New Phytologist

Research

Are any primroses (*Primula*) primitively monomorphic?

Austin R. Mast[1], Sylvia Kelso[2] and Elena Conti[3]

[1]Department of Biological Science, Florida State University, Tallahassee, FL 32306, USA; [2]Department of Biology, Colorado College, Colorado Springs, CO 80903, USA; [3]Institute of Systematic Botany, University of Zürich, Zollikerstrasse 107, 8008 Zürich, Switzerland

Author for correspondence:
Austin Mast
Tel: +1850 6451500
Fax: +1850 6449829
Email: amast@bio.fsu.edu

Received: *7 December 2005*
Accepted: *24 January 2006*

Summary

• *Primula* (*c.* 430 species) and relatives (Primulaceae) are paradigmatic to our understanding of distyly. However, the common co-occurrence of distyly and monomorphy in closely related groups within the family has made the interpretation of its evolution difficult.

• Here, we infer a chloroplast DNA (cpDNA) phylogeny for 207 accessions, including 51% of the species and 95% of the sections of *Primula* with monomorphic populations, using Bayesian methods. With this tree, we infer the distribution of ancestral states on critical nodes using parsimony and likelihood methods.

• The inferred cpDNA phylogeny is consistent with prior estimates. The most recent common ancestor (MRCA) of *Primula* is resolved as distylous using both methods of inference. However, whether the distyly in *Primula*, *Hottonia*, and *Vitaliana* arose once or three independent times is not clear.

• We conclude that monomorphism in descendants of the MRCA of *Primula* is derived from distyly in all cases. Thus, scenarios for the evolution of distyly that rely on the persistence of primitive monomorphy (such as in *Primula* section *Sphondylia*) require re-evaluation.

Key words: distyly, evolution, heteromorphic incompatibility, heterostyly, *Hottonia*, phylogeny, *Primula*, *Vitaliana*.

New Phytologist (2006) **doi**: 10.1111/j.1469-8137.2006.01700.x

Introduction

The flowers of primroses and relatives (family Primulaceae) have played a central role in our understanding of the form, function, genetics, and evolution of distyly (e.g. Darwin, 1877; Charlesworth & Charlesworth, 1979b; Ganders, 1979; Barrett, 1992; Richards, 2003). However, the common co-occurrence of distyly and monomorphy (the term is here used to describe populations or species fixed for a single floral morphology) in closely related groups within the family has made interpretation of the evolution of distyly difficult. Ninety-two per cent of the *c.* 430 species of *Primula* have distylous populations (Richards, 2003), as do members of the closely related genera *Hottonia* (one of two species), *Dionysia* (40 of 41 species), and *Vitaliana* (one of one species). The 45 species of *Primula* known to have monomorphic populations outside cultivation have been placed into 19 of the 38 sections (Richards, 2003; Table 1), and they share 18 of these 19 sections with distylous species. Here, we infer a chloroplast DNA (cpDNA) phylogeny for the family using the largest taxonomic and character sampling published to date and present the optimal states for key ancestral nodes, including the most recent common ancestor (MRCA) of *Primula*. This permits us to assess an important evolutionary hypothesis that some of the monomorphic species in *Primula*, particularly in *Primula* sections *Proliferae* and *Sphondylia*, represent a persistence of the monomorphy that existed before the evolution of distyly and thus perhaps reveal the suite of floral traits preceding that complex adaptation.

In distylous plants, a genetic polymorphism produces two floral types (morphs) among individuals of a population. The two floral morphs have their anthers and stigma at reciprocal heights (reciprocal herkogamy; Webb & Lloyd, 1986). This structural difference is often accompanied by a sporophytically controlled, diallelic incompatibility system that makes intermorph crosses more successful than intramorph crosses. Reciprocal herkogamy reduces pollen wastage, and thus increases male fitness, whereas the diallelic incompatibility protects

Table 1 Distribution of monomorphic taxa[1] in Richards' (2003) classification of *Primula* with presence (+) or absence (–) in sampling indicated

Subgen. *Sphondylia*	+ *P. eximia*
Sect. *Sphondylia*	Sect. *Cordifoliae*
+ *P. verticillata*	Sect. *Fedtschenkoana*
+ *P. simensis*	Sect. *Proliferae*
+ *P. floribunda*[2]	+ *P. prolifera*[2]
Subgen. *Auriculastrum*	+ *P. chungensis*[2]
Sect. *Auricula*	+ *P. cockburniana*
Sect. *Cuneifolia*	+ *P. japonica*
+ *P. cuneifolia* ssp. *saxifragifolia*[3]	– *P. miyabeana*
Sect. *Suffrutescens*	+ *P. prenantha*
Sect. *Amethystina*	– *P. polonensis*
Sect. *Parryi*	Sect. *Sikkimensis*
Subgen. *Primula*	– *P. morsheadiana*
Sect. *Primula*	Sect. *Oreophlomis*
+ *P. vulgaris*[2]	Sect. *Armerina*
Sect. *Sredinskya*	+ *P. egaliksensis*
+ *P. grandis*	Sect. *Glabra*
Subgen. *Auganthus*	– *P. macrocarpa*
Sect. *Auganthus*	Sect. *Yunnanensis*
Sect. *Monocarpicae*	– *P. homogama*
Sect. *Obconicolisteri*	– *P. clutterbuckii*
+ *P. sinolisteri* var. *aspera*[3]	Sect. *Aleuritia*
– *P. filipes*	– *P. frondosa*[2]
– *P. dumicola*	+ *P. halleri*[2]
– *P. listeri*	– *P. scotica*
Sect. *Malvacea*	+ *P. scandinavica*
Sect. *Pycnoloba*	– *P. stricta*
Sect. *Reinii*	+ *P. incana*
Sect. *Cortusoides*	+ *P. laurentiana*
+ *P. mollis*	– *P. magellanica*
– *P. septemloba*	– *P. yuparensis*
Sect. *Bullatae*	Sect. *Pulchella*
Sect. *Dryadifolia*	Sect. *Minutissimae*
Subgen. *Pinnatae*	– *P. annulata*
Sect. *Pinnatae*	+ *P. muscoides*
+ *P. cicutariifolia*	– *P. subularia*
Subgen. *Carolinella*	– *P. praetermissa*
Sect. *Carolinella*	Sect. *Denticulata*
– *P. larsenii*	Sect. *Capitatae*
Subgen. *Aleuritia*	Sect. *Muscarioides*
Sect. *Chartacea*	+ *P. watsonii*
Sect. *Davidii*	– *P. concholoba*
Sect. *Petiolares*	+ *P. bellidifolia*[2]
– *P. hookeri*	Sect. *Soldanelloides*
Sect. *Crystallophlomis*	– *P. sherriffae*

[1]Taxa at the rank of species or below are listed when monomorphic populations have been reported outside of cultivation.
[2]Distylous populations of these species have also been reported outside of cultivation.
[3]Other taxa of this species are distylous outside of cultivation.
P., *Primula*.

against self-fertilization and inbreeding depression, and thus increases female fitness (Barrett, 2002). Additional morphological features, such as pollen size and stigmatic papillae length, differ between floral morphs in *Primula*, *Hottonia*, and *Dionysia* (Darwin, 1877; Schaeppi, 1934; Wendelbo, 1961a,c; Richards, 2003), although not in *Vitaliana* (Schaeppi, 1934). These might serve to further promote intermorph pollination (reviewed in Dulberger, 1992). Distyly, and the functionally similar tristyly, are currently known from 28 families scattered throughout the angiosperm phylogeny, suggesting many independent origins (Barrett *et al.*, 2000; Barrett, 2002).

Within *Primula*, distyly is thought to be controlled by at least three tightly linked genes (Dowrick, 1956), which are collectively referred to as the 'heterostyly supergene' (e.g. Ganders, 1979; Barrett, 2002). The style length, stigmatic papillae length, and female mating type are thought to be controlled by locus G, the pollen size and male mating type by locus P (or two separate loci Pp and Pm, respectively; Kurian & Richards, 1997), and the anther height by locus *A*. 'Pin' plants, with their stigma positioned high in the flower and their anthers low, are homozygous recessive for these loci (gpa/gpa); 'thrum' plants, with their stigma and anther positions reversed, are heterozygous (GPA/gpa; Bateson & Gregory, 1905; Lewis & Jones, 1992). The rarity of homozygote thrums might be attributable to the presence of recessive sublethal alleles linked to GPA (Mather & de Winton, 1941; Kurian & Richards, 1997; Richards, 1998). The loci are likely ordered GPA or GAP based on the frequency of observed recombinations (Dowrick, 1956; Lewis & Jones, 1992; Richards, 2003; cf. Charlesworth & Charlesworth, 1979a).

Monomorphy in members of *Primula* has been interpreted as either primitive (representing the condition before the origin of distyly in the genus; 'primary homostyly') or derived from distyly ('secondary homostyly'). The most common route to monomorphy from distyly is thought to be via a recombination of the heterostyly supergene to produce the genotype gPA/Gpa with subsequent fixation of the gPA alleles in a population (Charlesworth & Charlesworth, 1979a). This produces a single, self-compatible morph with anthers at the high position of the thrum producing pollen that is the large size of the thrum and a stigma at the high position of the pin ('long homostyle'). The term 'homostyly' has a long history in discussions of heterostyly (e.g. Darwin, 1877). Today, it is typically reserved for species or populations that have a single floral morphology, have their anthers and stigma at the same height in the flower, are thought to be derived from heterostylous ancestors, and/or are very closely related to extant heterostylous species (e.g. Ganders, 1979). Here, we use the term 'monomorphy' as a description of floral homogeneity in a population or species, as it does not imply a particular spatial relationship between anthers and stigma or a particular evolutionary scenario.

It has long been noted that there is differential reproductive success when interspecific crosses are made within and between morphs of two close distylous relatives (e.g. Darwin, 1877). It is also known that, upon crossing some monomorphic species with the two morphs of a close distylous relative, the monomorphic species will act as thrums when the pollen donor and pins when the pollen recipient in these crosses

(e.g. Ernst, 1943), suggesting the persistence of a diallelic incompatibility system. This has been interpreted as evidence for recombination (Dowrick, 1956; Barrett & Shore, 1987; Wedderburn & Richards, 1992).

Ernst (1943, 1955) found that crosses between two monomorphic species (*Primula cockburniana* and *Primula chungensis*) and the two morphs of a close distylous relative (*Primula pulverulenta*) within *Primula* section *Proliferae* did not result in the differential success expected immediately after recombination. He concluded from this that these two species are primitively monomorphic. Although the predicted incompatibility responses are absent, these two monomorphic species are similar in morphology to what is expected if the typical recombinational route to monomorphy is taken (long style, anthers high in the corolla tube, and large pollen).

Examined species of *Primula* section *Sphondylia* (six of eight species) lack a diallelic incompatibility system (Al Wadi & Richards, 1993), and thus patterns in crossing success cannot be used to argue for the primitive or derived nature of monomorphy in the section. However, section *Sphondylia* displays a number of putatively primitive character states (e.g. inflorescences with superimposed whorls, and diploidy) that suggest an early divergence of the lineage (or lineages, if the section is not monophyletic) from other species in *Primula* before the origin of more derived character states (Al Wadi & Richards, 1993; Richards, 1993). Al Wadi & Richards (1993) assumed that two monomorphic species in the section, *Primula simensis* and *Primula verticillata,* are primitively so, 'representing the original condition in the genus before distyly evolved, and that various distylous conditions in other species of the [section] can be considered to be representative of the evolutionary stages by which the distylous syndrome evolved' (p. 337). In Al Wadi & Richards' (1993) view, each species in the section is a 'palaeo-endemic relict evolved from more widespread ancestors, which once ranged widely from eastern Africa to western India., [and] five species appear to represent isolates from a morphological and geographical continuum' (p. 330). They asserted that interspecific comparisons involving these five species represent a complementary approach to previous theoretical work on the evolution of heterostyly (Charlesworth & Charlesworth, 1979b; Lloyd & Webb, 1992), and they constructed a scenario for it based upon what Richards (1993) called the 'intermediate stages in the evolution of "full" heterostyly, "frozen" in evolutionary time' (internal quotation marks his; p. 65). It is unclear whether Richards and Al Wadi preferred the view that the species of *Sphondylia* represent the intermediate steps in the origin of the distyly that is seen throughout *Primula* (e.g. last paragraph of Al Wadi & Richards, 1993), suggesting the paraphyly of the section with respect to the rest of the genus, or the view that they represent a build-up of distyly in section *Sphondylia* independent of the evolutionary origin(s) responsible for the distyly seen elsewhere in the genus (e.g. p. 65 in Richards, 1993). However, they clearly consider the monomorphic condition present in *P. simensis* and *P. verticillata* to represent the primitive combination of features for the genus, and thus a good starting point for a general scenario of the origin of distyly in *Primula*. Like *P. cockburniana* and *P. chungensis* from section *Proliferae* and monomorphic species that appear to have arisen following a recombinational route, *P. simensis* and *P. verticillata* have long styles, anthers high in the corolla tube, and large pollen.

In our interpretation of the Al Wadi & Richards (1993) scenario for the evolution of distyly, there are five major features that involve the introduction of new alleles to the heterostyly supergene.

1 The primitive condition for the genus is represented by *P. simensis* (in the Ethiopian highlands) and *P. verticillata* (in the south-west corner of the Arabian Peninsula). This involves a long style (g), large pollen size (P), anthers high in the corolla tube (A), and the absence of diallelic incompatibility.

2 The product of the first step, in which a dominant allele for short style length (G) arises in a population with genotype gPA, is represented by the thrum morph of *Primula boveana* (on the Sinai peninsula), *Primula gaubeana* (in west Iran), and *Primula davisii* (in south-east Turkey).

3 The product of the second step, in which a recessive allele for small pollen size (p) becomes linked to g in a population with genotypes of gPA and GPA, is represented by the other morph of *P. boveana* [called the 'pin' morph by Al Wadi & Richards (1993), although it has anthers high, rather than low, in the corolla tube].

4 The product of the third step, in which a recessive allele for low anther height (a) becomes linked to gp in a population with genotypes of GPA, gpA, and perhaps gPA, is represented by the pin morph of *P. gaubeana* and *P. davisii*.

5 The final step, the emergence of a diallelic incompatibility system, is represented by the remaining distylous species in the genus *Primula*.

Our focus here is on feature (1) of the scenario, and we will reserve detailed considerations of the remaining features of the scenario for elsewhere.

Recent DNA sampling in the Primulaceae has provided results critical to our understanding of the phylogeny and evolution of distyly and monomorphy in the family. Both cpDNA and nuclear DNA data suggest that Primulaceae, as traditionally circumscribed (e.g. Cronquist, 1981), is not monophyletic (Martins *et al.*, 2003; Anderberg *et al.*, 1998; Källersjö *et al.*, 2000; Mast *et al.*, 2001; Trift *et al.*, 2002), and some genera – including *Cyclamen*, *Lysimachia*, and *Samolus* – are more closely related to Theophrastaceae and Mrysinaceae. With their extensive taxonomic sampling of cpDNA in the Primulaceae *sensu stricto* (s.str.), Mast *et al.* (2001) and Trift *et al.* (2002) inferred a bifurcation at the root of the family, with *Androsace* and genera nested in it (*Vitaliana*, *Douglasia*, and *Pomatosase*) sister to the remaining genera. Mast *et al.*'s (2001) sampling of the ribosomal protein L16 (*rpl*16) gene and tRNA-Leu (*trn*L) gene introns resolved a bifurcation in this latter clade that

produced a clade of *Primula* and genera nested in it [*Dionysia*, *Dodecatheon*, *Cortusa*, and *Sredinskya* (treated as *Primula grandis* here)] sister to a clade formed by *Hottonia*, *Omphalogramma*, and *Soldanella*. Trift *et al.*'s (2002) sampling of the ribulose-1,5-bisphosphate carboxylase/oxygenase large subunit (*rbcL*) gene resolved this latter clade as part of a polytomy with three clades of *Primula* and nested genera. The only extensive nuclear DNA sampling of the family to date (Martins *et al.*, 2003; using the internal transcribed spacers of the nuclear ribosomal DNA) inferred poor support (≤ 62% bootstrap frequencies) for the relevant nodes, although it inferred strong support (100% bootstrap frequency) for the position of the previously unsampled monomorphic genus *Bryocarpum*, sister to the single sampled species of *Omphalogramma*.

The objectives of our study were (1) to infer the cpDNA phylogeny for an expanded taxonomic sampling within Primulaceae s.str. in order (2) to test by ancestral state inference the hypothesis that monomorphic members of *Primula*, particularly taxa in sections *Sphondylia* and *Proliferae*, are primitively monomorphic [i.e. represent lineages that diverged from the group before the origin(s) of distyly] and (3) to determine by ancestral state inference the number of origins of distyly in Primulaceae. We consider phylogenetic inference from cpDNA sequence variation to be a good starting point for estimating organismal relationships and character evolution in the family. However, wholesale acceptance of the cpDNA phylogeny as capturing all relevant evolutionary relationships is premature while fine-resolution taxonomic samplings of independent 'linkage partitions' (*sensu* Slowinski & Page, 1999) from the nuclear or mitochondrial genomes are unavailable. Comparison of independent linkage partitions can uncover biological processes (e.g. ancient hybridization and introgression) that would reduce the predictive utility of phylogenies inferred from any single linkage partition (Rieseberg & Soltis, 1991; Rieseberg & Brunsfeld, 1992).

Materials and Methods

Taxonomic sampling

We sequenced four cpDNA regions from 207 accessions (Supplementary Table S1). These represent 147 (34%) of the 430 species, 34 of the 38 sections, and six of the seven subgenera of *Primula* recognized by Richards (2003), as well as at least one species from each of the remaining 11 genera recognized as part of Primulaceae s.str. (Källersjö *et al.*, 2000). Mast *et al.* (2001) inferred strong support (≥ 99% bootstrap frequencies) for a clade composed of four of these 11 genera (*Androsace*, *Douglasia*, *Vitaliana*, and *Pomatosace*) and a sister clade composed of the remaining sampled taxa in the Primulaceae. Consequently, we show the root of the family between this clade of four genera and the clade of remaining genera.

The sampling at the rank of species in *Primula* included 23 (51%) of the 45 species of *Primula* documented to have monomorphic populations outside of cultivation. Furthermore, it included 18 of the 19 sections of *Primula* with monomorphic species and seven of the eight subgenera of *Primula* with monomorphic species (Table 1). Data on the presence of monomorphic populations outside cultivation were taken from Wedderburn & Richards (1992) and Richards (2003). For many taxa of *Primula*, comparisons were not made to determine whether a diallelic incompatibility system exists or whether there are differences in ancillary characters (e.g. pollen size and stigmatic papillae length) in populations (Richards, 2003). Here, we consider evidence for the presence or absence of reciprocal herkogamy to be necessary and sufficient to classify a species as distylous or monomorphic. This is a common requirement for application of the term 'distyly' in the literature (e.g. Lloyd & Webb, 1992; Barrett, 2002).

Generation of the cpDNA data set

When available, we used the previously published DNA sequence data from Mast *et al.* (2001, 2004) for the specimens. To these pre-existing data, we added 138 new sequences for the maturase K (*matK*) gene, 89 for the *rpl16* intron, 89 for the *trnL* intron, and 142 for the tRNA-Leu gene/tRNA-Phe gene (*trnL/F*) spacer (Supplementary Table S1). Our generation of new DNA sequences for the four cpDNA regions followed the protocols outlined in Mast *et al.* (2001, 2004). To detect mistakes and correct uncertainties in the computer-generated sequence, we compared aligned trace-files in SEQUENCHER 4.1 (Gene Codes Corporation, Ann Arbor, MI, USA).

The manually aligned sequence matrices are available from the first author (ARM). We determined the borders of each region from comparisons with the complete cpDNA sequence of *Nicotiana tabacum* L. (GenBank accession NC_001879). Each dataset represents the complete sequence for each region and does not include portions of adjacent regions, with one exception. The *rpl16* intron dataset is missing 50 nt on the 5′ end of the intron because of the position of the priming site. The dataset is > 99% complete; there are *c.* 7250 question marks in the *c.* 964 kb of data. Each taxon has sequence data for each of the four regions.

Phylogenetic inference

We calculated Bayesian posterior probabilities for branches with MRBAYES 3.1 (Huelsenbeck & Ronquist, 2001; Ronquist & Huelsenbeck, 2003). MRBAYES used a general-time-reversible (GTR; Lanave *et al.*, 1984; Tavare, 1986; Rodriguez *et al.*, 1990) substitution model with among-site rate heterogeneity assumed to follow a gamma distribution (Γ; Yang, 1994) and a proportion of sites invariant (*I*). It estimated the parameters of the substitution model independently for each of the four cpDNA regions. MRBAYES spawned two independent Markov Chain Monte Carlo (MCMC) runs, each with one cold and three heated chains. It sampled the two runs

every 1000th generation for 1.5×10^6 generations. By 1.5×10^6 generations, the average standard deviation of split frequencies in the two runs was < 0.02, the potential scale reduction factor (Gelman & Rubin, 1992) for every parameter (a 'rough guide to convergence' according to the MrBayes manual) was within 0.06 of 1, and there was nearly complete congruence between the two runs in the branches resolved with ≥ 0.95 posterior probability (with a burn-in of 1.0×10^5 generations excluded, as suggested by a plot of the log likelihoods).

Ancestral state inference

MESQUITE OSX version 1.05 (Maddison & Maddison, 2004) inferred the presence/absence of distyly at ancestral nodes in a parsimony framework using the topology that included only branches with posterior probabilities ≥ 0.95, an assumption that polytomies are 'soft' (will later be resolved with more data), and alternate gain:loss weightings of 1 : 1, 2 : 1, 3 : 1, and 20 : 1. Weighting the loss of distyly more than the gain of it (e.g. with a weighting of 1 : 2) appears unrealistic to us, given the complexity of the character and the relative ease with which it appears to be lost (by recombination and fixation of the recombinant). Others (e.g. Kohn *et al.*, 1996) have also used the complexity of heterostyly to argue for a greater weighting of its gain over its loss.

MESQUITE calculated the likelihood of the data under the alternative models of equal rates of gain and loss (a one-parameter Markov k-state model; Lewis, 2001) and different rates of gain and loss (an asymmetrical two-parameter Markov k-state model). MESQUITE does not accept polymorphic terminal taxa for the calculations, and we converted these to the absent state as this seemed most conservative given the hypothesis to be tested. The likelihood values were compared using the likelihood ratio test (Goldman, 1993) with one degree of freedom to determine whether the more complex asymmetrical model produced a significantly higher likelihood than the simpler equal rates model. For the likelihood calculations, MESQUITE used a tree with branches of > 0.50 posterior probability and branch lengths that are the mean from the posterior probability density. MrBayes generated this tree with the 'sumt' command. MESQUITE reported the results as proportional likelihoods for the data given the alternative states at each node, with a threshold of significance set at 2. This is equivalent to the threshold of a log likelihood ratio of 7.4 : 1 advocated by Edwards (1972). It is a threshold used elsewhere for likelihood inference of ancestral states as a 'rough minimum' (Schluter *et al.*, 1997).

Results

Phylogenetic inference

The topology that includes only branches with posterior probabilities ≥ 0.95 in one of the two MCMC runs is given in Fig. 1 (arithmetic mean of the –log likelihood (L) of trees sampled after the burn-in = −31265.46; harmonic mean = −31367.36). All but three of the branches in Fig. 1 (shown as dotted lines) also had ≥ 0.95 posterior probability in the other MCMC run.

The topology is in broad agreement with the relationships previously inferred in Primulaceae using smaller taxonomic samplings of cpDNA (Mast *et al.*, 2001, 2004; Trift *et al.*, 2002). Membership in the six clades of *Primula* recognized by Mast *et al.* (2001; their fig. 3) is congruent between the current and previous cpDNA studies (ignoring the additional taxa sampled here), as are the relationships inferred among the clades where these were previously resolved. In particular, the inferred position of *Primula* sections *Sphondylia* and *Proliferae* with the current cpDNA sampling is congruent with previous results. Section *Sphondylia* is separated from the MRCA of sampled taxa of *Primula* by the same four nodes (ignoring the additional taxa sampled here) as in Mast *et al.* (2001) and Trift *et al.* (2002), and *Proliferae* is separated from the MRCA by three of the four nodes inferred in Mast *et al.* (2001). Also relevant to our inference of the evolution of distyly is the position of the previously unsampled monomorphic *Hottonia inflata*, sister to the distylous *Hottonia palustris* (Fig. 1).

Ancestral state inference

With equal weightings of gain:loss, distyly is unequivocally inferred to have arisen four times in Primulaceae s.str. (Fig. 2a): (1) in the lineage of *Vitaliana primuliflora*, (2) in the lineage of *H. palustris*, (3) in the lineage leading to the MRCA of *Primula*, and (4) in the lineage of *Primula prolifera*. If the gain of distyly is weighted ≥ 3 times more than the loss of distyly (the 20 : 1 weighting results in the same results as in Fig. 2c), a single origin of distyly is inferred before the MRCA of Primulaceae s.str. and all cases of distyly are inferred to represent the uninterrupted transmission of distyly from this single origin (Fig. 2c). If the gain of distyly is weighted twice as much as the loss of distyly (Fig. 2b), then these two scenarios for the origin(s) of distyly at the base of Primulaceae s.str. (Figs 2a,c) are equally optimal. In each of these three scenarios, the MRCA of *Primula* (Fig. 2) is inferred to be distylous, and thus all origins of monomorphy in the descendants are inferred as derived from distyly.

The likelihood of the data given the more complex, asymmetrical model ($-\log L$ = 87.5748; rate of gain = 0.6737; rate of loss = 0.9303) was not significantly better (at a threshold of P = 0.05) than that for the equal rates model ($-\log L$ = 87.6982; rate of change = 0.9286) as judged with the likelihood ratio test. Thus, we report here the proportional likelihoods of the data under alternative state assumptions for select nodes of interest near the base of the tree using the equal rates model (Fig. 3). The likelihoods calculated assuming the alternative states for the MRCA of *Primula* and immediately descendant nodes in that clade were significantly different using our threshold of significance, and at each of these nodes

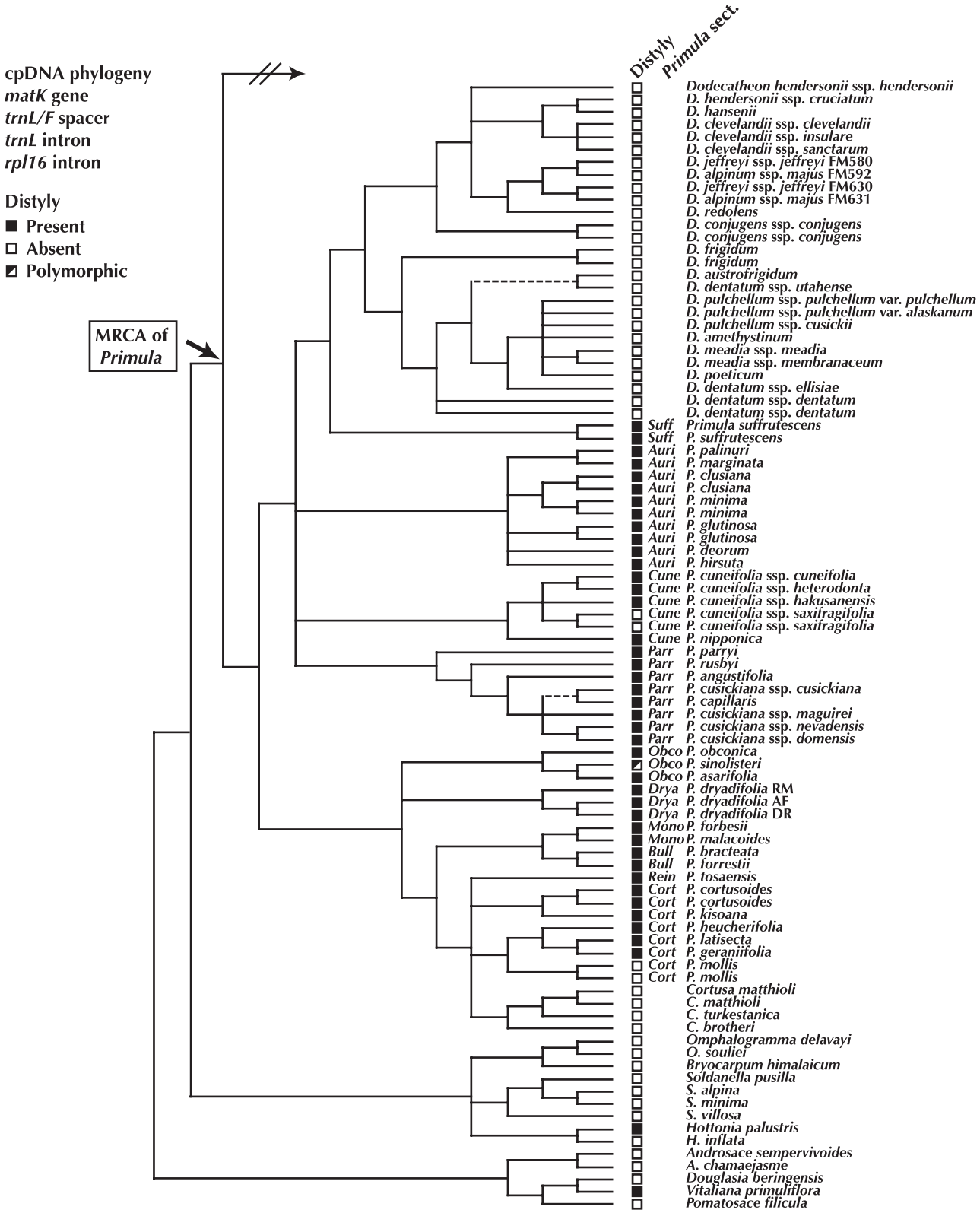

Fig. 1 Chloroplast DNA (cpDNA) phylogeny inferred using Bayesian methods. All branches shown had a posterior probability of ≥ 0.95 in both of two independent Markov Chain Monte Carlo (MCMC) runs except the three shown as dotted lines. Membership in a section of *Primula* is indicated where appropriate using the first four letters of the section name as an abbreviation (see Table 1 for section names). The most recent common ancestor (MRCA) of *Primula* is indicated. *Matk*, maturase K; *trn*L/F, tRNA-Leu gene/tRNA-Phe gene; *trn*L, tRNA-Leu gene; *rpl*16, ribosomal protein L16.

Fig. 1 *Continued*

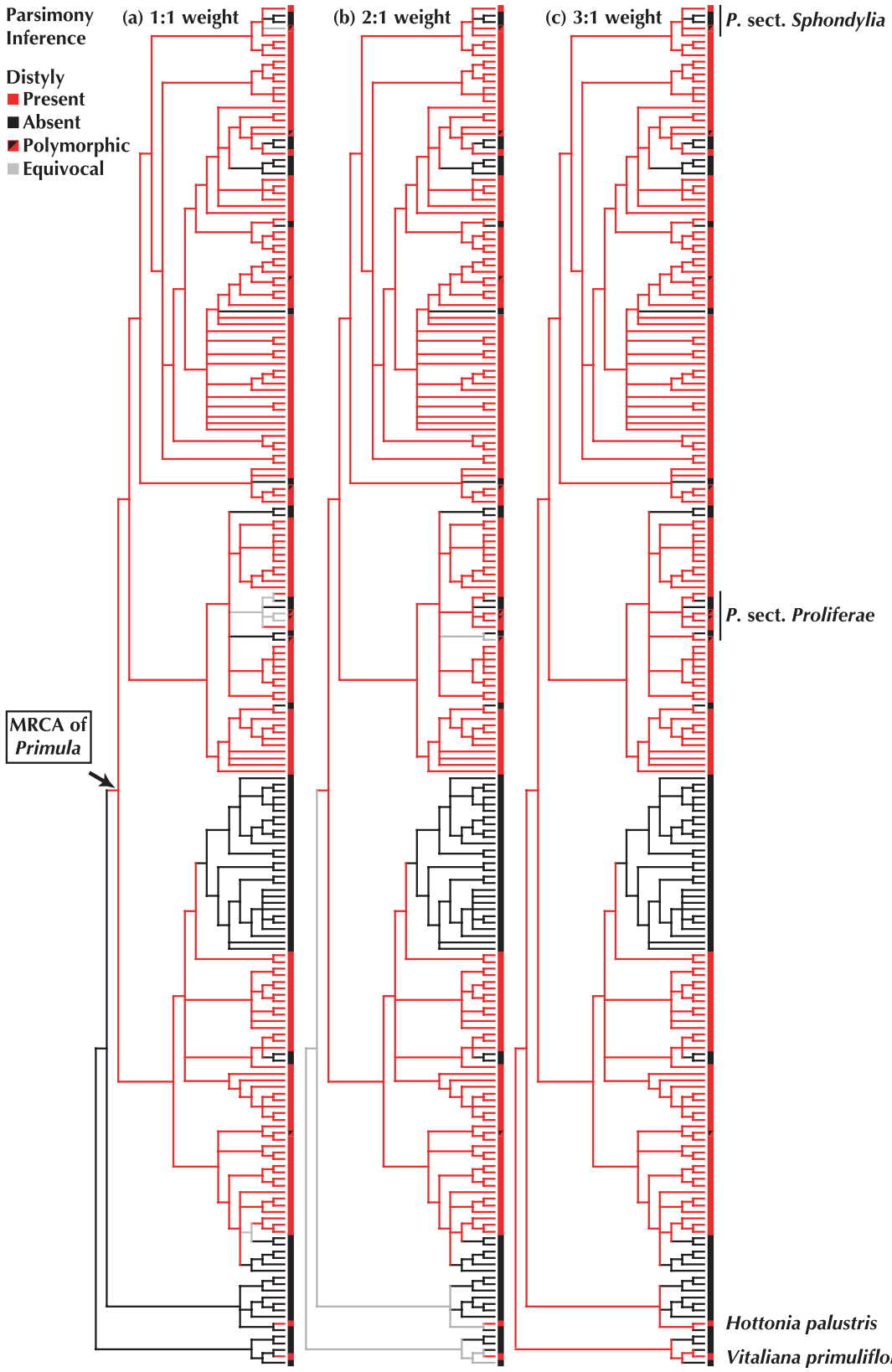

Fig. 2 Parsimony inference of ancestral states on the chloropast DNA (cpDNA) topology shown in Fig. 1 using alternative weightings of gain:loss of distyly. The inference using a weighting of 20 : 1 is identical to that using 3 : 1. The positions of important taxa are indicated on the right. MRCA, most recent common ancestor.

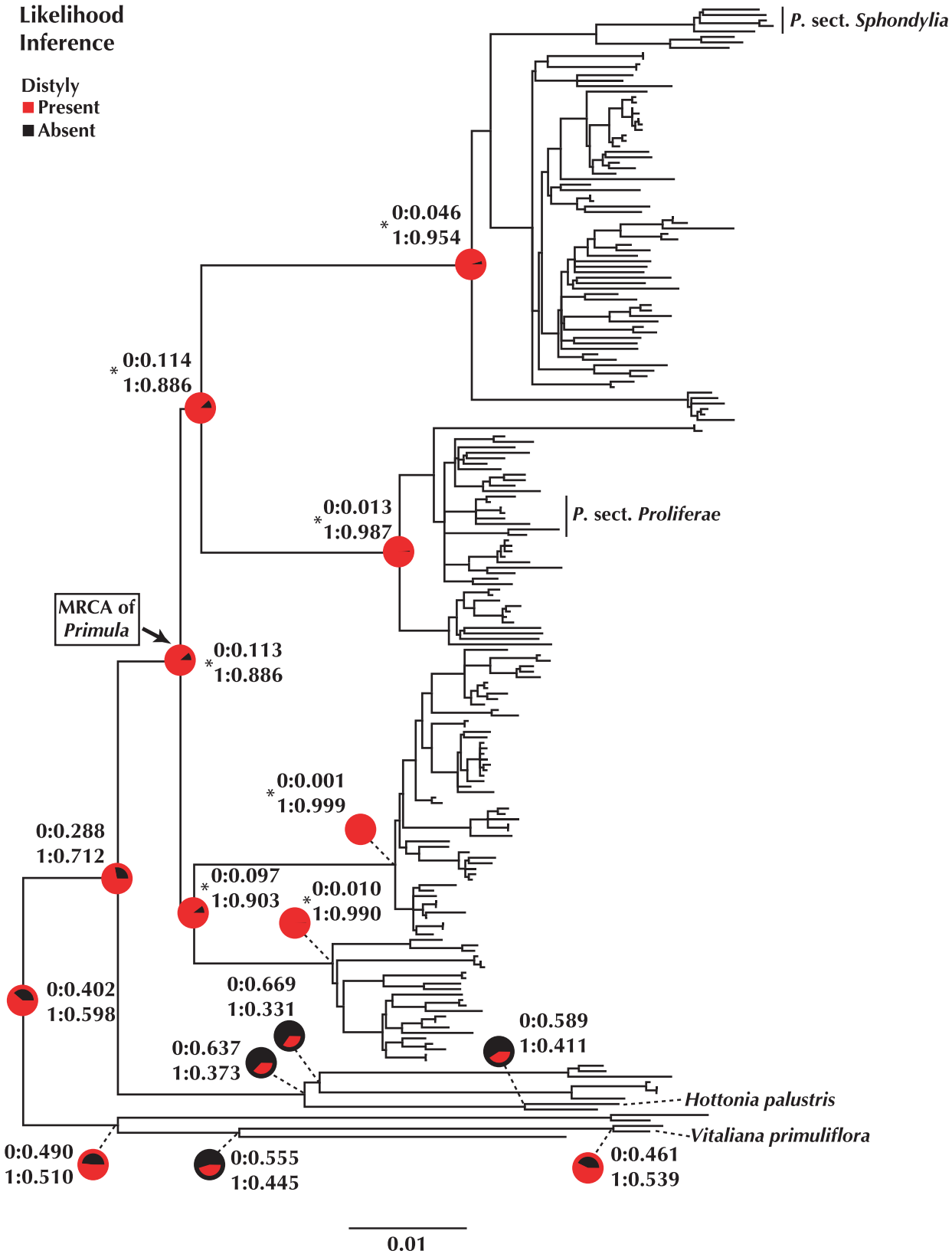

Fig. 3 Likelihood inference of ancestral states on the chloroplast DNA (cpDNA) topology using a one-parameter model. Branches shown have a > 0.50 posterior probability, and the branch lengths are the mean from the posterior probability density. The likelihoods of the data given the alternative states at each node are shown as proportional likelihoods. Significantly different likelihoods are indicated with an asterisk (*).

the state conferring the greatest likelihood was the presence of distyly (Fig. 3). The likelihoods calculated assuming alternative states for nodes of interest outside of that clade were not significantly different (e.g. the node resolving *Vitaliana* as sister to *Pomatosace* had a proportional likelihood of 0.461 : 0.539).

Discussion

The most recent common ancestor of *Primula* was distylous

Bayesian analysis of sequence variation in the four cpDNA regions provided a well-resolved, phylogenetic hypothesis for the 207 accessions of *Primula* and related genera (Fig. 1). Relationships inferred here are congruent with previous results inferred using fewer taxa and cpDNA sequence characters (Mast *et al.*, 2001; Trift *et al.*, 2002). Using alternative weightings of gain:loss in a parsimony framework suggests that the distyly of Primulaceae s.str. arose one or more times (Fig. 2). In each of these weighting scenarios, the MRCA of *Primula* is unequivocally inferred to have been distylous. Furthermore, the likelihood of the data given the alternative states (distylous vs monomorphic) is significantly better when the MRCA of *Primula* is constrained to be distylous (Fig. 3). This implies that every monomorphic lineage in *Primula* is derived from distyly. In this light, previous evidence for the existence of primitive monomorphy in extant taxa that purportedly diverged before the origin(s) of distyly merits re-examination.

A reconsideration of the evidence for primitive monomorphy

Ernst's (1943, 1955) conclusion that *Primula* section *Proliferae* contains primitively monomorphic species was derived from the results of crossing experiments between monomorphic taxa (*P. cockburniana* and *P. chungensis*) and the two morphs of a close distylous relative (*P. pulverulenta*). He interpreted the 'absence of incompatibility in what should be illegitimate combinations' to be 'evidence of the primary nature of monomorphism of these species relative to the dimorphism in other species of [*Primula* section *Proliferae*]' (Ernst, 1955, pp. 435–6).

However, two points are relevant to this interpretation of the data. First, it is unclear what selective mechanism would maintain the alleles involved in the pre-existing diallelic incompatibility system in monomorphic lineages following recombination (Baker, 1966). Thus, their loss over time should not be surprising. Secondly, one of these two monomorphic species, *P. chungensis*, has both monomorphic and distylous populations (Smith & Fletcher, 1941; Richards, 2003). The assumption that the monomorphic populations of *P. chungensis* represent a persistence of monomorphy in a lineage that diverged before the origin of distyly seems to require one to assume either that the distyly of the remaining populations arose independent of that in other distylous species of *Primula* or that the species is paraphyletic with respect to the other distylous species of *Primula*. To our knowledge, neither of these scenarios has been previously suggested in the literature. A third explanation, that the ancestral lineage in which distyly arose maintained primitively monomorphic populations over cladogenetic events up to the present, seems impossible given that the interbreeding between distylous and monomorphic populations necessary to maintain the identity of the diverging lineages would introduce the derived states of the heterostyly supergene to the monomorphic populations.

Al Wadi & Richards (1993) considered two species of *Primula* section *Sphondylia* (*P. verticillata* and *P. simensis*) to be primitively monomorphic, 'representing the original condition in the genus before distyly evolved' (p. 337), and thus a good starting point for their scenario for the evolution of distyly in *Primula*. This assumption stems from the conclusion of Wendelbo (1961b) and Richards (1993) that section *Sphondylia* is an early diverging lineage (or lineages, if it is not monophyletic) in *Primula*. These authors based their conclusion on putatively primitive morphological and cytological character states, with some disagreement over which character states should indeed be considered primitive (in the case of leaf vernation). Mast *et al.* (2001) found that some of these putatively primitive states are inferred to be advanced or particularly labile on the cpDNA phylogeny.

Rather than appearing as an early diverging lineage on the cpDNA phylogeny, the MRCA of section *Sphondylia* is separated from the MRCA of *Primula* by four nodes (Fig. 1), each of which is resolved as distylous (Figs 2, 3). With our current taxonomic sampling (five of eight species) and resolution of relationships in section *Sphondylia* (Fig. 1), it would be premature to rule out the possibility that this primitive distyly was lost before the MRCA of section *Sphondylia* or the MRCA of some clade within section *Sphondylia* and that the extant taxa represent steps in the re-emergence of distyly. Similarly, it would be premature to rule out the possibility that the extant taxa represent steps in the break-down (rather than build-up) of distyly in the section. However, our results do imply that none of the species in section *Sphondylia* represents an uninterrupted transmission of monomorphy from the ancestors that preceded the origin of distyly in *Primula*. This weakens Al Wadi & Richards' (1993) argument that the starting conditions of the scenario for the path to distyly represent the conditions before the origin of distyly in *Primula* and calls into question the general utility of the scenario for understanding the evolution of distyly in the genus.

More generally, we can conclude from the results of our ancestral state inference that all of the monomorphic lineages descended from the MRCA of *Primula* are derived from distyly. As pointed out elsewhere (Lloyd & Webb, 1992), there is no *a priori* reason to believe that species or populations representing fixation of a recombination in the heterostyly supergene (in most cases self-compatible plants with long styles, anthers

high in the corolla tube, and large pollen) represent the condition before the origin of distyly. Thus, taxonomic and character state sampling outside of the lineage of *Primula* and embedded genera will be crucial to inferring the floral states in monomorphic ancestors immediately preceding the origin of distyly in *Primula*. With colleagues, we are assembling a dataset with a more complete taxonomic sampling of *Androsace* to reconstruct relevant floral states in ancestors that preceded the origin(s) of distyly for comparison with theoretical models (Charlesworth & Charlesworth, 1979b; Lloyd & Webb, 1992) and similar analyses in other distylous groups (e.g. Graham & Barrett, 2004).

Effects of taxon sampling on ancestral state inference

The effect of taxon sampling on conclusions drawn from ancestral state inference can be substantial (Salisbury & Kim, 2001). In the present application, we expect taxon sampling to have a smaller effect on the inference of distyly in the MRCA of *Primula* than on our inference of the number of origins of distyly in Primulaceae s.str. If one or more early diverging, monomorphic lineages of *Primula* are unsampled, then our conclusion of distyly in the MRCA of *Primula* indeed might not be robust to future samplings of these lineages. However, if each of the sections with unsampled monomorphic members (Table 1) proves to be monophyletic [or each of the lineages grouped within polyphyletic sections (e.g. section *Minutissimae*) are sampled here], then this problem is unlikely to arise, because all but one of the sections with monomorphic members was sampled in this study. The single reported case of a monomorphic species (*Primula larsenii*) in the unsampled section with a monomorphic taxon (section *Carolinella*) would introduce uncertainty into the parsimony inference of the presence/absence of distyly in the MRCA of *Primula* only if *P. larsenii* proved to be sister to all other species of *Primula* (including the distylyous members of section *Carolinella*). Such a scenario is unlikely, given the strong morphological similarity of *P. larsenii* to a distylous member of the same section (*Primula kwangtungensis*; Richards, 2003).

We anticipate that our inference of the number of origins of distyly in Primulaceae s.str. will prove to have been more sensitive to taxonomic sampling. Inference of shifts from distyly to monomorphy to distyly again in descendants of the MRCA of *Primula*, as seen in *P. prolifera*, is likely a result of incomplete sampling in the sections with a large fraction of monomorphic taxa. Outside *Primula* we also predict sensitivity to taxon sampling. The more complete taxon sampling of cpDNA data in *Androsace* by Schneeweiss *et al.* (2004) resolved two additional internal nodes between the distylous *Vitaliana primuliflora* and the MRCA of *Androsace*. If we add two additional monomorphic taxa to our tree as placeholders for these two missing nodes, the distyly of *Vitaliana* is unequivocally inferred to have an origin independent of that for *Primula*, *Dionysia*, and *Hottonia* with weightings of gain:loss of 1 : 1, 2 : 1, and 3 : 1. The inference is equivocal at a weighting of 4 : 1, and it resolves a single origin of the distyly in the Primulaceae s.str. at a weighting of 5 : 1. Estimating the effect of these two additional nodes on the likelihood inference is not as straightforward because the data would modify the estimated model parameter. However, we expect that inclusion of these additional monomorphic clades would result in an increase in the proportional likelihood of the data with monomorphy at the MRCA of *Androsace*.

Acknowledgements

We thank the many collectors, botanical gardens, and herbaria that generously provided material for this study and whose names are listed in the supplementary material and appendices of this paper and Mast *et al.* (2001, 2004). We extend special thanks to A.J. Richards for providing valuable assistance and to S. Wagen, D. Lang, and E. Jones for generating many of the sequences. Two anonymous reviewers provided valuable comments. This project was funded by grants from the Swiss National Science Foundation (no. 3100-061674.00/1) and the Office for Equal Opportunity of the University of Zurich to EC and from the Florida State University Council on Research and Creativity to ARM.

References

Al Wadi H, Richards AJ. 1993. Primary homostyly in *Primula* L. subgenus *Sphondylia* (Duby) Rupr. and the evolution of distyly in *Primula*. *New Phytologist* **124**: 329–338.

Anderberg AA, Ståhl B, Källersjö M. 1998. Phylogenetic relationships in the Primulales inferred from *rbcL* sequence data. *Plant Systematics and Evolution* **211**: 93–102.

Baker HG. 1966. The evolution, functioning and breakdown of heteromorphic incompatibility systems. I. The Plumbaginaceae. *Evolution* **20**: 349–368.

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

Barrett SCH. 2002. The evolution of plant sexual diversity. *Nature Genetics* **3**: 274–284.

Barrett SCH, Jesson LK, Baker AM. 2000. The evolution and function of stylar polymorphisms in flowering plants. *Annals of Botany* **85**: 253–265.

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

Bateson W, Gregory RP. 1905. On the inheritance of heterostylism in *Primula*. *Proceedings of the Royal Academy of London, Series B* **76**: 581–586.

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

Charlesworth D, Charlesworth B. 1979b. A model for the evolution of distyly. *American Naturalist* **114**: 467–498.

Cronquist A. 1981. *An integrated system of classification of flowering plants*. Columbia University Press, New York, NY, USA.

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

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

Dulberger R. 1992. Floral polymorphisms and their functional significance in the heterostylous syndrome. In: Barrett SCH, ed. *Evolution and function of heterostyly*. Berlin, Germany: Springer-Verlag, 51–84.
Edwards AWF. 1972. *Likelihood*. Cambridge, UK: Cambridge University Press.
Ernst A. 1943. Die Nachkommenschaften aus Kreuzungen zwischen dimorphen und monomorphen *Primula*-Arten der Sektion *Candelabra*. *Archiv der Julius Klaus-Stiftung* **18**: 87–233.
Ernst A. 1955. Self-fertility in monomorphic primulas. *Genetica* **27**: 391–448.
Ganders FR. 1979. The biology of heterostyly. *New Zealand Journal of Botany* **17**: 607–635.
Gelman A, Rubin DB. 1992. Inference from iterative simulation using multiple sequences. *Statistical Science* **7**: 457–511.
Goldman N. 1993. Statistical tests of models of DNA substitution. *Journal of Molecular Evolution* **36**: 182–198.
Graham SW, Barrett SCH. 2004. Phylogenetic reconstruction of the evolution of stylar polymorphisms in *Naracissus* (Amaryllidaceae). *American Journal of Botany* **91**: 1007–1021.
Huelsenbeck JP, Ronquist F. 2001. MRBAYES: Bayesian inference of phylogenetic trees. *Bioinformatics* **17**: 754–755.
Källersjö M, Bergqvist G, Anderberg AA. 2000. Generic realignment in primuloid families of the Ericales s.1. A phylogenetic analysis based on DNA sequences from three chloroplast genes and morphology. *American Journal of Botany* **87**: 1325–1341.
Kohn JR, Graham SW, Morton B, Doyle JJ, Barrett SCH. 1996. Reconstruction of the evolution of reproductive characters in Pontederiaceae using phylogenetic evidence from chloroplast restriction-site variation. *Evolution* **50**: 1454–1469.
Kurian V, Richards AJ. 1997. A new recombinant in the heteromorphy 'S' supergene in *Primula*. *Heredity* **78**: 383–390.
Lanave C, Preparata G, Saccone C, Serio G. 1984. A new method for calculating evolutionary substitution rates. *Journal of Molecular Evolution* **20**: 86–93.
Lewis PO. 2001. A likelihood approach to estimating phylogeny from discrete morphological character data. *Systematic Biology* **50**: 913–925.
Lewis D, Jones DA. 1992. The genetics of heterostyly. In: Barrett SCH, ed. *Evolution and function of heterostyly*. Berlin, Germany: Springer-Verlag, 129–150.
Lloyd DG, Webb CJ. 1992. The evolution of heterostyly. In: Barrett SCH, ed. *Evolution and function of heterostyly*. Berlin, Germany: Springer-Verlag, 151–178.
Maddison WP, Maddison DR. 2004. *Mesquite, a modular system for evolutionary analysis*, version 1.05. http//mesquiteproject.org
Martins L, Oberprieler C, Hellwig FH. 2003. A phylogenetic analysis of Primulaceae s.1. based on internal transcribed spacer (ITS) DNA sequence data. *Plant Systematics and Evolution* **237**: 75–85.
Mast AR, Feller MS, Kelso S, Conti E. 2004. Buzz-pollinated *Dodecatheon* originated from within the heterostylous *Primula* subgenus *Auriculastrum* (Primulaceae): a seven-region cpDNA phylogeny and its implications for floral evolution. *American Journal of Botany* **91**: 926–942.
Mast AR, Kelso S, Richards AJ, Lang DJ, Feller DMS, Conti E. 2001. Phylogenetic relationships in *Primula* L. and related genera (Primulaceae) based on noncoding chloroplast DNA. *International Journal of Plant Sciences* **162**: 1381–1400.
Mather K, De Winton D. 1941. Adaptation and counter-adaptation of the breeding system in *Primula*. *Annals of Botany* **5**: 297–311.
Richards J. 1993. *Primula* First Ed. Portland, OR, USA: Timber Press.
Richards AJ. 1998. Lethal linkage and its role in the evolution of plant breeding systems. In: Owens SJ, Rudall PJ, eds. *Reproductive biology*. Kew, UK: Royal Botanic Gardens, Kew, 71–83.
Richards J. 2003. *Primula* Second Ed. Portland, OR, USA: Timber Press.
Rieseberg LH, Brunsfeld SJ. 1992. Molecular evidence and plant introgression. In: Soltis PS, Soltis DE, Doyle JJ, eds. *Molecular systematics of plants*. New York, NY, USA: Chapman & Hall, 151–176.
Rieseberg LH, Soltis DE. 1991. Phylogenetic consequences of cytoplasmic gene flow in plants. *Evolutionary Trends in Plants* **5**: 65–84.
Rodriguez F, Oliver JL, Marin A, Medina JR. 1990. The general stochastic model of nucleotide substitution. *Journal of Theoretical Biology* **142**: 485–501.
Ronquist F, Huelsenbeck JP. 2003. MRBAYES 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* **19**: 1572–1574.
Salisbury BA, Kim J. 2001. Ancestral state estimation and taxon sampling density. *Systematic Biology* **50**: 557–564.
Schaeppi H. 1934. Untersuchungen über die Narben- und Antherenstellung in den Blüten der Primulaceen. *Archiv Julius-Klaus-Stiftung Vererbungsforschung, Sozialanthropologie Rassenhygiene* **10**: 133–236.
Schluter D, Price T, Mooers Ø, Ludwig D. 1997. Likelihood of ancestor states in adaptive radiation. *Evolution* **51**: 1699–1711.
Schneeweiss GM, Schönswetter P, Kelso S, Niklfeld H. 2004. Complex biogeographic patterns in *Androsace* (Primulaceae) and related genera: Evidence from phylogenetic analyses of nuclear internal transcribed spacer and plastid *trnL-F* sequences. *Systematic Biology* **53**: 856–876.
Slowinski JB, Page RDM. 1999. How should species phylogenies be inferred from sequence data? *Systematic Biology* **48**: 814–825.
Smith WW, Fletcher HR. 1941. The genus *Primula*: Section *Candelabra*. *Transactions and Proceedings of the Botanical Society of Edinburgh* **33**: 122–181.
Tavare S. 1986. Some probabilistic and statistical problems on the analysis of DNA sequences. *Lectures on Mathematics in the Life Sciences* **17**: 57–86.
Trift I, Källersjö M, Anderberg AA. 2002. The monophyly of *Primula* (Primulaceae) evaluated by analysis of sequences from the chloroplast gene *RbcL*. *Systematic Botany* **27**: 396–407.
Webb CJ, Lloyd DG. 1986. The avoidance of interference between the presentation of pollen and stigmas in angiosperms II. Herkogamy. *New Zealand Journal of Botany* **24**: 163–178.
Wedderburn FM, Richards AJ. 1992. Secondary homostyly in *Primula* L. Evidence for the model of the S-supergene. *New Phytologist* **121**: 649–655.
Wendelbo P. 1961a. Studies in Primulaceae I. A monograph of the genus *Dionysia*. *Arbok for Universitetet I Bergen, Series Mathematica Rerumque Naturalium* **3**: 1–88.
Wendelbo P. 1961b. Studies in Primulaceae II. An account of *Primula* subgenus *Sphondylia* (syn. sect. *Floribundae*) with a review of the subdivisions of the genus. *Arbok for Universitetet I Bergen, Series Mathematica Rerumque Naturalium* **11**: 1–46.
Wendelbo P. 1961c. Studies in Primulaceae III. On the genera related to *Primula* with special reference to their pollen morphology. *Arbok for Universitetet I Bergen, Series Mathematica Rerumque Naturalium* **19**: 1–31.
Yang Z. 1994. Maximum likelihood phylogenetic estimation from DNA sequences with variable rates over sites: approximate methods. *Journal of Molecular Evolution* **39**: 306–314.

Supplementary Material

The following supplementary material is available for this article online:

Table S1 Species in *Primula* are organized into the subgenera and sections of Richards (2003). Vouchers are herbarium specimens unless otherwise noted.

This material is available as part of the online article from http://www.blackwell-synergy.com

