A complex pattern of postdivergence expansion, contraction, introgression and asynchronous responses to Pleistocene climate changes in two Dipelta sister species from western China

Research Article

Title

A complex pattern of post-divergence expansion, contraction, introgression and asynchronous responses to Pleistocene climate changes in two *Dipelta* sister species from western China

Running title

Formation of allopatry of two *Dipelta* species

Bin Tian[1,2], Yi Fu[2], Richard I. Milne[3], Kangshan Mao[4], Yongshuai Sun[5,6*], Xiangguang Ma[1], Hang Sun[1*]

[1]Key Laboratory for Plant Diversity and Biogeography, Kunming Institute of Botany, Chinese Academy of Sciences, Kunming 650204, China

[2]Key Laboratory of Biodiversity Conservation in Southwest China, State Forestry Administration, Southwest Forestry University, Kunming 650224, China

[3]Institute of Molecular Plant Sciences, University of Edinburgh, Edinburgh EH9 3JH, UK

[4]Key Laboratory of Bio-Resource and Eco-Environment of Ministry of Education, College of Life Sciences, Sichuan University, Chengdu 610065, China

[5]Key Laboratory of Tropical Forest Ecology, Xishuangbanna Tropical Botanical Garden, Chinese Academy of Sciences, Mengla 666303, Yunnan, China

[6]Centre for Plant Ecology, Core Botanical Gardens, Chinese Academy of Sciences, Mengla 666303, Yunnan, China

*Author for correspondence. E-mail: sunhang@mail.kib.ac.cn, sunyongshuai@xtbg.ac.cn

Received 13 November 2019; Accepted 11 June 2019

Abstract

The well-known vicariance and dispersal models dominate in understanding the allopatric pattern for related species and presume the simultaneous occurrence of speciation and biogeographic events. However, the formation of allopatry may postdate the species divergence. We examined this hypothesis using DNA sequence data from 3 chloroplast fragments and 5 nuclear loci of *Dipelta floribunda* and *D. yunnanensis*, two shrub species with the circum Sichuan Basin distribution, combining the climatic niche modeling approach. The best-fit model supported by the approximate Bayesian computation (ABC) analysis indicated that, *D. floribunda* and *D. yunnanensis* diverged during the mid-Pleistocene period, consistent with the largest glacial period in the Qinghai-Tibet Plateau (QTP). The historically inter-specific gene flow was identified, but seemed to have ceased after the last interglacial period (LIG), when the range of *D. floribunda* moved northward from the south of the Sichuan Basin. Further, populations of *D. floribunda* had expanded obviously in the north of the Sichuan Basin after the last glacial maximum (LGM). Relatively, the range of *D. yunnanensis* expanded before the LGM, and reduced during the post-LGM especially in the north of the Sichuan Basin, reflecting the asynchronous responses of related species to the contemporary climate changes. Our results suggested that complex topography should be considered in understanding the distributional patterns even for closely related species and their demographic responses.

Keywords: allopatric pattern, asynchronous demographic responses, hABC, introgression, the Pleistocene climate change, the Sichuan Basin

1 Introduction

It is important to test hypotheses regarding to the biogeographic drivers and processes of the distribution pattern for related species (Macarthur, 1972; Crisp *et al.,* 2011; Usinowicz *et al.,* 2017). The vicariance and long-distance dispersal (LDD) models which are tested and employed in many studies, provide two major hypotheses that can explain disjunct distributions (Ball, 1975; Crisp *et al.,* 2011). In the vicariance model, geographic barriers develop and divide a large population into separate parts, and prevent gene flow between them (Ball, 1975). In contrast, the dispersal model requires that organisms overcome geographic barriers to migration and establish new populations on the other side of that barrier (Nathan, 2008). Both will ultimately lead to allopatric speciation, hence; the speciation event should be timed as occurring soon after the disjunction was established. However, for plants that occur in regions with complex topographies, neither of these two traditional paradigms may be appropriate to explain the pattern of allopatry. Complex topography can provide ecological gradients, for example along mountain slopes (Badgley *et al.,* 2017), and in theory parapatric speciation may occur along such a gradient, followed by allopatry at a later time if one or both species then moves its native range. In the present study, we aim to test this hypothesis using a case study of two shrub species of *Dipelta* endemic to western China.

Mountains and valleys which formed accompanying with the uplift of the Qinghai-Tibet Plateau (QTP) in western China (Clark *et al.,* 2005; Wang *et al.,* 2012, 2014), can restrict dispersal and lead to species divergence and allopatric pattern (Endler 1977; Smith *et al.,* 2014; Steinbauer *et al.,* 2016). Recent studies suggest that the uplifts of the QTP and adjacent mountains has played important roles in the diversification of highland plants in western China (Wang *et al.,* 2005; Qiu *et al.,* 2011; Wen *et al.,* 2014; Favre *et al.,* 2015; Sun *et al.,*

2017; Xing & Ree, 2017). The landscape complexity provides steep ecological gradients along the mountain slope (Favre *et al.,* 2015; Liu *et al.,* 2014), and potential refugia for plants during climatic extremes such as the Pleistocene glacial cycles. It therefore provides opportunities for both retaining high levels of plant diversity, and generating new lineages (Qiu *et al.,* 2011; Liu *et al.,* 2012). However, the biogeographical processes underlying divergence are still unclear for most plants here, restricting our ability to explain the origin and distribution of plant diversity in western China (Liu *et al.,* 2014).

Dipelta Maxim. (Caprifoliaceae) includes three species endemic to the west of China (Fig. 1; Table 1), and occur in mid/high-elevation montane forests but never in valleys (Yang & Landrein, 2011). *Dipelta elegans* Batal. is an endangered species and thus was not focused presently. Populations of *Dipelta floribunda* Maxim. and *D. yunnanensis* Franch constitute a near-circular distribution surrounding the Sichuan Basin (Fig. 1), a region with an area larger than 260,000 m^2 and an elevation ~400 m at the bottom of the basin. *Dipelta floribunda* occurs around the Qin-Ba Mountains to the north and east of the Sichuan Basin, whereas *D. yunnanensis* occurs to the south and west of the basin, in most of the Hengduan Mountains (Fig. 1). The annual mean temperature and precipitation in the habitats of both species are much lower than those at the bottom of the Sichuan Basin (Wang *et al.,* 2013), constituting a geographic barrier to their dispersal. Althrough the distribution of these two species is close (nearest population less 100km), our three-years field investigations (2015-2017) haven't found hybrids and contact zones between species. These two species were combined into a system to investigate the biogeographic role of a medium scale geographic barrier, specifically a climatically unavailable low-altitude region, in the formation of a local disjunction between closely related species.

97 Three hypotheses might explain how this species pair speciated and how the allopatric
98 pattern formed. First, as the vicariance model suggested (Ball, 1975; Crisp *et al.,* 2011), if the
99 formation of the Sichuan Basin had driven the initial divergence between *D. floribunda* and *D.*
100 *yunnanensis*, the divergence time would be consistent with the formation of the Sichuan Basin
101 during the Neogene period (Shi *et al.,* 1998; Clark *et al.,* 2005; Wang *et al.,* 2012; He *et al.,*
102 2013; Wang *et al.,* 2014). Second, species divergence could have been initiated through a
103 dispersal event from one side of the Sichuan Basin to the other, as the LDD model proposed
104 (Nathan, 2008; Crisp *et al.,* 2011). The divergence time thus would be later than the formation
105 of the basin but consistent with the LDD event. In the 1st and 2nd models, the formation of
106 allopatric pattern would be in synchrony with the divergence. Third, speciation occurred
107 without the present geographic barriers, possibly by local ecological speciation along a
108 gradient within the western China. If so, a consistent difference in ecological preference
109 between the species would exist, and these different preferences may form before the
110 formation of allopatric pattern. In the 3rd model, the initial divergence between *D. floribunda*
111 and *D. yunnanensis* could be independent of the formation of the Sichuan Basin and the
112 allopatric pattern.
113 To evaluate and compare these hypotheses, we aimed to assess the following questions: 1)
114 When did the divergence between *D. floribunda* and *D. yunnanensis* occur? 2) Are there
115 ecological differences between the two species? 3) What role did the Sichuan Basin play in
116 the formation of allopatric pattern? 4) Has the basin influenced their responses to climatic
117 changes, after their divergence?
118 **2 Material and Methods**
119 **2.1 Sampling**

We collected a total of 547 individuals from 56 populations throughout the ranges of *D. floribunda* and *D. yunnanensis* (Table 1). The number of individuals collected from each population was between 1 and 20, and these were always spaced at least 100m apart. Fresh leaves were collected and dried immediately using silica gel. In addition, *Dipelta elegans*, *Diabelia serrata*（Sieb. & Zucc.) Land. and *Kolkwitzia amabilis* Graebn. were collected and used as outgroups in our analyses below. All voucher specimens collected from each population were deposited in Southwest Forestry University Herbarium (SWFC). The latitude, longitude and altitude of each sampling site were recorded using an eTrex GPS (Garmin).

2.2 DNA extraction, amplification and sequencing

We used EZ-10 Spin Column Plant Genomic DNA Purification Kits (Sangon Biotech, Shanghai, China) to extract total genomic DNA from 547 individuals, and 1% agarose gels was used for testing the quantity of genomic DNA isolating from the all of individuals.

To identify cpDNA regions with sufficient variation, we randomly selected 12 shrubs of *D. yunnanensis* and 12 of *D. floribunda* to conduct preliminary screening of primer pairs for three highly variable regions: *psbA-trnH*, *psbB-psbF* and *trnL-trnF*. All were determined to be useful. For nuclear markers, fresh leaves of *D. yunnanensis* were gathered in Lijiang (population YL) to transcriptome sequencing. The sequencing was performed on HiSeq sequencing platforms at BGI-Shenzhen. For further details on RNA extractions, transcriptome sequencing, and assembly, see Ju et al. 2015. Then we developed primers of 5 species-specific low-copy nuclear loci (23311, 38541, 41398, 45367, 56546) following the procedures described by Ye et al. (2017). Therefore, a total of eight DNA fragments from chloroplast and nuclear genomes were sequenced to determine the genetic variation of *D. floribunda* and *D. yunnanensis* (Table S1).

Polymerase chain reaction experiments were performed using the S1000 Thermal Cycler (Applied Biosystems, Foster City, California, USA) in a volume of 25 μL containing 1 μL (~10 ng) DNA template, 12.5 μL Taq PCR Mix (Sangon Biotech), 9.5 μL double-distilled H2O, and 1 μL (5 pmol) of each primer. The PCR program consisted of 5 min of initial denaturation at 94 °C; followed by 30 cycles of denaturation at 94 °C for 30 s, annealing at specific temperature (52 °C-58 °C, Table S1) for 45 s, extension at 72 °C for 1 min, and a final extension at 72 °C for 10 min. We also used 1% agarose gels to check the quantity of all PCR products. Finally, 223 and 547 individuals were amplified successfully in each of the 5 nuclear DNA loci and of the 3 cpDNA loci, respectively. All DNA fragments that were amplified successfully were sequenced using the amplified forward primer with an ABI 3730 XL genetic analyzer (Applied Biosystems, Foster City, USA). We used program MEGA version 5.0 (Tamura *et al.,* 2011) to check whether the SNPs and indels (insertions and deletions) were consistent with the chromatogram peaks manually, and to proofread variable sites. For nrDNA diploid sequences, we used DnaSP version 5.0 (Librado & Rozas, 2009) to determine the phases of each heterozygous sites. All sequences have been deposited into GenBank (NO. MG993626-MG994796).

2.3 Analyses of cpDNA sequences

We aligned sequences at each cpDNA fragment independently, and deleted indels using MEGA version 5.0 (Tamura *et al.,* 2011). The numbers of polymorphic sites for each cpDNA fragment were counted by manual. Based on the concatenated cpDNA sequences, we calculated the average gene diversity within populations (H_S), total gene diversity (H_T), and the coefficients of genetic differentiation (G_{ST} and N_{ST}) for each *Dipelta* species using PERMUT (available at http://www.pierroton.inra.fr/genetics/labo/Software/Permut/). To test the chloroplast genomic differentiation among populations and between species, the analysis

of molecular variance (AMOVA) was performed using ARLEQUIN version 3.5 with significance tested using 10,000 permutations (Excoffier & Lischer, 2010).

We used DnaSP v5.0 to determine haplotypes based on concatenated cpDNA, and counted the number of haplotypes for each of 56 populations. Then we inferred the genealogical relationships of all cpDNA haplotypes using NETWORK version 5.0.0.1 (available at http://www.fluxus-engineering.com/sharenet.htm) and dated the divergence between species using BEAST version 1.7.5 (Drummond *et al.,* 2012). For BEAST analysis, we employed a Yule speciation prior and a uncorrelated lognormal relaxed clock model. We used jmodeltest version 2.1.10 (Darriba *et al.,* 2012) to choose the appropriate nucleotide substitution model which was the HKY+I model. The Monte Carlo Markov chain was set for 50 million generations with parameters sampled every 10000 generations in BEAST analysis. The substitution rate (μ) of three cpDNA loci was estimated to be 4.18×10^{-8} - 4.61×10^{-8} by the ABC toolbox (see detail below). Tracer version 1.6 was used to assess the convergence and effective sample sizes (ESS) for all parameters. After discarding the first 3000 trees as burn-in, the rest of trees were summarized in a maximum clade credibility (MCC) tree with TreeAnnotator version 1.7.5. Finally, the MCC tree was visualized in FigTree version 1.4.2 (available at http://tree.bio.ed.ac.uk/software/figtree/).

2.4 Analyses of nrDNA sequences

Sequences also were edited and aligned manually using MEGA5 (Tamura et al., 2011). All polymorphic and heterozygous sites were visually confirmed and separated. For each of 5 nuclear loci, and within each species, we computed the number of segregating sites (S), Watterson's θ_w (Watterson, 1975), nucleotide diversity π (Tajima, 1983), and the minimum number of recombinant events *R*m (Hudson & Kaplan, 1985), Tajima's *D* (Tajima, 1989), number of haplotypes (N_h) and haplotype diversity (H_e), Fu and Li's *D** and *F** (Fu & Li,

191 1993; Fu, 1997), and Fay and Wu's *H* (Fay & Wu, 2000) using DnaSP v5.0 (Librado & Rozas,
192 2009). Meanwhile, the multi-locus Hudson–Kreitman–Aguade test (Hudson *et al.,* 1987) was
193 used to evaluate the fit of data to the neutral model. The sequences of *Diabelia serrate* were
194 used as the outgroup. For each nuclear locus, we used NETWORK version 5.0.0.1 (available
195 at http://www.fluxus-engineering.com/sharenet.htm) to construct median-joining networks of
196 nuclear haplotypes determined by DnaSP v5.0.
197 To examine the population structure, we used two approaches. First, the Wright's fixation
198 index (F_{ST}) was estimated for each locus using ARLEQUIN version 3.5 (Excoffier & Lischer,
199 2010). Second, the admixture model implemented in STRUCTURE version 2.3.4 (Pritchard
200 *et al.,* 2000) was used to assess individual clustering. In STRUCTURE analysis, polymorphic
201 sites with $r > 0.7$ after Bonferroni correction (Fisher's exact test) were deleted due to likely
202 linkage disequilibrium. Twenty independent runs were performed for each number of
203 populations (K) from 1 to 10 with 1×10^5 MCMC steps of burn-in, followed by 1×10^6 steps
204 using an admixture model with correlated allele frequencies. The best number of clusters was
205 inferred using the original method (Pritchard *et al.,* 2000) and the *ΔK* statistic of (Evanno *et*
206 *al.,* 2005). Finally, DISTRUCT version 1.1 (Rosenberg, 2004) was employed to draw the
207 graphics.
208 **2.5 Testing hypotheses of historical gene flow**
209 We tested five models of species divergence between *D. floribunda* and *D. yunnanensis*
210 based on sequences at all nuclear and chloroplast DNA loci from all samples of 56
211 populations, using the approximate Bayesian computation (ABC) approach implemented in
212 the ABCTOOLBOX software package (Wegmann *et al.,* 2010). All models began with the
213 divergence of *D. yunnanensis* from *D. floribunda* at a time point labeled "T". Model 1
214 assumed no gene flow after divergence (Fig. 2a), whereas Models 2-5 all assumed historical

gene flow following divergence. Models 2 and 3 assumed that gene flow continued after divergence, and in Model 2 it continued until the present, whereas in Model 3 it ceased at time $T1$ as required by the ancient migration model (Roux *et al.,* 2016). Models 4 and 5 assumed secondary contact between the species from the time $T2$ onwards; in Model 4 this continued until the present, whereas in Model 5 it ceased at time $T1$.

For each species, we computed five statistics to summarize population genetic information: the number of polymorphic sites (S) and private S, Tajima's D, Fu's Fs and nucleotide diversity (π). For the two species together, we computed three more statistics: the total S, index of population differentiation (F_{ST}) and π_{xy} using ARLEQUIN version 3.5 (Excoffier & Lischer, 2010). All of these statistics were calculated independently for both cpDNA and nrDNA loci, making 26 statistical values in total. We used the R function 'pls' in ABCTOOLBOX package (Wegmann *et al.,* 2010) to extract 11 partial least-squares (PLS) components based on the summary statistics generated by simulation under each of 5 models, for decreasing the redundancy of statistics. Conversion equations were inferred from the 10,000 samples simulated by a standard simulating algorithm for each of five models.

The simulator fastsimcoal (Excoffier & Foll, 2011) was employed to simulate samples for each of 5 models. A total of 5,000,000 simulated samples were generated. For each model, the best 10,000 simulated samples were retained and used to compute the marginal density and Bayes Factor (BF), which was used to determine which model is the best. The regression adjustment general linear model (GLM) was used to generate posterior distributions of all parameters in the best model.

2.6 Testing current gene flow

We measured migration rate (m_c) using BAYESASS 1.3 (Wilson and Rannala, 2003) to estimate short-term gene flow between *D. floribunda* and *D. yunnanensis* based on sequences

239 at all 5 nuclear loci. This software estimates migration rates over the past 2-3 generations
240 using Markov chain Monte Carlo techniques and does not assume that populations are in
241 migration-drift or Hardy-Weinberg equilibrium. Initial runs showed that convergence was
242 reached using 5×10^6 Markov Chain Monte Carlo (MCMC) iterations. We ran the program
243 for 5×10^7 MCMC iterations with a sampling interval of 1000, following the burn-in of $5 \times$
244 10^6. We used the Brownian motion model with F_{ST} calculations of θ and M as starting
245 parameters, and Metropolis-Hastings sampling and uniform prior distributions to estimate θ
246 (range, 0-100; delta, 10) and M (range, 0-00; delta, 0).

247 **2.7 Testing synchronous changes of population sizes**

248 We tested the hypothesis that these two species had shifted their population sizes
249 synchronously using ABCTOOLBOX software (Wegmann *et al.,* 2010), based on sequence
250 variation at the nuclear and chloroplast DNA loci. Because STUCTURE analysis of variation
251 at 5 nuclear loci above revealed two clusters within *D. floribunda*, to reduce the effects of
252 intra-specific substructure, we should analyze each cluster respectively. However, the sample
253 sizes of HL and NZ are too low to do test of population expansion, thus we deleted HL and
254 NZ. Finally, samples from 30 populations of *D. yunnanensis* and 24 populations of *D.*
255 *floribunda* were used in this hypothesis test of synchronous changes. To assess the
256 synchronization in population size change, we introduce a new parameter φ which is a scale
257 factor used to alter the timing parameter of population expansion (Fig. 2b). The null
258 hypothesis (model A) here was that these two species shifted their population sizes
259 synchronously ($\varphi = 1$). When $\varphi > 1$, the population size of *D. yunnanensis* shifted earlier than
260 the shift of *D. floribunda* (model B) and the reverse scenario ($\varphi < 1$) represented a later
261 change of the population size of *D. yunnanensis* (model C). Therefore, our test here is a

simplified version of the hierarchical approximate Bayesian computation (hABC) model for two species, which can allow species-specific parameters to vary independently (Chan *et al.,* 2014). The simulation and estimation procedures are similar to those in testing historical gene flow above.

2.8 Species distribution modeling (SDM)

To explore the niche differentiation and distributional changes of *D. yunnanensis* from *D. floribunda*, we used the MAXENT program (Phillips & Dudík, 2008) to conduct testing of the ecological niche models of either species and projected their potential distributions during three periods: the present, the last glacial maximum (LGM, 21 kya), the last interglacial (LIG, 120-140 kya). Distribution information including 100 localities from *D. yunnanensis* and 170 localities from *D. floribunda* was gleaned from our field records and the Chinese Virtual Herbarium (CVH, available at http://www.cvh.ac.cn/). We downloaded the environmental dataset of 19 climate variables with spatial resolutions of 30 arc seconds from the WorldClim database (http://www.worldclim.org, CCSM) as environmental layers. To reduce the correlation between environmental variables, we examined pairwise correlations among the 19 variables and deleted variables with Pearson correlation coefficient (r) > 0.7. This reduced to 8 the number of environmental variables. These 8 variables (Table S2) were used to model the distributional ranges of each evolutionary lineage. We used 80% of the species records for training and 20% for testing the model in Maxent analysis. The accuracy of the model's performance was evaluated based on the area under the receiver operating characteristic curve (AUC; Fielding & Bell, 1997) and the true skill statistic (TSS; Allouche et al., 2006) using an ensemble modelling approach in BIOMOD2, and graphics were drawn using DIVA-GIS7.5.

3 Results

3.1 Genetic variation of cpDNA sequences

We successfully sequenced three cpDNA fragments (*psbA-trnH*, *psbB-psbF*, *trnL-trnF*) across all sampled individuals from 56 populations without missing sites. All indels were excluded from subsequent analyses because of difficulty in alignment. The total length of concatenated sequences was 1937 bps (*psbA-trnH*: 245 bps; *psbB-psbF*: 779 bps; *trnL-trnF*: 913 bps) after deleting indels. We identified a total of 29 polymorphic sites and 30 haplotypes (Table S2). Twenty-one haplotypes (H10-H30) were present in *D. yunnanensis* and 11 haplotypes (H1-H11) were present in *D. floribunda*. Two haplotypes (H10, H11) were shared by *D. yunnanensis* and *D. floribunda* (Figs. 1, 2). The total genetic diversity is higher in *D. yunnanensis* (H_T=0.908, H_S=0.144) than in *D. floribunda* (H_T=0.633, H_S=0.113). The coefficient of genetic differentiation *Nst* is significantly larger than *Gst* for each of *D. yunnanensis* and *D. floribunda*, indicating significant spatial genetic structure within species (Table S3).

The AMOVA analysis (Table S4) revealed that 52.29% of molecular variation was distributed between species (F_{CT} = 0.52, $P < 0.01$). The intra-specific population fixation indexes (F_{ST}) were 0.92 and 0.83 ($P < 0.01$) for *D. yunnanensis* and *D. floribunda*, respectively. These high levels of differentiation indicated restricted movements of chloroplast genomes among intra-specific populations, and also between species. The genealogy of 30 haplotypes showed that most sampled individuals were grouped into two clades comprising 22 haplotypes (clades 1 & 2; Figs. 3, 4). The remaining 8 haplotypes formed a third, more weakly supported group (Grade 3; Fig. 4), which was sister to clade 2. However, the haplotype network gives what may be a clearer picture, with haplotypes H8-H15 forming Grade 3, from which clades 1 and 2 are independently derived. Clade 1 (H15-H30) occurs only in *D. yunnanensis*, and comprises 277 out of 309 individuals of that species examined. Likewise, Clade 2 (H1-H7) occurs only in *D. floribunda,* comprising 195 of 238

individuals of *D. floribunda* examined (Figs. 3, 4). The remaining 32 and 43 individuals of the two species comprised the Grade 3 (Figs. 3, 4). The dating tree inferred by BEAST suggested that the first (crown) divergence among these haplotypes occurred 430 Kya years ago (Ma; 95% HPDI: 0.26-0.66), assuming a generation time of 10 years.

3.2 Genetic diversity at nuclear loci

The diploid sequences were aligned and phased for each of the five nuclear loci. No indels was found in any of the nrDNA loci examined. The total length of alignments was 2069 bps and the length of each locus ranged from 315 bps to 577 bps, with mean length 414 bps. The neutrality tests for each locus indicated no significant signal of selection (Table 2, S5).

The average value of total nuclear nucleotide variation was slightly higher in *D. floribunda* (θw = 0.0062, π = 0.0057) than in *D. yunnanensis* (θw = 0.0054, π = 0.0055). The minimum number of recombination events (Rm) was from 2 to 3 in *D. floribunda* and from zero to 6 in *D. yunnanensis*. For *D. yunnanensis*, the mean Tajima's *D* values (-0.0070) were negative, and the average Fu's *F** (1.08) and Li's *D** (0.80) were positive. For *D. floribunda*, the mean Tajima's *D* value (-0.25) was negative, whereas the mean Fu's *F** (1.08) and Li's *D** (0.70) were positive.

Networks for each of the five nuclear loci did not detect any polymorphic sites with a fixed difference between the species, and were some shared haplotypes found (Fig. S1). Significant population differentiations within and between species were found (Table 3). The STRUCTURE analysis revealed that the likely number of clusters across all sampled individuals was *K* = 2 (Fig. 5). The first cluster comprised individuals from 24 populations of *D. floribunda*, and the second cluster was composed of the remaining two populations of *D. floribunda* (NZ, HL) and all populations of *D. yunnanensis*.

3.3 Inter-specific divergence and gene flow

Based on both chloroplast and nuclear DNA sequences, model comparison by ABCtoolbox showed that two models bear BFs larger than 3.0, relative to the model 2 which assumed continual gene exchange between species from splitting to the present. The model 5 was the best fit to our data with the highest BF = 3.90 (Fig. 2a). The second best model is model 3, of which BF = 3.78 was slightly lower than model 5. Both models identified gene exchange after divergence and recent reduction or even cessation of inter-specific gene flow. However, model 5 assumed a period of primary isolation between *D. yunnanensis* and *D. floribunda*. The divergence time (T_{div}) between *D. yunnanensis* and *D. floribunda* was estimated by ABCtoolbox at 628 029 – 1 023 500 years ago (assuming 10 years per generation), consistent with the mid-Pleistocene climatic transition between 700 000 – 1 250 000 years ago. Taking into account the younger divergence estimate from BEAST (see above), this gives an age range of 430 – 1024 ka (thousand years ago) for the divergence event. The cessation of inter-specific gene flow ($T1$) was dated at 48 – 6734 years ago, in the Holocene period. The estimated parameters indicated that the effective population size of *D. yunnanensis* was slightly larger (not significantly so) than that of *D. floribunda* (Table 4).

Recent migration rates (m) by the BAYESASS showed that gene flow are low either from *D. floribunda* to *D. yunnanensis* (0.0267, 95% CI 0.005-0.043) or opposite direction (0.0026, 95% CI 0.0005-0.042). This estimation indicated rare gene exchange between these two species, consistent with the ABCtoolbox test above.

3.4 Asynchronous changes of population sizes

The simulations in the hABC framework showed that model B ($\varphi > 1$, BF = 50595.4) was better supported than model A ($\varphi = 1$, BF = 1.0) and model C ($\varphi < 1$, BF = 3×10^{-134}), indicating that *D. yunnanensis* and *D. floribunda* responded asynchronously to the Pleistocene climate changes (Fig. 2b). For both species, signals of population expansion were detected.

For *D. floribunda*, the estimated timing of population expansion was 16.68 thousand years ago (ka; 95% HPDI: 1.35 - 394.29), during the post-glacial period. The estimated φ was 2.01 (95% HPDI: 1.00 - 73.93), indicating that populations of *D. yunnanensis* expanded much earlier, at around 33.58 ka, before the LGM period but not earlier than the Last Interglacial (LIG) period.

3.5 The distributional prediction of the two species during three periods

AUC and TSS values indicated high levels of predictive performance for both species (Table 2). For *D. floribunda*, AUC and TSS values were 0.98 and 0.89, respectively. For *D. yunnanensis* were 0.95 and 0.84, respectively. The results of ecological niche modeling (Fig. 6) showed that the similarities (*D* and *I*) between the climatic niches occupied by *D. floribunda* and *D. yunnanensis* were significantly lower than would be expected from random sampling. The projected distributions of these two species at present encompassed most of sampling locations. During the LGM period, the range of *D. floribunda* was narrow and scattered, relative to the current distribution, and it seemed to have been restricted mainly to the north and east of the Sichuan Basin. Conversely, *D. yunnanensis* was mainly distributed in the west and south of the basin as far south as Myanmar and Laos, but might have occupied some areas to the north of the basin. During the LIG period, these two species were likely distributed adjacently in the south and west of the basin.

From the LGM to the present, the range of *D. floribunda* expanded but the range of *D. yunnanensis* seems to have either remained stable or reduced, following expansion during the LIG-LGM period. Surprisingly, the distribution of *D. floribunda* appears to have integrally moved northwards by some distance during the LIG-LGM period, spanning the Sichuan Basin. Conversely, *D. yunnanensis* experienced *in situ* expansion in the southwest of the basin from the LIG to the present.

4 Discussion

It is important to examine the role of geographic barriers in the process of species divergence (Endler, 1977; Abbott *et al.,* 2008; Avise, 2012; Grant & Grant, 2017). In the present study, we tested the effects of the Sichuan Basin on the divergence of two montane species, *D. floribunda* and *D. yunnanensis*. The analyses of chloroplast and nuclear sequence variation showed high differentiation between species and among intra-specific populations (Fig. 1; Tables 3, S3, S4), indicating limited dispersal ability for both species. The divergence event between species was dated during the mid-Pleistocene period, between 430 and 1,024 Ka depending on the analysis used (Table 4; Fig. 4); hence they diverged long after the Sichuan Basin formed, which was during the Neogene. Species distribution modeling (SDM) suggested that the two taxa might have shared a range during the LIG, meaning that allopatry between the species formed, or was resumed, during the LIG and LGM, continuing until the present (Fig. 6). Consistent with this, reduction of interspecific gene flow after the LIG was supported by the ABC analysis (Fig. 2a; Table 4).

4.1 Asynchronous responses to climate change

Demographic analyses based on the chloroplast and nuclear sequence variation recovered signals of asynchronous population expansion (Fig. 2b). hABC and SDM analysis together (Figs. 2b, 6) suggested that *D. floribunda* expanded in the north of the Sichuan Basin at around 16.68 ka, i.e. after the LGM (~20 ka), although ENM suggests it could have occupied parts of that range during the LGM (Fig. 6). Such post-glacial range expansion is seen in many other plants from western China (Qiu *et al.,* 2011; Liu *et al.,* 2012).

405 In contrast, the last detectable population expansion in *D. yunnanensis* was ~33580 years
406 ago, a little before the LGM began, following which SDM suggested that it maintained a
407 near-stable distribution in the south of the basin and the Hengduan Mountains (Fig. 6).
408 Consistent with this, hABC analysis suggested that the expansion timing of *D. floribunda* was
409 more recent than the expansion of *D. yunnanensis*, as shown in model B ($\varphi > 1$). Furthermore,
410 the greater number of haplotypes, and steps between them, in Clade 1 relative to Clade 2,
411 likewise is consistent with expansion within the former (and hence *D. yunnanensis*) having
412 occurred somewhat earlier. Hence *D. floribunda*'s last major expansion was after the LGM,
413 whereas that for *D. yunnanensis* was before it, indicating profoundly different and
414 asynchronous demographic responses to Pleistocene climate changes. That the range of *D.*
415 *yunnanensis* changed little after the LGM could be explained if *D. yunnanensis* responded to
416 the climate changes of the time by shifting altitudes (Fig. 6).
417 ENM suggests that the Sichuan Basin would have remained unavailable to these species
418 through the LIG and LGM as well as the present, forming a constant barrier. Especially
419 during the LGM, both species seemed distributed in the north of the basin, despite *D.*
420 *yunnanensis* not occurring there at present, indicating a profound post-LGM range shift for
421 that species. Genetic similarity to *D. yunnanensis* in population HL and NZ of *D. floribunda*
422 (Figs. 1, 5), might be the result of genetic swamping of *D. yunnanensis* by immigrant material
423 of *D. floribunda*.
424 The presence of the basin likely reduced the area available for contact between these
425 species whether they were distributed on opposite sides of it, as during the present. Without it,
426 there could have been many more contact points towards the centre of the species' shared
427 range. Hence the basin potentially restricts contact, gene flow and competition between these
428 species, but thereby also might promote genetic swamping for isolated populations.

Moreover, by reducing available routes from north to south, it might have restricted and delayed recolonization, perhaps enhancing asynchronous demographic responses.

4.2 The effect of basin isolation on the divergence between *D. floribunda* and *D. yunnanensis*

The estimated time of divergence between *D. floribunda* and *D. yunnanensis* (430 – 1,024 ka) is consistent with the onset of the Naynayxungla Glaciation (0.5 – 0.8 Ma) in the Qinghai-Tibet Plateau (Zheng & Rutter, 1998; Zhang *et al.,* 2000; Shi 2002; Zheng *et al.,* 2002), and also broadly consistent with the mid-Pleistocene climatic transition 0.7 – 1.25 Ma (Ciaranfi *et al.,* 2005; Head *et al.,* 2008). Between species divergence and the LIG, it is possible that gene flow between the species was intermittent or even continuous (Model 3; Fig. 2a; Table 4). However, gene flow during and after the LIG appears highly likely (Models 3 or 5; Figs. 2a, 6; Table 4).

The nature of ABC analysis is to assign relative probabilities to different models, meaning that in this case less supported models cannot be rejected entirely based on this analysis alone. Despite this, evidence from haplotype relationships and STRUCTURE analysis provide further insight into gene flow between these species, and can be used to assess these models. The SDM analysis showed that these two species might have been co-distributed in the southwest of the Sichuan Basin before and during the LIG period (Fig. 6a), providing opportunities for hybridization and gene exchange, in which case that the Sichuan basin was less of a barrier to them then than it is now.

Theoretical and simulated studies suggest that geographic isolation would contribute to speciation even in the presence of gene flow (Nosil, 2008; Abbott *et al.,* 2013; Sousa & Hey, 2013). Nevertheless, the biogeographic processes of speciation in most plants are still unclear. In the present study, we compared the models allowing primary or secondary contacts (Fig.

2a). If geographic isolation contributed to the differentiation between *D. floribunda* and *D. yunnanensis*, then models that predict a complete cessation of gene flow for some period after speciation should perform better than those that predict ongoing gene flow following speciation; our analysis showed consistent results. Indeed, the best performing model was that predicting gene flow for a period, but ceasing some time before the present (Model 3, Fig. 2a). Models allowing recent gene flow (2 and 4) were not supported, which fits well with ENM analyses that indicate sympatry during the LIG, but not afterwards (Fig. 6). Genetic migration estimates by BAYESASS also indicated that current gene flow is rare detectable between two species.

4.3 Range expansion and interspecific gene flow

Taking the two species together, cpDNA haplotypes fall into three clear groups: two large, well-supported monophyletic clades, 1 and 2, comprise only material of *D. yunnanensis* and *D. floribunda*, respectively (Figs. 3, 4, S2). The remaining eight haplotypes comprise the Grade 3, whose relationships are poorly supported; this comprises four haplotypes from *D. yunnanensis*, two from *D. floribunda*, and two that are shared. Notably, all three of these haplotype groups exhibits a very distinctive geographical range: the Grade 3 comprises the four most northerly populations of *D. yunnanensis* plus neighbouring populations from the far west of and *D. floribunda*'s range, plus two southeastern outliers of *D. floribunda*. All remaining material from the centre of *D. floribunda*'s range has Clade 2 haplotypes (except for a few plants from population SNJ), whereas all remaining populations of *D. yunnanensis* have Clade 1 haplotypes (Fig. 1).

Such a pattern, with admixture among early branching haplotypes, could suggest lineage sorting, but this alone cannot explain the strong geographical structuring of clades. However, the haplotype network (Fig. 3) shows a pattern where two particular haplotypes (H25 for

Clade 2, and either H22 or H23 for Clade 1) were ancestral to a burst of cpDNA haplotype divergence. Effectively, these particular haplotypes diverged many daughter haplotypes, while those from the Grade 3 diverged few or none. This can be explained if haplotypes H25 and H22/H23 were the only haplotypes present in material that was undergoing range expansion, which in turn implies a biogeographic barrier limiting the number of within-species lineages that could move past it. Therefore, the Sichuan Basin might have acted as a filter during these range shifts, causing founder or extreme leading edge effects, reducing within-species diversity.

The situation in *D. yunnanensis* may be more complex, with a clade within Clade 1 possibly indicating more than one wave of expansion. Nonetheless, the existence of the monophyletic, geographically well-defined clades within each species is highly consistent with episodes of range expansion, as indicated by our other analyses. Based on ENM, the expansion in *D. yunnanensis* might have been southward, following the LIG (Fig. 6), but the picture is less clear for *D. floribunda*. The Sichuan basin might have separated the Clade 2 material of this species from Grade 3 material during the LGM (Fig. 6).

What gene flow there appears to have been between these two species involves mainly, but not only, those populations that have the Grade 3 haplotypes. Haplotype H3 diverged from H4 around 600000 years ago (Figs. 3, 4), yet both are shared between the species (Figs. 3, 4), indicating that at least one has jumped between species since that time. Otherwise, some haplotype admixture across the Grade 3 could be attributable to lineage sorting, especially H2, which occurs well away from other Grade 3 haplotypes in population SNJ of *D. floribunda*. With haplotype data alone, one could infer that material of both species to the NW of the Sichuan Basin was ancestral, that material to the west (*floribunda*) and south (*yunnanensis*)

resulted from later waves of expansion, and that very limited gene flow had followed, involving the older populations.

STRUCTURE (Fig. 5) reveals two populations (HL and NZ) that match *D. floribunda* for morphology and geographic range, but cluster with *D. yunnanensis*, probably indicating past hybridization between species (Muir & Schlotterer, 2005; Petit & Excoffier, 2009). The two populations are distant from each other, and NZ is well separated from *D. yunnanensis* by the basin, implying that it received *floribunda* germplasm either via a dispersal event across the basin, or a relict population left over from when it was distributed on the north side during the LGM (Fig. 6). Either way, the fact that no neighbouring populations are affected suggests that introgression occurred after the most recent episode of range expansion. This, plus the two shared haplotypes between the species, provides evidence for sporadic gene flow between them, and an indication that some of it may have been post-LGM. From this, ABC model 1 (allopatric speciation with no subsequent gene flow) can be confidently rejected. Conversely, the rarity of interspecific gene flow according to our data also indicates that the current allopatric pattern surrounding the Sichuan Basin at least minimizes inter-specific gene flow (Fig. 2a; Table 4). Overall, both allopatric and other speciation modes are possible, such as ecological niche divergence, but complex-post-divergence history would obscure their signal.

5 Conclusion

We used a case study of two *Dipelta* species to test the hypothesis of the basin isolation on the species evolution. The ABC, hABC and SDM analyses all supported the post-divergence formation of allopatric distribution and asynchronous demographic shifts. The extreme northward movements of *D. floribunda* from the south to the north of the Sichuan Basin after the LIG, causing the formation of allopatric pattern of these two species, occurred much later than the species divergence event. Subsequently, these two species responded to

the Pleistocene climate changes asynchronously because the Sichuan Basin increased the difficulty in colonizing suitable habitats for *D. floribunda*. CpDNA haplotype patterns within both species are consistent with independent demographic expansions within each of them, whereas cpDNA and nuclear evidence reveal occasional instances of gene flow between them. Species-specific biological attributes have been repeatedly indicated to be the main determinants of diversification and demographic patterns (Smith *et al.,* 2014; Papadopoulou & Knowles, 2016; Prates *et al.,* 2016). However, our results highlight that complex topography should be considered in understanding the distributional pattern and asynchronous responses of closely related species.

Acknowledgements

We thank Dr. Dongrui Jia, Dr. Yuanwen Duan, Dr. Zhiqiang Lu and Dr. Qianlong Liang for their assistance in this study. This study was supported by the Major Program of the NSFC (31590823 to H.S.), the National Key R & D Program of China (2017YF0505200 to H.S.) and The Open Project of the Key Laboratory of Biodiversity and Biogeography (KLBB201206 to B.T.).

References

Abbott R, Albach D, Ansell S, Arntzen JW, Baird SJE, Bierne N, Boughman JW, Brelsford A, Buerkle CA, Buggs R, Butlin RK, Dieckmann U, Eroukhmanoff F, Grill A, Cahan SH, Hermansen JS, Hewitt G, Hudson AG, Jiggins C, Jones J, Keller B, Marczewski T, Mallet J, Martinez-Rodriguez P, Most M, Mullen S, Nichols R, Nolte AW, Parisod C, Pfennig K, Rice AM, Ritchie MG, Seifert B, Smadja CM, Stelkens R, Szymura JM, Vainola R, Wolf JBW, Zinner D. 2013. Hybridization and speciation. *Journal of Evolutionary Biology*, 26: 229-246.

Abbott RJ, Ritchie MG, Hollingsworth PM. 2008. Introduction. Speciation in plants and animals: pattern and process. *Philosophical transactions of the Royal Society of London B. Biological Sciences*, 363: 2965-2969.

Allouche O, Tsoar A, Kadmon R. 2006. Assessing the accuracy of species distribution models: prevalence, kappa and the true skill statistic (TSS). *Journal of Applied Ecology*, 43: 1223-1232.

Avise JC. 2012. *Molecular markers, natural history and evolution.* New York: Springer Science & Business Media.

Badgley C, Smiley TM, Terry R, Davis EB, Desantis LR, Fox DL, Hopkins SS, Jezkova T, Matocq MD, Matzke N. 2017. Biodiversity and Topographic Complexity: Modern and Geohistorical Perspectives. *Trends in Ecology & Evolution*, 32: 211-226.

Ball IR. 1975. Nature and Formulation of Biogeographical Hypotheses. *Systematic Zoology*, 24: 407-430.

Chan YL, Schanzenbach D, Hickerson MJ. 2014. Detecting concerted demographic response across community assemblages using hierarchical approximate bayesian computation. *Molecular Biology and Evolution*, 31: 2501-2515.

561 Ciaranfi N, van Kolfschoten T, Coltorti M. 2005. The Plio-Pleistocene boundary and the Lower/Middle
562 Pleistocene transition: type areas and sections - an introduction. *Quaternary International*, 131: 1-3.
563 Clark MK, House M, Royden L, Whipple K, Burchfiel B, Zhang X, Tang W. 2005. Late Cenozoic uplift of
564 southeastern Tibet. *Geology*, 33: 525-528.
565 Crisp MD, Trewick SA, Cook LG. 2011. Hypothesis testing in biogeography. *Trends in Ecology & Evolution*, 26:
566 66-72.
567 Darriba D, Taboada GL, Doallo R, Posada D. 2012. jModelTest 2: more models, new heuristics and parallel
568 computing. *Nature Methods*, 9: 772-772.
569 Drummond AJ, Suchard MA, Xie D, Rambaut A. 2012. Bayesian phylogenetics with BEAUti and the BEAST
570 1.7. *Molecular Biology and Evolution*, 29: 1969-1973.
571 Endler JA. 1977. *Geographic variation, speciation, and clines.* Princeton: Princeton university press.
572 Evanno G, Regnaut S, Goudet J. 2005. Detecting the number of clusters of individuals using the software
573 STRUCTURE: a simulation study. *Molecular Ecology*, 14: 2611-2620.
574 Excoffier L, Lischer HE. 2010. Arlequin suite ver 3.5: a new series of programs to perform population genetics
575 analyses under Linux and Windows. *Molecular Ecology Resources*, 10: 564-567.
576 Excoffier L, Foll M. 2011. Fastsimcoal: a continuous-time coalescent simulator of genomic diversity under
577 arbitrarily complex evolutionary scenarios. *Bioinformatics*, 27: 1332-1334.
578 Favre A, Packert M, Pauls SU, Jahnig SC, Uhl D, Michalak I, Muellner-Riehl AN. 2015. The role of the uplift of
579 the Qinghai-Tibetan Plateau for the evolution of Tibetan biotas. *Biological Reviews*, 90: 236-253.
580 Fay JC, Wu C-I. 2000. Hitchhiking under positive Darwinian selection. *Genetics*, 155: 1405-1413.
581 Fielding AH, Bell JF. 1997. A Review of Methods for the Assessment of Prediction Errors in Conservation
582 Presence/Absence Models. *Environmental Conservation*, 24: 38-49.
583 Fu YX. 1997. Statistical tests of neutrality of mutations against population growth, hitchhiking and background
584 selection. *Genetics*, 147: 915-925.
585 Fu YX, Li WH. 1993. Statistical tests of neutrality of mutations. *Genetics*, 133: 693-709.
586 Grant BR, Grant PR. 2017. Watching speciation in action. *Science*, 355: 910-911.
587 He TN, Liu SW, Liu JQ. 2013. A New Qinghai Tibet Plateau Endemic Genus Sinoswertia and Its Pollination
588 Mode. *Plant Diversity,* 35: 393-400.
589 Head MJ, Pillas B, Farquhar SA. 2008. The Early-Middle Pleistocene transition: characterization and proposed
590 guide for the defining boundary. *Episodes*, 31: 255-259.
591 Hudson RR, Kaplan NL. 1985. Statistical properties of the number of recombination events in the history of a
592 sample of DNA sequences. *Genetics*, 111: 147-164.
593 Hudson RR, Kreitman M, Aguade M. 1987. A test of neutral molecular evolution based on nucleotide data.
594 *Genetics*, 116: 153-159.
595 Ju MM, Ma HC, Xin PY., Zhou ZL,Tian B. 2015. Development and characterization of EST-SSR markers in
596 Bombax ceiba (Malvaceae). *Applications in Plant Sciences* 3: 1500001.
597 Librado P, Rozas J. 2009. DnaSP v5: a software for comprehensive analysis of DNA polymorphism data.
598 *Bioinformatics*, 25: 1451-1452.
599 Liu JQ, Duan YW, Hao G, Ge XJ, Sun H. 2014. Evolutionary history and underlying adaptation of alpine plants
600 on the Qinghai–Tibet Plateau. *Journal of Systematics and Evolution*, 52: 241-249.
601 Liu JQ, Sun YS, Ge XJ, Gao LM, Qiu YX. 2012. Phylogeographic studies of plants in China: Advances in the
602 past and directions in the future. *Journal of Systematics and Evolution*, 50: 267-275.
603 Macarthur RH. 1972. *Geographical ecology: patterns in the distribution of species.* Princeton, New Jersey.
604 Princeton University press.
605 Muir G, Schlotterer C. 2005. Evidence for shared ancestral polymorphism rather than recurrent gene flow at
606 microsatellite loci differentiating two hybridizing oaks (Quercus spp.). *Molecular Ecology*, 14: 549-
607 561.
608 Nathan R. 2008. Long-distance dispersal of plants. *Science*, 322: 375-375.
609 Nosil P. 2008. Speciation with gene flow could be common. *Molecular Ecology*, 17: 2103-2106.
610 Papadopoulou A, Knowles LL. 2016. Toward a paradigm shift in comparative phylogeography driven by trait-
611 based hypotheses. *Proceedings of the National Academy of Sciences USA*, 113: 8018-8024.
612 Petit RJ, Excoffier L. 2009. Gene flow and species delimitation. *Trends in Ecology & Evolution*, 24: 386-393.
613 Phillips SJ, Dudík M. 2008. Modeling of species distributions with Maxent: new extensions and a
614 comprehensive evaluation. *Ecography*, 31: 161-175.
615 Prates I, Xue AT, Brown JL, Alvaradoserrano DF, Rodrigues MT, Hickerson MJ, Carnaval AC. 2016. Inferring
616 responses to climate dynamics from historical demography in neotropical forest lizards. *Proceedings of*
617 *the National Academy of Sciences USA*, 113: 7978-7985.

Pritchard JK, Stephens M, Donnelly P. 2000. Inference of population structure using multilocus genotype data. *Genetics*, 155: 945-959.
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 Evolution*, 59: 225-244.
Rosenberg NA. 2004. *DISTRUCT: a program for the graphical display of population structure*. *Molecular Ecology Notes*, 4: 137-138.
Roux C, Fraisse C, Romiguier J, Anciaux Y, Galtier N, Bierne N. 2016. Shedding light on the grey zone of speciation along a continuum of genomic divergence. *PLoS Biology*, 14: e2000234.
Shi Y. 2002. Characteristics of late Quaternary monsoonal glaciation on the Tibetan Plateau and in East Asia. *Quaternary International*, 97-98: 79-91.
Shi YF, Li JJ, Li BY. 1998. *Uplift and Environmental Changes of Qinghai-Tibet Plateau in the Late Cenozoic.* Guangzhou, Guangdong Science and Technology Press..
Smith BT, McCormack JE, Cuervo AM, Hickerson MJ, Aleixo A, Cadena CD, Pérez-Emán J, Burney CW, Xie X, Harvey MG. 2014. The drivers of tropical speciation. *Nature*, 515: 406-409.
Sousa V, Hey J. 2013. Understanding the origin of species with genome-scale data: modelling gene flow. *Nature Reviews Genetics*, 14: 404-414.
Steinbauer MJ, Field R, Grytnes JA, Trigas P, Ah‐Peng C, Attorre F, Birks HJB, Borges PAV, Cardoso P, Chou CH. 2016. Topography-driven isolation, speciation and a global increase of endemism with elevation. *Global Ecology & Biogeography*, 25: 1097-1107.
Sun H, Zhang J, Deng T, Boufford DE. 2017. Origins and evolution of plant diversity in the Hengduan Mountains, China. *Plant Diversity*, 39: 161-166.
Tajima F. 1983. Evolutionary relationship of DNA sequences in finite populations. *Genetics*, 105: 437-460.
Tajima F. 1989. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics*, 123: 585-595.
Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S. 2011. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Molecular Biology and Evolution*, 28, 2731-2739.
Usinowicz J, Changyang CH, Chen YY, Clark JS, Fletcher C, Garwood NC, Hao Z, Johnstone J, Lin Y, Metz MR. 2017. Temporal coexistence mechanisms contribute to the latitudinal gradient in forest diversity. *Nature*, 550: 105-108.
Wang A, Yang M, Liu J. 2005. Molecular phylogeny, recent radiation and evolution of gross morphology of the rhubarb genus Rheum (Polygonaceae) inferred from chloroplast DNA trnL-F sequences. *Annals of Botany*, 96: 489-498.
Wang C, Dai J, Zhao X, Li Y, Graham SA, He D, Ran B, Meng J. 2014. Outward-growth of the Tibetan Plateau during the Cenozoic: A review. *Tectonophysics*, 621: 1-43.
Wang E, Kirby E, Furlong KP, Van Soest M, Xu G, Shi X, Kamp PJ, Hodges K. 2012. Two-phase growth of high topography in eastern Tibet during the Cenozoic. *Nature Geoscience*, 5: 640-645.
Wang S, Jiao S, Xin H. 2013. Spatio-temporal characteristics of temperature and precipitation in Sichuan Province, Southwestern China, 1960–2009. *Quaternary International*, 286: 103-115.
Watterson GA. 1975. On the number of segregating sites in genetical models without recombination. *Theoretical Population Biology*, 7: 256-276.
Wegmann D, Leuenberger C, Neuenschwander S, Excoffier L. 2010. ABCtoolbox: a versatile toolkit for approximate Bayesian computations. *BMC Bioinformatics*, 11: 116.
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.
Wilson GA, Rannala B. 2003. Bayesian inference of recent migration rates using multilocus genotypes. *Genetics*, 163: 1177-1191.
Xing YW, Ree RH. 2017. Uplift-driven diversification in the Hengduan Mountains, a temperate biodiversity hotspot. *Proceedings of the National Academy of Sciences USA*, 114: 3444-3451.
Yang Q, Landrein S. 2011. *Linnaeaceae. Flora of China (ed. by Z.Y. Wu and P.H. Raven).* Beijing, Science Press.
Ye JW, Li Q, XY,Bao L, Wang HF, Ge JP. 2017. Twenty-Seven Low-Copy Nuclear Primers for Lindera obtusiloba (Lauraceae): A Tertiary Relict Species in East Asia. *Applications in Plant Sciences* 5(12): 1700120
Zhang DF, Fengquan L, Jianmin B. 2000. Eco-environmental effects of the Qinghai-Tibet Plateau uplift during the Quaternary in China. *Environmental Geology*, 39: 1352-1358.

Zheng B, Xu Q, Shen Y. 2002. The relationship between climate change and Quaternary glacial cycles on the Qinghai–Tibetan Plateau: review and speculation. *Quaternary International*, 97: 93-101.
Zheng BX, Rutter N. 1998. On the problem of quaternary glaciations, and the extent and patterns of Pleistocene ice cover in the Qinghai-Xizang (Tibet) Plateau. *Quaternary International*, 45-46: 109-122.

Supplementary item legends:

Fig. S1. Network analyses of genotypes at each nuclear locus

Fig. S2. Geographical distribution of 3 Clades identified from Phylogenetic relationships among cpDNA haplotypes and 2 Clusters from STRUCTURE (K=2) based on five low-copy nuclear gene dataset.

Table S1. List of primers for the five nuclear loci and three cpDNA loci, putative function, primer sequences and according references.

Table S2. The 8 environmental variables used for ecological niche modeling in this study.

Table S3. Genetic diversity and differentiation analyses for cpDNA variations in *D. yunnanensis* and *D. floribunda*.

Table S4. Analysis of molecular variance (AMOVA) of cpDNA data.

Table S5. HKA test statistic for *D. yunnanensis* and *D. floribunda*.

Data Accessibility Statement

"Data Accessibility:

- DNA sequences: GenBank accessions MG993626-MG994796

- Climate data and MaxEnt input files: Climate data were downloaded from the WorldClim database (http://www.worldclim.org, CCSM) and MaxEnt input files Dryad doi.org/10.5061/dryad.b52n66j

- Sampling locations: In this manuscript (Table 1)

699 **Tables**

700 Table 1 The information of sampling locations of *Dipelta floribunda* and *D. yunnanensis*. The
701 numbers of individuals used for chloroplast and nuclear DNA sequencing are represented by *N*1 and
702 *N*2, respectively. *N* represents the number of chloroplast haplotypes identified in each population.

Population	Sample location (all in China)	Longitude	Latitude	Altitude	*N*1	*N*2	Haplotypes (*N*)
Dipelta yunnanensis							
BS	Tianlinxian GX	106°13.5′	24°17.833′	1292	14	5	H24(14)
CH	Wuchuanxian GZ	108°07′	28°40.783′	1108	5	5	H24(5)
LLX	Longlixian GZ	106°54.667′	27°18.733′	1350	8	5	H27(2) H25(6)
ML	Mulixian SC	101°16.7′	27°55.933′	2269	13	5	H29(13)
YXX	Yuexixian SC	102°27.083′	28°45.35′	2588	9	2	H28(9)
HY	Hongyaxian SC	102°51.517′	29°29.3′	2100	12	6	H30(9) H29(3)
ZJ	Zhaojuexian SC	102°33.65′	27°49.867′	2824	2	2	H24(2)
LJS	Pugexian SC	102°25.933′	27°19.25′	3050	8	6	H29(8)
HD	Ludingxian SC	102°02′	29°43.95′	2310	15	3	H29(15)
YE	Emeishan SC	103°28.933′	29°36.217′	2433	10	2	H30(10)
WC	Wenchuanxian SC	103°35.167′	31°36.817′	2337	7	6	H10(2) H9(1) H8(4)
DJY	Dujiangyan SC	103°33′	31°03′	1986	6	1	H11(6)
BX	Baoxingxian SC	102°50.383′	30°36.567′	2100	4	4	H15(3) H14(1)
DY	Xilingxueshan SC	103°09′	30°40.2′	2250	15	6	H15(15)
CWL	Chayuxian XZ	98°27.8′	28°28.583′	1920	16	5	H22(16)
JZ	Jiaozixueshan YN	102°53.717′	26°05′	2730	12	6	H18(12)
YL	Yulongxueshan YN	100°15.967′	27°02′	2800	8	3	H29(8)
JZS	Jizushan YN	100°22.05′	25°28.783′	2800	16	6	H23(7) H16(5) H17(4)
HTX	Hutiaoxia YN	99°57.4′	27°21.3′	2737	15	4	H29(15)
BR	Wengshuixiang YN	99°42.244′	28°00′	3105	11	5	H29(5)H21(6)
BZL	Benzilan YN	99°09′	28°17.29′	3131	8	3	H21(5) H19(3)
LP	Lanpingxian YN	99°24.361′	26°27.682′	2650	13	5	H23(13)
BD	Yezhizhen YN	99°04′	27°40.459′	2754	8	0	H21(1) H20(7)
LD	Langduxiang YN	99°41.983′	27°49.983′	3282	7	4	H29(7)
MS	Meilixueshan YN	98°51.24′	28°28.73′	2875	9	2	H21(9)
MD	Gongshanxian YN	98°19.383′	28°10.35′	2390	14	6	H22(14)
YG	Huapingxian YN	101°25.483′	26°37.95′	1320	11	6	H29(11)
YM	Yimenxian YN	102°16.396′	24°61.886′	1600	7	2	H23(7)
YMX	Yanmenxiang YN	98°53.569′	28°04′	2910	14	4	H21(14)
JF	Jinfoshan CQ	107°11.017′	28°58.7′	1350	12	6	H26(2) H24(10)
D. floribunda							
LX	Lixian GS	105°02′	33°41.567′	1563	8	5	H3(8)
ZKQ	Tielouxiangzhaikeqiao GS	104°27.833′	32°54.4′	1743	13	4	H3(13)
TLX	Tielouxiangcaoheba GS	104°27.833′	32°54.4′	1650	4	4	H9(4)
BKZ	Bikouzhen GS	105°14.75′	32°44.983′	1659	20	1	H3(20)
DBZ	Danbaozhen GS	104°44.814′	32°51.099′	1208	13	5	H9(6) H8(1) H3(6)
ZQX	Zhouquxian GS	105°23.449′	33°34.182′	1928	17	5	H3(17)

DCX	Tianshuitaohuagou GS	105°43.25′	34°34.917′	1169	4	1	H3(4)
DC	Tianshuidangchuan GS	106°08′	34°20.15′	1596	4	3	H3(2) H1(1)H7(1)
CX	Chengxian GS	105°49.811′	33°43.322′	1460	8	3	H3(8)
HX	Huixian GS	105°45.365′	34°03′	1413	7	1	H3(7)
XYS	Guchengxian HB	111°18.783′	32°07′	611	1	1	H4(1)
YRZ	Shenlongjiayangrizhen HB	110°50′	31°45.4′	864	5	5	H3(5)
SNJ	Shenlongjiasongbaizhen HB	110°38.617′	31°45.35′	978	7	6	H3(5) H12(2)
JS	Jishou HN	109°35.433′	28°19.917′	584	7	5	H13(7)
FH	Fenghuangxian HN	109°30.15′	28°15.6′	824	12	5	H13(12)
XXX	Xixiangxian SX	107°32.033′	32°42.567′	1299	12	6	H3(10)H6(1)H5(1)
NZ	Nanzhengxian SX	106°57.45′	32°45.1′	1050	13	6	H3(13)
XY	Xunyangxian SX	109°34.667′	32°58.617′	1290	12	2	H7(12)
YX	Yangxian SX	107°40.917′	33°26.45′	830	15	5	H3(15)
PL	Pinglixian SX	109°14.917′	32°05′	1201	9	3	H7(9)
BJ	Baojishi SX	107°13.967′	34°21.867′	867	2	2	H7(2)
NS	Ningshanxian SX	108°18.567′	33°18.733′	882	13	4	H3(11) H2(2)
GY	Guangyuanxibeixiang SC	105°44.083′	32°33.817′	800	13	6	H3(13)
WCX	Wangcangxain SC	106°29.55′	32°32.5′	690	8	5	H3(8)
HL	Huanglong SC	103°49.25′	32°45.05′	3301	5	1	H9(4) H8(1)
PW	Pingwuxian SC	104°31.2′	32°37.6′	1407	6	4	H9(6)

703 Abbreviations: GX, Guangxi; GZ, Guizhou; SC, Sichuan; XZ, Xizang; YN, Yunnan; CQ, Chongqing;
704 GS, Gansu; HN, Hunan; HB, Hubei; SX, Shaanxi.

Table 2 Nucleotide variation, nucleotide diversity, haplotype diversity and neutrality tests at five nuclear loci for *Dipelta yunnanensis* and *D. floribunda*.

Species	Locus	Total					Haplotype diversity		Recombination		Neutrality tests			
		N	*L*	*S(singl).*	*θwt*	*πt*	N_h	H_e	R_m	*4Ner*	*D*	*D**	*F**	*H*
D. yunnanensis	23311	250	343	14(1)	0.00669	0.00681	16	0.7496	2	3.00	0.04088	0.85712	0.65612	-3.53157
	38541	250	416	13(2)	0.00473	0.00443	18	0.821	3	18.00	-0.15088	0.70147	0.45926	-0.0808
	41398	250	315	13(0)	0.00677	0.00587	15	0.800	1	2.00	-0.61670	0.92647	0.39977	0.8303
	45367	250	577	23(0)	0.00654	0.00798	31	0.911	6	8.00	0.45828	1.86375**	1.55371	1.7741
	56546	250	418	6(0)	0.00235	0.00263	7	0.677	0	0.00	0.23349	1.04717	0.91315	0.4662
	Average	250			0.005472	0.005544					-0.006986	1.079196	0.796402	-0.108354
D. floribunda	23311	196	343	15(0)	0.00747	0.00722	19	0.869	2	6.00	-0.24294	1.60808*	1.07781	-1.2089
	38541	196	416	10(0)	0.00411	0.00213	13	0.424	2	14.00	-1.12587	1.33137	0.52624	-0.9970
	41398	196	315	18(3)	0.00976	0.00814	20	0.813	2	9.00	-0.43758	0.04962	-0.16831	1.2018
	45367	196	577	14(1)	0.00415	0.00440	15	0.800	2	11.00	0.15406	0.89179	0.73558	1.6819
	56546	196	418	14(0)	0.00572	0.00669	13	0.709	3	8.00	0.42674	1.52684	1.33319	1.0822
	Average	196			0.006242	0.005716					-0.245118	1.08154	0.700902	0.352

Abbreviations: *N*, sample size; *L*, length in base pairs; *S*, number of segregating sites; π, nucleotide diversity; θ, Watterson's parameter; R_m, the minimum number of recombinant events; N_h, number of haplotypes; H_e, Nei's haplotypic diversity; *D*, Tajima's D statistic; *H*, Fay and Wu's H; *D**, *F**, Fu and Li's D*, F* test; Significant level: *$0.01 \leq P < 0.05$; ** $0.001 \leq P < 0.01$; ***$P < 0.001$.

709

710 Table 3 Genetic differentiation of the five nuclear loci for *Dipelta yunnanensis* and *D.*
711 *floribunda*.

Species	Locus					Average
	23311	38541	41398	45367	56546	
DY	0.39081***	0.34968***	0.53696***	0.36509***	0.59930***	0.448368***
DF	0.40922***	0.47459***	0.29449***	0.54062***	0.43416***	0.430616***
DY vs. *DF*	0.68504***	0.70627***	0.61164***	0.57469***	0.67537***	0.650602***

712 Abbreviations: *DY* indicates *Dipelta yunnanensis*, *DF* indicates *Dipelta floribunda*.
713 Significant level: * $P < 0.05$, ** $P < 0.01$ and *** $P < 0.001$.
714

715 Table 4 Estimates of the posterior distributions of all parameters for the best model (Model 5).

Model	Parameter	N_a	*Ny*	*Nf*	*T1*	*T2*	*T*	*Myf*	*Mfy*
ancient SC	**Mode**	24484	29492	21295	48.63	19155.79	62802.94	1.26E-07	2.06E-06
	HPD 95% Lower	3810	7655	5527	10.00	45.40	5463.87	1.00E-09	1.45E-09
	HPD 95% Upper	217916	150193	113621	29831.84	557557.90	999700.71	7.39E-03	1.56E-02
PC	**Mode**	24484	24484	16876	6734.11	-	102350.51	1.32E-05	3.85E-07
	HPD 95% Lower	3637	8402	5790	14.51	-	5463.49	1.45E-09	1.20E-09
	HPD 95% Upper	207994	136842	103533	52127.87	-	999746.75	1.56E-02	1.56E-02

716

717

718

719 **Figure legends**

720 Fig. 1 Geographical distribution of 30 cpDNA haplotypes identified from the two *Dipelta*
721 species. The pie charts reflect the frequency of haplotype occurrence in each population.
722 Haplotype colours were shown in legend. The maps were made using DIVA-GIS 7.5
723 (www.diva-gis.org).
724

725 Fig. 2 Schematic diagram of five models designed for testing the most likely speciation
726 patterns (a) and synchronous changes of population sizes (b) with Approximate Bayesian
727 Computation (ABC). Bayes-Factors (BFs) are shown in top left corner of each panel. The
728 black arrows represent migration rate between the two *Dipelta* species, T indicate divergence
729 time of the two *Dipelta* species, T1 in model 3 indicate a time point that there is no gene flow
730 after this time point, T1 in model 4 denote a time point that there is no gene flow before this
731 time point, T2 and T1 in model 5 indicate two time point that there is gene flow between
732 these two time point.
733 Abbreviations are as follows: *DY*, *D. yunnanensis*; *DF*, *D. floribunda*; *Na*, effective
734 population size of ancestral species; migration between diverging lineages (*Mfy, Myf*); *Ndy*
735 and Ndf, long-term equilibrium effective population size of *D. yunnanensis* and *D. floribunda*,
736 respectively.
737

738 Fig. 3 Median-joining network of cpDNA haplotypes inferred by NETWORK.

739

740 Fig. 4 Phylogenetic relationships among cpDNA haplotypes and divergence time estimation
741 generated from BEAST. Numbers above the branches were posterior probabilities (PP) for
742 main clades. A-E indicate main node ages.
743

744 Fig. 5 Population cluster analysis with plot of the delta K (ΔK) (a) and the Ln P(D) ± SD (b)
745 using STRUCTURE (K=2, 3 and 4) based on five low-copy nuclear gene dataset (c).
746

747 Fig. 6 Ecological niche modelling predicted distributional range for each of the two *Dipelta*
748 species at three periods: (a) The Last Interglacial (LIG), (b) The Last Glacial Maximum
749 (LGM) (c) The Present time;. (d) The background tests.
750