Loss of floral polymorphism in heterostylous *Luculia pinceana* (Rubiaceae): a molecular phylogeographic perspective

WEI ZHOU,*†‡ SPENCER C. H. BARRETT,§ HONG WANG,*† and DE-ZHU LI*†

*Key Laboratory of Biodiversity and Biogeography, Kunming Institute of Botany, Chinese Academy of Sciences, 132 Lanhei Road, Kunming, Yunnan 650201, China, †Plant Germplasm and Genomics Center, Germplasm Bank of Wild Species, Kunming Institute of Botany, Chinese Academy of Sciences, 132 Lanhei Road, Kunming, Yunnan 650201, China, ‡Graduate University of the Chinese Academy of Sciences, 19 Yuquan Road, Beijing 100049, China, §Department of Ecology and Evolutionary Biology, University of Toronto, 25 Willcocks Street, Toronto, Ontario M5S 3B2, Canada

Abstract

Both deterministic and stochastic forces determine the representation and frequency of floral morphs in heterostylous plant populations. Phylogeographic analysis of molecular variation can provide information on the role of historical factors, including founder events, in affecting population morph structure. Here, we investigate geographical patterns of floral morph variation in a distylous shrub *Luculia pinceana* (Rubiaceae) by examining the relations between floral polymorphism and molecular (*cp*DNA and microsatellite) variation in 25 populations sampled throughout the distribution of the species in southwest China and adjacent countries. In 19 of the 25 populations, the frequency of floral morphs was not significantly different from the expected 1:1 ratio. The remaining populations were either L-morph biased (2) or monomorphic (4) for this form and were morphologically differentiated from the remaining populations in several floral traits, that is, corolla tube length, sex organ position and stigma-anther separation. Phylogeographic analysis supports the hypothesis that *L. pinceana* was initially split into west-central and eastern lineages in the Early Pleistocene (~1.982 Mya). A centrally located lineage composed of morph-biased and monomorphic populations appears to have been subsequently derived from the west-central lineage, perhaps by a founder event after the last glacial maximum. Hypotheses to explain why these populations have not returned to equilibrium morph frequencies are considered.

Keywords: *cp*DNA, distyly, floral polymorphism, founder events, *Luculia pinceana*, morph-ratio variation, microsatellites

Received 23 November 2011; revision received 24 May 2012; accepted 25 May 2012

Introduction

Negative frequency-dependent selection resulting from disassortative mating should lead to equal frequencies of floral morphs (isoplethy) in heterostylous populations (Ganders 1979; Heuch 1979; Barrett 1993; Eckert *et al.* 1996; Thompson *et al.* 2003). But surveys of natural populations of heterostylous plants commonly report floral morph ratios that deviate significantly from

this theoretical expectation, including populations with strongly biased morph frequencies (anisoplethy) and stylar monomorphism (Pailler & Thompson 1997; Kéry *et al.* 2003; Wang *et al.* 2005; Brys *et al.* 2007, 2008). Most workers investigating the causes of biased morph ratios in heterostylous species have focused on the role of selective forces associated with differences between the floral morphs in pollination and mating (e.g. Barrett *et al.* 1983, 2004; Pérez-Barrales & Arroyo 2010; Sosenski *et al.* 2010). However, historical processes, including founder events, may also play an important role in shaping patterns of morph-frequency variation, and

Correspondence: De-Zhu Li, Fax: +86 871 5217791; E-mail: dzl@mail.kib.ac.cn

knowledge of the phylogeography of heterostylous species may therefore aid in interpreting contemporary patterns of floral variation.

Morph and sex ratios in polymorphic plant populations may bear the signature of historical contingency associated with founder events (Morgan & Barrett 1988; Eckert & Barrett 1995; Barrett *et al.* 2010). Indeed, there is considerable evidence that stochastic processes influence both the representation and relative frequency of floral morphs in heterostylous populations (Ornduff 1972; Barrett & Forno 1982; Eckert & Barrett 1992; Husband & Barrett 1992; Wang *et al.* 2005; Luo *et al.* 2006; Castro *et al.* 2007). Unfortunately, few studies have used molecular markers to examine historical events in plant species with sexual polymorphism (but see Dorken & Barrett 2004; Hodgins & Barrett 2007; Pérez-Alquicira *et al.* 2010) and hence the role of migration in affecting the geographical distribution of sexual morphs is not well understood.

Here, we investigate the phylogeography of *Luculia pinceana* Hook. (Rubiaceae), a heterostylous shrub distributed from the south-east margin of the Tibetan Plateau in south-west China to adjacent north-east Myanmar and northern Vietnam (Lou *et al.* 1999). This species is of particular interest because populations vary significantly in floral morph frequency and composition. Populations are either distylous containing long- and short-styled morphs (hereafter L- and S-morphs) or are monomorphic containing the L-morph. Distylous populations are either isoplethic or L-morph biased. In this article, we refer to L-biased populations and those containing only the L-morph as 'nonisoplethic populations'. Despite extensive field surveys, we have not observed populations that are S-morph biased or are monomorphic for this morph (and see Zhou & Wang 2009; Zhou *et al.* 2010).

Patterns of morph-frequency variation in *L. pinceana* are not randomly distributed throughout south-west China and adjacent countries but instead are geographically structured. However, in contrast to many heterostylous species, in which monomorphic populations occur at the geographical margins of the species range, or have markedly disjunct distributions (e.g. Baker 1959; Barrett 1985; Barrett & Shore 1987), in *L. pinceana* monomorphic (and L-biased) populations are centrally located between western and eastern isoplethic populations (Fig. 1a). This curious distribution raises the question of how these populations might have originated and why they are currently in a nonisoplethic state.

Luculia pinceana is also atypical for a distylous species in possessing striking differences between the floral morphs in the expression of heteromorphic self-incompatibility. Most distylous species are self- and intra-morph incompatible, and full seed set is only obtained from inter-morph crosses (Ganders 1979). However, controlled pollinations of *L. pinceana* indicate that whereas the S-morph exhibits the expected self-incompatibility, the L-morph is self-compatible (Ma *et al.* 2009; but misspelled as *L. pinciana*). Difference in compatibility status between floral morphs has been reported in other heterostylous species (reviewed in Barrett & Cruzan 1994) and could have important consequences for the likelihood of establishment of floral morphs following dispersal.

In this study, we use chloroplast (*cp*) DNA and microsatellite (SSR) variation to investigate the historical relationships of 25 populations sampled throughout the range in south-west China and adjacent countries. Our sampling included the three primary floral morph structures (isoplethic, L-biased and monomorphic populations). We also compared floral traits among these populations to determine whether patterns of floral differentiation might be associated with the contrasting morph structures. Our study addressed the following specific questions: (i) to what extent is the geographical distribution of morph structures reflected by patterns of molecular differentiation and, specifically, is there evidence that founder events may have played a role in the origin of monomorphic populations? (ii) Is there evidence that floral traits differ among populations with contrasting morph structures? Specifically, we predicted that if nonisoplethic populations were characterized by weak disassortative mating in anisoplethic populations, or increased selfing rates in monomorphic populations, this should be reflected in modifications to floral traits promoting outcrossing, including flower size and the relative positions of sexual organs.

Materials and methods

Study species

Luculia pinceana is an evergreen perennial shrub found in mountainous and hilly regions of south-west China and adjacent countries at altitudes between 350 and 1800 m (Lou *et al.* 1999). The species flowers from August to December and produces compact inflorescences composed of showy pink or white tubular flowers (mean = 7.8 flower per inflorescence) that last for up to 8 days. Pollinators of *L. pinceana* are mainly long-tongued insects (bumblebees, moths and butterflies) and pollen-gathering bees (*Apis florae*). The fruits of *L. pinceana* are two-valved and dehisce to release seeds after maturity (about 2 months) and seeds are small (diameter: 0.05 ± 0.01 mm), light (thousand seed weight: 62.9 ± 0.44 mg) and possesses wings (length: 1.92 ± 0.18 mm) (Zhou *et al.* unpublished data) and are thus adapted for wind dispersal. Based on our study of
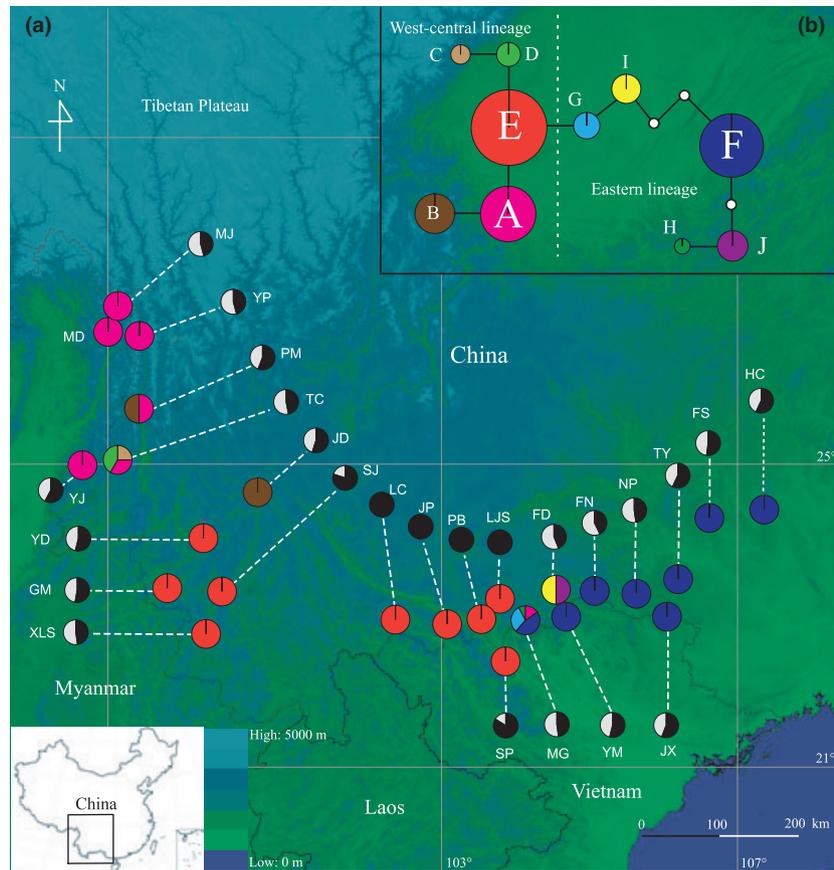
Fig. 1 Geographical patterns of floral morph frequencies and molecular variation in 25 populations of distylous *Luculia pinceana* sampled from south-western China and neighbouring countries (for population codes see Table 1) (a) Coloured circles represent population locations, and within each population, the distribution of *cp*DNA (*trn*L-*trn*F and *rpl*20-*rps*12) haplotypes (HapA-J); black-white circles represent floral morph frequencies in each population, with closed segments indicating the L-morph and open segments the S-morph. (b) The 95% confidence network of 10 *cp*DNA haplotypes (HapA-J). The size of circles corresponds to the frequency of each haplotype. Each solid line represents one mutational step that interconnects two haplotypes for which parsimony is supported at the 95% level. The small open circles indicate hypothetical missing haplotypes.

herbarium specimens and extensive fieldwork, we treated *L. yunnanensis* Hu (Hu 1951) as conspecific with *L. pinceana*. The distribution of *L. yunnanensis* is extremely restricted (Lou *et al.* 1999) and two populations (MJ and YP) of it were sampled in the present study, and the molecular results we obtained (data not presented) were consistent with its conspecific status.

Population sampling and morph ratios

During August 2007–December 2009, we sampled leaf material of *L. pinceana* from a total of 321 plants in 25 populations from throughout the range including 200 and 121 individuals of the L-morph and S-morph, respectively. Identification of genets in *L. pinceana* is straightforward because this species does not reproduce by clonal propagation. The average number of individuals sampled per population was 12.8, range 12–18. Leaf material was dried in silica gel and stored at room temperature. Voucher specimens from all populations, including the outgroup *L. gratissima* (Wall.) Sweet, are lodged at the Herbarium of the Kunming Institute of Botany, Chinese Academy of Sciences (KUN).

In each population, we also recorded the representation and frequencies of floral morphs by randomly sampling flowering individuals (mean sample size per population 75, range: 24–285). The range of sample sizes reflects variation in population size, except in L-monomorphic populations in which larger sample sizes were used to confirm the absence of individuals of the S-morph. We tested floral morph ratios in each dimorphic population against the expected equilibrium ratio of 1:1 using *G* tests (Sokal & Rohlf 1995).

Molecular procedures

We extracted genomic DNA using a modified cetyl trimethyl ammonium bromide (CTAB) protocol (Doyle

1991). Quantification of DNA was carried out with a SmartSpec^TM Plus Spectrophotometer (Bio-Rad, Hercules, CA). Working stocks of DNA were then prepared based on these estimates and stored in $0.1 \times$ TE buffer. We conducted an initial screen for DNA sequence variability at various chloroplast markers using universal primers on one individual from each population. After a preliminary screening of 12 intergenic spacer (IGS) fragments, we chose *trn*L-*trn*F IGS and *rpl*20-*rps*12 IGS (Shaw *et al.* 2005) for the full survey because they contained the most polymorphic sites.

We performed DNA reactions using the following protocol: 25 μL of reaction volume containing 25 ng of genomic DNA, 0.1 μM of each primer, 1.5 mM MgCl$_2$, 0.2 mM of each dNTP, 1.25 U Taq polymerase (Fermentas) and $10 \times$ PCR buffer. PCR amplifications were conducted on thermocycler (Perkin-Elmer) under following conditions: 94 °C for 4 min followed by 32 cycles at 94 °C for 1 min for denaturing, 52 °C (for the *trn*L-*trn*F) or 57 °C (for the *rpl*20-*rps*12) for 1 min for anneal, 72 °C for 1.5 min for extension and a final extension step at 72 °C for 10 min. PCR products were checked on 1% agarose gels stained with ethidium bromide and purified with a PCR purification kit (Shanghai Bio-engineering Company, Shanghai, China). We sequenced purified DNA in both directions by standard methods on an ABI PRISM 3730 automated sequencer. Sequencing primers were identical with those used in the PCR and were conducted in forward and reverse reactions individually. We used seven primer pairs: LP58, LP45, LP4, LP18, LP154, LP107 and LP198 to assess population genetic structure at microsatellite loci in *L. pinceana* (Zhou *et al.* 2010). We performed the PCR reactions and amplifications on a GeneAmp 9700 DNA Thermal Cycler using the same program settings and procedures as described in Zhou *et al.* (2010). PCR products were separated and visualized using a QIAxcel capillary gel electrophoresis system (Qiagen, Irvine, USA) with an internal 10–300 bp size standard. An individual was declared null (nonamplifying) at a locus and treated as missing data after at least two amplification failures.

Data analysis of cpDNA sequences

We edited and assembled sequences in SeqManPro using Lasergene version 8.0 (DNASTAR Inc., Madison, WI, USA). Variable sites in the data matrix were double-checked against the original electrophorogram. Multiple alignments of the *cp*DNA sequences were made manually with the assistance of Clustal X version 1.83 (Thompson *et al.* 1997), with subsequent adjustment in Bioedit 7.0.4.1 (Hall 1999).

We assessed phylogenetic relationships between *cp*DNA haplotypes of *L. pinceana* using maximum parsimony (MP) in PAUP* version 4.0 beta 10 (Swofford 2002) with *L. gratissima* as the outgroup. Indels and a single inversion were coded separately as binary states (0 or 1) using the SEQSTATE program version 1.32 (Müller 2005). We performed full heuristic tree searches with 100 replications of 'random' sequence entries, using the TBR (tree bisection reconnection) branch-swapping and 'MulTrees' options. We generated bootstrap support by performing 1000 iterations of the original data set with tree calculations and then a majority rule consensus tree was created. We also calculated Bayesian posterior probabilities for this data set using the software program MrBayes version 3.1.2 (Ronquist & Huelsenbeck 2003) with a burn-in of 1×10^6 MCMC repetitions. In addition, the degree of relatedness among *L. pinceana cp*DNA haplotypes was estimated with 95% statistical parsimony criteria using the TCS program version 1.21 (Clement *et al.* 2000) with indels treated as single mutation events and coded as substitutions (A or T).

To infer the possible divergence times of the main *cp*DNA lineages within *L. pinceana*, we initially estimated the time to the most recent common ancestor (tMRCA) of the genus *Luculia* by using a coalescence approach as implemented in BEAST version 1.5.4 (Drummond & Rambaut 2007). Forty-two *rbc*L sequences from three main subfamilies of Rubiaceae and two *rbc*L sequences from out of the family were downloaded from GenBank (Table S1, Supporting information). In addition, two *rbc*L sequences from *L. pinceana* and *L. gratissima* were included in the data set. Three fossils were used to calibrate the phylogenetic tree. The oldest known fossils (54 Mya) of Rubiaceae (Roth & Dilcher 1979) were used as a prior for the stem group age (F1) of the family. The fossils represent an extinct leaf type, which cannot be identified to subfamily, tribe or genus of Rubiaceae. Pollen of *Faramea* Aublet (Graham 1985) and fruits of *Emmenopterys* Oliver (Manchester 1999) dated to 45 and 40 Mya, respectively, were used as priors for the minimum age of *Faramea* (F2) and *Emmenopterys* (F3). Using the fossil calibrations of the *rbc*L data set, we obtained the prior age of the *Luculia* crown group. Then, we estimated the timing of lineage splits within *L. pinceana* from the combined plastid data set of *trn*L-*trn*F and *rpl*20-*rps*12, with the age of the genus assumed as a normal distribution around the mean estimated from the *rbc*L data set. For these analyses, we used the common GTR model, and two independent MCMC chains were run for 2×10^8 generations, and the parameters were sampled every 2000 generations. We implemented the program Tracer version 1.5 (Drummond & Rambaut 2007) to check the stationarity and convergence of the parameter estimation. A tree with ages for each node and their 95%

HPD were displayed and modified in FigTree version 1.3.1 (Drummond & Rambaut 2007).

To investigate evidence of recent population expansion, we plotted the mismatch distribution of the west-central group and the eastern group as the observed number of differences between pairs of haplotypes. The goodness-of-fit of the observed mismatch distribution to the predicted theoretical distribution involving a rapid (stepwise) expansion (Rogers & Harpending 1992) was tested using the sum of squared deviations (SSD) between the observed and expected mismatch distributions with the raggedness index ($HRag$) of Harpending (1994). We employed a parametric bootstrap approach (Schneider & Excoffier 1999) with 1000 replicates to test the observed mismatch distribution's fit to the rapid expansion model, and to test the significance of $HRag$. We also performed Fu's F_s test of selective neutrality (Fu 1997) for each population group. Significant negative values indicate an excess of low-frequency haplotypes, as would be expected from a recent population expansion (Fu 1997). We estimated the values from distributions by performing 1000 random permutations in ARLEQUIN version 3.0.

Data analysis of SSR marker

We calculated Hardy–Weinberg equilibrium (HWE) for each locus and each population using Genepop version 1.2 (Raymond & Rousset 1995). We calculated the number of alleles (N_A), effective number of alleles (N_E), expected heterozygosity (H_E), observed heterozygosity (H_O), allelic richness (A_R) and inbreeding index (F_{IS}) using FSTAT version 2.9.3 (Goudet 1995). We also assessed significant differences in these indices of genetic diversity among three geographical/morph structure groupings (i.e. western isoplethic, central monomorphic and anisoplethic, and eastern isoplethic populations) using a one-way ANOVA (Tukey's test) as implemented in SPSS version 13.0 (SPSS Inc., Chicago, IL, USA).

To visualize genetic relationships among all sampled individuals, we conducted an individual-based principal coordinate (PCO) analysis in the program MVSP version 1.3 (Kovach 1999) using Euclidean distances among SSR phenotypes. We also calculated Nei's (1987) unbiased genetic distance (D) among all possible pairs of populations from allele frequencies estimated in the procedure GENDIST of the program PHYLIP version 3.63 (Felsenstein 2005). We then used the procedures NEIGHBOR and CONSENSE of the same program to generate a neighbour-joining (NJ) tree and to infer bootstrap confidence intervals on tree branches with 1000 replicates. To test for genetic admixture among popula-

tions, we conducted a Bayesian analysis of SSR population structure on the entire data set using the program STRUCTURE version 2.2 (Pritchard et $al.$ 2000). The number of population clusters (K) was set from 1 to 10, and 10 independent runs were performed at 1×10^6 Markov chain Monte Carlo samplings after a burn-in period of 5×10^5 iterations. The combination of 'no admixture' and 'uncorrelated allele frequencies' model was used for the analysis. We estimated the most likely number of clusters according to the model value (ΔK) based on the second-order rate of change of the likelihood function with respect to K, following the procedure outlined by Evanno et $al.$ (2005). Finally, the BARRIER program version 2.2 (Manni et $al.$ 2004) was used to identify geographical locations where major genetic barriers among populations might occur, based on their genetic distance (D). We tested the significance of barriers by means of 1000 bootstrapped distance matrices.

Floral morphology

To investigate differentiation of floral traits among isoplethic, L-biased and L-monomorphic populations of $L.$ $pinceana$, we obtained a random sample of flowers from populations in each group and classified them according to style morph. We collected a total of 233 flowers, including 156 L-morph flowers with 66 from isoplethic populations (mean number of flowers per population 4.1), 36 from L-biased populations (18) and 54 from L-monomorphic populations (13.5), and 77 S-morph flowers with 60 from isoplethic populations (3.8) and 17 from L-biased populations (8.5). Only fresh flowers in which styles were fully mature were sampled, and these were preserved in FAA (10% formalin, glacial acetic acid and 70% ethanol; 5/5/90 v/v/v) for subsequent measurements. Corolla tube length, stigma height and anther height were measured from the base of the ovary. We also measured the distance between the stigma height and the nearest anther (stigma–anther separation), as this measure is often closely correlated with selfing rates (e.g. Barrett & Shore 1987). All floral measurements were taken to the nearest 0.01 mm using digital callipers. We compared the floral traits of two morphs between isoplethic and nonisoplethic populations by using a two-factor nested ANOVA (Model III) where populations were nested within morph structure (Zar 1999). Satterthwaite's approximation was used to calculate degrees of freedom in this mixed model design, where population was treated as a random effect and morph structure was treated as a fixed effect (Satterthwaite 1946). This analysis was performed using the GLM procedure within SPSS.

Results

Morph structure

Of the 25 populations that we sampled, 19 were isoplethic and these were located in the western (MJ, YP, PM, MD, YJ, TC, JD, YD, GM and XLS) or eastern (MG, FD, YM, FN, NP, JX, TY, FS and HC) portions of the range. The remaining six populations were distributed at the south-eastern periphery of the western concentration of isoplethic populations and were either L-biased populations (SJ and SP) or L-monomorphic (LC, JP, PB and LJS). Figure 1a illustrates the geographical distribution of populations with contrasting morph structures

and Table 1 provides data on the L-morph frequencies of populations.

CpDNA data

Of the two *cp*DNA regions sequenced in our sample of 301 individuals from 25 populations, *trn*L-*trn*F had the greatest variability with six polymorphisms detected in 607 aligned positions (0.98%), followed by *rpl*20-*rps*12 (4/381; 1.04%). These combined sequences (986–991 bp) were aligned with a consensus length of 990 bp and generated 10 haplotypes (HapA–J) from the 10 polymorphisms (Table S2, Supporting information). Nucleotide sequences for the two *cp*DNA regions have been

Table 1 Summary statistics for *Luculia pinceana* populations sampled for *cp*DNA and SSR variation from south-west China and adjacent countries (including the outgroup *L. gratissima*): latitude and longitude of sampling location; frequencies of haplotypes (A-J); and sample size (*n*) for *cp*DNA/SSR. The frequency of the L-morph is also provided for each population. A dash indicates missing data or not applicable

Region/population code	Locations	Latitude (N), longitude (E)	Altitude (m)	L-morph frequency (sample size)[†]	Haplotypes (number of individuals)	*n* (*cp*DNA/ SSR)
Western isoplethic populations						
MJ	China: Yunnan, Maji	27°27', 98°30'	1370–1680	0.46 (98)	A	12/10
YP	China: Yunnan, Yaping	27°01', 98°49'	1740	0.46 (84)	A	12/12
PM	China: Yunnan, Pianma	25°59', 98°48'	1870	0.56 (50)	A(6),B(6)	12/12
MD	Myanmar: Kachin, Hkawng Lamhpu	27°05', 98°22'	1600	— (24)	A	12/12
YJ	China: Yunnan, Yinjiang	25°10', 97°59'	1830	0.58 (29)	A	12/14
TC	China: Yunnan, Tengchong	25°15', 98°30'	1750	0.47 (51)	A(4),C(3), D(5)	12/12
JD	China: Yunnan, Jindong	24°47', 100°31'	1790	0.55 (40)	B	12/15
YD	China: Yunnan, Yongde	24°07', 99°44'	1400–1600	0.54 (251)	E	12/17
GM	China: Yunnan, Gengma	23°24', 99°12'	1225	0.53 (63)	E	12/12
XLS	China: Yunnan, Lancang	22°46', 99°46'	1760	0.48 (285)	E	12/12
Central nonisoplethic populations (L-monomorphic or L-biased)						
SJ	China: Yunnan, Shuangjiang	23°22', 100°00'	1750	**0.80 (46)**	E	12/12
LC	China: Yunnan, Lvchun	22°58', 102°28'	1840	1.00 (163)	E	12/12
JP	China: Yunnan, Jinping	22°54', 103°13'	1470	1.00 (214)	E	12/18
PB	China: Yunnan, Pingbian	22°57', 103°42'	1610	1.00 (156)	E	12/12
LJS	China: Yunnan, Laojunshan	23°16', 103°58'	1850	1.00 (129)	E	12/12
SP	Vietnam: Laocai, Sa Pa	22°22', 104°03'	1600	**0.84 (38)**	E	12/12
Eastern isoplethic populations						
MG	China: Yunnan, Maguan	22°57', 104°19'	1410	0.48 (31)	E(2),F(6), G(4),H(1)	13/12
FD	China: Yunnan, Fadou	23°23', 104°46'	1490	0.44 (27)	I(6),J(6)	12/11
YM	Vietnam: Hagiang, Banbo	23°06', 104°55'	940	0.54 (24)	F	12/12
FN	China: Yunnan, Funing	23°22', 105°20'	1100	0.43 (53)	F	12/12
NP	China: Guanxi, Napo	23°21', 105°54'	1170	0.48 (41)	F	12/17
JX	China: Guanxi, Jinxi	23°01', 106°22'	730	0.56 (39)	F	12/12
TY	China: Guanxi, Tianyang	23°33', 106°32'	890	0.57 (135)	F	12/12
FS	China: Guanxi, Fengshan	24°25', 106°58'	750	0.52 (42)	F	12/15
HC	China: Guanxi, Hechi	24°33', 107°45'	330	0.57 (73)	F	12/12
L.gratissima	Nepal: Makwanpur, Daman	27°40', 85°05'	1970	— (30)		1/—

[†]Boldfaced type indicates morph frequencies that were significantly different from 1:1 following *G* tests.

deposited in GenBank (accession nos. HQ174524– HQ174534). The geographic distribution of haplotypes (HapA-J) among the 25 populations of *L. pinceana* is illustrated in Fig. 1b, while the frequencies of haplotypes at each location are presented in Table 1. The most frequent haplotype E occurred in 110 individuals (36.5% of all samples), but was restricted to the six nonisoplethic populations and three of the 10 western isoplethic populations. Haplotypes F and A were also common, but were restricted to eastern and western isoplethic populations, respectively. Haplotype B was shared by two western isoplethic populations PM and JD. Haplotypes C and D, G and H, and I and J were unique to population TC, MG and FD, respectively.

Maximum parsimony analysis of the 10 *cp*DNA haplotypes of *L. pinceana*, with *L. gratissima* used as an outgroup, resulted in three most parsimonious trees each with a length of 26 steps, a consistency index (CI) of 0.786, and a retention index (RI) of 0.800. The Bayesian (BI) tree agreed with the strict consensus MP tree at nearly all nodes. In the strict consensus tree (Fig. 2), with posterior probabilities and bootstrap values assigned, haplotypes F, H and J formed a clade with 1.00 posterior probability (PP) and 66% bootstrap support (BS), while the other seven haplotypes formed another clade (0.92 PP, 52% BS). Although the clade including haplotypes A, B, C, D and E was not well supported (bootstrap value < 50%), it did show a 0.93 posterior probability value.
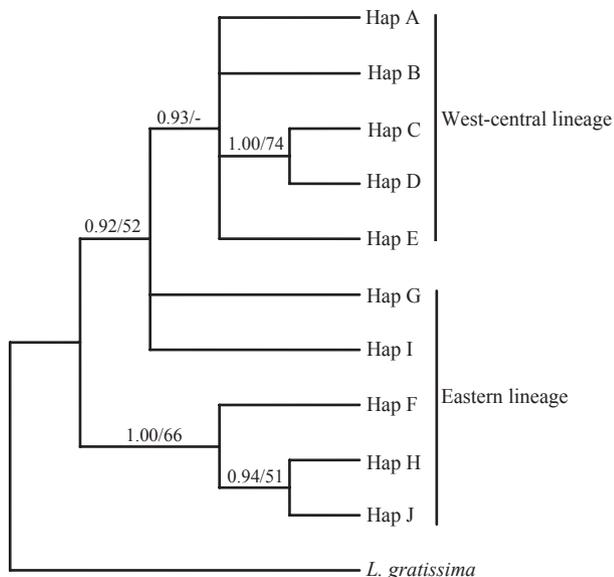
Fig. 2 Strict consensus tree obtained by analysis of 10 *cp*DNA (*trn*L-*trn*F and *rpl*20-*rps*12) haplotypes of *Luculia pinceana*, with *L. gratissima* used as the outgroup. The numbers on the left of branches indicate posterior probabilities from the Bayesian analysis; the numbers on the right indicate bootstrap values (> 50%).

The network analysis (Fig. 1c) obtained using TCS yielded a similar structure to the phylogenetic tree (Fig. 2) and organized all haplotypes into two groups (i.e. the west-central group vs. eastern group). On the basis of their location in the network, haplotypes E and F appear to be ancestral (interior) (Clement *et al.* 2000) in the west-central and the eastern groups, respectively, with haplotypes A and D derived from haplotype E, and haplotypes I and J from haplotype F.

By applying a fossil-calibrated relaxed molecular clock to the Rubiaceae *rbc*L data set (Fig. S1, Supporting information), the *Luculia* crown group age was estimated to be 5.462 Mya (95% HPD 0.219–15.198 Mya, median 4.013 Mya). According to the estimated prior tMRCA for the genus, the estimated divergence of the west-central and eastern lineages in *L. pinceana* was dated to c. 1.982 Mya (95% HPD 0.134–4.775 Mya, median 1.642 Mya) (Fig. S2, Supporting information).

The mismatch distribution of *cp*DNA haplotypes in the eastern clade was clearly bimodal and differed strongly from that predicted by a model of sudden population expansion (Fig. 3a). This prediction was also reflected in a significant SSD statistic (SSD = 0.027, $P = 0.011$), a high and significant $HRag$ value ($HRag = 0.175$, $P = 0.009$), as well as a positive and nonsignificant Fu's F_s value (Fu's $F_s = 1.775$, $P = 0.802$). Together, these data suggest relatively stable population sizes in the east. By contrast, the unimodal mismatch distribution for the west-central clade is consistent with the expected distribution from expansion (Fig. 3b), as supported by uniformly nonsignificant SSD and $HRag$ values (SSD = 0.071, $P = 0.118$; $HRag = 0.404$, $P = 0.425$), and a marginally significant negative Fu's F_s value (Fu's $F_s = -0.092$, $P = 0.052$) at the lineage-wide scale.

SSR data

The total number of alleles per locus in our sample of 321 individuals ranged from seven (LP4) to 36 (LP198) with a total of 135 alleles for the seven loci. Departure from HWE was not consistent across loci; however, many populations departed from HWE as a result of an excess number of homozygotes. Diversity estimates varied among microsatellite loci, among populations and among regions. Values for numbers of alleles (N_A) ranged from 3.00 to 6.71, effective numbers of alleles (N_E) from 1.88 to 4.60, observed heterozygosity (H_O) from 0.06 to 0.46 and expected heterozygosity (H_E) from 0.40 to 0.78. Allelic richness (A_R) estimates among all populations ranged from 2.91 to 6.16, with an average (±SD) of 4.05 ± 0.83. Genetic diversity (N_A, N_E, H_O, H_E and A_R) was significantly greater (all $P < 0.01$) in the western isoplethic populations compared to the central nonisoplethic populations. Similar patterns of genetic
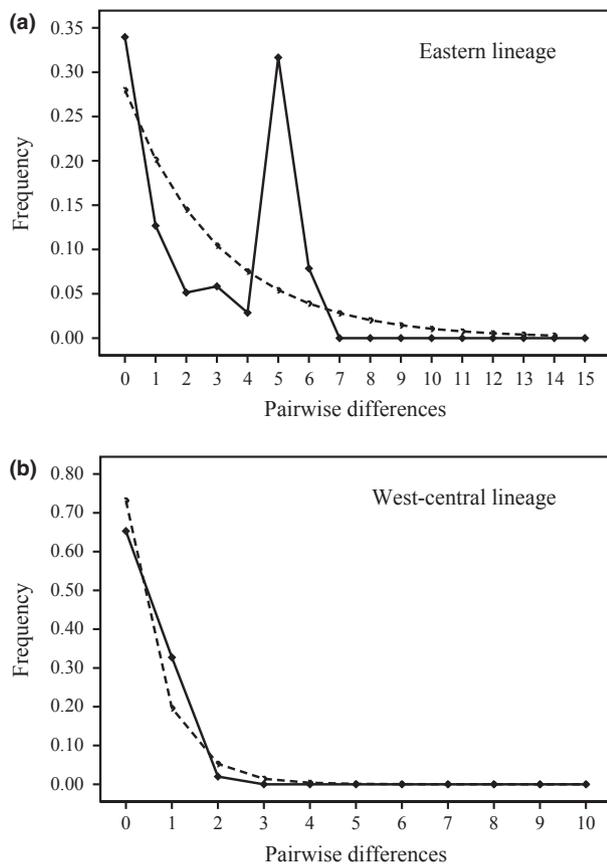
Fig. 3 Distribution of the number of pairwise nucleotide differences for *cp*DNA sequence data in: (a) the eastern lineage (b) the west-central lineage of *Luculia pinceana*. The solid line shows observed values, whereas the dashed line represents expected values under a model of sudden (stepwise) population expansion (see Rogers & Harpending 1992).

diversity were found between the eastern isoplethic populations and the central nonisoplethic populations, except with respect to H_O and H_E (Table 2). F_{IS} values ranged from 0.34 to 0.89, with an average of 0.54 ± 0.15. This indicates considerable variation in the intensity of inbreeding among populations. Further analysis indicated that F_{IS} values were significantly lower ($P < 0.01$)

in the western and eastern isoplethic populations than in central nonisoplethic populations (Table 2).

Two-dimensional PCO of SSR phenotypes of *L. pinceana* (Fig. 4) separated all individuals from eastern and western populations along the first axis (PCO1, explaining 44.85% of the total variance), and both isoplethic clusters were separated from the central nonisoplethic samples along PCO2 (16.46%). The population-based NJ tree (Fig. 6) confirmed this strong pattern of regional differentiation by a high bootstrap value (96%) separating west-central and eastern clusters, and a relative lower bootstrap value (60%) between western and central clusters. In the SSR admixture analysis employing STRUCTURE (Fig. 5), the highest likelihood of the data was obtained when samples were clustered into two groups ($K = 2$). Using $K = 3$ in STRUCTURE, all central nonisoplethic populations were separated from the west-central group and formed an independent group (i.e. central group). The BARRIER analysis revealed two major genetic boundaries, with 96% and 97% mean bootstrap support in Barrier one (B I) and Barrier two (B II), respectively, which separated the central nonisoplethic populations from the eastern and western isoplethic populations (Fig. S3, Supporting information).

Floral variation

Patterns of floral variation differed significantly between isoplethic and nonisoplethic populations: corolla tube length, stigma height and anther height were all significantly larger in nonisoplethic populations compare to isoplethic populations (all $P < 0.01$). However, these differences were only evident in the L-morph; the floral morphology of the S-morph did not differ between the two types of population morph structure (Table 3). Stigma-anther separation in the L-morph was significantly reduced ($P < 0.01$) in L-monomorphic populations (mean = 3.13) relative to isoplethic populations (mean = 3.98) with values in anisoplethic populations intermediate (mean = 3.58).

Table 2 A comparison of genetic diversity based on SSR loci among populations of *Luculia pinceana* sampled from south-west China and adjacent countries: number of alleles (N_A); effective number of alleles (N_E); expected (H_E) and observed heterozygosity (H_O); inbreeding index (F_{IS}); allelic richness (A_R)

Region	Morph ratios	Population number	N_A	N_E	H_E	H_O	F_{IS}	A_R
Western	Isoplethic	10	4.99 ± 0.97[a]	3.31 ± 0.68[a]	0.65 ± 0.10[a]	0.33 ± 0.08[a]	0.47 ± 0.08[a]	4.71 ± 0.88[a]
Central	Nonisoplethic	4/2	3.76 ± 0.68[b]	2.46 ± 0.41[b]	0.48 ± 0.04[b]	0.13 ± 0.07[b]	0.73 ± 0.16[b]	3.55 ± 0.52[b]
Eastern	Isoplethic	9	3.90 ± 0.59[b]	2.58 ± 0.37[b]	0.50 ± 0.06[a]	0.25 ± 0.05[a]	0.48 ± 0.07[a]	3.68 ± 0.44[b]
Total		25	4.30 ± 0.94	2.84 ± 0.63	0.55 ± 0.10	0.25 ± 0.10	0.54 ± 0.15	4.05 ± 0.83

Different letters indicate significant differences (Turkey's test) at $P < 0.01$.
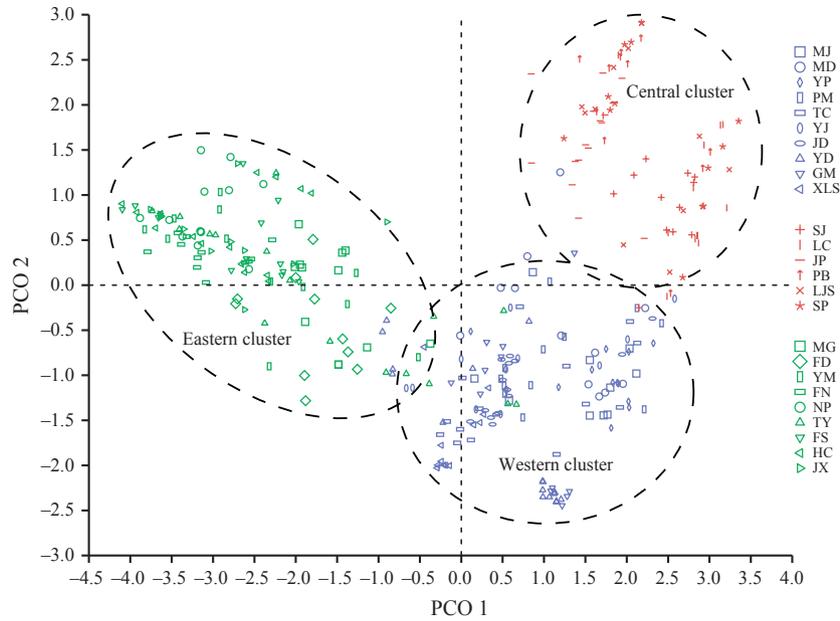
Fig. 4 Principal coordinates analysis of SSR phenotypes from 25 populations and 321 individuals of *Luculia pinceana*. The first and second axes extracted 44.85% and 16.46% of the total genetic variance, respectively. The symbols MJ-JX on the right side of the figure represent population codes, which are identified in Table 1. Colour coding corresponds to populations clusters (*K* = 3) identified by the STRUCTURE analysis (see Fig. 5) and neighbour-joining tree (see Fig. 6).
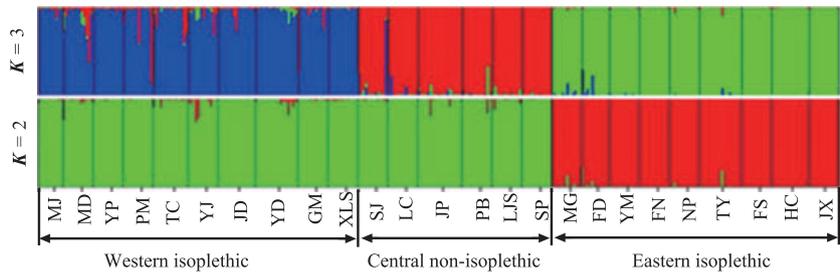
Fig. 5 Estimated genetic clustering (*K* = 2 and *K* = 3) obtained with the STRUCTURE program for 25 populations of *Luculia pinceana* (321 individuals) based on SSR variation. Each vertical bar represents an individual and its estimated membership fractions in *K* clusters. Black lines separate different populations. Population codes are identified in Table 1.

Discussion

The major finding of this study is the occurrence of marked phylogeographic structure in the distribution of distylous *L. pinceana* among the populations from south-west China and adjacent countries. The patterns of molecular variation are consistent with a splitting of the geographical distribution of *L. pinceana* during the Early Pleistocene into two major lineages that today are both characterized by populations with equilibrium (1:1) floral morph frequencies. A cluster of L-morph-biased and L-monomorphic populations are distributed between these two isoplethic lineages and are characterized by reduced genetic diversity and differentiated floral morphology. The populations appear to have been derived from the western lineages by a later dispersal event. Below we discuss these findings and propose a hypothesis to account for the breakdown in stylar dimorphism and the origin of monomorphism.

Phylogeographic and demographic history of Luculia pinceana

The tree topology (Fig. 2) and network analysis (Fig. 1c) of *cp*DNA data are consistent with the hypothesis that *L. pinceana* was initially split into west-central and eastern lineages. However, the results of SSR genotyping indicated that the species has differentiated into three distinct groups: the western isoplethic group (blue coded in Figs 4–6), the central nonisoplethic group (red

Table 3 Tests of differences in floral traits between populations of *Luculia pinceana* with isoplethic and nonisoplethic morph structure using nested analysis of variance (ANOVA). We calculated error degrees of freedom between morph structures using Satterthwaite's approximation. Measurements of floral traits (mean ± SE) are in mm. A dash indicates not applicable data

	Stigma height		Anther height		Stigma-anther separation		Corolla tube length	
	L-morph	S-morph	L-morph	S-morph	L-morph	S-morph	L-morph	S-morph
Between morph structures (isoplethic and nonisoplethic)[†]	$F_{1,\,17.75} = 105.60^{**}$	$F_{1,\,11.49} = 0.27^{NS}$	$F_{1,\,16.22} = 247.86^{**}$	$F_{1,\,6.43} = 0.00^{NS}$	$F_{1,\,18.68} = 34.29^{**}$	$F_{1,\,12.24} = 0.13^{NS}$	$F_{1,\,18.14} = 32.82^{**}$	$F_{1,\,6.43} = 0.27^{NS}$
Populations within morph structure	$F_{20,\,134} = 2.865^{**}$	$F_{16,\,59} = 2.80^{*}$	$F_{20,\,134} = 1.67^{NS}$	$F_{16,\,59} = 1.18^{NS}$	$F_{20,\,134} = 1.96^{NS}$	$F_{16,\,59} = 1.40^{NS}$	$F_{20,\,134} = 3.48^{**}$	$F_{16,\,59} = 1.18^{NS}$
Morph structure means								
Isoplethic	30.75 ± 0.25	19.80 ± 0.35	22.74 ± 0.18	29.75 ± 0.47	3.98 ± 0.17	4.40 ± 0.15	25.43 ± 0.22	28.89 ± 0.31
L-biased	35.92 ± 0.17	20.12 ± 0.29	27.66 ± 0.35	29.64 ± 0.33	3.58 ± 0.11	4.05 ± 0.12	29.73 ± 0.30	29.07 ± 0.34
L-monomorphic	36.62 ± 0.41	—	28.57 ± 0.43	—	3.13 ± 0.20	—	29.36 ± 0.34	—

$^{**}P < 0.01$, $^{*}P < 0.05$, NS not significant.
[†]L-biased populations were pooled with monomorphic populations as the nonisoplethic group.

coded) and the eastern isoplethic group (green coded). These pronounced phylogeographic divisions are also indicated by the spatial BARRIER analysis (Fig. S3, Supporting information). The contrast in lineage resolution between the chloroplast data and microsatellite data undoubtedly result, in part, from asymmetrical gene flow as a result of different transmission modes and rates of divergence for plastid vs. nuclear DNA (Du *et al.* 2009; Bai *et al.* 2010). However, they may also reflect contemporary events associated with gene flow and population migration.

According to our fossil-calibrated relaxed molecular clock analysis, the separation between the west-central and eastern lineages detected in *cp*DNA most likely occurred at the beginning of the Early Pleistocene, about 1.982 Mya (95% HPD 0.134–4.775 Mya, median 1.642 Mya). This major split in the species distribution colocalizes with a biogeographic boundary, the Tanaka Line, which is located approximately from 28°N, 98° E southward to 19°N, 108°E (Li & Li 1997 and see Fig. 1). The same pattern involving the Tanaka Line has also been demonstrated in recent phylogeographic analyses of *Tacca* (Zhang *et al.* 2006) and *Dipentodon* (Yuan *et al.* 2008) in this region. On either side of the Tanaka Line, populations of *L. pinceana* were characterized by unique *cp*DNA haplotypes (except haplotype E in MG) and reciprocal SSR monophyly. This pattern suggests that the significant levels of gene flow have not occurred between these lineages for ~1.982 My. While this region encompasses the well-known boundary of the Honghe River (and the paleo-Honghe), the reduction in gene flow may not necessarily be entirely associated with the river itself, and geographical isolation resulting from habitat fragmentation may have played a more prominent role in the divergence of the two lineages.

Biome reconstructions in south-west China since the last glacial maximum (LGM: ~0.018 Mya) using fossil pollen data (Harrison *et al.* 2001) indicate that the warm-temperate evergreen forests favoured by *L. pinceana* in this region were separated by the emergence of nonforested areas at central sites in this study (west of the Tanaka Line), including northern Laos. Therefore, we presume that conditions favouring the persistence of centrally distributed populations of *L. pinceana* deteriorated during the LGM or earlier, resulting in the eradication of populations in this area. The later reappearance of forest during the post- or inter-glacial periods (Harrison *et al.* 2001) most likely provided an opportunity for recolonization of populations and the formation of a new connection between the two lineages.

Postglacial migration has had a profound influence on the population structure of numerous plant species. In south-west China, this process has played an
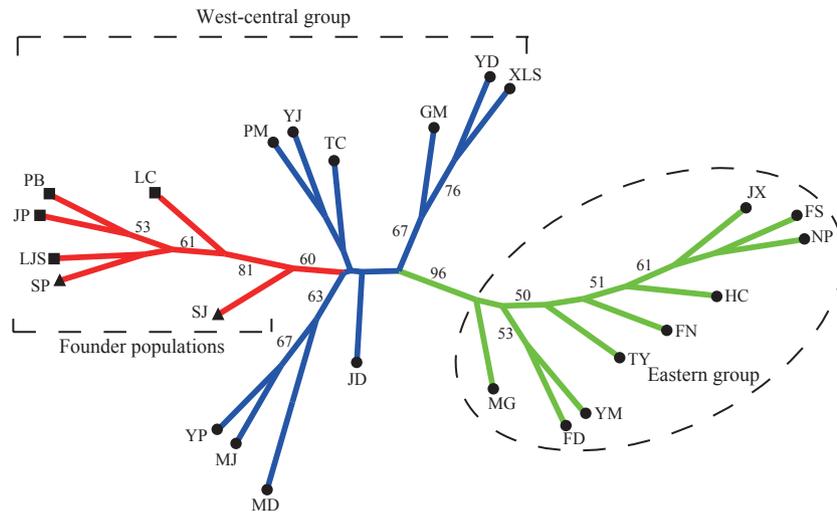
Fig. 6 Neighbour-joining tree illustrating the genetic relationships among 25 populations of *Luculia pinceana*, based on Nei's (1987) unbiased genetic distance (*D*) calculated from SSR data. Population codes are identified in Table 1, and branches are colour coded as in Figs 4 and 5 (*K* = 3). Circles represent isoplethic populations, whereas squares and triangles represent L-monomorphic and L-biased populations, respectively. Numbers by nodes are bootstrap values (> 50%) from 1000 replicates.

important role in affecting contemporary patterns of genetic variation in several taxa (reviewed in Qiu *et al.* 2011). As would be predicted if the central nonisoplethic lineage was recently derived from the western isoplethic lineage, the haplotypes found in nonisoplethic populations represent a restricted subset of those found in this isoplethic region. Indeed, all individuals from nonisoplethic populations have the same haplotype (HapE), and this haplotype is widely shared by the western isoplethic lineage, for example, populations YD, GM and XLS (Fig. 1b). The distribution pattern of haplotype E strongly suggests that eastward dispersal occurred relatively recently. Accordingly, the STRUCTURE (Fig. 5) and NJ tree (Fig. 6) analyses based on the SSR data also indicate that the six populations of the central group are more similar genetically to the western group, despite the fact that most of these populations are geographically closer to the eastern isoplethic lineage. The derived condition of the central lineage is also supported by a significant loss of genetic diversity during dispersal eastwards from the western cluster of populations (Table 2).

The eastern migration of *L. pinceana* to form the central lineage may have involved founder events, in which only a fraction of the genetic variation present in the source population establishes in the colonized region (e.g. Nei *et al.* 1975; Dlugosch & Parker 2008; Zhang *et al.* 2010). Furthermore, according to the mismatch analysis for the west-central lineage, both the unimodal mismatch distribution (Fig. 3b) and Fu's F_S provide evidence for a relatively recent demographic expansion of this lineage. The expansion eastward may

be directly responsible for the origin of central populations and also resulted in secondary contact between the west-central and eastern isoplethic lineages, which had earlier experienced a longer period of isolation.

Loss of stylar polymorphism

What population level processes might explain the loss of stylar dimorphism and origin of monomorphic populations in *L. pinceana*? Differences in the self-compatibility status of floral morphs and the pattern of inheritance for distyly provide a plausible explanation for the origin of stylar monomorphism in this species. Experimental pollinations of floral morphs in *L. pinceana* indicate that while the S-morph is self-incompatible, the L-morph is self-compatible (Ma *et al.* 2009). It is well established that self-compatibility can promote establishment following long distance dispersal, even for single propagules (Baker 1955), and, in heterostylous species, derived floral monomorphism is often associated with self-compatibility (reviewed in Barrett 1989; Weller 1992). Thus, founder events involving the L-morph rather than the S-morph would be more likely to establish new populations because of its capacity for self-fertilization. The inheritance of distyly in *L. pinceana* has not been investigated; however, if the common pattern for distylous species (L-morph *ss*; S-morph *Ss*; Lewis & Jones 1992) also occurs, colonizing plants of the L-morph would only produce individuals of the L-morph thus creating monomorphic populations.

Another advantage that the L-morph of *L. pinceana* might experience in founding new populations may be

associated with differences in the pollination biology of the floral morphs. The concealed stigma of the S-morph requires long-tongued pollinators for successful pollination and seed set. In contrast, the exserted stigma of the L-morph is more accessible to a wide range of pollinators, particularly bees. Thus, in small founding populations, the L-morph would be more likely to maintain fertility than the S-morph. Such differences in pollination biology have been documented in stylar dimorphic *Narcissus papyraceus* and explain the high incidence of L-monomorphism and L-biased populations in this species (Arroyo *et al.* 2002; Pérez-Barrales & Arroyo 2010).

Two of the nonisoplethic populations of *L. pinceana* that we sampled were L-morph biased. These were geographically separated at the western (SJ) and eastern (SP) edges of the central cluster of low diversity populations fixed for the E haplotype. It is not clear whether the presence of the S-morph in these populations has resulted from later gene flow events, involving dispersal of the S-morph from nearby distylous population, or whether selective processes in long-established dimorphic populations favour the L-morph at the expense of the S-morph. Regardless of the residency time of the S-morph in these populations, disassortative mating should establish 1:1 morph ratios, as occur elsewhere in the range. The fact that this has not occurred suggests that mating patterns in these populations are not strictly disassortative, perhaps because of selfing in the L-morph or that because of demographic factors associated with low seedling establishment and perenniality the approach to an isoplethic equilibrium is slow.

The differentiation in floral morphology that we observed between isoplethic and nonisoplethic populations was restricted to the L-morph only (Table 3). The flowers of this morph were generally larger in size in nonisoplethic populations, with higher values for the dimensions of female and male sex organs and longer corolla tubes. In contrast, the morphology of S-morph flowers was not significantly different between isoplethic and nonisoplethic populations, although because of the limited representation of the S-morph in anisoplethic populations, this result should be treated cautiously. The apparent difference between the floral morphology of the L- and S-morphs in anisoplethic populations would be predicted if, as discussed earlier, the S-morph was more recently introduced from dimorphic to monomorphic populations, rather than being resident in dimorphic populations since their initiation. According to this hypothesis, the differentiation between isoplethic and nonisoplethic populations in floral traits of the L-morph results from morphological adjustment to the prevailing pollination conditions and opportunities for disassortative mating.

In heterostylous species, populations with stylar monomorphism commonly exhibit reductions in flower size and stigma–anther separation and increased selfing rates (e.g. Barrett *et al.* 1989). Although stigma–anther separation was reduced in nonisoplethic populations of *L. pinceana*, implicating increased rates of self-pollination, this was not associated with a reduction in flower size. Indeed, as discussed earlier, sex organs and floral tubes were larger in flowers from monomorphic populations than in those from distylous populations. Further research is required to explain these unexpected patterns of floral variation. In particular, studies of mating patterns would be valuable to determine the extent to which the reproductive biology of populations contributes towards their morph structure, in addition to the historical factors that we have identified.

Acknowledgements

References

Arroyo J, Barrett SCH, Hidalgo R, Cole WW (2002) Evolutionary maintenance of stigma-height dimorphism in *Narcissus papyraceus* (Amaryllidaceae). *American Journal of Botany*, **89**, 1242–1249.

Bai WN, Liao WJ, Zhang DY (2010) Nuclear and chloroplast DNA phylogeography reveal two refuge areas with asymmetrical gene flow in a temperate walnut tree from East Asia. *New Phytologist*, **188**, 892–901.

Baker HG (1955) Self-compatibility and establishment after long-distance dispersal. *Evolution*, **9**, 347–349.

Baker HG (1959) The contribution of autecological and genecological studies to our knowledge of the past migration of plants. *American Naturalist*, **93**, 255–272.

Barrett SCH (1985) Floral trimorphism and monomorphism in continental and island populations of *Eichhornia paniculata* (Pontederiaceae). *Biological Journal of the Linnean Society*, **25**, 41–60.

Barrett SCH (1989) The evolutionary breakdown of heterostyly. In: *The Evolutionary Ecology of Plants* (eds Linhart Y, Bock J), pp. 151–169. Westview Press, Boulder, Colorado.

Barrett SCH (1993) The evolutionary biology of tristly. In: *Oxford Surveys in Evolutionary Biology* (eds Futuyma D, Antonovics J), pp. 283–326. Oxford University Press, Oxford.

Barrett SCH, Cruzan MB (1994) Incompatibility in heterostylous plants. In: *Genetic Control of Self-Incompatibility*

and *Reproductive Development in Flowering Plants* (eds Williams EG, Knox RB and Clarke AE), pp. 189–219. Kluwer Academic Publishers, Dordrecht.

Barrett SCH, Forno IW (1982) Style morph distribution in new world populations of *Eichhornia crassipes* (Mart.) Solms-Laubach (water hyacinth). *Aquatic Botany*, **13**, 299–306.

Barrett SCH, Shore JS (1987) Variation and evolution of breeding systems in the *Turnera ulmifolia* L. complex (Turneraceae). *Evolution*, **41**, 340–354.

Barrett SCH, Price SD, Shore JS (1983) Male fertility and anisoplethic population structure in tristylous *Pontederia cordata* (Pontederiaceae). *Evolution*, **37**, 745–759.

Barrett SCH, Morgan MT, Husband BC (1989) Dissolution of a complex genetic polymorphism: the evolution of self-fertilization in tristylous *Eichhornia paniculata* (Pontederiaceae). *Evolution*, **43**, 1398–1416.

Barrett SCH, Harder LD, Cole WW (2004) Correlated evolution of floral morphology and mating-type frequencies in a sexually polymorphic plant. *Evolution*, **58**, 964–975.

Barrett SCH, Yakimowski SB, Field DL, Pickup M (2010) Ecological genetics of sex ratios in plant populations. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, **365**, 2549–2558.

Brys R, Jacquemyn H, Hermy M (2007) Impact of mate availability, population size, and spatial aggregation of morphs on sexual reproduction in a distylous, aquatic plant. *American Journal of Botany*, **94**, 119–127.

Brys R, Jacquemyn H, Beeckman T (2008) Morph-ratio variation, population size and female reproductive success in distylous *Pulmonaria officinalis* (Boraginaceae). *Journal of Evolutionary Biology*, **21**, 1281–1289.

Castro S, Loureiro J, Santos C, Ater M, Ayensa G, Navarro L (2007) Distribution of flower morphs, ploidy level and sexual reproduction of the invasive weed *Oxalis pes-caprae* in the western area of the Mediterranean region. *Annals of Botany*, **99**, 507–517.

Clement M, Posada D, Crandall KA (2000) TCS: a computer program to estimate gene genealogies. *Molecular Ecology*, **9**, 1657–1659.

Dlugosch KM, Parker IM (2008) Founding events in species invasions: genetic variation, adaptive evolution, and the role of multiple introductions. *Molecular Ecology*, **17**, 431–449.

Dorken ME, Barrett SCH (2004) Chloroplast haplotype variation among monoecious and dioecious populations of *Sagittaria latifolia* (Alismataceae) in eastern North America. *Molecular Ecology*, **13**, 2699–2707.

Doyle J (1991) DNA protocols for plants: CTAB total DNA isolation. In: *Molecular Techniques in Taxonomy* (eds Hewitt GM and Johnston A), pp. 283–293. Springer-Verlag, Berlin.

Drummond AJ, Rambaut A (2007) BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evolutionary Biology*, **7**, 214.

Du FK, Petit RJ, Liu JQ (2009) More introgression with less gene flow: chloroplast vs. mitochondrial DNA in the *Picea asperata* complex in China, and comparison with other conifers. *Molecular Ecology*, **18**, 1396–1407.

Eckert CG, Barrett SCH (1992) Stochastic loss of style morphs from populations of tristylous *Lythrum salicaria* and *Decodon verticillatus* (Lythraceae). *Evolution*, **46**, 1014–1029.

Eckert CG, Barrett SCH (1995) Style morph ratios in tristylous *Decodon verticillatus* (Lythraceae): selection vs. historical contingency. *Ecology*, **76**, 1051–1066.

Eckert CG, Manicacci D, Barrett SCH (1996) Frequency-dependent selection on morph ratios in tristylous *Lythrum salicaria* (Lythraceae). *Heredity*, **77**, 581–588.

Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology*, **14**, 2611–2620.

Felsenstein J (2005) *PHYLIP (Phylogeny Infrence Package) Version 3.6*. Department of Genome Sciences, University of Washington, Seattle, WA.

Fu YX (1997) Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. *Genetics*, **147**, 915–925.

Ganders FR (1979) The biology of heterostyly. *New Zealand Journal of Botany*, **17**, 607–635.

Goudet J (1995) FSTAT (version 1.2): a computer program to calculate F-statistics. *Journal of Heredity*, **86**, 485–486.

Graham A (1985) Studies of Neotropical paleobotany. IV. The Eocene communities of Panama. *Annals of the Missouri Botanical Garden*, **72**, 504–534.

Hall TA (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symposium*, **41**, 95–98.

Harpending HC (1994) Signature of ancient population growth in a low-resolution mitochondrial DNA mismatch distribution. *Human Biology*, **66**, 591–600.

Harrison SP, Yu G, Takahara H, Prentice IC (2001) Palaeovegetation: diversity of temperate plants in East Asia. *Nature*, **413**, 129–130.

Heuch I (1979) Equilibrium populations of heterostylous plants. *Theoretical Population Biology*, **15**, 43–57.

Hodgins KA, Barrett SCH (2007) Population structure and genetic diversity in tristylous *Narcissus triandrus*: insights from microsatellite and chloroplast DNA variation. *Molecular Ecology*, **16**, 2317–2332.

Hu SY (1951) Notes on the Flora of China. *Journal of the Arnold Arboretum*, **32**, 390–402.

Husband BC, Barrett SCH (1992) Genetic drift and the maintenance of the style length polymorphism in tristylous populations of *Eichhornia paniculata* (Pontederiaceae). *Heredity*, **69**, 440–449.

Kéry M, Matthies D, Schmid B (2003) Demographic stochasticity in population fragments of the declining distylous perennial *Primula veris* (Primulaceae). *Basic and Applied Ecology*, **4**, 197–206.

Kovach WL (1999) *MVSP: A Multivariate Statistical Package for Windows, Version 3.1*. Kovach Computing Services, Pentraeth.

Lewis D, Jones DA (1992) The genetics of heterostyly. In: *Evolution and Function of Heterostyly* (ed Barrett SCH), pp. 129–148. Springer-Verlag, Berlin.

Li XW, Li J (1997) The Tanaka-Kaiyong line – an important floristic line for the study of the flora of East Asia. *Annals of the Missouri Botanical Garden*, **84**, 888–892.

Lou XR, Gao WZ, Chen WQ, Xu XH, Wu H (1999) *Flora Republicae Popularis Sinicae Vol. 71/1*. Science Press, Beijing.

Luo S, Zhang D, Renner SS (2006) *Oxalis debilis* in China: distribution of flower morphs, sterile pollen and polyploidy. *Annals of Botany*, **98**, 459–464.

Ma H, Wang Y, Li ZH, Wan YM, Liu XX, Liang N (2009) A study on the breeding system of *Luculia pinciana*. *Forest Research*, **22**, 373–378 (In Chinese)

Manchester SR (1999) Biogeographical relationships of North American Tertiary floras. *Annals of the Missouri Botanical Garden*, **86**, 472–522.

Manni F, Guerard E, Heyer E (2004) Geographic patterns of (genetic, morphologic, linguistic) variation: How barriers can be detected by using Monmonier's algorithm. *Human Biology*, **76**, 173–190.

Morgan M, Barrett SCH (1988) Historical factors and anisoplethic population structure in tristylous *Pontederia cordata* L. *Evolution*, **42**, 496–504.

Müller K (2005) SeqState: primer design and sequence statistics for phylogenetic DNA datasets. *Applied Bioinformatics*, **4**, 65–69.

Nei M (1987) *Molecular Evolutionary Genetics*. Columbia University Press, New York.

Nei M, Maruyama T, Chakraborty R (1975) The bottleneck effect and genetic variability in populations. *Evolution*, **29**, 1–10.

Ornduff R (1972) The breakdown of trimorphic incompatibility in *Oxalis* section Corniculatae. *Evolution*, **26**, 52–65.

Pailler T, Thompson JD (1997) Distyly and variation in heteromorphic incompatibility in *Gaertnera vaginata* (Rubiaceae) endemic to La Reunion Island. *American Journal of Botany*, **84**, 315–327.

Pérez-Alquicira J, Molina-Freaner FE, Piñero D *et al.* (2010) The role of historical factors and natural selection in the evolution of breeding systems of *Oxalis alpina* in the Sonoran desert 'Sky Islands'. *Journal of Evolutionary Biology*, **23**, 2163–2175.

Pérez-Barrales R, Arroyo J (2010) Pollinator shifts and the loss of style polymorphism in *Narcissus papyraceus* (Amaryllidaceae). *Journal of Evolutionary Biology*, **23**, 1117–1128.

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 and Evolution*, **59**, 225–244.

Raymond M, Rousset F (1995) GENEPOP (version 1.2): population genetics software for exact tests and ecumenicism. *Journal of Heredity*, **86**, 248–249.

Rogers AR, Harpending H (1992) Population growth makes waves in the distribution of pairwise genetic differences. *Molecular Biology and Evolution*, **9**, 552–569.

Ronquist F, Huelsenbeck JP (2003) MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics*, **19**, 1572–1574.

Roth JL, Dilcher DL (1979) Investigations of angiosperms from the Eocene of North America: stipulate leaves of the Rubiaceae including a probable polyploid population. *American Journal of Botany*, **66**, 1194–1207.

Satterthwaite FE (1946) An approximate distribution of estimates of variance components. *Biometrics Bulletin*, **2**, 110–114.

Schneider S, Excoffier L (1999) Estimation of past demographic parameters from the distribution of pairwise differences when the mutation rates very among sites: application to human mitochondrial DNA. *Genetics*, **152**, 1079–1089.

Shaw J, Lickey EB, Beck JT *et al.* (2005) The tortoise and the hare II: Relative utility of 21 noncoding chloroplast DNA sequences for phylogenetic analysis. *American Journal of Botany*, **92**, 142–166.

Sokal RR, Rohlf FJ (1995) *Biometry: The Principles and Practice of Statistics in Biological Research*. Freeman, New York.

Sosenski P, Fornoni J, Molina-Freaner FE, Weller SG, Domínguez CA (2010) Changes in sexual organ reciprocity and phenotypic floral integration during the tristyly–distyly transition in *Oxalis alpina*. *New Phytologist*, **185**, 829–840.

Swofford DL (2002) *PAUP. Phylogenetic Analysis Using Parsimony (*and other methods), Version 4.0b10*. Sinauer Associates, Sunderland, Massachusetts.

Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG (1997) The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Research*, **25**, 4876–4882.

Thompson JD, Barrett SCH, Baker AM (2003) Frequency-dependent variation in reproductive success in *Narcissus*: implications for the maintenance of stigma-height dimorphism. *Proceedings of the Royal Society of London Series B-Biological Sciences*, **270**, 949–953.

Wang Y, Wang QF, Guo YH, Barrett SCH (2005) Reproductive consequences of interactions between clonal growth and sexual reproduction in *Nymphoides peltata*: a distylous aquatic plant. *New Phytologist*, **165**, 329–335.

Weller SG (1992) Evolutionary modifications of tristylous breeding systems. In: *Evolution and Function of Heterostyly* (ed Barrett SCH), pp. 247–270. Springer-Verlag, Berlin.

Yuan QJ, Zhang ZY, Peng H, Ge S (2008) Chloroplast phylogeography of *Dipentodon* (Dipentodontaceae) in southwest China and northern Vietnam. *Molecular Ecology*, **17**, 1054–1065.

Zar JH (1999) *Biostatistical Analysis*. Prentice-Hall, London.

Zhang L, Li QJ, Li HT, Chen J, Li DZ (2006) Genetic diversity and geographic differentiation in *Tacca chantrieri* (Taccaceae): an autonomous selfing plant with showy floral display. *Annals of Botany*, **98**, 449–457.

Zhang YY, Zhang DY, Barrett SCH (2010) Genetic uniformity characterizes the invasive spread of water hyacinth (*Eichhornia crassipes*), a clonal aquatic plant. *Molecular Ecology*, **19**, 1774–1786.

Zhou W, Wang H (2009) Heterostyly in angiosperms and its evolutionary significance. *Chinese Bulletin of Botany*, **44**, 742–751 (In Chinese)

Zhou W, Wang H, Li DZ, Yang JB (2010) Isolation and characterization of 13 microsatellite loci from *Luculia pinceana* (Rubiaceae), a typical distylous species. *HortScience*, **45**, 840–841.

Data accessibility

DNA sequences: Genbank accessions nos. HQ174524–HQ174534; haplotype sequence and microsatellite data provided as supplemental material with the online version of this manuscript.

Supporting information

Additional supporting information may be found in the online version of this article.

Fig. S1 Phylogeny and divergence time estimates of *Luculia* inferred from 42 *rbc*L sequences representing species from all three subfamilies of Rubiaceae.

Fig. S2 Phylogeny and divergence time estimates of west-central and eastern lineages in *Luculia pinceana* inferred from the combined plastid dataset of *trn*L-*trn*F and *rpl*20-*rps*12, with the prior age of the genus set to a normal distribution around the mean (5.462 Mya) estimated in the *rbc*L analysis.

Fig. S3 The boundaries (red lines) detected using the BARRIER program based on 1000 bootstrapped matrices of Nei's (1987) unbiased genetic distance (*D*) in a sample of 25 populations of *Luculia pinceana*.

Table S1 GenBank accession numbers for *rbc*L sequences used to date the age of *Luculia*.

Table S2 *Cp*DNA sequence polymorphisms detected in two intergenic spacer regions (*trn*L-*trn*F and *rpl*20-*rps*12) of *Luculia pinceana* identifying 10 haplotypes (HapA-J).

Data S1 Haplotype sequence.

Data S2 SSR data for *Luculia pinceana*.

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting information supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.