Demographical history and genetic differentiation of an endemic and endangered Ulmus lamellosa (Ulmus)

Huimin Hou
Shanxi Normal University
Hang Ye
Shanxi Normal University
Zhi Wang
Shanxi Normal University
Jiahui Wu
Shanxi Normal University
Yue Gao
Shanxi Normal University
Wei Han
Shanxi Normal University
Dongchen Na
Shanxi Normal University
Genlou Sun
Saint Mary's University
Yiling Wang (ylwangbj@hotmail.com)
https://orcid.org/0000-0002-0358-0851

Research article

Keywords: Aat, demographical history, ITS, genetic differentiation, Ulmus lamellosa

Posted Date: March 23rd, 2020

DOI: https://doi.org/10.21203/rs.3.rs-18212/v1

Version of Record: A version of this preprint was published at BMC Plant Biology on November 17th, 2020. See the published version at https://doi.org/10.1186/s12870-020-02723-7.

Demographical history and genetic differentiation of an endemic and endangered *Ulmus lamellosa* (*Ulmus*)

Huimin Hou[1#], Hang Ye[1#], Zhi Wang[1], Jiahui Wu[1], Yue Gao[1#], Wei Han[1], Dongchen Na[1], Genlou Sun[2], Yiling Wang[1*]

Huimin Hou: houhm0308@outlook.com; Hang Ye: 715610810@qq.com; Zhi Wang: 772578103wangzhi@gmail.com; Jiahui Wu: 18303478207@163.com; Yue Gao: 2090148512@qq.com; Wei Han: 2864537981@qq.com; Dongchen Na: 443708198@qq.com; Genlou Sun: genlou.sun@smu.ca

1 College of Life Science, Shanxi Normal University, Linfen, China

2 Saint Mary's University, Halifax, Canada

# These authors have contributed equally to this work

Correspondence: *Yiling Wang, Shanxi Normal University, Linfen, 041000, PR China; Tel.: +86-183-0347-9941;Fax: +86-357-2051630; Email: ylwangbj@hotmail.com

Abstract

Background: *Ulmus lamellosa*, an endemic and endangered species and one ancient species of *Ulmus*, has undergone the climate oscillations and geographical changes. Elucidating its demographical and genetic differentiation is very important for understanding the evolutionary history and ecological adaption of forests in Northern China.

Results: According to ITS and *Aat* sequences, eighteen and twenty-three haplotypes were detected respectively. The haplotype distribution was polymorphic in most populations. All levels-phylogeographic clades were geographically structured, namely THM, YM and YSM groups. Within *U. lamellose*, higher genetic diversity and significant genetic variation were present, YSM and THM had a relatively higher diversity than that of YM. The divergence of *U. lamellosa* intraspecies haplotypes occurred during Miocene-Pliocene that associated with Tertiary major geological and/or climatic events, which was supported by the gene exchanges among three groups. During the glaciation, YSM and THM regions might be regarded as refugia for *U. lamellose*. A range expansion was not expected at the evolutionary process, except the THM group based on ITS data. The series Mountains uplift (such as Yanshan Mountains and Taihang Mountains) in North China after Miocene-Pliocene and subsequently Quaternary climatic oscillations further promoted the divergence among populations.

Conclusions: Geographical topology and climate change played a critical role on the phylogeographic structure of *U. lamellosa*, leaded to the current pattern of *U.*

lamellosa. These results would provide important information and clues for the demographical study of the trees in Northern China.

Key words: *Aat*; demographical history; ITS; genetic differentiation; *Ulmus lamellosa.*

Background

Examining how historical and contemporary ecological factors contribute to demographical history and genetic differentiation of plants is a hotspot question in ecology and evolution. Climatic oscillations have greatly influence on the demographic history and genetic diversification patterns of many plant species, especially during the Pleistocene [1-8]. In Asia, at least four major glaciations are thought to have occurred, although the glacial advance was not as extensive as that in Europe and North America, which was believed to have affected the flora and fauna in the area. Particularly, the Northern China experienced severe climatic oscillations throughout the Quaternary, despite it was never covered by massive ice-sheets [9-11]. The climatic oscillations caused by repeated glaciations shifted species ranges and isolated populations, which could generate a distinct population structure, namely, "southern richness" and "northern purity" [12-14].

Meanwhile, a possible role of geology in facilitating the population divergence and species diversification of plants has been revealed [9, 11, 15, 16]. The geographical and physiognomy complexity due to the mountains in China [17, 18], which often offered dispersal corridors allowing range expansions/contractions, population

connection, and gene flow between fragmented populations [19, 20], or might serve as shelters for range shifts of plants in response to climatic oscillation [21-26]. Some mountains, such as Taihang, Yanshan, Yinshan and Qinling mountains located in North China, are an important part of the south-north vegetation transect, and crucial biodiversity areas covered with diverse plant biomes ranging from tropical to cold forests and taiga [27, 28]. These mountains formed diverse topography and a wide spectrum of environmental conditions, generally provided refuge for species during the Quaternary global climate changes [29, 30]. According information, different refugia patterns during glacial period have been revealed for the forest trees in Northern China, such as single refugium, multiple refugia, microrefugia and crytic refugia [31-33]. Nevertheless, most studies still has focused on the population or species in southern, southwest and northwest regions of China [3, 24, 34]. The possible roles of ecological factors (geology and climatic oscillations) in promoting the population divergence of plants in Northern China are unclear.

Ulmus lamellose (*Ulmus*), mainly distributed and endemic in North China, is an important tree in many communities and terrestrial ecosystems. Abundant decylic acid in seeds makes this species as a critical medicinal and light industrial raw material [35], and the trunk shows the rich ornamental characteristics [36]. The population size of this species is decreasing constantly with the environmental degradation and human destruction, and it has been classified as the Class II State-Protected Endangered Plant Species in China [37, 38]. Being an ancient and long-lived tree species [39], *U. lamellose* might have existed before the uplift of Taihang, Yinshan

and Yanshan mountains, underlay the climatic oscillations of Quaternary, even the Tertiary, and was believed an ideal species for studying the influence of geological and climatic dynamics on species divergence in North China.

Using chloroplast DNA sequences, Liu et al. (2017) preliminary investigated the molecular phylogeographical structure and biogeographic history of this tree species. A significant phylogeographic pattern was found in *U. lamellose* and the intraspecific divergence of all cpDNA haplotypes began to occur in the late Miocene. Most studied populations of *U. lamellose* had not experienced recent expansions. Across the distribution ranges of this species, multiple refuge areas were identified in North China [40]. These results provided some lights on the phylogeography and demographical history of *U. lamellose*. However, chloroplast genes have the relatively slower evolutionary ratio than nucleus genes, are generally maternal inheritance, and cannot reflect the pollen flow in angiosperms [41-43]. Furthermore, the results based on chloroplast genes only provide limit information on phylogeographic structure and evolutionary history of plants, especially intra-species [25, 44].

In this study, we point the hypothesis of *U. lamellose* having multiple refuges in Norther China and the geological changes and climatic oscillations playing an important role on demographical history for this species. So, our main goals are to comparatively investigate the demographical history and genetic differentiation of *U. lamellose*, analyze the phylogeography pattern of this boreal tree, identify the refugia patterns during glacial period and estimate the intra-special divergent time during evolutionary process of this species. These results would shed light on the influence

of climatic oscillations and orogeny of Mountains (Taihang, Yanshan and Yinshan mountains) of North China, and offer a theoretical foundation for studying the phylogeography and demographical history of other forest tree in North China.

Methods

List of abbreviations

Abbreviations	Full name
ITS	Internal transcribed spacer
Aat	Single-copy gene
THM	Taihang Mountains
YM	Yanshan Mountains
YSM	Yinshan Mountains
PCR	Polymerase chain reaction
AMOVA	A hierarchical analysis of molecular variance
SSD	Sum of squared deviations
HRag	Harpending's raggedness index
ABC	Approximate Bayesian Computation
IBD	Isolation by distance
IBE	Isolation by environment

Plant materials

Our field work was approved by the Chinese government and carried out in accordance with the laws of the People's Republic of China. All participants held a letter of permission from the School of Life Sciences of Shanxi Normal University to collect samples.The samples were classified and identified by the associate professor of systematic botany (Junxia Su) of Shanxi Normal University based on their phenotypic characteristics. Voucher specimen were kept in the Herbarium of College of Life Sciences, Shanxi Normal University (No: 20180703422 - 20180703436).

Two hundred and thirty-two individuals from fourteen natural populations of *U.*

lamellose were used in this study. Based on the wild investigation and samples location, three main geographical groups were assigned based on mountain ranges, which were marked as THM (Taihang Mountains), YSM (Yanshan Mountains) and YM (Yinshan Mountains) respectively. All sampling plots covered most of its entire distribution range in North China. Individuals within each population were separated at least 30m to avoid sampling clones of the same individual. The detail geographical location of the populations and the number of sampled individuals were given in Additional file 1: Table S1 and Figure 1. Leaves of each sample using for DNA extraction were collected and stored in silica gel at room temperature.

DNA extraction, amplification and sequencing

The genomic DNA was extracted from fresh leaves using CTAB methods [45, 46]. The quality and concentration of the DNA samples were determined by spectrophotometer and electrophoresis on 0.8% agarose gel.

The regions ITS4, ITS5 of Ribosomal DNA internal transcribed spacer (nrDNA ITS) and Single-copy gene (*Aat*) [47, 48] were chosen for this study because of high sequence variation within and among populations of *U. lamellosa.* Polymerase chain reaction (PCR) was performed in a 25μL volume, with 4 μL of template DNA at approximately 30 ng/μL, 0.5 μL of each forward and reverse primers, 12.5 μL of 2 × MasterMix (0.2 mM dNTPs, 3 mM $MgCl_2$, 1× PCR buffer and 0.1 U *Taq* DNA polymerase), and 7.5 μL of ddH_2O. The PCR was run in a PTC-200 Peltier Thermal Cycler (MJ Research) using the following program: 3 min at 94 °C, followed by 34 cycles of 30 s, 1 min at 55 °C (for ITS and *Aat*), and 2 min at 72 °C, and a final

146 elongation step of 10 min at 72 °C. The PCR products were visualized via 1.5%
147 agarose gel electrophoresis and purified using a Wizard PCR Preps DNA purification
148 system (Promega, Madison, WI, USA). The purified PCR products were sequenced
149 using on an ABI Prism 310 DNA sequencer (Applied Bio systems Inc., Foster City,
150 CA, USA) by Shanghai Sangon Biotech Co., Ltd.

151 ***Data analysis***

152 ***Genetic diversity and Differentiation Analysis***

153 The DNA sequences from nrDNA ITS and *Aat* were aligned using CLUSTAL X
154 software [49] and then adjusted manually. All indels were coded as unordered and
155 equal weight [50, 51]. The nucleotide diversity (π) and haplotype diversity (H_d) were
156 calculated for each population by using DNASP software [52-55]. Genetic
157 differentiation among populations was estimated by computing a distance matrix
158 based on the number of mutational steps between haplotypes (N_{ST}) and by using
159 haplotype frequencies (G_{ST}). The presence of a phylogenetic component in the
160 phylogeographic structure was assessed by testing whether N_{ST} was significantly
161 higher than G_{ST}, on the basis of 10,000 permutations in PERMUT software [56-58].

162 In order to study the relationship between the differentiation and evolution of
163 haplotypes, the haplotype network diagram was constructed by using the maximum
164 parsimony median-joining model in NETWORK software [59].

165 The software STRUCTURE was used to infer the population structure [60-62]. A
166 given number of genetic clusters (K) were used to perform Bayesian assignment of
167 individuals assuming that each cluster in Hardy–Weinberg and linkage equilibrium,

168 and ΔK was calculated to determine the optimal K (from K=1-10) [63].

169 A hierarchical analysis of molecular variance (AMOVA) [64] was conducted to
170 quantify the proportion of total genetic variance explained by the differences between
171 groups and among populations within groups. The AMOVA using pairwise
172 differences was conducted using the ARLEQUIN software [54, 65], with the
173 significance tested by 1000 permutations.

174 ***Mismatch distribution***

175 Mismatch distribution analyses [66] were carried out in ARLEQUIN [65] to infer
176 the historic demographic expansion events within the entire species and in each group.
177 The goodness-of-fit under a sudden expansion model was tested with the sum of
178 squared deviations (SSD) and Harpending's raggedness index (HRag) [67]. A
179 parametric bootstrap approach with 1000 replicates was used to test the observed
180 mismatch distribution's fit to the sudden expansion model. Tajima's D (Tajima 1989)
181 [68] and Fu's F_S (Fu 1997) [69] were also calculated to assess possible expansions
182 using ARLEQUIN [65]. When the expansions were detected, the parameter value for
183 the mode of the mismatch distribution (τ) was converted to an estimate of time (T)
184 since the expansion began by using $T = \tau/2u$ [66, 70]. The value u was calculated as
185 $u = \mu kg$, where μ is the number of substitutions per site per year (s/s/y), k is the
186 average sequence length under study, and g is the generation time in years [71, 72].
187 The parameter value α (maturity age of species, about 30 years) and $[S/1-S]$ (about
188 100; S, annual survival rate of adult individuals) was converted to an estimate of time
189 (g, in number of generations) by using $g = \alpha + [S/(1-S)]$ [73, 74].

Divergence time

The divergence time among haplotypes was estimated to elucidate the historic events by calibrating a molecular clock implemented in BEAST [44, 75-79]. The HKY model was selected as the optimal model based on the results of model test for nucleotide substitutions using JMODELTEST software [80, 81], and coalescence with constant size was set as the prior tree. Because the mutation rate has not been estimated for ITS in *U. lamellosa* or members of the Ulmaceae family, we used minimum and maximum values of a range of average mutation rates reported of ITS in angiosperm [3.2×10^{-10} and 9.0×10^{-9} substitutions per site per year (s/s/y)] [82]. The analysis was run for 50,000,000 generations, sampling every 1000th generation. The log file was analyzed in TRACER [83], with 1,000,000 generations discarded as burn in to ensure that the ESS values for all parameters reached a sufficient level (>200). The trees were analyzed using TREEANNOTATOR, and visualized in FIGTREE [79, 84].

Approximate Bayesian Computation

Demographic history was modeled through coalescent-based Approximate Bayesian Computation (ABC) [76, 85-87] by DIY-ABC software [88]. In order to reduce the number of model settings and operation time, we used three geographical groups of *U. lamellose* populations. First, seven population divergence scenarios were tested: either a Simple split model (scenario 1), Hierarchical split model (scenarios 2-4), or Isolation with admixture model (admixture, scenarios 5-7). The prior distributions of all parameters were uniform, and the same range was used for

common parameters between models. A total of 500,000 coalescent simulations were run for each scenario. There was up to a 95% confidence level through logistic regression, and the posteriori probability of each evolutionary event was estimated using the observation data closest to 1% of the simulated data. An optimal evolutionary scenario was selected according to the maximum posterior probability (PP) value. Based on the scenario with the highest support from the population divergence data, optimal scenarios with effective population size changes were then tested. Finally, the ancestral group of the optimal scenarios was further inferred.

Historical and contemporary gene flow

The coalescent approach implemented in MIGRATE-N software was used to estimate the historical gene flow pattern among groups of *U. lamellose* [89-92]. This method through Bayesian method estimates for both effective population size (θ =4$Ne\mu$, where μ = mutation rate) and historical migration rates ($M = m/\mu$, where m = migration rate) among groups [89, 93]. The parameters θ and M could be used to estimate the number of migrants per generation (Nm) into each population using the equation $4Nm = \theta * M$ [89, 93, 94]. The analyses were run for five replicates using a F84 mutation model, with constant mutation rates and starting parameters based on F_{ST} calculations. For each replicate, we used 10 short chains (10,000 trees) and three long chains (200,000) with 100,000 trees discarded as an initial 'burn-in' and a static heating scheme at four temperatures (1, 1.2, 1.5 and 3, 1000,000). The historical gene flow was the average of the results of five independent operations, with a confidence interval of 95%.

234 The BAYESASS software (Wilson and Rannala 2003) that carries assignment
235 tests in a Bayesian framework and a Markov coupled Markov chain (MCMC) was
236 used to estimate contemporary gene flow [90, 92] between the above three groups of
237 *U. lamellose*. After a burn-in of 10^6, the analyses with a sampling frequency of 10^3
238 were ran 10^7 iterations. Delta values for the migration (m), allele frequencies (a) and
239 inbreeding (f) switching proposals were adjusted so that the accepted numbers of
240 changes were 20–60% of the total number of iterations. The delta values for a, m, and
241 f were 0.5, respectively. The analyses were performed 20 runs (each with a different
242 seed) and the one with the lowest deviance was selected for further analyses. The 95%
243 confidence interval was calculated by mc ±1.96 × SD [74].

244 ***Isolation by distance and environment (IBD and IBE)***

245 The correlations between pairwise F_{ST} estimated from the ITS /*Aat* sequences
246 and geographic distance, as well as between pairwise F_{ST} and environmental distance,
247 were calculated to assess the effect of geographic distances and environmental
248 conditions (IBD and IBE) on the observed genetic differentiation [54, 95, 96]. A
249 matrix of pairwise F_{ST} and geographic distance between populations were calculated
250 using GENALEX software [97]. The matrices of the environmental distance were
251 computed using PASSAGE software [98].

252 A multiple matrix regression with randomization (MMRR) was further
253 performed to test whether the genetic distance responded to variations in the
254 geographic and/or environmental distances using the R package 'PopGenReport' [99,
255 100]. Regression coefficients of the Mantel test (r) and MMRR (r^2) and their

256 significance were determined based on 9,999 random permutations.

257

258 **Results**

259 ***Phylogeographical structure and genetic differentiation***

260 Alignment of nrDNA ITS and *Aat* sequences from 232 *U. lamellose* individuals
261 produced a consensus sequence of 652 and 392 base pairs, respectively.

262 For ITS, eighteen haplotypes were defined with 27 polymorphic sites, including
263 17 parsimony informative sites (Table 1). Among the 18 haplotypes, the most
264 common haplotype was H1 (in 13 of 14 populations), followed by H2 (in 3 of 14
265 populations). Thirteen of the 14 populations were polymorphic, whereas the only one
266 population exhibited one haplotype (MS population; Table 1 and Fig. 1a). All
267 haplotypes presented a 'star-like' network, the common haplotype (H1) was in the
268 central position regarded as an ancient one that connected other descendant
269 haplotypes (Fig. 1b).

270 For *Aat*, twenty-three haplotypes (H1-H23) were defined containing 33
271 polymorphic sites, of which 26 were parsimony informative sites (Table 1). Among
272 the 23 haplotypes, 13 haplotypes were unique for populations and 10 haplotypes were
273 shared by different populations, and the most common haplotype was H4 (in 5 of 14
274 populations). Most of the population contained both shared and unique haplotypes.
275 Eleven of the 14 populations had high haplotype diversity, and only one haplotype
276 was found in MS, LS and WLS population. The geographical distribution of
277 haplotypes H1 to H23 and their occurrence at each location were shown in Fig. 1c and

278 Table 1. Compared other haplotypes, H4 had a more connection and relatively
279 situated in the middle of the network, was an ancient one (Fig. 1d).

280 High genetic diversity in *U. lamellosa* was detected. At the species level, the
281 estimated haplotype diversity ($H_{d\ ITS}$, $H_{d\ Aat}$) were 0.424 and 0.941, and the nucleotide
282 diversity (π_{ITS}, π_{Aat}) were 0.00212 and 0.01601, respectively. For three groups, the
283 highest values of $H_{d\ ITS}$ (0.909) and π_{ITS} (0.00939) existed in the YSM group,
284 followed by the THM and YM groups. On the contrary, the highest values of $H_{d\ Aat}$
285 and π_{Aat} existed in the THM and YM groups. Among populations, the haplotype
286 diversity ($H_{d\ ITS}$, $H_{d\ Aat}$) was 0-0.929 and 0-1.000, and the nucleotide diversity (π_{ITS}, π
287 $_{Aat}$) within populations was 0-0.00852 and 0-0.01197(Table 1). However, the highest
288 value of H_d was detected in different populations for different data ($H_{d\ ITS}$=0.798 for
289 SFS population; $H_{d\ Aat}$=1.000 for SGS, HHG, SFS, MLG populations). The highest π
290 $_{ITS}$ value also existed in the SFS population (0.00583), the highest π_{Aat} value was
291 found in SGS and HHG populations (Table 1).

292 No matter which molecular markers were used, genetic differentiation among
293 populations was significantly higher when computed using a distance matrix ($N_{ST\ ITS}$
294 = 0.435; $N_{ST\ Aat}$ = 0.738) than when using haplotype frequencies ($G_{ST\ ITS}$ = 0.117; G_{ST}
295 $_{Aat}$ = 0.192; $P < 0.05$), indicating a significant phylogeographic structure within *U.*
296 *lamellosa*. For ITS, the STRUCTURE output showed that ΔK supported the existence
297 of three clusters (Additional file 3: Fig. S2), which corresponded roughly with our
298 geographically assignment to *U. lamellosa* (Table 1 and Fig. 1). Nevertheless, the
299 STRUCTURE output of *Aat* data showed that ΔK supported the existence of four

clusters (Additional file 3: Fig. S2), but we prefer to divide the *U. lamellosa* into three clusters because the THM group could be further classified into two clusters.

Higher inter-population differentiation of this species was confirmed by AMOVA analysis ($\Phi_{ST\ ITS} = 0.550$, $\Phi_{ST\ Aat} = 0.674$, $P < 0.01$, Table 2). A large amount of variation (ITS: 40.60%, *Aat*: 41.96%) occurred among the three groups, and only 14.43% (ITS) and 25.40% (*Aat*) differences presented among populations within groups, whereas 44.97% (ITS) and 32.64% (*Aat*) of the variation was within populations (Table 2).

Phylogenetic relationships and divergence time estimations

The phylogenetic relationships of the ITS haplotypes (H1–H18) reconstructed using Bayesian approach was shown in Fig. 2. All haplotypes were clustered into three major clades I-III, which was basically consistent with the geographical distribution of the populations (namely YSM, THM and YM group). Clade I contained H9-H12 in the SFS population, which were found in YSM group. Clade II contained H8 and H16, which were found in THM and YSM group. Clade III consisted of twelve haplotypes (H1-7, H13-15, H17 and H18) found in THM and YM group.

For *Aat* data, all haplotypes were also clustered into three major clades: I, II and III (Additional file 2: Fig. S1), which was basically consistent with that of haplotype network. Clade I and II contained most of haplotypes, which were found in THM and YSM group. Clade III contained H5-H7, H15-H17 and H23, which were found in THM and YM group.

322 The phylogenetic tree based on ITS data showed that, when *Ulmus_pumila* and
323 *Ulmus_rubra* were used as outgroups, all haplotypes from *U. pumila* and *U. rubra*
324 formed a monophyletic group, the haplotypes of *U. lamellose* formed another
325 monophyletic group (Fig. 2). The ESS values were greater than 200 for all the nodes
326 discussed below. The haplotype identified in outgroups was divergent from all other
327 haplotypes at around 14.50 Ma (95% HPD: 8.47–22.20 Ma). The eighteen haplotypes
328 (H1–H18) of *U. lamellosa* formed a lineage that was divergent at around 13.10 Ma
329 (95% HPD: 7.79–20.20 Ma). This suggested that the intraspecific divergence of all
330 haplotypes most probably began in the period between Middle Miocene to the early
331 stage of the Late Miocene. Within *U. lamellose*, all haplotypes could be roughly
332 classified into three lineages, and the divergence time from clade I to clade III were
333 4.56 Ma (95% HPD: 1.77–8.09), 2.69 Ma (95% HPD: 0.05–7.74) and 10.18 Ma (95%
334 HPD: 5.69-15.90), respectively.

335 ***Demographical history***

336 The SSD and HRag values were non-significant ($P > 0.05$) of all groups based
337 on ITS sequences. Nevertheless, the unimodal mismatch distribution for THM was a
338 perfect fit to the expected expansion distribution (Table 3), as supported by a
339 significant Tajima's *D* statistic value. Thus, *U. lamellosa* experienced a recent
340 demographic expansion in the Taihang Mountains and the demographic expansion
341 occurred approximately 0.0038 Ma. In addition, mismatch distributions and neutrality
342 test for *Aat* haplotypes were multimodal and differed strongly from the predictions
343 under a model of sudden population expansion, which was evidenced by

non-significant values ($P > 0.01$) of Tajima's D statistics for three groups (Table 3).

An ABC framework permitted an evaluation of speciation scenarios for *U. lamellosa.* The scenario 2 (Fig. 3a ,3b) had the highest posterior probability (ITS: Direct estimate: 0.4540, 95% CI: 0.0176-0.8904; Logistic regression: 0.5913, 95% CI: 0.5694-0.6132; *Aat*: Direct estimate: 0.2360, 95% CI: 0-0.6082; Logistic regression: 0.4695, 95% CI: 0.4584-0.4807) and did not overlap with those obtained for the other scenarios (Fig. 3c, 3d, 3e, 3f). And this scenario was fitted to the observed data confirmed by PCA results. According to the model checking, error for the scenario 2 was low (ITS: Direct approach: 0.061, Logistic approach: 0.075; *Aat*: Direct approach: 0.071, Logistic approach: 0.077). Consequently, the scenario 2, namely, YSM group might be a common ancestor of THM and YM group diverged at time t2. Subsequently, YM group diverged from THM group at the time t1 (Fig. 3a, 3b). Under the scenario 2, posterior media parameter estimates indicated that the first divergence occurred in 10.36×10^6 (ITS: 95% CI: 4.23×10^6 – 12.88×10^6) and 8.25×10^6 (*Aat*: 95% CI: 2.56×10^6 – 12.73×10^6) years with a generation time of 130 years.

The effective population size was estimated based on the sequencing data. For ITS data, the THM group has the largest effective population size (5.92×10^5, 95% CI: 1.88×10^5– 9.52×10^5) among groups, followed by YSM group (4.56×10^5, 95% CI: 1.06×10^5– 9.49×10^5) and YM group (1.74×10^5, 95%CI: 1.96×10^4– 8.59×10^5). For *Aat*, the median values of the effective population size were 2.35×10^5 (95% CI: 4.65×10^4– 8.70×10^5) for THM group, 2.79×10^5 (95% CI: 4.50×10^4– 9.21×10^5) for YSM (the ancestral group), and 2.17×10^5 (95% CI: 1.56×10^4– 9.27×10^5) for YM

366 group. There was no significant difference in effective population size among all
367 groups.

368 ***Historical and contemporary gene flow***

369 In general, the historical gene flow was asymmetric. For ITS sequences, the
370 historical gene flow (0.204-1.752) were higher than the modern gene flow. The
371 highest level of migration occurred from YSM to THM (1.752, 95% CI: 0.394-3.797),
372 and the lowest level of gene flow occurred in the opposite direction (THM to YSM:
373 0.204, 95% CI: 0.013-0.245, Fig. 4a). The higher level of migration occurred between
374 THM and YM (YM to THM: 0.923, 95% CI: 0–2.931; THM to YM: 0.839, 95% CI:
375 0.122-1.999, Fig. 4a) and the lower level occurred from YSM to YM (0.714, 95% CI:
376 0.117-1.999). According to *Aat* sequences, the historical gene flow was generally high.
377 The higher level of migration also occurred from THM to YM (6.184, 95% CI:
378 0-46.437; Fig. 4c) than the opposite direction. Nevertheless, a relative high level of
379 migration occurred from YSM to YM (3.477, 95% CI: 0–4.744).

380 In addition, modern gene flow was low and asymmetrical. The BAYESASS
381 analysis showed that the mean contemporary gene flow (m) of ITS among three
382 groups ranged from 0.008 to 0.279 (Fig. 4b). Similarly, the higher level of migration
383 occurred from YSM -THM-YM (YSM-THM: 0.058, 95% CI: 0-0.138, THM-YM:
384 0.264, 95% CI: 0.104-0.423). However, a strong migration occurred from YM to
385 THM (0.279, 95% CI: 0.208-0.349). For *Aat* sequences, the mean contemporary gene
386 flow (m) among three groups ranged from 0.008 to 0.260 (Fig. 4d). The highest level
387 of migration occurred from YM to THM (0.260, 95% CI: 0.175-0.345). And the

higher level of migration also occurred from YSM-THM-YM (YSM-THM: 0.030, 95% CI: 0-0.084, THM-YM: 0.036, 95% CI: 0.006-0.065).

Isolation by distance and environment (IBD and IBE)

Pairwise F_{ST} values between populations used as genetic distance were plotted against the respective pairwise geographic distance and environmental distance (Additional file 4: Fig. S3). The positive correlation between the matrix of genetic distances and geographical distances was significant (ITS: r = 0.594, $P < 0.05$; *Aat*: r = 0.391, $P < 0.05$). The correlations between pairwise F_{ST} estimated from the ITS /*Aat* and environmental distance were non-significant ($P > 0.05$).

The multiple matrix regression with randomization (MMRR) analyses also revealed a significant effect of IBD on the divergence of this species. Furthermore, IBD explained more genetic variation than IBE (ITS: MMRR ≈ 12.4 times as much; *Aat*: MMRR ≈ 3.7 times as much).

Discussion

Phylogeographic structure and genetic differentiation of U. lamellose

A high level of genetic diversity was observed at *U. lamellose* ($H_{d\ ITS}$=0.424, π_{ITS}=0.00212; $H_{d\ Aat}$=0.941, π_{Aat}=0.01601), which was similar with other plant, such as *Gentiana lawrencei* var. *farreri* (H_d=0.414, π=0.0026) and was consistent with the results of Liu et al. (2017)[40]. Compared with cpDNA data (H_d=0.783, π=0.00891)[40], the genetic diversity index of *U. lamellose* was higher than that of ITS sequences and lower than that of *Aat* sequences (Table 1). The different

evolutionary trajectories and rates of nuclear and plastid markers may be the cause of different degrees of variation and reflect different biogeographic events such as pollen flow and seed dispersal [7, 16, 18, 44, 57, 101, 102]. And the difference between ITS and *Aat* data might be due to that the single copy gene was more conservative.

High genetic diversity might reflect the accumulation of nucleotide mutations over long evolutionary time-scales for a wind-pollinated, dioecious and perennial of *U. lamellosa* species [103, 104]. In addition, the geographical distribution of this species may also be one of the factors determining high genetic diversity. Widespread species generally may have higher levels of genetic variability [105]. *U. lamellose* widely distributed in North China (Additional file 1: Table S1 and Fig. 1).

In YSM group, higher genetic diversity was than in THM/YM group (Table 1). Being an ancestor for other two groups populations (Fig. 3), this group could have a higher diversity and variation. Meanwhile, it suggested a smaller influence of historical climatic oscillations on this group. The Yanshan Mountains, running roughly east-west, is an important mountain in North China [106]. The heterogeneous geology in this mountain might have facilitated the existence of multiple glacial refugia for the plants [28, 107]. In our study, higher haplotype diversity was detected in YSM group. The populations in this group contained five unique haplotypes (H8-H12, H16). Therefore, YSM group could be regarded as a refugia for *U. lamellose* at the glacial period. In addition, although THM group having lower genetic diversity, QLY, SGS, JM and HLS populations in this group (Table 1) presented highest genetic diversity. THM should be believed another refugia for *U. lamellose*.

YS group, was not regarded as a refugia in this study, inconsistent with Liu et al. (2017), not only having the lowest genetic diversity but also being a descendent of THM and YSM (Fig. 3).

Significant genetic differentiation among groups and among populations within groups ($\Phi_{ST\ ITS}$= 0.542, $\Phi_{ST\ Aat}$=0.620, $\Phi_{CT\ ITS}$ =0.406, $\Phi_{CT\ Aat}$ =0.562, Table 2) were revealed in the studied populations, which suggested a long period of isolation in the distribution of *U. lamellose*. Our results also found asymmetrical and lower contemporary gene flow (ITS: 0.0058-0.0690, *Aat*: 0.0083-0.1636, Fig. 4) among three groups. Meanwhile, a significant isolation by distance (r_{ITS} = 0.594, r_{Aat} = 0.39, $P < 0.05$) was found in populations of the species. Such findings pinpointed that mountains as barriers may have driven the population differentiation through long-term geographic isolation. More strikingly, most haplotypes were private to a single group. This pattern was more likely resulted from long-term in situ diversification among isolated montane habitats.

Moreover, a strong phylogeographic structure (ITS: N_{ST}=0.435 > G_{ST}=0.117, $P <$ 0.05, *Aat*: N_{ST}=0.738 > G_{ST}=0.192, $P < 0.05$) was founded, which was consistent with Liu et al. (2017). Theoretically, phylogeographic structure often arises as a consequence of restricted gene flow among populations and genetic drift within populations over relatively long periods of time [56, 108, 109]. According to field investigation, the habitats of *U. lamellosa* were natural fragmentation [38]. With the uplift of Mountains in North China, the populations were isolated and fragmented. The climatic fluctuations throughout inter-/postglacial events would further accelerate

454 the fragmentation of *U. lamellosa*. Such isolation made it difficult for pollinators to
455 locate flowers that limiting dispersal ranges and form potential barriers for gene flow
456 to occur. Due to decay and animal consumption, the dispersal of seeds of *U. lamellosa*
457 was also restricted [36, 57]. The historical and modern restricted gene flow (Fig. 4)
458 among populations supported this view. In addition, fragmentation is a consequence
459 of isolation that can be from either geographic or environmental factors [40]. The
460 fragmented populations of *U. lamellosa* cannot counteract the effects of genetic drift.
461 Under the limited gene flow and genetic drift, the phylogeographical structure
462 presented in this species.

463 ***Demographical history***

464 No matter what molecular markers were used, most of the populations have not
465 experienced expansion in *U. lamellosa*, which supported by the mismatch distribution
466 (Table 3) as well as non-significance Tajima's *D* and Fu's F_S values, was consistent
467 with Liu et al. (2017). Interestingly, *U. lamellosa* experienced a recent demographic
468 expansion in the THM group (ITS) that occurred approximately 3800 years ago. This
469 was strongly supported by ABC analyses that revealed the THM group has the larger
470 effective population size (5.92×10^5, 95% CI: 1.88×10^5–9.52×10^5) than ancestral
471 populations (YSM group). Furthermore, a higher gene flow occurred from THM to
472 YM (Fig. 4).

473 In the northern China, despite the absence of major Quaternary glaciations,
474 significant climatic oscillations still occurred [110-112]. During this period, the plant
475 community might have experienced the upward and downward migration of the local

476 altitude, rather than the southward retreat and northward advance in latitude [113].
477 Therefore, *U. lamellosa* might take refuge in situ and only experience limited
478 expansion during the interglacial period. The geographical distribution of haplotypes
479 also supported this expansion in the historical demography of *U. lamellosa*, with the
480 occurrence of often frequent haplotype H1 (ITS data) and H4 (*Aat* data) in the
481 populations (Fig. 1a, 1c).

482 The intraspecific divergence of all *U. lamellosa* ITS haplotypes (13.10 Ma, 95%
483 HPD: 7.79-20.20 Ma) probably began during the period between Middle Miocene and
484 the early stage of the Late Miocene (Fig. 2). This estimated divergence time was
485 earlier than that based on cpDNA data (9.27 Ma, 95% HPD: 5.17-13.33 Ma), about
486 the Late Miocene [40].

487 The period from Miocene to Pliocene is an important differentiation epoch for
488 the tree species in China. And Miocene was a period of climatic instability and abrupt
489 environmental changes. A global cooling event reoccurred in the Miocene (14 Ma) [11,
490 14, 114]. Before and after 16~12 Ma, the monsoon circulation was significantly
491 enhanced in Asia [14, 115-117]. The severe climate fluctuations prompted the *U.*
492 *lamellosa* populations to migrate from YSM (ancestral group) to southern and then
493 colonized in Taihang Mountains (THM group, descendent one). The deformation of
494 Taihang Mountains occurred with the significant uplift of the Tibetan-Himalayan
495 Plateau [14, 118-120]. And the uplift height of the northern section of Taihang
496 Mountains is only half of the western part of Yanshan Mountains during Miocene
497 [118]. Low height of Taihang Mountains did not become a geographical barrier and

498 did not blocked *U. lamellose* migration to northwest and northeast, which supported
499 by the gene exchange among YSM, THM and YM groups (Fig. 4 and Additional file 3:
500 Fig. S2).

501 In ABC, the divergence between THM and YM group was dated back to 3.54 Ma
502 (ITS) and 3.25 Ma (*Aat*) (ITS: 95% CI: 5.36×10^5- 6.24×10^6; *Aat*: 95% CI: $3.82 \times$
503 10^5- 6.28×10^6; Fig. 3), corresponding to the Pliocene. And the YM group diverged
504 from the THM group (Fig. 3a, 3b), was a descendant of THM. During the Pliocene,
505 most of Taihang Mountains had only < 800m elevation [118]. From Eocene to the end
506 of Pliocene, Yinshan Mountains rapidly raised. The uplift of Yinshan Mountains
507 might hinder the glacial activities and climatic oscillations. Consequently, the
508 populations in Taihang Mountains began to migrate toward north, forming YM group.
509 This also was supported by the higher gene flow from THM to YM group (Fig. 4).

510 In addition, the demographic history of *U. lamellosa* also had been influenced by
511 climatic oscillations and historical processes during the Quaternary [14, 121], which
512 was supported by the divergence time of Clade II (2.69 Ma, 95% HPD: 0.05-7.74, Fig.
513 2). In this period, the Taihang Mountains and Yanshan Mountains were uplifted
514 sharply along with the Himalayan orogeny movement and finally formed in the
515 Pleistocene (about 2.5 Ma B.P.) [16]. Meanwhile, a repeated glacial period had
516 occurred since the Quaternary. These would further promote the differentiation among
517 *U. lamellosa* populations and finally lead to the current pattern of this species.

518

519 **Conclusions**

In summary, the studies found that both geographical topology and past climatic changes have contributed to demographical history and genetic differentiation pattern of *U. lamellosa* in the Northern China. Higher genetic diversity and significant genetic variation occurred in this species. The limited range expansion and the existence of refugia in Northern China, allowed *U. lamellosa* to persist in situ and to maintain high diversity. The divergence of haplotypes occurred during Miocene-Pliocene, which associated with Tertiary major geological or climatic events. Subsequently, the series Mountains (for example Yanshan Mountains and Taihang Mountains) uplift of North China after Tertiary and Quaternary climatic oscillations might block the gene exchange and further promote the divergence of populations. These results in the present study not only revealed the roles of geology and climatic oscillations in driving the population demography and divergence of *U. lamellosa,* but also provide important information and clues for demographical study of the trees in Northern China.

Declarations

Ethics approval and consent to participate: Not applicable.

Consent for publication: Not applicable.

Availability of data and materials: All data generated or analysed during this study are included in this published article [and its supplementary information files].

Competing interests: The authors declare that they have no competing interests.

Funding: This study was funded by the National Natural Science Foundation of China (31970358). Funds are used for plant materials collection and basic experimental material acquisition costs.

Authors' contributions: HHM was a major contributor in writing the manuscript. YH and WZ analyzed and interpreted the raw data. WJH, GY and HW completed the plant materials collection and processing. NDC finished the work of language polishing. SGL and WYL are the providers of the main ideas of manuscript and revised the content of the article. All authors read and approved the final manuscript.

Acknowledgements: Not applicable.

References

1. Wang J, Wu Y, Ren G, Guo Q, Liu J, Lascoux M. Genetic differentiation and delimitation between ecologically diverged *Populus euphratica* and *P. pruinosa*. PLoS One. 2011;6:e26530.

2. Husemann M, Guzman NV, Danley PD, Cigliano MM, Confalonieri VA. Biogeography of *Trimerotropis pallidipennis* (Acrididae: Oedipodinae): deep divergence across the A mericas. Journal of Biogeography. 2013;40:261-273.

3. Wang J, Källman T, Liu J, Guo Q, Wu Y, Lin K, Lascoux M. Speciation of two desert poplar species triggered by Pleistocene climatic oscillations. Heredity. 2014;112:156.

4. Ivory SJ, Blome MW, King JW, McGlue MM, Cole JE, Cohen AS: Environmental change explains cichlid adaptive radiation at Lake Malawi over the

563 past 1.2 million years. Proceedings of the National Academy of Sciences 2016,
564 113:11895-11900.
565 5. Naidina O, Richards K. Pollen evidence for Late Pliocene–Early Pleistocene
566 vegetation and climate change in the North Caucasus, North-Western Caspian Region.
567 Quaternary international. 2016;409:50-60.
568 6. Zhou W, Jin J, Wu J, Chen H, Yang J, Murphy RW, Che J. Mountains too high
569 and valleys too deep drive population structuring and demographics in a
570 Qinghai–Tibetan Plateau frog *Nanorana pleskei* (Dicroglossidae). Ecology and
571 evolution. 2017;7:240-252.
572 7. Ye JW, Zhang ZK, Wang HF, Bao L, Ge JP. Phylogeography of *Schisandra*
573 *chinensis* (Magnoliaceae) Reveal Multiple Refugia With Ample Gene Flow in
574 Northeast China. Frontiers in Plant Science. 2019;10.
575 8. Zhao Y, Pan B, Zhang M. Phylogeography and conservation genetics of the
576 endangered *Tugarinovia mongolica* (Asteraceae) from Inner Mongolia, Northwest
577 China. PloS one. 2019;14:e0211696.
578 9. Hewitt G. The genetic legacy of the Quaternary ice ages. Nature.
579 2000;405:907-913.
580 10. Qiu YX, Fu CX, Comes HP. Plant molecular phylogeography in China and
581 adjacent regions: tracing the genetic imprints of Quaternary climate and
582 environmental change in the world's most diverse temperate flora. Molecular
583 phylogenetics and evolution. 2011;59:225-244.
584 11. Zeng YF, Zhang JG, Abuduhamiti B, Wang WT, Jia ZQ. Phylogeographic

585 patterns of the desert poplar in Northwest China shaped by both geology and climatic
586 oscillations. BMC Evol Biol. 2018;18:75.
587 12. Donoghue M, Bell C, Li J. Phylogenetic Patterns in Northern Hemisphere Plant
588 Geography. Int J Plant Sci. 2001;162:41-52.
589 13. Milne RI, Abbott RJ. The origin and evolution of tertiary relict floras. Advances
590 in Botanical Research. vol. 38. Academic Press; 2002.
591 14. Ye J, Zhang Y, Wang X. Phylogeographic breaks and the mechanisms of their
592 formation in the Sino-Japanese floristic region. Chinese Journal of Plant Ecology.
593 2017;41:1003-1019.
594 15. Bruch A, Utescher T, Mosbrugger V, Gabrielyan I, Ivanov D. Late Miocene
595 climate in the circum-Alpine realm-a quantitative analysis of terrestrial palaeofloras.
596 Palaeogeography, Palaeoclimatology, Palaeoecology. 2006;238:270-280.
597 16. Yang J, Vázquez L, Feng L, Liu Z, Zhao G. Climatic and Soil Factors Shape the
598 Demographical History and Genetic Diversity of a Deciduous Oak (Quercus
599 liaotungensis) in Northern China. Frontiers in Plant Science. 2018;9.
600 17. Kohler T, Giger M, Hurni H, Ott C, Wiesmann U, von Dach SW, Maselli D.
601 Mountains and climate change: a global concern. Mountain Research and
602 Development. 2010;30:53-56.
603 18. Tian S, Kou Y, Zhang Z, Yuan L, Li D, López-Pujol J, Fan D, Zhang Z.
604 Phylogeography of Eomecon chionantha in subtropical China: the dual roles of the
605 Nanling Mountains as a glacial refugium and a dispersal corridor. BMC evolutionary
606 biology. 2018;18:20.

19. Swenson NG, Howard DJ. Clustering of contact zones, hybrid zones, and phylogeographic breaks in North America. The American Naturalist. 2005;166:581-591.

20. Chung MY, López-Pujol J, Chung MG. The role of the Baekdudaegan (Korean Peninsula) as a major glacial refugium for plant species: A priority for conservation. Biological Conservation. 2017;206:236-248.

21. Cun YZ, Wang XQ. Plant recolonization in the Himalaya from the southeastern Qinghai-Tibetan Plateau: Geographical isolation contributed to high population differentiation. Molecular Phylogenetics and Evolution. 2010; 56:972-982.

22. Pyron RA, Burbrink FT. Hard and soft allopatry: physically and ecologically mediated modes of geographic speciation. Journal of Biogeography. 2010; 37:2005-2015.

23. Wang P, Zhang X, Tang N, Liu J, Xu L, Wang K. Phylogeography of *Libanotis buchtormensis* (Umbelliferae) in disjunct populations along the deserts in Northwest China. PloS one. 2016;11:e0159790.

24. Zhang YH, Wang IJ, Comes HP, Peng H, Qiu YX. Contributions of historical and contemporary geographic and environmental factors to phylogeographic structure in a Tertiary relict species, Emmenopterys henryi (Rubiaceae). Scientific reports. 2016; 6:24041.

25. Fu PC, Ya HY, Liu QW, Cai HM, Chen SL. Out of Refugia: Population Genetic Structure and Evolutionary History of the Alpine Medicinal Plant *Gentiana lawrencei* var. *farreri* (Gentianaceae). Frontiers in Genetics. 2018;9.

629 26. Liu H, Gao Q, Zhang F, Khan G, Chen S. Westwards and northwards dispersal
630 of Triosteum himalayanum (Caprifoliaceae) from the Hengduan Mountains region
631 based on chloroplast DNA phylogeography. PeerJ. 2018;6:e4748.
632 27. Gao Q, Li X, Yang X. Responses of vegetation and primary production in
633 north-south transect of eastern China in global change under land use constraint. Acta
634 Botanica Sinica. 2003;45:1274-84.
635 28. Jiang Y, Kang M, Zhu Y, Xu G. Plant biodiversity patterns on Helan Mountain,
636 China. Acta Oecologica. 2007;32:125-133.
637 29. López-Pujol J, Zhang FM, Sun HQ, Ying TS, Ge S. Centres of plant endemism
638 in China: places for survival or for speciation? Journal of Biogeography.
639 2011;38:1267-1280.
640 30. Feliner GN. Patterns and processes in plant phylogeography in the
641 Mediterranean Basin. A review. Perspectives in Plant Ecology, Evolution and
642 Systematics. 2014;16:265-278.
643 31. Tian B, Liu R, Wang L, Qiu Q, Chen K, Liu J. Phylogeographic analyses
644 suggest that a deciduous species (Ostryopsis davidiana Decne., Betulaceae) survived
645 in northern China during the Last Glacial Maximum. Journal of Biogeography. 2009;
646 36:2148-2155.
647 32. Liu HZ, Harada K. Geographic distribution and origin of the chloroplast T/C -
648 type in Quercus mongolica var. crispula in northeastern Japan. Plant species biology.
649 2014;29:207-211.
650 33. Stewart JR, Lister AM, Barnes I, Dalén L. Refugia revisited: individualistic

651 responses of species in space and time. Proceedings of the Royal Society B:
652 Biological Sciences. 2010;277:661-671.
653 34. Favre A, Päckert M, Pauls SU, Jähnig SC, Uhl D, Michalak I, Muellner‐Riehl
654 AN: The role of the uplift of the Qinghai-Tibetan Plateau for the evolution of Tibetan
655 biotas. Biological Reviews. 2015;90:236-253.
656 35. Chen H, Huang C. Fagaceae, Ulmaceae and Rhoipteleaceae. Flora of China.
657 1998;22:334-350.
658 36. Ru W, Zhang G, Bi R, Zhang F, Zhang J. Population structure and pattern of
659 endangered *Ulmus lamellosa* in Shanxi. 2007;13:14-17.
660 37. Gu S, Du G. Relationships between geographica1 distribution of *Ulmus*
661 *lamellosa* and climate in China. J Inner Mongolia For Coll. 1994;16:28-32.
662 38. Wang X, Zhang Q, Bi R,Bai Y. Intra-and inter-specific competition of rare and
663 endangered plant *Ulmus lamellosa* in Shanxi Province of China. Chinese Journal of
664 Ecology. 2013;32:1756-61.
665 39. Bi Rc, Zhang J, Su Jx. Ecological characters of rare-endangered plant Ulmus
666 lamellose in Shanxi Province. Journal of Plant Resources and Environment.
667 2002;11:45-50.
668 40. Liu L, Chen W, Yan DT, Li J, Liu L, Wang YL. Molecular phylogeography and
669 paleodistribution modeling of the boreal tree species *Ulmus lamellosa* (T. Wang et SL
670 Chang)(Ulmaceae) in China. Tree genetics & genomes. 2017;13:11.
671 41. Wen Z, Xu Z, Zhang H, Feng Y. Chloroplast phylogeography of a desert shrub,
672 Calligonum calliphysa (Calligonum, Polygonaceae), in arid Northwest China.

673 Biochemical systematics and ecology. 2015;60:56-62.
674 42. Wang MN, Duan L, Qiao Q, Wang ZF, Zimmer EA, Li ZC, Chen HF:
675 Phylogeography and conservation genetics of the rare and relict Bretschneidera
676 sinensis (Akaniaceae). PloS one. 2018;13:e0189034.
677 43. Wang Y, Liu K, Bi D, Zhou S, Shao J. Molecular phylogeography of East Asian
678 Boea clarkeana (Gesneriaceae) in relation to habitat restriction. PloS one.
679 2018;13:e0199780.
680 44. Ye JW, Guo XD, Wang S, Bai WN, Bao L, Wang HF, Ge JP. Molecular evidence
681 reveals a closer relationship between Japanese and mainland subtropical specimens of
682 a widespread tree species, Acer mono. Biochemical Systematics and Ecology.
683 2015;60.
684 45. Zhou Z, Miwa M, Hogetsu T. Analysis of genetic structure of a Suillus grevillei
685 population in a Larix kaempferi stand by polymorphism of inter-simple sequence
686 repeat (ISSR). The New Phytologist. 1999;144:55-63.
687 46. Stefenon VM, Nodari R, Guerra MP. The genetics and conservation of Araucaria
688 angustifolia: III. DNA extraction protocol and informative capacity of RAPD markers
689 for analysis of genetic diversity in natural populations. Biotemas. 2004;17:47-63.
690 47. White TJ, Bruns T, Lee S, Taylor J. Amplification and direct sequencing of
691 fungal ribosomal RNA genes for phylogenetics. PCR protocols: a guide to methods
692 and applications. 1990;18:315-322.
693 48. Strand A, Leebens-Mack J, Milligan B: Nuclear DNA-based markers for plant
694 evolutionary biology. Molecular Ecology. 1997;6:113-8.

695 49. Larkin MA, Blackshields G, Brown N, Chenna R, McGettigan PA, McWilliam
696 H, Valentin F, Wallace IM, Wilm A, Lopez R. Clustal W and Clustal X version 2.0.
697 bioinformatics. 2007;23:2947-8.
698 50. Simmons MP, Ochoterena H. Gaps as characters in sequence-based phylogenetic
699 analyses. Systematic biology. 2000;49:369-381.
700 51. Chen SC, Zhang L, Zeng J, Shi F, Yang H, Mao YR, Fu CX. Geographic
701 variation of chloroplast DNA in *Platycarya strobilacea* (Juglandaceae). Journal of
702 Systematics and Evolution. 2012;50:374-385.
703 52. Rozas J, Sánchez-DelBarrio JC, Messeguer X, Rozas R. DnaSP, DNA
704 polymorphism analyses by the coalescent and other methods. Bioinformatics.
705 2003;19:2496-7.
706 53. Librado P, Rozas J. DnaSP v5: a software for comprehensive analysis of DNA
707 polymorphism data. Bioinformatics. 2009;25:1451-2.
708 54. Excoffier L, Lischer HE. Arlequin suite ver 3.5: a new series of programs to
709 perform population genetics analyses under Linux and Windows. Molecular ecology
710 resources. 2010;10:564-7.
711 55. Rozas J, Ferrer-Mata A, Sánchez-DelBarrio JC, Guirao-Rico S, Librado P,
712 Ramos-Onsins SE, Sánchez-Gracia A. DnaSP 6: DNA sequence polymorphism
713 analysis of large data sets. Molecular biology and evolution. 2017;34:3299-3302.
714 56. Pons O, Petit R. Measwring and testing genetic differentiation with ordered
715 versus unordered alleles. Genetics. 1996;144:1237-1245.
716 57. Petit J, Duminil J, Fineschi S, Hampe A, Salvini D, Ven-dramin G. Comparative

717 organization of chloroplast, mitochondrial and nuclear diversity in plant populations.
718 Mo-lecular Ecology. 2005;14:689-701.
719 58. Lin N, Deng T, Moore MJ, Sun Y, Huang X, Sun W, Luo D, Wang H, Zhang J,
720 Sun H. Phylogeography of *Parasyncalathium souliei* (Asteraceae) and its Potential
721 Application in Delimiting Phylogeoregions in the Qinghai-Tibet Plateau
722 (QTP)-Hengduan Mountains (HDM) Hotspot. Frontiers in Genetics. 2018;9.
723 59. Bandelt H-J, Forster P, Röhl A. Median-joining networks for inferring
724 intraspecific phylogenies. Molecular biology and evolution. 1999;16:37-48.
725 60. Pritchard JK, Stephens M, Donnelly P. Inference of population structure using
726 multilocus genotype data. Genetics. 2000;155:945-959.
727 61. Falush D, Stephens M, Pritchard JK. Inference of population structure using
728 multilocus genotype data: dominant markers and null alleles. Molecular ecology notes.
729 2007;7:574-8.
730 62. Pritcharda JK, Wena X, Falushb D. Documentation for structure software:
731 Version 2.3. 2009.
732 63. Evanno G, Regnaut S, Goudet J. Detecting the number of clusters of individuals
733 using the software STRUCTURE: a simulation study. Molecular ecology.
734 2005;14:2611-20.
735 64. Excoffier L, Smouse PE, Quattro JM. Analysis of molecular variance inferred
736 from metric distances among DNA haplotypes: application to human mitochondrial
737 DNA restriction data. Genetics. 1992;131:479-491.
738 65. Excoffier L, Laval G, Schneider S. Arlequin (version 3.0): an integrated

739 software package for population genetics data analysis. Evolutionary bioinformatics.
740 2005;1:117693430500100003.
741 66. Rogers AR, Harpending H. Population growth makes waves in the distribution
742 of pairwise genetic differences. Molecular biology and evolution. 1992;9:552-569.
743 67. Harpending H. Signature of ancient population growth in a low-resolution
744 mitochondrial DNA mismatch distribution. Human biology. 1994;591-600.
745 68. Tajima F. The effect of change in population size on DNA polymorphism.
746 Genetics. 1989;123:597-601.
747 69. Fu YX. Statistical tests of neutrality of mutations against population growth,
748 hitchhiking and background selection. Genetics. 1997;147:915-925.
749 70. Rogers AR. Genetic evidence for a Pleistocene population explosion. Evolution.
750 1995;49:608-615.
751 71. Graur D, Li WH. Fundamentals of molecular evolution. Dynamics. 2000;20:38.
752 72. Qiu YX, Guan BC, Fu CX, Comes HP. Did glacials and/or interglacials promote
753 allopatric incipient speciation in East Asian temperate plants? Phylogeographic and
754 coalescent analyses on refugial isolation and divergence in Dysosmaversipellis.
755 Molecular Phylogenetics and Evolution. 2009;51:281-293.
756 73. Chen J, Huang R, Li G, Hou X, Qin Y, Zhang G, Liu H, Vishnyakova O. Radial
757 growth characteristics of *Ulmus pumila* var. *sabulosa* in Otindag sand land, northern
758 China. Journal of Beijing Forestry University. 2014;36:41-7.
759 74. Feng L, Zheng QJ, Qian ZQ, Yang J, Zhang YP, Li ZH, Zhao GF. Genetic
760 Structure and Evolutionary History of Three Alpine Sclerophyllous Oaks in East

761 Himalaya-Hengduan Mountains and Adjacent Regions. Frontiers in Plant Science.
762 2016;7.
763 75. Drummond AJ, Rambaut A. BEAST: Bayesian evolutionary analysis by
764 sampling trees. BMC evolutionary biology. 2007;7:214.
765 76. Bertorelle G, Benazzo A, Mona S. ABC as a flexible framework to estimate
766 demography over space and time: some cons, many pros. Molecular ecology.
767 2010;19:2609-25.
768 77. Csilléry K, Blum MG, Gaggiotti OE, François O. Approximate Bayesian
769 computation (ABC) in practice. Trends in ecology & evolution. 2010;25:410-8.
770 78. Drummond AJ, Suchard MA, Xie D, Rambaut A. Bayesian phylogenetics with
771 BEAUti and the BEAST 1.7. Molecular biology and evolution. 2012;29:1969-1973.
772 79. Bouckaert R, Heled J, Kühnert D, Vaughan T, Wu CH, Xie D, Suchard MA,
773 Rambaut A, Drummond AJ. BEAST 2: a software platform for Bayesian evolutionary
774 analysis. PLoS computational biology. 2014;10:e1003537.
775 80. Posada D. jModelTest: phylogenetic model averaging. Molecular biology and
776 evolution. 2008;25:1253-6.
777 81. Darriba D, Taboada GL, Doallo R, Posada D. jModelTest 2: more models, new
778 heuristics and parallel computing. Nature methods. 2012;9:772.
779 82. Richardson JE, Pennington R, Pennington T, Hollingsworth P, Richardson JE,
780 Pennington RT, Pennington TD, Hollingsworth PM. Rapid diversification of a
781 species-rich genus of neotropical rain forest trees. Science. 2001;293:2242-5.
782 83. Rambaut A, Drummond A, Xie D, Baele G, Suchard M. Tracer v1. 7, 2018.

783 2018.

784 84. Rambaut A. FigTree v1. 4.3 software. Institute of Evolutionary Biology,

785 University of Edinburgh. 2016.

786 85. Beaumont MA, Zhang W, Balding DJ. Approximate Bayesian computation in

787 population genetics. Genetics. 2002;162:2025-35.

788 86. Cornuet JM, Santos F, Beaumont MA, Robert CP, Marin JM, Balding DJ,

789 Guillemaud T, Estoup A. Inferring population history with DIY ABC: a user-friendly

790 approach to approximate Bayesian computation. Bioinformatics. 2008;24:2713-2719.

791 87. Cornuet JM, Ravigné V, Estoup A. Inference on population history and model

792 checking using DNA sequence and microsatellite data with the software DIYABC (v1.

793 0). Bmc Bioinformatics. 2010;11:401.

794 88. Cornuet JM, Pudlo P, Veyssier J, Dehne-Garcia A, Gautier M, Leblois R, Marin

795 JM, Estoup A. DIYABC v2. 0: a software to make approximate Bayesian computation

796 inferences about population history using single nucleotide polymorphism, DNA

797 sequence and microsatellite data. Bioinformatics. 2014;30:1187-9.

798 89. Beerli P. Comparison of Bayesian and maximum-likelihood inference of

799 population genetic parameters. Bioinformatics. 2005;22:341-5.

800 90. Johnson JA, Tingay RE, Culver M, Hailer F, Clarke ML, Mindell DP. Long-term

801 survival despite low genetic diversity in the critically endangered Madagascar

802 fish-eagle. Molecular Ecology. 2009;18:54-63.

803 91. Beerli P, Palczewski M. Unified framework to evaluate panmixia and migration

804 direction among multiple sampling locations. Genetics. 2010;185:313-326.

805 92. Chiucchi JE, Gibbs H. Similarity of contemporary and historical gene flow
806 among highly fragmented populations of an endangered rattlesnake. Molecular
807 Ecology. 2010;19:5345-58.
808 93. Beerli P, Felsenstein J. Maximum likelihood estimation of a migration matrix
809 and effective population sizes in n subpopulations by using a coalescent approach.
810 Proceedings of the National Academy of Sciences. 2001;98:4563-8.
811 94. Udupa S, Baum M. High mutation rate and mutational bias at $(TAA)_n$
812 microsatellite loci in chickpea (*Cicer arietinum* L.). Molecular Genetics and
813 Genomics. 2001;265:1097-1103.
814 95. Oksanen J, Blanchet F, Friendly M, Kindt R, Legendre P, McGlinn D, Minchin P,
815 O'Hara R, Simpson G, Solymos P. vegan: Community Ecology Package 2017. R
816 package version 2.4–4. 2017.
817 96. Li M, Xie D, Xie C, Deng Y, Zhong Y, Yu Y, He X. A Phytogeographic Divide
818 Along the 500 mm Isohyet in the Qinghai-Tibet Plateau: Insights From the
819 Phylogeographic Evidence of Chinese Alliums (Amaryllidaceae). Frontiers in Plant
820 Science. 2019;10.
821 97. Peakall R, Smouse PE. GENALEX 6: genetic analysis in Excel. Population
822 genetic software for teaching and research. Molecular ecology notes. 2006;6:288-295.
823 98. Rosenberg MS, Anderson CD. PASSaGE: pattern analysis, spatial statistics and
824 geographic exegesis. Version 2. Methods in Ecology and Evolution. 2011;2:229-232.
825 99. Wang IJ. Examining the full effects of landscape heterogeneity on spatial
826 genetic variation: a multiple matrix regression approach for quantifying geographic

827 and ecological isolation. Evolution. 2013;67:3403-11.

828 100. Adamack AT, Gruber B. PopGenReport: simplifying basic population genetic

829 analyses in R. Methods in Ecology and Evolution. 2014;5:384-7.

830 101. Sakaguchi S, Qiu YX, Liu YH, Qi X, Kim SH, Han J, Takeuchi Y, Worth J,

831 Yamasaki M, Sakurai S, et al. Climate oscillation during the Quaternary associated

832 with landscape heterogeneity promoted allopatric lineage divergence of a temperate

833 tree *Kalopanax septemlobus* (Araliaceae) in East Asia. Molecular ecology.

834 2012;21:3823-38.

835 102. Bai WN, Wang WT, Zhang DY. Contrasts between the phylogeographic

836 patterns of chloroplast and nuclear DNA highlight a role for pollen-mediated gene

837 flow in preventing population divergence in an East Asian temperate tree. Molecular

838 phylogenetics and evolution. 2014;81.

839 103. Wu H, Xu S, Cui S. The four new records of plant species in Shanxi Province.

840 Shanxi For Sci Tech. 1999;2:20-1.

841 104. Teixeira H, Rodríguez-Echeverría S, Nabais C. Genetic diversity and

842 differentiation of Juniperus thurifera in Spain and Morocco as determined by SSR.

843 PLoS One. 2014;9:e88996.

844 105. Hamrick JL, Godt MW. Effects of life history traits on genetic diversity in

845 plant species. Philosophical Transactions of the Royal Society of London Series B:

846 Biological Sciences. 1996;351:1291-8.

847 106. Bai QQ, Ren GD. Phylogeographic analysis of *Episyrphus balteatus* (Diptera:

848 Syrphidae) in Yanshan-Taihan. Chinese Journal of Ecology. 2018;37:157-163.

849 107. Ye Z, Zhu G, Damgaard J, Chen X, Chen P, Bu W. Phylogeography of a
850 semi-aquatic bug, *Microvelia horvathi* (Hemiptera: Veliidae): an evaluation of
851 historical, geographical and ecological factors. Scientific reports. 2016;6:21932.
852 108. Frankham R, Briscoe DA, Ballou JD. Introduction to conservation genetics:
853 Cambridge university;2002.
854 109. Stefenon VM, Klabunde G, Lemos RPM, Rogalski M, Nodari RO.
855 phylogeography of plastid DNA sequences suggests post-glacial southward
856 demographic expansion and the existence of several glacial refugia for Araucaria
857 angustifolia. Scientific reports. 2019;9:2752.
858 110. Yu G, Chen X, Ni J, Cheddadi R, Guiot J, Han H, Harrison SP, Huang C, Ke M,
859 Kong Z. Palaeovegetation of China: a pollen data-based synthesis for the
860 mid-Holocene and last glacial maximum. Journal of Biogeography. 2000;27:635-664.
861 111. Liu JQ, Sun YS, Ge XJ, Gao LM, Qiu YX. Phylogeographic studies of plants
862 in China: advances in the past and directions in the future. Journal of Systematics and
863 Evolution. 2012;50:267-275.
864 112. Liu Y, Dietrich CH, Wei C. Genetic divergence, population differentiation and
865 phylogeography of the cicada *Subpsaltria yangi* based on molecular and acoustic data:
866 an example of the early stage of speciation? BMC evolutionary biology. 2019;19:5.
867 113. Hewitt GM. Some genetic consequences of ice ages, and their role in
868 divergence and speciation. Biological journal of the Linnean Society.
869 1996;58:247-276.
870 114. Zachos JC, Mo P, Sloan LC, Thomas E, Billups K. Trends, Rhythms, and

871 Aberrations in Global Climate 65 Ma to Present. Science (New York, NY).
872 2001;292:686-693.
873 115. Sun X, Wang P. How old is the Asian monsoon system? Paleobotanical records
874 from China. Palaeogeography, Palaeoclimatology, Palaeoecology. 2005;222:181-222.
875 116. Heermance RV, Chen J, Burbank DW, Wang C. Chronology and tectonic
876 controls of Late Tertiary deposition in the southwestern Tian Shan foreland, NW
877 China. Basin Research. 2007;19:599-632.
878 117. Ge J, Dai Y, Zhang Z, Zhao D, Li Q, Zhang Y, Yi L, Wu H, Oldfield F, Guo Z.
879 Major changes in East Asian climate in the mid-Pliocene: Triggered by the uplift of
880 the Tibetan Plateau or global cooling? Journal of Asian Earth Sciences.
881 2013;69:48-59.
882 118. Wu C, Zhang X, Ma Y. The Taihang and Yan mountains rose mainly in
883 Quaternary. Norht China Earthquake Sciences. 1999;17:1-7.
884 119. Zhou S, Wang X, Wang J, Xu L. A preliminary study on timing of the oldest
885 Pleistocene glaciation in Qinghai-Xizang Plateau. Quaternary International.
886 2006;154–155:44-51.
887 120. Royden LH, Burchfiel BC, van der Hilst RD. The geological evolution of the
888 Tibetan Plateau. Science. 2008;321:1054-8.
889 121. Cao YN, Comes HP, Sakaguchi S, Chen LY, Qiu YX. Evolution of East
890 Asia's Arcto-Tertiary relict *Euptelea* (Eupteleaceae) shaped by Late Neogene
891 vicariance and Quaternary climate change. BMC Evol Biol. 2016;16:66.
892

893

894

895

896 **Tables**:

897 Table 1 The estimated diversity indexes of *Ulmus lamellosa* populations

Groups	Populations	Sample sizes	ITS sequences			*Aat* sequences		
			Number haplotypes	Haplotype diversity (H_d)	Nucleotide diversity (π)	Number haplotypes	Haplotype diversity (H_d)	Nucleotide diversity (π)
THM	CC	17	H1(13), H6(2), H15(2)	0.396	0.00066	H2(3), H3(1), H4(5), H10(2), H15(4), H16(2)	0.857	0.00720
	HS	22	H1(20), H2(2)	0.318	0.00052	H12(11), H13(5), H14(6)	0.833	0.00808
	QLY	22	H1(18), H6(2), H7(2)	0.517	0.00140	H3(11), H4(6), H8(5)	0.833	0.00638
	SGS	15	H1(10), H2(5)	0.667	0.00103	H14(5), H19(5), H20(5)	1	0.01190
	MS	11	H1(11)	0	0	H17(11)	0	0
	JM	18	H1(12), H2(6)	0.667	0.00103	H10(18)	0	0
	HHG	15	H1(12), H2(3)	0.439	0.00073	H15(5), H21(5), H22(5)	1	0.01190

	LS	14	H17(7), H18(7)	1	0.00309	H17(7), H23(7)	1	0.00513
	TGS	14	H1(10), H8(4)	0.5	0.00155	H4(9), H9(5)	0.667	0.00680
	HLS	18	H1(10), H3(2), H4(2), H5(4)	0.636	0.00127	H1(5), H2(2), H3(6), H4(5)	0.81	0.00705
				0.359	0.00084		0.926	0.01506
	SFS	20	H1(3), H8(2), H9(6), H10(7), H11(1), H12(1)	0.929	0.00839	H5(7), H6(7), H7(6)	0.8	0.00272
YSM	WLS	18	H1(12), H16(6)	0.667	0.00309	H5(18)	0	0
				0.909	0.00939		0.714	0.00219
	JFS	12	H1(11), H5(1)	0.275	0.00063	H4(8), H18(4),	0.667	0.00680
YM	MLG	16	H1(12), H13(3), H14(1)	0.423	0.00109	H11(5), H12(5), H13(6)	1	0.00510
				0.332	0.00079		0.933	0.00969
Total		232	18	0.424	0.00212	23	0.941	0.01601

898 Table 2 Analysis of molecular variance (AMOVA) based on ITS and *Aat* sequences for *Ulmus lamellosa*

Source of variance	d.f[a]	SS[b]	Variance components	Variance percentage (%)	Fixation indices ($P < 0.01$)
Among groups	2	26.684(39.147)	0.346(0.981)	40.60(41.96)	Φ_{CT} =0.406(Φ_{CT} =0.562)
Among populations within groups	11	15.287(83.274)	0.123(1.621)	14.43(25.40)	Φ_{SC} =0.243(Φ_{SC} =0.254)
Within populations	231	43.631(55.476)	0.383(1.261)	44.97(32.64)	Φ_{ST} =0.550(Φ_{ST} =0.674)

899 In parentheses--the results based on *Aat* data; Outside parentheses--the results based on ITS data.

900

901

Table 3 Neutrality tests (Tajima's *D* and Fu's F_S tests) and mismatch distribution analyses

Groups	ITS sequences				*Aat* sequences			
	Mismatch distribution analyses				Mismatch distribution analyses			
	Tajima's *D*	Fu's F_S	SSD (P_{SSD})	HRag (P_{HRag})	Tajima's *D*	Fu's F_S	SSD (P_{SSD})	HRag (P_{HRag})
THM	-2.211	-10.024	0.00028	0.19533	-0.296	-2.907	0.010	0.013
	(*P*<0.01)	(*P*<0.01)	(0.800)	(0.630)	(*P*>0.1)	(*P*>0.1)	(0.250)	(0.880)
YSM	0.498	0.132	0.404	0.072	0.413	-0.071	0.048	0.286
	(*P*>0.1)	(*P*>0.1)	(0.350)	(0.420)	(*P*>0.1)	(*P*>0.1)	(0.170)	(0.240)
YM	-1.751	-0.999	0.009	0.234	1.394	-0.839	0.054	0.191
	(*P*>0.1)	(*P*>0.1)	(0.540)	(0.560)	(*P*>0.1)	(*P*>0.1)	(0.350)	(0.330)
Total	-0.605	-0.044	0.069	0.336	0.015	0.449	0.105	0.307
	(*P*>0.1)	(N.A.)	(0.274)	(0.393)	(*P*>0.1)	(N.A.)	(0.258)	(0.445)

Figture:

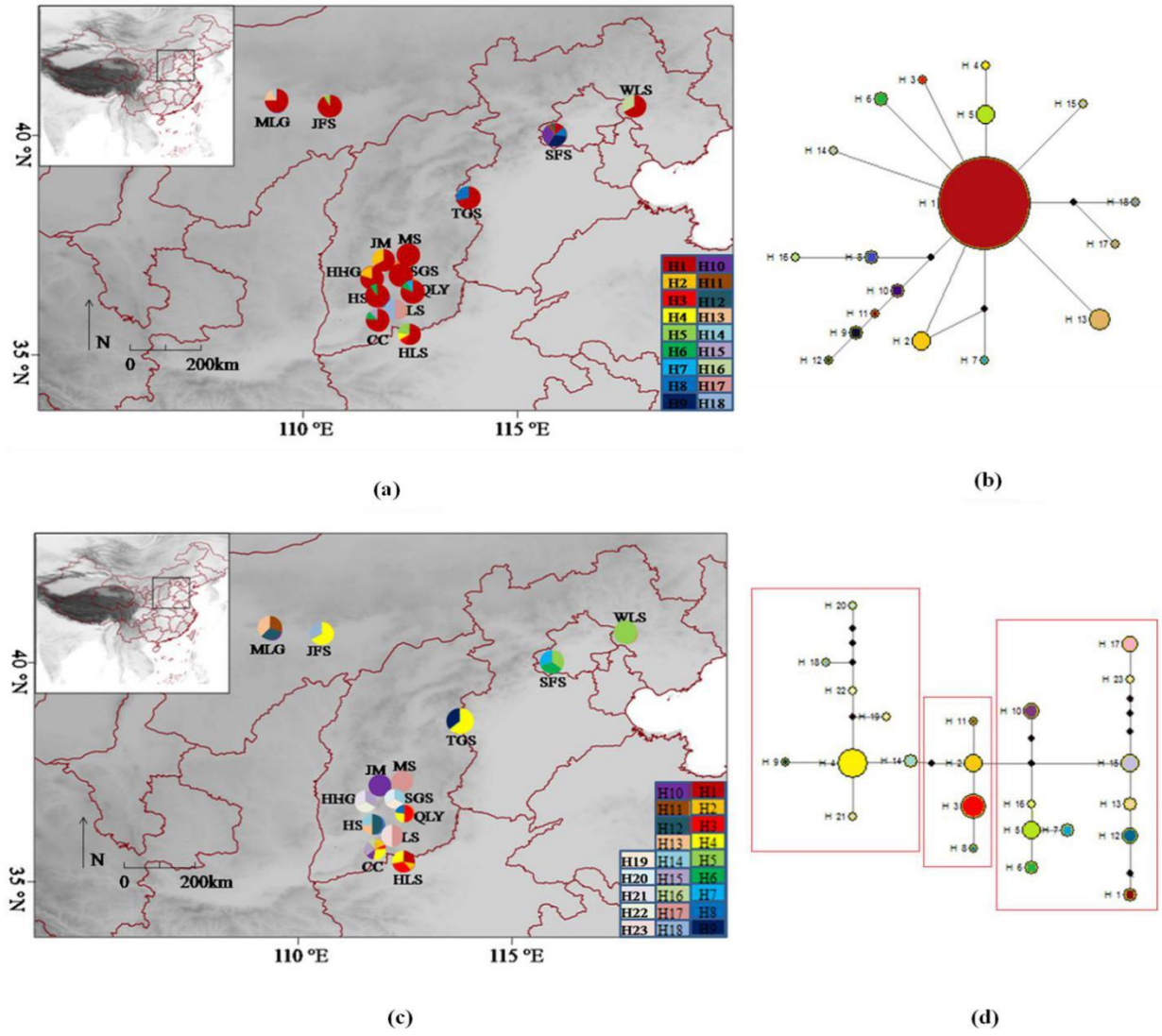
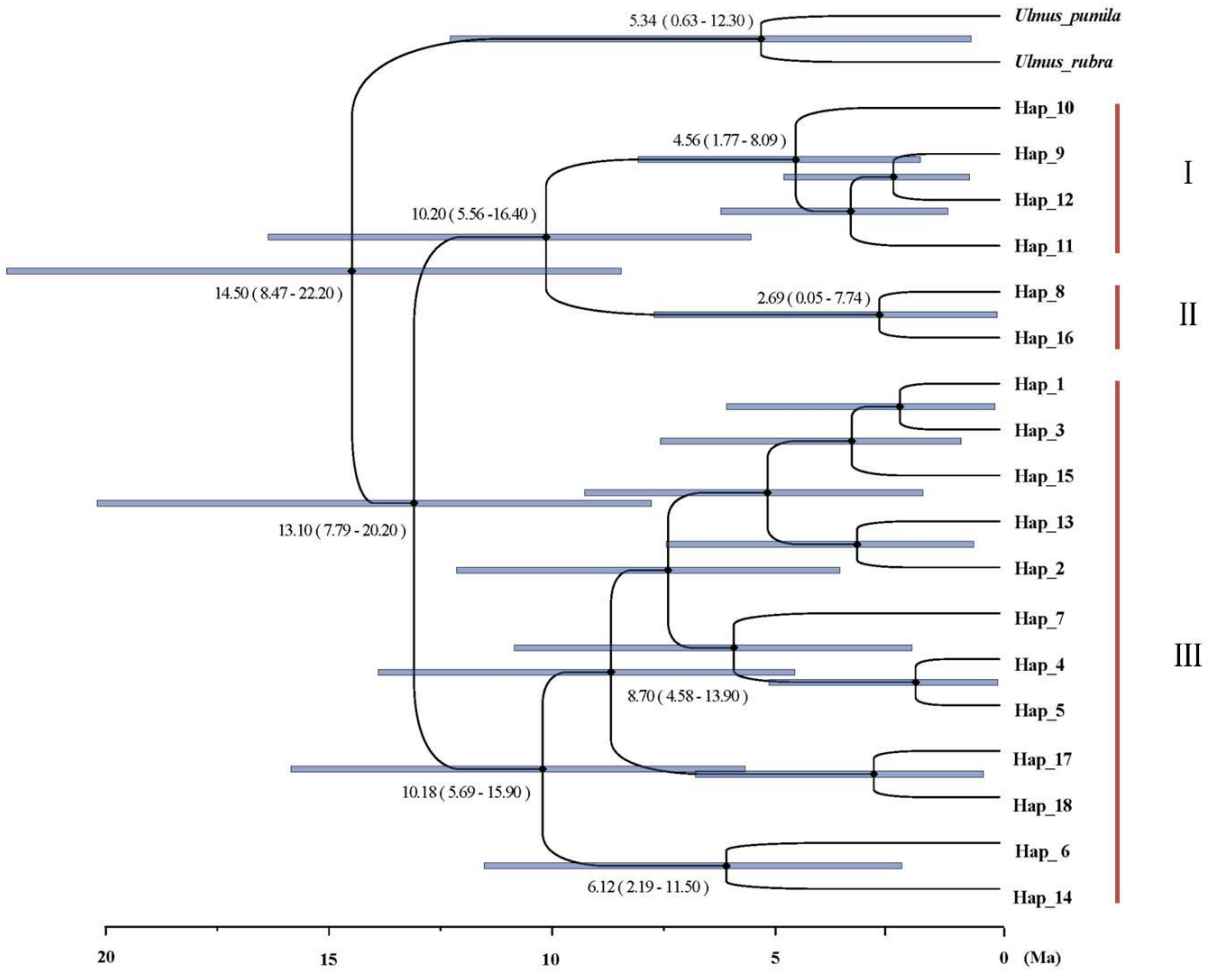


Fig. 1 Geographical distribution and the median-joining network of haplotypes. (a) Geographical distribution of 18 ITS haplotypes; (b) The median-joining network of 18 ITS haplotypes; (c) Geographical distribution of 23 *Aat* haplotypes; (d) The median-joining network of 23 *Aat* haplotypes



912

913 Fig. 2 The chronogram for *Ulmus lamellose* based on BEAST analysis of ITS sequences.

914 Positions of the fossil calibrations were indicated by a five-pointed star. Divergence times were

915 labeled on each node. Blue bars at nodes represented the 95% highest probability density (HPD)

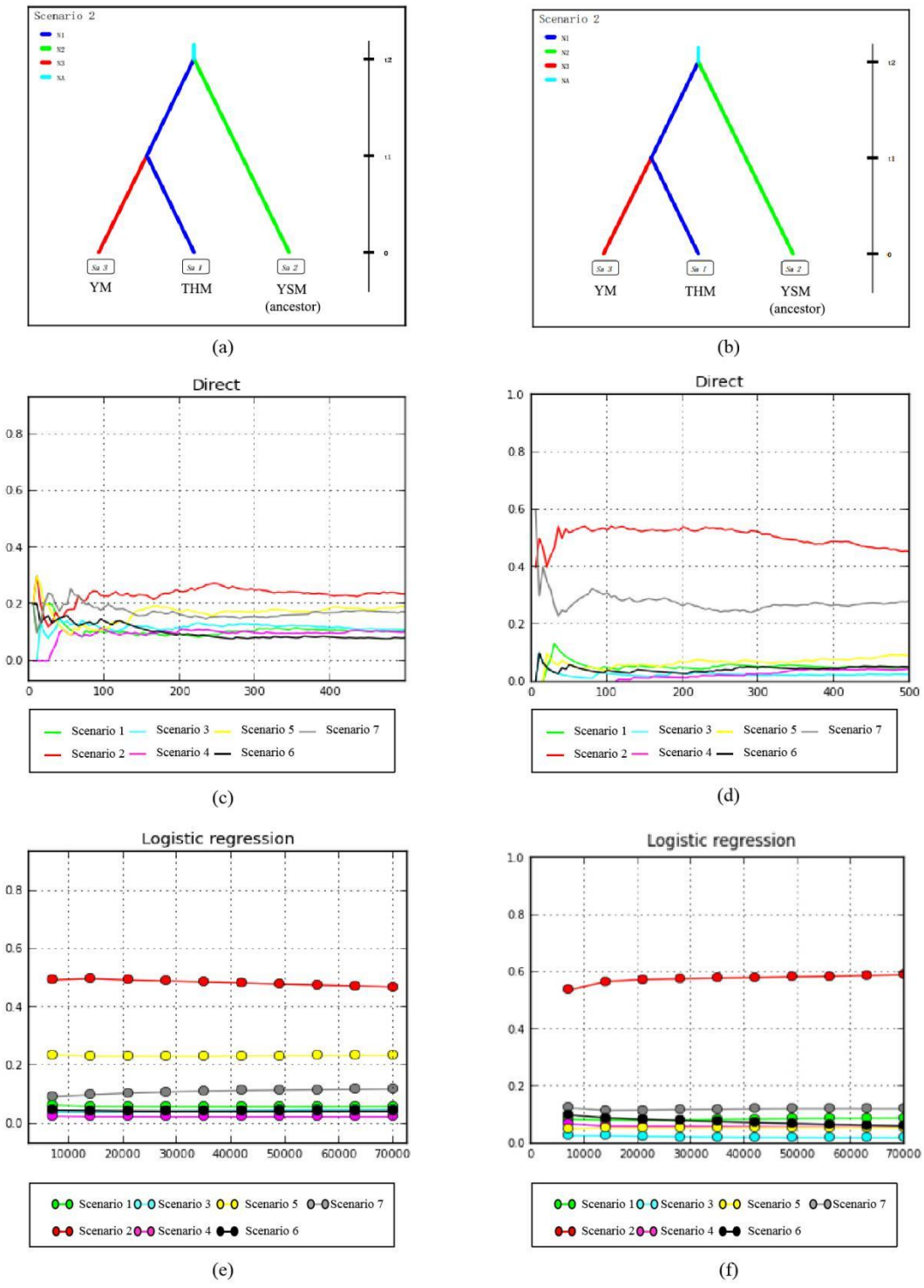
916 for the age of that node.

917

918

919

920



921

922 Fig. 3 The Approximate Bayesian Computation analysis of *Ulmus lamellose* assessed using

923 DIYABC software. Time in generations was t (t2⩾t1). (a) optimal model (ITS), (c)(e) posterior

probabilities (ITS), (b) optimal model (*Aat*), (d)(f) posterior probabilities (*Aat*).

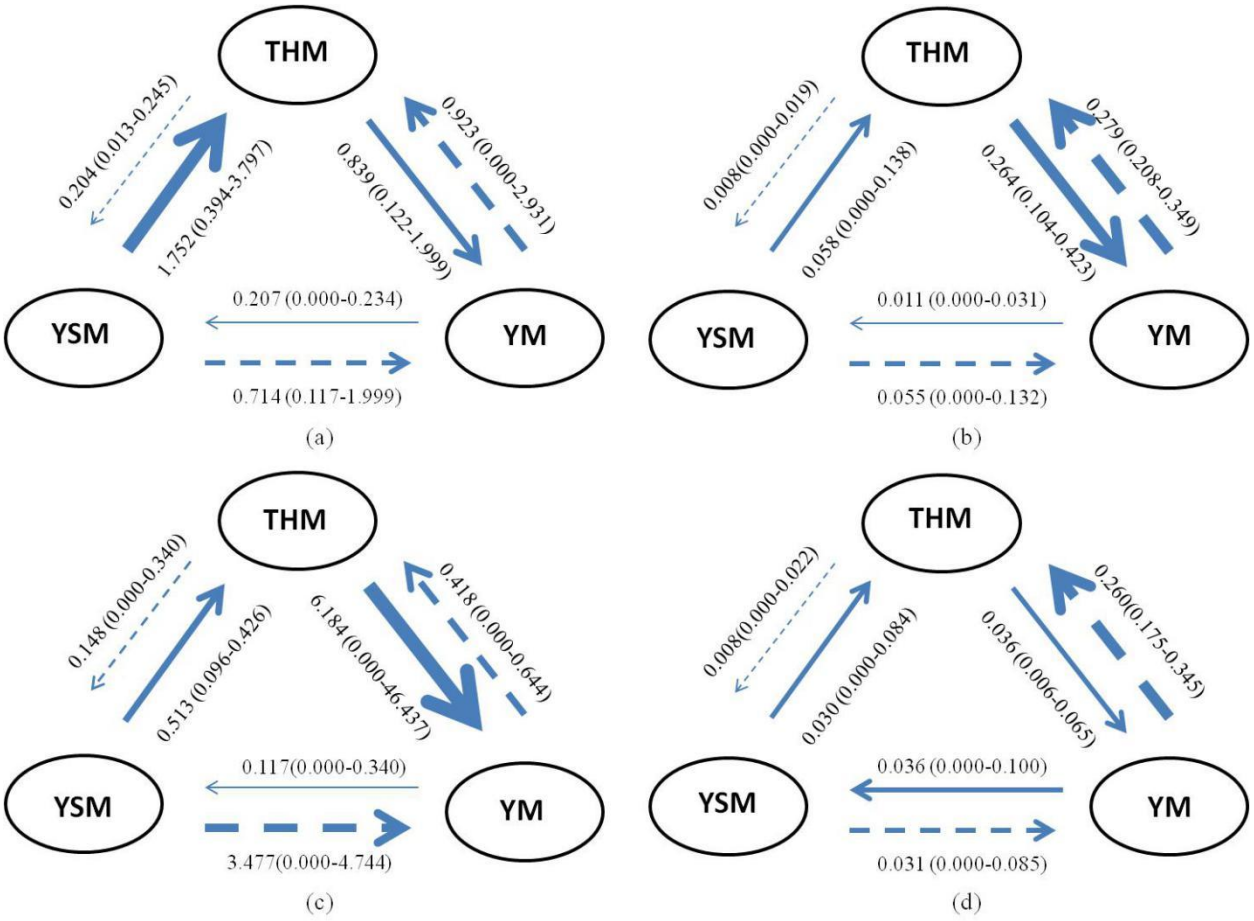


Fig. 4 Estimates of historical gene flow using MIGRATE, contemporary gene flow using BEYASASS and 95% confidence intervals (CI) (in parentheses) within (a) historical gene flow based on ITS, (b) contemporary gene flow based on ITS, (c) historical gene flow based on *Aat*, and (d) contemporary gene flow based on *Aat*.

Figures

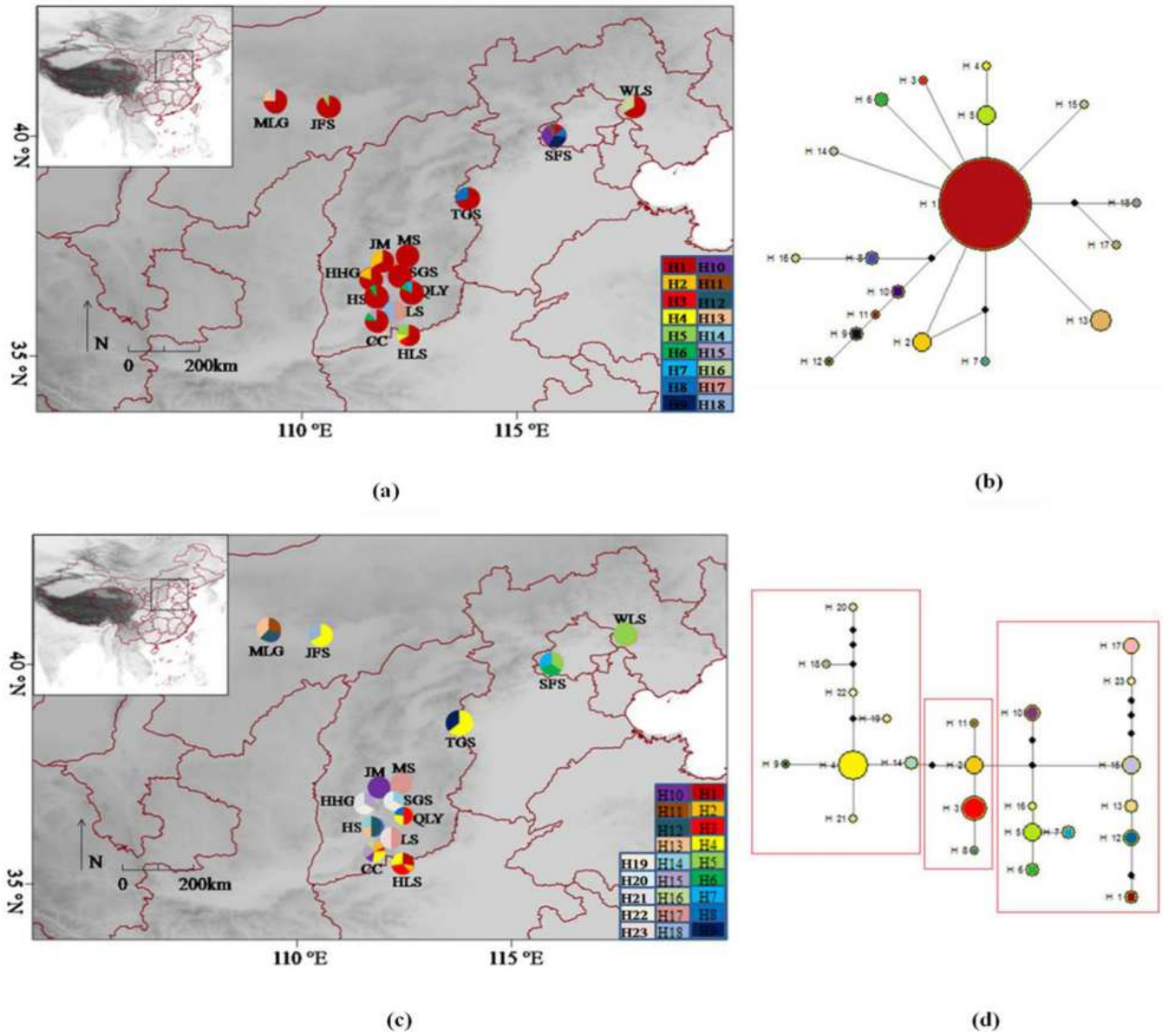


Figure 1

Geographical distribution and the median-joining network of haplotypes. (a) Geographical distribution of 18 ITS haplotypes; (b) The median-joining network of 18 ITS haplotypes; (c) Geographical distribution of 23 Aat haplotypes; (d) The median-joining network of 23 Aat haplotypes

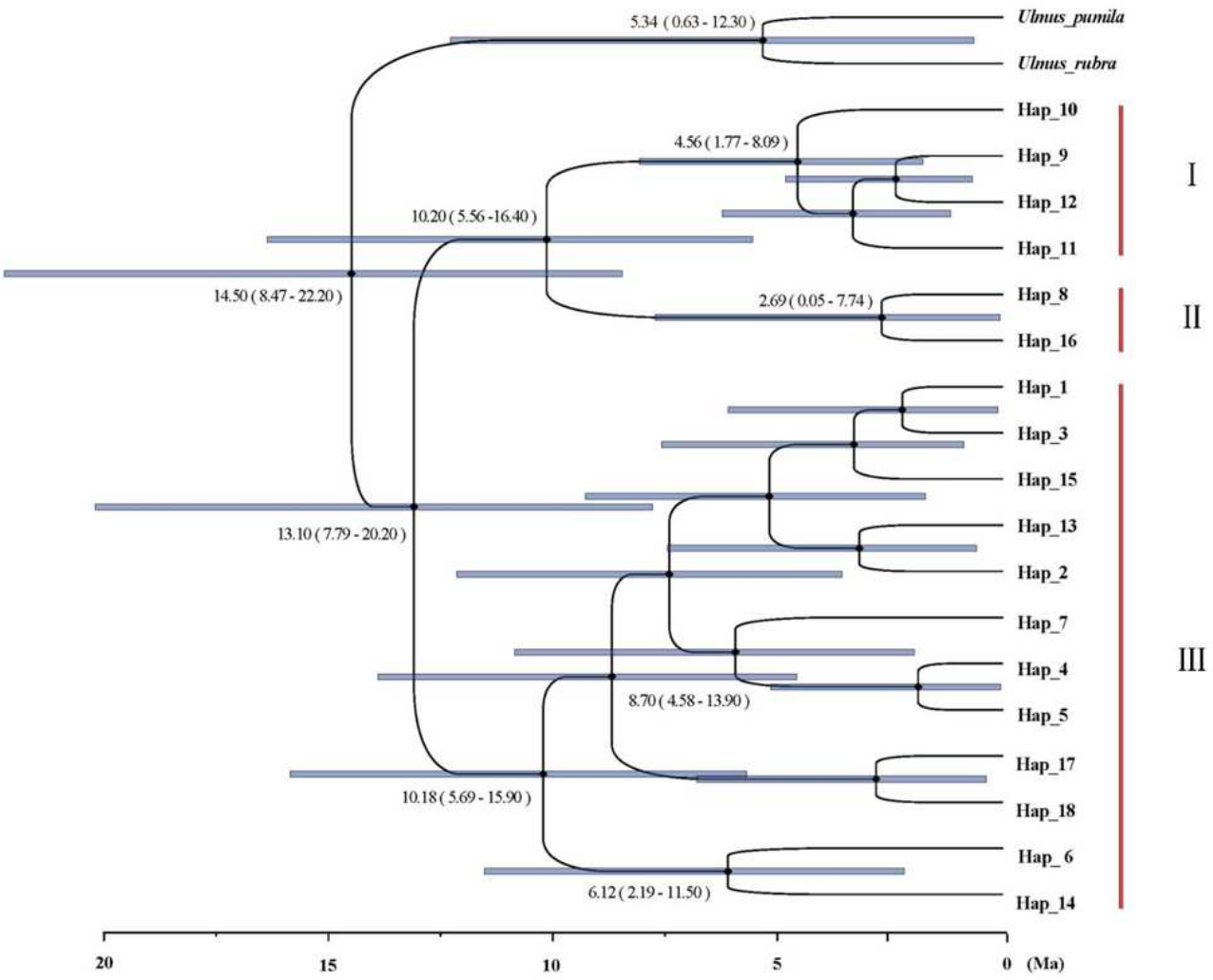


Figure 2

The chronogram for Ulmus lamellose based on BEAST analysis of ITS sequences. Positions of the fossil calibrations were indicated by a five-pointed star. Divergence times were labeled on each node. Blue bars at nodes represented the 95% highest probability density (HPD) for the age of that node.

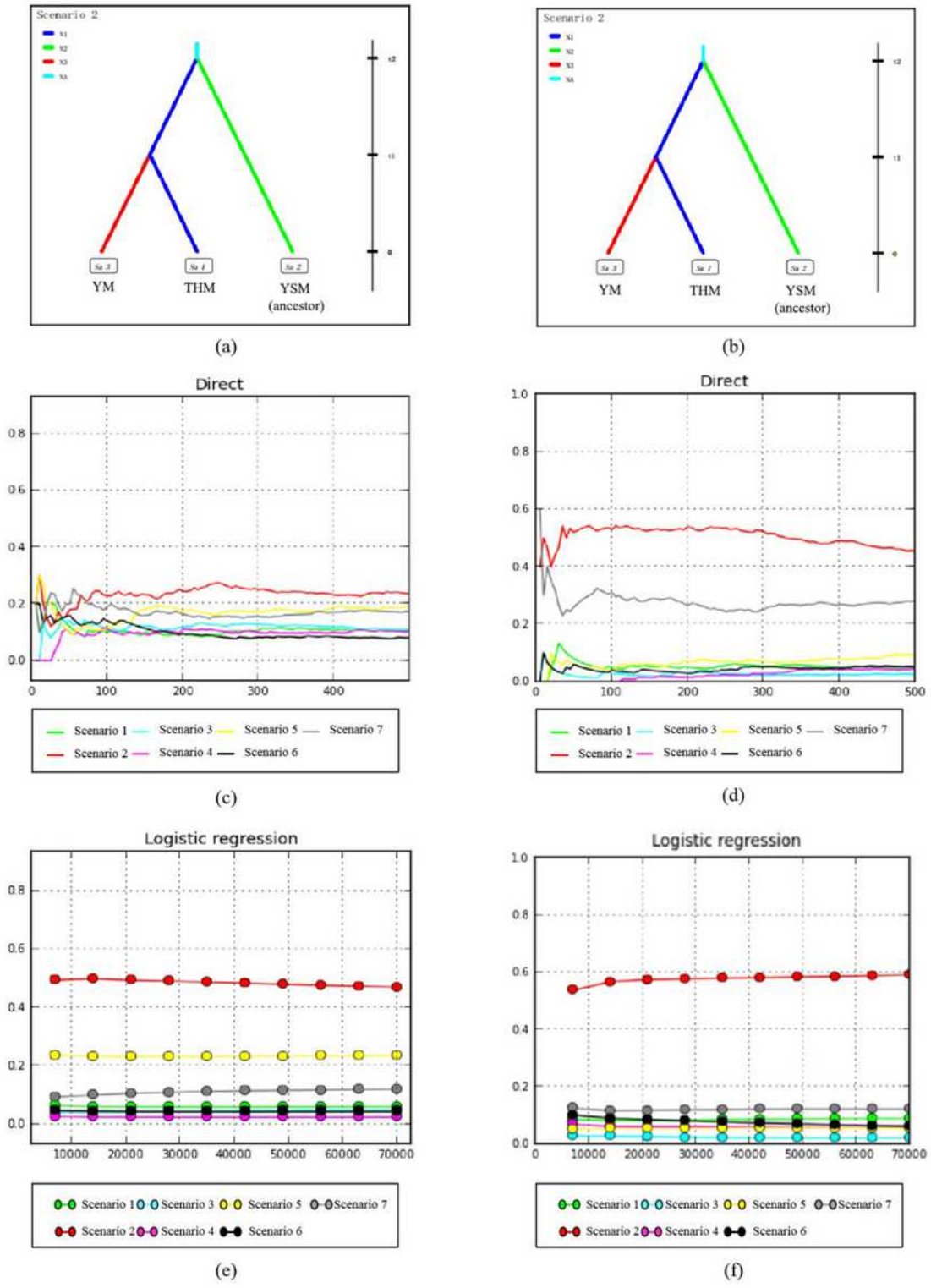


Figure 3

The Approximate Bayesian Computation analysis of Ulmus lamellose assessed using DIYABC software. Time in generations was t (t2≥t1). (a) optimal model (ITS), (c)(e) posterior probabilities (ITS), (b) optimal model (Aat), (d)(f) posterior probabilities (Aat).

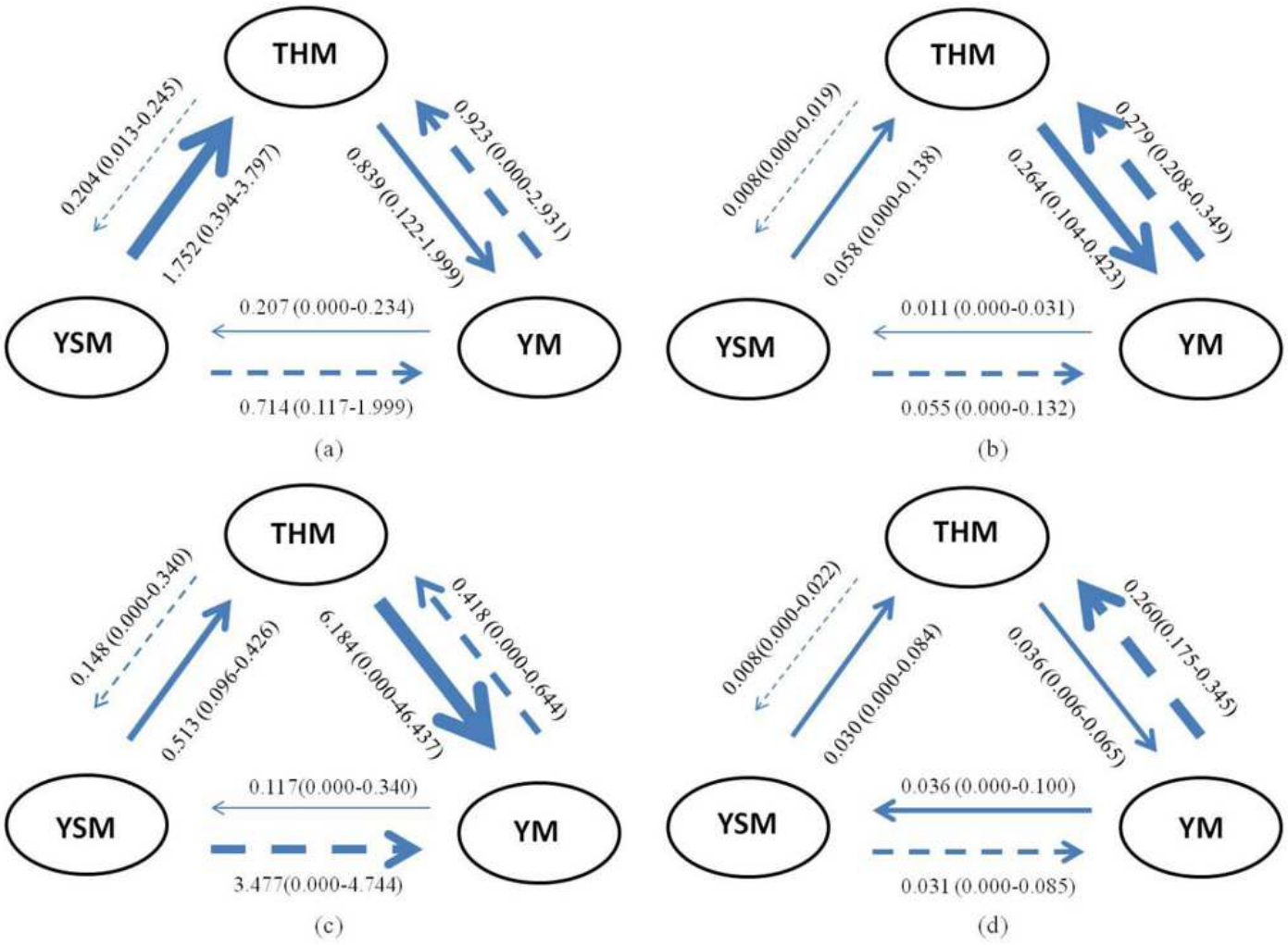


Figure 4

Estimates of historical gene flow using MIGRATE, contemporary gene flow using BEYASASS and 95% confidence intervals (CI) (in parentheses) within (a) historical gene flow based on ITS, (b) contemporary gene flow based on ITS, (c) historical gene flow based on Aat, and (d) contemporary gene flow based on Aat.

Supplementary Files

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

- Additionalfile1.xls
- Additionalfile3.doc
- Additionalfile2.doc
- Additionalfile4.doc