Research

New Phytologist

Phylogenomic analysis reveals multiple evolutionary origins of selfing from outcrossing in a lineage of heterostylous plants

Li Zhong[1,2], Spencer C. H. Barrett[3], Xin-Jia Wang[1,2], Zhi-Kun Wu[4], Hua-Ying Sun[1], De-Zhu Li[5], Hong Wang[1] and Wei Zhou[5]

[1]CAS Key Laboratory for Plant Diversity and Biogeography of East Asia, Kunming Institute of Botany, Chinese Academy of Sciences, 132 Lanhei Road, Kunming, Yunnan 650201, China; [2]University of Chinese Academy of Sciences, 19 Yuquan Road, Beijing 100049, China; [3]Department of Ecology and Evolutionary Biology, University of Toronto, Toronto, Ontario M5S 3B2, Canada; [4]Department of Pharmacy, Guiyang University of Traditional Chinese Medicine, Guiyang, Guizhou 550002, China; [5]Plant Germplasm and Genomics Centre, Germplasm Bank of Wild Species, Kunming Institute of Botany, Chinese Academy of Sciences, 132 Lanhei Road, Kunming, Yunnan 650201, China

Authors for correspondence:
Wei Zhou
Tel: +86 871 65223503
Email: zhouwei@mail.kib.ac.cn

Hong Wang
Tel: +86 871 65223534
Email: wanghong@mail.kib.ac.cn

Received: *11 February 2019*
Accepted: *1 May 2019*

New Phytologist (2019) **224:** 1290–1303
doi: 10.1111/nph.15905

Key words: floral evolution, genetic diversity, heterostyly, homostyly, mating systems, phylogenomics, *Primula*, selfing.

Summary

- Evolutionary transitions from outcrossing to selfing often occur in heterostylous plants. Selfing homostyles originate within distylous populations and frequently evolve to become reproductively isolated species. We investigated this process in 10 species of *Primula* section *Obconicolisteri* using phylogenomic approaches and inferred how often homostyly originated from distyly and its consequences for population genetic diversity and floral trait evolution.
- We estimated phylogenetic relationships and reconstructed character evolution using the whole plastome comprised of 76 protein-coding genes. To investigate mating patterns and genetic diversity we screened 15 microsatellite loci in 40 populations. We compared floral traits among distylous and homostylous populations to determine how phenotypically differentiated homostyles were from their distylous ancestors.
- Section *Obconicolisteri* was monophyletic and we estimated multiple independent transitions from distyly to homostyly. High selfing rates characterised homostylous populations and this was associated with reduced genetic diversity. Flower size and pollen production were reduced in homostylous populations, but pollen size was significantly larger in some homostyles than in distylous morphs.
- Repeated transitions to selfing in section *Obconicolisteri* are likely to have been fostered by the complex montane environments that species occupy. Unsatisfactory pollinator service is likely to have promoted reproductive assurance in homostyles leading to subsequent population divergence through isolation.

Introduction

The evolution of self-fertilization from obligate outcrossing enforced by self-incompatibility has occurred on numerous occasions among herbaceous angiosperms (Stebbins, 1974; Igic *et al.*, 2006). Two principal selective forces are most often invoked to explain this transition (Busch & Delph, 2012): reproductive assurance due to unsatisfactory pollinator service and/or a scarcity of mates (Darwin, 1876), or the transmission advantage associated with selfing compared with outcrossing (Fisher, 1941). The evolution of selfing from outcrossing is of particular interest to evolutionary biologists because it has biogeographical (Grossenbacher *et al.*, 2015; Moeller *et al.*, 2017), ecological (Moeller, 2006; Levin, 2012), demographic (Barrett *et al.*, 2014; Pannell, 2015), phenotypic (Goodwillie *et al.*, 2010; Sicard & Lenhard, 2011), genetic (Ness *et al.*, 2010; Wright *et al.*, 2013) and evolutionary consequences (Igic & Busch, 2013; Barrett & Harder, 2017). Of particular value for investigating this mating-system transition are lineages in which this process has occurred on numerous occasions.

High selfing rates in flowering plants are commonly associated with changes to floral morphology and sex allocation resulting in the convergent evolution of the selfing syndrome (Sicard & Lenhard, 2011), including alterations in sex-organ deployment and reduced investment in floral display and pollen production (Lloyd, 1965; Morgan & Barrett, 1989). Genome-wide polymorphism is generally diminished in selfers as high rates of selfing lower the effective population size (N_e) by up to two-fold due to the reduced effective number of alleles in populations (Charlesworth & Wright, 2001). Moreover, because of greater linkage disequilibrium in selfing populations, genetic diversity may be further reduced because of the action of linked selection and other processes in the genome, including selective sweeps and background selection which combine to reduce the efficacy of selection (Nordborg, 2000; Arunkumar *et al.*, 2015). Patterns of genetic diversity may also be influenced by the demographic

history of species with colonization–extinction cycles resulting in genetic bottlenecks (Pannell, 2015). Therefore, both genetic and demographic processes reduce genetic variation within populations of selfers in comparison with related outcrossers (Ness *et al.*, 2010).

Heterostylous taxa provide opportunities for investigating evolutionary transitions from outcrossing to selfing and their consequences. Populations with this mating polymorphism are usually composed of two (distyly) or three (tristyly) style morphs differing reciprocally in stigma and anther height and possessing self- and intramorph incompatibility (Darwin, 1877; Barrett, 1992). Heterostyly is reported from 28 angiosperm families and, in the majority, the polymorphism has broken down giving rise to homostylous populations comprised of a single, self-compatible floral form with anthers and stigmas located close together within a flower (Ganders, 1979; Weller, 1992; Barrett & Shore, 2008). Homostylous plants usually have the capacity for autonomous self-pollination and as a result often possess high selfing rates (Piper *et al.*, 1984; Ganders *et al.*, 1985), although some degree of herkogamy in homostyles can result in mixed mating (Belaoussoff & Shore, 1995; de Vos *et al.*, 2018).

Homostyles in distylous taxa usually have either long styles and long-level stamens ('long homostyles') or short styles and short-level stamens ('short homostyles'), with the former more commonly observed in nature either as intraspecific variants (Crosby, 1949; Yuan *et al.*, 2017; Zhou *et al.*, 2017), or as derived monomorphic varieties or species in otherwise distylous lineages (Schoen *et al.*, 1997; Truyens *et al.*, 2005; de Vos *et al.*, 2014). In contrast with the relatively gradual evolution of selfing via mixed mating that is characteristic of most angiosperm groups, high selfing in distylous species can arise in a single generation by genetic changes at the *S* locus linkage group governing the heterostylous syndrome (Kappel *et al.*, 2017).

Primula (Primulaceae) is the most intensively studied heterostylous taxon, with over a century of sustained attention since Darwin's classic work (Darwin, 1877). Most research has focused on European and North American taxa despite the fact that the centre of diversity resides in the Himalayas and southwest China (Richards, 2003). The vast majority (92%) of *c.* 400–500 species are distylous, with all remaining species monomorphic for stylar condition and homostylous. Homostyles are dispersed among 19 of the 38 sections and, in all but one, they co-occur with distylous species (Mast *et al.*, 2006; de Vos *et al.*, 2014). Ancestral state reconstructions of *Primula* demonstrate a single origin of distyly and multiple breakdown events to homostyly (Mast *et al.*, 2006). Moreover, intraspecific investigations of several *Primula* species have provided further evidence that homostyles are derived from heterostylous morphs (Crosby, 1949; Charlesworth & Charlesworth, 1979; Yuan *et al.*, 2017; Zhou *et al.*, 2017).

Primula section *Obconicolisteri* Balf. f. is a relatively small section comprised of *c.* 12 species, with a typical Sino-Himalayan distribution in eastern Asia. The section is of particular interest because the number of homostylous taxa is disproportionally higher than other sections of the genus, with at least five of the 12 species either exclusively homostylous or with both distylous and homostylous populations (e.g. *P. oreodoxa*; Yuan *et al.*, 2017, 2019). The high proportion of homostylous taxa in section *Obconicolisteri* could result from a single transition followed by speciation within homostylous lineages, or it could arise from multiple independent transitions or some combination of these scenarios. A major goal of this study was to investigate which of these hypotheses best explains the observed patterns of mating-system variation.

Here, we employed phylogenomic approaches to reconstruct a comprehensive phylogeny of *Primula* section *Obconicolisteri* and used the resulting trees to investigate the evolutionary history of distyly and homostyly. Our study addressed the following questions: (1) Is there evidence for multiple origins of homostyly in section *Obconicolisteri*? And if so, what types of homostyles have evolved? Specifically, are homostyles exclusively long homostyles, as predicted by theory or do short homostyles also occur, as occasionally reported in *Primula* (Charlesworth & Charlesworth, 1979)? (2) Has the transition to homostyly resulted in high selfing rates and if so is this associated with reduced genetic diversity within populations? Despite a considerable literature on the evolution of homostyly in *Primula,* few studies have estimated mating patterns using genetic markers (but see Piper *et al.*, 1984; de Vos *et al.*, 2018), and no study has examined the population genetic consequences of shifts to selfing. (3) What changes in floral morphology and sex allocation have accompanied transitions to selfing and are convergent patterns evident in floral traits? Although transitions to homostyly are commonly associated with the evolution of the selfing syndrome, de Vos *et al.* (2014) found that, in *Primula*, this was not necessarily the case and that strong morphological divergence was evident, often among related homostylous species.

Materials and Methods

Study system

Section *Obconicolisteri* Balf. f. (Primulaceae) is one of 38 sections traditionally recognised in *Primula* (Richards, 2003). It has a centre of distribution in southwest China occurring between 2000 and 3000 m (Hu & Kelso, 1996) and is comprised of *c.* 12 species (Richards, 2003) of which 10 (*P. ambita*, *P. asarifolia*, *P. barbicalyx*, *P. densa*, *P. dumicola*, *P. obconica*, *P. oreodoxa*, *P. rubifolia*, *P. sinolisteri* and *P. vilmoriniana*; Supporting Information Table S1) are endemic to China, mainly occurring in Yunnan and Sichuan provinces (Fig. S1). Distyly is common among species in the section with five exceptions: *P. dumicola* is recorded as homostylous; *P. oreodoxa* exhibits both distylous and homostylous populations; *P. sinolisteri* is comprised of two allopatric varieties, one distylous (var. *sinolisteri*) and the other homostylous (var. *aspera*); and *P. obconica* has five subspecies two of which are homostylous (*P. obconica* subsp. *fujianensis* and *P. obconica* subsp. *parva*) (Fig. 1). The 10 species of section *Obconicolisteri* endemic to southwest China are reported as diploid with chromosome numbers of $2n = 22$, except *P. sinolisteri* (Richards, 2003) and *P. oreodoxa* which are $2n = 24$, but with a single reported tetraploid homostylous population in the latter (Yuan *et al.*, 2017).

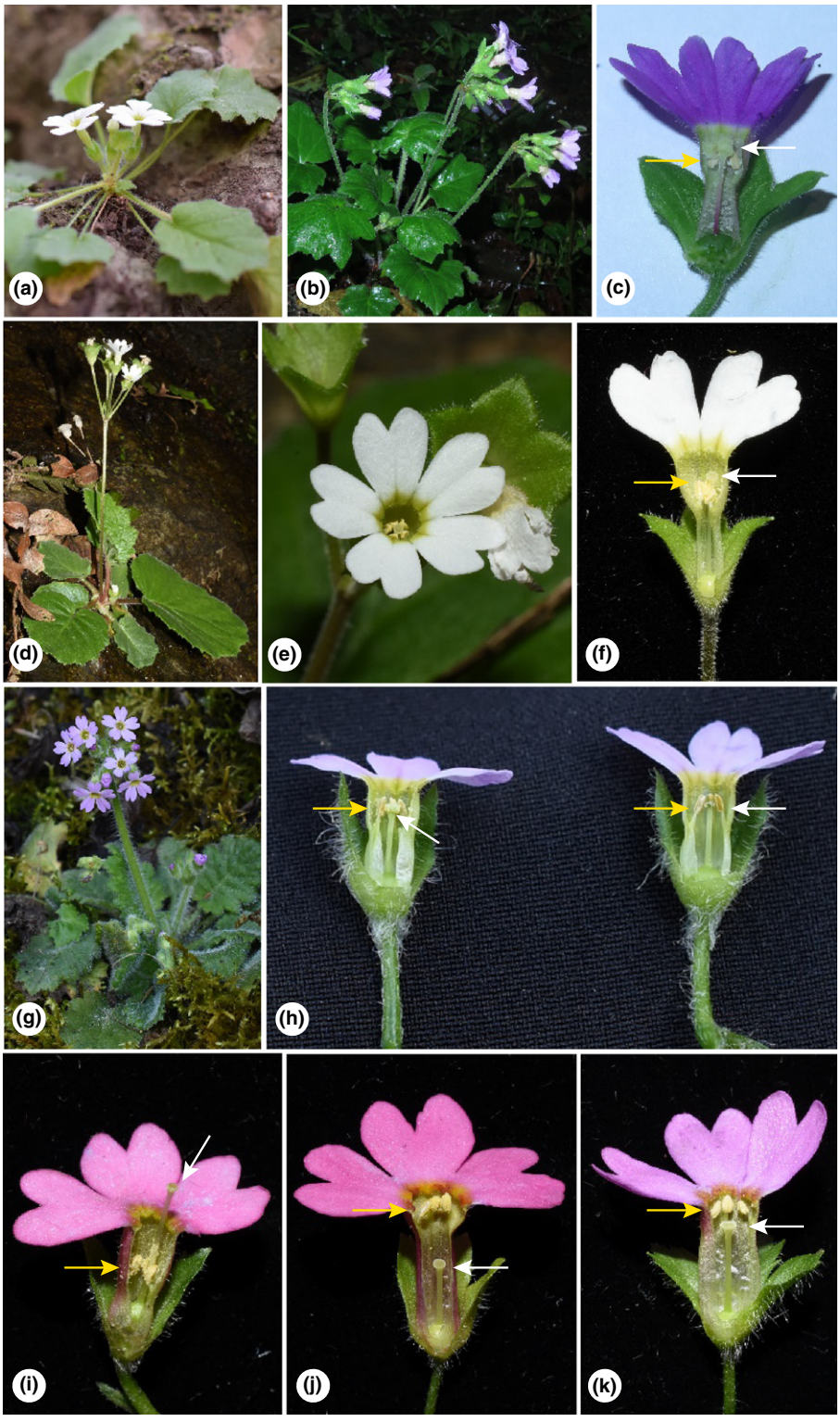

Fig. 1 Variation in floral morphology and sex-organ position in five species of *Primula* section *Obconicolisteri* all of which are either homostylous or are comprised of distylous and homostylous floral morphs: (a) homostylous *P. obconica* subsp. *parva*, (b, c) homostylous *P. sinolister* var. *aspera*, (d–f) homostylous *P. obconica* subsp. *fujianensis*, (g, h) homostylous *P. dumicola*, (i, j) long-styled and short-styled morphs of *P. oreodoxa,* (k) homostylous morph of *P. oreodoxa*. Stigma and anthers are indicated by white and yellow arrows, respectively.

Population sampling

To investigate phylogenetic relationships between distylous and homostylous taxa we collected 37 samples representing 10 of the 12 species belonging to section *Obconicolisteri* and three outgroup taxa (Table S1). To investigate population genetic diversity and variation in floral traits we sampled 40 populations including 12 distylous/homostylous species or varieties (Table S2). Of the 40 populations, 25 were distylous and 15 were homostylous. In each population, we sampled leaf tissue from 15 randomly sampled individuals that was subsequently dried in silica gel for sequencing and genotyping.

Molecular procedures

We extracted total genomic DNA from dried leaf tissue using a modified cetyl trimethyl ammonium (CTAB) protocol (Doyle, 1991). Quantification of DNA was conducted using agarose gel electrophoresis (0.8%), and we determined concentrations using a SmartSpec™ Plus Spectrophotometer (Bio-Rad, Hercules, CA, USA). The purified genomic DNA was then sheared into *c.* 500-bp fragments to construct a paired-end (PE) library according to the Nextera XT sample preparation procedures (Illumina, San Diego, CA, USA). We generated the PE reads of 90 bp or 300 bp using an Illumina HiSeq 2500 sequencer at the Beijing Genomic Institute (Shenzhen, Guangdong, China) or Illumina MiSeq instrument at the Laboratory of Molecular Biology of Germplasm Bank of Wild Species (Kunming Institute of Botany, CAS). We obtained *c.* 2 Gb of sequence data for each of 40 individuals, with two samples obtained from each taxon. We used 15 primer pairs: PBOB50, PBOB51, PBOB83, PBOB92, PBOB 97, PB29, PB40, PB46, PB62, PB120, c8899, c15305, c27661, c45817, c62889, and c80317 to assess population genetic diversity and structure at microsatellite loci for 600 individuals sampled from 40 populations following Wang *et al.* (2019) and Yuan *et al.* (2018).

Plastome assembly

We performed sequence quality trimming using the NGS QC TOOL Kit (Patel & Jain, 2012) with default parameters. We obtained *de novo* assembly of high-quality short reads using CLC Genomics Workbench v.8.5.1 (CLC Inc., Arhus, Denmark) with a k-mer of 63. Reads were mapped back to assemblies using Bowtie v.2.2.3 (Langmead & Salzberg, 2012) to verify the assembly quality. We annotated the complete plastomes using the Dogma pipeline (Wyman *et al.*, 2004), which implements Blast and a fully annotated plastome database to predict protein-coding genes, transfer RNA (tRNA), and ribosome RNA (rRNA). We confirmed the boundaries of intron/exon manually by aligning their orthologs using other published information from congeneric species. We determined transfer RNAs using tRNAscan-SE (Schattner *et al.*, 2005) and graphical maps of annotated genomes were generated by using Organellar Genome Draw (Lohse *et al.*, 2013).

Phylogenetic inference

We generated two data sets for phylogenetic analyses; a whole plastome alignment performed using Muscle (Edgar, 2004) and a concatenated data set comprised of coding sequence (CDS) of the 76 protein-coding genes in the plastome. For the CDS data set, we extracted exons of individual genes and aligned them using the program Mafft v.7.388 (Katoh & Standley, 2013) implemented in Geneious v.8.1.4 (Biomatters Ltd, Auckland, New Zealand). A custom Perl script was written to concatenate the aligned genes in the final data matrix.

We investigated the phylogenetic relationships among species with the whole plastome and CDS alignments using maximum likelihood (ML) and Bayesian inference, with *P. effusa* and *P. duclouxii* as outgroups. We used the program jModelTest2 (Darriba *et al.*, 2012) with the Akaike information criterion (AIC) (Posada & Buckley, 2004) to determine the most appropriate model of nucleotide substitution for each dataset. We conducted ML analysis using the program RAxML v.8.1.1180 (Stamatakis, 2006) with 1000 pseudoreplicates of rapid bootstrap, and used MrBayes v.3.2.1 (Ronquist & Huelsenbeck, 2003) to find the optimal tree topology and calculate Bayesian posterior probabilities for the dataset. Markov chain Monte Carlo was implemented using two independent runs and four chains, sampling every 1000 generations for 20 million generations. Chains were determined to have converged as 50% majority-rule consensus trees from two independent runs when they exhibited the same topology. Trees generated before the four Markov chains reached stationarity were removed as burn-in with the remaining trees used to construct a majority-rule consensus tree.

Reproductive character evolution

We reconstructed ancestral character states and traced the evolutionary history of distyly and homostyly in section *Obconicolisteri* using Mesquite v.3.50 (Maddison & Maddison, 2016). We coded the character states for stylar condition as categorical data, with distyly as '0' and homostyly as '1'. The dataset for this trait was generated and converted into a Nexus format, and the evolution of ancestral states was optimised onto the ML tree inferred from the whole plastome data set using the trace character history function in Mesquite by following the maximum likelihood ancestral state reconstruction method. Initially, we assumed an equal rate for losses and gains of heterostyly and conducted the analysis using a symmetrical one-parameter Markov k-state (MK1) model, in which there is equal probability of switching between discrete character states. Next, we considered a model in which the transition rates to and from homostyly (0 to 1 and 1 to 0) were treated as independent parameters. We then tested whether this two-rate (asymmetrical) model had a significantly better fit to the data than the one-rate (symmetrical) model using a likelihood ratio test with one degree of freedom. Given the complexity of the evolution of the distylous syndrome, compared with the relative ease with which it is lost (Barrett & Shore, 2008; Huu *et al.*, 2016; Li *et al.*, 2016), we examined a range of gain and loss ratios (that is rate of gain of 1 vs rates of loss of 1.5, 2, 3, 4, 5, 10, 100, 1000 and 10 000) and calculated the likelihood for each combination of rates by using an asymmetrical two-parameter Markov k-state (MK2) model. The best estimate of the character state at each node was determined by the threshold of an absolute difference of 2.0 units in the log likelihoods for each state. If the log likelihoods of two states differed by this threshold value or more, the state with the lower likelihood was rejected, and if the difference was smaller (that is < 2.0), the reconstruction was considered ambiguous. We used the option of 'Root State Frequencies Equal' for the treatment of the root in the asymmetrical MK2 model. Mesquite treats all terminal taxa as separate species and multiple samples with the same character state may

influence ancestral reconstructions, we therefore used one accession for each species or variety. However, we retained homostylous accessions of diploid and tetraploid lineages of *P. oreodoxa* because they did not form a monophyletic species.

Selfing rates, genetic diversity and population structure

We used microsatellite data to estimate population-level selfing rates with the software RMES. This method is relatively insensitive to null alleles and scoring errors, which otherwise may lead to overestimation of selfing rates (see David *et al.*, 2007). RMES implements two methods for estimating selfing rates: two-locus heterozygosity disequilibrium values (g_2, with associated selfing rate $s(g_2)$) and the other maximises the log-likelihood of the multilocus heterozygosity structure of samples (s(ML)). To evaluate the robustness of RMES estimates of selfing rates, we obtained significance tests for $s(g_2)$ and s(ML) from 10 000 iterations by randomly re-assorting single-locus heterozygosities among individuals, with a maximum of 10 generations of selfing allowed for a heterozygote, with a precision for 0.00001 log-likelihood in each case. 95% confidence intervals (CI) were provided by RMES for both methods (David *et al.*, 2007). We estimated allele frequencies at microsatellite loci, as well as the mean number of alleles per locus (N_A), observed and expected heterozygosity (H_O, H_E), and effective number of alleles (N_E) using the program FSTAT v.2.9.3 (Goudet, 1995).

To obtain evidence for the multiple origin of homostylous lineages and determine relationships between monomorphic and dimorphic populations, we investigated patterns of differentiation among populations with different morph structures using three approaches. First, we used the original genotypic data to calculate a Euclidean distance matrix between all samples and created a visual representation of genetic relationships among individuals using principal coordinate analysis (PCoA) in the program MVSP v.1.3 (Kovach, 1999). Second, we calculated Nei's (1987) unbiased genetic distance (D) among all possible pairs of individuals from allele frequencies estimated in the program MSA (Dieringer & Schlötterer, 2003). We then constructed a consensus neighbour-joining (NJ) tree based on pairwise estimates of genetic distance, using 1000 bootstrap trees and random input order in Phylogeny Inference Package (PHYLIP) v.3.63 (Felsenstein, 2005). Finally, we inferred the genetic structure of population using the program INSTRUCT (Gao *et al.*, 2007), which uses a model similar to STRUCTURE (Pritchard *et al.*, 2000), but does not enforce random mating therefore making it suitable for the study of populations with inbreeding (homostyly) or disassortative mating (distyly).

Variation in floral traits: herkogamy, flower size and pollen production and size

To measure herkogamy among distylous and homostylous taxa, we obtained a random sample of 967 flowers (L-morph: $n = 219$; S-morph: $n = 226$; H-morph: $n = 522$) from 31 populations during peak flowering of each species from 2016 to 2018 (Fig. S2). We sampled an average of 31.2 flowers per population (range 24–40) at peak anthesis. We classified flowers according to floral morph before they were preserved in 70% ethanol for subsequent measurement. For all sampled flowers, we measured stigma–anther separation, stigma and anther height and the length of the style from the base of the ovary using digital callipers to the nearest 0.01 mm. We then estimated the normal functions distribution of stigma and anther height by using a ML-based parameter fitting method in R v.3.41 (R Development Core Team, 2014) following methods detailed in Zhou *et al.* (2017).

To determine whether additional reproductive traits differed between distylous and homostylous populations, we compared floral syndromes represented by measurements of floral-tube length, corolla diameter, petal biomass, sepal length and sepal width in the samples discussed above. We obtained dry weights (DWs) for petal biomass by drying samples at 55°C for 2 d and used these measurements as an indirect estimate of flower size. We also collected four to six flower buds from each floral morph per population for pollen and ovule counting. Pollen grains were counted using a haemocytometer and further confirmed on a Micromeritics Elzone II 5390-Partical Size Analyser (Micromeritics Instrument Corp., Norcross, GA, USA). To count ovule number per flower, pistils fixed in ethanol : acetic acid (9 : 1) were cut at the receptacle and placed onto a microscope slide under a dissecting microscope. We explored differences in floral traits between distylous and homostylous populations using linear mixed models as implemented in the *lme* function in the NLME package in R v.3.41, where stylar condition (distyly vs homostyly) was treated as a fixed effect and populations nested within species were treated as random effects.

To confirm that homostyles were long homostyles and not L-morph plants with modified 'short-level' stamens or short homostyles we used pollen size to discriminate between these alternative hypotheses (see Zhou *et al.*, 2017). The pollen size of long homostyles should be significantly larger than pollen from the L-morph because stamens of this morph are of a fundamentally different origin than short-level anthers of the L-morph (Barrett & Shore, 2008). To prepare pollen grains for a scanning electron microscope (SEM), acetolyzed grains were suspended in 95% ethanol for dehydration, then sputter-coated with gold–palladium on specimen stubs. We measured the polar axis of pollen grains from 14 populations of nine species (L-morph: $n = 422$; S-morph: $n = 458$; putative H-morph: $n = 537$) under an SEM (S-4800; Hitachi, Tokyo, Japan) to the nearest 0.01 μm. Pollen size differentiation among the floral morphs was determined by one-way ANOVA using *post hoc* Tukey tests to identify statistically significant differences between morphs.

Results

Phylogenetic inferences on mating-system evolution

We obtained 40 newly generated complete plastid genomes of *Primula* (Table S1); these displayed the typical size, structure, gene content and gene order expected for plastid genomes (Fig. S3). The whole plastome data set had an aligned length of 158, 324 bp with 7638 parsimony informative characters (PICs, 4.82%). By removing

the intergenic regions and a few introns, the CDS data set had an aligned length of 66, 292 nucleotides with a reduced number of PICs of 2862 (4.31%). Genome sequences are deposited in GenBank with accession nos. MK344725–MK344764.

We reconstructed four separate phylogenetic trees based on ML and Bayesian analyses of the whole plastome and CDS data set. The tree topologies obtained from ML and Bayesian analyses were identical for each data set and the different data sets produced largely congruent topologies (Fig. 2a). Section *Obconicolisteri* was strongly supported as monophyletic (ML bootstrap value: ML = 100%; Bayesian posterior: PP = 1.00) and two main clades were recovered both with very strong nodal support (Fig. 2a). The clade *oreo-dumi* included samples from distylous *P. densa* that clustered with homostylous *P. dumicola* and distylous–homostylous *P. oreodoxa*. The remaining samples formed another well supported clade *obco-sino*, but the monophyly of each species was not supported, as the varieties of *P. obconica* and *P. sinolisteri* were nested within this clade in a dispersed manner (Fig. 2a).

Likelihood-based ancestral character state reconstructions are depicted in Fig. S4. The asymmetrical two-rate model ($-\log_e L = 31.95$) had a significantly better fit to the data than the symmetrical one-rate model ($-\log_e L = 34.92$), based on a likelihood ratio test (statistic = 5.94, df = 1, $P < 0.05$). Under asymmetrical models, the calculated likelihoods varied among different combinations of rates and the model with the highest log-likelihood value was a 3 : 1 weighting (gain rate × 3 = loss rate), but this weighting was not significantly better than models with 4 : 1, 5 : 1 and 10 : 1 weighting (Fig. S4a). A similar pattern was also obtained when gains and losses of heterostyly were allowed to evolve under a range of rates (Fig. S4). The most recent common ancestor of section *Obconicolisteri* was unequivocally inferred to be distylous. Under the asymmetrical models, three independent origins of homostyly (*P. obconica* subsp. *parva, P. obconica* subsp. *fujianensis, P. sinolisteri* var. *aspera*) were inferred within the *obco-sino* clade for all combination of rates. However, it is equivocal whether homostyly has arisen once or twice within the clade *oreo-dumi* depending on the degree of model asymmetry (for example two independent origins of homostyly were inferred when the loss/gain rate was 10 : 1 or higher) (Fig. S4b). Regardless of which weighting scheme is used, we infer that distyly has broken down to homostyly on multiple (at least four) occasions within the section.

Selfing rates, genetic diversity and population structure

Selfing rate estimates were substantially higher in all homostylous populations compared with distylous populations, with values for the two methods for estimating mating patterns yielding similar results (Fig. 3a; Table S3). Selfing rates were strongly associated with differences in the degree of herkogamy between distylous and homostylous populations (Fig. 3a; Table S3).

The total number of alleles per locus in our sample of 600 individuals of 10 *Primula* species ranged from 18 (locus: PB46) to 61 (locus: c15305), with 2528 alleles distributed among the 15 microsatellite loci. Departures from Hardy–Weinberg equilibrium varied among loci as a result of an excess number of homozygotes. Genetic diversity estimates varied among species (Table S3) but overall levels of genetic diversity were significantly lower in homostylous than distylous population (Fig. 3b,c).

Two-dimensional PCoA of SSR phenotypes (Fig. 4) separated all individuals of the clade *oreo-dumi* (*P. oreodoxa* and *P. dumicola*) from the remaining clade along the first axis, with PCoA1 explaining 35.56% of the total variation. The homostylous individuals of the clade *obco-sino* (*P. obconica* subsp. *parva, P. obconica* subsp. *fujianensis* and *P. sinolisteri* var. *aspera*) were partially separated from the distylous samples along PCoA2 (12.04%), but this pattern was not evident between distylous and homostylous morphs in the clade *oreo-dumi*. The individual-based NJ tree (Fig. 5) provided support for multiple origins of homostyles in both clades. Analyses in InStruct recovered 17 population clusters. The log-likelihood of these reconstructions saturated at $K = 17$, with a local peak in the ΔK value. Homostylous taxa were strongly supported as five differentiated clusters (Fig. S5) consistent with the hypothesis of multiple origins of homostyly.

Differences in floral traits between mating systems

In the five distylous species, the heights of stigmas and anthers in the L- and S-morphs exhibited a clear bimodal distribution, whereas these organs had a unimodal distribution in homostylous populations (Fig. 2b). In each distylous species, there was near perfect spatial matching between corresponding compatible stigma and anther heights of the floral morphs, and no significant overlap between incompatible stigmas and anthers (Fig. 2b). The only exception was *P. obconica* subsp. *obconica* in which S-morph stigmas were significantly higher ($\mu = 4.20$, $\sigma = 0.30$) than L-morph anthers ($\mu = 3.19$, $\sigma = 0.54$). As expected, stigma–anther separation in homostyles was significantly reduced in comparison with the distylous morphs (mean = 1.26 mm, SD = 1.530; $F = 1543$, $P < 0.001$). All homostylous species exhibited weak approach herkogamy, except *P. obconica* subsp. *parva* in which stigmas were at the same level as anthers (Fig. 2b).

Floral traits in distylous populations were significantly larger in size compared with homostylous populations. This pattern was evident for floral-tube length ($F_{2,842} = 260.8$, $P < 0.001$; Fig. 6a), corolla diameter ($F_{2,939} = 502.3$, $P < 0.001$; Fig. 6b), and petal biomass ($F_{2,836} = 306.3$, $P < 0.001$; Fig. 6c). Homostyles had significantly shorter and narrower sepals compared with distylous morphs (sepal length: $F_{2,950} = 25.56$, $P < 0.001$; sepal width: $F_{2,950} = 77.55$, $P < 0.001$; Fig. 6g,h). The ovule number of homostyles was similar to the L-morph ($P = 0.799$), but significantly different from the S-morph ($F_{2,698} = 13.39$, $P < 0.001$) (Fig. 6d). On average, homostyles had a 66.8% reduction in pollen number per flower compared with distylous morphs ($F_{2,88} = 16.53$, $P < 0.001$; Fig. 6e) and this was associated with significantly lower P : O ratios for homostyles compared with distylous morphs ($F_{2,88} = 11.27$, $P < 0.001$; Fig. 6f).

There was a significant difference in the mean pollen size between the L- and S-morphs in all five distylous species (Fig. 7). Pollen grains of the S-morph (mean = 13.375 μm, SD = 0.998) were on average *c.* 1.36 times larger than the L-morph (mean = 9.786 μm, SD = 0.993) with respect to their polar axes.

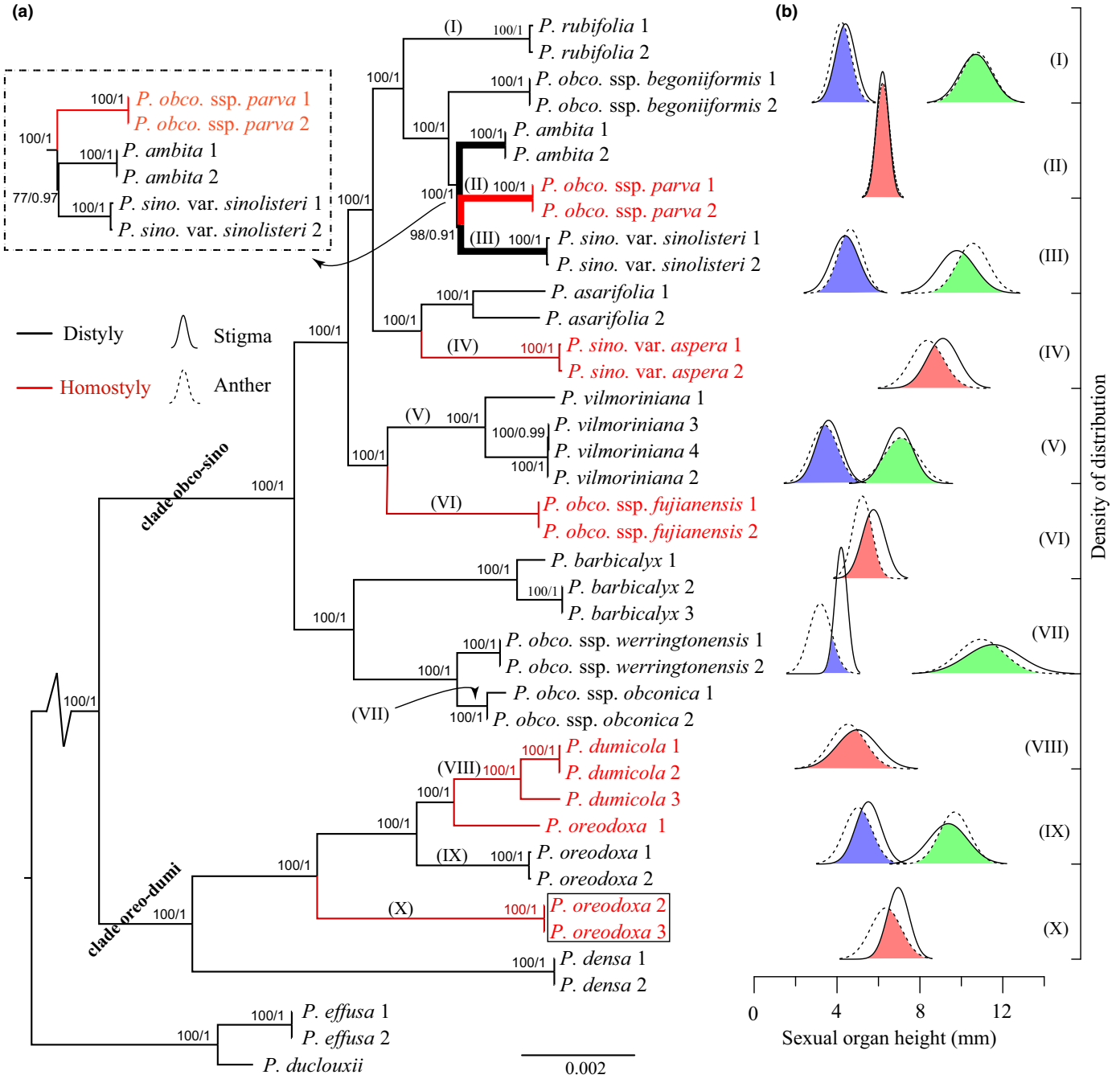

Fig. 2 Phylogenetic tree of taxa in *Primula* section *Obconicolisteri* and the distribution of sex-organ heights in distylous and homostylous species. (a) The maximum likelihood (ML) tree inferred from the exons of protein-coding genes. The thicker branches indicate the position in which the topology was different from the ML tree inferred from the whole plastome data set, as shown in the left-hand dotted line box. Colour codes: black, distylous species; red, homostylous species; boxed taxa, tetraploid species. Numbers on the left side of branches indicate posterior probabilities from the Bayesian analysis; numbers on the right indicate bootstrap values (> 90%). (b) Distributions and overlap of sex organs in five distylous species and five homostylous species from section *Obconicolisteri*. Solid and dotted lines indicate density distribution of stigmas and anthers, respectively. Coloured areas indicate the proportion of overlap between female and male sex organs: the blue and green areas indicate overlap between L-morph anthers and S-morph stigmas, and S-morph anthers and L-morph stigmas, respectively, in distylous species and the pink area indicates the overlap of stigmas and anthers in homostylous species. Species coded in Roman numerals are identified in the left phylogenetic tree branch.

The mean pollen size of homostyles exhibited wide variation among species ranging from 12.873 μm (SD = 0.970) in *P. sinolisteri* var. *aspera* to 19.895 μm (SD = 1.315) in *P. obconica* subsp. *fujianensis*. Interestingly, in three of the five homostylous taxa there was a significant increase in pollen size with an overall average increase of 16.375 μm (SD = 2.505), which was significantly larger than pollen of the L- and S-morphs of distylous taxa ($F_{2,1414} = 1713$, $P < 0.001$) (Fig. 7).

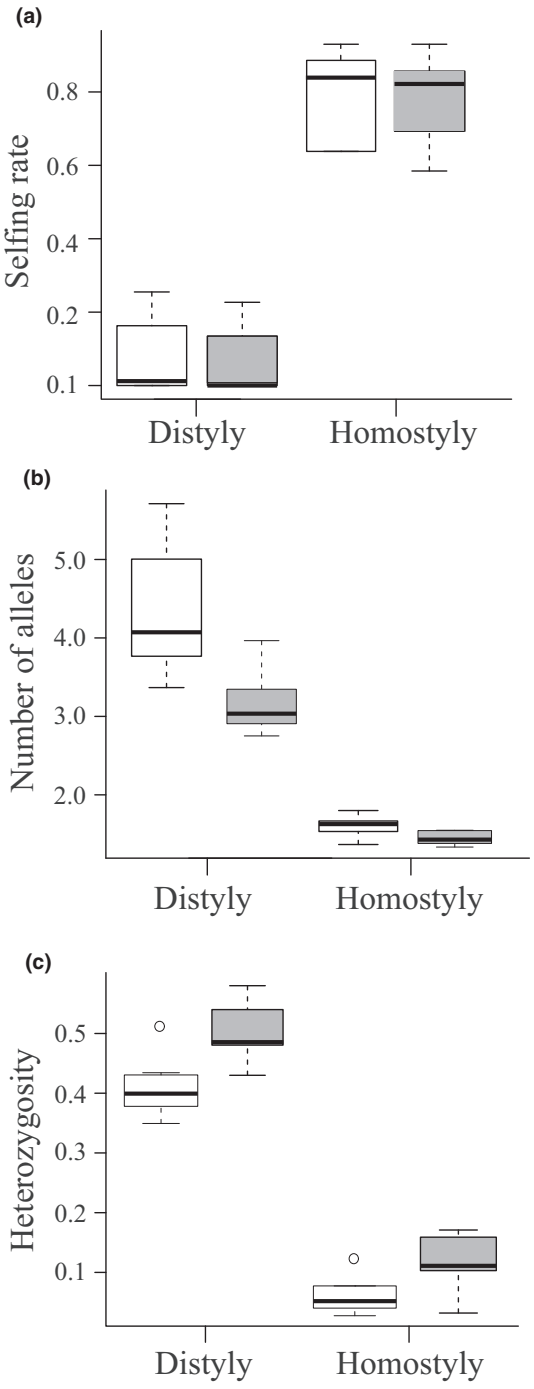

Fig. 3 Box plots illustrating the grand mean and variance of maternal population-level selfing rates and measures of genetic diversity in distylous and homostylous populations of distylous and homostylous taxa of *Primula* section *Obconicolisteri*. (a) Selfing rate: open box indicates *s*(g2) and closed box *s*(ML); (b) allele number: open box indicates number of alleles (N_A) and closed box effective number of alleles (N_E); (c) heterozygosity: open box indicates observed heterozygosity (H_O) and closed box expected heterozygosity (H_E). Boxplot legend: upper horizontal line of box 75th percentile, lower horizontal line of box 25th percentile, horizontal bar within box median value and vertical dotted line minimum–maximum value. Outliers are plotted as open circles.

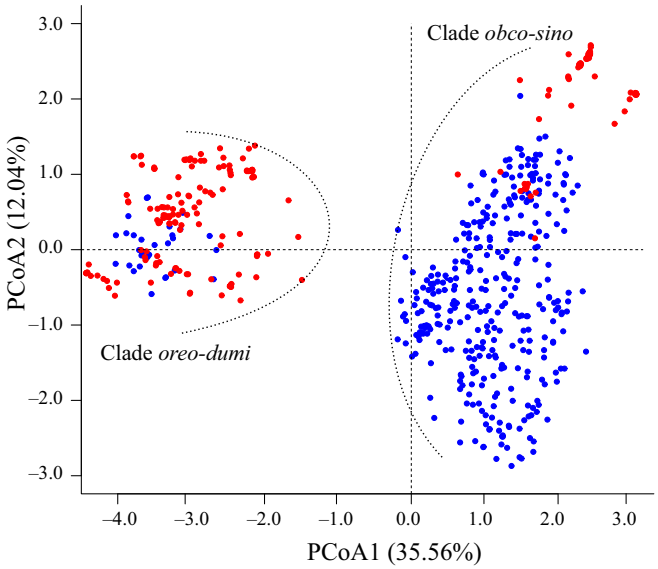

Fig. 4 Principal coordinates analysis (PCoA) of the SSR genotypes of 600 individuals sampled from 40 populations of 12 taxa of *Primula* section *Obconicolisteri*. Blue dots, distylous plants, red dots, homostylous plants.

Discussion

Using phylogenomic analyses, we have demonstrated that the evolutionary breakdown of distyly in *Primula* section *Obconicolisteri* has occurred on multiple occasions. Our character reconstructions estimated four to five independent transitions among the 10 species, depending on the models of character evolution we investigated. These numbers represent minimum estimates as more than one transition can occur within species, as has been shown for one of the species in the section (Yuan *et al.*, 2017). The shifts from distyly to homostyly are associated with significant changes to floral traits with profound consequences for the mating biology of populations; distylous populations were largely outcrossing whereas homostylous populations were characterised by high selfing rates. Our estimates of genetic diversity in populations reflected these differences; homostylous populations exhibited substantially lower levels of allelic richness and heterozygosity than distylous populations. Below we consider why the pathway from outcrossing to selfing has been followed frequently in section *Obconicolisteri* and evaluate historical factors and ecological mechanisms that have likely played a role.

Multiple transitions to homostyly and selfing

The distylous genetic polymorphism is susceptible to evolutionary breakdown, most often resulting in transitions to homostyly and selfing. Homostyles have originated on numerous occasions among virtually all distylous families (Barrett & Shore, 1987; Schoen *et al.*, 1997; Mast *et al.*, 2006; Wu *et al.*, 2017). Several intrinsic features of distyly may account for this recurrent pattern involving loss of floral polymorphism. First, in comparison with homomorphic self-incompatibility the occurrence of only two

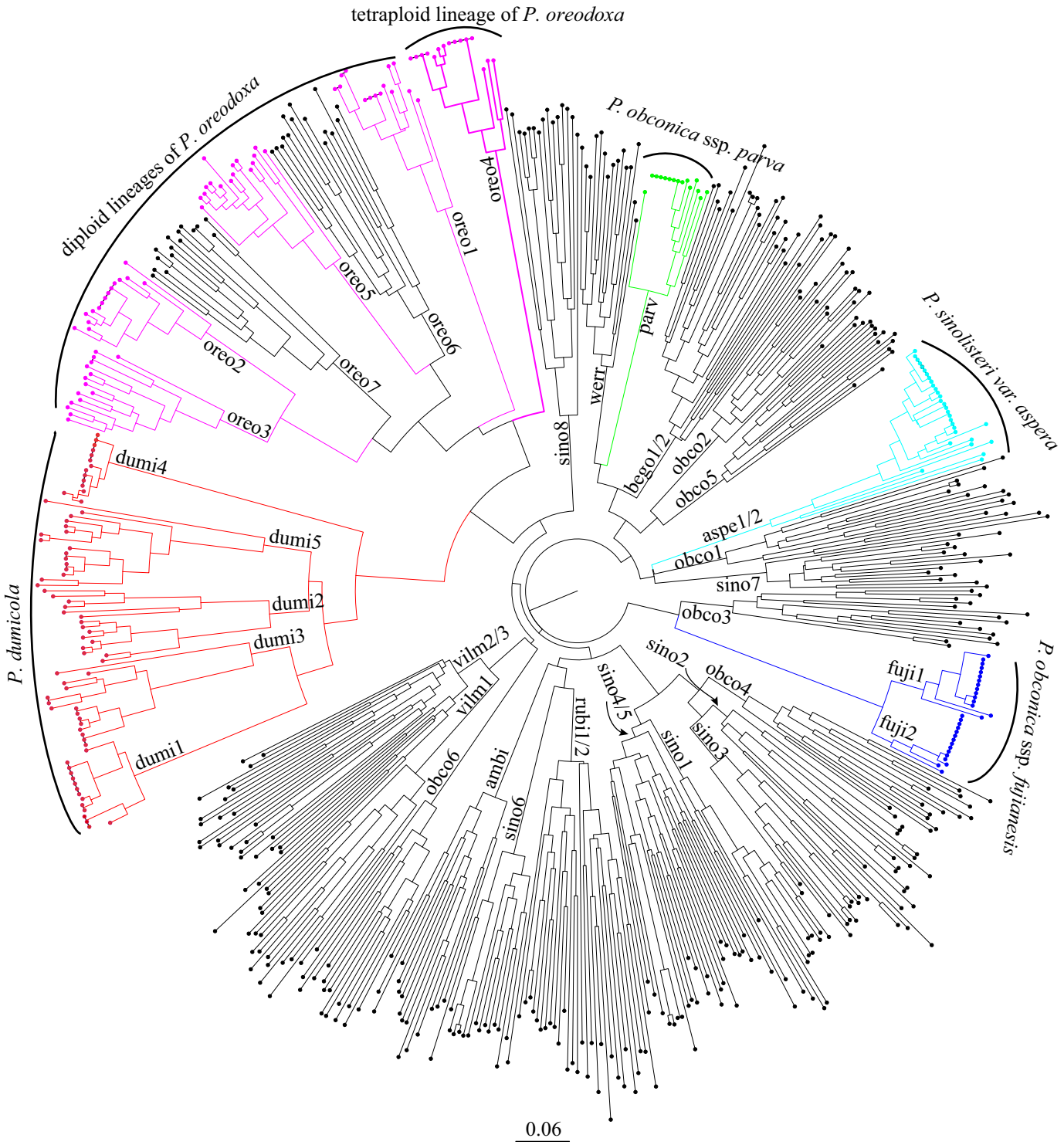

Fig. 5 Neighbour-joining tree illustrating the genetic relationships among 600 individuals from 40 populations of 12 taxa of *Primula* section *Obconicolisteri* based on Nei's (1987) unbiased genetic distance calculated from SSR data. Coloured circles at the terminal ends of branches indicate homostylous individuals and black circles distylous individuals.

mating types in populations may often act as a constraint to successful reproduction favouring the spread of selfing mutations (Lloyd & Webb, 1992). Second, the relatively high frequencies with which homostylous forms arise by mutation and/or possibly recombination may enable appropriate phenotypes to take advantage of ecological conditions hostile to the maintenance of heterostyly. For example, the occurrence of homostyles in *Primula* (at frequencies above 10^{-3}) appears to be higher than the known per locus mutation rates of 10^{-4} to 10^{-7} (Lewis & Jones, 1992; Kappel *et al.*, 2017). Third, the occurrence of both

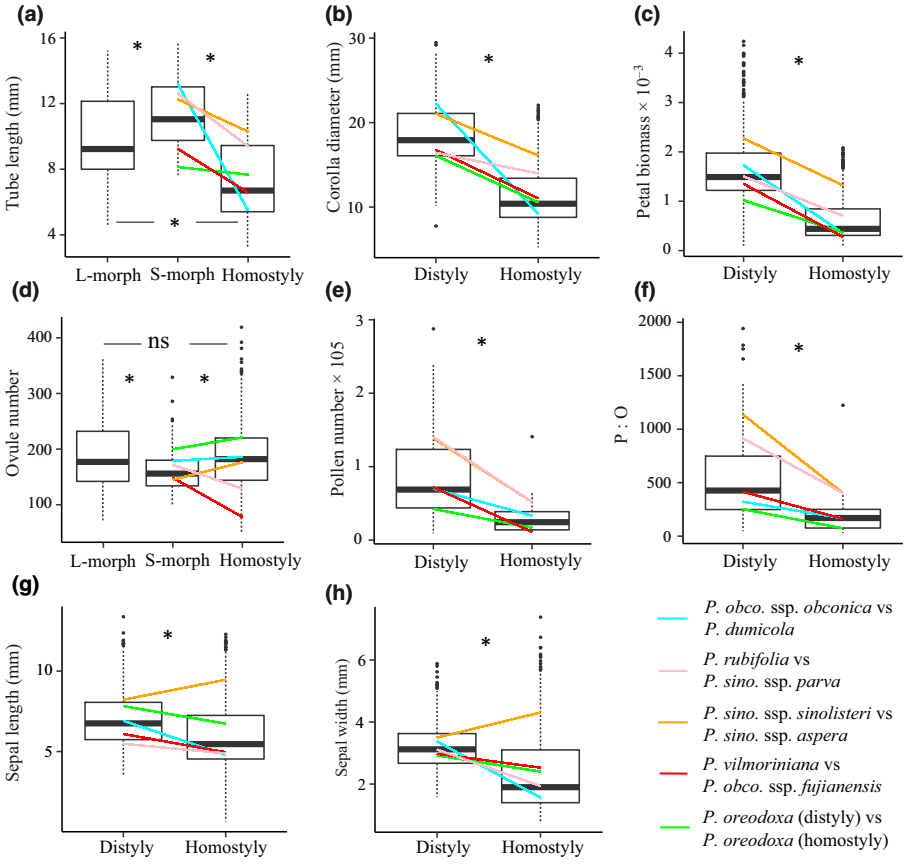

Fig. 6 Box plots illustrating the grand mean and variance of floral traits in distylous and homostylous taxa of *Primula* section *Obconicolisteri*. The data are presented for the three morphs when trait values were significantly different between the L- and S-morphs, otherwise data were pooled for the distylous morphs. (a) Floral-tube length; (b) corolla diameter; (c) petal biomass; (d) ovule number; (e) pollen number; (f) pollen : ovule (P : O) ratio; (g) sepal length; and (h) sepal width. Boxplot legend: upper horizontal line of box 75th percentile, lower horizontal line of box 25th percentile, horizontal bar within box median value and vertical dotted line minimum-maximum value. Outliers are plotted as points. Coloured lines indicate predicted slopes for the means of distylous and homostylous species pairs. *, $P < 0.01$; ns, not significant.

partial heteromorphic incompatibility and complete self-compatibility in some distylous groups (Barrett & Cruzan, 1994), including *Primula* (Wedderburn & Richards, 1990), may often allow selfing and initiate early stages in the breakdown process (Yuan *et al.*, 2017; Zhou *et al.*, 2017). Finally, the necessity for long-tongued pollinators that can mediate segregated pollen dispersal and disassortative mating may, in some environments, be too demanding and favour homostyles because they are capable of autonomous self-pollination.

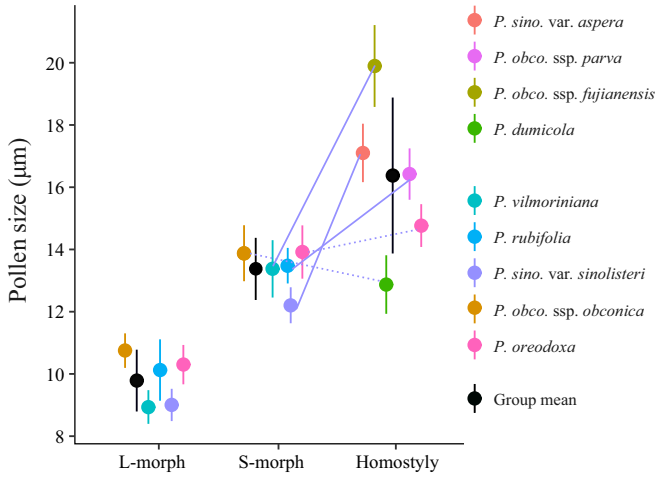

Fig. 7 Pollen size variation among distylous and homostylous floral morphs of nine taxa of *Primula* section *Obconicolisteri*. Dots indicate the mean and bars indicate ± SD. Measurements involve the polar axis of pollen grains. Coloured lines join ancestor–descendant relationships based on the phylogenetic evidence.

The relative high frequency of transitions to selfing in section *Obconicolisteri* may be associated with the biogeography of this section. Its centre of diversity is in the Hengduan mountains of China adjacent to the Qinghai-Tibetan Plateau. This region is characterised by tectonic uplift driven biotic diversification (Favre *et al.*, 2015), and the flora of the Hengduan mountains has been assembled disproportionately by *in situ* diversification (Xing & Ree, 2017), with many endemic taxa, including all members of section *Obconicolisteri*. Changes to topography and climate associated with historical orogenesis can cause fragmentation of species distributions and limit gene flow between isolated populations therefore initiating allopatric divergence and speciation. Indeed, these processes have been invoked to explain the patterns of interspecific divergence of *Primula* section *Armerina* in this region (Ren *et al.*, 2015), and also in *Primula* section *Reinii* in Japan (Yamamoto *et al.*, 2017) and may have also influenced differentiation in section *Obconicolisteri*. The reproductive isolation accompanying transitions to selfing and the ability of homostyles to establish geographically isolated populations following dispersal is likely to have fostered evolutionary divergence from their distylous ancestors leading to speciation, as originally proposed

by Baker (1961). Significantly, *Primula* exhibits particularly high levels of species diversity in the Qinghai-Tibetan Plateau and adjacent mountains with *c.* 60% of all species in the genus occurring in this region. The floral biology of most of these species has not been investigated thoroughly and many more transitions to homostyly probably await to be discovered.

Although historical orogenesis and associated habitat change may have played an important historical role in driving divergence among species, the specific ecological mechanisms favouring homostyles remain to be determined. The most plausible hypothesis is that deterioration of the pollination environment favoured selfing homostyles through reproductive assurance. This hypothesis is the most widely invoked mechanism to account for transitions from outcrossing to selfing and there is biogeographical and experimental evidence indicating an advantage to selfers when pollinator service limits outcross pollen delivery (Lloyd, 1980; Moeller, 2006; Busch & Delph, 2012). In *Primula*, homostylous species are generally favoured over distylous species in alpine and arctic environments where pollinator service often limits seed production due to harsh conditions (Richards, 2003; Guggisberg *et al.*, 2006; Carlson *et al.*, 2008; de Vos *et al.*, 2012). Moreover, a recent intraspecific investigation of *P. oreodoxa* of section *Obconicolisteri* provides support for the reproductive assurance hypothesis (Yuan *et al.*, 2017). This self-compatible species is comprised of both distylous and homostylous populations (Fig. 1i–k). Observations of pollinators indicated that visitation by long-tongued pollinators capable of mediating disassortative mating declined with elevation, and overall levels of visitation were significantly lower at sites occupied by homostylous compared with distylous populations.

Recent molecular genetic studies in *P. vulgaris* and *P. veris* have begun to identify and characterise the causal distylous genes comprising the *S*-locus supergene governing the polymorphism (reviewed in Kappel *et al.*, 2017). Recent studies indicate that naturally occurring long and short homostyles have arisen by mutation of the genes CYP^T and GLO^T, respectively (Huu *et al.*, 2016; Li *et al.*, 2016; Burrows & McCubbin, 2017), rather than through recombination as previously thought (Dowrick, 1956; Barrett & Shore, 2008). In section *Obconicolisteri* all homostyles have long styles and long-level anthers (and see Yuan *et al.*, 2019) and pollen is of similar size to long-level anthers of the S-morph, demonstrating that they are long homostyles, a pattern consistent with theoretical models (Charlesworth & Charlesworth, 1979). These authors showed that the preferential spread and fixation of long homostyles can be explained by the dominance and compatibility relations of distylous and homostylous morphs resulting in long homostyles having a transmission advantage through male function compared with short homostyles, despite the fact that both forms should originate at similar frequencies. If the origin of long homostyles in section *Obconicolisteri* results from the same genetic mechanism as long homostyles of *P. vulgaris* and *P. veris* then these phenotypes should have arisen by a mutation at the style-length gene (CYP^T) and are therefore S-morph plants with elongated styles. This hypothesis could be tested by future molecular investigations.

Reproductive and genetic consequences of transitions to selfing

An earlier comparative study of floral traits in distylous and homostylous species of *Primula* revealed complex patterns of phenotypic evolution in homostyles that were not in accord with more conventional expectations of a deterministic path to the selfing syndrome (de Vos *et al.*, 2014). These authors found that, in some instances, flower size increased in homostyles compared with their outcrossing distylous relatives. Similar patterns were also previously reported in the *Turnera ulmifolia* complex (Barrett & Shore, 1987). de Vos *et al.* (2014) suggested that due to the reduced effective population size of homostyles, phenotypic evolution was more likely to be affected by genetic drift and relaxed selective constraints than directional selection towards a selfing syndrome. As a result, they suggested that, after the transition to homostyly, diverse trajectories of trait evolution could have occurred in *Primula,* with a reduction in flower size only one of several possible outcomes.

The results of our comparison of distylous and homostylous taxa revealed a more predictable pattern of floral evolution. Although there was considerable variation among species in overall trait values, all homostyles exhibited floral traits consistent with being at different stages of the evolution of the selfing syndrome. In all five instances, homostyles had smaller flowers with reduced herkogamy and produced less pollen than distylous flowers. Some of the observed variation may have also resulted from neutral processes, relaxed selective constraints and the history of selfing in each lineage. The occurrence of a small degree of herkogamy in homostyles may promote some outcrossing (and see de Vos *et al.*, 2018) and explain why floral evolution has not proceeded to the extent it has in some highly autogamous species of flowering plants with very small flowers and much lower pollen : ovule (P : O) ratios ratios.

Consistent with sex allocation theory, homostylous populations had lower P : O ratios than related distylous populations, due to a much reduced investment in pollen production (Fig. 6e, f). Interestingly, in three of the five homostylous taxa, lower pollen production was associated with an increase in pollen size (Fig. 7). This result was unexpected as reduced pollen size might have been predicted because in all transitions homostyles exhibited smaller flowers (Fig. 6a–c) and shorter styles in comparison with their distylous relatives (mean style length of L-morph = 9.67 mm; mean style length of H-morph = 6.6 mm). Elsewhere, increased pollen size was also evident in two independently derived lineages of long homostyles in *P. chungensis*, a species with both distylous and homostylous populations (Zhou *et al.*, 2017). Among four European species of *Primula*, pollen volume was substantially higher in each of the three homostylous species in comparison with a single distylous species, although this may be because of the polyploid status of homostyles (Mazer & Hultgård, 1993).

It is difficult to envision an adaptive hypothesis to explain the larger pollen sizes of homostyles in section *Obconicolisteri*. The most probable explanation is that a trade-off occurs between pollen size and pollen production and that allocation to pollen

number has decreased more rapidly than pollen size. Pollen size–number trade-offs have been demonstrated in comparative studies (Vonhof & Harder, 1995), including in *Primula* (Mazer & Hultgård, 1993), and also directly through artificial selection experiments (Sarkissian & Harder, 2001). A variety of factors could explain the variation among homostyles in pollen size and its relation to pollen number, including the relative age of homostyles, relaxed selective constraints associated with selfing, and the strength of genetic correlations between these pollen traits. Comparisons of floral traits in outcrossers and selfers have neglected consideration of pollen size despite numerous studies of pollen production. Future work on the influence of mating-system change on pollen size evolution to determine if general patterns are evident would be worthwhile.

In each of the transitions from distyly to homostyly we found significant reductions in allelic diversity and heterozygosity in populations (Fig. 3b,c). These patterns are consistent with other studies on molecular diversity in related outcrossing and selfing populations, indicating a loss of variation in populations with high selfing rates (Ness *et al.*, 2010; Pettengill *et al.*, 2016). Demographic factors also play an important role in reducing diversity in selfing populations, due to genetic bottlenecks and reduced effective population size and, at this stage, our existing data do not enable us to partition out the relative influence of mating system and demography on patterns of genetic diversity.

Finally, what is the likely evolutionary fate of selfing homostyles in *Primula* section *Obconicolisteri*? The success of homostyles following their origin will be determined by their capacity for local adaptation and geographical spread. Comparisons of the overall range sizes of distylous and homostylous taxa in section *Obconicolisteri* do not indicate substantial differences; most of these endemic species regardless of mating system have relatively restricted distributions (Fig. S1). This contrasts with other lineages of *Primula* in which polyploid homostyles have expanded their ranges northward into previously glaciated areas in comparison with their diploid distylous ancestors and, as a result, are more frequent in arctic regions (Guggisberg *et al.*, 2006). More generally, comparative surveys of flowering plants indicate that selfing species tend to have larger geographical ranges than their outcrossing sister species (Grossenbacher *et al.*, 2015). Inbreeding depression associated with rapid transitions to selfing, limited genetic diversity in populations, and the primarily diploid status of most members of *Primula* section *Obconicolisteri* may limit genetic and ecological opportunities for future geographical spread of homostylous taxa.

Acknowledgements

We thank Hai-Fei Yan, Shuai Yuan and Qing-Jun Li for providing samples, Lawrence Harder and Susan Mazer for valuable discussions on pollen size evolution and Mark Rausher for editorial advice on character reconstructions. This work was supported by the Strategic Priority Research Program of the Chinese Academy of Sciences (XDB31000000), the National Natural Science Foundation of China (31570384, 31770417, U1502261), the Yunling Scholarship of Yunnan Province (YLXL20170001), Light of West China Program of the Chinese Academy of Sciences, and a Discovery Grant from the Natural Sciences and Engineering Research Council of Canada.

Author contributions

WZ, SCHB, HW and D-ZL planned and designed the research. LZ, WZ, X-JW, Z-KW and H-YS performed experiments, conducted field work, and WZ analysed the data. WZ and SCHB wrote the manuscript.

ORCID

Spencer C. H. Barrett https://orcid.org/0000-0002-7762-3455
Li Zhong https://orcid.org/0000-0001-7776-5042
Wei Zhou https://orcid.org/0000-0001-5527-3776

References

Arunkumar R, Ness RW, Wright SI, Barrett SCH. 2015. The evolution of selfing is accompanied by reduced efficacy of selection and purging of deleterious mutations. *Genetics* **199**: 817–829.

Baker HG. 1961. Rapid speciation in relation to changes in the breeding systems of plants. In: *Recent advances in botany. Section 9. Proceedings of the International Botanical Congress (IX), Montreal*. Toronto, ON, Canada: University of Toronto Press, 881–885.

Barrett SCH. 1992. *Evolution and function of heterostyly*. Berlin, Germany: Springer-Verlag.

Barrett SCH, Arunkumar R, Wright SI. 2014. The demography and population genomics of evolutionary transitions to self-fertilization in plants. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences* **369**: 20130344.

Barrett SCH, Cruzan MB. 1994. Incompatibility in heterostylous plants. In: Williams EG, Knox RB, Clarke AE, eds. *Genetic control of self-incompatibility & reproductive development*. Dordrecht, the Netherlands: Kluwer Academic, 189-219.

Barrett SCH, Harder LD. 2017. The ecology of plant mating and its evolutionary consequences in seed plants. *Annual Review of Ecology, Evolution, and Systematics* **48**: 135–157.

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

Barrett SCH, Shore JS. 2008. New insights on heterostyly: comparative biology, ecology and genetics. In: Franklin-Tong V, ed. *Self-incompatibility in flowering plants: evolution, diversity and mechanisms*. Berlin, Germany: Springer-Verlag, 3–32.

Belaoussoff S, Shore JS. 1995. Floral correlates and fitness consequences of mating-system variation in *Turnera ulmifolia*. *Evolution* **49**: 545–556.

Burrows W, McCubbin AG. 2017. Sequencing the genomic regions flanking *S*-linked *PvGLO* sequences confirms the presence of two *GLO* loci, one of which lies adjacent to the style-length determinant gene *CYP734A50*. *Plant Reproduction* **30**: 53–67.

Busch JW, Delph LF. 2012. The relative importance of reproductive assurance and automatic selection as hypotheses for the evolution of self-fertilization. *Annals of Botany* **109**: 553–562.

Carlson ML, Gisler SD, Kelso S. 2008. The role of reproductive assurance in the Arctic: a comparative study of a homostylous and distylous species pair. *Arctic, Antarctic and Alpine Research* **40**: 39–47.

Charlesworth B, Charlesworth D. 1979. The maintenance and breakdown of distyly. *American Naturalist* **114**: 499–513.

Charlesworth D, Wright SI. 2001. Breeding systems and genome evolution. *Current Opinion in Genetics and Development* **11**: 685–690.

Crosby JL. 1949. Selection of an unfavourable gene-complex. *Evolution* **3**: 212–230.

Darriba D, Taboada GL, Doallo R, Posada D. 2012. jModelTest 2: more models, new heuristics and parallel computing. *Nature Methods* **9**: 772.

Darwin C. 1876. *The effects of cross- and self-fertilization in the vegetable kingdom*. London, UK: John Murray.

Darwin C. 1877. *The different forms of flowers on plants of the same species*. London, UK: John Murray.

David P, Pujol B, Viard F, Castella V, Goudet J. 2007. Reliable selfing rate estimates from imperfect population genetic data. *Molecular Ecology* **16**: 2474–2487.

Dieringer D, Schlötterer C. 2003. Microsatellite analyser (MSA): a platform independent analysis tool for large microsatellite data sets. *Molecular Ecology Notes* **3**: 167–169.

Dowrick VPJ. 1956. Heterostyly and homostyly in *Primula obconica*. *Heredity* **10**: 219–236.

Doyle J. 1991. DNA protocols for plants: CTAB total DNA isolation. In: Hewitt GM, Johnston A, eds. *Molecular techniques in taxonomy*. Berlin, Germany: Springer-Verlag, 283–293.

Edgar RC. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research* **32**: 1792–1797.

Favre A, Päckert M, Pauls SU, Jähnig SC, Uhl D, Michalak I, Muellner-Riehl AN. 2015. The role of uplift of the Qinghai-Tibetan Plateau for the evolution of Tibetan biotas. *Biological Reviews* **90**: 236–253.

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

Fisher RA. 1941. Average excess and average effect of a gene substitution. *Annals of Eugenics* **11**: 53–63.

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

Ganders FR, Denny SK, Tsai D. 1985. Breeding systems and genetic variation in *Amsinckia spectabilis* (Boraginaceae). *Canadian Journal of Botany* **63**: 533–538.

Gao H, Williamson S, Bustamante CD. 2007. A Markov chain Monte Carlo approach for joint inference of population structure and inbreeding rates from multilocus genotype data. *Genetics* **176**: 1635–1651.

Goodwillie C, Sargent RD, Eckert CG, Elle E, Geber MA, Johnston MO, Kalisz S, Moeller DA, Ree RH, Vallejo-Marín M *et al.* 2010. Correlated evolution of mating system and floral display traits in flowering plants and its implications for the distribution of mating system variation. *New Phytologist* **185**: 311–321.

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

Grossenbacher D, Briscoe Runquist R, Goldberg EE, Brandvain Y. 2015. Geographical range size is predicted by plant mating system. *Ecology Letters* **18**: 706–713.

Guggisberg A, Mansion G, Kelso S, Conti E. 2006. Evolution of biogeographic patterns, ploidy levels, and breeding systems in a diploid-polyploid species complex of *Primula*. *New Phytologist* **171**: 617–632.

Hu CM, Kelso S. 1996. Primulaceae. In: Wu ZY, Raven PH, eds. *Flora of China*. Beijing, China: Science Press, 99–185.

Huu CN, Kappel C, Keller B, Sicard A, Takebayashi Y, Breuninger H, Nowak MD, Bäurle I, Himmelbach A, Burkart M *et al.* 2016. Presence versus absence of CYP734A50 underlies the style length dimorphism in primroses. *eLife* **5**: e17956.

Igic B, Bohs L, Kohn JR. 2006. Ancient polymorphism reveals unidirectional breeding system shifts. *Proceedings of the National Academy of Sciences, USA* **103**: 1359–1363.

Igic B, Busch JW. 2013. Is self-fertilization an evolutionary dead end? *New Phytologist* **198**: 386–397.

Kappel C, Huu CN, Lenhard M. 2017. A short story gets longer: recent insights into the molecular basis of heterostyly. *Journal of Experimental Botany* **68**: 5719–5730.

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

Kovach WL. 1999. *MVSP 3.1: a multivariate statistical package for Windows*. Pentraeth, UK: Kovach Computing Services.

Langmead B, Salzberg SL. 2012. Fast gapped-read alignment with Bowtie 2. *Nature Methods* **9**: 357–359.

Levin DA. 2012. Mating system shifts on the trailing edge. *Annals of Botany* **109**: 614–620.

Lewis D, Jones DA. 1992. The genetics of heterostyly. In: Barrett SCH, ed. *Evolution and function of heterostyly*. Berlin, Germany: Springer-Verlag, 129–148.

Li J, Cocker JM, Wright J, Webster MA, McMullan M, Dyer S, Swarbreck D, Caccamo M, van Oosterhout C, Gilmatrin PM. 2016. Genetic architecture and evolution of the *S* locus supergene in *Primula vulgaris*. *Nature Plants* **2**: 16188.

Lloyd DG. 1965. Evolution of self-compatibility and racial differentiation in *Leavenworthia* (Cruciferae). *Contributions from the Gray Herbarium of Harvard University* **195**: 3–134.

Lloyd DG. 1980. Demographic factors and mating patterns in angiosperms. In: Solbrig OT, ed. *Demography and evolution in plant populations*. Oxford, UK: Blackwell, 67–88.

Lloyd DG, Webb CJ. 1992. The evolution of heterostyly. In: Barrett SCH, ed. *Evolution and function of heterostyly*. Berlin, Germany: Springer-Verlag, 151–178.

Lohse M, Drechsel O, Kahlau S, Bock R. 2013. OrganellarGenomeDRAW – a suite of tools for generating physical maps of plastid and mitochondrial genomes and visualizing expression data sets. *Nucleic Acids Research* **41**: W575–W581.

Maddison WP, Maddison DR. 2016. Mesquite: a modular system for evolutionary analysis. Version 3.5. [WWW document] URL http://www.mesquiteproject.org [accessed 1 June 2019].

Mast AR, Kelso S, Conti E. 2006. Are any primroses (*Primula*) primitively monomorphic? *New Phytologist* **171**: 605–616.

Mazer SJ, Hultgård UM. 1993. Variation and covariation among floral traits within and among four species of northern European *Primula* (Primulaceae). *American Journal of Botany* **80**: 474–485.

Moeller DA. 2006. Geographical structure of pollinator communities, reproductive assurance, and the evolution of self-pollination. *Ecology* **87**: 1510–1522.

Moeller DA, Briscoe Runquist RD, Moe AM, Geber MA, Goodwillie C, Cheptou P, Eckert CG, Elle E, Johnston MO, Kalisz S *et al.* 2017. Global biogeography of mating system variation in seed plants. *Ecology Letters* **20**: 375–384.

Morgan MT, Barrett SCH. 1989. Reproductive correlates of mating system evolution in *Eichhornia paniculata* (Spreng.) Solms (Pontederiaceae). *Journal of Evolutionary Biology* **2**: 183–203.

Nei M. 1987. *Molecular evolutionary genetics*. New York, NY, USA: Columbia University Press.

Ness RW, Wright SI, Barrett SCH. 2010. Mating-system variation, demographic history and patterns of nucleotide diversity in the tristylous plant *Eichhornia paniculata*. *Genetics* **184**: 381–392.

Nordborg M. 2000. Linkage disequilibrium, gene trees and selfing: an ancestral recombination graph with partial self-fertilization. *Genetics* **154**: 923–929.

Pannell JR. 2015. Evolution of the mating system in colonizing plants. *Molecular Ecology* **24**: 2018–2037.

Patel RK, Jain M. 2012. NGS QC Toolkit: a toolkit for quality control of next generation sequencing data. *PLoS ONE* **7**: e30619.

Pettengill JB, Runquist RDB, Moeller DA. 2016. Mating system divergence affects the distribution of sequence diversity within and among populations of recently diverged subspecies of *Clarkia xantiana* (Onagraceae). *American Journal of Botany* **103**: 99–109.

Piper JG, Charlesworth B, Charlesworth D. 1984. A high rate of self-fertilization and increased seed fertility of homostyle primroses. *Nature* **310**: 50–51.

Posada D, Buckley TR. 2004. Model selection and model averaging in phylogenetics: advantages of Akaike information criterion and Bayesian approaches over likelihood ratio tests. *Systematic Biology* **53**: 793–808.

Pritchard JK, Stephens M, Donnelly P. 2000. Inference of population structure using multilocus genotype data. *Genetics* **155**: 945–959.

R Development Core Team. 2014. *R: a language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical computing. [WWW document] URL http://www.R-project.org/ [accessed 1 June 2019].

Ren G, Conti E, Salamin N. 2015. Phylogeny and biogeography of *Primula* sect. *Armerina*: implications for plant evolution under climate change and the uplift of the Qinghai-Tibet Plateau. *BMC Evolutionary Biology* **15**: 161.

Richards J. 2003. *Primula*. Portland, OR, USA: Timber Press.

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

Sarkissian TS, Harder LD. 2001. Direct and indirect responses to selection on pollen size in *Brassica rapa* L. *Journal of Evolutionary Biology* **14**: 456–468.

Schattner P, Brooks AN, Lowe TM. 2005. The tRNAscan-SE, snoscan and snoGPS web servers for the detection of tRNAs and snoRNAs. *Nucleic Acids Research* **33**: W686–W689.

Schoen DJ, Johnston MO, L'Heureux AM, Marsolais JV. 1997. Evolutionary history of the mating system in *Amsinckia* (Boraginaceae). *Evolution* **51**: 1090–1099.

Sicard A, Lenhard M. 2011. The selfing syndrome: a model for studying the genetic and evolutionary basis of morphological adaptation in plants. *Annals of Botany* **107**: 1433–1443.

Stamatakis A. 2006. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics* **22**: 2688–2690.

Stebbins GL. 1974. *Flowering plants: evolution above the species level*. Cambridge, MA, USA: Belknap Press.

Truyens S, Arbo MM, Shore JS. 2005. Phylogenetic relationships, chromosome and breeding system evolution in *Turnera* (Turneraceae): inferences from ITS sequence data. *American Journal of Botany* **92**: 1749–1758.

Vonhof MJ, Harder LD. 1995. Size-number trade-offs and pollen production by papilionaceous legumes. *American Journal of Botany* **82**: 230–238.

de Vos JM, Keller B, Isham ST, Kelso S, Conti E. 2012. Reproductive implications of herkogamy in homostylous primroses: variation during anthesis and reproductive assurance in alpine environments. *Functional Ecology* **26**: 854–865.

de Vos JM, Keller B, Zhang LR, Nowak MD, Conti E. 2018. Mixed mating in homostylous species: genetic and experimental evidence from an alpine plant with variable herkogamy, *Primula halleri*. *International Journal of Plant Sciences* **179**: 87–99.

de Vos JM, Wüest RO, Conti E. 2014. Small and ugly? Phylogenetic analyses of the "selfing syndrome" reveal complex evolutionary fates of monomorphic primrose flowers. *Evolution* **68**: 1042–1057.

Wang XJ, Zhong L, Wu ZK, Sun HY, Wang H, Li DZ, Barrett SCH, Zhou W. 2019. Characterization of 30 microsatellite markers in distylous *Primula sinolisteri* (Primulaceae) using HiSeq sequencing. *Applications in Plant Sciences* 7: e1208.

Wedderburn F, Richards AJ. 1990. Variation in within-morph incompatibility inhibition sites in heteromorphic *Primula* L. *New Phytologist* **116**: 149–162.

Weller SG. 1992. Evolutionary modifications of tristylous breeding systems. In: Barrett SCH, ed. *Evolution and function of heterostyly*. Berlin, Germany: Springer-Verlag, 247–272.

Wright SI, Kalisz S, Slotte T. 2013. Evolutionary consequences of self-fertilization in plants. *Proceedings of the Royal Society B: Biological Sciences* **280**: 20130133.

Wu LY, Wang B, Schoen DJ, Huang SQ. 2017. Transitions from distyly to homostyly are associated with floral evolution in the buckwheat genus (*Fagopyrum*). *American Journal of Botany* **104**: 1232–1240.

Wyman SK, Jansen RK, Boore JL. 2004. Automatic annotation of organellar genomes with DOGMA. *Bioinformatics* **20**: 3252–3255.

Xing Y, Ree RH. 2017. Uplift-driven diversification in the Hengduan mountains, a temperate biodiversity hotspot. *Proceedings of the National Academy of Sciences, USA* **114**: E3444–E3451.

Yamamoto M, Ohtani M, Kurata K, Setoguchi H. 2017. Contrasting evolutionary processes during Quaternary climatic changes and historical orogenies: a case study of the Japanese endemic primroses *Primula* sect. *Reinii*. *Annals of Botany* **120**: 943–954.

Yuan S, Barrett SCH, Duan TT, Qian X, Shi MM, Zhang DX. 2017. Ecological correlates and genetic consequences of evolutionary transitions from distyly to homostyly. *Annals of Botany* **120**: 775–789.

Yuan S, Barrett SCH, Li C, Li X, Xie K, Zhang DX. 2019. Genetics of distyly and homostyly in a self-compatible *Primula*. *Heredity* **122**: 110–119.

Yuan S, Zeng G, Zhang DX. 2018. Development of microsatellite markers for *Primula oreodoxa* (Primulaceae), a distylous-homostylous species. *Applications in Plant Sciences* **6**: e1150.

Zhou W, Barrett SCH, Li HD, Wu ZK, Wang XJ, Wang H, Li DZ. 2017. Phylogeographic insights on the evolutionary breakdown of heterostyly. *New Phytologist* **214**: 1368–1380.

Supporting Information

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

Fig. S1 The geographical distribution of 12 taxa in *Primula* section *Obconicolisteri* investigated in this study.

Fig. S2 Total sample of flowers of species in *Primula* section *Obconicolisteri* measured in this study ranked by style length.

Fig. S3 Gene maps of species of *Primula* section *Obconicolisteri* investigated in this study.

Fig. S4 Floral conditions in *Primula* section *Obconicolisteri* reconstructed by using the program Mesquite.

Fig. S5 Population structure inferred from InStruct analysis of polymorphism at SSR loci in *Primula* section *Obconicolisteri*.

Table S1 The plastid genome features of the sequenced species of *Primula* section *Obconicolisteri* investigated in this study.

Table S2 Taxa of *Primula* section *Obconicolisteri* investigated in this study and the summary information of populations sampled.

Table S3 Comparison of genetic diversity and selfing rate based on SSR loci among species of *Primula* section *Obconicolisteri*.

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 *New Phytologist* Central Office.