**Phylogenomic insights into the divergent history and hybridization within the East Asian endemic *Abelia* (Caprifoliaceae)**

Qing-Hui Sun[1], Diego F. Morales-Briones[2,3], Hong-Xin Wang[1,4], Jacob B. Landis[5,6] Jun Wen[7], Hua-Feng Wang[1]*

[1] *Key Laboratory of Tropical Biological Resources of Ministry of Education, College of Tropical Crops, Hainan University, Haikou 570228, China*

[2] *Department of Plant and Microbial Biology, College of Biological Sciences, University of Minnesota, 140 Gortner Laboratory, 1479 Gortner Avenue, Saint Paul, MN 55108, USA*

[3] *Systematics, Biodiversity and Evolution of Plants, Department of Biology I, Ludwig-Maximilians-Universität München, Menzinger Str. 67, 80638, Munich, Germany*

[4] *Zhai Mingguo Academician Work Station, Sanya University, Sanya 572022, China*

[5] *School of Integrative Plant Science, Section of Plant Biology and the L.H. Bailey Hortorium, Cornell University, Ithaca, NY 14850, USA*

[6] *BTI Computational Biology Center, Boyce Thompson Institute, Ithaca, NY 14853, USA*

[7] *Department of Botany, National Museum of Natural History, MRC-166, Smithsonian Institution, PO Box 37012, Washington, DC 20013-7012, USA*

*Authors for correspondence:

Hua-Feng Wang, E-mail: hfwang@hainanu.edu.cn

## Abstract:

• **Background and Aims** *Abelia* (Caprifoliaceae) is a small genus with six species (including one artificial hybrid). The genus has a disjunct distribution across mainland China, Taiwan, and the Ryukyu islands, providing a model system to explore mechanisms of species dispersal in the East Asian flora. However, the current phylogenetic relationships within *Abelia* remain uncertain.

• **Methods** In this study, we reconstructed phylogenetic relationships within *Abelia* using nuclear loci generated by target enrichment and plastomes from genome skimming. Divergence time estimation, ancestral area reconstruction, and ecological niche modelling (ENM) were used to examine the biogeographic diversification of *Abelia*.

• **Key Results** We found extensive cytonuclear discordance across the genus. Based on nuclear and plastid phylogenies we proposed to merge *A. macrotera* var. *zabelioides* into *A. uniflora*. Network analyses suggested that *A chinensis*, *A. macrotera* and *A. uniflora* are from complex hybrid origins. These hybridization events may explain the complicated taxonomy of the group. Our results show that *Abelia* originated in southwest China, with the nuclear haplotype network suggesting an ancestral haplotype from southwest and central China. The diversification of *Abelia* began in the late Eocene, followed by *A. chinensis* var. *ionandra* colonizing the island of Taiwan in the Middle Miocene. The ENM results suggested an expansion of climatically suitable areas during the Last Glacial Maximum and range contraction during the Last Interglacial. Disjunction between the Himalayan-Hengduan Mountain region (HHM) and the island of Taiwan is most likely the consequence of topographic isolation and postglacial contraction.

• **Conclusions** We used genomic data to reconstruct the phylogeny of *Abelia* and found a clear pattern of reticulate evolution in the group. Overall, our results supported postglacial range contraction together with topographic heterogeneity resulted in the mainland China-Taiwan island disjunction. This study provides important new insights into the speciation process and taxonomy of *Abelia*.

**Key words:** *Abelia*; Gene tree discordance; Hybridization; Mainland China-Taiwan island

56    disjunction.

## INTRODUCTION

The underlying causes of disjunct distributions in various organisms have been a central issue in biogeography (e.g., Xiang et al., 1998; Wen, 2001; Qian and Ricklefs, 2004; Wang et al., 2012; Wen et al., 2013, 2018). Much discussion has focused on two mechanisms, vicariance and long-distance dispersal (Givnish and Renner, 2004; Yoder and Nowak, 2006; Šarhanová et al., 2017; Wang et al., 2020). The emergence of geographic barriers leads to vicariance, resulting in decreased gene flow; while long-distance dispersal in plants usually involves rare events driven by dispersal vectors such as animals, wind or water currents (Nathan et al., 2008). Recent molecular systematic studies have shown that many plant disjunctions are too young to be explained by continental drift or major geographic barriers (Mummenhoff et al., 2004; Yoder and Nowak, 2006; Linder et al., 2013; Jiang et al., 2019). Hence, long-distance dispersal has emerged as a major mechanism to explain many discontinuous distributions in historical biogeography (Wen et al., 2013; Ye et al., 2019; Wen and Wagner, 2020).

Biogeographic disjunctions between mainland China-Taiwan island have received considerable attention with the separation of the two regions by 130 km (Wu and Wang, 2001; Xiang and Soltis, 2001; Qiu et al., 2011; Chen et al., 2012; Ye et al., 2014, Jin et al., 2020). The island of Taiwan is the largest subtropical mountainous island in the monsoonal western Pacific region, located on the Tropic of Cancer, and originated in the late Miocene (ca. 9-5 million years ago; Sibuet and Hsu, 2004). Molecular phylogenetic analyses have shown that several plant lineages have disjunct distributions between mainland China and Taiwan (e.g., *Sassafras* J. Presl, Nie et al., 2007; *Pseudotsuga* Carrière, Wei et al., 2010; *Triplostegia* Wall. ex DC., Niu et al., 2018; *Prinsepia*., Jin et al., 2020). Geologically, the island of Taiwan and its adjacent continental margin of mainland China belong to the Eurasian plate; in the Late Cretaceous both were part of the continental block. The continental block has undergone many extensions and rifting events since the Paleocene (Wei, 2010; Suo et al., 2015). Two hypotheses have been proposed to explain the formation of current plant disjunctions in this region: long-distance dispersal and postglacial contraction (Chou et al., 2011; Wang et al.,

2013; Ye et al., 2014; Niu et al., 2018). A phylogeographic study of *Cunninghamia* R.Br. ex

A.Rich. (Taxodiaceae) revealed that populations of Taiwan derived from continental Asia and

colonized the island via long-distance seed dispersal (Lu et al., 2001). Factors that have major

impacts on population dynamics and genetic variation patterns in plants in this region include

past geological and climatic changes, uplift of the Qinghai-Tibet Plateau, global optimum

temperature since the Miocene and the Quaternary glacial-interglacial transition (Xu et al.,

2015; Cao et al., 2016; Chen et al., 2020). Recent studies have suggested that although the ice

sheet was not ubiquitous from the Qinling Mountains and Huai River (at ca. 34° N) to the

tropical south (≤ 22° N) in China during the Last Glacial period, the temperature in most

parts of mainland China was lower than today (Shi et al., 2006). Temperature changes

influence species distributions, population sizes, and intraspecific genetic differentiation. The

postglacial contraction hypothesis states that many species had a continuous distribution from

mainland China to Taiwan prior to the Quaternary glacial period, sometimes with disjunctions

between the Himalayan-Hengduan Mountain region and the island of Taiwan occurring when

samples migrated to higher altitudes as global temperatures increased (Chen and Ying, 2012).

*Abelia* R.Br. belongs to the subfamily Linnaeoideae Raf. within Caprifoliaceae *s.l.*

(Dipsacales; Yang and Landrein, 2011; Caprifoliaceae; Wang et al., 2021). The genus

includes *A. chinensis* (four varieties), *A. forrestii* (two varieties), *A. × grandiflora*, *A.*

*macrotera* (eight varieties), *A. schumannii* and *A. uniflora* (Landrein et al., 2017; 2020; Fig.

1). *Abelia* is a typical East Asian endemic genus with a disjunct distribution in southwestern

China, extending to the eastern portion of the island of Taiwan and the Ryukyu islands in

Japan. The genus includes an important ornamental species, *Abelia × grandiflora,* widely

cultivated in yards, parks and roadsides of temperate cities (e.g., Beijing, London,

Washington D.C., and Wuhan). However, the phylogenetic relationships within *Abelia* remain

unresolved due to limited DNA markers and sampling, and hence the biogeographic history

and process of speciation are still uncertain. Kim (1998) used one nuclear (ITS) and one

plastid (*matK*) marker to infer the phylogenetic relationships within *Abelia*. However, species

relationships could not be clarified because only two *Abelia* species were included, and a

polytomy existed between the two *Abelia* species and those of *Diabelia*. Three *Abelia* species

were included (*A. chinensis*, *A. engleriana* and *A. schumannii*) in Jacobs et al. (2010), but no

information about relationships can be inferred since the three formed a polytomy. Using ITS

and nine plastid regions, Wang et al. (2015) found that *A. forrestii* was sister to all other

*Abelia* species, while *A. chinensis* was sister to *A. × grandiflora* with moderate support.

Although Landrein (2017) included 46 samples (representing five species) of *Abelia* (22 from

natural samples and 24 of horticultural breeds), using AFLP, ITS and three plastid genes, the

species phylogeny was still unclear due to lack of resolution and low to moderate bootstrap

support for many nodes.

Here, we explored the phylogenetic relationships and biogeographic diversification of

*Abelia* using a target enrichment dataset for nuclear loci and plastomes from genome

skimming. We aim to (1) clarify the species-level phylogeny of *Abelia* based on nuclear

genes and plastome data; and (2) reconstruct the biogeographic history of *Abelia* and explore

the evolution of its distribution patterns related to the disjunction between mainland China

and the island of Taiwan.

## MATERIALS AND METHODS

*Sampling, genomic data generation*

We sampled 38 individuals from all six species of *Abelia* (Landrein, 2017; 2020). Six

Linnaeoideae species [*Diabelia serrata*, *D. sanguinea*, *Kolkwitzia amabilis*, *Dipelta*

*floribunda* (two individuals), *Linnaea borealis*, and *Vesalea occidentails*] were used as

outgroups. Voucher specimens were deposited in the herbarium of the Institute of Tropical

Agriculture and Forestry of Hainan University (HUTB), Haikou, China. The complete list of

samples and information of data sources are provided in Table 1.

DNA was extracted from dried leaf tissue using a modified cetyltrimethylammonium

bromide (CTAB) method, with chloroform: isoamyl alcohol separation and isopropanol

precipitation at -20□ (Doyle and Doyle, 1987). We checked the concentration of each

extraction with a Qubit 2.0 Fluorometer (Thermo Fisher Scientific, Waltham, MA, USA) and

sonicated 400 ng of DNA using a Covaris S2 (Covaris, Woburn, MA) to produce fragments

~150 - 350 bp in length for making sequencing libraries. To ensure that genomic DNA was

sheared to the appropriate fragment size, we evaluated all samples on a 1.2% (w/v) agarose

gel.

We used a Hyb-Seq approach (Weitemier et al., 2014) combining target enrichment and

genome skimming, which allows simultaneous data collection for low-copy nuclear genes

and high-copy genomic targets. To obtained nuclear data from target enrichment, we ran

MarkerMiner v.1.2 (Chamala et al., 2015) with default settings to identified putative single

copy nuclear (SCN) genes, using the transcriptomes of *Lonicera japonica* from 1KP, and the

genome of *Arabidopsis thaliana* (L.) Heynh. (Gan et al., 2011) as a reference. Then, we used

GoldFinder (Vargas et al., 2019) to further filter the SCN genes identified by MarkerMiner.

Finally, we obtained 428 putatively single-copy orthologous genes in total for phylogenetic

analysis (Wang et al., 2021). Hybridization, enrichment, and sequencing methods followed

those of Wang et al. (2021). We obtained full plastomes from genome skimming data (Wang

et al., 2021) with detailed descriptions of plastome sequencing methods in Wang et al. (2020).

*Read processing and assembly*

Raw sequencing reads were cleaned using Trimmomatic v.0.36 (Bolger et al., 2014) by

removing adaptor sequences and low-quality bases (ILLUMINACLIP: TruSeq_ADAPTER:

2:30:10 SLIDINGWINDOW: 4:5 LEADING: 5 TRAILING: 5 MINLEN: 25). Nuclear loci were

assembled with HybPiper v.1.3.1 (Johnson et al., 2016). Individual exon assemblies were

carried out to avoid chimeric sequences in multi-exon genes produced by potential paralogy

(Morales-Briones et al., 2018) and only exons with a reference length of more than 150 bp

were assembled. Paralog detection was performed for all exons with the 'paralog_

investigator' warning. To obtain 'monophyletic outgroup'(MO) orthologs (Yang and Smith,

2014), all assembled loci (with and without paralogs detected) were processed following

Morales-Briones et al. (2021).

For plastome assembly, raw reads were trimmed using SOAPfilter_v2.2 (BGI, Shenzhen,

China). Adapter sequences and low-quality reads with Q-value ≤ 20 were removed. Clean

reads were assembled against the plastome of *Kolkwitzia amabilis* (KT966716.1) using

MITObim v.1.8 (Hahn et al., 2013). Assembled plastomes were preliminarily annotated with

Geneious R11.04 (Biomatters Ltd., Auckland, New Zealand) to detect possible

inversions/rearrangements, with corrections of start/stop codons based on published *Abelia*

plastomes (*A. chinensis*, accession No. MN384463.1; *A. forrestii*, accession No. MN524636.1;

*A.* × *grandiflora*, accession No. MN524635.1; *A. macrotera*, accession No. MN524637.1).

*Phylogenetic analyses*

*Nuclear dataset.* We analyzed the nuclear data set using concatenation and

coalescent-based methods. Individual exons were aligned with MAFFT v.7.037b (Katoh and

Standley, 2013) and aligned columns with more than 90% missing data were removed using

Phyutility (Smith and Dunn, 2008) prior to concatenation. Maximum likelihood (ML) analysis

was conducted using IQ-TREE v.1.6.1 (Nguyen et al., 2015) searching for the best partition

scheme (Lanfear et al., 2012) followed by ML gene tree inference and 1000 ultrafast

bootstrap replicates (Hoang and Chernomor 2018). Additionally, we used Quartet Sampling

(QS; Pease et al., 2018) to distinguish strong conflict from weakly supported branches. We

carried out QS with 1000 replicates. We used ASTRAL-III v.5.7.1 (Zhang et al., 2018) for

species tree estimation, with the input being individual exon trees generated with RAxML

using a GTRGAMMA model.

*Plastome dataset*. Contiguous contigs were aligned with MAFFT v.7.407 (Katoh and

Standley, 2013) and columns with more than 90% missing data were removed using Phyutility.

We performed extended model selection (Kalyaanamoorthy et al., 2017) followed by ML

gene tree inference and 200 non-parametric bootstrap replicates for branch support in

IQ-TREE. Additionally, we evaluated branch support with QS using 1,000 replicates.

*Species network analysis*

Due to computational restrictions and given our primary focus on potential reticulation between different species, we reduced the 45-sample data set to one outgroup and 14 ingroup samples representing all major clades based on the nuclear analyses. We looked for evidence of hybridization within *Abelia* via species network search carried out with the InferNetwork_MPL function in PhyloNet (Wen et al., 2018). Network searches were carried out using only nodes in the gene trees that had at least 50 percent bootstrap support, enabling up to five hybridization events and optimizing the branch lengths and probabilities of inheritance of the returned species networks under the full probability. To disentangle nested hybridization, we also created a reduced, 12-sample data set by removing three known hybrids (*A.× grandiflora* E159, *A.× grandiflora* E227 and *A. chinensis* E206; see results) and carried out analyses as previously described. We used the command 'CalGTProb' in PhyloNet to calculate the probability scores of the concatenated RAxML, ASTRAL, and plastid trees, given individual nuclear gene trees, to estimate the optimal number of hybridizations and to test if the network model explained our gene trees better than a purely bifurcating tree. We selected the best-fit-species network model using the lowest values for bias-corrected Akaike information criterion (AICc; Sugiura, 1978), Akaike information criterion (AIC; Akaike, 1998), and the Bayesian information criterion (BIC; Schwarz, 1978).

*Genetic diversity and population structure*

Because increasing the quantity of missing data in the 'median' ($Fst \approx 0.15$) and 'maximum' ( $Fst > 0.15$ ) SNP matrices can lead to skewed estimations of $Fst$ (Hodel et al., 2017), we only used the 'minimum' data set for genetic diversity analysis. We used BAYESCAN v.2.1 (Foll and Gaggiotti, 2008; Cao et al., 2016) to calculated genetic distance ($Fst$) for nuclear gene loci, to define the 'minimum' dataset ($Fst < 0.15$). All parameters were kept at default settings except for modifying the *a priori* odds of the neutral model to 250 (Foll and Gaggiotti, 2008).

DnaSP v.5 (Librado and Rozas, 2009) was used to analyze nucleotide polymorphisms

and haplotypes from population of each species in the 'minimum' data set. NETWORK

v.5.0.0.3 (Bandelt et al., 1999) was used to visualize the haplotype network diagrams of each

species. We estimated genetic diversity indices including nucleotide diversity ($\pi$), expected

heterozygosity (*He*), observed heterozygosity (*Ho*) and inbreeding coefficients (*Fis*) using

GenAlEx v.6.1(Peakall and Smouse, 2006).

Bayesian clustering of individuals was conducted using STRUCTURE v. 2.3.4

(Pritchard et al., 2000) for *Abelia* using the 'minimum' data set. Ancestral genetic populations

(K) were determined using an admixture model assuming correlated allele frequencies. The

length of the burn-in period and the MCMC (Markov Chain Monte Carlo algorithm) runs

were set to 50,000 and 500,000, respectively. Three runs for each K value (K = 1-10) were

used to estimate the most likely value of K using $\Delta$K (Evanno et al., 2005) after combining

with Clumpp v.1.1.2b (Jakobsson and Rosenberg, 2007) and analyzing with Structure

Harvester v.0.6.94 (Earl, 2012).

*Divergence time estimation*

To examine the historical biogeography of *Abelia*, we performed a dated phylogenetic

analysis using BEAST v.1.8 (Drummond et al., 2012). As there is potential ancient

hybridization in *Abelia*, we estimated dates separately using the nuclear and plastid data. We

selected the fruit fossil of *Diplodipelta* from the late Eocene Florissant flora of Colorado

(Manchester, 2000; Bell and Donoghue, 2005) to calibrate the node containing *Diabelia* and

*Dipelta*, as the fossil shows morphological similarities to the two genera. We set a lognormal

prior for the divergence at the stem of *Diabelia* and its sister clade *Dipelta* (offset 36 Ma,

log-normal prior distribution 34.07-37.20 Ma). The ancestral node of Linnaeoideae (the root

of our tree) was constrained to 50.18 Ma (mean 50.18 Ma) from results of previous studies

(Wang et al., 2021) despite the lack of fossil evidence. The BEAST analyses were performed

using an uncorrelated log normal (UCLN) relaxed clock with a Yule tree prior, random

starting trees, and GTR + G as model of sequence evolution (as selected by MrModelTest

v.2.3; Nylander et al., 2004). Each MCMC run was conducted for 100,000,000 generations

with sampling every 5,000 generation. The first 10% of trees were discarded as burn-in. We

used Tracer v.1.6 to assess convergence and effective sample size (ESS) > 200 (Rambaut et

al., 2014). Tree Annotator v.1.8.0 (Drummond et al., 2012) was used to summarize the set of

post-burn-in trees and their parameters in order to produce a maximum clade credibility

chronogram showing the mean divergence time estimates with 95% highest posterior density

(HPD) intervals. FigTree v.1.4.2 (Drummond et al., 2012) was used to visualize the resulting

divergence times.

*Ancestral area reconstructions*

We reconstructed the historical biogeography of *Abelia* using the Bayesian binary

MCMC (BBM) analysis implemented in RASP v.3.21 (Yu et al., 2015). To account for

phylogenetic uncertainty, 5,000 out of 20,000 post-burn-in trees from the BEAST analyses

(nuclear data) were randomly chosen for the BBM analyses. We compiled distribution data

for the species of *Abelia* and assigned the included taxa to their respective ranges.

Considering the areas of endemism in *Abelia* and tectonic history of the continent, we used

three areas: (A) Southwest China, (B) Central-East China, and (C) Taiwan and Ryukyu

islands. Each sample was assigned to its respective area according to its contemporary

distribution range (Table 1). We set the root distribution to null, applied 10 MCMC chains

each run for 100 million generations, sampling every 10,000 generations using the JC + G

model. The MCMC samples were examined in Tracer v.1.6 (Drummond et al., 2012) to

confirm sampling adequacy and convergence of the chains to a stationary distribution.

*Ecological niche modelling*

We performed Ecological Niche Modelling (ENM) on *Abelia* using MAXENT v.3.2.1

(Phillips et al., 2006). For the input we obtained 200 species occurrence data points from the

Chinese Virtual Herbarium (CVH, https://www.cvh.ac.cn/), Global Biodiversity Information

Facility (http://www.gbif.org/) and field observations. The details of these occurrence data are

listed in Table S1. We used 19 bioclimatic variables as environmental data for ENM

representing five different time periods. Past and current data for these variables were

available from the WorldClim database (http://worldclim.org/). The data included baseline

climate (BioClim layers for the period 1950-2000 at a spatial resolution of 30 s arc), data for

the last glacial maximum (LGM; ~22 000 years BP) with spatial resolution at 2.5 min arc

resolution simulated by Community Climate System Model (CCSM), Max Planck Institute

Model (MPI) and Model for Interdisciplinary Research on Climate (MIROC), and the last

interglacial period. Replicate runs (25) of MAXENT were performed to ensure reliable

results. Model performance was assessed using the area under the curve (AUC) of the

receiver operating characteristic. AUC scores may range between 0.5 (randomness) and 1

(exact match), with those above 0.9 indicating good performance of the model (Swets, 1988).

The MAXENT jackknife analysis was used to evaluate the relative importance of the 19

bioclimatic variables employed. All ENM predictions were visualized in ArcGIS v.10.2

(ESRI, Redlands, CA, USA).

*Data accessibility*

Raw Illumina data from target enrichment is available at the Sequence Read Archive

(SRA) under accession SUB7674585 (see Table 1 for individual sample SRA accession

numbers). DNA alignments, phylogenetic trees and results from all analyses and datasets can

be found in the Dryad data repository.

**RESULTS**

*Exon assembly*

The number of assembled exons per species ranged from 253 (*A. macrotera* var.

*macrotera* E45) to 992 (*A. uniflora* E46) (with > 75% of the target length), with an average of

851 exons per sample (Table S2; Fig. S1). The number of exons with paralog warnings

ranged from 6 in *A. macrotera* var. *macrotera* E45 to 698 in *A. uniflora* E46 (Table S2). The

concatenated alignment had a length of 365,330 bp, including 22,045 parsimony-informative

sites, and a matrix occupancy of 83% (Table 2). The smallest locus had a length of 150 bp,

while the largest locus had a length of 3,998 bp. The plastid alignment resulted in a matrix of

163,130 bp with 3,850 parsimony-informative sites and a matrix occupancy of 96% (Table 2).

*Phylogenetic reconstruction*

We inferred the Abelia phylogeny based on nuclear and plastid alignments (Figs. 2-3 and

S2-3). Despite the generally well-supported topology, the positions of clades representing *A.*

× *grandiflora*, *A. uniflora* and *A. macrotera* clade were unstable across all analyses.

Phylogenetic reconstructions based on nuclear concatenation data showed all of the major

clades with relationships well supported (BS ≥ 75%; Fig. 2, Left). *Abelia schumannii* is

nested within *A. macrotera*, and sister to *A. macrotera* var. *macrotera* E45, with the QS

analysis showing moderate support (0.5/0.92/1, Fig. S2, Left). *Abelia chinensis* forms a clade

with *A.* × *grandiflora*; and *A. macrotera* var. *zabelioides* is sister to *A. uniflora* with the clade

having strong QS support but also a signal of a possible alternative topology (0.97/0/1, Fig.

S2, Left). The *A. chinensis* + *A.* × *grandiflora* clade is then sister to the *A. macrotera* var.

*zabelioides* + *A. uniflora* clade (Fig. 2, Left).

The ASTRAL analyses (Fig. 3) recovered all major clades with strong support (local

posterior probabilities, LPP ≥ 0.95). Similar to the ML concatenation nuclear tree, *A.*

*schumannii* is nested within *A. macrotera*. The *A. chinensis* + *A.* × *grandiflora* clade is sister

to the remaining species. *Abelia forrestii* is sister to *A. macrotera* + *A. schumannii* clade,

which is then sister to *A. macrotera* var. *zabelioides* + *A. uniflora*. The clade composed of *A.*

*chinensis* + *A.* × *grandiflora* has moderate QS support with signal for a possible alternative

topology (0.33/0.06/0.99). The clade composed of *A. macrotera* var. *zabelioides* + *A. uniflora*

has strong QS support but a signal of a possible alternative topology (0.98/0/1). The sister

relationship of the *A. macrotera* var. *zabelioides* + *A. uniflora* has QS counter support as a

clear signal for alternative topology (-0.21/0.32/0.99).

For the plastid tree (Fig. 2, Right and S3), *Abelia schumannii* was nested within *A.*

*macrotera*, and sister to *A. macrotera* var. *macrotera* E45, with the QS analysis showing

moderate support (0.27/0.18/0.91). *Abelia uniflora* is sister to *A. macrotera* var. *zabelioides*,

with the clade having strong QS support signaling an alternative topology (0.51/0.35/0.98).

The *A. chinensis* + *A.* × *grandiflora* clade shows a sister relationship with the *A. macrotera*

var. *zabelioides* + *A. uniflora* + *A. macrotera* + *A. schumannii* clade, with strong QS support

(0.75/0.51/0.98). The results strongly support *A. forrestii* from Yunnan as sister to the

remainder of the genus with the QS analysis providing full support (1/–/1, Fig. S3).

In our species network analyses (including both 15 and 12 samples), all networks with

one to five reticulations events were shown to be a better model than a strictly bifurcating tree

(Table 3; Figs. 4 and S4-5). Four reticulation events were inferred for the best networks in the

15-sample tree, the two reticulation-events involving *A.* × *grandiflora*, the third reticulation

event revealed *A. chinensis* and the fourth reticulation event revealed ancestral gene flow in

the *A.* × *grandiflora* clade (Fig. 4A). For *A.* × *grandiflora* E159, the inheritance probability

analysis showed that the major genetic contributor (57.2%) came from the ancestor of *A.* ×

*grandiflora* and a minor contributor was from the *A. macrotera* clade (42.8%). For *A.* ×

*grandiflora* E227, the major genetic contributor came from *A. uniflora* (53.6%). For *A.*

*chinensis* E206, the major genetic contributor (53.0%) came from the ancestor of *A.* ×

*grandiflora*. For the ancestor of the *A.* × *grandiflora* clade, the major genetic contributor was

a putative ancestor of *A. chinensis* (74.1%) with a minor contribution from *A. uniflora*

(25.9%). In the 12-sample tree (Fig. 4B), we removed the known hybrids from the previous

analyses, excluding *A.* × *grandiflora* E159, *A.*× *grandiflora* E227 and *A. chinensis* E206, and

then discovered that the best network had three reticulation events. There appears to be

abundant gene flow in the backbone of *Abelia* and the 12-sample analyses showed that *A.*

*macrotera*, *A. uniflora* and *A. chinensis* were each derived from complex hybrid origins.

*Population structure*

We used BAYESCAN v.2.1 to obtain a total of 788 genes (*Fst* < 0.15, Fig. S6). Fifteen

haplotypes (H1-15) were identified across the 12 populations of *Abelia* (Fig. 5). Five

haplotypes (H3, H5, H9, H12 and H13) were found to be shared by multiple species. H3 was

a shared haplotype between *A. macrotera* var. *engleriana* and *A. chinensis*. H5 was a shared

haplotype between *A. uniflora* and *A. macrotera*. H9 was shared by *A. macrotera* var. *mairi*,

*A. macrotera* var. *deutziaefoli*a, *A. macrotera* var. *zabelioides* and *A. macrotera* var.

*engleriana*. H12 is unique to *Abelia chinensis* var. *ionandra*. H13 was a shared haplotype

between *A. chinensis* var. *ionandra* and *A. chinensis*. There were three haplotypes (H3, H8

and H9), mainly distributed in southwestern China (e.g., Chongqing, Guizhou, Sichuan and

Yunnan), with the diversity ranging from 0.139 to 0.176 (Fig. 5, Table. S3). There were six

haplotypes (H2, H7, H9, H10, H11 and H15) distributed in central China (e.g., Hubei), with

the diversity is 0.477. There were six haplotypes (H1, H5, H7, H12, H13 and H14)

distributed in eastern China (e.g., Fujian), southern portion of the island of Taiwan, and

southern Japan (Fig. 5).

The results of the STRUCTURE analyses based on the BAYESCAN markers are

presented in Fig. 5, the values of the log-likelihood of the data, log Pr(K), increased

progressively with the increase in K. According to Evanno's approach, K$=$4 was the most

likely number of genetic clusters for 0.5 million generations. The results of STRUCTURE

suggests that there is gene flow among the four species (Fig. 5).

*Divergence time estimation*

Divergence time estimates based on the nuclear tree (Fig. 6) suggested that the crown

group of *Abelia* originated 33.57 Ma (95% HPD: 29.16-37.55 Ma, node 1). The *A. chinensis*

clade was dated to have diverged from the remaining *Abelia* species at 31.82 Ma (95% HPD:

27.57-37.10 Ma, node 2). The the split between *A.* $\times$ *grandiflora* and all varieties of *A.*

*chinensis* was dated to 25.09 Ma (95% HPD: 19.00-33.80 Ma, node 4). The split between *A.*

*macrotera* var. *deutziaefolia* and *A. macrotera* var. *engleriana* was dated to 18.53 Ma (95%

HPD: 16.72-23.31 Ma, node 8). The split between *A. uniflora* and *A. macrotera* var.

*zabelioides* was dated to 18.33 Ma (95% HPD: 15.09-22.12 Ma, node 9). The split time

between *A. macrotera* and *A. schumannii* was 6.78 Ma (95% HPD: 4.92-11.29 Ma, node 12).

The crown of *A. forrestii* was 4.61 Ma (95% HPD:1.88-6.71 Ma, node 14). Most

diversification events within *Abelia* dated to the Early to Middle Miocene (Fig. 6). A

comparison of the time estimates using the chloroplast tree is shown in Fig. S7. For instance,

the age of the crown group of *Abelia* was dated to 45.10 Ma (95% HPD: 31.82-50.09 Ma,

node 1). The split between *A. uniflora* and *A. macrotera* var. *zabelioides* was 35.07 Ma (95%

HPD: 25.00-39.88 Ma, node 4). The split between *A. macrotera* and *A. schumannii* was 31.79

Ma (95% HPD: 18.39-38.15 Ma, node 5).

*Ancestral area reconstruction*

According to our ancestral area reconstructions paired with divergence time estimation,

*Abelia* originated around the late Eocene (Fig. 6) in mainland China (Fig. 7). Our analyses

revealed four dispersal events and two vicariance events among the three defined

biogeographic areas (Fig. 7). The ancestor of *Abelia* was inferred to have been present

throughout region A (Yunnan, Chongqing, Sichuan, Guizhou and Hubei) and B (Jiangxi and

Fujian), with subsequent dispersal or vicariance across southeast China. The results indicate a

southwest China origin for the genus *Abelia*.

*Ecological niche modelling and niche identity tests*

The AUC test result (mean ± SD) for the ENM, averaged across all 25 runs, was

significant (0.996 ±0.000 for *A. chinensis* and 0.993± 0.001 for *A. uniflora*) (Figs. 8 and S8),

which supports the predictive power of the model. The MaxEnt analyses show that the

current distribution prediction is largely identical to the recorded distribution of *Abelia*, and

expansion of distribution beyond the existing range is also reasonable, including *A. uniflora*

populations along the narrow coastline of eastern China and Japan (Fig. S8). In our analysis,

the LGM climate and the present time climate species distribution areas show different

distribution patterns. We used CCSM, MIROC and MPI models to simulate the LGM climate

for five species of *Abelia* with the distribution of the two LGM models (CCSM and MPI

models) being highly similar. For the CCSM and MIROC models (the MIROC model area

simulation during the LGM period is warmer and drier than the CCSM model), the

distribution density and range of *Abelia* under the MPI model was slightly lower than the

distribution density and range under the CCSM model, but the differences were not

significant (Fig. S8). Compared with the LGM climate, the potential distribution range of

*Abelia* species in mainland China will be much smaller, while the habitat remains stable in

most of the present time climate distributions.

**DISCUSSION**

*Phylogenetic relationships within Abelia*

This study reconstructed well-supported phylogenetic relationships for *Abelia* using both

nuclear and plastid data (Figs. 2-3 and S2-3). The inferred topologies of the four recognized

species in the nuclear and plastid trees were different from each other (Figs. 2-3 and S2-3),

but also share many commonalities. For instance, *Abelia schumannii* is nested within *A.

macrotera*, and *A. macrotera* var. *zabelioides* is sister to the *A. uniflora* clade. In addition,

these nodes have strong QS support values. Cytonuclear discordance has been widely

detected in plants. The three main reasons for cytonuclear discordance are gene

duplication/extinction, horizontal gene transfer/hybridization, and incomplete lineage sorting

(ILS), but sources of such discordance remain hard to disentangle, especially when multiple

processes co-occur (Degnan and Rosenberg, 2009).

*Abelia macrotera* var. *zabelioides* and *A. uniflora* share a similar origin (Landrein et al.,

2017), which indicates that *A. macrotera* × *chinensis* formed from different hybridization

events at different times as shown in our species network analyses (Fig. 4). We also found

that *A. macrotera* var. *zabelioides* and *A. uniflora* show many similarities in morphology (i.e.,

single-flowered inflorescence, red corolla, and two sepals). Landrein (2020) reported that the

*A. uniflora* group contains *A. macrotera* var. *zabelioide*s.

Our results are consistent with Jacobs et al. (2010) in terms of *A. schumannii* being nested

within *A. macrotera*. Additionally, *A. schumannii* has no independent ecological niche, and its

distribution overlaps with *A. macrotera* (Landrein et al., 2020). However, this does not

necessitate that *A. schumannii* is a synonym of *A. macrotera* since we have only sampled one

accession of *A. schumannii*. A single accession is likely not enough since we are not

capturing any genetic variation within *A. schumannii*. Therefore we need to expand the

sampling of *A. schumannii* in future studies.

*Hybridization and introgression*

Over many decades, evolutionary studies have focused on hybridization, the

crossbreeding of individuals from different species, and introgression, the transfer of genes

across species mediated largely by backcrossing (Rieseberg and Carney, 1998; Grant PR and

Grant BR, 2019; Rendón-Anaya et al., 2019; Nwankwo et al., 2019). The *A. uniflora*

complex is the most enigmatic complex in the genus, though it was only briefly described by

Wallich (1829). Members of the *A. uniflora* complex are often difficult to identify using

morphological characters alone. The morphological differences among entities within the *A.*

*uniflora* complex can be explained by one or several historical or ongoing evolutionary

processes. Specifically, the null taxonomic hypothesis is that the *A. uniflora* complex is a

single hyperdiverse species (Link-Pérez, 2010), while the alternative is that speciation has

occurred or is occurring. We expect to detect clear genetic differentiation which may be

stronger if speciation is complete, or weaker if it is ongoing or confounded by introgression

(Nosil, 2008; Liu et al., 2018). Based on morphological evidence and AFLP data, Landrein et

al. (2017) indicated that *A. uniflora* shows evidence of formation due to hybridization and

introgression between *A. macrotera* and *A. chinensis*. Examples of similar speciation events

in different families were cited in Li et al. (2011), and the hypothesis put forward was

glaciation survival east and west of the Mekong-Salween divide. Glaciation may have led to

the isolated distribution of *A. forrestii*. Repeated hybridization and introgression tend to blur

the lines between hybrids and their parents. Landrein et al (2017) showed that *A. forrestii*

occurs in the contact zone between *A. macrotera* and *A. chinensis*, leading to the hypothesis

that hybridization and introgression between *A. macrotera* and *A. chinensis* led to the

emergence of *A. forrestii.*

We find that *A. × grandiflora* is sister to *A. chinensis* in the inferred phylogenies (Figs. 2,

3) and our network analyses further confirm the hybrid origin of *A. × grandiflora* from *A.*

*chinesis* and *A. uniflora* (Fig. 4A). Landrein et al. (2017) found that *A. uniflora* formed a

clade with *A. × grandiflora* and they reported that the hybrid origin of cultivated *A. ×*

*grandiflora* involving *A. uniflora* as one progenitor, but the other progenitor could not

previously be identified. They speculated that the second progenitor was *A. chinensis* or to a

lesser extent *A. macrotera* based on their phylogenetic analysis, or even a hybrid individual

between the two (*A. chinensis* and *A. macrotera*).

Landrein et al. (2017) highlighted the genetic distinction between *A. uniflora* and *A.*

*macrotera*, two taxa that are often difficult to identify based on morphology alone (Fig. 1).

Our results indicate that a clear disjunction in their distribution exists (Fig. 3). West-to-east

disjunctions across China have also been described in other genera and species, with some

being hypothesized to form from allopatric incipient speciation (Qiu et al., 2009a). In our

study, the identity and origin of *A. × grandifolia*, a hybrid between *A. chinensis* and *A.*

*uniflora*, is weakly supported, while *A. uniflora* shows evidence of allopatric speciation due

to hybridization and introgression between *A. macrotera* and *A. chinensis* (Table 3; Figs. 4,

S4-5). In the STRUCTURE analysis, we found clear evidence of hybridization among the

four species (Fig. 5), however, the full extent of hybridization is difficult to assess as it is

represented by relatively few samples. In addition, a substantial amount of discordance in the

main branches of *Abelia* exist, and the network (12-sample) analyses (Fig. 4B) showed that *A.*

*macrotera*, *A. uniflora*, and *A chinensis* all have complex hybrid origins. This is a potential

reason why the taxonomy of this group has been historically complicated.

*Biogeographic diversification history of Abelia*

In this study, divergence time estimates from inferred from nuclear (Fig. 6) and

chloroplast data (Fig. S7) show stark differences. Simulation studies have revealed that

methods for inferring species trees are likely to be relatively resistant to intra-locus

recombination (Lanier and Knowles, 2012; Wang and Liu, 2016). In our analysis, the

molecular clock estimates of the species tree indicate that the crown of *Abelia* dates to the

late Eocene 33.57 Ma (95% HPD: 29.16-37.55 Ma) (Fig. 6). Our phylogenetic analyses

indicate that *A. chinensis* var. *ionadra* from the island of Taiwan is sister to *A. chinensis* from

Jiangxi province, China (Figs. 2-3 and S2-3). In addition, previous phylogenetic and

phylogeographic studies have generally shown that existing populations of many species on

the island of Taiwan are descendants of mainland China (e.g., *Taiwania*, Chou et al., 2011;

*Sassafras*, Nie et al., 2007; *Dichocarpum arisanense*, Xiang et al., 2017; *Triplostegia*, Niu et

al., 2018), with no reported cases supporting a reverse direction of dispersal. Based on our

time estimates, the stem age of *A. chinensis* var. *ionadra* was dated to the Middle Miocene, ca.

18.59 Ma (Fig. 6, node 8), which greatly predates the formation of the island of Taiwan,

which originated in the Late Miocene (Sibuet and Hsu, 2004). A similar disjunction predating

the formation of Taiwan is shown in *Tinospora dentata* which split from its Japanese relatives

in the Early Oligocene around 38.97 Ma (95% HPD: 26.45-52.52) (Wang et al., 2017).

The island of Taiwan is located at the eastern edge of the Eurasian plate and the Taiwan

Strait could have first occurred in the Late Mesozoic (Ye, 1982; Suo et al., 2015). Owing to

the intense tectonic movements, contacts between the island of Taiwan and mainland China

occurred during the Late Cretaceous-Early Paleocene and Late Eocene-Early Oligocene

(Teng, 1992; Zhang, 1995). One hypothesis that has been put forward is that tropical

rainforests in the Indo-Malayan region may have appeared near the Cretaceous-

Paleogene (K-Pg) boundary (Wang et al., 2012). Macrofossils of certain plants, closely

related to present subtropical species, have been recorded from Taiwan during the Early

Miocene, but pollen fossils date to the Oligocene (Li, 2000), although their geographic

origins are still debated (Huang, 2011). Due to the landmass being subjected to stretching and

rifting (Ye, 1982; Teng, 1992; Suo et al., 2015), part of the mountain ranges on the island of

Taiwan represents a relict area of the 'Taiwan-Sinzi Folded Zone' (Ye, 1982; Teng, 1992; Li

and Wen, 2016), which helped facilitate an increasingly cooler climate in Asia during the

Late Eocene (Zachos et al., 2001). Around this time, a vicariance event between Taiwan and

mainland China may have occurred. Based on the results of our *Abelia* sequencing, the

Taiwan (Yilan) flora may contain relatively ancient components (Li and Wen, 2016). The

flora of mainland China has been acknowledged to have contributed to Taiwan's floristic

assembly (Li and Keng, 1950). The Luzon arc did not begin to collide with the Eurasian plate

until the Late Miocene (9-6 Ma) (Sibuet et al., 2002). Thus, we hypothesize that the relatively

young gene flow in the flora of Taiwan may be derived from plants in mainland China,

whereas the ancient components are derived solely from plants in mainland China. This

hypothesis needs further validation by combining phylogenetic and molecular dating methods,

as well as by studying other plant groups, particularly more ancient lineages.

The crown group of *Abelia* is dated to have formed during the late Eocene (35.07 Ma,

25.00-39.88 Ma; node 1), later than the early uplift of the Qinghai-Tibet Plateau (QTP).

However, ancestral area reconstructions revealed that the *A. uniflora* complex might have

colonized southwest China, Central-East China, and Taiwan during the Eocene (Fig. 8).

*Potential Refugia and Demographic History*

Quaternary climate shifts and associated environmental changes may have made the

diversification of *Abelia* relatively complicated. On one hand, diversification would have

been sustained by Pleistocene glaciations because species distributions underwent expansion

and contraction, leading to population migration or extinction. On the other hand, no clear

geographical isolation, differentiation and subsequent population expansion (Comes et al.,

2008; Qiu et al., 2011) occurred.

Our ENM analysis provides a detailed picture of the last glacial cycle (Figs. 8 and S8).

The results indicate that species differentiation in *Abelia* occurred in the Oligocene,

suggesting that extant populations likely differentiated well before the LGM and also provide

strong evidence for survival of *Abelia* in southwest China during the LGM period (Fig. S8).

Therefore, the climatic shock in the LGM may have had a significant impact on the recent

history of *Abelia* at the population level compared with the earlier period of species

differentiation (Qiu et al., 2011; Qi et al., 2012). Potential suitable habitat for *A. forrestii* was

distributed mainly over southwest China during the LGM period. However, its current

distribution is only in Yunnan and Sichuan of southwest China from 1900 to 3300 m in

altitude.

Based on genetic diversity, although the *Abelia* distribution area is highly discontinuous

(Fig. 5), the populations still show similar or even higher diversity at haplotypes (nuclear data)

(Table S3). *A.* × *grandiflora* has the highest genetic diversity ($\pi$=0.478), with the expected

heterozygosity (*He*) ranging from 0.351-0.474 and the observed heterozygosity (*Ho*) ranging

from 0.436-0.778. *Abelia schumannii* has the highest expected heterozygosity (*He*=0.474),

while the highest observed heterozygosity (*Ho*=0.778) is in *A. macrotera* var. *deutziaefolia*.

Based on our nuclear haplotype networks (Fig. 5), the populations of mainland China,

Taiwan and Japan have significant genetic differentiation (Fig. 5). Based on ENM analysis, *A.*

*chinensis* var. *ionandra* expanded to the coastal regions of East China and reached Taiwan

via land bridges across the Taiwan Strait during the glacial period. Due to repeated

glacial-interglacial cycles, mainland China and Taiwan populations may have experienced

frequent gene flow through a continental bridge during the Quaternary.

Notably, we found that the populations in Jiangxi Province, which have not been

dispersed to other areas, have a unique haplotype (H10, Fig. 5) and high genetic diversity (JX,

$\pi$: 0.596; Table S3). After the LGM, *Abelia* populations in southeastern China may have

experienced a large-scale population extinction event, and Jiangxi may be a glacial refuge for

*A. chinensis*. The H13 haplotype can be found in the eastern coastal areas of China and Japan

(e.g., Hubei, Taiwan and Ryukyu islands), however, no bottleneck effects (i.e., high genetic

diversity and high genetic differentiation) were observed as would be expected for a refugial

population (Hewitt, 2000). In addition, both the southwest and central China lineages were

found to harbor the ancestral haplotype (H9). The predicted results of ENM during the LGM

also support this conclusion of a refugial area (Fig. S8). The uplift of the QTP was a

consequence of the collision of the Indian subcontinent with the Eurasian plate 40-50 million

years ago (Yin and Harrison, 2000; Fu et al., 2020). Because of multiple asynchronous

expansions, multiple independent refugia *in situ* and complex topography in this region might

explain why accurate analyses of species distribution patterns are difficult (Cun and Wang,

2010; Yan et al., 2012; Shao et al., 2015).

Our findings, together with those from other studies in this region (Han et al., 2016),

suggest that, apart from geographic barriers, local adaptation to different climatic conditions

appear to be the major drivers for lineage diversification of plants endemic to East Asian

land-bridge islands.

**CONCLUSIONS**

Robust phylogenetic relationships were consistently reconstructed within *Abelia* based

on comprehensive taxon sampling, using both nuclear and plastid data. The nuclear and

plastid phylogenies show that *A. schumannii* is nested within *A. macrotera*, and that *A.*

*macrotera* var. *zabelioides* is sister to *A. uniflora*. Also, there was substantial discordance in

the backbone of *Abelia*, and the network (12-sample) analyses suggested that *A chinensis*, *A.*

*macrotera* and *A. uniflora* are all from complex hybrid origins.

We used molecular dating, biogeographic and ENM methods to explore how geology

and climate together influenced the evolutionary history of *Abelia*, an ecologically sensitive

and cold-adapted genus. High genetic divergences were detected among *Abelia* populations

in southwest China, Central China, and Taiwan, and several interglacial refugia were

recognized in the central Hengduan Mountains. Ancestral area reconstruction revealed that

*Abelia* originated in southwest China, and its diversification began during the late Eocene,

colonizing Central China in the mid-Oligocene, and finally dispersing to the island of Taiwan

via a land bridge at approximately 19.33 Ma. Based on phylogeographic and ENM analyses,

frequent gene flow might have existed between populations of mainland China and the island

of Taiwan with repeated continuous dispersal between the two regions. Our research

represents one of the first comprehensive phylogeographic studies investigating plant

disjunctions between the island of Taiwan and mainland China, while supporting hypotheses

of postglacial range contraction together with topographic heterogeneity being responsible for

the observed disjunction.

**TAXONOMIC TREATMENT**

*Abelia uniflora* R. Br. in Wall., Pl. As. Rar. 1. 15. 1830. -Type: China, Fujian,

['Ngan-Ke-Kyen'Black Tea Country], *Reeves s.n.* (holotype CGE!, isotype BM!).

=*Abelia macrotera* (Graebn. & Buchw.) Rehder var. *zabelioides* Landrein, Kew Bull.

74(4)-70: 93. 2019. Type: China, Sichuan, Leshan Pref., E'mei Shan, 21 June 1938, *H. C.*

*Chow 7573* (holotype A!). **Synonym nov.**

**Author contributions**

H.F.W. and J.W. conceived the study. H.F.W. and D.F.M-B. performed the research and

analyzed the data. Q.H.S., D.F.M-B., H.X.W., J.W. J.B.L, and H.F.W. wrote and revised the

manuscript.

**ACKNOWLEDGEMENTS**

The work was funded by National Science Foundation of China (31660055). We appreciate

Gabriel Johnson for his help with the target enrichment experiment, and the United States

National Herbarium for collection access. We acknowledge the staff in the Laboratories of

Analytical Biology at the National Museum of Natural History, the Smithsonian Institution

for technical support and assistance.

**REFERENCES**

**Akaike H. 1998.** Information theory and an extension of the maximum likelihood principle.
*Selected papers of hirotugu akaike* 199-213.

**Bell CD, Donoghue MJ. 2005.** Phylogeny and biogeography of Valerianaceae (Dipsacales)
with special reference to the South American valerians. *Organisms Diversity and*

*Evolution* **5**: 147-159.

**Cao YN, Wang IJ, Chen LY, Ding YQ, Liu LX, Qiu YX. 2016.** Inferring spatial patterns

and drivers of population divergence of *Neolitsea sericea* (Lauraceae), based on

molecular phylogeography and landscape genomics. *Molecular Phylogenetics and*

*Evolution* **126**: 162-172.

**Chen XD, Yang J, Feng LI, Zhou T, Zhang H, Li HM, Bai GQ, Meng X, Li ZH, Zhao**

**GF. 2020.** Phylogeography and population dynamics of an endemic oak (Quercus

fabri Hance) in subtropical China revealed by molecular data and ecological niche

modeling. *Tree Genetics Genomes* **16**: 1-132.

**Chen ZD, Ying JS, Lu AM. 2012.** Disjunct distribution of seed plants between southwestern

China and Taiwan Island island of China. *Chinese Journal of Botany* **47**: 551-570.

**Chou YW, Thomas PI, Ge XJ, LePage BA, Wang CN. 2011.** Refugia and phylogeography

of Taiwania in East Asia. *Journal of Biogeography* **38**: 1992-2005.

**Comes HP, Tribsch A, Bittkau C. 2008.** Plant speciation in continental island floras as

exemplified by *Nigella* in the Aegean Archipelago. *Philosophical Transactions of the*

*Royal Society B: Biological Sciences* **363**: 3083-3096.

**Cun YZ, Wang XQ. 2010.** Plant recolonization in the Himalayan from the southeastern

Qinghai-Tibetan Plateau: geographical isolation contributed to high population

differentiation. *Molecular Phylogenetics and Evolution* **56**: 972-982.

**Degnan JH, Rosenberg NA. 2009.** Gene tree discordance, phylogenetic inference and the

multispecies coalescent. *Trends in Ecology and Evolution* **24**: 332-340.

**Doyle JJ, Doyle J. 1987.** A rapid DNA isolation procedure for small quantities of fresh leaf

tissue. *Phytochemical Bulletin* **19**: 11-15.

**Drummond AJ, Suchard MA, Xie D, Rambaut A. 2012.** Bayesian phylogenetics with

BEAUti and the BEAST 1.7. *Molecular Biology Evolution* **29**: 1969-1973.

**Earl DA. 2012.** STRUCTURE HARVESTER: a website and program for visualizing

STRUCTURE output and implementing the Evanno method. *Conservation Genetics*

*Resources* **4**: 359-361.

**Evanno G, Regnaut S, Goudet J. 2005.** Detecting the number of clusters of individuals using

the software STRUCTURE: a simulation study. *Molecular Ecology* **14**: 261-

2620.

**Fu PC, Sun SS, Khan G, Dong XX, Tan JZ, Favre A, Zhang FQ, Chen SL. 2020.**

Population subdivision and hybridization in a species complex of Gentiana in the

Qinghai-Tibetan Plateau. *Annals of botany* **125**: 677-690.

**Foll M, Gaggiotti O. 2008.** A genome scan method to identify selected loci appropriate

for both dominant and codominant markers: a Bayesian perspective. *Genetics* **180**:

977-993.

**Gan X, Stegle O, Behr J, Steffen JG, Drewe P, Hildebrand KL, Lyngsoe R, Schultheiss**

**SJ, Osborne EJ, Sreedharan VT. 2011.** Multiple reference genomes and

transcriptomes for *Arabidopsis thaliana*. *Nature* **477**: 419-423.

**Givnish TJ, Renner SS. 2004.** Tropical intercontinental disjunctions: Gondwana breakup,

immigration from the boreotropics, and transoceanic dispersal. *International Journal*

*of Plant Sciences* **165**: S1-S6.

**Grant PR, Grant BR. 2019.** Hybridization increases population variation during adaptive

radiation. *Proceedings of the National Academy of Sciences* **116**: 23216-23224.

**Hahn C, Bachmann L, Chevreux B. 2013.** Reconstructing mitochondrial genomes directly

from genomic next-generation sequencing reads: a baiting and iterative mapping

approach. *Nucleic Acids Research* **41**: e129-e129.

**Huang S. 2011.** Historical biogeography of the flora of Taiwan. *Journal of the National*

*Taiwan Island Museum* **64**: 33-63.

**Hoang DT, Chernomor O. 2018.** UFBoot2: improving the ultrafast bootstrap approximation.

*Molecular Biology and Evolution* **35**: 518-522.

**Hodel RG, Chen SC, Payton AC, Mcdaniel SF, Soltis P, Soltis DE. 2017.** Adding loci

improves phylogeographic resolution in red mangroves despite increased missing

data: comparing microsatellites and RAD-Seq and investigating loci filtering.

*Scientific Reports* **7**: 17598.

**Jacobs B, Pyck N, Smets E. 2010.** Phylogeny of the Linnaea clade: are *Abelia* and *Zabelia*
693            closely related?. *Molecular Phylogenetics and Evolution* **57**: 741-752.

694 **Jakobsson M, Rosenberg NA. 2007.** CLUMPP: a cluster matching and permutation program
695            for dealing with label switching and multimodality in analysis of population structure.
696            *Bioinformatics* **23**: 1801-1806.

697 **Jiang XL, Hipp AL, Deng M, Su T, Zhou ZK, Yan MX. 2019.** East Asian origins of
698            European holly oaks (*Quercus* section *Ilex* Loudon) via the Tibet☐Himalaya. *Journal*
699            *of Biogeography* **46**: 2203-2214.

700 **Jin WY, Li HW, Wei R, Huang BH, Liu B, Sun TT, Mabberley DJ, Liao PC, Yang Y.**
701            **2020.** New insights into biogeographical disjunctions between Taiwan and the Eastern
702            Himalayas: The case of *Prinsepia* (Rosaceae). *Taxon* **69**: 278-289.

703 **Kalyaanamoorthy S, Minh BQ, Wong TK, Von Haeseler A, Jermiin LS. 2017.**
704            ModelFinder: fast model selection for accurate phylogenetic estimates. *Nature*
705            *Methods* **14**: 587-589.

706 **Katoh K, Standley DM. 2013.** MAFFT multiple sequence alignment software version 7:
707            improvements in performance and usability. *Molecular Biology and Evolution* **30**:
708            772-780.

709 **Kim TJ, Sun BY, Park CW, Suh Y. 1998.** Phylogenetic implications of *matK* sequences in
710            Caprifoliaceae. *American Journal of Botany* **48**: 1-48.

711 **Landrein S, Buerki S, Wang HF, Clarkson JJ. 2017.** Untangling the reticulate history of
712            species complexes and horticultural breeds in *Abelia* (Caprifoliaceae). *Annals of*
713            *Botany* **120**: 257-269.

714 **Landrein S, Farjon A. 2020.** A monograph of Caprifoliaceae: Linnaeeae. *Kew Bulletin* **74**:
715            1-197.

716 **Lanfear R, Calcott B, Ho SYW, Guindon S. 2012.** PartitionFinder: combined selection of
717            partitioning schemes and substitution models for phylogenetic analyses. *Molecular*
718            *Biology and Evolution* **29**: 1695-1701.

**Lanier HC, Knowles LL. 2012.** Is recombination a problem for species-tree analyses?.

*Systematic Biology* **61**: 691-701.

**Li J, Milne RI, Ru D, Miao J, Tao W, Zhang L, Mao K. 2000.** 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 H. 2011.** A statistical framework for SNP calling, mutation discovery, association

mapping and population genetical parameter estimation from sequencing data.

*Bioinformatics* **27**: 2987-2993.

**Li HL, Keng H. 1950.** Phytogeographical affinities of southern Taiwan. *Taiwania* **1**: 103-128.

**Li R, Wen J. 2016.** Phylogeny and diversification of Chinese Araliaceae based on nuclear

and plastid DNA sequence data. *Journal of Systematics and Evolution* **54**: 453-467.

**Librado P, Rozas J. 2009.** DnaSP v5: a software for comprehensive analysis of DNA

polymorphism data. *Bioinformatics* **25**: 1451-1452

**Liu YP, Ren ZM, Harris AJ, Peterson PM, Wen J, Su X. 2018.** Phylogeography of *Orinus*

(Poaceae), a dominant grass genus on the Qingha-Tibet Plateau. *Botanical Journal of*

*the Linnean Society* **186**: 202-223.

**Linder PH, Antonelli A, Humphreys AM, Pirie MD, Wüest RO. 2013.** What determines

biogeographical ranges? Historical wanderings and ecological constraints in the

danthonioid grasses. *Journal of Biogeography* **40**: 821-834.

**Link-Pérez MA, Dollo VH, Weber KM, Schussler EE. 2010.** What's in a Name:

Differential labelling of plant and animal photographs in two nationally syndicated

elementary science textbook series. *International Journal of Science Education* **32**:

1227-1242.

**Lu SY, Peng CI, Cheng YP, Hong KH, Chiang TY. 2001.** Chloroplast DNA

phylogeography of *Cunninghamia konishii* (Cupressaceae), an endemic conifer of

Taiwan. *Genome* **44**: 797-807.

**Manchester SR. 2000.** Update on the megafossil flora of Florissant, Colorado. *Denver*

*Museum of Nature and Science* **4**:137-162.

**Morales-Briones DF, Liston A, and Tank DC. 2018.** Phylogenomic analyses reveal a deep

history of hybridization and polyploidy in the Neotropical genus *Lachemilla*

(Rosaceae). New Phytologist **218**: 1668-1684.

**Morales-Briones DF, Gehrke B, Huang CH, Liston A, Ma H, Marx HE, Tank DC and**

**Yang Y. 2021.** Analysis of paralogs in target enrichment data pinpoints multiple

ancient polyploidy events in *Alchemilla* s.l. (Rosaceae). *Systematic Biology* syab032.

**Mummenhoff K, Linder P, Friesen N, Bowman JL, Lee JY, Franzke A. 2004.** Molecular

evidence for bicontinental hybridogenous genomic constitution in *Lepidium* sensu

stricto (Brassicaceae) species from Australia and New Zealand. *American Journal of*

*Botany* **91**: 254-261.

**Nathan R, Schurr FM, Spiegel O, Steinitz O, Trakhtenbrot A, Tsoar A. 2008.**

Mechanisms of long-distance seed dispersal. *Trends in ecology evolution* **23**:

638-647.

**Nguyen LT, Schmidt HA, von Haeseler A, Minh, BQ. 2015.** IQ-TREE: a fast effective

stochastic algorithm for estimating maximum-likelihood phylogenies. *Molecular*

*Biology and Evolution* **32**: 268-274.

**Nie ZL, Sun H, Li H, Wen J. 2007.** Intercontinental biogeography of subfamily

Orontioideae (*Symplocarpus*, *Lysichiton*, and *Orontium*) of Araceae in eastern Asia

and North America. *Molecular Phylogenetics and Evolution* **40**: 155-165.

**Niu YT, Ye JF, Zhang JL, Wan JZ, Yang T, Wei XX, Lu LM, Li JH, Chen ZD. 2018.**

Long-distance dispersal or postglacial contraction? Insights into disjunction between

Himalayan-Hengduan Mountains and Taiwan Island in a cold-adapted herbaceous

genus, *Triplostegia*. *Ecology and Evolution* **8**: 1131-1146.

**Nosil P. 2008.** Speciation with gene flow could be common. *Molecular ecology* **1**7:

2103-2106.

**Nwankwo EC, Mortega KG, Karageorgos A, Ogolowa BO, Papagregoriou G, Grether**

**GF, Ara Monadjem, Kirschel AN. 2019.** Rampant introgressive hybridization in

*Pogoniulus* tinkerbirds (Piciformes: Lybiidae) despite millions of years of divergence.

*Biological Journal of the Linnean Society* **127**: 125-142.

**Nylander JAA, Ronquist F, Huelsenbeck JP, Nieves-Aldrey JL. 2004.** Bayesian

phylogenetic analysis of combined data. *Systematic Biology* **53**: 47-67.

**Peakall R, Smouse PE. 2006.** GENALEX 6: genetic analysis in Excel. Population genetic

software for teaching and research. *Molecular Ecology Notes* **6**: 288-295.

**Phillips SJ, Anderson RP, Schapire RE. 2006.** Maximum entropy modeling of species

geographic distributions. *Ecological Modelling* **190**: 231-259.

**Pritchard JK, Stephens M, Donnelly P. 2000.** Inference of population structure using

multilocus genotype data. *Genetics* **155**: 945-959.

**Qian H, Ricklefs RE. 2004.** Geographical distribution and ecological conservatism of

disjunct genera of vascular plants in eastern Asia and eastern North America. *Journal

of Ecology* **92**: 253-265.

**Qiu YX, Fu CX, Comes HP, 2011.** Plant molecular phylogeography in China and adjacent

regions: Tracing the genetic imprints of Quaternary climate and environmental change

in the world's most diverse temperate flora. *Molecular Phylogenetics and Evolution* **59**:

225-244.

**Qiu YX, Guan BC, Comes HP. 2009a.** Did glacials and/or interglacials promote allopatric

incipient speciation in East Asia temperate plants? Phylogeographic and coalescent

analyses on refugial isolation and divergence in Dysosma versipellis. *Molecular

Phylogenetics and Evolution* **51**: 281-293.

**Rambaut A, Suchard MA, Xie D, Drummond AJ. 2014.** Tracer, version 1.6.

http://beast.bio.ed.ac.uk/Tracer/

**Rendón-Anaya M, Ibarra-Laclette E, Méndez-Bravo L, Lan T, Zheng C, Carretero-

Paulet L, Herrera-Estrella L. 2019.** The avocado genome informs deep angiosperm

phylogeny, highlights introgressive hybridization, and reveals pathogen-influenced gene

space adaptation. *Proceedings of the National Academy of Sciences* **116**: 17081-17089.

**Rieseberg LH, Carney SE.1998.** Tansley review no. 102 Plant hybridisation. *New Phytol

140**: 599-624.

**Šarhanová P, Sharbel TF, Sochor M, Vašut RJ, Dančák M and Trávníček B. 2017.**

Hybridization drives evolution of apomicts in *Rubus* subgenus *Rubus*: evidence from

microsatellite markers. *Annals of Botany* **120**: 317-328.

**Schwarz G. 1978.** Estimating the dimension of a model. *Annals of Statistics* **6**: 461-464

**Shao YZ, Chen Y, Zhang XC, Xiang QP. 2015.** Species delimitation and phylogeography of

the *Abies chensiensis* complex inferred from morphological and molecular data.

*Botanical Journal of the Linnean Society* **177**: 175-188.

**Shi Y, Cui Z, Su Z. 2006.** The Quaternary glaciations and environmental variations in China.

*Hebei Science and Technology Press* **65**: 101.

**Sibuet JC, Hsu SK, Le Pichon X, Le Formal JP, Reed D, Morre G, Liu CS. 2002.** East

Asia plate tectonics since 15 Ma: constraints from the Taiwan Island region.

*Tectonophysics* **344**: 103-134.

**Sibuet JC, Hsu SK. 2004.** How was Taiwan Island created?. *Tectonophysics* **379**: 159-181.

**Smith SA, Dunn CW. 2008.** Phyutility: a phyloinformatics tool for trees, alignments and

molecular data. *Bioinformatics* **24**: 715-716.

**Sugiura N. 1978.** Further Analysis of the Data by Akaike's Information Criterion and the

Finite Corrections. *Communications in Statistics-Theory and Methods* **7**: 13-26.

**Suo YH, Li SZ, Zhao SJ, Somerville ID, Yu S, Dai LM, Xu LQ, Cao XZ, Wang PC. 2015.**

Continental margin basins in East Asia: tectonic implications of the Meso- Cenozoic

East China Sea pull-apart basins. *Geological Journal* **50**: 139-156.

**Swets JA. 1988.** Measuring the accuracy of diagnostic systems. *Science* **240**: 1285-1293.

**Teng LS. 1992.** Geotectonic evolution of Tertiary continental margin basins of Taiwan.

*Petroleum Geology of Taiwan Island* **27**: 1-19.

**Vargas OM, Heuertz M, Smith SA, Dick CW. 2019.** Target sequence capture in

the Brazil nut family (Lecythidaceae): Marker selection and in silico capture

from genome skimming data. *Molecular Phylogenetics and Evolution* **135**:

98-104.

**Wallich N. 1829.** *Plantae Asiaticae rariores*. London, Paris, Strasburgh: Treuttel and Würtz

**Wang HF, Landrein S, Dong WP, Nie ZL, Kondo K, Funamoto T, Wen J, Zhou SL. 2015.**

Molecular phylogeny and biogeographic diversification of Linnaeoideae

(*Carprifoliaceae s.l.*) disjunctively distributed in Eurasia, North American and Mexico.

*PLoS One* **10**: e0116485.

**Wang HX, Liu H, Moore MJ, Landrein S, Liu B, Zhu ZX, Wang HF. 2020.** Plastid

phylogenomic insights into the evolution of the Caprifoliaceae *s.l.* (Dipsacales).

*Molecular Phylogenetics and Evolution* **142**: 106641.

**Wang HX, Morales-Brionesb DF, Moorec MJ, Wen J, Wang HF. 2021.** A phylogenomic

perspective on gene tree conflict and character evolution in *Caprifoliaceae* using target

enrichment data, with Zabelioideae recognized as a new subfamily. *Journal of*

*Systematics and Evolution* https://doi.org/10.1111/jse.12745.

**Wang IJ. 2013.** Examining the full effects of landscape heterogeneity on spatial genetic

variation: a multiple matrix regression approach for quantifying geographic and

ecological isolation. *Evolution* **67**: 3403-3411.

**Wang J, Feng C, Jiao T, Von Wettberg EB, Kang M. 2017.** Genomic Signature of Adaptive

Divergence despite Strong Nonadaptive Forces on Edaphic Islands: A Case Study of

*Primulina juliae*. *Genome Biology and Evolution* **9**: 3495-3508.

**Wang W, Ortiz RDC, Jacques FM, Xiang XG, Li H., Lin L, Chen ZD. 2012.**

Menispermaceae and the diversification of tropical rainforests near the Cretaceous-

Paleogene boundary. *New Phytologist* **195**: 470-478.

**Wang Z, Liu KJ. 2016.** A performance study of the impact of recombination on species tree

analysis. *BMC Genomics* **17**: 165-174.

**Wei XX, Yang ZY, Li Y, Wang XQ. 2010.** Molecular phylogeny and

biogeography of Pseudotsuga (Pinaceae): insights into the floristic relationship

between Taiwan Island and its adjacent areas. *Molecular Phylogenetics and Evolution*

**55**: 776-785.

**Weitemier K, Straub SCK, Cronn RC, Fishbein M, Schmickl R, McDonnell A, Liston A.**

**2014.** Hyb-Seq: combining target enrichment and genome skimming for plant

phylogenomics. *Applications in Plant Sciences* **2**: 1400042.

**Wen J. 2001.** Evolution of eastern Asian-eastern North American biogeographic disjunctions:

a few additional issues. *International Journal of Plant Sciences* **162**: S117-S122.

**Wen J, Nie ZL, Ickert☐Bond SM. 2018.** Advances in biogeography in the age of a new

modern synthesis. *Journal of Systematics and Evolution* **57**: 543-546.

**Wen J, Ree RH, Ickert-Bond SM, Nie Z, Funk V. 2013.** Biogeography: where do we go

from here?. *Taxon* **62**: 912-927.

**Wen J, Wagner WL. 2020.** Collections-based systematics and biogeography in the 21st

century. *Journal of Systematics and Evolution* **58**: 743-750.

**Wu PC, Wang MZ. 2001.** Relationship of the tropical elements of the bryophytes between

Hengduan Mts., southwest China and Taiwan Island province, southeast China.

*Guizhou Science* **19**: 6-9.

**Xiang CL, Dong HJ, Landrein S, Zhao F, Yu WB, Soltis DE, Peng H. 1998.** Revisiting the

phylogeny of Dipsacales: new insights from phylogenomic analyses of complete

plastomic sequences. *Journal of Systematics and Evolution* **58**: 103-117.

**Xiang KL, Zhao L, Erst AS, Yu SX, Jabbour F, Wang W. 2017.** A molecular phylogeny of

*Dichocarpum* (Ranunculaceae): Implications for eastern Asian biogeography.

*Molecular Phylogenetics and Evolution* **107**: 594-604.

**Xiang QY, Soltis DE. 2001.** Dispersal-vicariance analyses of intercontinental disjuncts:

historical biogeographical implications for angiosperms in the Northern Hemisphere.

*International Journal of Plant Sciences* **162**: S29-S39.

**Xu J, Deng M, Jiang XL, Westwood M, Song YG, Turkington R. 2015.** Phylogeography

of *Quercus glauca* (Fagaceae), a dominant tree of East Asian subtropical evergreen

forests, based on three chloroplast DNA interspace sequences. *Tree Genet Genomes*

**11**: 805.

**Yan HF, Zhang CY, Wang FY, Hu, CM, Ge XJ, Hao G. 2012.** Population expanding with

the phalanx model and lineages split by environmental heterogeneity: a case study of

*Primula obconica* in subtropical China. *PLoS One* **7**: e41315.

**Yang Q, Landrein S. 2011.** *Linnaeaceae. Flora of China* **19**: 642-648.

**Yang Y, Smith SA. 2014.** Orthology inference in nonmodel organismsusing transcriptomes

and low-coverage genomes: improving accuracy and matrix occupancy for

phylogenomics. *Molecular Biology and Evolution* **31**: 3081-3092.

**Ye J, Lu L, Liu B, Yang T, Zhang J, Hu H, Chen Z. 2019.** Phylogenetic delineation of

regional biota: A case study of the Chinese flora. *Molecular Phylogenetics and*

*Evolution* **135**: 222-229.

**Ye XQ. 1982.** On the formation and development of the geology and geomorphology of

Taiwan. *Journal Central China Teachers College* 83-89.

**Ye XY, Ma PF, Yang GQ, Guo C, Zhang YX, Chen YM, Li DZ. 2019.** Rapid

diversification of alpine bamboos associated with the uplift of the Hengduan

Mountains. *Journal of Biogeography* **46**: 2678-2689.

**Ye Z, Zhu G, Chen P, Zhang D, Bu W. 2014.** Molecular data and ecological niche

modelling reveal the Pleistocene history of a semi-aquatic bug (*Microvelia douglasi*

*douglasi*) in East Asia. *Molecular Ecology* **23**: 3080-3096.

**Yin A, Harrison TM. 2000.** Geologic evolution of the Himalayan Tibetan orogen. *Annual*

*Review of Earth and Planetary Sciences* **28**: 211-280.

**Yoder AD, Nowak MD. 2006.** Has vicariance or dispersal been the predominant

biogeographic force in Madagascar? Only time will tell. *Ecology Evolution and*

*Systematics* **37**: 405-431.

**Yu Y, Harris AJ, Blair C, He X. 2015.** RASP (Reconstruct Ancestral State in Phylogenies):

a tool for historical biogeography. *Molecular Phylogenetics and Evolution* **87**: 46-49.

**Zachos J, Pagani M, Sloan L, Thomas E, Billups K. 2001.** Trends, rhythms, and

aberrations in global climate 65 Ma to present. *Science* **292**: 686-693.

**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.

**Figure Legends**

Figure 1. Photographs of *Abelia* taxa. (A) *A. chinensis*; (B) *A. chinensis* var. *ionandra*; (C) *A.*

*forrestii*;(D) *A.* × *grandiflora*; (E) *A.* × *grandiflora* 'Francis Mason'; (F) *A. macrotera*;

(G) *A. macrotera* var. *macrotera*; (H) *A. uniflora*; (I) *A. schumannii.*

Figure 2. Tanglegram of the nuclear concatenated (left) and plastid (right) phylogenies.

Dotted lines connect taxa between the two phylogenies. Maximum likelihood bootstrap

support values are shown above branches. The asterisks indicate maximum likelihood

bootstrap support of 100%. Major taxonomic groups or main clades in the family as

currently recognized are indicated by branch colors as a visual reference to relationships.

Figure 3. ASTRAL-III species tree; Clade support is depicted as: Quartet Concordance

(QC)/Quartet Differential (QD)/Quarte Informativeness (QI)/Local posterior

probabilities (LPP).

Figure 4. Best supported species networks inferred with PhyloNet for the (A) 15 samples, (B)

12 samples. Numbers next to the inferred hybrid branches indicate inheritance

probabilities. Red lines represent minor hybrid edges (edges with an inheritance

contribution < 0.50).

Figure 5. Distribution of 12 Hyb-seq haplotypes in *Abelia.* (A) Geographical distribution of

haplotypes among four *Abelia* species, (B) The network of 15 chloroptypes, (C) Results

of STRUCTURE analysis for all individuals of *Abelia* studied based on *Fst* data.

Figure 6. BEAST analysis of divergence times based on nuclear data. Calibration points are

indicated by A, B. Numbers 1-13 represent major divergence events in *Abelia*; mean

divergence times and 95% highest posterior densities are provided for each.

Figure 7. Ancestral area reconstruction for *Abelia*. Areas of endemism are as follows: (A)

Southwest China, (B) Central-East China, (C) Taiwan and Ryukyu Islands. The

numbered nodes represent crown nodes of important colonization events.

Figure 8. Comparison of potential distributions as probability of occurrence for *A. chinensis*,

and *A. uniflora* at the CCSM climatic scenarios of the Last Glacial Maximum (LGM, ca.

21,000 years BP). The maximum training sensitivity plus specificity logistic threshold

has been used to discriminate between suitable (cutline 0.1-1 areas) and unsuitable

habitat. The darker color indicates a higher probability of occurrence.

Fig. S1 Heatmaps showing gene recovery efficiency for the nuclear gene in 45 species of

*Abelia*. Columns represent genes, and each row is one accession. Shading indicates the

percentage of the reference locus length coverage.

Fig. S2 Nuclear concatenated RAxML tree; Clade support is depicted as: Quartet

Concordance (QC)/Quartet Differential (QD)/Quarte Informativeness (QI).

Fig. S3 Plastid RAxML tree; Clade support is depicted as: Quartet Concordance (QC)/Quartet

Differential (QD)/Quarte Informativeness (QI)/Local posterior probabilities (LPP).

Fig. S4 Best species networks of the selective nuclear data estimated with PhyloNet for the

15-sample data set. A: One hybridization event; B: Two hybridization events; C: Three

hybridization events; D: Four hybridization events; E: Five hybridization events. Blue

branches connect the hybrid nodes. Numbers next to blue branches indicate inheritance

probabilities.

Fig. S5 Best species networks of the selective nuclear data estimated with PhyloNet for the

12-sample data set. A: One hybridization event; B: Two hybridization events; C: Three

hybridization events; D: Four hybridization events; E: Five hybridization events. Blue

branches connect the hybrid nodes. Numbers next to blue branches indicate inheritance

probabilities.

Fig. S6 Corresponding value of *Fst*.

Fig. S7. BEAST analysis of divergence times based on the plastid matrix. Calibration points

are indicated by A, B. Numbers 1-8 represent major divergence events in *Abelia*; mean

divergence times and 95% highest posterior densities are provided for each.

Fig. S8 Potential distribution probability (in logistic value) of occurrence for *Abelia* in East

Asia generated form MAXENT.

Table. S1 The 19 bioclimatic variables for ENM analyses

Table. S2 HybPiper assembly statistics.

Table. S3 Summary of genetic diversity of *Abelia*.

**Table 1. List of species and vouchers used in this study.**

| Order | Ingroup/ Outgroup | Taxon | Coding | Locality | Voucher specimen number | SRA accession |
|---|---|---|---|---|---|---|
| 1 | Ingroup | *Abelia chinensis* | B | Jiujiang, Jiangxi, China | HUTB, E119 | SRR13890741 |
| 2 | Ingroup | *Abelia chinensis* | B | Jiujiang, Jiangxi, China | HUTB, E111 | SRR13890742 |
| 3 | Ingroup | *Abelia chinensis* | B | Jiujiang, Jiangxi, China | HUTB, E206 | SRR13705765 |
| 4 | Ingroup | *Abelia chinensis* var. *ionandra* | C | Keelung Islet,Taiwan, China | HUTB, E30 | SRR13705776 |
| 5 | Ingroup | *Abelia chinensis* var. *ionandra* | C | Amami Islands, Japan | HUTB, E33 | SRR13890725 |
| 6 | Ingroup | *Abelia chinensis* var. *ionandra* | C | Okinawa Island, Okinawa, Japan | HUTB, E166 | SRR13890726 |
| 7 | Ingroup | *Abelia chinensis* var. *ionandra* | C | Yilan, Taiwan, China | HUTB, E167 | SRR13890744 |
| 8 | Ingroup | *Abelia chinensis* var. *ionandra* | C | Yilan,Taiwan, China | HUTB, E229 | SRR13890743 |
| 9 | Ingroup | *Abelia chinensis* var. *ionandra* | C | Yilan, Taiwan, China | HUTB, E230 | SRR13890751 |
| 10 | Ingroup | *Abelia chinensis* var. *ionandra* | C | Yilan, Taiwan, China | HUTB, E231 | SRR13890745 |
| 11 | Ingroup | *Abelia forrestii* | A | Nujiang, Yunnan, China | HUTB, E36 | SRR13890748 |
| 12 | Ingroup | *Abelia forrestii* | A | Nujiang, Yunnan, China | HUTB, E37 | SRR13705787 |
| 13 | Ingroup | *Abelia grandiflora* | B | Zhengzhou, Henan, China | HUTB, E159 | SRR13890724 |
| 14 | Ingroup | *Abelia grandiflora* | B | Zhengzhou, Henan, China | HUTB, E227 | SRR13890730 |
| 15 | Ingroup | *Abelia macrotera* var. *mairei* | A | Emei, Sichuan, China | HUTB, E110 | SRR13705799 |
| 16 | Ingroup | *Abelia macrotera* var. *deutziaefolia* | A | Fuyang, Guizhou, China | HUTB, E153 | SRR13890758 |
| 17 | Ingroup | *Abelia macrotera* var. *deutziaefolia* | A | Fuyang, Guizhou, China | HUTB, E154 | SRR13890757 |
| 18 | Ingroup | *Abelia macrotera* var. *deutziaefolia* | A | Fuyang,Guizhou, China | HUTB, E155 | SRR13890727 |
| 19 | Ingroup | *Abelia macrotera* var. *deutziaefolia* | A | Fuyang, Guizhou, China | HUTB, E156 | SRR13890731 |
| 20 | Ingroup | *Abelia macrotera* var. *deutziaefolia* | A | Fuyang, Guizhou, China | HUTB, E157 | SRR13890728 |
| 21 | Ingroup | *Abelia macrotera* var. *deutziaefolia* | A | Fuyang, Guizhou, China | HUTB, E158 | SRR13890729 |
| 22 | Ingroup | *Abelia macrotera* var. *engleriana* | A | Nanchuan, Chongqing, China | HUTB, E34 | SRR13890756 |
| 23 | Ingroup | *Abelia macrotera* var. *engleriana* | A | Nanchuan, Chongqing, China | HUTB, E149 | SRR13890753 |
| 24 | Ingroup | *Abelia macrotera* var. *engleriana* | A | Nanchuan, Chongqing, China | HUTB, E347 | SRR13890755 |
| 25 | Ingroup | *Abelia macrotera* var. *engleriana* | A | Nanchuan, Chongqing, China | HUTB, E349 | SRR13890754 |

| 26 | Ingroup | *Abelia macrotera* var. *zabelioides* | A | Emei, Sichuan, China | HUTB, E3 | SRR13890733 |
| 27 | Ingroup | *Abelia macrotera* var. *zabelioides* | A | Emei, Sichuan, China | HUTB, E43 | SRR13890740 |
| 28 | Ingroup | *Abelia macrotera* var. *zabelioides* | A | Emei, Sichuan, China | HUTB, E211 | SRR13890732 |
| 29 | Ingroup | *Abelia macrotera* var. *macrotera* | A | Yichang, Hubei, China | HUTB, E6 | SRR13890752 |
| 30 | Ingroup | *Abelia macrotera* var. *macrotera* | A | Yichang, Hubei, China | HUTB, E45 | SRR13890750 |
| 31 | Ingroup | *Abelia uniflora* | B | Wuyishan, Fujian, China | HUTB, E46 | SRR13890738 |
| 32 | Ingroup | *Abelia uniflora* | B | Wuyishan, Fujian, China | HUTB, E47 | SRR13890739 |
| 33 | Ingroup | *Abelia uniflora* | B | Wuyishan, Fujian, China | HUTB, E48 | SRR13890734 |
| 34 | Ingroup | *Abelia uniflora* | B | Wuyishan, Fujian, China | HUTB, E51 | SRR13705798 |
| 35 | Ingroup | *Abelia uniflora* | B | Wuyishan, Fujian, China | HUTB, E52 | SRR13890736 |
| 36 | Ingroup | *Abelia uniflora* | B | Wuyishan, Fujian, China | HUTB, E53 | SRR13890735 |
| 37 | Ingroup | *Abelia uniflora* | B | Wuyishan, Fujian, China | HUTB, E54 | SRR13890737 |
| 38 | Ingroup | *Abelia schumannii* | A | Ganzi, Sichuan, China | HUTB, E44 | SRR13890749 |
| 39 | Outgroup | *Diabelia serrata* | | Takakuma, Kagoshima, Japan | HUTB, E321 | SRR13890747 |
| 40 | Outgroup | *Diabelia sanguinea* | | Numadzu, Shizuoka, Japan | HUTB, E303 | SRR13890746 |
| 41 | Outgroup | *Dipelta floribunda* | | Cambridge, UK. | HUTB, E55 | SRR13705794 |
| 42 | Outgroup | *Dipelta floribunda* | | Cambridge, UK. | HUTB, E56 | SRR13705795 |
| 43 | Outgroup | *Kolkwitzia amabilis* | | Weinan, Shanxi, China | HUTB, E9 | SRR13705797 |
| 44 | Outgroup | *Linnaea borealis* | | Yili, Xinjiang, China | HUTB, E14 | SRR13705786 |
| 45 | Outgroup | *Vesalea occidentalis* | | Durango, Mexico | HUTB, E96 | SRR13705793 |

**Table 2. Dataset statistics, including the number of samples, number of characters, number of PI characters, missing data.**

| Alignment | No. of samples | No. sites | Missing data (%) | No. of variable/Parsimony informative sites |
|---|---|---|---|---|
| Nuclear | 45 | 365,330 | 16.88 | 43,248/22,045 |
| Plastid | 45 | 163,130 | 4.35 | 7,926/3,850 |

**Table 3. Model selection of the different species networks and bifurcating trees.**

| Topology | lnL | Parameters | Loci | Number of hybridizations | Information criteria | | |
|---|---|---|---|---|---|---|---|
| | | | | | AIC | AICc | BIC |
| **15 samples** | | | | | | | |
| nuclear ASTRAL | -4177.949631 | 27 | 167 | N/A | 8409.899262 | 8420.776960 | 8494.085095 |
| nuclear RAxML | -4148.054160 | 27 | 167 | N/A | 8350.108319 | 8360.986017 | 8434.294152 |
| Plastid | -4231.190608 | 27 | 167 | N/A | 8516.381215 | 8527.258913 | 8600.567048 |
| Network 1 | -4088.228145 | 29 | 167 | 1 | 8234.456290 | 8247.157020 | 8324.878111 |
| Network 2 | -4050.799135 | 31 | 167 | 2 | 8163.598270 | 8178.294566 | 8260.256078 |
| Network 3 | -3948.351402 | 33 | 167 | 3 | 7962.702804 | 7979.574984 | 8065.596600 |
| **Network 4** | **-3932.455932** | **35** | **167** | **4** | **7934.911864** | **7954.148505** | **8044.041647** |
| Network 5 | -3935.844300 | 37 | 167 | 5 | 7945.688601 | 7967.487050 | 8061.054372 |
| **12 samples** | | | | | | | |
| nuclear ASTRAL | -2716.350474 | 21 | 162 | N/A | 5474.700949 | 5481.300949 | 5539.540472 |
| nuclear RAxML | -2724.448876 | 21 | 162 | N/A | 5490.897753 | 5497.497753 | 5555.737276 |
| Plastid | -2723.043552 | 21 | 162 | N/A | 5488.087104 | 5494.687104 | 5552.926627 |
| Network1 | -2660.433072 | 23 | 162 | 1 | 5366.866143 | 5374.866143 | 5437.880859 |
| Network2 | -2601.023466 | 25 | 162 | 2 | 5252.046932 | 5261.605756 | 5329.236841 |
| **Network3** | **-2549.319250** | **27** | **162** | **3** | **5152.638500** | **5163.922082** | **5236.003601** |
| Network4 | -2552.995453 | 27 | 162 | 3 | 5159.990906 | 5171.274488 | 5243.356007 |
| Network 5 | -2606.591029 | 27 | 162 | 3 | 5267.182058 | 5278.465640 | 5350.547159 |

**A**

- *A. forrestii* E36
- *A. macrotera* var. *macrotera* E45
- *A. schumannii* E44
- *A. macrotera* var. *deatziaefolia* E154
- *A. macrotera* var. *deutziaefolia* E157
- *A. macrotera* var. *engleriana* E149
- *A. uniflora* E48
- *A. macrotera* var. *zabelioides* E43
- *A. uniflora* E46
- *A. × grandiflora* E159
- *A. × grandiflora* E227
- *A. chinensis* E206
- *A. chinensis* E111
- *A. chinensis* var. *ionandra* E33
- *D. serrata* E321

0.428
0.536
0.572
0.259
0.53
0.741
0.464
0.47

**B**

- *A. uniflora* E48
- *A. uniflora* E46
- *A. macrotera* var. *zabelioides* E43
- *A. macrotera* var. *engleriana* E149
- *A. macrotera* var. *deutziaefolia* E157
- *A. macrotera* var. *deatziaefolia* E154
- *A. schumannii* E44
- *A. macrotera* var. *macrotera* E45
- *A. forrestii* E36
- *A. chinensis* E111
- *A. chinensis* var. *ionandra* E33
- *D. serrata* E321

0.661
0.339
0.762
0.238
0.448
0.552

**(A)**

**(B)**

*A. macrotera*

*A. chinensis*

*A. uniflora*

A: *A. macrotera*    B: *A. forrestii*

**(C)**

*A. macrotera*      *A. uniflora*      *A. forrestii*      *A. chinensis*      *A. × grandiflora*

(A)*A. macrotera* var. *deutziaefolia* E156
(A)*A. macrotera* var. *deutziaefolia* E157
(A)*A. macrotera* var. *deutziaefolia* E158
(A)*A. macrotera* var. *deutziaefolia* E155
(A)*A. macrotera* var. *deatziaefolia* E153
(A)*A. macrotera* var. *deatziaefolia* E154
(A)*A. macrotera* var. *engleriana* E347
(A)*A. macrotera* var. *engleriana* E149
(A)*A. macrotera* var. *engleriana* E349
(A)*A. macrotera* var. *macrotera* E6
(A)*A. macrotera* var. *engleriana* E34
(A)*A. schumannii* E44
(A)*A. macrotera* var. *mairei* E110
(A)*A. macrotera* var. *macrotera* E45
(A)*A. forrestii* E36
(A)*A. forrestii* E37
(B)*A. uniflora* E46
(B)*A. uniflora* E54
(B)*A. uniflora* E47
(B)*A. uniflora* E53
(B)*A. uniflora* E52
(B)*A. uniflora* E51
(A)*A. macrotera* var. *zabelioides* E3
(A)*A. macrotera* var. *zabelioides* E43
(A)*A. macrotera* var. *zabelioides* E211
(B)*A. uniflora* E48
(C)*A. chinensis* var. *ionandra* E231
(C)*A. chinensis* var. *ionandra* E230
(C)*A. chinensis* var. *ionandra* E167
(C)*A. chinensis* var. *ionandra* E30
(A)*A. chinensis* var. *ionandra* E229
(A)*A. chinensis* var. *ionandra* E33
(C)*A. chinensis* var. *ionandra* E166
(B)*A. chinensis* E111
(B)*A. chinensis* E119
(B)*A. chinensis* E206
(B)*A. × grandiflora* E159
(B)*A. × grandiflora* E206

Legend
*
A
AB
B
BC
C

⭕ Dispersal
⚫ Vicariance

47 45    36      27      18      9      0