PLANT SCIENCE

Taxon-specific, phased siRNAs underlie a speciation locus in monkeyflowers

Mei Liang[1]*, Wenjie Chen[1,2], Amy M. LaFountain[1], Yuanlong Liu[3], Foen Peng[4,5], Rui Xia[3], H. D. Bradshaw[4], Yao-Wu Yuan[1,6]*

Taxon-specific small RNA loci are widespread in eukaryotic genomes, yet their role in lineage-specific adaptation, phenotypic diversification, and speciation is poorly understood. Here, we report that a speciation locus in monkeyflowers (*Mimulus*), *YELLOW UPPER* (*YUP*), contains an inverted repeat region that produces small interfering RNAs (siRNAs) in a phased pattern. Although the inverted repeat is derived from a partial duplication of a protein-coding gene that is not involved in flower pigmentation, one of the siRNAs targets and represses a master regulator of floral carotenoid pigmentation. *YUP* emerged with two protein-coding genes that control other aspects of flower coloration as a "superlocus" in a subclade of *Mimulus* and has contributed to subsequent phenotypic diversification and pollinator-mediated speciation in the descendant species.

Related organisms share developmental regulatory genes; morphological diversity is often generated by changes in function or expression patterns of these shared genes (*1*–*5*). However, whole-genome sequencing across the tree of life has revealed that new genes do evolve in specific groups of organisms (*6*–*13*). What is less clear is whether these "taxon-restricted" genes affect the subsequent evolution of the organisms that inherit them, including lineage-specific adaptation, morphological diversification, and speciation. Here, we report that a speciation locus in monkeyflowers (*Mimulus*) produces phased small interfering RNAs (siRNAs) that regulate floral carotenoid pigmentation. This noncoding gene evolved with two adjacent protein-coding genes that control other aspects of flower coloration as a "superlocus" that is restricted to a subclade of *Mimulus*.

The *Mimulus lewisii* complex contains several closely related species with three distinct pollination syndromes, including the bumble bee–pollinated *M. lewisii*, the self-pollinated *Mimulus parishii*, and the hummingbird-pollinated *Mimulus cardinalis* (Fig. 1, A to C). Where *M. lewisii* and *M. cardinalis* coflower in sympatry, they are reproductively isolated by pollinator preference (*14*, *15*). The flower color locus *YELLOW UPPER* (*YUP*) was shown to be a major contributor to pollinator choice in the natural habitat: Substitution of the *M. lewisii YUP* allele by the *M. cardinalis* version in a near-isogenic line (NIL) decreased bumble bee visitation and increased hummingbird visitation, producing a pollinator shift (*16*). As such, *YUP* is an example of a single locus causing pollinator-mediated reproductive isolation and speciation.

YUP determines the presence or absence of yellow carotenoid pigments in the petal upper epidermis (*17*). Previous genetic analyses using crosses between *M. lewisii* and *M. cardinalis* (*17*–*20*) suggested that the dominant *M. lewisii* YUP^L allele suppresses carotenoid accumulation in the pink corolla, except in the yellow nectar guides (Fig. 1A) where *YUP* is presumably not expressed. The recessive *M. cardinalis* yup^C allele allows carotenoid deposition in the entire corolla; this produces the bright red color in conjunction with a high concentration of anthocyanins (Fig. 1C), which is regulated by the *R2R3-MYB* gene *PETAL LOBE ANTHOCYANIN* (*PELAN*). Fine-scale genetic mapping of *YUP* using the *M. lewisii*–*M. cardinalis* species pair has been impeded by the lack of recombination around the *YUP* region between the two species (*19*, *20*).

Results

YUP *is a noncoding locus that produces siRNAs*

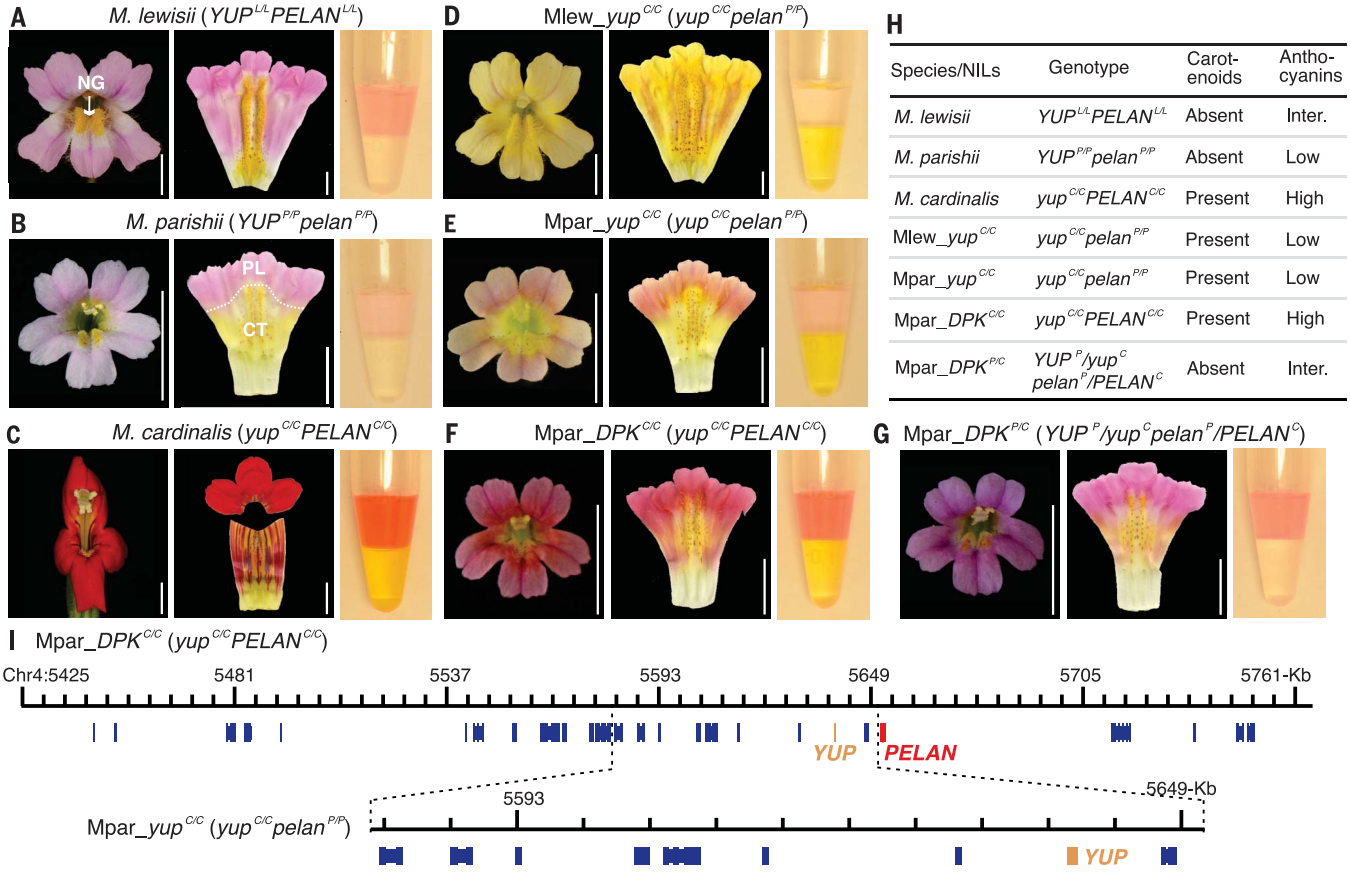
Here, we explored the possibility of using *M. parishii* as one of the parents for genetic mapping. As in *M. lewisii*, the YUP^P allele in *M. parishii* is dominant and the F_1 hybrid between *M. parishii* and *M. cardinalis* lacks carotenoid accumulation in the petal lobes (fig. S1). We first generated a crude $yup^{C/C}$ NIL in the *M. parishii* background by introgressing the *M. cardinalis* yup^C allele into *M. parishii* through two rounds of backcrossing and selfing. A resulting BC_2S_1 individual that was most similar to *M. parishii* in both vegetative and reproductive morphology but with yellow flowers was selected for the third round of backcrossing and selfing. Genotyping ~3000 BC_3S_1 and BC_3S_2 individuals delimited *YUP* to a 70-kb interval (Fig. 1I and fig. S2) and resulted in a refined Mpar_$yup^{C/C}$ NIL with petal lobe carotenoid accumulation (Fig. 1E).

This 70-kb candidate interval is part of the 337-kb chromosome fragment introgressed from *M. cardinalis* to *M. parishii* in the previously generated *DARK PINK* (*DPK*) NIL (*21*). The Mpar_$DPK^{C/C}$ NIL (Fig. 1F) accumulates more petal lobe anthocyanins than the wild-type *M. parishii* because it contains the semi-dominant *M. cardinalis PELAN* allele. The colocalization of *YUP* and *PELAN* in the *DPK* fragment explains the color hue difference between the *DPK* NILs in the heterozygous and homozygous states (dark pink versus orange) because the latter accumulates carotenoids in the petal lobes, whereas the former does not (Fig. 1, F to H). The 70-kb introgressed fragment in the Mpar_$yup^{C/C}$ NIL resulted from a rare recombination between *YUP* and *PELAN* (Fig. 1I). As such, this NIL carries the *M. cardinalis* yup^C allele but still has the *M. parishii* $pelan^P$ allele (Fig. 1H). We also introgressed the yup^C-$pelan^P$ fragment from the Mpar_$yup^{C/C}$ NIL into *M. lewisii*, which resulted in a Mlew_$yup^{C/C}$ NIL that resembles *M. lewisii* but with yellow flowers (Fig. 1D).

The 70-kb candidate interval contains eight protein-coding genes (Fig. 1I). However, RNA interference experiments of these genes in *M. lewisii* produced no phenotypic changes in carotenoid pigmentation (fig. S3), which prompted us to search for noncoding genes in this region. Examination of RNA sequencing reads mapped to this interval revealed a ~1.3-kb transcript (fig. S4A) that contains no open reading frames conserved between the dominant YUP^L and YUP^P alleles (fig. S4B) but includes an inverted repeat that could fold into a stable hairpin structure with an arm length of ~250 base pairs (fig. S5). Comparing this transcript against the genome using the Basic Local Alignment Search Tool (BLAST) revealed that the 5′-tail and left arm contiguously matched the 5′ untranslated region (5′UTR) and the first exon of a *CYP450* gene that is ~5 Mb away (Fig. 2A and fig. S6). Given that hairpin structures are often substrates for small RNA (sRNA) production (*22*), we sequenced sRNA libraries prepared from the whole corollas of wild-type *M. lewisii* and *M. parishii* as well as the Mpar_$yup^{C/C}$ NIL.

The sRNA sequencing data showed that this inverted repeat region in *M. lewisii* and *M. parishii* produced abundant 21–nucleotide (nt) siRNAs in a phased pattern (Fig. 2A), with the siRNAs generated from a specific starting position in a regular, head-to-tail arrangement (fig. S5) (*23*, *24*). By contrast, the siRNAs from the *M. cardinalis* allele, as expressed in the Mpar_$yup^{C/C}$ NIL, showed a different profile—the most abundant siRNA produced by the

[1]Department of Ecology and Evolutionary Biology, University of Connecticut, Storrs, CT 06269, USA. [2]Qinghai Provincial Key Laboratory of Crop Molecular Breeding and Key Laboratory of Adaptation and Evolution of Plateau Biota, Northwest Institute of Plateau Biology, Chinese Academy of Sciences, Xining 810008, Qinghai, China. [3]State Key Laboratory for Conservation and Utilization of Subtropical Agro-Bioresources, South China Agricultural University, Guangzhou, Guangdong 510642, China. [4]Department of Biology, University of Washington, Seattle, WA 98195, USA. [5]Department of Biology, Haverford College, Haverford, PA 19041, USA. [6]Institute for Systems Genomics, University of Connecticut, Storrs, CT 06269, USA.
*Corresponding author. Email: mei.liang@uconn.edu (M.L.); yaowu.yuan@uconn.edu (Y.-W.Y.)



H

Species/NILs	Genotype	Carot-enoids	Antho-cyanins
M. lewisii	$YUP^{L/L}PELAN^{L/L}$	Absent	Inter.
M. parishii	$YUP^{P/P}pelan^{P/P}$	Absent	Low
M. cardinalis	$yup^{C/C}PELAN^{C/C}$	Present	High
Mlew_*yup*$^{C/C}$	$yup^{C/C}pelan^{P/P}$	Present	Low
Mpar_*yup*$^{C/C}$	$yup^{C/C}pelan^{P/P}$	Present	Low
Mpar_*DPK*$^{C/C}$	$yup^{C/C}PELAN^{C/C}$	Present	High
Mpar_*DPK*$^{P/C}$	YUP^{P}/yup^{C} $pelan^{P}/PELAN^{C}$	Absent	Inter.

Fig. 1. Flower color phenotypes of different genotypic combinations of *YUP* and *PELAN*. (**A** to **G**) The front view (left image) and dissected view (middle image) of the corollas of the three parental species [(A) to (C)] and various NILs [(D) to (G)]. The dissected view was produced by cutting the corolla along the junction between the two dorsal (upper) petals 1 day before anthesis. The white arrow in (A) indicates the yellow nectar guides (NG) of *M. lewisii*; the dashed line in (B) marks the boundary between the petal lobes (PL) and corolla tube (CT). Phase separation of pigment extracts from petal lobes of equal area are also shown (right image). Anthocyanins and carotenoids were separated into the upper and lower layer, respectively, following the protocol described in (*25*). The Mpar_*DPK*$^{C/C}$ homozygous NIL (F) accumulates more anthocyanins than the Mpar_*DPK*$^{P/C}$ heterozygous NIL (G) in the petal lobes because the *PELAN*^{C} allele is semidominant (*21*). Scale bars are 10 mm. (**H**) Genotypic combinations of *YUP* and *PELAN* and pigment compositions of the parental species and NILs shown in (A) to (G). Inter, intermediate. (**I**) The 70-kb candidate interval for *YUP* is part of the 337-kb introgressed fragment in the Mpar_*DPK*$^{C/C}$ NIL. Chr4, chromosome 4.

M. parishii YUP^{P} allele had much lower accumulation in the Mpar_*yup*$^{C/C}$ NIL (Fig. 2A and fig. S5). We named the most abundant siRNA from *YUP*^{P} as *YUP_siR1* and used its accumulation as a proxy to assess the total siRNA expression level from the inverted repeat region. Quantitative reverse transcription polymerase chain reaction (qRT-PCR) across multiple stages of flower development confirmed that the abundance of *YUP_siR1* in *M. parishii* was higher than that in *M. cardinalis* (Fig. 2B). Further qRT-PCR of different floral organs of the wild-type *M. parishii* at the peak expression stage (9 mm) showed that *YUP_siR1* accumulation is restricted to the corolla and is ~7-fold higher in the petal lobes than in the corolla tube (Fig. 2C).

To test whether this sRNA locus is *YUP*, we transformed the orange-flowered Mpar_*DPK*$^{C/C}$ NIL with the ~1.3-kb transcript from *M. lewisii*, *M. parishii*, and *M. cardinalis* driven by the constitutive *CaMV 35S* promoter. All of the 30 independent transgenic lines that overexpressed the *M. lewisii* transcript showed dark pink flowers that resembled those of the heterozygous Mpar_*DPK*$^{P/C}$ NIL, with no carotenoid accumulation (Fig. 2D and figs. S7A and S8A). Transformation of the *M. parishii* transcript (35 lines) produced the same results (Fig. 2E and figs. S7A and S8B). By contrast, no phenotypic change was observed among the 41 lines that were transformed with the *M. cardinalis* allele (Fig. 2F and figs. S7A and S8C). These results confirmed that this noncoding gene is indeed *YUP*.

Although the *35S:yup*^{C} lines had similar transcript levels of the transgene as the *35S:YUP*^{L} and *35S:YUP*^{P} lines (fig. S7B), the *35S:yup*^{C} lines produced conspicuously less *YUP_siR1* (Fig. 2G). Taking this finding together with the observation that the endogenous *yup*^{C} allele does produce abundant total siRNAs (Fig. 2A), just not *YUP_siR1* per se, we conclude that the recessive *yup*^{C} allele is due to mutations that disrupt the original phased pattern (fig. S5), which leads to accumulations of different siRNA species than those produced by the dominant *YUP* alleles.

YUP *represses carotenoid accumulation through* RCP2

To explore the molecular mechanism through which the dominant *YUP* allele represses carotenoid accumulation, we queried the siRNAs derived from the dominant *YUP* allele against the sequences of all known carotenoid biosynthesis genes and their regulators in *Mimulus*. Three forms of *YUP_siR1*, two 21-nt forms and one 22-nt form, were predicted to target *REDUCED CAROTENOID PIGMENTATION2* (*RCP2*) (Fig. 3A and fig. S9A), which encodes a tetratricopeptide repeat protein that is required for carotenoid accumulation in *Mimulus* (*25*). RNA ligase–mediated 5′ rapid amplification of cDNA ends (5′ RLM RACE) confirmed that *RCP2* transcripts could be cleaved at the predicted target site in *M. lewisii* (Fig. 3A).

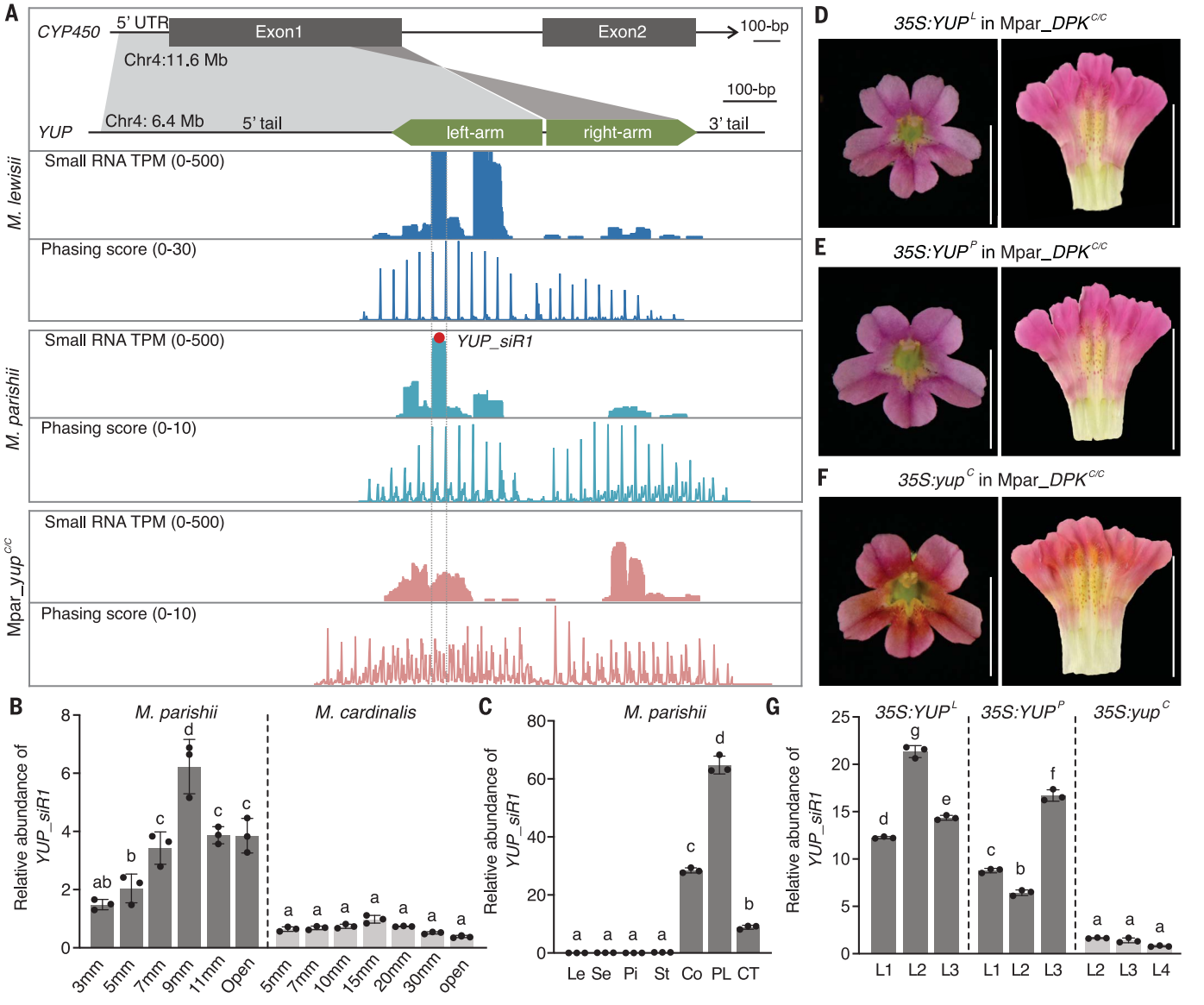


Fig. 2. *YUP* is an inverted repeat that produces phased siRNAs and represses carotenoid accumulation. (**A**) Integrative Genomics Viewer (IGV) view of the sRNA reads mapped to *YUP* and the siRNA phasing score. Sequencing data were generated from the 15-mm corolla of *M. lewisii* and the 5-mm corollas of *M. parishii* and the Mpar_*yup*$^{C/C}$ NIL and were mapped to the *M. lewisii*, *M. parishii*, and *M. cardinalis* genomes, respectively. The phasing score was calculated as in (*35*) and indicates the strength of the phasing pattern of siRNA distribution (higher score and less noise around the peaks indicate a stronger phasing pattern). The red dot indicates *YUP_siR1*, the most abundant siRNA from the *YUP*^{P} allele. TPM, tags per million. (**B**) Relative abundance of *YUP_siR1* across multiple corolla developmental stages, as measured by qRT-PCR. (**C**) Relative abundance of *YUP_siR1* in the leaf and different floral organs of *M. parishii* at the 9-mm corolla stage. Le, leaf; Se, sepal; Pi, pistil; St, stamen; Co, corolla; PL, petal lobe; CT, corolla tube. (**D** to **F**) The front view (left) and dissected view (right) of the corollas of *35S:YUP*^{L} (D), *35S:YUP*^{P} (E), and *35S:yup*^{C} (F) transgenic lines in the Mpar_*DPK*$^{C/C}$ NIL background. Scale bars are 10 mm. Also see pigment quantification in fig. S8, A to C. (**G**) Relative abundance of *YUP_siR1* in various transgenic lines. Error bars in (B), (C), and (G) are one standard deviation from three technical replicates. Different letters denote significant differences [$P < 0.01$, one-way analysis of variance (ANOVA)].

Furthermore, by examining the *M. lewisii* sRNA sequencing reads, we found that *YUP_siR1* also triggered the production of secondary siRNAs from the 5′ cleavage fragment of *RCP2* (fig. S9B). These results demonstrate that *YUP_siR1* indeed targets *RCP2*, and hereafter we will refer to it as *siR-RCP2*.

To determine whether *siR-RCP2* is sufficient to suppress carotenoid accumulation, we transformed the orange-flowered Mpar_*DPK*$^{C/C}$ NIL with an artificial *siR-RCP2* overexpression construct (*35S:amiR-RCP2*; fig. S9C). Multiple *35S:amiR-RCP2* transgenic lines (6 of 21 total lines) displayed dark pink flowers without carotenoid accumulation (Fig. 3B and fig. S8D) that were indistinguishable from those of the *35S:YUP*^{L} or *35S:YUP*^{P} lines. Further qRT-PCR analysis of three strong *35S:amiR-RCP2* lines showed a 17- to 43-fold increase in *siR-RCP2* abundance compared with the Mpar_*DPK*$^{C/C}$ NIL (Fig. 3C). However, we found that the transcript levels of *RCP2* in these lines only moderately decreased (Fig. 3D), indicating that *siR-RCP2* may also act on *RCP2* through translational inhibition.

To test the translational inhibition hypothesis, we generated transgenic *M. parishii* that overexpressed the coding sequences of *RCP2* with a C-terminal yellow fluorescent protein (YFP) fusion (fig. S10A). None of the 71 resulting transgenic lines that expressed *35S:RCP2-YFP* showed

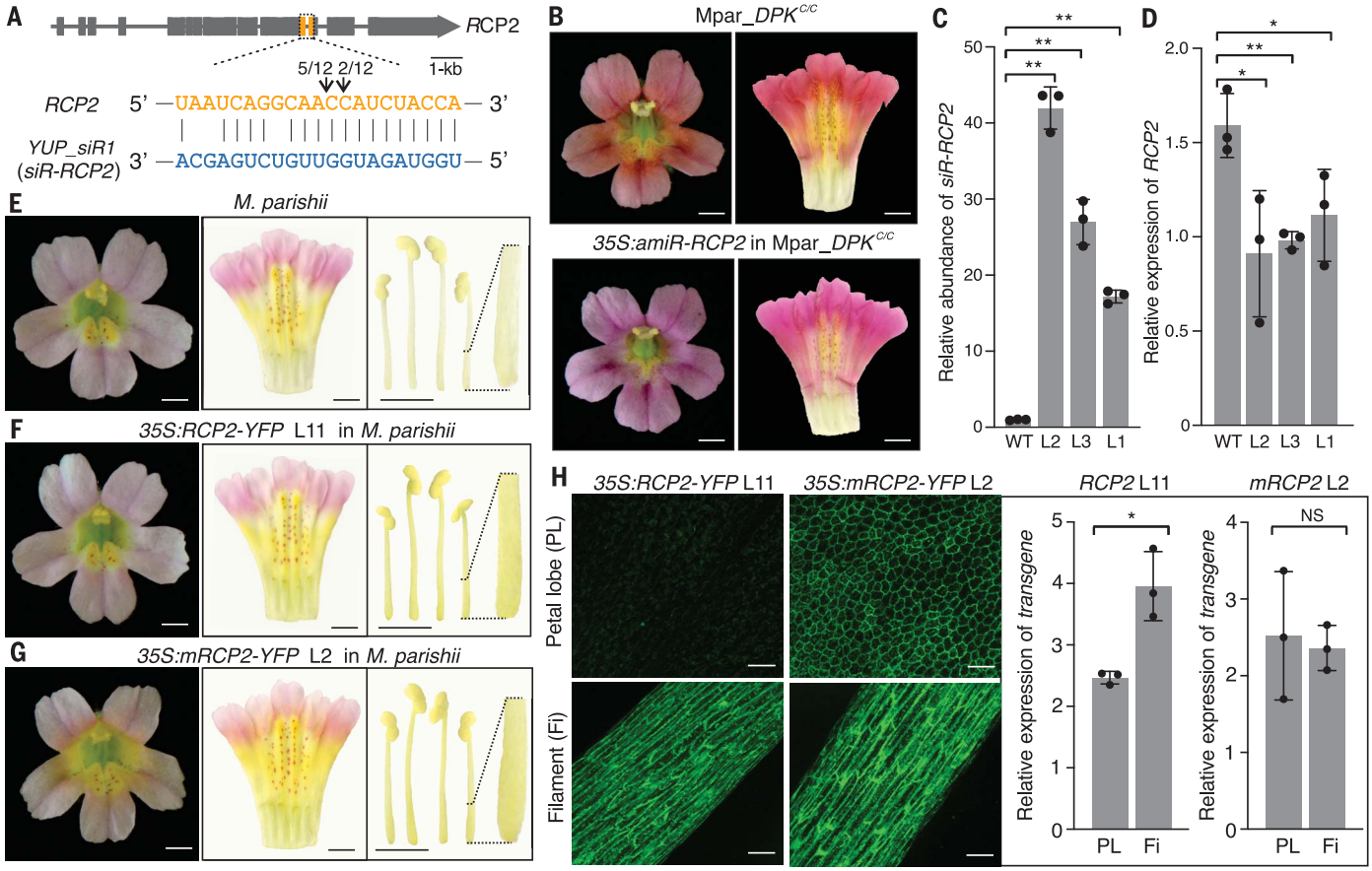


Fig. 3. *siR-RCP2* targets and represses *RCP2* through both transcript cleavage and translational inhibition. (**A**) The *siR-RCP2* target site on *RCP2* identified by 5′ RLM RACE assay. The cleavage positions are indicated by arrows, with the proportion of sequenced clones noted. (**B**) Overexpression of *siR-RCP2* in the Mpar_*DPK*$^{C/C}$ NIL resulted in a flower color shift from orange to pink. Scale bars are 2 mm. Also see pigment quantification in fig. S8D. (**C** and **D**) Relative abundance of *siR-RCP2* (C) and the transcript level of *RCP2* (D) in 9-mm corollas of the wild type (Mpar_*DPK*$^{C/C}$ NIL) and *35S:amiR-RCP2* transgenic lines, as measured by qRT-PCR. (**E** to **G**) Overexpression of the wild-type *RCP2* in *M. parishii* (E) resulted in ectopic accumulation of carotenoids in the stamen filaments but not in the corolla (F), whereas the *siR-RCP2*–insensitive *RCP2* (*mRCP2*) increased carotenoid accumulation in both organs (G). Scale bars are 2 mm. Also see pigment quantification in fig. S8, E and F, and the light microscope images of the filaments in fig. S10B. (**H**) The *35S:RCP2-YFP* transgene showed a slight decrease in petal lobes (PL) compared with the filament (Fi) at the transcript level, as measured by qRT-PCR, but a marked decrease at the protein level, as reflected by YFP fluorescent signals (left image and left graph). By contrast, the *35S:mRCP2-YFP* transgene showed similar levels of expression in petal lobes and filament at both the transcript and protein levels (right image and right graph). The YFP signals were confirmed in at least three flowers per line. Scale bars are 50 μm. Error bars in all qRT-PCR graphs [(C), (D), and (H)] are one standard deviation from three biological replicates. NS is not significant, $^{*}P < 0.05$, and $^{**}P < 0.01$ (Student's *t* test).

any change in corolla pigmentation, whereas many of the lines had increased carotenoid pigmentation in the stamen filament (Fig. 3, E and F, and figs. S8E and S10B). This was expected because, in the wild-type *M. parishii*, *siR-RCP2* is expressed in the corolla but not in stamens (Fig. 2C). The relative transcript and protein levels of the *RCP2-YFP* transgene in petal lobes and filaments were then assessed using qRT-PCR and confocal microscopy, respectively. qRT-PCR detected only a moderate difference in the transcript level between petal lobes and filaments (Fig. 3H). By contrast, confocal imaging revealed a conspicuous difference in the YFP fluorescence signal, which was hardly detectable in the petal lobes but was prominent in the filaments (Fig. 3H). These observations suggest that the *RCP2-YFP* reporter gene was translated in the filaments in the absence of *YUP* expression but that the translation was repressed in the petal lobes, which supports the translational inhibition hypothesis. This hypothesis also predicts that a *siR-RCP2*–resistant version of *RCP2* should result in protein translation in both the petal lobes and filaments. We thus transformed *M. parishii* with a *35S:mRCP2-YFP* construct by replacing the *siR-RCP2* target site with silent mutations (fig. S10A). Indeed, 11 of the 41 resulting lines showed increased carotenoid accumulation in both corolla and filament (Fig. 3G and fig. S8F) and comparable YFP signals in the two organs (Fig. 3H). In addition, the petal lobe of a *35S:mRCP2-YFP* line showed threefold higher YFP intensity than that of a *35S:RCP2-YFP* line with a very similar steady-state mRNA level of the transgene (fig. S10C), further suggesting that the *mRCP2-YFP* transcript is translated more efficiently than the *siR-RCP2*–sensitive version.

RCP1 *patterns nectar guides by repressing* YUP

The spatial segregation of carotenoids in the corolla is pronounced in the bumble bee–pollinated *M. lewisii*, where the yellow pigments accumulate only in the ventral part of the corolla tube (Fig. 1A), which serves as the nectar guides for bumble bee visitation (*26*). We found that *siR-RCP2* accumulation in *M. lewisii* is high in the petal lobes but low in the yellow nectar guides (Fig. 4A). This spatial pattern of

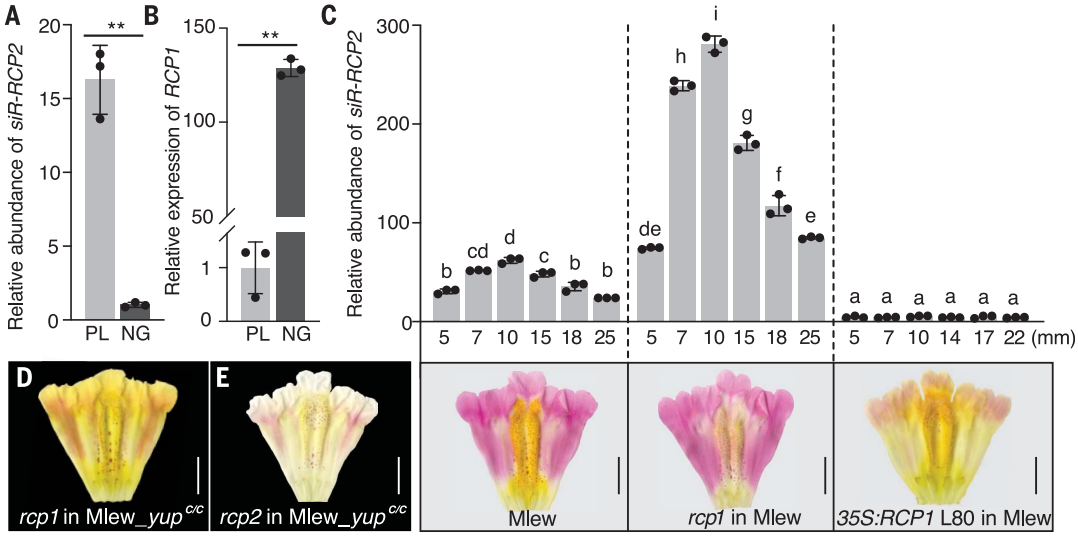


Fig. 4. *RCP1* represses *YUP* expression. (**A** and **B**) qRT-PCR assays of *siR-RCP2* (A) and *RCP1* (B) on the petal lobe (PL) and nectar guide (NG) tissue of a 10-mm *M. lewisii* corolla. Error bars are one standard deviation from three biological replicates. $^{**}P < 0.01$ (Student's *t* test). (**C**) qRT-PCR assay of *siR-RCP2* across multiple corolla stages in the wild-type *M. lewisii*, the *rcp1* mutant, and an *RCP1* overexpression line. Error bars are one standard deviation from three technical replicates. Different letters denote significant differences ($P < 0.01$, one-way ANOVA). The dissected corollas are shown below each bar graph. Scale bars are 5 mm. (**D** and **E**) Carotenoid accumulation in the corolla of Mlew_$yup^{C/C}$ NIL is not affected by the *rcp1* loss-of-function mutation (left) but is abolished by the *rcp2* mutation (right). Scale bars are 5 mm.

YUP expression is complementary to that of *REDUCED CAROTENOID PIGMENTATION1* (*RCP1*) (Fig. 4B), a previously characterized *R2R3-MYB* gene that is required for carotenoid pigmentation in the nectar guides of *M. lewisii* (*27*). We thus reasoned that *RCP1* may repress the expression of *YUP*. We performed qRT-PCR of *siR-RCP2* across multiple stages of corolla development in the wild-type *M. lewisii*, the previously characterized *rcp1* mutant, and a *35S:RCP1* overexpression line (*27*). We found that *siR-RCP2* accumulation increased in the *rcp1* mutant compared with the wild type at all corresponding stages and decreased in the *RCP1* overexpression line (Fig. 4C). These results suggest that *RCP1* represses *YUP* expression in the nectar guides, thereby allowing carotenoid accumulation in a tissue-specific fashion to promote bumble bee pollination in *M. lewisii*.

To further confirm the regulatory relationship among *RCP1*, *YUP*, and *RCP2* in the same genetic background, we crossed the loss-of-function *M. lewisii* mutants, *rcp1* and *rcp2* (*25*, *27*), with the Mlew_$yup^{C/C}$ NIL (Fig. 1B) to generate the "double mutants" *rcp1* $yup^{C/C}$ and *rcp2* $yup^{C/C}$; the former had yellow flowers with carotenoids in both the petal lobes and nectar guides (Fig. 4D), similar to the Mlew_$yup^{C/C}$ NIL, whereas the latter showed little carotenoid accumulation (Fig. 4E). These observations support the model that *RCP1* acts upstream of *YUP* and *YUP* acts upstream of *RCP2*.

Origin and evolution of the YUP-SOLAR-PELAN *superlocus*

Sequence homology indicates that *YUP* has originated from a partial duplication of the *CYP450* gene (Fig. 2A), although the source gene itself has no known function in carotenoid pigmentation and has negligible expression in the corolla of both *M. parishii* and the orange-flowered Mpar_$yup^{C/C}$ NIL (fig. S11). To determine the evolutionary timing of *YUP* origin, we first compared the synteny around *YUP* among our focal species and two distantly related *Mimulus* species with chromosome-level genome assemblies, *Mimulus guttatus* and *Mimulus aurantiacus* (*28*, *29*). Whereas the genomic region encompassing ~10 genes upstream and ~10 genes downstream of *YUP* showed great collinearity among these species, *M. guttatus* and *M. aurantiacus* lacked *YUP*, *PELAN*, and the one gene between them (Fig. 5A). The gene between *YUP* and *PELAN* is paralogous to *LIGHT AREAS1* (fig. S13), which is a subgroup-7 *R2R3-MYB* known to activate *FLAVONOL SYNTHASE* in *M. lewisii* (*30*), and is thus named *SISTER OF LIGHT AREAS1* (*SOLAR*) here. Knockdown of *SOLAR* by RNA interference resulted in reduced floral anthocyanin pigmentation (fig. S3E), suggesting that *SOLAR* is also involved in the regulation of flower coloration. We note that the source genes for *YUP*, *SOLAR*, and *PELAN* are far apart from each other in the *Mimulus* genomes (fig. S13).

The synteny analysis indicates that the *YUP-SOLAR-PELAN* "superlocus" originated after the split between the lineage that gave rise to the *M. guttatus* group and the one that gave rise to the *M. lewisii* complex (section Erythranthe) and the three closely related sections (*31*) (Fig. 5A). To pin down the evolutionary origin of this superlocus, we sequenced the genomes of four additional species that represent sections Monimanthe, Monantha, and Paradantha. Examining contigs that contain the ~20 genes surrounding *YUP* revealed that the three genes are only present in section Monimanthe, suggesting that the *YUP-SOLAR-PELAN* superlocus arose in the common ancestor of sections Erythranthe and Monimanthe ~5 million years ago (Fig. 5A).

Overexpression of the dominant YUP^{L} allele in four additional *Mimulus* species representing Erythranthe and the distantly related *M. guttatus* group all resulted in reduced flower carotenoid accumulation (Fig. 5, B to E, and fig. S14, A to D), indicating that the common ancestor of these *Mimulus* species was already predisposed to *YUP* action. Indeed, the *siR-RCP2* target site in the *RCP2* orthologs is highly conserved not only within *Mimulus* but also among a wide range of angiosperm species (fig. S14E).

Discussion

In this study, we isolated the speciation locus in *Mimulus*, *YUP*, as a taxon-restricted noncoding gene that produces phased siRNAs. One of these siRNAs represses *RCP2*, a master regulator of carotenoid accumulation, through both transcript cleavage and translational inhibition. Since its origin ~5 million years ago, the *YUP-SOLAR-PELAN* superlocus has contributed to flower color diversification and adaptation to different pollination modes in the descendant species (Fig. 5F).

Another example of a regulatory sRNA locus derived from inverted duplication underlying phenotypic variation with clear ecological importance is the *SULF* locus in the snapdragon *Antirrhinum majus* (*22*). In monkeyflowers, *YUP* shows the following distinguishing features: (i) the target of *YUP* siRNA (i.e., *RCP2*) has no phylogenetic affinity to the *CYP450* source gene; (ii) production of this *RCP2*-targeting siRNA relies on a particular mode

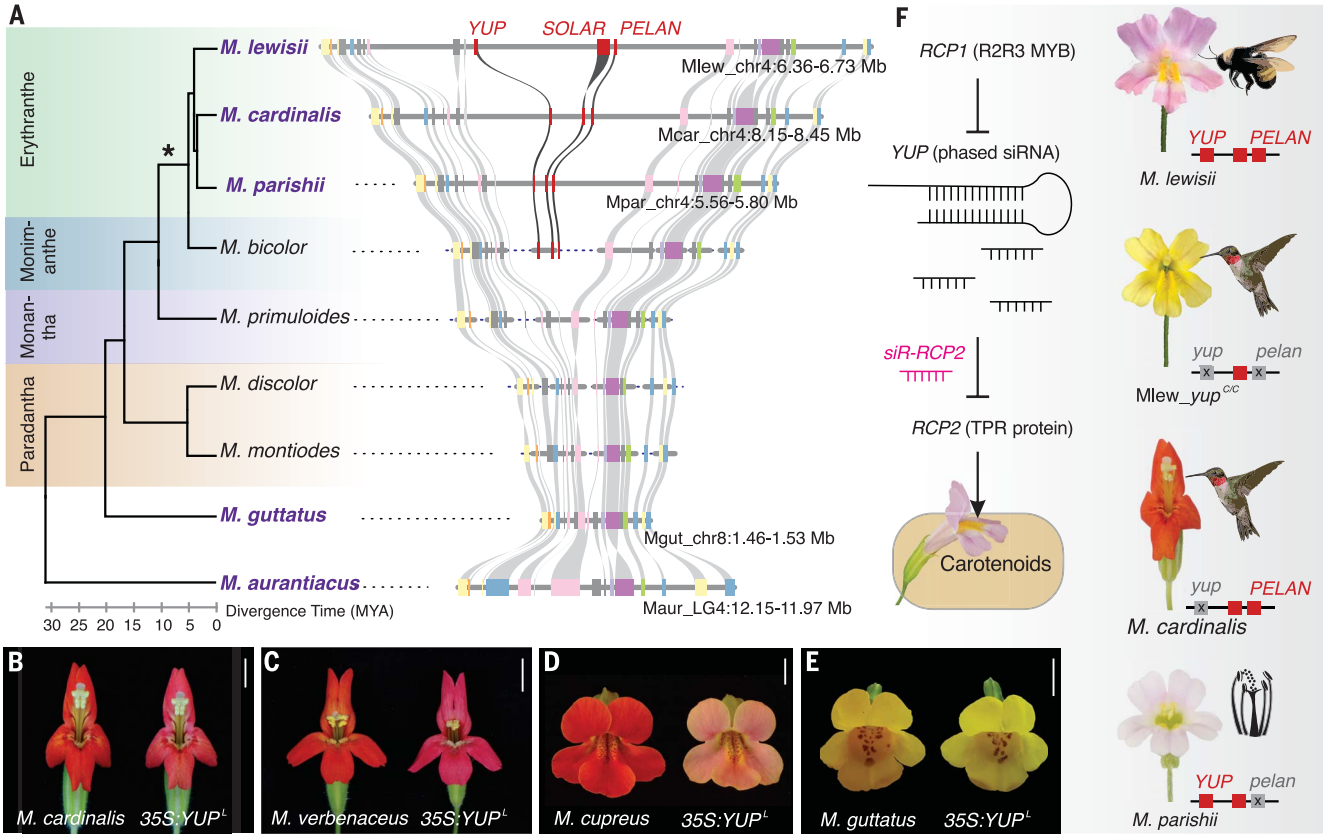


Fig. 5. The *YUP-SOLAR-PELAN* superlocus evolved in a subclade of *Mimulus* and contributed to flower color diversification. (**A**) Syntenic relationships of the *YUP* region encompassing ~10 flanking genes on each side of the *YUP-SOLAR-PELAN* superlocus. The time tree shown on the left is based on the maximum likelihood phylogeny (fig. S12) with the divergence time between the *M. guttatus* group and "*M. lewisii* + *M. bicolor*" clade as the calibration point. The species highlighted in bold have chromosome-level genome assemblies, where the *YUP* region is contiguously assembled. The other four species have fragmentary genome assemblies, and their contigs that match the *YUP* region are manually linked, as indicated by the dashed lines. The asterisk marks the node where *YUP*, *SOLAR*, and *PELAN* originated. The exceptionally large size of the *M. lewisii SOLAR* gene is due to a long intron. MYA, million years ago. (**B** to **E**) Overexpression of the *M. lewisii* YUP^{L} allele in various *Mimulus* species results in reduced petal carotenoid accumulation. Scale bars are 10 mm. Also see pigment quantification in fig. S14. (**F**) Summary of the regulatory relationships among *RCP1*, *YUP*, and *RCP2* (left) and flower color diversity generated by different *YUP* and *PELAN* genotypic combinations (right). Flower size in (F) is not scaled in proportion. All species listed in (A) to (F) are genus *Mimulus*. TPR, tetratricopeptide repeat.

of hairpin processing to generate siRNAs in a phased pattern, likely by DICER-LIKE 4 on the basis of our current understanding of the DICER-like proteins in *Arabidopsis* (*24*); and (iii) *YUP* evolved as part of the *YUP-SOLAR-PELAN* superlocus. It is difficult to envision how these three genes were brought together in the first place because their source genes are far apart in all the available *Mimulus* genomes (fig. S13). Given that all three genes are involved in flower coloration, it is tempting to speculate that this superlocus was maintained by pollinator-mediated selection after its emergence. Each of these three features alone seems to render *YUP* origination improbable; the odds are even smaller with all three combined. Yet, not only was *YUP* generated de novo, but it also has been subsequently maintained in the descendant species, contributing to flower color variation and speciation by evolving tissue specific expression patterns (e.g., as in *M. lewisii*) or loss-of-function mutations that disrupt the phased pattern of siRNA production (e.g., as in *M. cardinalis*).

Our results highlight the role of taxon-restricted regulatory sRNAs in species diversification. That these taxon-restricted sRNAs have not garnered as much attention as transcription factors and structural genes is perhaps due more to the practical challenges of their identification and functional characterization than to any difference in their relevance to phenotypic evolution.

The *M. lewisii* complex has served as a laboratory and field model in our quest to understand the molecular basis of phenotypic diversification, adaptation, and speciation (*32*). The common ancestor of the *M. lewisii* complex most likely had dark pink flowers (*21*, *33*). The bumble bee–pollinated *M. lewisii* evolved light pink flowers because of cis-regulatory mutations that up-regulated the transcription of *ROSE INTENSITY1* (*ROI1*), a conserved anthocyanin-repressing *R3-MYB* gene (*34*), which conforms to the widely accepted paradigm for the molecular basis of phenotypic evolution (*2*–*5*). By contrast, the self-pollinated *M. parishii* convergently evolved pale pink flowers because of a single nucleotide substitution in the 5′UTR of the anthocyanin-activating *R2R3-MYB* gene *PELAN* that inhibits protein translation (*21*), a type of mutation rarely considered when investigating phenotypic variation. Our results here show that the hummingbird-pollinated *M. cardinalis* evolved red flowers using a third mechanism that involved taxon-restricted phased siRNA derived from an inverted duplication of a gene that is not even part of the pigmentation pathway. We expect that further studies of this system will likely reveal additional molecular mechanisms of adaptive genetic variation.

REFERENCES AND NOTES

1. M. C. King, A. C. Wilson, *Science* **188**, 107–116 (1975).
2. H. E. Hoekstra, J. A. Coyne, *Evolution* **61**, 995–1016 (2007).

3. S. B. Carroll, *Cell* **134**, 25–36 (2008).
4. D. L. Stern, V. Orgogozo, *Science* **323**, 746–751 (2009).
5. V. Courtier-Orgogozo, L. Arnoult, S. R. Prigent, S. Wiltgen, A. Martin, *Nucleic Acids Res.* **48**, D696–D703 (2020).
6. K. Khalturin, G. Hemmrich, S. Fraune, R. Augustin, T. C. G. Bosch, *Trends Genet.* **25**, 404–413 (2009).
7. H. Kaessmann, *Genome Res.* **20**, 1313–1326 (2010).
8. S. Chen, B. H. Krinsky, M. Long, *Nat. Rev. Genet.* **14**, 645–660 (2013).
9. R. Neme, D. Tautz, *BMC Genomics* **14**, 117 (2013).
10. M. E. Santos, A. Le Bouquin, A. J. J. Crumière, A. Khila, *Science* **358**, 386–390 (2017).
11. J. Paps, P. W. H. Holland, *Nat. Commun.* **9**, 1730 (2018).
12. S. Xia *et al*., *PLOS Genet.* **17**, e1009654 (2021).
13. L. Wu, J. D. Lambert, *Semin. Cell Dev. Biol.* S1084-9521 (22) 00183-5 (2022).
14. D. W. Schemske, H. D. Bradshaw Jr., *Proc. Natl. Acad. Sci. U.S.A.* **96**, 11910–11915 (1999).
15. J. Ramsey, H. D. Bradshaw Jr., D. W. Schemske, *Evolution* **57**, 1520–1534 (2003).
16. H. D. Bradshaw Jr., D. W. Schemske, *Nature* **426**, 176–178 (2003).
17. W. M. Hiesey, M. A. Nobs, O. Björkman, *Experimental Studies on the Nature of Species: V. Biosystematics, Genetics and Physiological Ecology of the Erythranthe Section of the Erythranthe Section of* Mimulus (Publication 628, Carnegie Institute of Washington, 1971).
18. H. D. Bradshaw Jr., S. M. Wilbert, K. G. Otto, D. W. Schemske, *Nature* **376**, 762–765 (1995).
19. C. R. Pince, thesis, University of Washington (2009).
20. L. Fishman, A. Stathos, P. M. Beardsley, C. F. Williams, J. P. Hill, *Evolution* **67**, 2547–2560 (2013).
21. M. Liang, C. E. Foster, Y. W. Yuan, *Sci. Adv.* **8**, eabo1113 (2022).
22. D. Bradley *et al*., *Science* **358**, 925–928 (2017).
23. Q. Fei, R. Xia, B. C. Meyers, *Plant Cell* **25**, 2400–2415 (2013).
24. Y. Liu, C. Teng, R. Xia, B. C. Meyers, *Plant Cell* **32**, 3059–3080 (2020).
25. L. E. Stanley *et al*., *Plant Cell* **32**, 1536–1555 (2020).
26. C. R. Owen, H. D. Bradshaw, *Arthropod-Plant Interact.* **5**, 235–244 (2011).
27. J. M. Sagawa *et al*., *New Phytol.* **209**, 1049–1057 (2016).
28. U. Hellsten *et al*., *Proc. Natl. Acad. Sci. U.S.A.* **110**, 19478–19482 (2013).
29. S. Stankowski *et al*., *PLOS Biol.* **17**, e3000391 (2019).
30. Y. W. Yuan, A. B. Rebocho, J. M. Sagawa, L. E. Stanley, H. D. Bradshaw Jr., *Proc. Natl. Acad. Sci. U.S.A.* **113**, 2448–2453 (2016).
31. W. R. Barker, G. L. Nesom, P. M. Beardsley, N. S. Fraga, *Phytoneuron* **39**, 1–60 (2012).
32. Y. W. Yuan, *New Phytol.* **222**, 694–700 (2019).
33. T. C. Nelson *et al*., *PLOS Genet.* **17**, e1009095 (2021).
34. Y. W. Yuan, J. M. Sagawa, R. C. Young, B. J. Christensen, H. D. Bradshaw Jr., *Genetics* **194**, 255–263 (2013).
35. R. Xia *et al*., *Plant Cell* **25**, 1555–1572 (2013).

ACKNOWLEDGMENTS

We thank M. Opel, M. Moriarty, A. Garchow, and C. Liu for excellent plant care in the University of Connecticut EEB research greenhouses. We thank D. Grossenbacher for providing the *Mimulus bicolor* seeds; A. Cooley for the *Mimulus cupreus* seeds; L. Fishman for the *Mimulus primuloides* tissue; and L. Stanley for the *Mimulus discolor* and *Mimulus montioides* tissues. We thank members of the Yuan laboratory for comments on the manuscript and thank D. Schemske, B. Watson, J. Sagawa, L. Fishman, L. Stanley, Q. Lin, and B. Ding for discussion and assistance during the hunt for *YUP* over the years. **Funding:** This work was supported by an NSF grant (IOS-1827645) and an NIH grant (R01GM140092) to Y.-W.Y. and an NIH grant (R01GM088805) to H.D.B. **Author contributions:** M.L., H.D.B., and Y.-W.Y. designed the research; M.L., W.C., A.M.L., F.P., H.D.B., and Y.-W.Y. performed the research; M.L., Y.L., R.X., and Y.-W.Y. analyzed data; and M.L. and Y.-W.Y. wrote the paper with input from all authors. **Competing interests:** The authors declare that they have no competing interests. **Data and materials availability:** All the data required to evaluate and extend the conclusions in the paper are present in the paper and/or the supplementary materials. The Illumina sequencing data (sRNA sequencing, whole-genome sequencing) have been deposited to the National Center for Biotechnology Information (NCBI) under BioProject numbers PRJNA882815 and PRJNA882787. The described biological material can be obtained from Y.-W.Y.

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.adf1323
Materials and Methods
Figs. S1 to S14
Table S1
References (*36*–*56*)
MDAR Reproducibility Checklist

View/request a protocol for this paper from *Bio-protocol*.

Submitted 30 September 2022; accepted 6 December 2022
10.1126/science.adf1323