CrossMark

Ancient genetic bottleneck and Plio-Pleistocene climatic changes imprinted the phylobiogeography of European Black Pine populations

Krassimir D. Naydenov[1] · Michel K. Naydenov[2] · Alexander Alexandrov[3] ·
Kole Vasilevski[4] · Georgi Hinkov[3] · Vlado Matevski[5] · Biljana Nikolic[6] ·
Venceslas Goudiaby[7] · Dave Riegert[8] · Despina Paitaridou[9] · Andreas Christou[10] ·
Irina Goia[11] · Christopher Carcaillet[12,13] · Adrian Escudero Alcantara[14] ·
Cengiz Ture[15] · Suleyman Gulcu[16] · Veselka Gyuleva[3] · Srdjan Bojovic[17] ·
Lorenzo Peruzzi[18] · Salim Kamary[19] · Anatoly Tsarev[20] · Faruk Bogunic[21]

Abstract The historical changes in European Black Pine population size across the whole natural distribution in Europe and Asia Minor were analyzed facing the Plio-Pleistocene climatic fluctuations. Thirteen chloroplast SSRs and SNPs markers have been studied under the assumptions of "*neutral evolution.*" Populations and meta-populations had different histories of migration routes, and they were strongly affected by complex patterns of isolation, fragmentation, speciation, expansion (1.88–4.28 Ma), purification selection (2.09–21.41 Ma) and bottleneck (1.85–21.76 Ma). A significant number of populations (min. 29–41%) were in equilibrium for very long periods. Generally, the bottleneck revealed by chloroplast DNA is weaker than the bottleneck revealed by nuclear DNA. The N_e immediately after the bottleneck reaches between 1820 and 3640 individuals. Generally, the historical effective population sizes shrink significantly for the Tertiary period from 10–15 up to 2.5 Ma in Western Europe (by 82%), followed by Asia Minor (69%) and the Balkan Peninsula (28%), likely resulting from important climatic changes. The rates and frequencies of stepwise westwards migration waves have been not sufficient to prevent isolation between the meta-populations and to suppress "*sympatric speciation.*" The migration was weak for the Pliocene, but was maximal for the Pleistocene, and finally *silent* for the present interglacial period, namely the Holocene.

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

Michel K. Naydenov
Michel92KN@hotmail.com

Alexander Alexandrov
forestin@bas.bg; Alexandrov_38@abv.bg

Kole Vasilevski
kvasilevski@ukim.edu.mk

Georgi Hinkov
georgihi@abv.bg

Vlado Matevski
vladom@iunona.pmf.ukim.edu.mk

Biljana Nikolic
smikitis@gmail.com

Venceslas Goudiaby
venceslas.goudiaby@gmail.com

Dave Riegert
driegert@gmail.com; david.riegert@queensu.ca

Despina Paitaridou
d.paitaridou@prv.ypeka.gr

Andreas Christou
achristou@fd.moa.gov.cy

Irina Goia
igoia@yahoo.com; irina.goia@ubbcluj.ro

Christopher Carcaillet
christopher.carcaillet@ephe.sorbonne.fr

Adrian Escudero Alcantara
a.escudero@escet.urjc.es

Springer

Introduction

One of the main population genetics research goals of the last 50 years is to determine the role of demography, migration and selection on the evolution of organisms (Hartl and Clark 1997). The estimation of genetic and demographic properties is important for species from isolated populations and long-time fragmented natural areas of distribution. From the "*microevolution*" theory, isolated populations are expected to show low genetic variability, low gene flow and low ability to cope with environmental changes due to drift and inbreeding (Chetverikov 1961; Dobzhansky 1982). This is particularly important for tree populations, due to their fundamental roles in the ecosystem under climatic change and anthropogenic pressure (Kamari et al. 2010; Linares and Tiscar 2010; Bijlsma and Loeschcke 2012; Naydenov et al. 2014, 2015).

Available data have shown contrasting results from the "*microevolution*" theory for multiple organisms, especially in the natural distribution of forest tree populations (Lowe et al. 2005; Petit and Hampe 2006; Rafii and Dodd 2007). Trees are long-lived plant species, show large ecological adaptation capacity and apparent resilience to disturbance and finally show multiple generations co-occurring on single and isolated sites. The latter feature helps to delay the loss of genetic diversity (Bacles and Jump 2010; Davies et al. 2010). The high environmental heterogeneity in very large areas of natural distribution and repetitive disturbances, especially fire, is considered as essential for positive selective pressures and strong long-term local adaptation (Lowe 2005; Naydenov et al. 2006, 2012; Savolainen et al. 2007; Kramer et al. 2008). Persistent speciation might attenuate the loss of genetic diversity and extinction rates (Crow and Morton 1955; Eyre-Walker et al. 2002). It is important to mention the role of migration as a factor for a marked change in allele frequencies. The gene migration forces act in opposition to speciation by attenuating the ongoing genetic differentiation among different isolated populations and, thus, prevent the evolution process of new taxa.

Effective population size (N_e) is one of the basic population genetics parameters (Wright 1938, 1942, 1969), and it determines the rate of change in the composition of a population, i.e., the level of variability in a population and the effectiveness of selection caused by genetic drift in a population. A limited effective population size may cause (1) an alteration of frequency distribution of alleles across different sites and (2) genetic hitchhiking and background selection (Kimura 1983; Charlesworth 2009). For this reason, a population bottleneck, selective sweep, expansion, equilibrium or background selection must be investigated with special ad hoc models and not by simply documenting a reduction or increase in N_e (Braverman

Cengiz Ture
cture@anadolu.edu.tr

Suleyman Gulcu
suleymangulcu@sdu.edu.tr

Veselka Gyuleva
gjuleva33v@abv.bg

Srdjan Bojovic
bojovic@ibiss.bg.ac.rs; biosbojos@gmail.com

Lorenzo Peruzzi
lperuzzi@biologia.unipi.it; lorenzo.peruzzi@unipi.it

Salim Kamary
kamarisalim@yahoo.fr

Anatoly Tsarev
tsarev@karelia.ru; tsarev@psu.karelia.ru;
antsa_55@yahoo.com

Faruk Bogunic
faruk_bogunic@yahoo.com

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

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

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

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

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

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

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

[8]   Department of Mathematics and Statistics, Queen's University, Jeffery Hall, University Ave., Kingston, ON K7L 3N6, Canada

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

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

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

[12] Ecole Pratique des Hautes Etudes PSLResearch University Paris, Paris, France

[13] LEHNA (UMR5023) CNRS, Université Lyon 1, 43 Bd 11 Novembre 1918, 69622 Villeurbanne Cedex, France

et al. 1995; Simonsen et al. 1995; Gordo et al. 2002). The effective population size varies across the genome, from organelles (i.e., plastid: chloroplast, mitochondria, ribosomes, etc.) and chromosomes, as a result of the *"recombination"* versus *"non-recombination"* pattern.

Information about historical effective population size and immigration pattern is crucial to better establish the effective forest genetic resource management and protection, but also to predict the impact of climate change, reforestation and effect of domestication of forest tree species. Three main genetic lineages (i.e., meta-populations) of *P. nigra*, from Europe and Asia Minor have been analyzed (Naydenov et al. 2016). These analyses highlighted the existence of three meta-populations. The meta-populations are called, based on geography, (1) the western Mediterranean (AA), (2) the Balkan Peninsula (BB) and (3) the Asia Minor—Turkey (CC). However, no information on the effective size historical fluctuations was provided. Conservation protocols have to pay special attention on the migration and on the N_e fluctuation for long geological period of climatic changes, as also described by Hofmann et al. (2014), Frantz et al. (2014), Bueno et al. (2016), Thesing et al. (2016) and Yuan et al. (2016). If a response of European Black Pine to climatic changes is expected, this species, however, appears rather resilient to fires based on paleoecological and dendrochronological studies (Christopoulou et al. 2013; Leys et al. 2014).

Accordingly, we hypothesize that the different spatial and temporal responses at the meta-population level played an important role in effective population size fluctuation of European Black Pine, i.e., the N_e fluctuated in the context of the *"interglacial refugia"* (20,000 years of warm and wet geological period) versus *"macro-long-term refugia"* (over multiglacial/interglacial cycle, i.e., 100,000 years). To explore some elements of the population's demography as a historical effective population size fluctuation under the assumptions of *"neutral evolution,"* we generated new break through results, which provided new information about bottleneck, selective sweep, purification selection, expansion and equilibrium and the impact of the migration rates (gene flow) between the meta-populations (i.e., formations). Totally 106 natural populations were sampled, with and an average of 30 individuals per population (Naydenov et al. 2016). The aim of our study is to provide answers to the four following questions: (a) What was the N_e in populations after bottleneck, expansion, positive selection and equilibrium? (b) When were significant historical effective population size fluctuations observed? (c) Are there N_e-based differences between meta-populations? (d) What is the migration pattern between main meta-populations and how it contributed to shape the demographic and phylogeographic structure?

Materials and methods

The full description of sample collection from 106 natural Black Pine populations (3154 individuals), population geographic data, DNA extraction, laboratory manipulations and statistical methods used are presented in Naydenov et al. (2016). To sum up, the sampled area covers the entire range of the *Pinus nigra*, from 5.1°N to 48.1°N latitude (from Morocco to Austria) and 4.7°W to 39.1°E longitude (from Morocco to Turkey). The data are from specific hotspot regions of chloroplast genome as described by Vendramin et al. (1996). Sequence variation in 13 regions was investigated: seven SSRs (single sequence repeats, i.e., size variations) and six flanking regions with single nucleotide polymorphisms (SNPs); for more information, please see the DATA-S (Electronic supplementary material).

Here, we we tested whether the historical effective population size varied through time and across the populations and meta-populations from the western Mediterranean (AA), the Balkan Peninsula (BB) and Asia Minor—Turkey (CC). The meta-populations were determined by the admixture model of spatial population genetic analyses (BYM algorithm; Fig. 1). To avoid confusion, the terms *"meta-population," "formation"* or *"admixture cluster"* were used hereafter as synonyms in sensu lato (*s.l.*). There is a great deal of discussion regarding their definitions, which is beyond the scope of the current study (Table 1).

Historical migration rate (M)

We obtained coalescent-based estimates of maximum posterior values of historical effective population size and migration rate between the European Black Pine

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

[15] Department of Biology, Ecology Program, Science Faculty, Anadolu University, 26470 Eskisehir, Turkey

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

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

[18] Dipartimento di Biologia, Università di Pisa, 56126 Pisa, Italy

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

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

[21] Faculty of Forestry, University of Sarajevo, 20 Zagrebacka, Sarajevo, Bosnia and Herzegovina
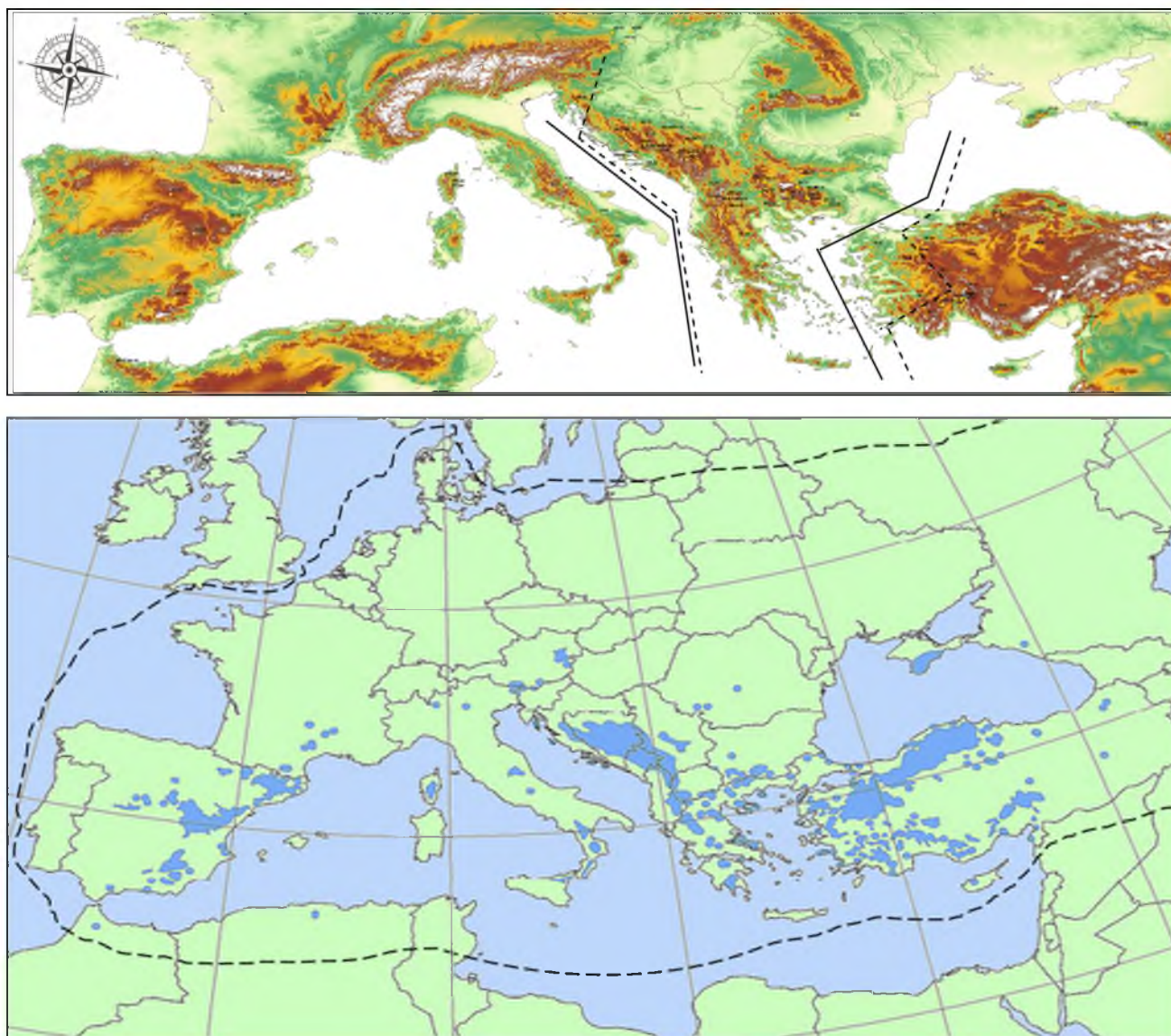
Fig. 1 The studied natural populations of *Pinus nigra* (Arn.) and the principal edges of genetic diversity determined by the BARRIER program (*dashed line*) and the model-based method of STRUCTURE/TESS programs (*solid line*) according Naydenov et al. (2016) on the top. The present natural distribution of the European Black Pine (*blue polygons*) according EUFORGEN; the *dashed line* is the probable limit of the area of distribution of the ancestors of the European Black Pine (*Pinus thomasiana* Goepp. and Reich; and *Pinus laricioides* Menzel) in the Tertiary epoch (i.e., Eocene, Oligocene, Miocene and Pliocene) according the paleobotanical artefacts (Stefanov 1941/1942, 1943; Gorbunov 1958; Dijkstra 1973; Palamarev 1989; Arslan et al. 2012; Ehrendorfer 2013). (Color figure online)

formations by using Bayesian and Maximum likelihood inferences from MIGRATE-n v.3.0 program (Beerli 2006). The time scaling is in "*time units,*" i.e., in units of mutation rate scaled to the generation (1 unit ≈ 1.2 million years [hereafter Ma]; number of generations = $1/\mu$). To discard any misinterpretation, we used the MIGRATE algorithm using the coalescent theory of Kingman (1982a, b), but the time scaling in output files of the program was in the "*time unit,*" which is independent of effective population size, and the "*time unit*" was different from the "*coalescent unit.*" One "*coalescent unit*" is equal to $1N_e$ for haploid and $2N_e$ for diploid DNA in generations. In the present manuscript, we have only used "*time unit*" as defined by (Beerli 2006). The MIGRATE program was used for the chloroplast DNA data for a large number of species (Edh et al. 2007; Clark and Carbone 2008; Bai et al. 2010; Wang et al. 2011; Naydenov et al. 2014).

Preliminary tests were performed to adjust the method of analysis to the different groups of populations: AA-1 to AA-5; BB-center, northern, southern, eastern and western; CC-1, 2, 3, center, northern, southern, eastern, western (Table 2; see Naydenov et al. 2016). The initial test with

Table 1 Summary of historical effective size meta-population fluctuation from 105 natural populations and 3154 individuals determined by chloroplast DNA polymorphism of *Pinus nigra* (Arn.)

Meta-pop.	Average time		%	Time (Ma)	N_e*	Test	Program
	Gen* (SD)	Ma/SD (Av. ± SD) duration		min/max [duration]			
Pure bottleneck model (26 pops/936 individuals)							
AA	50.67 (23.9)	3.45/1.6 (1.85/5.05) 3.20	21	2.10/5.79 [3.69]	2.73	1	3
BB	233.25 (87.4)	15.86/5.9 (9.96/21.76) 11.80	32	1.27/26.51 [25.24]	3.64	1	3
CC	92.81 (32.8)	6.31/2.2 (4.11/8.51) 4.40	15	4.32/9.76 [5.44]	1.82	1	3
Purification selection (29 pops/1044 individuals)							
AA	57.19 (26.8)	3.89/1.8 (2.09/5.69) 3.60	42	1.78/7.90 [6.12]	4.54	1	3
BB	251.62 (63.2)	17.11/4.3 (12.81/21.41) 8.60	21	11.0/27.31 [16.31]	3.64	1	3
CC	98.03 (20.5)	6.67/1.4 (5.27/8.07) 2.80	29	3.94/8.79 [4.85]	2.73	1	3
Permanent population genetic equilibrium (37 pops/1332 individuals)							
AA			32				
BB			41				
CC			29				
Pure expansion model (18 pops/648 individuals)							
AA	36.41 (9.5)	2.48/0.6 (1.88/3.08) 1.20	10	2.02/2.93 [0.91]	0.004/119[†]	2	4
BB	44.00 (8.2)	2.99/0.6 (2.39/3.59) 1.20	4	2.60/3.39 [0.79]	0.027/663[†]	2	4
CC	49.74 (13.0)	3.38/0.9 (2.48/4.28) 1.80	41	1.86/4.89 [3.03]	1.800/484[†]	2	4
Mixed expansion model (16 pops/576 individuals)							
AA	54.78 (10.4)	3.72/0.7 (3.02/4.42) 1.40	10	3.02/4.43 [1.41]	0.027/189[†]	2	4
BB	48.16 (8.7)	3.27/0.6 (2.67/3.87) 1.20	17	2.41/4.33 [1.92]	0.018/328[†]	2	4
CC	43.57 (8.6)	2.96/0.6 (2.36/3.56) 1.20	15	2.38/4.01 [1.63]	2.400/754[†]	2	4

† is N_{e-o}/N_{e-1}, i.e., N_e before/after expansion; * $x10^3$ from Tables 1S and 2S; T is time of occurrence of the bottleneck measured in units of $2N_e$ generations, i.e., Gen = $2TN_e$ at N_e are historical effective meta-population sizes 23,437; 93,750 and 40,178 individuals, respectively, for AA, BB and CC formations; the historical effective meta-population sizes are from n-MIGRATE program (Beerli 2006), from Tables 3S, 4S and 5S; Tau (τ) is time of most recent range expansion—$\tau = 2\mu t$, t is the time in generations (Gen = $\tau/2\mu$ at $\mu = 5.6 \times 10^{-5}$); the number of years for one generation is 68 from Naydenov et al. (2016); [1]G–T test (Griffiths and Tavare 1994a, b); [2]Tajima's D/Fu's Fs tests (Tajima's 1989; Fu 1997); [3]SweepBOOT v.1 (Galtier et al. 2000); [4]ARLEQUIN v. 3.1 (Excoffier et al. 2002) and DNAsp v. 5.10.1 (Rozas et al. 2003)

700 cores of high-performance computing (HPC) system suggested the use of a dataset with no more 90–100 individuals for each scenario (i.e., for two admixing edges AB and BC). This calibration is important to determine the reasonable time-consuming calculation versus stable results (at SE ≤ 5% for 20 independent replicates). Similar samples sizes, in agreement with the MIGRATE-n v.3.0 program manual recommendations, have been used for chloroplasts and mitochondrial DNA in other plant (see above) and animal case studies (Carreras et al. 2007; Juste et al. 2009; and Lin et al. 2009). The MIGRATE-n v.3.0 program (Beerli 2006) was used in more than 500 similar studies for large numbers of organisms (please see the program credits from http://popgen.sc.fsu.edu/Migrate/Info.html).

After that, we used two sets of data (AA/BB and BB/CC). Each set of data has two representative populations of 36 individuals for a total of 72 individuals. To avoid imbalance between sample size of formations and to examine the most recent migration patterns, we used the most frequent haplotypes with the minimum frequency of

5% for the simulation, using initial "*theta*" and migration parameters derived from F_{st}-like calculations, as described by Beerli and Felsenstein (1999), a free-to-vary custom migration model without predetermined geographic distances, 5×10^5 burn-in interactions and 5×10^4 sample outputs, one long chain (long chain ε = infinity) and four heating temperatures (between 1 and 50), and sum-Gelman convergence for the MCMC search strategy.

The haplotype choice for the MIGRATE program is very important and depends on the objective of analysis. There are two principal strategies: (a) to use the most frequent haplotypes; and (b) to use haplotypes related to the MRCA. According to the type of basic data available and study aims, we selected the former strategy. The use of the most frequent haplotypes allows to: (1) complete analysis with N_e historical fluctuations at population and meta-population levels of large number of "*neutral tests*" as G–T, R–H and Tajima's D/Fu's Fs tests (all method assuming "*neutral mutation*"); (2) DNA fragments can be moderately long (<5 kb); (3) obtained results of historical migration patterns are, for recent times, related to the

Table 2 Summary of the historical effective size population (HPD-10%) and the historical effective size meta-population (HPD-50%) fluctuation from 71 natural populations and 2338 individuals determined by nuclear DNA (microsatellites) polymorphism of four Mediterranean *Pinus* species

Treats	Mediterranean population of *Pinus* species			
	nigra (9 pops/324 ind.)	*halepensis* (23 pops/828 ind.)	*sylvestris* (12 pops/432 ind.)	*pinaster* (27 pops/972 ind.)
Bottleneck end/start in generations BP; [Ma], i.e., $t_f N_e$ = generations				
HPD-10%	6146/9426 [0.418/0.641]	7346/10,743 [0.500/0.730]	343/358 [0.023/0.024]	98/102 [0.007/0.007]
HPD-50%	2751/26,658 [0.187/1.813]	3612/25,135 [0.246/1.709]	317/387 [0.021/0.026]	90/111 [0.006/0.007]
$\ln(t_f)$; $[t_f]$; max/min				
HPD-10%	2.19/2.62 [8.93/13.74]	4.74/5.12 [114.43/167.33]	0.16/0.20 [1.17/1.22]	0.12/0.16 [1.13/1.17]
HPD-50%	1.39/3.66 [4.01/38.86]	4.03/5.97 [56.26/391.51]	0.08/0.28 [1.08/1.32]	0.04/0.25 [1.04/1.28]
Bottleneck duration in generations; [My]				
HPD-10%	3280 [0.223]	3397 [0.231]	15 [0.001]	4 [0.0003]
HPD-50%	23,907 [1.626]	21,523 [1.463]	70 [0.005]	21 [0.001]
Present effective population size N_e ($\approx N_1$)				
	686.0	64.2	293.2	86.8
Ancestral population size no max/min; [% of N_1], i.e., $r N_e = N_o$				
HPD-10%	39,767/24,607 [1.72/2.79]	22,971/15,709 [0.28/0.41]	9807/7485 [2.99/3.92]	622/568 [13.95/15.28]
HPD-50%	125,600/9810 [0.55/6.99]	53,210/7495 [0.12/0.86]	24,365/4451 [1.20/6.59]	799/475 [10.86/18.27]
$\ln(r)$; $[r]$; max/min				
HPD-10%	−4.06/−3.58 [−57.97/−35.87]	−5.88/−5.50 [−357.81/−244.69]	−3.51/−3.24 [−33.45/−25.53]	−1.97/−1.88 [−7.17/−6.55]
HPD-50%	−5.21/−2.66 [−183.09/−14.30]	−6.72/−4.76 [−828.82/−116.75]	−4.42/−2.72 [−83.10/−15.18]	−2.22/−1.70 [−9.21/−5.47]

$t_f = t_d/N_o$; $r = N_o/N_1$; N_1—present effective population size; r—rate of effective population change; t_f—time in generation; N_o—ancestral effective population size; the number of years for one generation is 68 from Naydenov et al. (2016); the used data are from Kamari et al. (2010) and Naydenov et al. (2011, 2014, 2015)

results from G–T and Tajima's D/Fu's Fs tests; (4) the sample size is more close to reality—i.e., mixing of haplotypes is close to observed population structure.

MIGRATE implements a likelihood analysis with Metropolis algorithm using random coalescence to calculate genetic parameters. The Metropolis Monte Carlo sampling technique allows the concentration of sampling in regions that contribute to the final result. For the Markov chain Monte Carlo (MCMC) search strategy approach in the maximum likelihood (ML) method, the "*start parameters*" must be very close to the maximum likelihood values. For this reason, it is necessary to run multiple short chains. The result of the last chain was used as the "*starting value*." The Bayesian and maximum likelihood inferences from MIGRATE have produced similar results. The service of CLUMEQ (Compute Canada-McGill University, a nationwide network of HPC installations; http://support.clumeq.ca) was used for computing support. To avoid confusion, for all interpretation in the present publication, we used the result of MIGRATE-Bayesian inference.

Historical effective population size fluctuation (N_e)

The effective population size fluctuations are the result of two principal events: demographic and selective, i.e., interaction of population with its environment. The demographic events include expansion, bottleneck and founder effects, and the selective events include positive, neutral, purification and directional selection (*selective sweep*). The more popular methods for studying the N_e fluctuation for variants of a nucleotide sequence in a population are: (1) the test *D*-statistics from Tajima's D/Fu's Fs test (Tajima's 1989; Fu 1997), (2) the methods based on pairwise differences of Rogers and Harpending (1992) and Harpending (1994) and (3) the method to compute likelihoods under the coalescent model of G–T test (Griffiths and Tavare 1994a, b). The first two methods are more adapted for populations under the assumptions of constant population size and non-migration (or low migration). The third method is better for panmictic populations.

Historically important changes in population size were studied by Tajima's D/Fu's Fs test and the Rogers and Harpending test (i.e., the Harpending's Raggedness index, H-R), using the ARLEQUIN v. 3.1 (Excoffier et al. 2002) and DNAsp v. 5.10.1 (Rozas et al. 2003) software. The test computed the difference between the mean number of pairwise differences (*Theta-Pi*) and the number of segregated sites (*Theta-k*), i.e., the test determines the difference between DNA sequences evolving randomly (i.e., *neutrally*), or under a non-random process. We used Fu's simulation from Fu's Fs test (similar to Tajima's D) because it was a more sensitive indicator of population
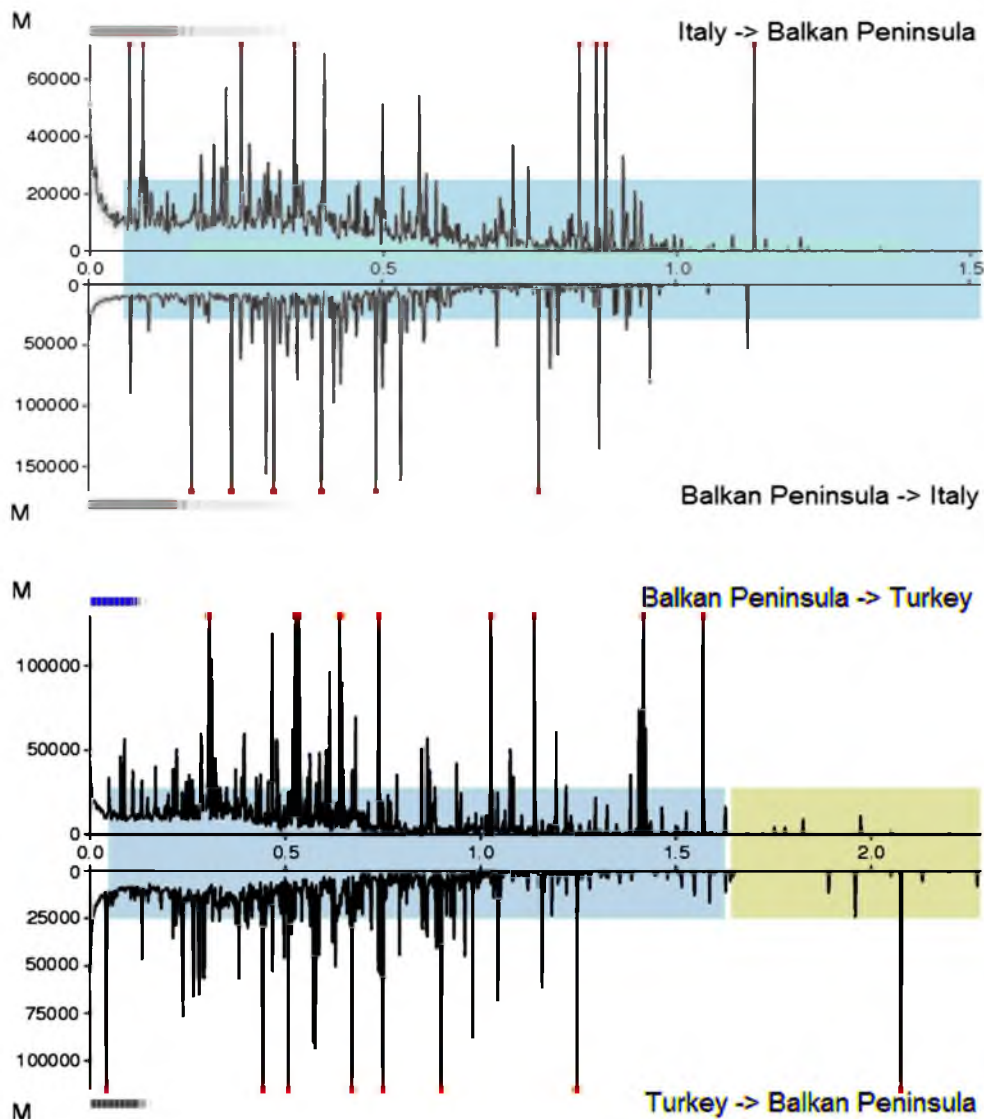
Fig. 2 The migration flow (*M*) in the number of individuals (*vertical axes*) from coalescence simulation of the two principal edges of "*admixture.*" The admixture edges are reported from Naydenov et al. (2016). The simulations are performed using Bayesian inference from n-MIGRATE v.3.0 program (Beerli 2006) as described in the methods. The *gray lines* indicate the frequency of events, the *black lines* are the rate of migration flow back in time, and the *red dots* suggest that both the upper quintile and the main value were higher (i. e., more sure events). Time is scaled by mutation rate per generation (*horizontal axes*), and one "*times unit*" is ≈1.2 Ma, i.e., 17,857 generations at 68 years by generation and μ = 5.6 × 10⁻⁵ (number of generations = 1/μ). The approximate time of the Holocene epoch (0–0.1 Ma), Pleistocene epoch, i.e., ice age (0.1–1.8 Ma), and Pliocene epoch (1.8–5.3 Ma) is shown by *white*, *blue* and *yellow colors* on the *horizontal axis*, respectively. *Note*: the calculated from n-MIGRATE v.3.0 historical meta-population effective size (N_e) for AA (western Mediterranean formation), BB (Balkan Peninsulas formation) and CC (Asia Minor formation) is 23,437; 93,750; and 40,178, respectively (n.b.: Watterson estimator Theta = $2N_e\mu$ for haploid data; and $\Theta = 4N_e\mu$ for diploid data). (Color figure online)

expansion and genetic hitchhiking (Fu 1997). According to these authors: (1) a population evolving as per mutation-drif "*equilibrium with no-evidence of selection*" shows $D = 0$ with $p \leq 0.05$; (2) "*balancing selection*" or "*recent sudden population contraction*" shows $D > 0$ with $p \leq 0.05$; (3) "*recent selective sweep linked to a swept gene*" or "*population expansion after recent bottleneck*" shows $D < 0$ with $p \leq 0.05$. In case of $D \neq 0$, the use of an additional "*neutral test*" as G–T (i.e., Griffiths and Tavare 1994a, b) is recommendable to discern among "*balancing selection*" and "*recent sudden population contraction*" (both for $D > 0$ with $p \leq 0.05$); and among "*recent selective sweep linked to a swept gene*" and "*population expansion after recent bottleneck*" (both for $D < 0$ with $p \leq 0.05$). The number of years per generation was estimated 68 (Naydenov et al. 2016), and it was used in the calculation of occurrence time for the demographic fluctuation of European Black Pine. The Fu's Fs test has been widely used with chloroplast

DNA data (Liu et al. 2013; Gao et al. 2015; Du and Wang 2016).

To determine how "*recent bottlenecks*" or "*selective sweep*" (maximum $10N_e$ back in time, i.e., maximum 6.8 Ma if $N_e = 10,000$ and 68 years is one generation) occurred in the populations' demographic history according to their chloroplast DNA sequences for the Fu's Fs test $D \neq 0$ ($p \leq 0.05$), we used the maximum likelihood method implemented in SweepBOOT v.1 (Galtier et al. 2000), which excludes sequences violating the assumptions of the "*infinite site model*," under the hypotheses of neutrality and constant population size, with 10^4 interactions. The algorithm used to identify apparent bottlenecks at specific loci is described in Griffiths and Tavare (1994a, b). The prior population size (*theta*), time of occurrence of the bottleneck (*T*) and coalescence pressure (*S*, strength of bottleneck) was set at 0.1/10.0 units (for mono-parental DNA one unit of *T* is equal to N_e). The same algorithm was used for chloroplast data by Taylor and Keller (2007) and Brooks (2012). All N_e historical fluctuation methods (i.e., Tajima's D/Fu's Fs; R–H; and G–T tests) and historical migration pattern from MIGRATE program assumed *neutral mutation/selection model* from the Neutral Theory of Molecular Evolution (Kimura 1983). The correlations of the parameter from the five sub-groups of demographic fluctuations from Tables 1S–5S versus population longitude and latitude data are according to the method of Naydenov and Alexandrov (1999) and Naydenov et al. (2002).

To avoid confusion related with the name of geological epochs, especially for Pleistocene and Pliocene, we used terminology provided by Gradstein et al. (2004) and accepted by the International Commission on Stratigraphy as International Stratigraphic Chart-2008 (http://www.stra tigraphy.org/ICSchart/StratChart2008.pdf). In the present manuscript, we have set the base of the Pleistocene (Quaternary) at 2.6 Ma. The traditionally used Tertiary comprises the Pliocene (2.6–5.3 Ma), the Miocene (5.3–23.0 Ma) and the Paleogene (23.0–65.5 Ma).

Results

Historical migration rate (M)

The average N_e values are 23,437, 93,750 and 40,178 individuals for formations AA (the western Mediterranean), BB (the Balkan Peninsula) and CC (Asia Minor—Turkey), respectively (at SE $\leq$ 5% for 20 independent replicates). The variability of historical effective population size inside of the 3 meta-populations is very low (data not shown). The maximum population size in the past appears to have been in the admixture cluster BB, followed by the CC and AA formations. The average historical effective population sizes are for the period from minimum of the time of the detected migration signal from same region (i.e., 1.7–2.5 Ma, early Pleistocene). The historical waves of migrations represented in Fig. 2 show the recent migration pattern.

The migration flow and "*theta*" (Watterson estimator "*theta*" $= 2N_e\mu$ for haploid DNA) values obtained here indicate that past genetic flow via pollen was not simultaneous and equal between meta-populations from the AA, BB and CC regions. The most intensive gene exchange was between the Balkan Peninsula (BB) and Asia Minor (CC) populations, i.e., across the BC edge that corresponds to the Aegean and Marmara Seas. Nine major migration waves involving more than 150 thousand individuals were identified from the BB to CC admixture clusters (Fig. 2, bottom), while migration in the opposite direction appears to have been less intensive, with eight major migration waves involving 120 thousand individuals. Evidence of multiple, more restricted migration waves involving approximately 25–50 thousand individuals was also obtained, likely with a less pronounced contribution to the genetic pool. One migration wave (i.e., one pic) shown in Fig. 2 bottom represents migration events for a minimum of 12,000 years. The signals of the historic waves of migration between formations BB and CC maximum are 2.5 Ma Plio-Pleistocene transition.

The migration pattern between formations in the region west of the Adriatic and Ionian Sea basins (AA) and Balkan Peninsula (BB) admixture clusters, i.e., across edge AB, was less balanced. Six major waves of migration occurred, involving a minimum of 175 thousand individuals and multiple smaller waves involving 75 thousand individuals from BB to AA. In the opposite direction (from AA to BB), the migration flow was weaker, as eight waves occurred, involving a maximum of 75 thousand individuals and multiple, very small waves of 10 thousand individuals. One migration wave (i.e., one pic) shown in Fig. 2 (top) represents migration events for a minimum of 8000 years. The signal of historic migration through the AB edge dates to a maximum of 1.7 Ma, early Pleistocene (Fig. 2, top). The oldest and youngest major waves of effective pollen transfer were from Asia Minor (CC) to the Balkan Peninsula (BB).

Historical effective population size fluctuation (N_e)

G–T TEST. The results of historic demographic population fluctuations based on sequences that do not violate "*infinite site model*" assumptions indicate the presence of a bottleneck that generally occurred a long time ago (≥ 0.5 units) of a $2N_e$ generation over the natural distribution area of the European Black Pine (Tables 1S–2S from Appendix,

Electronic supplementary material). The method of Griffiths and Tavare (1994a, b) showed diversity reduction in 55 populations ($p \leq 5\%$). The Shapiro and Wilk (1965) test of normal distributions for the same population for the historical effective population size immediately after the bottleneck (N_{e-1}) showed the deviation of Gaussian distribution. We tested different data transformations according to Naydenov and Alexandrov (1999, 2000), but the normal distribution was impossible to reach. The normal probability plots (Q–Q plot) and histograms showed a *polymodal distribution* for any data transformation (data not shown). This is an indication that results represent different demographic scenarios for the diversity reduction of N_{e-1} in 55 populations. Therefore, we decided to split the data in two sub-groups according to the logarithm of maximum likelihood probability (*L*-ratio); i.e., $0.01 < p \leq 0.05$ and $p \leq 0.01$ (Tables 1S–2S from Appendix, Electronic supplementary material).

Pure bottleneck group

The high-probability sub-group with $p \leq 0.01$ showed normal distribution for biological data after removing the population AT-03 (*S-W*-statistic = 0.909 at $p < 0.05$). This sub-group is named the *pure bottleneck group* with average $N_{e-1} = 3200$ (SD = 1.484). Bottlenecks were detected, with statistical support *L*-ratio ($p \leq 0.01$), in 24.5% of the populations (i.e., 26 populations; Table 1S from Appendix, Electronic supplementary material), generally a *long time ago* (according to the time limit in the SweepBOOT software). Evidence of a *strong bottleneck* was found for 5 populations, with no indications of a specific geographic location, such as Iberian Peninsula, Balkan Peninsula or Asia Minor (Fig. 1S; Table 1S from Appendix, Electronic supplementary material). In this sub-group, the past reductions in effective population size were not limited to specific parts of the natural distribution area of *P. nigra*. Diversity reduction was observed in 4, 5 and 17 populations from the western Mediterranean, Asia Minor and the Balkan Peninsula formations, respectively (Table 1S). Average observed bottlenecks were weak (i.e., $S < 8.0$) and less pronounced for CC ($S_{av} = 0.759$), BB ($S_{av} = 0.569$) and AA ($S_{av} = 0.336$) formations; please see Table 1. The time (min/max) for the bottleneck signal at the meta-population level was different: 2.10/5.79 Ma (Pliocene) for Western Europe, 1.27/26.51 Ma (from late Oligocene to early Pleistocene) for the Balkan Peninsula and 4.32/9.76 Ma (late Miocene to early Pliocene) for Asia Minor (Table 1). The average times for N_e reduction are 3.45 Ma (SD = 1.6), 15.86 Ma (SD = 5.9) and 6.31 Ma (SD = 2.2), respectively. The bottleneck was observed in 21, 32 and 15% of the natural populations from AA, BB and CC meta-populations, respectively, and the historical effective

population size immediately after the bottleneck was 2730 (SD = 1482), 3640 (SD = 1518) and 1820 (SD = 814) individuals, respectively (Table 1).

Purifying selection (s.l.)

The second sub-group (low-probability sub-group) with $0.01 < p \leq 0.05$ did not fit well with the bottleneck model of Griffiths and Tavare (1994a, b) implanted in Sweep-BOOT v.1 program and probably represented the natural populations in the final step of *purifying selection*, i.e., very strong stabilizing selection (*s.l.*; Table 2S). The purifying selection is one kind of stabilizing selection by the selective purging of deleterious variations. In this sub-group, there were 29 natural populations from all geographic regions (Fig. 1S; Table 2S from Appendix, Electronic supplementary material). The N_{e-1} data showed a *polymodal distribution* for any data transformation, which indicates the complex and different demographic scenarios in the studied populations (data not shown). For this reason, the obtained data must be interpreted with precautions. The *purifying selection* was observed in 42, 21 and 29% of the populations from the regions of the western Mediterranean (AA), the Balkan Peninsula (BB) and Asia Minor—Turkey (CC), respectively (Table 1). The time of occurrence min/max differed between the meta-populations as follows: 1.78/7.90 Ma (late Miocene to early Pleistocene), 11.0/27.31 Ma (late Oligocene up to middle Miocene) and 3.94/8.79 Ma (late Miocene up to middle Pliocene), respectively; please see Table 1.

Fu's Fs TEST: Fu's simulations showed negative values (Fs; $D < 0.0$), implying that *Theta-Pi* was less than *Theta-k*, i.e., fewer haplotypes than numbers of segregating sites (Tables 3S–4S). The relative low frequency of polymorphisms to the expected polymorphisms indicates population size expansion after bottleneck and/or purifying selection. To avoid speculative interpretation, we applied the same strategy as for the *pure bottleneck* and *purifying selection* analysis. For this reason, we decided to split the results into two sub-groups, according the Harpending's Raggedness index (H-R) probability, i.e., $0.01 < p \leq 0.05$ and $p \leq 0.01$ (Tables 3S–4S). The high-probability sub-group was named the *pure expansion* model, and the low-probability sub-group was named the *mixed expansion* model (*s.l.*), i.e., expansion with very strong positive selection.

Pure expansion

The analysis of the "*pure expansion*" model showed that 18 populations evolved under non-random processes generally present in 41% of Asia Minor (CC), fewer than 10% in western Mediterranean (AA), and less 4% in Balkan

Peninsula (BB) formations. Detailed results at the population level are in Fig. 1S and Table 3S. The N_{e-1} data showed the *polymodal distribution* for any data transformation, i.e., the data varied significantly from the expected pattern if the data were drawn from a group (population stat. meaning) with a normal distribution (data not shown). The time of occurrence (min/max) was very similar between all meta-populations: 2–5 Ma (Table 1). The level of maximum historical effective population size after expansion was very different in each region: N_{e-1} was 119 thousand individuals in Western Europe, 484 thousand individuals for Asia Minor and 663 thousand individuals for the Balkan Peninsula (Table 3S).

Mixed expansion (s.l.)

The *mixed expansion* model sub-group for 16 populations was observed in 10% from AA, 17% from BB and 15% from CC meta-populations (Fig. 1S; Table 4S from Appendix, Electronic supplementary material). The data do not fit with a normal distribution (data not shown). The occurrence was between 2 and 4 Ma (Pliocene to early Pleistocene); N_{e-1} was 189, 328 and 754 thousand individuals for the western Mediterranean (AA), the Balkan Peninsula (BB) and the Asia Minor (CC) meta-populations, respectively (Table 1).

Equilibrium (s.l.)

The populations that did not belong to the above-mentioned four sub-groups of demographic fluctuations ("*pure bottleneck,*" "*purifying selection,*" "*pure expansion*" and "*mixed expansion*") have been interpreted as populations in "*equilibrium*" (*s.l.*; Table 5S). The gene frequencies reached an equilibrium between mutation and selection pressures in this sub-group and was logically between the reduction and expansion of historical effective population size. For the population in *equilibrium*, the historical N_{e-1} fluctuations were within sustainable range for a long period (Table 1). The obtained data from N_{e-1} showed strong deviations of normal distribution (data not shown). The populations number from the *equilibrium* sub-group fluctuated through time and was a minimum of 32, 41 and 29% of populations from principal formations, respectively (Tables 1, 5S).

The Mann–Whitney U statistic (i.e., rank sum test; Mann and Whitney 1947) has been applied to compare the five sub-groups of demographic fluctuations. The difference in the median values between some sub-groups is not relevant enough to exclude the possibility that the difference is due to random sampling variability ($p > 0.05$). On the contrary, differences between groups 1 ("*pure bottleneck*" and "*purifying selection*") and 2 ("*pure expansion,*"

"*mixed expansion*" and "*equilibrium*"), i.e., the "*reduction events*" versus "*expansion events,*" are statistically significant (at $p < 0.05$). The correlations of the parameter from the five sub-groups of demographic fluctuations from Tables 1S–5S versus population longitude and latitude data did not show specific patterns, i.e., cline variations were not observed (data not shown).

Discussion

Historical migration rate (M)

The historical wave of migration through the AB edge (i.e., between the Balkan and the Apennine Peninsula) suggests that there was a non-equal exchange of gene, with a strong tendency for stepwise westward migration during the Pleistocene period (see Fig. 2, top). Migration waves from the Balkan to the Apennine Peninsula (Italy) occurred approximately twice over the last 1.75 Ma (Pleistocene). Migration toward the BC edge appears to have been balanced over the last 2.5 Ma (Pleistocene) and has occurred in multiple waves. The origin of this result is (1) a relatively short period of contact zone and (2) relatively restricted areas. This conclusion applies strictly to *P. nigra*, but other trees with similar biology distribution range may have been subjected to a similar history. No references were found for trees describing the relationship between the migration pattern, TMRCA and the most recent split. This question is important to establish a management strategy to protect biodiversity respectful from the population genetic dynamics in natural populations.

The oldest major wave of migration identified by coalescent analysis originated from Asia Minor toward the Balkan Peninsula (i.e., edge BC; Fig. 2, bottom). The obtained migration history and strong signal of range expansion from Balkan Peninsula support the hypothesis of a westward migration for some trees from the Mediterranean region (Mirov 1967; Stebbins 1974; Petit et al. 2005) and possible eastern origin for the ancestors of European Black Pine and some Neogene Mediterranean pines (Vidakovic 1991). The documentation of similar patterns of westwards migration of plants in the Mediterranean area was reported for *Fagus sylvatica* and *F. orientalis* (Comps et al. 2001), *Quercus ilex* (Lumaret et al. 2002), *Androcymbium gramineum* (Caujape-Castells and Jansen 2003), *Frangula alnus* (Hampe et al. 2003) and *Pinus halepensis* (Fady-Welterlen 2005). The strong signal of migration pattern from Asia Minor (Turkey) to the Balkan Peninsula was not observed for the Pleistocene and late Pliocene epoch (Fig. 2, bottom).

The present data showed that migration between principal formations occurred in multiple waves (and not as a

constant flow) and that it occurred in both directions, from west to east and vice versa. Only the accumulated differences in the *migrant* number over a significant time period (combined with natural selection and drift) created the present phylogeographic structure. The migration pattern obtained here is from the Holocene and the Pleistocene (max 2.6 Ma; Fig. 2). More intensive migration patterns were inferred from the Pleistocene, which is rich in multiple glacial–interglacial cycles lasting one million years, with large temperature fluctuations (Fig. 3S, top). The 0–0.1 Ma signal of European Black Pine migration is silent for both edges of admixing (Fig. 2). Chloroplast DNA evidences shown a period of significant rarefaction for Black Pine likely started in the last 100,000 years (0.1 Ma), the upper Pleistocene that includes the Late Glacial Maximum (25,000–18,000 years) where forests and tree populations tremendously suffered from spatial contraction in glacial refugia. However, this period further contains the birth of agriculture during the Neolithic (since 12,000 years depending on regions in the Mediterranean basin), which was an important cause of deforestation through pastoralism and eventually fires, and of tree transplantation during modern times (Flannery 1994; Brown 1997; Lee and Daly 1999).

For plants, which pollen and seeds are dispersed by wind, the effective radius of distribution of seed and pollen has significant differences and depends on multiple factors that can vary greatly among populations. Many conifer trees have effective distribution of 100 m maximum for seed and 1 km maximum for pollen per generation (Gonzalez-Martınez et al. 2002; Chybicki and Dzialuk 2014; Heredia et al. 2015). For this reason, a *"contact zone"* should be interpreted for the long- and short-distance gene transfer and separated and non-separated (i.e., close contact) populations, respectively. For the same gene flow, a close contact needs three–fourfold less area.

Migration waves occurred during a period of 8000–12,000 years; for each period, the migrant number was between 25 and 175 thousand individuals. The minor waves of migrations occurred more often (Fig. 2). The major migration waves represent 1.0–1.5 thousand individuals per generation of 68 years. If we speculate that a *contact zone* was 300 km^2, then the migrant number was between 3 and 5 individuals per km^2 per generation. This is similar to *P. nigra* results obtained over the last 30 years based on enzymes and DNA (Goncharenko and Silin 1997; Cengel et al. 2012; Rubio-Moraga et al. 2012; Bonavita et al. 2015). The recently published data regarding *Pinus pinaster* migration across the genetic barriers of the Straits of Gibraltar and the Pyrenees Mountains showed similar results (Fig. 2; Naydenov et al. 2014). Note that one *time unit* for nuclear DNA is approximately 120,000 years and for chloroplast DNA is approximately 1.2 Ma due to the

different mutation rates (i.e., μ_{nu} vs μ_{ch}; $\approx 5.6 \times 10^{-4}$ vs 5.6×10^{-5}).

This calculation allows us to conclude that the intensive migration between the principal meta-populations was non-permanent, short (less of 5% of scanning period) and in a restricted area of 100–300 km^2. The multiple minor waves of migrations are likely the result of long-distance gene transfer or very restricted area of close contact, such as 10–30 km^2. Multiple migration waves between the main formations did not seem to have significant impact on the TMRCA, the more recent split, and continuous *"sympatric speciation."* This speciation is very slow, rich in intermediate phenotypes and incorporates isolation with migration model, in which the principal meta-populations diverge despite that some gene flow continued. The obtained data of TMRCA and the more recent split from Naydenov et al. (2016) are the result of more intensive migration in the past, followed by very long isolation, i.e., the most recent split appeared after a long period of isolation. In Fig. 2, the isolation for the edge BC is confirmed by a small migration score in the 1.5–2.5 Ma period (late Pliocene up to early Pleistocene). We can speculate that this isolation was longer, but we cannot document older patterns of migration (>2.5 Ma), because of methodological constraints. It is important to determine that 8–9 *contact zones* of reduced area (100–300 km^2 for a period of 8000–12,000 years) were not sufficient to compensate for the isolation effect between the Black Pine meta-populations and continuing intensive sympatric in situ speciation at each geographic location. Generally, the TMRCA should be less than the divergence time of the meta-populations (Beerli 2006, 2009).

Pinus nigra and historical climate change in Europe

Starting from the Eocene epoch, the average global temperature had started to progressively decrease, according to Scotese and Golonka (1992) and Scotese (2001). This likely caused a decline in occurrence extent of the Black Pine ancestors. One of the first palaeobotanic studies in Europe to confirm the presence of *P. nigra* ancestors in the Neogene over a large area outside of the Mediterranean region is that by Mirov (1967). The presence of this species outside of the Mediterranean region probably declined as a result of its inability to follow climactic fluctuations during the period of humidity and temperature changes. The simulations of *forest communities space fluctuation* (Eder et al. 2008) based on multiple palaeobotanical remains in Europe in the last 15 Ma (Miocene) confirm this scenario, which supports by the present study the existence of a long and ancient bottleneck. The palaeobotanical remains showed larger areas of distribution of ancestors of European Black Pine (Fig. 1, bottom) in the Tertiary epoch, i.e.,

from the Eocene to the Pliocene (Stefanov 1941/1942, 1943; Palamarev 1989; Arslan et al. 2012). Starting from this period, the ancestors of *P. nigra* had lost competition in this part in Europe, which can explain the large reduction of the historical effective population size in meta-populations from Western Europe, and the relatively long bottleneck and purification selection over its modern area of distribution (Gorbunov 1958; Dijkstra 1973; Ehrendorfer 2013).

European Black Pine ancestors relatively quickly abandoned large areas of Western Europe from the present region of Denmark. The quick migration reached the area limit of the southern part of Western Europe, likely close to central France. These fragmented populations had small N_e. We speculate that, since this period, the meta-population AA has kept smaller effective population size. The European Black Pine ancestors that progressively moved southwards may be in relationships to changes in climate. The most adapted soil moisture regimes for *P. nigra*, as probably for its ancestors, would be semiarid climate or under cool winters and dry–warm summers (Vidakovic 1991; Mihailov 1993). The southern migration of European Black Pine was the result of humid pressing in a double front for western Europe (Atlantic and Northern climatic components) at narrow geographic regions with high speed for short period about 10 Ma, middle Miocene (i.e., 4–5 Ma before the signal of N_e fluctuations from same region) and significant loss of gene pool. Then, the *P. nigra* ancestors would remain in the southwestern Europe generally in *"interglacial refugia."* During the last one million years, the earth climate had experienced between 10 and 12 glacial periods (Fig. 3S, top) according Milankovitch (1920, 1930, 1941) and Richmond and Fullerton (1986), in which *P. nigra* populations were subjected to expansions of their distribution area during the glacial periods, and contractions during the interglacial.

In Central and Eastern Europe, the humid ecological conditions were generally in a single front from the northern climatic component, at a large geographic region with slow speed (less than average) and moderate loss of gene pool, for a long period about 25–30 Ma, Oligocene (i. e., 4–5 Ma before the signal of N_e fluctuations from same region). In the main territory from Central and Eastern Europe, loss of area was very slow, and the Black Pine ancestors had close concentrations toward the present Balkan Peninsula and northern and eastern Black Sea regions. Starting from this period, the region was under significant topographic and tectonic transformation related to the Tethys and Parathethys Seas and Balkan–Pontides microplates (Fig. 3; and Figs. 2S–3S from Appendix, Electronic supplementary material). The disappearance of the *"Pannonian Sea"* and the development of European mountain complex under favorable climatic transformation

likely allowed a larger effective population size, and larger variability, supporting the idea of an eastern origin for this species (Naydenov et al. 2016). During the Pleistocene, multiple glaciations were responsible for the transformation of large regions of Southeastern Europe: the Pannonian, Dacia, Northern Black, Caspian, Aral and Azov Seas watershed in wetlands (Fig. 3S). That region is well known as the *"Sarmatian sea"* from the last glacial depression (Belousov et al. 1988; Popov et al. 2004; Jipa and Olariu 2009) The repetition of glaciations fragmented a significant portion of the European Black Pine natural population area in that region since 1 Ma. Today, *Pinus nigra* persisted only in well-drained and dry spots as the Crimea Peninsula, eastern Black Sea coasts and the Carpathians. The wetlands regions did not reach Balkan Peninsulas southern of Danube River and Asia Minor (Turkey). The *P. nigra* ancestors' region from Turkey and the Balkan Peninsulas remained better protected from climatic fluctuations and cataclysms as *"macro-long-term refugia."* Differences between past scenarios in the Western and Central–Eastern Europe since 20 Ma likely imprinted the *P. nigra* phylogeographic structure. These phylogeographic evidences match the results for other plants and animals (Finch et al. 2009; Hofmann et al. 2014; Frantz et al. 2014; Bueno et al. 2016; Yuan et al. 2016).

The Quaternary and late Tertiary climatic fluctuations paralleled the tectonic history of the Alps Mountain complex. The raising of Alpine orogeny in Europe stopped before 2.5–7.0 million years, after non-constant altitude fluctuations during the Paleogene and Neogene periods (2.5–65.0 Ma; Moores and Fairbridge 1998; Poulsen et al. 2015). The first topographic on setting of low-elevation Alpine mountain complexes dates back to the Rupelian age (approximately 28.1–33.9 Ma; HAGRC 2009; Champagnac et al. 2009). Consequently, the climate became more humid over Europe, especially the Mediterranean region and Western Europe (SGN 2012; Campani et al. 2012).

The Plio-Pleistocene climatic fluctuations interacted with the frequency and intensity of forest fire. The disturbance of forest fire had a major impact on the pine's phylogeographic structure, morphology and physiology (Naydenov et al. 2006, 2012; He et al. 2012). Although *Pinus nigra* is less adapted to fire than other thermophilous Mediterranean pines, such as *P. halepensis* and *P. pinaster*, it appears rather resilient based on paleoecological and dendrochronological studies (Christopoulou et al. 2013; Leys et al. 2014). It is possible that forest fires played special and different roles in the context of the *"interglacial refugia"* versus *"macro-long-term refugia"* by changing biogeochemical cycles and biotic interaction. This impact lasted for long periods and probably had an impact on ancestral forms, subspecies and varieties of European Black Pine. The significant population shrink observed
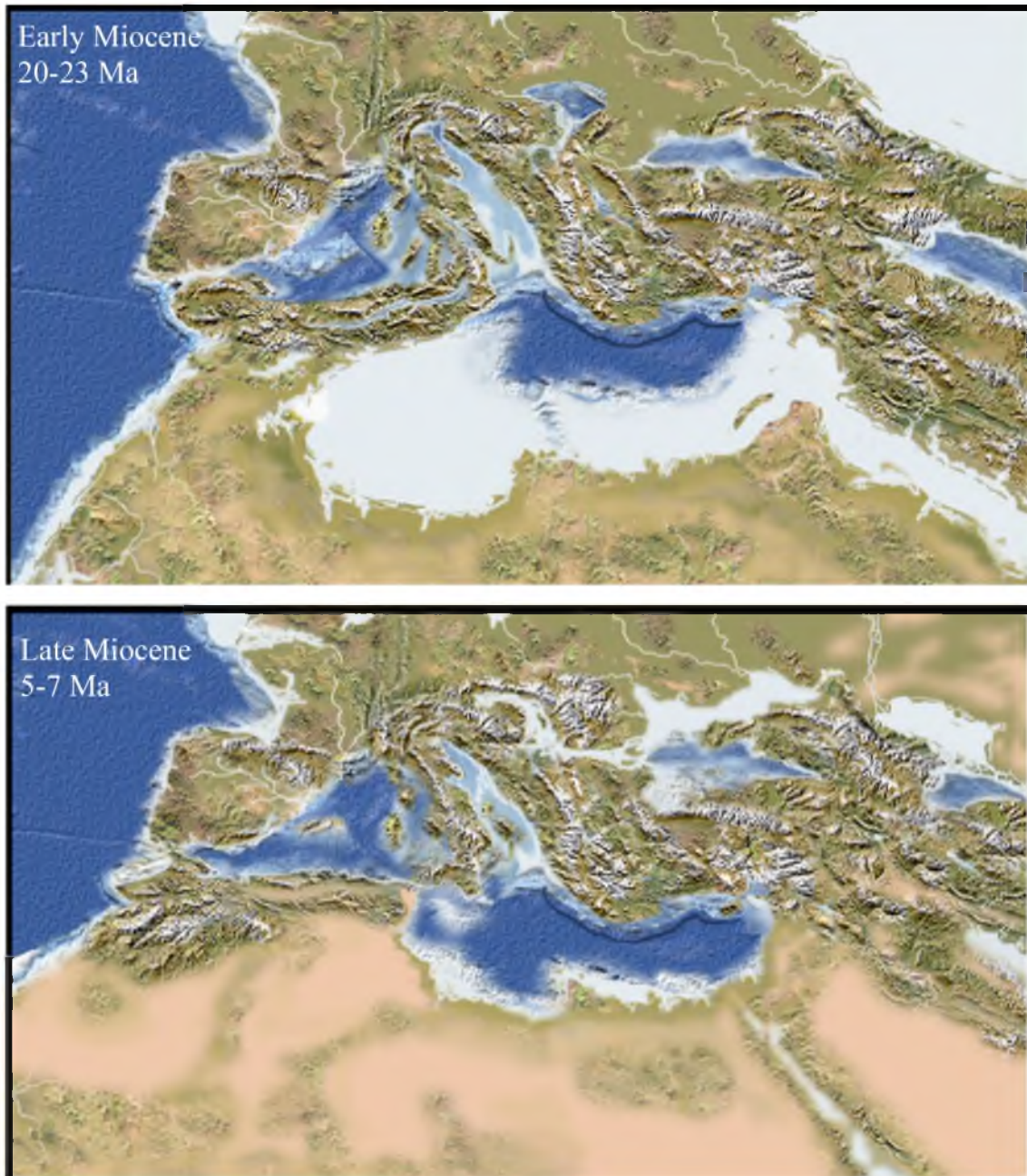
Fig. 3 Tectonic development in the studied region for the Miocene epoch. The general plate data are from Scotese et al. (1998) and the PALEOMAP Project, Department of Geology, University of Texas at Arlington, Arlington, Texas. The paleogeographic, sedimentologic and tectonic data are compiled from many sources, but are especially influenced by Ziegler (1988), Geary et al. (1989) and Sengor and Natalin (1996)

from 10 to 25 Ma is the result of: (a) climatic changes from dry to more humid and cool climates, (b) the change of forest fire disturbance sequences and (c) changes in biotic communities. The combination of forest fire and forest succession patterns are probably responsible for the different effective population size historical fluctuation sequence patterns observed from Asia Minor (Turkey) versus the remaining areas of *P. nigra*. More details regarding N_e fluctuations are discussed in the following paragraph.

Further progresses in the study of the complex relationship between the Quaternary and late Tertiary climatic change, tectonic dynamic, ancestor's structure and the obtained historical genetic structure of the European Black Pine may come from population genetic simulation tools, which allow to validate inferences analysis from empirical studies.

Historical effective population size fluctuation (N_e)

At the meta-population level, the N_e for the last 1.7–2.5 Ma is 23,437; 93,750 and 40,178 individuals for formations AA (western Mediterranean), BB (Balkan Peninsula) and CC (Asia Minor), respectively (at SE $\leq$ 5% for 20 independent replicates). The N_e is relatively stable for this period in the limit of each formation and confirms similar observations (Wakeley and Aliacar 2001; Woolfit and Bromham 2005; Charlesworth and Eyre-Walker 2007; Charlesworth 2009). For the period 10–15 Ma (middle Miocene), N_e was approximately 250 thousand individuals (Table 3; Naydenov et al. 2016). These N_e differences likely result from a long bottleneck and purification selection through a complex of biotic and abiotic factors following the global climatic change and topographic/tectonic transformation in Europe and the Mediterranean basin. Generally, the chloroplast effective population sizes shrunk significantly by 82% in Western Europe, by 69% in Asia Minor, and by 28% in the Balkan Peninsula. Black Pine lost vast portions of its range during the last 100,000 years (measured value), i.e., during the Holocene probably linked to human activities.

Fluctuations in the effective population size of the three large *P. nigra* formations (AA, BB and CC) show different patterns. For the western European meta-population (AA), the demographic fluctuation revealed by chloroplast DNA reveals that purification selection had started approximately 7.9 Ma, and progressively reached approximately 42% of all populations at 5.7 Ma (late Miocene; Table 1). Significant presence for this kind of selection was at 3.6 Ma, started to decline at 3.9 Ma and ended at 2.1 Ma (early Pleistocene). In the final step of this period of accelerating presence of purification selection, some populations started to experiment bottleneck (5.8 Ma, late Miocene). The bottleneck reached maximum presence in 21% of the pine area in this region at 5.1 Ma (early Pliocene) and continued for a period of 3.2 Ma (Table 1). The bottleneck started to decline at 2.1 Ma and finished relatively quickly at 1.9 Ma. The purification selection and bottleneck acted together for long time. In the end of this period, very slow expansion mixed with strong positive selection at 4.43 Ma and reached its maximum shortly at 4.42 Ma for 10% of all populations (early Pliocene; Table 1). This slow expansion continued for a period of 1.4 Ma and declined shortly 3.0 Ma (late Pliocene). Almost immediately after this period, the populations from same geographic area (western European meta-population) experimented pure expansion of the N_e, which started 2.9–3.1 Ma in 10% of populations (late Pliocene). The pure expansion had started to decline 2.0 Ma and totally disappeared at 1.9 Ma (early Pleistocene; Table 1). The average individuals number for the demographic reducing events of the effective population size was 2730–4540; for demographic expansion events, N_e reached 119–189 thousand individuals. The equilibrium population number was a minimum of 32% and for some period increased to 58% (Table 1).

The effective population size in the meta-population from the Balkan Peninsula (BB) had started to experiment purification selection 27.3 Ma (Oligocene) and reached a maximum of 21% of the population at 21.4 Ma (early Miocene). This period lasted for long (8.6 Ma) and declined shortly between 11.0 and 12.8 Ma (Table 1). The total period of purifying selection was 16.3 Ma. The bottleneck appeared shortly after the purification selection, at 26.5 Ma (Oligocene) and reached its maximum at 21.8 Ma up to 10.0 Ma (early to middle Miocene) in 32% of populations. After this period of 10.8 Ma, the bottleneck slowly pushed down until 1.3 Ma—early Pleistocene (total period of 25.2 Ma; Table 1). In the final period of bottleneck presence, in some population of *P. nigra* from meta-population BB, the demographic mixed expansion combined with strong positive selection appeared at 4.3 Ma (early Pliocene). This phenomenon reached 17% of populations and reached maximum at 3.9 Ma (middle Pliocene); after a very short period between 2.7 and 2.4 Ma, the mixed expansion with strong positive selection had declined (Table 1). The pure expansion had very little presence in the Black Pine from the Balkan Peninsulas at only 4% of the populations; it started at 2.4–3.6 Ma and finished at 2.4–2.6 Ma (early Pleistocene), with total duration of 0.8–1.2 Ma. The number of populations in permanent equilibrium was 41% and, for some period, increased to 68%. The historical effective population size after the bottleneck and purification selection was av. 3640 individuals; for mixed expansion with strong selection and pure expansion events, N_e reached 323 and 663 thousand individuals, respectively (Table 1).

The sequence of the demographic events for the meta-populations from Western Europe (AA) and the Balkan Peninsula (BB) is similar: The purification selection predated the pure bottleneck, followed by mixed expansion with positive selection and pure expansion. In general, the reducing events predate expansion. For the meta-population from Asia Minor (CC), the sequence of the demographic scenarios is little different. The bottleneck started before purification selection at 9.8 Ma and reached its maximum at 8.5 Ma for 15% of the populations (late Miocene; Table 1). The decline started at 4.3 Ma and quickly finished at 4.1 Ma. The bottleneck lasted in this region was 4.4–5.4 Ma. The purification selection appeared approximately one Ma after the bottleneck at 8.9 Ma (late Miocene) and after a short period reached up to 29% of populations (8.1 Ma; Table 1). The decline started at 5.3 Ma and finished at 3.9 Ma (Pliocene). The total

duration of purification selection was 2.8–4.9 Ma. The pure expansion predated the mixed expansion, started at 4.9 Ma and reached maximum presence for 41% of populations at 4.3 Ma. After that, the pure expansion started to decline at 2.5 Ma and definitively ended at 1.9 Ma (early Pleistocene; Table 1). The total duration of the pure expansion was 1.8–3.0 Ma. The mixed expansion with strong positive selection started after the pure expansion for the population from the Asia Minor at 4.0 Ma and reached a maximum 15% of the population at 3.6 Ma. The decline finished shortly at 2.4 Ma (early Pleistocene; Table 1). The total period of mixed expansion was 1.2–1.6 Ma. The effective population sizes after diversity-reducing events were between 1820 and 2730 individuals and between 484 and 754 thousand individuals after the expansion events, respectively. As far as we are aware, no data are available for *P. nigra* to compare with those reported here (Table 1).

The European Black Pine meta-populations generally experienced a weak bottleneck with an east–west gradient. This phenomenon occurred earlier than the last significant range expansion. The effective population size values obtained here are between 1820 and 3640 individuals immediately after the bottlenecks, which are close to the average value of 3190 individuals for the modern N_e of *Pinus contorta* Douglas ex Loudon in North America (Marshall et al. 2002). Long-term isolation and stability is probably the origin of the small number of haplotypes in some *P. nigra* populations, as observed in *Rhamnus* populations (Hampe et al. 2003) from the same region.

The comparison between the cpDNA and nuDNA bottlenecks shows that nuclear DNA supports a strong bottleneck for *P. nigra* (Table 2). The recently published data from nuclear DNA support the hypothesis of an old reduction in population size (Naydenov et al. 2015). Some Black Pine N_e declines started 1.8 Ma, and others ended 0.2 Ma (middle to late Pleistocene; at HPD-50%). Subsequently, effective population size remained relatively stable until 60,000 years BP (upper Pleistocene, during the last glaciation), and then human activity probably pushed N_e to further decline during the Holocene to the present day. The results obtained for the eastern part of the Balkan Peninsula from MSVR v.04 and M-ratio algorithms (Beaumont 1999; Garza and Williamson 2001) present a strong bottleneck. It is normal for cpSSR to show a weaker bottleneck than nuSSR due to their large effective population size, low mutation rate (μ) and wider/large effective radius of distribution. The ancient bottleneck of Black Pine populations from the Balkan Peninsula is not surprising, compared to the 14.0–18.4 Ma bottleneck recorded for *Iochrominae* taxa (Paape et al. 2008). Bottleneck is not peculiar to Black Pine, among the Mediterranean's pines (Kamari et al. 2010; Naydenov et al. 2011, 2014, 2015). All of them suffered significant reductions of their effective population size throughout different periods in their history (Table 2). Nuclear DNA (nuSSR) analysis showed bottlenecks for *P. halepensis* (≈0–0.005 and ≈0.2–1.7 Ma; Pleistocene), *P. pinaster* (≈0–0.007 Ma; Holocene) and *P. sylvestris* (≈0–0.026 Ma, since the Pleistocene). For this period, the same species have lost nuclear effective population size more of 81% (at HPD-50%). It seems that European Black Pine, Aleppo pine and Scots pine bottlenecks are independent and thus result from independent evolutionary mechanisms.

The observed pure and mixed expansion started with large differences in the numbers of individuals (N_{e-o}), from 4 up to 12,000 (Tables 3S–4S from Appendix, Electronic supplementary material), and indicates an important adaptation capacity and significant variants of principal demographic scenarios. It is likely that (a) the restricted number of N_{e-o} is the result of new mutations or natural hybridization very well adapted to local biotic and abiotic conditions. Further, (b) the large N_{e-o} number is the result of sudden climatic change or migration in new geographic area accompanied by resonance with the successional species in the new ecosystem. The result of population expansion increased the effective population size (N_{e-1}) generally between 37 and 909 thousand individuals (Tables 3S–4S).

Adaptation capacity and subspecies structure

European Black Pine adaptation capacity and subspecies structure probably resulted from a strong positive natural selection with selective sweep (reduction or elimination of variation among the nucleotides in the neighboring DNA of a mutation). A selective sweep occurs with a new mutation, which increases the fitness of the new individual to other members of the population. After natural selection, individuals with stronger fitness will increase in frequency. Then, all genetic neutral variations (hitchhikers) linked to the new mutation will increase at the same time. Hitchhiking may possibly result in part of the European Black Pine's morphological variability and retained ancestral polymorphism. The linked background selection and selective sweep are parts/cases of the general process known at the present day as the Hill–Robertson effect (Cameron et al. 2008).

The significant presence of the populations in equilibrium

There is a significant presence of the populations in equilibrium: 32–58% for meta-population AA, 41–79% for meta-population BB and 29–85% for meta-population CC (Table 1). Additionally, the European Black Pine has the capacity for populations to very quickly reach equilibrium

after expansion or bottleneck. We have observed that the majority of single populations (not the meta-populations), approximately 80%, are under historical expansion or reducing demographic scenarios, after which they reached genetic equilibrium and remained there for a very long time (i.e., observed only two demographic steps). Exceptions from this rule have been observed for: TR-14, TR-22 and TR-27 from bottleneck to purification selection and finish in equilibrium; BG-7, CY-02, ES-03, GR-02 and RO-03 from bottleneck through equilibrium to mixed expansion and finish in equilibrium; CY-01, ES-02, RO-02, TR-03, TR-05, TR-06, TR-07 and TR-11 from purification selection through equilibrium to pure expansion and finish in equilibrium; and BG-03, ES-05, GR-01 and RO-04 from pure to mixed expansion and finish in equilibrium (Tables 1S–5S from Appendix, Electronic supplementary material). The full cycle of historical genetic demographic fluctuation in natural populations as "*purification selection* → *bottleneck* → *mixed expansion* → *equilibrium* → *pure expansion* → *purification selection* →" has been generally observed at the meta-population level (i.e., AA, BB and CC, see Table 2). This is very important and suggests that the future conservation strategies must reach the meta-population level to better protect biodiversity in each region. The single population strategy for conservation management appears more appropriate for a short-term solution (i.e., non-sustainable). For this reason, the protection of biodiversity must be a priority of international efforts and not only for local national indicatives. The collective international effort to prevent forest resources from forest fire should continue, and protection from illegal harvesting and large pan-European programs for reforestation and protection of biodiversity are urgent.

Conclusions

Periodically, different European Black Pine populations from their fragmented natural area of distribution went through a bottleneck in the limit of each of the meta-population. The period of bottleneck (and diversity-reducing events) for chloroplast DNA is relatively long and can differ between different meta-populations; in contrast, for nuclear DNA, the bottleneck is shorter and stronger. Small and isolated populations may not accurately represent ancestral population structure in different parts of its natural distribution. This type of bottleneck has a founder effect, and multiple varieties and subspecies over areas of distribution confirm this statement (Ivanov 1971; Kostov 1974; Mihailov 1983, 1987, 1993, 1998). The multiple specific forms and phenotypic varieties are reported over full areas of the natural distribution of European Black Pine (Gulcu and Ucler 2008; Akkemik et al. 2010; Bogunic

et al. 2010; Sivacioglu and Ayan 2010). The phenotype variability exists in micro- and macro-geographic schedules. Later, the founder effect can lead to individual-based speciation (Crow and Morton 1955; Eyre-Walker et al. 2002). The speciation and migration are opposite and complimentary for the historic development of populations. The low migration pattern and multiple phenotypic forms confirm the permanent "*sympatric speciation*" for European Black Pine. The very long bottleneck is possibly due to incomplete isolation; the observed spatial distribution today has changed significantly through time (for millions of years), and some isolated populations today were probably related in the past (and vice versa). This is a possible relationship between the observed reduction of effective population size and multiple subspecies and variety structure of *P. nigra*.

References

Akkemik U, Yılmaz H, Oral D et al (2010) Some changes in taxonomy of pines (*Pinus* L.) native to Turkey, Istanbul University. J Fac For 61(1):63–78

Arslan M, Tosun S, Ok K et al. (2012) Seben Fosil Ormanı'nın Doğal ve Kültürel Değerlerinin Saptanması ve Uygun Yönetim Planının Geliştirilmesi (Determination of Natural and Cultural Values and Development of Management Plan in the Seben Petrified Forest), Batı Karadeniz Ormancılık Araştırma Enstitüsü Müdürlüğü (The Western Black Sea Forestry Research Institute), Press No: 31, ISBN 978-605-393-124-9

Bacles C, Jump A (2010) Taking a tree's perspective on forest fragmentation genetics. Trends Plant Sci 16:13–18

Bai W, Liao W, Zhang D (2010) Nuclear and chloroplast DNA phylogeography reveal two refuge areas with asymmetrical gene flow in a temperate walnut tree from East Asia. New Phytol 188:892–901

Beaumont M (1999) Detecting population expansion and decline using microsatellites. Genetics 153:2013–2029

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 M, Hauffe H, Rizzoli A, Vernesi C (eds) Population genetics for animal conservation. Cambridge University Press, Cambridge. doi:10.1017/CBO9780511626920

Beerli P, Felsenstein J (1999) Maximum-likelihood estimation of migration rates and effective population numbers in two populations using a coalescent approach. Genetics 152:763–773

Belousov V, Volvovsky S, Arkhipov V et al (1988) Structure and evolution of the earth's crust and upper mantle of the Black Sea Boll. Geofis Teor Appl 30:109–196

Bijlsma R, Loeschcke V (2012) Genetic erosion impedes adaptive responses to stressful environments. Evol Appl 5:117–129

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

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

Braverman J, Hudson R, Kaplan N et al (1995) The hitchhiking effect on the site frequency spectrum of DNA polymorphism. Genetics 140:783–796

Brooks A (2012) Phylogeography and Species distribution modeling of the flowering dogwood, *Cornus florida* L. (*Cornaceae*). Thesis. http://www.lib.ncsu.edu/resolver/1840.16/8880

Brown T (1997) Clearances and clearings: deforestation in mesolithic/neolithic britain. Oxf J Archaeol 16(2):133. doi:10.1111/1468-0092.00030

Bueno M, Pennington R, Dexter K et al (2016) Effects of quaternary climatic fluctuations on the distribution of neotropical savanna tree species. Ecography 39:001–012. doi:10.1111/ecog.01860

Cameron J, Williford A, Kliman R (2008) The Hill-Robertson effect: evolutionary consequences of weak selection in finite populations. Heredity 100:19–31

Campani M, Mulch A, Kempf O et al (2012) Miocene paleotopography of the Central Alps. Earth Planet Sci Lett 174:337–338. doi:10.1016/j.epsl.2012.05.017

Carreras C, Pascual M, Cardona L et al (2007) The genetic structure of the loggerhead sea turtle (*Caretta caretta*) in the Mediterranean as revealed by nuclear and mitochondrial DNA and its conservation implications. Conserv Genet 8:761–775

Caujape-Castells J, Jansen R (2003) The influence of the Miocene Mediterranean desiccation on the geographical expansion and genetic variation of *Androcymbium gramineum* (Cav.) McBride (*Colchicaceae*). Mol Ecol 12:1515–1525

Cengel B, Tayanc Y, Kandemir G et al (2012) Magnitude and efficiency of genetic diversity captured from seed stands of *Pinus nigra* (Arnold) subsp. *pallasiana* in established seed orchards and plantations. New For 43(3):303–317

Champagnac J, Schlunegger F, Norton K et al (2009) Erosion-driven uplift of the modern Central Alps. Tectonophysics 474(1–2):236–249. doi:10.1016/j.tecto.2009.02.024

Charlesworth B (2009) Effective population size and patterns of molecular evolution and variation. Nat Rev Genet 10:195–205. doi:10.1038/nrg2526

Charlesworth J, Eyre-Walker A (2007) The other side of the nearly neutral theory, evidence of slightly advantageous back-mutations. Proc Natl Acad Sci USA 104:16992–16997

Chetverikov S (1961) On certain aspects of the evolutionary process from the standpoint of modern genetics. (transl. of 1921 paper by Malina Parker; ed I.M. Lerner). Proc Am Philos Soc 105(2):167–195

Christopoulou A, Fulé PZ, Andriopoulos P, Sarris D, Arianoutsou M (2013) Dendrochronology-based fire history of *Pinus nigra* forests in Mount Taygetos, Southern Greece. For Ecol Manage 293:132–139

Chybicki I, Dzialuk A (2014) Bayesian approach reveals confounding effects of population size and seasonality on outcrossing rates in a fragmented subalpine conifer. Tree Genet Genomes 10:1723–1737. doi:10.1007/s11295-014-0792-3

Clark C, Carbone I (2008) Chloroplast DNA phylogeography in long-lived Huon pine, a Tasmanian rain forest conifer. Can J For Res 38:1576–1589

Comps B, Gomory D, Letouzey J et al (2001) Diverging trends between heterozygosity and allelic richness during postglacial colonization in the European beech. Genetics 157:389–397

Crow J, Morton N (1955) Measurement of gene frequency drift in small populations. Evolution 9:202–214

Davies S, Cavers S, Finegan B et al (2010) Genetic consequences of multigenerational and landscape colonisation bottlenecks for a neotropical forest pioneer tree, *Vochysia ferruginea*. Trop Plant Biol 3:14–27

Dijkstra S (1973) Fossilium catalogus.: II Plantae., Part 84, Alexander Doweld Publisher, ISBN: 9061933188, 9789061933182

Dobzhansky T (1982) Genetics and the origin of species, Reprint edn. Columbia University Press, New York

Du Z-Y, Wang Q-F (2016) Allopatric divergence of *Stuckenia filiformis* (*Potamogetonaceae*) on the Qinghai-Tibet Plateau and its comparative phylogeography with *S. pectinata* in China. Sci. Rep. 6:20883. doi:10.1038/srep20883

Eder J, Jechorek H, Kvacek Z et al (2008) The integrated plant record: an essential tool for reconstructing neogene zonal vegetation in Europe. Palaios 23:97–111. doi:10.2110/palo.2006.p06-039r

Edh K, Widen B, Ceplitis A (2007) Nuclear and chloroplast microsatellites reveal extreme population differentiation and limited gene flow in the Aegean endemic *Brassica cretica* (*Brassicaceae*). Mol Ecol 16:4972–4983

Ehrendorfer F (2013) Woody plants—evolution and distribution since the tertiary: proceedings of a symposium organized by Deutsche Akademie der Naturforscher LEOPOLDINA in Halle/Saale, German Democratic Republic, 9–11 Oct 1986. Second publication In: Springer, 11 Nov 2013—Science

Excoffier L (2004a) Special issue: analytical methods in phylogeography and genetic structure. Mol Ecol 13:727. doi:10.1111/j.1365-294X.2004.02170.x

Excoffier L (2004b) Analysis of population subdivision. Handbook of statistical genetics, vol 4. Wiley, p 24. doi:10.1002/0470022620.bbc25

Excoffier L, Schneider S, Roessli D (2002) Arlequin ver 2.001: a software for population genetics data analysis. Department of Anthropology and Ecology, University of Geneva, Geneva. http://cmpg.unibe.ch/software/arlequin/software/2.001/doc/whatsnew/whatsnew.html

Eyre-Walker A, Keightley P, Smith N et al (2002) Quantifying the slightly deleterious mutation model of molecular evolution. Mol Biol Evol 19:2142–2149

Fady-Welterlen B (2005) Is there really more biodiversity in Mediterranean forest ecosystems? Taxon 54:905–910

Finch J, Leng M, Marchant R (2009) Late quaternary vegetation dynamics in a biodiversity hotspot, the Uluguru Mountains of Tanzania. Quatern Res 72(1):111–122. doi:10.1016/j.yqres.2009.02.005

Flannery T (1994) The future eaters. Reed Books, Melbourne (ISBN: 0-7301-0422-2)

Frantz L, Madsen O, Megens H et al (2014) Testing models of speciation from genome sequences: divergence and asymmetric admixture in Island South-East Asian Sus species during the Plio-Pleistocene climatic fluctuations. Mol Ecol 23(22):5566–5574. doi:10.1111/mec.12958

Fu T (1997) Statistical tests of neutrality of mutations against population growth, hitchhiking, and background selection. Genetics 147:915–925

Gaggiotti O, Excoffier L (2000) A simple method of removing the effect of a bottleneck and unequal population sizes on pairwise genetic distances. Proc R Soc Lond B 267:81–87

Galtier N, Depaulis F, Barton N (2000) Detecting bottlenecks and selective sweeps from DNA sequence polymorphism. Genetics 155:981–987

Gao Y-D, Zhang Y, Gao X-F et al (2015) Pleistocene glaciations, demographic expansion and subsequent isolation promoted

morphological heterogeneity: a phylogeographic study of the alpine *Rosa sericea* complex (*Rosaceae*). Sci Rep 5:11698. doi:10.1038/srep11698

Garza J, Williamson E (2001) Detection of reduction in population size using data from microsatellite DNA. Mol Ecol 10:305–318

Geary D, Rich J, Valley J et al (1989) Stable isotopic evidence of salinity change: influence on the evolution of melanopsid gastropod in the late Miocene Pannonian basin. Geology 17:981–985

Goncharenko G, Silin A (1997) Populyatsionnaya i evolyutsionnaya genetika sosen Vostochnoi Evropy i Sibiri. [Population and Evolutionary Genetics of Pine in Eastern Europe and Siberia.] Tekhnalogiya, Minsk, Belarus (**in Russian**)

Gonzalez-Martınez S, Gerber S, Cervera M et al (2002) Seed gene flow and fine scale structure in a Mediterranean pine (*Pinus pinaster* Ait.) using nuclear microsatellite markers. Theor Appl Genet 104:1290–1297

Gorbunov M (1958) Tretichnye sosny Zapadnoj Sibiri- *Pinus thomasiana* Varietas *tomskiana and Pinus thomasiana* Varietas *kasparanica*. Tomsk State University. Bot Z 3(43):349–350 **(Russian)**

Gordo I, Navarro A, Charlesworth B (2002) Muller's ratchet and the pattern of variation at a neutral locus. Genetics 161:835–848

Gradstein F, Ogg J, Smith A et al (2005) A geologic time scale 2004. Cambridge University Press, Cambridge, 589 pp

Griffiths R, Tavare S (1994a) Simulating probability distributions in the coalescent. Theor Popul Biol 46:131–159

Griffiths R, Tavare S (1994b) Sampling theory for neutral alleles in a varying environment. Philos Trans R Soc Lond B Biol Sci 344:403–410

Gulcu S, Ucler A (2008) Genetic variation of Anatolian black pine (*Pinus nigra* Arnold. subsp. *pallasiana* (Lamb.) Holmboe) in the Lakes district of Turkey. Silvae Genet 57:1–5

HAGRC-Helmholtz Association of German Research Centers (2009) Are the Alps growing or shrinking? Science Daily. www.sciencedaily.com/releases/2009/11/091105121207.htm

Hammor G, Halmai J (1988) Neogene paleogeographic atlas of central and Estern Europe. Geological Institute, Budapest

Hampe A, Arroyo J, Jordano P et al (2003) Rangewide phylogeography of a bird-dispersed Eurasian shrub: contrasting Mediterranean and temperate glacial refugia. Mol Ecol 12:3415–3426

Harpending H (1994) Signature of ancient population growth in a low-resolution mitochondrial DNA mismatch distribution. Hum Biol 66(4):591–600

Hartl D, Clark A (1997) Principles of population genetics, 3rd edn. Sinauer Associates Inc, Sunderland

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

Heredia U, Nanos N, Rey E et al (2015) High seed dispersal ability of *Pinus canariensis* in stands of contrasting density inferred from genotypic data. For Syst 24(1):e-015. doi:10.5424/fs/2015241-06351

Hofmann S, Kraus S, Dorge T et al (2014) Effects of Pleistocene climatic fluctuations on the phylogeography, demography and population structure of a high-elevation snake species, *Thermophis baileyi*, on the Tibetan Plateau. J Biogeogr 41:2162–2172. doi:10.1111/jbi.12358

Ivanov I (1971) Form variability of *Pinus nigra* (Arn.) in Western Rhodope Mountain. Ph.D. thesis, Forest Research Institute, Bulgarian Academy of Sciences, Sofia

Jipa D, Olariu C (2009) Dacian basin depositional architecture and sedimentary history of a Papatethys sea. National Institute of Marine Geology and Geo-Ecology (GeoEcoMar) - Romania, Special Publication No. 3

Juste J, Bilgin R, Munoz J et al (2009) Mitochondrial DNA signatures at different spatial scales: from the effects of the Straits of Gibraltar to population structure in the meridional serotine bat (*Eptesicus isabellinus*). Heredity 103:178–187

Kamari S, Naydenov KD, Benyounes H et al (2010) Genetic signals of ancient decline in Aleppo pine populations at the species' southwestern margins in the Mediterranean Basin. Hereditas 147:165–175. doi:10.1111/j.1601-5223.2010.02176.x

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

Kingman J (1982a) On the genealogy of large populations. In: Gani J, Hannan EJ (ed) Essays in statistical science. Applied Probability Trust, London, pp 27–43. (Also as J Appl Probab 19A: 27–43)

Kingman J (1982b) The coalescent. Stoch Process Appl 13:235–248

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

Kramer A, Ison A, Ashley M et al (2008) The paradox of forest fragmentation genetics. Conserv Biol 22:878–885

Lee R, Daly R (1999) Cambridge encyclopedia of hunters and gatherers. Cambridge University Press, Cambridge, 534 pp

Leys B, Finsinger W, Carcaillet C (2014) Historical range of fire frequency is not the Achilles' heel of the Corsican black pine ecosystem. J Ecol 102:381–395

Li W (1977) Stochastic models in population genetics. Benchmark papers in genetics, vol 7. Dowden, Hutchinson & Ross, Stroudsburg, Pennsylvania. 484 S Biomed J 21: 297. doi:10.1002/bimj.4710210311

Lin H, Sanchez-Ortiz C, Hastings P (2009) Colour variation is incongruent with mitochondrial lineages: cryptic speciation and subsequent diversification in a Gulf of California reef fish (*Teleostei: blennioidei*). Mol Ecol 18:2476–2488

Linares J, Tiscar P (2010) Climate change impacts and vulnerability of the southern populations of *Pinus nigra* subsp. *salzmannii*. Tree Physiol 30(7):795–806. doi:10.1093/treephys/tpq052

Liu J, Möller M, Provan J et al (2013) Geological and ecological factors drive cryptic speciation of yews in a biodiversity hotspot. New Phytol 199:1093–1108

Lowe A (2005) Population genetics of neotropical trees focus issue. Heredity 95:243–245

Lowe A, Boshier D, Ward M et al (2005) Genetic resource loss following habitat fragmentation and degradation; reconciling predicted theory with empirical evidence. Heredity 95:255–273

Lumaret R, Mir C, Michaud H et al (2002) Phylogeographical variation of chloroplast DNA in holm oak (*Quercus ilex* L.). Mol Ecol 11:2327–2336

Magyar I, Geary D, Muller P (1999) Paleogeographic evolution of the Late Miocene Lake Pannon in Central Europe. Palaeogeogr Palaeoclimatol Palaeoecol 147:151–167

Mann H, Whitney D (1947) On a test of whether one of two random variables is stochastically larger than the other. Ann Math Stat 18(1):50–60. doi:10.1214/aoms/1177730491

Marshall H, Newton C, Ritland K (2002) Chloroplast phylogeography and evolution of highly polymorphic microsatellites in lodgepole pine (*Pinus contorta*). Theor Appl Genet 104:367–387

Mihailov V (1983) Sur la variabiliteˊ endoge'ne des aiguilles et son importance pour la taxonomie du pin noir (*Pinus nigra* Arn.). For Sci (Bulgaria) 1:3–20

Mihailov V (1987) Variability of European Black pine (*Pinus nigra* Arn.) in size, weight and form of the seeds in Pirin and Slavianka mountains. For Sci (Bulgaria) 6:26–37

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, (Bulgaria), Sofia

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 (Bulgaria) 1 (2):24–37

Milankovitch M (1920) Theorie Mathematique des Phenomenes Thermiques produits par la Radiation Solaire. Gauthier-Villars et Cie., Paris, 338 pp

Milankovitch M (1930) Mathematische Klimalehre und Astronomische Theorie der Klimaschwankungen, Handbuch der Klimalogie Band 1 Teil A Borntrager Berlin

Milankovitch M (1941) Canon of Insolation and the ice age problem. Zavod za Udžbenike i Nastavna Sredstva, Belgrade. ISBN: 86-17-06619-9

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

Moores E, Fairbridge R (1998) Encyclopedia of European and Asian regional geology. Encyclopedia of Earth Sciences Series, London

Naydenov KD, Alexandrov A (1999) Geographic variability of some of the Monoterpenes (α-pinene, camphene and β-pinene) in Authohtonious population of Scots Pine (*Pinus sylvestris* L.) in Rila-Rhodopses massif. Diagnosis Press. J Biotechnol Biotechnol Equip 13(2):14–18

Naydenov KD, Alexandrov A (2000) Application of transformed data on terpenes in *Pinus sylvestris* L. populations for phenotypic studies. Diagnosis Press. J Biotechnol Biotechnol Equip 14 (1):40–46

Naydenov KD, Alexandrov A, Tremblay F (2002) Terpene composition of Scots pine (*Pinus sylvestris* L.) in the eastern part of the Balkan Peninsula. 1. Provenance tests. Diagnosis Press. J Biotechnol Biotechnol Equip 16(2):99–108

Naydenov KD, Tremblay F, Bergeron Y et al (2006) Germination response to forest fire-related charcoal active effect of Jack pine (*Pinus banksiana* Lamb.) seeds. Can J For Res 36:761–767

Naydenov KD, Naydenov MK, Tremblay F et al (2011) Patterns of genetic diversity that result from bottlenecks in Scots Pine and the implications for local genetic conservation and management practices in Bulgaria. New For 42:179–193. doi:10.1007/s11056-010-9245-5

Naydenov KD, Alexandrov A, Naydenov M et al (2012) Impact of activated charcoal on germination and initial growth of some pine species. J Balk 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. doi:10.1111/j.1601-5223.2013.00027.x

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

Naydenov KD, Naydenov MK, 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. doi:10.1007/s11295-016-1022-y

NordNordWest Trust-Germany (2005) exhibition—Ausstellung "Meeresstrand am Alpenrand" der Niederösterreichischen Landesmuseum, under licence. http://creativecommons.org/licenses/by-sa/3.0/de/legalcode, ISBN: 978-3-85252-644-7

Overpeck J, Otto-Bliesner B, Miller G et al (2006) Paleoclimatic evidence for future ice-sheet instability and rapid sea-level rise. Science 311(5768):1747–1750. doi:10.1126/science.1115159

Paape T, Igic B, Smith S et al (2008) A 15-Myr-old genetic bottleneck. Mol Biol Evol 25(4):655–663. doi:10.1093/molbev/msn016

Palamarev E (1989) Paleobotanical evidences of the tertiary history and origin of the Mediterranean sclerophyll dendroflora. Plant Syst Evol 162(1/4):93–107

Papaianopol I, Marinescu F, Popescu A et al (1987) Paleogeographie du Pontien du Bassin Dacique, insistant sur le developement du facies charboneux. D S Inst Geol Geofiz Bucarest 72–73(4):261–275

Petit R, Hampe A (2006) Some evolutionary consequences of being a tree. Annu Rev Ecol Evol Syst 37:187–214

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

Popov S, Rogl F, Rozanov A et al (2004) Lithological paleogeographic maps of Paratethys. 10 maps Late Eocene to Pliocene. Cour Forsch Senckenberg 250:1–46

Poulsen T, Veyret P, Diem A (2015) The Alps. Encyclopedia Britannica

Rafii Z, Dodd R (2007) Chloroplast DNA supports a hypothesis of glacial refugia over post-glacial recolonization in disjunct populations of black pine (*Pinus nigra*) in Western Europe. Mol Ecol 16:723–736

Richmond G, Fullerton D (1986) Summation of quaternary glaciations in the United States of America. Quatern Sci Rev 5:183–196. doi:10.1016/0277-3791(86)90184-8

Rogers A (1995) Genetic evidence for a pleistocene population explosion. Evolution 49(4):608–615. doi:10.2307/2410314

Rogers A, Harpending H (1992) Population growth makes waves in the distribution of pairwise genetic differences. Mol Biol Evol 9:552–569

Rozas J, Sanchez-Del-Barrio J, Messeguer X et al (2003) Dna SP, DNA polymorphism analysis by the coalescent and other methods. Bioinformatics 19:2496–2497

Rubio-Moraga A, Candel-Perez D, Lucas-Borja M et al (2012) Genetic diversity of *Pinus nigra* Arn. populations in southern Spain and northern Morocco revealed By inter-simple sequence repeat profiles. Int J Mol Sci 13:5645–5658

Saulea E, Popescu I, Săndulescu J (1969) Atlas litofacial. VI – Neogen, 1:200.000 (in Romanian and in French). Institute Geologic, Bucureşti

Savolainen O, Pyhajarvi T, Knurr T (2007) Gene flow and local adaptation in trees. Annu Rev Ecol Evol S 38:595–619

Schneider S, Excoffier L (1999) Estimation of past demographic parameters from the distribution of pairwise differences when the mutation rates vary among sites: application to human mitochondrial DNA. Genetics 152(3):1079–1089

Scotese C (2001) Atlas of earth history. PALEOMAP Project, Arlington, p 52

Scotese C, Golonka J (1992) PALEOMAP paleogeographic Atlas, PALEOMAP Progress Report No. 20. Department of Geology, University of Texas at Arlington, Arlington

Scotese C, Kazmin V, Natapov L et al (1998) The paleogeographic atlas of northern Eurasia. Institute of Tectonics of Lithospheric Plates, Russian Academy of Sciences, Moscow **(26 maps)**

Sengor A, Natalin B (1996) Palaeotectonics of Asia: fragments of a synthesis, the tectonic evolution of Asia. In: Yin A, Harrison M (eds) Rubey Colloquium. Cambridge University Press, Cambridge, pp 486–640

SGN-Senckenberg Research Institute and Natural History Museum (2012) Outstanding for the past 15 million years: swiss Alps have influenced Europe's climate since the Miocene. Science Daily. www.sciencedaily.com/releases/2012/07/120710093407.htm

Shapiro S, Wilk M (1965) An analysis of variance test for normality. Biometrika 52(3–4):591–611. doi:10.1093/biomet/52.3-4.591

Simonsen K, Churchill G, Aquadro C (1995) Properties of statistical tests of neutrality for DNA polymorphism data. Genetics 141:413–429

Sivacioglu A, Ayan S (2010) Variation in cone and seed characteristics in a clonal seed orchard of Anatolian black pine [*Pinus nigra* Arnold subsp. *pallasiana* (Lamb.) Holmboe]. J Environ Biol 31:119–123

Stebbins G (1974) Flowering plants. Evolution above the species level. Harvard University Press, Cambridge

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

Stefanov B (1943) The phyto-geographical elements of Bulgaria. Thesis of Bulgarian Academy of Sciences, Faculty of Nature and Mathematics, Sofia, vol. XXXIX, N 19

Tajima F (1989) Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. Genetics 123(3):585–595

Taylor D, Keller S (2007). Historical range expansion determines the phylogenetic diversity introduced during contemporary species invasion. Evolution 61(2): 334–345. http://www.jstor.org/stable/4621290

Thesing B, Noye R, Starkey D et al (2016) Pleistocene climatic fluctuations explain the disjunct distribution and complex phylogeographic structure of the southern red-backed Salamander, *Plethodon serratus*. Evol Ecol 30:89. doi:10.1007/s10682-015-9794-3

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

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

Wakeley J, Aliacar N (2001) Gene genealogies in a meta-population. Genetics 159:893–905

Wang J, Wu Y, Ren G et al (2011) Genetic differentiation and delimitation between ecologically diverged *Populus euphratica* and *P. pruinosa*. PLOS ONE 6(10):e26530. doi:10.1371/journal.pone.0026530

Woolfit M, Bromham L (2005) Population size and molecular evolution on islands. Proc R Soc B 272:2277–2282

Wright S (1938) Size of population and breeding structure in relation to evolution. Science 87:430–431

Wright S (1942) Statistical genetics and evolution. Bull Am Math Soc 48(4):223–246. doi:10.1090/S0002-9904-1942-07641-5

Wright S (1969) Evolution and the genetics of populations, vol 2. University Chicago Press, Chicago

Yuan Z-Y, Suwannapoom C, Yan F et al (2016) Red River barrier and Pleistocene climatic fluctuations shaped the genetic structure of *Microhyla fissipes* complex (Anura: *Microhylidae*) in southern China and Indochina. Curr Zool 62(6):531–543. doi:10.1093/cz/zow042

Ziegler P (1988) Evolution of the Arctic-North Atlantic and the western Tethys. AAPG Memoir 43:164–196