Received: 20 July 2022 | Revised: 10 October 2022 | Accepted: 14 October 2022

DOI: 10.1111/mec.16746

MOLECULAR ECOLOGY WILEY

NEWS AND VIEWS

PERSPECTIVE

# Primrose homostyles: A classic case of possible balancing selection revisited

## Deborah Charlesworth

Institute of Ecology and Evolution, School of Biological Sciences, Ashworth Laboratories, University of Edinburgh, Edinburgh, UK

**Correspondence**
Deborah Charlesworth, Institute of Ecology and Evolution, School of Biological Sciences, University of Edinburgh, Ashworth Laboratories, Edinburgh, UK.
Email: deborah.charlesworth@ed.ac.uk

**Handling Editor:** Emily Warschefsky

In this issue of *Molecular Ecology*, Mora-Carrera et al. (2022) revisit a case of the loss of an outcrossing system in primroses, which has been studied as an example of balancing selection in the wild since the 1940s. Molecular variants in the gene involved in the mutant self-fertile phenotype, which is now known, help towards understanding this textbook example of breakdown of an outcrossing system. However, as often happens, new information also raises further questions.

Most natural primrose (*Primula vulgaris*) populations are outcrossing, as they are distylous. Distylous plants have two flower "morphs," each able to fertilize only the other morph, though both are hermaphroditic, as Darwin (1877) discovered. The two types of primrose individuals are "pin" (flowers have long styles, with anthers positioned low in the corolla tube) and "thrum" (short styles, with anthers near the top of the corolla tube, see figure 1 in Mora-Carrera et al., 2022). The morphs are controlled by a single genetic locus; long-styled plants are homozygous for the recessive s allele (s/s), and short-styled individuals are S/s. Pollinations of s/s by S/s, or the reverse, produce seeds, while other combinations are termed "illegitimate" and produce at most very few seeds, due to an incompatibility system. Both legitimate crosses produce a 1:1 ratio of the two morphs (much like an XY sex chromosome system controlling male and female individuals), maintaining a 1:1 ratio every generation.

It has long seemed likely that these three intermorph differences—in style length, anther position, and whether the pollen is compatible with pin or thrum pistils—might be controlled by at least three separate genes (Figure 1a). Recent studies combining genetics with genome sequencing indeed showed that the distyly locus region includes a set of five genes, and that these are present only in the S allele (Li et al., 2016). The so-called "S haplotype" appears to have evolved by duplications of progenitor genes originally present in other genome locations (Potente et al., 2022). The gene that controls the short style length is called CYP$^T$. Loss of this gene produces a "long homostyle" phenotype (Figure 1a). Such mutant individuals are self-compatible because CYP$^T$ also controls the stigma incompatibility type (Huu et al., 2022), while the pollen incompatibility type remains the same as that of short-styled plants (incompatibility is controlled by another gene, whose identity remains mysterious). Not surprisingly, the long homostyle mutation, termed S$^H$ (with lost CYP$^T$ function) is recessive to the wild-type short-styled "allele."

Distyly was probably present in the common ancestor of the more than 400 *Primula* species (Mast et al., 2006; Vos et al., 2006), but many extant species are not distylous, and consist of secondarily homostylous individuals. Theoretical modelling predicts that mutations causing self-compatibility should readily spread throughout populations, since self-fertilization can transmit two copies of the selfing allele to offspring, whereas outcrossed offspring inherit only one maternal allele (Fisher, 1941). This "transmission advantage" is often opposed by inbreeding depression (low survival or fertility of progeny from selfing). If, however, conspecifics or pollinators become scarce, self-fertilization may be the only way to produce seeds, and even strong inbreeding depression cannot oppose the further advantage of "reproductive assurance" provided by selfing. It is therefore not surprising that secondary loss of outcrossing systems

---

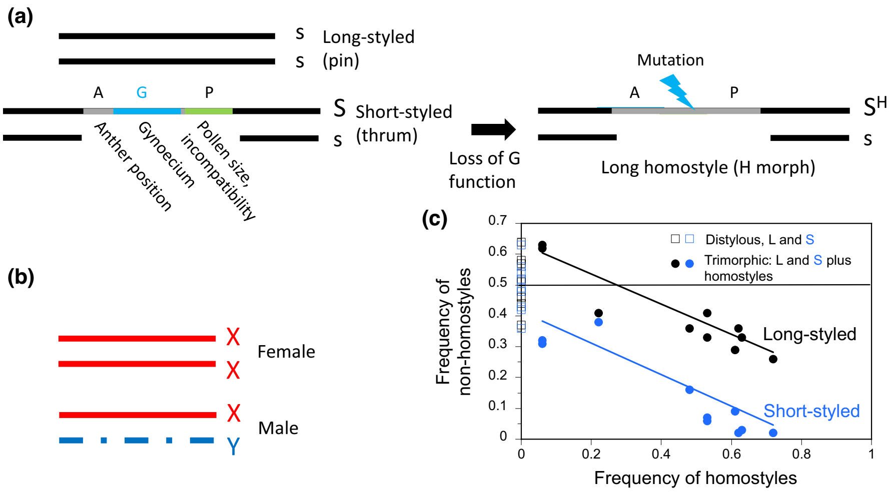

FIGURE 1 The genetics and evolution of distyly and an XY sex chromosome pair, showing the two haplotypes in both situations. (a) the distyly genome region, or S-locus. At least three S-locus genes probably control the phenotype differences between the long- and short-styled morphs in distylous plants. The S-locus region is hemizygous in the S/s genotype (short-styled), and five genes have been identified in the roughly 280-kb region, none of which is present in the s haplotype (long-styled). The S haplotype genes are therefore hemizygous in the short-styled genotype. Long homostyles can arise by a mutation causing loss of a functional $CYP^T$ gene, which controls the gynoecium character of the short-styled morph; the other $s^H$ haplotype genes are hemizygous in the initial mutant ($S^H/s$), but the haplotype can become homozygous through self-fertilization or after the haplotype spreads to a high frequency in the population. (b) X-linked genes are present in diploid coverage in females, but may be hemizygous in males, due to loss of copies from the Y, so that homozygotes for the Y chromosome are lethal. (c) Frequencies of homostyled individuals and the two other morphs in primrose natural populations sampled in Somerset; distylous populations, with no homostyles, are also shown on the *y*-axis

is common in flowering plants, which are at the mercy of environmental changes causing low density of their own species, or their pollinators. Reversions from separate sex systems back to hermaphroditism can readily occur by loss of a dominant Y-linked femaleness suppressor, and are known in several plant genera, such as papaya (VanBuren et al., 2016). Similarly, several *Primula* species from arctic environments, and *Primula scotica*, from the north of Scotland, are entirely homostylous.

Mora-Carrera et al. (2022) restudied an intriguing set of primrose populations with both heterostylous and homostylous individuals in Somerset, England. New samples of around 100 plants per site confirmed earlier findings (Crosby, 1949) that, while many populations are distylous, and have roughly 50% of each morph, others also include long homostyles, though their frequencies in nature rarely exceed 75% (Figure 1c), unlike the homostyled species mentioned above. The multilocus outcrossing rate, estimated using microsatellite genotyping in progeny of homostyles in one entirely homostylous natural population, was 15%, versus 98% for the self-incompatible morphs in a distylous population. This is consistent with these homostyles' self-compatibility providing an advantage through self-fertilization, especially in years with bad weather when reproductive assurance is important (Piper et al., 1986).

Does the persistence of populations that are not completely homostylous (Figure 1c) imply that balancing selection is maintaining the more outcrossing ancestral long-styled morph? Theoretical modelling assuming that homostyles reproduce entirely by selfing (Crosby, 1949) predicted that populations with homostyle frequencies around 75% can be maintained by balancing selection if homozygous homostyles' survival rate is 35% below that of other genotypes. Low survival of $S^H/S^H$ plants was long considered unsurprising, since distyly shares with many sex chromosome systems the property of hemizygosity of a nonrecombining region, and Y chromosomes are often lethal when made homozygous (in papaya, for example, self-fertilized XY hermaphrodites produce no YY progeny). However, there is a fundamental difference. YY lethality reflects gene losses from a multigene Y-linked region (genetic degeneration), so that males are hemizygous for many X-linked genes (Figure 1b). In contrast, hemizygosity of the distyly locus in S/s heterozygotes is due not to gene losses (as in Y chromosome degeneration), nor to mutational differences in genes present both the S and s haplotypes (with the s-linked alleles recessive, as was long thought to be the case), but instead to a history of gene duplications and movements (Potente et al., 2022) that added a small (roughly 280 kb) region with the $CYP^T$ gene along with the other four genes present in the S and $S^H$ haplotypes, while the s haplotype of normal long-styled plants completely lacks these. Low $S^H/S^H$ homozygote survival is thus puzzling. Crosby's idea is now potentially testable in progeny generated by selfing homostyles (at least in greenhouse or botanic garden conditions). Homostyles must be either $s/S^H$ heterozygotes (hemizygous for the S locus) or $S^H/S^H$ homozygotes (since $S^H$ is recessive to the S haplotype, they cannot be $S/S^H$, as such heterozygotes are short-styled). The different genotypes are thus distinguishable using sequence coverage estimates of S locus genes.

Hemizygosity greatly simplifies interpretation of sequence data. Assuming that homozygotes do indeed survive poorly, and are rare in natural populations, most homostyled individuals should be $s/S^H$

heterozygotes (lacking functional CYP$^T$, and hemizygous for a mutant copy unless it is deleted). In 27 homostylous individuals from 10 different populations, Mora-Carrera et al. (2022) found 21 CYP$^T$ sequences that encode proteins that are probably nonfunctional (as expected): 12 sequences, found in multiple populations, had a mutation creating a premature stop codon, while others, mostly from single populations, had indel or nonsynonymous mutations. In contrast, and again as expected, all 17 sequences from short-styled plants (hemizygous) encode an identical functional amino acid sequence. However, six sequences (from homostyles in two different populations) also had this sequence, and may have mutations in noncoding regions. Since no coding sequence mutation is shared by all homostyles, an earlier noncoding mutation may have disabled CYP$^T$'s function, so homostyly may not have evolved independently in the different Somerset populations. If it is older, it may have been followed by secondary mutations (neutral, because purifying selection no longer operated). It is difficult to test this by sequencing, as this gene's noncoding regions are very long. If the seven diverse mutations in the coding sequence are indeed secondary, neutral mutations, the initial homostyly mutation must have occurred long enough ago for them to have become established in the sequenced region (at most 1597 bases) in the six lineages sequenced. Homostyly may therefore not have arisen recently in Somerset, and may not be maintained by balancing selection. It could instead be a relic of an ancient homostyled primrose race, like the arctic species mentioned above, that could have colonized Britain after the ice retreated.

## DATA AVAILABILITY STATEMENT

Data sharing is not applicable to this article as no new data were created or analyzed in this study.

## ORCID

*Deborah Charlesworth* https://orcid.org/0000-0002-3939-9122

## REFERENCES

Crosby, J. L. (1949). Selection of an unfavourable gene-complex. *Evolution*, *3*, 212–230. https://doi.org/10.2307/2405559

Darwin, C. R. (1877). *The different forms of flowers on plants of the same species*. John Murray.

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.tb02272.x

Huu, C., Plaschil, S., Himmelbach, A., Kappel, C., & Lenhard, M. (2022). Female self-incompatibility type in heterostylous *primula* is determined by the brassinosteroidinactivating cytochrome P450 CYP734A50. *Current Biology*, *32*, 671–676. https://doi.org/10.1016/j.cub.2021.11.046

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

Mast, A., Kelso, S., & Conti, E. (2006). Are any primroses (*primula*) primitively monomorphic? *New Phytologist*, *171*(12), 605–616. https://doi.org/10.1111/j.1469-8137.2006.01700.x

Mora-Carrera, E., Stubbs, R. L., Keller, B., Léveillé-Bourret, E., Vos, J. M., Szövényi, P., & Conti, E. (2022). Different molecular changes underlie the same phenotypic transition: Origins and consequences of independent shifts to homostyly within species. *Molecular Ecology*, *32*, 61–78. https://doi.org/10.1111/mec.16270

Piper, J., Charlesworth, B., & Charlesworth, D. (1986). Breeding system evolution in *Primula vulgaris* and the role of reproductive assurance. *Heredity*, *56*, 207–217.

Potente, G., Léveillé-Bourret, É., Yousefi, N., Choudhury, R. R., Keller, B., Diop, S. I., Duijsings, D., Pirovano, W., Lenhard, M., Szövényi, P., & Conti, E. (2022). Comparative genomics elucidates the origin of a supergene controlling floral heteromorphism. *Molecular Biology and Evolution*, *39*, msac035. https://doi.org/10.1093/molbev/msac035

VanBuren, R., Wai, C., Zhang, J., Han, J., Arro, J., Lin, Z., Liao, Z., Yu, Q., Wang, M. L., Zee, F., Moore, R. C., Charlesworth, D., & Ming, R. (2016). Extremely low nucleotide diversity in the X-linked region of papaya caused by a strong selective sweep. *Genome Biology*, *17*(1), 230–211. https://doi.org/10.1186/s13059-016-1095-9

Vos, J. M. D., Wüest, R. O., & Conti, E. (2006). 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

**How to cite this article:** Charlesworth, D. (2023). Primrose homostyles: A classic case of possible balancing selection revisited. *Molecular Ecology*, *32*, 30–32. https://doi.org/10.1111/mec.16746