Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Back to roots: a phylogeographic assessment of the refugial mosaic of Western Eurasian Fagus L. (Fagaceae) in the Caucasus

Jana Ekhvaia

zhana.ekhvaia@iliauni.edu.ge

Institute of Ecology, 45 Ilia Chavchavadze Avenue, Ilia State University

Tamar Jolokhava

Institute of Ecology, 45 Ilia Chavchavadze Avenue, Ilia State University

Arsen Bakhia

Institute of Ecology, 45 Ilia Chavchavadze Avenue, Ilia State University

Aleksandre Gavashelishvili

Institute of Ecology, 45 Ilia Chavchavadze Avenue, Ilia State University

Marco Cosimo Simeone

University of Tuscia, 4 Via Santa Maria in Gradi

Guido W Grimm

No affiliation

Thomas Denk

Department of Paleobiology, Swedish Museum of Natural History, Box 50007, 10405

Mariam Osepashvili

Institute of Ecology, 45 Ilia Chavchavadze Avenue, Ilia State University

Tornike Pankvelashvili

Institute of Ecology, 45 Ilia Chavchavadze Avenue, Ilia State University

Mariam Kikvidze

Institute of Ecology, 45 Ilia Chavchavadze Avenue, Ilia State University

Natalia Togonidze

Niko Ketskhoveli Institute of Botany, 1 Botanikuri Street, Ilia State University

Khabagin Aliev

Mountain Botanical Garden of the Federal Dagestan Research Center of the RAN, 45 Gadzhieva Street

Oleksiy Yanchukov

Institute of Ecology, 45 Ilia Chavchavadze Avenue, Ilia State University

Tsira Mikautadze-Pantsulaia

Tbilisi Botanical Garden, 4 Sanapiro Street

James R. P. Worth

Department of Forest Molecular Genetics and Biotechnology, Forestry and Forest Products Research Institute

Georgi Fayvush
Department of Conservation of Genetic Resources of Armenian Flora, After Takhtajan Institute of Botany, Armenian National Academy of Sciences

Lilit Khachatryan
Department of Conservation of Genetic Resources of Armenian Flora, After Takhtajan Institute of Botany, Armenian National Academy of Sciences

Annush Nersesyan
Department of Conservation of Genetic Resources of Armenian Flora, After Takhtajan Institute of Botany, Armenian National Academy of Sciences

Sergey Trepet
Shaposhnikov Caucasian State Nature Biosphere Reserve

Research Article

Keywords: Fagus, Chloroplast DNA, Caucasus, Diversity, Evolution, Climate refugia, Isolation-by-Environment

Posted Date: October 23rd, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11207603/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License. Read Full License

Additional Declarations: No competing interests reported.

Abstract

Assessing the refugial history of western Eurasian *Fagus* requires looking back to its roots in the Caucasus – a key evolutionary crossroads and primary refugium. We used a fine-scale, multi-locus plastid sequencing approach across highly dissected mountain terrain spanning the Black and Caspian Sea drainage systems. Analyzing 252 beech populations (446 individuals), we investigated historical isolation, secondary contact, and introgression to reconstruct maternal lineage evolution and demographics. Our analysis revealed robust regional population structure and high divergence between drainage systems. The plastid chronogram confirmed a pre-Pliocene maternal sequence divergence that clearly separates Caucasian *Fagus* lineages from Quaternary-dominated European *F. sylvatica*, likely reflecting a contemporary footprint of Miocene Paratethys vicariance. This demonstrates how ancient cytoplasmic architecture can be preserved across complex landscapes, even as nuclear genomes undergo more recent adaptive radiation and speciation. Our results resolved undocumented Species-Level Plastome Types, including ancient Euxinian lineages linked to European types, a chimeric lineage highlighting historical introgression, and a trans-regional lineage demonstrating a shared Euxinian-Caucasian history. The Caucasus functions as a complex refugial mosaic; the Black Sea drainage exhibited long-term demographic stability as an ancient melting pot, whereas the Caspian Sea drainage underwent massive, recent post-glacial expansion with a founder effect. Driven by plastome sweeps, environmental selection restricted gene flow across sharp altitudinal belts, reflecting chronological divergence histories of distinct plastome types despite narrow secondary contact zones along low-elevation rivers. Chloroplast DNA screening remains vital for identifying and conserving this complex refugial mosaic, which preserved unique maternal lineages for millions of years.

Introduction

Species are dynamic entities, and the divergence and structure of their constituent populations are moulded by both historical and contemporary population dynamics and gene flow (Schaal et al. 1998; Widmer and Lexer 2001; Excoffier et al. 2009; Orsini et al. 2013). Historical fluctuations during the Miocene orogenies, together with increasingly cool and dry Pliocene climatic conditions (Willis et al. 2000; Tiffney and Manchester 2001), followed by Quaternary glaciations (Amarilla et al. 2015), have strongly shaped the population genetic structure of most western Eurasian tree and shrub species, promoting population migrations that often resulted in primary and secondary contact events, triggering hybridization and introgression (Xing and Ree 2017). Typical plant responses included adaptive evolution and range shifts that altered geographic distributions and genetic combinations, as well as the persistence of small populations within glacial refugia (Etterson and Shaw 2001; Petit et al. 2005). A better understanding of the roles played by these processes in shaping the intraspecific genetic diversity, evolutionary history, and biogeography of species and biomes is therefore necessary to inform forest biodiversity conservation in the face of both habitat fragmentation and climate change (Milesi et al. 2024; Leroy and Heuertz 2026; Wang et al. 2026).

In Europe, present-day temperate forest vegetation originated from the classical western refugia of the Iberian, Italian, and Balkan peninsulas (Médail and Diadema 2009; Gavin et al. 2014), and multiple eastern refugia surrounding the Black and Caspian Sea basins (Leroy and Arpe 2007). In this scenario, the Caucasus Mountain system was an important area of eastern refugia likely functioning as a complex refugial mosaic, generating multiple divergent lineages and triggering different expansion patterns across the region (Tarkhnishvili et al. 2012). For instance, in the case of deciduous subtropical to temperate tree species, several genetic studies, revealed unique genotypes and notable intraspecies variability and interspecies divergence in Pontic-Caucasian-Hyrcanian oaks (Denk and Grimm 2010), beeches (Denk et al. 2001, 2024; Cardoni et al. 2022; 2026), maples (Grimm et al. 2007; Grimm and Denk 2014), and zelkovas (Denk and Grimm 2005; Cozzolino et al. 2014), further highlighting genetically distinct endemic species such as *Acer orthocampestre* G.W.Grimm & Denk (a cryptic ancient sister species of *A. campestre* L.), *Quercus castaneifolia* C.A.Mey., (a Hyrcanian isolate of *Q.* subsection *Cerris*), and *Zelkova carpinifolia* (Pall.) K.Koch.

The Caucasus is traditionally viewed as a ~ 800 km wide stretch of land situated between the Black and Caspian Seas, including Armenia, Azerbaijan, Georgia, the North Caucasus region of the Russian Federation (with the republics of Dagestan, Chechnya, Ingushetia, North Ossetia, Kabardino-Balkaria, Karachay-Cherkessia and Adygea, and the Krasnodar and Stavropol administrative regions), northeastern Turkey, and parts of northwestern Iran (Williams et al. 2006). The region lies at the crossroads of different biogeographic regions and serves as a significant biogeographic corridor (Tarkhnishvili et al. 2014). Biogeographically, this complex landscape is fundamentally partitioned into distinct botanical zones: the coastal, humid Euxinian Colchic and Hyrcanian provinces, and the expanded, continental Caucasian province spanning the central and eastern territories of both the Greater and Lesser Caucasus (Takhtajan 1986). Furthermore, the windward slopes of barrier mountains along the Black and Caspian Seas have also retained relatively humid, thermally buffered microclimates since the Neogene. This suggests that the Caucasus served as a shelter for thermophilous relicts during the Quaternary ice ages (Nakhutsrishvili et al. 2013; Jacińca et al. 2015; Ekhvaia et al. 2018; Beridze et al. 2023), facilitating the persistence of two major warm and humid refugia. The Colchic, with the Euxine-Colchic deciduous forest, situated in the Black Sea drainage system and formed mainly since the end of the middle Sarmatian, ca. 11–12 million years ago (Kolakovsky 1961; Shatilova et al. 2011), and the Hyrcanian represented by the Caspian Hyrcanian mixed forests (Nakhutsrishvili et al. 2011), at the southeastern tip of the Caucasus, surrounding the western slopes of the Alborz Mountains and the southern coast of the Caspian Sea, for which a Late Miocene age has been proposed (Safarov 1979). The modern legacy of these ancient refugia, combined with the region’s steep topography, has created sharp environmental gradients. This suggests that spatial genetic structure is often driven not only by physical barriers and geographic distance (Isolation-by-Distance, IBD) but also more strongly by ecological filtering along steep environmental and altitudinal gradients – a process known as Isolation-by-Environment (IBE; Orsini et al. 2013; Wang and Bradburd 2014).

Beeches are the dominant component of montane forest ecosystems in the temperate zones of Europe and Western Asia, where they play a crucial role in the Temperate Broadleaf and Mixed Forests biome

(Olson et al. 2001) and occupy the climax niche in many forest communities (Maycock 1994; Peters 1997; Denk and Grimm 2009). Among western Eurasian broadleaved deciduous trees, the Caucasian beeches are probably the most ecologically important broad-leaved deciduous forest-forming taxa in the region, covering extensive forested areas, especially in Georgia (51%; Dolukhanov 2010) and Azerbaijan (32%; Hasanov et al. 2016). However, the number and classification of *Fagus* taxa have long been debated. Deciphering their evolutionary relationships is like solving a genetic puzzle, as they are all products of complex processes of divergent and reticulate speciation (Schulze and Grimm 2022; Cardoni et al. 2022). These processes involved transcontinental vicariance, incomplete sorting of genetic and morphological traits, and repeated secondary contacts accompanied by interspecies gene flow, all driven by re-radiation and dynamic dispersal. Traditionally, most sources, including Flora Europaea (Tutin et al. 1993), treated all western Eurasian beech populations as a single species, *Fagus sylvatica* L., divided into two subspecies: the European beech (*F. sylvatica* subsp. *sylvatica*), which is distributed throughout most of Europe, and the Oriental beech (*F. sylvatica* subsp. *orientalis*), ranging from the southeastern Balkan Peninsula (along the Black Sea coast of Bulgaria and northeastern Greece) through Turkey and the Caucasus to Iran's Alborz Mountains. This traditional view was challenged by floristic and ecological observations (Gulisashvili et al. 1975; Mayer et al. 1986) and over the last two decades, an increasing number of molecular studies (Renner et al. 2016; Gömöry et al. 2018) culminating in the formalization of four species based on morphological and molecular evidence: the European *F. sylvatica* L., the chiefly northeastern Greek to northwestern Turkish *F. orientalis* Lipsky, the Caucasian *F. hohenackeriana* Palib., and the Hyrcanian *F. caspica* Denk & G.W.Grimm (Denk et al. 2024).

Two recent studies sketched the main patterns of genetic diversity of Caucasian beech at the nuclear genome level. Thirty-two natural populations of *F. hohenackeriana* and likely *F. caspica* ("*F. orientalis*") from Georgia, Armenia, and Azerbaijan, investigated with nuclear microsatellites (Sękiewicz et al. 2022), revealed a west-east gradient of genetic differentiation strictly corresponding to the Colchis and Hyrcanian glacial refugia, suggesting that the species' post-glacial recolonization history is the most important driver of its current genetic diversity. Capblanq et al. (2024) also identified two genetic clusters corresponding to the same glacial refugia (Colchis in the west and Hyrcanian in the east) through whole-genome sequencing of 165 beech individuals (33 populations), with a secondary contact zone in the eastern Greater Caucasus.

The plastid genome is a key marker in angiosperm phylogenetics (Ruhfel et al. 2014; Shaw et al. 2014). Because it is haploid, predominantly maternally inherited, non-recombinant, and has a relatively low effective population size, chloroplast DNA (cpDNA) is particularly susceptible to haplotype fixation through genetic drift (Pham et al. 2017). Seed-mediated dispersal is generally more restricted than pollen- and seed-mediated nuclear gene flow, allowing cpDNA to retain strong phylogeographic signals and making it especially valuable for tracing maternal lineages, identifying glacial refugia, and reconstructing postglacial colonization routes (Petit and Excoffier 2009). In European *Fagus*, cpDNA markers have played an important role in elucidating spatial genetic structure and historical migration patterns, as supported by the integration of palaeobotanical and genetic evidence (e.g., Magri et al. 2006). Recently, Worth et al. (2025, work in progress), using a whole-plastome phylogeny integrated with

nuclear genomic data and fossil calibrations, revealed five deeply divergent chloroplast lineages within the genus, appearing poorly sorted among East Asian species, and including an exclusive western Eurasian lineage ('Lineage V') formed by two so-called "Species-Level Plastome Types" (SLPTs): one shared by all currently available *F. sylvatica* complete plastomes (V-Sy) markedly different to that of *F. caspica* (V-EO; no data available on *F. hohenackeriana* and *F. orientalis*). Three further SLPTs, two unique to *F. hohenackeriana* and one to *F. caspica* from the Talysh Mountains, were subsequently detected in an extended western Eurasian plastome dataset extracted from the total DNA data of Capblancq et al. (2024) (Worth et al., 2025, work in progress; cf. Figure 1). Gaining momentum from this evolutionary and phylogeographic background, the present study aims to improve our understanding of the phylogeography and genetic structure of western Eurasian *Fagus* distributed across the Caucasus and adjacent regions. Specifically, we expanded the number of species and populations analyzed to document all relevant taxonomic and regional diversity and used cpDNA sequence analyses to integrate previous assessments. Our objectives were to: (i) investigate the relative contributions of historical isolation, secondary contact, and introgression in western Eurasian maternal lineages; (ii) reconstruct the evolutionary history of western Eurasian *Fagus* and assess the signatures of past demographic events preserved in contemporary patterns of genetic diversity; (iii) evaluate the role of the Caucasus as a long-term center of persistence, evolutionary diversification, and lineage differentiation in *Fagus*.

Material and Methods

Sampling material and molecular analysis

A total of 252 beech populations from four (notho-)species, including *F. orientalis* s.str. Lipsky, *F. hohenackeriana* Palib., *F. caspica* Denk & G.W.Grimm, and *F.* x *taurica* Popl., were collected across the Caucasus (spanning the Georgian, Armenian, Azerbaijani, Turkish, and Russian Federation parts of the Greater and Lesser Caucasus), the Crimean Peninsula, and the Black Sea and Caspian Sea coastal regions of Georgia, the Russian Federation, Turkey, Azerbaijan, and Iran (Fig. 2). Additionally, data for 25 European beech (*F. sylvatica* L.) were extracted from the complete plastome sequence dataset of Worth et al. (2025, work in progress) and used as an outgroup. Data on four *F. caspica* and nine *F. hohenackeriana* individuals included in the emended complete plastome data were added for comparison and correlation with SLPTs (Species-Level Plastome types) identified by Worth et al. (2025, work in progress; cf. Figure 1). Detailed information for the newly studied and added samples is provided in Supplementary File S1 (worksheet "Dataset").

Five trees (> 100 meters apart) were sampled in each population. One to three individuals per population —usually the first, third, and fifth trees—were selected for DNA extraction, PCR amplification, and sequencing. In populations showing multiple haplotypes or a unique haplotype restricted to that population ('private' haplotype), all five samples were analyzed to capture full genetic diversity. Total genomic DNA was extracted from silica-gel-preserved leaf tissues of 446 trees using the DNeasy Plant MiniKit (Qiagen) according to the manufacturer's instructions. Based on mutational patterns observed in their Western Eurasian *Fagus* samples (two identical *F. caspica* plastomes, *F. sylvatica* plastomes

covering the species' range), two plastid intergenic spacers (*trnK-trnQ* and *petA-psbJ*) were selected from a whole plastome alignment (Worth et al. 2025, work in progress). Primers for these cpDNA loci were designed using PrimerBLAST (https://www.ncbi.nlm.nih.gov/tools/primer-blast/). Since the expected length of *trnK-trnQ* exceeded 2 kb, it was split into two overlapping fragments, which were amplified as separate markers. Detailed information on selected plastid primers and PCR amplification protocols is provided in Supplementary File S1 (worksheet "PCR primers and conditions"). PCR product purification and sequencing were performed at BGI Genomics (Hong Kong, China). Electropherograms were edited with Chromas ver. 2.3 (http://www.technelysium.com.au) and checked visually. Multiple sequence alignments (MSAs) were generated with MEGA ver. 12.0 (Kumar et al. 2024). In the final MSA, all indel events were recoded as single binary characters, whereas indels with internal nucleotide variation were represented as ternary characters to capture the additional mutational state (Supplementary File S1, worksheet "Mutations").

Molecular data analysis

Sampling material was divided into the two main catchments: the Black Sea drainage basin, which includes populations distributed across the gorges of key river systems, including the Ajaristskali, Bziphi, Chorokhi, Enguri, Filyos, Kodori, Kuban, Machakhela, Natanebi, Rioni, and Uzen rivers, and the Caspian Sea drainage basin, covering populations from the gorges of the Alazani, Aragvi, Aras, Dzhalganchai, Gamri-Ozen, Gorganrud, Iori, Lenkeran, Mtkvari, Samur, Sulak, and Terek rivers (Supplementary File S1, worksheet "Dataset").

DnaSP ver. 6 (Rozas et al. 2017) was used to calculate the main haplotype and nucleotide diversity parameters and standard deviations (*SD*) for the two joined cpDNA IGS regions across the whole dataset and separately for the Black Sea and Caspian Sea drainage systems, including the total number of haplotypes (*H*), haplotype (gene) diversity (*Hd*), nucleotide diversity (π), the number of singleton variants (*SVs*), the average number of nucleotide differences (*k*), and the stochastic, sampling, and total variances of *k*. Further, to investigate signals of demographic dynamics of *Fagus* populations in the Black Sea and Caspian Sea drainage systems, mismatch distribution analysis was used to compare observed distributions with expectations under a model of sudden population expansion, and graphs were generated to show distribution patterns across drainage basins. The smoothness of the model was evaluated using the raggedness statistics (*r*; Harpending 1994) and the more powerful Ramos-Onsins and Rozas's R^2 statistics (Ramos-Onsins and Rozas 2003) across 1000 parametric bootstrap replicates. Discrimination of a mutation/drift equilibrium and evaluation of the hypothesis of population expansion through the significant excess of low-frequency haplotypes were assessed with Fu and Li's (Fu and Li 1993) *D* and *F* neutrality test statistics. Analysis of genetic differentiation and gene flow between drainage basins was also evaluated with DnaSP using the χ^2 test (Hudson et al 1992a) with the degrees of freedom (*df*) equal to 13 for a total of 14 haplotypes, haplotype-based estimates between populations within (*Hs*) and among basins (*Hst*), sequence-based estimates showing the proportion of nucleotide divergence between populations (*Kst*, *Kst**), and rank-based estimates via comparison of average

difference within *vs* between populations among basins (*Z*, *Z**) together with the nearest-neighbor statistic (*Snn*). Gene flow was assessed using indirect estimates of the effective number of migrant chloroplasts per generation (*Nm*), and coefficients of gene differentiation between drainage systems relative to total genetic variation were computed in both haplotype-based (*Gst*) and sequence-based (*γst*) analyses, as well as distance-based approaches (*Fst*).

To estimate maternal lineage divergence and visualize relationships among cpDNA haplotypes, an unrooted haplotype network was initially created using the median-joining algorithm in NETWORK ver. 10.2 (Bandelt et al. 1999). This network served as the starting topology for reconstructing a weighted (quasi-)full median network, as described by Bandelt et al. (2000), identifying ancestral states of SNPs and length polymorphisms and subsequently the direction of point mutations (transitions, transversions), indels (insertions, deletions, duplications, inversions; each one treated as a single mutational events), and polynucleotide length polymorphisms (PNLPs) on the background of the emended complete plastome phylogeny of Worth et al. (cf. Figure 1; 2025, work in progress). PNLPs were given lower weight, being more susceptible to convergence, backmutations, and incomplete phylogenetic sorting (Supplementary File S1, worksheets “Pivot Summary” and “Mutations”). Among the fourteen haplotypes obtained, H01 showed a chimeric (mosaic) structure and was excluded from the divergence-time analysis because it combines clearly distinct sequence patterns characteristic of two different plastome lineages; its placement within a strictly bifurcating phylogenetic framework would have thus resulted in incorrect or ambiguous placements. However, the H01 haplotype was retained in the median network to illustrate its intermediate position and potential role in historical plastome exchange between the lineages.

Bayesian molecular clock analysis was performed with BEAST ver. 2.7.7 (Bouckaert et al. 2014). To build a time-calibrated phylogeny (chronogram), we used a single fossil calibration for the root age, treating the European *F. sylvatica* as the outgroup. The root age was constrained to ≥ 28 Ma based on the earliest fossil cupules and nuts of *F. castaneifolia* Unger, considered an ancestor of the western Eurasian *Fagus* lineage (Denk and Grimm 2009; Worth et al. 2025, work in progress). A log-normal distribution prior was chosen for this calibration node because it is the most commonly used for fossil-based minimum age constraints (Ho and Phillips 2009). To simplify computation and take into account the complete plastome phylogeny (cf. Figure 1), in addition to the fossil-calibrated age prior, three topological priors were used: (*i*) a clade comprising individuals of Caucasian *F. hohenackeriana* - Crimean *F.* x *taurica* complex (H06-H10), and Hyrcanian *F. caspica* (H04-05); (*ii*) Caucasian *F. hohenackeriana* - Crimean *F.* x *taurica* complex (H06-H10), and Hyrcanian *F. caspica* from the Talysh Nature Reserve (H04); and (*iii*) Caucasian *F. hohenackeriana* - Crimean *F.* x *taurica* complex (H06-H10). The HKY + I substitution model (Hasegawa et al. 1985) was selected in MEGA ver. 12.0 based on the Akaike Information Criterion (AIC; Possada and Crandall 1998), which accounts for non-variable positions and improves estimates of branch lengths and divergence times. Due to the limited number of cpDNA loci and sequence variation, a strict molecular clock was employed, as it is more appropriate and statistically stable for datasets with limited phylogenetic signal (Drummond et al. 2006; Ho and Duchêne 2014). Additionally, a Birth-Death speciation process (Gernhard 2008) was used to model lineage diversification within western Eurasian

Fagus. The Markov Chain Monte Carlo (MCMC) was run for 100 million generations, sampling every 1000th generation, resulting in 100 000 trees, of which 10 000 trees were kept using a burn-in of 90% to generate a lineage maximum- credibility tree with median heights in TreeAnnotator ver. 1.8.2. BEAST log files were checked in Tracer v1.6 (http://tree.bio.ed.ac.uk/software/tracer) to ensure adequate sampling and convergence to stationarity (Supplementary File S1, worksheet “BEAST Trace log”). The chronograms were visualized in FigTree ver. 1.4.2 (http://tree.bio.ed.ac.uk/software/figtree). Moreover, to evaluate the mosaic structure of haplotype H01 (which shows contrasting phylogenetic signals across the *trnK-trnQ* IGS region), two strict-clock chronograms were constructed: one based on 13 of the 14 haplotypes (excluding H01) and another including all 14 haplotypes. In the latter, H01 was partitioned into 5' and 3' regions, which were represented as distinct operational taxonomic units (OTUs; ‘tips’), each with its own independent priors, within a single BEAST analysis. Missing regions were coded as missing data ('?'), allowing each sequence partition to be placed independently in the phylogeny.

Additionally, to address inflated 95% Highest Posterior Density (HPD) intervals for internal nodes in the BEAST chronogram from the initial Bayesian MCMC analysis, which indicated significant uncertainty in the divergence time estimates due to limited polymorphic sites and reliance on a single fossil calibration point, we applied a semiparametric Penalized Likelihood (PL) method (Sanderson 2002). To achieve this, the maximum clade credibility (MCC) tree from TreeAnnotator ver. 1.8.2 was analyzed in *R* ver. 4.6.1 using the packages “treeio” (Wang et al. 2020) and “ape” (Paradis 2013). The original Bayesian topology and posterior probabilities were strictly retained and mapped onto the nodes, while the branch lengths were re-estimated in absolute time units using the “chronos” function. To conduct the divergence dating, we used a single reliable fossil calibration solely to the root node (defined as the MRCA of Sy1 and H11 from the first chronogram, and the MRCA of Sy1 and H12 for the second chronogram) within a strict interval of 28.52 to 30.18 Ma, avoiding secondary calibrations or internal node priors from the BEAST analysis. We used the "correlated" rate model in 'chronos' to evaluate rate variation across branches. The smoothing parameter (lambda) was optimized through cross-validation to balance the parametric clock and non-parametric rate-smoothing models. This method used the sequence alignment length (2487 bp) and a fixed root constraint to effectively smooth and resolve internal-node ages, addressing over-parameterization in fully Bayesian methods under data-limited conditions.

Ecological data analysis

Differences among environments based on forest-type composition were quantified using *R* ver. 4.5.2. The “ggplot2” package was used to assess ecological differentiation among species-level plastome types (SLPTs) along the altitudinal and annual precipitation gradients. A nonparametric analysis was conducted using chi-square (χ^2) and Kruskal-Wallis (*K-W*) tests, followed by a median test to evaluate shifts in central tendency relative to the overall grand median. Altitudinal (m a.s.l.) and annual precipitation (mm) data for each sampled population were extracted from the georeferenced CHELSA bioclimatic datasets (1981–2010, ver. 2.1, https://www.chelsa-climate.org/datasets/chelsa_bioclim) using study-site coordinates in QGIS 3.4 (Supplementary File S1, worksheet “Samples Stands for

Mapping"). This approach has higher accuracy under mountain-specific conditions (Brown et al. 2018) and allowed us to reduce the effects of the complex Caucasian topography on microclimate. The distributions of environmental variables were visualized through violin plots, accompanied by a summary table that included the number of populations, mean values (*± SD*), ranges (*min–max*), and medians for each SLPT, illustrating distributional patterns, overlap, and ecological amplitude among plastome types. Differences among plastome types along ecological gradients of altitude, annual precipitation, and forest type were further examined using principal coordinates analysis (PCoA) based on the Bray-Curtis dissimilarity metric, implemented in the "vegan" package. Graphical representations of forest-type composition among SLPTs were generated using the "ggplot2" and "scales" packages in *R* ver. 4.5.2. Vegetation data were obtained from phytosociological surveys following the Braun-Blanquet (1964) method. The resulting syntaxa were subsequently grouped into broader forest-type categories for ecological analyses, including Colchic humid and hyper-humid forests, low-mid beech and beech-hornbeam forests, montane beech and beech-hornbeam forests, montane conifer forests, and secondary and thermophilous forests (detailed information provided in Supplementary File S1, worksheet "Ecological Profile Data"), to facilitate comparisons among plastome types. Ecological data analysis was limited to natural stands of *F. hohenackeriana* totaling 229 stands where phytosociological surveys were conducted.

Results

Genetic diversity, demographic patterns and population structure in Fagus populations

The final dataset comprised two chloroplast intergenic spacer (IGS) regions, *trnK–trnQ* and *petA–psbJ*, from 468 western Eurasian *Fagus* individuals (446 individuals from the Caucasus and adjacent regions, and 25 European *F. sylvatica* sequences). The Caucasus and adjacent regions included data for individuals from the Greater and Lesser Caucasus (409), Northwestern Anatolia (8), the Crimean Peninsula (6), and the Hyrcanian forests (23), all of which are part of the Black and Caspian Sea drainage systems (Fig. 2; Supplementary File S1, worksheet "Dataset").

The multiple alignment of the two concatenated loci resulted in a matrix of 2610 characters, reduced to 2487 bp after coding indels as single characters. Nucleotide and haplotype diversity statistics were estimated for the total dataset (N = 446 individuals/252 populations) and separately for the Black Sea (N = 260 individuals/142 populations) and Caspian Sea (N = 186 individuals/110 populations) drainage systems (Table 1). A total of 14 haplotypes were identified across both drainage systems, yielding high overall haplotype diversity (*Hd* = 0.804) and moderate nucleotide diversity (π = 0.0017), with a mean number of pairwise nucleotide differences (*k*) of 3.985. Within the two basins, the Black Sea drainage system, with 10 haplotypes, showed higher genetic variability and variance than the Caspian Sea drainage system, with 9 haplotypes (Table 1).

Table 1
Chloroplast DNA (cpDNA) sequence nucleotide and haplotype diversity parameters of *Fagus* populations from the Caucasus and adjacent regions for the Black Sea and Caspian Sea drainage systems. *H* - number of haplotypes, *Hd* - haplotype (gene) diversity, *π* - nucleotide diversity, *SVs* - number of singleton variants, *k*– average number of nucleotide differences, *Vst*(*k*)– stochastic variance of *k* under the no-recombination model, *Vs*(*k*)– total variance of *k*, *V*(*k*) – stochastic variance of *k*.

Nucleotide and Haplotype variation parameters	**Black Sea drainage system (# of individuals/ populations) 260/142**	**Caspian Sea drainage system (# of individuals/ populations) 186/110**	**Total (# of individuals/ populations) 446/252**
H	10	9	14
Hd	0.752 ± 0.016	0.585 ± 0.042	0.804 ± 0.009
π	0.0019 ± 0.0000	0.0009 ± 0.0000	0.0017 ± 0.0000
SVs	0	6	6
k	4.54	2.078	3.985
Vst(k)	4.98	1.349	3.97
Vs(k)	0.039	0.016	0.018
V(k)	5.018	1.365	3.989

In terms of haplotype composition, the Black Sea drainage system was dominated by haplotypes H02 (16.15%), H06 (27.31%), and H08 (38.09%), whereas the Caspian Sea drainage system was characterized by the predominance of haplotype H14, which accounted for 62.9% of all samples (Fig. 2). The two basins also differed significantly in their unique haplotypes' composition: H01, H03, H07, H11, and H12 were restricted to the Black Sea basin, whereas H04, H05, H09 and H13 were exclusively found in the Caspian Sea basin. Additionally, five haplotypes were shared between the both drainage systems: H02 predominated in the Black Sea system, with only a single individual crossing into the Caspian drainage via the Mtkvari River main system in the Lesser Caucasus adjacent to the Zekari Range near Ajara in western Georgia; H06 primarily covered river gorges within the Black Sea basin but extended into the Caspian Sea basin along the slopes of the central-eastern and northern Greater Caucasus (representing 17 individuals out of 88); H08 also predominated in the Blacks Sea drainage system with 5 Caspian individuals recorded on the eastern and northern slopes of the Greater Caucasus; H10 belonged mostly to Caspian Sea drainage system (11 individuals) but converged into the Black Sea basin (6 individuals) within the central-western Greater Caucasus; and H14, though highly dominant in the Caspian Sea drainage system, was found in the Black Sea drainage system for 14 out of 202 individuals, localized along the slopes of the western Lesser Caucasus and the central-eastern Greater Caucasus (Fig. 2; Supplementary File S1, worksheet "Dataset"). Furthermore, we identified stands with diverse internal

haplotypes across both drainage systems, confirming the presence of secondary contact zones. This pattern occurred in the Greater Caucasus populations, specifically in Imereti and Racha-Lechkhumi (central-western Georgia, the Black Sea basin), Kakheti (eastern Georgia, the Caspian Sea Basin), Ingushetia and Dagestan (the Russian Federation, the Caspian Sea basin), and the Lesser Caucasus population in Guria (south-western Georgia; the Black Sea Basin). For example, Racha-Lechkhumi-Kvemo Svaneti harbored three diverse stands with H06/H08, H08/H10, and H08/H14 haplotype combinations. Single diverse stands were recorded in Imereti (H08/H14) and in Guria (H02/H08), while Kakheti contained two stands (H06/H14 and H08/H14). In Dagestan and Ingushetia, isolated stands were also found exhibiting H10/H14 and H08/H10 haplotype combinations (Fig. 2; Supplementary File S1, worksheet “Dataset”).

Mismatch distribution analysis showed contrasting patterns between the two drainage systems. The Caspian Sea drainage system shows lower genetic diversity and a smoother, unimodal distribution than the Black Sea drainage system (Fig. 3a). The observed distribution showed a dominant low-divergence peak and reduced variance in pairwise differences, closely approaching the expected value under demographic expansion (Table 2). Similarly, the observed coefficient of variation was moderately higher than expected. Furthermore, the low Ramos–Onsins and Rozas R^2 statistic, together with significantly negative Fu and Li’s neutrality test values, support the hypothesis of recent demographic expansion and reduced phylogeographic structure. In contrast, the Black Sea drainage system population exhibited a clearly ragged, multimodal mismatch distribution (Fig. 3b). The observed variance of pairwise differences was much higher than expected under the expansion model (Table 2). The observed coefficient of variation was also nearly twice the expected value. The relatively high Ramos–Onsins and Rozas R^2 statistic, together with significantly positive Fu and Li’s test statistics, suggested that the mismatch distribution is more consistent with historical population structure, admixture, or the survival of distinct maternal lineages in multiple refugial areas. Additionally, these results were accompanied by an absence of singleton variants (*SVs*) in the Black Sea drainage system (Table 2) and, in contrast, an excess of *SVs* in the Caspian Sea drainage system (6 *SVs*), indicating different demographic histories for the two drainage systems

Table 2
Summary of demographic parameters inferred from observed and expected mismatch distribution analysis and neutrality tests based on cpDNA sequence nucleotide and haplotype variation observed in *Fagus* populations from the Caucasus and adjacent regions for Black Sea and Caspian Sea drainage systems.

Nucleotide and haplotype variation parameters	**Caspian Sea drainage system**	**Black Sea drainage system**
Observed variance of pairwise differences k	4.01	17.69
Observed coefficient of variation (CV) of k	0.96	0.93
Expected variance of k under no-recombination model	1.36	5.02
Expected CV of k under no-recombination model	0.56	0.44
Raggedness statistic, r	0.26	0.23
Ramos-Onsins and Rozas R^2 statistics	0.08	0.14
Fu and Li's F	-3.46, $p \leq 0.02$	1.6, $p \leq 0.05$
Fu and Li's D	-2.76, $p \leq 0.05$	2.04, $p \leq 0.02$

Genetic differentiation between drainage systems was highly significant across all estimates and strongly supported by a chi-square test ($\chi^2 = 27.16$, $df = 13$, $p < 0.001$). Haplotype-based statistics showed significantly higher haplotype diversity between populations within drainage systems (Table 3) and lower, still significant, diversity among them, with up to 15% of haplotype diversity that might be explained by differences between the Black and Caspian Sea drainage systems. Further, both sequence-based and rank-based statistics were also statistically significant (Table 3) and showed pronounced separation between populations in sequence space. Additionally, the nearest-neighbor estimates (*Snn* = 0.82, $p < 0.001$) confirmed the strong geographic structuring and genetic divergence (Hudson 2000). Gene flow analysis using both haplotype and sequence data based on indirect estimates (Whitlock and McCauley 1999) of the effective number of migrant chloroplasts per generation (*Nm*) ranged from 1.83 to 3.99 (Table 3). Coefficients of gene differentiation between drainage systems relative to total genetic variation in both haplotype- and sequence-based analyses were low (Table 3), with relatively high gene diversity, indicating that most genetic diversity is found within drainage systems rather than between them. Distance-based approaches yielded higher differentiation estimates (*Fst*) and a lower number of migrant chloroplasts per generation (*Nm*), indicating moderate historical gene flow over evolutionary time.

Table 3
Estimates of genetic structure and gene flow in *Fagus* populations from the Caucasus and adjacent regions for the Black Sea and Caspian Sea drainage systems. *Hs*– haplotype diversity within populations, *Hst*– proportion of total haplotype diversity due to differences between populations among basins, *Kst*– proportion of nucleotide divergence between populations, *Kst**– corrected version of *Kst*, *Z*– average divergence within *vs* between populations among basins, *Z**– corrected version of *Z*, *Snn*– nearest-neighbor statistic, *Gst* – genetic differentiation among populations based on haplotype frequencies, *γst*– genetic differentiation among populations based on nucleotide sequence data, *Fst*– sequence-level differentiation between basins based on genetic distances between haplotypes, *Nm*– effective number of migrant chloroplasts per generation.

Haplotype-based estimates	
Hs	0.69, *p < 0.0001*
Hst	0.15, *p < 0.001*
Sequence-based estimates	
Kst	0.11, *p < 0.001*
Kst *	0.14, *p < 0.001*
Rank-based estimates	
Z	41337.89, *p < 0.001*
Z*	10.3, *p < 0.001*
Snn	0.82, *p < 0.001*
Gene flow estimates	
Gst/Nm	0.15/2.86
γstNm	0.11/3.99

Haplotype-based estimates	
Fst/Nm	0.21/1.83

Phylogeography of Fagus in the Caucasus and adjacent regions

To estimate maternal lineage divergence and visualize relationships among cpDNA haplotypes of western Eurasian *Fagus*, including European *F. sylvatica*, we constructed a weighted (quasi-)full median network (Fig. 4) following Bandelt et al. (2000), and conducted a Bayesian strict-clock analysis followed by divergence time estimations modified using the Penalized Likelihood method (Figs. 5, 6).

A clear phylogeographic structure emerged among the main *Fagus* Species-Level Plastome Types (SLPTs; as defined in Worth et al. 2025, work in progress) in the Caucasus and adjacent regions. Our data allowed us to map the distribution of the previously identified SLPTs within the eastern sublineage of Lineage V: '(V-)HoA,' 'HoB,' 'CaA,' and 'CaB' (Fig. 5; Table 4; cf Fig. 1). In addition, we found several potentially new SLPTs with different phylogenetic affinities (Figs. 4, 5) and distribution areas (Fig. 2, 5). The potential SLPT '(V-) Or' differs from all other beeches by two unique mutational patterns, one PNLP and one transition (Fig. 4, Supplementary File S1, worksheet "Mutations"). It is restricted to the studied Western Pontic populations of *F. orientalis* and resolved as an early-diverged (Early Miocene; ≥ 18 Ma) sister type of the eastern SLPTs (Fig. 5). The plastome types 'EuA' and 'EuB' form an Euxinian lineage closer to the SLPT of *F. sylvatica* than any other haplotype found in the Black Sea and Caspian Sea drainage systems (a possible 'Sy' sister type). Within this group, 'EuB' has drifted further from the common ancestor than 'EuA'. In our penalized-likelihood strict clock-rooted-dated tree (Fig. 5), 'EuA' and 'EuB' are resolved as a Miocene-diverged (≥ 14 Ma) sister lineage to the *F. sylvatica* (Sy) plastomes. The 'HoX' is recognized here as a potentially new SLPT within the 'CaB-HoB' sublineage. It consistently lacks mutations diagnostic of either 'CaB' or 'HoB', while showing a genus-unique transversion (Fig. 4); it may have been isolated during the Late Miocene (≥ 6 Ma).

Additionally, Euxinian SLPT 'Eu×' shows a chimeric *trnK-trnQ* sequence, with the 5' part showing mutations unique to 'EuB' (and 'EuA') but a 3' part indistinguishable from the (reconstructed) ancestral sequence of the 'HoB(-HoX)-CaA-CaB' sublineage (Fig. 6).

Within the focal area of our study, the *F. hohenackeriana* SLPT 'HoA' is predominantly associated with the key river gorges of the Caspian Sea drainage system throughout the central and eastern parts of the Greater Caucasus and the Lesser Caucasus. SLPT 'HoB' is chiefly Caucasian and brackets 'HoA' to the west and east while overlapping with it locally in the central Greater Caucasus and central Georgia. Within 'HoB', the ancestral and most common haplotype H08 is predominant in the western-central Greater Caucasus, centered on the Colchis-facing slopes, and extending locally onto the adjacent northern slopes of key river gorges in the Black Sea drainage system, while its satellite haplotypes H09 and H10 are more peripheral. Westwards, 'HoB' is subsequently replaced by its potential sister SLPT 'HoX'. 'HoX' is distributed across the northern and southern slopes of the central and western Greater Caucasus, extending from the far west through the central regions into the interior Caucasian highlands,

and, notably, also occurs in *F.* × *taurica* from the Crimean Mountains, indicating a broader Euxinian-Caucasian distribution. In contrast, ‘EuA’, ‘EuB’ (and ‘Eu×’) are restricted to the Ajaristskali, Chorokhi, Machakkhela, and Natanebi river system. Within *F. caspica* plastomes, a clear split is seen between the Talysh Mountains (‘CaB’, more closely related to ‘HoB’ and ‘HoX’) and the eastern Alborz Mountains (‘CaA’; Table 4; Fig. 2).

Table 4
Composition and distribution of earlier (Worth et al., 2025) and newly found Species-Level Plastome Types (SLPTs) in the Caucasus and adjacent regions, sorted by estimated divergence ages. Abbreviations: *Ma*- million years, *Lat* – Latitude, *Long* – Longitude, *BS* – Black Sea drainage system, *CS* – Caspian Sea drainage system.

SLPTs	**Age (Ma)**	**Comprised haplotypes**	**Distribution range:** **Lat/Long (in decimals)/** **Altitude (m a.s.l.)**	**Mountain Chain**	**Main river system/Sea Drainage System: BS, CS** **(# of individuals)**
Lineage Va					
‘Or’	18.48	H11, H12	41.0466–41.4974 N/ 31.6644-31.71739 E/ 162–1246	Western Pontides	Filyos/BS (8)
‘HoA’	14.27	H13, H14	39.06506–43.03033 N/ 42.76781-48.13265E/ 401–1851	Greater Caucasus, Lesser Caucasus	Alazani-Iori/CS (35); Aragvi/CS (12); Aras/CS (4); Dzhalganchai/CS (3); Gamri-Ozen/CS (3); Mtkvari/CS (56); Rioni/BS (12); Samur/CS (3); Sulak/CS (6); Terek/CS (3)
‘CaA’	9.38	H05	36.70711–36.683 N/ 54.111131 E/ 740, 1606	Eastern Alborz	Gorganrud/CS (7)
‘CaB’	7.84	H04	38.63846-38.75741N/ 48.37886-48.69906E/ 229–1432	Western Alborz (Talysh Mountains)	Lenkeran/CS (16)
‘HoB’	6.48	H08- H10	41.68429-43.22432N/ 41.83449-45.08444E/ 91-2170	Greater Caucasus, Lesser Caucasus	Alazani/CS (6); Enguri/BS (23); Natanebi/BS (18); Kintrishi/BS (14), Rioni/BS (42); Terek/CS (1)
‘HoX’	6.48	H06, H07	41.35851–44.82532 N/	Central Pontides, Crimean	Alazani-Iori/CS (3); Aragvi/CS (3), Bziphi/BC (6); Kodori/BC (5);

SLPTs	Age (Ma)	Comprised haplotypes	Distribution range: Lat/Long (in decimals)/ Altitude (m a.s.l.)	Mountain Chain	Main river system/Sea Drainage System: BS, CS (# of individuals)
Lineage Va					
			34.92765-45.60841E/ 193–1968	Mountains, Greater Caucasus, Lesser Caucasus,	Kuban/BS (39); Mtkvari/CS (8); Rioni/BS (17); Terek/BS (3); Ulu-Uzen/BS (6)
Lineage Vb					
'EuA + EuB+Eu×'	14.94	H01, H02, H03	41.35403–41.87864 N/ 41.50796–42.57853 E/ 104–1731	Eastern Pontides, Lesser Caucasus	Ajaristskali/BS (24); Chorokhi/BS (18); Machakhela/BS (6); Mtkvari/CS (1); Natanebi/BS (7)

In the network, the placement of the putative ancestral node (indicated by the black arrow, Fig. 3) near the split between the Western Euxinian and Eastern Caucasian plastome groups was guided by mutation mapping onto the complete plastome phylogeny and supported by the [1499ff] insertion within the *trnK-trnQ* IGS region, which serves as a genetic synapomorphy linking the Eastern Caucasian plastome group with lineage V, and the transversion within the [1566ff] motif at *trnK-trnQ* acts as another genetic synapomorphy characterizing the sister lineage leading to the *F. sylvatica* plastomes (Worth et al. 2025, work in progress). This rooting scenario was consistent with the penalized likelihood strict-clock chronogram and the available fossil evidence, supporting an ancient divergence between the Western Euxinian and Eastern Caucasian-Hyrcanian plastome lineages.

The Va lineage includes haplotypes currently distributed across the Pontides, Caucasus, Crimean Mountains, and Alborz Mountains, the 'proto-Pontides-Caucasus-Alborz plastome type' (Fig. 5; Supplementary File S1, worksheet "Sample Stands for Mapping"), characterized by a [1499ff] insertion within *trnK-trnQ* (Fig. 3), while another lineage (Vb; Fig. 5) comprises haplotypes now predominantly distributed in Europe, henceforth, 'proto-European plastome type', which shows a transversion within the [1566ff] motif at *trnK-trnQ* (Fig. 4).

Prior to the Burdigalian (Early Miocene; node divergence time was 18.48 Ma; Fig. 5), the 'proto-Pontides-Caucasus-Alborz plastome type' diverged into a western Pontic *F. orientalis* lineage (SLPT 'Or'), characterized by a unique [1334] transition at *trnK-trnQ* (Fig. 3), and a second maternal lineage whose descendants are distributed across the Caucasus, the Crimean Mountains, and the Alborz Mountains, the 'proto-Caucasus-Alborz plastome type 1'.

By the Langhian (Middle Miocene; node divergence time was 14.94 Ma; Fig. 5), the 'proto-European plastome type' split into the plastome lineage now found in *F. sylvatica* (SLPT 'Sy') and the southwestern Georgian, Eastern Pontides *F. hohenackeriana* ('EuA/EuB'). Shortly after (node divergence time – 14.27 Ma; Fig. 5), the 'proto-Caucasus-Alborz plastome type 1' diverged into a Caucasian *F. hohenackeriana* lineage ('HoA'), now predominantly found in the Eastern Caucasus, and into another lineage, the 'proto-Caucasus-Alborz plastome type 2,' whose descendants are distributed across the Caucasus, the Crimean Mountains, and the Alborz Mountains and share an 8-nt duplication within *trnK-trnQ* [1566ff] (Fig. 4). This divergence coincides with the Middle Miocene Climatic Transition, a period of pronounced global cooling and ice-sheet expansion.

By the mid-Tortonian (Late Miocene; node divergence time- 9.38 Ma; Fig. 5), the 'proto-Caucasus-Alborz plastome type 2' had split into the Hyrcanian *F. caspica* ('CaA') lineage, which has a unique [551] transversion within the more conservative *petA-psbJ* IGS region (Fig. 3) and currently occurs in the eastern Alborz Mountains, and into another Caucasian-Hyrcanian clade, whose descendants remained distributed across the Caucasus, the Crimean Mountains, and the Alborz Mountains, the 'proto-Caucasus-Alborz plastome type 3,' which shares a [206] transition at *trnK-trnQ* (Fig. 3). Circa 1.5 myrs later (node divergence times- 7.84 and 6.48 Ma; Fig. 5) this 'proto-Caucasus-Alborz plastome type 3' disintegrated into the Hyrcanian *F. caspica* ('CaB') lineage, which has a unique [517] transversion at the *trnK-trnQ* (Fig. 3) and is currently found in the Talysh Mountains (western Alborz), a Caucasian *F. hohenackeriana* lineage ('HoB') and a lineage ('HoX') ranging from the Caucasus into the Crimean Mountains.

Ecological structuring of Euxinian-Caucasian F. hohenackeriana plastome lineages across environmental gradients

We compared Euxinian-Caucasian *F. hohenackeriana* plastome types from 229 natural stands using ecological profiles based on altitudinal range, annual precipitation, and forest-type composition (Supplementary File S1, worksheet "Ecological Profile Data"). Both altitudinal and annual precipitation gradients differed markedly across plastid lineages. The SLPTs show clear altitudinal differentiation (Fig. 7), supported by the Kruskal–Wallis test ($\chi^2 = 14.78$, $p < 0.002$) and the Median test ($\chi^2 = 7.98$, $p < 0.05$). The overall grand median altitude was 1022 m a.s.l (Fig. 7a). The Euxinian 'EuA/EuB/Eu×' lineage occupied the lowest elevations, with most occurrences concentrated in the low-montane belt at 300–800 m a.s.l., and median values clearly below the overall grand median (Fig. 7b). In contrast, the Caucasian 'HoA' showed the highest altitudinal distribution, mainly concentrated between approximately 900–1500 m a.s.l. The other two Caucasian lineages, 'HoX' and 'HoB', showed mixed or intermediate positions; 'HoX' was centered mostly around mid-montane elevations (700–1400 m a.s.l.), whereas 'HoB' showed the broadest altitudinal range, extending from low elevations to upper montane conditions (200–2000 m a.s.l.)

The annual precipitation gradient showed the greatest differentiation among SLPTs (Fig. 8); both the Kruskal–Wallis test ($\chi^2 = 108.83$, $p < 0.001$) and the median test ($\chi^2 = 114.17$, $p < 0.001$; Fig. 8a) indicated

highly significant ecological differentiation. The overall grand median precipitation was 1426 mm. The Euxinian ‘EuA/EuB/Eu×’ lineages occupied the wettest environments, with most populations concentrated between approximately 1500–3000 mm per year and several sites reaching hyper-humid conditions above 3000 mm (Fig. 8b). In contrast, ‘HoA’ occupied substantially drier climates, mostly within mesic to moderately humid environments with around 700–1400 mm of annual precipitation. The mixed/intermediate lineages, ‘HoX’ and ‘HoB’, occupied broad transitional precipitation ranges; ‘HoX’ showed a bimodal tendency, spanning both relatively dry montane Caucasian-like and humid Colchic-like conditions (about 800–3000 mm), whereas ‘HoB’ displayed the widest hydrological amplitude, ranging from mesic to extremely humid environments exceeding 3500 mm per year.

The dissimilarity patterns within the Euxinian-Caucasian *F. hohenackeriana* plastomes, including forest-type composition, are well reflected by the Bray–Curtis Principal Coordinates Analysis (PCoA) shown in Fig. 9a, b. The PCoAs showed strong ecological differentiation among SLPTs, with the first two axes explaining 92.5% of the total ecological variance (PCoA1 = 59.86%, PCoA2 = 32.6%; Fig. 9a). The ordination clearly separated the ‘EuA/EuB/Eu×’ and HoA lineages and highlighted the intermediate ecological position of the ‘HoB’ and ‘HoX’ lineages. Along PCoA1, the ‘EuA/EuB/Eu×’ occupied the positive extreme and were strongly associated with humid and hyper-humid Colchic Forest types (Fig. 9b). In contrast, ‘HoA’ occupied the negative extreme and was associated primarily with low-mid, montane beech, beech–hornbeam, montane coniferous, and thermophilous forest types characteristic of comparatively drier mesic montane conditions. The mixed/intermediate lineages ‘HoB’ and ‘HoX’ occupied transitional positions between the two lineages; ‘HoB’ was placed closer to ‘HoA’ along PCoA1, showing affinities in forest-type composition, but also showed affinity with ‘EuA/EuB/Eu×’ through the presence of Colchic humid forests. ‘HoX’ was differentiated mainly along PCoA2 and appeared to be associated with Colchic hyper-humid forests, combined with mixed montane broadleaf forests. Additionally, the eigenvector structure further emphasized this ecological contrast. Positive values along Vector1 were associated mainly with Colchic humid forests, whereas strongly negative values corresponded to montane coniferous, low–mid beech, thermophilous, and secondary forests. Vector2 primarily reflected a gradient between Colchic hyper-humid and low–mid beech–hornbeam forests.

Discussion

In this study, we investigated plastome diversity in western Eurasian beech populations located between the Black Sea and Caspian Sea drainage systems in the Caucasus and adjacent regions. As the most ecologically dominant deciduous broad-leaved tree species native to the Caucasus, these populations preserve an ancient evolutionary history of range shifts, disruptions, and adaptations to diverse local environmental conditions since the Late Miocene. Tectonic uplift and the subsequent formation of mountain systems increased landscape heterogeneity, driving ecological adaptations and shaping migration routes by creating dispersal corridors and geographic barriers (Hoorn et al. 2018; Tarkhnishvili 2014). Specifically, we examined how historical landscape oscillations during the Miocene orogenies, alongside increasingly cool and dry climatic conditions in the Pliocene and subsequent Quaternary

glaciations, shaped haplotype diversity and biogeographical distribution of modern western Eurasian *Fagus* populations across this topographically complex region.

Deep Miocene maternal lineage divergence and clear phylogeographic structure of Caucasian Fagus

Our plastid penalized-likelihood strict-clock chronogram and spatial mapping (Fig. 5) provided valuable insights into these historical questions. These findings revealed significant temporal and spatial patterns supporting a pre-Pliocene maternal sequence divergence (more than 5.3 Ma), which underlies the current diversity in Caucasian and adjacent Hyrcanian beech populations. This divergence differed notably from results obtained from nuclear genome-based phylogenies, aligning with a holistic evolutionary framework in which plastid phylogenies and nuclear genealogies capture fundamentally different dimensions of a plant group's history. Such discrepancies are often interpreted as transcontinental vicariance, incomplete sorting of genetic and morphological traits, and repeated secondary contacts, all driven by re-radiation and dynamic dispersal accompanied by different rates of pollen- and seed-mediated interspecies gene flow. Based on nuclear-encoded isozyme markers and fossil evidence from existing western Eurasian beech, modern species differentiation in this area is estimated to be much younger, likely not exceeding 2 Ma (Gömöry et al. 2018). While these biparentally inherited nuclear markers capture the relatively recent population divergence and speciation events, our analysis of maternal plastomes revealed a much deeper evolutionary history, indicating that these deeply diverged plastid lineages originated within ancient fossil taxa long before modern species boundaries stabilized. Over time, through subsequent cycles of reticulate evolution, these ancient plastomes were captured and preserved within modern populations, serving as molecular "living witnesses" of Miocene vicariance and past geographic disruptions.

This deep evolutionary framework is clearly illustrated by our weighted (quasi-)full median network of plastid haplotypes representing the major western Eurasian *Fagus* lineages, constructed from weighted mutational characters derived from an emended complete plastome alignment (Fig. 4). Our estimate of distance-based differentiation (*Fst* = 0.21, Table 3) was notably higher than the frequency-based estimates (*Gst* = 0.15; *yst* = 0.11), suggesting that differences between the Black Sea and Caspian Sea drainage systems are not simply due to minor contemporary changes in haplotype frequencies within modern species. Instead, it highlighted substantial, sequence-level divergence caused by unique, ancient evolutionary lineages within each drainage system. The predominant maternal lineages in each basin have accumulated these specific mutations over extended periods of geographical separation. These lineages survived Quaternary environmental glaciations and introgressed into modern species pools, highlighting how ancient cytoplasmic architecture can be maintained across complex landscapes, even as nuclear genomes undergo more recent adaptive radiation and speciation.

These data provide robust, quantitative validation for a Miocene vicariance scenario, linking the divergence of ancient Lineages Va ('the proto-European plastome type') and Vb ('the proto-Pontides-Caucasus-Alborz plastome type') to the complex palaeogeography of the ancient Paratethys Sea. Rather than tracking modern species boundaries, these ancient plastid lineages effectively preserve an

ancestral cytoplasmic architecture, locking in the precise palaeogeographic boundaries of the Miocene epoch. The spatial segregation of these two main western Eurasian maternal lineages during the Early-to-Middle Miocene began around ~ 18.48 Ma, marked by the early divergence of SLPT 'V(Or)', a unique type restricted to the Western Pontic populations of *F. orientalis* that acts as a sister type to the eastern *F. hohenackeriana*, and continued through ~ 14.94 Ma. This segregation corresponded to the initial intense phases of the Caucasian Island orogeny driven by the Arabian-Eurasian collision (Rögl 1998; Saintot et al. 2006). Instead of spreading across a continuous modern landmass, ancestral, pre-extant beech populations were fragmented among an emerging proto-Caucasian and Pontic Island archipelago separated by marine barriers in the Eastern Paratethys (Popov et al. 2004). The subsequent intra-lineage diversification events during the Middle-to-Late Miocene (e.g., splits at ~ 14.27 Ma, ~ 9.38 Ma, ~ 7.84 Ma, and ~ 6.48 Ma leading to highly localized SLPTs like 'HoA', 'HoB', 'HoX', 'CaA', and 'CaB') were influenced by episodic structural oscillations and sea-level fluctuations characteristic of this land-sea system, rather than Quaternary climate shifts. This pre-Pliocene genetic architecture definitively separates Caucasian *Fagus* from the Quaternary-dominated history of European *F. sylvatica*, proving that the micro-geographic endemism observed today is a living footprint of Miocene vicariance (Demesure et al. 1996; Magri et al. 2006).

Notably, our fine-scale multi-locus plastid sequencing revealed potentially new SLPTs with distinct phylogenetic affinities (Figs. 4, 5) and distribution areas, providing evidence of reticulate evolution. Identifying the Euxinian lineage ('EuA, EuB, Eu×') represents a major shift. Shared genetic synapomorphies confirm that plastome types 'EuA' and 'EuB' are Miocene-diverged (≥ 14Ma) sister lineages to European *F. sylvatica* (possible sister types of 'Sy'), rather than to any other Caucasian haplotypes found in the Black Sea and Caspian Sea drainage systems. Additionally, the chimeric (mosaic) lineage 'Eu×' within the Euxinian SLPT, which shows mutations unique to 'EuB' (and 'EuA') in its 5' European-like part, has a 3' Caucasian-like part (Fig. 6) that is indistinguishable from the (reconstructed) ancestral sequence of the *F. hohenackeriana-F. caspica* 'HoB(-HoX)-CaA-CaB' sublineage. This provides molecular evidence of historical introgressive hybridization, implying that these ancestral lineages actively hybridized when their ranges overlapped during past climate shifts. Furthermore, the trans-regional lineage 'HoX', recognized as a distinct Late Miocene type (≥ 6.48 Ma), exhibits a wide distribution across the central-western Greater Caucasus and a restricted occurrence in the eastern Greater Caucasus. Its presence in *F.* × *taurica* from the Crimean Mountains supports a shared Euxinian-Caucasian history consistent with nuclear SSR evidence (Aliev et al. 2025).

Contrasting demographic histories in the Black Sea and the Caspian Sea drainage systems, molecular evidence of a complex refugial mosaic of Fagus in the Caucasus and secondary contact zones

Within the Caucasus, distinct geographic partitioning and contrasting demographic histories in the Black Sea and Caspian Sea drainage systems provided robust molecular evidence of a complex refugial mosaic, highlighting multiple micro-refugia. The Black Sea drainage system exhibited a highly complex demographic history, showing a ragged, multimodal mismatch distribution accompanied by significantly positive Fu and Li's neutrality test statistics, and rejecting the hypothesis of a simple, recent population

expansion. Indeed, these patterns indicate long-term stability and historical population structure. The high haplotype diversity and absence of *SVs* suggest that the Colchis region and adjacent Black Sea watersheds acted as a massive, ancient refugial melting pot (Connor and Kvavadze 2009; Shatilova et al. 2011; Dagtekin et al. 2020), where multiple distinct lineages likely persisted in localized microrefugia throughout past glaciations (Capblancq et al. 2024). For instance, the Caucasian *F. hohenackeriana* 'HoB' and 'HoX' plastome types (haplotypes H06-H08) were predominantly constrained to the slopes of the central-western Greater Caucasus, reflecting an ancient colonization zone that remained isolated from eastern dynamics.

In contrast, the Caspian Sea drainage system displays a signature of recent demographic expansion, evidenced by a smooth, unimodal mismatch distribution with reduced pairwise variance, significantly negative Fu and Li's test statistics, and the presence of *SVs*. A single haplotype dominated: H14 (part of the Caucasian *F. hohenackeriana* 'HoA' SLPT), which accounted for 62.9% of all samples, suggesting an initial founder effect at colonization, consistent with findings by Sękiewicz et al. (2022) using nuclear SSR markers. Post-glacial recolonization of the eastern Greater Caucasus and Lesser Caucasus likely proceeded via rapid expansion from a restricted maternal source. This sudden expansion established the dominance of haplotype H14 across the river gorges of the eastern Greater Caucasus and extended down into the Lesser Caucasus in the Caspian Sea drainage system. This pronounced east-west genetic gradient, established during the late Miocene oscillations (from ~ 9.38 Ma to ~ 6.48 Ma), is physically demarcated today by the low-lying Rioni and Mtkvari depressions, which have historically acted as major topographical and ecological barriers preventing gene flow between these micro-regions (Tarkhnishvili 2014).

However, despite the topographic barriers of the Greater and Lesser Caucasus mountains, our data reveal that these ridges do not form absolute barriers to seed dispersal, but function as semi-permeable corridors and secondary contact zones (Tarkhnishvili et al. 2012; Sękiewicz et al. 2022), as confirmed by the presence of five shared haplotypes pointing to historical migration corridors. Biogeographically, the Meskheti Range within the Lesser Caucasus acts as the primary watershed between the Colchic (Black Sea drainage system) and Caucasian (Caspian Sea drainage system) floristic provinces, historically regulating migration between the Rioni and Mtkvari main river systems (Tarknishvili et al. 2012). Both river systems serve as key low-elevation gateways in both directions, facilitating eastward migrations of haplotypes H06 and H08 and westward migration of H14. Identifying stands with internal haplotype diversity supports evidence for secondary contact zones. Regions such as Racha-Lechkhumi and Imereti (the Black Sea drainage system), alongside Kakheti, Dagestan, and Ingushetia (the Caspian Sea drainage system), may be considered evolutionary suture zones (Tarkhnishvili et al. 2012; Sękiewicz et al. 2022; Capblancq et al. 2024). In these areas, formerly isolated maternal lineages from distinct refugial origins have converged and blended during recent postglacial expansions.

Isolation-by-environment as a trigger for ecological filtering of Euxinian-Caucasian F. hohenackeriana plastome types

Further, the observed high level of structuring in Euxinian-Caucasian *F. hohenackeriana* plastome types is maintained not only by the complicated landscape of the Caucasus, but also by ecological filtering of plastome types along steep environmental and altitudinal gradients. In this context, the Likhi Range serves as an important geomorphological and bioclimatic boundary, collecting maritime precipitation from the west and separating the Euxinian-Colchic flora from the mainland Caucasian province assemblies to the east (Zazanashvili et al. 2000). The significant differences in altitude, precipitation, and dissimilarity patterns in forest-type composition are well reflected by both univariate tests (such as the Kruskal-Wallis and Median tests) and multiple analysis (such as Bray-Curtis Principal Coordinate Analysis (PCoA)), proving that these maternal lineages occupy distinct environmental spaces. For instance, the ancient Euxinian lineages ('EuA/EuB/Eu×') are strictly confined to the low montane belt (300–800 m a.s.l.) within the Ajaristskali, Chorokhi, Machakahela, and Natanebi river systems, reflecting a highly specialized ecological profile with a strong affinity for humid, buffered maritime conditions associated with humid and hyper-humid Colchic Forest types. These include beech forests with Colchic underwood (*Fageta fruticosa colchica*), regarded as a unique forest formation endemic to the Caucasus (Dolukhanov, 2010). In contrast, the Caucasian 'HoA' plastome type exhibited the highest altitudinal distribution, primarily ranging from 900 to 1500 m a.s.l. and was associated with substantially drier climates, typically found in mesic to moderately humid environments. This lineage is largely characterized by low- to mid-montane beech, beech-hornbeam, montane coniferous, and thermophilous forest types, which are typical of the relatively drier conditions found in mesic montane areas. This ecological filtering limits gene flow, isolating maternal lineages by environment even when they are geographically close, supporting the idea that local adaptation is still shaping the genetic structure of Caucasian beech.

Additionally, because most cpDNA evolves neutrally and is strictly seed-mediated, this pronounced ecological divergence was likely driven by historic selective sweeps. Strong positive selection on localized adaptive traits in the non-recombinant plastome likely fixed specific maternal haplotypes through genetic hitchhiking, effectively preserving neutral genetic variation within distinct Euxinian and Caucasian maternal lineages. The evolutionary patterns of these lineages closely reflect their divergence times; both the Euxinian ('EuA/EuB/Eu×') and Caucasian ('HoA') plastome types diverged early in the Middle Miocene (~ 14.94 and 14.25 Ma; Fig. 5), providing prolonged environmental stability for intense, long-term ecological filtering and localized selective sweeps to guide highly specialized, contrasting profiles. The Euxinian 'EuA/EuB/Eu×' lineages were confined to hyper-humid, buffered maritime refugia, and the Caucasian 'HoA' lineage to the drier, higher-altitude montane zones. Conversely, the mixed-intermediate 'HoX' and 'HoB' lineages, which diverged much later (6.48 Ma; Fig. 5) during a period of intense Late Miocene climatic fluctuations and Caucasian orogeny, display broad altitudinal and annual precipitation ranges. Rather than undergoing strict environmental filtering, these younger lineages are likely characterized by high phenotypic plasticity, enabling them to successfully span different, transitional environments across the changing landscape.

Regional structure of Caucasian beech populations contrasts with European Fagus

At a broad geographic scale, this profound regional structure observed in Caucasian beech populations contrasts with that of European *Fagus*. While the European lineages occupy a highly uniform, low-diversity space across central and western Europe due to massive post-glacial climate bottlenecks and rapid post-glacial expansions from multiple interglacial cycles (Margi et al. 2006), the Caucasian plastome types remain deeply compartmentalized into a refugial mosaic. Rather than forming a single continuous forest pool, these mountains acted as independent evolutionary islands, locking unique maternal lineages into separate valleys and gorges, where they have persisted without significant inter-range homogenization for millions of years. This highlights that the Caucasus is not just a valuable Tertiary repository but a vital and irreplaceable source of highly localized genetic evolutionary potential that requires focused conservation strategies for this unique forest biodiversity.

Additionally, tracking non-recombinant maternal lineages is an essential resource for biodiversity management, because the conservative nature of cpDNA is preserved in the structural geography of maternal migrants and seed-mediated isolation over millions of years. These provide conservation agencies with a reliable, empirical blueprint of where long-term environmental stability has historically persisted. Although most Caucasian beech populations in this study are preserved within existing protected areas or are geographically isolated, gaps in preservation remain. For instance, in the Lapankuri Gorge (Kakheti, eastern Greater Caucasus), one of the secondary contact zones containing distinct haplotypes - H06 (trans-regional 'HoX' *F. hohenackeriana* lineage) and H14 ('HoA' *F. hohenackeriana* lineage predominantly associated with Caspian Sea drainage gorges across the Greater Caucasus and the Lesser Caucasus) - persists within secondary forests, indicating ongoing anthropogenic pressure. Therefore, cpDNA screening is a valuable tool for spatial prioritization, helping identify and legally protect these complex refugial mosaics with contrasting catchment patterns in the Caucasus.

Conclusion

The plastome divergence in Caucasian *Fagus* is fundamentally shaped by Miocene tectonic and palaeographic dynamics within the ancient Paratethys archipelago rather than Quaternary glacial cycles, in contrast to the bottleneck-dominated evolutionary history of European *Fagus*. These maternal lineages persisted through Quaternary glaciations and contributed to the gene pools of modern species, demonstrating how ancient cytoplasmic architecture can be preserved across complex landscapes, even as nuclear genomes undergo more recent adaptive radiation and speciation. Our fine-scale multi-locus plastome sequence analysis successfully resolved undocumented SLPTs, including ancient Euxinian lineages linked to European types, a chimeric (mosaic) lineage highlighting historical chloroplast introgressive hybridization, and a trans-regional lineage verifying a shared Euxinian-Caucasian history. The Caucasus itself operates as a complex refugial mosaic with highly contrasting catchment patterns. The Black Sea drainage system exhibits long-term demographic stability, acting as an ancient melting pot. In contrast, the Caspian Sea drainage system mirrors a massive, recent post-glacial expansion with a founder effect dominated by a single H14 haplotype. While major geographic barriers restrict broad-scale mixing, low-elevation gateways such as the Mtkvari and Rioni river basins facilitate rare seed

dispersal westward and eastward, resulting in secondary contact zones. Isolation-by-environment triggers the high level of spatial structuring of unique Euxinian-Caucasian *F. hohenackeriana* lineages. Driven by historic selective sweeps within the non-recombinant plastome, this ecological filtering is actively reinforced by environmental selection, which restricts immigrant gene flow as effectively as physical mountain barriers and reflects deep chronological divergence histories. Early diverging Middle Miocene lineages were guided by highly specialized, contrasting niches, while younger Late Miocene lineages showed high phenotypic plasticity across transitional eco-climatic gradients. The Caucasian mountains, as independent evolutionary islands, have preserved unique maternal lineages for millions of years, highlighting the Caucasus as a long-term center of persistence, evolutionary diversification, and lineage differentiation. Chloroplast DNA screening is an effective method for spatial prioritization, helping identify and legally safeguard these complex refugial mosaics with contrasting catchment patterns across the region.

Declarations

Author Contribution

J.E.: Conceptualization, Formal analysis, Funding acquisition, Investigation, Methodology, Project Administration, Supervision, Visualization, Writing – original draft; T.J.: Investigation, Project Administration, Resources, Supervision; A.B.: Investigation, Formal analysis, Resources; A.G: Formal analysis, Writing – original draft; M.C.S: Conceptualization, Methodology, Validation, Supervision, Writing – review and editing; G.W.G.: Formal Analysis, Validation, Visualization, Writing – review and editing; T.D.: Validation, Writing – review and editing; M.O.: Investigation, Resources; T.P.: Investigation, Resources; M.K.: Visualization; N.T.: Formal analysis, Investigation, Resources; Kh.A.: Investigation, Resources; O.Y.: Investigation, Resources; T.M-P..: Data Curation; J.R.P.W.: Resources, Writing – review and editing; G.F.: Investigation, Resources; L.Kh.: Investigation, Resources; A.N.: Investigation, Resources; S.T.: Investigation, Resources.

Acknowledgement

This research was carried out within the framework of the project FR-22-889 (“The phylogeography and genetic structure of the most widespread and ecologically essential tree species (Fagus orientalis Lipsky) in Georgia, South Caucasus”) funded by the Shota Rustaveli National Science Foundation of Georgia (SRNSFG).

Data Availability

All generated chloroplast DNA sequences are available on NCBI under accession numbers ... . All other relevant data and files needed to reproduce the reported results are contained within the manuscript and its supporting materials.

References

Aliev KhU, Kutsev MG, Tuniyev BS (2025) Population genetic structure of *Fagus orientalis* with biogeographical notes. Turczaninowia 28(3): 165-181.

https://www.doi.org/10.14258/turczaninowia.28.3.19

Amarilla LD, Anton AM, Chiapella JO, Manifesto MM, Angulo DF, Sosa V (2015) *Munroa argentina*, a grass of the South American transition zone, survived the Andean uplift, aridification and glaciations of the Quaternary. PLOS ONE 10(6).

https://doi:10.1371/journal.pone.0128559

Bandelt HJ, Forster P, Röhl A (1999) Median-joining networks for inferring intraspecific phylogenies. Mol Biol Evol 16(1):37–48

Bandelt HJ, Macaulay V, Richards M (2000) Median networks: speedy construction and greedy reduction, one simulation, and two case studies from human mtDNA. Mol Phyl Evol 16(1):8–28.

https://doi:10.1006/mpev.2000.0792. PMID: 10877936

Beridze B, Sękiewicz K, Walas Ł, Thomas PA, Danelia I, Kvartskhava G, Fazaliyev V, Bruch A A, Dering M (2023) Evolutionary history of *Castanea sativa* Mill. in the Caucasus driven by Middle and Late Pleistocene paleoenvironmental changes. AoB Plants 15(5).

https://doi.org/10.1093/aobpla/plad059

Bouckaert R, Heled J, Kühnert D, Vaughan T, Wu C-H, Xie D, Suchard MA, Rambaut A, Drummond AJ (2014) BEAST 2: A Software Platform for Bayesian Evolutionary Analysis. PLoS Comput Biol 10(4).

https://doi:10.1371/journal.pcbi.100353

Braun-Blanquet J (1964) Pflanzensoziologie: Grundzüge der Vegetationskunde. 3rd Edition, Springer, Berlin.

Brown CD, Macias-Fauria M, Johnstone J F (2018) Topoclimate versus macroclimate: How does climate mapping methodology affect species distribution models and climate change projections? *Divers Distrib 20(80):952-963.*

https://doi.org/10.1111/ddi.12216

Budde B, Kurz B, Papageorgiou AC, Gailing O, Leinmann L, Finkeldey R (2023) Bidirectional gene flow between *Fagus sylvatica* L. and *F. orientalis* Lipsky despite strong genetic divergence. For Ecol Manage 538:120984.

https://doi.org/10.106/j.foreco.2023.120984

Capblancq T, Sękiewicz K, Dering M (2024) Forest genomics in the Caucasus through the lens of its dominant tree species – *Fagus orientalis*. Mol Ecol 33:e17475.

https://doi.org/10.1111/mec.17475

Cardoni S, Piredda R, Denk T, Grimm GW, Papageorgious AC, Schulzes ED, Scoppola A, Shanjjani PS, Suyama Y, Tomaru N, Worth JRP, Simeone MS (2022) 5S-IGS rDNA in wind-pollinated trees (*Fagus* L.) encapsulates 55 million years of reticulate evolution and hybrid origins of modern species. Plant J 109:909-926.

https://doi.org/10.1111/tpj.15601

Cardoni S, Grimm GW, Denk T, Csilléry K, Schulze ED, Simeone MS (2026) A new subspecies of *Fagus orientalis* (Fagaceae) in Turkey: *F. orientalis* subsp. *nurensis*, a relict close to the ancestor of modern European and Oriental Beech. npj Sci Plants 2:25.

https://doi.org/10.1038/s44383-026-00039-3

Connor S E, and Kvavadze E V (2009) Modelling late Quaternary changes in plant distribution, vegetation and climate using pollen data from Georgia, Caucasus. J Biogeogr 36(3):529–545.

https://doi.org/10.1111/j.1365-2699.2008.02019.x

Cozzolino S, Vendramin GG, Chrietenhusz S, Galf G, Pasta P, Fazan L, Kozlowski G (2014) Footprints of fast intensive diversification and structuring in the layers of western Eurasian *Zelkova* taxa. J Biogeogr 41(6)1081-1092.

https://doi.org:10.1111/jbi.12276

Dagtekin D, Şahan E., Denk , Köse, N, Dalfes HN (2020) Past, present and future distributions of Oriental beech (*Fagus orientalis*) under climate change projections. PLoS One 15(11).

https://doi.org/10.1371/journ al.pone.0242280

Demesure B, Comps B, Petit RJ (1996) Chloroplast DNA phylogeography of the common beech (*Fagus sylvatica* L.) in Europe. Evol 50(6):2515–2520

Denk T (1999a) The taxonomy of *Fagu*s in western Eurasia 1: *Fagus sylvatica* subsp. *orientalis* (= *F. orientalis*). Feddes Repert 110:177–200.

https://doi.org/10.1002/fedr.4921100305

Denk T (1999b) The taxonomy of *Fagus* in western Eurasia 2: *Fagus sylvatica* subsp. *sylvatica*. Feddes Repert 110:381–412.

https://doi.org/10.1002/fedr.19991100510

Denk T, Frotzler N, Davitashvili N (2001) Vegetation patterns and distribution of relict taxa in humid temperate forests and wetlands of Georgia (Transcaucasia). Biol J Linn Soc 72:287-332.

https://doi.org/10.1006/BIJL.2000.050

Denk T, Grimm GW (2005) Phylogeny and biogeography of *Zelkova* (Ulmaceae sensu stricto) as inferred from leaf morphology, ITS sequence data and the fossil record. Bot J Linn Soc147(2):129-157.

https://doi.org/10.1111/j.1095-8339.2005.0034.x

Denk, T., and Grimm, G. W. (2009) The biogeographic history of beech (*Fagus*) in the Northern Hemisphere. Rev Palaeobot Palynol 158(1-2):83-100.

https://doi.org/10.1016/j.revpalbo.2009.08.007

Denk T, Grimm GW (2010) The oaks of western Eurasia: Traditional classifications and evidence from two nuclear markers. Taxon 59:351-366.

https://doi.org/10.1002/tax.592002

Denk T, Grimm GW, Cardoni S, Gsilléry K, Kurz M, Scultze ED, Simeone MC, Worth JRP (2024) A subgeneric classification of *Fagus* (Fagaceae) and revised taxonomy of western Eurasian beeches. Willdenowia 54(2-3):151-181.

https://doi.org/10.3372/wi.54.54301

Dolukhanov AG (2010) Georgian forest vegetation (Lesnaia rastitelnost Gruzii). Publisher house “Universal”, Tbilisi, Georgia (in Russian).

Drummond AJ, Ho SYW, Phillips MJ, Rambaut A (2026) Relaxed phylogenetics and dating with confidence. PLOS Biol 4(5):e88.

https://doi.org/10.371/journal.pbio.0040088

Ekhvaia J, Simeone MC, Silakadze N, Abdaladze O (2018) Morphological diversity and phylogeography of the Georgian durmast oak (*Q. petraea* subsp. *iberica*) and related Caucasian oak species in Georgia (South Caucasus). Tree Genet Genomes 14(2):1-15.

https://doi.org/10.1007/s11295-018-1232-6

Etterson JR, Shaw RG (2001) Constraints to adaptive evolution in response to global warming. Science 294(5540):151- 154

Excoffier L, Foll M, Petit RJ (2009) Genetic consequences of range expansions. Annu Rev Ecol Evol Systemat 40:481–501

Fu YX, Li WH (1993) Statistical tests of neutrality of mutations. Genetics 133:693-709.

https://doi.org:10.1093/genetics/133.3.693

Gavin DG, Fitzpatrick MG, Gugger PF, Health KD, Rodríguez-Sánchez F, Dobrowski SZ et al (2014) Climate refugia: joint inference from fossil record, species distribution models and phylogeography. Phytologist 204(1):37- 54.

https://doi.org/10.1111/nph.12929

Gerhard T (2008) The conditioned reconstructed process. J Theor Biol 253:769- 778.

https://doi.org/10.1016/j.jtbi.2008.04.005

Gömöry D, Paule L, Macejovsky V (2018) Phylogeny of beech in western Eurasia as inferred by approximate Bayesian computation. Acta Soc Bot Pol 87:3582.

https://doi.org/10.5586/asbp.3582

Grimm GW, Denk T, Hemleben V (2007) Evolutionary history and systematics of Acer section *Acer* - a case study of low-level phylogenetics. Plant Syst Evol 267:215-253.

https://doi.org/10.1007/s00606-007-0572-8

Grimm GW, Denk T (2014) The Colchic region as refuge for relict tree lineages: cryptic speciation in field maples. Turk J Bot *38* (6):1050-1066.

https://doi.org/10.3906/bot-1403-87

Gulisashvili VZ, Mahatadze LB, Prilipko LI (1975) Vegetation of the Caucasus (Rastitelnost Kavakaza). Moscow: Nauka.

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

Hasanov ZM, Ibrahimov ZA, Nabiyev VR (2016) Beech forests of Azerbaijan: The modern condition, age structure and regeneration. Ann Agrar Sci 15(4):563-457.

https://doi.org/10.1016/j.aasci.2016.08.003

Hasegawa M, Hirohisa K, Taka-aki Y (1985) Dating of the Human-Ape Splitting by a Molecular Clock of Mitochondrial DNA. J Mol Evol 22:160-174

Ho SYW, Phillips MJ (2009) Accounting for calibration uncertainty in phylogenetic estimation of evolutionary divergence times. Syst Biol 58(3):367- 380

Ho SYW, Duchêne S (2014) Molecular-clock methods for estimating evolutionary rates and timescales. Mol Ecol 23(24):5947- 5965

Hoorn C, Perrigo A, Antonelli A. (Eds.) (2018) Mountains, Climate and Biodiversity: An Introduction. Wiley-Blackwell.

Hudson RR, Boos DD, Kaplan NN (1992a) A statistical test for detecting population subdivision. Mol. Biol. Evol. 9: 138-151

Hudson R (2000) A New Statistic for Detecting Genetic Differentiation. Genetics, 155:2011-2014.

https://doi.org/10.1093/genetics/155.4.2011.

Jasińca AK, Rucińca B, Kozlowski G, Bétrisey H, Safarov H, Boratyńca K, Boratyński A (2015) Morphological differentiation of leaves in the relict tree *Zelkova carpinifolia* (Ulmaceae). Dendrobiology, 74:109-122

https://doi.org/10.12657/denbio.074.011

Kolakovsky AA (1961) The Flora and Vegetation of Colchida. – Moscow: Moscow University Press, (Russian with English summary).

Kumar S, Stecher G, Suleski M, Sanderford M, Sharma S, Tamura K (2024) Molecular Evolutionary Genetics Analysis Version 12 for adaptive and green computing. Mol Biol Evol 41:1-9

Kurz M, Kolz A, Gorges J, Carmona BP, Brang P, Vitasse Y, Kohler M, Rezzonico F et al (2023) Tracing the origin of Oriental beech stands across Western Europe and reporting hybridization with European beech – implications for assisted gene flow. For Ecol Manage 531(120801)

Leroy SA, Arpe K (2007) Glacial refugia for summer-green trees in Europe and south-west Asia as proposed by ECHAM3 time-slice atmospheric model simulations. J Biogeogr 34(12):2115-2128

Leroy T, M. Heuertz M (2026) Adaptive Introgression in the Context of Climate Adaptation. Mol Ecol 35(8): e70333.

https://doi.org/10.1111/mec.70333

Margi D, Vendramin GG, Comps B, Dupanloup I, Geburek T, Gömöry D, Latałowa M, Litt T et al (2006) A new scenario for the quaternary history of European beech populations: palaeobotanical evidence and

genetic consequences. New Phytol. 171(1):199-221.

https://doi.org/10.1111/j.1469-8137.2006.01740.x.

Maycock PF (1994) The ecology of beech (*Fagus grandifolia* Ehrh.) forests of the deciduous forests of southeastern North America, and a comparison with the beech (*Fagus crenata*) forests of Japan. Pp. 351–410 in: Miyawaki A, Iwatsuki K, Grandtner MM (ed.), Vegetation in eastern North America. Tokyo: University of Tokyo Press.

Mayer H, Aksoy H (1986) Forests of Turkey (Walder der Turkei). Stuttgart: Gustav Fischer.

Médail F, Diadema K (2009) Glacial refugia influence plant diversity patterns in the Mediterranean Basin. J Biogeogr 36(7):1333-1345.

https://doi.org/10.1111/j.1365-2699.2008.02051.x

Milesi P, Kastally C, Dauphin B et al (2024) Resilience of genetic diversity in forest trees over the Quaternary. Nat Commun 15: 8538.

https://doi.org/10.1038/s41467-024-52612-y

Nakhutsrishvili G, Zazanashvili N, Batsatsashvili K (2011) Regional profile: Colchic and Hyrcanic Temperate Rainforest of the Western Eurasian Caucasus. Pp. 214-221 in: Della Sala D A (Ed.), Temperate and boreal rainforests of the world: Ecology and conservation. – Washington, Covelo, London.

Nakhutsrishvili G, Zazanashvili N, Batsatsashvili K, Montalvo CS (2015) Colchic and Hyrcanian forests of the Caucasus: similarities, differences and conservation status. Fl Medit 25 (Special Issue).

Olson DM, Dinerstein E, Wikramanayake E.D, Burgess ND, Powell GV, Underwood EC et al (2001) Terrestrial ecoregions of the world: a new map of life on earth. Bioscience, 51:933– 938.

https://doi.org /10.1641/0006-3568(2001)051[0933:TEOTWA]2.0.CO;2

Orsini L, Vanoverbeke J, Swillen I, Mergeay J, Dee Meester L (2013) Drivers of population genetic differentiation in the wild: Isolation by dispersal limitation, isolation by adaptation, and isolation by colonization. Mol Ecol 22(24):5983- 5999

Paradis E (2013) Molecular dating of phylogenies by likelihood methods: A comparison of models and a new information criterion. Mol Phylogen Evol 67(2):436–444

Peters R (1997) Beech forests. Geobot 24:1–169.

https://doi.org/10.1007/978-94-015-8794-5

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

https://doi.org/10.2307/25065474

Pham KP, Hipp AL, Manos PS, Cronn RC (2017) A time and a place for everything: phylogeographic history and geography as joint predictors of oak plastome phylogeny. Genome 60: 720–732.

https://doi.org/10.1139/gen-2016-0191

Popov SV, Rozanov AY, Rögl F, Steininger FF, Shcherba IG, Kovac M (2004) Lithological-paleogeographic maps of Paratethys. CFS Courier Forschungsinstitut Senckenberg, 250:1– 46.

Possada D, Crandall KA (1998) Model test: testing the model of DNA substitution. Bioinform 14(9):817-818.

https://doi.org/10.1093/bioinformatics/14.9.817

Ramos-Onsins SE, Rozas J (2003) Statistical properties of new neutrality tests against population growth. Mol Biol Evol 19(12):2092- 2100.

https://doi.org/10.1093/0xfordjournal.molbev.a004034

Renner SS, Grimm GW, Kapli P, Denk T (2016) Species relationships and divergence times in beeches: new insights from the inclusion of 53 young and old fossils in a birth–death clock model. Philosophical Transactions of the Royal Society Publishing B, 371:20150135.

http://dx.doi.org/10.1098/rstb.2015.0135

Rögl F (1998) Palaeogeographic Mesozoic–Cenozoic development of the Central Paratethys. Annalen des Naturhistorischen Museums in Wien, 100A, 523–544.

Rozas J, Ferrer-Mata A, Sánchez-DelBarrio JC, Guirao-Rico S, Librado P, Ramos-Onsins SE, Sánchez-Gracia A (2017) DnaSP v.6: DNA Sequence Polymorphism Analysis of Large Datasets. Mol Biol Evol 34:3299- 3302

Ruhfel BR, Gitzendanner MA, Soltis PS, Soltis DE, Burleigh JG (2014) From algae to angiosperms—inferring the phylogeny of green plants (Viridiplantae) from 360 plastid genomes. BMC Evol Biol 14: 23.

https://doi.org/10.1186/1471- 2148-14-23. PMID:24533922

Saintot A, Brunet, M-F, Yakovlev F, Sébrier M, Stephenson R, Ershov A, Chalot-Prat FE, McCann T (2006) The Mesozoic-Cenozoic tectonic evolution of the Greater Caucasus. Geological Society, London, Memoirs. 32. 277-289.

https://doi.org/10.1144/GSL.MEM.2006.032.01.16

Safarov IS (1979) Subtropical forests of Tallish (Subtropicheskie lesa Talysha). Baku, Elm (In Russian)

Sanderson MJ (2002) Estimating absolute rates of molecular evolution and divergence times: a penalized likelihood approach. Mol Biol Evol 19, 101–109.

https://doi.org/10.1093/ 562 oxfordjournals.molbev.a003974

Schaal BA, Hayworth DA, Olsen KM, Rauscher JT, Smith WA (1998) Phylogeographic studies in plants: problems and prospects. Mol Ecol 7(4):465- 474

Shatilova I, Mchedlishvili N, Rukhadze L, Kvavadze E (2011) The history of flora and vegetation of Georgia. Georgian National Museum, Institute of Paleobiology, Tbilisi, Georgia.

Shaw J, Shafer HL, Leonard OR, Kovach MJ, Schorr M, Morris AB (2014) Chloroplast DNA sequence utility for the lowest phylogenetic and phylogeographic inferences in angiosperms: the tortoise and the hare IV. Am J Bot 101(11):1987– 2004.

https://doi.org/10.3732/ajb.1400398. PMID:25366863.

Schulze ED, Grimm GW (2022) Alles Bastarde. Die Buche, ein eurasisches Art-Mosaik. Biologie in Unserer Zeit, 52: 340–351.

https://doi.org/10.11576/biuz-5864

Sękiewicz K, Danelia I, Farzaliyev V, Gholizadeh H, Iszkuło G, Naqinezhad A, Ramezani E et al (2022) Past climatic refugia and landscape resistance explain spatial genetic structure in Oriental beech in the South Caucasus. Ecol Evol 12(e9320).

https://doi.org/10.1002/ece3.9320

Takhtajan AL (1986) Floristic Regions of the World (TJ Crovello, Trans.; A. Cronquist, Ed.). University of California Press

Tarkhnishvili D, Gavashelishvili A, Mumladle L (2012) Paleoclimatic models help to understand current distribution of Caucasian Forest species. Biol J the Lin Soc 105(1):231-248.

https://doi.org/10.1111/j.1095-8312.2011.01738.x

Tarkhnishvili D, Gavashelishvili A, Murtskhvaladze M, Mumladle L (2014) Historical biogeography of the Caucasus: a synthesis of paleoecological, geological and paleogeographic evidence. Biol Rev 89(2):281-301.

https://doii.org/10.1111/brv.12063

Tarkhnishvili D (2014) Historical Biogeography of the Caucasus. Nova Science Publishers.

Tiffney BH, Manchester SR (2001) The use of geological and paleontological evidence in evaluating plant phylogeographic hypotheses in the Northern Hemisphere Tertiary. Int J Plant Sci 162(S6), pp. S3-S17

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

Wang IJ, Bradburd GS (2014) Isolation by environment. Mol Ecol 23(23):5649–5662

Wang LG, Lam TTY, Xu S, Dai Z, Zhou L, Feng T, Guo P, Dunn CW, Jones BR, et al (2020) “treeio”: an R package for phylogenetic tree input and output with richly annotated and associated data. Mol Biol Evol 37(2):599-603.

https://doi.org/10.1093/molbev/msz240

Wang J, Oliveira BF, Moore FC, Kozar DJ, Fu Y, Dong X (2026) Climate-induced range shifts support local plant diversity but don't reduce extinction risk. Science, 392:648- 654.

https://doi.org/10.1126/science.aea1676

Widmer A, Lexer C (2001) Glacial refugia: sanctuaries for allelic richness, but not for gene diversity. Trends Ecol Evol 16(6): 267–269

Williams L, Zazanashvili N, Sanadiradze G, Kandaurov A (Eds) (2006) Ecoregional conservation plan for the Caucasus. Tbilisi.

Willis KJ, Rudner E, Sümegi P (2000) The full-glacial forests of central and southeastern Europe. Quat Res 53(2): 203-213

Whitlock M, McCauley D (1999) Indirect measures of gene flow and migration: FST≠1/(4Nm+1). Heredity 82, 117–125.

https://doi.org/10.1038/sj.hdy.6884960

Worth J, Grimm G, Tokuko I-U, Li P, Aristotelis CP, Simeone MC, Shanjani PS et al (2025) Whole chloroplast genomes reveal a complex genetic legacy of lost lineages, past radiations and secondary contacts in the dominant temperate deciduous tree genus *Fagus*. bioRxiv 2025.06.03.653586

https://doi.org/10.1101/2025.06.03.653586

Xing Y, Ree RH (2017) Uplift-driven diversification in the Hengduan Mountains, a temperate biodiversity hotspot. Proc Natl Acad Sci USA 114:3444-E3451.

https://doi.org/10.1073/pnas.1616063114

Zazanashvili N, Gagnidze R, Nakhutrishvili G (2000) Main types of vegetation zonations on the mountains of the Caucasus. Acta Phytogeogr Suec 85:7-16

Figures

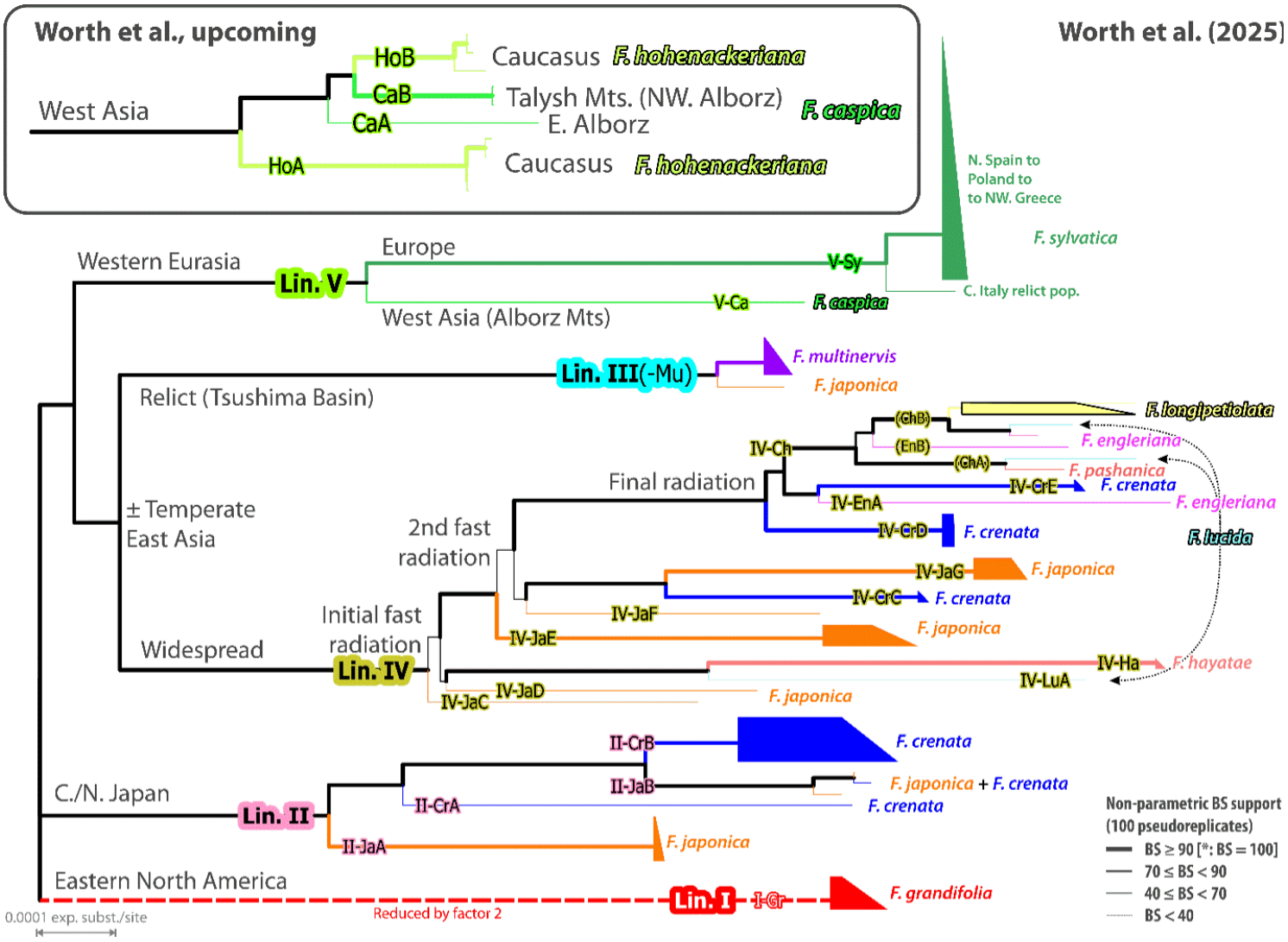


Figure 1

Plastome genealogy of *Fagus* as inferred from complete chloroplast genomes (modified from Worth et al., 2025, work in progress). The five main plastid lineages (I–V) are mostly decoupled from morphological and nuclear-genetic coherent species and subgenera, but show a species-level internal differentiation, some of which predates the formation of the modern-day species (I-Gr, *F. grandifolia*; II-JaA, *F. japonica*;... denote 'species-level plastid types', SLPTs, as defined by Worth et al., 2025, work in progress). The Western Eurasian Lineage V represents one out of two lineages restricted to species sharing a common ancestor exclusive to them; its SLPTs appear to be strictly sorted by species: V-Sy in the European *F. sylvatica,* V-HoA/B in *F. hohenackeriana*, and V-CaA/B in *F. caspica* (no complete plastome data available for *F. orientalis*).

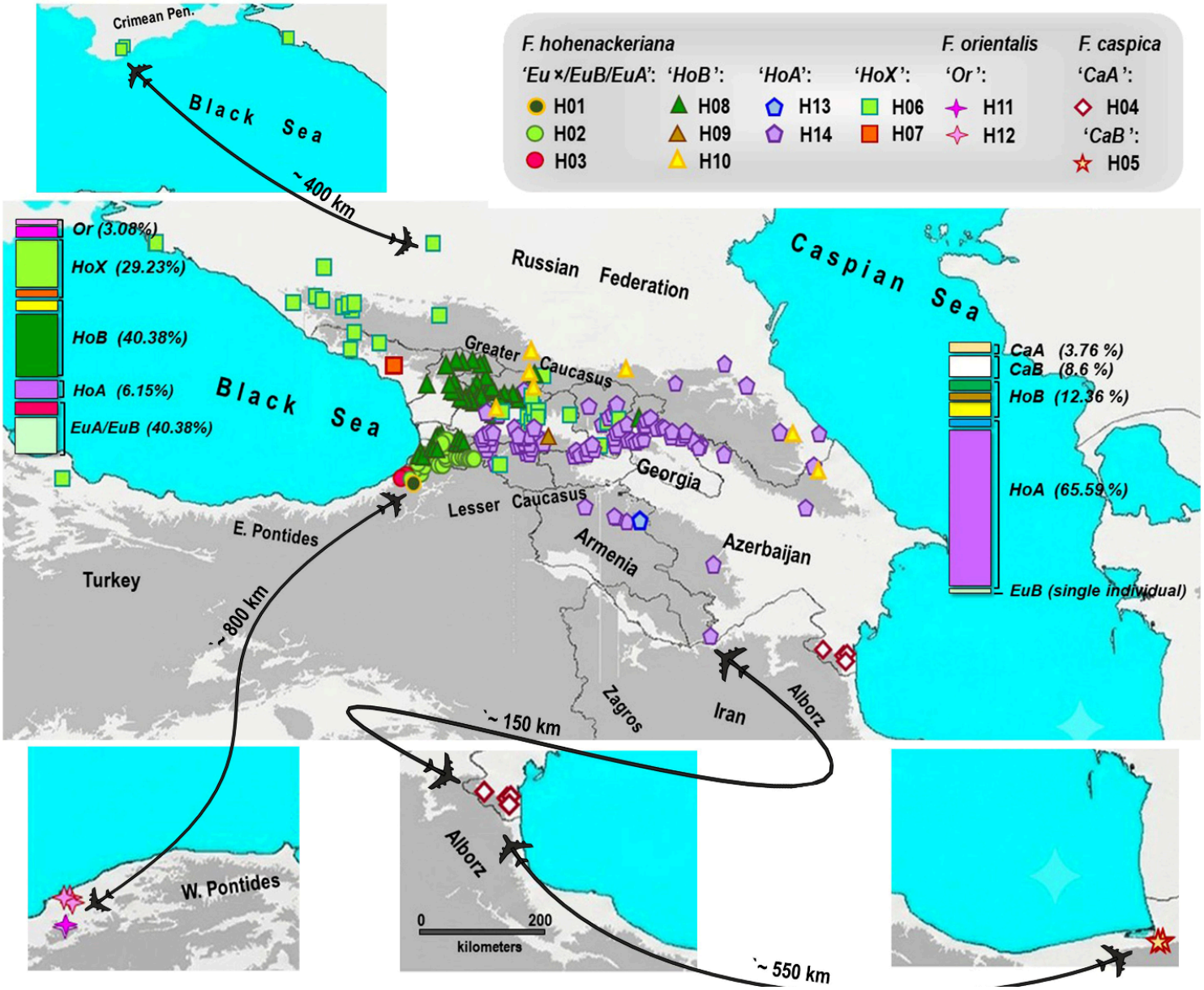


Figure 2

Geographic distribution of sampled *Fagus* stands (colored symbols) across the Caucasus and adjacent regions, together with haplotype composition within the Black Sea and Caspian Sea drainage systems. Colored symbols represent distinct haplotypes and haplogroups. Stacked bar charts show the relative haplotype distribution within each drainage system, with percentages indicating the proportion of *F. hohenackeriana, F. orientalis,* and *F. caspica* Species-Level Plastome Types (SLPTs) across the respective drainage systems. Approximated geographic distances (in kilometers, km) between regions and lineages are indicated by black arrows.

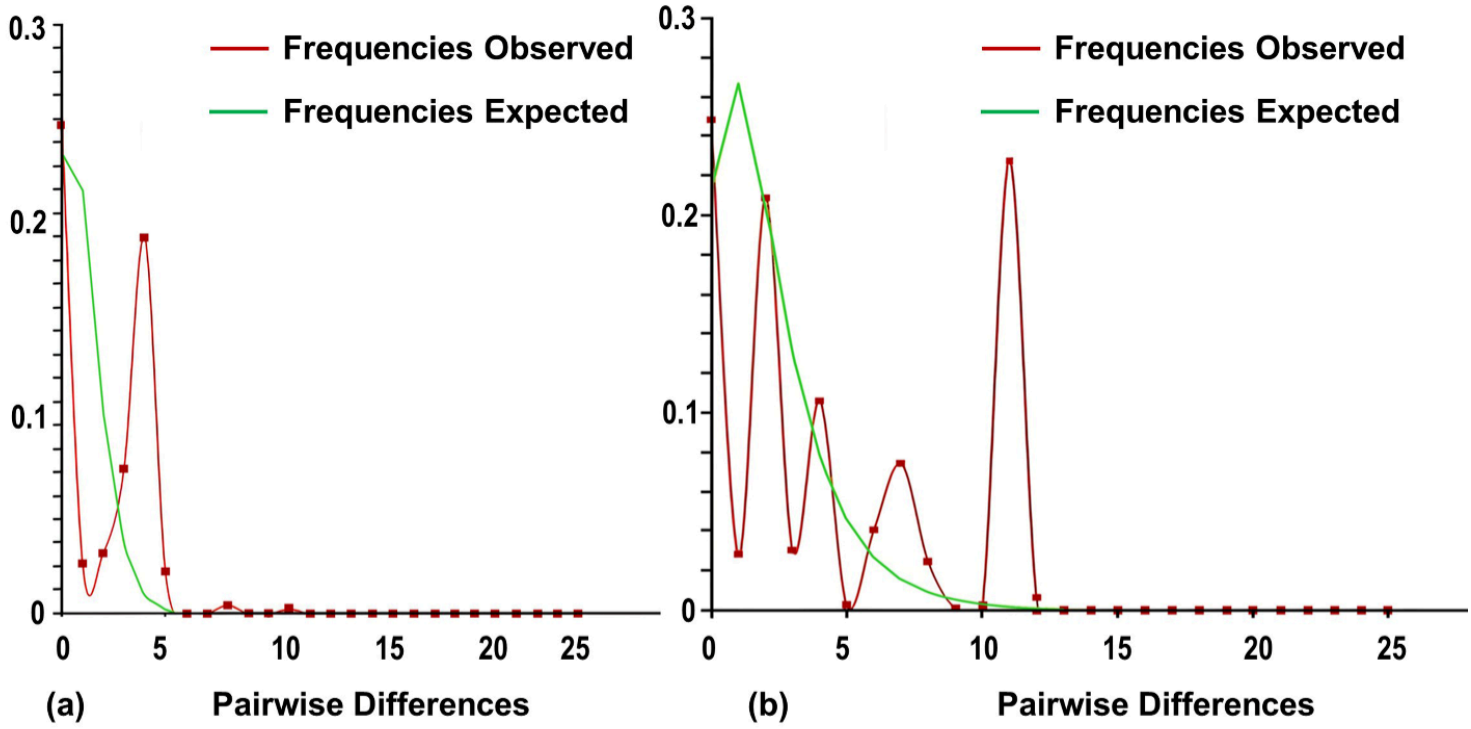


Figure 3

Observed and expected mismatch distribution patterns based on pairwise cpDNA sequence differences in *Fagus* populations from the Caucasus and adjacent regions for the Black Sea (**a**) and Caspian Sea (**b**) drainage systems.

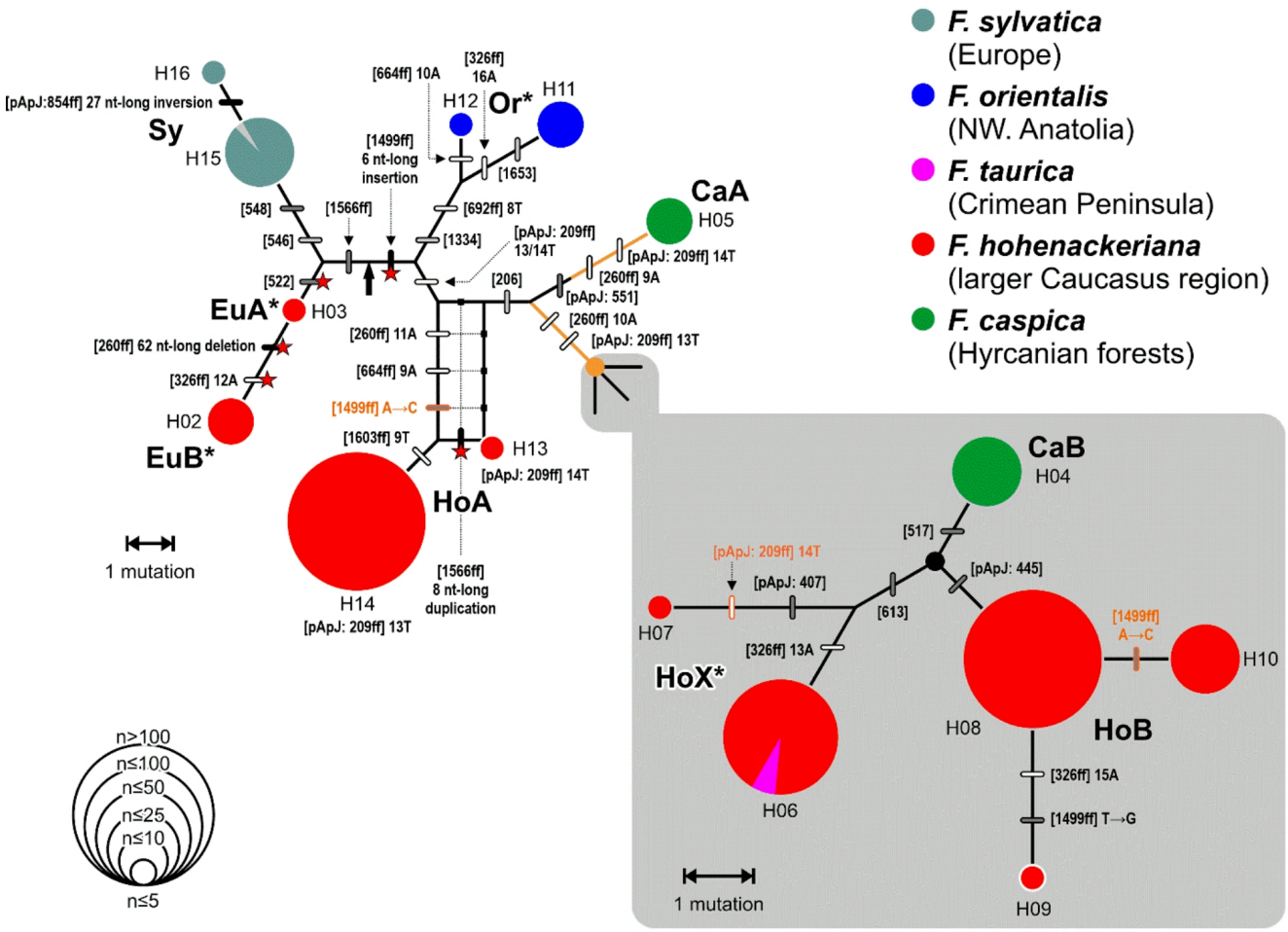


Figure 4

Weighted (quasi-)full median network of plastid haplotypes representing the major western Eurasian *Fagus* lineages constructed from weighted mutational characters derived from an emended complete plastome alignment. Circle size is proportional to haplotype frequency, and colors indicate taxonomic and geographic groups: *F. sylvatica* (Europe), *F. orientalis* s. str. (Northeastern Anatolia), *F.* x *taurica* (Crimean Peninsula), *F. hohenackeriana* (Caucasus region), and *F. caspica* (Hyrcanian forests). Single-nucleotide polymorphisms (SNPs) are classified as transitions (light gray bars), transversions (dark gray bars), indels (insertions, deletions, duplications, and inversions; black bars), and polynucleotide length polymorphisms (PNLPs; white bars). Red stars indicate mutations present in the excluded recombinant *Eu*^{X} haplotype; orange labels and outlines indicate non-sorted mutations, including a putative convergent change and a possible back mutation; and dark yellow edges indicate alternative sorting of polymorphic ancestral states. The black arrow indicates the inferred ancestral node, placed according to mutation mapping onto the complete plastome phylogeny. Asterisks indicate Species-Level Plastome Types (SLPTs), which are newly recognized and not covered by whole plastid genome data (Worth et al., 2025, work in progress).

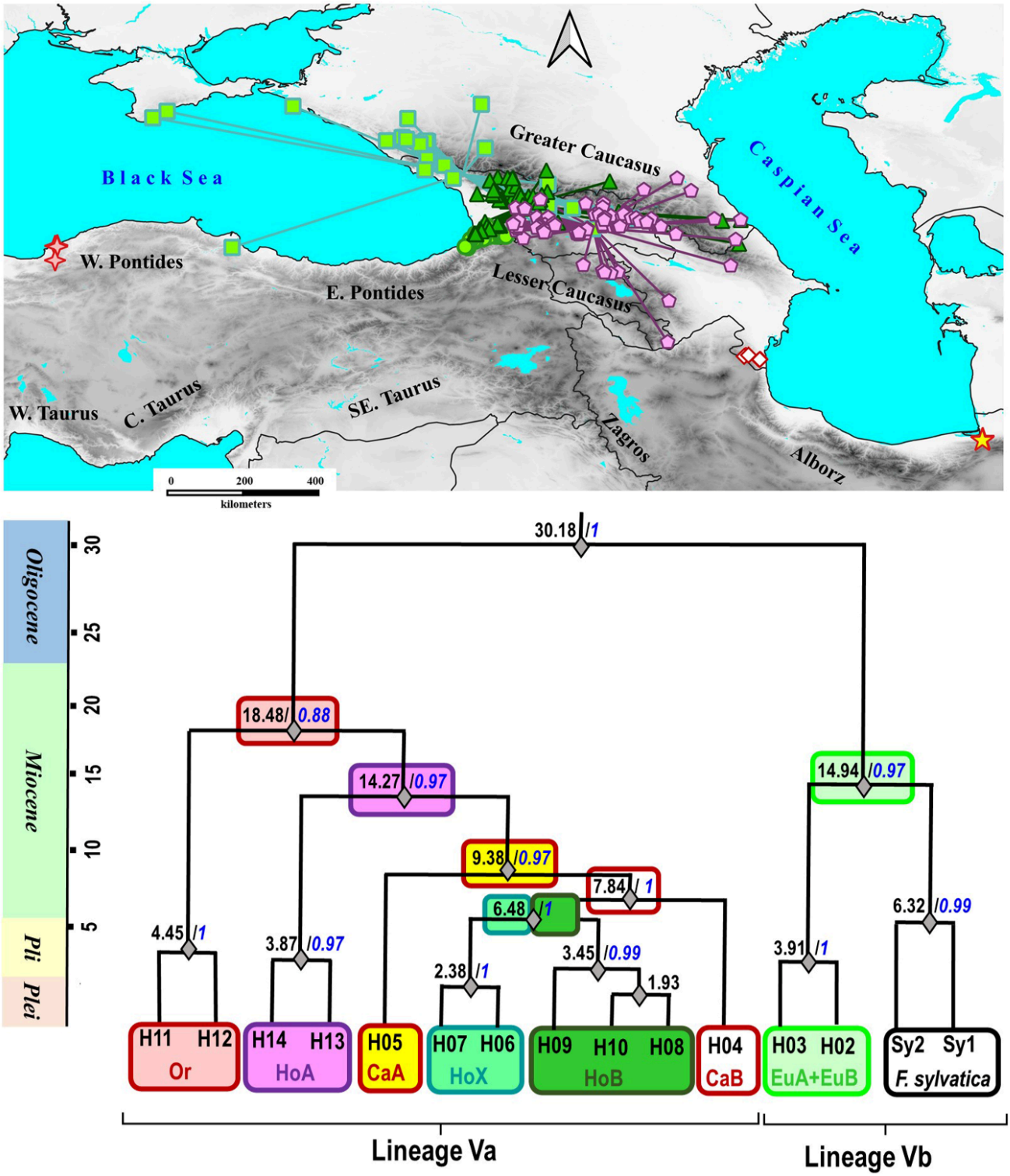


Figure 5

Combined Bayesian BEAST phylogeny of *Fagus*Species-Level Plastome Types (SLPTs; see details in Fig. 4) and the contemporary spatial mapping of sampled *Fagus* stands (colored symbols) across the Caucasus and adjacent regions. Major mountain systems are labeled. SLPTs are uniquely colored and symbol-assigned as follows: '**CaA'**- yellow five-point star with dark-red outline; '**CaB'** – white rhombus with dark-red outline; **'EuA'**, **'EuB'** – lime green cycle with forest-green contour; **'HoA'**- light violet

pentagon with dark-violet outline; '**HoB'**- forest-green triangle with conifer-green outline; '**HoX'**- lime-green quadrat with deep sea green outline; '**Or'**- salmon-pink four-point star with dark-red outline. The node ages are indicated as height divergence time estimates (Ma, million years ago) derived from a penalized likelihood modification of the initial BEAST median-height strict-clock chronogram. Abbreviations: ***Pli,*** Pliocene; ***Plei,*** Pleistocene.

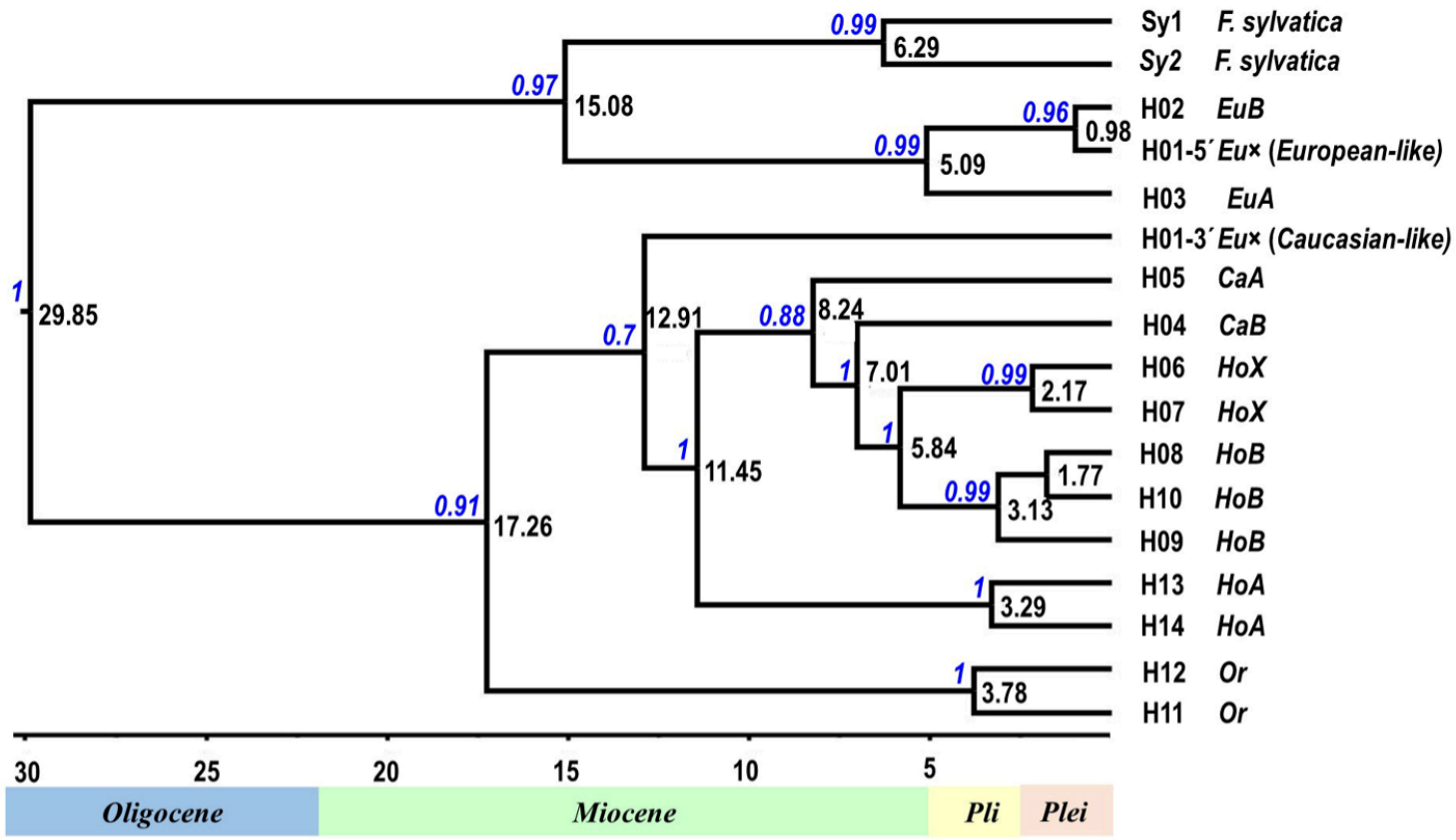
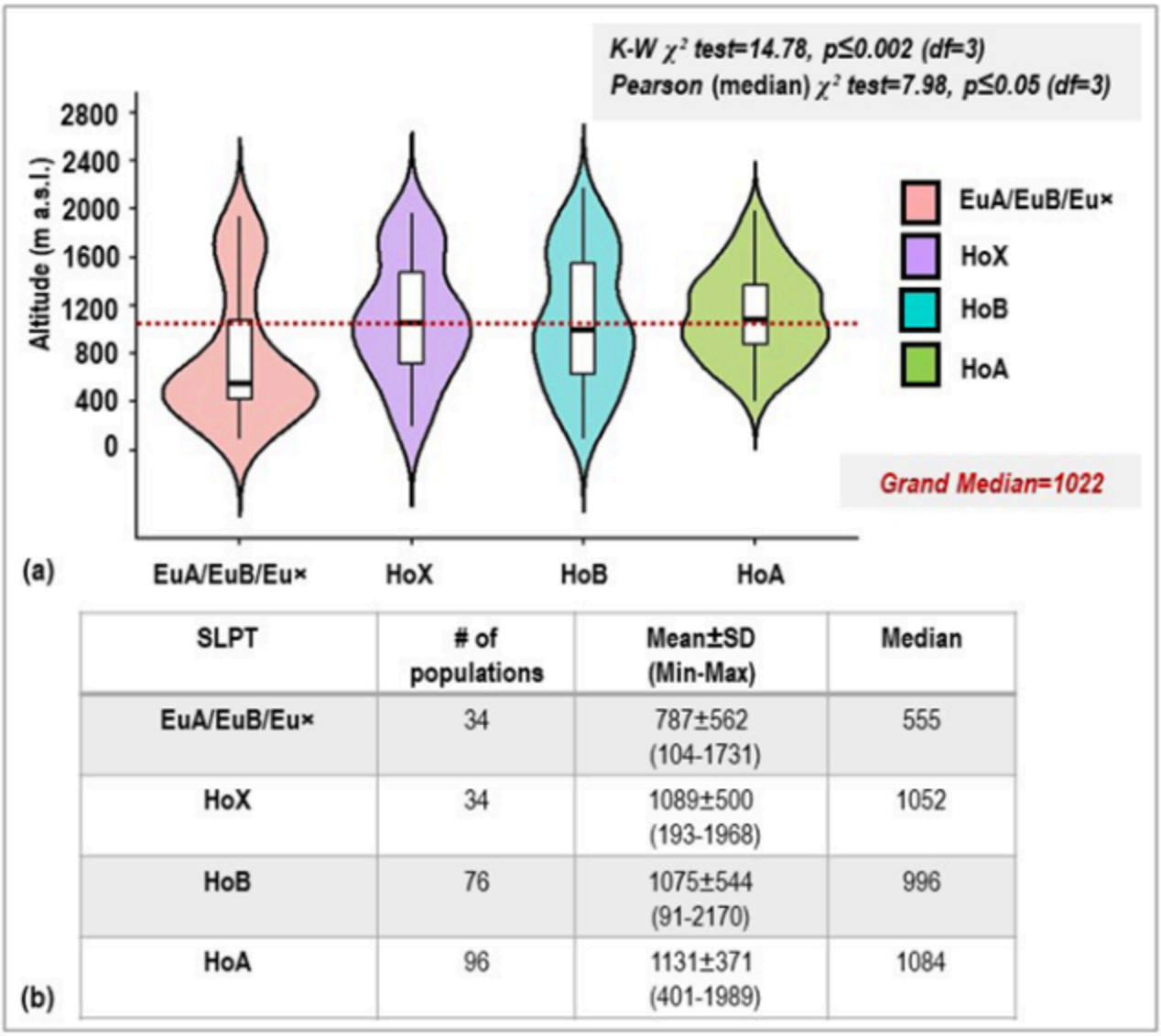


Figure 6

Modified Bayesian BEAST strict-clock chronogram inferred from all 14 plastid haplotypes representing the major western Eurasian *Fagus* maternal lineages across the Caucasus and adjacent regions. Haplotypes are grouped into Species-Level Plastome Types (SLPTs). The height divergence time estimates (Ma, million years ago) derived from a penalized likelihood modification of the initial BEAST median-height strict-clock chronogram are shown in black below the nodes, and posterior probabilities are shown in blue above the nodes. To adapt the mosaic (recombinant) structure of H01, it was divided into 5' and 3' regions and analyzed as separate operational taxonomic units (OTUs) within a single BEAST analysis. The missing regions were coded as question marks, allowing each fragment to be placed independently within the phylogeny.

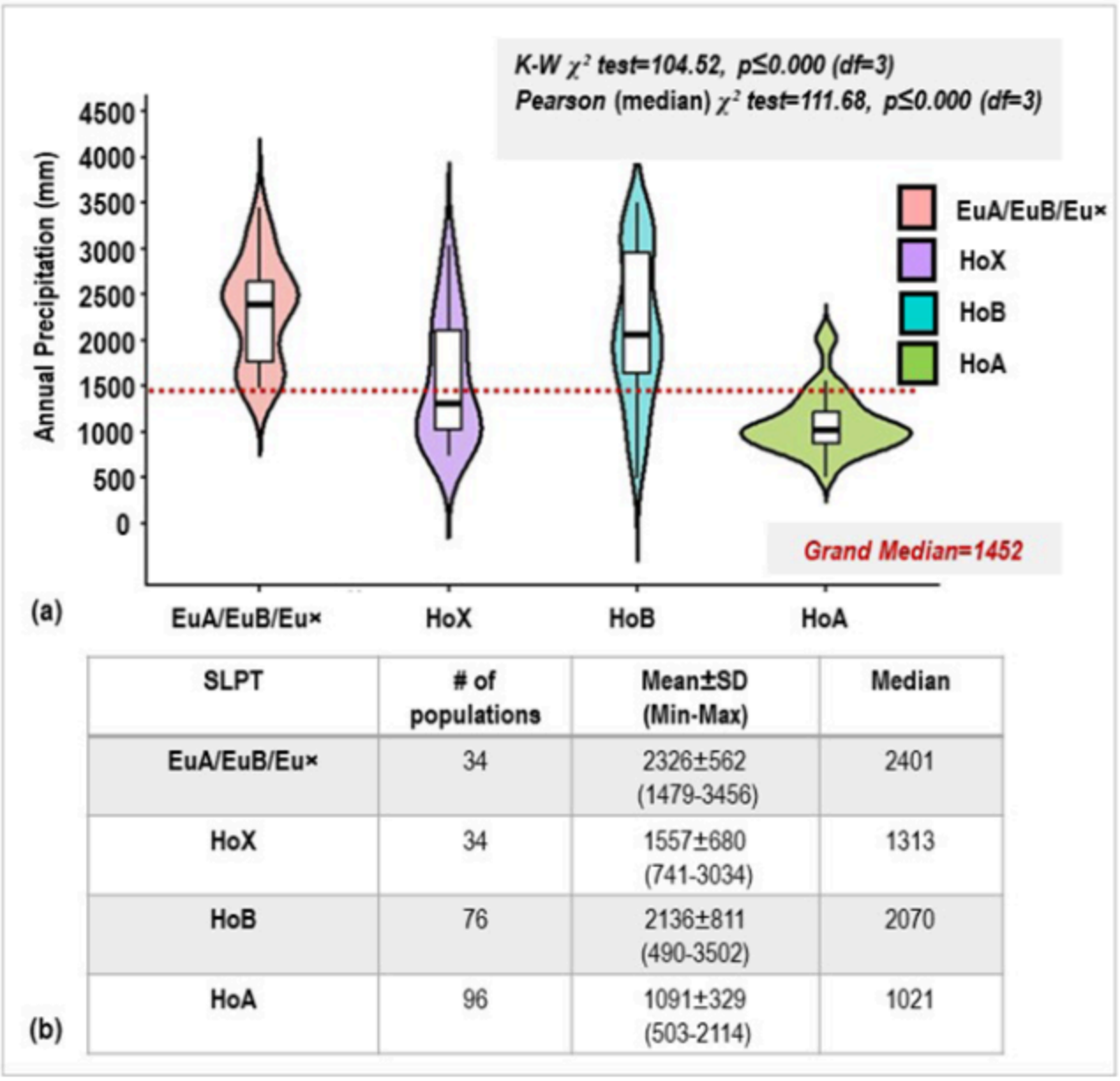


(b)

SLPT	# of populations	Mean±SD (Min-Max)	Median
EuA/EuB/Eu×	34	787±562 (104-1731)	555
HoX	34	1089±500 (193-1968)	1052
HoB	76	1075±544 (91-2170)	996
HoA	96	1131±371 (401-1989)	1084

Figure 7

(a) Ecological differentiation among the Euxinian-Caucasian *F. hohenackeriana* Species-Level Plastome Types (SLPTs) along the altitudinal gradient. Violin plots show the density of altitude values across populations assigned to each SLPT, with inset boxplots indicating the median and dispersion. The dark-red dashed horizontal line represents the grand median along an altitudinal gradient (m a.s.l.) across populations. Significant differences among SLPTs were assessed using both the Kruskal–Wallis test and Pearson's median test; **(b)** a summary table presents the number of populations, mean ± standard deviation (*SD*), range (*min–max*), and median for each SLPT along the altitudinal gradient.



(b)

SLPT	# of populations	Mean±SD (Min-Max)	Median
EuA/EuB/Eu×	34	2326±562 (1479-3456)	2401
HoX	34	1557±680 (741-3034)	1313
HoB	76	2136±811 (490-3502)	2070
HoA	96	1091±329 (503-2114)	1021

Figure 8

(**a**) Ecological differentiation among the Euxinian-Caucasian *F. hohenackeriana*Species-Level Plastome Types (SLPTs) based on annual precipitation gradient (mm). Violin plots show the density of annual precipitation values across populations assigned to each SLPT, with inset boxplots indicating the median and dispersion ranges.

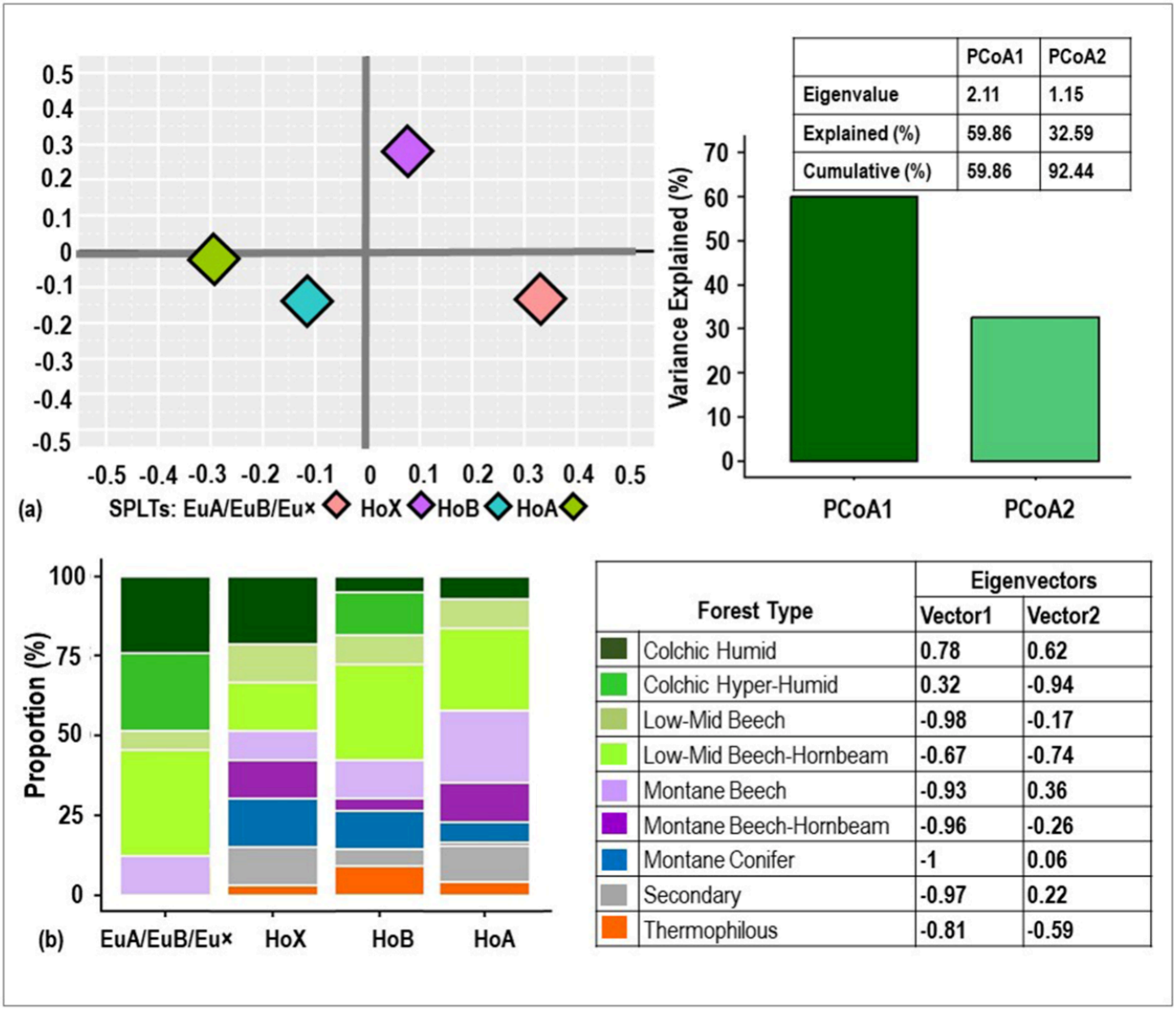


Figure 9

(**a**) Principal Coordinates Analysis (PCoA) ordination based on the Bray–Curtis dissimilarity algorithm among the Euxinian-Caucasian *F. hohenackeriana* Species-Level Plastome Types (SLPTs) derived from ecological gradient composition, including eigenvalues and the percentage of variation explained by the first two PCoA axes; (**b**) stacked bar plots showing the proportional composition of ecological gradients within each SLPT, together with the eigenvector table summarizing correlations between ecological variables and the eigenvectors.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupplementaryFileS1Ekhvaiaetal2026.xlsx