Acta Societatis Botanicorum Poloniae

ORIGINAL RESEARCH PAPER

Phylogeny of beech in western Eurasia as inferred by approximate Bayesian computation

Dušan Gömöry*, Ladislav Paule, Vladimír Mačejovský

Faculty of Forestry, Technical University in Zvolen, TG Masaryka 24, 960 53 Zvolen, Slovakia

* Corresponding author. Email: gomory@tuzvo.sk

Abstract

The *Fagus sylvatica* L. species complex in Europe and Western Asia comprises two commonly recognized subspecies, *F. sylvatica* subsp. *sylvatica* [= *F. sylvatica* sensu stricto (s. str.)] and *F. sylvatica* subsp. *orientalis* (= *F. orientalis*), and two putatively hybridogenous or intermediate taxa, "*F. moesiaca*" and "*F. taurica*". The present study aimed to examine the demographic history of this species complex using 12 allelic loci of nine allozymes scored in 279 beech populations in western Eurasia. Three sets of phylogenetic scenarios were tested by approximate Bayesian computation: one dealing with the divergence of subspecies and/or regional populations within the whole taxonomical complex, and two others focusing on the potential hybrid origin of "*F. moesiaca*" and "*F. taurica*". The best-supported scenario within the first set placed the time of divergence of regional populations of *F. orientalis* in the Early Pleistocene (1.18–1.87 My BP). According to this scenario, the Iranian population was the ancestral lineage, whereas *F. sylvatica* s. str. was the lineage that diverged most recently. "*Fagus taurica*" was found to have originated from hybridization between the Caucasian population of *F. orientalis* and *F. sylvatica* s. str. at 144 ky BP. In contrast, there was no evidence of a hybrid origin of "*F. moesiaca*". The best-supported scenario suggested that the Balkan lineage is a part of *F. sylvatica* s. str., which diverged early from *F. orientalis* in Asia Minor (817 ky BP), while both the Italian and Central-European lineages diverged from the Balkan one later, at the beginning of the last (Weichselian) glacial period.

Keywords

Fagus sylvatica L.; *Fagus orientalis* Lipsky; phylogenetic scenario; allozymes

Introduction

Beech trees represent the most important component of montane forests in the temperate zones of Europe and Western Asia. Beech forests with the admixture of local spruce and fir species are the most productive and stable native forest communities in the mountains of Central Europe, as well as the Near and Middle East. As such, various taxa of the genus *Fagus* L. have long been the focus of the interest of botanists. However, the number and rank of *Fagus* taxa in western Eurasia have been controversial, and their phylogeny is still not resolved. The state-of-the-art view, accepted by *Flora Europaea* [1] and recent systematic revisions [2,3], considers all beech populations in western Eurasia as belonging to a single species, *Fagus sylvatica* L., with two subspecies: *F. s.* subsp. in most of the species' European distribution [denoted hereafter as *F. sylvatica* sensu stricto (s. str.) for simplicity] and *F. s.* subsp. *orientalis* (denoted hereafter as *F. orientalis*), whose range begins in the southeastern Balkan Peninsula (Black Sea coast of Bulgaria and northeastern Greece) and continues through Turkey and the Caucasus to the Alborz Mountains in Iran. However, the latter taxon was described by Lipsky [4] as a separate

species, *F. orientalis* Lipsky, and this opinion is generally shared by local botanists [5,6] and forestry practitioners [7]. Palibin [8] distinguished beech in the Caucasus and the Alborz Mountains as representing a separate species, *F. hohenackeriana*, but this view has not been generally accepted, and Greuter and Burdet [9] included it within their *F. sylvatica* subsp. *orientalis*. Two other taxa are mostly considered to be intermediates between the above subspecies: "*F. taurica*" Popl. distributed in the mountains of the Crimean peninsula, and "*F. moesiaca*" (Domin, Maly) Czeczott of the southern Balkans. Both taxa were originally described as separate species ([10] and [11], respectively), but again views about their taxonomical statuses are quite controversial, and they are sometimes considered to represent lineages that originated from hybridization between the subspecies [4,8,12,13]. The subspecies also differ in the patterns of their geographical distributions: in Europe, the range of beech is almost continuous, except in peripheral regions like the Cantabrian range, Apennine Peninsula, or southern Balkans, whereas the range of oriental beech is much more fragmented into several large populations that are relatively isolated from each other (northern Turkey, the Amanus Mountains, the Caucasus, and the Alborz) [14,15].

Knowledge of the phylogeny of *F. sylvatica* is also limited. Complex studies based on internal transcribed spacer (ITS) ribosomal DNA sequences and morphology of this taxon plus related Asian and North American *Fagus* species were carried out by Denk et al. [16,17]. Thorough insights into beech phylogeny based on the fossil record, biogeography, and geological history were provided in a study by Denk and Grimm [18]. *Fagus sylvatica* was represented in all three of these studies, but they did not focus specifically on this species. Generally, phylogenetic studies within the genus *Fagus* have included a broad variety of species covering the whole range of the genus, but they typically lacked the detail necessary to reveal relationships among subspecific taxa or regional populations within *F. sylvatica* s. l. [19–21].

Marker-based studies focusing on beech in western Eurasia, including those based on the same dataset as the current study [22,23] (and see below), usually inferred the evolutionary past of beech populations from the results of clustering methods, which were either distance-based or based on simulations (mostly Bayesian Monte Carlo–Markov chain approaches, such as those of Pritchard et al. [24] or Corander et al. [25]). Recently, more direct and more powerful tools based on coalescent theory have been developed, such as those using approximate Bayesian computation (ABC), which allow simultaneous estimation of the times of divergence of genetic lineages and their effective population sizes [26,27]. In this study, we applied the ABC approach to verify the conclusions about the phylogenetic relationships among beech taxa and regional populations from earlier studies [22,23], and to make inferences about the past demography of beech in western Eurasia.

Material and methods

This study examined 279 beech populations covering the majority of the geographic range of the whole taxonomical complex (hereafter denoted as *F. sylvatica* s. l.), each represented by a minimum of 50 adult trees. Twelve polymorphic gene loci of nine isozyme systems were analyzed: glutamate oxaloacetate transaminase (*Got-B*), isocitrate dehydrogenase (*Idh-A*), leucine aminopeptidase (*Lap-A*), malate dehydrogenase (*Mdh-A*, *Mdh-B*, *Mdh-C*), menadione reductase (*Mnr-A*), phosphoglucomutase (*Pgm-A*), phosphoglucose isomerase (*Pgi-B*), peroxidase (*Per-A*, *Per-B*), and shikimate dehydrogenase (*Skdh-A*). Details of enzyme extraction, electrophoresis, and staining procedures are given in Gömöry et al. [22].

Interpretation of zymograms and genotyping followed Merzeau et al. [28] and Müller-Starck and Starke [29]. Alleles at each locus were identified by the relative migration rates of their respective protein fractions, expressed as a percent of the migration rate of the most frequent allele of that gene in the Central- and Western-European beech populations.

Speciation scenarios were tested by applying ABC analyses. However, simulations based on the complete dataset consisting of 15,194 genotypes would last excessively long (estimated duration in the range of several weeks per simulation), so the dataset

had to be trimmed in such a way that a certain number of randomly chosen trees per population were retained. In the dataset of *F. sylvatica* s. str., trimming was the strongest, and only two trees per population were retained except for populations that originated from minor glacial refugia and were underrepresented in the collection, specifically the Apennine Peninsula (all samples retained) and "*F. moesiaca*" (southern Balkans; 14 samples per population). For *F. orientalis*, 24 samples per population were retained from the Alborz and Asia Minor regions, 28 from the Caucasus, and 27 of "*F. taurica*" (Crimea). In this way, the total sample size was reduced to 2,659 trees (resulting in a reduction of computing time to few days per simulation), while all studied taxa and/or regional populations were represented by comparable sample sizes.

The computations were performed in DIYABC v. 2.05 [30]. In all runs, 100,000 datasets were simulated for each scenario. As we have no information on the distribution of mutation rates of allozyme loci in *Fagus*, these were set to vary between 10^{-7} and 10^{-5}, which is analogous to observed values in other organisms [31,32]. The 1% of the simulated datasets that were the most similar to the observed data were used to estimate relative posterior probabilities. Simulations applied the generalized stepwise mutation (GSM) model. No prior assumptions were made of divergence times. As the ABC algorithms use the number of generations to estimate the divergence time, a generation turnover time of 100 years was chosen for the conversion of generations into a time scale of years. Isolated beech trees may start flowering at 15–20 years of age, and trees in a canopy at approximately 50 years [14,33]; however, the average age of offspring-producing trees is certainly higher in a natural forest ecosystem. At the local scale, abundant regeneration repeatedly occurs every 100–120 years, when the first gaps occur in the closed canopy after a stage of one-layer stand associated with the culmination of the standing stock [34]. The lower limit of this interval was taken as the estimated generation time, as the earliest survivors are likely to develop the largest dimensions and thus contribute the most through both male and female gametes to the gene pool of the offspring.

Model checking was done by ranking the summary statistics of the observed data (mean number of alleles and mean heterozygosity across loci, FST) against the posterior predictive distributions, as well as by principal component analysis (PCA), which placed the observed data onto the cloud of datasets simulated with the prior distributions of parameters and datasets from the posterior predictive distribution.

The scenarios tested are shown in Fig. 1, and they are displayed geographically in Fig. S1. The first set of scenarios focused on the phylogeny within *F. sylvatica* s. l., namely on the divergence of regional populations within *F. orientalis* (Asia Minor, Caucasus, and Alborz) and the potential paraphyly of *F. sylvatica* s. str. In the second round of simulations, we tested whether Crimean beech ("*F. taurica*") is a lineage within *F. orientalis* or a taxon that originated from hybridization between *F. sylvatica* s. str. and *F. orientalis*. Finally, the last set of scenarios addressed the potential hybrid origin of Balkan beech ("*F. moesiaca*") and the phylogenetic position of beech populations in the Apennine Peninsula. Only scenarios with realistic biogeographic settings were considered.
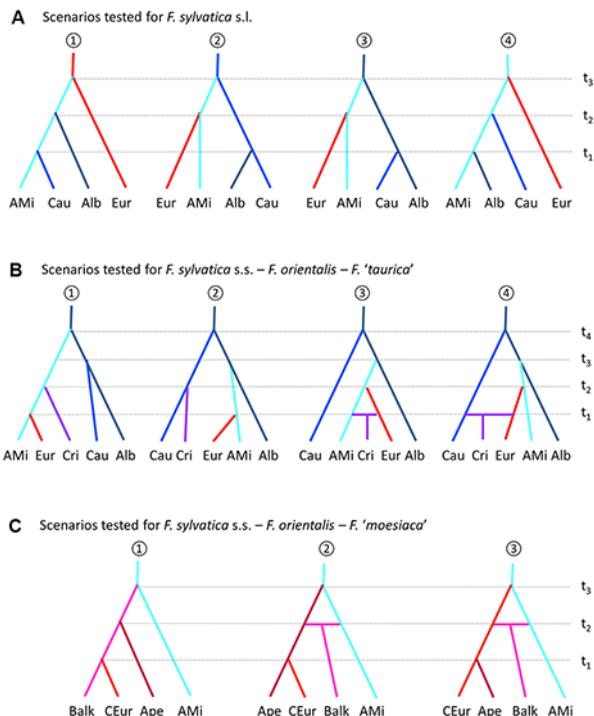
Fig. 1 Evolutionary scenarios tested in the ABC simulations. Abbreviations of regional populations: Alb – Alborz; Ami – Asia Minor; Ape – Apennine Peninsula; Balk – southern Balkans ("*F. moesiaca*"); Cau – Caucasus; CEur – Central Europe; Cri – Crimea ("*F. taurica*"); Eur – Europe (*F. sylvatica* s. str.). Lines: shades of red – *F. sylvatica* s. str.; shades of blue – *F. orientalis*; shades of violet – potential hybrid taxa.

Results

Allelic variation within *F. sylvatica* s. l.

Allele frequency distributions in both subspecies of *F. sylvatica* are shown in Tab. 1. Even though differences between subspecies were sometimes substantial, most alleles

Tab. 1 Allelic frequencies in subspecies of *Fagus sylvatica* s. l. across all regional populations. Alleles are labeled by their relative migration rates.

Locus	Allele	*F. sylvatica* s. str.	*F. orientalis*
Per-A	105	0.35724	0.32377
	100	0.64276	0.67623
Per-B	52	0.00072	0.00103
	39	0.17966	0.49362
	26	0.74256	0.43821
	13	0.07707	0.06598
	10	-	0.00115
Lap-A	106	0.01899	0.01363
	100	0.51255	0.63448
	97	0.30979	0.21413
	94	0.15866	0.13756
	91	-	0.00019
Got-B	54	0.00009	0.00121
	36	0.33814	0.05692
	18	0.66166	0.93887
	6	0.00009	0.00300
Mnr-A	152	-	0.00153
	139	-	0.00035
	126	0.02101	0.23546
	100	0.91569	0.73794
	74	0.00469	0.01704
	63	0.05857	0.00766
	52	0.00004	-
Idh-B	132	-	0.00018
	116	0.28021	0.26672
	100	0.71458	0.73258
	84	0.00517	0.00053
Mdh-A	135	0.00005	-
	130	0.00449	0.02238
	125	0.99496	0.81711
	113	0.00014	0.00693
	101	0.00037	0.15358
Mdh-B	118	0.05825	0.03575
	109	0.00250	0.00105
	100	0.93677	0.92932
	78	0.00195	0.03013
	56	0.00054	0.00375
Mdh-C	22	0.22599	0.01888
	18	0.77388	0.98112
	14	0.00014	-
Pgi-B	126	-	0.00034
	113	0.00792	0.07257
	100	0.97899	0.89924
	87	0.01300	0.02136
	74	0.00008	0.00649

were shared between them. In the cases when an allele was absent in one subspecies, it was found at very low frequency in the other one (always <0.2%). In multiallelic loci, relative migration rates of the respective protein fractions mostly changed by regular steps. Moreover, the more the migration rate differed from that of the most common protein fraction, the lower the respective allele frequency was, with a few exceptions (*Mdh-A*, *Mdh-B*, and partly *Skdh-A*).

Phylogeny simulations

ABC simulations were done under three sets of scenarios. The performance of simulations was quite good, as the observed data were placed fairly well in the center of the point clouds produced from simulated datasets in the case of all three selected scenarios on PCA plots (Fig. S2), and (with two exceptions) summary statistics were always within the interval of 5% to 95% of simulated values (Appendix S1).

The parameter estimates for the most probable simulated scenarios are presented in Tab. 2. In most simulations, the posterior estimates of the mutation rate were around 6×10^{-6}.

The first set of simulations comprised four scenarios, successively testing *F. sylvatica* s. str. and each of the major regional populations of *F. orientalis* as the ancestral lineage from which the others diverged. Scenario 3, in which the Alborz population of beech was treated as the ancestral one, received the best support (posterior probability of 0.63). However, the probability of Scenario 1, supporting the current taxonomical classification of the two subspecies, was also relatively high (0.22), which means that the odds for Scenario 3 compared to this one were approximately 3:1. Divergence times for all major regional populations under Scenario 3 ranged between 11,800 and 18,700 generations, placing the divergence events in the Early or Middle Pleistocene. Effective population size estimates for the Caucasus and Asia Minor populations were approximately threefold of those from the Alborz and for *F. sylvatica* s. str.

Simulations in the second round of scenarios rendered Crimean beech a product of an ancient hybridization between *F. sylvatica* s. str. and a regional population of *F. orientalis*. The geographically more plausible Scenario 4, wherein the second parental population was that in the Caucasus, received higher support (posterior probability of 0.69), but again the probability of Scenario 3 (hybridization with the beech population in Asia Minor) was also relatively high (0.25). The timing of hybridization in Scenario 4 was quite late, 1,440 generations, meaning that the event would have occurred at the beginning of the Holocene, but since the upper limit of the 95% confidence interval was high, 6,400 generations, there was much uncertainty in this estimate. The genetic contribution of *F. sylvatica* s. str. to the gene pools of Crimean beech in Scenario 4 was 30%.

In the last set of simulations, focusing on the origin of Balkan beech, the odds in favor of the best scenario of successive splits were from 16:1 to 74:1 compared to scenarios involving hybridization, so only this scenario was taken into consideration. A hybrid origin of Balkan beech was thus ultimately not supported, and it seemed to rather represent

Tab. 1 Continued

Locus	Allele	*F. sylvatica* s.str.	*F. orientalis*
Pgm-A	112	0.00181	0.02895
	100	0.99508	0.96952
	97	0.00312	0.00154
Skdh-A	114	0.00410	0.00671
	100	0.96727	0.73705
	86	0.01128	0.02426
	72	0.01729	0.23163
	58	0.00005	0.00034

a lineage that diverged from *F. orientalis* in Asia Minor approximately 8,200 generations ago.

Discussion

Suitability of the stepwise mutation model for allozymes

The choice of an appropriate mutation model is an essential issue in model-based phylogeny reconstruction. For protein markers such as allozymes, the infinite allele model of Kimura and Crow [35] is mostly considered suitable. Under this model, each mutation is expected to produce a new allele, while all alleles are considered equally different from each other. However, we used a different model from this, which assumed stepwise mutations, i.e., alleles are ranked by some property of themselves or their products and a mutation can only change a particular allelic state into either of the two possible adjacent states. Stepwise models of mutations are generally preferred in studies employing repetitive sequences, such as minisatellites or microsatellites, among which mutations mostly consist in the addition or deletion of a short sequence or motif and alleles thus differ in their total sequence length (i.e., number of repetitions).

However, the stepwise mutation model was originally developed for electrophoretically detectable variants differing in electric charge [36,37]. Brown et al. [38] found good support for the occurrence of stepwise manner of mutations in the available experimental allozyme data. In our case, the relative migration rates of protein fractions produced by different isozyme alleles differed at quite regular intervals. Even though the migration rate is not necessarily a linear function of charge and depends also on the shape and size of a protein molecule, such regularity is a strong indication of stepwise changes in the net charge [39]. Moreover, the frequency distributions of protein fractions are unimodal for most allozyme loci, which also suggests that mutations mostly change protein charge by one unit and backward mutations may occur. Taking this into account, we consider the mutation model used in our simulations to be untraditional but adequate.

Phylogeny of beech in western Eurasia

Formulating a priori hypotheses about the phylogeny of *F. sylvatica* s. l. is not easy: molecular studies have mostly covered only a part of the range of *F. orientalis* [40,41], and those based on more representative sampling did not yield a clear picture of the relationships among genetic lineages within the species (e.g., the ITS study of Denk et al. [17]).

In earlier studies [22,23] based on the material examined in this one, a much more detailed view of beech phylogeny in terms of the number (24) of small regional populations sampled was attained. However, such a detailed subdivision would result in an enormous number of potential evolutionary scenarios with equal or comparable posterior probabilities. Therefore, the small regions were merged into four geographic areas in the present study: Europe (i.e., *F. sylvatica* s. str.), Asia Minor (including the Ponthic range, Amanus Mountains, and the southeastern tip of the Balkan Peninsula), the Caucasus (including the Greater and the Lesser Caucasus and the Colchis), and the Alborz.

The results of ABC simulations in this study did not differ substantially from the conclusions of earlier studies derived from distance-based (neighbor-joining tree, reticulogram) and model-based (Bayesian) analyses; however, they were not identical to these either. The Structure analysis [24] yielded two clusters within *F. sylvatica* s. l.: one corresponding to *F. sylvatica* s. str., and the other (a more homogeneous one) to *F. orientalis*, with the gene pools of "*F. moesiaca*" and "*F. taurica*" each containing a mixture of samples from both of these clusters [22]. In terms of the evolutionary scenarios tested herein by ABC, this corresponds to an early split of the *F. sylvatica*

Tab. 2 Posterior estimates of the parameters of the demographic inference based on the approximate Bayesian computation for the best-supported scenarios in different constellations of taxa and regional populations of *Fagus sylvatica* s. l.

Parameter	Best-supported scenario		Alternative scenario	
	Mode	95% confidence interval	Mode	95% confidence interval
F. sylvatica s. l.	Scenario 3		Scenario 1	
Probability of scenario	0.6329	0.6004–0.6654	0.2179	0.1879–0.2478
N_e (Alb)	26,500	12,600–63,300	30,200	13,200–84,200
N_e (Cau)	82,700	41,700–98,600	71,300	27,300–97,400
N_e (AMi)	88,000	42,400–98,300	74,100	30,900–97,900
N_e (Eur)	30,100	12,500–90,000	25,300	10,800–57,600
t_1 (divergence Cau from Alb)	16,200 (1.62 My)	6,870–26,300		
t_2 (divergence Eur from AMi)	11,800 (1.18 My)	2,590–25,300		
t_3 (divergence AMi from Alb)	18,700 (1.87 My)	8,270–29,400		
t_1^* (divergence Cau from AMi)			15,200 (1.67 My)	3,580–26,400
t_2^* (divergence Alb from AMi)			18,700 (2.06 My)	6,190–27,700
t_3^* (divergence AMi from Eur)			25,800 (2.84 My)	8,060–29,300
μ	6.46×10^{-6}	$2.48 \times 10^{-7} – 9.92 \times 10^{-6}$	6.35×10^{-6}	$2.88 \times 10^{-6} – 9.91 \times 10^{-6}$
F. sylvatica s. str. – *F. orientalis* – "*F. taurica*"	Scenario 4		Scenario 3	
Probability of scenario	0.6925	0.6431–0.7420	0.2456	0.1989–0.2923
N_e (Alb)	31,700	19,800–69,300	27,700	15,200–62,500
N_e (Cau)	127,000	60,600–148,000	118,000	54,400–147,000
N_e (AMi)	95,200	38,700–146,000	91,900	39,500–145,000
N_e (Cri)	44,000	13,700–95,200	54,100	17,100–94,600
N_e (Eur)	38,600	18,400–89,000	36,600	16,200–87,900
t_1 (hybridization Cau × Eur → Cri)	1,440 (144 ky)	532–6,400		
Hybridization rate (Eur)	0.307	0.0757–0.855		
t_1^* (hybridization AMi × Eur → Cri)			3,070 (307 ky)	820–10,500
Hybridization rate* (Eur)			0.484	0.304–0.953
t_2 (divergence Eur from AMi)	6,740 (674 ky)	2,300–9,760	9,330 (933 ky)	3,640–9,860
t_3 (divergence AMi from Alb)	18,200 (1.82 My)	9,720–27,400	15,600 (1.56 My)	6,990–25,400
t_4 (divergence Cau from Alb)	22,700 (2.27 My)	8,630–29,500	27,300 (2.73 My)	10,500–29,400
μ	5.80×10^{-6}	$1.14 \times 10^{-6} – 9.45 \times 10^{-6}$	7.05×10^{-6}	$1.91 \times 10^{-6} – 9.89 \times 10^{-6}$
F. sylvatica s. str. – *F. orientalis* – "*F. moesiaca*"	Scenario 1			
Probability of scenario	0.9297	0.9153–0.9442		
N_e (AMi)	108,000	51,900–146,000		
N_e (Balk)	46000	14,600–138,000		
N_e (Apen)	131,000	39,900–476,000		
N_e (CEur)	74,900	22,700–145,000		
t_1 (divergence CEur from Balk)	963 (96 ky)	257–7,190		
t_2 (divergence Ape from Balk)	702 (70 ky)	193–5,700		
t_3 (divergence Balk from AMi)	8,170 (817 ky)	3,320–28,300		
μ	1.5×10^{-6}	$2.66 \times 10^{-7} – 4.11 \times 10^{-6}$		

* Parameters of the alternative scenario. t – timing of the event (number of generations before present); μ – mutation rate.

s. str. clade from some of the regional populations of *F. orientalis*, and is also in ac-cordance with morphology-based relationships previously determined among taxa and regional populations [17]. ABC support for such a scenario in this study was not negligible (posterior probability of 0.22), but was still much lower than the support for the scenario in which the easternmost beech population in Iran was the ancestral one. The best-supported scenario placed the divergence of regional populations in the Early Pleistocene, with *F. sylvatica* s. str. being the latest-diverging lineage. This timing corresponds with the history of beech in western Eurasia as reflected in the fossil record [18]. This would mean that barriers to gene flow (probably due to range fragmenta-tion associated with climate fluctuations) appeared earlier within the Asian range of beech than those between Asia Minor and Europe arose. Fossils reveal a permanent presence of the genus *Fagus* in Europe and Western Asia since at least the Oligocene [42], and a wide distribution during the Late Tertiary [16,18,43,44]. These fossils have been described using many different names, but those that were suggested to be linked with *F. sylvatica* s. l. can probably all be subsumed under the name *F. haidingeri* Kováts, which was documented in the Miocene and Pliocene of Europe and Georgia [18]. Although the distribution of *F. sylvatica*-like fossils is not completely identical with the current range of the extant species, their number and geographic coverage suggest that the range of beech in the Pliocene was quite continuous. Fragmentation, which began with the onset of Pleistocene climatic fluctuations, may have disrupted gene flow among regional "islands" and promoted genetic differentiation, while later re-established contacts during Pleistocene warm phases [16] were too short to homogenize the gene pools again. Such a course of events is compatible with both scenarios tested for the *F. sylvatica* s. l. complex in this study. Alternatively, beech may have effectively disappeared from Europe at the end of the Pliocene, with a subsequent recolonization of Europe from Asia Minor, while remnant populations surviving in Europe until the Pleistocene (if any) were overlaid by this recolonization event.

In the case of Crimean beech, ABC simulations confirmed a hybrid origin of "*F. taurica*", with *F. sylvatica* s. str. as one of the parental lineages and the Caucasus/Colchis regional population of *F. orientalis* as the second one. Alternatively, beech in Asia Minor may have participated in the hybridization, as the posterior probability of this scenario (0.25) is too large for it to be ignored. The timing of hybridization (mode of 144 ky BP) was placed at the end of the Saalian glacial period, but the confidence interval is broad enough to include the whole Eemian interglacial (130 to 114 ky BP), and is thus not unrealistic. Beech was certainly present in Crimea during the Pleistocene [45,46], and it was also present (although not dominant) in the forests on the other side of the cur-rent Kerch Strait [47]. A contact between the Crimean and the Caucasian population is thus easy to imagine. The more problematic aspect is a contact with Balkan populations of *F. sylvatica* s. str. *Fagus* pollen was discovered in Late Pleistocene sediments along the northern Black Sea coast [48], but a continuous presence of beech in this area is not sufficiently documented. A certain affinity of Moldovan beech populations to the Crimean ones [23,49] may support the existence of such contact, but this would require a permanent presence of beech in southern Romania throughout the Weichselian glacial period, which cannot be excluded but to date has not been corroborated by the fossil record (see [50]). No final conclusions can thus be made concerning the plausibility of a hybrid origin of Crimean beech.

In contrast, there is no support for a hybrid origin of Balkan beech. The simple divergence scenario without reticulations received the best support (posterior prob-ability of 0.92), and the odds against the next-best scenario were 16:1. This scenario again suggests an early divergence of *F. sylvatica* s. str. from the Turkish lineage of *F. orientalis* occurring during the early Pleistocene (the mode of 817 ky BP falls into the Cromerian interglacial). On the other hand, the divergence of both the Italian and Central-European lineages from the Balkan one was suggested to occur at the begin-ning of the last (Weichselian) glacial. In light of the Quaternary history of vegetation in Europe, such a course of events is plausible: the range of beech expanded and con-tracted along with glacial/interglacial cycles, with a maximum extension during the Holsteinian interglacial [51], and such fluctuations may have contributed to recurrent contacts among local populations and thus have maintained the integrity of gene pools. The expansion during the Eemian was limited to southern Europe, and the onset of the Weichselian glacial may have led to range fragmentation, interruption of gene flow

among refugial areas, and, consequently, formation of the present-day genetic lineages. This view was also supported by the discovery of pollen, the chloroplast haplotype of which was identical to that of the Balkan lineage ("*F. moesiaca*"), in the Venice lagoon from the last pleniglacial [52]; apparently, the formation of the present-day lineages within *F. sylvatica* s. str. may have begun quite late.

Conclusions

The number of scenarios tested in this study was limited, but not primarily for technical reasons: we chose only those that make sense in terms of biogeography, and which were supported by fossil evidence. For instance, the splitting of the European lineage from the Caucasian one was not considered, as it would suggest the migration of beech into Europe along the northern Black Sea coast, and there is not enough fossil evidence to support such a scenario. Taking this limitation into account, this study partly corroborated earlier views on the phylogeny of beech in the western Eurasian area. ABC simulations confirmed that *F. sylvatica* subsp. *orientalis* is a paraphyletic subspecies, as it excludes *F. s.* subsp. *sylvatica*, which diverged from the Asia-Minor lineage. A hybrid origin was confirmed for one transitional taxon (Crimean "*F. taurica*"), but not for the other (Balkan "*F. moesiaca*").

The applicability of the stepwise mutation model to allozyme data is a crucial point in judging whether the above conclusions are realistic. Posterior model checking indicated that both the model and the choice of prior parameters were appropriate for these data in the present study. It seems that the mutation rates of allozymes fit well within the time frame of the processes under study herein. This does not imply that simulations using this model on longer temporal scales (e.g., phylogeny of the family or the genus as a whole) or shorter scales (e.g., postglacial migration) would be equally successful. Nevertheless, since a broad-scale application of allozyme analysis in the 1970s, plenty of data have been gathered in many plant taxa. Our study suggests that a reanalysis of these data, employing new methodological approaches such as ABC and revision of earlier views of phylogeny, may be meaningful.

Supplementary material

The following supplementary material for this article is available at http://pbsociety.org.pl/journals/index.php/asbp/rt/suppFiles/asbp.3582/0:

Fig. S1   Geographical display of the scenarios tested in ABC analyses.

Fig. S2   Model checking for ABC simulations of the chosen scenarios based on principal component analysis (PCA) of summary statistics for: *F. sylvatica* s. l., *F. sylvatica* s. str. – *F. orientalis* – "*F. taurica*", *F. sylvatica* s. str. – *F. orientalis* – "*F. moesiaca*".

Appendix S1   Outcomes of the ABC simulations.

References

1. Tutin TG, Heywood VH, Burges NA, Valentine, DH, Moore DM, editors. Flora Europaea. Volume 1, Psilotaceae to Platanaceae. Cambridge: Cambridge University Press; 1993.

2. Denk T. The taxonomy of *Fagus* in western Eurasia, 1: *Fagus sylvatica* subsp. *orientalis* (= *F. orientalis*). Feddes Repert. 1999;110:177–200. https://doi.org/10.1002/fedr.19991100305

3. Denk T. The taxonomy of *Fagus* in western Eurasia, 2: *Fagus sylvatica* subsp. *sylvatica*. Feddes Repert. 1999;110:381–412. https://doi.org/10.1002/fedr.19991100510

4. Wulff EV. The beech in Crimea, its systematic position and origin. In: Rübel E, editor. Die Buchenwälder Europas. Bern: Hans Huber; 1932. p. 223–261.

5. Гулисашвили [Gulisashvili] ВЗ [VZ], Махатадзе [Mahatadze] ЛБ [LB], Прилипко [Prilipko] ЛИ [LI]. Растительность Кавказа [Rastitel'nosť Kavkaza]. Москва [Moskva]: Наука [Nauka]; 1975.

6. Mobayen S, Tregubov V. La carte de la végétation naturelle de l'Iran. Tehran: Université de Tehran; 1970.

7. Mayer H, Aksoy H. Wälder der Türkei. Stuttgart: Gustav Fischer; 1986.

8. Czeczott H. Distribution of *Fagus orientalis* Lipsky. In: Rübel E, editor. Die Buchenwälder Europas. Bern: Hans Huber; 1932. p. 362–387.

9. Greuter W, Raus T. Med-checklist notulae, 4. Willdenowia. 1981;11:271–280.

10. Poplavskaja GI. Contribution to experimental studies of the systematics of Crimean beech. Trudy Leningradskogo Obshchestva Estestvoispytatelej. 1936;3:353–371.

11. Czeczott H. Studium nad zmiennością liści buków: *F. orientalis* Lipsky, *F. sylvatica* L. i form pośrednich. Cz. I. Rocznik Dendrologiczny. 1933;5:45–121.

12. Mišić V. Varijabilitet i ekologija bukve u Jugoslaviji. Beograd: Biološki Institut N. R. Srbije; 1957.

13. Borza A. Le genre *Fagus* dans la République Populaire Roumaine. Biológia. 1965;20:367–373.

14. von Wühlisch G. EUFORGEN technical guidelines for genetic conservation and use for European beech (*Fagus sylvatica*). Rome: Bioversity International; 2008.

15. Kandemir G, Kaya Z. EUFORGEN technical guidelines for genetic conservation and use of oriental beech (*Fagus orientalis*). Rome: Biodiversity International; 2009.

16. Denk T, Grimm GW, Hemleben V. Patterns of molecular and morphological differentiation in *Fagus* (Fagaceae): phylogenetic implications. Am J Bot. 2005;92:1006–1016. https://doi.org/10.3732/ajb.92.6.1006

17. Denk T, Grimm G, Stögerer K, Langer M, Hembelen V. The evolutionary history of *Fagus* in western Eurasia: the evidence from genes, morphology and the fossil record. Plant Syst Evol. 2002;232:213–236. https://doi.org/10.1007/s006060200044

18. Denk T, Grimm GW. The biogeographic history of beech trees. Rev Palaeobot Palynol. 2009;158:83–100. https://doi.org/10.1016/j.revpalbo.2009.08.007

19. Göker M, Grimm GW. General functions to transform associate data to host data, and their use in phylogenetic inference from sequences with intra-individual variability BMC Evol Biol. 2008;8:86. https://doi.org/10.1186/1471-2148-8-86

20. Grimm GW, Denk T, Hemleben V. Coding of intraspecific nucleotide polymorphisms: a tool to resolve reticulate evolutionary relationships in the ITS of beech trees (*Fagus* L., Fagaceae). Syst Biodivers. 2007;5:291–309. https://doi.org/10.1017/S1477200007002459

21. Renner SS, Grimm GW, Kapli P, Denk T. Species relationships and divergence times in beeches: new insights from the inclusion of 53 young and old fossils in a birth–death clock model. Philos Trans R Soc Lond B Biol Sci. 2016;371:20150135. https://doi.org/10.1098/rstb.2015.0135

22. Gömöry D, Paule L, Vyšný J. Patterns of allozyme variation in western-Eurasian beeches. Bot J Linn Soc. 2007;154:165–174. https://doi.org/10.1111/j.1095-8339.2007.00666.x

23. Gömöry D, Paule L, Reticulate evolution patterns in western Eurasian beeches. Botanica Helvetica. 2010;120(1):63–74. https://doi.org/10.1007/s00035-010-0068-y

24. Pritchard JK, Stephens M, Donnelly P. Inference of population structure from multilocus genotype data. Genetics. 2000;155:945–959.

25. Corander J, Waldmann P, Sillanpää MJ. Bayesian analysis of genetic differentiation between populations. Genetics. 2003;163:367–374.

26. Beaumont MA, Zhang WY, Balding DJ. Approximate Bayesian computation in population genetics. Genetics. 2002;162:2025–2035.

27. Csilléry K, Blum MGB, Gaggiotti OE, Francois O. Approximate Bayesian computation (ABC) in practice. Trends Ecol Evol. 2010;25:410–418. https://doi.org/10.1016/j.tree.2010.04.001

28. Merzeau D, Di Giusto F, Comps B, Thiébaut B, Letouzey J, Cuguen J. The allozyme

variants of beech (*Fagus sylvatica* L.): inheritance and application to a study of the mating system. Silvae Genet. 1989;38:195–201.

29. Müller-Starck G, Starke R. Inheritance of isoenzymes in European beech (*Fagus sylvatica* L.). J Hered. 1993;84:291–296.

30. Cornuet JM, Pudlo P, Veyssier J, Dehne-Garcia A, Gautier M, Leblois R, et al. DIYABC v2.0: a software to make approximate Bayesian computation inferences about population history using single nucleotide polymorphism, DNA sequence and microsatellite data. Bioinformatics. 2014;30:1187–1189. https://doi.org/10.1093/bioinformatics/btt763

31. Voelker RA, Schaffer HE, Mukai T. Spontaneous allozyme mutations in *Drosophila melanogaster*: rate of occurrence and nature of the mutants. Genetics. 1980;94:961–968.

32. Baer CF, Miyamoto MM, Denver DR. Mutation rate variation in multicellular eukaryotes: causes and consequences. Nat Rev Genet. 2007;8:619–631. https://doi.org/10.1038/nrg2158

33. Houston Durrant T, de Rigo D, Caudullo G. *Fagus sylvatica* and other beeches in Europe: distribution, habitat, usage and threats. In: San-Miguel-Ayanz J, de Rigo D, Caudullo G, Houston Durrant T, Mauri A, editors. European atlas of forest tree species. Luxembourg: Publications Office of the European Union; 2016. p. e012b90+.

34. Korpeľ Š. Die Urwälder der Westkarpaten. Jena: Gustav Fischer Verlag; 1995.

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

36. Ohta T, Kimura M. A model of mutation appropriate to estimate the number of electrophoretically detectable alleles in a finite population. Genet Res. 1973;22:201–204. https://doi.org/10.1017/S0016672300012994

37. Barbadilla A, King LM, Lewontin RC. What does electrophoretic variation tell us about protein variation? Mol Biol Evol. 1996;13(2):427–432. https://doi.org/10.1093/oxfordjournals.molbev.a025602

38. Brown AHD, Marshall DR, Weir BS. Current status of the charge state model for protein polymorphism. In: Gibson JB, Oakeshot JG, editors. Genetic studies of *Drosophila* populations. Canberra: Australian National University Press; 1981. p. 15–43.

39. Richardson BJ, Baverstock PR, Adams M. Allozyme electrophoresis: a handbook for animal systematics and population studies. Sydney: Academic Press; 1986.

40. Demesure B, Comps B, Petit R. Phylogeography of the common beech (*Fagus sylvatica* L.) in Europe inferred by restriction studies of PCR amplified chloroplast DNA fragments. Evolution. 1996;50:2515–2520. https://doi.org/10.1111/j.1558-5646.1996.tb03638.x

41. Gailing O, von Wühlisch G. Nuclear markers (AFLPs) and chloroplast microsatellites differ between *Fagus sylvatica* and *F. orientalis*. Silvae Genet. 2004;53:105–110. https://doi.org/10.1515/sg-2004-0019

42. Kvaček Z, Walther H. Revision der mitteleuropäischen tertiären Fagaceen nach blattepidermalen Charakteristiken. IV. Teil *Fagus* Linné. Feddes Repert. 1991;102:471–534. https://doi.org/10.1002/fedr.19911020702

43. Tralau H. Die spättertiären *Fagus*-Arten Europas. Botaniska Notiser. 1962;115:147–177.

44. Denk T. Phylogeny of *Fagus* L. (Fagaceae) based on morphological data. Plant Syst Evol. 2003;240:55–81. https://doi.org/10.1007/s00606-003-0018-x

45. Gerasimenko N. Environmental changes in the Crimean mountains during the Last Interglacial–middle pleniglacial as recorded by pollen and lithopedology. Quat Int. 2007;164–165:207–220. https://doi.org/10.1016/j.quaint.2006.12.018

46. Cordova CE, Gerasimenko NP, Lehman PH, Kliukin AA. Late Pleistocene and Holocene paleoenvironments of Crimea: pollen, soils, geomorphology, and geoarchaeology. In: Buynevich IV, Yanko-Hombach V, Gilbert AS, Martin RE, editors. Geology and geoarchaeology of the Black Sea Region: beyond the flood hypothesis. Boulder, CO: Geological Society of America; 2011. p. 133–164. (Special Papers, Geological Society of America; vol 473). https://doi.org/10.1130/2011.2473(09)

47. Zubakov VA. Climatostratigraphic scheme of the Black-Sea Pleistocene and its correlation with the oxygen isotope scale and glacial events. Quat Res. 1988;29:1–24. https://doi.org/10.1016/0033-5894(88)90067-1

48. Dodonov AE, Zhou LP, Markova AK, Tchepalyga AL, Trubikhin VM, Aleksandrovski AL, et al. Middle-Upper Pleistocene bio-climatic and magnetic records of the northern Black Sea coastal area. Quat Int. 2006;149:44–54.

https://doi.org/10.1016/j.quaint.2005.11.017

49. Gömöry D, Paule L, Shvadchak IM, Popescu F, Sulkowska M, Hynek V, et al. Spatial patterns of the genetic differentiation in European beech (*Fagus sylvatica* L.) at allozyme loci in the Carpathians and the adjacent regions. Silvae Genet. 2003;52:78–83.

50. Magri D, Vendramin GG, Comps B, Dupanloup I, Geburek T, Gömöry D, et al. Palaeobotanical and genetic data outline the Quaternary history of European beech populations. New Phytol. 2006;171:199–222. https://doi.org/10.1111/j.1469-8137.2006.01740.x

51. Peters R. Beech forests. Dordrecht: Springer Science & Business Media; 1997. (Geobotany; vol 24). https://doi.org/10.1007/978-94-015-8794-5

52. Paffetti D, Vettori C, Caramelli D, Vernesi C, Lari M, Paganelli A, et al. Unexpected presence of *Fagus orientalis* complex in Italy as inferred from 45,000-year-old DNA pollen samples from Venice lagoon. BMC Evol Biol. 2007;7(2 suppl):S6. https://doi.org/10.1186/1471-2148-7-S2-S6