Bulk Segregant Analysis of an Induced Floral Mutant Identifies a *MIXTA*-Like R2R3 *MYB* Controlling Nectar Guide Formation in *Mimulus lewisii*

Yao-Wu Yuan,[1] Janelle M. Sagawa, Verónica S. Di Stilio, and H. D. Bradshaw, Jr.[1]
Department of Biology, University of Washington, Seattle, Washington 98195

ABSTRACT The genetic and developmental basis of many ecologically important floral traits (*e.g.*, carotenoid pigmentation, corolla tube structure, nectar volume, pistil and stamen length) remains poorly understood. Here we analyze a chemically induced floral mutant of *Mimulus lewisii* through bulk segregant analysis and transgenic experiments and identify a *MIXTA*-like R2R3 *MYB* gene that controls nectar guide formation in *M. lewisii* flowers, which involves epidermal cell development and carotenoid pigmentation.

THE rapid adaptive radiation of the >250,000 species of flowering plants has produced an astonishing diversity of flower morphology. Uncovering the genetic basis (*i.e.*, genes and genetic pathways/networks) of floral trait variation is a fundamental step toward understanding the origin and evolution of these "endless forms" (Darwin 1859). Floral trait diversification is often thought to be driven principally by plant–pollinator interactions (Darwin 1862; Grant and Grant 1965; Stebbins 1970; Fenster *et al.* 2004; Harder and Johnson 2009). Therefore, an ideal experimental system to study the genetic basis of flower diversification should include diverse phenotypes that interact with different pollinators and be amenable to rigorous genetic and developmental analysis.

The foremost plant genetic model system, *Arabidopsis thaliana*, has been instrumental in unraveling the genes and pathways involved in making flowers from leaves (*i.e.*, the origin of the first flower) (Coen and Meyerowitz 1991; Theissen 2001; Glover 2007). However, being a self-fertilizing species, it has little variation in floral traits that are important for pollinator interactions (*e.g.*, color, shape, rewards, display). Another important plant model system, *Antirrhinum majus*, has been invaluable for our understanding of the genetic and developmental basis of floral organ identity, flower symmetry, and anthocyanin pigmentation (Coen and Meyerowitz 1991; Schwarz-Sommer *et al.* 2003; Glover 2007), largely thanks to its endogenous active transposable elements (TEs) that allow gene isolation by transposon tagging. However, the lack of standard genomic resources (*e.g.*, genome assembly) and a routine stable transformation protocol has impeded exploitation of this system to study floral traits for which TE-induced mutants are not available (*e.g.*, carotenoid pigmentation, corolla tube formation and elaboration, and stamen and pistil length).

Mimulus (monkeyflowers) represents an emerging model system that complements the aforementioned, well-established study systems, especially for exploring the diversification of flower morphology. The 160–200 species in the genus exhibit tremendous variation in floral traits and interact with a diverse array of pollinators (Wu *et al.* 2008). Of particular interest to us are *Mimulus lewisii* and *M. cardinalis*, sister species that are genetically very similar but display dramatically different flower phenotypes and are pollinated by bumblebees and hummingbirds, respectively (Hiesey *et al.* 1971; Bradshaw *et al.* 1995; Ramsey *et al.* 2003). These species have several features that greatly facilitate genetic analysis, including high fecundity (~1000 seeds per fruit), short generation time (3 months), and relatively small genome size (~500 Mb).

doi: 10.1534/genetics.113.151225
Manuscript received January 18, 2013; accepted for publication March 26, 2013
Supporting information is available online at http://www.genetics.org/lookup/suppl/doi:10.1534/genetics.113.151225/-/DC1.
Short read data have been deposited in the NCBI Short Read Archive (SRA056512); annotated *GUIDELESS* and other *MIXTA*-like gene sequences have been deposited in GenBank (KC139356 and KC692454–KC692460).
[1]Corresponding authors: Department of Biology, University of Washington, Box 355325, Seattle, WA 98195. E-mail: yuan.colreeze@gmail.com; and Department of Biology, University of Washington, Box 351800, Seattle, WA 98195. E-mail: toby@uw.edu

Recently, we have developed genomic resources for *M. lewisii* and *M. cardinalis* (Yuan *et al.* 2013), in conjunction with community resources developed for the other model species in the genus, *M. guttatus* (http://www.mimulusevolution.org/; http://www.phytozome.net/cgi-bin/gbrowse/mimulus/). More importantly, we have established an efficient *in planta* transformation system for *M. lewisii*, which allows transgenic experiments to be performed to characterize gene function and developmental processes rigorously. In the previous study (Yuan *et al.* 2013), we have demonstrated that these genomic resources and functional tools enable fine dissection of the genetic basis of flower color variation between *M. lewisii* and *M. cardinalis*. However, using this system to understand the genetics and development of flower diversification in other angiosperms—at the most fundamental level—is limited by the existing natural floral trait variation between *M. lewisii* and *M. cardinalis*. To overcome this limitation, we initiated a large-scale ethyl methanesulfonate (EMS) mutagenesis experiment using *M. lewisii* inbred line LF10, to generate novel flower phenotypes that have potential ecological relevance (Owen and Bradshaw 2011). Studying the developmental genetic basis of these mutant phenotypes presumably will generate useful knowledge for understanding the genetic basis of similar phenotypes found in natural species across the angiosperm phylogeny. Here we present an exemplar case, describing the discovery of a *MIXTA*-like R2R3 *MYB* gene that controls the formation of nectar guides in *M. lewisii* by analyzing an EMS mutant.

Results and Discussion

The ventral petal of the pink-flowered *M. lewisii* has two yellow hairy ridges as nectar guides for bumblebees (Figure 1A). This contrasting color pattern is typical of bee-pollinated flowers (Daumer 1958), including *A. majus*, although in *Antirrhinum* the yellow color is due to aurones (Jorgensen and Geissmann 1955), a type of flavonoid pigment, whereas in *M. lewisii* it is due to carotenoid pigments (Supporting Information, Figure S1). The ecological function of the nectar guides in attracting and properly orienting bumblebees into the flower during pollination has been demonstrated in *M. lewisii* by using an EMS mutant, *guideless* (Owen and Bradshaw 2011). This mutant displays a novel phenotype, lacking the yellow color and the brushy hairs (trichomes) in the nectar guides (Figure 1B), but without pleiotropic effects outside the flower. *guideless* was observed to segregate as a Mendelian recessive trait (Owen and Bradshaw 2011), but the gene identity remained unknown.

To identify the *GUIDELESS* gene, we carried out a bulk segregant analysis coupled with deep sequencing (Lister *et al.* 2009). We first crossed *guideless* (in the LF10 genetic background) with another *M. lewisii* inbred line, SL9, and pooled DNA samples from 100 F_2 segregants with the mutant phenotype (*i.e.*, homozygous for the LF10 *guideless* allele). We then sequenced the pooled DNA sample to an average coverage of 55-fold (277 million 100-bp Illumina paired-end reads), and mapped the short reads to the SL9 genome using CLC Genomics Workbench. The *GUIDELESS* gene and tightly linked regions are expected to be homozygous for the LF10 genotype among all individuals displaying the mutant phenotype (Figure S2), which means that these regions are highly enriched in homozygous single nucleotide polymorphisms (SNPs) in the "F_2 reads–SL9 genome" alignment.

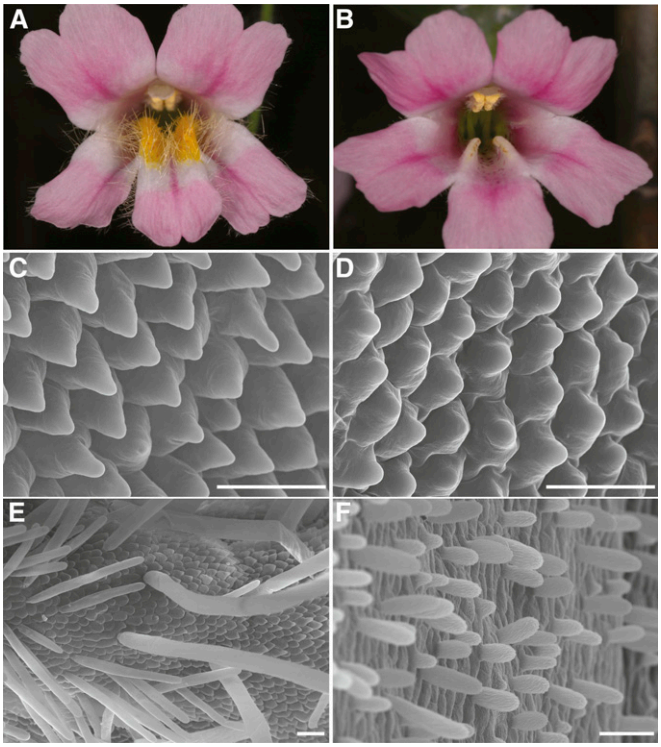

Figure 1 Phenotypic characterization of wild-type *M. lewisii* LF10 and the *guideless* mutant. Wild-type LF10 has two yellow ridges with brushy hairs (trichomes) on the ventral petal (A), conical cells on the inner epidermis of all petal lobes (C), and long (1–3 mm) single-celled trichomes in the nectar guides (E). In *guideless* mutants, there are neither yellow pigment nor brushy hairs on the ventral petal (B); the conical cells on the inner epidermis of petal lobes are much less elaborated (D), and the vestigial trichomes in the nectar guides are short (<50 μm) and stumpy (F). Bars on the SEM micrographs, 50 μm.

To generate the reference SL9 genome, we sequenced SL9 to an average coverage of 12-fold (82 million 75-bp Illumina paired-end reads), and *de novo* assembled the short reads into 86,563 contigs with an N50 of 2.3 kb, using CLC Genomics Workbench. We then aligned these contigs against the 14 chromosome-level superscaffolds of the *M. guttatus* genome using the software package MUMmer 3.0 (Kurtz *et al.* 2004), assuming gene collinearity between *M. lewisii* and *M. guttatus*. The *M. lewisii* and *M. guttatus* genomes are sufficiently diverged at nucleotide level that only the coding regions are readily alignable; therefore, only the genic regions of SL9 were captured in this genome alignment, with essentially all of the intergenic noncoding sequences being left out. This resulted in

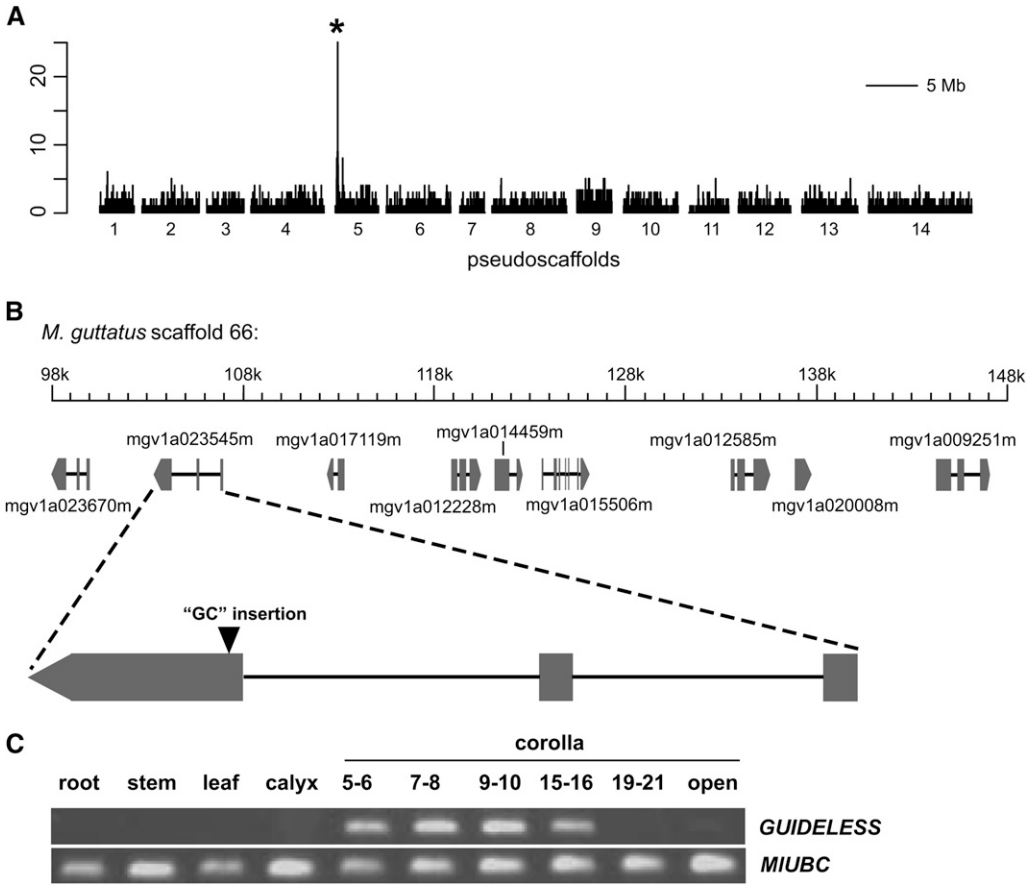

Figure 2 Identification of the *GUIDELESS* gene by bulk segregant analysis and deep sequencing. (A) Genome scan for regions that are enriched in homozygous SNPs. Each pseudoscaffold of the *M. lewisii* SL9 genome was binned into 20-kb intervals, and the number of homozygous SNPs in each 20-kb interval was plotted in a bar graph. (B) The 50-kb candidate interval contains nine genes, among which a *MIXTA*-like R2R3 *MYB* gene, *mgv1a023545m*, has a 2-bp insertion in the beginning of the third exon that disrupts the reading frame. (C) RT–PCR of *GUIDELESS* and the reference gene *MlUBC* (Yuan *et al.* 2013) in various tissue types and six stages of corolla development. *GUIDELESS* expression is restricted to the floral tissue and increases during early corolla development, with the highest level at the 9- to 10-mm stage and then decreases as the corolla matures.

14 "pseudoscaffolds" of SL9, which together contain ~70 Mb of genic sequences.

We scanned all 14 pseudoscaffolds in 20-kb intervals for enrichment of homozygous SNPs and found one sharp peak at the beginning of pseudoscaffold 5 (Figure 2A). This peak corresponds to a 50-kb region on *M. guttatus* scaffold 66, which contains only nine genes (Figure 2B). A manual inspection of the pooled mutant sample sequences that match each of the nine genes revealed neither nonsense nor nonsynonymous mutations nor mutations that potentially affect intron splicing. Instead, we found a 2-bp frameshift insertion in the beginning of the third exon of *mgv1a023545* (Figure 2B), which is a *MIXTA*-like R2R3 *MYB* gene and is the most promising candidate for *GUIDELESS* (Figure S3).

MIXTA-like genes are known to positively regulate trichome development and epidermal cell differentiation in *Antirrhinum* and other plants (Glover *et al.* 1998; Perez-Rodriguez *et al.* 2005; Baumann *et al.* 2007), which is consistent with the aborted trichome and epidermal conical cell development in the *guideless* mutant (Figure 1). However, to our knowledge, *MIXTA*-like genes have never before been associated with the regulation of carotenoid pigmentation, as the absence of yellow carotenoids in the ventral petal of *guideless* mutants would indicate. The expression of this *MYB* gene is restricted to floral tissue in LF10, and peaked at the 9–10 mm stage of corolla development (Figure 2C). This is consistent with the observation that the *guideless* mutant has no phenotypic effect outside the flower.

MIXTA-like R2R3 MYBs can be conveniently identified by a conserved signature motif, "HMAQWESARLEAEARLx-RxS" (Stracke *et al.* 2001; Brockington *et al.* 2013) (Figure S3). A TBLASTN search against the *M. guttatus* genome assembly (http://www.phytozome.net/cgi-bin/gbrowse/mimulus/) using this motif as query with an *E*-value cutoff of 1 retrieved the same set of *MIXTA*-like *MYB* genes identified in a previous study (Scoville *et al.* 2011) (Figure S4). Using the same search strategy for the *M. lewisii* genome assembly, we retrieved 10 putative *MIXTA*-like sequences, 2 of which contain multiple nonsense and frameshift mutations and are most likely pseudogenes. The other 8, including the *GUIDELESS* candidate (GenBank: KC139356) and *MlMYBML1*–*MlMYBML7* (KC692454–KC692460), were annotated as *bona fide MIXTA*-like R2R3 *MYB* genes (Figure S4).

To confirm that the candidate *MIXTA*-like *MYB* is *GUIDELESS*, we wanted to rescue the *guideless* phenotype by transforming a genomic copy of the wild-type LF10 allele into the mutant background. However, a transposable element of unknown size, located 190 bp upstream of the ATG translation initiation codon, rendered our attempts to clone the promoter region of the wild-type allele unsuccessful. Therefore, we took an alternative approach—transforming the wild-type LF10 with an RNAi construct. Knocking down

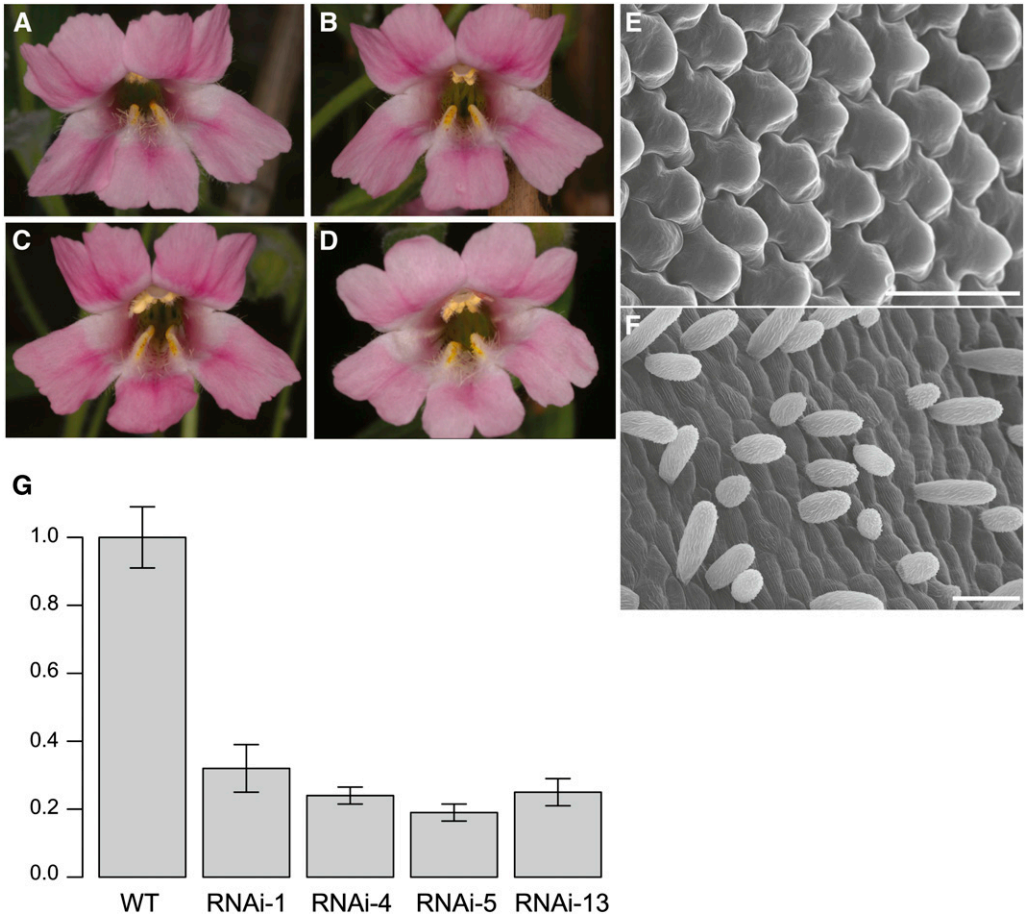

Figure 3 Characterization of the *GUIDELESS* RNAi transgenic lines. (A–D) Flower images of RNAi line 1, 4, 5, and 13, showing much reduced trichomes and carotenoid pigmentation in the nectar guides. (E and F) Aborted conical cell development on the petal lobes and aborted trichome elongation in the nectar guides of RNAi-4. Bars on the SEM micrographs, 50 μm. (G) qRT–PCR of *GUIDELESS* at the 10-mm corolla stage. *GUIDELESS* expression was knocked down 70–80% in the four RNAi transgenic lines. Bars, 1 SE from three biological replicates.

the expression of this gene in LF10 was expected to recapitulate the *guideless* phenotype.

An RNAi plasmid was constructed with a 339-bp fragment from the third exon of the candidate *MYB* gene (File S1). This fragment was BLASTed against the LF10 genome assembly to ensure target specificity. We obtained four independent RNAi lines that closely resemble the *guideless* mutant not only in gross morphology (*i.e.*, much reduced trichome development and carotenoid pigmentation), but also in the fine structure of petal lobe and nectar guide epidermal cells (Figure 3 A–F). Presence of the transgene in these RNAi lines was verified by PCR using transgene-specific primers (Figure S5 and Table S1). Quantitative reverse-transcription PCR (qRT–PCR) showed a 70–80% knockdown of the candidate *MYB* gene in these transgenic lines (Figure 3G). We have also further verified that no other *MIXTA*-like genes were inadvertently knocked down in these RNAi lines (Figure S5), which was expected as the 339-bp region used in the RNAi construct is so divergent among the *M. lewisii MIXTA*-like paralogs that no obvious sequence similarity exists at the nucleotide level.

Taken together, all three lines of evidence led to the conclusion that *GUIDELESS* is a *MIXTA*-like R2R3 *MYB* gene necessary for the development of nectar guides in *M. lewisii*: (i) in the nine-gene interval mapped by the bulk segregant analysis, only this *MIXTA*-like gene contained a mutation that could severely interfere with protein function; (ii) the specific knockdown of this gene in the wild-type genetic background recapitulated the mutant phenotype; and (iii) elements of the *guideless* phenotype are consistent with known functions of previously characterized *MIXTA*-like genes in other plants.

It is worth noting that the *guideless* mutant does not produce completely flat cells on the petal inner epidermis (Figure 1D), nor does it completely lack trichomes in the nectar guides (Figure 1F). Instead, it produces less elaborated conical cells and very short, stumpy hairs. This indicates that the primary function of *GUIDELESS* is to promote cell elaboration (*i.e.*, unidirectional cell expansion) once the outgrowth of a cell has been initiated, rather than determine cell fate in the first place. In this sense, *GUIDELESS* is functionally more similar to *AmMYBML2/PhMYB1/AtMYB16* (Baumann *et al.* 2007) than *MIXTA* (Glover *et al.* 1998) or *AmMYBML1* (Perez-Rodriguez *et al.* 2005), although phylogenetically *GUIDELESS* seems more closely related to the latter genes (Figure S4).

A somewhat similar phenotype with loss of yellow aurones and the mass of trichomes in the corolla throat has been described in the *divaricata* mutant of *A. majus* (Perez-Rodriguez *et al.* 2005). DIVARICATA is responsible

for determining ventral petal identity (Galego and Almeida 2002) and is likely to directly regulate *AmMYBL1*, thereby restricting the expression of *AmMYBL1* to the ventral petal (Perez-Rodriguez *et al.* 2005). The loss of trichomes in the corolla throat in *divaricata* is probably mediated by the down-regulation of *AmMYBL1*. The *guideless* mutant differs from *divaricata* in that the ventral petal identity is not affected in *guideless*, as the pair of ventral petal specific ridges is still present (Figure 1B). Furthermore, the *DIVARICATA* ortholog in *M. lewisii* is unlikely to directly regulate *GUIDELESS*, because *GUIDELESS* not only affects the formation of nectar guides in the ventral petal, but also regulates conical cell development in the dorsal and lateral petals (Figure 1).

Finally, an unexpected but intriguing observation of the *guideless* mutant is the loss of carotenoid pigmentation. This implies that *GUIDELESS* might directly regulate carotenoid production or deposition in the nectar guides, independent of unidirectional cell expansion during trichome and conical cell development. Alternatively, the loss of carotenoids could be the consequence of a defect in cell elaboration. Distinguishing these possibilities will require identifying the downstream target genes of GUIDELESS.

The *GUIDELESS* example highlights the potential of our collection of chemically induced *M. lewisii* mutants for contributing new knowledge of floral morphogenesis and diversification. The developmental genetics of many ecologically important floral traits (*e.g.*, carotenoid pigmentation, corolla tube structure, touch-sensitive stigma, nectar concentration and volume, and various petal lobe ornaments) remains poorly understood, simply because these traits do not exist in the conventional plant genetic model system, *A. thaliana*. The induced mutants in *M. lewisii* furnish the raw materials to study these traits. Here we outline a general strategy to use these induced floral mutants for further rapid progress in understanding the genetic and developmental basis of floral trait variation. First, one can rapidly identify the genes underlying a particular *M. lewisii* mutant phenotype by bulk segregant analysis and manipulate the genes to study their function by stable transformation. Starting from these genes, one can discover other genes in the same genetic pathway/network by three complementary approaches: (i) characterizing nonallelic mutants with similar phenotypes; (ii) yeast two-hybrid screening to detect genes whose protein products physically interact with the newly discovered protein; and (iii) comparing transcriptomes of wild-type and mutants to identify downstream target genes. Once the genetic network underlying a particular floral trait (*e.g.*, corolla tube formation and elaboration) is understood in *Mimulus*, one can then apply the *Mimulus* "gene toolbox" to dissect the developmental genetic basis of similar floral trait variation in nonmodel systems across the flowering plant phylogeny.

The defining characteristic of classical genetic model systems is the ability to go from phenotype to gene (or the reverse) with a high standard of experimental evidence. The advent of massively parallel DNA sequencing now makes it possible to develop—quickly and inexpensively—a sophisticated genetics/genomics toolkit for "emerging" model systems. Induced mutants have proven indispensable for unraveling genetic pathways and networks and must be part of the toolkit. Finally, stable transgenesis is required for rigorous testing of genetic hypotheses and precise characterization of developmental mechanisms. The rapid identification of the *GUIDELESS* gene through analyzing a chemically induced mutant, together with our recent work on fine dissection of the genetic basis of natural flower color variation between *M. lewisii* and *M. cardinalis* (Yuan *et al.* 2013), suggest that *Mimulus* is becoming such a "classical" genetic model system that is particularly suitable for studying flower diversification.

Acknowledgments

We are grateful to Brian Watson, James Vela, Doug Ewing, Jeanette Milne, Paul Beeman, and Erin Forbush for plant care. We also thank two anonymous reviewers for valuable comments on the manuscript. Piotr Mieczkowski at the University of North Carolina High Throughput Sequencing Facility supervised the Illumina sequencing. Wai Pang Chan (University of Washington Biology Imaging Center) helped with the scanning electron microscopy. The *Mimulus guttatus* genome sequencing consortium and Department of Energy Joint Genome Institute provided the chromosome-level assembly of the *M. guttatus* genome. This work was supported by National Science Foundation Frontiers in Integrative Biological Research grant 0328636 and National Institutes of Health grant 5R01GM088805.

Literature Cited

Baumann, K., M. Perez-Rodriguez, D. Bradley, J. Venail, P. Bailey *et al.*, 2007 Control of cell and petal morphogenesis by R2R3 *MYB* transcription factors. Development 134: 1691–1701.

Bradshaw, H. D., S. M. Wilbert, K. G. Otto, and D. W. Schemske, 1995 Genetic-mapping of floral traits associated with reproductive isolation in monkeyflowers (*Mimulus*). Nature 376: 762–765.

Brockington, S. F., R. Alvarez-Fernandez, J. B. Landis, K. Alcorn, R. H. Walker *et al.*, 2013 Evolutionary analysis of the *MIXTA* gene family highlights potential targets for the study of cellular differentiation. Mol. Biol. Evol. 30: 526–540.

Coen, E. S., and E. M. Meyerowitz, 1991 The war of the whorls: genetic interactions controlling flower development. Nature 353: 31–37.

Darwin, C., 1859 *On the Origin of Species by Means of Natural Selection, or the Preservation of Favoured Races in the Struggle for Life*. John Murray, London.

Darwin, C., 1862 *On the Various Contrivances by Which British and Foreign Orchids Are Fertilised by Insects, and on the Good Effects of Intercrossing*. John Murray, London.

Daumer, K., 1958 Flower colors, how they see the bees. Z. Vgl. Physiol. 41: 49–110 (in German).

Fenster, C. B., W. S. Armbruster, P. Wilson, M. R. Dudash, and J. D. Thomson, 2004 Pollination syndromes and floral specialization. Annu. Rev. Ecol. Evol. Syst. 35: 375–403.

Galego, L., and J. Almeida, 2002 Role of *DIVARICATA* in the control of dorsoventral asymmetry in *Antirrhinum* flowers. Genes Dev. 16: 880–891.

Glover, B. J., M. Perez-Rodriguez, and C. Martin, 1998 Development of several epidermal cell types can be specified by the same *MYB*-related plant transcription factor. Development 125: 3497–3508.

Glover, B. J., 2007 *Understanding Flowers and Flowering: an Integrated Approach*, Oxford University Press, Oxford.

Grant, V., and K. A. Grant, 1965 *Flower Pollination in the Phlox Family*, Columbia University Press, New York.

Harder, L. D., and S. D. Johnson, 2009 Darwin's beautiful contrivances: evolutionary and functional evidence for floral adaptation. New Phytol. 183: 530–545.

Hiesey, W. M., M. A. Nobs, and O. Björkman, 1971 *Experimental Studies on the Nature of Species. V. Biosystematics, Genetics, and Physiological Ecology of the Erythranthe Section of Mimulus*. Pub 628, Carnegie Institute, Washington, DC.

Jorgensen, E. C., and T. A. Geissmann, 1955 The chemistry of flower pigmentation in *Antirrhinum majus* color genotypes. III. Relative anthocyanin and aurone concentration. Arch. Biochem. Biophys. 55: 389–402.

Kurtz, S., A. Phillippy, A. L. Delcher, M. Smoot, M. Shumway *et al.*, 2004 Versatile and open software for comparing large genomes. Genome Biol. 5: R12.

Lister, R., B. D. Gregory, and J. R. Ecker, 2009 Next is now: new technologies for sequencing of genomes, transcriptomes, and beyond. Curr. Opin. Plant Biol. 12: 107–118.

Owen, C. R., and H. D. Bradshaw, 2011 Induced mutations affecting pollinator choice in *Mimulus lewisii* (Phrymaceae). Arthropod-Plant Interact. 5: 235–244.

Perez-Rodriguez, M., F. W. Jaffe, E. Butelli, B. J. Glover, and C. Martin, 2005 Development of three different cell types is associated with the activity of a specific *MYB* transcription factor in the ventral petal of *Antirrhinum majus* flowers. Development 132: 359–370.

Ramsey, J., H. D. Bradshaw, and D. W. Schemske, 2003 Components of reproductive isolation between the monkeyflowers *Mimulus lewisii* and *M. cardinalis* (Phrymaceae). Evolution 57: 1520–1534.

Schwarz-Sommer, Z., B. Davies, and A. Hudson, 2003 An everlasting pioneer: the story of *Antirrhinum* research. Nat. Rev. Genet. 4: 657–666.

Scoville, A. G., L. L. Barnett, S. Bodbyl-Roels, J. K. Kelly, and L. C. Hileman, 2011 Differential regulation of a MYB transcription factor is correlated with transgenerational epigenetic inheritance of trichome density in *Mimulus guttatus*. New Phytol. 191: 251–263.

Stebbins, G. L., 1970 Adaptive radiation of reproductive characteristics in angiosperms, I: pollination mechanisms. Annu. Rev. Ecol. Syst. 1: 307–326.

Stracke, R., M. Werber, and B. Weisshaar, 2001 The R2R3-MYB gene family in Arabidopsis thaliana. Curr. Opin. Plant Biol. 4: 447–456.

Theissen, G., 2001 Development of floral organ identity: stories from the MADS house. Curr. Opin. Plant Biol. 4: 75–85.

Wu, C. A., D. B. Lowry, A. M. Cooley, K. M. Wright, Y. W. Lee *et al.*, 2008 *Mimulus* is an emerging model system for the integration of ecological and genomic studies. Heredity 100: 220–230.

Yuan, Y. W., J. M. Sagawa, R. C. Young, B. J. Christensen, and H. D. Bradshaw, Jr., 2013 Genetic dissection of a major anthocyanin QTL contributing to pollinator-mediated reproductive isolation between sister species of *Mimulus*. Genetics 194: 255–263.

Communicating editor: S. Poethig

GENETICS

Supporting Information
http://www.genetics.org/lookup/suppl/doi:10.1534/genetics.113.151225/-/DC1

Bulk Segregant Analysis of an Induced Floral Mutant Identifies a *MIXTA*-Like R2R3 *MYB* Controlling Nectar Guide Formation in *Mimulus lewisii*

Yao-Wu Yuan, Janelle M. Sagawa, Verónica S. Di Stilio, and H. D. Bradshaw, Jr.

DOI: 10.1534/genetics.113.151225

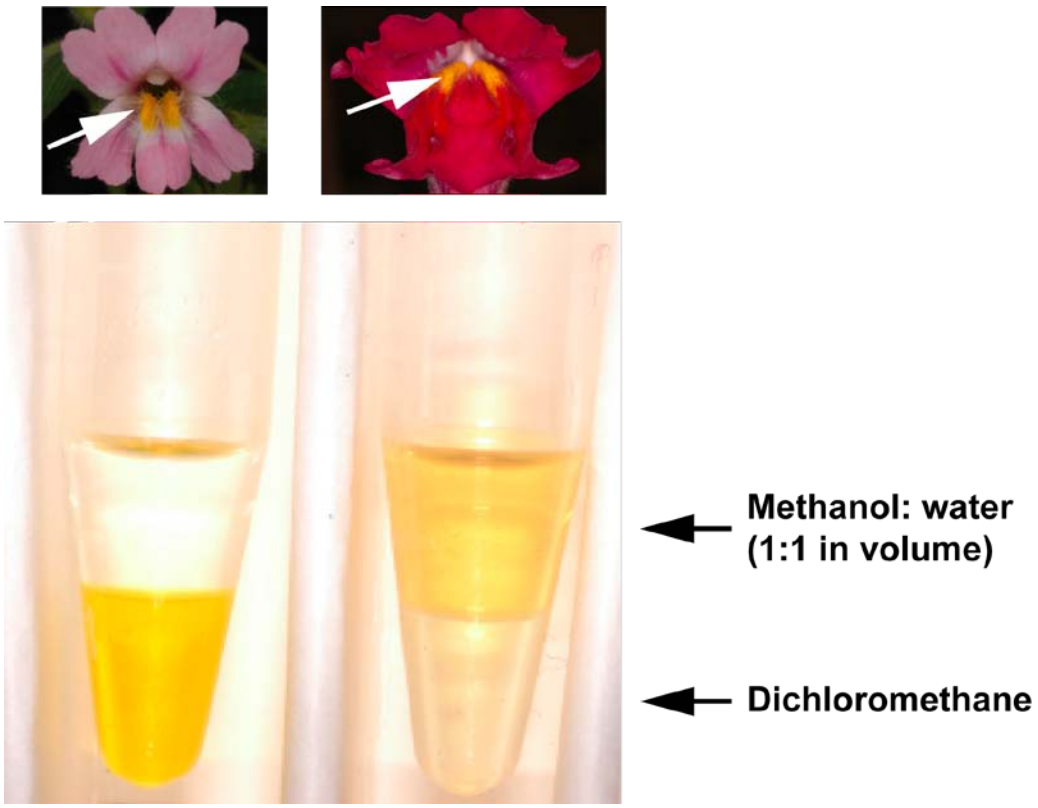

Figure S1 The yellow color of the *Mimulus lewisii* nectar guides is due to hydrophobic carotenoids. In contrast, the yellow color of *Antirrhinum majus* nectar guides is due to hydrophilic flavonoid pigments, aurones (Jorgensen and Geissmann 1955). The yellow trichomes from four flowers of each species were ground in 150 μl methanol, which dissolves both carotenoids and flavonoids. Then an equal volume of water (polar solvent) and dichloromethane (nonpolar solvent) were mixed thoroughly with the methanol extract. Centrifuging at 13,000 rpm for 2 min separated flavonoids and carotenoids to the aqueous and non-aqueous phase, respectively.

Figure S2 Bulk segregant analysis of *guideless*. Schematic representation of the expected short-read distribution, when mapping the Illumina reads of the pooled DNA sample from the F_2 segregants displaying the *guideless* phenotype to the SL9 reference genome. The LF10 and SL9 genotype are represented by blue and red, respectively. Within the *GUIDELESS* gene or very tightly linked regions, all of the reads should be from LF10, the progenitor of the mutant line. The greater the distance from *GUIDELESS*, the more SL9 reads will be found, until reaching a point that is completely unlinked with *GUIDELESS*, when the two genotypes will be randomly segregating in an expected proportion of 50%:50%.

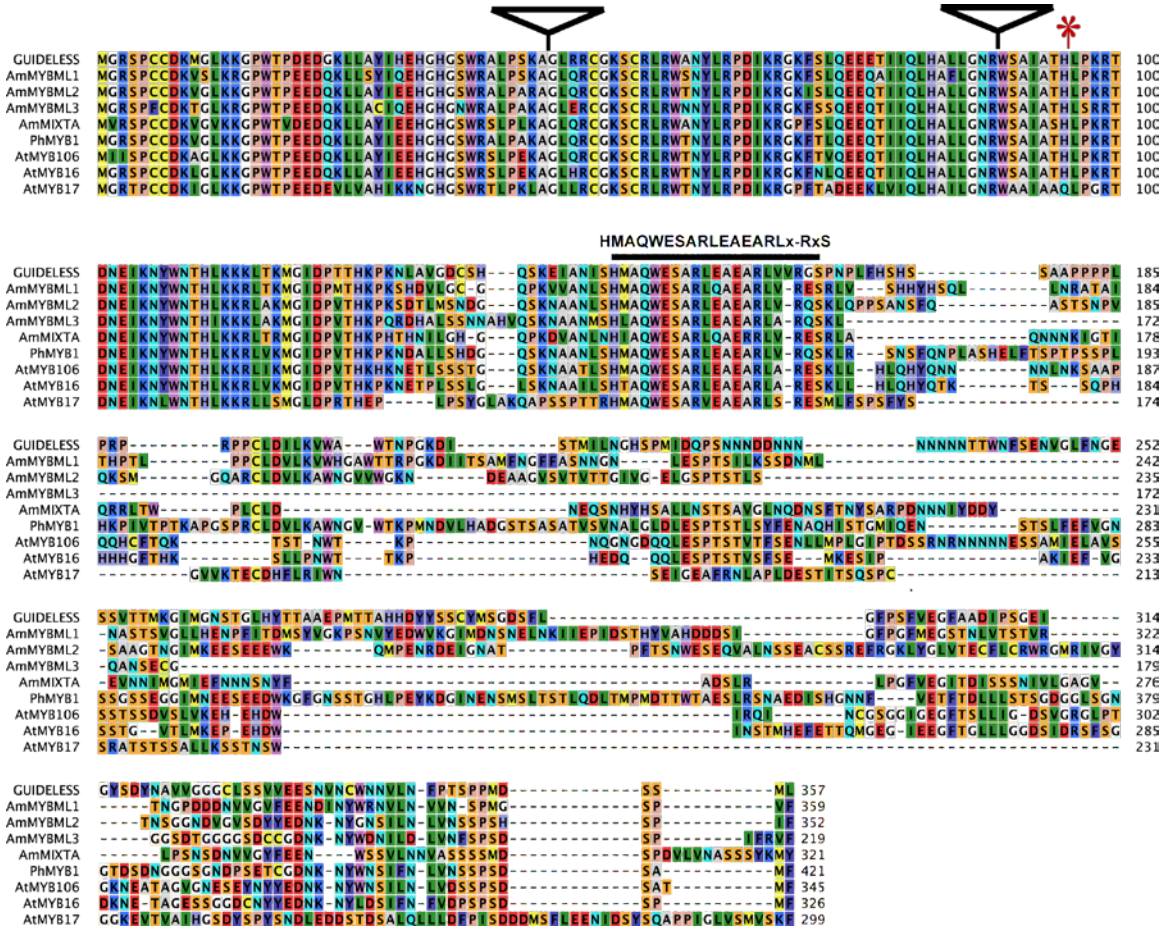

Figure S3 Alignment of *Mimulus lewisii* GUIDELESS and its closely related MIXTA-like R2R3 MYBs from *Antirrhinum*, *Petunia*, and *Arabidopsis*. The signature motif that defines the MIXTA-like MYB clade is labeled by a black bar. The positions of the two introns are indicated by the black triangles. The 2-bp (GC) insertion found in *guideless* is represented by the red asterisk. Sequences of AtMYB16, AtMYB17, and AtMYB06 were retrieved from the TAIR site (http://www.arabidopsis.org/); *Antirrhinum* and *Petunia* sequences were retrieved from GenBank (AmMIXTA: X79108; AmMYBML1: CAB43399; AmMYBML2: AAV70655; AmMYBML3: AAU13905; PhMYB1: CAA78386).

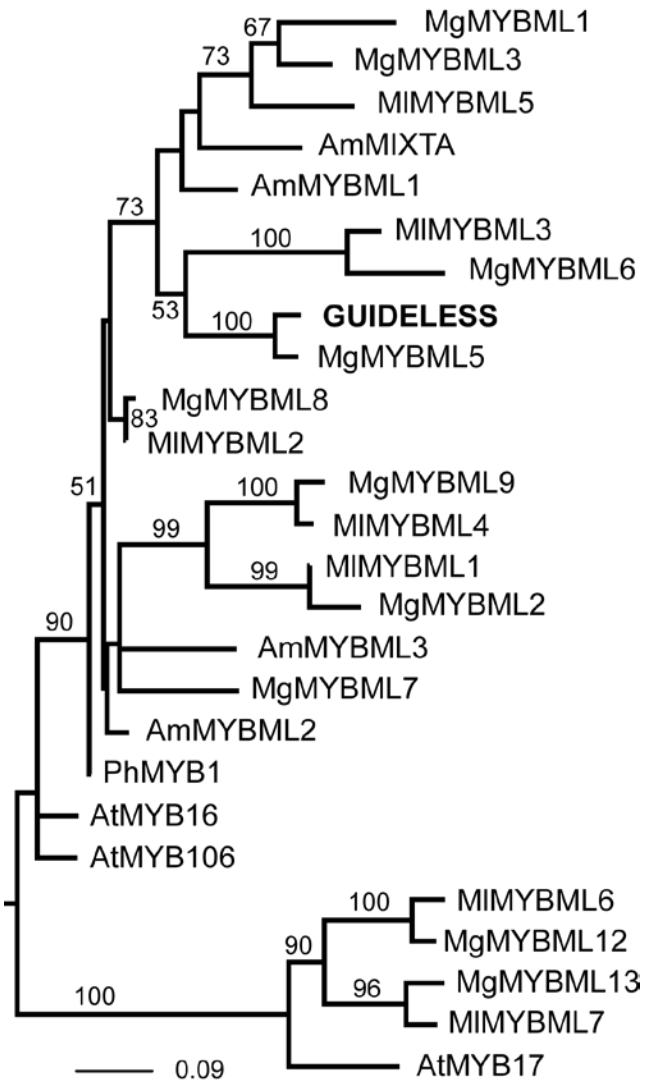

Figure S4 A maximum likelihood phylogeny of MIXTA-like proteins based on an alignment of the R2R3 MYB domains. Phylogenetic analyses were conducted using RAxML 7.0.4 (Stamatakis 2006), with the BLOSUM62 amino acid substitution matrix and CAT approximation. Bootstrap values greater than 50% are indicated along the branches. *Mimulus lewisii* sequences have been deposited in GenBank (KC139356, KC692454-KC692460). *M. guttatus* sequences were retrieved from Scoville et al. (2011). Information on other sequences was given in Figure S3.

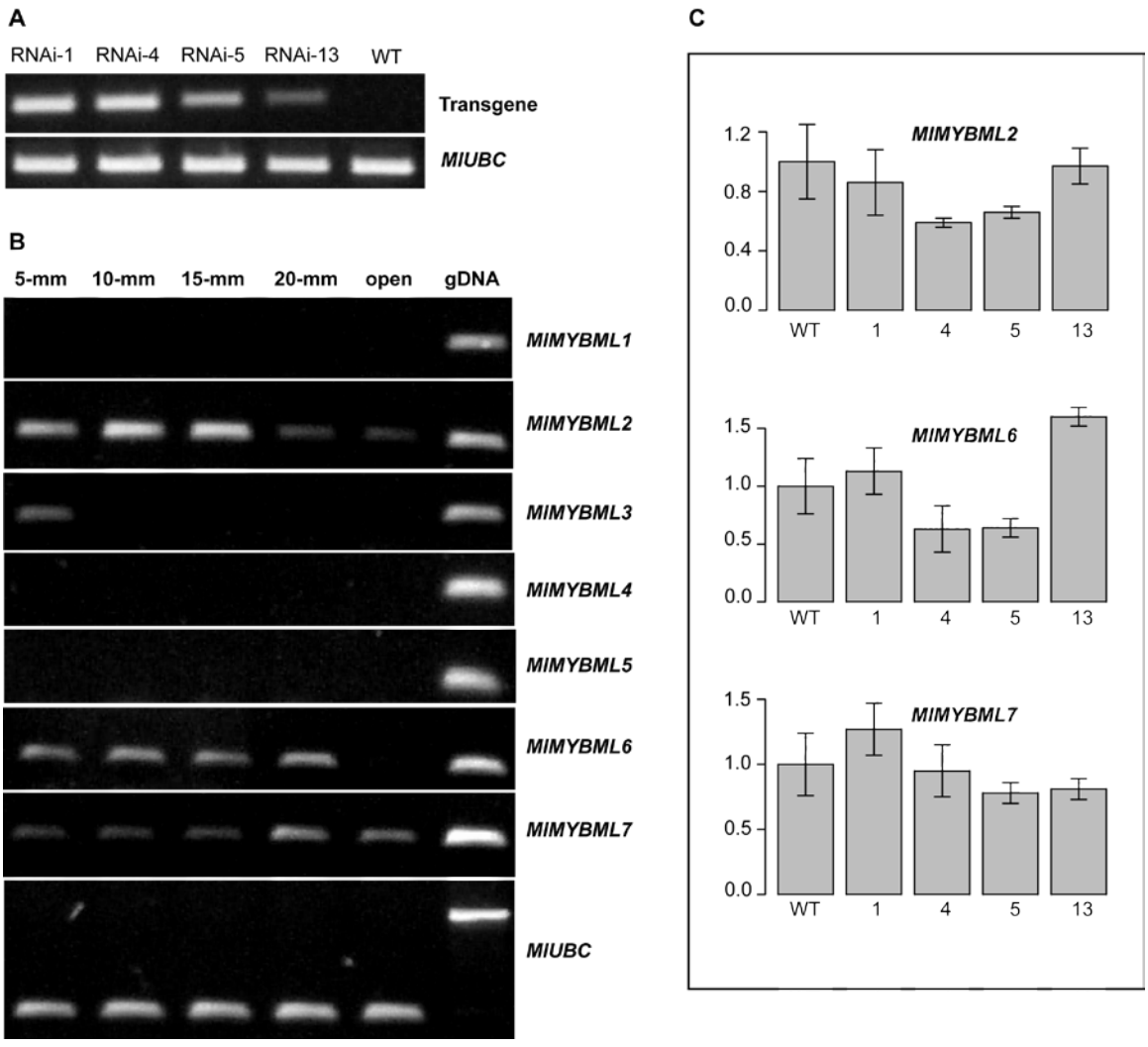

Figure S5 Molecular characterization of the RNAi transgenic lines. (A) A pair of transgene-specific primers (pFGC5941_342F and pFGC5941_1214R, Table S1) amplified a single product that is present in the genomic DNA of RNAi lines but absent from the wild-type plant. The endogenous *MlUBC* gene was used as a control for genomic DNA quality. (B) Qualitative RT-PCR to test which *MlXTA*-like genes are expressed in the wild-type corolla during flower development—genes that are not expressed in the corolla should not affect the nectar guide phenotype. Only *MlMYBL2*, *MlMYBL6*, and *MlMYBL7* were detectable at multiple stages and were subjected to further quantitative RT-PCR (qRT-PCR) analyses. The *MlUBC* gene was used for cDNA quality control, and genomic DNAs were used for primer quality control. The larger size of the *MlUBC* genomic amplicon is due to the presence of an intron in the amplified fragment. (C) qRT-PCR of *MlMYBL2, 6, 7* at the 10-mm corolla stage. Expression level of none of the three genes is significantly different between wild-type and each of the four RNAi lines (two-tailed *t*-test: $p > 0.1$). Bars represent 1 SE from three biological replicates.

File S1

Supporting Materials and Methods

Plant materials and growth condition: The *Mimulus lewisii* inbred line LF10 was described in Owen and Bradshaw (2011) and Yuan *et al*. (2013). The inbred line SL9 was developed from another individual in the same population as the parent of LF10. The *guideless* mutant is from Owen and Bradshaw (2011). Greenhouse conditions are as described in Yuan *et al*. (2013)

Genome sequencing and assembly of SL9: To produce the reference SL9 genome assembly for the bulk segregant analysis, we generated 82 million 75-bp Illumina paired-end reads (12-fold average coverage) at the University of North Carolina High-Throughput Sequencing Facility (UNC-HTSF). We assembled these reads into 86,563 contigs (N50 = 2.3 kb) using CLC Genomics Workbench. These contigs were aligned to the 14 chromosomal-level super-scaffolds of *Mimulus guttatus* with the *nucmer* module from the MUMmer 3.0 package (Kurtz *et al*. 2004), assuming gene collinearity between *M. lewisii* and *M. guttatus*. A customized perl script "MUMmer_parser.pl" (available upon request) was written to connect the SL9 contigs into 14 "pseudoscaffolds" based on the MUMmer output.

Bulk segregant analysis of the *guideless* mutation by deep sequencing: An F_2 population was produced from the cross between SL9 and a *guideless* mutant in the LF10 background, and 500 F_2 individuals were grown to flowering. One hundred F_2 segregants displaying the mutant phenotype were collected, and total genomic DNA was isolated from each of them using the BIO 101 System FastDNA kit (Qbiogene, Inc., Carlsbad CA). The concentration of each DNA sample was determined by using the PicoGreen dsDNA Quantitation Reagent (Invitrogen). The 100 DNA samples were then pooled together with equal representation from each segregant. A small-insert library (200-400 bp) was prepared for the pooled sample at UNC-HTSF, and 100-bp paired-end reads were generated by an Illumina HiSeq 2000.

The 277 million resulting reads (55-fold average coverage) were mapped to the 14 SL9 "pseudoscaffolds" with CLC Genomics Workbench, and 157,551 raw SNPs were detected. The *GUIDELESS* gene and tightly linked regions are expected to be homozygous for the LF10 genotype among all individuals displaying the mutant phenotype, which means that these regions are highly enriched in homozygous SNPs (Figure S2). The greater the distance from *GUIDELESS*, the more SL9 reads will be found. Upon reaching a point that is completely unlinked with *GUIDELESS*, the two genotypes (LF10 and SL9) will be randomly segregating in an expected proportion of 50%:50%.

The 157,551 raw SNPs were first filtered by depth of coverage. SNPs with >120-fold coverage were discarded because these regions are highly repetitive and, therefore, the reads were likely to be mapped incorrectly. The remaining 135,297 SNPs were then filtered by their tendency to cluster. The average SNP density between LF10 and SL9 is less than 0.002 (1 SNP every 500 bp); therefore, the highly clustered SNPs (3 or more SNPs in a 100-bp region) were likely to be caused by incorrect mapping. A total of 36,219 high quality SNPs were kept after filtering out clustered SNPs. The third step is to filter out heterozygous SNPs. SNPs with variant frequency less than 95% were considered as heterozygous. As a result, 3,450 high quality, homozygous SNPs were retained. To search for regions that are highly enriched in homozygous SNPs, the SL9 pseudoscaffolds were binned into 20-kb intervals, and the numbers of homozygous SNPs in each 20-kb interval were plotted in a bar graph (Figure 2A).

Two customized perl scripts, "SNP_filter.pl" and "HomoSNP_enrichment.pl" (available upon request), were written to automate the process of filtering SNPs and searching for homozygous SNP enrichment.

Plasmid construction and plant transformation: An RNAi plasmid was constructed with a 339-bp fragment from the third exon of *GUIDELESS*, essentially following the protocol described in Yuan *et al*. (2013). The primer pair *GUIDELESS*_RNAi_F and *GUIDELESS*_RNAi_R (Table S1) was used to amplify the 339-bp fragment. This fragment was BLASTed against the LF10 genome assembly with an E-value cutoff of 0.1 to ensure that no other genomic regions perfectly match this fragment for a contiguous block longer than 16 bp. The final plasmid construct was verified by sequencing and then transformed into *Agrobacterium tumefaciens* strain GV3101 for subsequent plant transformation, as described in Yuan *et al*. (2013).

Qualitative RT-PCR: Total RNA was isolated from root, stem, leaf, calyx and 6 stages of corolla development of LF10. RNA extraction and cDNA synthesis followed Yuan *et al*. (2013). The gene-specific primers *GUIDELESS*_SP3F and *GUIDELES*_SP3R (Table S1) were used to amplify a 208-bp fragment of the third exon, to examine *GUIDELESS* expression in the wild-type LF10 across different tissue type and different stages of corolla development.

Gene- specific primers for the other seven *MIXTA*-like genes, *MlMYBML1-7* (Table S1), were used to examine their expression at five different stages of corolla development (Figure S5). *MlUBC* was used as a reference gene.

Quantitative RT-PCR: qRT-PCR was performed to quantify expression levels of *GUIDELESS*, *MlMYBML2*, *MlMYBML6*, and *MlMYBML7* in the 10-mm corolla of the wild-type LF10 and the four RNAi transgenic lines (RNAi-1, -4, -5, and -13). *MlUBC* was used as a reference gene. The same gene-specific primers were used as qualitative RT-PCR. Three independent biological replicates of each line were analyzed, essentially following the procedure described in Yuan *et al.* (2013). We determined amplification efficiencies for each primer pair using critical threshold values obtained from a dilution series (1:4, 1:20, 1:100, 1:500).

Scanning electron microscopy: Flower petal lobes and the nectar guides (the part of the ventral petal without the petal lobe) were dissected, fixed overnight in Formalin-Acetic-Alcohol (FAA) at 4°C, dehydrated for 30 min through a 50%, 60%, 70%, 95%, and 100% alcohol series. Samples were then critical-point dried, mounted, and sputter coated before being observed in a JEOL JSM-840A scanning electron microscope (University of Washington Biology Imaging Facility).

Supporting References

Jorgensen, E. C., and T. A. Geissmann, 1955 The chemistry of flower pigmentation in *Antirrhinum majus* color genotypes. III. Relative anthocyanin and aurone concentration. Arch. Biochem. Biophys. 55: 389–402.

Kurtz, S., A. Phillippy, A. L. Delcher, M. Smoot, M. Shumway *et al.*, 2004 Versatile and open software for comparing large genomes. Genome Biol. 5.

Owen, C. R., and H. D. Bradshaw, 2011 Induced mutations affecting pollinator choice in *Mimulus lewisii* (Phrymaceae). Arthropod-Plant Interact. 5: 235-244.

Scoville, A. G., L. L. Barnett, S. Bodbyl-Roels, J. K. Kelly and L. C. Hileman, 2011 Differential regulation of a MYB transcription factor is correlated with transgenerational epigenetic inheritance of trichome density in *Mimulus guttatus*. New Phytol. 191: 251-263.

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

Yuan, Y. W., J. M. Sagawa, R. C. Young, B. J. Christensen, H. D. Bradshaw, Jr., 2013. Genetic dissection of a major anthocyanin QTL contributing to pollinator-mediated reproductive isolation between sister species of *Mimulus*. Genetics doi:10.1534/genetics.112.146852

Table S1 Primers used in this study

The sequences highlighted in red indicate restriction sites.

Primer	Sequence (5'-3')
GUIDELESS_RNAi_F	GTTCTAGACCATGGTACGACGTGGAATTTCTCGGAA
GUIDELESS_RNAi_R	GTGGATCCGGCGCGCCGGCGAAGTCGGGAAATTCAGTA
pFGC5941_342F	TTGCCAACATGGGAGTCCAAGA
pFGC5941_1214R	TCGGCGTGTAGGACATGGCAA
GUIDELESS_SP3F	TAGCCGTTGGTGATTGCAGCCA
GUIDELESS_SP3R	GCCCACACTTTGAGTATGTCCA
MlMYBML1_SP3F	AGTCGCCGACTTCTACTCTGAAC
MlMYBML1_SP3R	CATGTCATGAAGCAGCTGAGTTGA
MlMYBML2_SP4F	TCGTTCGGGTCGTGTGAAGACA
MlMYBML2_SP4R	AGTTCCCTCTCATCGATTCCGAC
MlMYBML3_SP3F	TCCGAGCTTTATTGCAGATGTTGC
MlMYBML3_SP3R	TCGATTGGAGAAGGCCACGTCAT
MlMYBML4_SP3F	ACGTCATGCCCACGAACAACTAC
MlMYBML4_SP3R	GTCGTCGTAGAACTCTGGGTTAT
MlMYBML5_SP3F	CGGCGATCTTCCAACTAGTGTTC
MlMYBML5_SP3R	GCCAATTATCCATCGGTGGCGAA
MlMYBML8_SP3F	AGACGAGAAAGAGTGCAGGAGCA
MlMYBML8_SP3R	CAGGGTACAAGGTATAGCAATCAC
MlMYBML7_SP3F	CAGAGGTCGGAGAGACGTTTCGA
MlMYBML7_SP3R	ATCCAGAAGCAGCTGCAATGAAGA
MlUBC_SP3F	GGCTTGGACTCTGCAGTCTGT
MlUBC_SP4R	TCTTCGGCATGGCAGCAAGTC