











Integrated phylogenomic analyses unveil reticulate evolution in *Parthenocissus* (Vitaceae), highlighting speciation dynamics in the Himalayan–Hengduan Mountains

Jinren Yu^{1,2*} , Yanting Niu^{1,2,3*} , Yichen You^{1,2} , Cymon J. Cox⁴ , Russell L. Barrett⁵ ,
Anna Trias-Blasi⁶ , Jing Guo⁷ , Jun Wen⁸ , Limin Lu¹  and Zhiduan Chen¹ 

¹State Key Laboratory of Systematic and Evolutionary Botany, Institute of Botany, Chinese Academy of Sciences, Beijing 100093, China; ²University of Chinese Academy of Sciences, Beijing 100049, China; ³China National Botanical Garden, Beijing 100093, China; ⁴Centro de Ciências do Mar, Universidade do Algarve, Gambelas, Faro 8005-319, Portugal; ⁵National Herbarium of New South Wales, Australian Botanic Garden, Locked Bag 6002, Mount Annan, 2567 NSW, Australia; ⁶Royal Botanic Gardens, Kew, Richmond, TW9 3AB, UK; ⁷State Key Laboratory of Genetic Engineering and Collaborative Innovation Center of Genetics and Development, Ministry of Education Key Laboratory of Biodiversity and Ecological Engineering, Institute of Plant Biology, Center of Evolutionary Biology, School of Life Sciences, Fudan University, Shanghai 200433, China; ⁸Department of Botany, National Museum of Natural History, MRC-166, Smithsonian Institution, Washington, DC 20013-7012, USA

Summary

Authors for correspondence:

Limin Lu

Email: liminlu@ibcas.ac.cn

Zhiduan Chen

Email: zhiduan@ibcas.ac.cn

Received: 9 August 2022

Accepted: 19 October 2022

New Phytologist (2022)

doi: 10.1111/nph.18580

Key words: biogeography, comprehensive approach, Himalayan–Hengduan Mountains region, *Parthenocissus*, phylogenetic discordance, reticulation, speciation reversal.

- Hybridization caused by frequent environmental changes can lead both to species diversification (speciation) and to speciation reversal (despeciation), but the latter has rarely been demonstrated. *Parthenocissus*, a genus with its trifoliolate lineage in the Himalayan–Hengduan Mountains (HHM) region showing perplexing phylogenetic relationships, provides an opportunity for investigating speciation dynamics based on integrated evidence.
- We investigated phylogenetic discordance and reticulate evolution in *Parthenocissus* based on rigorous analyses of plastome and transcriptome data. We focused on reticulations in the trifoliolate lineage in the HHM region using a population-level genome resequencing dataset, incorporating evidence from morphology, distribution, and elevation.
- Comprehensive analyses confirmed multiple introgressions within *Parthenocissus* in a robust temporal–spatial framework. Around the HHM region, at least three hybridization hot spots were identified, one of which showed evidence of ongoing speciation reversal.
- We present a solid case study using an integrative methodological approach to investigate reticulate evolutionary history and its underlying mechanisms in plants. It demonstrates an example of speciation reversal through frequent hybridizations in the HHM region, which provides new perspectives on speciation dynamics in mountainous areas with strong topographic and environmental heterogeneity.

Introduction

Hybridization has been recognized as an important process during biological evolution and is particularly common in plants, creating reticulate structures in the Tree of Life (Rieseberg & Wendel, 1993; Soltis & Soltis, 2009; Abbott *et al.*, 2013; Payseur & Rieseberg, 2016). Interspecific hybridization can lead to either speciation or speciation reversal (Abbott *et al.*, 2013; Wu *et al.*, 2022). If reproductive isolation is newly established, speciation may occur through allopolyploidization or homoploid hybrid speciation and thus promoting biodiversity (Abbott *et al.*, 2013; Wu *et al.*, 2022). However, when reproductive isolation is weak or eroded by secondary contact, parental lineages can be replaced by proliferated hybrids, leading to genetic swamping

and eventually, speciation reversal (Seehausen, 2006; Seehausen *et al.*, 2008; Todesco *et al.*, 2016; Zhang *et al.*, 2019). Differing from most introgressions that typically form narrow hybrid zones, speciation reversal usually results in hybrid lineages with mosaic genomes replacing their distinct parental lineages and is recognized as a cause of biodiversity decline (Seehausen, 2006; Seehausen *et al.*, 2008; Todesco *et al.*, 2016). Previous studies suggested speciation reversal can be induced by fluctuating environments (Gilman & Behm, 2011; Zhang *et al.*, 2019) and is more likely to occur in regions with complex topography where the effects of climatic and elevational shifts are accentuated.

The Himalayan–Hengduan Mountains (HHM) region is well-known as a natural laboratory for investigating the processes of rapid allopatric and ecological speciation (Wen *et al.*, 2014; Favre *et al.*, 2015; Meng *et al.*, 2017; Manish & Pandit, 2018; Niu *et al.*, 2018). Continuous geological activities and climatic

*These authors contributed equally to this work.

oscillations in this region have facilitated secondary contact and increased the chances of hybridization (Liu *et al.*, 2013; Wen *et al.*, 2014; Jiang *et al.*, 2018; Dong *et al.*, 2021). While numerous examples have shown evidence for speciation and even rapid diversification through hybridization in the HHM (Gao *et al.*, 2012; Ru *et al.*, 2018; J. L. Li *et al.*, 2020), few cases of speciation reversal have been reported (Wu *et al.*, 2022). Lineages in the HHM region showing complex diversification processes driven by antagonistic forces are ideal to investigate the role of geographic and climatic changes in reticulate evolution and can enhance our understanding of speciation dynamics in areas with strong topographic heterogeneity.

Parthenocissus Planch. is a monophyletic genus in the grape family (Vitaceae) with nine recognized species from Asia (Chen *et al.*, 2007) and three from North America (Moore & Wen, 2016), reflecting the well-known Asian–North American disjunct distribution pattern (Fig. S1) (Chen & Manchester, 2007; Wen *et al.*, 2007; Nie *et al.*, 2010; Lu *et al.*, 2012, 2018; Hu *et al.*, 2022). Three lineages within the Asian clade are recognized, namely the henryana, dalzielii, and trifoliolate lineages (Fig. S1a). However, relationships among and/or within the three lineages and the North American clade have confounded researchers due to cytonuclear discordance, especially in the trifoliolate lineage (Lu *et al.*, 2012; Ma *et al.*, 2021). The four species of the trifoliolate lineage (*P. chinensis*, *P. feddei*, *P. heterophylla*, and *P. semicordata*) co-occur within the HHM region, occupying an altitudinal range from 500 to 3800 m (Chen *et al.*, 2007), with *P. heterophylla* extending to India and Indonesia (Fig. S1b). Species in the trifoliolate lineage are morphologically similar, particularly *P. semicordata* and *P. chinensis*. *Parthenocissus chinensis* was once recognized as a narrow endemic to a dry habitat area in the upper Yangtze River valley, but molecular evidence suggests more widely distributed individuals formerly identified as *P. semicordata* form a strongly supported clade with *P. chinensis* (Nie *et al.*, 2010). No consistent morphological characters can distinguish individuals of the *P. chinensis* clade from those of the *P. semicordata* clade, with only minor size differences in leaflets and inflorescences (Chen *et al.*, 2007; Lu *et al.*, 2012). Despite their morphological similarity, phylogenetic analyses have revealed that *P. semicordata* and *P. chinensis* are well-diverged lineages, resolving *P. semicordata* as sister to *P. heterophylla*, rather than to *P. chinensis* (Nie *et al.*, 2010). However, *P. heterophylla* is a distinct species based on morphological, geographical, and molecular evidence, differing from *P. chinensis* and *P. semicordata* in possessing larger inflorescences and more obvious appendages on the inner side of petals (Lu *et al.*, 2012). We therefore suggest that gene flow, and potentially ongoing speciation reversal, may exist between *P. semicordata* and *P. chinensis*. The plastid capture hypothesis should also be considered; that is, the similarity between *P. chinensis* and *P. semicordata* may result from their recent divergence, and multiple plastid capture events might have caused cytonuclear discordance in the trifoliolate lineage. These alternative scenarios can be tested by investigating the reticulate evolutionary history of the trifoliolate lineage.

Based on phylogenetic signals left by past hybridization, it is possible to restore complex reticulate evolutionary histories

involving speciation and speciation reversal. However, it remains challenging to reconstruct a robust reticulate evolutionary history of lineages enduring frequent hybridizations and rapid diversification. Besides hybridization, gene tree estimation error and incomplete lineage sorting (ILS) also can cause phylogenetic discordance (Morales-Briones *et al.*, 2018; Cai *et al.*, 2021). Gene tree estimation error can be quantified by multiple approaches (Budenhagen *et al.*, 2016; Arcila *et al.*, 2017; Cai *et al.*, 2021). However, species network inference methods that tease apart the interplay between ILS and hybridization (e.g. Mirarab *et al.*, 2014; Yu *et al.*, 2014; Yu & Nakhleh, 2015; Solís-Lemus *et al.*, 2017) sometimes generate inaccurate results due to limited numbers of taxa and/or complex reticulation events (Hejase & Liu, 2016; Blair & Ané, 2020). In such cases, rigorous statistical analyses, such as the ABBA–BABA test (Green *et al.*, 2010), can help determine gene flow (e.g. Rose *et al.*, 2021). Recently, multiple empirical studies of animals have illustrated that comprehensive approaches integrating population-level analyses, morphology, and distribution are effective in reliably elucidating reticulations (e.g. Andersen *et al.*, 2021; Ferreira *et al.*, 2021; Pavón-Vázquez *et al.*, 2021). However, such integrative studies of reticulate evolution in plants are rare, even though plants tend to show more complex evolutionary histories than animals due to frequent hybridizations (Barraclough & Nee, 2001; Linder & Rieseberg, 2004; Soltis & Soltis, 2009). Therefore, additional studies of plant lineages, applying a comprehensive analytical approach, are required to build robust phylogenetic networks revealing complex reticulate evolution.

In this study, we use *Parthenocissus* as an exemplar taxon to test how geographic and climatic changes facilitated reticulate evolution with a comprehensive set of methodological approaches combining phylogenetic reconstruction, phylogenetic conflict assessment, gene flow inference, biogeographic reconstruction, and population structure analysis. We focus on demonstrating a reticulate evolutionary history of *Parthenocissus* within a robust temporal–spatial framework by employing transcriptome orthologs and complete plastomes of all 12 *Parthenocissus* species. We then investigate speciation dynamics and underlying mechanisms in the HHM region using a population-level genome resequencing dataset including 212 individuals from 37 populations of the trifoliolate lineage of *Parthenocissus*. We specifically test the following two conditions to support potential speciation reversal in *P. chinensis* and *P. semicordata*: (1) *P. chinensis* and *P. semicordata* are deeply diverged lineages and (2) hybrids with admixed genomes have replaced or are replacing their parents. We also demonstrate that plastid capture cannot fully explain cytonuclear discordance in the trifoliolate lineage.

Materials and Methods

Taxon sampling

Taxon sampling was designed to resolve phylogenetic relationships among major lineages of *Parthenocissus* with plastome and transcriptome data, and investigate the evolutionary history of the trifoliolate lineage with genome resequencing data. For

plastome sequencing, we sampled 14 individuals representing 12 species of *Parthenocissus* and one sister taxon (*Yua austro-orientalis* (F.P. Metcalf) C.L. Li) as outgroup (Table S1). The same individuals of *Parthenocissus* were used for transcriptome sequencing but with *Y. thomsonii* (M.A. Lawson) C.L. Li as outgroup (Table S1) as the attempts at transcriptome sequencing of *Y. austro-orientalis* failed. To make use of reliable calibration points in Vitaceae, we included 27 plastomes representing other clades of Vitaceae to generate a family-level dataset including 40 plastomes (hereinafter referred to as the 40-plastome dataset) and downloaded transcriptomes of two *Vitis* species (Table S2).

We collected a total of 212 individuals from 37 populations of the trifoliolate lineage throughout the geographic distribution ranges of *P. chinensis* C.L. Li, *P. feddei* (H. Lév.) C.L. Li, *P. heterophylla* (Blume) Merr., and *P. semicordata* (Wall.) Planch. (Table S3; populations of the four species were marked as CH, FE, HE, and SE, respectively). Voucher specimens were deposited at the Herbarium of Institute of Botany, Chinese Academy of Sciences (PE; herbaria codes follow Thiers, 2017).

Plastome sequencing, assembly, and annotation

Library construction and sequencing were based on DNA extracted from silica gel-dried leaves. Paired-end reads of 150 bp for all samples were generated in a single lane on an Illumina HiSeq2500 sequencer. For each newly sequenced sample (including samples representing species of *Parthenocissus* and *Yua* and the population samples), we filtered raw bases using TRIMMOMATIC 0.3.2 (Bolger *et al.*, 2014) with the following settings: SLIDINGWINDOW:4:20 LEADING:10 TRAILING:10 MINLEN:70. We assembled plastomes with filtered data using both *de novo* assembly method and a reference genome of *Vitis vinifera* L. (NC007957; Jansen *et al.*, 2006). We mapped quality-controlled reads against the reference genome in GENEIOUS 9.1.4 (Kearse *et al.*, 2012) with medium–low sensitivity using five iterations. Reads were exported and plastomes were *de novo* assembled using VELVET 7.0.4 (Zerbino & Birney, 2008) with auto-adjusted coverage cutoffs and *k*-mer sizes from 97 to 145. Final plastome scaffolds were manually adjusted to eliminate errors and ambiguities. We annotated plastomes of *Parthenocissus* and *Yua* using DOGMA (Wyman *et al.*, 2004) and GENEIOUS, with the plastome of *V. vinifera* as a reference.

Transcriptome sequencing, assembly, and ortholog inference

Fresh leaves of 13 *Parthenocissus* individuals whose plastomes were sequenced and one individual of *Yua thomsonii* were used for RNA extraction, library construction, and sequencing (individuals used are shown in Table S1). For all samples, 150 bp paired-end reads were generated in a single lane on an Illumina HiSeq2500 sequencer. The raw bases were filtered using TRIMMOMATIC with the same settings above. Transcriptome for each sample was *de novo* assembled using TRINITY 2.5.1 (Grabherr *et al.*, 2011; Haas *et al.*, 2013). We used TRANSDCODER 5.0.0 (Haas *et al.*, 2013) to predict peptide sequences longer than 100

amino acids with an open read frame. Nonredundant, representative sequences were retained by CD-HIT 4.6.5 (Li & Godzik, 2006) with a threshold value of 0.98. Finally, 2587 single-copy orthologs of the 13 species (the individual of *P. heterophylla* from India was excluded to keep one individual per species) were identified with ORTHOFINDER 2.2.6 (Emms & Kelly, 2015) based on the output of CD-HIT (hereinafter referred to as 13taxa-2587nu dataset).

To further test potential reticulations within lineages, we divided the genus into three lineage subsets: (1) the North American lineage (*P. vitacea* (Knerr) Hitchc., *P. quinquefolia* (L.) Planch., and *P. heptaphylla* (Planch.) Britton), (2) the dalzielii lineage (*P. dalzielii* Gagnep., *P. suberosa* Hand.-Mazz., and *P. tricuspidata* (Siebold & Zucc.) Planch.), and (3) the trifoliolate lineage (*P. chinensis*, *P. feddei*, *P. heterophylla*, and *P. semicordata*). To test whether network analysis is sensitive to taxon sampling (Karimi *et al.*, 2020), individuals of *P. heterophylla* from India and Indonesia were both or separately included in the trifoliolate subset, and different outgroups were selected for each subset. We also generated orthologs from 12 *Parthenocissus* species, 1 *Yua* species, and 2 *Vitis* species (hereinafter referred to as 15taxa-1101nu) for divergence time estimation using fossil calibrations. Orthogroups of each dataset were identified using ORTHOFINDER (Table S4).

Alignment and phylogenetic analyses

Transcriptome orthologs (i.e. 13taxa-2587nu), and plastomes with one inverted repeat removed, were aligned using L-INS-i algorithm in MAFFT 7.450 (Katoh & Standley, 2013) and trimmed with TRIMAL 1.2 with the option ‘automated1’ (Capella-Gutiérrez *et al.*, 2009).

A phylogenetic tree based on the 40-plastome dataset was constructed using IQ-TREE 2.1.2 (Nguyen *et al.*, 2015) with 5000 ultrafast bootstrap replicates (Hoang *et al.*, 2018) under edge-unlinked partitioned model (Chernomor *et al.*, 2016), where the best model scheme was selected by MODELFINDER (Kalyaanamoorthy *et al.*, 2017). Both concatenation and coalescent-based methods were used to reconstruct phylogenetic relationships of *Parthenocissus* for the nuclear data. We used IQ-TREE to construct the phylogeny of *Parthenocissus* based on the concatenated 13taxa-2587nu dataset with 5000 ultrafast bootstrap replicates, and the best model was selected by MODELFINDER. We constructed gene trees of each ortholog with maximum likelihood (ML), Bayesian inference (BI), and maximum parsimony (MP) approaches. Maximum likelihood gene trees were constructed using IQ-TREE with the best model selected by MODELFINDER and 100 bootstrap replicates. Bayesian inference gene trees were constructed by MRBAYES 3.2.7a (Ronquist *et al.*, 2012), and the best model for each gene was selected by MRMODELTEST2 (Nylander, 2004) using the AIC metric. Maximum parsimony gene trees were constructed by PAUP* v.4.0a163 (Swofford, 2002) with 1000 bootstrap replicates using heuristic search. We then generated multispecies coalescent (MSC) trees in ASTRAL 5.6.3 (Mirarab *et al.*, 2014) with the ML, BI, and MP gene trees, respectively, as input. As gene trees with poorly

supported clades can mislead MSC tree-based methods (Zhang *et al.*, 2018), we collapsed nodes of gene tree with <10%, 50%, or 75% bootstrap support (BS) to determine the best cutoff for ASTRAL species tree inference. We also used SVDQUARTETS (Chifman & Kubatko, 2014) implemented in PAUP* to reconstruct phylogeny of *Parthenocissus* with the concatenated 13taxa-2587nu alignment. All possible quartets were evaluated, and trees were selected using the QUARTET FM (QFM) method with 1000 bootstrap replicates.

Phylogenetic conflict assessment

We used PHYPARTS (Smith *et al.*, 2015) to calculate the number of conflicting and concordant bipartitions and the ‘internode certainty all’ (ICA) values representing the degree of certainty for internodes considering the frequency of the bipartition (Salichos *et al.*, 2014). We also calculated gene concordance factors (gCF) using IQ-TREE, with the ASTRAL species tree as the mapping tree for ML, BI, and MP gene trees. As gene trees constructed with different methods showed similar phylogenetic conflicts, we only used the ML gene trees in subsequent analyses. We compared the topologies of 2587 rooted ML gene trees, and the 10 most frequently occurring topologies were documented. To investigate conflicts within *Parthenocissus*, we also compared the topologies of 2587 unrooted single gene trees. We assessed all three possible topologies for the four species in the trifoliolate lineage. SPLITSTREE 4.13.1 (Huson & Bryant, 2006) was used to visualize conflicts within *Parthenocissus* among 2587 rooted single gene trees. We collapsed all nodes below the best BS cutoff to avoid overestimating conflicts among single gene trees in SPLITSTREE.

Tests for the effects of ILS and gene tree estimation error on gene tree conflicts

We investigated whether the observed gene tree discordance could be explained by ILS or gene tree estimation error as described by Cai *et al.* (2021). We carried out coalescent simulations to test whether ILS alone can explain gene tree discordance. The population mutation parameter ‘theta’ of each internal branch was used to estimate the level of ILS, which was calculated by dividing branch length in mutation units by length in coalescent units. We also simulated 2587 gene trees under the MSC model with the R package PHYBASE 1.4 (Liu & Yu, 2010). Thereafter, the distributions of Robinson–Foulds distances (Robinson & Foulds, 1981) between the species tree and the simulated gene trees, and those between the species tree and the empirical gene trees, were compared. The effect of gene tree estimation error was quantified by bipartition information of the species tree based on simulated gene trees generated by RAXML v.8.2.12 (Stamatakis, 2014) with the ‘-f b’ option, which shows how often each node of the species tree is recovered by the simulated gene trees.

Gene flow inference

To construct a framework of reticulate evolution in *Parthenocissus*, and to test whether plastid capture events left traces in the

genome, we identified gene flow using a phylogenetic network inference and the ABBA–BABA test. We first used the 13taxa-2587nu dataset (Table S4) to infer the optimal network in *Parthenocissus* with Species Networks applying Quartets (SNAQ; Solís-Lemus *et al.*, 2017) implemented in PHYLONETWORKS (<https://github.com/crsl4/PhyloNetworks.jl>), by calculating the maximum pseudolikelihood of a network from four-taxon concordance factors (CFs). ASTRAL MSC species tree was selected as the starting tree. Concordance factor tables were generated with ML gene trees (Ané *et al.*, 2007; Larget *et al.*, 2010) following the Tree Incongruence Checking in R (TICR) pipeline (Stenz *et al.*, 2015). We ran SNAQ with possible maximum hybrid node number (hmax) from 0 to 10 for 100 searches to determine the optimal hmax based on the slope of a plot of $-\log\text{plik}$ against maximum hybridization number (Solís-Lemus & Ané, 2016). The network with the optimal hmax was selected as the starting network for bootstrap analysis with 100 replicates and 100 SNAQ searches. Phylogenetic networks for different subsets were estimated as described above, but with the hmax varying from 0 to 5. To test whether outgroup selection affected the network inference, each subset was analyzed with different outgroups (Table S4).

ABBA–BABA tests were conducted to identify gene flow between and within clades using the ‘CalcD’ function in the R package EVOBIR 1.3 (Blackmon & Adams, 2015). We tested all 102 possible combinations of three species from each of the trifoliolate, dalzielii, henryana lineages, and the North American clade, with *Y. thomsonii* as the outgroup. In addition, we used 17 quartets representing all possible combinations of three species within each lineage-subset with different outgroups, to test introgression within each lineage. The taxon compositions of all the quartets are shown in Table S5. We conducted the ABBA–BABA test using gene alignments obtained by concatenating and pruning the alignments of the subset datasets and the 13taxa-2587nu dataset. To obtain a Z-score for the null hypothesis of no introgression, 1000 bootstrap replicates were conducted in each test. Z-scores statistically deviating from the null expectation ($Z > 3$) indicate introgression events within the quartet. P-values were corrected with the Benjamini–Hochberg method (Benjamini & Hochberg, 1995) to control multiple comparisons.

Divergence time estimation

To investigate the divergence history of *Parthenocissus*, especially of the trifoliolate lineages, divergence time estimation was conducted with the 40-plastome dataset and 15taxa-1101nu dataset using MCMCTREE in the PAML 4.9j package (Yang, 2007). For the plastome dataset, *Leea indica* (Burm. f.) Merr. was used as the outgroup. A secondary calibration point was applied to constrain the stem age of Vitaceae, and two fossils were assigned to internal nodes (see stars in Fig. S2). Calibration strategies are detailed in Methods S1. MCMCTREE was run with the following settings: birth–death model, correlated rates, and HKY85 substitution model with $\alpha = 0.5$. We combined the results of two independent MCMC chains, where samples were drawn every 10 000 generations with the first 20% iterations discarded as burnin,

until effective sample size (ESS) > 200 for all parameters. We also estimated divergence time using 15taxa-1101nu dataset partitioned by codon position (first/second/third), with calibration points in Methods S1 (see stars in Fig. S3), and the same settings of MCMCTREE. We also ran MCMCTREE without sequence data and compared marginal prior distributions (e.g. distribution results from prior interactions) to posterior distributions to check prior interactions (Fig. S4).

Ancestral area reconstruction

We used the Bayesian Binary MCMC (BBM) method in RASP 4.1 (Yu *et al.*, 2015) to estimate the ancestral areas of *Parthenocissus* with the topologies based on 40-plastome dataset and 15taxa-1101nu dataset (Table S4). Four geographic regions representing the current distribution of *Parthenocissus* were defined: A, Qinghai–Tibet Plateau (QTP); B, East Asia (including eastern China, Japan, and the Korean Peninsula); C, Southeast Asia (including Indochina, Java, and the Indian subcontinent); and D, North America. We set the root distribution to null, applied 10 MCMC chains with the F81 + G model, and sampled the posterior distribution every 1000 generations discarding the first 20% as burnin, until ESS for all parameters > 200.

Distribution data for *Parthenocissus* were mainly assembled based on specimens from the Chinese Virtual Herbarium (<http://www.cvh.org.cn>), Global Biodiversity Information Facility (GBIF.org., 2016), and other herbarium specimens we examined. We confirmed the identification of individual specimens and assessed the altitudinal range of species in the trifoliolate lineage to evaluate whether topography was involved in species diversification in the HHM region. The strategy for retaining natural distribution records and ensuring each species' distribution range is further described in Methods S2.

Genetic structure and phylogenetic analyses of the trifoliolate lineage

We used a population-level dataset to investigate the hybridization zones of the trifoliolate lineage species. Coding sequences of 212 individuals from 37 populations were extracted by HYBPIPER 1.3.1 (Johnson *et al.*, 2016) with the 2587 transcriptome orthologs as references and were aligned by MAFFT. For each gene alignment, sites with missing data > 50% were excluded. Then, we extracted SNPs from the remained 1091 ortholog alignments with SNP-SITES (Page *et al.*, 2016) and conducted linkage disequilibrium (LD) pruning using PLINK 1.9 (Purcell *et al.*, 2007) to filter out highly correlated SNPs for genetic structure analysis, with option '--indep-pairwise' and parameters '200 100 0.2'. We identified genetic subgroups with a Bayesian analysis in STRUCTURE 2.3.4 (Pritchard *et al.*, 2000) using the admixture model with $K = 2-8$ and chose the optimal K value using STRUCTURE HARVESTER (Earl & VonHoldt, 2012) based on the maximum Delta K value (Evanno *et al.*, 2005). To statistically evaluate the spatial concentration of gene flow, we performed Pearson correlation analyses between introgression intensity and geographic distance for three species pairs using the R package HMIST 4.6-0

(Harrell, 2021). Introgression intensity was denoted by the mean f_d statistic between population pairs calculated by 'ABBABABAwindows.py' script (Martin *et al.*, 2015). We obtained the population MSC tree with ASTRAL using 1021 gene trees constructed by IQ-TREE. We also constructed an ML population tree in IQ-TREE with a plastome dataset including 212 individuals.

Results

Phylogenetic relationships based on the plastome and transcriptome data

To resolve phylogenetic relationships among species of *Parthenocissus*, 13 whole plastomes of *Parthenocissus* were assembled (Table S1). Phylogenetic analyses based on plastomes showed four well-supported clades within *Parthenocissus* (all with BS = 100%; Fig. 1a). The North American clade of three species formed a sister group to the Asian clade (Fig. 1a). Within the Asian clade, three well-supported lineages were identified (Fig. 1a): the henryana lineage with 5-foliolate leaves including *P. henryana* and *P. laetevirens*; the dalzielii lineage with both simple and trifoliolate leaves including *P. dalzielii*, *P. tricuspidata*, and *P. suberosa*; and the trifoliolate lineage with trifoliolate leaves including *P. chinensis*, *P. feddei*, *P. heterophylla*, and *P. semicordata*. Moreover, relationships among these three Asian lineages were well-resolved, with the henryana lineage diverging first and the dalzielii and trifoliolate lineages forming sister-groups.

The ML and SVDQUARTETS methods conducted on the 13taxa-2587nu dataset yielded well-supported and identical topologies. We also obtained identical ASTRAL MSC species trees with high supporting values regardless of the BS cutoff, but we selected the species tree constructed from gene trees with node BS values $\geq 10\%$ for subsequent analysis to reduce potential noises, as suggested by Zhang *et al.* (2018). Well-supported cytonuclear phylogenetic incongruences were detected within the North American clade and the trifoliolate lineage (Fig. 1). Within the North American clade, the plastome data supported a closer relationship between *P. vitacea* and *P. quinquefolia*, while the transcriptome data placed *P. vitacea* together with *P. heptaphylla* (Fig. 1). Within the trifoliolate lineage, the plastome data supported *P. heterophylla* as sister to *P. chinensis*, while the transcriptome data identified a closer relationship between *P. chinensis* and *P. semicordata* (Fig. 1).

Topological conflicts among transcriptome orthologs

The gene tree topologies of 2587 single-copy orthologs were highly variable (Figs S5, S6; Table S6). In the trifoliolate lineage, most gene trees were congruent with the transcriptome phylogeny. Topology supporting *P. chinensis* and *P. heterophylla* as sister groups is secondarily dominant (Fig. S7).

SPLITS TREE detected conflicts from deep to shallow nodes of *Parthenocissus*, and the trifoliolate lineage showed particularly significant conflicts (Fig. S8). Gene trees constructed with different methods showed similar patterns of low ICA value and gene

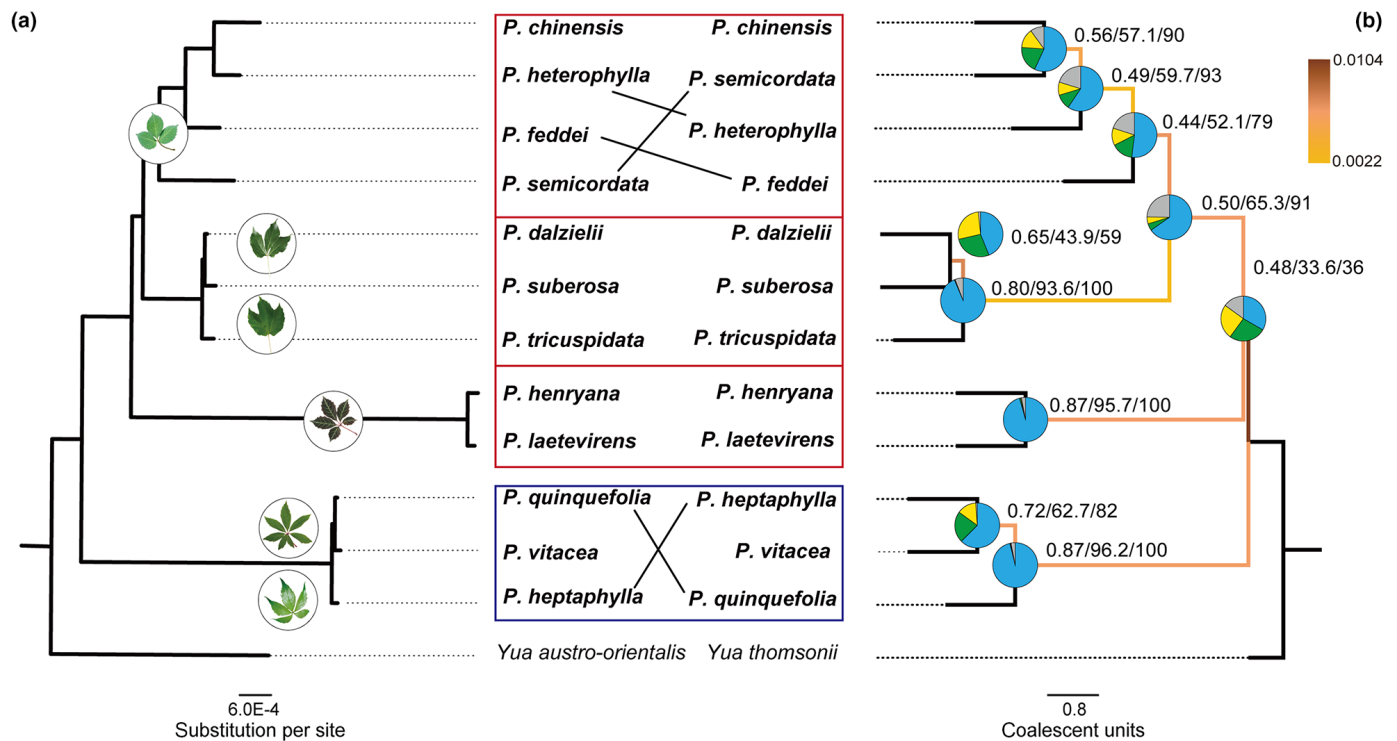


Fig. 1 Topological discordance between (a) the phylogenetic relationships of *Parthenocissus* based on the 40-plastome dataset with all nodes supported by 100% maximum likelihood bootstrap support (BS) value and (b) the phylogenetic relationships of *Parthenocissus* based on 13taxa-2587nu dataset with internode certainty all values/gene concordance factors/bipartition information restored from simulated gene trees shown above the branches. *ASTRAL* and *SVDQUARTETS* generated identical topologies with all the branches of 100% BS value. The multispecies coalescent species tree generated by *ASTRAL* is presented to show branch lengths. The pie charts at each node represent the proportion of genes supporting congruent relationships (blue), the dominant alternative (green), the remaining conflicting alternatives (yellow), and genes with <75% BS value for that node (gray). Internal branches are colored to show population mutation parameter theta.

concordance factors, strongly supporting potential conflicts with the species tree (Figs 1b, S9).

Effects of ILS and gene tree estimation on conflicts and identification of reticulation events

The difference between distribution of Robinson–Foulds distances calculated with simulated and empirical gene trees indicates that, in addition to ILS, hybridization can also explain gene tree discordance (Fig. S10). Theta, calculated from the ML tree and the MSC species tree, showed weak signals of ILS for most branches, except for the stem of the Asian clade with a short internal branch (Fig. 1b). The stem of the Asian clade was also affected by gene tree estimation error, as suggested by the low recovery of simulated gene trees with this lineage (Fig. 1b).

SNAQ runs performed on the whole genus suggested the optimal value of $h_{max} = 4$, but only three hybridization events have $BS > 10\%$ (Figs 2, S11a). The highest supported introgression was detected within the North American clade between *P. vitacea* and *P. quinquefolia* ($BS = 46\%$; Fig. 2a). Introgression between *P. suberosa* and the common ancestor of the dalzielii lineage and minor gene flow from *P. chinensis* to *P. feddei* were also detected. SNAQ searches conducted on the subsets obtained better-supported results (Figs 2b–f, 3). Two introgression events between (or within) the dalzielii and trifoliolate lineages were identified (Fig. 2b): (1) gene flow from *P. suberosa* to the

common ancestor of *P. heterophylla*–*P. semicordata*–*P. chinensis* ($BS = 95\%$) and (2) gene flow from *P. heterophylla* to *P. semicordata* ($BS = 100\%$). When the henryana lineage was used as an outgroup, no gene flow with $BS > 10\%$ was identified (Fig. 2c). The introgression between *P. heptaphylla* and the ancestor of the North American clade was identified despite using different outgroups (Fig. 2d–f). In the trifoliolate lineage, gene flow from *P. semicordata* to *P. heterophylla* (Fig. 3a,d–f) was identified with the henryana lineage as outgroup, but the direction of gene flow reversed when *P. dalzielii* was used as outgroup and two *P. heterophylla* individuals were separately included (Fig. 3b–c).

ABBA–BABA tests supported introgression events between the North American clade and the trifoliolate lineage or the henryana lineage with significant Z -scores (Fig. 4; Table S5). Introgressions were identified within all three subsets. In the trifoliolate lineage, introgressions between *P. semicordata* and *P. heterophylla* were strongly supported, whereas introgressions between *P. chinensis*–*P. feddei* or *P. semicordata*–*P. feddei* were not consistently supported. In the dalzielii lineage and the North American clade, introgressions were detected irrespective of the selected outgroup (Fig. 4; Table S5).

Divergence times and biogeographic reconstruction

The genus *Parthenocissus* was estimated to have split with *Yua* in East Asia (Node 1 in Figs 5, S12) during the early Eocene (Node

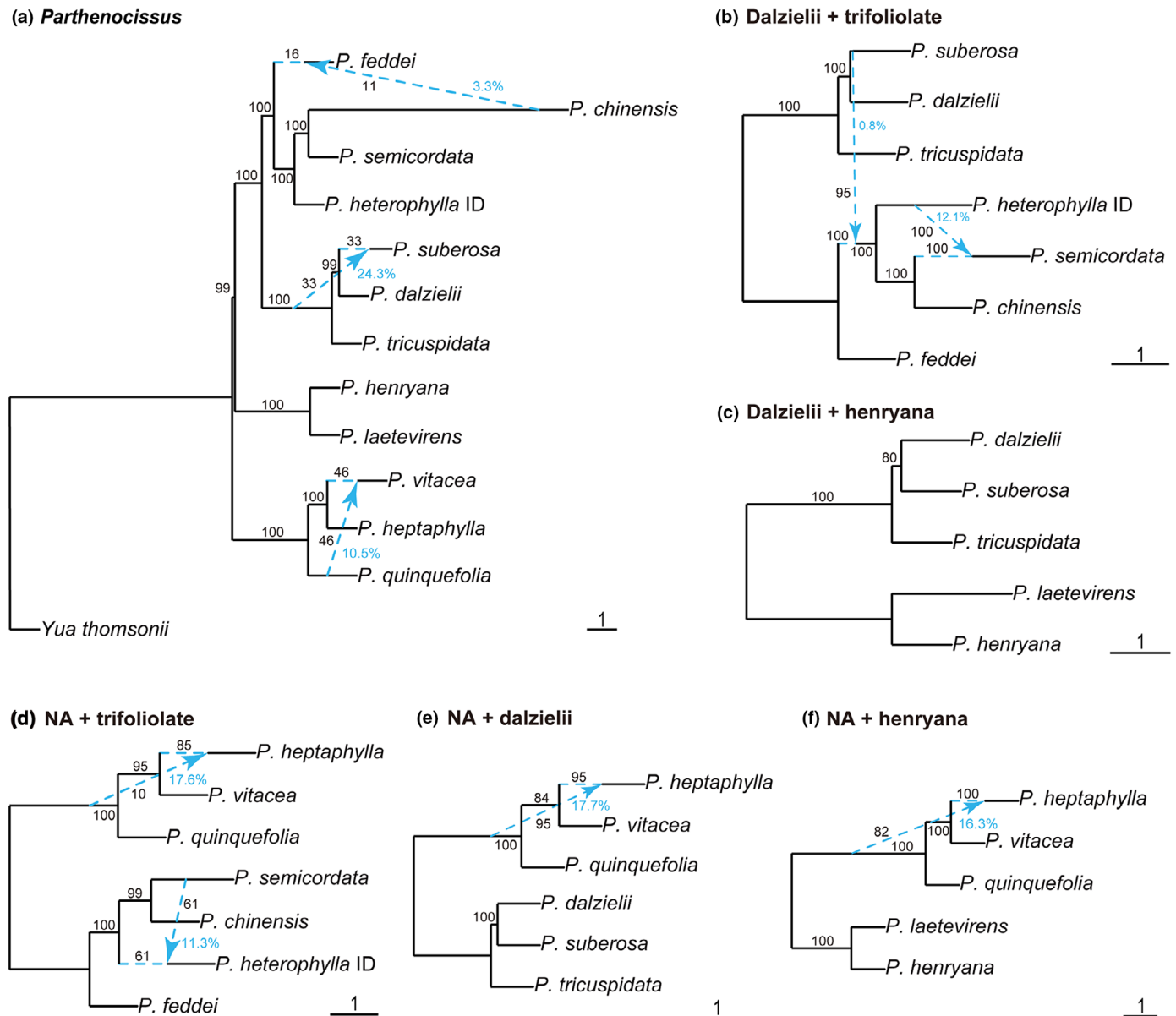


Fig. 2 Reticulations in *Parthenocissus* estimated with SNAQ on (a) 13taxa-2587nu dataset of *Parthenocissus*, (b, c) the dalzielii subset with different outgroups, and (d–f) the North American (NA) subset with different outgroups. Hybrid edges with bootstrap support (BS) values $\geq 10\%$ are denoted by blue dashed lines. Minor inheritance probabilities and BS values are shown in blue and black numbers along the edges, respectively. Edge lengths are shown in coalescent units, with the length of each terminal branch set as 1. 'ID' represents the individual from Indonesia.

1 in Figs S2, S3) based on both plastome and nuclear gene datasets. According to the 15taxa-1101nu dataset, *Parthenocissus* (the crown group) started to diversify during the early Eocene (Node 2 in Figs 5, S3). Extant species of the North American clade did not diversify until the middle Miocene (Node 3 in Figs 5, S3). Crown age of the Asian clade was estimated to be *c.* 43.06 Ma in the middle Eocene (Node 4 in Figs 5, S3). The trifoliolate lineage diversified during the Oligocene (Node 8 in Figs 5, S3). For most nodes within *Parthenocissus*, the divergence times calculated from the two datasets were similar (Figs 5, S2, S3), with only Nodes 3–5 being much younger based on the 40-plastome dataset, indicating later diversifications of plastomes in the crown groups of

the North American lineage (Node 3 in Fig. S2), the Asian clade (Node 4 in Fig. S2), and the henryana lineage (Node 5 in Fig. S2). As concerted evolution between genes can result in slower nucleotide substitution rates in plastid genes (Okuyama *et al.*, 2005; Dong *et al.*, 2021), we primarily used the divergence times estimated from the 15taxa-1101nu dataset for discussion.

Phylogeographic analyses of the trifoliolate lineage

After filtering, the final occurrence dataset of the trifoliolate lineage included 809 records. The altitudinal ranges of *P. chinensis*, *P. feddei*, *P. heterophylla*, and *P. semicordata* were 1000–

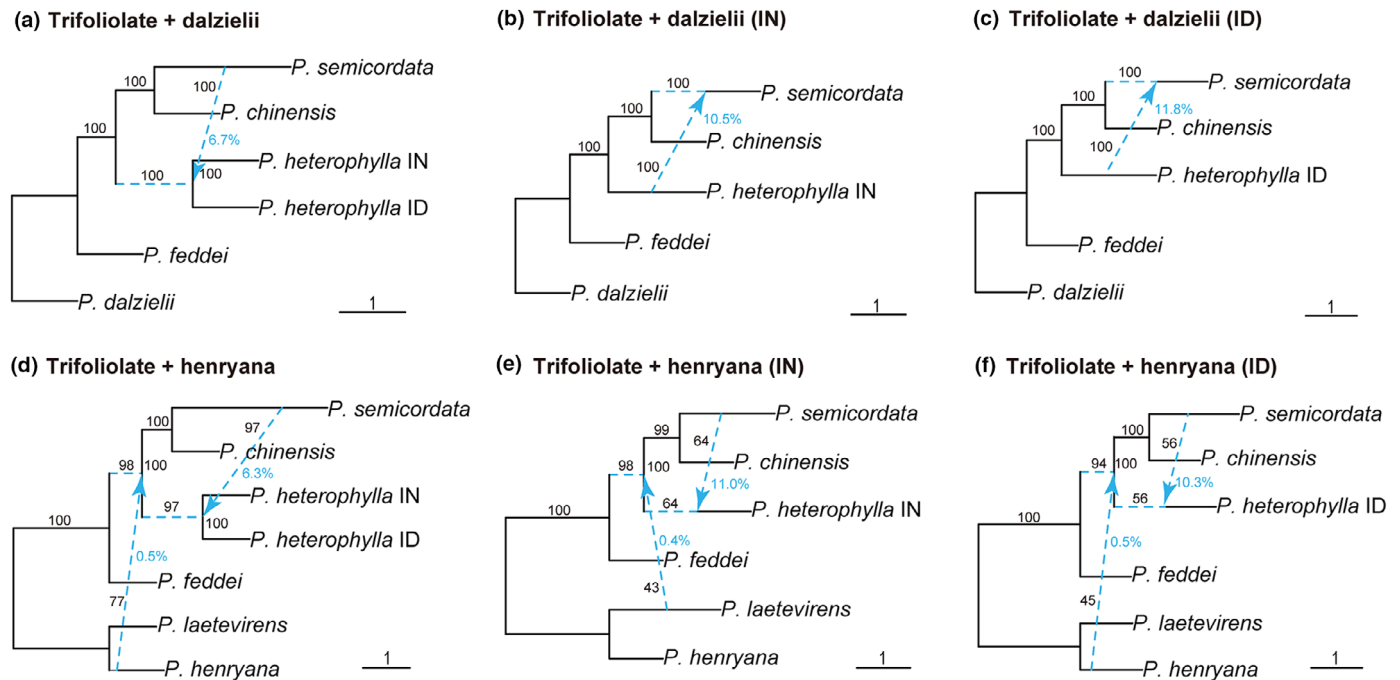


Fig. 3 Reticulations in *Parthenocissus* estimated with SNAQ on the trifoliolate subset, respectively with (a–c) *P. dalzielii* and (d–f) the henryana lineage as outgroup. Individuals from India (IN) and Indonesia (ID) were both or separately included to test sampling sensitivity of the method. Hybrid edges with bootstrap support (BS) values are denoted by blue dashed lines. Minor inheritance probabilities and BS values are shown in blue and black numbers along the edges, respectively. Edge lengths are shown in coalescent units, with the length of each terminal branch set as 1.

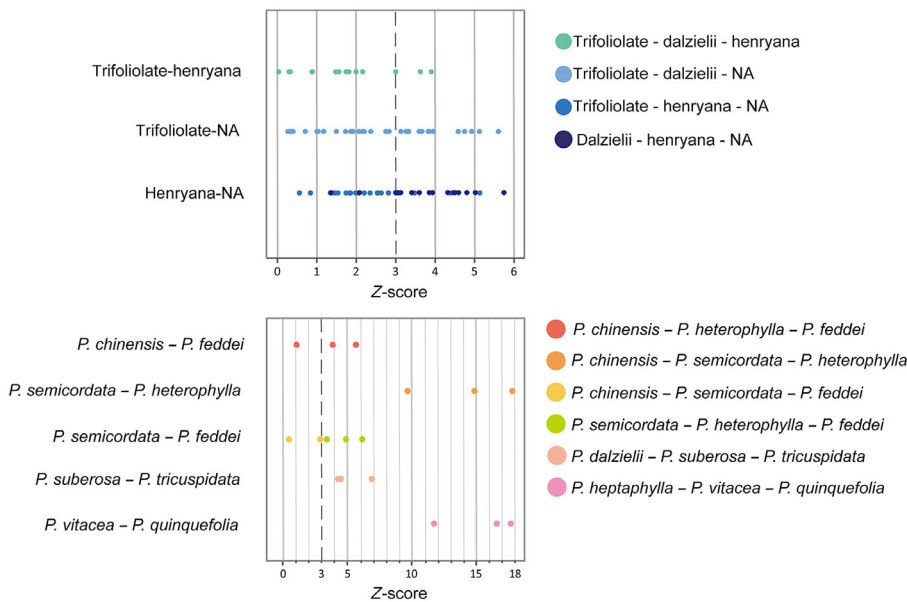


Fig. 4 Z-scores of gene flow (denoted on the x-axis) between or within three lineages (trifoliolate, dalzielii, and henryana) and the North American clade of *Parthenocissus* that were recognized as statistically significant ($Z > 3$) in at least one quartet. Z-scores were obtained from 1000 bootstrap replicates. Combinations of species or the lineages/clade they belong to are listed on the right, as indicated by different colors. Names of lineages were simplified; 'NA' represents the North American clade.

3800, 330–1650, 550–2600, and 1600–4065 m, respectively (Fig. 6b; Table S3). A total of 1448 filtered independent SNPs were used as input for population structure analysis. At optimal $K = 3$, individuals of *P. heterophylla* and *P. feddei* were clustered as separate groups, while individuals of *P. chinensis* and *P. semicordata* were clustered together with some populations showing genetic admixture from *P. heterophylla* (Fig. 6a).

At $K = 4$, individuals of *P. chinensis* and *P. semicordata* were identified as two distinct clusters, though obvious admixture still exists between the two clusters (Fig. 6a). The Pearson r correlation showed that the mean f_d statistic was negatively correlated with geographic distance for population pairs of *P. semicordata*–*P. heterophylla*, *P. chinensis*–*P. heterophylla*, and *P. semicordata*–*P. chinensis* (Fig. S13).

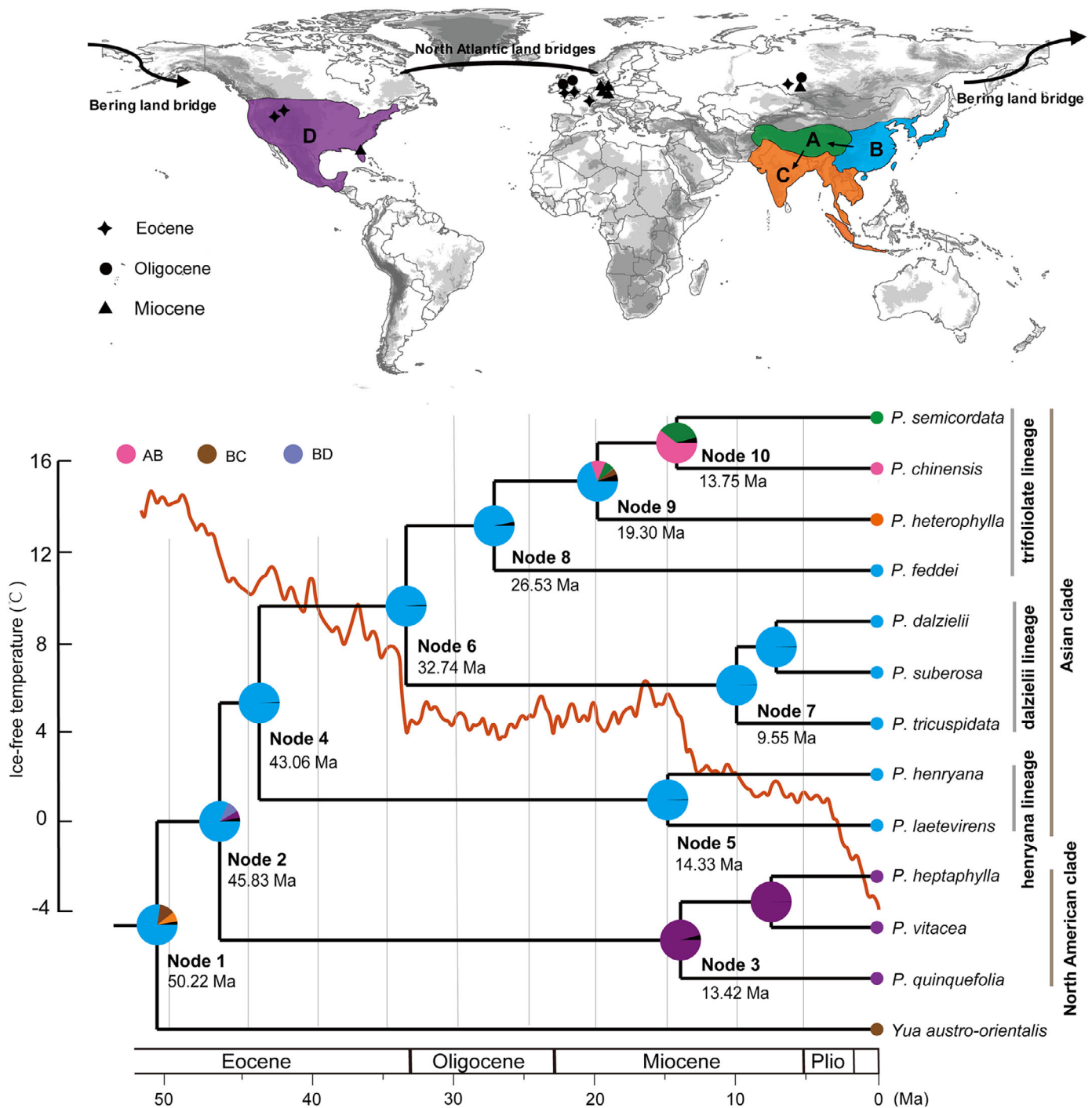


Fig. 5 Ancestral area reconstructions for *Parthenocissus* in RASP using the MCMCTREE-derived chronogram based on nuclear genes (15taxa-1101nu dataset). Node numbers and mean node ages are displayed at the corresponding nodes. The red line represents the global temperature changes (from Westerhold *et al.*, 2020). The black symbols represent locations of fossils for *Parthenocissus* summarized by Chen (2009). The pie charts indicate the relative possibilities of ancestral areas estimated. A, Qinghai–Tibet Plateau; B, East Asia (including eastern China, Japan, and the Korean Peninsula); C, Southeast Asia (including Indochina, Java, and the Indian subcontinent); and D, North America. Ma, million years ago.

Phylogenetic analyses based on the 212 individuals-1021 nu and 212-plastome datasets resulted in conflicting relationships, with a few individuals clustered in unexpected positions compared with topologies obtained from the 13-taxa-2587 nu and 40-plastome datasets (Fig. S14). In the nuclear phylogeny, six individuals from SE populations distributed along the Himalaya

formed a sister clade to the clade of *P. chinensis* and the remaining individuals of *P. semicordata* (Fig. S14a). The plastid phylogeny showed that individuals of CH1, HE6, and some SE populations at the southern edge of the HHM formed a clade sister to the *P. chinensis* clade (Fig. S14b). In addition, the individuals of CH13 and one individual of CH14 in the southern border

of the Hengduan Mountains were nested within the *P. heterophylla* clade (Fig. S14b).

Discussion

Reticulate evolution in *Parthenocissus* in a temporal–spatial framework

We establish a well-supported temporal–spatial evolutionary history of *Parthenocissus* with reticulation events between and within lineages. *Parthenocissus* is estimated to have originated from East Asia during the Early Eocene Climatic Optimum (Node 1 in Figs 5, S2, S3). The Asian and North American clades of *Parthenocissus* split in the Eocene (Node 2 in Figs 5, S2, S3), but ABBA–BABA tests and extensive gene tree conflicts elucidate gene flow between the North American clade and the henryana or trifoliolate lineages (Fig. 4). This suggests that the ancestors of lineages in the North American clade and the Asian clade might have come into contact after their early divergence at the beginning of the first major glaciation of the Oligocene (Oi-1 Glaciation; Figs 5, S12). Both the North Atlantic land bridges (Tiffney, 1985) and the Bering land bridge (Wen *et al.*, 2016) might have played important roles in intercontinental dispersal and range expansion of *Parthenocissus* between Asia and North America. Fossils similar to extant *Parthenocissus* species have been recorded from both Europe and western North America (Fig. 5; Tiffney & Manchester, 2001; Morley, 2003), suggesting climate-driven extinctions in these regions might have led to the intercontinental disjunct pattern between Asia and primarily eastern to southwestern North America in *Parthenocissus*. However, it is worth noting that the extinctions in North America and Europe may bias the ancestral area reconstruction of *Parthenocissus*.

In the North American clade, the plastid phylogeny suggests a sister relationship between *P. vitacea* and *P. quinquefolia* (Fig. 1a) with overlapping distributions throughout North America, whereas *P. heptaphylla*, the sister group of *P. vitacea* according to the transcriptome dataset (Fig. 1b), is narrowly endemic to central Texas (Brizicky, 1965; Moore & Wen, 2016). Our results are concordant with the phenomenon of plastid capture, where plastid phylogeny is usually strongly correlated with geographic distribution (Rieseberg & Soltis, 1991; Whittemore & Schaal, 1991; Cristina Acosta & Premoli, 2010). The large divergence time gap in the North American clade based on plastid and nuclear data, with the plastid estimate much younger than that of the nuclear, also implies plastid capture as suggested by previous studies (Drew & Sytsma, 2013; Huang *et al.*, 2014). Plastid capture in the North American clade might have occurred in the

Pliocene after the diversification of the clade during the Miocene (Figs 5, S12). Furthermore, our SNAQ searches detect gene flow between *P. heptaphylla* and the ancestor of the North American clade (Fig. 2d–f). Although gene flow between the ancestor of a clade and its descendants might be an artifact due to limitations of SNAQ on detecting multiple introgressions (Karimi *et al.*, 2020), it could also be explained by a ‘ghost’ lineage (Norrell, 1993), diverged from the stem group that contributed genetic material to *P. heptaphylla*, which also leads to a ‘BABA’ pattern in the ABBA–BABA test (Durand *et al.*, 2011; Zheng & Janke, 2018). The ‘ghost’ lineage may have gone extinct during the aridification of the west and central North America that has occurred since the Eocene (Sheldon & Retallack, 2004), as evidenced by fossil seeds of *Parthenocissus* from these regions dating to the Eocene and Miocene (Tiffney & Manchester, 2001; Morley, 2003; Chen, 2009).

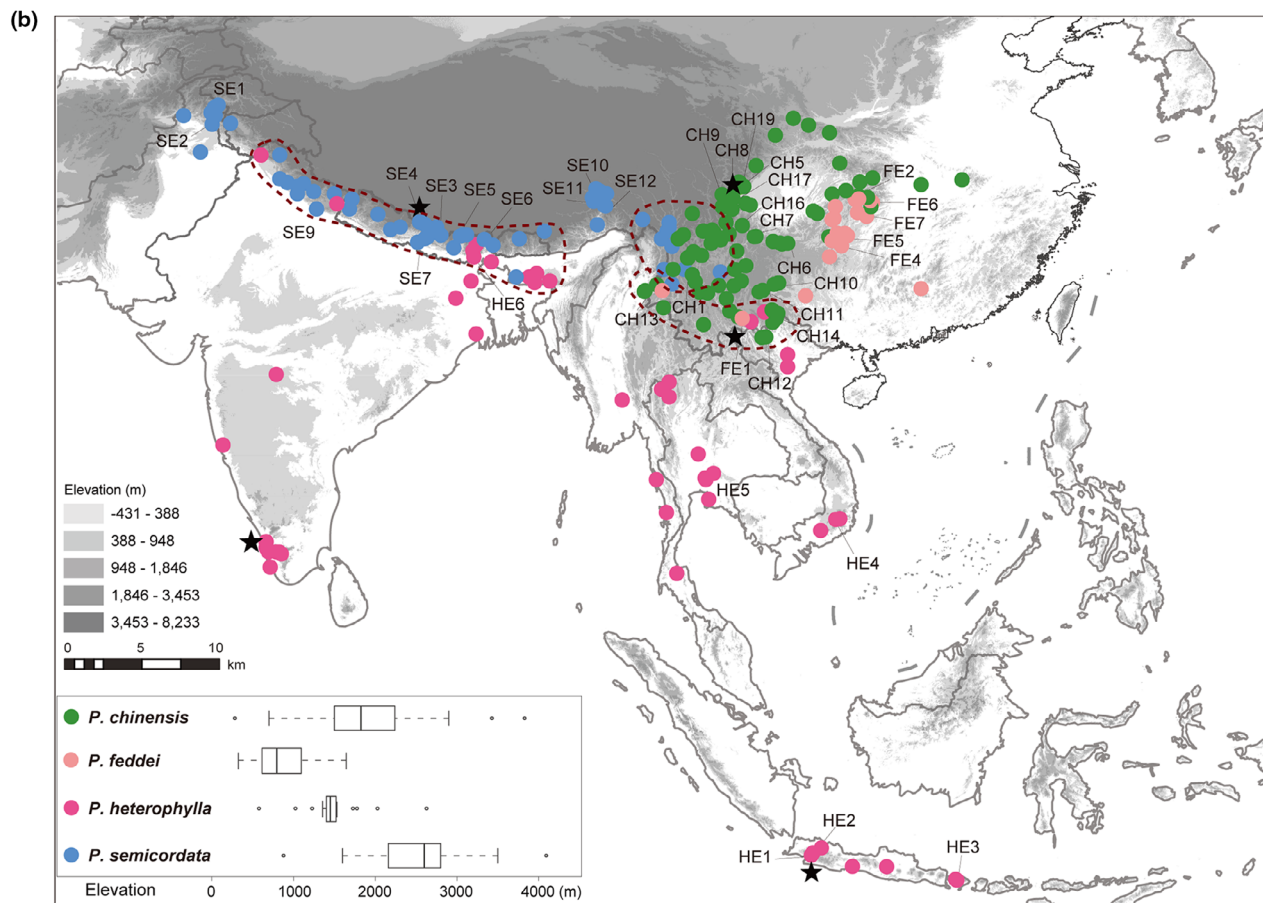
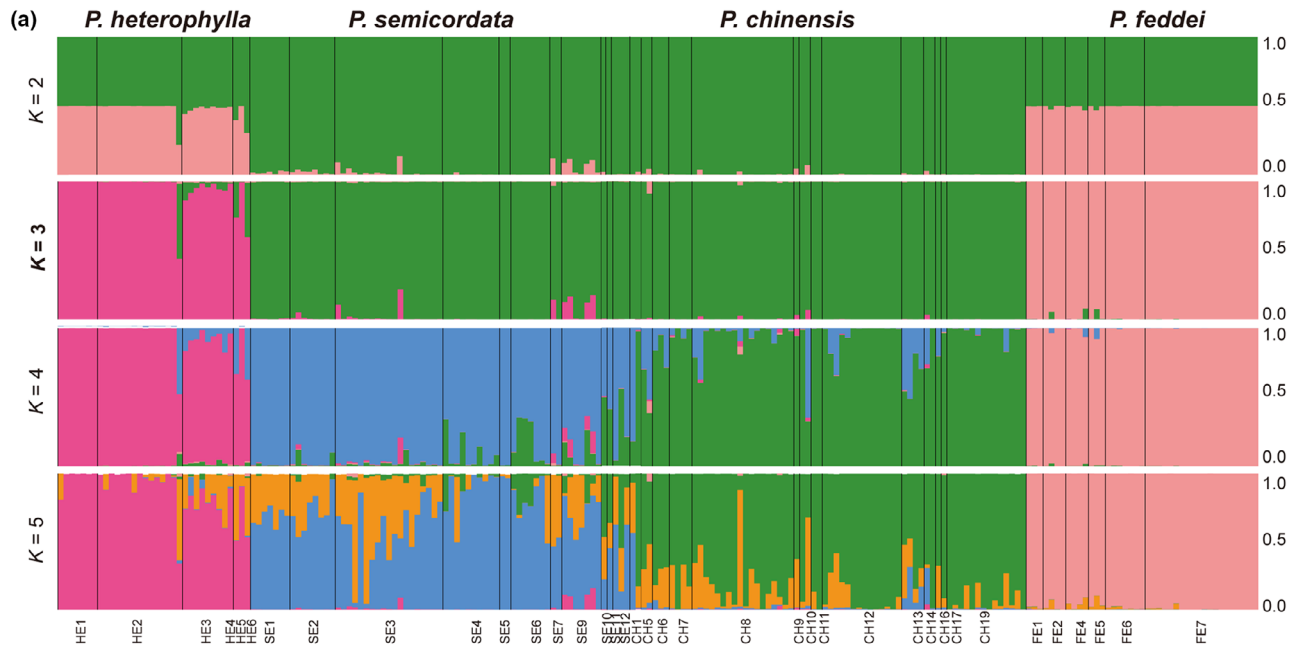
In the Asian clade, especially in the trifoliolate lineage, abundant gene flow across species/lineages has been detected by SNAQ searches and ABBA–BABA tests (Figs 2–4). As frequent hybridizations may obscure signals of network inference and rigorous statistical analyses (Karimi *et al.*, 2020), the evolutionary history of the trifoliolate lineage in the HHM region will be discussed below incorporating population-level evidence.

Hybridization hot spots and ongoing speciation reversal in the HHM region

The trifoliolate lineage is inferred to have split from the *dalzielii* lineage in East Asia during the Oi-1 Glaciation (Node 6 in Figs 5, S12). Diversification and habitat expansion within the trifoliolate lineage began during the Oligocene, with the warm-adapted *P. heterophylla* now distributed southward to Java through Indochina and the cold-adapted *P. semicordata* westward to the HHM (Figs 5, S12). Before the middle Miocene, most parts of the Hengduan Mountains have achieved their near-modern elevation (S. H. Li *et al.*, 2020; Spicer *et al.*, 2020). Erosion (e.g. river incision) in uplifted regions may have been enhanced by the intensification of the East Asian monsoon and Indian monsoon in the warming Miocene (Nie *et al.*, 2018), which caused greater heterogeneity in geology, elevation, and soil (Antonelli *et al.*, 2018). The emergence of abundant new ecological niches might have facilitated diversification of the trifoliolate lineage in the HHM.

The STRUCTURE results and the population-level phylogenies identify at least three potential hybridization hot spots around the HHM where multiple species co-occur (Figs 6, S14). The negative correlation between introgression intensity and

Fig. 6 Population structure and biogeographic distribution of the four species, *Parthenocissus chinensis* (CH), *P. feddei* (FE), *P. heterophylla* (HE), and *P. semicordata* (SE), in the trifoliolate lineage. (a) Bayesian clustering of population structure analysis in STRUCTURE based on a dataset of 1448 SNPs for 37 populations of the trifoliolate lineage with $K = 2–5$ (optimal $K = 3$ shown in bold). Each color represents one ancestral population, and each bar represents a single accession. The length of each colored segment in the bar represents the admixture proportion contributed by that ancestral population. (b) The distributions of the four species in the trifoliolate lineage based on 809 occurrence records. Population codes are provided in Table S3. Black stars indicate the localities of individuals with sequenced transcriptomes. Three potential hybridization hot spots are indicated with red dashed lines. Boxplots on the bottom left show the elevation range of each species, denoting the median (vertical line), 25th and 75th percentiles (left and right bounds of the box), and limits of the 95% minimum or maximum values (lower and upper whiskers), with outliers shown as dots.



geographic distance also reveals a spatial concentration of gene flow and therefore further supports the hybridization hot spots we propose (Fig. S13).

One hybridization hot spot involves the Himalaya, with great differences in elevation, where *P. semicordata* and *P. heterophylla* co-occur (Fig. 6b). SNAQ searches and ABBA–BABA tests strongly support gene flow between *P. semicordata* and *P. heterophylla*. However, it is difficult to determine the direction of gene flow, as subsets with different outgroups suggest opposite directions (Fig. 3) probably due to sensitivity of SNAQ to taxon sampling (Karimi *et al.*, 2020). We therefore recommend being cautious when selecting outgroup taxa for network inference. However, the admixed genomes of both species in the STRUCTURE analysis (Fig. 6a) tend to support bidirectional gene flow with stronger signals from *P. semicordata* to *P. heterophylla*. The phylogeny based on the 212-plastome dataset also indicates introgressions between the two species (Fig. S14). Generally, *P. semicordata* occurs in the HHM at an average elevation of *c.* 2500 m, but the lower altitude in the southern edge of the Himalaya allows some populations of *P. semicordata* to hybridize with *P. heterophylla*, which occupies a warmer zone with an average elevation of *c.* 1600 m (Fig. 6b).

The second hybridization hot spot is located at the southern border of the Hengduan Mountains where *P. chinensis* and *P. heterophylla* co-occur (Fig. 6b). Two populations of *P. chinensis* from Yingjiang and Malipo, Yunnan, China (CH13 and CH14, near the distribution range of *P. heterophylla* in Vietnam and Myanmar), were nested in the *P. heterophylla* clade in the plastid phylogeny, indicating hybridizations at low frequency (Fig. S14). This introgression was not detected by SNAQ or ABBA–BABA tests, possibly due to plastid capture or sampling bias. Notably, our population-level analyses do not support any introgressions involving *P. feddei* (Figs 6a, S14) as the SNAQ searches and ABBA–BABA tests suggest (Figs 2, 4), although some populations of *P. feddei* co-occur with *P. chinensis* and *P. heterophylla* in this area. The introgressions involving *P. feddei* might be an artifact caused by gene flow forming intersect loops on one quartet (Karimi *et al.*, 2020). The unique limestone habitat of *P. feddei* may have played a role in minimizing gene flow with other species. The lower elevation and less-complex topographic heterogeneity in this area may explain a stronger correlation between introgression intensity and geographic distance in *P. chinensis*–*P. heterophylla* than other species pairs (Fig. S13b).

The third hybridization hot spot is located at the southern edge of the HHM with a complex topography where *P. semicordata* and *P. chinensis* co-occur (Fig. 6b). Although SNAQ analysis and ABBA–BABA tests struggle to detect introgressions between sister groups, the population structure analyses and the population-level phylogenies show admixture of the two species (Figs 6, S14). Our population structure analysis clusters *P. semicordata* and *P. chinensis* together when employing the optimal $K=3$, suggesting highly admixed genomes of the two species (Fig. 6). When $K=4$, genetic differences between the two species, involving populations from a wide range, increase along with geographic distances (Figs 6, S13), resulting in a classical

clinal pattern commonly caused by frequent secondary contact between diverged lineages (Gompert & Buerkle, 2016). Notably, most populations of *P. chinensis* have received genetic contributions from *P. semicordata* and *vice versa* (Fig. 6), leading to the ongoing replacement of the two pure genomes with hybridized ones. This geographic structure might be a consequence of random interbreeding for generations without strong selection against hybrids (Todesco *et al.*, 2016). The conflicting gene trees further suggest mosaic genomes of the two species, with two dominant topologies, respectively, supported by plastomes and reflecting introgressions (Fig. S7). Recent speciation with gene flow may also cause admixed genomes and conflicting topologies, but is less likely here, as our plastome data show that *P. semicordata* and *P. chinensis* are distinct lineages persisting through geological time (Fig. S2). If plastid capture caused such patterns, we would expect historical gene flow between *P. semicordata* and species outside the trifoliolate lineage, or between *P. chinensis* and *P. heterophylla*, and *P. chinensis* and *P. feddei*. However, these introgressions are not consistently supported, and evidence for this hypothesis remains lacking. To summarize, the wide hybridization zone, tendency to replace parental lineages, and mosaic genomes of *P. semicordata* and *P. chinensis* all support possible ongoing speciation reversal between *P. semicordata* and *P. chinensis* (Jacobsen & Omland, 2011; Kearns *et al.*, 2018).

The ongoing speciation reversal unveiled by this study is most likely driven by geographic and climatic changes via natural processes, that is, the active paleoenvironmental changes in the HHM since the late Miocene (S. H. Li *et al.*, 2020; Spicer *et al.*, 2020), although human factors cannot be excluded. While speciation reversal can cause biodiversity decline (Seehausen, 2006; Zhang *et al.*, 2019), the hybrid derivative lineages can adapt to a changing heterogeneous environment as new taxa with recombinant genomes and even exhibit an advantage over their parental lineages when facing environmental change (von-Holdt *et al.*, 2016; Edelman *et al.*, 2019). With the surviving genomic fractions in the gene pool of extant species, the fused lineages can potentially diverge into multiple species under certain circumstances. In the HHM region, besides species diversification that has been intensively studied, the largely underestimated speciation reversal processes might also have played an important role in the formation of biodiversity cradles. Therefore, further consideration of protecting hybrid populations and their living habitats is required to sustain the potential of generating future biodiversity in response to accelerated climatic changes.

Acknowledgements

We are grateful to Jianfei Ye, Bing Liu, Zhangjian Shan, Viet-Cuong Dang, Jingbo Zhang, Tuo Yang, Miao Sun, Anming Lu, Haihua Hu, Van-Du Nguyen, Viet-Hung Dang, Xintang Ma, Qiang Wang, Shengxiang Yu, Yong Yang and Xiaoya Yu for field work assistance and sample collection, Meixia Wang, Yuchang Yang, Romer Narindra Rabarijaona, Zhuzhi Zhang, Zhe Cai, and Yanlei Feng for experiment assistance and data analyses, and the staff at the herbaria of A, BM, CAL, E, HN, IBSC, K, KUN, L, LE, MO, NY, P, PE, PH, U and US for their

help and for the loan of or access to specimens. We thank Jana Vamosi, AJ Harris, and three anonymous reviewers for their constructive comments. This work was supported by the Strategic Priority Research Program of the Chinese Academy of Sciences (XDB31000000, XDA19050103); the National Natural Science Foundation of China (31870197, 32122009, 32200177); the International Partnership Program of CAS (151853KYSB2019 0027); the Youth Innovation Promotion Association CAS (2020080); the Sino-Africa Joint Research Center, CAS International Research and Education Development Program (SAJC202101); the Biodiversity Survey and Assessment Project of the Ministry of Ecology and Environment, China (2019 HJ2096001006); the Portuguese Foundation for Science and Technology (FCT; UIDB/04326/2020, UIDP/04326/2020, and LA/P/0101/2020 to CJC); and the K.C. Wong Education Foundation (GJTD-2020-05).











Competing interests

None declared.

Author contributions

ZC and LL conceived the project and designed the research. ZC, LL, JW and AT-B collected the samples. YN and JG performed the laboratory work. JY, YN and YY conducted the analyses. JY, YN and LL drafted the manuscript. YY, CJC, RLB, JG, JW and ZC revised the manuscript. All authors read and improved the manuscript. JY and YN contributed equally to this work.

ORCID

Russell L. Barrett  <https://orcid.org/0000-0003-0360-8321>
 Zhidian Chen  <https://orcid.org/0000-0002-7716-4834>
 Cymon J. Cox  <https://orcid.org/0000-0002-4927-979X>
 Jing Guo  <https://orcid.org/0000-0003-3689-6374>
 Limin Lu  <https://orcid.org/0000-0001-6973-7877>
 Yanting Niu  <https://orcid.org/0000-0001-8548-4742>
 Anna Trias-Blasi  <https://orcid.org/0000-0001-9745-3222>
 Jun Wen  <https://orcid.org/0000-0001-6353-522X>
 Yichen You  <https://orcid.org/0000-0002-4004-1902>
 Jinren Yu  <https://orcid.org/0000-0002-8157-3673>

Data availability

The data that support the findings of this study are openly available in the NCBI Sequence Read Archive, BioProject PRJNA888447 (BioSamples SAMN31217733–SAMN31217746, SAMN31329148–SAMN31329149, and SAMN31359324–SAMN31359537).

References

Abbott R, Albach D, Ansell S, Arntzen JW, Baird SJE, Bierne N, Boughman J, Brelsford A, Buerkle CA, Buggs R *et al.* 2013. Hybridization and speciation. *Journal of Evolutionary Biology* 26: 229–246.

- Andersen MJ, McCullough JM, Gyllenhaal EF, Mapel XM, Haryoko T, Jönsson KA, Joseph L. 2021. Complex histories of gene flow and a mitochondrial capture event in a nonsister pair of birds. *Molecular Ecology* 30: 2087–2103.
- Antonelli A, Kissling WD, Flantua SGA, Bermúdez MA, Mulch A, Muellner-Riehl AN, Kreft H, Linder HP, Badgley C, Fjeldså J *et al.* 2018. Geological and climatic influences on mountain biodiversity. *Nature Geoscience* 11: 718–725.
- Ané C, Larget B, Baum DA, Smith SD, Rokas A. 2007. Bayesian estimation of concordance among gene trees. *Molecular Biology and Evolution* 24: 412–426.
- Arcila D, Ortí G, Vari R, Armbruster JW, Stiassny MLJ, Ko KD, Sabaj MH, Lundberg J, Revell LJ, Betancur RR. 2017. Genome-wide interrogation advances resolution of recalcitrant groups in the tree of life. *Nature Ecology and Evolution* 1: 0020.
- Barraclough TG, Nee S. 2001. Phylogenetics and speciation. *Trends in Ecology & Evolution* 16: 391–399.
- Benjamini Y, Hochberg Y. 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society: Series B (Methodological)* 57: 289–300.
- Blackmon H, Adams RA. 2015. EVOBIR: tools for comparative analyses and teaching evolutionary biology. *Zenodo*. doi: [10.5281/zenodo.30938](https://doi.org/10.5281/zenodo.30938).
- Blair C, Ané C. 2020. Phylogenetic trees and networks can serve as powerful and complementary approaches for analysis of genomic data. *Systematic Biology* 69: 593–601.
- Bolger AM, Lohse M, Usadel B. 2014. TRIMMOMATIC: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30: 2114–2120.
- Brizicky GK. 1965. The genera of Vitaceae in the southeastern United States. *Journal of the Arnold Arboretum* 46: 48–67.
- Budenhagen C, Lemmon AR, Lemmon EM, Bruhl J, Cappa J, Clement WL, Donoghue M, Edwards EJ, Hipp AL, Kortyna M *et al.* 2016. Anchored phylogenomics of angiosperms I: assessing the robustness of phylogenetic estimates. *bioRxiv*. doi: [10.1101/086298](https://doi.org/10.1101/086298).
- Cai LM, Xi ZX, Lemmon EM, Lemmon AR, Mast A, Buddenhagen CE, Liu L, Davis CC. 2021. The perfect storm: gene tree estimation error, incomplete lineage sorting, and ancient gene flow explain the most recalcitrant ancient angiosperm clade, Malpighiales. *Systematic Biology* 70: 491–507.
- Capella-Gutiérrez S, Silla-Martínez JM, Gabaldón T. 2009. TRIMAL: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics* 25: 1972–1973.
- Chen IJ, Manchester SR. 2007. Seed morphology of modern and fossil *Ampelocissus* (Vitaceae) and implications for phytogeography. *American Journal of Botany* 94: 1534–1553.
- Chen IJ. 2009. *Biogeographic history of Vitaceae inferred from morphological phylogeny and the fossil record of seeds*. PhD thesis, University of Florida, Gainesville, FL, USA.
- Chen ZD, Ren H, Wen J. 2007. Vitaceae. In: Wu ZY, Hong DY, Raven PH, eds. *Flora of China*. Beijing, China; St Louis, MO, USA: Science Press; Missouri Botanical Garden, 173–222.
- Chernomor O, Von Haeseler A, Minh BQ. 2016. Terrace aware data structure for phylogenomic inference from supermatrices. *Systematic Biology* 65: 997–1008.
- Chifman J, Kubatko L. 2014. Quartet inference from SNP data under the coalescent model. *Bioinformatics* 30: 3317–3324.
- Cristina Acosta M, Premoli AC. 2010. Evidence of chloroplast capture in South American *Nothofagus* (subgenus *Nothofagus*, Nothofagaceae). *Molecular Phylogenetics and Evolution* 54: 235–242.
- Dong SS, Wang YL, Xia NH, Liu Y, Liu M, Lian L, Li N, Li LF, Lang XA, Gong YQ *et al.* 2021. Plastid and nuclear phylogenomic incongruences and biogeographic implications of *Magnolia* s.l. (Magnoliaceae). *Journal of Systematics and Evolution* 60: 1–15.
- Drew BT, Sytsma KJ. 2013. The South American radiation of *Lepechinia* (Lamiaceae): phylogenetics, divergence times and evolution of dioecy. *Botanical Journal of the Linnean Society* 171: 171–190.
- Durand EY, Patterson N, Reich D, Slatkin M. 2011. Testing for ancient admixture between closely related populations. *Molecular Biology and Evolution* 28: 2239–2252.

- Earl DA, VonHoldt BM. 2012. STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources* 4: 359–361.
- Edelman NB, Frandsen PB, Miyagi M, Clavijo B, Davey J, Dikow RB, García-acinelli G, Van Belleghem SM, Patterson N, Neafsey DE *et al.* 2019. Genomic architecture and introgression shape a butterfly radiation. *Science* 366: 594–599.
- Emms DM, Kelly S. 2015. ORTHOFINDER: solving fundamental biases in whole genome comparisons dramatically improves orthogroup inference accuracy. *Genome Biology* 16: 157.
- Evanno G, Regnaut S, Goudet J. 2005. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology* 14: 2611–2620.
- Favre A, Päckert M, Pauls SU, Jähmig SC, Uhl D, Michalak I, Muellner-Riehl AN. 2015. The role of the uplift of the Qinghai–Tibetan Plateau for the evolution of Tibetan biotas. *Biological Reviews* 90: 236–253.
- Ferreira MS, Jones MR, Callahan CM, Farello L, Tolesa Z, Suchentrunk F, Boursot P, Mills LS, Alves PC, Good JM *et al.* 2021. The legacy of recurrent introgression during the radiation of hares. *Systematic Biology* 70: 593–607.
- Gao J, Wang BH, Mao JF, Ingvarsson P, Zeng QY, Wang XR. 2012. Demography and speciation history of the homoploid hybrid pine *Pinus densata* on the Tibetan Plateau. *Molecular Ecology* 21: 4811–4827.
- GBIF.ORG. 2016. GBIF occurrence download. doi: [10.15468/dl.5tcbur](https://doi.org/10.15468/dl.5tcbur).
- Gilman RT, Behm JE. 2011. Hybridization, species collapse, and species reemergence after disturbance to premating mechanisms of reproductive isolation. *Evolution* 65: 2592–2605.
- Gompert Z, Buerkle CA. 2016. What, if anything, are hybrids: enduring truths and challenges associated with population structure and gene flow. *Evolutionary Applications* 9: 909–923.
- Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, Adiconis X, Fan L, Raychowdhury R, Zeng QD *et al.* 2011. Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nature Biotechnology* 29: 644–652.
- Green RE, Krause J, Briggs AW, Maricic T, Stenzel U, Kircher M, Patterson N, Li H, Zhai WW, Fritz MHY *et al.* 2010. A draft sequence of the Neandertal genome. *Science* 328: 710–722.
- Haas BJ, Papanicolaou A, Yassour M, Grabherr M, Blood PD, Bowden J, Couger MB, Eccles D, Li B, Lieber M *et al.* 2013. *De novo* transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. *Nature Protocols* 8: 1494–1512.
- Harrell FE Jr. 2021. *Hmisc: Harrell miscellaneous*. R package v.4.6-0. [WWW document] URL <https://CRAN.R-project.org/package=Hmisc> [accessed 24 September 2022].
- Hejase HA, Liu KJ. 2016. A scalability study of phylogenetic network inference methods using empirical datasets and simulations involving a single reticulation. *BMC Bioinformatics* 17: 422.
- Hoang DT, Chernomor O, Von Haeseler A, Minh BQ, Vinh LS. 2018. UFBOOT2: improving the ultrafast bootstrap approximation. *Molecular Biology and Evolution* 35: 518–522.
- Hu HH, Ye JF, Liu B, Mao LF, Smith SA, Barrett RL, Soltis PS, Soltis DE, Chen ZD, Lu LM. 2022. Temporal and spatial comparisons of angiosperm diversity between eastern Asia and North America. *National Science Review* 9: nwab199.
- Huang DI, Hefer CA, Kolosova N, Douglas CJ, Cronk QCB. 2014. Whole plastome sequencing reveals deep plastid divergence and cytonuclear discordance between closely related balsam poplars, *Populus balsamifera* and *P. trichocarpa* (Salicaceae). *New Phytologist* 204: 693–703.
- Huson DH, Bryant D. 2006. Application of phylogenetic networks in evolutionary studies. *Molecular Biology and Evolution* 23: 254–267.
- Jacobsen F, Omland KE. 2011. Increasing evidence of the role of gene flow in animal evolution: hybrid speciation in the yellow-rumped warbler complex. *Molecular Ecology* 20: 2236–2239.
- Jansen RK, Kaitanis C, Saski C, Lee S-B, Tomkins J, Alverson AJ, Daniell H. 2006. Phylogenetic analyses of *Vitis* (Vitaceae) based on complete chloroplast genome sequences: effects of taxon sampling and phylogenetic methods on resolving relationships among rosids. *BMC Evolutionary Biology* 6: 32.
- Jiang XL, An M, Zheng SS, Deng M, Su ZH. 2018. Geographical isolation and environmental heterogeneity contribute to the spatial genetic patterns of *Quercus kerrii* (Fagaceae). *Heredity* 120: 219–233.
- Johnson MG, Gardner EM, Liu Y, Medina R, Goffinet B, Shaw AJ, Zerega NJC, Wickett NJ. 2016. HYBPIPER: extracting coding sequence and introns for phylogenetics from high-throughput sequencing reads using target enrichment. *Applications in Plant Sciences* 4: 1600016.
- Kalyaanamoorthy S, Minh BQ, Wong TKF, Von Haeseler A, Jermini LS. 2017. MODELFINDER: fast model selection for accurate phylogenetic estimates. *Nature Methods* 14: 587–589.
- Karimi N, Grover CE, Gallagher JP, Wendel JF, Ané C, Baum DA. 2020. Reticulate evolution helps explain apparent homoplasy in floral biology and pollination in baobabs (*Adansonia*; Bombacoideae; Malvaceae). *Systematic Biology* 69: 462–478.
- Katoh K, Standley DM. 2013. MAFFT multiple sequence alignment software v.7: improvements in performance and usability. *Molecular Biology and Evolution* 30: 772–780.
- Keams AM, Restani M, Szabo I, Schröder-Nielsen A, Kim JA, Richardson HM, Marzluff JM, Fleischer RC, Johnsen A, Omland KE. 2018. Genomic evidence of speciation reversal in ravens. *Nature Communications* 9: 906.
- Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, Buxton S, Cooper A, Markowitz S, Duran C *et al.* 2012. GENEIOUS BASIC: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* 28: 1647–1649.
- Larget BR, Kotha SK, Dewey CN, Ané C. 2010. BUCKY: gene tree/species tree reconciliation with Bayesian concordance analysis. *Bioinformatics* 26: 2910–2911.
- Li JL, Milne RI, Ru DF, Miao JB, Tao WJ, Zhang L, Xu JJ, Liu JQ, Mao KS. 2020. Allopatric divergence and hybridization within *Cupressus chengiana* (Cupressaceae), a threatened conifer in the northern Hengduan Mountains of western China. *Molecular Ecology* 29: 1250–1266.
- Li SH, Ji XP, Harrison T, Deng CL, Wang SQ, Wang LR, Zhu RX. 2020. Uplift of the Hengduan Mountains on the southeastern margin of the Tibetan Plateau in the late Miocene and its paleoenvironmental impact on hominoid diversity. *Palaeogeography, Palaeoclimatology, Palaeoecology* 553: 109794.
- Li WZ, Godzik A. 2006. CD-HIT: a fast program for clustering and comparing large sets of protein or nucleotide sequences. *Bioinformatics* 22: 1658–1659.
- Linder CR, Rieseberg LH. 2004. Reconstructing patterns of reticulate evolution in plants. *American Journal of Botany* 91: 1700–1708.
- Liu J, Möller M, Provan J, Gao LM, Poudel RC, Li DZ. 2013. Geological and ecological factors drive cryptic speciation of yews in a biodiversity hotspot. *New Phytologist* 199: 1093–1108.
- Liu L, Yu LL. 2010. PHYBASE: an R package for species tree analysis. *Bioinformatics* 26: 962–963.
- Lu LM, Cox CJ, Mathews S, Wang W, Wen J, Chen ZD. 2018. Optimal data partitioning, multispecies coalescent and Bayesian concordance analyses resolve early divergences of the grape family (Vitaceae). *Cladistics* 34: 57–77.
- Lu LM, Wen J, Chen ZD. 2012. A combined morphological and molecular phylogenetic analysis of *Parthenocissus* (Vitaceae) and taxonomic implications. *Botanical Journal of the Linnean Society* 168: 43–63.
- Ma ZY, Nie ZL, Ren C, Liu XQ, Zimmer EA, Wen J. 2021. Phylogenomic relationships and character evolution of the grape family (Vitaceae). *Molecular Phylogenetics and Evolution* 154: 106948.
- Manish K, Pandit MK. 2018. Geophysical upheavals and evolutionary diversification of plant species in the Himalaya. *PeerJ* 6: e5919.
- Martin SH, Davey JW, Jiggins CD. 2015. Evaluating the use of ABBA–BABA statistics to locate introgressed loci. *Molecular Biology and Evolution* 32: 244–257.
- Meng HH, Su T, Gao XY, Li J, Jiang XL, Sun H, Zhou ZK. 2017. Warm–cold colonization: response of oaks to uplift of the Himalaya–Hengduan Mountains. *Molecular Ecology* 26: 3276–3294.
- Mirarab S, Reaz R, Bayzid MS, Zimmermann T, Swenson SM, Warnow T. 2014. ASTRAL: genome-scale coalescent-based species tree estimation. *Bioinformatics* 30: 541–548.
- Moore MO, Wen J. 2016. Vitaceae. In: Flora of North America Editorial Committee, ed. *Flora of North America, North of Mexico, Magnoliophyta: Vitaceae to Garryaceae*. New York, NY, USA: Oxford University Press, 3–23.

- Morales-Briones DF, Liston A, Tank DC. 2018. Phylogenomic analyses reveal a deep history of hybridization and polyploidy in the Neotropical genus *Lachemilla* (Rosaceae). *New Phytologist* 218: 1668–1684.
- Morley RJ. 2003. Interplate dispersal paths for megathermal angiosperms. *Perspectives in Plant Ecology, Evolution and Systematics* 6: 5–20.
- Nguyen LT, Schmidt HA, Von Haeseler A, Minh BQ. 2015. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Molecular Biology and Evolution* 32: 268–274.
- Nie JS, Ruetenik G, Gallagher K, Hoke G, Garzzone CN, Wang WT, Stockli D, Hu XF, Wang Z, Wang Y *et al.* 2018. Rapid incision of the Mekong River in the middle Miocene linked to monsoonal precipitation. *Nature Geoscience* 11: 944–948.
- Nie ZL, Sun H, Chen ZD, Meng Y, Manchester SR, Wen J. 2010. Molecular phylogeny and biogeographic diversification of *Parthenocissis* (Vitaceae) disjunct between Asia and North America. *American Journal of Botany* 97: 1342–1353.
- Niu YT, Jabbar F, Barrett RL, Ye JF, Zhang ZZ, Lu KQ, Lu LM, Chen ZD. 2018. Combining complete chloroplast genome sequences with target loci data and morphology to resolve species limits in *Triplostegia* (Caprifoliaceae). *Molecular Phylogenetics and Evolution* 129: 15–26.
- Norell MA. 1993. Tree-based approaches to understanding history: comments on ranks, rules, and the quality of the fossil record. *American Journal of Science* 293 A: 407–417.
- Nylander JAA. 2004. *MrMODELTEST, h.2. Program distributed by the author.* Uppsala, Sweden: Evolutionary Biology Centre, Uppsala University. [WWW document] URL <https://github.com/nylander/MrModeltest2> [accessed 16 July 2022].
- Okuyama Y, Fujii N, Wakabayashi M, Kawakita A, Ito M, Watanabe M, Murakami N, Kato M. 2005. Nonuniform concerted evolution and chloroplast capture: heterogeneity of observed introgression patterns in three molecular data partition phylogenies of Asian *Mitella* (Saxifragaceae). *Molecular Biology and Evolution* 22: 285–296.
- Page AJ, Taylor B, Delaney AJ, Soares J, Seemann T, Keane JA, Harris SR. 2016. SNP-SITES: rapid efficient extraction of SNPs from multi-FASTA alignments. *Microbial Genomics* 2: e000056.
- Pavón-Vázquez CJ, Brennan IG, Keogh JS. 2021. A comprehensive approach to detect hybridization sheds light on the evolution of earth's largest lizards. *Systematic Biology* 70: 877–890.
- Payseur BA, Rieseberg LH. 2016. A genomic perspective on hybridization and speciation. *Molecular Ecology* 25: 2337–2360.
- Pritchard JK, Stephens M, Donnelly P. 2000. Inference of population structure using multilocus genotype data. *Genetics* 155: 945–959.
- Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MAR, Bender D, Maller J, Sklar P, De Bakker PIW, Daly MJ *et al.* 2007. PLINK: a tool set for whole-genome association and population-based linkage analyses. *American Journal of Human Genetics* 81: 559–575.
- Rieseberg LH, Soltis DE. 1991. Phylogenetic consequences of cytoplasmic. *Evolutionary Trends in Plants* 5: 65–84.
- Rieseberg LH, Wendel JF. 1993. Introgression and its consequences in plants. In: Harrison RG, ed. *Hybrid zones and the evolutionary process*. New York, NY, USA: Oxford University Press, 70–100.
- Robinson DF, Foulds LR. 1981. Comparison of phylogenetic trees. *Mathematical Biosciences* 53: 131–147.
- Ronquist F, Teslenko M, Van Der Mark P, Ayres DL, Darling A, Höhna S, Larget B, Liu L, Suchard MA, Huelsenbeck JP. 2012. MRBAYES 3.2: efficient bayesian phylogenetic inference and model choice across a large model space. *Systematic Biology* 61: 539–542.
- Rose JP, Toledo CAP, Lemmon EM, Lemmon AR, Sytsma KJ. 2021. Out of sight, out of mind: widespread nuclear and plastid-nuclear discordance in the flowering plant genus *Polemonium* (Polemoniaceae) suggests widespread historical gene flow despite limited nuclear signal. *Systematic Biology* 70: 162–180.
- Ru DF, Sun YS, Wang DL, Chen Y, Wang TJ, Hu QJ, Abbott RJ, Liu JQ. 2018. Population genomic analysis reveals that homoploid hybrid speciation can be a lengthy process. *Molecular Ecology* 27: 4875–4887.
- Salichos L, Stamatakis A, Rokas A. 2014. Novel information theory-based measures for quantifying incongruence among phylogenetic trees. *Molecular Biology and Evolution* 31: 1261–1271.
- Seehausen O, Takimoto G, Roy D, Jokela J. 2008. Speciation reversal and biodiversity dynamics with hybridization in changing environments. *Molecular Ecology* 17: 30–44.
- Seehausen O. 2006. Conservation: losing biodiversity by reverse speciation. *Current Biology* 16: 334–337.
- Sheldon ND, Retallack GJ. 2004. Regional paleoprecipitation records from the Late Eocene and Oligocene of North America. *Journal of Geology* 112: 487–494.
- Smith SA, Moore MJ, Brown JW, Yang Y. 2015. Analysis of phylogenomic datasets reveals conflict, concordance, and gene duplications with examples from animals and plants. *BMC Evolutionary Biology* 15: 1–15.
- Soltis PS, Soltis DE. 2009. The role of hybridization in plant speciation. *Annual Review of Plant Biology* 60: 561–588.
- Solís-Lemus C, Ané C. 2016. Inferring phylogenetic networks with maximum pseudolikelihood under incomplete lineage sorting. *PLoS Genetics* 12: 1–21.
- Solís-Lemus C, Bastide P, Ané C. 2017. PHYLONETWORKS: a package for phylogenetic networks. *Molecular Biology and Evolution* 34: 3292–3298.
- Spicer RA, Farnsworth A, Su T. 2020. Cenozoic topography, monsoons and biodiversity conservation within the Tibetan Region: an evolving story. *Plant Diversity* 42: 229–254.
- Stamatakis A. 2014. RAXML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* 30: 1312–1313.
- Stenz NWM, Larget B, Baum DA, Ané C. 2015. Exploring tree-like and non-tree-like patterns using genome sequences: an example using the inbreeding plant species *Arabidopsis thaliana* (L.) Heynh. *Systematic Biology* 64: 809–823.
- Swofford DL. 2002. *PAUP*. Phylogenetic analysis using parsimony (* and other methods)*, v.4. Sunderland, MA, USA: Sinauer Associates.
- Thiers B. 2017. *Index Herbariorum a global directory of public herbaria and associated staff* New York Botanical Garden's Virtual Herbarium. [WWW document] URL <http://sweetgum.nybg.org/ih> [accessed 24 May 2019].
- Tiffney BH, Manchester SR. 2001. The use of geological and paleontological evidence in evaluating plant phylogeographic hypotheses in the Northern Hemisphere Tertiary. *International Journal of Plant Sciences* 162: S3–S17.
- Tiffney BH. 1985. The Eocene North Atlantic land bridge: its importance in Tertiary and modern phylogeography of the Northern Hemisphere. *Journal of the Arnold Arboretum* 66: 243–273.
- Todesco M, Pascual MA, Owens GL, Ostevik KL, Moyers BT, Hübner S, Heredia SM, Hahn MA, Caseys C, Bock DG *et al.* 2016. Hybridization and extinction. *Evolutionary Applications* 9: 892–908.
- vonHoldt BM, Kays R, Pollinger JP, Wayne RK. 2016. Admixture mapping identifies introgressed genomic regions in North American canids. *Molecular Ecology* 25: 2443–2453.
- Wen J, Nie ZL, Ickert-Bond SM. 2016. Intercontinental disjunctions between eastern Asia and western North America in vascular plants highlight the biogeographic importance of the Bering land bridge from late Cretaceous to Neogene. *Journal of Systematics and Evolution* 54: 469–490.
- Wen J, Nie ZL, Soejima A, Meng Y. 2007. Phylogeny of Vitaceae based on the nuclear *GAI1* gene sequences. *Canadian Journal of Botany* 85: 731–745.
- Wen J, Zhang JQ, Nie ZL, Zhong Y, Sun H. 2014. Evolutionary diversifications of plants on the Qinghai–Tibetan Plateau. *Frontiers in Genetics* 5: 4.
- Westerhold T, Marwan N, Drury AJ, Liebrand D, Agnini C, Anagnostou E, Barnett JSK, Bohaty SM, De Vleeschouwer D, Florindo F *et al.* 2020. An astronomically dated record of earth's climate and its predictability over the last 66 million years. *Science* 369: 1383–1388.
- Whittemore AT, Schaal BA. 1991. Interspecific gene flow in sympatric oaks. *Proceedings of the National Academy of Sciences, USA* 88: 2540–2544.
- Wu SD, Wang Y, Wang ZF, Shrestha N, Liu JQ. 2022. Species divergence with gene flow and hybrid speciation on the Qinghai–Tibet Plateau. *New Phytologist* 234: 392–404.
- Wyman SK, Jansen RK, Boore JL. 2004. Automatic annotation of organellar genomes with DOGMA. *Bioinformatics* 20: 3252–3255.
- Yang ZH. 2007. PAML 4: phylogenetic analysis by maximum likelihood. *Molecular Biology and Evolution* 24: 1586–1591.
- Yu Y, Dong JR, Liu KJ, Nakhleh L. 2014. Maximum likelihood inference of reticulate evolutionary histories. *Proceedings of the National Academy of Sciences, USA* 111: 16448–16453.

- Yu Y, Harris AJ, Blair C, He XJ. 2015. RASP (Reconstruct Ancestral State in Phylogenies): a tool for historical biogeography. *Molecular Phylogenetics and Evolution* 87: 46–49.
- Yu Y, Nakhleh L. 2015. A maximum pseudo-likelihood approach for phylogenetic networks. *BMC Genomics* 16: 1–10.
- Zerbino DR, Birney E. 2008. VELVET: algorithms for *de novo* short read assembly using de Bruijn graphs. *Genome Research* 18: 821–829.
- Zhang C, Rabiee M, Sayyari E, Mirarab S. 2018. ASTRAL-III: polynomial time species tree reconstruction from partially resolved gene trees. *BMC Bioinformatics* 19: 15–30.
- Zhang L, Thibert-Plante X, Ripa J, Svanbäck R, Brännström Å. 2019. Biodiversity loss through speciation collapse: mechanisms, warning signals, and possible rescue. *Evolution* 73: 1504–1516.
- Zheng YC, Janke A. 2018. Gene flow analysis method, the *D*-statistic, is robust in a wide parameter space. *BMC Bioinformatics* 19: 10.

Supporting Information

Additional Supporting Information may be found online in the Supporting Information section at the end of the article.

Fig. S1 Phylogeny and distribution of *Parthenocissus* according to previous study.

Fig. S2 Chronogram of *Parthenocissus* inferred from MCMCTREE in the PAML package based on a dataset including 40 plastomes of Vitaceae.

Fig. S3 Chronogram of *Parthenocissus* inferred from MCMCTREE in the PAML package based on 15taxa-1101nu dataset.

Fig. S4 Posterior distributions and marginal prior distributions of the calibrated nodes in divergence time estimation based on plastid and nuclear gene datasets.

Fig. S5 Ten most frequently occurring topologies of 2587 rooted maximum likelihood gene trees.

Fig. S6 Ten most frequently occurring unrooted topologies of 2587 unrooted maximum likelihood gene trees.

Fig. S7 Topology types of 2587 maximum likelihood gene trees when focusing on the trifoliolate lineage.

Fig. S8 A supernetwork constructed with SPLITSTREE based on 2587 rooted maximum likelihood gene trees.

Fig. S9 Phylogenies generated by ASTRAL with gene trees constructed using maximum parsimony and Bayesian inference approach.

Fig. S10 Distribution of Robinson–Foulds distances calculated with simulated and empirical gene trees.

Fig. S11 Negative log pseudolikelihood (–logplik) score profiles obtained by SNAQ in PHYLONETWORKS.

Fig. S12 Ancestral area reconstructions for *Parthenocissus* in RASP using the MCMCTREE-derived chronogram based on plastome sequences.

Fig. S13 Pearson correlation analyses showing negative correlation between introgression intensity and geographic distance.

Fig. S14 Phylogenetic relationships within the trifoliolate lineage with two individuals of *Parthenocissus dalzielii* as the outgroup.

Methods S1 Detailed information of calibration strategies.

Methods S2 Strategies for retaining only natural distribution records and ensuring each species' distribution range.

Table S1 Detailed information for transcriptome and plastome sequencing species.

Table S2 Species and sources of the plastome and transcriptome data for divergence time estimation.

Table S3 Details of location and size of each population for the four species, *P. chinensis* (CH), *P. feddei* (FE), *P. heterophylla* (HE), and *P. semicordata* (SE), in the trifoliolate lineage.

Table S4 Datasets used and analyses conducted for this study.

Table S5 ABBA–BABA test results on quartets of the combinations of species from each Asian lineage or the North American clade, and combinations of three species within each lineage-subset.

Table S6 Maximum likelihood bootstrap values for single gene trees supporting the 10 most frequently occurring topologies.

Please note: Wiley is not responsible for the content or functionality of any Supporting Information supplied by the authors. Any queries (other than missing material) should be directed to the *New Phytologist* Central Office.