Research paper

Evolutionary history of a desert perennial Arnebia szechenyi (Boraginaceae): Intraspecific divergence, regional expansion and asymmetric gene flow

Meng-Jiao Fu a, Hai-Yang Wu a, Dong-Rui Jia b, **, Bin Tian a, c, *

a Key Laboratory for Forest Resources Conservation and Utilization in the Southwest Mountains of China, Ministry of Education, Southwest Forestry University, Kunming 650224, China
b School of Ecology and Environmental Science & Yunnan Key Laboratory for Plateau Mountain Ecology and Restoration of Degraded Environments, Yunnan University, Kunming 650091, China
c CAS Key Laboratory for Plant Diversity and Biogeography, Kunming Institute of Botany, Chinese Academy of Sciences, Kunming 650204, China

** Corresponding author.
* Corresponding author. Key Laboratory for Forest Resources Conservation and Utilization in the Southwest Mountains of China, Ministry of Education, Southwest Forestry University, Kunming 650224, China.

E-mail addresses: jia.dru@ynu.edu.cn (D.-R. Jia), tianbin@swfu.edu.cn (B. Tian).

1. Introduction

Over the past three decades, technological and conceptual advances have led to an explosive growth in plant phylogeographic studies (Beheregaray, 2008; Hickerson et al., 2010; Morris and Shaw, 2018). Numerous studies have shown that past geological events and climate fluctuations play a major role in shaping the distributions and genetic structures of extant organisms (Comes and Kadereit, 2003; Qiu et al., 2011; Liu et al., 2012; Tian et al., 2020). Although only a small fraction of the existing taxa have been investigated, phylogeography has revealed various patterns and mechanisms for diversification across a wide range of taxonomic groups (Zamudio et al., 2016).

Drylands, which occupy about 40% of the Earth’s land surface, are characterized by low and highly variable precipitation, intense solar radiation, extreme air temperatures and seasonally high potential evaporation (Reynolds et al., 2007; Maestre et al., 2012). These climatic characteristics, coupled with the relatively low soil fertility, impose important limitations on the biota of drylands, and make them highly sensitive and vulnerable to global climate change (Le Houerou, 1996; Körner, 2000; Sala et al., 2000; Reynolds et al., 2007; Maestre et al., 2012). Drylands encompass diverse ecosystems and host about 20% of the major centers of plant
Two congeneric species, *Arnebia* guttata Bge. and *A. szechenyi* Maxim., located in Zhangye, Gansu (38°42′35.99″N, 105°39′24″E) were also sampled as outgroups for phylogenetic analysis.

2.1. Study site and sampling

We collected leaves of *Arnebia szechenyi* from 16 natural populations covering the whole range of *A. szechenyi*. We sampled 6–12 individuals per population with sampled individuals located more than 100 m apart in a population. Fresh leaves without disease were collected and dried immediately using silica gel in the field. For each population, one sample was deposited as a voucher specimen for further reference in the Southwest Forestry University Herbarium (SWFC). The latitude, longitude and elevation of each specimen for further reference in the Southwest Forestry University field. For each population, one sample was deposited as a voucher specimen for further reference in the Southwest Forestry University Herbarium (SWFC). The latitude, longitude and elevation of each specimen for further reference in the Southwest Forestry University Herbarium (SWFC). The latitude, longitude and elevation of each specimen for further reference in the Southwest Forestry University Herbarium (SWFC). The latitude, longitude and elevation of each specimen for further reference in the Southwest Forestry University Herbarium (SWFC). The latitude, longitude and elevation of each specimen for further reference in the Southwest Forestry University Herbarium (SWFC). The latitude, longitude and elevation of each specimen for further reference in the Southwest Forestry University Herbarium (SWFC).

2.2. DNA extraction, amplification and sequencing

We used DNeasy™ Tissue Kit (Qiagen, Hilden, Germany) to isolate total genomic DNA from dried leaves. We amplified and sequenced six nuclear LCG loci (28490, 33938, 38060, 40789, 40947, 43173), the nuclear ribosomal internal transcribed spacer (nrITS), and three chloroplast DNA (cpDNA) intergenic spacers (psbA-trnH, psbB-psbF and trnL-trnF) for all individuals following previous studies (Rock et al., 1987; White et al., 1990; Taberlet et al., 1991; Sang et al., 1997; Fu et al., 2016; see Appendix S1 for details). MEGA v.7.0 (Kumar et al., 2016) was used to adjust and align generated sequences. DnaSP v.5.0 (Librado and Rozas 2009) was used to phase heterozygous sequences of the nuclear genes. The newly identified sequences were deposited in GenBank under the accession numbers MW229731-MW231889 and MW229777-MW229730. Sequences from the three cpDNA regions were concatenated into a matrix.

2.3. Nuclear DNA

To assess selective neutrality of loci from biparentally inherited nuclear data (nrITS and LCGs), we first calculated Tajima’s D (Tajima, 1989), Fu and Li’s D* (Fu and Li, 1993) and Fu and Li’s F* (Fu, 1997) in DnaSP v.5.0 (Librado and Rozas, 2009). Nucleotide
diversity ($\pi$) and allelic richness ($A_R$) for each population were calculated using SPADS v.1.0 (Dellicour and Mardulyn, 2014). Interpopulation differences in nucleotide diversity were then calculated in the same software. Pearson correlations between latitude/longitude and genetic diversity were performed in R 4.0.1 (R Core Team, 2013). Potential population structure was examined using STRUCTURE v.2.3.4 (Falush et al., 2007) with an admixture model and assuming allele frequencies to be correlated among populations ($K = 1$ to 10, with 1,000,000 Markov chain Monte Carlo (MCMC) steps and a burn-in of the first 10,000 steps). The most likely number of clusters was determined by combining LnP(D) and $\Delta K$. Additionally, we performed a distance-based clustering method through principal coordinate analysis (PCoA) based on pairwise Fst values among populations using SPADS.

To detect any potential demographic expansions within the identified genetic groups, temporal changes in the effective population size ($N_e$) were inferred with extended Bayesian skyline plots (EBSP) as implemented in BEAST v.2.4. The best-fit models of nucleotide substitution for each locus were determined using MEGA (Available at https://www.megasoftware.net). The substitution models implemented in the analysis were determined using the program Permut 1.0 with 1000 permutations in Arlequin v.3.5 (Excoffier and Lischer, 2010).

2.4 Chlroplast DNA

For each population, we calculated the genetic diversity (nucleotide diversity $\pi$ and haplotype diversity $H_d$) of maternally inherited cpDNA using DnaSP v.5.0 (Librado and Rozas, 2009). The total genetic diversity ($H_t$), within-population genetic diversity ($H_s$), and between-population genetic differentiation ($N_st$) were estimated using the program Permut 1.0 with 1000 permutation tests (Pons and Petit, 1996; available at http//www.pierroton.inra.fr/genetics/labo/Software/Permut/). The partitioning of genetic variation at different levels was estimated by a hierarchical analysis of molecular variance (AMOVA; Excoffier et al., 1992) using 1000 permutations in Arlequin v.3.5 (Excoffier and Lischer, 2010).

To examine the demographic expansions within the genetic groups identified, we conducted a mismatch distribution analysis (MDA; Rogers and Harpending, 1992) and an EBSP analysis as implemented in DnaSP v.5.0 (Librado and Rozas, 2009) and BEAST v.2.4, respectively. The EBSP analysis was run as described above. We also calculated Tajima’s $D$ (Tajima, 1989) and Fu & Li’s $D^*$ test (Fu and Li, 1993) to assess possible expansions. If Arnebia szechenyi genetic groups had undergone expansions, Tajima's $D$ and Fu & Li's $D^*$ test should generate large negative values due to excess rare mutations. We calculated significance of the tests with 1000 replicates in DnaSP v.5.0 (Librado and Rozas, 2009).

We analyzed phylogenetic relationships of the different cpDNA and estimated genetic divergence through Bayesian analyses using BEAST v.2.4 (Bouckaert et al., 2014). For most angiosperms, cpDNA substitution rates have been estimated to vary between 1 and $3 \times 10^{-3}$ substitutions per site per year (s/s/y) (Wolfe et al., 1987). In this study, we used a normal distribution prior with a mean cpDNA substitution rate of $2.0 \times 10^{-3}$ s/s/y and a SD of $6.08 \times 10^{-10}$ s/s/y (Jia et al., 2012; Wang et al., 2016). Best-fit models of nucleotide substitution for cpDNA matrices were determined to be HKY by the AIC using ModelTest v.2.1.10 (Darriba et al., 2012). A strict clock
model and Yule process with a piecewise linear and constant population size model were applied. The length of the MCMC algorithm was set to 10,000,000 steps. We sampled all parameters once every 1000 steps out of 10,000,000 MCMC steps, with the first 10% of samples discarded as burn-in. The convergence of chains to a stationary distribution (effective population sizes > 200) was examined in Trace v.1.7 (Rambaut et al., 2018). Trees were then compiled into a maximum clade credibility (MCC) tree using TreeAnnotator v.2.4 (Bouckaert et al., 2014) to display posterior probabilities (PP) and mean node ages for each node. The program FigTree v.1.7.1 (available at http://tree.bio.ed.ac.uk/software/figtree/) was used to visualize the MCC tree.

To detect genealogical relationships among sequences with shallow genetic divergences, we also constructed the most parsimonious cpDNA haplotype networks using Network v.5.0 (available at http://www.fluxus-engineering.com/sharenet.htm).

2.5. Ecological niche modelling

We used a maximum-entropy modelling technique to predict the potential distribution of Arnebia szechynyi during the Last Glacial Maximum (LGM) and at present as implemented in MAXEnt v.3.4.1 (Phillips et al., 2006). A total of 34 occurrences, including the 16 sampled locations, were obtained from the Chinese Virtual Herbarium (http://www.cvh.org.cn/), the Global Biodiversity Information Facility (http://data.gbif.org) and records reported in grassland community studies. To reduce the influence of sampling deviation on the accuracy of simulation results, we used a spatial filtering method to randomly reserve one of any two distribution coordinates with a distance less than 1 km, and a total of 34 coordinate points were reserved after filtering. Environmental variables, including 19 bioclimatic and one elevation variable, eight soil variables and six ultraviolet rays, were downloaded from the world biological climate database (https://www.worldclim.org/), HWSD database (http://www.fao.org/) and world soil GIUV database (http://www.ufz.de/giuv/). To avoid model overfitting caused by strong correlations between environmental variables, Pearson correlation coefficients were used to screen variables. When each pair of environmental variables was correlated (|r| > 0.75), the variable that contributed the most to the model was retained (Appendix S2). To better project the simulation results to the LGM period, only linear (L) and quadratic (Q) were selected for the modeling characteristic (AUC; Fawcett, 2006). Finally, Arcgis v.10.5 was used to reclassify the simulation results and map the suitability of species distribution in each period.

3. Results

3.1. Nuclear DNA

Measures of genetic variation varied across populations, with nucleotide diversity (π) ranging from 1.62 × 10⁻³ to 6.90 × 10⁻³, and allele richness (A̅k) varying from 1.97 to 4.26 (Table 1). The highest levels of genetic diversity were found in population YBL (A̅k = 4.26 and π = 6.90 × 10⁻³) (Table 1, Fig. 1). Neutrality tests detected significant negative D* values and a large negative D̄ value for locus 33938 (Appendix S3), indicating a departure from neutral expectations. We therefore excluded this locus from further analyses. Bayesian clustering analysis implemented in STRUCTURE clearly showed two distinct genetic groups (Fig. 2a and c; Appendix S4), corresponding to the two geographic regions separated by the Tengger Desert. The western genetic group was composed of five populations (SD, ZY, JC, YBL and NRG) located between the Badain Jaran Desert and the Tengger Desert, while the eastern genetic group consisted of the remaining eleven populations located southeast of the Tengger Desert and along the Helan Mountains. However, PCoA yielded a different pattern of genetic clustering, with the six northern populations from the western genetic group (populations YBL and NRG) and the eastern genetic group (populations DT020, W18, S1 and Q18) clustered into another genetic group (Fig. 2b). The genetic differentiation between the two genetic groups was supported by a large Fst value (0.71). Migrate analysis indicated that historical migration between eastern and western genetic groups was significantly asymmetrical. Unidirectional gene flow was considerably higher from eastern to western genetic groups (Nm = 32.90; 95% HPD: 26.64–34.41) than from western to eastern groups (Nm = 0.64; 95% HPD: 0.52–0.68). Nevertheless, we did not observe any significant correlation between genetic diversity (A̅k and π), latitude, or longitude (Appendix S5).

Extended Bayesian skyline plots (EBSP) (see models in Appendix S6) suggest that the eastern genetic group and all populations combined have undergone historical expansion episodes, and that the effective population size has increased over the last 0.4–0.5 Mya (Fig. 3a and c).

3.2. Chloroplast DNA

After combining the three cpDNA fragments, we identified a total of 42 chlorotypes (H1–H42; Table 1; see details in Appendix S7). Within the 16 populations, nucleotide diversity (π) ranged from 0 to 72.60 × 10⁻⁴ and haplotype diversity (Hd) varied from 0 to 0.95 (Table 1). The highest levels of haplotype diversity were found in populations NRG and JT (0.95), while population DT015 showed the highest nucleotide diversity (72.60 × 10⁻⁴). The total genetic diversity (Ht) (0.98 ± 0.02) throughout all populations was much higher than the average within-population diversity (Hs) (0.50 ± 0.09), indicating that the majority of the cpDNA diversity was distributed among populations. The difference between the coefficient of genetic differentiation for non-ordered alleles (Gst = 0.49) and that for ordered alleles (Nst = 0.48) was non-significant (P = 0.10), thus revealing the absence of phylogeographical structure in Arnebia szechynyi (Appendix S8). The hierarchical AMOVA attributed most of the total genetic variation (64%) within populations, while a relatively low percentage (7%) to differences among the two genetic groups (Appendix S9).

Neither the eastern nor western genetic groups show signs of unimodal mismatch distribution under models of population expansion, nor do all population combined (Appendix S10 and Appendix S11). However, the EBSP analysis inferred that the effective population underwent slight fluctuations in size around 0–0.05 Mya (Appendix S12).

All haplotypes from Arnebia szechynyi formed a monophyletic lineage with four major clades (C1–C4) (Fig. 4a and b). The haplotypes placed in clades C2 and C3 were exclusively distributed in the western genetic groups and the eastern genetic groups, respectively; the haplotypes place in C1 or C4 were mainly distributed in either the western (C1) or eastern (C4) genetic groups (Fig. 4c). The haplotype network was largely congruent with the phylogenetic tree, although it depicted relationships between haplotypes in more detail. The dating analysis estimated that A. szechynyi originated in the Late Miocene (10.69 Mya; 95% HPD: 8.32–13.36 Mya), and started to diversify in the early Pleistocene (2.66 Mya; 95% HPD: 1.79–3.64 Mya) (Fig. 4a).
3.3. Ecological niche modelling

Ecological niche modelling indicated that the current distribution of *Arnebia szechenyi* is a good fit to its potential distribution (AUC value = 0.97 ± 0.03). Our analysis also indicated that the potential range of *A. szechenyi* underwent a considerable contraction during the Last Glacial Maximum (LGM) (Fig. 5). However, compared with the current distribution, the most suitable habitat remained in the areas of the Hexi corridor, south and east of the Tengger Desert and Helan Mountains (Fig. 5).

4. Discussion

4.1. High level of genetic diversity

Our analyses based on both nuclear and chloroplast data revealed a considerably high level of genetic variation within *Arnebia szechenyi*. From the 181 individuals sampled, we recovered a total of 42 chlorotypes. The total genetic diversity (Ht = 0.98) across all populations was much higher than the average within-population diversity (H_s = 0.50) indicating that the majority of cpDNA diversity was distributed among populations. This pattern of genetic variation has also been observed in other herbaceous plants occurring in the drylands of Northwest China (e.g. *Lagochilus ilicifolius*: Meng and Zhang, 2011; *Allium mongolicum*: Zhang et al., 2017). This genetic variation is probably associated with the particular ecological conditions in the drylands and the long evolutionary histories of desert plants (Meng and Zhang, 2011). The biophysical characteristics and landscapes of drylands likely imposed important limitations on their biota and produced discrete vegetation patches (Maestre et al., 2012). As a consequence, isolated populations may have limited gene flow to maintain genetic similarity among populations. Additionally, the distyly and heteromorphic self-incompatibility of *A. szechenyi*, which is considered a mechanism that evolved to promote outcrossing and avoid selfing (Lloyd and Webb 1992; Barrett et al., 2000; Zhou and Wang, 2009), ensures higher levels of genetic diversity within populations, conducive to evolution and adaptation to the harsh environment (Zhang et al., 2014).

4.2. Intraspecific divergence and chloroplast-nuclear discordance

The chloroplast genome, which is maternally inherited and does not undergo recombination, has lower rates of gene flow and smaller effective population sizes than nuclear DNA (Du et al., 2009). These differences in inheritance underlie how the two genomes have responded to past demographic fluctuations and biogeographic events (Despres, 2019; Hinojosa et al., 2019). Our Bayesian clustering analysis based on nDNA showed a clear split between the western and the eastern genetic groups, which

---

*Fig. 2.* (a) Frequency distribution of the two genetic clusters as revealed in STRUCTURE analysis with K = 2 in the 16 *Arnebia szechenyi* populations. The eastern genetic group (EAST) which consists of eleven populations located southeast of the Tengger Desert and along the Helan Mountains; the western genetic group (WEST), which is composed of five populations (SD, ZY, JC, YBL and NRG) located between the Badain Jaran Desert and the Tengger Desert; (b) Principle component analysis based on pairwise genetic differentiation (Fst) of populations. (c) Histogram of the STRUCTURE assignments for all individuals at K = 2 and K = 3.
corresponds to the two geographic regions separated by the Tengger Desert (Fig. 2a; Appendix S4). However, analyses based on cpDNA revealed a different pattern. Although the haplotypes in clade C2 and in C3 were exclusively distributed in the western and eastern genetic groups, respectively, the haplotypes in C1 or C4 were only mostly distributed in either the western or the eastern genetic groups, with haplotypes from each clade distributed in the eastern and western genetic groups (Fig. 4c).

Fig. 3. BEAST-derived extended Bayesian skyline plot (EBSP) based on nDNA for (a) the eastern genetic group; (b) the western genetic group and (c) all populations combined.
Fig. 4. (a) BEAST-derived chronogram of divergence times within *Arnebia szechenyi* based on three chloroplast DNA fragments (*psbA-trnH, psbB-psbF, trnL-trnF*). Blue bars on each node indicate 95% highest posterior density (HPD) intervals for the TMRCA estimates. Posterior probabilities (PP) > 0.95 are shown above nodes. (b) The most parsimonious network of the haplotypes. Circles sizes are proportional to chlorotype frequencies within the total sample. Small red dots represent hypothetical unsampled or extinct ancestral haplotypes, and lines between circles represent single mutational steps. Pie colors correspond the different clades identified. (c) Geographical distribution of the four cpDNA clades. Pie charts show the relative proportion of each cpDNA clade within each population.
Cytonuclear discordance can be caused by incomplete lineage sorting, introgression, selection or gene flow (Barrett and Schluter, 2008; Sloan et al., 2017). In this study, incomplete lineage sorting and selection seem unlikely to have caused such a discordance, due to the high sequence variation among the samples and rapid progression of lineage sorting in plastid DNA (Funk and Omland, 2003). Further, because Arnebia szechenyi was recovered as monophyletic with respect to the closely related congener A. guttata, A. linearifolia and A. finibriata (Wang et al., 2015), the role of introgression can also be neglected.

Our estimations based on cpDNA suggest that genetic divergence within Arnebia szechenyi occurred 2.66 Mya (95% HPD: 1.79–3.64 Mya), corresponding to the late Pliocene. The estimated timescales for this divergence fit well with the uplift of the Helan Mountains and the Qilian Mountains during the middle-late Pliocene (2.60–3.55 Mya; Li, 2013), which blocked water vapor from the Pacific Ocean and the Arctic Ocean (Xie and Zhang, 2013; Meng et al., 2015). Between 2.85 and 2.60 Mya, the deposition rate of alluvial-diluvial deposits and fluvial rocks in the Tengger desert area increased. At the same time, the high Himalayan Mountains and the Qinghai-Tibet Plateau (QTP) impeded warm and moist air flow from the Indian Ocean (Zhang et al., 2000). Together, these factors intensified the aridification in the Tengger desert region (Li, 2013). The severe drought conditions combined with habitat loss and destruction could have caused local or even regional extinction of A. szechenyi in the hinterland, and/or blocked gene flow between surviving geographically isolated populations and therefore promoted intraspecific divergence. After the Tengger Desert formed during the middle Pleistocene (1.1–0.9 Mya; Li, 2013; Li et al., 2014; Wang et al., 2015), its repeated expansions and contractions further accelerated differentiation within each genetic group (Fig. 4a).

During repeated geographical and demographic expansions and contractions, the nuclear genomes may have come into contact and fully recombined (in absence of reproductive isolation); however, divergent nonrecombinant chlorotypes may have been retained, as drift plays no role in large expanding populations. In addition, we found no phylogeographical structure across all populations of the species (Nst = 0.48 vs. Gst = 0.49; P = 0.10). Such a result might also indicate a complex history of colonization by gene pools that had been genetically distinct for a long period and spread across the whole distribution range. In light of this evidence, the most likely explanation for the chloroplast-nuclear discordance observed in the Arnebia szechenyi populations is that these populations experienced long periods of geographic isolation followed by range expansions precipitated by the formation of the Tengger Desert area, which homogenized the nuclear genome (which evolves rapidly) but not the chloroplast genome.

### 4.3. Regional expansion

In addition to intraspecific divergence and diversification due to long-term isolation, geological processes and climatic fluctuations may have driven range shifts in regions that were not continuously isolated (Avise, 2004). Generally, range expansions result in a series of founder events leading to loss of genetic diversity and wide fixatures of single chlorotypes (Avise et al., 1987; Hewitt, 1996, 1999, 2000; Comes and Kadereit, 1998). However, genetic signs of expansion were not found in either eastern or western genetic groups, indicating that this species did not undergo a recent rapid range expansion. This result was supported by our ENM analysis, which suggested an overall retention of the species’ potential distribution range, with the most suitable habitat remaining relatively stable since the LGM (Fig. 5).

Although neither the neutrality test nor mismatch distribution analysis based on cpDNA indicated that the populations in the western, eastern or combined regions underwent sudden demographic expansions, EBSP analysis based on nDNA suggested historical expansion episodes for the eastern group and all populations combined. According to these analyses, the effective population size of the eastern genetic group and all populations combined have been increasing since around 0.4–0.5 Mya (Fig. 3a and c), which is highly consistent with the period during when the Tengger Desert began to contract dramatically (0.47–0.12 Mya) after a large-scale expansion (Wang et al., 2015). This desert contraction likely provided an opportunity for Arnebia szechenyi to expand northward from south of the Tengger Desert and along the Helan Mountains. Previous studies have also shown that the Helan Mountains acted as a migration corridor for recolonization of desert plants in this area (Meng et al., 2015; Zhang et al., 2017; Zhao et al., 2019).

### 4.4. Asymmetric gene flow

Additionally, we detected noteworthy asymmetry in historical migration from the east to the west (32.90 >> 0.64). Historical gene
flow mainly occurred among the six northern populations from the western genetic group (populations YBL and NRG) and eastern genetic group (populations DT020, W18, S1 and Q18). This pattern was also supported by PCoA on pairwise Fst, which grouped the six populations into another group (Fig. 2b). This secondary gene exchange between two genetic groups was reflected by the nuclear DNA, implying that pollen-mediated gene flow was limited but not blocked entirely in spite of the presence of the desert. Located at the margin of the East Asian monsoon region, the Tengger Desert area is influenced dramatically by strong anomalous northwesterly and westerly winds. It is likely that the northwestward movement of the East Asian monsoon intensified the spread from the eastern to western genetic groups. The important role of East Asian monsoon system in shaping population genetic structure has also been demonstrated in other plants in this region (Yin et al., 2015, 2020; Zhang et al., 2017b).

5. Conclusions

Our analyses revealed that *Arnebia szechyeni* have high levels of genetic diversity distributed among its populations. Genetic variation of the species is not geographically structured across its whole distribution, and two nuclear genetic groups of populations were revealed, corresponding to the two geographic regions that are separated by the Tengger Desert and the Helan Mountain. However, the results from cpDNA do not support the boundary of these two geographic groups. We suggest that this chloroplast-nuclear discordance may have resulted from long periods of geographic isolation followed by lineage fusion processes that occurred in the geologic history of the Tengger Desert area. A significant asymmetry in historical migration between the two genetic groups was found, with a considerably higher rate of unidirectional gene flow from eastern to western genetic groups than from the opposite. Additionally, western genetic group showed signs of long-term local persistence and stability, while eastern genetic group showed signs of an expansion process. Our study highlights that cpDNA and nDNA may follow different mechanisms of evolution, and that their respective evolutionary history could reflect complementary aspects of past demographic and biogeographic events.

Author contributions

B.T. designated the study and managed the project. D.R.J. and B.T. collected samples and completed the species identification. B.T. prepared DNA samples and performed sequencing. M.J.F., H.Y.W., performed the molecular phylogenetic, population structure, demographic history and niche modeling analyses. M.J.F. wrote the manuscript. D.R.J. and B.T. modification and improvement the manuscript. All authors read and approved the final manuscript.

Declaration of competing interest

We believe that the results presented here are of general interests to the readers of *Plant Diversity*, and we promise that our interests to the readers of *Plant Diversity*, and we promise that our results have been acknowledged by another journal. All people contributed to this work have been acknowledged or co-authored according to his/her contributions. The authors have no conflict of interest to declare. All the authors agree to the submission of this manuscript.

Acknowledgments

We are grateful to Dr. Jun-Wei Ye for assistant on data analysis. We thank Dr. Yuan-Wen Duan for help during the field survey. This study was supported by the National Natural Science Foundation of China (41861008), Science Foundation of Yunnan Education Department (2018J3477), and the Ten-thousand Talents Program of Yunnan Province (YNWR-QNBJ-2020).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.pld.2021.04.002.

References

Avise, J.C., 2004. Molecular Markers, Natural History, and Evolution. Sinauer Associates, Sunderland, MA.
Avise, J.C., Arnold, J., Ball, R.M., et al., 1987. Intraspecific phylogeography: the mitochondrial DNA bridge between population genetics and systematics. Annu. Rev. Ecol. Syst. 18, 489–522.
Barrett, R.D.H., Schluter, D., 2008. Adaptation from standing genetic variation. Trends Ecol. Evol. 23, 38–44.
Barrett, S.C.H., Jesson, L.K., Baker, A.M., 2000. The evolution and function of stylar polymorphisms in flowering plants. Annu. Bot. 85, 253–265.
Beehler, P., 2006. Comparison of Bayesian and maximum-likelihood inference of population genetic parameters. Bioinformatics 22, 341–345.
Beheragary, L.B., 2008. Twenty years of phylogeography: the state of the field and the challenges for the Southern Hemisphere. Mol. Ecol. 17, 3754–3774.
Bouckaert, R., Heled, J., Kuhner, M., et al., 2014. Beast 2: a software Platform for Bayesian evolutionary analysis. PLoS Comput. Biol. 10, e1003537.
Comes, H.P., Kaderet, J.W., 1998. The effect of Quaternary climatic changes on plant distribution and evolution. Trends Plant Sci. 3, 432–438.
Comes, H.P., Kaderet, J.W., 2003. Spatial and temporal patterns in the evolution of the flora of the European Alpine System. Taxon 52, 451–462.
Darriba, D., Taboada, G.L., Doallo, R., et al., 2012. jModelTest 2: more models, new heuristics and parallel computing. Nat. Methods 9, 772-772.
Dellicour, S., Mardulyn, P., 2014. Spads 1.0: a toolbox to perform spatial analyses on DNA sequence data sets. Mol. Ecol. Resour. 14, 547–551.
Després, L., 2019. One, two or more species? Mitonuclear discordance and species delimitation. Mol. Ecol. 28, 3845–3847.
Drouin, G., Daoud, H., Xia, J.N., 2008. Relative rates of synonymous substitutions in the mitochondrial, chloroplast and nuclear genomes of seed plants. Mol. Phylogenet. Evol. 49, 827–831.
Du, F.K., Petit, R.J., Liu, Q.J., 2009. More introgression with less gene flow: chloroplast vs. mitochondrial DNA in the Picea aspera complex in China, and comparison with other Conifers. Mol. Ecol. 18, 1396–1407.
Excoffier, L., Lischer, H.E.L., 2010. Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. Mol. Ecol. Resour. 10, 564–567.
Excoffier, L., Smouse, P.E., Quattro, J.M., 1992. Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. Genetics 131, 479–491.
Falush, D., Stephens, M., Pritchard, J.K., 2007. Inference of population structure using multilocus genotype data: dominant markers and null alleles. Mol. Ecol. Notes 7, 574–578.
Fawcett, T., 2006. An introduction to ROC analysis. Pattern Recogn. Lett. 27, 861–874.
Field, C.B., Behrenfeld, M.J., Randerson, J.T., et al., 1998. Primary production of the biosphere: integrating terrestrial and oceanic components. Science 281, 237–240.
Fu, Y., He, C.Z., Tian, B., 2016. Development of low copy nuclear gene primers for *Arnebia szechyeni* based on transcriptome. J. S.W. Forest. Univ. 36, 86–90.
Fu, Y.X., Fu, Y.X., 1993. Statistical tests of neutrality of mutations. Genetics 133, 693–709.
Funk, D.J., Omland, K.E., 2003. Species-level paraphyly and polyphyly: frequency, causes, and consequences, with insights from animal mitochondrial DNA. Annu. Rev. Ecol. Evol. Syst. 34, 397–423.
Hewitt, G., 2000. The genetic legacy of the Quaternary ice ages. Nature 405, 907–913.
Hewitt, G.M., 1996. Some genetic consequences of ice ages, and their role in divergence and speciation. Biol. J. Linn. Soc. Lond. 58, 247–276.
Hewitt, G.M., 1999. Post-glacial re-colonization of European biota. Biol. J. Linn. Soc. Lond. 68, 87–112.
Hickerson, M.J., Carstens, B.C., Cadeneva-Bares, J., et al., 2010. Phylogeography’s past, present, and future: 10 years after Avise, 2000. Mol. Phylogenet. Evol. 54, 291–301.
Hinojosa, J.C., Kobuninová, D., Stenteczki, M.A., et al., 2019. A mirage of cryptic species: genomics uncover striking mitonuclear discordance in the butterfly Thymelicus sylvestris. Mol. Ecol. 28, 3857–3868.
Jia, D., Abbott, R., Liu, T., et al., 2012. Out of the Qinghai-Tibet plateau: evidence for the origin and dispersal of Eurasian temperate plants from a phylogeographic study of *Hippophae rhamnoides* (Elagagnaceae). New Phytol. 194, 1123–1133.
