ORIGINAL ARTICLE

MOLECULAR ECOLOGY WILEY

Different molecular changes underlie the same phenotypic transition: Origins and consequences of independent shifts to homostyly within species

Emiliano Mora-Carrera[1] 🔟 | Rebecca L. Stubbs[1] | Barbara Keller[1] |
Étienne Léveillé-Bourret[1,2] | Jurriaan M. de Vos[3] | Peter Szövényi[1] | Elena Conti[1]

[1]Department of Systematic and Evolutionary Botany, University of Zurich, Zurich, Switzerland

[2]Département de Sciences Biologiques, Institut de Recherche en Biologie Végétale, Université de Montréal, Montréal, Québec, Canada

[3]Department of Environmental Sciences – Botany, University of Basel, Basel, Switzerland

Correspondence
Emiliano Mora-Carrera and Elena Conti, Department of Systematic and Evolutionary Botany, University of Zurich, Zurich, Switzerland.
Emails: emiliano.mora@systbot.uzh.ch; elena.conti@systbot.uzh.ch

Funding information
Schweizerischer Nationalfonds zur Förderung der Wissenschaftlichen Forschung, Grant/Award Number: 3100-061674.00/1, 310030_185251 and 31003A_175556; Swiss Government Excellence Scholarship, Grant/Award Number: 2018.0475; Universitat Zurich

Abstract

The repeated transition from outcrossing to selfing is a key topic in evolutionary biology. However, the molecular basis of such shifts has been rarely examined due to lack of knowledge of the genes controlling these transitions. A classic example of mating system transition is the repeated shift from heterostyly to homostyly. Occurring in 28 angiosperm families, heterostyly is characterized by the reciprocal position of male and female sexual organs in two (or three) distinct, usually self-incompatible floral morphs. Conversely, homostyly is characterized by a single, self-compatible floral morph with reduced separation of male and female organs, facilitating selfing. Here, we investigate the origins of homostyly in *Primula vulgaris* and its microevolutionary consequences by integrating surveys of the frequency of homostyles in natural populations, DNA sequence analyses of the gene controlling the position of female sexual organs (CYP^T), and microsatellite genotyping of both progeny arrays and natural populations characterized by varying frequencies of homostyles. As expected, we found that homostyles displace short-styled individuals, but long-style morphs are maintained at low frequencies within populations. We also demonstrated that homostyles repeatedly evolved from short-styled individuals in association with different types of loss-of-function mutations in CYP^T. Additionally, homostyly triggers a shift to selfing, promoting increased inbreeding within and genetic differentiation among populations. Our results elucidate the causes and consequences of repeated transitions to homostyly within species, and the putative mechanisms precluding its fixation in *P. vulgaris*. This study represents a benchmark for future analyses of losses of heterostyly in other angiosperms.

KEYWORDS
heterostyly, intraspecific, loss-of-function mutations, mating system, *Primula*, selfing

1 | INTRODUCTION

A central topic in evolutionary biology is the repeated transition from outcrossing to selfing (Cutter, 2019). A main advantage of selfers over outcrossers is the increase in the number of gene copies transmitted to the next generation (i.e., automatic selection advantage; Fisher, 1941; reviewed in Busch & Delph, 2012). Additionally, selfers can reproduce when mates or pollinators, in the case of plants, are scarce (i.e., reproductive assurance; Darwin, 1876). These advantages are counteracted by fitness reduction in selfed progeny due to inbreeding depression (Schemske & Lande, 1985). In plants, transitions to selfing are enabled by the loss of self-incompatibility (SI). The genetic basis of shifts from SI to self-compatibility (SC) has been investigated in systems where the genes responsible for self-pollen recognition have been characterized (Shimizu & Tsuchimatsu, 2015). Repeated transitions to SC via the inactivation of SI genes are often studied by comparing closely related species with contrasting mating systems (Vekemans et al., 2014). Independent losses of SI have also been reported within species (Busch, Joly & Schoen, 2011; Foxe et al., 2010; Goodwillie, 1999), but only a few studies have investigated the molecular basis of these losses (Chantha et al., 2013; Shimizu et al., 2008).

A classic model for the study of mating system transitions is the shift from heterostyly to homostyly (Barrett, 2019). Heterostyly, occurring in at least 28 angiosperm families, is a floral polymorphism that promotes outcrossing, whereby two (distyly) or three (tristyly) floral morphs differing in the position of reproductive organs co-occur in a population (Darwin, 1877). In distylous species, flowers of the long-styled morph, henceforth L-morph (also known as pin), exhibit the stigma above the anthers, whereas flowers of the short-styled morph, henceforth S-morph (also known as thrum), show the reciprocal arrangement (Ganders, 1979; Figure 1). The reciprocal position of male and female sexual organs in flowers of heterostylous taxa simultaneously promotes pollen export to individuals of the opposite floral morph and reduces self-pollination (Keller et al.,2014; Lloyd & Webb, 1992). Moreover, heterostyly is often, but not always, associated with a heteromorphic self-incompatibility system that reduces self- and intramorph fertilization and with differences in the size of pollen grains and stigma papillae of S- and L-morphs (Dulberger, 1992). Thus, heterostyly is commonly described as a complex trait (Barrett, 2019). Heterostyly is controlled by a single Mendelian locus known as the heterostyly supergene or S-locus (Lewis & Jones, 1992). Recent genomic and functional studies in *Primula* indicated that the S-locus is entirely absent from the L-morph and hemizygous in the S-morph, where it consists of five tightly linked genes (Huu et al., 2016; Li et al., 2016). Two genes in the S-locus, GLOBOSA2 (GLO^T) and CYP734A50 (CYP^T), have been experimentally demonstrated to control anther and stigma positions, respectively (Huu et al., 2016, 2020).

As is typical for complex traits, heterostyly has been repeatedly lost. Conforming to predictions about mating system transitions, losses of heterostyly are usually associated with the decline or absence of specialized pollinators (Pérez-Barrales & Arroyo, 2010;
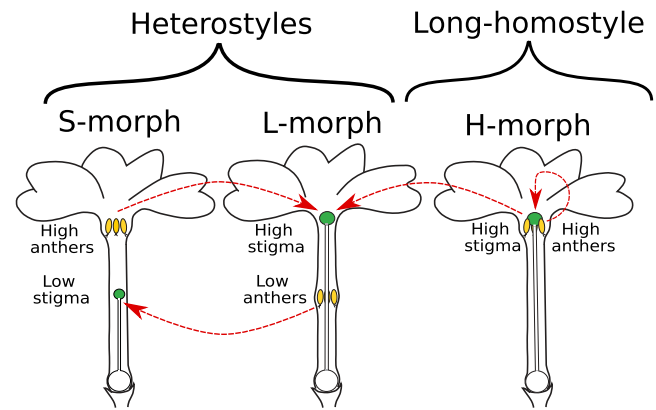
FIGURE 1 Heterostyly and homostyly in *Primula vulgaris*. Heterostylous floral morphs can be either short-styled (S-morphs) or long-styled (L-morphs). They are characterized by reciprocal placement of male (anthers) and female (stigmas) sexual organs (i.e., reciprocal herkogamy), promoting outcrossing. Additionally, a physiological self-incompatibility mechanism prevents self- and intramorph fertilization. Homostylous floral morphs (H-morphs) are characterized by reduced or no herkogamy and are self-compatible, favoring selfing. Most homostyles have both stigma and anthers at the mouth of the corolla tube (i.e., long-homostyly); short-homostyly is rare (not shown). Long-homostyles stem from S-morphs and retain their male compatibility type, thus can fertilize only L-morphs. Conversely, S-morphs cannot fertilize long-homostyles due to stigma clogging in the latter. Red arrows indicate compatible, likely pollinations between floral morphs. Anthers are represented in yellow and stigmas in green

Yuan et al., 2017) or the absence of plant mates, as in colonization after long-distance dispersal (Ness et al., 2010) and recolonization after glacial retreat (Guggisberg et al., 2006). The breakdown of heterostyly can occur either through the loss of one or more of the heterostylous floral morphs (e.g., transitions from tristyly and distyly to monomorphism; Ness et al., 2010; Pérez-Barrales et al., 2009) or the invasion of homostylous plants (Barrett, 2019). Homostylous flowers are characterized by reduced distance between male and female sexual organs (i.e., reduced herkogamy) and, typically, loss of SI (de Vos, Hughes, et al., 2014; de Vos, Wüest, et al., 2014). Both long- and short-homostyles are known, the latter being rare (Lewis & Jones, 1992). Long- and short-homostyles, respectively, bear both anthers and stigmas either higher (H-morph in Figure 1) or lower in the corolla tube. Homostyly enables reproductive assurance through autonomous self-pollination, as demonstrated in *Primula* (Carlson et al., 2008; de Vos et al., 2012; Yuan et al., 2017), and increases selfing rates (Husband & Barrett, 1993; Schoen et al., 1997; Zhong et al., 2019). Repeated shifts from heterostyly to homostyly have been reported both between (Costa, Torices, & Barrett, 2019; Kissling & Barrett, 2013; Kohn et al., 1996; Ruiz-Martín et al., 2018; de Vos, Hughes, et al., 2014) and within species (Ness et al., 2010; Shao et al., 2019; Zhou et al., 2017). However, none of these studies identified the molecular basis of such transitions, probably because the necessary genomic resources have only recently become available.

Recent studies in *Primula* revealed that the transition from heterostyly to homostyly is achieved via disruption of either GLO^T or CYP^T. For example, a transposable element insertion probably compromising the function of GLO^T was detected in one short-homostyle of a cultivated *P. vulgaris* variety (Li et al., 2016). Additionally, silencing GLO^T caused a reduction of anther height in S-individuals, leading to the expression of short-homostyly in *Primula forbesii* (Huu et al., 2020), although heteromorphic SI was retained, preventing self-fertilization. Conversely, silencing CYP^T in S-individuals of *P. veris* increased stigma height, causing the shift to long-homostyly (Huu et al., 2016). Correspondingly, the expression of CYP^T was reduced in long-homostyles compared to S-individuals of both *P. vulgaris* and *P. oreodoxa* (Huu et al., 2016; Zhao et al., 2019). Furthermore, all five exons of CYP^T were probably lost in long-homostyles of the ancestrally heterostylous *P. forbesii*, whereas exon three of CYP^T was not detected by PCR assays in the long-homostylous *P. grandis*, *P. halleri*, and *P. scotica* (Huu et al., 2016). Moreover, two loss-of-function mutations in CYP^T were reported in self-compatible long-homostyles of *P. vulgaris*, suggesting that CYP^T might additionally be involved in the control of SI (Li et al., 2016), although this has not been experimentally demonstrated (Kappel et al., 2017). The fact that mutations in CYP^T are associated with both long-homostyly and loss of SI, whereas mutations in GLO^T are associated with short homostyly but retention of SI, and the fact that GLO^T has a lower mutation rate than CYP^T (Huu et al., 2020) might explain the lower frequencies of short- than long-homostyles observed in nature. Altogether, the findings above indicate that any genetic changes disrupting the function or expression of CYP^T, a gene present only in S-individuals, cause long-homostyly. However, comprehensive analyses of the different types of mutations associated with shifts to homostyly based on extensive sampling of heterostyles and homostyles from multiple natural populations have never been performed.

After homostyles originate, their reproductive advantages over heterostyles should prompt changes in floral morph composition and genetic diversity of populations. Previous theoretical models predicted that homostyles should increase in frequency by replacing S-individuals (Crosby, 1949, 1960). Concomitantly, crosses between homostyles and L-individuals and reduced fitness of homostyles, presumably caused by homozygosity of genes at the S-locus, were expected to prevent the fixation of homostyly (Crosby, 1949, 1960). Both mechanisms should ultimately result in the co-occurrence of homostyles with L-individuals, the latter remaining at low frequency. The predictions above were supported by two surveys of the frequency of L-, S- and homostylous individuals from natural populations of *P. vulgaris* performed at a 36-year interval (Crosby, 1949; Curtis & Curtis, 1985). Subsequent studies testing whether homostyles were affected by inbreeding depression found no evidence of seed-mass differences between heterostyles and homostyles (Piper et al., 1986). However, inbreeding depression at later life-cycle stages has never been investigated. At the population genetic level, higher selfing in homostyles should increase homozygosity, thus decreasing diversity within and increasing differentiation among populations (Hamrick & Godt, 1996). These microevolutionary consequences of mating system transitions have been quantified only in a few homostylous species (Ness et al., 2010; Yuan et al., 2017; Zhong et al., 2019; Zhou et al., 2017).

Primula vulgaris Huds. (the common primrose) represents a classical model for the study of heterostyly and intraspecific transitions to homostyly (Bodmer, 1958; Cocker et al., 2018; Crosby, 1958; Piper et al., 1986). While distylous populations consisting exclusively of L- and S-individuals are common throughout Eurasia, mixed populations with varying frequencies of long-homostylous, S-, and L-individuals have been reported and widely documented for decades only in Somerset, UK (Crosby, 1940; Curtis & Curtis, 1985). The occurrence of homostyles in multiple populations and their wide frequency variation in this region prompted Crosby (1949) to suggest that homostyles migrated from an initial center of origin to neighbouring populations. However, two different loss-of-function mutations in CYP^T were recently detected in two long-homostyles from different natural populations of *P. vulgaris* in England (Li et al., 2016), suggesting that homostyly might arise independently rather than *via* migration. Therefore, population genetic analyses aimed at clarifying the role of migration in the spread of homostyly are needed.

This study investigates the origins of homostyly, its effects on the mating system, and its population genetic consequences using the well-characterized *P. vulgaris* model. We combined surveys of floral morph frequencies from multiple natural populations of the ancestrally heterostylous *P. vulgaris* with both DNA sequence analyses of CYP^T and microsatellite data from multiple individuals to ask the following questions: (1) Are increases of homostyle frequency associated with reductions of S-individuals and maintenance of L-individuals at low frequencies, as predicted by the model of Crosby (1949) model and found in earlier studies (Curtis & Curtis, 1985)? (2) Are transitions to homostyly associated with a single or multiple molecular changes in CYP^T and which ones? (3) Are losses of heterostyly accompanied by shifts from outcrossing to selfing? If so, is there evidence of inbreeding depression as well as increased inbreeding within and genetic differentiation among populations? Our study generates new knowledge on the molecular basis of transitions from outcrossing to selfing, its evolutionary consequences, and the putative mechanisms precluding the fixation of homostyly within species.

2 | MATERIALS AND METHODS

2.1 | Study species

The ancestrally heterostylous *Primula vulgaris* is a perennial, rosette-forming diploid (2n = 22) blooming in early spring (February–April) with pale-yellow, fused corollas reaching up to 40 mm in width formed by broad, overlapping, v-notched lobes (Richards, 2003). *Primula vulgaris* has dimorphic pollen grains and stigma papillae (Mast et al., 2006): pollen grains are small in L-individuals but large in S-individuals, whereas stigma papillae are long in L-individuals but

short in S-individuals (Richards, 2003). The species reproduces mainly sexually, but clonal reproduction, although uncommon in the wild, has been reported under cultivation (Boyd et al.,1990). Pollination in natural populations is mostly performed by long-tongued bees, bumblebees, and bee-flies, but butterflies, pollen-gathering bees and moths have been occasionally reported (Boyd et al., 1990; Keller et al., 2020; Woodell, 1960). Seeds, produced in capsules, are dispersed by ants and rodents (Valverde & Silvertown, 1995). *Primula vulgaris* has a patchy distribution with populations ranging from dozens to hundreds of individuals mostly found in woodlands and hedgerows and less frequently in grasslands (Jacquemyn et al., 2009). Previous studies indicated that long-homostyles of *P. vulgaris* are self-compatible and produce seeds in the absence of pollinators (Piper et al., 1986). Analyses of the mating system of homostyles estimated a range of outcrossing rates from 0.05 to 0.8 (Bodmer, 1958; Crosby, 1958; Piper et al., 1984). It has been suggested that reduced pollinator service due to intensive pastoral and agricultural activities could be the selective pressure favouring the increase of self-fertilizing homostyles in Somerset, England (Piper et al., 1986). Recently, long-homostyles have been reported in one population of an agriculturally fragmented area in the Netherlands (Barmentlo et al., 2017).

2.2 | Frequency variation of floral morphs in natural populations of *Primula vulgaris*

To test whether homostyles increase concomitantly with the decrease of S-individuals, while L-individuals are maintained at low frequencies, we quantified floral morph composition in 22 natural populations of the Somerset region (England) in the spring of 2019 (Figure 2a). The approximate location of these populations was based on Curtis and Curtis (1985). The frequency of floral morphs was estimated by randomly scoring 100 individuals whenever possible (Table 1). Flowering plants were visually assigned to S-, L- and H-morphs (Figure 1). To determine whether S- and L-morphs segregated at equal frequencies in both dimorphic heterostylous (i.e., populations consisting of S- and L-individuals) and trimorphic populations (i.e., populations consisting of S-, L- and H-individuals), we tested whether morph ratios deviated significantly from isoplethy (i.e., equal frequencies of S- and L-morphs) using G-tests of goodness-of-fit as implemented in the R package DescTools v0.99.38 (Signorell et al., 2020). We estimated a ternary plot using the R package GGTERN v3.1.0 (Hamilton & Ferry, 2018) to assess whether the increase in the frequency of H-morphs occurred at the cost of only one or both S- and L-morphs.
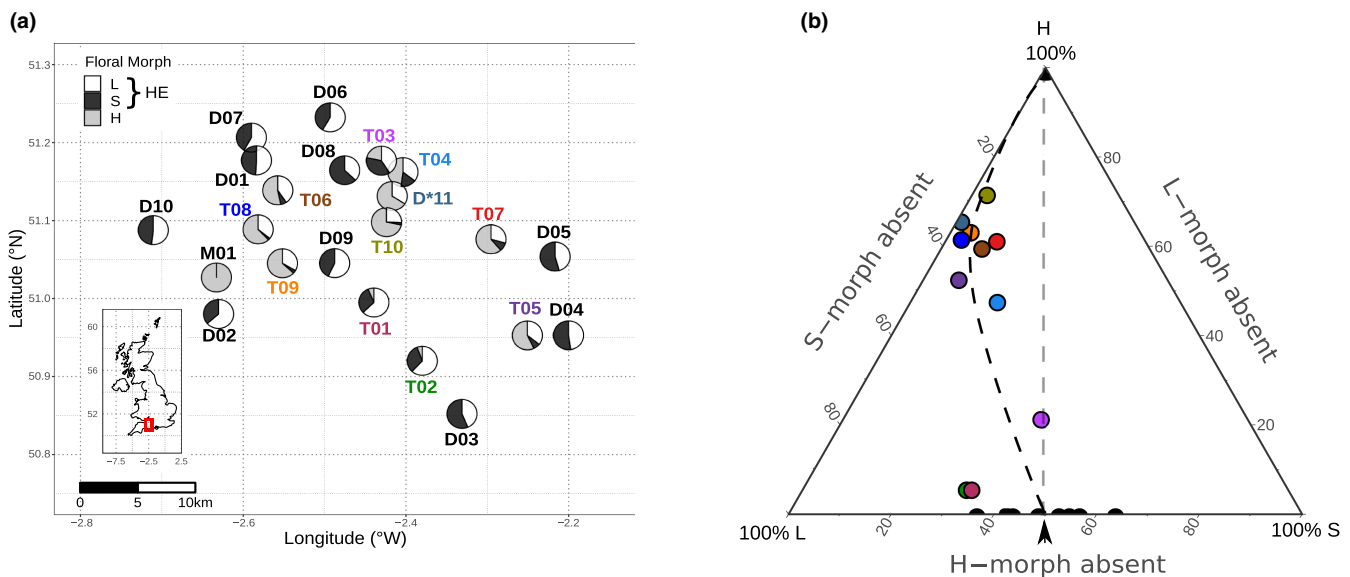
FIGURE 2 Distribution and frequencies of morph types across the 22 sampled populations of *Primula vulgaris* in Somerset, England. (a) Geographic distribution. Pie charts represent intra-population frequencies of long-styled (L; white), short-styled (S; black) and homostylous (H; grey) floral morphs; dimorphic (D) populations consist of L- and S-individuals, except for population D*11, which consists of L- and H-individuals; trimorphic (T) populations consist of L-, S- and H-individuals; the single monomorphic (M) population consists of H-individuals. (b) Ternary plot representing the frequencies of S-, L- and H-individuals sampled from the same 22 populations. Each vertex of the triangle represents complete population monomorphy for each floral morph; each side of the triangle represents different levels of population dimorphism (bottom: H-morph absent, only L- and S-morphs; left side: S-morph absent, only L- and Hmorphs; right side: L-morph absent, only S- and H-morphs); each point inside the triangle represents trimorphic populations; arrow pointing to the base of the triangle represents equal frequencies of S- and Lmorphs (i.e., isoplethy); the black dashed line represents the trajectory from isoplethy in distylous populations to complete homostyly observed in the sampled populations; the leftward shift of the line is caused by the greater reduction of S- than L-individuals; the grey dashed line represents the theoretical trajectory from isoplethy in distylous populations to complete homostyly with equal reduction of S- vs. L-individuals

TABLE 1 Summary data for 22 natural populations of *Primula vulgaris* sampled in Somerset, England, with floral morph frequencies, levels of inbreeding (F_{is}), population-level estimates of selfing rates ($s(BES)$) and mean pairwise genetic differentiation among populations (G'_{st}) inferred using 12 microsatellites. Dimorphic (D) populations consist of long-styled (L) and short-styled (S) morphs, except for population D*11, which consists of L- and homostylous (H) morphs; trimorphic (T) populations consist of L-, S- and H-morphs; the single monomorphic (M) population consists of H-morphs

| Population | Latitude (N), Longitude (E) | Population size | Sample size (morph freq) | Floral Morph | | | Sample size (pop genetics) | F_{is} | $s(BES)$ | G'_{st} |
				L	S	H				
D01	−2.58°, 51.18°	273	100	0.51	0.49	0	30	0.131	0.180	0.06
D02	−2.63°, 50.98°	~300	96	0.64	0.36	0	30	0.066	0.154	0.07
D03	−2.33°, 50.85°	~500	110	0.44	0.56	0	30	0.071	0.131	0.08
D04	−2.20°, 50.95°	65	65	0.48	0.52	0	24	0.143	0.212	0.08
D05	−2.22°, 51.05°	260	105	0.46	0.54	0	30	0.105	0.318	0.07
D06	−2.49°, 51.23°	~350	101	0.58	0.42	0	30	0.009	0.122	0.09
D07	−2.59°, 51.21°	~250	100	0.58	0.42	0	30	0.149	0.348	0.07
D08	−2.48°, 51.16°	150	94	0.37	0.63	0	29	−0.01	0.222	0.07
D09	−2.49°, 51.05°	~300	100	0.57	0.43	0	29	0.002	0.168	0.07
D10	−2.71°, 51.09°	~500	104	0.52	0.48	0	29	−0.085	0.052	0.08
D*11	−2.42°, 51.13°	58	50	0.34	0	0.66	30	0.065	0.218	0.07
T01	−2.44°, 51.00°	94	94	0.63	0.31	0.06	30	0.117	0.295	0.08
T02	−2.38°, 50.92°	~150	108	0.62	0.32	0.06	24	0.124	0.255	0.07
T03	−2.43°, 51.18°	88	88	0.41	0.38	0.22	30	0.052	0.219	0.07
T04	−2.40°, 51.16°	~200	104	0.36	0.16	0.48	31	0.389	0.529	0.10
T05	−2.25°, 50.95°	20	15	0.33	0.07	0.53	17	0.284	0.473	0.10
T06	−2.56°, 51.14°	79	79	0.41	0.06	0.53	28	0.225	0.422	0.07
T07	−2.30°, 51.08°	103	85	0.29	0.09	0.61	30	0.227	0.394	0.08
T08	−2.58°, 51.09°	161	95	0.36	0.02	0.62	30	0.126	0.314	0.15
T09	−2.55°, 51.05°	87	87	0.33	0.03	0.63	31	0.175	0.318	0.10
T10	−2.42°, 51.10°	~150	82	0.26	0.02	0.72	30	0.251	0.367	0.09
M01	−2.63°, 51.03°	19	19	0	0	1	11	0.272	0.461	0.23
							Mean (± SD)	0.131 (± 0.11)	0.281 (± 0.13)	0.088 (± 0.03)

Note: F_{is} varies from −1 (heterozygotes excess) to +1 (homozygotes excess due to inbreeding): values closer to 0 indicate random outcrossing under HWE; population level estimates of selfing rates vary from 0 (maximum outcrossing) to 1 (maximum self-fertilization).

2.3 | Sampling for genetic analyses

In the spring of 2019, we randomly collected leaf tissue from 11 to 30 individuals in each of the 22 natural populations described above for a total of 613 individuals. Leaf tissue was dried in silica gel. We sampled the same number of each floral morph where possible; sampled individuals were at least 1–2 m apart from each other. We additionally collected leaf tissue and 2–3 capsules with seeds from nine individuals (four pins and five thrums) of one dimorphic heterostylous population (D07) and seven individuals from the fully homostylous population (M01). Collected seeds from each individual were sowed in a single pot; seedlings from a single individual formed a family. All pots were placed in water containing gibberellic acid (70 ppm) for 24 h and later placed in growth chambers at 18°C with a 12:12 h light-dark period. The proportion of germinated seeds per family was recorded after eight weeks, and leaf tissue from a total of ten seedlings per family was collected whenever possible and dried in silica gel. DNA of individuals from natural populations and progeny arrays was extracted with a modified CTAB protocol (Doyle & Doyle, 1987) and used for subsequent genetic analyses.

2.4 | Molecular basis of the transition to homostyly

2.4.1 | Data generation

To determine whether the transition to homostyly is associated with a single or multiple molecular changes in CYP^T, we Sanger-sequenced all five CYP^T exons in 38 individuals from different Somerset populations. Specifically, we sampled 27 H-individuals from 10 populations (T03-T10, D*11 and M01) and 11 S-individuals from five populations (T03, T06, T08, T09 and T10), corresponding to 2–3 randomly selected individuals per population (CYP^T is absent in L-individuals). Additionally, we obtained CYP^T sequences from whole genome resequencing data of six S-individuals from populations T04 and T07 (three individuals per population; EMC, unpublished data), for a total of 44 newly sequenced individuals. Finally, we complemented our data set with previously published sequences (CYP^T-S^{LH1} and -S^{LH2}) of two H-individuals from Somerset and Chiltern Hills, respectively (Li et al., 2016). In total, we thus analysed CYP^T sequences from 46 individuals (29 H- and 17 S-morphs).

All five CYP^T exons were amplified and Sanger-sequenced in the 38 newly sampled individuals using previously designed primers (Huu et al., 2016). Additionally, reverse primers for exons 1, 2, and 3 were newly designed to obtain longer sequences (Table S1). Detailed PCR conditions for the amplification of all CYP^T exons are included in Supporting Information. Sanger sequencing was performed in an ABI Prism 3130 genetic analyser (Applied Biosystems). Forward and reverse sequences were visually inspected and aligned with MUSCLE, as implemented in MEGA X (Kumar et al., 2018). Polymorphisms in CYP^T were counted only if mutations showed unambiguous peaks in both forward and reverse chromatograms, had at least one Phred value above 30, and/or were present in more than one individual. In cases where a mutation in CYP^T was detected in a single individual, this individual was sequenced twice to confirm the presence of the polymorphism. All exon sequences were concatenated with MESQUITE v3.61 (Maddison & Maddison, 2019). Library preparation and high-throughput sequencing for whole genome resequencing data were performed by RAPiD GENOMICS using paired-end sequencing (~150 bp sequence reads) in NovaSeq 6000 (Illumina). We used HYBPIPER v1.3.1 (Johnson et al., 2016) with default parameters (except for the coverage-cutoff level for assemblies set to 4) to target and extract the sequence of CYP^T from whole genome resequencing data. To identify synonymous and nonsynonymous mutations, we used the Open Reading Frame (ORF) from the *P. vulgaris* CYP^T sequence deposited in GenBank (KT257665.1; Li et al., 2016).

2.4.2 | Analyses

To determine the extent of CYP^T variation in S-individuals and homostyles, we estimated nucleotide diversity at synonymous (πS) and nonsynonymous sites (πNS) using DNASP v6.12.03 (Rozas et al., 2017). To infer the relationships among CYP^T sequences from H- and S-individuals, we first estimated a haplotype network with the R package PEGAS v0.12 (Paradis, 2010). Such analysis is appropriate because CYP^T is located in the hemizygous, nonrecombining S-locus. In addition to single nucleotide substitutions, we scored each insertion or deletion (i.e., indel) as a character following the "simple indel contig" guidelines by Simmons and Ochoterena (2000). Briefly, each indel of any length was scored as a presence/absence character in every individual. The resulting presence/absence matrix was coded as pseudo-nucleotides using "A" as absence, "T" as presence and "-" as unknown and was concatenated to the CYP^T sequence alignment. Second, we estimated the CYP^T phylogeny using a partitioned maximum likelihood (ML) analysis in RAxML v8.01 (Kozlov et al., 2019) with a GTR-GAMMA substitution model for nucleotide substitutions and a binary (BIN) model for indels. Branch support was estimated with 1000 standard bootstrap resamplings. A publicly available CYP^T sequence from *P. veris* (KX589238; Huu et al., 2016) was used as an outgroup to root the CYP^T tree.

2.5 | Evolutionary consequences of transitions to homostyly

Using microsatellites, we tested whether outcrossing rates from progeny arrays are higher in heterostyles than homostyles. Additionally, we estimated the strength of inbreeding depression at the seed-germination stage. Finally, we tested whether increasing frequencies of homostyles are associated with higher population-level selfing rates, inbreeding coefficients within populations, and genetic differentiation among populations.

2.5.1 | Outcrossing rates from progeny arrays

Outcrossing rates were estimated from nine heterostylous families (stemming from the dimorphic heterostylous population D07) with a mean number of 10 seedlings per family and seven homostylous families (stemming from the fully homostylous population M01) with 7.86 ± 1.06 (mean ± SE) seedlings per family. Both maternal and seedling plants were genotyped with the same microsatellites as in section 2.5.2. Multilocus (t_m) and single-locus (t_s) outcrossing rates and their associated standard errors were estimated with MLTR v.3.4 (Ritland, 2002), using 10,000 bootstrap resamplings of maternal families. We used a one-sided Wilcoxon signed-rank test to assess whether homostylous individuals have a lower outcrossing rate than distylous individuals. Moreover, the germination rate of heterostylous and homostylous families was used as a fitness proxy to estimate the strength of inbreeding depression (δ) as 1 – (W_{HO}/W_{HE}) (Goodwillie et al., 2005), where W_{HO} and W_{HE} are the mean germination rates of homostyles and heterostyles, respectively.

2.5.2 | Population-level estimates of selfing rates

Data generation
We genotyped all 613 individuals collected from the 22 sampled, natural populations using 12 microsatellites previously developed for *P. vulgaris* (Triest et al., 2015). Briefly, the loci were amplified in two multiplex reactions with a Qiagen Multiplex PCR kit. Detailed conditions for the amplification of microsatellites are included in Supporting Information. Fragment size estimations were performed using an ABI Prism 3130xl genetic analyser (Applied Biosystems) with GeneScan 500 LIZ (Applied Biosystems) as internal standard. Alleles were scored using GeneMapper v4.1 (Applied Biosystems). The presence of null alleles was estimated with Micro-checker v2.2.3 (Van Oosterhout et al., 2004).

Analyses
Population-level selfing rates were inferred using model-based Bayesian estimation (BES v0.1.3; Redelings et al., 2015) from microsatellite data of 613 individuals from 22 natural populations (see section 2.3). The main advantage of this method compared to estimates generated by other programs such as RMES (David et al., 2007) is that BES utilizes all loci regardless of whether they are homozygous or heterozygous. Selfing rates and 95% CIs were inferred using 10,000 bootstrap iterations. To test whether higher frequencies of homostyles in natural populations are associated with higher selfing rates, we used generalized additive models where the frequency of homostyles was used as an independent variable and by specifying a beta distribution for the estimates of selfing rates as implemented in the R package GAMLSS v5.1.6 (Rigby et al., 2005).

2.5.3 | Population genetic analyses

Allelic richness (A_r), observed (H_o) and expected heterozygosity (H_e), and inbreeding coefficient (F_{is}) were calculated with the R package DIVERSITY v9.90 (Keenan et al., 2013) from microsatellite data of 613 individuals from 22 populations (see section 2.3). Each locus and population were tested for Hardy-Weinberg equilibrium using the R package GENEPOP v1.1.7 (Rousset, 2008). To assess whether an increase in the frequency of homostyly is associated with an increase in F_{is}, we used a standard linear model with homostyle frequency as the independent variable. Given that population size can influence F_{is}, we included census size as covariable with no interaction among independent variables.

Pairwise genetic differentiation measured as F_{ST} (Weir & Cockerham, 1984) and its normalized counterpart G'_{ST} (Meirmans & Hedrick, 2011) were estimated with the R package DIVERSITY v9.90 (Keenan et al., 2013). To discover whether higher homostyle frequencies increase genetic differentiation, we tested whether pairwise differences in homostyle frequencies between populations are positively correlated with pairwise G'_{ST} values. To test the significance of association

Consequences of shifts to homostyly on mating system and population geneticsbetween these two matrices, we used a Mantel test as implemented in the R package ADEGENET v2.1.3 (Jombart, 2008).

2.6 | Genetic structure and gene flow

To estimate genetic structure among *P. vulgaris* populations, we used INSTRUCT v3.2.09 (Gao et al., 2007), which extends the Bayesian model implemented in Structure (Pritchard et al., 2000) to situations of partial self-fertilization and recurrent inbreeding. We ran the analyses with 100,000 burnin iterations and 20 independent MCMC chains per run from 2 to 22 *k*, as suggested by Gilbert et al. (2012). The most likely number of clusters (*k*) was detected with the *Δk* method (Evanno et al., 2005) implemented in CLUMPAK (Kopelman et al., 2015). Additionally, a Discriminant Analysis of Principal Components (DAPC) was performed with the R package ADEGENET v2.1.1 (Jombart, 2008). To determine whether there is isolation by distance (*IBD*), we performed a Mantel test implemented in adegenet using both F_{ST} and G'_{ST} values of genetic differentiation.

Finally, to determine whether patterns of genetic structure are explained by gene flow, we estimated the size-corrected effective number of migrants per generation (N_m) using the frequency of private alleles with the R package GENEPOP v1.1.7 (Rousset, 2008). Moreover, we used approximate Bayesian computations (ABC), as implemented in POPABC v1.0 (Lopes et al., 2009), to compare models of genetic divergence with and without gene flow using the coalescent framework developed by Nielsen and Wakely (2001) and Hey and Nielsen (2004). Specifically, we compared a scenario of isolation without gene flow (i.e., proportion of migrants per generation,

$m = 0$) with seven scenarios of increasing migration rates ($m = 0.001$, 0.01, 0.1, 0.2, 0.3, 0.4 and 0.5). Priors used for this analysis are specified in Table S2. Each model was simulated with 100,000 iterations and tolerance for the rejection step was set to 0.01. Model selection was based on the posterior probability estimated by categorical regression (Beaumont et al., 2002) using custom R scripts included in the popABC documentation. A major assumption of the models implemented in popABC is that there is no genetic structure within populations. Given that H-individuals may cause within-population structure due to selfing, we decided to use only the ten dimorphic heterostylous populations for popABC analyses because no genetic substructure is expected in populations consisting entirely of outcrossing individuals. No differences in pollen and seed dispersal are known for H-, S- and L-individuals, thus excluding populations with H-individuals from popABC analyses should not affect estimates of migration between populations.

3 | RESULTS

3.1 | Frequency variation of floral morphs in natural populations of *Primula vulgaris*

The frequency of homostyles ranged from 0% to 100% across the 22 sampled populations of *P. vulgaris* in Somerset, UK (Table 1; Figure 2a). Ten populations were dimorphic heterostylous (L- and S-individuals; D01–D10), one was dimorphic with L- and H-individuals (D*11), 10 were trimorphic (L-, S- and H-individuals; T01–T10), and one was monomorphic for the H-morph (M01). Morph ratios in dimorphic heterostylous populations did not deviate from isoplethy except in D02 and D08, which had an excess of L- and S-plants ($G = 7.13$, $p = 0.007$ and $G = 6.11$, $p = 0.01$, respectively; Table S3). In trimorphic populations, L-individuals were significantly more frequent than S-individuals except in T03, where S- and L-morph frequencies did not differ significantly ($G = 0.13$, $p = 0.71$; Table S3). Congruently, the ternary plot suggested a skewed trajectory towards a reduction of S-individuals when H-individuals were present (Figure 2b).

3.2 | Molecular basis of the transition to homostyly

The CYP^T data matrix comprised ~97% of the 1587 bp of five exons from a total of 44 individuals. We detected a total of seven CYP^T alleles (CYP^T-1 to -7; Figure 3a). All 17 S-individuals shared the same allele (CYP^T-1), identical to the functional copy of CYP^T previously reported in *P. vulgaris* (GenBank: KT257665.1; Li et al., 2016). Unexpectedly, the putatively functional CYP^T-1 allele also occurred in six H-individuals. In the 21 remaining H-individuals, we detected six additional CYP^T alleles (CYP^T-2 to -7), comprising a single synonymous mutation, four nonsynonymous mutations, and two deletions. Specifically, CYP^T-4 contained two private mutations (one synonymous and one nonsynonymous) and five CYP^T

alleles contained a single private mutation (three different nonsynonymous mutations in CYP^T-2, -3, -7 and two different deletions in CYP^T-5 and -6). The nonsynonymous mutation in CYP^T-2 introduced a premature stop codon in exon 2, whereas a 31 bp deletion in exon 5 of CYP^T-5 and an 8 bp deletion in exon 1 of CYP^T-6 introduced frameshifts leading to early stop codons. Thus, CYP^T-2, 5 and 6, found in 15 out of 27 homostyles, contained inactivating mutations. Moreover, a mutation in exon 5 of CYP^T-3 changed a nonpolar (phenylalanine) to a polar (serine) amino acid, and a mutation in exon 3 of CYP^T-4 changed a nonpolar (asparagine) to a polar (serine) amino acid. Finally, a mutation in exon 4 of CYP^T-7 changed an amphipathic (i.e., both polar and nonpolar: arginine) to a polar (positively charged; histidine) amino acid. Thus, CYP^T-3, 4, and 7, found in six out of 27 homostyles, contained mutations likely to affect protein function. To summarize, the six CYP^T alleles (CYP^T-2 to -7) newly detected in our study, differing from the ones previously reported in two long-homostyles of *P. vulgaris* (CYP^T-S^{LH1} and -S^{LH2}), are all likely to compromise CYP^T function. Overall, nucleotide diversity in homostyles was $\pi S = 0.0005$ and $\pi NS = 0.001$, while S-morph individuals showed no genetic variation in CYP^T (both πS and πNS were zero).

In the haplotype network, CYP^T-1, the known functional allele, occupied a central node and was connected by a single mutational step to all other alleles except for CYP^T-5, which differed by two mutational steps (Figure 3b). Concordantly, our phylogenetic analyses recovered lack of structure in the backbone of the inferred tree, but individuals sharing the same CYP^T allele formed well-supported clades (Figure 4). Specifically, CYP^T-2 was shared among 12 homostyles from five different populations (T03, T04, T07, T10, and D*11; Figure 4); CYP^T-3 was unique to three individuals of population T05, and CYP^T-5 was unique to two individuals of population T09. Finally, two CYP^T alleles were detected in each of three populations: CYP^T-4 and CYP^T-6 were found in population T06, CYP^T-2 and CYP^T-4 in population T10, and CYP^T-2 and CYP^T-7 in population D*11 (Figure 4).

3.3 | Evolutionary consequences of transitions to homostyly

3.3.1 | Mating system in homostyles

Outcrossing rates estimated from progeny arrays were more variable and significantly lower in homostylous than heterostylous families (single locus [t_s]: mean $\pm$ SE; 0.24 ± 0.10 [homostylous] and 0.83 ± 0.05 [heterostylous]; multilocus [t_m]: 0.14 ± 0.06 and 0.98 ± 0.02, respectively; Wilcoxon signed-rank test $W = 63$, $p = 0.002$ and 0.001; Figure S1). As expected, the frequency of homostyles within populations was positively associated with population-level estimates of selfing rate (*pseudo-R^2* = 0.447, $p < .001$; Table 1, Figure 5a). Seed germination rate was significantly lower for homostyles than heterostyles (0.11 ± 0.03 vs. 0.26 ± 0.05; $W = 308$, $p = 0.04$) resulting in an estimated inbreeding depression (δ) of 0.58.
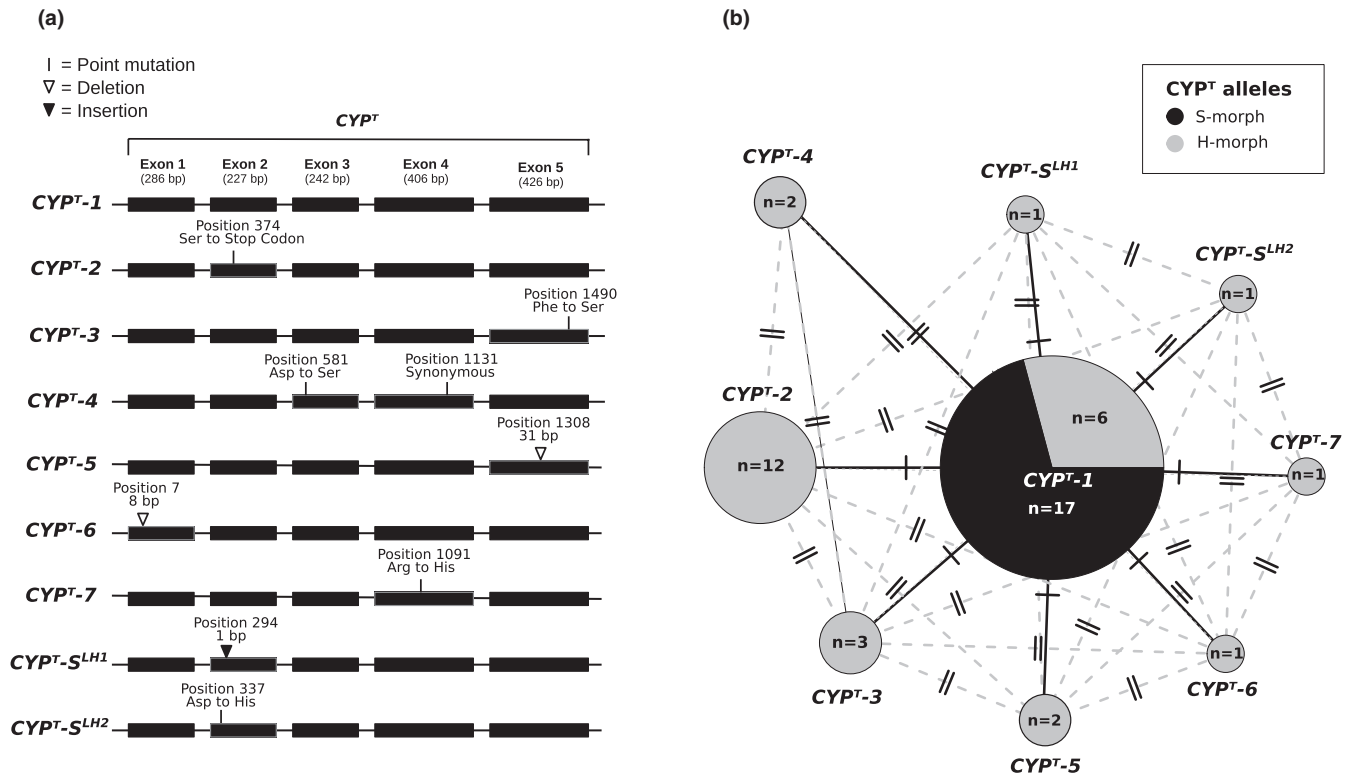
FIGURE 3 Variation of CYP^T sequences in 44 individuals of *Primula vulgaris* (17 short-styled and 27 homostylous individuals) from ten natural populations included in this study, plus two CYP^T alleles from two homostylous individuals previously reported as $CYPT$-S^{LH1} and -S^{LH2} by Li et al. (2016); CYP^T is a hemizygous gene comprising 1587 bps over five exons (base-pair lengths of each exon are indicated in parentheses). (a) Graphical representation of the nine CYP^T alleles with different types of mutations in CYP^T 2–9: vertical lines, white triangles, and black triangles represent point mutations, deletions, and insertions, respectively, with positions of each mutation reported above each exon; one synonymous and five nonsynonymous mutations, two deletions, and one insertion were found. (b) Haplotype network of the nine CYP^T alleles: each circle (i.e., node in the network) represents a different CYP^T allele (CYP^T 1–7 and CYP^T-S^{LH1} and -S^{LH2}, listed next to each circle), with circle size proportional to the number of homostylous (grey) and short-styled (black) individuals carrying that allele (reported inside the circle); nodes are connected by solid lines representing the lowest number of mutational steps between alleles, in this case always a single mutational step indicated by a single dash across the line, except for CYP^T-1 and CYP^T-4 connected by two mutational steps; grey dashed lines represent alternative connections among haplotypes with two mutational steps between alleles, indicated by a double dash across the line

## 3.3.2 | Patterns of genetic diversity

Eighteen out of 22 populations (except for D03, D06, D08 and D10) had significant deviations from Hardy-Weinberg equilibrium due to a heterozygote deficit (Table S4). Population-level estimates of allelic richness (A_r) varied from 1.58 to 3.78 (mean ± SE, 3.19 ± 0.09), expected heterozygosity (H_e) from 0.14 to 0.53 (0.42 ± 0.01), and observed heterozygosity (H_o) from 0.11 to 0.46 (0.37 ± 0.02). Inbreeding coefficient (F_{is}) ranged from −0.08 to 0.39 among populations (0.13 ± 0.02; Table 1) and had a significant positive relationship with the frequency of homostyly ($R^2 = .405$, $p = 0.03$), but not with population size ($p = 0.27$; Figure 5b).

Global values of genetic differentiation among populations were 0.083 (95% CI: 0.074–0.093) and 0.164 (95% CI: 0.147–0.181) for F_{ST} and G'_{ST}, respectively (Table S4), and were strongly correlated with each other (*pseudo-R^2* = 0.974, $p = 0.001$). Pairwise population values of F_{ST} and G'_{ST} are shown in Table S5. As expected, pairwise homostyle frequencies showed a significant

positive relationship with pairwise G'_{st} (Mantel test: $R^2 = 0.481$, $p < 0.001$).

3.4 | Genetic structure and gene flow

InStruct analyses recovered two clusters (Figure 6a) and showed that T08 differs from the remaining 21 populations. Identical results were found with Structure (results not shown). Accordingly, the DAPC analyses separated T08 from the rest of the populations (Figure 6b). Furthermore, we found no evidence of isolation by distance (Mantel test, $p = 0.457$). Together, these results suggest that genetic differentiation among populations is moderate, corroborating F_{ST} results presented above. Using the frequency of private alleles, we estimated that the effective number of migrants per generation (N_m) was 1.96. Comparison of different demographic scenarios lent higher support to models with interpopulation gene flow than to the isolation model. Among the models with varying
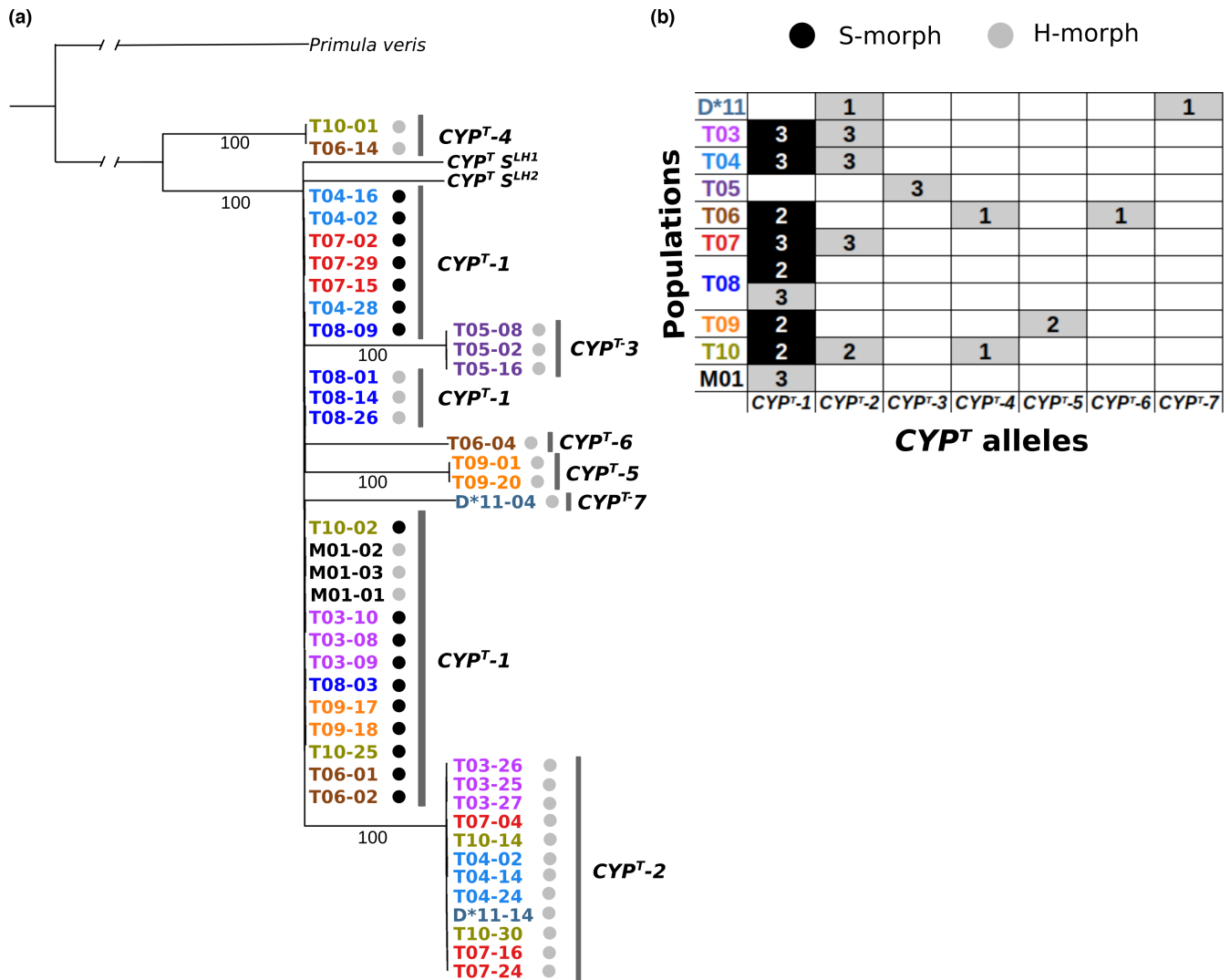
FIGURE 4 Clustering of CYP^T sequences. (a) Maximum likelihood tree of CYP^T sequences from 44 individuals of *Primula vulgaris* (17 short-styled and 27 homostylous) generated for this study, plus CYP^T sequences from two homostylous individuals previously reported as CYP^T-S^{LH1} and -S^{LH2} (Li et al., 2016), and a CYP^T sequence from an S-individual of *Primula veris* (GenBank: KX589238) as outgroup. Each accession at the tips is labeled from left to right with population number, individual number, and floral morph type (S and H represented as black and grey circles, respectively). Accessions from the same population have the same color (see also Figure 2). Bootstrap support values ≥80 are indicated below to the branches. (b) Counts of CYP^T alleles (x axis) per population (y axis) found in S- and H-individuals (represented as black and grey boxes, respectively)

migration rates (*m*), the one with *m* = 0.2 had the highest posterior probability, indicating moderate gene flow (Figure S2).

4 | DISCUSSION

The present analyses of shifts from heterostyly to homostyly provide novel insights into the molecular basis of intraspecific mating system transitions and their evolutionary consequences. Previous studies suggesting independent origins of homostyly used molecular markers not associated with the genetic control of heterostyly (Ness et al., 2010; Shao et al., 2019; Yuan et al., 2017; Zhou et al., 2017). By leveraging new knowledge on the molecular basis of heterostyly

and extensive sampling of populations with different frequencies of homostyles, this is the first study to demonstrate that heterostyly has been repeatedly lost probably via multiple, presumably loss-of-function mutations of the S-locus gene controlling stigma position (CYP^T). Additionally, we show that the transition to homostyly promotes a change of mating system from outcrossing to selfing, with ensuing population genetic consequences. Furthermore, we discuss the potential role of gene flow in the spreading of homostyly. Finally, we propose that inbreeding depression and gene flow might explain why homostyly, despite its apparent advantages, has not become fixed in *P. vulgaris*. Thus, the present contribution provides a benchmark for similar studies on the loss of heterostyly across angiosperms.
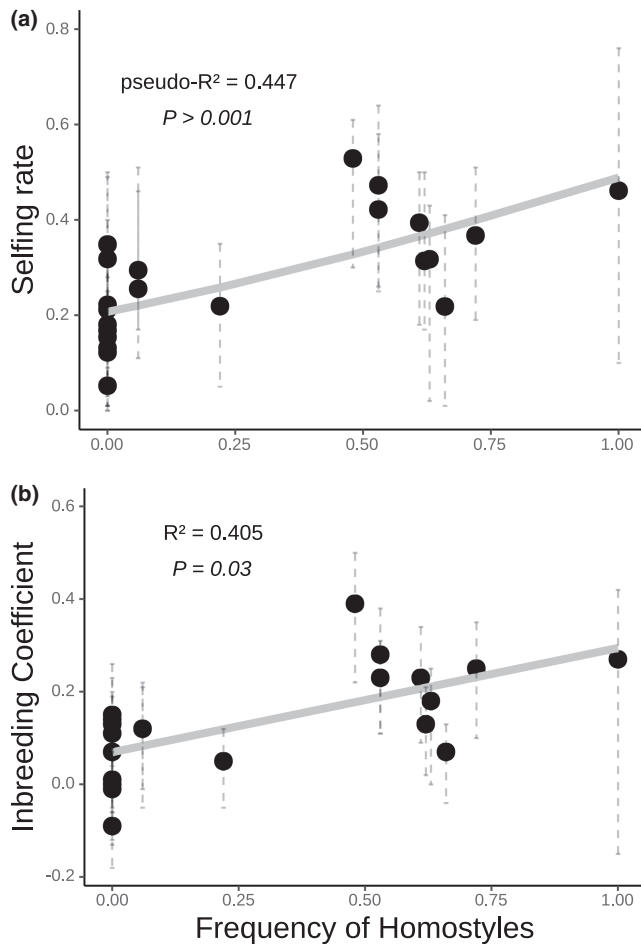
FIGURE 5 Correlations between population genetic parameters (*y* axis) and frequency of homostyles (*x* axis) in 22 natural populations of *Primula vulgaris* inferred from 12 microsatellites: (a) Population-level estimates ±95% CI of selfing rates based on multilocus linkage disequilibrium (See Methods) and (b) Inbreeding coefficient ±95% CI

## 4.1 | Homostyles replace short-styled individuals within populations

The variation in the frequency of floral morphs among populations provides an opportunity to elucidate how homostylous phenotypes originate and spread. According to previous evolutionary models, the origin of homostyles from S-individuals should cause an initial reduction of the latter and the concomitant increase of the former (Crosby, 1949, 1958). When homostyles first arise, their pollen can be used for both self- and cross-fertilization of L-individuals, but not for fertilization of S-individuals, due to lack of spatial reciprocity between high anthers of homostyles and low stigma of S-individuals and incompatibility between homostyle pollen and S-morph stigmas (Crosby, 1949; Figure 1). Conversely, homostyles can theoretically be fertilized by compatible pollen of S-morphs, but this cross is thought to occur rarely in nature due to stigma clogging with self-pollen in homostyles (Crosby, 1958; but see Bodmer, 1958). Hence, homostyles have a reproductive advantage over L- and S-morphs stemming from their ability to fertilize both themselves and the L-morphs

(Richards, 2003). Finally, self-fertilization of homostyles (Figure S3a and b) and crosses from homostyles to L-individuals produce homostylous and L-progeny, but no S-progeny (Figure S3c and d).

Congruently with the predictions of Crosby (1949) and findings from previous studies (Curtis & Curtis, 1985), our results demonstrate that, as homostyles increase, more S- than L-individuals are lost within populations (Table 1 and Figure 2a,b). After 72 and 36 years, respectively, since the observations of Crosby (1949) and Curtis and Curtis (1985), we also found that homostyly has not become fixed in Somerset populations. Furthermore, the fact that the self-incompatible L-individuals are maintained at low frequencies, even in the absence of S-individuals, suggests that they are occasionally fertilized by homostyles. Experiments aimed at quantifying pollen delivery and receipt among the three floral morphs would help to further clarify the selective regimes favouring the increase of homostyles. Overall, our results are consistent with the hypothesis that long-homostyles originate from S-individuals and explain why the latter occur at lower frequencies than L-individuals across populations with all three morphs.

4.2 | Multiple losses of heterostyly within species are associated with different molecular changes in the *CYP^T* gene

Recent advancements in genomics have enabled the identification of loci responsible for repeated losses of phenotypic traits (Albalat & Cañestro, 2016; Sharma et al., 2018). For instance, multiple loss-of-function mutations were detected in *SCR* or *SCR*-like genes corresponding to independent losses of self-incompatibility in *Arabidopsis thaliana* and *Laevenworthia alabamica* (Chantha et al., 2013; Shimizu et al., 2008). Similarly, recent experimental studies established that *CYP^T* disruption causes a shift to homostyly (Huu et al., 2016; Kappel et al., 2017). Accordingly, we found private alleles with nonsynonymous mutations or indels (*CYP^T-2* to -7) exclusively in homostyles of *P. vulgaris* (Figures 3a,b and 4), whereas all sampled short-styled individuals shared the same *CYP^T* allele (*CYP^T-1*), shown to control low stigma position and probably encoding a functional protein involved in the mechanism of self-incompatibility (Kappel et al., 2017). The mutations in *CYP^T-2* to -7 alleles caused either changes in the hydrophobic versus hydrophilic properties of the amino acid side chain (*CYP^T-3, -4, -7*) or premature stop codons (*CYP^T-2, -5* and *-6*), both types of mutation presumably inducing reduction or loss of function, respectively (Zhang, 2000). Notably, the alleles found in our study differ from the ones previously reported in two long-homostyles of *P. vulgaris* (*CYP^T-S^{LH1}* in Somerset and *CYP^T-S^{LH2}* in Chiltern Hills; Li et al., 2016). Thus, a total of eight different *CYP^T* alleles with putatively compromised protein function have been identified so far in homostyles of *P. vulgaris*. Future work should focus on elucidating how the detected nonsynonymous mutations affect *CYP^T* function. Finally, the groupings of the *CYP^T* haplotype network were congruent with those of the phylogeny (Figures 3b and 4), as expected given the lack of recombination in the S-locus. Overall, our results

(a)

$k = 2$

(b)
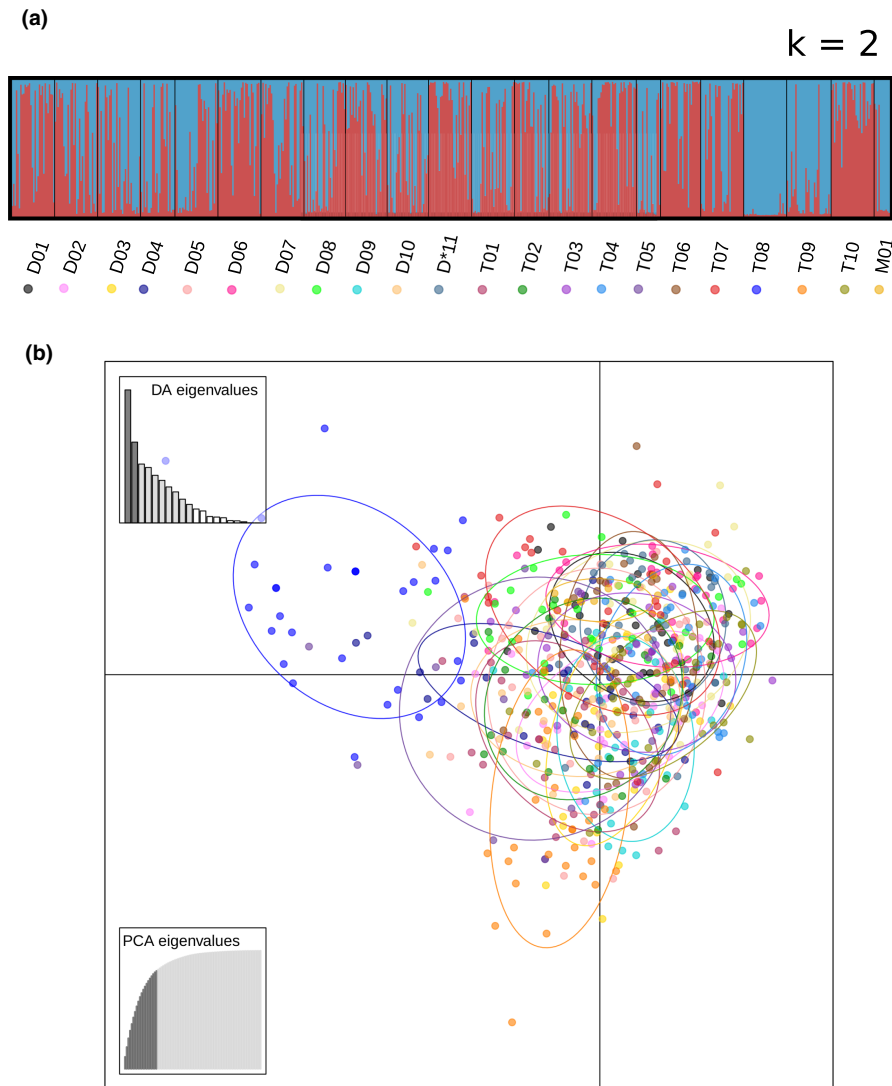
FIGURE 6 Patterns of genetic differentiation for the 22 natural populations of *Primula vulgaris*. (a) Results from InStruct analysis of microsatellites assigning all populations to two genetic clusters ($k = 2$): Optimal number of clusters was selected according to Evanno et al. (2005); (b) Plot of the two first PCs from Discriminant Analysis of Principal Components (DAPC) separating population T08 (blue circles) from the rest of the populations

show that a diversity of changes in the same gene is associated with multiple, independent origins of homostyly within species.

The occurrence of multiple, independent CYP^T mutations leading to homostyly in such a small geographic region is a puzzling result, and different evolutionary processes could be invoked to explain their origins. For instance, selective pressures favouring homostyly, such as lack of pollinators or mates, could have driven the independent origins of homostyles in different refugia during the Last Glacial Maximum and the subsequent colonization of the British Isles after glacial retreat. Under this scenario, homostyles of *P. vulgaris* should have a wider distribution in England and should also occur in southern glacial refugia such as the Italian and Iberian peninsulas (Hewitt, 1999), but this not the case. We thus suggest that the origin of the Somerset homostylous alleles is more recent, possibly explained by a history of fragmentation caused by land-use changes that started c. 2000 years ago (Kaplan et al., 2009), especially in southern England. Our microsatellite data do not enable us to determine whether mutations in CYP^T accumulated before or after the fragmentation of a hypothetical ancestral population. Organellar and whole genome sequences

of multiple heterostyles and homostyles from within and outside England are necessary to resolve the phylogeographic and demographic history of *P. vulgaris* homostyles.

Loss of phenotypic traits can be caused by mutations not only in coding regions of specific genes but also in their promoter regions or by structural rearrangements. For instance, some individuals of self-compatible *A. thaliana* were characterized by mutations in *cis*-regulatory regions, inversions, or splicing variants, causing loss of function in *SCR* or *SRK* genes controlling self-incompatibility (Dwyer et al., 2013; Shimizu et al., 2008; Shimizu & Tsuchimatsu, 2015). In *P. vulgaris*, a surprising result is that six out of 27 homostyles have the same CYP^T allele as the S-individuals (CYP^T-1; Figures 3a,b and 4). These six homostyles occur in two geographically close populations (M01 and T08; Figure 2a). The occurrence of a CYP^T allele with no mutations in homostyles enables us to suggest for the first time that additional mechanisms might be responsible for the inactivation of CYP^T, including mutations in its promoter region or structural rearrangements that cannot be detected via exon sequencing. In general, this result underscores that mutations outside genes controlling specific traits might represent essential sources for repeated trait losses.

The interpretation of CYP^T sequencing results presented above is compatible with multiple origins of homostyly in *P. vulgaris*, but alternative scenarios involving a single origin should be considered. For example, as discussed above, the initial homostylous phenotype could have arisen via mutations in the promoter region, structural rearrangements, or changes in the expression of CYP^T but not in the CYP^T coding region *per se*, as in the six homostyles carrying CYP^T-1, the allele also found in S-individuals. Such homostyle could have originated once in a large ancestral population that later became fragmented and/or migrated to neighboring populations, gaining non-synonymous mutations and indels in CYP^T after the gene had already been rendered functionless. However, in the scenario of a single origin of homostyly triggered by a mutation outside CYP^T, we would expect homostyles carrying CYP^T-1 to co-occur in the same population with other homostyles carrying mutated CYP^T alleles (i.e., CYP^T-2 to -7), but our findings do not match this prediction (Figure 4a,b). Thus, the most parsimonious interpretation of the haplotype network and phylogeny inferred from CYP^T sequences (Figures 3 and 4) is congruent with multiple origins of homostyly via independent mutations at CYP^T. Furthermore, approximately 200 km separate the seven CYP^T alleles sequenced in Somerset homostyles (CYP^T-2 to -7 in the present study, plus CYP^T-S^{LH1} from Li et al., 2016) from the CYP^T-S^{LH2} allele sequenced in a homostyle from the Chiltern Hills (Li et al., 2016). Given the limited dispersal abilities reported for *P. vulgaris* (Cahalan & Gliddon, 1985; Van Geert et al., 2010), this finding further supports independent origins of homostyly. Future transcriptomic or RT-PCR analyses of homostyles with different CYP^T alleles aimed at testing the expression of both mutated and unmutated CYP^T should further elucidate the origins of homostyly in *P. vulgaris*.

4.3 | Consequences of shifts to homostyly on mating system and population genetics

Changes in floral morphology can have profound effects on the mating system (Opedal, 2018). Specifically, the reduction of anther and stigma separation and loss of self-incompatibility of homostylous morphs should increase selfing compared to heterostyles. Evidence of this mating system transition in homostyles was reported for *Eichhornia* (Husband & Barrett, 1993), *Turnera* (Belaoussoff & Shore, 1995), *Amsinckia* (Schoen et al., 1997), and some *Primula* species (Yuan et al., 2017; Zhong et al., 2019; Zhou et al., 2017). In *P. vulgaris*, previous estimates of outcrossing rates in homostyles ranged broadly from very low (0.05) to quite high (0.80; Bodmer, 1958, 1984; Crosby, 1958; Piper et al., 1986).

Our results from progeny arrays indicate that outcrossing rates in homostyles are significantly lower than in heterostyles ($t_m = 0.14$ vs. 1.0, respectively), as expected. Such outcrossing rates were estimated based exclusively on germinated seeds. However, ungerminated seeds were probably those with higher homozygosity, hence higher inbreeding depression, thus their inclusion would further decrease estimates of outcrossing rates. Also as expected, the frequency of homostyles was associated with an increase in population-level selfing rates (Figure 5a). Nevertheless, the fully homostylous population (M01) had lower selfing rates than expected, possibly caused by increased error rates due to its small sample size (Redelings et al., 2015). The estimated outcrossing rates of homostyles firmly place them in the category of selfers (i.e., $t_m \leq 0.2$; sensu Goodwillie et al., 2005; Schemske & Lande, 1985). Overall, our findings clarify previous conflicting results about the effects of homostyly on the mating system of *P. vulgaris*, confirming that transitions to homostyly increase selfing.

Mating systems shape how genetic diversity is partitioned within and among populations (Barrett, 2010; Wright et al., 2013). In concordance with theoretical expectations, we found that intrapopulation homostyle frequency significantly increased inbreeding coefficients (Figure 5b). These results conform with the reduction of genetic diversity associated with transitions from heterostyly to homostyly documented in other systems (Husband & Barrett, 1993; Ness et al., 2010; Yuan et al., 2017; Zhong et al., 2019; Zhou et al., 2017). Moreover, our population genetic results indicate that increased frequency of homostyles in populations increases homozygosity, thus lowering intrapopulation genetic diversity and elevating inter-population genetic differentiation. Altogether, our findings confirm the population genetic consequences predicted for transitions to selfing.

4.4 | Does homostyly spread among populations and why has it not become fixed in *Primula vulgaris*?

Gene flow plays a central role in spreading advantageous mutations (Morjan & Rieseberg, 2004; Ralph & Coop, 2010) and could favour the migration of homostylous alleles among populations. In *P. vulgaris*, it has been proposed that homostyles could have migrated from initial places of origin to neighbouring populations through pollen or seed, the former being more likely (Crosby, 1949, 1960). Previous studies estimated that dispersal is restricted to a maximum of a few hundred metres from the parental plant in *P. vulgaris* (Cahalan & Gliddon, 1985). However, occasional pollen flow over 1–3 kilometers has been reported (Van Geert et al., 2010). Indeed, our population genetic results support a model of moderate gene flow among populations ($Nm = 1.96$; Figure S2) consistent with moderate genetic differentiation among populations ($F_{ST} = 0.083$; Figure 6a,b), the latter being lower than previous estimates for *P. vulgaris* ($F_{ST} = 0.086$–0.508; Van Geert, Van Rossum, & Triest, 2015). Furthermore, we did not detect isolation by distance (Mantel test, $p = 0.457$), further favouring the hypothesis of gene flow. Moreover, the geographic distribution of CYP^T alleles in our study indicates that 12 homostylous individuals from five populations (D*11, T03, T04, T07 and T10) separated by 3–18 km shared the same CYP^T allele (CYP^T-2; Figure 4), suggesting that some homostyles could have migrated between neighboring populations. To summarize, Crosby (1949, 1958) proposed a single origin of homostyles followed by their spreading via migration. Our results, while supporting multiple

origins of homostyly, imply that it can also spread via gene flow, providing partial corroboration of Crosby's hypothesis.

A crucial question remains as to why homostyles have not become fixed in *P. vulgaris* despite automatic selfing advantage (Fisher, 1941; Fisher, 1949) and reproductive assurance (Piper et al., 1986). In fact, the opposite has been found, for previous studies recorded the decrease of homostyles in revisited populations (Curtis & Curtis, 1985). Crosby (1949) proposed that, in addition to crosses between homostyles and L-individuals (Figure 1 and Figure S3), maintenance of L-individuals could be explained by the reduced fitness of homostyles that are homozygous for the S-locus, probably stemming from deleterious effects of dominant homozygosity in S-locus genes (see Figure S4 for patterns of inheritance of CYP^T and the S-locus). Evidence of such negative effects comes from crossing experiments using self-compatible S-morphs of *Primula sinensis* (Mather & de Winton, 1941) and *P. oreodoxa* (Yuan et al., 2018). However, whether homostyles with one or two copies of the S-locus have different fitness in *P. vulgaris* has not been experimentally tested.

In addition to Crosby's hypothesis, the maintenance of heterostyles could be explained by inbreeding depression (δ) slowing the increase of homostyles, and by gene flow reintroducing heterostyles into populations. A previous model suggested that, when $\delta < 0.5$, homostyles should replace heterostyles (Charlesworth & Charlesworth, 1979). However, homostyles could invade even with higher inbreeding depression if their higher reproductive assurance is considered, as proposed in a general model for the invasion of selfing (Lande & Schemske, 1985). Our estimate of inbreeding depression in *P. vulgaris* at seed germination stage was $\delta = 0.58$ and could be even higher, for it could also act at later life-cycle stages. Thus, our results suggest that homostyles can invade heterostylous populations even if $\delta \geq 0.5$, as expected if homostyly provides reproductive assurance. While theory proposes that increased selfing should purge inbreeding depression over time, facilitating the fixation of selfers (Lande & Schemske, 1985), recessive deleterious alleles could be reintroduced into populations via gene flow, thus slowing the purging process. Moreover, migration from neighbouring populations could also reintroduce heterostyles in populations. Indeed, our analyses estimated moderate levels of migration among populations of *P. vulgaris* ($Nm = 1.96$ and $m = 0.2$; Figure S2), implying that gene flow might play a crucial role in the reintroduction of both genetic load and heterostyles, thus precluding the fixation of homostyles in populations despite their reproductive advantages.

CONFLICT OF INTEREST

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

AUTHOR CONTRIBUTIONS

Emiliano Mora-Carrera, Barbara Keller, Jurriaan M. de Vos, Peter Szövényi and Elena Conti designed the research. Emiliano Mora-Carrera collected the data. Emiliano Mora-Carrera, Rebecca L. Stubbs and Barbara Keller performed laboratory work. Emiliano Mora-Carrera, Rebecca L. Stubbs, Étienne Léveillé-Bourret and Barbara Keller analysed the data. Emiliano Mora-Carrera and Elena Conti wrote the manuscript and all authors provided feedback.

DATA AVAILABILITY STATEMENT

Microsatellite genotypes from the 613 individuals from the 22 natural populations analysed in this study, DNA alignment for CYP^T haplotype network and phylogeny and scripts for analyses and figures in R format have been made available in Dryad (https://doi.org/10.5061/dryad.tdz08kq15).

ORCID

Emiliano Mora-Carrera https://orcid.org/0000-0001-8237-4265

REFERENCES

Albalat, R., & Cañestro, C. (2016). Evolution by gene loss. *Nature Reviews Genetics*, *17*, 379–391. https://doi.org/10.1038/nrg.2016.39

Barmentlo, S. H., Meirmans, P. G., Luijten, S. H., Triest, L., & Oostermeijer, J. G. B. (2017). Outbreeding depression and breeding system evolution in small, remnant populations of *Primula vulgaris*: consequences for genetic rescue. *Conservation Genetics*, *19*, 545–554. https://doi.org/10.1007/s10592-017-1031-x

Barrett, S. C. H. (2010). Understanding plant reproductive diversity. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *365*(1537), 99–109. https://doi.org/10.1098/rstb.2009.0199

Barrett, S. C. H. (2019). 'A most complex marriage arrangement': Recent advances on heterostyly and unresolved questions. *New Phytologist*, *224*(3), 1051–1067. https://doi.org/10.1111/nph.16026

Beaumont, M. A., Zhang, W., & Balding, D. J. (2002). Approximate Bayesian computation in population genetics. *Genetics*, *162*(162), 2025–2035. https://doi.org/10.1111/j.1937-2817.2010.tb01236.x

Belaoussoff, S., & Shore, J. S. (1995). Floral correlates and fitness consequences of mating-system variation in *Turnera ulmifolia*. *Evolution*, *49*(3), 545–556. https://doi.org/10.2307/2410278

Bodmer, W. F. (1958). Natural crossing between homostyle plants of *Primula vulgaris*. *Heredity*, *12*(3), 363–370. https://doi.org/10.1038/hdy.1958.35

Bodmer, W. F. (1984). Sex and generations of primroses. *Nature*, *310*, 731. https://doi.org/10.1038/310731a0

Boyd, M., Silvertown, J., & Tucker, C. (1990). Population ecology of heterostyle and homostyle *Primula vulgaris*: Growth, survival and

reproduction in field populations. *Journal of Ecology*, 78(3), 799–813. https://doi.org/10.2307/2260900

Busch, J. W., Joly, S., & Schoen, D. J. (2011). Demographic signatures accompanying the evolution of selfing in Leavenworthia alabamica. *Molecular Biology and Evolution*, 28(5), 1717–1729. https://doi.org/10.1093/molbev/msq352

Busch, J. W., & Delph, L. F. (2012). The relative importance of reproductive assurance and automatic selection as hypotheses for the evolution of self-fertilization. *Annals of Botany*, 109(3), 553–562. https://doi.org/10.1093/aob/mcr219

Cahalan, C. M., & Gliddon, C. (1985). Genetic neighbourhood sizes in *Primula vulgaris*. *Heredity*, 54, 65–70. https://doi.org/10.1038/hdy.1985.9

Carlson, M. L., Gisler, S. D., & 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(1), 39–47.

Chantha, S. C., Herman, A. C., Platts, A. E., Vekemans, X., & Schoen, D. J. (2013). Secondary evolution of a self-incompatibility locus in the Brassicaceae genus *Leavenworthia*. *PLoS Biology*, 11(5), e1001560. https://doi.org/10.1371/journal.pbio.1001560

Charlesworth, B., & Charlesworth, D. (1979). The maintenance and breakdown of distyly. *The American Naturalist*, 114(4), 499–513. https://doi.org/10.1086/283497

Cocker, J. M., Webster, M. A., Li, J., Wright, J., Kaithakottil, G., Swarbreck, D., & Gilmartin, P. M. (2015). Oakleaf: An S locus-linked mutation of *Primula vulgaris* that affects leaf and flower development. *New Phytologist*, 208(1), 149–161. https://doi.org/10.1111/nph.13370

Cocker, J. M., Wright, J., Li, J., Swarbreck, D., Dyer, S., Caccamo, M., & Gilmartin, P. M. (2018). *Primula vulgaris* (primrose) genome assembly, annotation and gene expression, with comparative genomics on the heterostyly supergene. *Scientific Reports*, 8(1), 17942. https://doi.org/10.1038/s41598-018-36304-4

Costa, J., Torices, R., & Barrett, S. C. H. (2019). Evolutionary history of the buildup and breakdown of the heterostylous syndrome in Plumbaginaceae. *New Phytologist*, 224(3), 1278–1289. https://doi.org/10.1111/nph.15768

Crosby, J. L. (1940). High proportions of homostyle plants in populations of *Primula vulgaris*. *Nature*, 145, 672–673. https://doi.org/10.1038/145672c0

Crosby, J. L. (1949). Selection of an unfavourable gene-complex. *Evolution*, 3(3), 212–230. https://doi.org/10.1111/j.1558-5646.1949.tb00022.x

Crosby, J. L. (1958). Outcrossing on homostyle primroses. *Heredity*, 13, 127–131. https://doi.org/10.1038/hdy.1959.9

Crosby, J. L. (1960). The use of electronic computation in the study of random fluctuations in rapidly evolving populations. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 242(697), 551–573. https://doi.org/10.1098/rstb.1960.0002

Curtis, J., & Curtis, C. F. (1985). Homostyle primroses re-visited. I. Variation in time and space. *Heredity*, 54, 227–234. https://doi.org/10.1038/hdy.1985.30

Cutter, A. D. (2019). Reproductive transitions in plants and animals: Selfing syndrome, sexual selection and speciation. *New Phytologist*, 224, 1080–1094. https://doi.org/10.1111/nph.16075

Darwin, C. (1876). *The effects of cross and self fertilization in the vegetable kingdom*. John Murray.

Darwin, C. (1877). *The different forms of flowers on plants of the same species*. 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(12), 2474–2487. https://doi.org/10.1111/j.1365-294X.2007.03330.x

de Vos, J. M., Hughes, C. E., Schneeweiss, G. M., Moore, B. R., & Conti, E. (2014). Heterostyly accelerates diversification *via* reduced extinction in Darwin's primroses. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 281(1784), 20140075. https://doi.org/10.1098/rspb.2014.0075

de Vos, J. M., Keller, B., Isham, S. T., Kelso, S., & Conti, E. (2012). Reproductive implications of herkogamy in homostylous primroses: Variation during anthesis and reproductive assurance in alpine environments. *Functional Ecology*, 26(4), 854–865. https://doi.org/10.1111/j.1365-2435.2012.02016.x

de Vos, J. M., Wüest, R. O., & Conti, E. (2014). Small and ugly? Phylogenetic analyses of the "selfing syndrome" reveal complex evolutionary fates of monomorphic primrose flowers. *Evolution*, 68(4), 1042–1057. https://doi.org/10.1111/evo.12331

Doyle, J., & Doyle, J. (1987). A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochemical Bulletin, 19*, 11–15.

Dulberger, R. (1992). Floral polymorphism and their functional significance in the heterostylous syndrome. In S. C. H. Barrett (Ed.), *Evolution and function of heterostyly* (pp. 41–84). Springer-Verlag.

Dwyer, K. G., Berger, M. T., Ahmed, R., Hritzo, M. K., Mcculloch, A. A., Price, M. J., & Nasrallah, M. E. (2013). Molecular characterization and evolution of self-incompatibility genes in *Arabidopsis thaliana*: The case of the *Sc* haplotype. *Genetics*, 193(3), 985–994. https://doi.org/10.1534/genetics.112.146787

Evanno, G., Regnaut, S., & Goudet, J. (2005). Detecting the number of clusters of individuals using the software STRUCTURE: A simulation study. *Molecular Ecology*, 14(8), 2611–2620. https://doi.org/10.1111/j.1365-294X.2005.02553.x

Fisher, R. A. (1941). Average excess and average effect of a gene substitution. *Annals of Eugenics, 11*, 53–63. https://doi.org/10.1111/j.1469-1809.1941.tb02

Fisher, R. A. (1949). A theoretical system of selection for homostyle *Primula*. *Sankhyā: The Indian Journal of Statistics, 9*(4), 325–342. https://doi.org/10.1002/0471667196.ess1229.pub2

Foxe, J. P., Stift, M., Tedder, A., Haudry, A., Wright, S. I., & Mable, B. K. (2010). Reconstructing origins of loss of self-incompatibility and selfing in North American *Arabidopsis lyrata*: A population genetic context. *Evolution, 64*, 3495–3510. https://doi.org/10.1111/j.1558-5646.2010.01094.x

Ganders, F. R. (1979). The biology of heterostyly. *New Zealand Journal of Botany, 17*(4), 607–635. https://doi.org/10.1080/0028825X.1979.10432574

Gao, H., Williamson, S., & Bustamante, C. D. (2007). A Markov Chain Monte Carlo approach for joint inference of population structure and inbreeding rates from multilocus genotype data. *Genetics, 176*(3), 1635–1651. https://doi.org/10.1534/genetics.107.072371

Gilbert, K. J., Andrew, R. L., Bock, D. G., Franklin, M. T., Kane, N. C., Moore, J.-S., & Vines, T. H. (2012). Recommendations for utilizing and reporting population genetic analyses: The reproducibility of genetic clustering using the program STRUCTURE. *Molecular Ecology, 21*(20), 4925–4930. https://doi.org/10.1111/j.1365-294X.2012.05754.x

Goodwillie, C. (1999). Multiple origins of self-compatibility in *Linanthus* section Leptosiphon (Polemoniaceae): Phylogenetic evidence from internal-transcribed-spacer sequence data. *Evolution, 53*, 1387–1395.

Goodwillie, C., Kalisz, S., & Eckert, C. G. (2005). The evolutionary enigma of mixed mating systems in plants: Occurrence, theoretical explanations, and empirical evidence. *Annual Review of Ecology, Evolution and Systematics, 36*(1), 47–79. https://doi.org/10.1146/annurev.ecolsys.36.091704.175539

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.

Hamilton, N. E., & Ferry, M. (2018). ggtern: Ternary Diagrams using ggplot2. *Journal of Statistical Software, 87*(3), 1–17. https://doi.org/10.18637/jss.v087.c03

Hamrick, J. L., & Godt, M. J. W. (1996). Effects of life history traits on genetic diversity in plant species. *Philosophical Transactions of the*

Royal Society B: Biological Sciences, *351*(1345), 1291–1298. https://doi.org/10.1098/rstb.1996.0112

Hewitt, G. M. (1999). Post-glacial re-colonization of European biota. *Biological Journal of the Linnean Society*, *68*, 87–112. https://doi.org/10.1111/j.1095-8312.1999.tb01160.x

Hey, J., & Nielsen, R. (2004). Multilocus methods for estimating population sizes, migration rates and divergence time, with applications to the divergence of *Drosophila pseudoobscura* and *D. persimilis*. *Genetics*, *167*(2), 747–760. https://doi.org/10.1534/genetics.103.024182

Husband, B. C., & Barrett, S. C. H. (1993). Multiple origins of self-fertilization in tristylous *Eichhornia paniculata* (Pontederiaceae): Inferences from style morph and isozyme variation. *Journal of Evolutionary Biology*, *6*(4), 591–608. https://doi.org/10.1046/j.1420-9101.1993.6040591.x

Huu, C. N., Kappel, C., Keller, B., Sicard, A., Takebayashi, Y., Breuninger, H., Nowak, M. D., Bäurle, I., Himmelbach, A., Burkart, M., Ebbing-Lohaus, T., Sakakibara, H., Altschmied, L., Conti, E., & Lenhard, M. (2016). Presence versus absence of *CYP734A50* underlies the style-length dimorphism in primroses. *eLife*, *5*, 1–15. https://doi.org/10.7554/eLife.17956

Huu, C. N., Keller, B., Conti, E., Kappel, C., & Lenhard, M. (2020). Supergene evolution via stepwise duplications and neofunctionalization of a floral-organ identity gene. *Proceedings of the National Academy of Science, USA*, *117*(37), 23148–23157. https://doi.org/10.1073/pnas.2006296117

Jacquemyn, H., Endels, P., Brys, R., Hermy, M., & Woodell, S. R. J. (2009). Biological flora of the British Isles: *Primula vulgaris* Huds. (*P. acaulis* (L.) Hill). *Journal of Ecology*, *97*(4), 812–833. https://doi.org/10.1111/j.1365-2745.2009.01513.x

Johnson, M. G., Gardner, E. M., Liu, Y., Medina, R., Goffinet, B., Shaw, A. J., Zerega, N. J. C., & Wickett, N. J. (2016). HybPiper: Extracting coding sequence and introns for phylogenetics from high-throughput sequencing reads using target enrichment. *Applications in Plant Sciences*, *4*(7), 1600016. https://doi.org/10.3732/apps.1600016

Jombart, T. (2008). Adegenet: A R package for the multivariate analysis of genetic markers. *Bioinformatics*, *24*, 1403–1405. https://doi.org/10.1093/bioinformatics/btn129

Kaplan, J. O., Krumhardt, K. M., & Zimmermann, N. (2009). The prehistoric and preindustrial deforestation of Europe. *Quaternary Science Reviews*, *28*, 3016–3034. https://doi.org/10.1016/j.quascirev.2009.09.028

Kappel, C., Huu, C. N., & Lenhard, M. (2017). A short story gets longer: Recent insights into the molecular basis of heterostyly. *Journal of Experimental Botany*, *68*(21–22), 5719–5730. https://doi.org/10.1093/jxb/erx387

Keenan, K., Mcginnity, P., Cross, T. F., Crozier, W. W., & Prodöhl, P. A. (2013). DiveRsity: An R package for the estimation and exploration of population genetics parameters and their associated errors. *Methods in Ecology and Evolution*, *4*(8), 782–788. https://doi.org/10.1111/2041-210X.12067

Keller, B., Ganz, R., Mora-Carrera, E., Nowak, M. D., Theodoridis, S., Koutroumpa, K., & Conti, E. (2020). Asymmetries of reproductive isolation are reflected in directionalities of hybridization: Integrative evidence on the complexity of species boundaries. *New Phytologist*, *229*(3), 1795–1809. https://doi.org/10.1111/nph.16849

Keller, B., Thomson, J. D., & Conti, E. (2014). Heterostyly promotes disassortative pollination and reduces sexual interference in Darwin's primroses: Evidence from experimental studies. *Functional Ecology*, *28*(6), 1413–1425. https://doi.org/10.1111/1365-2435.12274

Kissling, J., & Barrett, S. C. H. (2013). Variation and evolution of herkogamy in *Exochaenium* (Gentianaceae): Implications for the evolution of distyly. *Annals of Botany*, *112*(1), 95–102. https://doi.org/10.1093/aob/mct097

Kohn, J. R., Graham, S. W., Morton, B., Doyle, J. J., & Barrett, S. C. H. (1996). Reconstruction of the evolution of reproductive characters in Pontederiaceae using phylogenetic evidence from chloroplast DNA restriction-site variation. *Evolution*, *50*(4), 1454–1469. https://doi.org/10.1111/j.1558-5646.1996.tb03919.x

Kopelman, N. M., Mayzel, J., Jakobsson, M., Rosenberg, N. A., & Mayrose, I. (2015). Clumpak: A program for identifying clustering modes and packaging population structure inferences across K. *Molecular Ecology Resources*, *15*(5), 1179–1191. https://doi.org/10.1111/1755-0998.12387

Kozlov, A. M., Darriba, D., Flouri, T., Morel, B., & Stamatakis, A. (2019). RAxML-NG: A fast, scalable and user-friendly tool for maximum likelihood phylogenetic inference. *Bioinformatics*, *35*(21), 4453–4455. https://doi.org/10.1093/bioinformatics/btz305

Kumar, S., Stecher, G., Li, M., Knyaz, C., & Tamura, K. (2018). MEGA X: Molecular evolutionary genetics analysis across computing platforms. *Molecular Biology and Evolution*, *35*(6), 1547–1549. https://doi.org/10.1093/molbev/msy096

Lande, R., & Schemske, D. W. (1985). The evolution of self-fertilization and inbreeding depression in plants. I. Genetic models. *Evolution*, *39*(1), 24–40. https://doi.org/10.2307/2408514

Lewis, D., & Jones, D. A. (1992). The genetics of heterostyly. In S. C. H. Barrett (Ed.), *Evolution and function of heterostyly* (pp. 129–151). Springer-Verlag.

Li, J., Cocker, J. M., Wright, J., Webster, M. A., McMullan, M., Dyer, S., Swarbreck, D., Caccamo, M., Oosterhout, C. V., & Gilmartin, P. M. (2016). Genetic architecture and evolution of the S locus supergene in *Primula vulgaris*. *Nature Plants*, *2*(12), 16188. https://doi.org/10.1038/nplants.2016.188

Lloyd, D. G., & Webb, C. J. (1992). The selection of heterostyly. In S. C. H. Barrett (Ed.), *Evolution and function of heterostyly* (pp. 179–207). Springer-Verlag.

Lopes, J. S., Balding, D., & Beaumont, M. A. (2009). popABC: A program to infer historical demographic parameters. *Bioinformatics*, *25*(20), 2747–2749. https://doi.org/10.1093/bioinformatics/btp487

Maddison, W. P., & Maddison, D. R. (2019). *Mesquite: A modular system for evolutionary analysis*. Version 3.61. http://www.mesquiteproject.org

Mast, A. R., Kelso, S., & Conti, E. (2006). Are any primroses (*Primula*) primitively monomorphic? *New Phytologist*, *171*(3), 605–616. https://doi.org/10.1111/j.1469-8137.2006.01700.x

Mather, K., & DE Winton, D. (1941). Adaptation and counter-adaptation of the breeding system in Primula. *Annals of Botany*, *5*, 297–311.

Meirmans, P. G., & Hedrick, P. W. (2011). Assessing population structure: FST and related measures. *Molecular Ecology Resources*, *11*(1), 5–18. https://doi.org/10.1111/j.1755-0998.2010.02927.x

Morjan, C. L., & Rieseberg, L. H. (2004). How species evolve collectively: Implications of gene flow and selection for the spread of advantageous alleles. *Molecular Ecology*, *13*(6), 1341–1356. https://doi.org/10.1111/j.1365-294X.2004.02164.x

Ness, R. W., Wright, S. I., & Barrett, S. C. H. (2010). Mating-system variation, demographic history and patterns of nucleotide diversity in the tristylous plant *Eichhornia paniculata*. *Genetics*, *184*(2), 381–392. https://doi.org/10.1534/genetics.109.110130

Nielsen, R., & Wakely, J. (2001). Distinguishing migration from isolation: A Markov Chain Monte Carlo approach. *Genetics*, *158*(2), 885–896. https://doi.org/10.1093/genetics/158.2.885

Opedal, Ø. H. (2018). Herkogamy, a principal functional trait of plant reproductive biology. *International Journal of Plant Sciences*, *179*(9), 677–687. https://doi.org/10.1086/700314

Paradis, E. (2010). Pegas: An R package for population genetics with an integrated-modular approach. *Bioinformatics*, *26*(3), 419–420. https://doi.org/10.1093/bioinformatics/btp696

Pérez-Barrales, R., & Arroyo, J. (2010). Pollinator shifts and the loss of style polymorphism in *Narcissus papyraceus* (Amaryllidaceae).

Journal of Evolutionary Biology, *23*, 1117–1128. https://doi.org/10.1111/j.1420-9101.2010.01988.x

Pérez-Barrales, R., Pino, R., Albaladejo, R. G., & Arroyo, J. (2009). Geographic variation of flower traits in *Narcissus papyraceus* (Amaryllidaceae): Do pollinators matter? *Journal of Biogeography*, *36*, 1411–1422.

Piper, J. G., Charlesworth, B., & Charlesworth, D. (1984). A high rate of self-fertilization and increased seed fertility on homostyle primroses. *Nature*, *310*(5972), 50–51. https://doi.org/10.1038/309126a0

Piper, J. G., Charlesworth, B., & Charlesworth, D. (1986). Breeding system evolution in *Primula vulgaris* and the role of reproductive assurance. *Heredity*, *56*(2), 207–217. https://doi.org/10.1038/hdy.1986.33

Pritchard, J. K., Stephens, M., & Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics*, *155*(2), 945–959. https://doi.org/10.1111/j.1471-8286.2007.01758.x

Ralph, P., & Coop, G. (2010). Parallel adaptation: One or many waves of advance of an advantageous allele? *Genetics*, *186*(2), 647–668. https://doi.org/10.1534/genetics.110.119594

Redelings, B. D., Kumagai, S., Tatarenkov, A., Wang, L., Sakai, A. K., Weller, S. G., Culley, T. M., Avise, J. C., & Uyenoyama, M. K. (2015). A Bayesian approach to inferring rates of selfing and locus-specific mutation. *Genetics*, *201*(3), 1171–1188. https://doi.org/10.1534/genetics.115.179093

Richards, A. J. (2003). *Primula* (2nd ed.). Timber Press.

Rigby, R. A., Stasinopoulos, D. M., & Lane, P. W. (2005). Generalized additive models for location, scale and shape. *Applied Statistics*, *54*(3), 507–554. https://doi.org/10.1111/j.1467-9876.2005.00510.x

Ritland, K. (2002). Extensions of models for the estimation of mating. *Heredity*, *88*(4), 221–228. https://doi.org/10.1038/sj/hdy/6800029

Rousset, F. (2008). GENEPOP'007: A complete re-implementation of the GENEPOP software for Windows and Linux. *Molecular Ecology Resources*, *8*(1), 103–106. https://doi.org/10.1111/j.1471-8286.2007.01931.x

Rozas, J., Ferrer-Mata, A., Sánchez-DelBarrio, J. C., Guirao-Rico, S., Librado, P., Ramos-Onsins, S. E., & Sánchez-Gracia, A. (2017). DnaSP 6: DNA sequence polymorphism analysis of large datasets. *Molecular Biology and Evolution*, *34*, 3299–3302. https://doi.org/10.1093/molbev/msx248

Ruiz-Martín, J., Santos-Gally, R., Escudero, M., Midgley, J. J., Pérez-Barrales, R., & Arroyo, J. (2018). Style polymorphism in Linum (Linaceae): A case of Mediterranean parallel evolution? *Plant Biology*, *20*, 100–111.

Schemske, D. W., & Lande, R. (1985). The evolution of self-fertilization and inbreeding depression in plants. II. Empirical observations. *Evolution*, *39*(1), 41–52. https://doi.org/10.1111/j.1558-5646.1985.tb04078.x

Schoen, D. J., Johnston, M. O., L'Heureux, A.-M., & Marsolais, J. V. (1997). Evolutionary history of the mating system in *Amsinckia* (Boraginaceae). *Evolution*, *51*(4), 1090–1099. https://doi.org/10.1111/j.1558-5646.1997.tb03956.x

Shao, J. W., Wang, H. F., Fang, S. P., Conti, E., Chen, Y. J., & Zhu, H. M. (2019). Intraspecific variation of self-incompatibility in the distylous plant *Primula merrilliana*. *AoB Plants*, *11*(3), plz030. https://doi.org/10.1093/aobpla/plz030

Sharma, V., Hecker, N., Roscito, J. G., Foerster, L., Bjoern, E. L., & Hiller, M. (2018). A genomics approach reveals insights into the importance of gene losses for mammalian adaptations. *Nature Communications*, *9*, 1–9. https://doi.org/10.1038/s41467-018-03667-1

Shimizu, K. K., Shimizu-Inatsugi, R., Tsuchimatsu, T., & Purugganan, M. D. (2008). Independent origins of self-compatibility in *Arabidopsis thaliana*. *Molecular Ecology*, *17*(2), 704–714. https://doi.org/10.1111/j.1365-294X.2007.03605.x

Shimizu, K. K., & Tsuchimatsu, T. (2015). Evolution of selfing: Recurrent patterns in molecular adaptation. *Annual Review of Ecology, Evolution and Systematics*, *46*(1), 593–622. https://doi.org/10.1146/annurev-ecolsys-112414-054249

Signorell, A., Aho, K., Alfons, A., Anderegg, N., Aragon, A., Arachchige, A., Arppe, A., Baddeley, A., Barton, K., Bolker, B., Borchers, H. W., Caeiro, F., Champely, S., Chessel, D., Chhay, L., Cooper, N., Cummins, C., Dewey, M., & Zeileis, A. (2020). *DescTools: Tools for descriptive statistics*. R package version 0.99.39. https://cran.r-project.org/package=DescTools

Simmons, M. P., & Ochoterena, H. (2000). Gaps as characters in sequence-based phylogenetic analyses. *Systematic Biology*, *49*(2), 369–381. https://doi.org/10.1093/sysbio/49.2.369

Triest, L., Sierens, T., & Van Rossum, F. (2015). Multiplexing 15 microsatellite loci for European primrose (*Primula vulgaris*). *Conservation Genetic Resources*, *7*(1), 279–281. https://doi.org/10.1007/s12686-014-0357-7

Valverde, T., & Silvertown, J. (1995). Spatial variation in the seed ecology of a woodland herb (*Primula vulgaris*) in relation to light environment. *Functional Ecology*, *9*(6), 942–950. https://doi.org/10.2307/2389993

Van Geert, A., Van Rossum, F., & Triest, L. (2010). Do linear landscape elements in farmland act as biological corridors for pollen dispersal? *Journal of Ecology*, *98*(1), 178–187. https://doi.org/10.1111/j.1365-2745.2009.01600.x

Van Geert, A., Van Rossum, F., & Triest, L. (2015). Perspectives for genetic rescue of the extremely fragmented Primula vulgaris populations in The Netherlands: reflecting the future of Belgian populations?. *Plant Ecology and Evolution*, *148*(3), 329–334. https://doi.org/10.5091/plecevo.2015.1101

Van Oosterhout, C., Hutchinson, W. F., Wills, D. P. M., & Shipley, P. (2004). MICRO-CHECKER: Software for identifying and correcting genotyping errors in microsatellite data. *Molecular Ecology Notes*, *4*(3), 535–538. https://doi.org/10.1111/j.1471-8286.2004.00684.x

Vekemans, X., Poux, C., Goubet, P. M., & Castric, V. (2014). The evolution of selfing from outcrossing ancestors in Brassicaceae: What have we learned from variation at the S-locus? *Journal of Evolutionary Biology*, *27*, 1372–1385.

Weir, B. S., & Cockerham, C. C. (1984). Estimating F-statistics for the analysis of population structure. *Evolution*, *38*(6), 1358–1370. https://doi.org/10.1111/j.1558-5646.1984.tb05657.x

Woodell, S. R. J. (1960). What pollinates primulas? *New Scientist*, *8*, 568–571.

Wright, S. I., Kalisz, S., & Slotte, T. (2013). Evolutionary consequences of self-fertilization in plants. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *280*(1760), 20130133. https://doi.org/10.1098/rspb.2013.0133

Yuan, S., Barrett, S. C. H., Duan, T., Qian, X., Shi, M., & Zhang, D. (2017). Ecological correlates and genetic consequences of evolutionary transitions from distyly to homostyly. *Annals of Botany*, *120*(5), 775–789. https://doi.org/10.1093/aob/mcx09

Yuan, S., Barrett, S. C. H., Li, C., Li, X., Xie, K., & Zhang, D. (2018). Genetics of distyly and homostyly in a self-compatible Primula. *Heredity*, *1*–*10*, https://doi.org/10.1038/s41437-018-0081-2

Zhang, J. (2000). Rates of conservative and radical nonsynonymous nucleotide substitutions in Mammalian nuclear genes. *Journal of Molecular Evolution*, *50*(1), 56–68. https://doi.org/10.1007/s002399910007

Zhao, Z., Luo, Z., Yuan, S., Mei, L., & Zhang, D. (2019). Global transcriptome and gene co-expression network analyses on the development of distyly in *Primula oreodoxa*. *Heredity*, *123*(6), 784–794. https://doi.org/10.1038/s41437-019-0250-y

Zhong, L. I., Barrett, S. C. H., Wang, X.-J., Wu, Z.-K., Sun, H.-Y., Li, D.-Z., Wang, H., & Zhou, W. (2019). Phylogenomic analysis reveals multiple evolutionary origins of selfing from outcrossing in a lineage of

heterostylous plants. *New Phytologist*, *224*(3), 1290–1303. https://doi.org/10.1111/nph.15905

Zhou, W., Barrett, S. C. H., Li, H. D., Wu, Z. K., Wang, X. J., Wang, H., & Li, D. Z. (2017). Phylogeographic insights on the evolutionary breakdown of heterostyly. *New Phytologist*, *214*(3), 1368–1380. https://doi.org/10.1111/nph.14453

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.