Effects of mountain uplift and climate change on phylogeography and species divergence of East Asia Morella

cai zhao

zhaocai_11@163.com

Guizhou University

Yu Xia Lu

Guizhou University

Shan Shan He

Guizhou University

Chun Xue Jiang

Guizhou Agricultural College: Guizhou University

Jian Feng

Guizhou University

Li Hong Zhao

Guizhou University

Yue Li

Guizhou University

Yu Ting Chen

Guizhou University

Research Article

Keywords: Morella, Genetic diversity, Species differentiation, Demographic dynamics, Management and protection

Posted Date: May 20th, 2025

DOI: https://doi.org/10.21203/rs.3.rs-6558644/v1

Version of Record: A version of this preprint was published at Journal of Plant Research on November 17th, 2025. See the published version at https://doi.org/10.1007/s10265-025-01675-z.

Abstract

Mountain uplift and Quaternary climate oscillations have profoundly influenced plant species' distribution and diversification, yet their impacts on demographic history and biogeographic patterns remain unclear. This study investigates the effects of habitat fragmentation and climatic shifts on genetic diversity and phylogeographic distribution of four East Asian *Morella* species. Using chloroplast DNA (cpDNA) sequences and simple sequence repeats (SSR) were used to study the species divergence and genetic structure of *Morella* from 477 individuals of 63 populations. The whole-genome resequencing was also applied to ensure the accuracy of the estimation of species differentiation time and phylogenetic relationship. We identified species-specific haplotypes, only H2 haplotype was shared by *M. rubra* and *M. adenophora*, and H3 was shared by *M. esculenta* and *M. rubra* in cpDNA sequence. Phylogenetic analysis revealed a topology of *M. esculenta* + (*M. nana* (*M. rubra* + *M. adenophora*)), with significant gene flow among species. Its divergence occurring between 5.02 and 12.72 Ma was completed before the Quaternary period. Results suggest Late Miocene-Pliocene geological and climatic shifts drove speciation, while Quaternary climate fluctuations shaped their geographic distribution, with potential refugia maintaining genetic diversity. Our findings highlight the roles of orogeny and paleoclimate in speciation and range dynamics, providing insights into East Asia's history of lineage differentiation.

Introduction

East Asia is known as the "museum" of plants, which is rich in temperate flora diversity in the world (López-Pujol et al. 2011; Wang et al. 2015; Lu et al. 2018). At the same time, the central and southern mountainous areas of China are the main centres of endemic plants, and the stability of Tertiary may provide favorable conditions for the long-term reproduction of residual plant lineages(López-Pujol et al. 2011). The mountainous regions in southern China are complex and diverse, with vast territory and distinct seasons (Wu et al. 2023), which can provide a variety of habitats and maintain species diversity in a short distance, especially under unfavorable climatic conditions (Hoorn et al. 2013; Favre et al. 2016). What is important is that mountain uplift and climate change will greatly affect the geographical distribution and species divergence of plants (Liu et al. 2013; Hickerson et al. 2010; Zheng et al. 2021), and the systematic geographical pattern reflects the influence in the process of neutral evolution of species (Avise 2000; Hewitt 2004). The collision between the Indian plate and the Eurasian plate led to the uplift of the Qinghai-Tibet Plateau (QTP) (Zhao et al. 2025), which greatly affected the geological features and habitat structure of the region. The mountains, deep valleys and rivers formed with the uplift of the plateau may be obstacles to the migration and diffusion of plant species, which have profoundly affected the geographical distribution pattern and genetic structure of plant species in this area (Wen et al. 2014; Ickert-Bond and Renner 2016). Adjacent to the eastern part of the QTP, the changes of climate and topography will be influenced by the Qinghai-Tibet Plateau to a greater or lesser extent. Moreover, the geographical environment in southwest China is extremely complex, which has always been the focus of global biodiversity research. Complex topography and local selection will also lead to species divergence and genetic differences (López-Pujol et al. 2011; Wambulwa et al. 2022; Hu et

al. 2022). In addition, the violent movement of the QTP created the unique topographical and environmental characteristics of the Yunnan-Guizhou Plateau (Xie et al. 2017). The strong orogeny in Yunnan-Guizhou Plateau resulted in the landform pattern of high in the west and low in the east, which separated it from the surrounding areas (Dai et al. 2013; Wang et al. 2014).

Topographic heterogeneity and orogeny will form new habitats, promote the formation of species and affect the distribution pattern of species (Yin et al. 2021). The orogeny movement will lead to the regional transfer of species, which will lead to the fragmentation of a large area of habitats into smaller patches, that is, habitat fragmentation (Sun et al. 2003; Xie et al. 2017). Climate change usually leads to the expansion or contraction of the protected areas of plant habitat, which is an important factor affecting the geographical distribution pattern of species (Yang et al. 2022; Liu et al. 2023). The impact of climate change on each species is independent and not directly related. Species at the edge of geographical distribution are more sensitive to climate change (Taberlet et al. 1998; Matías et al. 2017). The origin and evolution of species are generally closely related to many historical events, such as the uplift of mountains and different evolutionary adaptations of species caused by climate change (Lowry et al. 2008; Liu et al. 2013). The uplift of the Qinghai-Tibet Plateau not only changed the topography of East Asia, but also led to great changes in the climate (Raymo and Ruddiman, 1992; Wang et al. 2010). It leads to the change of surrounding environment and affects the formation and divergence of species.

Morella is a perennial evergreen shrub or tree of Myricaceae, which grows mostly in valleys and forests. Only four species are distributed in East Asia: *Morella rubra* Lour, *Morella nana* (A. Chev.) J. Herb, *Morella adenophora* (Hance) J. Herb. and *Morella esculenta* (Buch. - Ham. ex D. Don) I. M. Turner (Herbert,2005; Liu, 2016). The research of Herbert (2005) confirmed that four species distributed in East Asia have high bootstrap support in phylogenetic clades. *M. esculenta* is at the base of four species, and the other three species present a complex group relationship (Figs. S1). Chen's (2014) research on nuclear gene ITS and cpDNA (*psb*A-*trn*H) has not resolved the relationship of the remaining three species. Therefore, in order to determine the relationship of four *Morella* species distributed in East Asia, it is necessary to study by various means. Liu (2016) resolved the relationship of four species, but did not resolve the issues of historical dynamics and genetic divergence.

Morella rubra is widely distributed in East Asia, and is often cultivated for its edible value. At present, a large number of cultivated varieties have been cultivated in China, such as "Crystal", "Dongkui" and "Big Leaf and Fine Pedicel". *M. esculenta* is mainly distributed in subtropical and tropical areas from East Asia to Southeast Asia, and is found in many provinces of China. *M. adenophora* is mainly distributed in Guangdong, Guangxi and Hainan Island. *M. nana* is mainly distributed in the higher altitude areas of Yunnan-Guizhou Plateau. *M. nana* and *M. adenophora* are endemic to China (Chen, 2014; Liu, 2016). The young clades of *M. esculenta* and *M. adenophora* are easily distinguished by dense hairs, the short clades and petioles of *M. nana* and *M. rubra* are glabrous or sparsely hairy, and there are dense golden glands on the back of *M. rubra* (Zhong and Xie, 2010; Chen, 2014; Jia, 2016). The plants of this genus have high medicinal and edible value, and can be eaten fresh and made into fruit wine. They can resist bacteria and inflammation, treat rheumatic pain, and resist oxidation and glycosylation (Zhou and Yang,

2000; Sun et al. 2013; Yang et al. 2015; Bai et al. 2020). With the beneficial value of species being known and valued by people, as well as environment change and human intervention, the wild germplasm resources of *Morella* have been severely degraded, and the number of individuals in the population is low. *M. adenophora* has been classified as a threatened species or a vulnerable (VU) species. Wild species play a vital role in the breeding programme, because of their wide variability in morphological structure, resistance and quantitative traits (Laidò et al. 2013; Wang et al. 2017). Genetic diversity is the basis of ecosystem diversity, and it is also the basis of species to survive, adapt to the environment and evolve in the process of evolution (Liu et al. 2022). Understanding the geographical distribution and genetic information of species is very important for the protection of wild germplasm resources (Shahzad et al. 2017). However, the previous studies on *Morella* mostly focused on one species, for example, Liu et al. (2013) and Ju et al. (2023) studied only one species of *M. rubra*. However, there is little research on the population history and genetic differences of *Morella*. Although Liu (2016) studied the genetic relationship of four species based on some individuals, this study is to expand the sample size of species and comprehensively analyze the reasons that affect the lineage geography and species divergence of four *Morella* species in East Asia by combining population genetics.

Here, this study uses SSR markers, cpDNA markers and whole-genome resequencing analysis to solve the following problems:1) to obtain the genetic structure and evolutionary history of four species;2) infer the history of gene flow and population statistics, and demonstrate the formation of the current distribution pattern;3) 3) to explore geological movement and climate change on the divergence of four *Morella* species; 4) to put forward some reasonable suggestions on the utilization of *Morella* resources and make protection strategies.

Material and methods

Sample collection

In the present study, 404 individuals from 62 populations of 4 species were collected in the south of China for cpDNA analysis (Table S1 and Fig. 1a), and Casuarina equisetifolia was taken as an outgroup. 477 individuals from 63 populations were collected for SSR analysis (Table S1). On the basis of cpDNA and SSR analysis, in order to further determine the phylogenetic relationship among the four species, On the basis of cpDNA and SSR analysis, in order to further determine the phylogenetic relationship among the four species, 12 individuals from 12 populations, including four species (each species contains three populations) (Table S1), were used for whole-genome resequencing analysis, with *Rhoiptelea chiliantha*, *Pterocarya stenoptera*, *Quercus fabri*, *Betula luminifera* and *Casuarina equisetifolia* as outgroups to further verify the phylogenetic relationship and divergence time. The collected materials cover the entire geographical distribution of four species. It was found that the wild population of *Morella* was small, so some natural populations and wild individuals that could be collected in the wild were relatively few, and *M. adenophora* was a threatened or vulnerable species (VU). Each population includes 2 ~ 13 individuals, and only GDTH has one individual. Each sample is spaced at least 20 meters apart, and healthy young leaves are dried with silica gel for DNA extraction. All specimens are kept in the herbarium of College of

Life Sciences, Guizhou University (sample numbers: zhaocaiM1 ~ zhaocaiM63). The plants collected in this study were identified by Dr. Zhao Cai (associate professor of Guizhou University, research on plant systematics and evolution). Figure 1a and Table S1 provide detailed information of latitude, longitude and altitude of all populations.

It should be emphasized that although the collected samples are unbalanced among populations, it will not affect our phylogeographic research results. The samples of similar phylogeographic studies are also unbalanced. For instance, sample sizes for *Tetrastigma hemsleyanum* (individuals from 6 to 15), *Notopterygium* (individuals from 2 to 17) and *Machilus thunbergii* (individuals from 2 to 8) were unbalanced (Wang et al. 2015; Shahzad et al. 2017; Fan et al. 2022). However, they can present the entire geographical distribution range and representative sample the species and can cover the geographical distribution of representative populations.

DNA extraction and sequencing

Total DNA was extracted by plant DNA extraction kit (Tiangen, Beijing, China). Put the processed materials in a mortar, add a small amount of silicon dioxide, and liquid nitrogen, quickly grind them into powder, and transfer them into a 2 mL centrifuge tube. For the extraction method, refer to the kit steps. We used a 1.5% agar gel to detect the quality of extracted DNA, and the ideal sample was used in the follow-up experiment. In this study, five chloroplast gene primers (*psb*A-*trn*H, *trn*D-*psb*M, *trn*L-*trn*F, *ycf1*1205-*ycf1*2402 and *ycf1*3125-*ycf1*4381) (Table S2) (Liu 2016), and five pairs of SSR primers (Table S3) (Zhang et al. 2009; Chen 2014; Liu 2016) were screened out by consulting relevant references and pre-testing (Firstly, some samples were randomly selected for PCR amplification, and the PCR products were detected by 1.5% agarose gel electrophoresis, and the primers with clear and single bands were preliminarily screened. Then, the primers were re-screened by polyacrylamide gel electrophoresis to determine whether the primers were stably amplified and polymorphic, and the qualified primers were screened for subsequent experiments.), which were used to amplify more than 400 samples of four *Morella* species. The sequencing reaction is carried out in the Beijing Tsingke Biotech Co, Ltd. The amplification conditions are: 2×Taq PCR Master Mix 12.5 uL, DNA template 1uL, upstream and downstream primers 1uL each, and the remaining volume was made up to 25 uL with ddH_2O. PCR amplification program: pre-denaturation at 94 ℃ for 2 min, denaturation at 94 ℃ for 30 s, annealing at 52 ℃-60 ℃ for 30 s, extension at 72 ℃ for 30 s, and final extension at 72 ℃ for 7 min, the reaction was set to 35 cycles.

Genetic diversity and genetic structure

A number of cpDNA fragments were spliced together with PhyloSuite 1.2.2 (Zhang et al. 2020). The haplotype (*N*), haplotype diversity (H_d) and nucleotide diversity index (π) of each population of the species were identified by DnaSP 6.0 (Rozas 2009). Genetic diversity (expected heterozygosity (*He*), observed heterozygosity (*Ho*), effective allele number (*Ne*), the mean number of alleles (*Na*), percentage

of polymorphic loci (*PPL*) and Shannon's information index (*I*) of *Morella* population were analyzed based on SSR using GenAlex6.5 (Peakall and Smouse 2012). The polymorphism information content (*PIC*) of SSR loci was measured by CERVUS 3.0 (Kalinowski et al. 2007). The Analysis of molecular variance (AMOVA) analysis of cpDNA and SSR was carried out with Arlequin 3.5 (Excoffier and Lischer 2010) to evaluate the genetic variation and divergence among species.

The Structure v2.3.4 (Pritchard et al. 2000) was used to analyze the genetic structure of SSR sequences by Bayesian clustering algorithm, aiming to identify genetic populations with similar allele frequencies. To determine the population number (*K* value), we set the range of *K* from 1 to 10, and perform 10 independent operations for each K value, using a Model called the "Admixture Model". For calculation parameters, set Length of burnin peroid to 100,000 and Number of MCMC Reps after Burnin to 1,000,000. After the operation is completed, we obtain the logarithmic likelihood (*L(K)*) rate of change corresponding to different *K* values. To determine the best *K* value, we follow the method proposed by Evanno et al. (2005) by calculating the Delta K value between successive *K* values. In addition, the PERMUT 2.0 software was used to calculate the mean in-population genetic diversity (*Hs*), total genetic diversity (*Ht*), and genetic divergence coefficients N_{ST} and G_{ST} based on chloroplast DNA data.

The genetic distance of Nei's was calculated by GenAlex 6.5 software, and the Neighbor-Joining (NJ) tree was constructed by cluster analysis by MEGA 6.0 software. In addition, Mantel test was carried out by GenAlex 6.5 software to explore the correlation between Nei's genetic distance and geographical distribution and altitude among species. Based on the genetic distance, the principal coordinate analysis (PCoA) of all populations of four species was further carried out by GenAlex 6.5 software to reveal the genetic structure and potential geographical patterns among populations.

Phylogenetic relationship and estimation of divergence time

The cpDNA sequence was analyzed by MEGA v 6.0 software, and the maximum likelihood (ML) model was bootstrap 1000 times to obtain the phylogenetic clade. The cpDNA intermediate connection network was constructed using Network10.2 (Bandelt et al. 1999) genealogical relationship analysis. ArcGIS v 10.2 was used to map the haplotype geographical distribution of the population. The software package BEAST 1.10.4 (Suchard et al. 2018) was used to estimate the divergence time of *Morella* species. In order to correct the divergence time more accurately, a calibration point was introduced in BEAST analysis. Based on Herbert's (2005) in-depth analysis of systematics and biogeography of Myricaceae, it was determined that the divergence time between *M. esculenta* and three other *Morella* species *M. nana*, *M. adenophora*, and *M. rubra* occurred at about 12.72 Ma (SD = 1.7). In order to determine the best nucleotide substitution model, according to Akaike information criterion, ModelFinder was used to select the model in PhyloSuite 1.2.2 (Kalyaanamoorthy et al. 2017). In the analysis, we adopted the Relaxed Clock Unorrelated Lognormal Method, set the Markov chain to run for 10,000 generations, take samples every 1,500 generations, and set the burn-in value to 10%. In order to ensure the high reliability of the evolutionary tree, Tracer v.1.7.2 is used to test the convergence of the Markov

chain. Finally, the Tree ANNOTATOR 1.10.4 (Suchard et al. 2018) is used to obtain the phylogenetic tree of Maximum Clade Credibility with the greatest reliability, and FigTree 1.4.2 is used to **check** the **beautification** tree. Finally, the differentiation time of *Morella* species was deduced.

In order to verify and compare the reliability of divergence time of four species based on cpDNA data, on this basis, three individuals of each species and five outgroups were selected for whole-genome resequencing, and the sequencing reaction was completed in Shanghai Personal Biotechnology Co, Ltd, and the results of 17 individuals were extracted from the genome sequence based on GeneMiner software (Xie et al. 2024), the angiosperm 353 gene (AGS) set was extracted from the genome sequence to construct phylogenetic tree and estimate the divergence time. AGS is composed of 353 universal low-copy nuclear genes, which were identified by systematic comparative analysis of more than 600 angiosperms, these genes can be widely used in phylogenetic research and population genetics of various taxonomic scales (Zhang et al. 2022). We used GeneMiner software (Xie et al. 2024) to analyze the sequencing data, which can effectively mine phylogenetic molecular markers from transcriptomics, genomics and other NGS data. Gene mining is carried out by GeneMiner software. reads of the target gene are screened from the original data, and the candidate genes are generated by sequence assembly. After quality control, the low-quality data are eliminated. The low-copy nuclear genes of each sample were merged and trimmed, and only the genes with the maximum difference rate (Max. Diff) less than 10% were retained, and then merged and trimmed again. Maximum Likelihood (ML) method was used to construct the phylogenetic tree, and the Bootstrap value was set to 1000 to analyze the statistical support.

Inference of Population history

The mismatch distribution of cpDNA sequences of four *Morella* species was detected by DnaSP 6.0, and the population expansion was tested by Arlequin v 3.5 and Tajima's D (Tajima, 1989), Fu and Li' s F* (Fu and Li, 1993) and Fu and Li' s D* (Fu and Li, 1993). In addition, in order to evaluate the continuity of species population dynamic changes and avoid errors caused by transient occurrences, we used BEAST v 2.7.6 (Bouckaert et al. 2019) to conduct Bayesian skyline plots (BSPs) for analysis, to obtain the dynamic changes of population size over time. In this analysis, we use the average base mutation rates (2×10^{-9} substitutions/site/year) previously approximated by Wolfe et al. (1987) for angiosperm chloroplast DNA ($1 \sim 3 \times10^{-9}$ substitutions/site/year), alongside a population coalescent Bayesian Skyride model for the prior tree and strict molecular clock. To ensure robust inference, we use random initial tree, linear model and setting the length of MCMC chain to make ESS ≥ 200 (five cpDNA haplotypes are 10,000,000 chains). To assess the reliability of the results, we conducted three separate analyses and combined their outcomes using TRACER 1.5, a tool designed for summarizing and visualizing the output of BEAST analyses. Ultimately, we leveraged the R programming language to analyze and visualize the BSPs, providing insights into the historical dynamics of the population size over time. We also used BARRIER 2.2 (Manni et al. 2004) to analyze whether there is an intergroup gene flow barrier in the four species distribution areas. In order to assess whether the four species suffer from

recent Bottleneck effects, the BOTTLENECK software (Piry et al. 1999) is used for analysis. The software provides three different mutation models to calculate bottleneck effects: the infinite allelic mutation model (IAM), the stepped-mutation model (SMM), and the biphasic mutation model (TPM). If the test results are statistically significant, it indicates that the group may have experienced a bottleneck.

Ecological niche modeling

In order to verify the niche differences among target species, we used MAXENT (Phillips et al. 2006) based on field specimen collection, literature report, Global Biodiversity Information Facility (GBIF; https://www.gbif.org/) and Chinese Virtual Herbarium (CVH; http://www.cvh.ac.cn/), delete the repeated geographical records, and predict the niche model of four species. The bioclimatic environment was obtained by downloading from WordClim database (http://www.worldclim.org/). At first, 19 climatic factors were screened to avoid the interaction among various factors affecting the prediction results. The variables were derived from monthly temperature and precipitation data, enabling the creation of more ecologically relevant parameters (Fava et al. 2020). These bioclimatic indicators capture yearly patterns (such as average yearly temperature and total yearly rainfall), seasonal variations (including yearly fluctuations in temperature and rainfall), and extreme or constraining environmental conditions (for example, minimum and maximum monthly temperatures) (Fava et al. 2020; Wu et al. 2024). Based on the relationship between the high Percent contribution and the high permeability, and Pearson correlation coefficient in SPSS, 10 climatic factors were screened out for niche model prediction, so as to improve the accuracy of prediction. The 10 climatic factors are as follows: *M. esculenta* (bio18, bio4, bio15, bio14, bio11, bio7, bio13, bio12, bio19 and bio16), *M. adenophora* (bio18, bio2, bio15, bio4, bio14, bio6, bio1, bio3, bio19 and bio8), *M. rubra* (bio18, bio4, bio15, bio2, bio6, bio14, bio1, bio11, bio9 and bio6), and *M. nana* please refer to our previous research (Wu et al. 2024). For MAXENT modeling, we use the default parameters for analysis. In order to evaluate the statistical performance of the model, we used the area under the “receiver operating characteristic (ROC) curve” (AUC; Fawcett, 2006). The model with AUC approximately 1 showed that the prediction ability was good. We use ArcMap v10.5 to draw a suitable distribution range.

Results

Genetic diversity analysis

By correcting and splicing five cpDNA fragments (*psb*A-*trn*H, *trn*D-*psb*M, *trn*L-*trn*F, *ycf*11205-*ycf*12402, and *ycf*13125-*ycf*14381) were used to analyze 404 individuals from 62 populations of the four *Morella* species. The total length of cpDNA fragment was 3157bp, of which *psb*A-*trn*H, *trn*D-*psb*M, *ycf1*1205-*ycf1*2402, *ycf1*3125-*ycf1*4381 and *trn*L-*trn*F fragments are 435, 587, 855, 887 and 393bp, respectively. Among the four species, 51 cpDNA haplotypes were detected (Fig. 1; Table 1), of which only H2 haplotype was shared by *M. rubra* and *M. adenophora*, and H3 was shared by *M. esculenta* and *M. rubra*. In addition, all haplotypes had species specificity. In *M. rubra* there were 14 haplotypes, and GXLG and GXSB haplotypes in Guangxi had the highest diversity. Among the 16 haplotypes of *M. nana*, the

population YNPL had the highest haplotype diversity. *Morella esculenta* has 18 haplotypes. The Guizhou populations (GZZJ, GZQL and GZQX) also had the highest haplotype diversity. The *M. adenophora* had five haplotypes, of which H27, H48, H49 and H50 were unique to the population, and H2 was a shared haplotype (Table 1). Genetic diversity analysis indicated that the total genetic diversity (*Ht*) was 0.913, and the average genetic diversity within populations (*Hs*) was 0.205. The genetic diversity of the four species of *Morella* was relatively high. At the species level, the haplotype diversity of the four species was the highest in *M. nana* (H_d =0.819) and the lowest in *M. adenophora* (H_d =0.234). The highest nucleotide diversity was found in *M. esculenta* (π = 0.00266) and the lowest in *M. adenophora* (π = 0.0001) (Table 2).

Genetic diversity analysis was conducted on 477 individuals from 63 populations of four species using five highly polymorphic SSR primers. The results showed that the average genetic diversity index of wild *Morella* species was at a relatively high level (*Na* = 3.53, *Ne* = 2.896, *I* = 1.07, *Ho* = 0.924, *He* = 0.597, *PIC* = 0.844, PPL = 93.33%) (Table S4). A higher *I* value means a higher genetic diversity, indicating that the genetic diversity of YNLY is higher (*I* = 1.671). Based on the above data analysis, the genetic diversity of the populations of YNLY, GXTE, GXTB and GXSS in this study is high, while the genetic diversity of FJCL and YNLC is low. The polymorphism information content (*PIC*) value was related to the degree of genetic variation at the locus, with *PIC* = 0.844, indicating that there might be variation within *Morella* populations (Table S4). The genetic diversity analysis of the five SSR microsatellite loci showed that SSR4 had the highest genetic diversity, while SSR2 showed a low level of genetic diversity. The genetic divergence coefficient (*Fst*) = 0.317, *F* = -0.591 (Table S4), suggesting that there might have been inbreeding or genetic drift among the four species.

Table 1
Population haplotype distribution and genetic diversity parameters of the four *Morella* species with cpDNA.

Species	Population	*N*	*Nh*	*Hd*	*π*	cpDNA Chlorotypes
M. rubra	GZDY	13	1	0.000	0.00000	H4(13)
	GXJA	4	1	0.000	0.00000	H4(4)
	AHGC	7	2	0.286	0.00036	H4(6),H28(1)
	GXLB	10	3	0.378	0.00026	H36(8),H37(1),H38(1)
	GXLG	8	4	0.821	0.00035	H36(3),H38(1),H39(2),H40(2)
	FJCL	5	1	0.000	0.00000	H43(5)
	GZDB	7	1	0.000	0.00000	H4(7)
	GZKY	9	1	0.000	0.00000	H4(9)
	FJXM	5	1	0.000	0.00000	H4(5)
	SCNJ	6	2	0.600	0.00038	H4(3),H44(3)
	ZJLS	5	2	0.600	0.00019	H4(3),H46(2)
	GDMX	3	1	0.000	0.00000	H4(3)
	AHFT	3	1	0.000	0.00000	H4(3)
	GXSB	2	2	1.000	0.00032	H2(1),H3(1)
	GXDC	6	2	0.333	0.00032	H2(5),H47(1)
	GDGN	3	1	0.000	0.00000	H2(3)
	GDTH	1	1	0.000	0.00000	H51(1)
	JXTH	3	1	0.000	0.00000	H4(3)
	ZJNL	5	1	0.000	0.00000	H4(5)
M. nana	GZQG	5	1	0.000	0.00000	H5(5)
	GZPZ	10	1	0.000	0.00000	H6(10)
	GZXR	4	1	0.000	0.00000	H6(4)
	YNPL	11	3	0.636	0.00058	H9(5),H10(1),H11(5)
	YNLQ	2	1	0.000	0.00000	H12(2)
	GZQX	2	2	1.000	0.00286	H5(1),H13(1)
	GZSC	4	1	0.000	0.00000	H5(4)

Species	Population	*N*	*Nh*	*Hd*	*π*	cpDNA Chlorotypes
	SCYB	10	1	0.000	0.00000	H16(10)
	SCRH	12	2	0.409	0.00013	H17(9),H18(3)
	YNDL	9	2	0.222	0.00007	H19(8),H20(1)
	YNNJ	9	1	0.000	0.00000	H19(9)
	YNQB	8	2	0.250	0.00016	H6(7),H21(1)
	YNYL	12	1	0.000	0.00000	H5(12)
	GZZJ	2	2	1.000	0.00032	H5(1),H6(1)
	YNHZ	11	1	0.000	0.00000	H6(11)
	GZWN	8	1	0.000	0.00000	H5(8)
	YNZY	12	1	0.000	0.00000	H5(12)
	YNES	10	2	0.200	0.00013	H22(9),H23(1)
	YNLC	5	1	0.000	0.00000	H12(5)
	GZSY	9	1	0.000	0.00000	H5(9)
	GZLZ	7	2	0.286	0.00037	H6(6),H24(1)
M. esculenta	GXTA	7	3	0.524	0.00036	H8(5),H25(1),H26(1)
	GXTB	9	1	0.000	0.00000	H8(9)
	YNYA	4	2	0.500	0.00080	H29(1),H30(3)
	GZQL	9	5	0.806	0.00048	H31(1),H32(4),H33(2),H34(1),H35(1)
	YNYB	7	1	0.000	0.00000	H1(7)
	SCQA	11	1	0.000	0.00000	H15(11)
	YNLY	9	2	0.222	0.00028	H41(8),H42(1)
	GXTE	5	1	0.000	0.00000	H8(5)
	GXSS	6	2	0.333	0.00021	H8(5),H45(1)
	GZZF	5	2	0.400	0.00026	H7(1),H8(4)
	YNXC	4	2	0.500	0.00144	H14(3),H15(1)
	GXSW	5	2	0.400	0.00153	H3(1),H8(4)
M. adenophora	GXXN	9	1	0.000	0.00000	H2(9)
	GXHZ	10	2	0.356	0.00011	H2(8),H27(2)

Species	Population	*N*	*Nh*	*Hd*	*π*	cpDNA Chlorotypes
	GXFA	3	1	0.000	0.00000	H2(3)
	GXFB	3	1	0.000	0.00000	H2(3)
	GXDA	5	1	0.000	0.00000	H2(5)
	GXDB	7	1	0.000	0.00000	H2(7)
	GXDD	2	1	0.000	0.00000	H2(2)
	GXQB	3	1	0.000	0.00000	H2(3)
	GXQA	11	4	0.673	0.00029	H2(6),H48(3),H49(1),H50(1)
	GXHP	3	1	0.000	0.00000	H2(3)

N, number of samples; *Nh*, Number of haplotypes; H_d, gene diversity; π, nucleotide diversity averaged across loci.

Table 2
Genetic diversity parameters of the four *Morella* species with cpDNA.

species	M	S	Ps	(θw)	H_d	π
M. rubra	105	17	0.005426	0.001038	0.611	0.00084
M. nana	162	26	0.008317	0.001469	0.819	0.00194
M. esculenta	81	39	0.012472	0.002512	0.807	0.00266
M. adenophora	56	4	0.001277	0.000278	0.234	0.0001
All	404	67	0.021433	0.003259	0.917	0.00407

M, sequence number; S, separation bit number; ps = S/m; θw = ps/α; H_d, gene diversity; π, nucleotide diversity averaged across loci.

Genetic divergence and Genetic structure analysis

The AMOVA analysis based on cpDNA data indicated that genetic variation mainly occurred among groups (61.23%) (Table 3). The genetic divergence coefficient Nst was greater than Gst (Nst = 0.804, Gst = 0.775; P < 0.05), suggesting a clear phylogeographic structure among *M. rubra* and *M. adenophora, M. esculenta and M. nana*. A series of analyses using SSR markers further revealed the genetic structure of the four species of *Morella*. The AMOVA analysis of SSR indicated that genetic variation primarily originated from within populations, accounting for 85% of the total variation (Table 3). The results of principal component analysis (PCoA) indicated that although there was a certain degree of genetic overlap among species (Fig. 2), the populations of the four species maintained a relatively clear genetic divergence overall. NJ tree shows that 63 populations can be clearly divided into four main genetic

groups (Fig. 3c). Corresponding to the species, *M. esculenta* and *M. nana* form different clades, and some populations of *M. rubra* and *M. adenophora* are mixed on the phylogenetic tree surface. Meanwhile, Mantel tests demonstrated a positive correlation between genetic distance and geographic distance in Morella (R2 = 0.0332, P = 0.000) (Fig. S2), and also a positive correlation between genetic distance and altitude (R2 = 0.0293, P = 0.000) (Fig. S3). The results suggest that the IBD model supports the role of geographic distance and altitude in driving the genetic divergence of the four species.

Bayesian clustering analysis based on SSR data determined the optimal *K* value, dividing the 63 populations of the four species of *Morella* into four groups: the first group was *M. rubra*, the second group was *M. nana*, the third group was *M. esculenta*, and the fourth group was *M. adenophora* (Fig. 3a and b). Under this cluster number, the genetic composition of each population exhibits specific characteristics, and also reveals the gene exchange between populations. The result of the SSR genetic structure indicate that there were genetic similarities between different species, and the mixing of groups (marked green, red, blue, and yellow, respectively) shows widespread gene exchange, genetic diversity, and potential hybridization events between species (Fig. 3b). Additionally, barrier analysis showed that there were genetic boundaries among the four species, and the boundaries a, b and c could divide the four species as a whole. But there was still gene exchange among some populations of different species, especially among the populations of *M. rubra* and *M. adenophora* (Fig. 4).

Table 3
Analysis of molecular variance (AMOVA) for *Morella* species based on cpDNA and SSR

Sign	Source of variation	Sum of squares	Variance components	Percentage of variation	Fixation Index (*F*st)
cpDNA	Among groups	2454.8	8.13776 Va	61.23	F_{SC}=0.94
	Among populations within groups	1829.99	4.86289 Vb	36.59	F_{ST}=0.98
	Within populations	99.142	0.28989 Vc	2.18	F_{CT}=0.61
	Total	4383.931	13.29053		
SSR	Among group	18.505	0.02321 Va	5.93	F_{SC}=0.10
	Among populations within groups	50.971	0.03551 Vb	9.07	F_{ST}=0.15
	Within populations	295.111	0.33271 Vc	85	F_{CT}=0.06
	Total	364.587	0.39143		

Phylogenetic relationship and estimation of divergence time

In order to determine the phylogenetic position of the four species, in the present study, cpDNA haplotype phylogenetic clade was constructed based on Maximum Likelihood (ML)and Maximum Parsimony (MP) using the genus *Casuarina equisetifolia* was used as outgroup(Fig. 1e). The analysis results indicate that the phylogenetic tree can be divided into four groups, which clearly reveals the interspecific relationship among them. *M. esculenta* is located at the base of the phylogenetic tree, followed by *M. nana*, in which *M. rubra* and *M. adenophora* form a sister group. The results of whole-genome resequencing showed that *M. esculenta* was at the base of the clade, *M. rubra* and *M. adenophora* a form sister group, and then form a sister group with *M. nana* (Fig. 1d). Among cpDNA haplotypes, H2, H4, H5, H6 and H8 with the highest distribution frequency are located at the central position of the network diagram and may be ancestral haplotypes. The remaining haplotypes with lower frequencies are located outside the network diagram, and we speculate that they may be young haplotypes that have diverged from ancestral haplotypes. Moreover, the haplotype sharing phenomenon between *M. rubra* and *M. adenophora* also reveals the close genetic relationship between them (Fig. 1c).

The results of BEAST analysis based on cpDNA data reveal a more detailed divergence timeline. The results showed that *M. nana* was differentiated from *M. rubra* and *M. adenophora* at about 10.45Ma (95%HDP), and began to intraspecific differentiate of *M. esculenta* at 8.38 Ma (95%HDP) in the late Miocene. In addition, further divergence within *M. nana* population began at 8.77 Ma (95%HDP) in the Late Miocene. the divergence of *M. rubra* and *M. adenophora* occurred at 5.02Ma (95%HDP) in the Pliocene, and the intraspecific divergence of *M. rubra* started at 3.8 Ma (95%HDP) in the early Pliocene, and the intraspecific divergence time of *M. adenophora* was the latest at 3.01Ma (95%HDP) in the late Pliocene (Fig. 1e). The time accuracy and relationships among the four species were further estimated using whole-genome resequencing of AGS. As shown in the figure (Fig. 1d), the four species diverged into two clades at 12.72 Ma in the late Miocene, and the *M. esculenta* diverged at the earliest, intraspecific divergence occurring around 6.65 Ma (95%HDP) in the late Miocene. The other three species form a complex group relationship, and *M. rubra* and *M. adenophora* form a sister group. The divergence of *M. nana* and the other two species began at about 10.12 Ma (95%HDP) in the late Miocene and intraspecific divergence occurred at 5.49 Ma (95%HDP) in the early Pliocene. The intraspecific divergence of *M. rubra* occurred at 4.19 Ma (95%HDP) in the Pliocene. The internal bifurcation of *M. adenophora* occurred in the Pliocene at about 4.17 Ma (95%HDP) (Fig. 1d). It is consistent with the species divergence time analyzed by cpDNA data, which further confirms the validity and reliability of these gene sets in studying plant phylogeny and historical evolution.

Inference of Population history

Based on the cpDNA sequence, we conducted various analyses to determine the population history of four *Morella* species. The neutrality test (Table S6) provided that the Tajima's D statistical value of the species was not significantly positive (0.71428, $P > 0.10$), and the Fu and Li's D values (1.85822, $P < 0.02$) were positive at the overall level, suggesting that the population expansion may was not significant in history. For individual species, the statistical values of Tajima's D and Fu and Li's D for *M. adenophora* and *M. rubra* were not significantly negative ($P > 0.10$), indicating that the species may have undergone

expansion during evolution. The statistical value of Tajima's D of *M. esculenta* was not significantly positive, but it cannot fully explain the trend of deviation from expansion in this region. Tajima's D of *M. nana* was not significantly positive (0.90941(P > 0.10)), the Fu and Li's D and Fu and Li's F were both negative and did not reach statistical significance (P > 0.10). These results further support that the species may follow the neutral evolution model and remain relatively stable on the scale of research time. The curve of mismatch distribution analysis (Fig. S4) shows a multimodal distribution pattern at the at the overall level and in an independent single species, which was inconsistent with the expected unimodal distribution, indicating that there was insufficient evidence to support that these species have experienced significant expansion. In summary, the results of neutrality test and mismatch distribution analysis revealed that the population size of the four species of *Morella* may remained relatively stable, and population may have expanded on a small scale and briefly in evolutionary history.

Because of the neutrality test and mismatch distribution analysis detection are instantaneous, which can only explain the state of species in a period of time, but can‘t clearly indicate whether the species has expanded during the evolutionary process. Therefore, the BSPs curve obtained by combining cpDNA data is shown in the figure (Fig. 5), which shows that the effective population of *Morella* has been in the amplification mode since 1.2 Ma. Among them, the expansion trend of *M. nana* mainly occurred in 0.4 ~ 0.05 Ma, and the population size of *M. esculenta* mainly expanded in 0.7 ~ 0.1 Ma, and the population of *M. rubra* expanded since 0.15 Ma. However, the overall expansion of *M. adenophora* was not obvious, and it expanded only after 0.04 Ma (Fig. 5**)**. Bottleneck effect (Table S7) detection showed that under IAM, TPM and SMM models, Bottleneck effect may occur in some populations of 4 *Morella* species, which is of great importance for understanding the genetic history and dynamics of species.

Present and past (Last Glacial Maximum) distribution

In this study, MAXENT has a good prediction effect on the potential distribution of four plants in East Asia (AUV > 0.990), which shows that the model and species distribution information have a high degree of fitting, and the current distribution range covers the existing range of this species. Niche simulation results (Fig. 6) indicated that during the Last Glacial Maximum (LGM), the climate was cold and the survival environment was lost. To adapt to environmental changes, the species formed fragmented habitats, with their distribution ranges contracting to form refuges and presenting a patchy distribution pattern while migrating towards suitable areas, mainly in a southward direction. The results showed that *M. esculenta* was mainly distributed in Yunnan and Guangxi, *M. rubra* was mainly distributed in Guangxi, *M. adenophora* was mainly distributed in Guangxi, and *M. nana* was distributed in Yunnan and Guizhou. This suggests that potential refuges existed in these areas during the ice age. Subsequently, the suitable areas of the species slightly expanded from small fragments to their current distribution ranges. The expansion of *Morella* mainly occurred in the southwestern region of China, suggesting that the main suitable area for the *Morella* is the southwestern region. The species distribution model indicated that the distribution ranges of *M. adenophora*, *M. esculenta*, and *M. rubra* were restricted during the LGM period, but rapidly expanded from the LGM to the present. Interestingly, the distribution range of *M. nana* did not change greatly from the LGM to the present. The knife cut test analysis of the selected

environmental variables indicated that warmest-season precipitation (bio18) was the most influential factor affecting *Morella*'s distribution in both the LGM and present eras, underscoring its significance in determining the species' geographic spread.

Discussions

Genetic diversity and genetic structure

Genetic diversity is the foundation and core of biodiversity, which reflects the adaptive changes made by species in response to environmental changes (Jia, 2016). Genetic diversity is a key factor for the survival and sustainable existence of species in their environment, and species in an unsuitable environment will lead to a decrease in genetic diversity (Wambulwa et al. 2022). Our research revealed that the genetic diversity of *Morella* was high (cpDNA: Hd = 0.917, π = 0.00407; SSR markers: *Na* = 3.53, *Ne* = 2.896, *I* = 1.07, *Ho* = 0.924, *He* = 0.597) (Table 1; Table S4). *Morella* is a perennial woody plant with a long-life span, and different generations can share their habitats, which may play an internal buffer role in preventing the loss of genetic diversity (Liu et al. 2023). At the species level, *M. nana* has the highest haplotype diversity, while *M. adenophora* has relatively low haplotype diversity. The average nucleotide diversity of *M. esculenta* is the highest, which further confirms the richness of its genetic diversity. The genetic diversity of four species in this study may be affected by many potential factors. Firstly, genetic diversity may be affected by the distribution range, and species with a wide distribution range may show high genetic diversity (Liu et al. 2022). In general, the widely distributed *M. esculenta* showed high genetic diversity, while that of *M. adenophora* with a narrow distribution, was lower. Secondly, the longer a species has evolved, the more genetic variation there was (Xu et al. 2021), This study indicated that the genetic variation of *M. esculenta* was high, it's evolutionary history is longer, as evidenced by the origin and differentiation time for the four *Morella* species. Furthermore, climate change and man-made destruction will also affect the genetic diversity of species, for example, the distribution range of *Morella* may shrink during LGM period (Fig. 6), and the effective population size is greatly reduced, which leads to the decrease of genetic diversity (Keppel et al. 2012). In addition, *M. rubra* is trained as a fruit, and it tastes delicious, so the wild species may be damaged to a higher degree, thus affecting genetic diversity.

Many factors work together on the genetic diversity of species, including but not limited to geographical distribution, gene flow and genetic communication among populations (Nybom, 2004). Four *Morella* species are mainly distributed in ecologically diverse areas such as Guangxi, Guizhou and Yunnan. The differences of climate and topography in different regions promote the generation of genetic variation to adapt to their respective niches and environmental pressures. The genetic diversity of *M. esculenta* may be related to its wide distribution in complex and changeable mountain habitats. *M. esculenta* was the first to differentiate, which may also create its rich genetic diversity. At the same time, the existence of shared haplotypes and the exchange of genetic species between species may also promote genetic diversity. In addition, bottleneck effect and population expansion events also play a great role in shaping genetic diversity (Yun et al. 2020). Our research shows that some populations of *Morella* have experienced bottleneck events, and their genetic diversity has also been affected. To sum up, the genetic

diversity of *Morella* is the result of many factors, and the interaction of these factors shapes the level of genetic diversity of *Morella*.

AMOVA analysis (Table 3) showed that the genetic variation of SSR markers mainly occurred in within populations, while that of cpDNA markers mainly occurred in among groups, which was consistent with previous analysis based on cpDNA sequences and SSR markers (Liu et al. 2022;Gao et al. 2022). There may be many reasons for the opposite results of AMOVA. Firstly, cpDNA is mainly maternal inheritance, with lower gene flow, low recombination and mutation rate, which makes the lineage inheritance of species clearer (Comes and Kadereit, 1998; Avise, 2000), whereas SSR is a parental inheritance, and the recombination and substitution rate of parental inheritance is high (Mort et al. 2007), so that the genetic differences between them are different. Moreover, the higher genetic variation between populations may be caused by the complex evolutionary processes that species have undergone over the course of history (Li et al. 2022). Thirdly, the phenomenon that genetic variation mainly occurs within populations mostly occurs in woody plants (Liu et al. 2013). Four *Morella* species are perennial woody plants, and the genetic variation is similar to other woody plants.

The genetic differentiation coefficient Nst based on cpDNA is greater than Gst, which indicates that the four *Morella* species have obvious lineage geographical structure, haplotypes with similar genetic distance are more likely to appear in geographically adjacent regions (Pons and Petit, 1996). Barrier analysis (Fig. 4) showed that there were genetic communication barriers among the four species, and three geographical (a, b and c) boundaries could separate the four species. However, there were genetic communication between *M. adenophora* and some populations of *M. rubra*, which also proved that *M. adenophora* and *M. rubra* were sister group with close genetic relationship. The results of structure showed that *Morella* was divided into four groups when $K = 4$, and some populations among the groups had gene exchange. NJ tree based on Nei's genetic distance method supported the clustering results of structure. Each group contains individuals from different populations, which reveals that there is a certain degree of gene exchange between populations. This genetic mixing may reflect that the genetic differences between populations are not completely independent, but overlap. These findings further confirm the genetic diversity among species and the existence of hybridization events.

Phylogenetic relationship of four Morella species

Accurate species classification and identification plays an important role in conservation biology, ecology and genealogy geography, which directly affects the reliability of phylogenetic research, the accuracy of species evolution analysis and the scientific formulation of conservation measures (Dagnino et al. 2017). In this study, the phylogenetic relationship between four *Morella* species distributed in East Asia was analyzed based on the data of cpDNA and SSR sequencing (Fig. 1e; Fig. 3c). The results showed that the four species formed a monophyletic clade with high bootstrap support. The genetic structure based on SSR sequencing data shows that 63 *Morella* populations can be divided into four distinct genetic groups, which is consistent with their taxonomic classification. The populations of *M. esculenta* and *M. nana* showed obvious divergence on the phylogenetic tree, and the genetic divergence

was significant. In contrast, the populations of *M. rubra* and *M. adenophora* showed some genetic mixing on the phylogenetic tree, suggesting that they may have a close phylogenetic relationship.

Phylogenesis based on cpDNA data further revealed the phylogenetic relationship of the four species. *M. rubra* and *M. adenophora* share a haplotype H2, and they have a common clade. *M. rubra* and *M. esculenta* share the H3 haplotype, *Morella* is a dioecious plant (Fig. 1e), which can be pollinated by wind, insects and animals (Pang, 2008). The geographical proximity may provide opportunities for the exchange of genetic materials between species. Phylogenetic tree divides four species into four groups, which clearly reveals the relationship between them. Phylogenetic analysis showed that *M. esculenta* differentiated the earliest, followed by *M. nana* and *M. rubra*. However, *M. adenophora* is classified into clades of *M. nana*, which is the closest to it in phylogenetic relationship. It may be because cpDNA is maternal inheritance, and the gene exchange of cpDNA marker is limited, so cpDNA marker is easier to infiltrate (Whittemore et al. 1991; Du et al. 2009; Wan et al. 2017). The NJ tree based on SSR sequencing data shows that *M. rubra* and *M. adenophora* have communication, but they are in different clades (Fig. 1e). The ML phylogenetic tree constructed by the results of whole-genome resequencing further clearly showed the phylogenetic relationship of four *Morella* species, and confirmed that *M. esculenta* was the earliest species to differentiate, and *M. nana*, *M. rubra* and *M. adenophora* formed a complex group relationship, in which *M. rubra* and *M. adenophora* formed a pair of closely related sister group (Fig. 1d). In a word, the four *Morella* species distributed in East Asia are monophyletic clades of each other, and *M. rubra* and *M. adenophora* are closely related sister group because of gene exchange. Our results are consistent with those of Liu (2016). We increased the sampling amount, verified and analyzed the phylogenetic relationship of four *Morella* species distributed in East Asia by SSR, cpDNA and whole-genome resequencing, and further proved their genetic relationship and phylogenetic position.

Species divergence and population history

The formation and diversity of species may be affected by mountain barriers, and complex topography will form different environmental heterogeneity and niche changes (Fjeldså et al. 2012). Molecular clock analysis revealed that four *Morella* species began to differentiate in the late Miocene (Fig. 1d and e), the first being *M. esculenta*, and then *M. nana* began to differentiate. The divergence between *M. rubra* and *M. adenophora* occurred in the Pliocene, suggesting that *Morella* had a long divergence process. The phylogenetic analysis based on whole-genome resequencing further demonstrates the reliability of the estimation of divergence time of four species. We believe that the divergence of four *Morella* species is related to the uplift of QTP. Previous studies and literature review have shown that the QTP uplift occurred in the middle Miocene (15 – 13 Mya), late Miocene (8 – 7 Mya) or Pliocene to Early Pleistocene (3.6–1.7 Mya), with the strongest uplift beginning at 3.6 Ma (Coleman and Hodges, 1995; Li and Fang, 1999; Shahzad et al. 2017). The divergence time of four *Morella* species in this study corresponds to the uplift time of the QTP, and the environmental differences caused by this long-term geological event may have caused species divergence. Terrain heterogeneity and climate fluctuation caused by the uplift of QTP have shaped the current species distribution pattern (Yu et al. 2019). This is consistent with previous research results (Lei et al. 2007; Lei et al. 2015; Shahzad et al. 2017). Due to the influence of

geographical barriers, gene exchange is limited, and there are differences in different ecological environments, which will promote adaptive evolution of species, thus intensifying genetic divergence. Terrain heterogeneity has driven the formation and maintenance of genetic structures of four *Morella* species (Li et al. 2024).

From late Miocene to Cenozoic, it is of great importance for the origin of biodiversity in East Asia (Sun et al. 2014). Significant climatic oscillations and geological events in the Late Miocene, such as continental drift, mountain range formation and climate change, especially the uplift of the Tibetan Plateau, led to great changes in the geomorphology and climate of neighboring areas. The interaction of these processes provides new habitats and niches for biodiversity in subtropical regions, and may also create favorable conditions for species divergence or the formation of new species. The climatic transition to a more suitable East Asian monsoon climate (Sun and Wang, 2005). It ensures the adaptation and evolution of species to the new environment and promotes the reconstruction of ecosystems and biological communities. For instance, the rate of speciation in *Primulina* was strongly linked on the East Asian monsoon climate (Kong et al. 2017). The East Asian monsoon climate leads to the increase of precipitation (Li et al. 2024), and the niche model shows that the Precipitation of Warmest Quarter has the greatest influence on species distribution, which further provides suitable habitats for *M. esculenta* and *M. nana*, and promotes species divergence. The unfavorable conditions in Pliocene led to the migration of species distribution and contraction into a low latitude refuge. Species divergence may be influenced by environmental changes and biological characteristics, and geographic isolation plays a role in promoting species formation (Givnish et al. 2011; Rabosky and Adams, 2012). These factors together shape the species formation and divergence of *M. rubra* and *M. adenophora*.

Mantel test shows (Fig. S3) that population altitude has a certain driving effect on genetic divergence of four species of *Morella*, which indicates that four *Morella* species are sensitive to climate change, and environmental heterogeneity guides four *Morella* species to adapt to local environmental changes (Zhang et al. 2016). Long-term environmental heterogeneity has shaped the existing lineage geographical structure of species. Neutrality test and mismatch analysis proved that the effective population of four *Morella* species had been in a relatively stable equilibrium state on the scale of research time. Niche simulation showed that *Morella* had at least three shelters in LGM period (Fig. 6). In LGM period, the temperature was low, the species could not survive, and the effective population size might be greatly reduced. Then, the temperature rose, the species' suitable area expanded and the effective population size became larger. The BSPs curve showed that *Morella* expanded from about 1.2 Ma (Fig. 5), and four species expanded in different periods before this period. The results showed that complex climate change may be an important factor affecting the historical and current geographical distribution of *Morella* (Zheng et al. 1998; Wu et al. 2024). Based on the analysis of cpDNA data, H2, H4 and H5 haplotypes are considered as ancestral haplotypes. Niche simulation showed that *M. esculenta* was mainly distributed in Yunnan and Guangxi, *M. rubra* was mainly distributed in Guangxi, *M. adenophora* was mainly distributed in Guangxi, and *M. nana* was mainly distributed in Yunnan and Guizhou during LGM period. Combined with SSR and cpDNA genetic diversity analysis, it is speculated that these areas may be ice age shelters for four species. These shelters may provide an environment

for species to survive and adapt during the ice age, thus shaping their genetic structure and existing geographical distribution model.

Management and conservation strategies for Morella resource

Biological diversity is an important part of maintaining ecological balance, and understanding genetic diversity of species is essential to understanding the evolution of biological diversity (Ren et al. 2024). With the loss of global biodiversity, especially the low genetic diversity of vulnerable species, it is essential to protect them (Tao et al. 2024). Habitat fragmentation is an important cause of species threat and endangerment (Yang et al. 2023). Genetic variation and habitat are a key factor in ensuring the continued presence of a species continued presence in the environment (Wambulwa et al. 2022). Only by ensuring sufficient genetic variation can we promote the long-term survival and evolution of species (Dai et al. 2015). In this study, the genetic diversity of 63 populations of 4 *Morella* species was analyzed based on cpDNA and SSR. The results showed that the genetic diversity of some populations was low, and our field investigation found that some populations were seriously damaged, which led to the gradual narrowing of the suitable area of species and the serious loss of wild resources. Therefore, it is urgent to effectively manage and protect these important wild resources of *Morella*.

The *M. rubra* is distributed in many areas and has strong adaptability. However, due to its delicious edible fruit, it has caused serious human destruction, forced the disappearance of wild resources, and sharply decreased genetic diversity. *M. adenophora* has been listed as a vulnerable species (IUCN Red List of Threatened Species) and its distribution area is narrow. Because the *M. nana* is only distributed at high altitude area of Yunnan-Guizhou Plateau, the growth range is limited. *M. esculenta* is mainly distributed in tropical and subtropical regions from East Asia to Southeast Asia, and in subtropical regions of China, it has a wild range, and early origin, and is at the base of clade. *Morella* is a perennial deciduous woody plant with high value in medicine, food, ecology and economy (Bai et al. 2020). For this reason, it is urgent to protect this resource. Combined with the niche model, Guangxi, Yunnan and Guizhou are ice age shelters of *Morella*, which shows that *Morella* has a suitable living environment in these areas. The populations of YNLY, YNDL, GXTB and GZQL contain high genetic diversity or abundant haplotypes, which can correspond to shelters. We suggest that nature reserves should be established in such areas for in-situ protection. Maintain the sustainability of the living environment, avoid habitat fragmentation, improve the adaptability of the population, and ensure species richness. Secondly, ex-situ conservation, for the populations whose genetic diversity has been lost and whose habitats have been destroyed, moves to an environment suitable for species growth for protection, such as botanical gardens, nature reserves and other effective protected areas. These two protective measures are effective and desirable. In addition, science and technology can be used to assist the propagation of *Morella*, such as tissue culture and artificial seedling raising, which can solve the problems related to plant regeneration, breeding and cultivation (Cabrera et al. 2012). In a word, the adaptive habitat of *Morella* wild resources has been seriously damaged, and it is difficult for some populations to survive in the original habitat, so effective measures must be taken for management and protection.

Conclusions

In this study, the genetic structure and systematic geographical distribution of 63 populations of four *Morella* species were studied by molecular systematics. Our study suggests that geological movement and climate change can well explain the species divergence model and genetic structure well. In addition, genetic diversity plays an important role in the adaptation of species to environmental change, and ice age shelters provide a suitable living environment for species. *Morella* has remained relatively stable in the evolutionary scale of the research time, and it did not begin to expand until 1.2 Ma. The results of this study will enhance our understanding of the evolution, phylogenetic relationships and distribution of the four species of *Morella* in East Asia. Based on our research, we propose that these four species are closely related and that there is gene exchange among the species. Whole-genome resequencing and cpDNA support the topology of ((*M. rubra* and *M. adenophora*) *M. nana*) *M. esculenta*. In a word, our study made a clear phylogenetic relationship of the four *Morella* species, and existing geographical distribution pattern of species was the common result of climate change and geological movement, combined niche simulation prediction with genetic diversity, and generalized the protection strategy of wild resources of *Morella*.

Declarations

Author Contributions

HSS came up with the original concept; HSS, LYX, ZC and JCX designed the study, selected species and conducted field investigations; FJ, ZLH, LY and CYT participate in field sample collection; HSS and LYX did the lab work, analyzed the genetic data and wrote the manuscript; ZC supervised the laboratory work and genetic analysis. All authors contributed to the final draft.

Funding

This work was supported by National Natural Science Foundation of China (32260252) and Guizhou science and technology support project ([2019]2451-2).

Data availability

The data generated or analyzed in our study were submitted to GenBank (https://www.ncbi.nlm.nih.gov/). The accession numbers are: PQ337779-PQ337829, PQ337830-PQ337880: PQ337881-PQ337931 and PQ337932-PQ337982.

Ethics approval and consent to participate

Our research did not involve any ethical issues, and our collection was in line with Chinese laws. The samples we collected were not privately owned, and we consulted the local forestry bureau and relevant laws and regulations when we collected them. We collected only a small number of leaves and did not

endanger the individual plants, in accordance with the relevant management regulations for the protection of vulnerable plants.

Competing interest statement

All authors certify that they have no affiliations with or involvement in any organization or entity with any financial interest or non-financial interest in the subject matter or materials discussed in this manuscript.

Supplementary materials

Supplementary data are available at Supplementary material.

References

1. Avise JC (2000) Phylogeography: the history and formation of species. Harvard university press.
2. Bai Y, Qie HL, Gao ZH (2020) A review for molecular biology of *Morella rubra*. J. Journal of Fruit Science 37: 764-772. https://doi.org/10.13925/j.cnki.gsxb.20190444
3. Bandelt HJ, Forster P, Röhl, A (1999) Median-joining networks for inferring intraspecific phylogenies. Molecular Biology and Evolution 16: 37-48. https://doi.org/10.1093/oxfordjournals.molbev.a026036
4. Bouckaert R, Vaughan TG, Barido-Sottani J, Duchêne S, Fourment M, Gavryushkina A, Heled J, Jones G, Kuhnets D, Maio ND, Matschiner M, Mendes F, Muller NF, Ogilvie HA, Plessis L, Popinga A, Rambaut A, Rasmussen D, Siveroni L, Suchard MA, Wu CH, Xie D, Zhang C, Stadler T, Drummond A J (2019) BEAST 2.5: An advanced software platform for Bayesian evolutionary analysis. PLoS computational biology 15: e1006650. https://doi.org/10.1371/journal.pcbi.1006650
5. Cabrera-Toledo D, González-Astorga J, Flores-Vázquez JC (2012) Fine-scale spatial genetic structure in two Mexican cycad species *Dioon caputoi* and *Dioon merolae* (Zamiaceae, Cycadales): Implications for conservation. Biochemical Systematics and Ecology 40: 43-48. https://doi.org/10.1016/j.bse.2011.09.004
6. Chen N (2014) Studies on population genetics and domestication of *Myrica rubra* (Myricaceae). Dissertation, Zhejiang University.
7. Coleman M, Hodges K (1995) Evidence for Tibetan plateau uplift before 14 Myr ago from a new minimum age for east–west extension. Nature 374: 49-52. https://doi.org/10.1038/374049a0
8. Comes HP, Kadereit JW (1998) The effect of Quaternary climatic changes on plant distribution and evolution. Trend in plant science 3: 432-438. https://doi.org/10.1016/S1360-1385(98)01327-2
9. Dagnino D, Minuto L, Casazza G (2017) Divergence is not enough: the use of ecological niche models for the validation of taxon boundaries. Plant Biology 19: 1003-1011. https://doi.org/10.1111/plb.12600
10. Dai S, Wang S, Ruan L, Zhou Y, Wu W, Liu Y, Shi S, Zhou R (2015) Extremely low genetic diversity and extensive genetic admixture at the northern range margins of *Bombax ceiba*. Biochemical Systematics and Ecology 60: 177-185. https://doi.org/10.1016/j.bse.2015.04.022

11. Dai CG, Wang M, Chen JS, Wang XH (2013) Tectonic Movement Characteristic and Its Geological Significance of Guizhou. Guizhou Geol. 30, 119-124.
12. Du FK, Petit RJ, Liu JQ (2009) More introgression with less gene flow: chloroplast vs. mitochondrial DNA in the *Picea asperata* complex in China, and comparison with other conifers. Molecular Ecology 18: 1396-1407. https://doi.org/10.1111/j.1365-294X.2009.04107.x
13. Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Molecular ecology 14: 2611-2620. https://doi.org/10.1111/j.1365-294x.2005.02553.x
14. Excoffier L, Lischer HE (2010) Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. Molecular ecology resources 10: 564-567. https://doi.org/10.1111/j.1755-0998.2010.02847.x
15. Fan D, Lei S, Liang H, Yao Q, Kou Y, Cheng S, Yang Y, Qiu Y, Zhang Z (2022) More opportunities more species: Pleistocene differentiation and northward expansion of an evergreen broadleaved tree species Machilus thunbergii (Lauraceae) in Southeast China. BMC Plant Biology 22: 35. https://doi.org/10.1186/s12870-021-03420-9
16. Fava WS, da Costa PC, Lorenz AP (2020) Ecological niche modelling and genetic analyses reveal lack of geographic differentiation of Leptolobium dasycarpum (Leguminosae, Papilionoideae) across the Brazilian savannah. Flora, 264: 151566. https://doi.org/10.1016/j.flora.2020.151566
17. Favre A, Michalak I, Chen CH, Wang JC, Pringle JS, Matuszak S, Sun H, Yuan YM, Struwe L, Muellner-Riehl AN (2016) Out-of-Tibet: The spatio-temporal evolution of *Gentiana* (Gentianaceae). Journal of Biogeography 43: 1967-1978. https://doi.org/10.1111/jbi.12840
18. Fawcett T (2006) An introduction to ROC analysis. Pattern recognition letters 27: 861-874. https://doi.org/10.1016/j.patrec.2005.10.010
19. Fjeldså J, Bowie RC, Rahbek C (2012) The role of mountain ranges in the diversification of birds. Annual Review of Ecology, Evolution, and Systematics 43: 249-265. https://doi.org/10.1146/annurev-ecolsys-102710-145113
20. Fu YX, Li WH (1993) Maximum likelihood estimation of population parameters. Genetics 134: 1261-1270. https://doi.org/10.1093/genetics/134.4.1261
21. Gao Y, Yan YH, Yin S, Yu L, Zhu L, Ding XM, Zhang YN, Chen HH, Tang LJ, Li ZM, Chu HL, Shi XD, Zou ZR, Tang LZ (2022) Fragmented habitats and Pleistocene climate shaped diversification of the hoary bamboo rat (*Rhizomys pruinosus*) in the mountainous plateau of SW China. Journal of Systematics and Evolution 60: 1358-1370. https://doi.org/10.1111/jse.12820
22. Givnish TJ, Barfuss MH, Van Ee B, Riina R, Schulte K, Horres R, Gonsiska PA, Jabaily RS, Crayn DM, Smith JAC, Winter K, Brown GK, EvansTM, Holst BK, Luther H, Till W, Zizka G, Berry PE, Sytsma KJ (2011) Phylogeny, adaptive radiation, and historical biogeography in Bromeliaceae: insights from an eight-locus plastid phylogeny. American journal of botany 98: 872-895. https://doi.org/10.3732/ajb.1000059

23. Herbert J (2005) New combinations and a new species in *Morella* (Myricaceae). Novon 293-295. https://doi.org/10.2307/3393340
24. Hewitt GM (2004) Genetic consequences of climatic oscillations in the Quaternary. Philosophical https://doi.org/10.1098/rstb.2003.1388
25. Hickerson MJ, Carstens BC, Cavender-Bares J, Crandall KA, Graham CH, Johnson JB, Rissler L, Victoriano PF, Yoder AD (2010) Phylogeography's past, present, and future: 10 years after. Molecular phylogenetics and evolution 54: 291-301. https://doi.org/10.1016/j.ympev.2009.09.016
26. Hoorn C, Mosbrugger V, Mulch A, Antonelli A (2013) Biodiversity from mountain building. Nature Geoscience 6: 154-154. https://doi.org/10.1038/ngeo1742
27. Hu H, Yang Y, Li A, Zheng Z, Zhang J, Liu J (2022) Genomic divergence of Stellera chamaejasme through local selection across the Qinghai–Tibet plateau and northern China. Molecular ecology 31: 4782-4796. https://doi.org/10.1111/mec.16622
28. Ickert-Bond SM, Renner SS (2016) The Gnetales: recent insights on their morphology, reproductive biology, chromosome numbers, biogeography, and divergence times. Journal of Systematics and Evolution 54: 1-16. https://doi.org/10.1111/jse.12190
29. Jia HM (2016) The draft genome of Chinese bayberry (*M. rubra*) and genetic analysis of sex determination between male and female plants. Dissertation, Zhejiang University.
30. Ju PJ, Zhu YF, Wang GY, Zhou CC, Yan LJ (2023) Construction of specific fluorescent-labeled SSR marker database of Chinese Bayberry (Morella rubra) varieties. Journal of Agricultural Biotechnology 31: 2209-2220. https://doi.org/10.3969/j.issn.1674-7968.2023.10.020
31. Kalinowski ST, Taper ML, Marshall TC (2007) Revising how the computer program CERVUS accommodates genoty** error increases success in paternity assignment. Molecular ecology 16: 1099-1106. https://doi.org/10.1111/j.1365-294X.2007.03089.x
32. Kalyaanamoorthy S, Minh BQ, Wong TK, Von Haeseler A, Jermiin LS (2017) ModelFinder: fast model selection for accurate phylogenetic estimates. Nature methods 14: 587-589. https://doi.org/10.1038/nmeth.4285
33. Keppe, G, Van Niel KP, Wardell-Johnson GW, Yates CJ, Byrne M, Mucina L, Schut AGT, Hopper SD, Franklin SE (2012) Refugia: identifying and understanding safe havens for biodiversity under climate change. Global ecology and biogeography 21: 393-404. https://doi.org/10.1111/j.1466-8238.2011.00686.x
34. Kong H, Condamine FL, Harris AJ, Chen J, Pan B, Möller M, Hoang VS, Kang M (2017) Both temperature fluctuations and East Asian monsoons have driven plant diversification in the karst ecosystems from southern China. Molecular Ecology 26: 6414-6429. https://doi.org/10.1111/mec.14367
35. Laidò G, Mangini G, Taranto F, Gadaleta A, Blanco A, Cattivelli L, Marone D, Mastrangelo A, Papa, R, De Vita P (2013) Genetic diversity and population structure of tetraploid wheats (*Triticum turgidum* L.) estimated by SSR, DArT and pedigree data. Plos one 8: e67280. https://doi.org/10.1371/journal.pone.0067280

36. Lei FM, Wei GA, Zhao HF, Yin ZH, Lu JL (2007) China subregional avian endemism and biodiversity conservation.Vertebrate Conservation and Biodiversit 293-304.https://doi.org/10.1007/s10531-006-9080-3
37. Lei F, Qu Y, Song G, Alström P, Fjeldså J (2015) The potential drivers in forming avian biodiversity hotspots in the East Himalaya Mountains of Southwest China. Integrative zoology 10: 171-181. https://doi.org/10.1111/1749-4877.12121
38. Li B, Liu C, Tang JR, Xin YX, Dong ZH, Bai B, Xin PY (2022) Genetic diversity analysis of Camellia fascicularis HT Chang based on SSR markers. Journal of Applied Research on Medicinal and Aromatic Plants 31: 100404. https://doi.org/10.1016/j.jarmap.2022.100404
39. Li J, Fang X (1999) Uplift of the Tibetan Plateau and environmental changes. Chinese science bulletin 44: 2117-2124. https://doi.org/10.1007/BF03182692
40. Li N, Wang Z, Deng Q, Wang T, Su YJ (2024) Effects of climatic fluctuations on the fragmented distribution pattern of a Tertiary relict plant, Pseudotaxus chienii (Taxaceae), in subtropical China. Botanical Journal of the Linnean Society 205: 55-74. https://doi.org/10.1093/botlinnean/boad064
41. Lin QH, Zhong QZ, Zhang ZH (2013) Genetic diversity analysis of 18 Chinese bayberry germplasm resources with SRAP. Chin. J. Trop. Crops 34, 1667-1671.
42. Liu G, Xue G, Zhao T, Li Y, Yue L, Song H, Liu Q (2023) Population structure and phylogeography of three closely related tree peonies. Ecology and Evolution 13: e10073. https://doi.org/10.1002/ece3.10073
43. Liu J, Möller M, Provan J, Gao LM, Poudel RC, Li DZ (2013) Geological and ecological factors drive cryptic speciation of yews in a biodiversity hotspot. New Phytologist 199: 1093-1108. https://doi.org/10.1111/nph.12336
44. Liu LX (2016) Studies of phylogeography and domestication origin of Chinese Bayberry (*Morella rubra*). Dissertation, Zhejiang University.
45. Liu YJ, Wu ML, Xu X, Zhu X, Dai ZX, Gou GQ (2022) Genetic diversity and phylogeography of the endemic species Chimonobambusa utilis growing in southwest China: Chloroplast DNA sequence and microsatellite marker analyses. Frontiers in Plant Science 13: 943225. https://doi.org/10.3389/fpls.2022.943225
46. López-Pujol J, Zhang FM, Sun HQ, Ying TS, Ge S (2011) Centres of plant endemism in China: places for survival or for speciation?. Journal of Biogeography 38: 1267-1280. https://doi.org/10.1111/j.1365-2699.2011.02504.x
47. Lowry DB, Modliszewski JL, Wright KM, Wu CA, Willis JH (2008) The strength and genetic basis of reproductive isolating barriers in flowering plants. Philosophical Transactions of the Royal Society B: Biological Sciences 363: 3009-3021. https://doi.org/10.1098/rstb.2008.0064
48. Lu LM, Mao LF, Yang T, Ye JF, Liu B, Li HL, Sun M, Miller JT, Mathews S, Hu HH, Niu YT, Peng DX, Chen YH, Smith SA, Chen M, Xiang KL, Le CT, Dang VC, Lu AM, Soltis PS, Soltis DE, Li JH, Chen ZD (2018) Evolutionary history of the angiosperm flora of China. Nature 554: 234-238. https://doi.org/10.1038/nature25485

49. Manni F, Guérard E, Heyer E (2004) Geographic patterns of (genetic, morphologic, linguistic) variation: how barriers can be detected by using Monmonier's algorithm. Human biology 76: 173-190. https://doi.org/10.1353/hub.2004.0034
50. Matías L, Linares JC, Sánchez-Miranda Á, Jump AS (2017) Contrasting growth forecasts across the geographical range of Scots pine due to altitudinal and latitudinal differences in climatic sensitivity. Global Change Biology 23: 4106-4116. https://doi.org/10.1111/gcb.13627
51. Mort ME, Archibald JK, Randle CP, Levsen ND, O'Leary TR, Topalov K, Wiegand CM, Crawford DJ (2007) Inferring phylogeny at low taxonomic levels: utility of rapidly evolving cpDNA and nuclear ITS loci. American Journal of Botany 94: 173-183. https://doi.org/10.3732/ajb.94.2.173
52. Nybom H (2004) Comparison of different nuclear DNA markers for estimating intraspecific genetic diversity in plants. Molecular ecology 13: 1143-1155. https://doi.org/10.1111/j.1365-294X.2004.02141.x
53. Pang H (2008) Resources survey and it's genetic diversity of wild *Myrica* in Guangxi. Dissertation, Guangxi University.
54. Peakall R, Smouse PE (2012) GenAlEx 6.5: genetic analysis in excel. population genetic software for teaching and research: an update. Bioinformatics 28: 2537–2539. https://doi.org/10.1111/j.1471-8286.2005.01155.x
55. Phillips SJ, Anderson RP, Schapire RE (2006) Maximum entropy modeling of species geographic distributions. Ecological Modelling 190: 231–259. https://doi.org/10.1016/j.ecolmodel.2005.03.026
56. Piry S, Luikart G, Cornuet JM (1999) BOTTLENECK: a computer program for detecting recent reductions in the effective population size using allele frequency data. https://doi.org/10.1093/jhered/90.4.502
57. Pons O, Petit R (1996) Measwring and testing genetic differentiation with ordered versus unordered alleles. Genetics 144: 1237-1245. https://doi.org/10.1093/genetics/144.3.1237
58. Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. Genetics 155: 945-959. https://doi.org/10.1093/genetics/155.2.945
59. Rabosky DL, Adams DC (2012) Rates of morphological evolution are correlated with species richness in salamanders. Evolution 66: 1807-1818. https://doi.org/10.1111/j.1558-5646.2011.01557.x
60. Raymo ME, Ruddiman WF (1992) Tectonic forcing of late Cenozoic climate. Nature 359: 117-122. https://doi.org/10.1038/359117a0
61. Ren Y, Zhang L, Yang X, Lin H, Sang Y, Feng L, Liu J, Kang M (2024) Cryptic divergences and repeated hybridizations within the endangered "living fossil" dove tree (Davidia involucrata) revealed by whole genome resequencing. Plant Diversity 46: 169-180. https://doi.org/10.1016/j.pld.2024.02.004
62. Rozas J (2009) DNA sequence polymorphism analysis using DnaSP. Bioinformatics for DNA sequence analysis 337-350. https://doi.org/10.1007/978-1-59745-251-9_17
63. Shahzad K, Jia Y, Chen FL, Zeb U, Li ZH (2017) Effects of mountain uplift and climatic oscillations on phylogeography and species divergence in four endangered *Notopterygium* herbs. Frontiers in Plant

Science 8: 1929. https://doi.org/10.3389/fpls.2017.01929

64. Suchard MA, Lemey P, Baele G, Ayres DL, Drummond AJ, Rambaut A (2018) Bayesian phylogenetic and phylodynamic data integration using BEAST 1.10. Virus evolution 4: vey016. https://doi.org/10.1093/ve/vey016
65. Sun C, Huang H, Xu C, Li X, Chen K (2013) Biological activities of extracts from Chinese bayberry (*Myrica rubra* Sieb. et Zucc.): a review. Plant Foods for Human Nutrition 68: 97-106. https://doi.org/10.1007/s11130-013-0349-x
66. Sun H, Li Z (2003) Qinghai-Tibet Plateau uplift and its impact on Tethys flora. Advances in Earth Science 18: 852. https://doi.org/10.11867/j.issn.1001-8166.2003.06.0852
67. Sun, X, Wang, P, 2005. How old is the Asian monsoon system?—Palaeobotanical records from China. Palaeogeo. Palaeoclim. Palaeoecol. 222, 181-222. https://doi.org/10.1016/j.palaeo.2005.03.005
68. Sun YX, Moore MJ, Yue LL, Feng T, Chu HJ, Chen ST, Ji YH, Wang HC, Li JQ (2014) Chloroplast phylogeography of the east Asian Arcto-Tertiary relict *Tetracentron sinense* (Trochodendraceae). Journal of Biogeography 41: 1721-1732. https://doi.org/10.1111/jbi.12323
69. Taberlet P, Fumagalli L, Wust-Saucy AG, Cosson JF (1998) Comparative phylogeography and postglacial colonization routes in Europe. Molecular ecology 7: 453-464. https://doi.org/10.1046/j.1365-294x.1998.00289.x
70. Tajima F (1989) Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. Genetics 123: 585-595. https://doi.org/10.1093/genetics/123.3.585
71. Tao T, Milne RI, Li J, Yang H, Wang S, Chen S, Mao K (2024) Conservation genomic investigation of an endangered conifer, Thuja sutchuenensis, reveals low genetic diversity but also low genetic load. Plant Diversity 46: 78-90. https://doi.org/10.1016/j.pld.2023.06.005
72. Wambulwa MC, Luo YH, Zhu GF, Milne R, Wachira FN, Wu ZY, Wang H, Gao LM, Li DZ, Liu J (2022) Determinants of genetic structure in a highly heterogeneous landscape in Southwest China. Frontiers in plant science 13: 779989. https://doi.org/10.3389/fpls.2022.779989
73. Wan, Q, Zheng, Z, Huang, K, Guichoux, E, Petit, R.J, 2017. Genetic divergence within the monotypic tree genus *Platycarya* (Juglandaceae) and its implications for species' past dynamics in subtropical China. Tree Genetics & Genomes 13: 1-11. https://doi.org/10.1007/s11295-017-1153-9
74. Wang CB, Wang T, Su YJ (2014) Phylogeography of *Cephalotaxus oliveri* (Cephalotaxaceae) in relation to habitat heterogeneity, physical barriers and the uplift of the Yungui Plateau. Molecular Phylogenetics and Evolution 80: 205-216. https://doi.org/10.1016/j.ympev.2014.08.015
75. Wang H, Qiong L, Sun K, Lu F, Wang Y, Song Z, Wu Q, Chen J, Zhang W (2010) Phylogeographic structure of *Hippophae tibetana* (Elaeagnaceae) highlights the highest microrefugia and the rapid uplift of the Qinghai-Tibetan Plateau. Molecular Ecology 19: 2964-2979. https://doi.org/10.1111/j.1365-294X.2010.04729.x
76. Wang YH, Jiang WM, Comes HP, Hu FS, Qiu YX, Fu CX (2015) Molecular phylogeography and ecological niche modelling of a widespread herbaceous climber, *Tetrastigma hemsleyanum*

(Vitaceae): insights into Plio–Pleistocene range dynamics of evergreen forest in subtropical China. New Phytologist 206: 852-867. https://doi.org/10.1111/nph.13261

77. Wang Z, Zeng Y, Zhang Z, Sheng S, Tian J, Wu R, Pang X (2017) Phylogeography study of the Siberian apricot (*Prunus sibirica* L.) in Northern China assessed by chloroplast microsatellite and DNA makers. Frontiers in plant science 8: 1989. https://doi.org/10.3389/fpls.2017.01989
78. Wen J, Zhang JQ, Nie ZL, Zhong Y, Sun H (2014) Evolutionary diversifications of plants on the Qinghai-Tibetan Plateau. Frontiers in genetics 5: 4. https://doi.org/10.3389/fgene.2014.00004
79. Whittemore AT, Schaal BA (1991) Interspecific gene flow in sympatric oaks. Proceedings of the National Academy of Sciences 88: 2540-2544. https://doi.org/10.1073/pnas.88.6.2540
80. Wolfe KH, Li WH, Sharp PM (1987) Rates of nucleotide substitution vary greatly among plant mitochondrial, chloroplast, and nuclear DNAs. Proceedings of the National Academy of Sciences 84: 9054-9058. https://doi.org/10.1073/pnas.84.24.9054
81. Wu M, Cheng Y, Jiang CX, Zhang MS, Shi T, Zhao C (2024) Phylogeography of *Morella nana*: The Wumeng Mountains as a natural geographical isolation boundary on the Yunnan-Guizhou Plateau. Ecology and Evolution 14: e11566. https://doi.org/10.1002/ece3.11566
82. Wu S, Ma S, Wang H, Wang L, Jiang J (2023) Spatiotemporal variations of ecosystem service value in the hill and mountain belt of southern China across different altitude gradients. Chinese Journal of Ecology 42: 966. https://doi.org/10.13292/j.1000-4890.202304.020
83. Xie DF, Li MJ, Tan JB, Price M, Xiao QY, Zhou SD, Yan Y, He XJ (2017) Phylogeography and genetic effects of habitat fragmentation on endemic *Urophysa* (Ranunculaceae) in YunGui Plateau and adjacent regions. PLoS One 12: e0186378. https://doi.org/10.1371/journal.pone.0186378
84. Xie P, Guo Y, Teng Y, Zhou W, Yu Y (2024) GeneMiner: A tool for extracting phylogenetic markers from next-generation sequencing data. Molecular Ecology Resources 24: e13924. https://doi.org/10.1111/1755-0998.13924
85. Xu J, F Zang, Q Wu, Y Wang, B Wang, P Huang, Zang D, Ma Y, Zheng Y (2021) Analysis of the genetic diversity and molecular phylogeography of the endangered wild rose (Rosa rugosa) in China based on chloroplast genes. Global Ecology and Conservation 28: e01653. https://doi.org/10.1016/j.gecco.2021.e01653
86. Yang L, Liang SQ, Pan JM, Wei JX, Ding T, Jiang RH, Shao YZ, Zhang XC, Liu YB, Xiang QP (2023) Species delimitation and genetic conservation of the endangered firs *Abies beshanzuensis* and *A. ziyuanensis*. Chinese Journal of Plant Ecology 47: 1629-1645. https://doi.org/10.17521/cjpe.2022.0295
87. Yang LF, Ding Q, Yang JL, Liu ZX, Zhang YX (2015) Study on the content of phenolic compound and its anti-glycation activity in 8 species of fruits. Cereal and Food Industry 22: 78-81.
88. Yang, Y, Wu Y, Chen Q, Liu C, Liu G, Cheng S, Wang L (2022) Potential spatiotemporal distribution changes and conservation recommendations of two connected endangered tree peony species (Paeonia decomposita & P. rotundiloba). Flora 294: 152131. https://doi.org/10.1016/j.flora.2022.152131

89. Yin X, Qian H, Sui X, Zhang M, Mao L, Svenning JC, Ricklefs RE, He FL (2021) Effects of climate and topography on the diversity anomaly of plants disjunctly distributed in eastern Asia and eastern North America. Global Ecology and Biogeography 30: 2029-2042. https://doi.org/10.1111/geb.13366
90. Yu H, Deane DC, Sui X, Fang S, Chu C, Liu Y, He F (2019) Testing multiple hypotheses for the high endemic plant diversity of the Tibetan Plateau. Global ecology and biogeography 28: 131-144. https://doi.org/10.1111/geb.12827
91. Yun SA, Son HD, Im HT, Kim SC (2020) Genetic diversity and population structure of the endangered orchid *Pelatantheria scolopendrifolia* (Orchidaceae) in Korea. Plos one 15, e0237546. https://doi.org/10.1371/journal.pone.0237546
92. Zhang Z, Xie P, Guo Y, Zhou W, Liu E, Yu Y (2022)Easy353: A Tool to Get Angiosperms353 Genes for Phylogenomic Research. Molecular Biology and Evolution 39: msac261. https://doi.org/10.1093/molbev/msac261
93. Zhang D, Gao F, Jakovlić I, Zou H, Zhang J, Li WX, Wang GT (2020) PhyloSuite: An integrated and scalable desktop platform for streamlined molecular sequence data management and evolutionary phylogenetics studies. Molecular ecology resources 20: 348-355. https://doi.org/10.1111/1755-0998.13096
94. Zhang SM, Xu CJ, Gao ZS, Chen KS, Wang GY (2009) Development and characterization of microsatellite markers for Chinese bayberry (*Myrica rubra* Sieb. & Zucc.). Conservation genetics 10:1605-1607. https://doi.org/ 10.1007/s10592-008-9804-x
95. Zhang YH, Wang IJ, Comes HP, Peng H, Qiu YX (2016) Contributions of historical and contemporary geographic and environmental factors to phylogeographic structure in a Tertiary relict species, *Emmenopterys henryi* (Rubiaceae). Scientific Reports 6: 24041. https://doi.org/10.1038/srep24041
96. Zhao YJ, Xi HH, Lin J, Yang R, Yin GS, Gong X (2025) Divergence of Two Diospyros Species Promoted by Geo-Climatic Processes Across Dry Valleys and Lowlands in Subtropical China. Journal of Biogeography e15136. https://doi.org/10.1111/jbi.15136
97. Zheng HY, Guo XL, Price M, He XJ, Zhou SD (2021) Effects of mountain uplift and climatic oscillations on phylogeography and species divergence of *Chamaesium* (Apiaceae). Frontiers in Plant Science 12: 673200. https://doi.org/10.3389/fpls.2021.673200
98. Zhong HM, Xie YY (2010) The cultivation technique of Chinese Bayberry. Jilin Agricultural 9: 124.
99. Zhou ZH, Yang CR (2000) Phenolic Constituents of the Fresh Leaves of *Morella nana*. Acta Botanica Yunnanica 22:219-224.
100. Zhuo Z, Baoyin Y, Petit-Maire N (1998) Paleoenvironments in China during the Last Glacial Maximum and the Holocene optimum. Episodes Journal of International Geoscience 21: 152-158. https://doi.org/10.18814/epiiugs/1998/v21i3/003

Figures

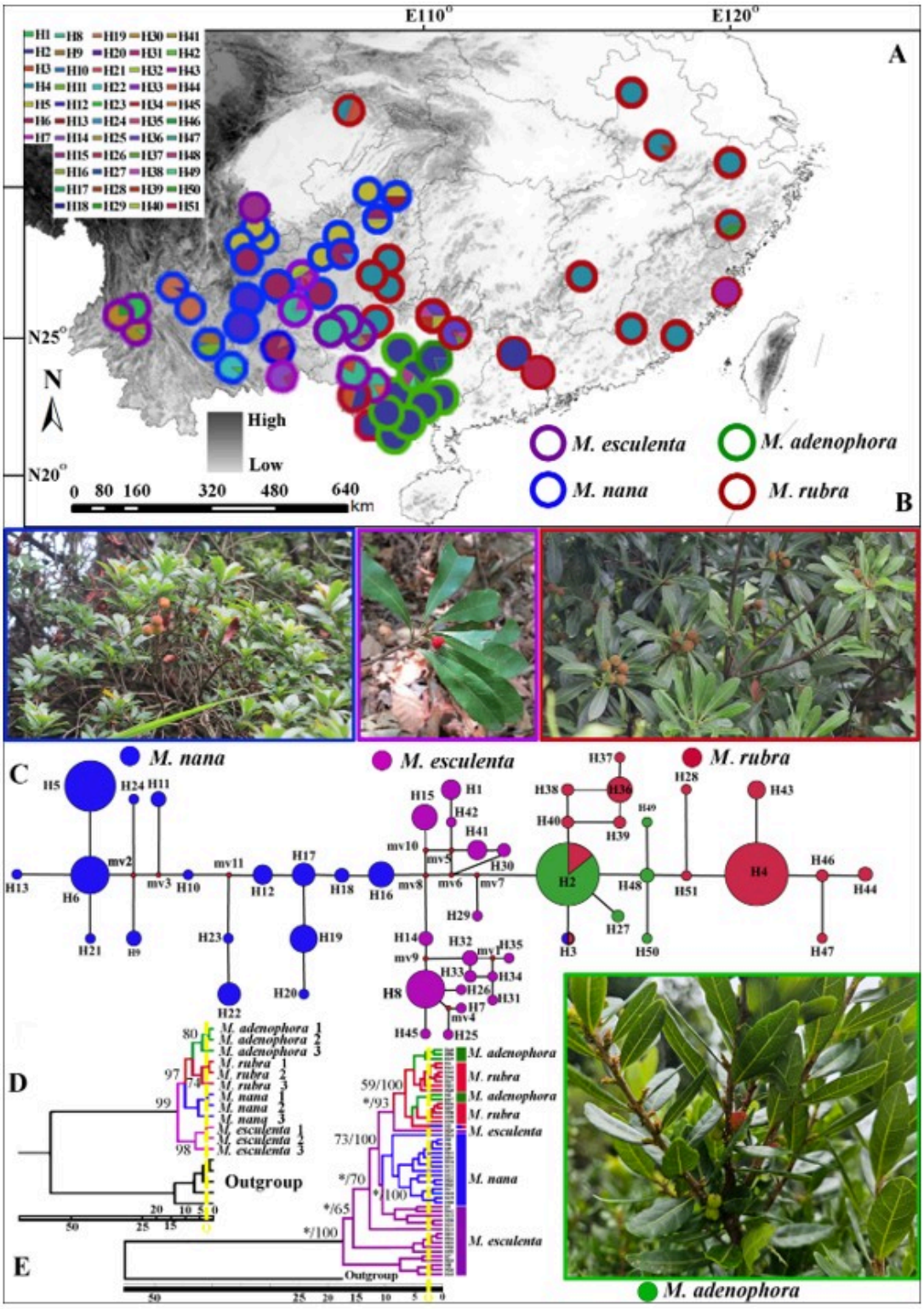

Figure 1

Species distribution, morphology, phylogeny, and interspecies relationships obtained from cpDNA. (a) Geographic distribution of cpDNA haplotypes for the four *Morella* species. Each circle represents a population and each color represents each haplotype. The colored outlines of the circles distinguish the four species, where purple indicates *M. esculenta*,blue indicates *M. nana*, green indicates *M. adenophora*, and red indicates *M. rubra*; (b) Morphological characteristics of four species of *Morella*; (c)

cpDNA haplotype intermediate connection network of four plants. Blue stands for *M. nana*, purple stands for *M. esculenta*, red stands for *M. rubra* and green stands for *M. adenophora*. The colors of the four species are consistent with the geographical distribution (Fig.1a) and morphological of haplotypes (Fig.1b). The size of the circle is proportional to the number of haplotypes; (d) Phylogenetic tree of 353 gene based on whole-genome resequencing and calibration time node, and the yellow dashed line represents the time boundary of the Quaternary period, the divergence time is at the bottom of the phylogenetic tree. The numerical values on the phylogenetic tree branches represent the bootstrap values (BS); (e) Phylogenetic tree of cpDNA haplotypes for the four *Morella*species and the yellow dashed line represents the time boundary of the Quaternary period, the divergence time is at the bottom of the phylogenetic tree. The numerical values on the left and right of the phylogenetic tree branches represent the bootstrap values (BS) of the Maximum Likelihood (ML) tree and the Maximum Parsimony (MP) tree, respectively, indicating slight discrepancies between the two methods. The asterisk "*”denotes branches with BS < 50%.

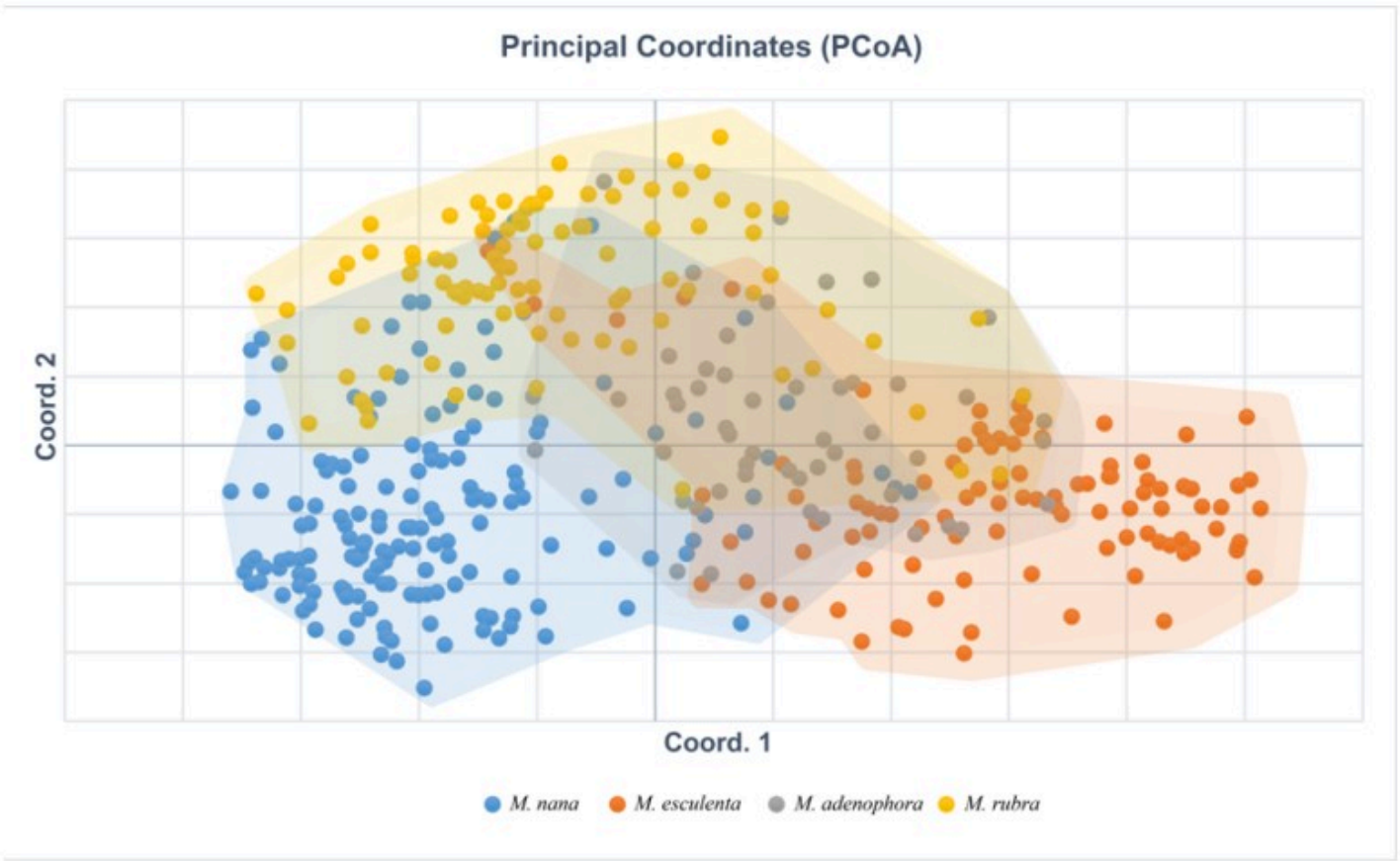

Figure 2

Principal Coordinate Analysis (PCoA) of four *Morella* species based on SSR Sequencing Data. Each point represents an individual. Among them, blue represents *M. nana*, red represents *M. esculenta*, gray represents *M. adenophora* and yellow represents *M. rubra*.

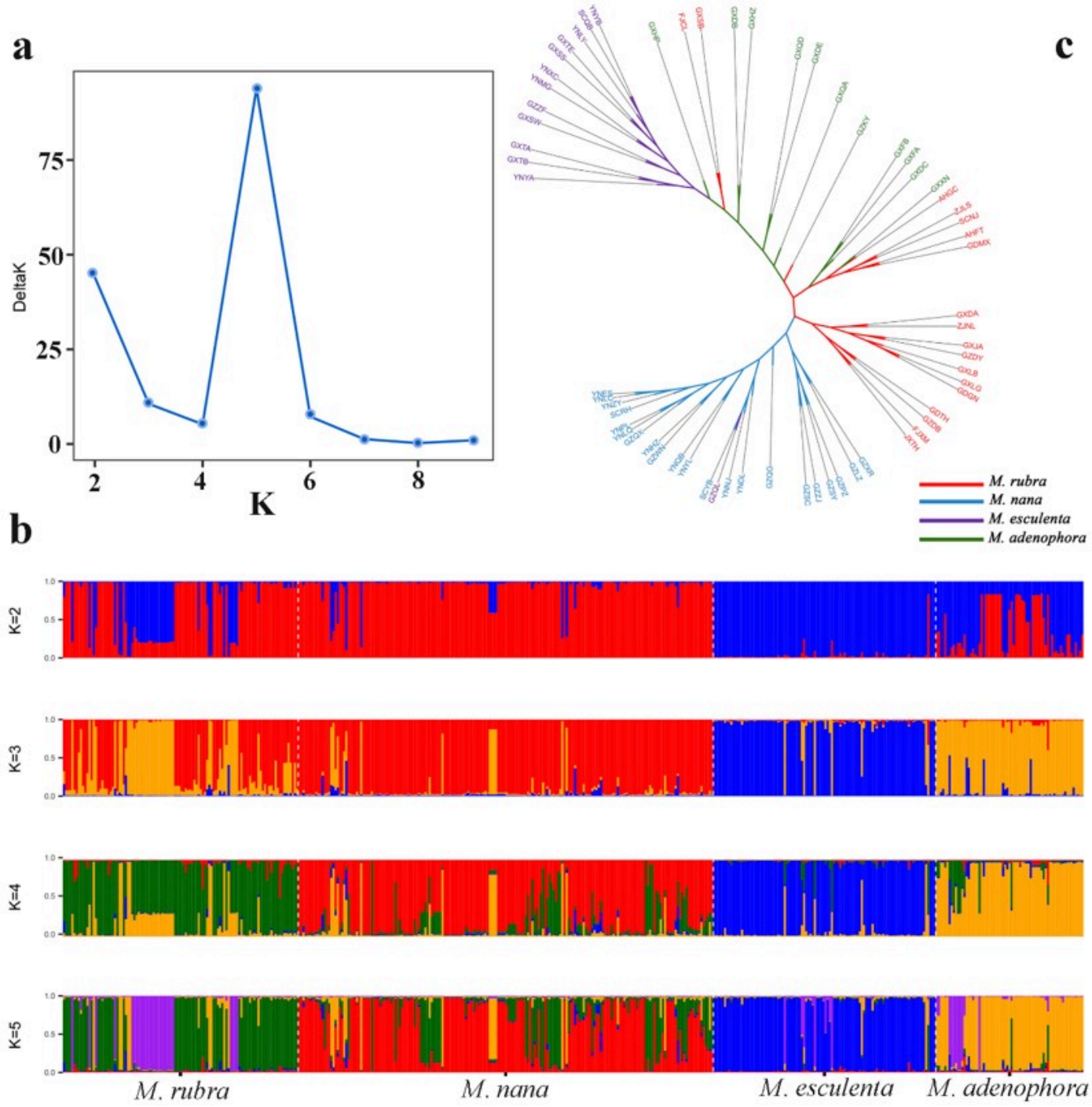

Figure 3

Analysis of genetic structure of 477 *Morella* individuals based on SSR sequencing data. (a) Presenting the Delta K; (b) The STRUCTURE diagram of 477 samples at *K*=4 shows that green represents *M. nana*, red represents *M. rubra*, blue represents *M. esculenta* and yellow represents *M. adenophora*, and showed the gene exchange among the four species; (c) Interpopulation NJ trees of *Morella* based on Nei's genetic distance.

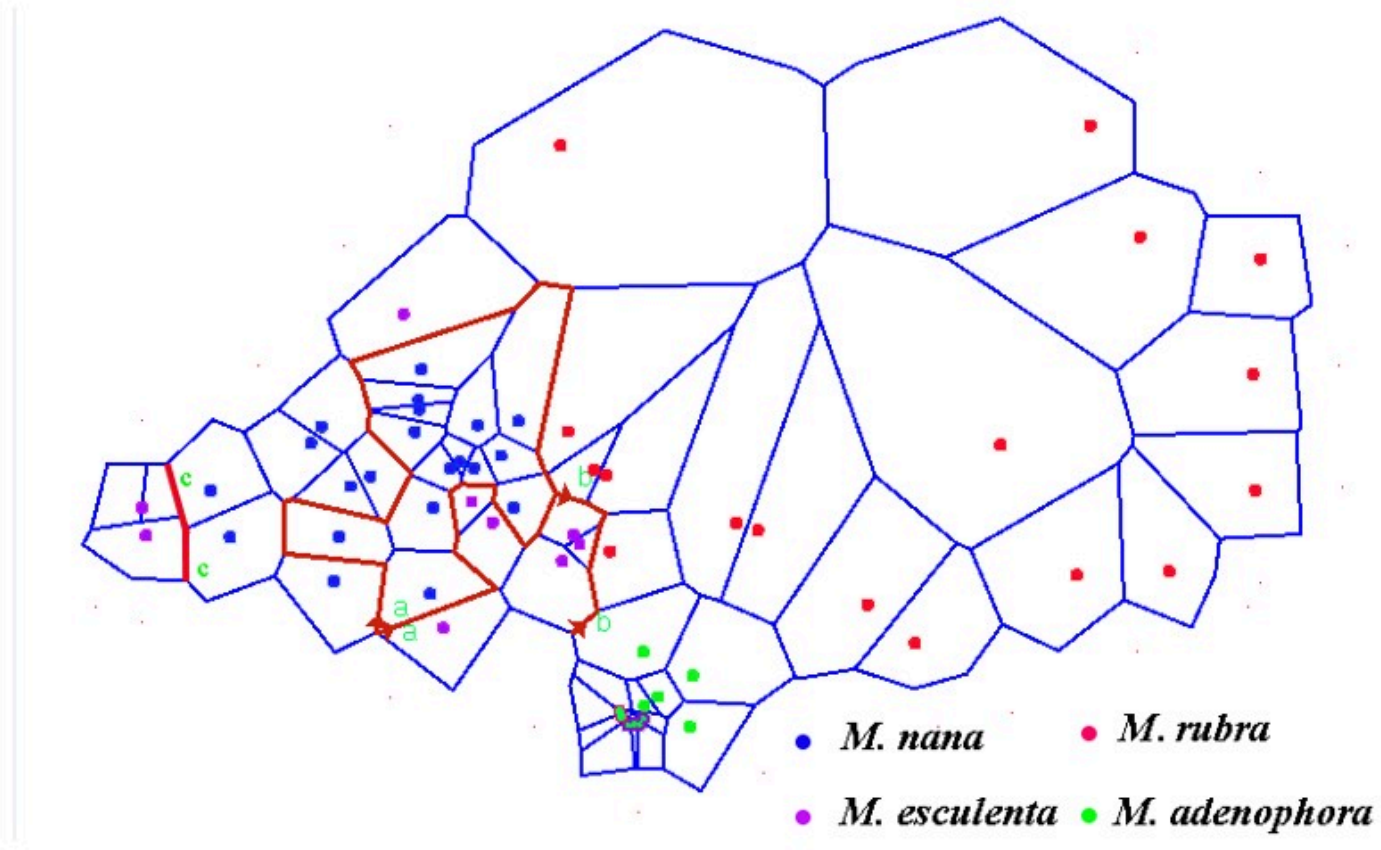

Figure 4

Genetic barrier analysis of four *Morella* species based on cpDNA data. Red represents the population location of *M. rubra*, blue represents the population location of *M. nana*, purple represents the population location of *M. esculenta*, and green represents the population location of *M. adenophora*. Blue lines indicate Delaunay triangulation, and red lines indicate Barrier robustness (a, b and c represent three genetic boundaries).

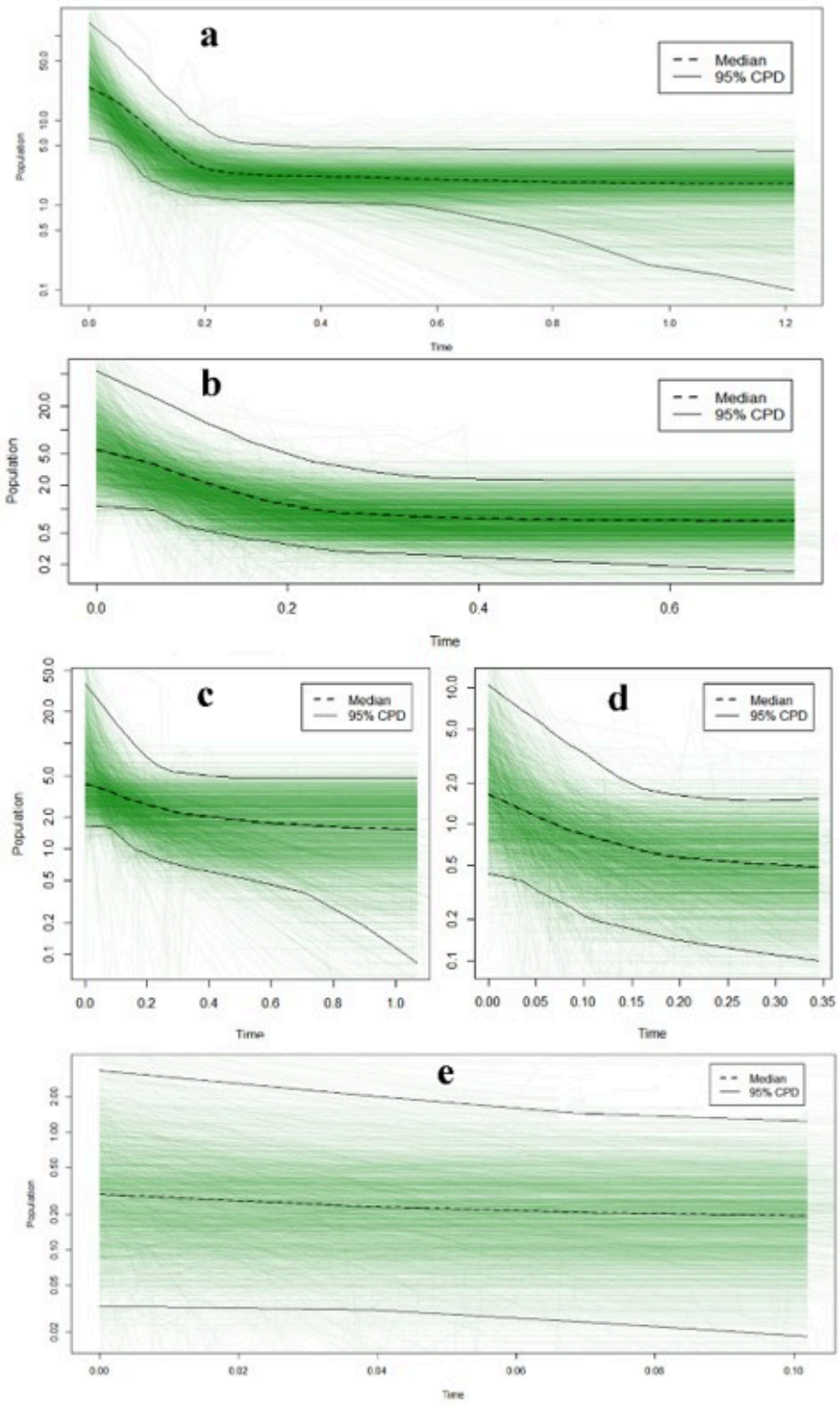

Figure 5

BSPs based on chloroplast DNA sequence. The y axis is the product between effective population size (*N*e) and the generation time and the x axis is time in millions of years. The dotted line represents the median of the *N*e. The upper and lower black lines represent the confidence interval of the highest 95% posterior density (HPD) interval of *N*e. (a) BSPs for *Morella* populations. (b) BSPs for *M.*

*nana*populations. (c) BSPs for *M. esculenta* populations. (d) BSPs for *M. rubra* populations. (e) BSPs for *M. adenophora* populations.

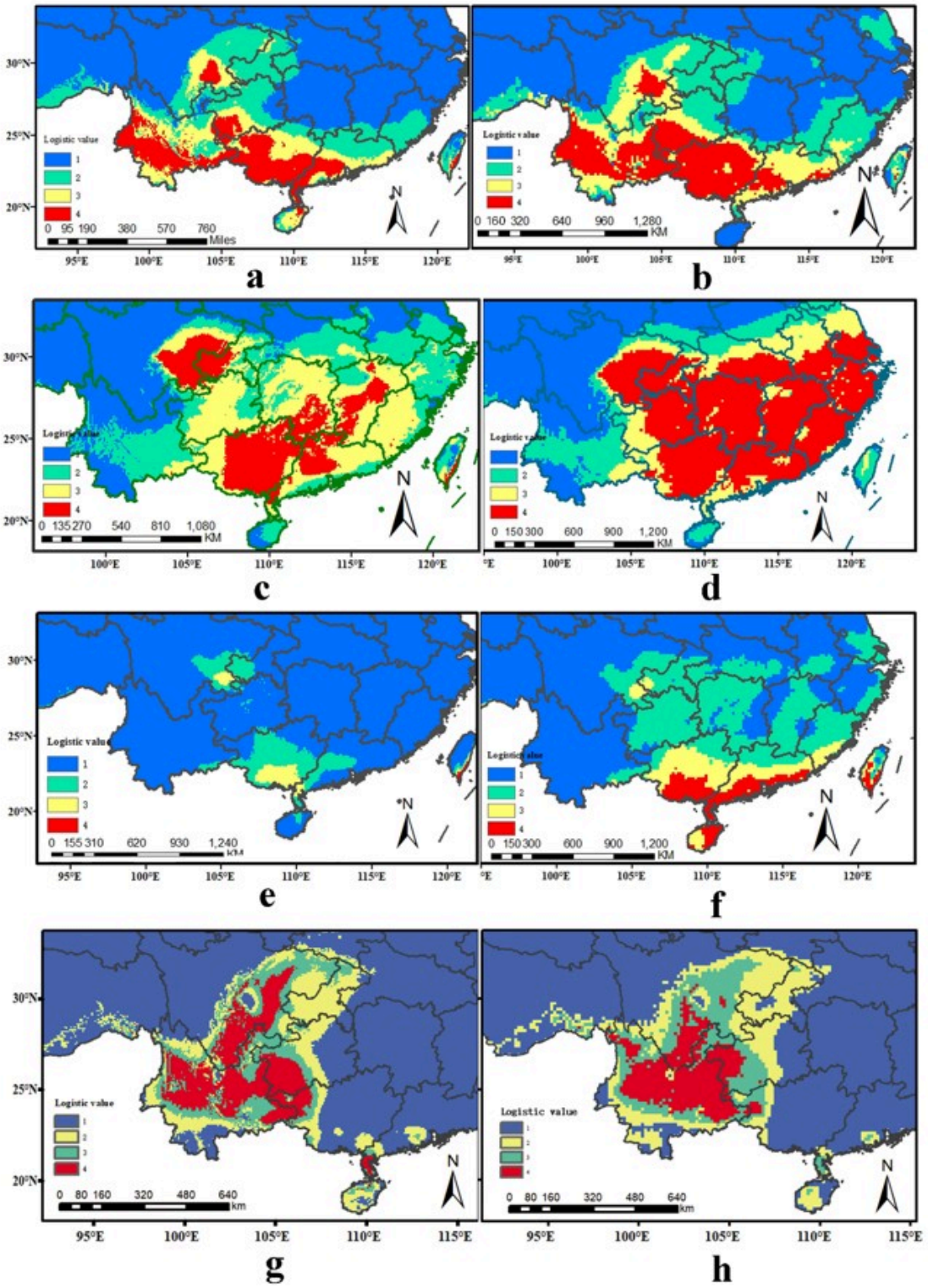

Figure 6

Potential distribution of *Morella* based on ecological niche modeling using MAXENT. (a) Predicted distribution of *M. esculenta* during the LGM. (b) Predicted distribution of *M. esculenta*during the current. (c) Predicted distribution of *M. rubra* during the LGM. (d) Predicted distribution of *M. rubra* during the

current. (e) Predicted distribution of *M. adenophora* during the LGM. (f) Predicted distribution of *M. adenophora* during the current. (g) Predicted distribution of *M. nana* during the LGM. (h) Predicted distribution of *M. nana* during the current. g and h quote the published articles of our team (Wu et al. 2024).

Supplementary Files

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

- Supplementarymaterial.docx