REVIEW ARTICLE

Speciation and historical migration pattern interaction: examples from *P. nigra* and *P. sylvestris* phylogeography

Krassimir D. Naydenov[1] · Michel K. Naydenov[2] · Alexander Alexandrov[3] · Todor Gurov[4] · Veselka Gyuleva[3] ·
Georgi Hinkov[3] · Sofiya Ivanovska[4] · Anatoly Tsarev[5] · Biljana Nikolic[6] · Venceslas Goudiaby[7] ·
Christopher Carcaillet[8] · Roman Volosyanchuk[9] · Srdjan Bojovic[10] · Kole Vasilevski[11] · Vlado Matevski[12] ·
Lorenzo Peruzzi[13] · Andreas Christou[14] · Despina Paitaridou[15] · Irina Goia[16] · Salim Kamary[17] · Suleyman Gulcu[18] ·
Cengiz Ture[19] · Faruk Bogunic[20]

Abstract
Here, from macrophylogeographic mtDNA empirical data, we propose a scenario for the evolution and speciation of two important forest trees, European black pine and Scotch pine, and their multiple subspecies and varieties. Molecular clock simulations revealed that *INDEL* variability in the *Pinus* mitochondrial genome is relatively old, i.e., from the Pliocene-Miocene epoch, and related to historical tectonic continental fluctuations rather than to climate change at a large geographic scale. For conservation and management biodiversity program recommendations, special attention is given to the relationships between different speciation models, historical migration patterns, and differences between peripheral and central populations. Species evolution involves the mixing of different speciation modes, and every speciation mode has different effects on different DNA types (e.g., mitochondrial vs. chloroplast vs. nuclear DNA). The misbalance between the contributions of different meta-population census sizes vs. effective population sizes to asymmetric migration patterns is the result of different genotypes (and subphylogenetic lines) responding to selection pressure and adaptive evolution. We propose initial minimal size of conservation unit (between 3 and 5 ha) from central and marginal natural area of distribution for both species in the dynamic management system for practical forest genetic diversity management. The proposed physical sizes were determined by the effective population size, effective radius of seed distribution data, forest density age dynamics, succession pattern, natural selection pressing and species biology [R-17].

Keywords Mitochondria · Phytogeography · Black pine · Scotch pine · Speciation and migration · Conservation and management

Introduction

Climatic change and anthropogenic pressures are important challenges facings natural ecosystems and their compositions [R-22]. Forest trees are an especially important component of most terrestrial vegetation and have an important impact on related flora and fauna. Most forest trees are commercially and economically valuable worldwide (Vidakovic 1974, 1991; Dobrinov 1983; Alexandrov et al. 1988). In the last century, protection and sustainable management have been crucial in preventing the loss of biodiversity and ensuring conservation program success [R-23].

Historical migration patterns and phylogeographic structure studies are important for better understanding species evolution. Migration is the principal factor opposing speciation and strong species phylogeographic structure. Migration sometimes accelerates the speciation process: reinforcement of premating isolation, introgression of an adaptive allele that boosts a further local adaptation [R-1]. The phylogenetic line of subspecies and varieties could be reinforced by low gene flow and vice versa [R-24]. A persistent significant migration rate could change a population's structure and genotypic diversity and could eliminate the isolation

---

Communicated by Oliver Gailing.

In Memory of Irena Naydenova.

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

Extended author information available on the last page of the article

barrier. Different scenarios of natural selection and genetic reinforcement interactions with partial or full back and forth natural population movement under specific ecological conditions underpin the majority of speciation models, which are allopatric (vicariance), peripatric, parapatric and sympatric models. The majority of speciation models "from the aspect of geographical classification of speciation" [R-2] using generally contemporary phenotype traits, paleo records from plants and animals or pure synthetic deductions, more theoretical speciation models have been developed by Mayr (1954), Anacker and Strauss (2013) for a geographic speciation; Brown (1957), Greenbaum et al. (1978), Briggs (2000), Bowen (2010) for a centrifugal speciation; Grant 1971 for a synthetic theory of plant speciation; Futuyma (1989) for an adaptive speciation; Gavrilets et al. (2000) for a parapatric speciation; Gottlieb (2003), Crawford (2010) for a derivative speciation; and Howard (2003) for an allopatric speciation [R-4].

Briefly, it is important, especially for plants, to note that first, migration dynamics are pronounced in time less for plants than for animals, and second, there is a closer relationship between physiological plant tolerance and the environment. This repeated generation after generation of breeding, mutating and carrying out natural selection (by individual physiological factors and sexual/mate selection) shapes population genetic structure, expands population areas into new habitats and guarantees species successional reproduction over time. We must bear in mind that plant persistence, evolution, speciation and migration are possible through successive genotypic changes that meet the historic pattern of the ecological conditions encountered. Long-term, significant evolutionary change is through species physiological changes, and species migration fluctuations and phylogeographic structure are often related to the presence of multiple phylogenetic lines and active speciation [R-5 and R-6].

The phylogeography of principal forest tree species has been relatively well studied by chloroplast and mitochondrial molecular markers over the last 25 years (Burban and Petit, 2003; Naydenov et al. 2007). In contrast, the role of historical migration patterns in determining phylogeographic structures remains poorly understood at present (Naydenov et al. 2014, 2016, 2017). The interaction between historical migration patterns and phylogeographic structure could allow us to hypothesize which classical evolution model is more appropriate for explaining long-term species persistence under different ecological conditions and how protected areas (surface, fragmentation and topography) and forest composition are preferable targets for conservation projects (Pant 1999).

Here, our principal concept is that a conservation and biodiversity strategy must guarantee that principal species such as forest trees have the opportunity to continue their evolution rate and speciation process, similar to the scenario present in the period before humans were active. This concept has never been used prior to this study and is a question of how modern conservation programs can ensure the sustainable existence of species. Maintaining species evolution rates and speciation processes is one of the more important factors of species protection and management programs. Similarly, only the effect of the long-term historical migration pattern on the present phylogeographic structure of a species throughout its natural distribution area could give us indirect indications of historical speciation model dynamics.

Historical migration pattern analysis depends on the molecular marker used and its phylogenetic memory. A low mutation rate reveals very old phylogeographic events and vice versa. For example, for the genus *Pinus*, mitochondria Indel DNA reaches middle *Miocene* epoch (from present date here); chloroplast DNA reaches late *Pliocene* epoch, and nuclear DNA with phylogenetic memory is relatively short, reaches late *Pleistocene* epoch years BP (Fig. 5; Naydenov et al. 2014, 2017). Other important factors for forest trees are differences between seed versus pollen distribution, chloroplast versus mitochondria distribution and paternal versus maternal status radiation in space. For this reason, we aimed to complete previously published chloroplast *SSR* genetic data and European black pine data with a mitochondrial analysis of the whole geographic area of this species. Thus, we compared historical migration fluctuations between meta-populations and determined where migration was the result of species area expansion under favorable global environmental changes. We speculated that symmetric migration waves would be the result of species natural range expansions, and nonsymmetric migration waves would be related to local speciation. The molecular marker used here is perfect for determining the most appropriate speciation model.

We considered the comparative phylogeography and comparative historical migration pattern between important forest trees from relatively close natural distribution areas as another important tool for determining modern conservation practices and validating species evolution hypotheses. A comparison of historical migration patterns between relatively close species has never been conducted before in this manner. We used mitochondrial data from the same molecular marker at the meta-population level for European black pine and Scotch pine to perform comparative phylogeographic and historical migration pattern analyses. The biological subspecies and phenotypic diversity variation in both species have been well described in the last century (Vidakovic 1974, 1991; Dobrinov 1983; Alexandrov et al. 1988). Theoretically, speciation and population migration interactions have been well discussed as two opposite factors of evolution, but direct empirical molecular genetic data have never been related to possible speciation models of

forest trees (i.e., *P. nigra* and *P. sylvestris*), and their iteration with historical migration patterns and phylogeographic structure remains poorly understood. We will compare migration waves based on their historical time of occurrence and their contribution. The difference between census size and effective population size has been hypothesized to be the result of historical genetic population structure differences and different selective pressures under specific environmental conditions. We compared data for *P. nigra* with those for *P. sylvestris*, with the aim of establishing species biological differences in relation to migration patterns and speciation model differences, and determined whether different conservation and management recommendations are needed.

In particular, we focused on contemporary phylogeography and population genetic structure and variability in the full natural distribution areas of these species to facilitate future conservation and management practices. For the reforestation programs in forests (i.e., introduction and reintroduction), it is important that we recognize that a new environment does not play a part in reproductive material preadaptation, and species must be able to persist in a given environment throughout the normal fluctuation in ecological conditions.

Materials and methods

This publication expands upon a previously published analysis of chloroplast microsatellites in European black pine (*Pinus nigra* Arn.) by the same author. The population name, geographic position and sample size of the analyzed plant material were the same as those in Naydenov et al. (Table 1: 2016, 2017). Briefly, we analyzed 106 natural populations of European black pine across its whole natural distribution. The more western populations were from Spain/Morocco, and the more eastern populations were from Crimea and Asia Minor (Turkey). The longitude distribution was 4.7°W to 39.1°E, and the latitude interval was 5.1°N to 48.1°N. In addition to recently collected material (needles and seeds), other reference materials from the herbarium collection of the Forest Research Institute—Bulgarian Academy of Science (BAS) in Sofia collected over the last 80 years from Austria, Crimea, France, Romania and the former Yugoslavia by Prof. D. Velkov and Acad. B. Stefanov were analyzed.

The plant material included 29.7 trees per population on average, and DNA extraction and storage, initial laboratory manipulations and fragment analysis via PCR amplification were the same as those in Naydenov et al. (2007). Population dynamics, phylogeographic structures and migration patterns were analyzed by fragment analysis of the mitochondrial region of *intron 1* from the *NAD7* gene (i.e., nicotinamide adenine dinucleotide: NADH dehydrogenase). The molecular markers and primer pairs were selected given

technical instrument restrictions, the molecular structure of the mitochondria for the *Pinus* genus and sufficient mutations with significant information for analyzing species phylogeography. The laboratory protocol was the same as that in Naydenov et al. (2007). The previously published GenBank sequences from current and former members of our team helped us identify a region with little risk of homoplasy that yielded sufficient capillary electrophoresis data under the limit of a maximal error of one base pair (i.e., 1 bp). The best fragment size range for the capillary electrophoresis technique used was 120–800 bp. The GenBank accession numbers are HQ185286.1, HQ185287.1, HQ185288.1, HQ185289.1, KT343352.1 and KT343353.1. There are currently no other reported sequences from the same DNA region for *P. nigra*.

The inter- and intrapopulation genetic structures were analyzed using the variation in the size of DNA variants as determined by multiple indices, such as the level of haplotype diversity, fixation index F_{st}, $F_{st\text{-coalescent}}$, G_{st}, R_{st}, Φ_{st} and N_{st}, and multiple model-based statistical methods. The size variant profile was analyzed as *INDEL*s of noncombined DNA (i.e., organellar origin). Standard population genetic parameters, such as mitotype diversity (H_{mt}), differentiation among populations and meta-populations in each group as indicated by the fixation index F_{st} (Wright 1943, 1951, 1965), $F_{st\text{-coalescent}}$ (Slatkin 1991), G_{st} (Nei 1973), R_{st} (Slatkin 1995), Φ_{st} (Excoffier et al. 1992) and N_{st} (Lynch and Crease, 1990), were analyzed, and ANOVA and MANOVA were performed with the ARLEQUIN (Excoffier et al. 2002), FSTAT (Goudent 2001), GENALEX (Peakall and Smouse 2006), GENETIX (Belkhir et al. 1996–2004), POPGENE (Yeh et al. 1997), PowerMARKER (Liu and Muse 2005), SAMOVA (Dupanloup et al. 2002) and SPAGeDi (Hardy and Vekemans, 2002) programs according to the author's instructions and were found to be similar to those previously published by Naydenov et al. (2014, 2016, 2017). Mitotype diversity robustness at the population and meta-population levels was tested with rarefaction analysis by the HP-RARE v.1.0 program (Kalinowski 2005).

Phylogeography structure, i.e., a geographically structured haplotype signal, was determined by multiple model-based statistical methods. Population structure was analyzed with distance-based and Bayesian statistical methods. A spatial analysis of genetic molecular variance was completed with the BARRIER program developed by Manni et al. (2004), and population/meta-population admixture was analyzed based on Bayesian statistics with the BAPS (Corander et al. 2004), STRUCTURE (Evanno et al. 2005) and TESS (Chen et al. 2007) programs. All among-population/ among-meta-population analyses were performed with 10 thousand resampling replications by bootstrapping, permutation or jackknife as recommended by Miller (1974), Efron (1979, 1982), Meirmans and Hedrick (2010) and the authors

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

Meta-pop[1]	Fst (SD)	N_{ind}	N_{hap}/N_{ind}	N_{hap}	H_{mt} (n.b)
Partition by principal geographic regions					
Morocco	0.295	50	0.06	3	52.08 (52.61)
Spain	0.304	50	0.12	6	65.44 (66.10)
France	0.285	50	0.14	7	76.56 (77.33)
Corsica	0.305	50	0.10	5	71.28 (72.00)
Italy	0.297	50	0.06	3	55.44 (56.00)
W-Balkan Peninsula	0.305	50	0.10	5	66.72 (67.39)
E-Balkan Peninsula	0.300	50	0.08	4	58.48 (59.07)
W-Greece	0.205	50	0.06	3	28.08 (28.36)
Crimea Peninsula	0.300	50	0.06	3	58.00 (58.59)
N-Turkey	0.185	50	0.04	2	18.00 (18.18)
Turkey	0.286	50	0.10	5	46.56 (47.03)
Mean	0.290 (0.128)	50	0.08	4.2	54.24 (54.79)
Best partition according BAPS spatial analysis, SAMOVA and BARRIER programs					
N-Turkey	0.185	50	0.04	2	18.00 (18.18)
W-Greece	0.205	50	0.06	3	28.08 (28.36)
France & Corsica	0.003	100	0.08	8	81.00 (81.41)
Spain & Morocco	0.088	100	0.07	7	59.98 (60.28)
Principal area (Crimea, Balkan and Apennine Peninsulas, Turkey) PPPeninsulas ()	0.076	250	0.03	7	58.74 (58.86)
Mean	0.111 (0.074)	110	0.06	5.4	49.16 (49.42)
SAMOVA		df	Sum of squares		Percentage of variation
	Among groups	4	108.639		38.02
	Within group	6	4.463		1.05
	Within pop	1089	298.320		60.93
AMOVA- Pairwise	Nst = 0.2027				Distance max = 40.20
	Gst = 0.2310				Distance mean = 17.83
	Rst = 0.1916				Distance mean ln = 2.68
AMOVA-Global	Fst = 0.2780				Rit = 0.2335
	Gst = 0.2780				Ris = 0.3971
	Nst = 0.2337				Rst = 0.0123

N is sample size; for N_{hap}/N_{ind} is the ratio between the number of haplotype of total number of analyzed individual per populations; H_{mt}—mitotype diversity; average haplotype per population is 2.4 (min/max: 1/5); Nuc. div.—nucleotide diversity is used to measure the degree of polymorphism from size variant data; SE – standard error (and standard deviation is SD = SE$\sqrt{n}$) is 0.012

The data are from population genetic model statistic basses on allele's identity, allele size and distance between alleles (GENETIX, PowerMARKER, IDB, GENEpop, ARLEQUIN and SPAGeDi programs)

[1] The signification is done by SD and SE in the note of Table 1

of the abovementioned programs. For the distance-based methods, the obtained list of the 10 thousand matrices in NEWICK or NEXUS format for each distance was imported into TreeVIEW (Page 1996) and FigTREE (Rambaut 2009) programs and combined with the TRACER (Rambaut and Drummond, 2009) program to draw a summary cluster tree with bootstrap support scores for all nodes. The resampling replication cutoff level was fixed at 2/3 (i.e., 66%).

Our initial genetic data were obtained from unknown mutation and mating models, and we decided to test a long list of accessible distance methods from the abovementioned programs. The genetic distances tested were as follows: Euclidean distance (DEU: Deza and Deza, 2009), Rogers's (1972) DR, Prevosti et al.'s (1975) CP, Cavalli-Sforza and Edwards' (1967) DC, Bhattacharyya's (1946) and Nei's (1987) θ^2, Sanghvi's (1953) X^2, Nei et al.'s (1983) DA, Cavalli-Sforza's (1969) chord distance (fv), Reynolds et al. (1983) θW, Nei's (1972) DS, Latter's (1972) $\phi*$, Latter's (1973) *DL*, Goldstein et al.'s (1995) $(\delta\mu)^2$, Slatkin's (1995) ASD, Shriver et al.'s (1995) DSW, Chakraborty and Jin's (1993) DSA and distances from Nei and Roychoudhury (1974), Takezaki and Nei (1996). From each summary distance matrix, we constructed phylogenetic trees using common approaches: (a) the neighbor-joining

(NJ) method according to Saitou and Nei, (1987) and (b) the unweighted pair-group method with arithmetic mean (UPGMA) approach proposed by Sneath and Sokal (1973), Nei and Kumar (2000). Different distance methods are often proposed for different mutation models and different mating models, i.e., population/subpopulation structure or population subdivision models, also known as partition models. More commonly used mutation models include the stepwise mutation model (*SMM*) proposed by Ohta and Kimura (1973), the infinite allele model (IAM) proposed by Kimura and Crow (1964), the infinite sites model (ISM), the K-alleles model (KAM), the nonmutation model including Euclidian distance/geometric distance and their multiple variants. Importantly, the ISM and KAM are derived from the IAM. The population subdivision models using the abovementioned distances include the stepping-stone model (Kimura and Weiss, 1964), the island model (Wright 1931) and the general model. Partition models provide insight into the spatial, ecological or reproductive barriers separating individuals and defining the relative probability of panmixia (i.e., random mating). Additionally, the simple Mantel test (Mantel 1967) and model-based analysis implemented in the IBD (Bohonak 2002) program can be used to identify the geographic pattern of the genetic distances, i.e., the correlation between the summary genetic and simple geographic distance matrices.

We analyzed historical migration patterns, time to the most recent common ancestor (MRCA) and historical effective population size (N_e) with the infinite allele model for haploid data by Bayesian (BY) and maximum likelihood (ML) analyses in the n-MIGRATE v.3.6.6 program as reported those reported in Naydenov et al. (2016). The program operates under the assumption of neutral evolution in Kimura (1983). The analysis was performed for multiple groups of 11 geographic regions as an overlapping problem to better describe effective population size fluctuations in space and time. Each region was represented by 50 individuals, and the haplotype structure reflected the proportions from single populations. We included 150 control samples from 3 well-studied Scotch pine regions with the same mitochondrial gene region (*NAD7-intron1*) from Naydenov et al. (2007). *Pinus sylvestris* data have never been studied in terms of historical effective population size fluctuations and asymmetric migration patterns. Comparing the two species is essential for evaluating speciation models and historical migration pattern interactions.

Complete and large geological time paleobotanical records (from the Pliocene and Miocene epochs) for the European black pine and its ancestors do not exist. Here, the calibration of the molecular clock for *INDEL*s and the *Pinus* genome was conducted due to the availability of more stable date records, such as those for tectonic and climatic changes. We speculated that tectonic dynamics have a significant impact on forest tree phylogeographic structure. The timings of historical events were calibrated according to the hypothesis that large climatic and tectonic changes in the region impacted the species migration history. Mutation speed (μ) is theoretically lower for large DNA fragments than for *SNP*s due to their different molecular masses (i.e., fragment size differences in bp). For this reason and according to multiple previous publications, the molecular clock time mark for *INDEL*s must predate the LGM and Plio-Pleistocene epochs (Naydenov et al. 2014, 2016, 2017). For the time mark, we used the Messinian salinity crisis (MSC) from the Mediterranean Basin. During this geological event that occurred in the later part of the Messinian age, the Mediterranean Sea dramatically shrank or was almost completely desiccated. The most pronounced change was from 5.96 to 5.33 Ma (i.e., the Miocene epoch). The opening of the Gibraltar Strait allowed a significant amount of Atlantic water to penetrate the Mediterranean Basin [R-26]. This rapid end to the MSC that occurred over approximately 100 years is known as the Zanclean flood (Gautier et al. 1994; Krijgsman 1996). We previously used the same geological time mark with success to calibrate historical migration patterns from nuclear *SSR*s for natural *Pinus pinaster* (Aiton) populations in the western Mediterranean Basin and chloroplast *SSR*s for *P. nigra* (Naydenov et al. 2014). For the present sample collection strategy, we used populations from the same region, including Morocco, Spain, Corsica and France. For the names of geological epochs, especially for the Pleistocene and Pliocene, we use terminology provided by Gradstein et al. (2004) and accepted by the International Commission on Stratigraphy according to the 2008 International Stratigraphic Chart (www.stratigraphy. org/ ICSchart/ StratChart2008.pdf).

Results

The detailed results of the molecular laboratory work are presented in Appendix: BOX-1. Briefly, we obtained a wide range of fragment sizes given the limits of the best separation effectiveness of the laboratory equipment. Black pine's fragments were between 224 and 598 bp with different frequencies. The eleven size variants were Mt-1 (224 bp), Mt-2 (236 bp), Mt-3 (239 bp), Mt-4 (248 bp), Mt-5 (253 bp), Mt-6 (255 bp), Mt-7 (281 bp), Mt-8 (284 bp), Mt-9 (534 bp), Mt-10 (564 bp) and Mt-11 (598 bp). Their geographic distribution showed a generally weak geographic structure (Fig. 1). Specific cryptic phylogeographic structure was observed in western Greece (Mt-10/564 bp) and the northern part of Asia Minor (Turkey: Mt-11/598 bp). The endemic mitotypes with a very low frequency (less than 1%) were Mt-2 (Bosnia, Bulgaria and Spain), Mt-3 (Serbia), Mt-4 (France), Mt-7 (France) and Mt-5 (France and Spain). The 248-bp fragment (Mt-4) from the herbarium collection
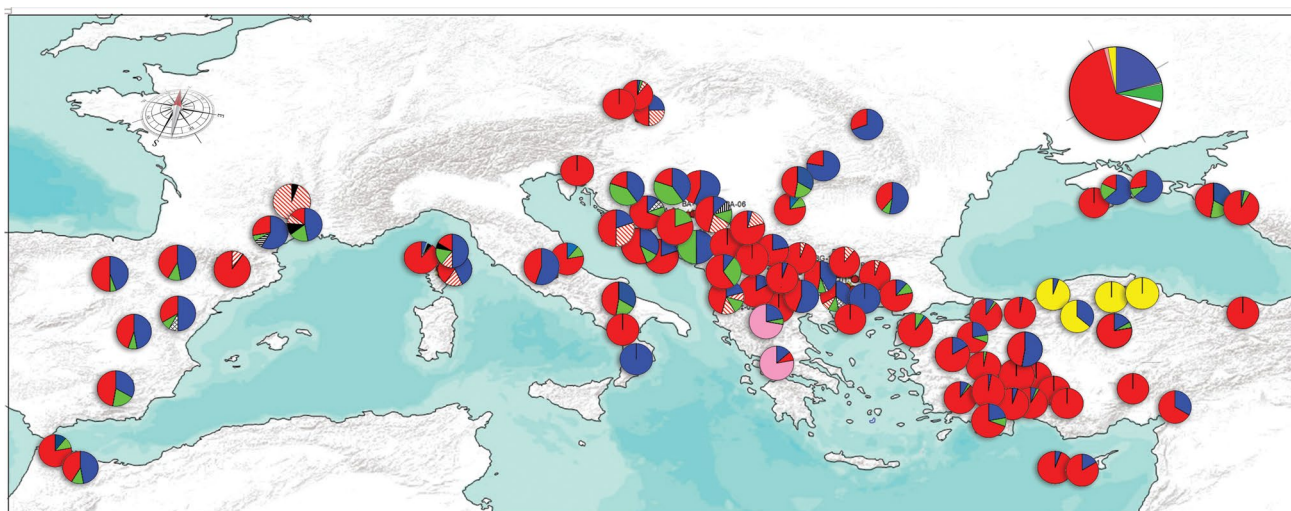
Fig. 1 European Black Pine Mitotype distribution by natural population on the full area of natural distribution. In the top right corner (i.e., the large pie chart) is a summary of the observed relative total mitotype diversity

(French origin) was not found in recently collected material (Appendix: Table 1S). The 534-bp fragment (Mt-9) was the most common, followed by the 224-bp fragment (Mt-1).

The average number of mitotypes per population was 2.4, with variation from 1 to 5, and the average N_{hap}/N_{ind} index was 0.09, with fluctuations between 0.03 and 0.20. The average haplotype diversity (H_{mt}) was 0.329 (Appendix: Table 2S). Low N_{hap}/N_{ind} index scores (≤ 0.03) for specific genetic diversity were observed for 21 natural populations from AT-03 (Austria); BA-06 (Bosnia); BG-03, 05 and 12 (Bulgaria); GR-01 (Greece); HR-04 (Hrvatska); IT-01, 02 (Italy); RS-02, 03 and 04 (Serbia); and TR-07, 08, 13, 17, 18, 19, 20, 25 and 27 (Turkey). High scores (≥ 0.15) were observed for 12 natural populations from AL-02 (Albania); AT-01 (Austria); CP-03 (Crimean Peninsula); ES-04 (Spain); HR-02 and 03 (Hrvatska); IT-03 (Italy); MK-04, 06 and 07 (Macedonia); and TR-10 and 24 (Turkey). [R-27, 28] Single haplotype population structure was observed at 20% for the following studied natural populations: AT-03 (Austria); BA-06 and 08 (Bosnia); BG-05 and 12 (Bulgaria); CP-01 (Crimean Peninsula); GR-01 (Greece); HR-04 (Hrvatska); IT-01 and 02 (Italy); RS-02, 03 and 04 (Serbia); and TR-07, 08, 13, 17, 18, 19, 20, 25 and 27 (Turkey). Low haplotype diversity (≤ 0.33) was observed at approximately 47% for the studded natural populations from AT-02 and 03 (Austria); BA-06, 07, 08, 09, 10 and 11 (Bosnia); BG-03, 04, 05, 06, 10, 11 and 12 (Bulgaria); CP-01 (Crimean Peninsulas); CY-01 and 02 (Cyprus island); FR-02 and 06 (France); GR-01 (Greece); HR-04 (Hrvatska); IT-01 and 02 (Italy); RS-02, 03 and 04 (Serbia); RU-01 (Russia); and TR-05, 07, 08, 09, 10, 11, 12, 13, 14, 16, 17, 18, 19, 20, 21, 22, 23, 25, 26 and 27 (Turkey). High correlation coefficients were observed for the number of haplotypes versus haplotype

diversity (0.81), followed by for the N_{hap}/N_{ind} index versus H_m index (0.71) and the N_{hap}/N_{ind} index versus the number of haplotypes (0.67).

Among-population structure

Population level: All distance methods showed low discrimination power with ten thousand permutation (or bootstrap) resamplings for all *P. nigra* populations (data not shown). The initial Bayesian method was able to divide all 106 studied populations into 11 principal geographic locations according to genetic variability (Fig. 2). The rarefaction analysis suggested using a sample size of 50 individuals (at $p \leq 0.01$) per population/meta-population to avoid false-positive results in the meta-population structure and to improve the robustness of statistical analysis (data not shown).

Meta-population level: The initial meta-population genetic analysis in each principal geographic location—western Mediterranean, eastern Mediterranean, southeastern Europe and Asia Minor—showed phylogeographic structure with statistical support for one to three groups (meta-populations). The distribution of the eleven black pine meta-populations from nonadmixing Bayesian analysis was as follows: (1) Morocco; (2) Spain; (3) France; (4) Corsica; (5) western Balkan Peninsula; (6) western Greece; (7) eastern Balkan Peninsula; (8) Turkey; (9) northern Turkey; (10) Italy; and (11) the Crimea-Eastern Black Sea coast (Fig. 2). Within each meta-population, the among-population structure was not significant, and the phylogeographic structure was weak (Mantel test, IBD and BARRIER programs: data not shown). The $N_{st} < G_{st}$ ratios obtained here (i.e., 0.203 vs. 0.231 and 0.234 vs. 0.278, from Table 1) confirmed the low signal of phylogeographic structure in *P. nigra* that was reported
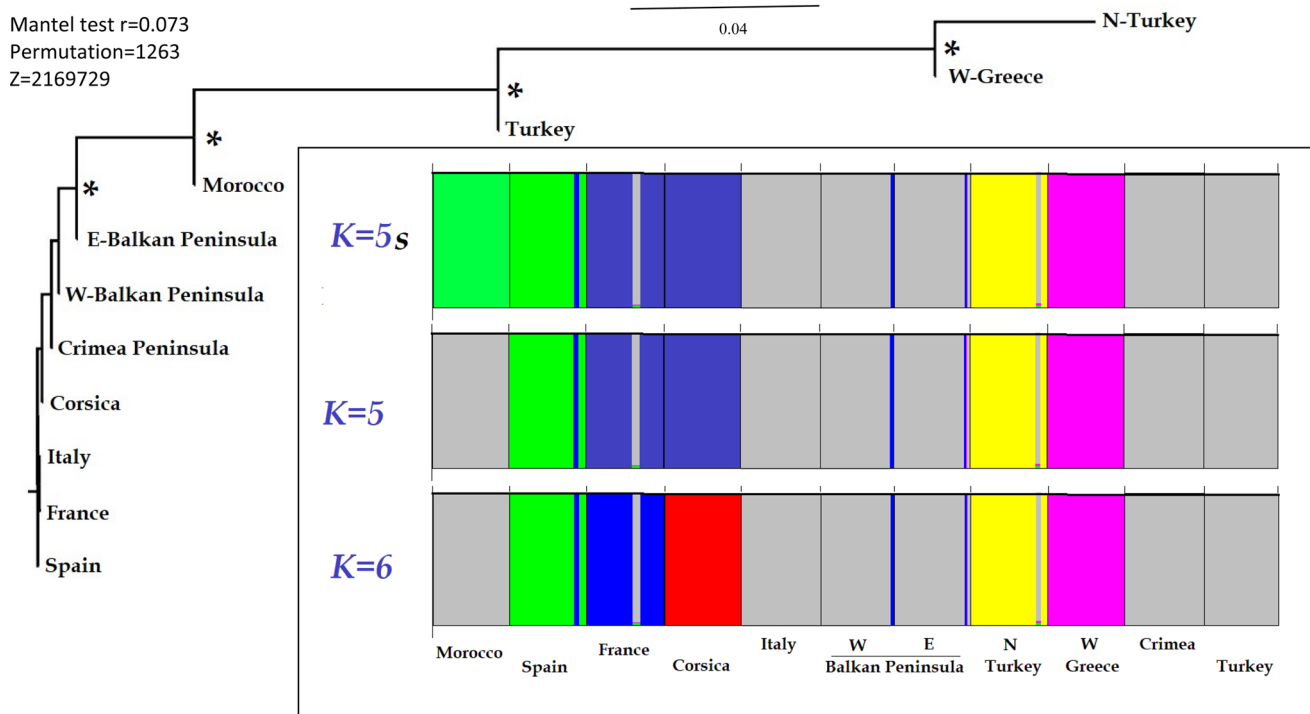
Fig. 2 The neighbor-joining (NJ) agglomerative construction clustering method of the Goldstein (1995) genetic distance among principal geographic meta-populations (top left). The asterisks are for bootstrap support by minimum 67%. The similar results are obtained by UPGMA (unweighted pair-group method with arithmetic mean) hierarchical clustering method by same distance (fig not shown). The colored chart is the result from BAPS's admixing Bayesian algorithm with spatial interference ($K=5$ s) and nonspatial interference $K=5$ and $K=6$

by Pons and Petit (1996). The meta-populations included a minimum of 2 populations (Morocco). In each meta-population, private mitotype presence was scored. We reanalyzed the black pine meta-populations by a sample size of 50 individuals with the proportional presence of all mitotypes. The nonadmixing meta-population number of distinct mitotypes was between 2 (N-Turkey) and 7 (France; av. 4.2). The nonadmixing meta-population mitotype diversity (H_m) was between 18.00 (N-Turkey) and 76.28 (France; av. 54.79, not biased). The nonadmixing meta-population fixation index (Fst) was between 0.185 (N-Turkey) and 0.305 (Corsica; av. 0.29 at 0.128 SD). The nonadmixing meta-population N_{hap}/N_{ind} index ranged from 0.04 (N-Turkey) to 0.14 (France; av. 0.08: Table 1).

Among all distance methods, only that of Goldstein et al. (1995) showed results similar to the nonspatial Bayesian admixture results with 10 thousand bootstrap replicates (Figs. 1 and 2). Poor results were obtained by the Mantel test and isolation by distance (IBD) analysis. The spatial Bayesian admixture analysis (BAPS program) split all meta-populations into 5 groups (at $K_{true}=5$): (A) Morocco-Spain; (B) France-Corsica; (C) western Greece; (D) northern Turkey; and (E) principal area (i.e., the remaining part of the natural area of black pine: Figs. 1 and 2). A similar result was

obtained by the spatial analysis performed by the STRUCTURE and TESS algorithms (at $K_{true}=5$; data not shown). The analysis was also confirmed by the good results of the BARRIER program. The more distinct black pine meta-populations were in western Greece and northern Turkey, as determined by the first two phylogenetic barriers (Fig. 3). The nonspatial Bayesian admixture analysis performed with the same programs split the *P. nigra* populations only between Morocco and Spain due to the presence of private mitotypes in the mountain region of the Iberian Peninsula. The meta-population analysis showed very good discrimination power for the molecular marker used, with $F_{st}=0.290/0.278$ (average/AMOVA-global), $G_{st}=0.231/0.278$ (AMOVA-pairwise/AMOVA-global), $R_{st}=0.192/0.233$ (AMOVA-pairwise/AMOVA-global), $\Phi_{st}=38.02$ (SAMOVA among groups), $N_{st}=0.203/0.234$ and $F_{st\text{-coalescent}}=0.111$ (average). The admixture meta-population number of distinct mitotypes was between 2 (N-Turkey) and 8 (France & Corsica; av. 5.4). The admixture meta-population mitotype diversity (H_m) was between 18.18 (N-Turkey) and 81.41 (France & Corsica; av. 49.42, not biased). The admixing meta-population fixation index (F_{st}) was between 0.003 (France & Corsica) and 0.205 (W-Greece; av. 0.11 at 0.07 SD). The admixing meta-population N_{hap}/N_{ind} index ranged from 0.03 (Principal area)
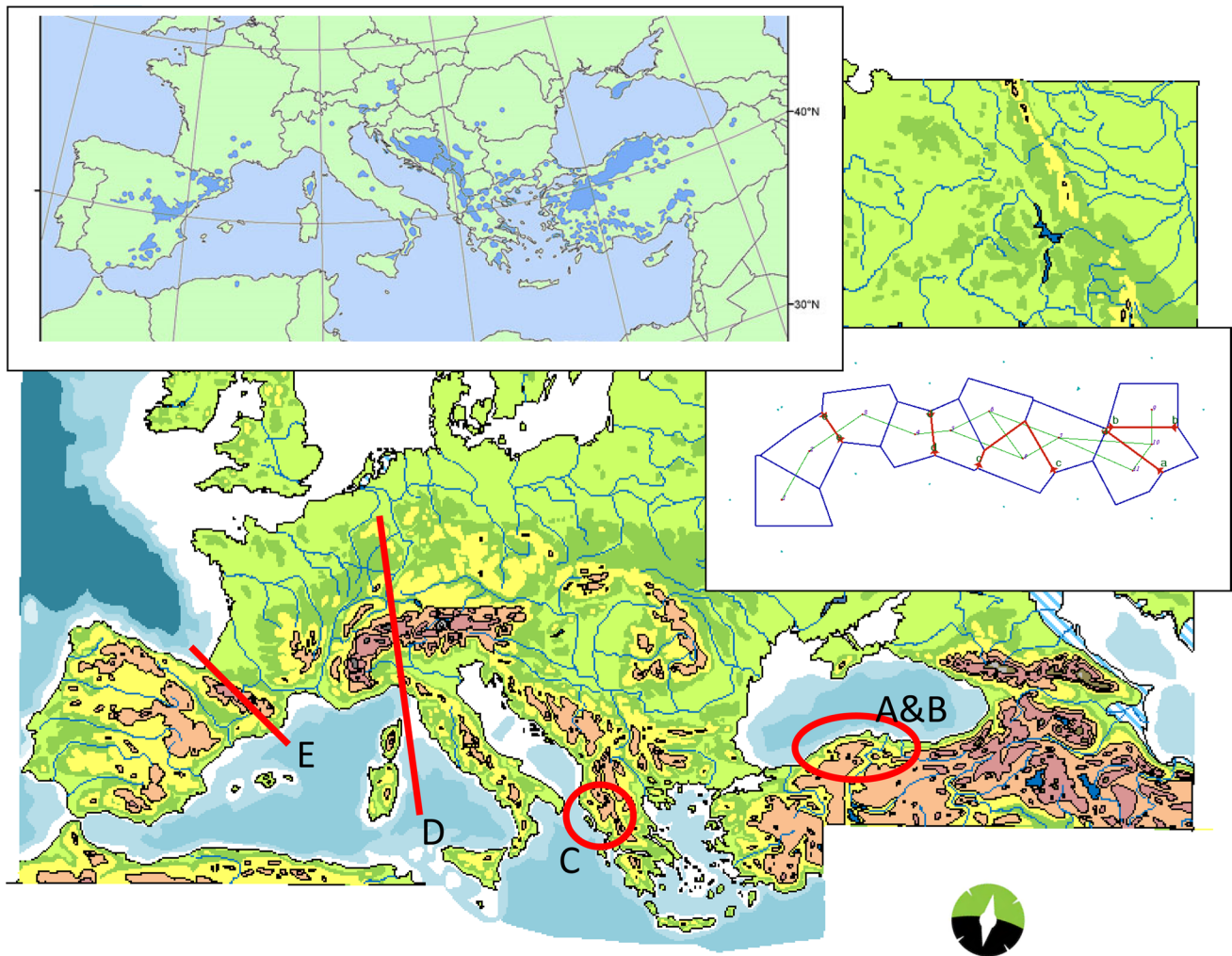
Fig. 3 The natural range of *Pinus nigra* according IUFRO (top left) and the principal five European Black Pine meta-population geographic interpolation determined by two different algorithms implemented in BAPS, BARRIER programs. The barrier range shown best support for barrier A&B (endemic region of Northern Turkey), flowed by C (western Greece endemic region), D (Italian- French barrier) and low support for E (French-Spain barrier)

to 0.08 (France & Corsica; av. 0.06: Table 1). The indices used here are complementary and reflect different aspects of admixing meta-population variability (Table 1). The admixing meta-population variability indices showed lower scores than those of the nonadmixing meta-population.

The migration pattern between the principal geographic centers showed that black pine migration generally occurred during the Pliocene (1.8–5.3 Ma) and Miocene (5.3–23.0 Ma) epochs and rarely exceeded the threshold of 15 thousand individuals (*cs*), i.e., census size (Appendix: Fig. 1S/1–16). The more recent significant migration waves (approximately ≥ 10 thousand individuals) were from the central part of the Balkan Peninsula to western Greece (approximately 1.8 Ma; early Pleistocene: Appendix: Fig. 1S/11). The migration waves from Western Europe (France) to the western part of the Balkan Peninsula were

even older (approximately 18.5 Ma; Early Miocene: Appendix: Fig. 1S/1), followed by those from the Balkan Peninsula to central Turkey, the Balkan Peninsula to Italy, the Balkan Peninsula to Crimea (and vice versa) and France to Spain (approximately 11.1 Ma; late Miocene). Generally, significant *P. nigra* migration waves were time- and volume-asymmetric. An exception was observed for time for the migration pattern between the Balkan Peninsula and Crimea. Appendix: Fig. 1S/1–16 shows simultaneous symmetric migration signals (fewer than 10 thousand individual's census size) of historical natural area expansions and contractions. The majority of migration waves involved fewer than 10 thousand individuals (*cs*), and one wave (i.e., one panel from Appendix: Fig. 1S/1–16) covered a period of approximately 300 thousand years. We speculate that each wave in the figure consisted of multiple migration waves of

at most 5–8 individuals per generation and per population in the historical contact zone. Generally, pine's effective radius of seed distribution is 3–5 times shorter than its effective radius of pollen distribution.

The average black pine historic effective population size was between 529 and 4015 individuals (Bayesian statistics: Appendix: Table 3S). The maximum likelihood (MCMC) method provided similar results of 2039 versus 1974 when using author program recommendations—from 78 up to 2778 individuals (Beerli 2006, 2009). The most recent common ancestor (MRCA) lived from 5.46 Ma to 1.97 Ma (av. 3.94 Ma, i.e., 57.93×10^3 generations). Corsica Island had the oldest MRCA signal from 5.06 Ma to 5.49 Ma (i.e., $74.36/80.33 \times 10^3$ generations). Logically, both meta-populations from 2 endemic regions showed the highest MRCA signal from 1.97 to 2.36 Ma (i.e., $29.00/34.78 \times 10^3$ generations for N-Turkey) and from 2.57 to 2.91 Ma (i.e., $37.75/42.75 \times 10^3$ generations for W-Greece). The contribution of migration waves to effective population size was between 22 and 880 individuals under maximal migration waves of 50–75 thousand individuals (Appendix: Table 3S and Fig. 1S). The mutation rate (μ) for *NAD7-intron1* was calculated to be approximately 3.6×10^{-6} per generation (i.e., $3.5–3.7 \times 10^{-6}$ per generation under a 5% SE), and the generation of one generation was 68 years on average for European black pine, the same as that determined by Naydenov et al. (2016). The model keys of effective population size fluctuations and migration patterns are presented in Appendix: Fig. 2S/4.

The control *P. sylvestris* metapopulation average effective population size (N_e) was 1434 individuals for Bayesian statistics and 1572 individuals for ML statistics. The minimal N_e values were observed from the endemic region of N-Turkey (i.e., 181–380 and 683–1014 individuals from ML and Bayesian statistics, respectively) and the maximal values from 788 to 2731 individuals from the C-Balkan Peninsula and western Russia. The contribution of migration waves to effective population size was between 63 and 993 individuals under maximal migration waves of 20 thousand individuals (Appendix: Table 3S and Fig. 2S). The average time of MRCA is 2.21 Ma (i.e., 32.54×10^3 generations); the oldest signal was from the C-Balkan Peninsula—between 2.38 and 2.48 Ma (i.e., between 35.03 and 36.42×10^3 generations), and the younger signal was from the N-Turkey endemic region—between 1.87 and 2.08 Ma (i.e., between 27.53 and 30.53×10^3 generations).

Historical migration pattern

Eastern area: Current natural European black pine populations are concentrated in the following regions: the Balkan Peninsula, Crimea and Asia Minor (Fig. 3). The obtained data showed that on the Balkan Peninsula, the migration

waves were very restricted and generally driven by in situ surface area fluctuations over highly variable topography. This historical scenario occurred during the late Miocene to Pliocene epoch. The positive effect of climatic change resulted in time-symmetric migration with a low intensity of fewer than a 3 thousand individual census size for the period of species range expansion [R-29]. In western Greece, a specific local cryptic mitotype was observed (pink color in the circles in Fig. 1). The migration pattern between western Greece and the eastern-central part of the Balkan Peninsula was different, with strong migration in the Pliocene epoch. The first wave of approximately 50 thousand individuals (*cs*) for one period of 300 thousand years occurred approximately 4.5 Ma (early Pliocene epoch) from western Greece, followed by two opposite, large, asymmetric waves from the central Balkan Peninsula to western Greece at approximately 3.7 and 1.6 Ma (i.e., in the Pliocene and late Pliocene/early Pleistocene epochs, respectively). Each migration wave involved at least 50 thousand individuals (*cs*) for one period of 300 thousand years. We speculate that the cryptic mitotype (Mt-10/pink color) arose after the large migration in the early Pliocene epoch (i.e., after 4.5 Ma: Appendix: Fig. 1S/11).

The migration between western Greece and the western Balkan Peninsula and between the central and western Balkan Peninsula was restricted to fewer than 5 thousand individuals. It is interesting that the census sample size of migration contrasted with the effective population size during the different migration waves. Although low-intensity migration was detected between western and eastern-central Balkan Peninsula meta-populations, the migration wave effective sample size was between 150 and 698 individuals, with a trend in the westward direction. The exchange between western Greece and the remaining part of the Balkan Peninsula was significantly misbalanced. The trend showed that the Ne transferred from migration waves to the remaining part of the Balkan Peninsula was modest at approximately 57–88 individuals, and in the opposite direction from the Balkan Peninsula to western Greece, it was 3–5 times more significant, i.e., an effective size of 148 to 493 individuals (Appendix-Table 3S). The effective population size of *P. nigra* was highest (1986 to 2778) in this region for the central and western Balkan Peninsula. In Greece, the cryptic mitotype had a low Ne of 186 to 250 individuals. The time to the most recent common ancestor was between 3.12 and 3.55 Ma for the majority of the area (i.e., middle Pliocene). The TMCA for the western Greece meta-population was relatively recent, at 2.57–2.91 Ma (i.e., middle Pliocene: Appendix: Fig. 1S/15).

The Asia Minor black pine populations showed similar population structures with the dominance of both principal mitotypes. An exception was noted in the northern part of Turkey, in the Ilgaz Mountains of the Pontic Mountains

system (with synonyms such as the Pontic Alps or North Anatolian Mountains complex), where one cryptic mitotype (Mt-11; yellow color in Fig. 1) was discovered. This region appears to be a hotspot of endemism because the same molecular marker showed a cryptic mitotype for Scotch pine (Naydenov et al., 2007; Fig. 1, yellow color), which will be discussed at the end of the present manuscript. The effective population size in this region was very small (78–106 individuals; ML), and TMCA was shortest of all those calculated here, at 1.97 to 2.36 Ma (i.e., late Pliocene: Appendix: Fig. 1S/6). The *P. nigra* migration from northern Turkey to the remaining part of Asia Minor was restricted to fewer than a 2500 individual census size per wave 9 Ma (i.e., from the late Miocene). Each wave covered a period of approximately 300 thousand years, and multiple waves almost showed a symmetric pattern, which is an indication of historical area fluctuations generally related to global climatic change. The effective population size contribution of all waves to migration to the remaining part of Turkey (i.e., central Turkey) was a very modest at 50–73 individuals (Appendix: Table 3S). In the opposite direction, the migration intensity was twice as high for the same period of time at approximately 9 Ma; i.e., there was a northern migration pattern with a historical migration peak at approximately 4 Ma (early Pliocene). The N_e contribution of all waves to the migration of black pine to northern Turkey was 22–337 individuals. The effective population size from central Turkey was very high, from 2431 to 2728 individuals, and the time to the most recent common ancestor was between 3.28 and 3.35 Ma (both in the middle Pliocene).

The *P. nigra* mitochondria results reported here showed that migration between the two areas (i.e., the Balkan Peninsula and Asia Minor) occurred at least 11.1 Ma (the Late Miocene). Three significant waves from the eastern-central Balkan Peninsula occurred at approximately 11, 7 Ma (both in the late Miocene) and 4 Ma (middle Pliocene: Appendix: Fig. 1S/7). Each wave lasted for one period of 300 thousand years. The first wave included approximately 55 thousand individuals; the remaining two included approximately 10 thousand individuals (census size) each. In the opposite direction, from central Turkey to the Balkan Peninsula, migration was more restricted and consisted of two waves at approximately 11.1 and 9.5 Ma (both in the late Miocene). The migrant number was approximately 10 thousand individuals for the first wave and approximately 20 thousand individuals (*cs*) for the second wave. The effective population sizes of black pine historical migrations were disproportional, with approximately 848–880 individuals from the Balkan Peninsula and 633–645 individuals from central Turkey (Appendix: Table 3S). We speculate that due to the nonsignificant presence of symmetric migration waves, the migration history between central Turkey and the Balkan Peninsula was associated with the tectonic history in

the region rather than global climatic change. The genetic exchange between northern Turkey and the Balkan Peninsula was generally low and featured fewer than 5 thousand individuals, with two exceptions—one wave of 15 thousand individuals to the Balkan Peninsula at approximately 5.8 Ma and one wave of a 18 thousand individual census size in the opposite direction at approximately 0.2–0.3 Ma (late Pleistocene, Appendix: Fig. 1S/5). The low *P. nigra* migration census size in both directions contrasts with the misbalance in the effective population size contribution; from the Balkan Peninsula, the migration wave was 752–822 individuals, and that in the opposite direction was 65 to 80 individuals. We speculate that this signal is from the tectonic history related to the Para-Tethys Sea in this region.

The coastal Black Sea *P. nigra* populations (Russian and Crimean peninsulas) are the northeastern natural populations at present. The results showed that the historical effective population size for this region fluctuated between 1736 and 1986 individuals (at ML) and that the TMCA ranged from 3.45 to 4.21 Ma (i.e., from the middle Pliocene to the early Pliocene). The migration pattern between the NE Black Sea coastal population and central Turkey was balanced and consisted of very few individuals, with no more than 7 thousand individuals (*cs*) in multiple waves in both directions. This small census size migration contrasts with the large effective population size contributions between 432 and 958 and 405 and 818 (Appendix: Table 3S & Fig. 1S/8). The symmetry of the black pine migration waves was the result of historical global climatic change in the natural area of the ancestral species in the same region. We believe that the natural area of the ancestral species distribution was larger than the present-day distribution. The migration history in this area was different than that on the Balkan Peninsula, with a strong imprint of the past tectonic history of the Black and Para-Tethys Seas. From the Balkan Peninsula to the coastal Black Sea *P. nigra* populations, two significant waves of approximately 85 thousand and 44 thousand individual census sizes were noted at approximately 14 Ma (middle Miocene) and 7 Ma (late Miocene), respectively. In the opposite direction, only one significant wave of approximately 22 thousand individuals (*cs*) at approximately 13.8 Ma (middle Miocene) was identified. Each *P. nigra* migration wave occurred for one period of 300 thousand years, and the effective population size contributions for the migration from the Balkan Peninsula to the coastal Black Sea and the opposite migration were 142–880 and 85–697 individuals, respectively (Appendix: Table 3S & Fig. 1S/9).

The historical black pine migration between the western Mediterranean and eastern areas within its natural distribution was related to the tectonic history of the Alps mountain complex and Apennine Peninsula; i.e., no symmetric migration waves were detected. The historical areas of *P. nigra* were connected between its largest areas by the northern

and southern routes from the present Alps mountain complex. The two meta-populations shared a common ancestor (i.e., TMCA) given their present population structure from 3.59 Ma (middle Pliocene) to 5.17 Ma (early Pliocene), and their historical effective population size fluctuation was between 1475 and 2717 individuals (at ML; Appendix: Table 3S). The northern route was probably a small contact zone with two limited waves from the west to the east of approximately 8–9 thousand individuals (*cs*) at 19.5 Ma (early Miocene) and 5.5 Ma (late Miocene: Appendix Fig. 1S/1). In the opposite direction, one wave of approximately 10 thousand individuals (*cs*) occurred 8.8 Ma, and the effective population size contributions in the two directions were 85–138 and 108–138 individuals, respectively. The southern black pine route showed a dominant western historical migration pattern with two waves of approximately 40 thousand and 8 thousand individuals based on census size 10.8 and 6.6 Ma (both in the late Miocene), respectively, with historical Ne contributions of 752–863 individuals. The opposite direction migration signal was very low and involved fewer than 5 thousand individuals (*cs*), with modest effective population size contributions of 85–138 individuals (Appendix: Table 3S & Fig. 1S/3).

Western Mediterranean area: European black pine populations from this region are fragmented at present (Critchfield and Little, 1966), and phytogeographic data are available from archived IUFRO maps online (Fig. 3). The restricted symmetric historical fluctuation resulted from large climate change-driven species range expansions between France and Italy at approximately 6 Ma (late Miocene); between Morocco and Spain at approximately 8 and 11 Ma (both in the late Miocene); between Spain and France at approximately 9 and 11 Ma (both in the late Miocene); and between Corsica and France at 3.7 Ma (middle Pliocene). These fluctuations generally contributed little to the mitochondrial population structure. The exchange involved no more than 10 thousand individuals over a period of approximately 300 thousand years (Appendix: Fig. 1S/2, 4, 10, 12, 14, and 16). Asymmetric migration generally results from tectonic topographic history, and we speculate in this case, this asymmetric migration of *P. nigra* population expansion occurred under the positive effects of climatic change, where more adaptive phenotypes were consistent with forest successional species patterns and low-frequency forest fire periods. Examples include multiple migration waves from Morocco to Spain at approximately 4, 6.5, 7.7, 7.9 and 11.1 Ma (all from the late Miocene to the early Pliocene); Spain to France 12 Ma (middle Miocene); France to Italy 5.5 and 8 Ma (both in the late Miocene); Italy to France 7.2 Ma (late Miocene); Corsica to France 5.5 Ma (late Miocene); and Spain to Corsica 5.5 Ma (late Miocene). The significant asymmetric migration pattern probably resulted from important tectonic regional historic events, such as the

Gibraltar Strait, Corsica Island, and rise of the Pyrenean and Alps mountain complexes. The obtained results showed a significant migration signal from France to Spain of approximately 50–60 thousand individuals at approximately 9.3 and 11.1 Ma (both in the late Miocene) that was related to historical Pyrenean topography dynamics. A similar signal was observed from the initial period of the Messinian salinity crisis (approximately 6 Ma: late Miocene) for migration from France to Corsica, with an approximate census size of 20 thousand individuals, and a smaller event (approximately 5 thousand individuals) was detected for the final period of the Messinian salinity crisis from Morocco to Spain at approximately 4 Ma (early Pliocene). The remaining four migration waves from the NW African coast to the Iberian Peninsula occurred before the Zanclean flood (late Miocene). The moderate historical black pine migration pattern of approximately 10 thousand individuals (*cs*) from Italy to Corsica (approximately 10.1 Ma: late Miocene) was probably related to the historical development of the Alps mountain complex (Appendix: Fig. 1S/2, 4, 10, 12, 14 and 16). Different migration waves provided different contributions to the effective population sizes. The relationship between the number of migrants and effective size contributions was not linear. For example, the effective size contribution of the migration waves from Morocco to Spain was 555–885 for a total of 5 waves of no more than 5 thousand individuals each. The opposite direction migration waves were difficult to distinguish, and they had low intensities, with fewer than 1.25 thousand individuals and effective population sizes of 240–410 individuals. A significant contribution to the effective population size of *P. nigra* was observed for the migration waves between Corsica and Italy and between France and Spain (Appendix: Table 3S).

The effective population sizes were maximal in France (2514–2778 at ML), followed by in Spain (2147–2756 at ML), Corsica (1147–2408 at ML), Italy (2181–2242 at ML) and Morocco (1464–2186 max at ML). The result from the Bayesian analysis ranked Spain, followed by France, Italy, Corsica and Morocco, as having the most pronounced differences. The comparison between the two methods is within the scope of the present manuscript and is well described by the authors of the program MIGRATE (Beerli 2006, 2009). The time to the most recent common ancestor (TMCA) of European black pine from the western Mediterranean region was approximately 4 Ma (early Pliocene), and the oldest ancestor was from Corsica (5.06–5.45 Ma: late Miocene–early Pliocene), which was exactly 2 Ma after the end of the Messinian salinity crisis (Appendix: Table 3S).

P. nigra versus *P. sylvestris* among species parallel analysis: The between-species parallel historical migration pattern comparison is very interesting and supports the obtained result. European black pine and Scotch pine from similar regions (northern Turkey and Eastern Europe) were chosen

from different altitudes. The biological limits of the first species (*P. nigra*) allow it to generally reach altitudes from sea level up to 800 m, and those of the second (*P. sylvestris*) range from 1200 up to 2000 m (Vidakovic [1974], [1991]; Dobrinov [1983]; Alexandrov et al. [1988]). Both species are from the same genus, *Pinus*, and subgenus (i.e., *Diploxylon*), but these two species have different biology, forest community succession patterns and forest fire frequency and intensity tolerances (Naydenov et al. [2012]). For the meta-population analysis presented here, the geographic region similarity was very good. The previous phylogeographic mitochondrial analysis of Scotch pine by the same team in Naydenov et al. ([2007]) did not perform statistical model-based analysis of historical effective population size fluctuation and the timing, frequency and intensity of asymmetric migrations. The control sample of Scotch pine showed historical effective population sizes from 683 to 2538 individuals; the MRCA lived between 2.48 and 1.87 Ma (late Pliocene), and the effective population size contribution from migration waves ranged from 63 to 993 individuals, with a maximal migration wave of a 15–20 thousand individual census size (Bayesian statistics, Appendix: Table 3S).

The *P. sylvestris* migration pattern was scored between the Balkan Peninsula and western Russia/central Europe (i.e., W Russia) in both directions for a total of 3–4 waves. Significant waves from W Russia to the Balkan Peninsula were small, with approximately 5 thousand to 10 thousand individuals (*cs*) 9.2, 8.8 and 5.5 Ma (in the late Miocene, Appendix: Fig. 2S/1). This historical migration pattern was predated in the opposite direction by one large wave of approximately 22 thousand individuals (*cs*) 9.5 Ma (late Miocene); i.e., we observed the northward migration of a Scotch pine census population of moderate size from the Balkan Peninsula to western Russia and central Europe. Their effective population size contributions to the Balkan Peninsula to W Russia were moderate at 680–793 and 960–993 individuals, respectively (Appendix: Table 3S).

The historical *P. sylvestris* migration pattern between western Russia/central Europe and the endemic genetic pool from northern Turkey was generally not symmetric. Migration to northern Asia Minor occurred through multiple low-intensity waves, and two low-intensity waves of approximately a 10 thousand individual census size were more significant 5.6 Ma (late Miocene) and 5.0 Ma (early Pliocene, Appendix: Fig. 2S /2). In the opposite direction, historical migration waves were very small, with fewer than 2 thousand individuals (*cs*). Their effective population size contributions of the wave to Asia Minor and that in the opposite direction were small and showed relatively similar patterns of 92–125 and 65–113 individuals, respectively (Appendix: Table 3S).

This particular result is related to that for the historical Scotch pine migration pattern between the Balkan Peninsula and N-Turkey. The migrations from the northern Asia Minor endemic genetic pool region to the Balkan Peninsula were very small, under the threshold line. In the opposite direction, two significant waves of approximately 15 thousand and 7 thousand individual census sizes were observed at approximately 5.5 Ma (late Miocene) and 2.2 Ma (late Pliocene, Appendix: Fig. 2S /3), respectively; their effective population size contributions were small and showed relatively similar patterns of 100–123 and 65–135 individuals, respectively. The TMCA was between 1.87 Ma (late Pliocene) and 2.48 Ma (late Pliocene), and the effective population size historical fluctuations were in the ranges of 181–380 individuals for N-Turkey, 2619–2731 individuals for the Balkan Peninsula, and 1167–2353 individuals for Western Russia and Central Europe (Appendix: Table 3S).

Interestingly, the Scotch pine migration wave 5.5 Ma (late Miocene) from the northern Asia Minor endemic genetic pool region to the Balkan Peninsula seems logical given the successional migration of black pine 5.8 Ma (late Miocene) with a similar census population size of 15 thousand individuals from the same region described above (Fig. 4). The two species have shown different directions of historical migration in successional patterns and probably have differences in biology related to altitudinal differences (such as forest community composition, soil conditions and forest fire disturbance adaptations) resulting from specific historical altitude tectonic changes (Naydenov et al. [2016]/Fig. 4; Naydenov et al. [2017]/Fig. 3 and Appendix Figs. 2S and 3S).

Discussion

Genotype versus phenotype

Chloroplast and mitochondrial life histories and origins are very long and complex. For all molecular markers, plastids or nuclear DNA is difficult to directly relate to a specific phenotype. Generally, tree phenotypes (i.e., physiological and morphological traits) are under polygenetic control. We believe that WGA (i.e., whole genomic analysis) combined with the epigenetic landscape (i.e., non-DNA heritability) is necessary to fully understand multiple morphologically distinct black pine (and Scotch pine) subspecies and varieties. Briefly, genomic analyses include high-resolution linkage maps (HR-LMs), high-resolution regulatory genetic networks (HR-GNs) and high-resolution genome-wide association studies (HR-GWASs). The large genomic dataset generally includes thousands and thousands of loci from noncoding DNA, and only machine learning analysis allows the construction of polygenetic predictors for specific traits. An epigenetic analysis will complete a genomic analysis, which is within the scope of the present manuscript.
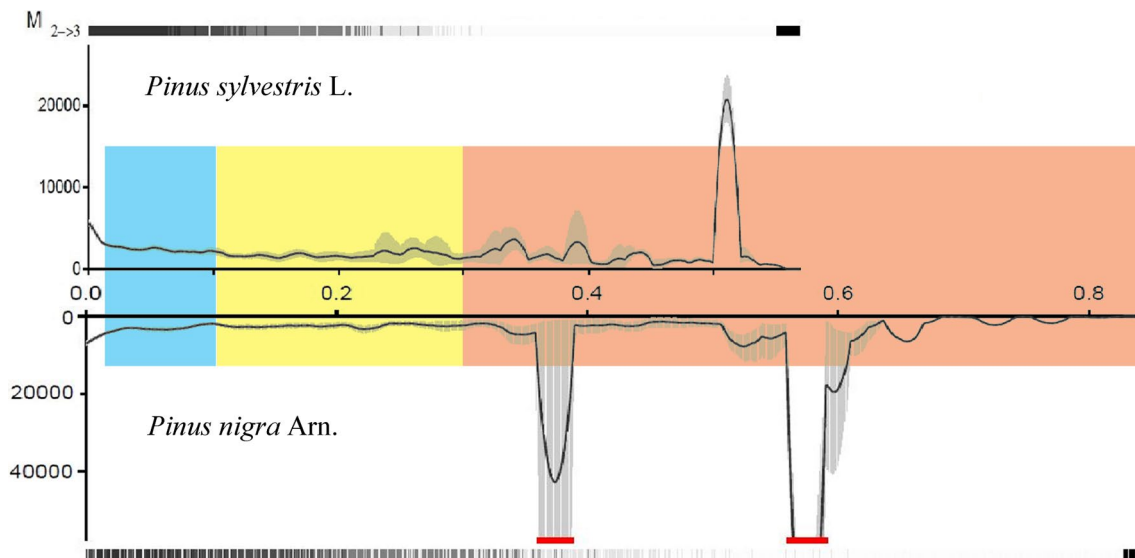
Fig. 4 Migration pattern from the Balkan Peninsulas to the Western Russia (at the top); and the migration pattern from the Balkan Peninsulas to the Crimea Peninsulas (at the bottom). The vertical colored bar limit is 15 thousand individuals; and the horizontal colored bar in the middle of the figure represents geological time grid as white color—*Holocene* epoch (up to 0.01 scanned units), blue light color—*Pleistocene* epoch (between 0.01 and 0.1 scanned units), yellow color—*Pliocene* epoch (between 0.1 and 0.3 scanned units), and orange color—*Miocene* epoch (≥ 0.3 scanned units). In the present manuscript, we have set the base of the *Pleistocene* (*Quaternary* Period) at 2.6 Ma. The traditionally used the *Tertiary* Period comprises the *Pliocene* epoch (2.6–5.3 Ma), the *Miocene* epoch (5.3–23.0 Ma) and the *Paleogene* epoch (23.0–65.5 Ma). For the reverse directions migration patterns see Appendix Fig. 1S and 2S

Microevolution, generally at the population level, is based on the concept that natural selection acts directly on morphological, biochemical and physiological phenotypes related to genetic variations. The accumulation of surviving genetic changes at some critical level leads to new morphological, biochemical and physiological changes; this is similar to Jacob's (1977) concept of tinkering evolution for a short time. Of particular interest are silent mutations; modifications of regulatory genes; chromosome aberrations, i.e., segment inversions; and other changes. Within-population polymorphisms from regulatory gene modifications could affect the quantitative aspect of structural gene expression, according to Purugganan (1998). Initially, the changes are unstable and at a low frequency in the population for some period of time; after the change are fixed, a new organismal trait is generated. Similarly, the concept of fitness landscapes (or adaptive landscapes) and the related Wright's dilemma were developed early by Wright (1932).

The multiple descriptions of the observed forest tree ancestral morphological traits, have been published at present day including those of *P. nigra* and *P. sylvestris* in this study, result from genomic fossils (i.e., pseudo-genes) imprint on cis- and/or trans-regulatory element transcription on the group of linked genes controlling specific phenotypic traits. It seems that not all ancestral morphological traits are under selective pressure and hold an adaptive evolution signal. The possibility that some silenced mutations will be reactivated into protein-coding sequences is known as Dollo's law (Marshall et al. 1994). The first hypothesis of reactivated genomic fossil sequences for forest trees was presented by Stefanov and Nedjalkov (1956), Stefanov (1958, 1972, 1973) as reversed mutations for the endemic *Quercus tracica* from Bulgaria.

Molecular clock

At present, it is not clear exactly how much an organelle's (mitochondrion or chloroplast) structure is affected by short-term natural selection at the microevolution level from our analysis. Organelles from plants, such as chloroplasts and mitochondria, are used as good molecular markers for macro- and microevolution, phylogenetic and phylogeographic analyses (Burban and Petit, 2003; Naydenov et al. 2007). These organelles generally show different life-history and evolutionary patterns. From our results, it seems that (1) in comparison with mitochondrial structure, chloroplast structure (i.e., *SNP*s and *SSR*s) is more affected by global long-term climatic changes and (2) mitochondrial structure (i.e., *INDEL*s) is more affected by tectonic history fluctuations (Naydenov et al. 2016, 2017: nuclear, chloroplast, mitochondria).

According to the results obtained here, the phylogeographic structure inferred from the mitochondrial *INDEL*s for the *Pinus* genus is more properly interpreted in light of the tectonic history in the Pliocene-Miocene epoch than that in the Last Glacial Maximum (LGM) or middle or late

Pleistocene (age). We agree with the *Pinus pinaster* discussion by Burban and Petit (2003) that the *SNPs* (single nucleotide polymorphisms) from mitochondrial DNA probably date back to the LGM. This result is supported by not only our result here but also the simple deduction that the molecular markers have very different molecular masses, i.e., significant differences in base pairs (bps), and that the mutation frequency for large DNA fragments is very low and requires more generations for population structure establishment and fixation, i.e., a relaxed molecular clock. For example, here, the mutation rate (μ) for *NAD7-intron1* size variants was calculated to be approximately 3.6×10^{-6} per generation (i.e., $3.5–3.7 \times 10^{-6}$ per generation at a 5% SE); in comparison, the average mutation rate (μ) for *P. nigra* from seven chloroplast DNA microsatellites (*SSRs*) with 13 regions of sequencing variation from our previous publications, Naydenov et al. (2016, 2017), was 5.6×10^{-5} (i.e., $5.4–5.8 \times 10^{-5}$) per generation at a 5% SE. Comparing the mutation rates, the difference occurred based on two factors: different fragment sizes and different organelle origins (i.e., chloroplast vs. mitochondrion). One generation was calculated to be an average of 68 years for the European black pine. Figure 5 compares the phylogenetic memory span in geological time for nuclei, chloroplasts and mitochondria from parallel historical migration analysis for some European pines. The migration pattern is an important factor for all plants and animals from an ecosystem perspective. At present, historical migration patterns are not integrated into conservation and biodiversity programs, although there is clear contribution potential based on theoretical development and accumulated empirical data from phylogeography.

For the *P. nigra* chloroplast analysis, historical effective population size fluctuations probably resulted from global climatic changes, forest fires and/or plant community successional patterns (Naydenov et al. 2017). For the mitochondria data and results here, it seems that forest disturbance and plant composition successional effects are nonsignificant. Similar conclusions were drawn for Scotch pine and other *Pinus* species. The obtained results and molecular clock calibration time mark in this study could be useful for research projects on *Pinus* species from regions without significant tectonic changes and a lack of time marks (e.g., North America). The impacts of global tectonic and climatic changes on different adaptive strategies and lineage evolution, especially for meta-phyla, have been reported by multiple authors, such as Chaloner and Hallam (1989), Knoll (1989), Martin and Barnosky (1993) and Shubin et al. (1997), based on paleo records.

Tectonic history and Pinus

Generally, phylogeography is related to global historical climatic and tectonic changes. The molecular marker used here and the historical meta-population migration pattern obtained are closely related to European orogenic architecture dynamics over the last 20 Ma from the early Miocene to the present (Fig. 4; Appendix Fig. 1S and Fig. 2S). More than two thousand scientific publications reported detailed information and possible scenarios for continental and regional historical tectonic change in southern and southeastern Europe and the Alpine folded regions (Muratov 1964; van Hinsbergen et al. 2020). We compare the reported results from European orogenic architecture dynamics to our results in the following part of the discussion.

West Mediterranean region: The historical western Mediterranean *P. nigra* natural distribution was probably affected by paleoclimatic dynamics, paleoflora distribution and paleoaltitude interactions with micro- and macrotectonic developments. The paleobotany record and palynological and carpological analyses have indicative and incomplete sources of information. The first important factor is the formation of the Pyrenean mountain complex and Gibraltar Strait as a result of the macrotectonic dynamics of interactions between the Iberian, African and European Plates (Appendix: Box-2). The second is the formation of the Apennine Peninsula, Alps mountain complex and Corsica Island, generally as a result of the collision between the African and European Plates (van Hinsbergen et al. 2020). For more information about tectonic history, please see Naydenov et al. (2016/Fig. 4) and Naydenov et al. (2017/Fig. 3 and Appendix: Fig. 2S, 3S).

Historical uplift from the Pyrenean mountain complex is consistent with the historical migration pattern between the French and Spanish *P. nigra* populations reported above (Appendix Fig. 1S/10, 12). The present results suggest that European black pine populations from Morocco and probably from northern Africa played important roles in species development and that they had long local historical patterns related to the Alps-Atlas orogenic history (i.e., geodynamic evolution). Similarly, 4 of the 5 migration waves from Morocco to Spain occurred before the reconvention of the Mediterranean Sea and the Atlantic Ocean (i.e., Zanclean flood: 5.33 Ma). Logically, before the Strait of Gibraltar opened, the principal black pine migration waves between Corsica and the European continental shelf (i.e., France/Spain/Italy) occurred (Appendix Fig. 1S/2, 4, 14, 16). The historical migration from Spain to Corsica detected here was also reported in our previous publications for *P. pinaster* and *P. nigra* and was determined with multiple nuclear and chloroplast *SSR* markers (Naydenov et al., 2014). For more information about paleoclimatology dynamics, please see Appendix: Box-2.

The shortened paleoaltitude dynamics from the Alps complex presented here support the hypothesis of historical northern Alps connections for the *P. nigra* populations from the W-Balkan Peninsula and France (Appendix Fig. 1S/1). The migration route was used in both directions over a period

Fig. 5 Comparative asymmetric migration pattern between the mitochondrial *INDEL*, chloroplast SSRs and nuclear SSRs from two Mediterranean Pines (*P. nigra* and *P. pinaster*). The red dashed line connected the red asterisk is the time mark of the Last Glacial Maximum (LGM ≈ 27,000 YBP). The horizontal colored bar in the middle of each chart represent geological time grid as white color—*Holocene* epoch (up to 0.0117 MY), blue light color—*Pleistocene* epoch (between 0.0117 and 1.806 MY), yellow color—*Pliocene* epoch (between 1.806 and 5.332 MY), and orange color—*Miocene* epoch (≥ 5.332 MY). The Pines seed radiation molecular markers as *INDEL* (top) and *nuSSR* (bottom) shown average low migration intensity of ≈ 10 thousand individuals per migration wave, but the pollen radiation marker as *cpSSR* (middle) average high migration intensity of ≈ 50 thousand individuals per migration wave. The used 3 molecular markers shown different phylogenetic memories: longest for *mt-INDEL*, following by *cpSSR* and shorten for *nuSSR*. *Note*: The horizontal axis for each chart is the time by mutation rate. The *time units* are different for the different molecular markers, respectively; average time span for each migration wave is different number of generations for each chart. The result from the middle and bottom charts is from Naydenov et al. (2014; 2017)
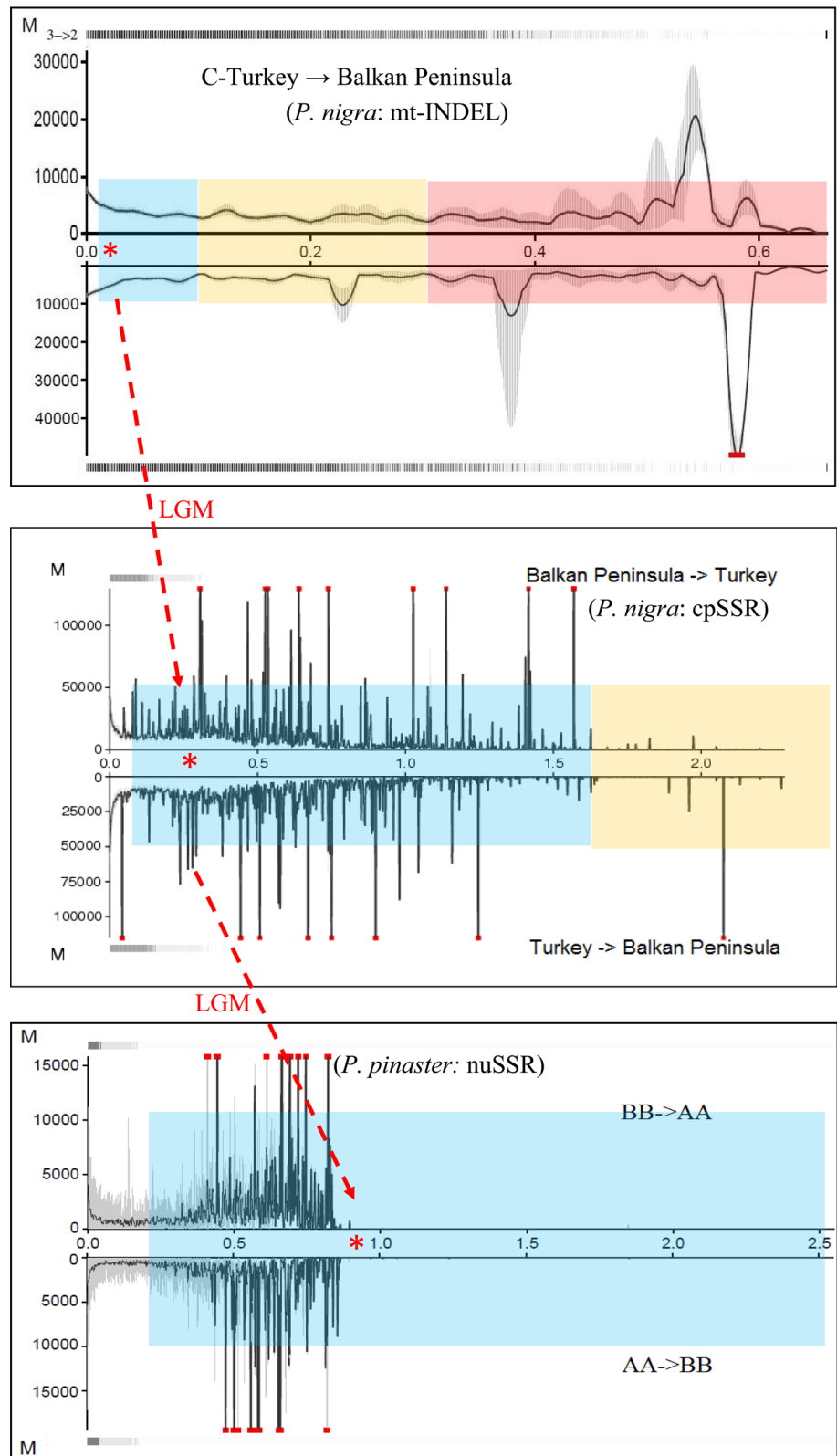
of 300 thousand years for each wave 5.5, 8.8 Ma (both in the late Miocene) and 19.5 Ma (early Miocene). A similar conclusion was reached for the historical European black pine migration route in the southern Alps from the W-Balkan Peninsula to Italy 10.8 and 6.6 Ma (both in the late Miocene: Appendix Fig. 1S/3). The historical *P. nigra* migration route

in the Southern Alps detected here was also reported in our previous publication based on 13 chloroplast *SSR* markers (Naydenov et al. 2016, 2017). For more information about paleoclimatology dynamics and forest tree migration, please see Appendix: Box-3. We speculate that early Miocene eastern migration waves from W France to the W-Balkan Peninsula were the result of the humidity attraction power of the Western Tethys and Para-Tethys Oceans. This migration tendency changed from the late Miocene to the present day due to the western migration tendency that occurred as frequent low-intensity migration waves. This scenario was the result of decreases in humidity attraction power generated from the Western Tethys and Para-Tethys Oceans in the past and the continental climatic conditions in Eastern Europe. The observed northern migration tendency in Asia Minor, i.e., from Central Turkey to northern Turkey, and in the Russian population also result from the humidity attraction power of the Black Sea currently and Para-Tethys Oceans in the past. Similar humidity forward migration with constant direction and low intensity in large geographic regions from Eastern and Southern Europe has been reported for multiple species, such as plants and animals (Petit et al. 2002; Margi et al. 2006; Sommer et al. 2008; Hewitt 2011). We speculate that this phenomenon is the result of the humidity attraction power from large sources such as oceans and seas. At the ecosystem level, animal migration is secondary and depends on the plant migration pattern.

Eastern European region: In Eastern Europe and Asia Minor, tectonic history is more complicated and is related to the collisions between the African, Euro-Asian and Arabic Plates; the history of the Western Tethys and Para-Tethys Oceans; the Caspian Depression; the Black, Pannonia and Sarmatia Seas; and orogenic history for all mountains in this region. In this region, paleobotanical artifacts are not well studied, but the region is considered to have a heterogenic topography with large variations in paleoaltitude (van Hinsbergen et al. 2020). Simulated paleomaps from different tectonic periods were presented in our previous publications on black pine phylogeographic structure inferred by chloroplast *SSR* variations for the same populations by the same authors (Naydenov et al. 2016/Fig. 4; Naydenov et al. 2017/ Fig. 3 and Appendix: Fig. 2S, 3S).

It is interesting that large historical migration waves from the Balkan Peninsula to the Crimean Peninsula for *P. nigra* have the same timing as historical *P. sylvestris* migration waves from the Balkan Peninsula to the European part of Russia, i.e., approximately 10 Ma (Appendix Fig. 1S/9 and Fig. 2S /1). In comparison with European black pine, Scotch pine fits the concept of macrolong-term refugia combined with interglacial refugia less. Naydenov et al. (2007) previously described scenarios of historical Scotch pine migration from the Balkan Peninsula as revealed by mitochondrial DNA variation, but time must be shifted to the late

Miocene (i.e., 9.5 Ma) and not the last glacial maximum (LGM: 26–8 thousand years ago). Similar suggestions were proposed by Vidiakin et al. (2012), and the present results agree that only *SNP*s from mitochondrial DNA can refer to the LGM, as discussed above. Historical northern radiation of *P. sylvestris* was the result of the disappearance of the Para-Tethys Ocean. The ancestors of this tree species were part of the mountain flora associated with deep-time paleo-topographical structure from the Atlas Mountains through the Alps, Carpathians, Balkan Peninsula, Asia Minor, Caucasia, Alborz Mountains and Hindu Kush to the Himalayan Mountain complex. The low variation in Scotch pine at *NAD7-intron1* suggests that its ancestors underwent parallel evolution in very large geographic regions. The seed migration of both *Pinus* species studied in this manuscript is very slow, probably less than 50 m per generation under favorable ecological conditions, i.e., 1.3–1.4 M years for a maximum migration of one thousand kilometers; or one million years for a migration distance of 710–765 km. A similar migration speed was calculated for *Pinus pinaster* in the western part of the Mediterranean Basin (Naydenov et al. 2014).

Speciation hypothesis

According to classical genetics, migration is opposite to speciation. High migration erodes population structure and homogenizes opposite forces. The final result will be that the general speciation process will be slowed, preventing the majority of isolation mechanisms, bottlenecks with founder effects and other processes. For both species in this study, it is very important to examine and speculate which evolution and speciation models are more appropriate using data on historical migration patterns, species biology and phylogeography and their spatial geographic structure. We believe that the species biology of currently existing *P. nigra* and *P. sylvestris* is similar to that of their ancestors from 10 to 15 Ma.

Pinus Nigra—European black pine slowly migrates by seeds, generally in the altitudinal direction in the same region and very rarely over long distances between regions. This species is best represented by Stewart et al.'s (2010) concept of macrolong-term refugia combined with interglacial refugia. Speciation as a multilevel effect at the population level from the interaction of natural selection, hitchhiking, drift and reproductive isolation is presented in principal evolution models as follows: (1) cladogenesis or (2) anagenesis (Cook 1906, 1908; Howard and Berlocher, 1998; Via 2009). We believe that both evolution within lineages and evolution by the splitting of lineages could persist in species history at different times, in different places and for different molecular markers. At the micro- and macro-evolution levels, principal speciation methods are known as phyletic gradualism and punctuated equilibrium, and the

mixed method is called punctuated gradualism. Morphogenesis and advanced large genomic analyses have shown for a large portion of organisms the presence of all evolutionary scenarios together, as described by Ricklefs (1980), Alters (1994), von Vaupel (1994), Hunt (2008), Landis and Schraiber (2017). The evolution model can be studied by in-silico simulation analysis and large-scale black pine genomic data. Additionally, the *P. nigra* results have shown that the principal geographic and ecological models of speciation, such as allopatric (vicariance), peripatric, parapatric and sympatric models, exist for each natural population in different historical periods, different sequences of mixing and different geographic regions. For *P. nigra*, it will be particularly interesting to determine the different evolutionary mechanisms of speciation between the central and peripheral populations under the different peripatric speciation model variants often reported for plants, such as centrifugal speciation, centripetal speciation, budding speciation and quantum speciation. Detailed descriptions of peripatric speciation model variants with multiple empirical examples of plant speciation worldwide have been published (Mayr 1954; Brown 1957; Grant 1971; Greenbaum et al. 1978; Futuyma 1989; Briggs 2000; Gavrilets et al. 2000; Gottlieb 2003; Howard 2003; Bowen 2010; Crawford 2010; Anacker and Strauss, 2013). The multiple historic symmetric black pine migration signals detected here support the existence of the abovementioned peripatric speciation model variants. It is well known from Schmalhausen's law (1949) that central and marginal populations have different tolerance limits, i.e., "a population at its limit of tolerance in one aspect is vulnerable to small differences in any other aspect". At the molecular level, Schmalhausen's adaptive evolution is governed by the quantity and level of deleterious mutation load. For marginal populations, the significant cost of a deleterious mutation load will restrict the number of functional loci (i.e., DNA sequences responsible for adaptive advantage) and generally will increase the selection pressure on each functional locus (i.e., locus under selection). If a forest tree is outside the physiological/biological limit of its existence, then the selection cost will cause the population to decline in fitness and become extinct. This scenario is very important for practical forest tree management and conservation recommendations for reproductive material transfer zone delimitation (which will be discussed in the next manuscript).

The multiple phenotypically different subspecies and varieties of European black pine are ideal for generally testing the many existing speciation concepts from the perspective of microevolution (Fukarek 1958a,b; Kostov 1974; Mihailov 1993, 1998; Bojovic 1995). For example, the long-term presence of phenotypic and genotypic variants in the same population is an indication of the possible presence of reproductive isolation described as the Wallace effect, which evolved later as the Bateson-Dobzhansky-Muller model in species evolution and speciation (Wallace 1889; Bateson 1909; Dobzhansky 1934, 1937; Orr 1996; Garvilets 2004). This theoretical phenomenon is essential to the existence of sexual selection, or for postzygotic isolation by epistatic gene interactions and or reinforcement [R-14]. We speculate that for *P. nigra*, selection against hybrids between subspecies and variations exists. This reinforcement process is currently reported for observed *Drosophila* hybrids in the late stage of speciation to be driven by a few genes causing reproductive isolation (Phadnis and Orr, 2009). At the population level, reinforcement will produce a disruptive selection pattern, which is a strong fixation signal (F_{st} significant); i.e., this process is bimodal for small populations or polymodal phenotypic and genotype distribution for large populations and meta-populations. According to Vidakovic (1974, 1991), Dobrinov et al. (1982), Dobrinov (1983), Alexandrov et al. (1988), European black pine has long existed in large geographic areas with more than 11–12 subspecies and more than 21–22 varieties, and some of them retain ancestral morphological traits. For conservation genetics, their presence in natural populations must be considered an evolutionarily significant unit (ESU). The existence of these taxa must occur because of reinforcement. The process must be predated by the significant accumulation of silent mutations (and/or regulatory gene modifications).

Another possible speciation mode for *P. nigra* is allopatry, where geographic isolation is essential, which is true for the *P. nigra* from northern Turkey and western Greece in our manuscript. The principal difference between sympatric speciation and allopatric speciation is the degree of spatial segregation, i.e., from zero to low landscape overlap between diverging populations. Both speciation modes seem to be more likely in areas with mountain topography for the natural distributions of *P. nigra* and *P. sylvestris*. In this region, interesting varieties living in very close areas and in areas that differ in their soil, bedrock, pH and hydrologic conditions have also been reported, which could indicate sympatric speciation and allopatric speciation. If the varieties are fixed in terms of morphological differences, then we could interpret them as an early step toward the emergence of a new subspecies. Under parapatric speciation, the level of spatial segregation is moderate to high and is probably more likely in the lowland areas of the natural distribution of Scotch pine (which will be discussed below).

The rate of speciation over geological time is under intensive scientific debate. The principal speciation rate trends are phyletic gradualism and punctuated equilibrium (Gould and Eldredge, 1977). The strong fragmentation of *P. nigra*'s natural area of distribution is a predisposition for punctuated equilibrium where the choice of mates is severely restricted. At a mutation rate of approximately 3.6×10^{-6}, the time necessary to fix one mitotype is a maximum of 139 thousand generations (9.4 Ma) for *INDEL*s in the present publication

according to the neutral theory of Kimura (1983). There is a chance of surviving genetic pressures such as population bottlenecks with founder effects, genetic drift, hitchhiking (i.e., genetic draft) and inbreeding. Population genetic structure could be fixed with new subspecies or new variants/varieties in small, isolated communities. At this point, stabilizing selection and adaptive evolution would then facilitate population growth (Schmalhausen 1949; Lieberman and Dudgeon, 1996; Levit et al. 2006). For pines, genetic events require a few thousand generations and could be different for nuclear, chloroplast and mitochondrial DNA and different locus types. If the restricted choice of mates is not too strong, then the population will express evolutionary stagnation (Eldredge and Gould, 1972), which is different from evolutionary stasis. Future large genomic datasets (WGA) for *P. nigra* will help determine where the limit between the principal models of evolution, phyletic gradualism and punctuated equilibrium and the borders between the biological (BCS) and phylogenetic (PSC) concepts of the species exist. Detailed descriptions of the BCS and PCS are provided in many scientific papers, from the modern evolutionary synthesis of species concepts in Dobzhansky (1934) and Mayr (1942) to present-day concepts (Ghiselin 1974; Mallet 2001; Pigliucci 2003). These works are known as essential parts of contemporary neo-Darwinism and post-Darwinism theory (Levit et al. 2006). It is important to know that the species definition is too multidimensional to be presented by a single concept. The difference between the two most popular species concept as PSC and BCS is the degree of reproductive compatibility. In summary, DNA evolution models, such as JC69, K80, F81, K81, HKY85, GTR, T92 and TN93, have been developed over the past half century by Jukes and Cantor (1969), Kimura (1980, 1981), Felsenstein, (1981), Hasegava et al. (1985), Tavare (1986), Tamura (1992), Tamura and Nei (1993) to explore the phylogenetic species concept (PSC). Generally, phylogenetic analysis is based on cladistics, which groups organisms based on the most recent common ancestor (not guaranty for reproductive compatibility); and is performed at the genus and subgenus levels from a very restricted sample size (Wang et al. 1999; Gernandt et al. 2005). The genetic species concept or biological species concept (BSC) defines species as a group of organisms or a population—i.e., group of reproductive compatible individuals (Mayr 1942). Genetic population analysis uses species genetically compatible and able to interbreed but are genetically isolated from other groups. Here sexually reproducing organisms exist in more or less distinct phenotypic and genetic clusters. The BSC effect of sexual reproduction on the dynamics of natural selection is essential for conservations and management of forest genetic resources than PSC [R-11].

For speciation, in comparison with those with low mutation rates, genetic variants with high mutation rates, such as *SNP*s and point mutations (i.e., single nucleotide polymorphisms), are more sensitive, and they are often negative and have small global contributions to species phylogenetic relationships and evolution. The European black pine genome is rich in chromosomal aberrations, such as translocations and rearrangements, as reported based on Giemsa banding analysis (i.e., florescent in situ hybridization-FISH), which confirms the intensive speciation process in this species (Bogunic et al. 2010). Over the last ten years, genomic analysis (WGA) techniques have evolved dramatically, and FISH will serve as an additional important tool for genomic architecture studies. We will discuss this topic in our next manuscript.

Pinus Sylvestris—The second species, Scotch pine, has fewer subspecies and variants than *P. nigra* (Vidakovic 1974, 1991; Dobrinov et al. 1982; Dobrinov 1983; Alexandrov et al. 1988; Zhelev 1992; Zhelev et al. 1994). The natural area of Scotch pine is divided into two main topographic regions: mountains and lowlands. The subspecies in lowland areas are occur in large geographic regions where reproductive isolation has a small effect, and this scenario is in contrast to that for black pine, which is generally governed by the principal geographic and ecological models of speciation, such as allopatric, peripatric, parapatric and sympatric speciation (Stenseth and Maynard, 1984). The Scotch pine model of evolution in this region appears to be a mix of contrasting punctuated equilibrium and phyletic gradualism with slow, gradual speciation (Eldredge and Gould, 1974; Eldredge 1989). The mixed variant, known as punctuated gradualism, consists of relative evolutionary stasis over a considerable amount of time combined with periodic, relatively rapid morphological change without lineage branching (Williams 1992). Mechanisms often related to stasis include stabilizing selection, Stenseth-Maynard Smith stability, normalizing clade selection, habitat tracking and constraints imposed by the nature of subdivided populations (Koeslag 1995). [R-16] Of particular interest for lowland Scotch pine is parapatric speciation and its variations, such as clinal (ecological gradient), stepping-stone (discrete populations), para-allopatric and stasipatric speciation. Parapatric speciation and its variants have principal characteristics of phenotype- and genotype-selective gradients, reproductive isolation, divergent selection, nonrandom mating, unequal gene flow and populations in continuous or less fragmented geographic and climatic ranges. Detailed descriptions of the abovementioned variations with multiple empirical examples are provided by Balkau and Feldman (1973), Caisse and Antonovics (1978), White (1978), Barton and Hewitt (1989), Bromham and Cardillo (2003), Doebeli and Dieckmann (2003), Coyne and Orr (2004), Polechova and Barton (2005). The multiple subspecies and variants indicate that the extreme variant of the phyletic gradualism model of evolution by chronospecies is probably impossible or time- and

space-restricted for *P. nigra* and *P. sylvestris*. A long-lived historically monotypic species is an example of a chronospecies, which is derived from a sequential development pattern. For example, some living fossils, such as *Gingko biloba* (*Ginkgoaceae* family), lost their known sister species, such as *G. yimaensis, G. apodes, G. adiantoides* and *G. cranei*, from the late Cretaceous to Miocene epochs (Begovic 2011). The botanical synonymous used are unispecific or monospecific species that involve continual and uniform changes from an extinct ancestral form at an evolutionary timescale.

P. sylvestris has great abilities to adapt and evolve depending on environmental circumstances. In mountainous regions, generally in southern Europe and Asia Minor, Scotch pine developed fewer subspecies and varieties than *P. nigra* (Vidakovic 1974, 1991; Dobrinov et al. 1982; Dobrinov 1983; Alexandrov et al. 1988). The evolution and speciation of mountain Scotch pine are similar to those of *P. nigra* but occurred at a slow speed and low intensity with a phenotypic effect and different sub-species extinction balance [R-13]. The combination of phenotypic richness and phylogeography is a very good indicator of the evolutionary process due to its polygenetic control and long phylogenetic memory. The results obtained with single genetic variations must be interpreted with caution, especially those from molecular markers with a low mutation range (n.b.: not low rates), such as *SNP*s and point mutations (i.e., with a high mutation frequency). Single nucleotide polymorphisms often show punctuated equilibrium and often have negative fitness and adaptation with a low chance of surviving adaptive selection to the next few generations (i.e., short phylogenetic memory). Empirical results in support of our hypothesis include the results for cancer reported by Cross et al. (2016), Notta et al. (2016), Sato et al. (2016), Graham and Sottoriva (2017). Generally, at the molecular level, *INDEL*s (as used in this study), inversions, translocations and other markers have a high mutation range and more often have positive fitness and a long phylogenetic memory (Fitch 1971; Drake et al. 1998; Papadopoulos et al. 1999; Penny et al. 2001; Ota and Penny, 2003). This hypothesis has been supported by multiple phylogenetic and phylogeographic analyses of a large number of species in the last 25 years. The molecular markers used here (i.e., *INDEL*s) are not exactly equivalent to morphological data but are very good indicators of historical migration patterns, phylogeography, speciation models and evolutionary tendencies at a large geological timescale. We speculate that the balance between the different evolution models (and speciation models) strongly depends on the molecular markers used, especially the mutation rate category (i.e., mutation range) and DNA sources at the molecular level. We will present in the next manuscript our complete and detailed concept of mutation rate variability and molecular evolution.

Additionally, the different numbers of subspecies, varieties and variants between the two species are a very good example of the theory of variable speedism from the large number of modern syntheses of evolution (Darwin 1859, 1972; Futuyma 2005; Levit et al. 2006). Variable speedism could involve discrete variable speedists and continuously variable speedists according to Dawkins (1996). This concept was very important to molecular clock theory development (Bromham and Penny, 2003). In the mountainous areas of natural Scotch pine distribution, Scotch pine numbers of subspecies and varieties are not as high as those of European black pine. The number of subspecies and varieties is an indicator of an initial step of speciation than an adaptive radiation [R-12]. A detailed analysis of which evolution model and speciation model are active in specific populations and their historical fluctuations is possible with intensive sampling efforts (i.e., a large sample size), large genomic analysis and forward- and backward in-silico simulations. In particular, forward and backward in-silico simulations could reveal the interaction between different genetic factors and their impact on evolution and speciation over large periods of time, which is currently impossible with available empirical data.

All evolution models, such as punctuated equilibrium, phyletic gradualism, stasis and punctuated equilibria, have resulted from the development of morphological evolution theory over the last two centuries, generally using paleoontological records. Evolution model definitions were initially developed based on biogeographical categories with strict geographic concepts. Later, evolution model definitions were developed with population genetic categories and concepts, including the fitness landscape, fixing, inbreeding, drift, draft and reproductive isolation. We attempted to connect the existing theoretical development of evolution and speciation to both species (*P. nigra* and *P. sylvestris*) by obtaining empirical data. Abundant empirical evidence has been accumulated over time for all forms of evolution models (Campbell 1990; Erwin and Anstey, 1995; Futuyma 2005; McCarthy and Rubridge, 2005; Benton and David, 2009; Bomfleur et al. 2014). The great discussion between scientists has been about the credibility of the fossil record, intermediate forms (known as Darwin's dilemma) and morphogenesis (Wyles et al. 1983; Bromham et al. 1998, 1999, 2002; Valentine et al. 1999; Smith and Peterson, 2002). For a more detailed description of the palynological pollen records and historical forest tree distribution and migration, see Appendix Box-3. We believe that at morphological stasis or near stasis, genetic change and evolution between one generation and the next are permanent. The accumulation of a critical quantity of silent mutation changes (and/or regulatory gene modifications) could act as a turn-on mechanism for morphological change. For this reason, molecular analysis is important for conservation genetics and biology.

Conclusions

Practical forest recommendations

The combination of mitotype diversity, effective radius of seed distribution, species biology and effective population size data allows us to propose initial minimal size of protected area for a relatively good forest tree density at reproductive age (density management diagram). However, nonfavorable soil conditions and humidity may increase this initial minimal size of protected area. The forest density changes with average ages of trees. In this case, the management of protected areas must include reforestation (natural or artificial) each 20 years with the seed from same sources. This dynamic strategy for sustainable forest genetic resource management ensures protected area with different age's class, density and succession pattern. Additional questions as interaction protected area vs management area, and anthropogenic movement of germplasm—i.e., effect of introduced provenance is very interesting and important. Here, the effective radius of pollen distribution could be a good indicator to determine the minimum physical distance to prevent a protected area from genetic contamination and will be discussed in next manuscript for the multilevel dynamic seed transfer zone [R-18].

Based on the possible speciation model discussed for both pine species, the comparative phylogeography, the comparative historical migration pattern and the large bibliography of the subspecies and varieties available at present, for forestry conservation and management biodiversity programs, we suggest following practical recommendations:

a.For distinct subspecies and varieties, it is necessary to determine their areas of natural distribution and their ecological specificity and then map these areas across large geographic regions to avoid the limitations imposed by administrative and political issues.

b.For *P. nigra*, protected areas should be encompass moderate to large areas (i.e., preferably more than 3 ha) in mixed ecosystem compositions, be established in middle and high altitudes across its local natural distribution, and include adequate protection from forest fire and illegal harvesting.

c.For *P. sylvestris*, protected areas should maintain the population, be less fragmented (i.e., preferably more than 5 ha), be established in the middle and high altitudinal range of its local natural distribution, and include adequate protection from forest fire and illegal harvesting.

d.Protected areas must be established after forest structure degradation, forest fire, recovery from illegal harvesting and insect/illness issues with cyclic reforestation (preferably) using reproductive material from the same natural origin. Using different age classes of principal forest trees in protected areas is important for ensuring natural stepping-stone succession patterns in ecosystems and the integration of all components (biotic and a-biotic).

e.For conservation and management biodiversity programs, peripheral natural populations must be better supported and monitored.

Molecular biology specificity

From the results obtained here for both pines, we confirmed the following:

(1) Historical migration inferred by mitochondrial DNA (*INDELs*) revealed the migration pattern from the Pliocene-Miocene epoch that was more closely related to local tectonic history over large geological time than to large-scale climatic change; historical migration inferred by chloroplast DNA (*SSRs*) revealed the migration pattern from the Pleistocene-Pliocene epoch that was related to tectonic history and climatic change.

(2) Historical migration waves inferred by mitochondrial DNA (*INDELs*) had a lower frequency, speed and intensity than those inferred from chloroplast DNA (*SSRs*), and the effective population sizes from migration waves were not proportional to the migration wave census population sizes.

(3) Speciation speed was different between the species *P. nigra* and *P. sylvestris*; speciation frequency at the molecular level was different among locus types and DNA types; and speciation speed depended on evolution mode, speciation mode, selection mode and species biology, among other factors;

(4) Effective population sizes inferred by mitochondrial DNA (*INDELs*) were small compared to those inferred by chloroplast DNA (*SSRs*); the time needed to fix one positive genetic change was greater for mitochondrial DNA (*INDELs*) than for chloroplast DNA (*SSRs*).

(5) Species evolution involves the mixing of different speciation modes rather than only one mode, and one speciation mode has different results/effects on different DNA types (e.g., mitochondrial vs. chloroplast vs. nuclear DNA).

(6) The misbalance between different meta-population census size contributions versus effective population size contributions for asymmetric migration patterns was the result of different genotypes (and subphylogenetic lines) responding to selection pressure and adaptive evolution.

Declarations

References

Alexandrov A, Rafailov G, Nedelin G, et al (1988) Coniferous forest in Bulgaria. Zemizdat, Sofia [in Bulgarian]

Alters BJ (1994) Modeling modes of evolution: Comparing phyletic gradualism & punctuated equilibrium. Am Biol Teach 56(6):354–360

Anacker BL, Strauss SY (2014) The geography and ecology of plant speciation: range overlap and niche divergence in sister species. Proc Royal Soc B Biol Sci 281(1778):20132980. https://doi.org/10.1098/rspb.2013.2980

Balkau BJ, Feldman MW (1973) Selection for migration modification. Genetics 74(1):171–174

Barton NH, Hewitt GM (1989) Adaptation, speciation and hybrid zones. Nature 341:497–503. https://doi.org/10.1038/341497a0

Bateson W (1909) Heredity and variation in modern lights. Darwin Modern Sci 85:101

Beerli P (2006) Comparison of Bayesian and maximum-likelihood inference of population genetic parameters. Bioinformatics 22:341–345

Beerli P (2009) How to use MIGRATE or why are Markov chain Monte Carlo programs difficult to use? In: Bertorelle G, Bruford MW, Hauffe HC et al (eds) Population genetics for animal conservation, of conservation biology. Cambridge University Press, Cambridge, pp 42–79

Begovic BM (2011) Nature's Miracle *Ginkgo biloba*, Book 1, Vols. 1–2, pp. 60–61

Belkhir K, Borsa P, Chikhi L, et al (1996–2004) GENETIX 4.05, logiciel sous Windows TM pour la génétique des populations. Laboratoire Genome, Populations, Interactions, CNRS UMR 5000, Université de Montpellier II, Montpellier, France

Benton M, David H (2009) Introduction to paleobiology and the fossil record. Wiley, New York, pp 123–124

Bhattacharyya A (1946) On a measure of divergence between two multinomial populations. Sankhya 7:401–406

Bogunic F, Yakovlev S, Muratovic E et al (2010) Different karyotype patterns among allopatric *Pinus nigra* (*Pinaceae*) populations revealed by molecular cytogenetics. Plant Biol 13(1):194–200

Bohonak AJ (2002) IBD (isolation by distance): a program for analyses of isolation by distance. J Hered 93:153–154

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

Bomfleur B, McLoughlin S, Vajda V (2014) Fossilized nuclei and chromosomes reveal 180 million years of genomic stasis in royal ferns. Science 343(6177):1376–1377. https://doi.org/10.1126/science.1249884

Bowen BW (2010) Interview with John C Briggs recipient of the 2005 Alfred Russel Wallace award. Front Biogeogr 2(3):78–80

Briggs JC (2000) Centrifugal speciation and centres of origin. J Biogeogr 27:1183–1188. https://doi.org/10.1046/j.1365-2699.2000.00459.x

Bromham L, Cardillo M (2003) Testing the link between the latitudinal gradient in species richness and rates of molecular evolution. J Evol Biol 16:200–207

Bromham L, Penny D (2003) The modern molecular clock. Nat Rev Genet 4:216–224. https://doi.org/10.1038/nrg1020

Bromham L, Rambaut A, Fortey R et al (1998) Testing the Cambrian explosion hypothesis by using a molecular dating technique. Proc Natl Acad Sci USA 95:12386–12389

Bromham L, Phillips M, Penny D (1999) Growing up with dinosaurs: molecular dates and the mammalian radiation. Trends Ecol Evol 14:113–118

Bromham L, Woolfit M, Lee MSY et al (2002) Testing the relationship between morphological and molecular rates of change along phylogenies. Evolution 56:1921–1930

Brown WL Jr (1957) Centrifugal speciation. Quart Rev Biol 32(3):247–277. https://doi.org/10.1086/401875

Burban C, Petit RJ (2003) Phylogeography of maritime pine inferred with organelle markers having contrasted inheritance. Mol Ecol 12(6):1339–1687

Caisse M, Antonovics J (1978) Evolution in closely adjacent plant populations. Heredity 40:371–384. https://doi.org/10.1038/hdy.1978.44

Campbell NA (1990) Biology. Benjamin Cummings Publishing Company, Redwood City CA

Cavalli-Sforza LL (1969) "Genetic Drift" in an Italian Population. Sci Am 221(2):30–37

Cavalli-Sforza LL, Edwards AWF (1967) Phylogenetic analysis: models and estimation procedures. Evolution 21:550–570

Chakraborty R, Jin L (1993) Determination of Relatedness between individuals using DNA fingerprinting, Human Biol 65(6): Article 2

Chaloner WG, Hallam A (1989) Evolution and extinction. Philos Trans R Soc Lond B Biol Sci 325:239–488

Chapman BB, Bronmark C, Nilsson JA et al (2011) The ecology and evolution of partial migration. Oikos 120:1764–1775

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(5):747–756. https://doi.org/10.1111/j.1471-8286.2007.01769.x

Cook OF (1906) Factors of species-formation. Science 23(587):506–507. https://doi.org/10.1126/science.23.587.506

Cook OF (1908) Evolution without isolation. Am Nat 42(503):727–731. https://doi.org/10.1086/279001

Corander J, Waldmann P, Marttinen P et al (2004) BAPS 2: enhanced possibilities for the analysis of genetic population structure. Bioinformatics 20:2363–2369

Coyne JA, Orr HA (2004) Speciation. Sinauer Associates, Massachusetts

Crawford DJ (2010) Progenitor-derivative species pairs and plant speciation. Taxon 59(5):1413–1423

Critchfield WB, Little EL (1966) Geographic distribution of the pines of the world. Forest Service Miscellaneous Publication, US, p 991

Cross WC, Graham TA, Wright NA (2016) New paradigms in clonal evolution: punctuated equilibrium in cancer. J Pathol 240(2):126–136

Dal Piaz GV (2001) History of tectonic interpretations of the Alps. J Geodyn 32:99–114

Darwin C (1859) On the origin of species, 1st edn. John Murray, London

Darwin C (1972) The origin of species by means of natural selection, 6th edn. Murray, London

Dawkins R (1996) The Blind Watchmaker. WW Norton & Company, New York

Deza MM, Deza E (2009) Encyclopedia of Distances. Springer, Berlin Heidelberg, p 590

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

Dobrinov I (1983) Genetic and selection of forest trees. Zemizdat, Sofia [in Bulgarian]

Dobzhansky T (1934) Studies on Hybrid Sterility. I. Spermatogenesis in pure and hybrid Drosophila pseudoobscura. Z Zellforsch Mikrosk Anat 21:169–221

Dobzhansky T (1937) Genetics and the origin of species. Columbia University Press, New York

Doebeli M, Dieckmann U (2003) Speciation along environmental gradients. Nature 421:259–264. https://doi.org/10.1038/nature01274

Drake J, Charlesworth B, Charlesworth D et al (1998) Rates of spontaneous mutation. Genetics 148:1667–1686

Dupanloup I, Schneider S, Excoffier L (2002) A simulated annealing approach to define genetic structure of populations. Mol Ecol 11:2571–2581

Efron B (1979) Bootstrap methods: another look at the jackknife. Ann Stat 7(1):1–26

Efron B (1982) The jackknife, the bootstrap and other resampling flans. SIAM, Philadelphia

Eldredge N (1989) Time frames. Princeton University Press

Eldredge N, Gould SL (1972) Punctuated equilibria: an alternative to phyletic gradualism. In: Schopt TJM (ed) Models in Paleobiology, Freeman, Cooper and Co. San Francisco, California, USA, pp 82–115

Eldredge N, Gould SJ (1974) Reply to Hecht. Evol Biol 7:305–306. https://doi.org/10.1007/978-1-4615-6944-2_8

Erwin DH, Anstey RL (1995) New approaches to speciation in the fossil record. Columbia University Press, New York

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

Excoffier L (2002) Human demographic history: Refining the recent African origin model. Curr Opin Genet Dev 12:675–682

Excoffier L, Smouse PE, Quattro JM (1992) Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. Genetics 1(131):479–491

Fauquette S, Suc JP, Médail F, et al (2018) The Alps: A geological, climatic and human perspective on vegetation history and modern plant diversity. Mount Climate Biodivers, 1–17

Fitch WM (1971) Rate of change of concomitantly variable codons. J Mol Evol 1:84–96

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])

Futuyma DJ (1989) Speciational trends and the role of species in macroevolution. Am Nat 134(2):318–321. https://doi.org/10.1086/284983

Futuyma DJ (2005) Evolution. Sinauer Associates, Sunderland MA

Gautier F, Clauzon G, Suc JP et al (1994) Age and duration of the Messinian salinity crisis. CR Acad Sci Paris (IIA) 318:1103–1109

Gavrilets S (2004) Fitness landscapes and the origin of species. Princeton University Press

Gavrilets S, Hai L, Vose MD (2000) Patterns of parapatric speciation. Evolution 54(4):1126–1134

Gernandt DS, López GG, García SO et al (2005) Phylogeny and classification of *Pinus*. Taxon 54:29–42. https://doi.org/10.2307/25065300

Ghiselin M (1974) A radical solution to the species problem. Syst Zool 23(4):536–544. https://doi.org/10.2307/2412471

Gokasan E, Demirbag E, Oktay FY et al (1997) On the origin of the Bosphorus. Mar Geol 140(1–2):183–199. https://doi.org/10.1016/S0025-3227(97)00022-4

Goldstein DB, Ruiz-Linares AL, Cavalli-Sforza II et al (1995) An evaluation of genetic distances for use with microsatellite loci. Genetics 139:463–471

Gottlieb LD (2003) Rethinking classic examples of recent speciation in plants. New Phytol 161:71–82. https://doi.org/10.1046/j.1469-8137.2003.00922.x

Goudet J (2001) FSTAT, a program to estimate and test gene diversities and fixation indices (version 2.9.3). Updated from Goudet (1995).

Gould SJ, Eldredge N (1977) Punctuated equilibria: the tempo and mode of evolution reconsidered. Paleobiology 3(2):115–151. https://doi.org/10.1017/s0094837300005224

Gradstein FM, Ogg JG, Smith AG (2004) A geologic time scale. Cambridge University Press, Cambridge

Graham TA, Sottoriva A (2017) Measuring cancer evolution from the genome. J Pathol 241(2):183–191

Grant V (1971) Plant speciation. Columbia University Press, New York

Greenbaum IF, Baker RJ, Ramsey PR (1978) Chromosomal evolution and the mode of speciation in three species of Peromyscus. Evolution 32(3):646–654. https://doi.org/10.1111/j.1558-5646.1978.tb04609.x

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

Hewitt GM (2011) Quaternary phylogeography: the roots of hybrid zones. Genetica 139:617–638. https://doi.org/10.1007/s10709-011-9547-3

Howard DJ (2003) Speciation: Allopatric. Wiley, Chichester. https://doi.org/10.1038/npg.els.0001748

Howard DJ, Berlocher SH (1998) Endless forms–species and speciation. Oxford University Press, Oxford, p 470

Hunt G (2008) Gradual or pulsed evolution: when should punctuational explanations be preferred? Paleobiology 34(3):360–377

Jacob F (1977) Evolution and tinkering. Science 196:1161–1166

Jiménez-Moreno G, Fauquette S, Suc JP (2008) Vegetation, climate and palaeoaltitude reconstructions of the Eastern Alps during the

Miocene based on pollen records from Austria, Central Europe. J Biogeogr 35(9):1638–1649. https://doi.org/10.1111/j.1365-2699.2008.01911.x

Kalinowski ST (2005) HP-RARE 1.0: a computer program for performing rarefaction on measures of allelic richness. Mol Ecol Notes 5:187–189

Kimura M (1983) The neutral theory of molecular evolution. Cambridge University Press, Cambridge

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

Kimura M, Weiss GH (1964) The stepping stone model of population structure and the decrease of genetic correlation with distance. Genetics 49:561–576

Knoll AH (1989) Evolution and extinction in the marine realm: some constraints imposed by phytoplankton. Philos Trans R Soc Lond B Biol Sci 325:279–290

Koeslag JH (1995) On the engine of speciation. J Theor Biol 177:401–409

Kondrashov AS, Rogozin IB (2004) Context of deletions and insertions in human coding sequences. Hum Mutat 23(2):177–185. https://doi.org/10.1002/humu.10312

Kostov K (1974) One form of *Pinus nigra* (Arn.) very resistant to insects in Bulgaria. For Manag 3:6–16

Krijgsman W (1996) A new chronology for the middle to late Miocene continental record in Spain. Earth Planet Sci Lett 142(3–4):367–380. https://doi.org/10.1016/0012-821X(96)00109-4

Landis MJ, Schraiber JG (2017) Pulsed evolution shaped modern vertebrate body sizes. PNAS 114(50):13224–13229

Latter BDH (1972) Selection in finite populations with multiple alleles. III. genetic divergence with centripetal selection and mutation. Genetics 70:475–490

Latter BDH (1973) The Island model of population differentiation: a general solution. Genetics 73:147–157

Levit GS, Hossfeld U, Olsson L (2006) From the "Modern Synthesis" to cybernetics: Ivan Ivanovich Schmalhausen (1884–1963) and his research program for a synthesis of evolutionary and developmental biology. J Exp Zool 306B:89–106. https://doi.org/10.1002/jez.b.21087

Lieberman BS, Dudgeon S (1996) An evaluation of stabilizing selection as a mechanism for stasis. Palaeogeogr Palaeoclimatol Palaeoecol 127:229–238. https://doi.org/10.1016/S0031-0182(96)00097-1

Liu K, Muse SV (2005) PowerMarker an integrated analysis environment for genetic marker analysis. Bioinformatics 21:2128–2129. https://doi.org/10.1093/bioinformatics/bti282

Lynch M, Crease TJ (1990) The analysis of population survey data on DNA sequence variation. Mol Biol Evol 7(4):377–394. https://doi.org/10.1093/oxfordjournals.molbev.a040607

Mallet J (2001) The speciation revolution. J Evol Biol 14(6):887–888. https://doi.org/10.1046/j.1420-9101.2001.00342.x

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(2):173–190

Mantel N (1967) The detection of disease clustering and a generalized regression approach. Can Res 27:209–220

Marshall CR, Raff EC, Raff RA (1994) Dollo's law and the death and resurrection of genes. PNAS USA 91(25):12283–12287. https://doi.org/10.1073/pnas.91.25.12283

Martin RA, Barnosky AD (1993) Morphological Change in Quaternary Mammals in North America. Cambridge University Press, Cambridge

Mason HL (1954) Migration and evolution in plants. Madrono 12(6):161–169

Mayr E (1942) Systematics and the origin of species from the viewpoint of a zoologist. Columbia University Press, New York

Mayr E (1954) Change of genetic environment and evolution. In: Huxley J, Hardy AC, Ford EB (eds) Evolution as a Process. Unwin Brothers, London

McCarthy T, Rubridge B (2005) The story of earth and life. Struik Publishers, Cape Town

Meirmans PG, Hedrick PW (2010) Assessing population structure: F_{ST} and related measures. Mol Ecol Res 11:5–18

Mihailov V (1998) Variability of European Black pine (Pinus nigra Arn.) according to the size, weight and form of the apophysis of the cones in Pirin and Slavianka mountains. For Sci 1(2):24–37

Mihailov V (1993) Biological and morphological study of the European Black pine's (*Pinus nigra* Arn.) seeds in different provenances and selection structure in Pirin and Slavianka mountains in Bulgaria. Ph. D. thesis, Forest Research Institute, Bulgarian Academy of Sciences, [in Bulgarian], Sofia

Miller RG (1974) The Jackknife-a review. Biometrika 61(1):1–15

Muratov MV (1964) History of the tectonic development of the Alpine folded region of southeastern Europe and Asia Minor–new data on the tectonic structure and history of development of the Alpine zone of Europe. Int Geol Rev 6(1):99–118

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

Naydenov KD, Alexandrov A, Naydenov M et al (2012) Impact of activated charcoal on germination and initial growth of some pine species. J Balkan Ecol 15(3):277–293

Naydenov KD, Alexandrov A, Matevski V et al (2014) Range-wide genetic structure of maritime pine predates the last glacial maximum: evidence from nuclear DNA. Hereditas 151:1–13. https://doi.org/10.1111/j.1601-5223.2013.00027.x

Naydenov KD, Naydenov MK, Alexandrov A et al (2016) Ancient split of major genetic lineages of European Black Pine: evidence from chloroplast DNA. Tree Genet Genomes 12(68):1–18. https://doi.org/10.1007/s11295-016-1022-y

Naydenov KD, Naydenov MK, Alexandrov A et al (2017) Ancient genetic bottleneck and Plio-Pleistocene climatic changes imprinted the phylobiogeography of European Black Pine populations. Eur J Forest Res 136(5–6):767–786. https://doi.org/10.1007/s10342-017-1069-9

Nei M (1972) Genetic distance between populations. Am Nat 106(949):283–292

Nei M (1973) Analysis of gene diversity in subdivided populations. PNAS 70(12):3321–3323

Nei M (1987) Molecular evolutionary genetics. Columbia University Press, New York

Nei M, Kumar S (2000) Molecular Evolution and Phylogenetics. Oxford University Press, New York

Nei M, Roychoudhury AK (1974) Sampling variances of heterozygosity and genetic distance. Genetics 76(2):379–390

Nei M, Tajima F, Tateno Y (1983) Accuracy of estimated phylogenetic trees from molecular data II Gene frequency data. J Mol Evol 19(2):153–170

Notta F, Chan-Seng-Yue M, Lemire M et al (2016) A renewed model of pancreatic cancer evolution based on genomic rearrangement patterns. Nature 538(7625):378

Ogurtsov AY, Sunyaev S, Kondrashov AS (2004) Indel-based evolutionary distance and mouse-human divergence. Genome Res 14(8):1610–1616. https://doi.org/10.1101/gr.2450504

Ohta T, Kimura M (1973) A model of mutation appropriate to estimate the number of electrophoretically detectable alleles in a finite population. Genet Res 22:201–204

Orr HA (1996) Dobzhansky, Bateson, and the genetics of speciation. Genetics 144:1331–1335

Ota R, Penny D (2003) Estimating changes in mutational mechanisms of evolution. J Mol Evol 57:S233–S240

Page RD (1996) TreeView: an application to display phylogenetic trees on personal computers. Comput Appl Biosci 12:357–358

Pant DD (1999) Biodiversity, conservation and evolution of plants. Curr Sci 76(1):21–23

Papadopoulos D, Schneider D, Meier-Eiss J et al (1999) Genomic evolution during a 10,000-generation experiment with bacteria. Proc Natl Acad Sci USA 96:3807–3812

Peakall R, Smouse PE (2006) GenAlEx 6 genetic analysis in excel. population genetic software for teaching and research. Mol Ecol Notes 6:288–295

Penny D, McComish BJ, Charleston MA et al (2001) Mathematical elegance with biochemical realism: the covarion model of molecular evolution. J Mol Evol 53:711–723

Petit R, Brewer S, Bordacs S et al (2002) Identification of refugia and post-glacial colonisation routes of European white oaks based on chloroplast DNA and fossil pollen evidence. For Ecol Manage 156:1–3. https://doi.org/10.1016/S0378-112

Phadnis N, Orr HA (2009) A single gene causes both male sterility and segregation distortion in drosophila hybrids. Science 323(5912):376–379. https://doi.org/10.1126/science.1163934

Pigliucci M (2003) Species as family resemblance concepts: The (dis-)solution of the species problem? BioEssays 25(6):596–602. https://doi.org/10.1002/bies.10284

Polechova J, Barton NH (2005) Speciation through competition: a critical review. Evolution 59(6):1194–1210. https://doi.org/10.1111/j.0014-3820.2005.tb01771.x

Pons O, Petit RJ (1996) Measuring and testing genetic differentiation with ordered versus unordered alleles. Genetics 144(3):1237–1245

Prevosti A, Ocara J, Alonso G (1975) Distances between populations of *Drosophila subobscura* based on chromosome arrangements frequencies. Theor Appl Genet 45:231–241

Purugganan MD (1998) The molecular evolution of development. BioEssays 20:700–711

Rambaut A, Drummond A (2009) Tracer v1.4 program- University of Edinburgh (Internet). http:// tree.bio.ed.ac.uk/ software/ tracer/

Rambaut A (2009) FigTree program-University of Edinburgh, (Internet). http:// tree.bio.ed.ac.uk/ software/ figtree/

Reynolds J, Weir BS, Cockerham CC (1983) Estimation for the coancestry coefficient Basis for a short-term genetic distance. Genetics 105:767–779

Ricklefs RE (1980) Phyletic gradualism vs. punctuated equilibrium: applicability of neontological data. Paleobiology 6(3):271–275

Rogers JS (1972) Measures of genetic similarity and genetic distance. In: Studies in Genetics VII, University of Texas Publication, Austin 7213: 145-153

Saitou N, Nei M (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol Biol Evol 4:406–425

Sanghvi LD (1953) Comparison of genetical and morphological methods for study of biological differences. Am J Phys Anthrop 11:385–404

Sato F, Saji S, Toi M (2016) Genomic tumor evolution of breast cancer. Breast Cancer 23(1):4–11

Schmalhausen II (1949) Factors of evolution: the theory of stabilizing selection. Blakiston, Philadelphia

Scotese CR, Golonka J (1992) PALEOMAP- Paleogeographic Atlas, PALEOMAP Progress Report `20, Dept. of Geology, University of Texas at Arligton

Scotese CR (2001) Atlas of earth history. Paleogeography PALEOMAP Project, Arlington, TX, Vol 1, p. 58

Shriver MD, Jin L, Boerwinkle E et al (1995) A novel measure of genetic distance for highly polymorphic tandem repeat loci. Mol Biol Evol 12:914–920

Shubin N, Tabin C, Carroll S (1997) Fossils, genes, and the evolution of animal limbs. Nature 388:639–646

Slatkin M (1991) Inbreeding coefficients and coalescence times. Genet Res 58(2):167–175

Slatkin M (1995) A measure of population subdivision based on microsatellite allele frequencies. Genetics 139(1):457–462

Smith AB, Peterson KJ (2002) Dating the time of origin of major clades: molecular clocks and the fossil record. Annu Rev Earth Planet Sci 30:65–88

Sneath PH, Sokal RR (1973) Numerical Taxonomy: the principles and practice of numerical classification, 1st edn. WH Freeman, San Francisco

Sommer R, Fahlke J, Schmölcke U et al (2008) Quaternary history of the European roe deer Capreolus capreolus. Mammal Rev 39:1–16. https://doi.org/10.1111/j.1365-2907.2008.00137.x

Stefanov B (1958) Über einige fälle warscheinlicher Umwandlung auf dem Wege der sprunghaften Veränderligkeit). Bull Int For Sofia 3:345–358 [**in Bulgarian with English and Russian summary**]

Stefanov B (1972) Further notes on the origin and taxonomy of Thracian oak, *Quercus thracica* Stef. et Ned. Forest Sci Sofia 6:3–17 [**in Bulgarian with English and Russian summary**]

Stefanov B (1973) Neueres zur Herkunft und Taxonomie der *Quercus thracica* Stef. Et Ned. Acta Bot Acad Sci Hung 19(1–4):343–357

Stefanov B, Nedjalkov S (1956) *Quercus thracica* nov. spec. - Arbor memorabilis. Bull Int Bot Sofia 5:39–53 [**in Bulgarian with English and Russian summary**]

Stenseth NC, Maynard SJ (1984) Coevolution in ecosystems: Red Queen evolution or stasis? Evolution 38(4):870–880. https://doi.org/10.2307/2408397

Stewart JR, Lister AM, Barnes I et al (2010) Refugia revisited: individualistic responses of species in space and time. Proc Biol Sci 277(1682):661–671

Takezaki N, Nei M (1996) Genetic distances and reconstruction of phylogenetic trees from microsatellite DNA. Genetics 144(1):389–399

Terrill SB, Able KP (1988) Bird migration terminology. Auk 105:205–206

Valentine J, Jablonski D, Erwin D (1999) Fossils, molecules and embryos, new perspectives on the Cambrian explosion. Development 126:851–859

Van Hinsbergen DJJ, Torsvik TH, Schmid SM et al (2020) Orogenic architecture of the Mediterranean region and kinematic reconstruction of its tectonic evolution sice the Triassic. Gondwana Res 81:79–229

Via S (2009) Natural selection in action during speciation. PNAS 106(S-1):9939–9946. https://doi.org/10.1073/pnas.0901397106

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

Vidakovic M (1974) Genetics of European Black pine (*Pinus nigra* Arn.). In: Ann For (Zagreb) 6: 57–86

Vidiakin AI, Semerikov VL, Polezhaeva MA et al (2012) Spread of mitochondrial DNA haplotypes in population of Scots pine (Pinus sylvestris L.) in northern European Russia. Russian J Genet 48(12):1267–1271

von Vaupel KJC (1994) Punctuated equilibria and phyletic gradualism: even partners can be good friends. Acta Biotheor 42(1):15–48

Wallace AR (1889) Darwinism an exposition of the theory of natural selection with some of its applications. Cosimo Inc, New York

Wang XR, Tsumura Y, Yoshimaru H et al (1999) Phylogenetic relationships of Eurasian pines (*Pinus*, Pinaceae) based on chloroplast rbcL, MATK, RPL20-RPS18 spacer, and TRNV intron sequences. Am J Bot 86:1742–1753. https://doi.org/10.2307/2656672

White MJD (1978) Modes of Speciation. WH Freeman and Company, San Francisco

Williams GC (1992) Natural selection: domains, levels and challenges. Oxford University Press, NY

Winker K (2000) Migration and speciation. Nature 404:36

Wright S (1931) Evolution in Mendelian populations. Genetics 16:97–159

Wright S (1932) The roles of mutation, inbreeding, crossbreeding, and selection in evolution. Proc Sixth Int Congr Genet 1(8):355–366

Wright S (1943) Isolation by distance. Genetics 28(2):114–138

Wright S (1951) The genetical structure of population. Ann Eugen 15:323–354

Wright S (1965) The interpretation of population structure by F-statistics with special regard to systems of mating. Evolution 19(3):395–420

Wyles JS, Kunkel JG, Wilson AC (1983) Birds, behavior, and anatomical evolution. Proc Natl Acad Sci USA 80:4394–4397

Yeh FC, Yang RC, Boyle TBJ, et al., (1997) POPGENE, the user-friendly shareware for population genetic analysis. Molecular Biology and Biotechnology Center, University of Alberta, Edmonton, Alberta, Canada

Zhelev P, Longauer R, Ladislav P et al (1994) Genetic variation of the indigenous Scots Pine (*Pinus sylvestris* L.) populations from the Rhodope Mountains. For Sci (bulgaria) 3:68–76 [**in Bulgarian**]

Zhelev P (1992) Ecolo-biological and selection—genetic Scots Pine investigations from the Rhodope Mountains. Ph.D. Thesis of forest genetic, selection and tree breeding. University of Forestry, Sofia, Bulgaria [**in Bulgarian**]

Authors and Affiliations

Krassimir D. Naydenov[1] · Michel K. Naydenov[2] · Alexander Alexandrov[3] · Todor Gurov[4] · Veselka Gyuleva[3] · Georgi Hinkov[3] · Sofiya Ivanovska[4] · Anatoly Tsarev[5] · Biljana Nikolic[6] · Venceslas Goudiaby[7] · Christopher Carcaillet[8] · Roman Volosyanchuk[9] · Srdjan Bojovic[10] · Kole Vasilevski[11] · Vlado Matevski[12] · Lorenzo Peruzzi[13] · Andreas Christou[14] · Despina Paitaridou[15] · Irina Goia[16] · Salim Kamary[17] · Suleyman Gulcu[18] · Cengiz Ture[19] · Faruk Bogunic[20]

Michel K. Naydenov
michel92kn@gmail.com

Alexander Alexandrov
Alexandrov_38@abv.bg

Todor Gurov
gurov@bas.bg; gurov@parallel.bas.bg

Veselka Gyuleva
gjuleva33v@abv.bg

Georgi Hinkov
georgihi@abv.bg

Sofiya Ivanovska
sofia@parallel.bas.bg

Anatoly Tsarev
antsa_55@yahoo.com

Biljana Nikolic
smikitis@gmail.com

Venceslas Goudiaby
venceslas.goudiaby@gmail.com;
venceslas.goudiaby@uqat.ca; venceslas.goudiaby@uqam.ca

Christopher Carcaillet
christopher.carcaillet@ephe.sorbonne.fr

Roman Volosyanchuk
volosyanchuk@yahoo.com

Srdjan Bojovic
bojovic@ibiss.bg.ac.rs

Kole Vasilevski
kvasilevski@ukim.edu.mk

Vlado Matevski
Vlado.Matevski@yahoo.com

Lorenzo Peruzzi
lorenzo.peruzzi@unipi.it

Andreas Christou
achristou@fd.moa.gov.cy

Despina Paitaridou
d.paitaridou@prv.ypeka.gr

Irina Goia
irina.goia@ubbcluj.ro

Salim Kamary
kamarisalim@yahoo.fr

Suleyman Gulcu
suleymangulcu@sdu.edu.tr

Cengiz Ture
cture@anadolu.edu.tr

Faruk Bogunic
faruk.bogunic@gmail.com

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

[2] UN Secretariat-UNOHR | DMSPC/GSPD/STMS/ODS, Headquarters, 405 E 42Nd St., New York, NY 10017, USA

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

[4] Institute of Information and Communication Technologies, Bulgarian Academy of Sciences, Acad. G. Bonchev Str., Bl. 25-A, 1113 Sofia, Bulgaria

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

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

[7] Department of Biological Sciences, University of Quebec at Montreal, 141 Président-Kennedy, Montréal, QC H2X 1Y4, Canada

[8] UMR5023 CNRS-Université Lyon 1, 43 Bd 11 Novembre 1918, 69622 Villeurbanne Cedex, France

[9] Ukrainian Research Institute of Forestry and Forest Amelioration, 86 Pushkinska St., Kharkiv 61024, Ukraine

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

[11] Ss. Cyril and Methodius University, Blvd. Edvard Kardelj Bb, Skopje 1000, Republic of Macedonia

[12] Faculty of Natural Sciences and Mathematics, Institute of Biology, Skopje 1000, Republic of Macedonia

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

[14] Departament of Forestry, 26 Loukis Akritas avenue, 1414 Nicosia, Cyprus

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

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

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

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

[19] Department of Biology, Science Faculty, Anadolu University, Biyoloji Bolumu, 26470 Eskisehir, Turkey

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