (vi-vi) appeared to be less important (0.61) and was located in the extreme eastern part of the distribution area of *P. nigra*. The AMOVA results confirmed the relative significance of the barriers with a Φ_{CT} of 0.018–0.020 for the strong barrier vs. 0.003 for the less pronounced edges of the combined genetic/geographic rates of change (Table S3 in the Appendix).

The UPGMA distance analysis of the population structure did not consider prior geographic data. Therefore, we complemented this analysis with an allele frequency test of the group-individual-population structure combined with geographic data in space via Bayesian analysis using two scenarios: without admixture (no-admixture) and with admixture (STRUCTURE and TESS programs). The with-admixture scenario indicated that the optimal number of clusters (K_{true}) was three for all trends (data not shown). The numbers of natural populations in the with-admixture clusters were approximately proportional to the present natural distribution area of the Black Pine, with 19, 53 and 34 populations in the three formations (designated AA—western Mediterranean, BB—the Balkan Peninsula, and CC—Asia Minor, respectively; Fig. 3). The population/individual membership was determined by estimating the admixture proportions at 90 % probability intervals.

Two large edges of the admixture were identified: (AB) in the Adriatic and Ionian Sea basins and (BC) in the basins of the Aegean, Marmara and Black Seas. The first with admixture cluster (AA) was predominantly comprised of individuals from regions west of the Adriatic and Ionian Sea basins. The second (BB) included individuals predominantly from the Balkan Peninsula and

Table 1 The studied natural populations of *Pinus nigra* (Arn.) along with details on genetic diversity

Country	Population name	Lat.N/ Long.	N	N_{hap}/N_{ind}	Nuc. div (SE)	H_{cp} (SE)
AL-01	Laithiza	42.083/20.167	36	0.39	0.034 (0.019)	0.308 (0.020)
AL-02	Belgaj	41.583/20.183	20	0.45	0.032 (0.018)	0.339 (0.029)
AT-01	St. Margarethen	47.769/16.609	20	0.30	0.031 (0.018)	0.326 (0.022)
AT-02	Stopfenreuth	48.143/16.863	36	0.28	0.033 (0.018)	0.275 (0.018)
AT-03	Wiena	48.133/16.100	30	0.37	0.039 (0.021)	0.262 (0.024)
BA-01	Visegrad	43.806/17.423	36	0.47	0.031 (0.017)	0.261 (0.020)
BA-02	Bugojno	44.026/17.413	22	0.32	0.026 (0.015)	0.331 (0.032)
BA-03	Olovo	44.188/16.572	36	0.61	0.030 (0.017)	0.281 (0.022)
BA-04	Zepce	43.367/18.000	20	0.40	0.032 (0.018)	0.307 (0.025)
BA-05	Vran	43.694/17.470	36	0.39	0.027 (0.015)	0.247 (0.023)
BA-06	Sjemec	43.817/19.250	22	0.45	0.030 (0.017)	0.328 (0.030)
BA-07	Bosansko Grahovo-1	44.250/18.417	36	0.75	0.024 (0.014)	0.223 (0.023)
BA-08	Borja	43.533/18.150	36	0.56	0.030 (0.017)	0.252 (0.018)
BA-09	Konjic	43.533/18.133	22	0.45	0.026 (0.015)	0.267 (0.028)
BA-10	Kladanj	44.269/18.584	20	0.40	0.033 (0.019)	0.269 (0.027)
BA-11	Bosansko Grahovo-2	44.200/16.567	20	0.40	0.033 (0.019)	0.286 (0.029)
BA-12	Donji Ugar	44.467/17.233	20	0.50	0.025 (0.015)	0.257 (0.031)
BG-01	Bansko	41.800/23.500	36	0.47	0.031 (0.017)	0.285 (0.018)
BG-02	Devin	41.900/24.667	36	0.69	0.034 (0.019)	0.201 (0.022)
BG-03	Hvoina	41.600/24.500	36	0.67	0.025 (0.014)	0.228 (0.028)
BG-04	Mesta	41.700/23.667	36	0.36	0.032 (0.018)	0.278 (0.022)
BG-05	Nevestino	42.250/22.600	22	0.36	0.026 (0.015)	0.267 (0.028)
BG-06	Panagurichte	42.500/24.200	20	0.35	0.023 (0.014)	0.328 (0.029)
BG-07	Pazardjik	42.012/24.359	36	0.44	0.023 (0.013)	0.301 (0.030)
BG-08	Sandanskii	41.500/23.333	36	0.42	0.032 (0.018)	0.297 (0.020)
BG-09	Smolian	41.500/25.326	36	0.47	0.020 (0.012)	0.362 (0.025)
BG-10	Trigrad	41.667/24.417	36	0.53	0.028 (0.016)	0.257 (0.028)
BG-11	Velingrad	42.083/24.000	30	0.63	0.025 (0.015)	0.304 (0.027)
BG-12	Zlatograd	41.417/25.083	36	0.53	0.025 (0.015)	0.267 (0.027)
CP-01	n.i.	44.665/33.834	18	0.56	0.028 (0.016)	0.367 (0.033)
CP-02	Pionerskaja	44.833/34.238	36	0.42	0.025 (0.014)	0.282 (0.020)
CP-03	Kyzylivka	45.000/34.688	20	0.50	0.028 (0.016)	0.367 (0.031)
CY†-01	Amargeti	34.833/32.583	36	0.33	0.039 (0.021)	0.316 (0.023)
CY†-02	Salamiou	34.834/32.697	36	0.28	0.032 (0.018)	0.265 (0.020)
ES-01	Huerta del Rey	41.799/−2.833	36	0.36	0.026 (0.015)	0.273 (0.020)
ES-02	Isérico	40.250/−1.833	36	0.19	0.026 (0.015)	0.286 (0.018)
ES-03	Sierra Majiuz	37.749/−2.664	36	0.33	0.027 (0.015)	0.259 (0.025)
ES-04	La Pobla de Denifassa	42.101/−0.331	15	0.33	0.019 (0.012)	0.343 (0.031)
ES-05	Antoli	40.800/−0.350	36	0.28	0.037 (0.020)	0.283 (0.022)
ES-06	Sur de Cuencia	42.199/1.183	24	0.38	0.022 (0.013)	0.288 (0.029)
FR†-04	Vallée de la Restonica	42.286/9.124	36	0.50	0.028 (0.016)	0.271 (0.027)
FR†-05	Col de Vizzavona	42.150/9.150	36	0.53	0.032 (0.018)	0.258 (0.027)
FR†-06	Fotêt de Bonifatu	42.442/8.858	37	0.49	0.033 (0.018)	0.266 (0.026)
FR-01	Montpellier	43.750/3.517	34	0.32	0.014 (0.009)	0.283 (0.026)
FR-02	St.Guilhem le desert	43.735/3.541	35	0.49	0.029 (0.016)	0.277 (0.027)
FR-03	Vallée de la Buége	44.167/3.683	36	0.53	0.030 (0.017)	0.273 (0.023)
GR†-01	Tasos	40.700/24.570	36	0.28	0.028 (0.016)	0.295 (0.025)
GR-02	Kolokithia	38.800/22.083	36	0.50	0.030 (0.017)	0.293 (0.027)
GR-03	Grevene	40.080/21.436	36	0.61	0.027 (0.015)	0.207 (0.022)

Table 1 (continued)

Country	Population name	Lat.N/ Long.	N	N_{hap}/N_{ind}	Nuc. div (SE)	H_{cp} (SE)
HR†-01	Vela Grablje, Hvar	43.159/16.616	36	0.28	0.039 (0.021)	0.281 (0.015)
HR†-03	Brac	43.300/16.600	20	0.35	0.037 (0.021)	0.283 (0.020)
HR-02	Rijeka	42.867/17.500	15	0.33	0.035 (0.020)	0.429 (0.023)
HR-04	Podgorje, Peljesac	45.427/14.468	20	0.35	0.040 (0.022)	0.276 (0.027)
IT-01	San Pietro Apostolo	39.000/16.500	24	0.38	0.023 (0.014)	0.337 (0.027)
IT-02	Castelluccio Superiore	40.004/15.999	20	0.40	0.023 (0.014)	0.342 (0.029)
IT-03	Monteverde	41.001/15.499	18	0.33	0.022 (0.013)	0.244 (0.033)
IT-04	Avezzano	42.051/13.418	36	0.53	0.030 (0.017)	0.254 (0.025)
IT-05	Fara	42.248/13.968	36	0.39	0.032 (0.018)	0.225 (0.023)
MA-01	Talassemtane	35.166/−4.635	36	0.42	0.031 (0.017)	0.316 (0.022)
MA-02	Chefchaonen	35.133/−4.767	36	0.39	0.026 (0.015)	0.268 (0.028)
MK‡-01	Karadzica	41.787/21.323	15	0.33	0.029 (0.017)	0.331 (0.021)
MK‡-02	Gostivar	41.692/21.034	36	0.36	0.032 (0.018)	0.248 (0.022)
MK‡-03	Kucevo	41.467/21.297	36	0.44	0.033 (0.018)	0.255 (0.025)
MK‡-04	Kriva Palanka	42.323/22.325	15	0.47	0.034 (0.020)	0.334 (0.036)
MK‡-05	Resen	41.199/21.063	15	0.40	0.036 (0.021)	0.336 (0.023)
MK‡-06	Stip	41.786/22.419	15	0.40	0.026 (0.016)	0.297 (0.034)
MK‡-07	Valandovo	41.285/22.502	15	0.47	0.025 (0.015)	0.289 (0.028)
RO-01	Scrovistea Dis.	47.200/25.083	36	0.47	0.032 (0.018)	0.259 (0.022)
RO-02	Dobra seed	45.669/23.250	36	0.44	0.027 (0.016)	0.245 (0.025)
RO-03	Bihor Country	44.750/26.167	36	0.39	0.033 (0.018)	0.315 (0.017)
RO-04	Domogled Mountain	45.470/22.715	36	0.31	0.027 (0.015)	0.322 (0.025)
RO-05	Cerna Mountain	44.924/22.446	36	0.39	0.029 (0.016)	0.262 (0.027)
RS-01	Zlatibor	43.853/19.623	36	0.56	0.028 (0.016)	0.275 (0.028)
RS -02	Pljevlje	43.358/19.457	36	0.56	0.035 (0.019)	0.284 (0.030)
RS -03	Kanjon Tare	43.151/19.294	36	0.69	0.034 (0.019)	0.268 (0.028)
RS -04	Kanjon Pive	43.301/18.848	36	0.69	0.026 (0.015)	0.192 (0.025)
RS-05	Zlatar	43.498/19.896	24	0.83	0.030 (0.017)	0.311 (0.022)
RU-01	Beregovoe	44.417/38.417	36	0.33	0.024 (0.014)	0.276 (0.018)
RU-02	Osipovka	44.400/38.517	36	0.31	0.025 (0.015)	0.269 (0.023)
TR-01	Cifteler	39.290/31.192	20	0.55	0.044 (0.024)	0.308 (0.034)
TR-02	Cumra	37.300/32.508	36	0.19	0.041 (0.022)	0.283 (0.027)
TR-03	Akcakent	39.745/34.094	36	0.36	0.030 (0.017)	0.207 (0.025)
TR-04	Tausanli-Balikoy	39.484/29.119	36	0.53	0.026 (0.015)	0.229 (0.027)
TR-05	B. Gamurlu	39.722/27.186	36	0.33	0.048 (0.025)	0.303 (0.023)
TR-06	Jezirkopru-Golkoy	36.850/36.383	35	0.46	0.034 (0.019)	0.220 (0.024)
TR-07	Osmaniye	41.167/35.050	36	0.28	0.041 (0.022)	0.261 (0.027)
TR-08	Kastamonu	37.628/35.244	36	0.58	0.025 (0.014)	0.239 (0.023)
TR-09	Balikesir	39.925/29.367	22	0.32	0.028 (0.016)	0.346 (0.030)
TR-10	Manisa	37.721/31.238	20	0.30	0.024 (0.014)	0.350 (0.031)
TR-11	Denizli	37.456/30.204	36	0.36	0.028 (0.016)	0.291 (0.025)
TR-12	Bursa	38.680/30.056	20	0.60	0.034 (0.019)	0.222 (0.031)
TR-13	Afyon	41.042/34.367	36	0.44	0.024 (0.014)	0.367 (0.025)
TR-14	Ankara	37.681/29.072	36	0.42	0.026 (0.015)	0.269 (0.023)
TR-15	Bolu	40.824/32.656	36	0.36	0.024 (0.014)	0.314 (0.018)
TR-16	Tota	40.956/32.283	18	0.28	0.018 (0.011)	0.320 (0.021)
TR-17	Tesildag	37.036/29.486	20	0.35	0.023 (0.014)	0.308 (0.025)
TR-18	Orenkoy	37.503/30.254	20	0.20	0.014 (0.009)	0.337 (0.000)
TR-19	Koyduzu	38.108/31.222	20	0.20	0.018 (0.011)	0.355 (0.020)

Table 1 (continued)

| Country	Population name	Lat.N/ Long.	N	N_{hap}/N_{ind}	Nuc. div (SE)	H_{cp} (SE)
TR-20	Sanli	37.691/31.044	20	0.20	0.025 (0.015)	0.442 (0.022)
TR-21	Y-Gokdere	37.431/31.053	20	0.35	0.025 (0.015)	0.329 (0.025)
TR-22	K-Belentepe	37.669/30.844	36	0.42	0.032 (0.018)	0.216 (0.027)
TR-23	Erikli	40.447/30.724	24	0.54	0.026 (0.015)	0.312 (0.029)
TR-24	Ercekbasi	37.564/31.131	18	0.39	0.023 (0.014)	0.305 (0.024)
TR-25	Aktas	37.756/31.082	30	0.23	0.027 (0.016)	0.301 (0.020)
TR-26	Sindirgi	39.142/28.432	36	0.47	0.028 (0.016)	0.274 (0.028)
TR-27	Trabzon	40.540/39.122	36	0.53	0.029 (0.016)	0.206 (0.027)
Mean			30	0.42	0.029 (0.000)	0.288 (0.000)

Latitude and longitude are in decimal

Abbreviation for populations (alphabetic list*) are as follows: *AL* Albania, *AT* Austria, *BA* Bosnia, *BG* Bulgaria, *CP* the Crime Peninsulas, *CY†* Cyprus island (SW-Cilician Sea), *ES* Spain, *FR* France, *FR†* Corsica island (SW-Ligurian Sea), *GR* Greece, *GR†* Thasos island (N-Aegean Sea or Thracian Sea), *HR* Hrvatska, *HR†* Hvar and Brac islands (E-Adriatic Sea), *IT* Italy, *MA* Morocco, *MK‡* Macedonia (all utilization of "Macedonia" should have been understood as short mentions of "FYROM" for "the Former Yugoslav Republic of Macedonia"), *RO* Romania, *RS* Rep. Serbia, *RS* Montenegro, *RU* Russia, *TR* Turkey, *†* all islands are in Mediterranean Sea, *N* sample size, *N_{hap}/N_{ind}* the ratio between the number of haplotype of total number of analysed individual per population, *H_{cp}* chlorotype diversity (average haplotype per population is 12.77 (min/max: 4/27)), *Nuc. div.* nucleotide diversity is used to measure the degree of polymorphism from sequence data; *SE* standard error (and standard deviation is SD = SE√n)

neighbouring islands, whereas the third (CC) included a high proportion of individuals from Asia Minor, the eastern Black Sea coast and the eastern Mediterranean Sea islands (Fig. 3 and Table 2). The partitioning of genetic variance among the formations was $\Phi_{CT} = 0.054$ and within the formations was $\Phi_{SC} = 0.119$. The optimum number of clusters (i.e. groups) detected assuming no admixture (BYM algorithm) was 22 (K_{true}) with a median interaction coefficient (ψ) of 1.2 and an estimated posterior marginal probability (tau) >88 % (data not shown). Between one and 15 natural populations were present in each no admixture cluster (Fig. 2). Six clusters (no admixture) were located in the region west of the Adriatic and Ionian Sea Basins (clusters 1, 2, 3, 4, 5 and 6); 11 in the Balkan Peninsula (clusters 8 to 17) and the neighbouring islands of Thasos (cluster 21) in the Aegean and Adriatic Seas; one in the region bordering the Eastern Black Sea coast in Russia, Turkey and the Crimean Peninsula (cluster 7); one on the eastern Mediterranean Sea islands Cyprus and Turkey (clusters 18, 19 and 20); and three in Asia Minor (clusters 18, 19 and 20). Hierarchical AMOVA indicated among-group and within-group $\Phi_{CT} = 0.079$ and $\Phi_{SC} = 0.081$ of the variation, respectively. The partitioning of genetic variance among the formations according to the with-admixture hierarchical analysis model was less pronounced than with the no admixture model. Strong phylogeographic patterns were detected in clusters 1, 2, 3, 5, 11, 20 and 22 of the no admixture model, (i.e. $N_{st} > G_{st(nc)}$ at $p < 0.05$ and 2×10^4 permutations; Table 2). The phylogeographic structure has been observed clearing physical barriers, such as the Adriatic

Sea (IT/HR), the Straits of Gibraltar (ES/MA), the western Alps (FR/IT) and lowlands with highly competitive flora in different regions. No significant differences were observed in the F_{st} (*RH* or *RH'*) and N_{st} among any of the clusters identified by the admixture model (AA, BB and CC formations) and among some of the natural populations identified by the no admixture model ($p < 0.05$; Table 2). These findings indicate that the mutation rate is negligible compared to the effects of migration and genetic drift.

The differences between the results obtained using all of the analytical strategies reflect the differences in the methodological algorithms, i.e. population vs. individual assessments, distance-based vs. model-based methods and non-model vs. model-based methods. The results of the model-based methods STRUCTURE and TESS agreed (full or partial match) with the results of the distance-based methods (Roger's geometric distances and the BARRIER program). Approximately 25 % of the clusters from the no admixture model did not match Roger's geometric distances. All clusters from the BARRIER program matched all clusters from the no admixture model with one partial exception (cluster 3). All clusters and their individuals (approximately 80 %) from Roger's geometric distances and from the BARRIER program matched clusters from the admixture model (Figs. 1, 2 and 3; Appendix-Fig S2). The presence of misclassified populations in Roger's geometric distances showed the limit of this method for a large population number. The best distance-based method was the Monmonier's algorithm from BARRIER program.
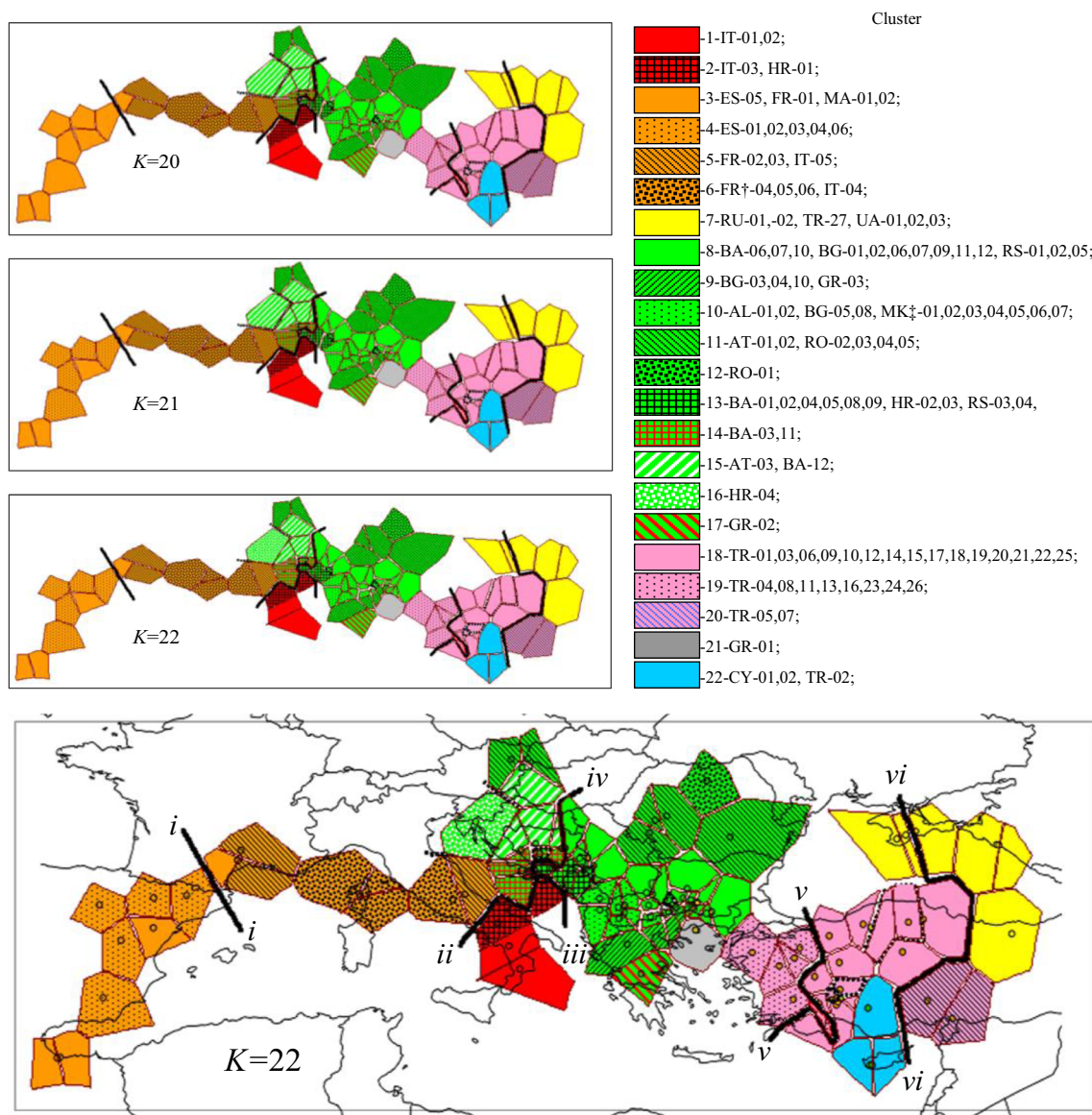
Fig. 2 Bayesian clustering (*coloured polygons*) of European Black Pine populations based of "*none admixing*" clusters algorithm implanted in STRUCTURE v.2.2 (K_{max} = 21 and K_{true} = 22) and TESS v.2.1 (K_{max} = 20 and K_{true} = 21). The *bold lines* are the genetic barriers detected with Monmonier's maximum difference algorithm of Manni et al. (2004; BARRIER v. 2.2 program)

Analysis of historic population demographics

The more recent coalescence events, i.e. time of the most recent split and time of the most recent common ancestor (TMRCA), were analysed for the populations of the three main admixture clusters (AA, BB and CC formations) determined by the Bayesian assignment test (Table 3). The results show that none of the 27 coalescent event scenarios were associated with the LGM (i.e. less than 20–25 thousand years ago). The best fit was observed for one scenario with 95 % quantiles, an average mutation rate of $\mu = 5.6 \times 10^{-5}$ per generation and a 68-year lapse of time per generation, i.e. calibration with the timing of the Early–Middle Pleistocene Transitions and the Messinian Salinity Crisis. All data

interpretation and the following discussion are based on this scenario. We interpreted the time of the paleo-climatic and tectonic events with caution because the different methodology used for the time calculation from the paleoclimatology and tectonic dynamics might vary to some extent. The most recent splits appear to have occurred between the natural populations of the Balkan Peninsula and Asia Minor (BC: 0.93 Ma the Early–Middle Pleistocene Transitions) and the natural populations of the Balkan and Apennine Peninsulas (BA: 4.37 Ma after the Messinian Salinity Crisis). The TMRCA follows the same pattern: 10.72 Ma for BB/CC and 14.75 Ma for AA/BB (i.e. from the early Pleistocene-Pliocene-late Miocene; former Tertiary period; Table 3). The posterior ancestral population size (N_a) is 26.0–26.5 $\times 10^4$
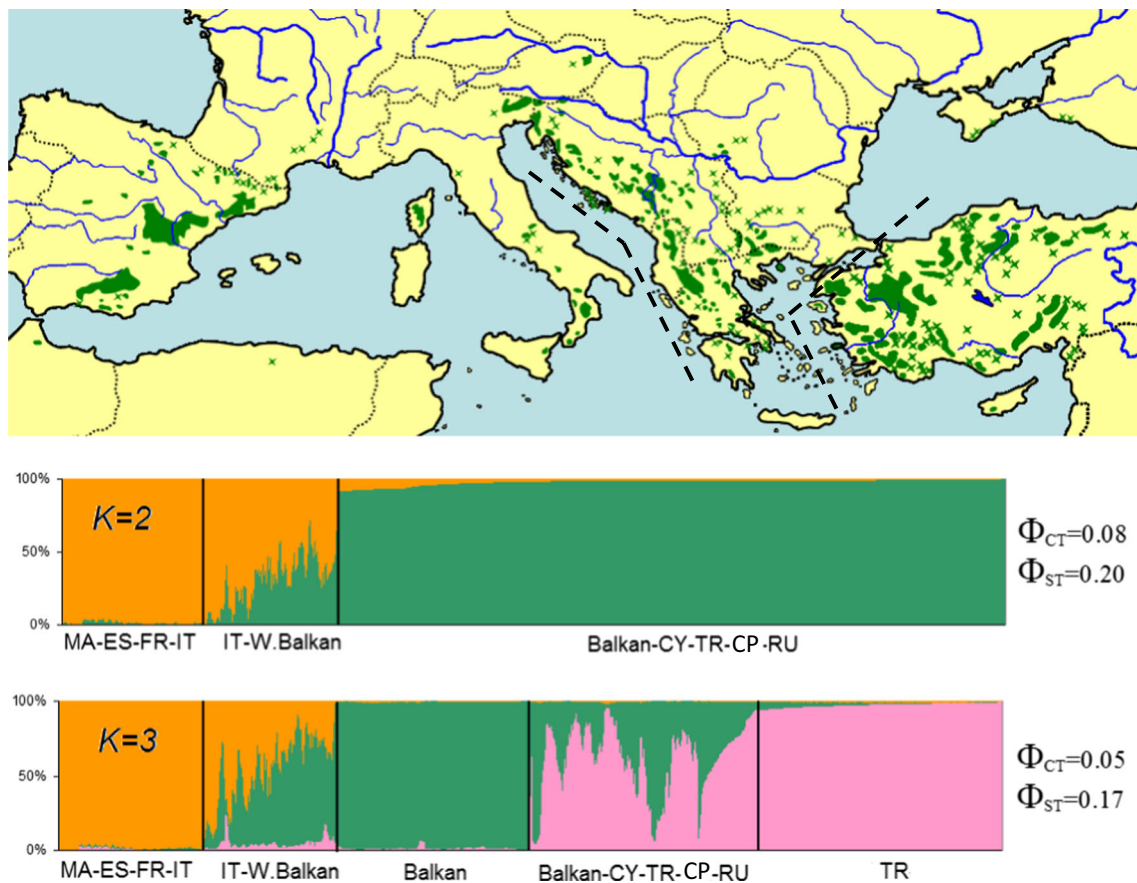
Fig. 3 Map of the natural distribution range of *P. nigra* (*top*) according Critchfield & Little (1966). *Dashed black lines* indicate the two big edges of "*admixture*", AB (*left*) and BC (*right*), according the algorithm implanted in STRUCTURE v.2.2 ($K_{true} = 3$) and TESS v.2.1 ($K_{true} = 3$). The average proportion of individual membership (*bottom*) for each individuals, and the relative sub-cluster size varies between 17.5 and 20.0 × 10⁴ individuals. "*admixture*" cluster AA (*brown*; Westerns Mediterranean Formation), BB (*green*; the Balkan Peninsula Formation) and CC (*rose*; Asia Minor Formation) and the regions of admixture AB (*brown/green*; IT-W.Balkan) and BC (*green/rose*; Balkan-CY-TR-CP-RU); Φ_{CT}—among formations; Φ_{ST}—within formations

individuals, and the relative sub-cluster size varies between 17.5 and 20.0 × 10⁴ individuals.

Discussion

Population demographic history

Our study showed a moderate level of association between the number of haplotypes and the total number of analysed individuals per population ($N_{hap}/N_{ind} = 0.42$), which was close to the results obtained using the same chloroplast microsatellites for Mediterranean pines (*Pinus pinea* and *Pinus halepensis*) by Gomez et al. (2002, 2005), *Pinus pinaster* by Bucci et al. (2007), *Pinus mugo* by Heuertz et al. (2010) and *P. nigra* by Bonavita et al. (2015) and Naydenov et al. (2006). The number of haplotypes per population is generally ≤0.50 for many pines worldwide, including *Pinus banksiana* and *Pinus resinosa* from North America (Walter & Epperson 2005; Naydenov et al. 2005) and *Pinus krempfii* from Asia (Wang et al.

2014). In the Iberian Peninsulas, slightly greater than average $N_{hap}/N_{ind} = 0.61$ (Table 2; nhp15 = 9.16) were reported for *P. nigra* by Soto et al. (2010). We and others using the same chloroplast microsatellites could not confirm the high level of Black Pine haplotypes per population (av. 0.91; Table 1) reported by Raffi & Dodd (2007).

The homoplasy is a phylogenetic phenomenon resulting from some cause other than common ancestry. In evolutionary studies, this phenomenon is of two main types: convergence and reversion. Homoplasy is contrasted with homology. For molecular analysis, the homoplasy is the presence of identical microsatellite electromorphs with heterogeneous DNA sequence (Doyle et al. 1998; Hale et al. 2004). In phylogeography studies, homoplasy has not been a substantial problem for low mutation rates ($\mu < 10^{-4}$) of microsatellites as chloroplast according Navascués & Emerson (2005), and it depends on species, molecular marker type, sample size and the used statistical method (Estoup et al. 2002; Adams et al. 2004). The population genetic studies here are less affected by homoplasy as compared to studies without sequencing control.

Table 2 Genetic diversity parameters and population differentiation estimates performed at 2×10^4 permutations using SPAGEDI v.1.2 (Hardy & Vekemans 2002) and GENETIX v.4.05.2 (Belkhir 2000)

Group/cluster	Populations (country)	Region	N[a]	hh[b]	A[c]	H_{cp}[d]	N_{st}/pN_{st}[e]	F_{st}[f]	R_H/R_H[g]	$G_{st}/G_{st}(nc)$[h]
Non-admixing model										
1	IT-01,02	Apennine Peninsulas	44	9	8.5	0.143	−0.044/−0.044	−0.043	−0.041/−0.029	0.002/−0.021
2	IT-03, HR-01	Apen. and W-Balkan Penin	54	16	8.0	0.213	0.195/0.195	0.163	0.096/0.140	0.128/0.107
3	ES-05, FR-01, MA-01,02	Western Mediterranean	142	24	12.5	0.236	0.098/0.086	0.062	0.081/0.112	0.068/0.048
4	ES-01,02,03,04,06	Iberian Peninsulas	147	23	9.2	0.208	0.068/0.066	0.050	0.052/0.088	0.069/0.039
5	FR-02,03, IT-05	Western-Central Medit.	107	32	16.7	0.272	0.092/0.086	0.050	0.051/0.097	0.053/0.034
6	FR-04,05,06, IT-04	Corse and Apennine Penin	145	34	18.5	0.257	−0.004/−0.004	−0.005	0.001/0.004	0.017/−0.004
7	RU-01,02, TR-27, CP-01,02,03	Eastern-Black See coastal	182	39	12.8	0.251	0.040/0.039	0.036	0.034/0.048	0.059/0.030
8	BA-06,07,10, BG-01,02,06,07,09,11,12, RS-01,02,05	Balkan Peninsulas (SE)	404	102	18.1	0.231	0.070/0.066	0.049	0.041/0.085	0.079/0.049
9	BG-03,04,10, GR-03	Balkan Peninsulas (SE)	144	52	19.5	0.244	0.073/0.071	0.058	0.043/0.185	0.065/0.044
10	AL-01,02, BG-05,08, MK-01,02,03,04,05,06,07	Balkan Peninsulas	261	35	9.8	0.226	0.031/0.029	0.033	0.027/0.040	0.077/0.034
11	AT-01,02, RO-02,03,04,05	Balkan Peninsulas (NW)	200	37	11.8	0.232	0.081/0.077	0.040	0.056/0.099	0.061/0.036
12	RO-01	Balkan Peninsulas (N)	36	17	17.0	0.260	n.a.	n.a.	n.a.	n.a.
13	BA-01,02,04,08,09, HR-02,HR-03, RS-03,04	Balkan Peninsulas (W)	279	74	13.8	0.225	0.081/0.075	0.064	0.064/0.201	0.099/0.064
14	BA-03,11	Balkan Peninsulas	56	28	15.0	0.276	0.103/0.103	0.172	0.104/0.149	0.114/0.094
15	AT-03, BA-12	Balkan Peninsulas (W)	50	20	10.5	0.262	0.022/0.022	0.010	0.027/0.043	0.027/0.005
16	HR-04	Balkan Peninsulas (W)	20	7	7.0	0.218	n.a.	n.a.	n.a.	n.a.
17	GR-02	Balkan Peninsulas (S)	36	18	18.0	0.270	n.a.	n.a.	n.a.	n.a.
18	TR-01,03,06,09,10,12,14,15,17,18,19,20,21,22,25	Asia Minor (C)	408	56	9.4	0.233	0.139/0.135	0.141	0.096/0.219	0.174/0.141
19	TR-04,08,11,13,16,23,24,26	Asia Minor (W)	224	59	13.9	0.248	0.102/0.100	0.064	0.050/0.103	0.108/0.078
20	TR-05,07	Asia Minor (E)	71	31	11.0	0.231	0.080/0.080	0.037	0.032/0.043	0.033/0.019
21	GR-01	Balkan Peninsulas (S)	36	10	10.0	0.236	n.a.	n.a.	n.a.	n.a.
22	CY-01,02, TR-02	Eastern Mediterranean island	108	16	9.7	0.293	0.117/0.108	0.184	0.100/0.150	0.148/0.131
All-*non-admixing*			3154	272	12.8	0.239	0.138/0.134	0.086	0.064/0.114	0.106/0.094
Admixing model										
AA-1	IT	Apennine Peninsulas	134	31	11.2	0.230	0.086/0.084	0.077	0.071/0.122	0.100/0.069
AA-2	ES	Iberian Peninsulas	183	28	9.3	0.228	0.075/0.073	0.066	0.069/0.137	0.084/0.054
AA-3	FR	Westerns Mediterranean isl.	109	26	18.3	0.255	−0.013/−0.013	−0.015	−0.011/−0.010	0.009/−0.010
AA-4	FR	France continental	105	24	15.7	0.230	0.042/0.046	0.031	0.028/0.037	0.040/0.021
AA-5	MA	North-West Africa coastal	72	16	14.5	0.228	−0.015/−0.015	−0.019	−0.017/−0.016	0.078/0.060
BB	AL, AT, BA, BG, HR, HR, GR, GR, MK, RO, RS	Balkan Peninsulas	1558	182	13.8	0.245	0.092/0.088	0.077	0.082/0.131	0.113/0.079
CC-1	TR	Asia Minor	775	85	11.1	0.245	0.125/0.121	0.141	0.096/0.161	0.177/0.146
CC-2	CY	Eastern Mediterranean island	72	12	11.0	0.273	−0.005/−0.005	−0.004	−0.007/−0.007	0.012/−0.002
CC-3	RU, CP	Eastern Black Sea coastal	146	25	11.6	0.245	0.039/0.038	0.035	0.048/0.069	0.055/0.025
All-*admixing*			3154	272	12.8	0.242	0.125/0.120	0.085	0.051/0.131	0.121/0.115

Estimates for group of populations of European Black pine delineated by TESS v.2.1 (Chen et al. 2007) for *non-admixing* (cluster 1 to 22) and *admixing* (cluster AA to CC) models and maximizing geographic structure

[a] Sample size

[b] Number of distinct multi-locus haplotypes (chlorotypes) observed in each group

[c] Average number of chlorotypes per population in each group

[d] Average chlorotypes diversity in each group

[e] Global N_{st}/permuted N_{st} (Pons & Petit 1996)

[f] F_{st} according Weir & Cockerham (1984)

[g] R_H according Robertson & Hill (1984) and R_H according Raufaste & Bonhomme (2000)

[h] $G_{st}(nc)$—unbiased estimators according Nei & Chesser (1983); R_H according Belkhir (2000) is the best index if $0 < F_{st} < 0.05$; R_H according Belkhir (2000) is a better index if $0.05 < F_{st} < 0.10$; the F_{st} of Weir & Cockerham (1984) is the best for values >0.10

Table 3 Posterior ancestral population size (N_a) in 10^4 individuals, most recent common ancestor in 10^4 years at generation time 20, 68[$] and 100[¥] years (TMRCA years); relative sub-cluster size in 10^4 individuals, time in 10^4 years at generation time 20, 68 and 100 years of the most recent split at prior (μ) mutation rate 3.2 and 8.0 × 10^{-5} per generation and sample size (4 up to 250,000 individuals) for coalescent analysis with Splitting Model at 11 × 10^4 interaction performed using Wilson et al. (2003). The AA, BB and CC are formation ("*admixing*") cluster from the haplotype frequency geographical distribution Bayesian method (BYM)

Parameter	Quantiles		
	5 % Min-max (Av.)	50 %	95 %
N_a			
×10^4 individuals			
(AA),(BB)	8–21 (14.5)	11–28 (19.5)	15–37 (26.0)
(BB),(CC)	8–21 (14.5)	11–28 (19.5)	15–38 (26.5)
TMRCA			
×10^4 years			
(AA),(BB)			
	49–116 (82)	103–243 (173)	262–606 (434)
[$]	167–394 (280)	350–826 (588)	891–2060 (1475)
[¥]	245–580 (412)	515–1215 (865)	1310–3030 (2170)
(BB),(CC)			
	42–102 (72)	84–206 (145)	183–448 (315)
[$]	143–346 (244)	285–700 (492)	622–1523 (1072)
[¥]	210–510 (360)	420–1030 (725)	915–2240 (1577)
Time of most recent split			
×10^4 years			
AB			
	6–18 (12)	19–53 (36)	62–195 (128)
[$]	20–61 (40)	65–180 (122)	211–663 *(437)*
[¥]	30–90 (60)	95–265 (180)	310–975 (642)
BC			
	3–9 (6)	7–18 (12)	15–40 (27)
[$]	10–31 (20)	24–61 (42)	51–136 *(93)*
[¥]	15–45 (30)	35–90 (62)	75–200 (137)
Relative sub-cluster size			
×10^4 individuals			
AA	3–8 (5.5)	6–14 (10.0)	10–25 (17.5)
BB	5–12 (8.5)	8–19 (13.5)	11–29 (20.0)
CC	3–8 (5.5)	6–15 (10.5)	10–26 (18.0)

The data in italics in the table for the quantile 95 % from the "time of most recent split" indicate a calibration match with the Early–Middle Pleistocene transitions (0.892–1.2 Ma) and the Messinian Salinity Crisis (5.33–5.96 Ma)

The level of differentiation among populations found in our study (Φ_{SC} = 0.07–0.13) was similar to the level reported for the Stone and Aleppo Pines by Gomez et al. (2002, 2005), the *P. pinaster* by Bucci et al. (2007), the Mountain Pine by Heuertz et al. (2010) and the Black Pine by Bonavita et al. (2015) for the same geographic region. Generally, the results for paternally inherited DNA (cpSSR) from coniferous species (*Abies*, *Picea* and *Pinus*) showed ≤0.15 among-population variability, including *Picea abies* by Scotti et al. (2006) from Europe, *Pinus sylvestris* by Semerikov et al. (2014) from Siberia-Asia and *Abies balsamea* by Cinget et al. (2015) from North America. It seemed that the best strategy is using the BARRIER program (distance-based method) as the principal method, combined with the complementary model-based method of STRUCTURE and TESS programs as necessary.

Genetic analyses identified three differentiated genetic formations that were consistent with the European Black Pine's natural geographic distribution. The ancestors of the principal formations began diverting relatively recently (14.75 Ma for AB and 10.72 Ma for the BC edges) compared with the TMRCA results from the *Haploxylon* and *Diploxylon* sections

of pines (89–90 Ma; He et al. 2012). None of these scenarios from Table 3 showed a population demographic signal from LGM. Paleobotanical research confirmed the presence of *P. nigra* (i.e. its ancestor) in the Balkan Peninsula during the Tertiary (Stefanov 1941/42, 1943; Palamarev 1989). In the period from 10–35 Ma, the Mediterranean region tectonic structure was different from the current structure (Gealey 1989; Stampfli et al. 2002), and a significant part of the Balkan Peninsula and Asia Minor and probably one very small part of Calabria in southern Italy were located in a massive long topographic structure known as the Balkan-Pontides microplate between the Western Tethys and the Paratethys Sea (Fig. 4). The region had multiple moderate elevations, including mountains, latitude variations and a stable tectonic structure for a 25-My period that was favourable for the speciation process. We speculate that the principal Black Pine ancestral populations were located in this area during this period (Fig. 4). In contrast to the Balkan Peninsula, Iberian Peninsula and Asia Minor with their relatively stable tectonic history, the Apennine Peninsula's tectonic history is a puzzle (i.e. non-monolith, unstable and collective over the last 25 My). The period of the speciation process was not sufficient for the Black Pine ancestor to evolve into multiple distinct species; this probably explains the lack of reproductive barriers between different Black Pine sub-species. In contrast to *P. nigra*, some ancestors of the *Pinus* species from Mexico and the western USA exhibit evolution into multiple separated species in the present day (more than 44 according Critchfield & Little 1966) in regions that have been tectonically stable for a very long period (>80 Ma). This stability has resulted in more pronounced isolation between separated populations in the large ecological gradient determined by mountain tops and large latitudinal ranges, i.e. species response with strong

glacial/interglacial cycles from oceanic-continental latitude-altitude gradients over a very long time.

The results of the population structure analysis discriminate between the natural populations of *P. nigra* along the edges of the Adriatic and Ionian Seas (AB), i.e. between the Apennine and Balkan Peninsulas. This edge (AB) divides the Mediterranean basin into two topographically and climatically distinct regions. Ancient climatic fluctuations most likely resulted in different migration histories and forest community associations with different successional patterns (i.e. *interglacial refugia* vs. *macro long-term refugia*). The refuge terminology and classification used here are based on the theory of Stewart et al. (2010).

Many non-pine tree genera are often present in Black Pine forest associations (including *Quercus, Acer, Fraxinus* and *Carpinus*), and their associated fauna follow similar phylogeographic patterns. The AB and BC edges seem to be a complementary edge of east/west admixing for multiple species with large distributions (i.e. not limited to the Mediterranean region). For example, Petit et al. (2002a, b) and Lumaret et al. (2005, 2009) used chloroplast DNA to illustrate the presence of eastern and western lineages of eight different species of oaks; similar results were published for hornbeam by Grivet & Petit (2003) and for three species of European ashes by Heuertz et al. (2006). The edge BC (Aegean and Marmara seas) between the Balkan Peninsula and Asia Minor (Turkey) determined here was ignored for many years because these areas have been sampled less intensely. The east/west phylogeographic structure was reported by Emerson & Hewitt (2005) using mtDNA for animals such as *Erinaceus europeus/concolor* and *Ursus arctos*, insects such as *Chorthippus parallelus* in Europe, and *P. sylvestris* (Naydenov et al. 2007). Interestingly, similar edges are
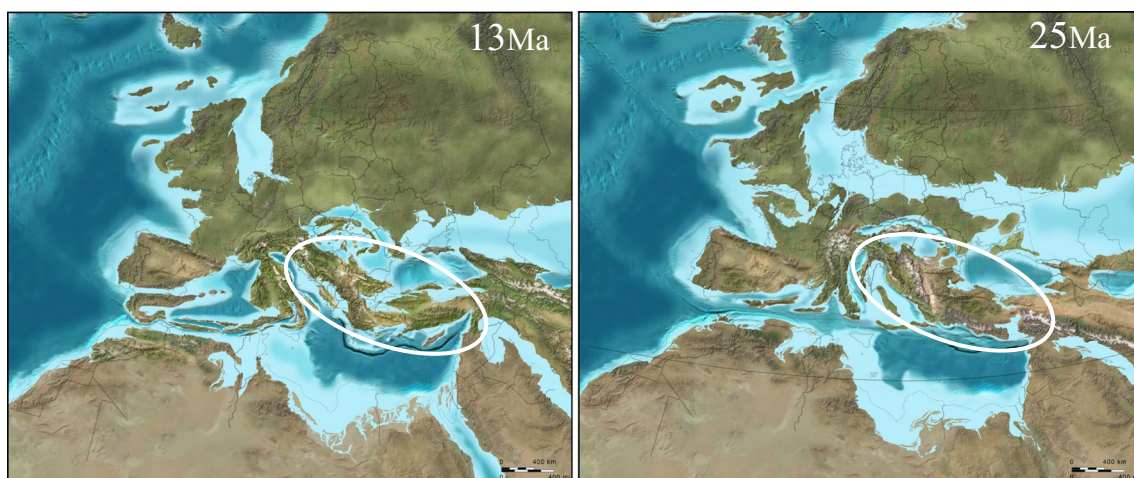
Fig. 4 The Balkan-Pontides microplate region of the most probable Black Pine ancestor distribution (*ellipses*) from the Paleogene and Neogene periods. This microplate today is fragmented and shared between the European and Anatolian tectonic plate. The Mediterranean's basin Paleogeographic reconstruction map from Miocene (*left*) up to the Oligocene epoch (*right*) are according to Gealey (1989) and Stampfli et al. (2002) and re-designed by Dr. Ron Blakey from Colorado Plateau Geosystems, USA. The Western Tethys and Paratethys Seas (i.e. Mediterranean Tethys and Paratethys) are in the southern and northern portion of the ellipse

marked by the distribution limits of other Mediterranean pines, such as *P. pinaster* and *P. brutia* (Critchfield & Little 1966).

The timing of the Black Pine's most recent split and the TMRCA occurred during the Pleistocene-Pliocene-late Miocene epochs during which there were multiple significant temperature fluctuations not typical of a warm period (Joos & Colin 2004). The maximal value of time of the more recent split was the Messinian Salinity Crisis (late Miocene epoch: 4.37 Ma) that occurred in the Mediterranean Basin between the Apennine and Balkan Peninsulas. This period resulted in the formation of a natural migration bridge between the Apennine and Balkan Peninsulas (and the continents and islands). The older split (4.37 Ma) between the Apennine and Balkan Peninsulas explained the more pronounced phylogeographic edge (AB) and indicated the beginning of significant fragmentation of the Black Pine's natural area.

The time of the most recent split and TMRCA between the Balkan Peninsula and Asia Minor do not predate those from the AB edge. The younger split (0.93 Ma) between the Balkan Peninsula and Asia Minor explains the less pronounced phylogeographic edge (BC) and indicates the relatively recent beginning with the significant fragmentation of the Black Pine's natural area. The timing of the younger split was a period of global climatic change and significant ice accumulation on both poles due to increasing climatic cycle duration that was described by Maslin & Ridgwell (2005) as the Early–Middle Pleistocene transitions (0.892–1.2 Ma). This ecological condition was most likely unfavourable for the Black Pines in this region (eastern Mediterranean) and contributed to a significant reduction in the species' distribution, fragmentation, isolation and the split between the Balkan Peninsula and Asia Minor. This hypothesis has been supported by the presence of severe bottlenecks a long time before the first human civilization developed in the region, or agricultural use of the land (Naydenov et al. 2011, 2015). We hope to obtain more information from future complementary mtDNA analyses.

The coalescence analysis did not show any signal from the LGM (BATWING program) in contrast to the results reported by Raffi & Dodd (2007). These authors reported *Black Pine postglacial recolonisation* (i.e. expansion); however, multiple independent paleobotanical studies have reported opposing data, i.e. Black Pine postglacial compression from the same region and time (García-Amorena et al. 2011; Roiron et al. 2013). The *P. nigra* natural populations from the western Mediterranean formation but not the Balkan Peninsula and Asia Minor formations are in habitats identified by fossil record from LGM as interglacial refugia (*cryptic southern refugia*). This hypothesis was supported by modern paleobotanical studies from the western area of the natural distribution of *P. nigra* which showed this species was one of the dominant forest trees in the LGM period in the Iberian Peninsula and

southern France (García-Amorena et al. 2011; Roiron et al. 2013). Only 4–7 °C was required to transform the western European Black Pine natural area from a dominant to an endangered species over a very short period of time (less than 10 thousand years).

The among-population structure of the Black Pine on the Balkan Peninsula and Asia Minor observed in the present study was generated by principal topographic structures, such as mountains and watersheds, and the repeated in situ survival of the populations during cool periods with little altitude migration as suggested by Naydenov et al. (2006, 2015). We did not find any fossil records for significant *P. nigra* area size fluctuations for either region during the LGM period (and before/after). Thus, the Balkan Peninsula and Asia Minor may be *areas of endemism* or macro long-term refugia rather than *glacial or post-glacial refugia* for the Black Pine natural populations according to the theory of Stewart (2003) and Stewart & Dalen (2008). Our phylogeographic and coalescence-based results support the hypothesis of individual spatial and temporal responses of different Black Pine formations (i.e. interglacial refugia vs. macro long-term refugia). Similar longitudinal species adaptation for contrasting climatic and soil conditions was observed for the sea buckthorn (*Hippophae rhamnoides*), which naturally persisted in the Asiatic step and the Atlantic seaboard (Iversen 1973). The phylogeographic structure of *P. nigra* revealed by cpDNA is the result of a long evolutionary history shaped by very longterm (15–16 Ma) climatic/ecological changes and tectonic historic structure changes. The obtained results show that chloroplast DNA is a very useful tool for phylogeographic studies with long historical imprints to present population structures.

Perspectives

Generally, isolated populations are believed to be capable of rapid evolutionary changes due to the high chance of random genetic drift, which generates variations on which selection can rapidly act (Lehman 1998). Mediterranean vegetation is considered to be very resilient because it has experienced intense and abrupt climatic changes. As a result, these taxa are very stable over time and consequently they have retained their ancestral characteristics. Furthermore, the vegetation tends to not fill its climatic range (Svenning & Skov 2004; Petit et al. 2005). The Black Pine fulfils the principal characteristics of Mediterranean taxa. Consequentially, it is necessary to mount a collective effort (national and international) to protect biodiversity in the region, and to manage the ecosystem with a better understanding of its genetic, morphological and physiological characteristics.

Much effort has gone into different meta-genomic forest tree projects (i.e. ProCoGen, EvolTree, TreeBreDex and NovelTree) over the last 10–15 years in Europe (EU) and

has achieved very good results. New genomics/epigenetics projects and collective effort are necessary to elucidate the sub-species structure of the Black Pine for future selection, management and conservation practices. The present results concerning the phylogeographic structure are complimentary and would have been very expensive and difficult to obtain using other techniques. In the near future, we will use mitochondrial DNA to complement the results reported here. This approach will help us develop a better understanding of the natural population structure and history of the Black Pine and confirm the importance of using large sampling areas for molecular analyses of periods of climatic change and anthropogenic pressure.

References

Adams RI, Brown KM, Hamilton MB (2004) The impact of microsatellite electromorph size homoplasy on multilocus population structure estimates in a tropical tree (*Corythophora alta*) and an anadromous fish (*Morone saxatilis*). Mol Ecol 13(9):2579–2588

Balaresque P, Bowden GR, Adams SM, et al. (2010) A predominantly Neolithic origin for European paternal lineages. PLoS Biol 8(1): e1000285. doi:10.1371/journal.pbio.1000285

Belkhir K (2000) Genetix ver 4.01: a software for population genetics data analysis. Loboratoire genome et populations. Universite de Montpellier, II, France

Bojović S (1995) Biodiversité du pin noir (*Pinus nigra* Arn.) en région méditerranéenne. Thèse doctorale. Université d'Aix-Marseille III, France, pp. 1–114

Bonavita S, Vendramin GG, Bernardini V, et al. (2015) The first SSR-based assessment of genetic variation and structure among *Pinus laricio* Poiret populations within their native area. Plant Biosyst Int J Dealing Asp Plant Biol. doi:10.1080/11263504.2015.1027316

Bucci G, Gonzalez-Martinez SC, Le Provost G, et al. (2007) Range-wide phylogeography and gene zones in *Pinus pinaster* Ait. revealed by chloroplast microsatellite markers. Mol Ecol 16(10):2137–2153. doi:10.1111/j.1365-294X.2007.03275.x

Chen C, Durand E, Forbes F, et al. (2007) Bayesian clustering algorithms ascertaining spatial population structure: a new computer program and a comparison study. Mol Ecol Notes 7:747–756

Cinget B, Gérardi S, Beaulieu J, et al. (2015) Less pollen-mediated gene flow for more signatures of glacial lineages: congruent evidence from Balsam Fir cpDNA and mtDNA for multiple refugia in eastern and Central North America. PLoS One 10(4):e0122815. doi:10.1371/journal.pone.0122815

Coart E, Van Labeke S, De Loose M, et al. (2006) Chloroplast diversity in the genus *Malus*: new insights into the relationship between the European wild apple (*Malus sylvestris* (L.) Mill.) and the domesticated apple (*Malus domestica* Borkh.). Mol Ecol 15:2171–2182. doi:10.1111/j.1365-294X.2006.02924.x

Critchfield WB, Little EL (1966) Geographic distribution of the pines of the world. U.S.D.A. Forest Service Miscellaneous Publication 991

Delplancke M, Alvarez N, Espíndola A, et al. (2012) Gene flow among wild and domesticated almond species: insights from chloroplast and nuclear markers. Evol Appl 5:317–329. doi:10.1111/j.1752-4571.2011.00223.x

Dobrinov I, Doykov G, Gagov V (1982) Forest genetic pool in Bulgaria. Zemizdat, Sofia [in Bulgarian]–

Doyle JJ, Morgante M, Tingey SV, et al. (1998) Size homoplasy in chloroplast microsatellites of wild perennial relatives of soybean (*Glycine subgenus* glycine). Mol Biol Evol 15(2):215–218

Duminil J, Heuertz M, Doucet J-L, et al. (2010) CpDNA-based species identification and phylogeography: application to African tropical tree species. Mol Ecol 19:5469–5483. doi:10.1111/j.1365-294X.2010.04917.x

Dzialuk A, Muchewicz E, Boratynski A, et al. (2009) Genetic variation of *Pinus uncinata* (*Pinaceae*) in the Pyrenees determined with cpSSR markers. Plant Syst Evol 277:197–205. doi:10.1007/s00606-008-0123-y

Eckenwalder JE (2009) Conifers of the world—the complete reference. Timber Press, Oregon, 720p

Emerson BC, Hewitt GM (2005) Phylogeography. Curr Biol 15(10): R367–R371

Estoup A, Jarne P, Cornuet JM (2002) Homoplasy and mutation model at microsatellite loci and their consequences for population genetics analysis. Mol Ecol 11:1591–1604

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

Excoffier L, Schneider S, Roessli D (2002) Arlequin ver 2.001: a software for population genetics data analysis. Department of Anthropology and Ecology, University of Geneva, Switzerland

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

Falush D, Wirth T, Linz B, et al. (2003b) Traces of human migrations in *Helicobacter pylori* populations. Science 299:1582–1585

Farjon A (2008) A natural history of conifer. Timber Press, Oregon, 304p

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

François O, Ancelet S, Guillot G (2006) Bayesian clustering using hidden Markov random fields in spatial population genetics. Genetics 174: 805–816

Fukarek P (1958a) Prilog poznavanju crnog bora (*Pinus nigra* ARN. S. lat.) Beitrag zur Kenntnis der systematischen Stellung, Gliederung und der rezenten Verbreitung der Schwarzkiefer. Rad Poljopr Sumersk Fak Univ Sarajevu 3:3–92 [in German]

Fukarek P (1958b) Standortrassen der Schwarzfhre (*Pinus nigra* ARN. S. lat.). Centralbl Gesamten Forstwes 75:203–207 [in German]

García-Amorena I, Rubiales JM, Amat EM, et al. (2011) New macrofossil evidence of *Pinus nigra* Arnold on the northern Iberian Meseta during the Holocene. Rev Palaeobot Palynol 63(3–4):281–288. doi:10.1016/j.revpalbo.2010.10.010

Gealey WK (1989) Plate tectonic evolution of the Mediterranean Middle East region. In: Scotese CR, Sager WW (eds) Mesozoic and Cenozoic plate reconstructions. Elsevier, Amsterdam, pp. 285–306

Godbout J, Fazekas A, Newton CH, et al. (2008) Glacial vicariance in the Pacific northwest: evidence from a lodgepole pine mitochondrial DNA minisatellite for multiple genetically distinct and widely separated refugia. Mol Ecol 17(10):2449–2462. doi:10.1111/j.1365-294X.2008.03761.x

Gomez A, Aguiriano E, Alía R, et al. (2002) Análisis de los recursos genéticos de *Pinus pinea* L. en España mediante microsatélites del cloroplasto. Invest Agr Sist Recur For 11(1):145–154 [in Spanisch]

Gomez A, Vendramin GG, González-Martínez SC, et al. (2005) Genetic diversity and differentiation of two Mediterranean pines (*Pinus halepensis* Mill. and *Pinus pinaster* Ait.) along a latitudinal cline using chloroplast microsatellite markers. Divers Distrib 11:257–263

Grant V (1980) Gene flow and the homogeneity of species populations. Biol Zbl 99:157–169

Grivet D, Petit RJ (2003) Chloroplast DNA phylogeography of the hornbeam in Europe: evidence for a bottleneck at the outset of postglacial colonization. Conserv Genet 4:47–56

Hale ML, Borland AM, Gustafsson MHG, et al. (2004) Causes of size homoplasy among chloroplast microsatellites in closely related *Clusia* species. J Mol Evol 58:182–190

Hardy OJ, Vekemans X (2002) spagedi: a versatile computer program to analyse spatial genetic structure at the individual or population levels. Mol Ecol Notes 2:618–620. doi:10.1046/j.1471-8278.2002.00305.x

He T, Pausas JG, Belcher CM, et al. (2012) Fire-adapted traits of *Pinus* arose in the fiery Cretaceous. New Phytol 194:751–759

Heuertz M, Carnevale S, Fineschi S, et al. (2006) Chloroplast DNA phylogeography of three European ashes, *Fraxinus* sp. (Oleaceae): roles of hybridization and contrasting mating systems. Mol Ecol 15: 2131–2140

Heuertz M, Teufel J, González-Martínez SC, et al. (2010) Geography determines genetic relationships between species of mountain pine (*Pinus mugo* complex) in Western Europe. J Biogeogr 37(3):541–556. doi:10.1111/j.1365-2699.2009.02223.x

Hodel RG, Gonzales E (2013) Phylogeography of sea oats (*Uniola paniculata*), a dune-building coastal grass in Southeastern North America. J Hered, June, doi:10.1093/jhered/est035

Hohn M, Gugerli F, Abran P, et al. (2009) Variation in the chloroplast DNA of Swiss stone pine (*Pinus cembra* L.) reflects contrasting post-glacial history of populations from the Carpathians and the Alps. J Biogeogr. doi:10.1111/j.1365-2699.2009.02122.x

Hudson RR (1990) Gene genealogies and the coalescent process. Oxf Surv Evol Biol 7:1–44

Iversen J (1973) The development of Denmark's nature since the last glacial. Dan Geol Undersogelse 7C:1–126

Jaramillo-Correa JP, Aguirre-Planter E, Khasa DP, et al. (2008) Ancestry and divergence of subtropical montane forest isolates: molecular biogeography of the genus *Abies* (Pinaceae) in southern Mexico and Guatemala. Mol Ecol 17(10):2476–2490. doi:10.1111/j.1365-294X.2008.03762.x

Joos F, Colin PI (2004) A paleo-perspective on changes in atmospheric CO$_2$ and climate. The global carbon cycle: integrating humans, climate, and the natural world. Scope 62:165–186

Kalinowski ST (2005) HP-Rare: a computer program for performing rarefaction on measures of allelic diversity. Mol Ecol Notes 5: 187–189

Kalinowski ST (2009) How well do evolutionary trees describe genetic relationships between populations? Heredity 102:506–513

Kayser M, Brauer S, Cordaux R, et al. (2006) Melanesian and Asian origins of Polynesians: mtDNA and Y chromosome gradients across the Pacific. Mol Biol Evol 23(11):2234–2244

Kingman JFC (1982) The coalescent. Stoch Process Appl 13(3):235–248. doi:10.1016/0304-4149(82)90011-4

Lande R, Engen S, Saether BE (1999) Spatial scale of population synchrony: environmental correlation vs. dispersal and density regulation. Am Nat 154:271–281

Lande R, Engen S, Saether BE (2003) Stochastic population dynamics in ecology and conservation. Oxford University Press, Oxford

Lehman N (1998) Conservation biology: genes are not enough. Curr Biol 8:R722–R724

Liu K, Muse SV (2005) PowerMarker: an integrated analysis environment for genetic marker analysis. Bioinformatics 21:2128–2129

Lumaret R, López de Heredia U, Soto A (2009) Origin and genetic variability of cork-oak. In: Aronson J, Pereira JS, Pausas J (eds) Cork oak woodlands on the edge: ecology, adaptive management and restoration. Island Press, Washington, DC, pp. 25–32

Lumaret R, Tryphon-Dionnet M, Michaud H, et al. (2005) Phylogeographical variation of chloroplast DNA in cork oak (*Quercus suber*). Ann Bot 96:853–861. doi:10.1093/aob/mci237

Manni F, Guerard E, Heyer E (2004) Geographic patterns of (genetic, morphologic, linguistic) variation: how barriers can be detected by using Monmonier's algorithm. Hum Biol 76:173–190

Maslin MA, Ridgwell AJ (2005) Mid-Pleistocene revolution and the 'eccentricity myth'. In: Head, M.J., Gibbard, P.L. (Ed), "Early–Middle Pleistocene transitions: the land–ocean evidence", Geol Soc Lond, Spec Publ, 247: 19–34.

Miller P (1768) The gardeners dictionary. Ed. 8. London

Mirov NT (1967) The genus *Pinus*. The Ronald Press, New York, 610 p

Navascués M, Emerson BC (2005) Chloroplast microsatellites: measures of genetic diversity and the effect of homoplasy. Mol Ecol 14:1333–1341. doi:10.1111/j.1365-294X.2005.02504.x

Naydenov KD, Mladenov I, Alexandrov A, et al. (2015) Patterns of genetic diversity resulting from bottlenecks in European Black Pine, with implications on local genetic conservation and management practices in Bulgaria. Eur J For Res 134(4):669–681. doi:10.1007/s10342-015-0881-3

Naydenov KD, Naydenov MK, Tremblay F, et al. (2011) Patterns of genetic diversity that result from bottlenecks in Scots Pine and the implications for local genetic conservation and management practices in Bulgaria. New For 42(2):179

Naydenov KD, Senneville S, Beaulieu J, et al. (2007) Glacial vicariance in Eurasia: mitochondrial DNA evidence from Scots Pine for a complex heritage involving genetically distinct refugia at mid-northern latitudes and in Asia Minor. BMC Evol Biol 7:233. doi:10.1186/1471-2148-7-233

Naydenov KD, Tremblay F, Fenton N (2005) Chloroplast microsatellites analysis revealed a high level of differentiation in Jack pine (*Pinus banksiana* Lamb.) population in Quebec. Belg J Bot 138(2):181–191

Naydenov KD, Tremblay F, 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

Naydenov KD, Tremblay F, Ganchev P (2003) Karyotype diversity in of European Black Pine (*Pinus nigra* Arn.) from Bulgarian provenances. Phyton 43(1):9–28

Nei M, Chesser RK (1983) Estimation of fixation indices and gene diversities. Ann Hum Genet 47:253–259

Palamarev E (1989) Paleobotanical evidences of the Tertiary history and origin of the Mediterranean sclerophyll dendroflora. Woody plants—evolution and distribution since the Tertiary. Plant Syst Evol 162(1/4):93–107

Perdereau AC, Kelleher CT, Douglas GC, Hodkinson TR (2014) High levels of gene flow and genetic diversity in Irish populations of *Salix caprea* L. inferred from chloroplast and nuclear SSR markers. BMC Plant Biol 14:202. doi:10.1186/s12870-014-0202-x

Petit RJ, Aguinagalde I, de Beaulieu J-L, et al. (2003) Glacial refugia: hotspots but not melting pots of genetic diversity. Science 300(5625):1563. doi:10.1126/science.1083264

Petit RJ, Brewer S, Bordaes S, et al. (2002a) Identification of refugia and post-glacial colonisation routes of European white oaks based on chloroplast DNA and fossil pollen evidence. For Ecol Manag 156: 49–74

Petit RJ, Csaikl UM, Bordaes S, et al. (2002b) Chloroplast DNA variation in European white oaks: phylogeography and

patterns of diversity based on data from over 2,600 populations. For Ecol Manag 156:5–26

Petit RJ, Excoffier L (2009) Gene flow and species delimitation. Trends Ecol Evol 24(7):386–393

Petit RJ, Hampe A, Cheddadi R (2005) Climate changes and tree phylogeography in the Mediterranean. Taxon 54(4):877–885

Pianka ER (1978) Evolutionary ecology. Harper and Row, New York

Pico FX, Mendez-Vigo B, Martinez-Zapater JM, et al. (2008) Natural genetic variation of *Arabidopsis thaliana* is geographically structured in the Iberian Peninsula. Genetics 180(2):1009–1021. doi:10.1534/genetics.108.089581

Pons O, Petit RJ (1996) Measuring and testing genetic differentiation with ordered versus unordered alleles. Genetics 144:1237–1245

Powell W, Morgante M, Andre C, et al. (1995b) Hypervariable microsatellites provide a general source of polymorphic DNA markers for the chloroplast genome. Curr Biol 5:1023–1029

Powell W, Morgante M, McDevitt R, et al. (1995a) Polymorphic simple sequence repeat regions in the chloroplast genome: applications to the population genetics of pines. Proc Natl Acad Sci U S A 92:7759–7763

Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. Genetics 155(2):945–959

Pritchard JK, Wen W (2003) Documentation for structure software Version 2 [online]. Available from http://pritch.bsd.uchicago.edu/software/readme_2_1/readme.html

Provan J, Bennett KD (2008) Phylogeographic insights into cryptic glacial refugia. Trends Ecol Evol 23(10):564–571. doi:10.1016/j.tree.2008.06.010

Provan J, Soranzo N, Wilson NJ, et al. (1999) A low mutation rate for chloroplast microsatellites. Genetics 153:943–947

Raffi ZA, 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. Mol Ecol 16(4):723–736

Raufaste N, Bonhomme F (2000) Properties of bias and variance of two multiallelic estimators of Fst. Theor Popul Biol 57:285–296

Robertson A, Hill WG (1984) Deviations from Hardy-Weinberg proportions: sampling variances and use in estimation of inbreeding coefficients. Genetics 107:703–718

Rogers JS (1972) Measures of similarity and genetic distance. In studies in genetics VII. University of Texas Publication 7213, Austin, Texas, p 145−153

Roiron P, Chabal L, Sueiral I, et al. (2013) Palaeobiogeography of *Pinus nigra* Arn. subsp. *salzmannii* (Dunal) Franco in the north-western Mediterranean Basin: a review based on macroremains. Rev Palaeobot Palynol 194:1–11. doi:10.1016/j.revpalbo.2013.03.002

Scotti I, Paglia G, Magni F, et al. (2006) Population genetics of Norway spruce (*Picea abies* Karst.) at regional scale: sensitivity of different microsatellite motif classes in detecting differentiation. Ann For Sci 63:485–491. doi:10.1051/forest:2006029

Semerikov VL, Semerikova SA, Dymshakova OS, et al. (2014) Microsatellite loci polymorphism of chloroplast DNA of Scots Pine (*Pinus sylvestris* L.) in Asia and Eastern Europe. Russ J Genet 50(6):577–585. doi:10.1134/S1022795414040127

Soto A, Robledo-Arnuncio JJ, Gonzalez-Martinez 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. doi:10.1111/j.1365-294X.2010.04571.x

Spellman GM, Klicka J (2006) Testing hypotheses of Pleistocene population history using coalescent simulations: phylogeography of the pygmy nuthatch (*Sitta pygmaea*). Proc R Soc B 273:3057–3063

Stampfli G, Borel G, Cavazza W, et al (2002) The paleotectonic atlas of the periTethyan domain: European Geophysical Society, CD-ROM

Stefanov B (1941/1942) Geographical distribution of coniferous species and their form in nature. Godichnik na Sofiiskia Darjaven Universitet, Sofia, (Bulgaria), XIX and XX, 1–88 [in Bulgarian]

Stefanov B (1943) The phyto-geographical elements of Bulgaria. Thesis of Bulgarian Academy of Sciences, Faculty of Nature and Mathematics, Sofia, (Bulgaria), vol. XXXIX, 19:1–121 [in Bulgarian]

Stewart JR (2003) Comment on 'Buffered tree population changes in a Quaternary refugium: evolutionary implications'. Science 299:825a. doi:10.1126/science.1079388

Stewart JR, Dalen L (2008) Is the glacial refugium concept relevant for northern species? A comment on Pruett and Winker 2005. Clim Chang 86:1–2. doi:10.1007/s10584-007-9366-9

Stewart JR, Lister AM, Barnes I et al (2010) Refugia revisited: individualistic responses of species in space and time. Proc R Soc B (277):661–671; DOI:10.1098/rspb.2009.1272.

Svenning JC, Skov F (2004) Limited filling of the potential range in European tree species. Ecol Lett 1:565–573

Taberlet P, Fumagalli L, Wust-Saucy AG, et al. (1998) Comparative phylogeography and postglacial colonization routes in Europe. Mol Ecol 7:453–464

Terrab A, Paun O, Talavera S, et al. (2006) Genetic diversity and population structure in natural populations of Moroccan Atlas cedar (*Cedrus atlantica*; *Pinaceae*) determined with cpSSR markers. Am J Bot 93:1274–1280. doi:10.3732/ajb.93.9.1274

Vendramin G, Lelli L, Rossi P, et al. (1996) A set of primer for the amplification of 20 chloroplast microsatellites in Pinaceae. Mol Ecol 5:585–598

Vidakovic M (1991) Conifers: morphology and variation. Graficki Zavod Hrvatske, Croatia

Vinceti B, Loo J, Gaisberger H, et al. (2013) Conservation priorities for *Prunus africana* defined with the aid of spatial analysis of genetic data and climatic variables. PLoS One 8(3):e59987. doi:10.1371/journal.pone.0059987

Walter R, Epperson BK (2005) Geographic pattern of genetic diversity in *Pinus resinosa*: contact zone between descendants of glacial refugia. Am J Bot 92(1):92–100

Wang B, Mahani MK, Ng WL, et al. (2014) Extremely low nucleotide polymorphism in *Pinus krempfii* Lecomte, a unique flat needle pine endemic to Vietnam. Ecol Evol 4(11):2228–2238. doi:10.1002/ece3.1091

Wang Y, Jiang H, Peng S, et al. (2011) Genetic structure in fragmented populations of *Hippophae rhamnoides ssp. sinensis* in China investigated by ISSR and cpSSR markers. Plant Syst Evol 295(1/4):97–107 http://www.jstor.org/stable/43558222

Weir BS, Cockerham CC (1984) Estimating F-statistics for the analysis of population structure. Evolution 38:1358–1370

Wilson IJ, Weale ME, Balding DJ (2003) Inferences from DNA data: population histories, evolutionary processes and forensic match probabilities. J R Stat Soc A 166:155–201

Wright S (1979) Evolution and the genetics of populations, vol. 4. Variability within and among natural populations. Chicago, 580 pp