

# The geography of divergence with gene flow facilitates multitrait adaptation and the evolution of pollinator isolation in *Mimulus aurantiacus*

Sean Stankowski,<sup>1,2</sup> James M. Sobel,<sup>3</sup> and Matthew A. Streisfeld<sup>1</sup>

<sup>1</sup>*Institute of Ecology and Evolution, University of Oregon, Eugene, Oregon, 97401*

<sup>2</sup>*E-mail: sstankow@uoregon.edu*

<sup>3</sup>*Department of Biological Sciences, Binghamton University, Binghamton, New York, 13902*

Received October 20, 2014

Accepted October 18, 2015

Ecological adaptation is the driving force during divergence with gene flow and generates reproductive isolation early in speciation. Although gene flow opposes divergence, local adaptation can be facilitated by factors that prevent the breakup of favorable allelic combinations. We investigated how selection, genetic architecture, and geography have contributed to the maintenance of floral trait divergence and pollinator isolation between parapatric ecotypes of *Mimulus aurantiacus*. Combining greenhouse, field, and genomic studies, we show that sharp clines in floral traits are maintained by spatially varying selection. Although adaptation breaks down where the ecotypes co-occur, leading to the formation of a hybrid zone, the largely non-overlapping distributions of the ecotypes shield them from immigrant genes, facilitating divergence across most of the range. In contrast to the sharp genetic discontinuities observed across most hybrid zones, we observed a gradual cline in genome-wide divergence and a pattern of isolation by distance across the landscape. Thus, contrary to a long period of allopatry followed by recent re-contact, our data suggest that floral trait divergence in *M. aurantiacus* may have evolved with locally restricted, but ongoing gene flow. Therefore, our study reveals how the geographic distribution of an organism can contribute to the evolution of pre-mating isolation in the early stages of divergence with gene flow.

**KEY WORDS:** Cline analysis, divergence with gene flow, divergent selection, hybrid zone, *Mimulus*, reproductive isolation.

A major goal of speciation research is to understand how reproductive barriers evolve with gene flow (Coyne and Orr 2004; Smadja and Butlin 2011; Nosil 2012; Abbott et al. 2013). Ecological adaptation is the driving force in models of divergence with gene flow and can be an important source of pre- and post-mating isolation early in the process of speciation (Sobel et al. 2010; Nosil 2012). Although isolation may result from selection on a single divergent trait (Orr 1991; Ueshima and Asami 2003), strong barriers usually consist of multiple divergent traits that each make a small contribution to the total reduction in gene flow (Coyne and Orr 2004; Nosil 2012). Thus, for strong isolation to evolve, ecologically based divergent selection must overcome or balance the homogenizing effects of gene flow that act to breakup

favorable genetic associations (Barton and de Cara 2009; Barton 2010). However, even very low levels of gene flow may be enough to suppress initial divergence or cause a previously established barrier to collapse (Barton 2010). Thus, while divergence with gene flow is feasible under a range of conditions (Lenormand 2002; Sinervo and Svensson 2002; Feder and Nosil 2009), detailed empirical studies are needed to identify the factors that ease the process in nature.

A key factor that can facilitate divergence with gene flow is the geographic distribution of the emerging taxa (Endler 1977; Feder and Nosil 2009; Barton 2013). Local adaptation can occur between populations that exist along a continuum of spatial isolation, varying from completely overlapping ranges with high gene



flow to geographically distant populations that only exchange occasional migrants (Feder and Nosil 2009; Barton 2013; Abbott et al. 2013). However, many natural systems include areas of spatial overlap and spatial separation of habitats (i.e., parapatry) (Butlin et al. 2008; Fitzpatrick et al. 2009), in some cases leading to the formation of hybrid zones (Barton and Hewitt 1985). In these situations, levels of gene flow between diverging populations will vary geographically, provided that the total range of the organism is larger than its dispersal range (Endler 1977; Butlin et al. 2008; Barton 2013). In or near areas of geographic overlap, high levels of gene flow may overwhelm selection, leading to the local breakdown of divergent adaptations (Clarke 1966; Endler 1977). However, individuals that are far from areas of overlap may be shielded from immigrant genes, which can facilitate adaptation over much of an organism's range (Barton 2013).

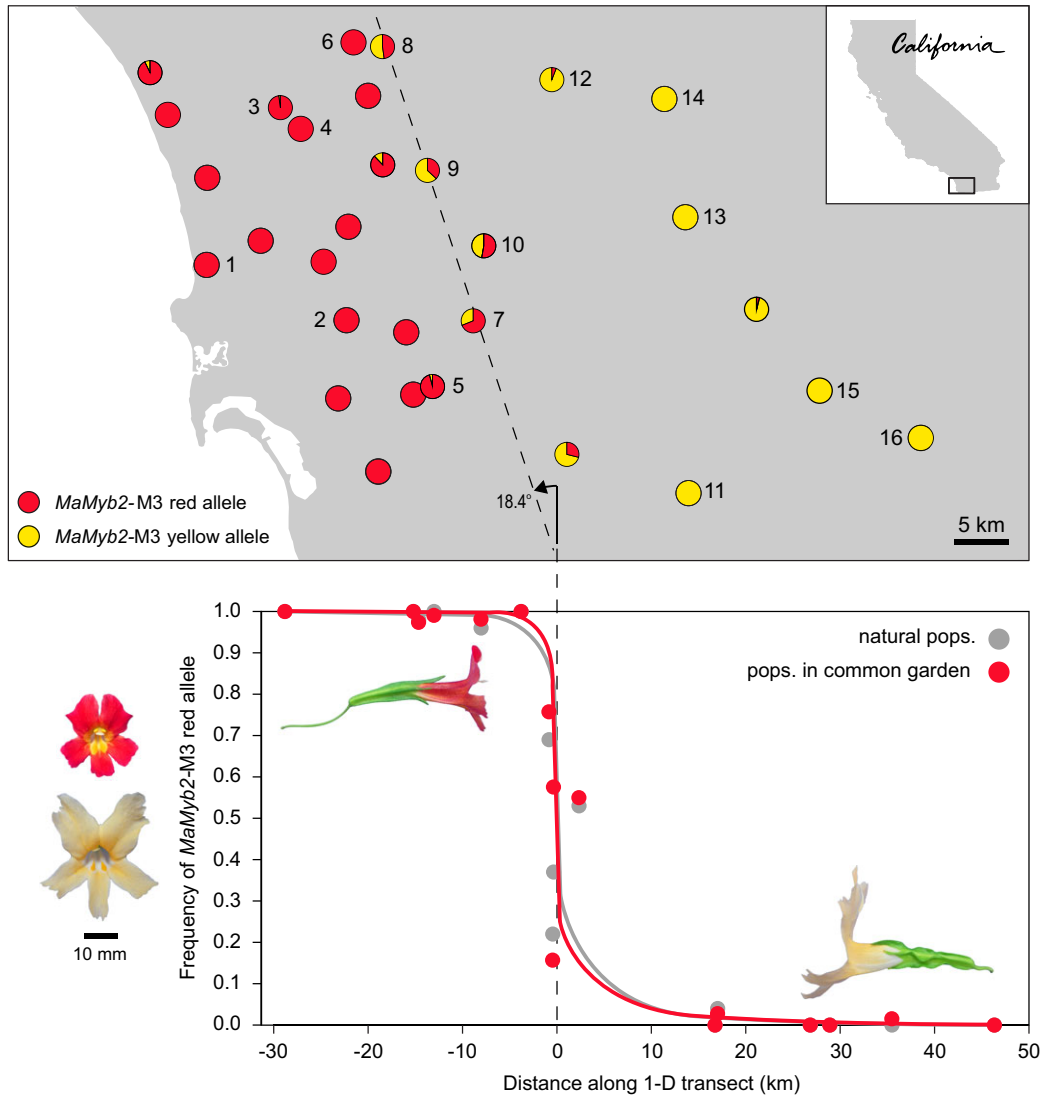
The geographic context can also change throughout the history of divergence (Butlin et al. 2008). For example, periods of parapatry may be punctuated by periods of geographic isolation (i.e., allopatry) during which gene flow is absent (Butlin et al. 2008). Although multitrait adaptation can occur during periods of uninterrupted gene flow across an ecological gradient (i.e., primary contact) (Barton and de Cara 2009; Barton 2010), periods of isolation can help initiate divergence, because associations between adaptive alleles can form more easily when gene flow is absent (Kirkpatrick and Ravigné 2002; Coyne and Orr 2004). Once they are established, divergent trait combinations may be maintained following re-contact (i.e., secondary contact) (Abbott et al. 2013). Distinguishing between alternative divergence histories based on contemporary patterns of molecular variation is notoriously difficult (Nosil 2008; Pinho and Hey 2010), because intermittent periods of allopatry may fail to leave a molecular signature. However, the presence of a sharp genome-wide discontinuity between parapatrically distributed taxa and a narrow zone of admixture suggest recent secondary contact of populations following a long period of divergence in allopatry. Alternatively, a genome-wide pattern of isolation by distance reflecting ongoing local dispersal and drift across the landscape is most consistent with a recent primary origin of parapatric taxa.

Divergence with gene flow can also be facilitated by the genetic architecture of the traits that generate isolation (Kirkpatrick 2010; Smadja and Butlin 2011; Yeaman and Whitlock 2011). This occurs when features of genetic architecture prevent the breakup of favorable trait combinations by segregation and recombination (Smadja and Butlin 2011). For example, chromosomal inversions have received considerable recent attention for their potential to shield sets of divergently adapted loci from recombination in hybrids (Noor et al. 2001; Rieseberg 2001; Butlin 2005; Kirkpatrick and Barton 2006; Feder and Nosil 2009; Kirkpatrick 2010). Similarly, pleiotropy and tight linkage of loci in collinear genomic regions can protect trait associations, but asso-

ciations due to linkage will decay over time if selection is relaxed (Smith 1966; Falconer and McKay 1996; Rieseberg 2001; Conner and Hartl 2004; Gavrillets 2004; Lovel et al. 2013). Any or all of these mechanisms can allow divergent natural selection to overcome the homogenizing effects of gene flow that would otherwise break trait combinations apart. Thus, to determine if aspects of genetic architecture may be important in facilitating divergence with gene flow, it is also necessary to examine whether traits segregate independently of each other in the absence of selection.

In this study, we investigate the roles that divergent selection, geography, genetic architecture, and history have played in the origin and maintenance of floral divergence between parapatrically distributed ecotypes of the *Mimulus aurantiacus* species complex (Phrymaceae). In San Diego County, California, there is a sharp geographic transition between red- and yellow-flowered ecotypes (Streisfeld and Kohn 2005). Phylogenomic analysis of the entire complex has shown that these ecotypes are extremely closely related to each other, with the red ecotype having evolved recently from an ancestral yellow-flowered population (Stankowski and Streisfeld 2015). The red ecotype occurs in the west and produces flowers that make red anthocyanin pigments in their petals, have short and narrow corollas, long pedicels, and exerted stigmas. The eastern, yellow ecotype produces yellow flowers with no anthocyanins, long and wide corollas, short pedicels, and inserted stigmas (Fig. 1; Waayers, 1996; Tulig, 2000; Streisfeld and Kohn 2005). Grant (1981, 1993a,b) interpreted these differences to be the result of selection to maximize visitation and pollen transfer by alternate pollinators. Consistent with these predictions, hummingbird and hawkmoth pollinators demonstrate opposing preferences and constancy for flowers of the red and yellow ecotypes, respectively, which results in strong, but incomplete reproductive isolation (Streisfeld and Kohn 2007; Sobel and Streisfeld 2015). However, it is not clear which floral traits are involved in pollinator isolation. Because postmating barriers are effectively absent (Sobel and Streisfeld 2015), there is opportunity for gene flow between the ecotypes, which hybridize in areas where they co-occur (Streisfeld and Kohn 2007).

Focusing on this classic model of pollinator-mediated divergence with gene flow, our study had two objectives. Our primary goal was to understand how selection, geography, and genetic architecture contribute to the maintenance of floral trait divergence between the red and yellow ecotypes. To assess this, we asked three questions: (1) which floral traits are experiencing divergent selection? (2) Are features of the genetic architecture that shield traits from recombination responsible for the maintenance of divergent floral morphologies? and (3) Does the spatial structure of populations play an important role in maintaining the ecotypes? To answer these questions, we measured a set of floral traits in a common garden and used cline analysis to identify traits experiencing divergent selection. We then studied trait covaria-



**Figure 1.** Geographic and clinal variation in an SNP marker tightly linked to the flower color locus *MaMyb2*. In the top panel, pie charts represent the frequency of the *red* or *yellow* allele at *MaMyb2*-M3, as defined in Streisfeld et al. (2013). The dashed line is the best-fitting two-dimensional cline center for the allele frequency cline. The 16 numbered locations are included in this study. In the bottom plot, each sampling location is collapsed onto a one-dimensional transect, based on its distance from the fitted two-dimensional center. The gray line is the maximum-likelihood one-dimensional cline for allele frequency data for the 16 natural populations. The red line is the same, but based on the allele frequencies of the seedlings from each of the 16 populations used in the common garden experiment.

tion in a greenhouse-grown  $F_2$  population to determine how gene flow between the ecotypes affects the segregation of floral traits in the absence of selection. To understand if geography facilitates local adaptation, we also analyzed trait variation in the hybrid zone, where levels of inter-ecotype gene flow and recombination are expected to be highest. An additional goal of this study was to shed light on the history of divergence in *M. aurantiacus* by asking whether there was evidence that initial floral divergence occurred during a long period of physical isolation, as proposed by Grant (1981, 1993a,b), or whether it appears that divergence evolved in the presence of ongoing gene flow? To do this, we generated a large panel of single nucleotide polymorphisms (SNPs)

and characterized patterns of genetic structure across the study area.

## Materials and Methods

### STUDY AREA AND CALCULATION OF A ONE-DIMENSIONAL TRANSECT

Although *M. aurantiacus* is distributed over a broad two-dimensional area (Fig. 1), the transition between the yellow and red ecotypes is narrow and occurs primarily in one dimension (roughly east–west). To enable subsequent one-dimensional cline analysis of floral traits, we used a widely adopted protocol to

collapse the two-dimensional location information onto a one-dimensional transect (Bridle et al. 2001). Specifically, we used the program *Analyse 1.3* (Barton and Baird 1995) to fit a two-dimensional sigmoid cline to an existing allele frequency dataset for a genetic marker in the gene *MaMyb2* from 30 sample sites (*MaMyb2*-M3; see Streisfeld et al. 2013). *MaMyb2*-M3 is tightly linked to the primary mutation responsible for the transition from yellow to red flowers in *M. aurantiacus* and shows a sharp spatial cline that matches the transition in qualitative flower color scores (Streisfeld et al. 2013). After testing models with up to five linear segments, we obtained the best-fitting two-dimensional cline (Bridle et al. 2001) and determined the minimum linear distance of each of the 30 sample sites to the two-dimensional cline center, which was set to position “0.” This polarized the new one-dimensional coordinates, so that sites in the territory of the red ecotype have negative values and sites in the territory of the yellow ecotype have positive values (Fig. 1).

### SAMPLING DESIGN

We collected seed and leaf tissue from 16 of the 30 sites from across the study area. These included six red-ecotype sites, six yellow-ecotype sites, and four “hybrid” sites as previously defined by Streisfeld et al. (2013) (Fig. 1, Table S1). Seeds were collected from 189 maternal families (range: 7–17 per sample site; mean = 11.8) in the summer of 2011. Leaves were also collected from 199 adult plants (mean  $n$  per population = 12.4; range 6–18) and stored in silica prior to DNA extraction. In June 2012, we raised plants from each of the 16 natural populations in a common garden. Seeds from each maternal family were sown on moist potting soil in plug trays and placed in a growth chamber under fluorescent light (16/8 h cycle) at 22°C. Approximately two weeks post-germination, two seedlings per family (368 seedlings) were transplanted into 2.25 inch pots and randomly placed in bottom-water trays. Plants were transported to the University of Oregon greenhouses and watered and fertilized as required.

To confirm that the samples from the 16 populations represented in the common garden experiment captured the sharp spatial transition in *MaMyb2*-M3 genotypes observed in nature (Fig. 1), we compared the shape of the marker clines in the experimental plants with those from the natural populations. We genotyped each seedling at *MaMyb2*-M3 according to Streisfeld et al. (2013), and fitted one-dimensional ML clines to each of three allele frequency datasets (the full dataset of 30 natural populations, the subset of 16 natural populations that were used in the common garden experiment, and the seedlings from the 16 populations grown in the common garden experiment). We then compared the best-fitting cline models using likelihood ratio tests (see “cline-fitting procedures” in the Supporting Information for more detail).

### PATTERNS OF TRAIT VARIATION ACROSS THE STUDY AREA

We phenotyped each greenhouse-raised plant by measuring 10 floral traits that are often associated with pollinator-mediated isolation (Grant 1993a,b; Schemske and Bradshaw 1999) (see Supporting Information for details). We then identified floral traits that showed evidence for spatial variation by performing linear regression on the mean trait values for each sample site against its distance along the transect. Those traits with a significant slope ( $P < 0.05$ ) were considered to show evidence for clinal variation.

For these traits, we used *Analyse 1.3* (Barton and Baird 1995) to fit maximum-likelihood cline models to the mean trait values for each population. Data were scaled between 0 and 1 as required by the program (see “cline-fitting procedures” in the Supporting Information for detail). We compared the fits of stepped (A-step and S-step) and “sigmoid” cline models and found that the sigmoid model was the optimum model for all six traits (see Results). The sigmoid model is described by four parameters: (1)  $P_{\min}$  and (2)  $P_{\max}$ , the mean trait values in the tails of the cline, (3)  $c$ , the geographic position of the cline center, and (4)  $w$ , the cline width, defined as the ratio between the change in the mean trait value in the tails ( $\Delta P$ ) and the maximum slope. For a single locus,  $w$  is inversely proportional to the effective strength of selection required to maintain the cline ( $s^*$ ), so that narrower widths indicate greater selection (Barton and Gale 1993). However, for quantitative traits,  $w$  may also be influenced by the number of genes underlying the trait, so we did not make inferences about relative strengths of selection based on cline width and instead focused primarily on variation in cline centers. We obtained ML estimates of each cline parameter and their two-unit support limits based on 10,000 random changes of the focal parameter.

To examine Grant’s (1981, 1993a,b) hypothesis that the floral traits are diverged due to selection by pollinators, we tested for the coincidence of cline centers among the floral trait clines. Assuming that a cline is at equilibrium and maintained by a balance between dispersal and environmental selection, the cline center ( $c$ ) represents the geographic position where the direction of divergent selection switches (Endler 1977; Barton and Gale 1993). Thus, if pollinator selection is responsible for divergence of floral traits, we expected the clines to share a common center. The coincidence of centers was tested using a likelihood profile method as described by Phillips et al. (2004) and implemented by Kawakami et al. (2009) and described in detail in the Supporting Information.

### EFFECT OF INTERE-COTYPE GENE FLOW ON FLORAL ARCHITECTURE

Although the floral morphologies may be maintained entirely by selection acting on a suite of genetically independent traits, it is also possible that trait associations are maintained by features of

the genetic architecture that protect loci from recombination in hybrids. If genetic architecture does not assist in maintaining trait associations, we expect traits to segregate independently following inter-ecotype gene flow in the absence of selection. To test this hypothesis, we examined trait covariation in an outcrossed  $F_2$  population grown in a common environment. Three hundred eighty plants were produced by crossing two  $F_1$  individuals, each the product of crossing different greenhouse-raised red- and yellow-ecotype plants collected from opposite ends of the collapsed transect (locations 1<sub>red</sub> and 16<sub>yellow</sub> in Fig. 1). To allow direct phenotypic comparison among plants grown in a common environment, we also raised 25 red-ecotype (location 1; Fig. 1), 31 yellow-ecotype (location 16), and 20  $F_1$  (1<sub>red</sub> × 16<sub>yellow</sub>) plants alongside the  $F_2$ s. Plants were raised and phenotyped as described above.

We analyzed trait covariation in two ways. First, we calculated Pearson's correlation coefficient between each pair of traits and tested the significance of the relationship against the null hypothesis that variation in trait  $i$  is inherited at random with respect to variation at trait  $j$ . Probability values were estimated by comparing the observed correlation coefficient to a null distribution of correlation coefficients generated by 1,000,000 random permutations of the data using custom scripts in R. A non-significant correlation provides evidence that a pair of traits can be broken apart by inter-ecotype segregation and recombination in the absence of selection. To provide a multivariate view of phenotypic variation, we also performed a principal components analysis (PCA) on the trait data, including the red, yellow,  $F_1$ , and  $F_2$  individuals. If traits are controlled by a small number of loci and co-localize to the same genomic region, we expected to see  $F_2$  individuals that routinely group with those of red- and yellow-ecotype plants based on PC scores. Alternatively, if  $F_2$  individuals do not group with the parents, we can conclude that the pure floral phenotypes are rarely re-created after a single generation of segregation and recombination in the absence of selection.

#### **PATTERNS OF FLORAL TRAIT VARIATION IN THE HYBRID ZONE**

The hybrid zone between the red and yellow ecotypes provides an opportunity to test if the divergent floral morphologies are maintained directly in the face of gene flow, because the rate of inter-ecotype re-combination is expected to be highest in areas where the ecotypes co-occur. This could occur if, for example, intermediate hybrid phenotypes are rarely produced and are visited less frequently than pure phenotypes, or if features of their genetic architecture protect trait associations from breakup in hybrid offspring. In this case, we expect to see a bimodal distribution of phenotypic traits that is consistent with the coexistence of pure red- and yellow-ecotype individuals and selection against intermediate phenotypes (Jiggins and Mallet 2000). Alternatively, the rate of gene flow between the ecotypes may be higher than the strength of selection, resulting in the breakdown of the divergent

morphologies. In this case, we expect to see a broad distribution of phenotypes inside the hybrid zone that falls roughly intermediate of the pure ecotypes (Jiggins and Mallet 2000).

To test these expectations, we measured floral traits in the field from 192 individuals from the four locations in the hybrid zone (48 individuals from locations 7 to 10 in Fig. 1). We also phenotyped individuals from four locations outside of the hybrid zone (locations 1<sub>red</sub>, 4<sub>red</sub>, 11<sub>yellow</sub>, and 16<sub>yellow</sub>) so that we could compare pure and hybrid phenotypes that were measured in the same flowering season. We focused only on traits that showed clinal variation (see Results). We then conducted PCA, including plants from inside and outside the hybrid zone. We saved the first two principal components and conducted a distributional analysis of each. Specifically, we fit unimodal (representing predominantly hybrids), bimodal (representing predominantly pure parental types), and trimodal (representing a combination of pure parental types and hybrids) distributions to each PC using maximum likelihood in JMP version 11, and used the Akaike Information Criterion (AIC) to select the optimum model.

#### **MOLECULAR METHODS AND ANALYSIS OF GENOME-WIDE DIVERGENCE**

To enable an analysis of genome-wide divergence across the distribution of the ecotypes, we identified SNPs by sequencing restriction site associated DNA tags (RAD-seq) generated from the 199 adult tissue samples collected across the 16 natural populations (Table S1). RAD-seq library preparation and Illumina HiSeq 2000 sequencing followed the methods described in Sobel and Streisfeld (2015). We used *Stacks version 1.12* (Catchen et al. 2013) to process the data and call genotypes, as described in the Supporting Information.

We used patterns of genome-wide variation to examine two hypotheses related to the history of gene flow between the ecotypes. First, we tested the hypothesis that clinal variation in the floral traits has been maintained by selection despite ongoing gene flow across the landscape. According to this hypothesis, we expect the average pattern of genome-wide divergence to show a cline that is wider than the cline in each floral trait. Although a large panel of anonymous SNPs may contain a mix of neutral, directly selected, and indirectly selected loci, a previous analysis of RAD data between these ecotypes revealed that most markers showed little or no differentiation (Sobel and Streisfeld 2015). Thus, the overwhelming signal in the data should reflect neutral evolutionary processes.

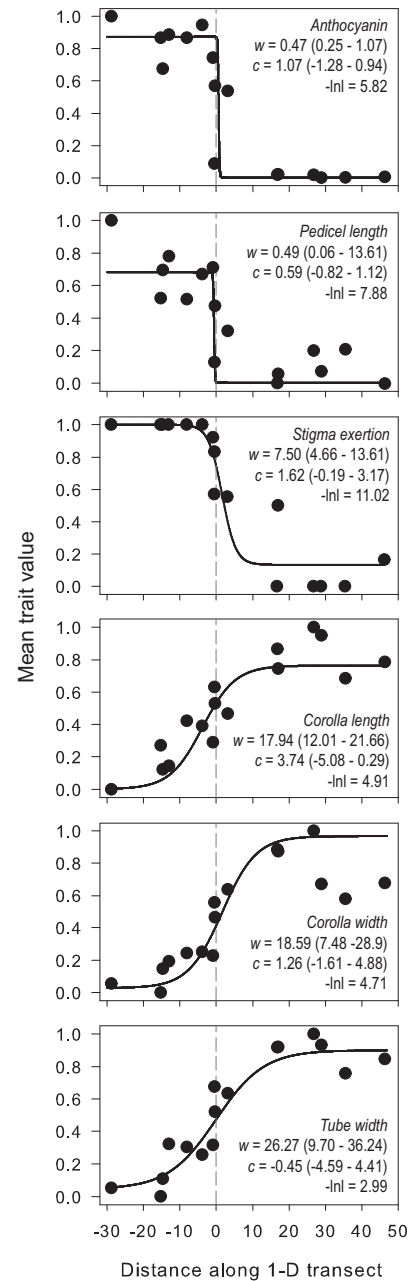
We obtained estimates of genome-wide divergence and population structure from two sources. First, we conducted a PCA on the genotype matrix using the R package *Adegenet* and calculated mean PC1 and PC2 scores for each population. Second, we used the Bayesian clustering algorithm implemented in *Structure 2.3.4* (Pritchard et al. 2000) to estimate the probability of assignment of each individual to one of two clusters (see “*Structure*



analysis” in the Supporting Information for more detail). As the mean PC1 and *Structure Q* scores were highly correlated with one another (see Results), cline analysis was only performed on the PC1 scores. Cline fitting was conducted using the quantitative trait model as described above, based on the mean and variance of scaled PC1 scores at each sample site. Pairwise likelihood profile tests were used to determine if each floral trait had a cline width that was narrower than the cline width obtained for the genomic PC1 scores, using methods described in the Supporting Information.

Second, we examined the hypothesis posed by Grant (1981, 1993a,b) that the floral traits diverged during an extensive period of geographic isolation followed by secondary contact. Under this scenario, we would expect to see: (1) a sharp cline in molecular genetic variation that is consistent with an abrupt genetic discontinuity between two divergent taxa; (2) elevated variance in molecular admixture scores in populations where the phenotypic hybrids are present compared to populations where pure red and yellow ecotypes are found; and (3) estimates of genome-wide divergence that are greater between the ecotypes than the levels of divergence observed within the ecotypes. The analyses of population structure and clinal variation described above provide tests of predictions 1 and 2. However, to examine point 3, it is important to correct for any effects of isolation by distance that arise because the geographic distances from inter-ecotype comparisons (i.e., red × yellow comparisons) are greater than intra-ecotype comparisons (i.e., red × red and yellow × yellow comparisons). If patterns of genomic differentiation between the ecotypes remain greater than differentiation within the ecotypes after accounting for the geographic distance between sample sites, this would support Grant’s (1993a,b) hypothesis that floral divergence evolved during a period of allopatry.

To test this final expectation, we first estimated pairwise  $F_{ST}$  (averaged over all SNP loci) between each pair of sample sites in *Arlequin* (Excoffier et al. 2005), excluding the four “hybrid” locations (sites 7–10 in Fig. 1). Mean estimates of pairwise  $F_{ST}$  within and between ecotypes were compared using permutation tests conducted with custom scripts in R (described in the Supporting Information). We then controlled for geography by examining the relationship between genomic differentiation and geographic distance. We regressed the pairwise estimates of  $F_{ST}$  against estimates of geographic distance between each pair of sample sites and tested the significance of the relationship using a Mantel test using the *Vegan* package in R. After confirming that the inter-ecotype and intra-ecotype comparisons had near identical scaling relationships, we then obtained the residual estimates of pairwise  $F_{ST}$  from the full regression and tested whether the mean residual variation of inter-ecotype comparisons was greater than intra-ecotype comparisons using a permutation test.



**Figure 2.** Clinal variation in six floral traits across the one-dimensional transect. The line in each plot is the maximum-likelihood one-dimensional cline fitted to scaled population mean estimates. The dashed vertical line shows the position of the cline center for the *MaMyb2-M3* marker. Maximum-likelihood estimates of cline width ( $w$ ), cline center ( $c$ ), and the model’s log likelihood ( $-\ln l$ ) are reported for each trait.

## Results

### ORIENTATION OF THE ONE-DIMENSIONAL TRANSECT AND CLINE IN *MAMYB2* ALLELE FREQUENCY

The best-fitting two-dimensional sigmoid cline model for the *MaMyb2-M3* allele frequency data ( $\text{LnL} = -21.66$ ) consisted

of a single linear segment with an orientation 18° anticlockwise of the north-south axis (Fig. 1). Cline models with more than one linear segment did not lead to a significant improvement in the log-likelihood score (likelihood ratio test (LRT);  $P > 0.05$ ). Fitting one-dimensional clines to the three allele frequency datasets resulted in cline models with nearly identical characteristics (Fig. 1; Table S2). In all cases, cline shapes were not significantly different among datasets ( $P > 0.05$ ), confirming that the samples in our common garden experiment captured the sharp spatial transition observed between the ecotypes in nature.

### CLINAL VARIATION IN THE FLORAL TRAITS

Linear regression of the phenotypic data in the common garden experiment revealed that six of the 10 floral traits we measured showed significant spatial variation across the study area (Table S3). Populations on the western end of the transect produced more anthocyanin, had narrower floral tubes, shorter and narrower corollas, longer pedicels, and more frequently had exerted stigmas than populations in the east.

As expected for traits maintained by a balance between gene flow and divergent selection, the clines in the floral traits were well described by a sigmoid model (Fig. 2). Of the four-, six-, and eight-parameter cline models implemented in *Analyse*, the four-parameter sigmoid model was the best fit for all six traits (LRT;  $P > 0.05$ ). Parameter estimates for the cline center along the collapsed transect were relatively consistent for each trait, ranging from  $-2.36$  km for corolla length to  $+1.65$  km for corolla width, with a mean estimate of  $-0.06 \pm 1.5$  km relative to the *MaMyb2-M3* cline center. Consistent with the hypothesis that these traits have diverged due to selection by pollinators, the cline centers were not significantly different from one another ( $\Delta ML = 8.24$ ,  $df = 5$ ,  $P = 0.144$ ). In contrast to the coincident cline centers, we observed striking variation in cline width (Fig. 2). Estimates of  $w$  varied more than 50-fold among traits, ranging from less than 1 km for floral anthocyanin and pedicel length to more than 26 km for tube width.

### TRAIT VARIATION IN EXPERIMENTAL CROSSES

One generation of segregation and recombination resulted in a broad and continuous distribution of floral phenotypes in the  $F_2$  generation (Fig. 3). Strong correlations were observed among traits affecting the size and shape of the corolla, including corolla length, corolla width, and corolla width (mean  $r = 0.58$ ;  $P < 10^{-6}$ ; Table 1). However, our analysis revealed greatly reduced correlations among the other trait pairs. Specifically, four of the five pairwise comparisons involving floral anthocyanin were substantially reduced (mean absolute  $r = 0.09$ ) and were not significantly different from our null expectation after correction for multiple comparisons (Bonferroni-corrected  $\alpha = 0.0034$ ). Nevertheless, six of the pairs of traits revealed statistically significant corre-

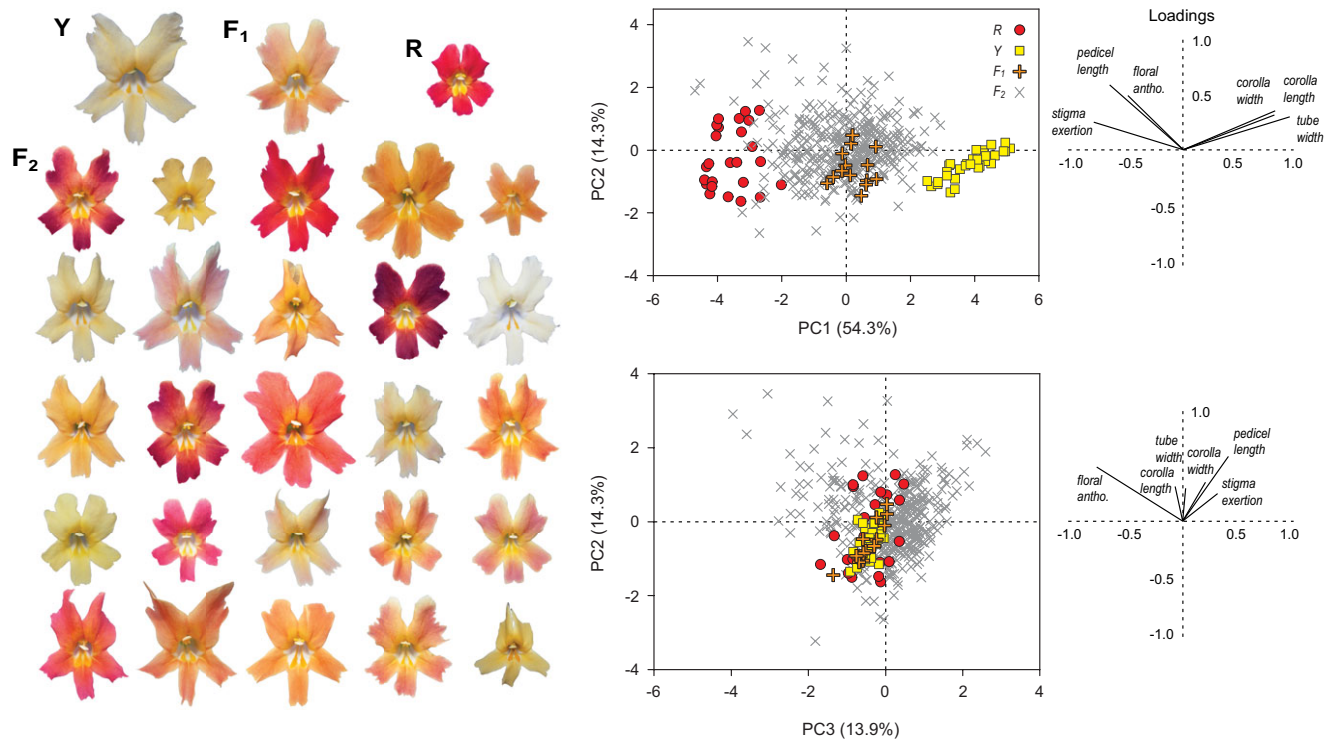
lations (Table 1). However, PCA of all six traits revealed that the parental combinations of phenotypes were rarely recovered in the  $F_2$ . Although the red- and yellow-ecotype plants had distinct, non-overlapping distributions along PC1,  $F_2$  individuals generally fell in the valley between them (Fig. 3). In addition to the abundance of intermediate phenotypes, we also observed transgressive variation in the  $F_2$  that was not present in the red-ecotype, yellow-ecotype, or  $F_1$  plants. This was evident in the bivariate plot of PCs 2 and 3, where the red-ecotype, yellow-ecotype, and  $F_1$  individuals overlapped considerably, but were surrounded by a broad cloud of  $F_2$  individuals.

### TRAIT VARIATION IN THE HYBRID ZONE

Although the six floral traits are maintained in distinct combinations across most of each ecotype's range, trait data collected from inside the hybrid zone are consistent with considerable breakdown of floral architectures in areas where the ecotypes co-occur (Fig. 4). The first two principal components together explain 73% of the variation in the dataset (Fig. 4). For individuals sampled outside the hybrid zone, frequency histograms and ML analysis revealed that each PC was bimodally distributed, with near complete separation of individuals from each ecotype (Fig. 4). In contrast, PC scores for individuals from inside the hybrid zone showed broad unimodal distributions that were centered roughly intermediate of the modes of the red and yellow ecotypes. Complete separation of the red and yellow ecotypes was observed when individuals were plotted in a two-dimensional PC space (Fig. 4), with the majority of individuals from the hybrid zone falling in the space between them. Some individuals from the hybrid zone had phenotypes that were consistent with the presence of "pure" red- and yellow-ecotype plants in the zone, or with later generation hybrids that have parental combinations of traits. However, the extensive trait variation observed in the hybrid zone reveals that the parental trait combinations are not broadly maintained directly in the face of gene flow.

### PATTERN OF GENOME-WIDE DIVERGENCE

Illumina sequencing of RAD tags from 199 samples collected from the 16 natural populations provided a total of 5382 SNPs that met our filtering requirements. The principal components and *Structure* analyses revealed very similar patterns of population structure. In the PC analysis, the first two principal components explained 14% of the variation in the SNP genotype matrix (9.1 and 5.1% for PCs 1 and 2, respectively). Although there was considerable variation in PC1 scores within sites, the PC1 axis separated the populations of each ecotype, which were bridged by the four hybrid populations (Fig. S1). PC2 did not contribute to the separation of sites by ecotype and was instead associated with variation among yellow-ecotype sites (Fig. S1). In the *Structure* analysis,  $K = 2$  was the optimum number of clusters ( $\Delta K = 3102$ ;



**Figure 3.** Standard photographs from a representative subset of  $F_2$  flowers and bivariate plots of the first three principal components extracted from a multivariate analysis of floral trait variation in red-ecotype, yellow-ecotype,  $F_1$ , and  $F_2$  plants raised in a common environment. The vectors in loading plots indicate the strength and direction of the correlation between each floral trait and the principal components.

**Table 1.** Correlations between each pair of floral traits in the  $F_2$  population.

	Floral anthocyanin	Pedicel length	Stigma exertion	Corolla length	Corolla width	Tube width
Floral anthocyanin	—	0.039	0.1019	0.0865	0.0007	0.0070
Pedicel length	0.092	—	$1.00 \times 10^{-6}$	0.0022	0.0671	$1.70 \times 10^{-5}$
Stigma exertion	0.066	<b>-0.334</b>	—	$1.00 \times 10^{-6}$	0.0004	$1.00 \times 10^{-6}$
Corolla length	-0.071	-0.147	<b>-0.331</b>	—	$1.00 \times 10^{-6}$	$1.00 \times 10^{-6}$
Corolla width	<b>-0.165</b>	-0.077	<b>-0.172</b>	<b>0.374</b>	—	$1.00 \times 10^{-6}$
Tube width	-0.130	<b>-0.214</b>	<b>-0.318</b>	<b>0.659</b>	<b>0.712</b>	—

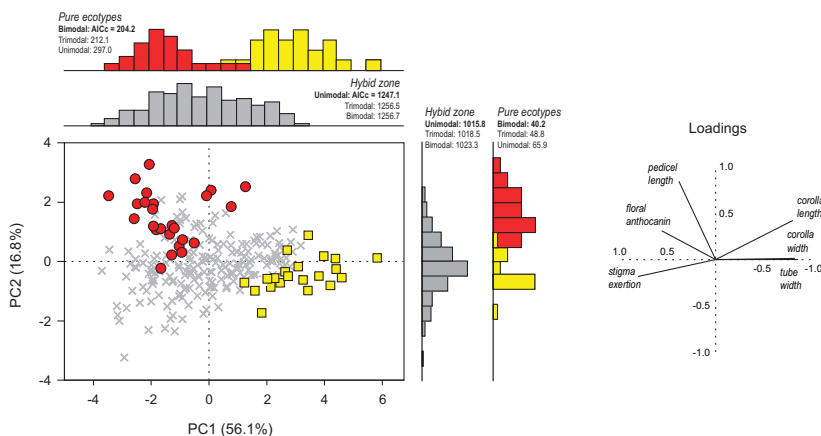
Pairwise correlation coefficients are in the lower diagonal.  $P$ -values, estimated based on 1,000,000 random permutations of the data, are in the upper diagonal. Correlation coefficients that are significant after correction for multiple comparisons (Bonferroni-corrected  $\alpha = 0.0034$ ) are bolded.

Table S4). However, rather than individuals of the “pure” red and yellow ecotypes showing strong assignment to the alternative clusters, with a sharp step in  $Q$  scores along the transect, we observed a gradual decrease in assignment scores from west to east. Also, there was no increase in the standard deviation of  $Q$  scores in the four hybrid populations (mean = 0.12; range = 0.06–0.18) relative to the pure populations (mean = 0.12; range = 0–0.33) (Wilcoxon test:  $z = 0.788$ ,  $P = 0.4306$ ) (Fig. S2).

As the mean *Structure* and PC1 scores were highly correlated ( $r^2 = 0.951$ ), we only performed cline analysis on the PC1 scores. Cline analysis revealed a gradual change in genome-wide variation across the landscape, rather than a

sharp genetic discontinuity between two genetic groups. The four-parameter sigmoid cline model was selected over the more complex six- and eight-parameter models (LRT;  $P > 0.05$ ) but gave an estimated cline width ( $w$ ) of 97 km, which was wider than the 75 km transect. As  $w$  is defined based on the maximum slope of the cline (Barton and Gale 1993), a width wider than the transect indicates that the cline is best described by a linear gradient. Subsequent linear regression revealed that the distance of sites across the one-dimensional transect explained 89% of the variation in mean PC1 scores (Fig. 5A; for mean *Structure* scores,  $r^2 = 0.91$ ; Fig. S2). Pairwise tests for concordance confirmed that the floral trait clines were significantly narrower





**Figure 4.** Principal components analysis of floral trait variation inside and outside the hybrid zone. Univariate histograms show the distributions of PC1 and PC2 scores based on the six floral traits that show clinal variation. The AIC scores for distribution models with one, two, or three normal modes are provided. Red and yellow bars show the distributions for the red and yellow ecotypes respectively. The scatter plot shows the distribution of individual scores within the bivariate PC space (red circles: red ecotype; yellow squares: yellow ecotype; gray x: hybrid). The vectors in the loading plot (below) indicate the strength and direction of the correlation between each floral trait and the principal components.

than the cline in genomic PC1 scores ( $P < 0.0003$  for all traits; Table S5).

Consistent with the broad linear cline in the genomic PC1 scores,  $F_{ST}$  analysis revealed a clear pattern of isolation by distance across the study area. There was a significant, positive relationship between pairwise  $F_{ST}$  and pairwise geographic distance between sites (Fig. 5B;  $r^2 = 0.281$ , Mantel test:  $P = 0.0002$ ). Separate analyses revealed nearly identical relationships for intra-ecotype (i.e., red  $\times$  red and yellow  $\times$  yellow) and inter-ecotype comparisons (i.e., red  $\times$  yellow) (Fig. S3). The mean estimate of inter-ecotype pairwise  $F_{ST}$  was significantly greater than the mean of intra-ecotype comparisons (Fig. 5C;  $P = 0.007$ ), but after correcting for geographic distance between sample sites, there was no difference in mean residual  $F_{ST}$  within or between the ecotypes (Fig. 5C;  $P = 0.759$ ). These results indicate that the elevated mean inter-ecotype  $F_{ST}$  can be attributed entirely to the greater geographic distance between red- and yellow-ecotype sites relative to intra-ecotype sites, which does not support a history of prolonged allopatry followed by recent secondary contact.

## Discussion

Although recent studies have shown that divergence with gene flow is common in nature (Coyne and Orr 2004; Nosil 2008; Smadja and Butlin 2011), the factors that facilitate the process are often unknown. In this study, we investigated the relative roles that selection, geography, genetic architecture, and history have played in the origin and maintenance of floral divergence between ecotypes of *M. aurantiacus*. These data illustrate

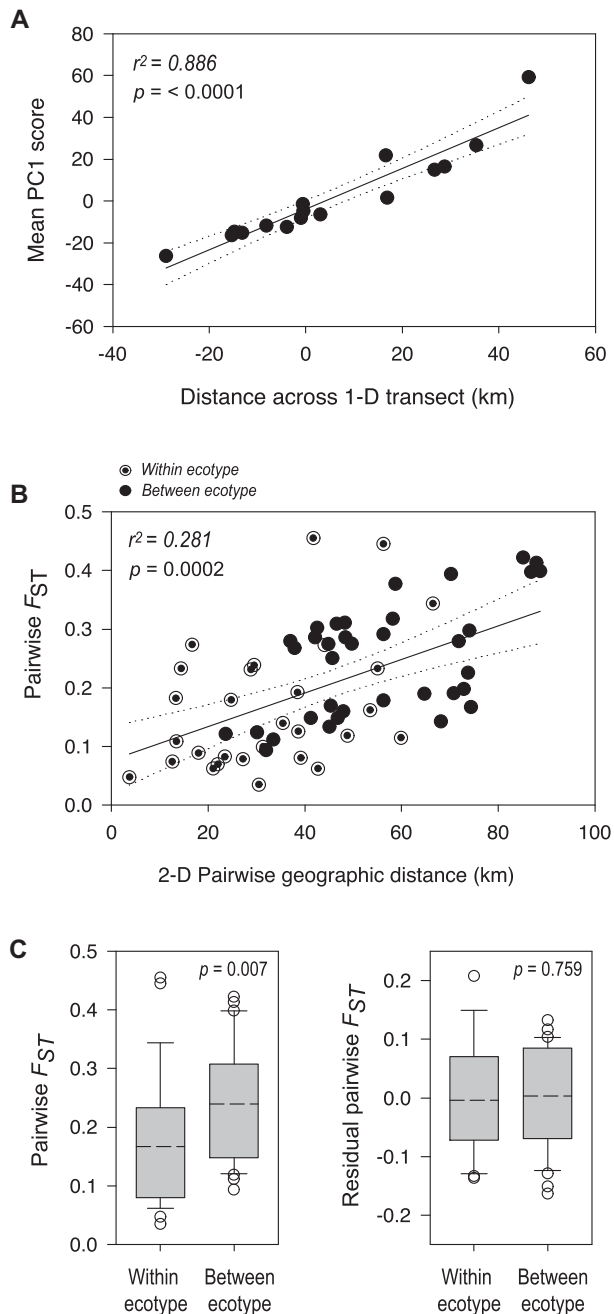
how the spatial pattern of populations can be critical for the evolution of pre-mating isolation during the early stages of speciation.

## SELECTION ON FLORAL TRAITS

We observed sharp geographic clines for six of the ten floral traits that we measured. Although clines can form in neutral traits following changes in historical distributions (Endler 1977; Barton and Hewitt 1985), all six clines are significantly narrower than the broad linear cline in genome-wide divergence. This pattern indicates that trait divergence has been maintained due to divergent selection despite a history of gene flow between the ecotypes (Barton and Hewitt 1985; Barton and Gale 1993).

In addition, our cline analyses support the hypothesis that these floral traits have diverged due to selection by different pollinators. Specifically, the divergent floral trait clines share a common center, as is expected when multiple traits diverge across a common selective gradient (Barton and Hewitt 1985; Barton and Gale 1993). Previous field experiments have shown that hummingbird and hawkmoth pollinators display strong, opposing preferences for flowers of each ecotype (Streisfeld and Kohn 2007), suggesting that floral traits are adapted for successful and efficient pollination (Grant 1949, 1981, 1994; Fenster et al. 2004).

Although the evidence for pollinator-mediated divergence is compelling, cline centers may become coincident for other reasons. For example, coincident cline centers can arise by neutral processes involving recent secondary contact (Barton and Hewitt 1985; Barton and Gale 1993). In addition, multiple selective gradients that all change in the same location may have caused floral trait divergence (Barton and Hewitt 1985; Barton and Gale 1993).



**Figure 5.** Genome-wide variation across the study area. (A) Mean genome-wide PC1 score across the collapsed transect, extracted from a principal components analysis conducted on the SNP genotype matrix (5382 loci). The line is the least-squares regression line and the dotted line is the 95% confidence interval. (B) Relationship between geographic and genetic distance ( $F_{ST}$ , averaged over loci) among sample locations. inter-ecotype (i.e., red ecotype  $\times$  yellow ecotype) and intra-ecotype (red  $\times$  red ecotype and yellow  $\times$  yellow ecotype) comparisons are coded using different symbols. (C) Levels of intra- and inter-ecotype  $F_{ST}$  before (left) and after (right) correcting for the relationship between geographic and genetic distance. The dashed line in the box represents the mean, the whiskers show the 2.5 and 97.5 quantiles, and the circles are outliers.

Clines also tend to be attracted to areas of low population density, causing them to become coincident (Barton and Hewitt 1985; Bierne et al. 2011). However, these alternative explanations seem unlikely to explain the coincidence of clines in this system. First, our genome-wide analyses do not support a history of recent secondary contact (see below), suggesting that the cline centers have not become associated as a result of recent distributional changes. In addition, there is currently little evidence to support other agents of selection affecting the floral traits. For example, it is often suggested that transitions in flower color are due to selection imposed by non-pollinator agents that act on pleiotropic effects of flower color genes (Strauss and Whittall 2006). However, little direct evidence exists for this conclusion (Rausher 2008), and previous field experiments were unable to detect fitness advantages for seedlings of each ecotype in their native environment (Streisfeld and Kohn 2007). Moreover, with six different floral traits all coinciding in the same geographic position, it seems very unlikely that multiple non-pollinator agents of selection drive the pattern. Finally, there is no obvious change in population density along the landscape that coincides with the hybrid zone, suggesting that the clines are unlikely to be positioned in a density trap (Barton 1979; Bierne et al. 2011). Thus, while future selection studies that experimentally manipulate floral traits are required to identify the direct targets and agents of selection across the transect, our data support Grant's (1981, 1993a,b) hypothesis and previous field experiments (Streisfeld and Kohn 2007) which suggest that these divergent floral architectures have evolved due to selection by pollinators.

In contrast to the coincident cline centers, we observed extensive variation in cline width, with  $w$  varying more than 50-fold among traits. For a single locus cline,  $w$  is inversely proportional to the effective selection acting on a locus ( $s^*$ ), which makes it possible to infer variation in the strength of selection among loci. Although it is tempting to attribute the variation in cline width among these floral traits to variation in the strength of selection acting among them, inferences made from quantitative traits must also take into account the number of genes affecting the trait. Thus, detailed field studies that measure the strength of selection on each trait will be the most informative way to determine how each trait contributed to pollinator isolation between the ecotypes.

### EFFECT OF INTER-ECOTYPE GENE FLOW ON FLORAL ARCHITECTURE

Recent theoretical and empirical studies have shown that divergence with gene flow can be facilitated by the genetic architecture of adaptive traits (Kirkpatrick 2010; Smadja and Butlin 2011; Yeaman and Whitlock 2011). This occurs when features of genetic architecture prevent the breakup of favorable trait combinations by segregation and recombination in hybrids (Smadja and Butlin 2011). For example, floral traits associated with pollinator

isolation between *M. lewisii* and *M. cardinalis* do not segregate independently, and comparative linkage mapping reveals that the loci controlling them reside primarily in a few inverted regions of the genome that show suppressed recombination (Fishman et al 2013). Thus, it is possible that genetic correlations among floral traits in *M. aurantiacus* could have helped to establish and maintain divergence.

Our data reveal that genetic architecture may have been important in the maintenance of some, but not all, divergent trait combinations. Trait pairs related to floral size and shape are tightly correlated with each other in the  $F_2$  population, suggesting that, on average, the same genes or genomic regions that make corollas longer are also responsible for making them wider. Moreover, correlations in six additional trait pairs are statistically significant after Bonferroni correction, suggesting the possibility that genetic architecture may have helped to establish initial floral trait divergence in this system. However, the maintenance of these associations likely requires extensive selection in the face of gene flow, as weak correlations imply that recombinant phenotypes are produced regularly. For the remaining trait combinations, particularly those involving floral anthocyanins, correlations are even weaker, revealing that associations can be broken down by one generation of segregation and recombination where selection is absent. Indeed, parental trait combinations involving all of the traits were rarely recovered in the  $F_2$ . Although this can be due to polygenic inheritance of these traits, the weak correlations and transgressive phenotypes observed in the  $F_2$ s suggest that low levels of inter-ecotype gene flow are sufficient to dismantle the parental trait associations found outside the hybrid zone. Although establishing the number of loci controlling these traits and determining the extent of genetic independence among traits will require future quantitative trait locus (QTL) mapping, the data presented here continue to support a prominent role for natural selection in the maintenance of divergent floral architectures.

### SPATIALLY MEDIATED DIVERGENCE WITH GENE FLOW

Even though divergent floral architectures are generally broken apart by recombination in the absence of selection, it is possible that selection may be strong enough to maintain the trait combinations in areas where they co-occur. However, rather than observing a bimodal distribution of phenotypes inside the hybrid zone, we observed a broad range of intermediate phenotypes that is consistent with the breakdown of “pure” red and yellow floral architectures following hybridization.

Because selection by pollinators seems to be the major force driving trait divergence in this system, there are two general explanations for this pattern. First, it is possible that selection is absent, relaxed, or operates differently within the hybrid zone, but is strong in other areas. Relaxed selection in the zone may

occur if pollinator preferences are influenced by the local frequencies of divergent phenotypes (i.e., selection is frequency dependent) (Mallet and Barton 1989). The second explanation is that the strength of pollinator-mediated selection on the floral traits is similar inside and outside of the hybrid zone, but where the ecotypes come into contact the rate of inter-ecotype gene flow is higher than the strength of selection. Although we do not have evidence to refute the first hypothesis, current and previous data support the second explanation. First, pollinator preferences are imperfect (Streisfeld and Kohn 2007; Sobel and Streisfeld 2015), which would result in some mating between the ecotypes where their ranges overlap. Second, the dismantling of floral architectures in our  $F_2$  population suggests that only very limited recombination is required to break apart many of the divergent trait combinations. Third, the absence of postmating barriers (Sobel and Streisfeld 2015) suggests that intrinsic selection against early generation hybrids is weak, implying that intermediate phenotypes will persist in nature even if they are less likely to be visited by pollinators.

Therefore, it appears that the geographic context of divergence plays a key role in the maintenance of these floral differences. The potential for geography to facilitate divergence with gene flow is well known (Coyne and Orr 2004; Butlin et al. 2008; Barton 2010, 2013). For example, in models of parapatric divergence, levels of gene flow between diverging populations are highest where habitats abut or overlap (Endler 1977; Barton 2010). In these areas, high levels of gene flow may overwhelm selection, leading to the erosion of local adaptation via gene swamping (Lenormand 2002). However, if the total area occupied by the organism is large relative to the average inter-generational dispersal distance, the level of gene flow between diverging populations decreases rapidly with increasing distance from contact zones (Barton 2010). Thus, demes that are far from areas of geographic overlap are shielded from immigrant genes, so that divergence can occur with only moderate or episodic selection (Barton 2010).

Although we do not have direct estimates of dispersal rates in *M. aurantiacus*, and cannot estimate dispersal using cline theory which assumes that individuals are outcrossing with purely diploid dispersal (Barton and Gale 1993), the pattern of isolation by distance that we observed across the broad two-dimensional landscape is consistent with geographically restricted gene flow. This suggests that most dispersal occurs between neighboring or nearby demes rather than over large distances. Consistent with these expectations, seeds in *M. aurantiacus* disperse due to gravity, which likely prevents frequent long-distance movement (Beeks 1962). Similarly, haploid pollen dispersal by hummingbirds and hawkmoths is likely to occur over relatively small spatial scales (tens to hundreds of meters) (Schlising and Turpin 1971; Webb and Bawa 1983; Linhart et al. 1987; Schlucke and Waser 2001; Johnson and Galloway 2008). Thus, while occasional

long-distance dispersal may occur, average dispersal distances are almost certainly much smaller than the total geographic range occupied by the ecotypes. These results indicate that the spatial arrangement of populations has facilitated floral divergence between the ecotypes despite the breakdown of adaptation in areas where there are opportunities for hybridization. Therefore, the stability of floral trait differences may also be influenced by other forms of local adaptation that limit the pattern of gene flow across this landscape, such as adaptation to the different abiotic environments that may cause ecogeographic isolation (Sobel and Streisfeld 2015).

### HISTORY OF FLORAL TRAIT DIVERGENCE

An additional goal of our study was to shed light on the history of trait divergence between these ecotypes. Although associations between adaptive loci can form during periods of uninterrupted gene flow across an ecological gradient (i.e., primary contact) (Barton and de Cara 2009; Barton 2010), periods of geographic isolation are thought to help establish divergent adaptations, which may be maintained following secondary contact. For this reason, Grant (1981, 1993a,b) believed that floral differences between the ecotypes likely evolved during a period of allopatry, with the hybrid zone representing a region of secondary contact. Distinguishing between primary and secondary contact is often extremely difficult because these alternative models can result in identical cline shapes (Endler 1977; Durrett et al. 2000) and similar patterns of genome-wide differentiation (Poelstra et al. 2014). In addition, the primary and secondary models represent extreme scenarios; current geographic distributions could reflect complex histories involving intermittent periods of allopatry that facilitated adaptive divergence. Although these additional scenarios may play an important role in driving initial divergence, they are also extremely difficult to detect (Barton and Hewitt 1985; Coyne and Orr 2004; Butlin et al. 2008). Although we cannot fully test these complex scenarios with current data, the pattern of genome-wide divergence between the ecotypes is in stark contrast to those observed across most well-studied hybrid zones, and it is not consistent with this hybrid zone having formed due to recent secondary contact following a long period of isolation.

Most well-studied zones of recent secondary contact are associated with a sharp molecular genetic discontinuity where the distributions of the divergent taxa abut and show clear evidence of admixture between these distinct genetic groups. In contrast to these patterns, we observed a broad, gradual cline in genome-wide variation and no clear molecular signature of hybridization in the region where morphologically intermediate samples are present. The absence of a zone of genome-wide admixture between the ecotypes is particularly striking, as most zones that show phenotypic evidence of hybridization also show molecular

evidence of admixture even when only a few neutral markers are examined (Pettengil and Moeller 2012; Stankowski 2013; Starr et al. 2013; Alder and Doadrio 2014; Noutsos et al. 2014). Moreover, our  $F_{ST}$  analysis does not support a conclusion of allopatric divergence followed by recent secondary contact. If populations were previously isolated in the east and west, we would expect to see elevated levels of inter-ecotype divergence relative to levels of intra-ecotype divergence. Even though mean  $F_{ST}$  is higher between the ecotypes, the pattern of isolation by distance reveals that the elevated divergence simply reflects the larger range of geographic distances of inter-ecotype comparisons compared to those made within ecotypes. Although the broad cline in genome-wide variation could reflect the decay of a cline following ancient secondary contact, the recent origin of the red ecotype within the broader *M. aurantiacus* species complex (Stankowski and Streisfeld 2015) does not support this scenario.

Thus, while additional analyses that provide further tests of these and alternative demographic histories will be necessary to parse the timing of various historical scenarios that may have generated this hybrid zone, our genomic analyses rule out divergence during a long period of allopatry followed by recent secondary contact. Moreover, the clear pattern of isolation by distance that we observed across the range of both ecotypes contrasts with the sharp genetic discontinuities often observed across other well-studied hybrid zones. This suggests that locally restricted dispersal and drift have shaped neutral patterns of genome-wide divergence, but the floral traits appear to have diverged in the presence of relatively constant gene flow between the ecotypes. This hypothesis predicts that only genomic regions involved in local adaptation will show sharp geographic clines across the landscape, while most markers will show weak spatial differentiation. Ongoing studies that combine QTL mapping with a locus-by-locus analysis of clinal variation from a genome-wide collection of SNPs will allow for a direct test of this hypothesis.

### ACKNOWLEDGMENTS

We would like to thank A. Royer and K. Karroly for useful comments, advice, and discussion. B. Winkleman, C. Benson, M. Chase, and L. Curran assisted with data collection. The project was supported by the University of Oregon and by a National Science Foundation grant: DEB-1258199.

### DATA ARCHIVING

Floral trait data have been archived at Dryad doi:10.5061/dryad.18796. Sequence data have been submitted to the Short Read Archive under bioproject ID: PRJNA299226.

### LITERATURE CITED

Abbott, R., D. Albach, S. Ansell, J. W. Arntzen, S. J. E. Baird, N. Bierne, and D. Zinner. 2013. Hybridization and speciation. *J. Evol. Biol.* 26:229–246.



- Alder, F., and I. Doadrio. 2014. Spatial genetic structure across a hybrid zone between European rabbit subspecies. *PeerJ* 2:e582. doi:10.7717/peerj.582
- Barton, N. H. 1979. The dynamics of hybrid zones. *Heredity* 43:341–359.
- . 2010. What role does natural selection play in speciation? *Philos. Trans. R. Soc. B Biol. Sci.* 365:1825–1840.
- . 2013. Does hybridization influence speciation? *J. Evol. Biol.* 26:267–269.
- Barton, N. H., and S. J. E. Baird. 1995. Analyse: an application for analysing hybrid zones. Freeware, Edinburgh.
- Barton, N. H., and M. A. R. De Cara. 2009. The evolution of strong reproductive isolation. *Evolution* 63:1171–1190.
- Barton, N. H., and K. S. Gale. 1993. Genetic analysis of hybrid zones. Pp. 13–45 in R. G. Harrison, ed. *Hybrid zones and the evolutionary process*. Oxford Univ. Press, Oxford, U.K.
- Barton, N. H., and G. M. Hewitt. 1985. Analysis of hybrid zones. *Annu. Rev. Ecol. Syst.* 16:113–148.
- Bierne, N., J. Welch., E. Loire, F. Bonhomme, and P. David. 2011. The coupling hypothesis: why genome scans may fail to map local adaptation genes. *Mol. Ecol.* 20:2044–2072.
- Beeks, R. M. 1962. Variation and hybridization in southern California populations of *Diplacus* (Scrophulariaceae). *El Aliso* 5:83–122.
- Bridle, J. R., S. J. Baird, and R. K. Butlin. 2001. Spatial structure and habitat variation in a grasshopper hybrid zone. *Evolution* 55:1832–1843.
- Butlin, R. K. 2005. Recombination and speciation. *Mol. Ecol.* 14:2621–2635.
- Butlin, R. K., J. Galindo, and J. W. Grahame. 2008. Sympatric, parapatric or allopatric: the most important way to classify speciation? *Philos. Trans. R. Soc. B Biol. Sci.* 363:2997–3007.
- Catchen, J., P. A. Hohenlohe, S. Bassham, A. Amores, and W. A. Cresko. 2013. Stacks: an analysis tool set for population genomics. *Mol. Ecol.* 22:3124–3140.
- Clarke, B. 1966. The evolution of morph-ratio clines. *Am. Nat.* 100:389–402.
- Conner, J. K., and D. L. Hartl. 2004. *A primer of ecological genetics*. Sinauer Associates Incorporated, Massachusetts.
- Coyne, J. A., and H. A. Orr. 2004. *Speciation*. Sinauer Associates, Sunderland, MA.
- Durrett, R., L. Buttel, and R. Harrison. 2000. Spatial models for hybrid zones. *Heredity* 84:9–19
- Endler J. A. 1977. *Geographic variation, speciation, and clines*. Princeton Univ. Press, Princeton.
- Excoffier, L., G. Laval, and S. Schneider. 2005. Arlequin (version 3.0): an integrated software package for population genetics data analysis. *Evol. Bioinform. Online* 1:47–50.
- Falconer, D. S., and T. F. Mackay. 1996. *Introduction to quantitative genetics*. Longmans Green, Harlow, U.K.
- Feder, J. L., and P. Nosil. 2009. Chromosomal inversions and species differences: when are genes affecting adaptive divergence and reproductive isolation expected to reside within inversions? *Evolution* 63:3061–3075.
- Fenster, C. B., W. S. Armbruster, P. Wilson, M. R. Dudash, and J. D. Thomson. 2004. Pollination syndromes and floral specialization. *Annu. Rev. Ecol. Syst.* 35:375–403.
- Fishman, L., A. Stathos, P. M. Beardsley, C. F. Williams, and J. P. Hill. 2013. Chromosomal arrangements and the genetics of reproductive barriers in *Mimulus* (monkeyflowers). *Evolution* 67:2547–2560.
- Fitzpatrick, B. M., J. A. Fordyce, and S. Gavrillets. 2009. Pattern, process and geographic modes of speciation. *J. Evol. Biol.* 22:2342–2347.
- Gavrilets, S. 2004. *Fitness landscapes and the origin of species, monographs in population biology series* Princeton Univ. Press, Princeton, NJ, 476.
- Grant, V. 1949. Pollination systems as isolating mechanisms in angiosperms. *Evolution* 3:82–97.
- . 1981. *Plant speciation*. Columbia Univ. Press, New York.
- . 1993a. Effect of hybridization and selection on floral isolation. *Proc. Natl. Acad. Sci. USA* 90:990–993.
- . 1993b. Origin of floral isolation between ornithophilous and sphingophilous plant species. *Proc. Natl. Acad. Sci. USA* 90:7729–7733.
- . 1994. Modes and origins of mechanical and ethological isolation in angiosperms. *Proc. Natl. Acad. Sci. USA* 91:3–10.
- Jiggins, C. D., and J. Mallet. 2000. Bimodal hybrid zones and speciation. *Trends Ecol. Evol.* 15:250–255.
- Johnson, L. M., and L. F. Galloway. 2008. From horticultural plantings to wild populations: movement of pollen and genes in *Lobelia cardinalis*. *Plant Ecol.* 197:55–67.
- Kawakami, T., R. K. Butlin, M. Adams, D. Paull, and S. J. Cooper. 2009. Genetic analysis of a chromosomal hybrid zone in the Australian morabine grasshoppers (*Vandiemenella*, viatica species group). *Evolution* 63:139–152.
- Kirkpatrick, M. 2010. How and why chromosome inversions evolve. *PLoS Biol.* 8:e1000501.
- Kirkpatrick, M., and N. Barton. 2006. Chromosome inversions, local adaptation and speciation. *Genetics* 173:419–434.
- Kirkpatrick, M., and V. Ravigné. 2002. Speciation by natural and sexual selection: models and experiments. *Am. Nat.* 159:S22–S35.
- Lenormand, T. 2002. Gene flow and the limits to natural selection. *Trends Ecol. Evol.* 17:183–189.
- Linhart, Y. B., W. H. Busby, J. H. Beach, and P. Feinsinger. 1987. Forager behavior, pollen dispersal, and inbreeding in two species of hummingbird-pollinated plants. *Evolution* 41:679–682.
- Lovell, J. T., T. E. Juenger, S. D. Michaels, J. R. Lasky, A. Platt, J. H. Richards, X. Yu, H. M. Easlon, S. Sen, J. K. McKay. 2013. Pleiotropy of *FRIGIDA* enhances the potential for multivariate adaptation. *Proc. R. Soc. Lond. B Biol. Sci.* 280:20131043.
- Mallet, J., and N. Barton. 1989. Inference from clines stabilized by frequency-dependent selection. *Genetics* 122:967–976.
- Noor, M. A., K. L. Grams, L. A. Bertucci, and J. Reiland. 2001. Chromosomal inversions and the reproductive isolation of species. *Proc. Natl. Acad. Sci. USA* 98:12084–12088.
- Nosil, P. 2008. Speciation with gene flow could be common. *Mol. Ecol. Oxford*, 17:2103–2106.
- . 2012. *Ecological speciation*. Oxford Univ. Press.
- Noutsos, C., J. O. Borevitz, and S. A. Hodges. 2014. Gene flow between nascent species: geographic, genotypic and phenotypic differentiation within and between *Aquilegia formosa* and *A. pubescens*. *Mol. Ecol.* 23:5589–5598
- Orr, H. A. 1991. Is single-gene speciation possible? *Evolution* 45:764–769.
- Pettengill, J. B. and D. A. Moeller. 2012. Phylogeography of speciation: allopatric divergence and secondary contact between outcrossing and selfing *Clarkia*. *Mol. Ecol.* 21:4578–4592.
- Phillips, B. L., S. J. Baird, and C. Moritz. 2004. When vicars meet: a narrow contact zone between morphologically cryptic phylogeographic lineages of the rainforest skink, *Carlia rubrigularis*. *Evolution* 58:1536–1548.
- Pinho, C., and J. Hey. 2010. Divergence with gene flow: models and data. *Annu. Rev. Ecol. Syst.* 41:215–230.
- Poelstra, J. W., N. Vijay, C. M. Bossu, H. Lantz, B. Ryll, I. Müller, and J. B. Wolf. 2014. The genomic landscape underlying phenotypic integrity in the face of gene flow in crows. *Science* 344:1410–1414.
- Pritchard, J. K., M. Stephens, and P. Donnelly. 2000. Inference of population structure using multilocus genotype data. *Genetics* 155:945–959.
- Rauscher, M. D. (2008). Evolutionary transitions in floral color. *Int. J. Plant Sci.* 169:7–21.



- Rieseberg, L. H. 2001. Chromosomal rearrangements and speciation. *Trends Ecol. Evol.* 16:351–358.
- Schemske, D. W., and H. D. Bradshaw. 1999. Pollinator preference and the evolution of floral traits in monkeyflowers (*Mimulus*). *Proc. Natl. Acad. Sci. USA* 96:11910–11915.
- Schlising, R. A., and R. A. Turpin. 1971. Hummingbird dispersal of *Delphinium cardinale* pollen treated with radioactive iodine. *Am. J. Bot.* 58:401–406.
- Schulke, B., and W. M. Waser. 2001. Long-distance pollinator flights and pollen dispersal between populations of *Delphinium nuttallianum*. *Oecologia* 127:239–245.
- Sinervo, B., and E. Svensson. 2002. Correlational selection and the evolution of genomic architecture. *Heredity* 89:329–338.
- Smadja, C. M., and R. K. Butlin. 2011. A framework for comparing processes of speciation in the presence of gene flow. *Mol. Ecol.* 20:5123–5140.
- Smith, J. M. 1966. Sympatric speciation. *Am. Nat.* 100:637–650.
- Sobel, J. M., and M. A. Streisfeld. 2015. Strong premating isolation exclusively drives incipient speciation in *Mimulus aurantiacus*. *Evolution* 69:447–461.
- Sobel, J. M., G. F. Chen, L. R. Watt, and D. W. Schemske. 2010. The biology of speciation. *Evolution* 64:295–315.
- Stankowski, S. 2013. Ecological speciation in an island snail: evidence for the parallel evidence of a novel ecotype and maintenance by ecologically dependent postzygotic isolation. *Mol. Ecol.* 22:2726–2741.
- Stankowski, S., and M. A. Streisfeld. 2015. Introgressive hybridization facilitates adaptive divergence in a recent radiation of monkeyflowers. *Proc. R. Soc. B* 282:20151666.
- Starr, T. N., K. E. Gadek, J. B. Yoder, J. B. Flatz, and C. I. Smith. 2013. Asymmetric hybridization and gene flow between Joshua trees (Agavaceae: *Yucca*) reflect differences in pollinator host specificity. *Mol. Ecol.* 22:437–449.
- Strauss, S. Y., and J. B. Whittall. 2006. Non-pollinator agents of selection on floral traits. In: Harder L, Barrett SCH (eds). *Ecology and Evolution of Flowers*. Oxford University Press: Oxford, Pp. 120–138.
- Streisfeld, M. A., and J. R. Kohn. 2005. Contrasting patterns of floral and molecular variation across a cline in *Mimulus aurantiacus*. *Evolution* 59:2548–2559.
- . 2007. Environment and pollinator-mediated selection on parapatric floral races of *Mimulus aurantiacus*. *J. Evol. Biol.* 20:122–132.
- Streisfeld, M. A., W. N. Young, and J. M. Sobel. 2013. Divergent selection drives genetic differentiation in an R2R3-Myb transcription factor that contributes to incipient speciation in *Mimulus aurantiacus*. *PLoS Genet.* 9:e1003385.
- Tulig, M. 2000. Morphological variation in *Mimulus* section *Diplacus* (*Scrophulariaceae*). Doctoral diss., California State Polytechnic University, Pomona, CA.
- Ueshima, R., and Asami, T. 2003. Evolution: single-gene speciation by left-right reversal. *Nature* 425:679–679.
- Waayers, G. M. 1996. Hybridization, introgression, and selection in *Mimulus aurantiacus* ssp. *australis* and *Mimulus puniceus*. Doctoral diss., San Diego State University, San Diego, CA.
- Webb, C. J., and K. S. Bawa. 1983. Pollen dispersal by hummingbirds and butterflies: a comparative study of two lowland tropical plants. *Evolution* 37:1258–1270.
- Yeaman, S., and M. C. Whitlock. 2011. The genetic architecture of adaptation under migration–selection balance. *Evolution* 65:1897–1911.

Associate Editor: D. Moeller  
Handling Editor: J. Conner

## Supporting Information

Additional Supporting Information may be found in the online version of this article at the publisher's website:

**Table S1.** Geographic coordinates for the 30 localities sampled in Streisfeld et al. (2013) and used to generate the *MaMyb2-M3 red* allele frequency dataset (*P*).

**Table S2.** Maximum-likelihood parameter estimates obtained from one-dimensional cline analysis of the *MaMyb2-M3* marker from the three datasets under the best-fitting cline model (A-step).

**Table S3.** Results of linear regression, testing for a relationship between variation in mean trait value for each population in the common garden experiment and distance along the one-dimensional transect.

**Table S4.** Results of the Evanno et al. (2005) method for inferring the optimum number of population clusters (*K*) following analysis in the program *Structure* (Pritchard et al. 2000).

**Table S5.** Composite likelihood tests comparing the cline width between the genomic PC1 and six floral traits.

**Figure S1.** Mean values from the first two principal components of a PC analysis conducted on the SNP genotype matrix (5382 loci).

**Figure S2.** Variation in population structure across the one-dimensional transect revealed by a Bayesian clustering analysis (5382 loci).

**Figure S3.** Relationship between geographic and genetic distance ( $F_{ST}$ , averaged over loci) among sample locations.