

1 **Integrating top-down and bottom-up approaches to understand the genetic** 2 **architecture of speciation across a monkeyflower hybrid zone**

3
4 Sean Stankowski^{1,2*}, Madeline A. Chase^{1,3}, Hanna McIntosh^{1,4} and Matthew A. Streisfeld^{1*}

5 ¹Institute of Ecology and Evolution, 5289 University of Oregon, Eugene, OR 97403, USA

6 ²Present address: IST Austria, AM1 Campus, Klosterneuburg, Austria.

7 ³Present address: Department of Evolutionary Biology, Evolutionary Biology Centre (EBC),
8 Uppsala University, Uppsala, Sweden

9 ⁴Present address: Department of Entomology, University of Wisconsin – Madison, Madison, WI

10 *Corresponding authors

11 **ORCID ids:**

12 SS: <https://orcid.org/0000-0003-0472-9299>;

13 MAC: <https://orcid.org/0000-0002-7916-3560>;

14 MAS: <https://orcid.org/0000-0002-2660-8642>

15 HRM: <https://orcid.org/0000-0003-2547-5347>

16 **Key words:** Barrier traits, Barrier loci, Cline, Ecotypes, *Mimulus*, Reproductive isolation, QTL

17 **Running title:** Speciation in a monkeyflower hybrid zone

18

19 **Abstract**

20 Understanding the phenotypic and genetic architecture of reproductive isolation is a longstanding
21 goal of speciation research. In many systems, candidate barrier traits and loci have been
22 identified, but causal connections between them are rarely made. In this study, we combine ‘top-
23 down’ and ‘bottom-up’ approaches with demographic modeling toward an integrated
24 understanding of speciation across a monkeyflower hybrid zone. Previous work in this system
25 suggests that pollinator-mediated reproductive isolation is a primary barrier to gene flow
26 between two divergent red- and yellow-flowered ecotypes of *Mimulus aurantiacus*. Several
27 candidate floral traits contributing to pollinator isolation have been identified, including a
28 difference in flower color, which is caused primarily by a single large-effect locus (*MaMyb2*).
29 Other anonymous SNP loci, potentially contributing to pollinator isolation, also have been
30 identified, but their causal relationships remain untested. Here, we performed demographic
31 analyses, which indicate that this hybrid zone formed by secondary contact, but that subsequent
32 gene flow was restricted in a large fraction of the genome by barrier loci. Using a cline-based
33 genome scan (our bottom-up approach), we demonstrate that candidate barrier loci are broadly
34 distributed across the genome, rather than mapping to one or a few ‘islands of speciation.’ A
35 QTL analysis (our top-down approach) revealed most floral traits are highly polygenic, with little
36 evidence that QTL co-localize, indicating that most traits are largely genetically independent.
37 Finally, we find little convincing evidence for the overlap of QTL and candidate barrier loci,
38 suggesting that some loci contribute to other forms of reproductive isolation. Our findings
39 highlight the challenges of understanding the genetic architecture of reproductive isolation and

40 reveal that barriers to gene flow aside from pollinator isolation may play an important role in this
41 system.

42

43 **Introduction**

44 Understanding the phenotypic and genetic architecture of reproductive isolation is a
45 major goal of modern speciation research [1-3]. Early studies took a ‘top-down’ approach by
46 using quantitative trait locus (QTL) mapping and other association methods to detect genomic
47 regions controlling barrier phenotypes or genetic incompatibilities [4-6]. More recently, ‘bottom-
48 up’ approaches, such as genome scans of genomic differentiation (e.g., F_{ST}) or admixture (e.g.,
49 f_d), have identified candidate barrier loci in numerous systems, including those where isolation is
50 thought to result from ecologically-based divergent selection or intrinsic incompatibilities [7-10].

51 Although both approaches have clear strengths, they also present significant challenges
52 [11]. Top-down methods require that traits involved in reproductive isolation have already been
53 identified, so our understanding of the genetic architecture of speciation can only ever be as
54 complete as our knowledge of the traits controlling reproductive isolation in the system. In
55 contrast, bottom-up approaches can provide a comprehensive view of the genomic landscape of
56 speciation without complete knowledge of the isolating traits (but see [3, 12]). However, even
57 though candidate barrier loci can be identified, their causal relationship with previously
58 identified barrier traits usually remains unclear. This is because speciation usually involves many
59 different isolating barriers (e.g., pre- and post-zygotic, extrinsic and intrinsic) [13, 14] that can
60 become coupled together through different aspects of the speciation process [15-17]. Although
61 the coupling of different barriers eases speciation by generating a stronger overall barrier [16,
62 17], the resulting linkage disequilibrium (LD) among barrier loci makes it difficult to understand
63 their individual contributions to barrier traits. For example, a barrier locus identified in a genome
64 scan might underlie an obvious phenotypic difference, or it may underlie a completely different
65 barrier that is less conspicuous or that has yet to be discovered.

66 Therefore, instead of relying on one approach, many researchers have advocated for the
67 integration of top-down and bottom-up methods [3, 11, 18]. However, this kind of integration is
68 missing from most studies of speciation, meaning that any links between candidate barrier traits
69 and barrier loci remain tentative. To date, some of the best efforts to integrate top-down and
70 bottom-up analyses have made use of natural hybrid zones between divergent populations [19].
71 Hybrid zones have been described as natural laboratories, because they allow us to understand
72 how reproductive isolation and barriers to gene flow play out in the real world [20]. In addition,
73 their presence provides compelling evidence for ongoing gene flow between the taxa being
74 studied, the relative duration of which can now be estimated using demographic inference
75 methods. Moreover, cline theory provides a rich, spatially explicit framework for studying
76 selection and gene flow across porous species boundaries [21-23]. Specifically, the shape and
77 position of geographic clines are impacted by the relative effects of selection and gene flow
78 across a hybrid zone [22]. Cline analysis has clear advantages over population genetic summary
79 statistics used in most selection scans (e.g., F_{ST}), and it can be applied to phenotypic traits [24].

80 However, it is only beginning to be applied to genome-scale datasets [24-28]. In this study, we
81 combine top-down and bottom-up analyses to investigate the phenotypic and genetic architecture
82 of pollinator isolation between hybridizing ecotypes of the bush monkeyflower (*Mimulus*
83 *aurantiacus*).

84 In San Diego County, California, there is a sharp geographic transition between red- and
85 yellow-flowered ecotypes of *M. aurantiacus* ssp. *puniceus* [29]. Despite being very closely
86 related ($d_a = 0.005$; [30], the ecotypes show extensive divergence across a suite of floral traits,
87 including color, size, shape, and placement of reproductive parts (Fig. 1; [29, 31, 32]). Previous
88 work suggests an important role of pollinators in driving floral trait divergence and reproductive
89 isolation in this system [29, 33-36]. Field experiments have shown that hummingbirds and
90 hawkmoths show strong preferences and constancy for the flowers of the red and yellow
91 ecotypes, respectively [34, 36]. In addition to providing a source of divergent selection,
92 pollinator behavior generates substantial pre-mating isolation, potentially reducing gene flow
93 between the ecotypes by 78% in sympatry [36]. Post-mating isolation is weak between the
94 ecotypes, suggesting that pollinator isolation is the primary barrier to gene flow in this system
95 [36].

96 Although the strength of pollinator-mediated reproductive isolation is strong, it is
97 incomplete, meaning that there is potential for gene flow between the ecotypes in locations
98 where their distributions overlap. This has led to the formation of a narrow hybrid zone,
99 characterized by extensive phenotypic variation and geographic clines in several floral traits. For
100 example, there is a steep cline in flower color that is centered on the hybrid zone and matches a
101 similarly steep cline in the gene *MaMyb2*, which controls much of the variation in pigmentation
102 [35]. Other floral traits and anonymous single-nucleotide polymorphisms (SNPs) also show
103 clinal variation, implying that multiple traits contribute to reproductive isolation [24, 33].
104 However, evidence for an association between these phenotypic traits and genotypic signatures
105 of selection is currently lacking.

106 In this study, we use demographic modeling, a cline-based genome scan, and QTL
107 mapping to investigate the history of divergence and the connection between phenotypic and
108 genomic signatures of selection in this system. One possible outcome is that QTL for the
109 divergent phenotypes will overlap with regions of the genome under selection, as predicted if
110 pollinator-mediated selection is the main barrier to gene flow between the ecotypes. These
111 regions may be abundant and widespread across the genome, reflecting polygenic divergence, or
112 they may consist of one or a few genomic regions enriched for loci that underlie multiple floral
113 traits. Under an alternate scenario, we may find that floral QTL only partially overlap with
114 genomic signatures of selection, which might reflect the spatial coupling of multiple different
115 kinds of barriers. Our findings highlight the challenges of understanding the genetic architecture
116 of reproductive isolation, and suggest that barriers to gene flow aside from pollinator isolation
117 may also contribute to speciation in this system.

118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159

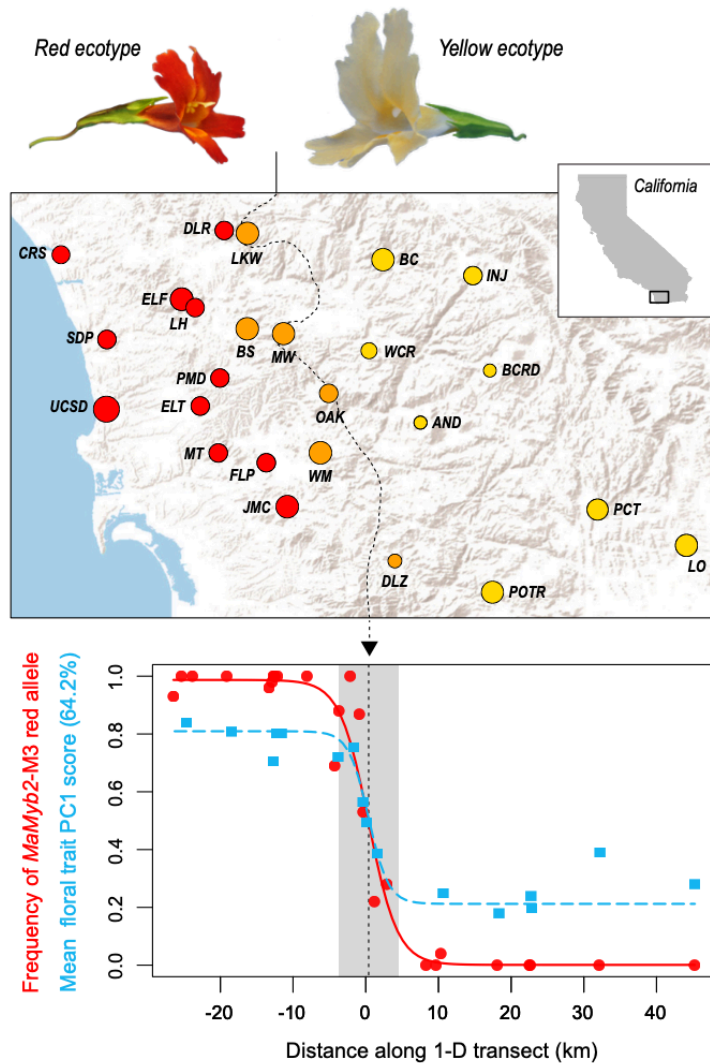


Figure 1. Clinal variation across a bush monkeyflower hybrid zone. (top) Typical flower phenotypes of the red and yellow ecotypes, and a map of the 25 sampling locations in San Diego County. The size of the circles shows variation in the sample sizes, which range from 4 to 18 individuals, totaling 292 individuals. The dashed line indicates the center of the hybrid zone, previously inferred from spatial variation in the frequency of alternative alleles at the *MaMyb2* locus. **(bottom)** Clines in allele frequency at the *MaMyb2* locus (red circles) and the mean floral trait PC1 score (blue squares) across the one-dimensional transect. The solid and dashed lines are the ML sigmoid cline models for *MaMyb2* allele frequency and trait PC1 score, respectively. The gray shaded rectangle represents the width of the hybrid zone.

Methods

RAD sequencing, read filtering, and SNP calling

We identified SNPs using previously sequenced restriction-site associated DNA sequences (RADseq) from 292 individuals sampled from 25 locations across the hybrid zone

160 (mean individuals per site = 12; range 4–18) [24]. These included 11 sites in the range of the red
161 ecotype, 8 sites in the range of the yellow ecotype, and 6 sites in the hybrid zone (Table S1).

162 We processed the raw sequences, identified SNPs, and called genotypes using *Stacks* v.
163 1.41 [37]. Reads were filtered based on quality, and errors in the barcode sequence or RAD site
164 were corrected using the *process_radtags* script in *Stacks*. Individual reads were aligned to the
165 *M. aurantiacus* genome [30] using *Bowtie 2* [38], with the *very_sensitive* settings. We identified
166 SNPs using the *ref_map.pl* function of *Stacks*, with two identical reads required to create a stack
167 and two mismatches allowed when processing the catalog. SNP identification and genotype calls
168 were conducted using the maximum-likelihood model ($\alpha = 0.01$) [39]. To include a SNP in
169 the final dataset, we required it to be present in at least 70% of all individuals; this resulted in a
170 final dataset of 219,152 SNPs.

171

172 *Demographic inference*

173 To gain a deeper understanding of the history of gene flow and selection in this system,
174 we performed demographic inference in $\partial\text{a}\partial\text{i}$ [40]. We calculated the unfolded joint site
175 frequency spectrum (JSFS) based on 19,902 SNPs, using subspecies *grandiflorus* as an outgroup
176 to polarize alleles as ancestral or derived [41]. SNPs were included if they were genotyped in
177 *grandiflorus* and in at least 70% of the red and yellow individuals. We included 124 individuals
178 from 10 sites of the red-flowered ecotype and 65 individuals from 7 sites of the yellow-flowered
179 ecotype, excluding sample sites that showed evidence of recent admixture (all hybrid sample
180 sites and populations DLR and BC). The JSFS was projected to a sample size of 85 to maximize
181 the number of segregating sites.

182 We fit nine two-population demographic models to the JSFS (Fig. S1): (i) strict isolation (SI),
183 (ii) ancient migration (AM), (iii) isolation with migration (IM), (iv) secondary contact (SC), and
184 (v) periods of secondary contact (PSC). The remaining four models—(vi) *AM2m*, (vii) *IM2m*,
185 (viii) *SC2m*, and (ix) *PSC2m*—are the same as models ii-v, except that migration rates are
186 inferred for two groups of loci to simulate the effect of a porous barrier to gene flow. For each
187 model, we performed 20 independent runs using randomly generated starting parameters, and we
188 reported the results for the run with the lowest log-likelihood. The goodness of fit of the models
189 was determined using the AIC. Parameter estimates were converted into biologically meaningful
190 values as described in [42], assuming a mutation rate of 7×10^{-9} [43].

191

192 *Admixture analysis*

193 We used the model-based clustering program *Admixture* [44] to characterize patterns of
194 genetic structure across the hybrid zone. We assigned the 292 individuals sampled from across
195 San Diego County into two clusters ($K=2$) based on the full dataset of 219,152 SNPs that met our
196 filtering requirements (note that $K=2$ was determined as the optimum number of clusters in [24]).
197 In addition to using the full dataset, we also pruned SNPs using the *--indep-pairwise* function in
198 *Plink* [45] to reduce linkage disequilibrium (LD) between neighboring SNPs (r^2 threshold of 0.1,
199 window size = 50 SNPs, step size = 10 SNPs).

200 We also ran *Admixture* separately for each chromosome, and for 2,173 non-overlapping
201 windows, each containing 100 SNPs (mean window size of 89.1 kb with 8 - 38 RAD tags per
202 window Fig. S2) The window-based analysis was automated using custom python scripts to
203 produce plink.map and .ped files for each consecutive window, which were then passed to
204 *Admixture*.

205 206 *Cline fitting*

207 To quantify the geographic variation in ancestry (Q), we fit a sigmoid cline model to the mean
208 ancestry scores from each site along a one dimensional transect, which was described in [24]
209 (Fig. 1). Clines were fitted using *HZAR* [46] using the quantitative trait model, with the variance
210 in the trait modeled separately on the left side, center, and right side of the cline. We estimated
211 the following parameters: the cline center (c), defined as the inflection point of the sigmoid
212 function; Q_{left} and Q_{right} , the mean ancestry scores on the left and right sides of the cline,
213 respectively; and the cline width (w), defined as the ratio between the total change in ancestry
214 across the cline (ΔQ) and the slope at the cline center (note that $\Delta Q = Q_{\text{left}} - Q_{\text{right}}$ because we
215 ensured that the mean ancestry score was higher on the left side before fitting). We conducted 3
216 independent fits with random starting values and retained the one with the highest log-likelihood.
217 All of the best fits were visually inspected to ensure a sensible fit.

218 219 *Summarizing clinal variation in windows*

220 After cline fitting, we calculated an *ad-hoc* statistic to identify genomic windows that had
221 clines with a similar shape and position to the genome-wide cline. We refer to this statistic as the
222 cline similarity score (cs score). Unlike individual parameters (e.g. the width or center), which
223 describe a single feature of a cline, the cs score describes the shape and relative position of a
224 cline with a single number. We calculate the cline similarity score as: $cs = [\Delta Q / (w + l)] e^{-(|c/l|)^2}$.
225 Briefly (but see Supplement S1 for more details), the total change in ancestry, ΔQ , is divided by
226 the sum of w and a scaling variable ($w + l$) to give an estimate of cline shape. The scaling
227 variable controls the spread of shape scores across the joint distribution of ΔQ and w . In our case,
228 $l =$ half the length of the transect ($0.5t$), which results in high shape scores when clines have high
229 ΔQ and low w , but low shape scores when clines have low ΔQ and high w . The shape score is
230 then scaled according to the position of the cline center, c , relative to a position of interest. This
231 could be a feature of the environment or a cline in a focal marker or trait. In our case, the
232 position of interest is the center of the genome-wide ancestry cline. If the cline center coincides
233 exactly with this point, then the shape score is equal to the cs score. However, the farther that the
234 cline center is shifted away from the point of interest, the more the shape score is downgraded,
235 resulting in a lower estimate of cs . Therefore, to have a high value of cs , a cline from a genomic
236 window must have its shape and position closely match the cline in genome-wide ancestry (as in
237 Fig. 3A). Finally, we scaled the cs score relative to the genome-wide ancestry cline, where 1 is
238 the cs score calculated for the genome-wide ancestry cline, and 0 is the minimum value of cs
239 observed for a window.

240 *Estimates of genetic differentiation in genomic windows*

241 To characterize levels of genetic differentiation in a more traditional way, we calculated
242 the population genetic statistic F_{CT} between the ecotypes in each 100 SNP window using the
243 program *Arlequin* [47]. This was done in an analysis of molecular variance framework that
244 partitioned genetic variation between the ecotypes, among populations within ecotypes, and
245 within populations. Populations were classified as coming either from the red or yellow ecotypes
246 based on the *Admixture* results. Samples from hybrid populations were excluded from this
247 analysis.

248

249 *QTL analysis*

250 We used QTL analysis to identify genomic regions underlying divergent floral traits. We
251 generated an outcrossed F_2 population that contained 292 offspring produced by crossing two F_1
252 individuals; each of these F_1 parents was produced by crossing different greenhouse-raised red
253 and yellow ecotype plants (from populations UCSD and LO; Table S1). To allow direct
254 phenotypic comparison among plants grown in a common environment, we raised 25 red ecotype
255 individuals (location UCSD), 31 yellow ecotype individuals (location LO) and 20 F_1 individuals
256 (LO x UCSD) alongside the F_2 s. For each plant, we measured 13 floral traits (Fig. S3). Plants
257 were raised as described in [33].

258 QTL mapping was conducted using *R/qtl* [48] and a previously published genetic map
259 [30] generated from the same mapping population using *Lep-MAP2* [49]. We used phase
260 information from *Lep-MAP2* to infer the grandparental origin of alleles in the F_2 s at 7574
261 mapped markers, which allowed us to recode them as coming either from a red or yellow
262 grandparent. This set of markers was then reduced down to 2631—one per map position—by
263 retaining the marker at each map position with the least missing data. Missing data for these
264 markers was inferred by imputation using phase information from the mapping software. For
265 each trait, we then used automated stepwise scanning for additive QTL and pairwise interactions.
266 QTL identified using this procedure were then incorporated into a multi-QTL model to refine
267 their positions, calculate 95% Bayes credible intervals, and estimate the percent phenotypic
268 variation explained (i.e., the effect size) of each QTL.

269

270 *Test for an excess of QTL overlap*

271 To test for co-localization among QTL, we used a permutation test to determine if there
272 was significantly more overlap among QTL than expected by chance. We first estimated the
273 observed number of overlaps based on the Bayes credible intervals among the 26 identified QTL
274 using the *findOverlaps* function of the *GenomicRanges* package [50] in R and determined the
275 average number of overlaps per QTL (n overlaps/ n QTL). To determine if this statistic was
276 significantly larger than expected by chance, we randomly generated new QTL positions while
277 maintaining the observed number and size of observed QTL. We made the probability of QTL
278 ‘landing’ on a given chromosome (Chr) a function of that chromosome’s length (L_i) relative to
279 the total genome length (L_T), $P(\text{Chr}_i) = L_i / L_T$, so that larger chromosomes were more likely to

280 have QTL i assigned to them. We calculated the mean number of overlaps per QTL for 9,999
281 random datasets and estimated a p -value for the observed value as the number of permuted
282 datasets where n overlaps/ n QTL was equal to or greater than the observed estimate + 1/number
283 of permutations +1.

284

285 *Test for overlap of QTL and outlier windows*

286 We also used a permutation test to determine if QTL regions were enriched for outliers
287 identified in our cline and F_{CT} based genome scans. We first counted the observed number of
288 outlier windows within the empirical QTL intervals. This was performed for both cs and F_{CT}
289 outliers, defined using two different cutoffs (top 1% and 5% of the empirical distributions). To
290 determine if these counts were significantly different from chance, we produced 9,999 datasets
291 where the genomic position of outlier windows was randomized, and we counted the number of
292 outliers falling inside the empirical QTL intervals. A p -value for the estimate was calculated as
293 described above.

294

295 **Results**

296 *Evidence for a history of heterogeneous gene flow*

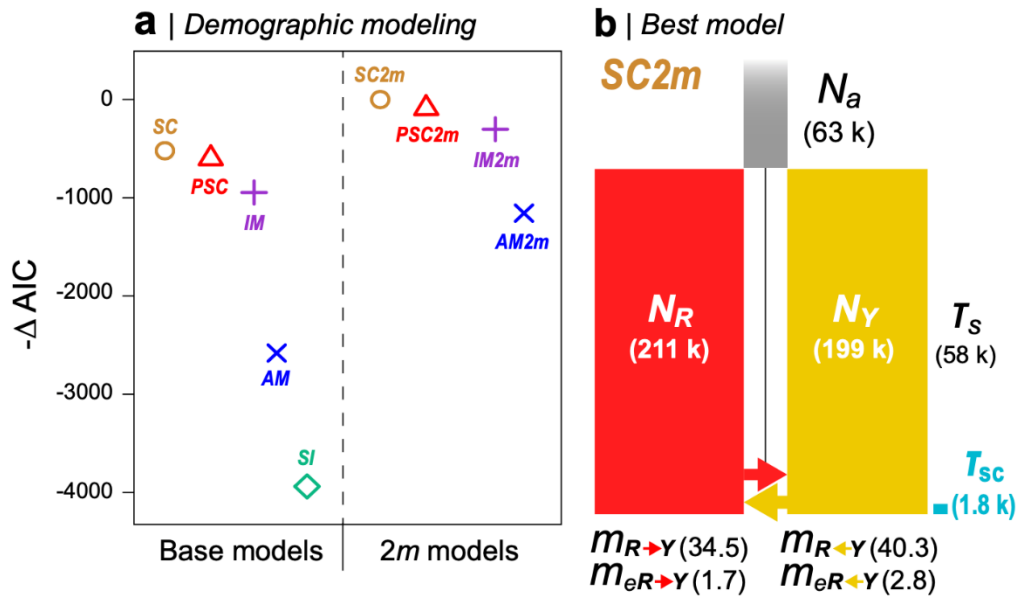
297 The results of our demographic modeling support a history of gene flow across the range
298 of the ecotypes. First, demographic models that included contemporary gene flow were far better
299 at recreating the observed JSFS than a model of divergence without gene flow (i.e., strict
300 isolation; $-\Delta AIC = 3938$), or the best model of ancient migration, which included historical but
301 not contemporary gene flow ($-\Delta AIC = 1157$) (Fig. 2; Fig. S4). Second, models that included
302 heterogeneous migration across the genome ($2m$) were always strongly favored over the
303 equivalent models, where gene flow was modeled with a single rate (Fig. 2). Third, the SC and
304 PSC models, which included periods of allopatry and secondary contact, were strongly favored
305 over the IM model, where divergence occurred without a period of geographic isolation. The
306 best-fitting model was the SC $2m$ model, indicating that divergence of the red and yellow
307 ecotypes included a period of allopatry followed by gene flow upon secondary contact (Fig. 2).
308 Assuming a mutation rate of 7×10^{-9} [43], the ML parameters indicate that the ecotypes have been
309 exchanging an average of 37 migrants per generation ($m_{YR} = 34.5$ per gen.; $m_{RY} = 40.3$ per gen.)
310 for the last 1,800 generations, which equates to roughly 3,600 years, assuming a two-year
311 generation time for these perennial plants. Despite evidence that gene flow between the ecotypes
312 has been extensive, the ML model suggests that 37.4% of loci have experienced a substantial
313 reduction in effective migration (15-20 fold; $m_{eYR} = 1.7$ per gen.; $m_{eRY} = 2.7$ per gen.) due to the
314 effects of selection against gene flow (Fig. 2).

315

316

317

318



335 **Figure 2. Demographic modeling reveals a history of gene flow following isolation.** (a) $-\Delta AIC$
 336 scores for the 9 demographic models fitted to the observed JSFS using $\partial a \partial i$. The base models (left of the
 337 dashed line) include a single migration parameter (m) for all loci, whereas the 2m models include
 338 separate migration parameters for neutral loci (m) and loci affected by a barrier to gene flow (m_e). The
 339 best model (SC2m) has a $-\Delta AIC$ of 0, with more negative values indicating that models were a poorer fit.
 340 (b) A graphical depiction of the SC2m model. The width of the columns is proportional to the population
 341 size estimates for the ancestral (N_a), red (N_R), and yellow (N_Y) populations. The height of the red and
 342 yellow bars is proportional to the total time in generations (T_S) that has passed since the split. The blue
 343 bar shows the period during which secondary gene flow (T_{Sc}) occurred. The difference in arrow size is
 344 proportional to the difference in the bi-directional migration rate, m . The rates of effective migration (m_e)
 345 are too small to show graphically.

346
 347

348 *Sharp clines in genome- and chromosome-wide ancestry*

349 The presence of sharp clines in multiple floral traits suggests that some fraction of the
 350 genome is impacted by selection against gene flow [33]. The results from *Admixture* support
 351 these findings, revealing genome-wide patterns of ancestry that closely match the ecotypic
 352 designations assigned based on floral phenotypes (Fig. S5). Specifically, red- and yellow-
 353 flowered individuals sampled from either side of the hybrid zone were strongly assigned to
 354 alternate clusters, while individuals from hybrid populations tended to show some assignment to
 355 both clusters, indicating their genomes are a mix of red and yellow ancestry. The results are
 356 nearly identical between independent runs of *Admixture* that include the full dataset or subsets of
 357 the data pruned to minimize LD between neighboring SNPs ($r^2 > 0.999$).

358 To compare these changes in ancestry to the observed geographic variation in floral traits,
 359 we used cline analysis to fit a sigmoid cline to the mean ancestry scores from each site. The best-
 360 fitting cline model provides an excellent summary of the change in ancestry across the transect
 361 (Fig. 3a) and has an extremely similar shape to cline models from the divergent floral traits and

362 molecular markers (Fig. 1). In addition, consistent with the increased variance we observed in
363 multiple phenotypic traits in hybrid populations [33], the standard deviation of ancestry scores is
364 higher in sample sites close to the cline center, thus providing genomic evidence for
365 hybridization (Fig. 3a).
366

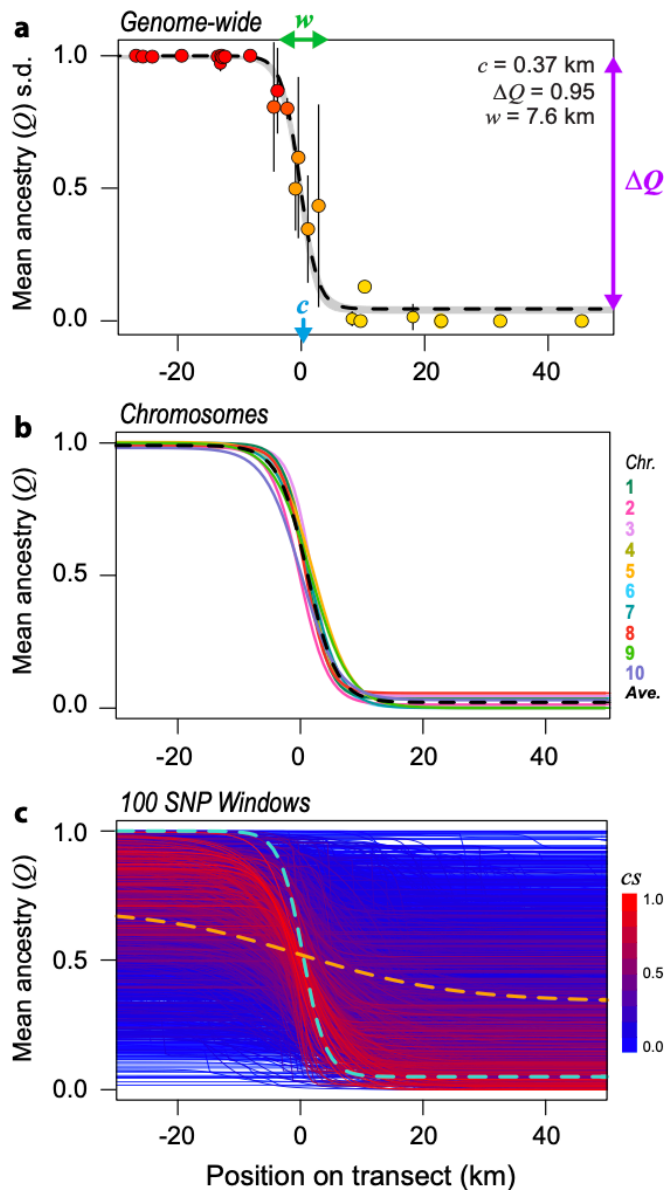


Figure 3. Clines in ancestry scores at different scales of genomic organization. (a) Genome-wide cline, inferred from the mean ancestry (Q) scores in each population along the 1-D transect. Position 0 on the horizontal axis corresponds to the cline center estimated from *MaMyb2* allele frequencies (see Fig. 1). The vertical bars show the standard deviation in ancestry scores for each population. The dashed line is the ML cline model, and the gray band is the two-unit support envelope. Three parameters of interest, including the cline center (c), width (w), and total change in ancestry across the cline (ΔQ), are indicated on the plot. **(b)** Ancestry clines estimated separately for each chromosome. Only the ML curves are shown for clarity (but see Fig. S6). The dashed line is the mean cline, estimated by taking the average of the ML parameters for all chromosomes. **(c)** Ancestry clines estimated for 2,173, 100-SNP windows. The dashed cyan line shows the cline shape for the genome-wide cline (as shown in panel a), while the dashed orange line is the mean cline shape, estimated by taking the average of the ML parameters obtained for all windows. Each solid line is the ML sigmoid curve for one of the genomic windows. The curves are colored according to the value of the cline similarity score (cs), which indicates how similar the shape and position of each cline is to the genome-wide cline. Redder clines are more similar to the genome-wide cline and bluer clines are less similar (See main text for more details).

401

402 The *Admixture* scores provide additional genetic evidence for restricted gene flow across
403 the hybrid zone, but they give us no indication as to the number of loci involved or their genomic
404 distribution. For example, the differences in ancestry could be driven by a small number of loci
405 that reside on a single chromosome, or they could reflect more widespread genomic divergence,
406 involving loci scattered across multiple chromosomes. By repeating the cline analysis of ancestry

407 scores separately for each chromosome, we find highly consistent clines in ancestry for all 10
408 chromosomes (Fig. 3b, Fig. S6).

409

410 *Heterogeneous clinal variation across the genome*

411 To understand how cline shape varies at a finer genomic scale, we fit clines to 2173 non-
412 overlapping 100-SNP windows. This analysis revealed broad variation in geographic patterns of
413 ancestry (Fig. 3c). Unlike each chromosome, the majority of windows show little or no spatial
414 change in ancestry between the red and yellow ecotypes, translating into very low *cs* scores
415 (mean *cs* = 0.15; s.d. = 0.15, Fig. 4).

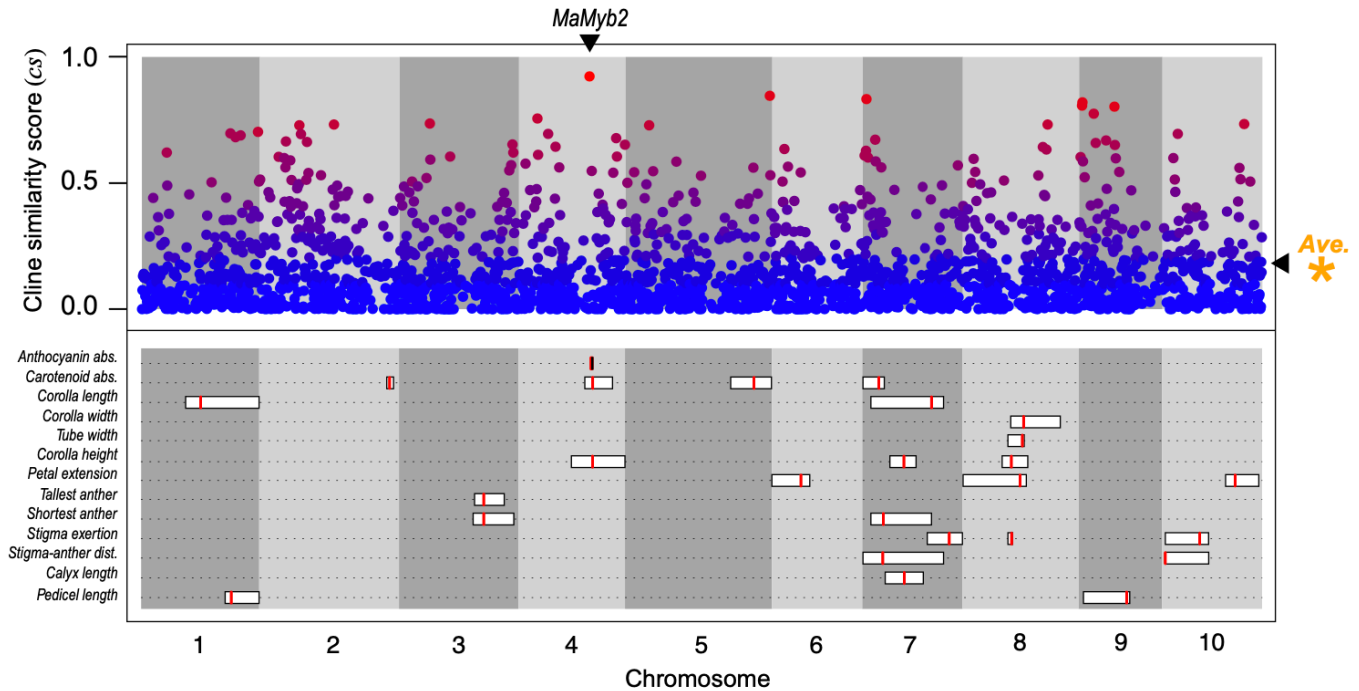
416 However, some genomic regions show clines in ancestry that strongly resemble the
417 genome wide cline, suggesting that they contain barrier loci. This includes the window with the
418 highest *cs* score (*cs* = 0.91), which contains the known barrier locus, *MaMyb2*. The shape and
419 position of this window-based ancestry cline ($\Delta Q = 0.95$, $w = 10.4$ km, $c = -0.3$ km) is highly
420 similar to the genome-wide cline in ancestry ($\Delta Q = 0.95$, $w = 7.6$ km, $c = 0.37$ km) and to the
421 cline in allele frequency for a SNP in *MaMyb2* (*MaMyb2*-M3 marker: $\Delta P = 0.99$, $w = 8.1$ km, c
422 = -0.07 km [24]. However, rather than a clear set of outliers, we observe a continuous
423 distribution of *cs* scores. Therefore, we use two arbitrary cutoffs (top 1% and 5% of the
424 distribution of *cs* scores) to define a set of candidate windows potentially containing barrier loci.

425 Regardless of which cutoff we use, these candidate barrier regions are broadly distributed
426 across the genome. For the 5% cutoff, they occurred on all 10 chromosomes (5 - 20 windows per
427 chromosome; for the 1% cutoff, they occur on 9 of the 10 chromosomes with 1 - 4 windows per
428 chromosome). There were only 12 cases where candidate windows were directly adjacent,
429 indicating that they also were broadly distributed within each chromosome. We also find that
430 genetic differentiation is higher for candidate regions than for the genomic background (1%
431 mean $F_{CT} = 0.31$, 5% mean $F_{CT} = 0.23$, overall mean $F_{CT} = 0.07$). However, F_{CT} explains only
432 38% of the variation in *cs* scores (Fig. S7).

433

434 *Most candidate barrier traits are polygenic*

435 We used quantitative trait locus (QTL) mapping to identify regions of the genome
436 associated with candidate barrier traits. The 13 floral traits showed significant differences
437 between pure red and yellow ecotype plants when grown in a common environment, with mean
438 trait values differing by 0.9 to 7.1 standard deviations (Fig. S8). A total of 26 QTL were
439 identified. For nine traits, we identified more than one QTL (range 2 - 4), and QTL were located
440 on all 10 linkage groups, with LG 7 containing QTL for seven different traits (Fig. 4, Figs. S9
441 and S10). On average, each QTL explained 9.9% of the variation in the F₂ population (range
442 1.82% - 62.6%) (Fig. S11), with an average total variation explained for each trait of 19.8%. The
443 exception was a large-effect QTL for anthocyanin content on LG 4 that explained 62.6% of the
444 variation and mapped to a region near the previously identified causal locus *MaMyb2* [35]. Thus,
445 despite clear heritable differences in these traits, QTL analysis was able to explain only a modest
446 amount of the segregating variation, indicating that most traits have a polygenic architecture.



447
448 **Figure 4. Cline-based genome scan and locations of QTL for floral traits.** (top) The scaled cline
449 similarity (cs score) score in each 100 SNP window plotted against the physical position of the window in
450 the bush monkeyflower genome. The points are colored as in Fig. 3c, with redder points containing
451 windows with cs scores that are more similar to the genome-wide pattern and bluer points are less similar
452 (See main text for more details). The orange asterisk denotes the average barrier score among all
453 windows. The position of the *MaMyb2* gene that controls differences in flower color is shown. (bottom)
454 The positions of the QTL for the 13 measured floral traits plotted along the physical position of the
455 genome. The red vertical line corresponds to the best estimate of the QTL peak, and the width of the
456 rectangles denotes the 95% Bayes credible intervals of the estimated QTL position.

457
458 The presence of multiple QTL occurring on the same chromosome indicates that some
459 regions may contribute to multiple traits, which would help maintain trait associations in hybrid
460 offspring [17]. Overall, we find that QTL do tend to co-localize more often than would be
461 expected by chance (mean observed overlap of 3.23 QTL; mean permuted overlap = 2.51 QTL; p
462 = 0.042, Fig. S12). However, the effects of this co-localization are seen most strongly only for
463 size-related traits (e.g. corolla length and height of the tallest anther), which remain highly
464 correlated in the F_2 generation ($r = 0.88$) (Fig. S13). By contrast, the average correlation
465 coefficient among all other pairs of traits was much lower (mean absolute value of $r = 0.26$). For
466 example, the three overlapping QTL on LG 4 that control anthocyanin and carotenoid
467 pigmentation, as well as corolla height, span a total physical distance of only 76 kb. However,
468 these traits show weak correlations in the F_2 population (anthocyanin vs carotenoid: $r = 0.18$;
469 anthocyanin vs corolla height: $r = -0.20$; carotenoid vs corolla height: $r = 0.19$). This shows that
470 the QTL overlap would have little effect on maintaining the phenotypic correlations where
471 hybridization occurs.

472 *Low concordance between QTL and outlier regions*

473 Finally, we tested for overlap between the floral trait QTL and the candidate barrier regions from
474 the cline-based and F_{CT} genome scans. Using a permutation test, we tested whether genomic
475 windows with higher cs scores tended to overlap with QTL more often than expected by chance.
476 Regardless of which cutoff we used (e.g. top 1% or top 5% of cs scores), we found that floral
477 trait QTL were not significantly enriched for candidate barrier regions ($p > 0.3$, Fig. S14). This
478 suggests a complex connection between the genetic and phenotypic architecture of reproductive
479 isolation. The results were the same when we defined candidate barrier regions based on F_{CT}
480 (e.g. top 1% or top 5% of the F_{CT} distribution; Fig. S14).

481 Given that wide QTL intervals reduce the power of the enrichment test, we also asked
482 how often the estimate of the QTL peak fell within a candidate barrier window. However, even
483 when using the 5% cutoff, we found that none of the QTL peaks occurred within candidate
484 barrier regions. This included the QTL for floral anthocyanin, where the QTL peak occurs 589
485 kb from the window containing the causal locus, *MaMyb2*.

486

487 **Discussion**

488 In this study, we used a combination of QTL mapping and population genomic analyses
489 to obtain a deeper understanding of the phenotypic and genetic architecture of reproductive
490 isolation in a hybrid zone. Past studies in this system have separately identified candidate floral
491 traits contributing to pollinator-mediated reproductive isolation [33, 34], and anonymous,
492 candidate barrier loci [24]. Here, we use top-down and bottom-up approaches in an effort to
493 connect phenotypic and genetic candidates. These results are discussed in the context of new
494 insights about the history of divergence revealed by demographic analysis and are aided by a
495 known, large-effect barrier locus (*MaMyb2*) with a clear phenotypic effect.

496

497 *The history of divergence: new insights from demographic analysis*

498 A firm understanding of the historical demography of speciation is essential when
499 interpreting divergence across hybrid zones [51, 52]. In zones that are at demographic
500 equilibrium, it is possible to interpret clines in terms of migration, selection, and drift and
501 sampling effects [22]. However, if hybrid zones formed recently, clines in neutral loci or traits
502 will be steep initially, but they will decay over time [22]. Previous work in this system revealed
503 patterns of isolation-by-distance across (and orthogonal) to the hybrid zone that were consistent
504 with a long-term ‘stepping-stone’ pattern of gene flow across the entire range of the study area
505 [33]. In fact, there was no evidence for substantial genome-wide differentiation between the
506 ecotypes after correcting for the effect of geography. Based on this result, Stankowski et al. [33]
507 concluded that the hybrid zone formed due to one of three possible scenarios: (i) a primary origin
508 with continuous gene flow during divergence, (ii) a secondary origin, where divergence occurred
509 in allopatry, followed by extensive gene flow after contact resumed, or (iii) a secondary origin
510 where the period of allopatry was short.

511 Our demographic analyses provide new evidence that this hybrid zone formed from
512 secondary contact after a long period of isolation. Indeed, both models that included periods of
513 geographic isolation (SC and PSC) were a better fit to the data than models of continuous gene
514 flow (IM). The parameter estimates for the preferred model (SC2 m) indicate a relatively long
515 period of isolation, followed by a period of secondary contact that began roughly 1,800
516 generations ago. It is important to note that we fit relatively simple models to the data that
517 excluded changes in population size in the ancestral and daughter populations and variation in N_e
518 along the genome. Recent work has shown that failure to model key parameters can result in
519 incorrect inference under some circumstances [53]. Although more sophisticated modelling may
520 arrive at different conclusions in the future, these results point to a secondary origin of this
521 hybrid zone.

522 In terms of the main goal of our paper, the key result of the demographic analysis was
523 that all models with two rates of migration (2 m models) fit the data better than those where
524 migration was modelled at a single rate, indicating a heterogeneous pattern of gene flow across
525 the genome. Moreover, the estimated parameters for the preferred model suggest that roughly
526 one third of loci (37%) have experienced migration at substantially reduced rate compared with
527 non-barrier loci. This result supports previous conclusions that candidate barrier traits and loci
528 are indeed impacted by natural selection [24], further motivating the need to connect the
529 phenotypic and genetic architecture of RI in more detail.

530

531 *From the 'bottom up': insights from the cline-based genome scan*

532 Starting with the bottom-up approach, we used a cline-based genome scan to identify
533 candidate loci underlying barriers to gene flow. Unlike traditional summary statistics calculated
534 between pre-defined groups (e.g., F_{ST}), geographic cline analysis is a more natural framework for
535 studying genetic divergence across hybrid zones, as it provides clearer insights into the strength
536 and nature of selection [24, 26]. Rather than fitting clines to allele frequencies for individuals
537 SNPs, we fit clines to model-based ancestry scores, treating them as a quantitative trait [21].

538 By conducting this analysis at different scales of genomic organization, we are able to
539 conclude that candidate barrier regions are widespread throughout the genome. At the genome-
540 wide scale, the cline in ancestry is centered on the hybrid zone and has a very similar shape to
541 the clines in floral traits. At the chromosome scale, all 10 chromosomes show clear sigmoid
542 clines in ancestry, with their shapes and positions being highly similar to the genome-wide cline.
543 The same conclusion can be drawn from the window-based analysis, as candidate barrier regions
544 are present on all 10 of the chromosomes. The window with the highest cs score is located on
545 chromosome 4 and contains the gene *MaMyb2*. Alternate alleles at this locus control the
546 difference in flower color, and other population genetic analyses indicate that it has been subject
547 to strong divergent selection in this system [35, 54]. Prior knowledge of this barrier locus
548 provides confidence that other windows with high cs scores also likely harbor barrier loci with
549 similarly large phenotypic effects.

550 However, rather than identifying a clear set of *cs* outliers, we observed a continuous
551 distribution of *cs* scores, indicating that clines show varying degrees of resemblance to the
552 genome-wide cline. Although it is tempting to interpret variation in the cline similarity score
553 exclusively in terms of the sieving effect of a porous species boundary (i.e., assuming that the *cs*
554 score is proportional to a local reduction in gene flow caused by associated barrier loci), the
555 observed variation in *cs* scores requires more conservative interpretation. First, neutral processes,
556 such as isolation-by-distance, can generate clines that are highly similar to clines generated by
557 selection [26]. Similarly, neutral clines generated by secondary contact can take a long time to
558 decay, making them hard to distinguish from selected ones [22, 51]. Localized drift (and
559 sampling effects) tends to distort cline shapes in a way that may lead to the discovery of false
560 positives [55, 56]. In addition, even though a genomic region may contain a large-effect barrier
561 locus, it might not show a cline if the genomic window is too broad to capture the relevant
562 signatures. Future efforts to help identify non-neutral clines may be accomplished using whole-
563 genome rather than reduced-representation sequencing, and by comparing results obtained from
564 multiple hybrid zones [27]. Simulations of cline formation could also help distinguish candidate
565 outliers (as in [26]). Even with these measures in place, the noise generated by background
566 processes and sampling effects may mean that we only have power to confidently detect large-
567 effect loci, which remains a general problem with all genome scan approaches [3].

568

569 *From the ‘top down’: insights from QTL analysis of candidate isolating traits*

570 We mapped QTL to identify genomic regions underlying floral trait divergence in this
571 system. Although we identified a small number of QTL for each trait (between 1 and 4), the
572 identified QTL explained only about 20% of the variation in each trait. Given that these traits are
573 under strong genetic control, the ‘missing’ variance implies that most of the floral traits are
574 polygenic, caused by many loci with effect sizes below our limit of detection. However, some
575 fraction of the unexplained variation also may be due to subtle environmental differences
576 experienced by each plant in the greenhouse.

577 Although finding many small-effect loci may be expected in studies of phenotypic
578 evolution [57], many analyses of adaptation and speciation have recovered distributions of effect
579 sizes skewed toward larger effects [58, 59]. Moreover, the identified regions often control more
580 than one trait, and in some cases, more than one type of isolating barrier (e.g., pre- and post-
581 mating barriers). For example, in another pair of *Mimulus* species, *M. cardinalis* and *M. lewisii*,
582 large-effect QTL for multiple traits associated with pollinator isolation and hybrid sterility occur
583 in a few genomic regions thought to harbor chromosomal inversions [4, 60]. In sunflowers
584 (*Helianthus*), multiple traits are associated with local adaptation to dune and non-dune habitats
585 and map to a small number of large, non-recombining haplotypes containing structural variants
586 [61] (see [62-64] for other examples).

587 Large, multiple effect loci are an expected outcome of local adaptation and speciation
588 [17, 65], because more ‘concentrated’ genetic architectures are favored in scenarios where gene
589 flow opposes adaptive divergence [65]. Not only do large effect loci make individual traits more

590 visible to selection, but tight linkage and pleiotropy enhance the coupling of different sets of
591 adaptive traits, meaning that they can remain associated despite gene flow [17]. This begs the
592 question: why do we see so few large effect loci and such little overlap among floral trait QTL in
593 the red and yellow ecotypes? One potential explanation is that divergence was initiated during a
594 period of geographic isolation—a hypothesis that is supported by our demographic analysis. If
595 trait divergence did occur during a phase of allopatry, the selection favoring certain combinations
596 of traits could build-up LD among many small-effect loci without opposition by gene flow.
597 Although the associations would decay rapidly upon secondary contact (as we see in the hybrid
598 zone; [24, 33]), this decay would be expected to occur over a spatial scale determined by the
599 strength of selection and the migration rate. If two adjacent habitats occur over a scale that is
600 many times larger than the dispersal distance of the organism (as is the case between the red and
601 yellow ecotypes; [33]), then hybridization has almost no bearing on adaptation occurring in
602 distant parts of the range [66]. This makes divergence in parapatry almost as easy as in allopatry
603 [67], meaning that local adaptation will persist far from the hybrid zone, and strong associations
604 among small effect loci can remain in all regions except for those closest to the hybrid zone.

605 Another factor that may have influenced the outcome of our QTL analysis is that the
606 parents we used for mapping came from populations located very far from hybrid zone. Because
607 concentrated genetic architectures evolve as a response to gene flow, theory predicts that the
608 genetic architecture of local adaptation may vary spatially in a way that reflects the local
609 hybridization risk. This was highlighted in a hybrid zone between two ecologically differentiated
610 subspecies of *Boechera stricta* [68]. In the vicinity of the hybrid zone, several phenological traits
611 map to a single locus containing a chromosomal inversion, where the ecotypes are differentially
612 fixed for the standard and inverted arrangements. However, in areas that are more distant to the
613 hybrid zone, both ecotypes harbor the standard arrangement, and QTL mapping with these
614 populations revealed distinct QTL. Although it is possible that the genetic architecture of
615 divergence also varies across the range of the red and yellow ecotypes in a way that might favor
616 divergence, our observations suggest that this is unlikely. Specifically, there is no evidence for
617 the phenotypic maintenance of two distinct ecotypic forms within the hybrid zone. Instead, we
618 see a continuum of phenotypic variation that resembles what we observe in the F₂ mapping
619 population [33], suggesting a similarly complex genetic architecture across the range.

620

621 *Inferences from integrating top-down and bottom up approaches*

622 Having identified a set of candidate barrier loci and QTL regions for putative barrier
623 traits, we next sought to understand how they were connected in relation to previous hypotheses
624 about reproductive isolation in this system [24]. Taken at face value, the two analyses seem
625 consistent, as they both suggest that divergence in this system is polygenic, involving regions
626 spread across the genome. However, when we intersect the regions identified by these
627 approaches, we find very little concordance. What does this tell us about divergence in this
628 system?

629 First, there are some technical and biological explanations that could account for these
630 findings. The first is that QTL analysis often has low resolution. Specifically, the QTL intervals
631 are very wide, substantially reducing our power to test if QTL regions are enriched for candidate
632 barrier loci. We see the same result if we focus on the estimated location of the QTL peaks,
633 which never fall inside a candidate barrier region. This is even true for the peak of the large-
634 effect QTL for floral anthocyanin pigmentation that is located 589 kb from the window with the
635 highest *cs* score and contains the *MaMyb2* gene, which we know is responsible for the major
636 difference in flower color [35]. Therefore, had we not had *a priori* knowledge about the position
637 of the gene from previous work, it is likely we would have failed to connect this strong signal of
638 selection with the underlying gene involved. In addition, our genome scan is based on RADseq
639 data, so the SNP density may have been too sparse to obtain sequences in LD with the loci under
640 selection. Moreover, from a biological perspective, the QTL analysis implies that most of the
641 traits studied are polygenic, meaning that selection on each locus is weak, making it difficult to
642 detect them using any genome scan. All of these factors probably contribute to the highly
643 complex pattern that we see.

644 Similarly, although many of the candidate barrier loci have clines that resemble the
645 window that contains the large effect locus *MaMyb2*, these windows are not associated with
646 large-effect QTL for floral traits. One possible explanation for this is that we may have failed to
647 measure the relevant floral traits contributing to pollinator isolation. This seems unlikely, given
648 how well floral morphology has been studied in this system [29, 31-33]. We therefore
649 hypothesize that other barriers to gene flow, besides pollinator isolation, play an important role
650 in the maintenance of this hybrid zone. For example, local adaptation of non-floral traits could be
651 an important source of pre- and post-mating isolation [13, 69]. We previously identified clines in
652 several eco-physiological traits that may be associated with habitat-based isolation [70], but
653 these show very shallow linear gradients rather than sharp sigmoid clines, and they are well
654 explained by variation in local climate variables, making them unlikely candidates [70].
655 However, other ecologically-based barriers may exist that remain to be characterized. Finally, it
656 is possible that some of the candidate barrier regions contain intrinsic incompatibilities that cause
657 reduced fitness in hybrids [71]. Although our previous work found little evidence for intrinsic
658 postzygotic isolation in the F_1 generation, we did detect partial male sterility in some inter-
659 ecotype crosses [36]. In addition, we have only surveyed plants under benign greenhouse
660 conditions and only through the F_1 generation. Genetic incompatibilities in later generations [72]
661 or under natural conditions [73] might play a larger role in the maintenance of the ecotypes than
662 anticipated—a prediction that has been made in relation to ‘ecological speciation’ more broadly
663 [15].

664 *Conclusions and implications for studying the architecture of speciation*

665 By combining top-down and bottom-up approaches with demographic modelling, our
666 study provides new insight into the history and genetic architecture of speciation between these
667 monkeyflower ecotypes. Our demographic analysis suggests that the hybrid zone formed by
668

669 secondary contact, but gene flow following contact is now heterogeneous across the genome due
670 to the effect of multiple barrier loci. A cline-based genome scan indicates that candidate barrier
671 loci are widespread across the genome, rather than being associated with one or a few ‘islands’
672 of speciation. A QTL analysis of floral traits identified many QTL of small effect, with limited
673 co-localization among QTL for different traits. We found limited evidence that QTL and
674 candidate barrier loci overlap, suggesting that other barriers to gene flow aside from pollinator
675 isolation may contribute to speciation.

676 In addition to providing knowledge about this system, our study has important
677 implications for efforts to understand the phenotypic and genetic architecture of isolating
678 barriers. For many study systems, candidate barrier traits and loci are identified in separate
679 studies, meaning that the link between them is not tested explicitly. However, any barrier locus
680 associated with an ecological gradient may underlie a completely different type of barrier. This
681 was highlighted by [15], who showed how intrinsic barriers can become spatially coupled with
682 ecologically-based barriers—a phenomenon that may cause researchers to erroneously identify
683 incompatibility loci as those underlying local adaptation. The same issue also arises if multiple
684 ecological gradients change in concert. We therefore advocate for additional studies that
685 integrate top-down and bottom-up approaches before drawing conclusions about causal
686 associations between candidate barrier traits and loci. Finally, our study shows that, even with a
687 concerted effort, understanding the phenotypic and genetic basis of speciation is extremely
688 difficult. Although emerging methods and data may help, this will likely remain a major
689 challenge for the field.

690

691 **Data availability**

692 Sequence data are available on the short read archive (SRA) under bio projects PRJNA299226
693 and PRJNA317601. All other datafiles are available on dryad. Scripts for performing custom
694 analyses are available at https://github.com/seanstankowski/Mimulus_cline_QTL_manuscript.

695

696 **Acknowledgments**

697 We thank Julian Catchen for making modifications to *Stacks* to aid this project. Peter L. Ralph,
698 Thomas Nelson and Roger K. Butlin and Anja M. Westram and Nicholas H. Barton provided
699 advice, stimulating discussion, and critical feedback. The project was supported by National
700 Science Foundation grant DEB-1258199.

701

702 **Author contributions**

703 SS and MAS designed the study and conducted the molecular work. SS and HM performed the
704 QTL analysis. SS, MAC, and MAS conducted all other analyses. SS and MAS wrote the
705 manuscript, with input from the other authors. Scripts for performing custom analyses are
706 available at https://github.com/seanstankowski/Mimulus_cline_QTL_manuscript.

707

708 **Competing interests**

709 The authors declare no competing financial interests

710

711 **References**

- 712 1. Nosil P. *Ecological Speciation*: Oxford University Press; 2012.
- 713 2. Seehausen O, Butlin RK, Keller I, Wagner CE, Boughman JW, Hohenlohe PA, et al.
714 Genomics and the origin of species. *Nat Rev Genet.* 2014;15(3):176-92. doi: 10.1038/nrg3644.
715 PubMed PMID: ISI:000331895300010.
- 716 3. Ravinet M, Faria R, Butlin RK, Galindo J, Bierne N, Rafajlovic M, et al. Interpreting the
717 genomic landscape of speciation: a road map for finding barriers to gene flow. *J Evolution Biol.*
718 2017;30(8):1450-77. doi: 10.1111/jeb.13047. PubMed PMID: ISI:000407280900001.
- 719 4. Schemske DW, Bradshaw HD. Pollinator preference and the evolution of floral traits in
720 monkeyflowers (*Mimulus*). *P Natl Acad Sci USA.* 1999;96(21):11910-5. doi: DOI
721 10.1073/pnas.96.21.11910. PubMed PMID: ISI:000083166800043.
- 722 5. Fishman L, Willis JH. Evidence for Dobzhansky-Muller incompatibilities contributing to
723 the sterility of hybrids between *Mimulus guttatus* and *M-nasutus*. *Evolution.* 2001;55(10):1932-
724 42. PubMed PMID: ISI:000172308600003.
- 725 6. Orr HA. The genetics of species differences. *Trends Ecol Evol.* 2001;16(7):343-50.
726 PubMed PMID: ISI:000169461000004.
- 727 7. Lamichhaney S, Berglund J, Almen MS, Maqbool K, Grabherr M, Martinez-Barrio A, et al.
728 Evolution of Darwin's finches and their beaks revealed by genome sequencing. *Nature.*
729 2015;518(7539):371-5. Epub 2015/02/18. doi: 10.1038/nature14181. PubMed PMID: 25686609.
- 730 8. Turner TL, Hahn MW, Nuzhdin SV. Genomic islands of speciation in *Anopheles*
731 *gambiae*. *Plos Biol.* 2005;3(9):1572-8. doi: ARTN e285
732 10.1371/journal.pbio.0030285. PubMed PMID: ISI:000231820900008.
- 733 9. Martin SH, Dasmahapatra KK, Nadeau NJ, Salazar C, Walters JR, Simpson F, et al.
734 Genome-wide evidence for speciation with gene flow in *Heliconius* butterflies. *Genome Res.*
735 2013;23(11):1817-28. doi: 10.1101/gr.159426.113. PubMed PMID: ISI:000326642500006.
- 736 10. Soria-Carrasco V, Gompert Z, Comeault AA, Farkas TE, Parchman TL, Johnston JS, et al.
737 Stick Insect Genomes Reveal Natural Selection's Role in Parallel Speciation. *Science.*
738 2014;344(6185):738-42. doi: 10.1126/science.1252136. PubMed PMID: ISI:000335912900039.
- 739 11. Barrett RDH, Hoekstra HE. Molecular spandrels: tests of adaptation at the genetic level
740 (vol 12, pg 767, 2011). *Nat Rev Genet.* 2012;13(1):70-. doi: 10.1038/nrg3138. PubMed PMID:
741 ISI:000298327300014.
- 742 12. Cruickshank TE, Hahn MW. Reanalysis suggests that genomic islands of speciation are
743 due to reduced diversity, not reduced gene flow. *Molecular ecology.* 2014;23(13):3133-57. Epub
744 2014/05/23. doi: 10.1111/mec.12796. PubMed PMID: 24845075.
- 745 13. Coyne JA, Orr HA. *Speciation*. Sunderland, MA: Sinauer Associates; 2004.
- 746 14. Stankowski S, Ravinet M. Defining the speciation continuum. *Evolution.*
747 2021;75(6):1256-73. doi: 10.1111/evo.14215. PubMed PMID: ISI:000647226400001.
- 748 15. Bierne N, Welch J, Loire E, Bonhomme F, David P. The coupling hypothesis: why
749 genome scans may fail to map local adaptation genes. *Molecular ecology.* 2011;20(10):2044-72.
750 doi: 10.1111/j.1365-294X.2011.05080.x. PubMed PMID: ISI:000290168900004.
- 751 16. Butlin RK, Smadja CM. Coupling, Reinforcement, and Speciation. *Am Nat.*
752 2018;191(2):155-72. doi: 10.1086/695136. PubMed PMID: ISI:000422944300003.

- 753 17. Smadja CM, Butlin RK. A framework for comparing processes of speciation in the
754 presence of gene flow. *Molecular ecology*. 2011;20(24):5123-40. doi: 10.1111/j.1365-
755 294X.2011.05350.x. PubMed PMID: ISI:000298089300004.
- 756 18. Faria R, Johannesson K, Stankowski S. Speciation in marine environments: Diving under
757 the surface. *J Evolution Biol*. 2021;34(1):4-15. doi: 10.1111/jeb.13756. PubMed PMID:
758 ISI:000608367500001.
- 759 19. Abbott R, Albach D, Ansell S, Arntzen JW, Baird SJE, Bierne N, et al. Hybridization and
760 speciation. *J Evolution Biol*. 2013;26(2):229-46. doi: 10.1111/j.1420-9101.2012.02599.x.
761 PubMed PMID: ISI:000313747600001.
- 762 20. Harrison RG. Hybrid zones: windows on evolutionary process. *Oxford Surveys in*
763 *Evolutionary Biology*. 1990;7:69-128.
- 764 21. Barton NH, Gale KS. Genetic analysis of hybrid zones. *Hybrid Zones and Evolutionary*
765 *Process*1993. p. 13-45.
- 766 22. Barton NH, Hewitt GM. Analysis of Hybrid Zones. *Annu Rev Ecol Syst*. 1985;16:113-
767 48. doi: DOI 10.1146/annurev.es.16.110185.000553. PubMed PMID: ISI:A1985AUL3900006.
- 768 23. Szymura JM, Barton NH. Genetic-Analysis of a Hybrid Zone between the Fire-Bellied
769 Toads, *Bombina-Bombina* and *Bombina-Variegata*, near Cracow in Southern Poland. *Evolution*.
770 1986;40(6):1141-59. doi: Doi 10.2307/2408943. PubMed PMID: ISI:A1986F052900004.
- 771 24. Stankowski S, Sobel JM, Streisfeld MA. Geographic cline analysis as a tool for studying
772 genome-wide variation: a case study of pollinator-mediated divergence in a monkeyflower.
773 *Molecular ecology*. 2017;26(1):107-22. doi: 10.1111/mec.13645. PubMed PMID:
774 ISI:000391940600008.
- 775 25. Singhal S, Bi K. History cleans up messes: The impact of time in driving divergence and
776 introgression in a tropical suture zone. *Evolution*. 2017;71(7):1888-99. doi: 10.1111/evo.13278.
777 PubMed PMID: ISI:000405888100013.
- 778 26. Westram AM, Rafajlovic M, Chaube P, Faria R, Larsson T, Panova M, et al. Clines on
779 the seashore: The genomic architecture underlying rapid divergence in the face of gene flow.
780 *Evol Lett*. 2018;2(4):297-309. doi: 10.1002/evl3.74. PubMed PMID: ISI:000446774400004.
- 781 27. Westram AM, Faria R, Johannesson K, Butlin R. Using replicate hybrid zones to
782 understand the genomic basis of adaptive divergence. *Molecular ecology*. 2021;30(15):3797-814.
783 doi: 10.1111/mec.15861. PubMed PMID: ISI:000669439700001.
- 784 28. Rafati N, Blanco-Aguiar JA, Rubin CJ, Sayyab S, Sabatino SJ, Afonso S, et al. A
785 genomic map of clinal variation across the European rabbit hybrid zone. *Molecular ecology*.
786 2018;27(6):1457-78. doi: 10.1111/mec.14494. PubMed PMID: ISI:000430183100010.
- 787 29. Streisfeld MA, Kohn JR. Contrasting patterns of floral and molecular variation across a
788 cline in *Mimulus aurantiacus*. *Evolution*. 2005;59(12):2548-59. PubMed PMID:
789 ISI:000234502400006.
- 790 30. Stankowski S, Chase MA, Fuiten AM, Rodrigues MF, Ralph PL, Streisfeld MA.
791 Widespread selection and gene flow shape the genomic landscape during a radiation of
792 monkeyflowers. *Plos Biol*. 2019;17(7). doi: ARTN e3000391
793 10.1371/journal.pbio.3000391. PubMed PMID: ISI:000478922100030.
- 794 31. Waayers GM. Hybridization, introgression, and selection in *Mimulus aurantiacus* ssp.
795 *australis* and *Mimulus puniceus*. San Diego, CA: San Diego State University; 1996.
- 796 32. Tulig M. Morphological variation in *Mimulus* section *Diplacus* (*Scrophulariaceae*).
797 Pomona, CA: California Polytechnic University; 2000.

- 798 33. Stankowski S, Sobel JM, Streisfeld MA. The geography of divergence with gene flow
799 facilitates multitrait adaptation and the evolution of pollinator isolation in *Mimulus aurantiacus*.
800 *Evolution*. 2015;69(12):3054-68. doi: 10.1111/evo.12807. PubMed PMID:
801 ISI:000368249000003.
- 802 34. Streisfeld MA, Kohn JR. Environment and pollinator-mediated selection on parapatric
803 floral races of *Mimulus aurantiacus*. *J Evolution Biol*. 2007;20(1):122-32. doi: 10.1111/j.1420-
804 9101.2006.01216.x. PubMed PMID: ISI:000242904600020.
- 805 35. Streisfeld MA, Young WN, Sobel JM. Divergent Selection Drives Genetic
806 Differentiation in an R2R3-MYB Transcription Factor That Contributes to Incipient Speciation
807 in *Mimulus aurantiacus*. *Plos Genet*. 2013;9(3). doi: ARTN e1003385
808 10.1371/journal.pgen.1003385. PubMed PMID: ISI:000316866700058.
- 809 36. Sobel JM, Streisfeld MA. Strong premating reproductive isolation drives incipient
810 speciation in *Mimulus aurantiacus*. *Evolution*. 2015;69(2):447-61. doi: 10.1111/evo.12589.
811 PubMed PMID: ISI:000348916200013.
- 812 37. Catchen J, Hohenlohe PA, Bassham S, Amores A, Cresko WA. Stacks: an analysis tool
813 set for population genomics. *Molecular ecology*. 2013;22(11):3124-40. Epub 2013/05/25. doi:
814 10.1111/mec.12354. PubMed PMID: 23701397; PubMed Central PMCID: PMC3936987.
- 815 38. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. *Nat Methods*.
816 2012;9(4):357-U54. doi: 10.1038/Nmeth.1923. PubMed PMID: ISI:000302218500017.
- 817 39. Hohenlohe PA, Bassham S, Etter PD, Stiffler N, Johnson EA, Cresko WA. Population
818 Genomics of Parallel Adaptation in Threespine Stickleback using Sequenced RAD Tags. *Plos*
819 *Genet*. 2010;6(2). doi: ARTN e1000862
820 10.1371/journal.pgen.1000862. PubMed PMID: ISI:000275262700012.
- 821 40. Gutenkunst RN, Hernandez RD, Williamson SH, Bustamante CD. Inferring the Joint
822 Demographic History of Multiple Populations from Multidimensional SNP Frequency Data. *Plos*
823 *Genet*. 2009;5(10). doi: ARTN e1000695
824 10.1371/journal.pgen.1000695. PubMed PMID: ISI:000272032100033.
- 825 41. Chase MA, Stankowski S, Streisfeld MA. Genomewide variation provides insight into
826 evolutionary relationships in a monkeyflower species complex (*Mimulus* sect. *Diplacus*). *Am J*
827 *Bot*. 2017;104(10):1510-21. doi: 10.3732/ajb.1700234. PubMed PMID: ISI:000413976200014.
- 828 42. Rougeux C, Bernatchez L, Gagnaire PA. Modeling the Multiple Facets of Speciation-
829 with-Gene-Flow toward Inferring the Divergence History of Lake Whitefish Species Pairs
830 (*Coregonus clupeaformis*). *Genome Biol Evol*. 2017;9(8):2057-74. doi: 10.1093/gbe/evx150.
831 PubMed PMID: ISI:000410677300007.
- 832 43. Ossowski S, Schneeberger K, Lucas-Lledo JI, Warthmann N, Clark RM, Shaw RG, et al.
833 The Rate and Molecular Spectrum of Spontaneous Mutations in *Arabidopsis thaliana*. *Science*.
834 2010;327(5961):92-4. doi: 10.1126/science.1180677. PubMed PMID: ISI:000273395400038.
- 835 44. Alexander DH, Novembre J, Lange K. Fast model-based estimation of ancestry in
836 unrelated individuals. *Genome Res*. 2009;19(9):1655-64. doi: 10.1101/gr.094052.109. PubMed
837 PMID: ISI:000269482200017.
- 838 45. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MAR, Bender D, et al. PLINK:
839 A tool set for whole-genome association and population-based linkage analyses. *Am J Hum*
840 *Genet*. 2007;81(3):559-75. doi: 10.1086/519795. PubMed PMID: ISI:000249128200012.
- 841 46. Derryberry EP, Derryberry GE, Maley JM, Brumfield RT. hzar: hybrid zone analysis
842 using an R software package. *Mol Ecol Resour*. 2014;14(3):652-63. doi: 10.1111/1755-
843 0998.12209. PubMed PMID: ISI:000334240000022.

- 844 47. Excoffier L, Lischer HEL. Arlequin suite ver 3.5: a new series of programs to perform
845 population genetics analyses under Linux and Windows. *Mol Ecol Resour.* 2010;10(3):564-7.
846 doi: 10.1111/j.1755-0998.2010.02847.x. PubMed PMID: ISI:000276407300020.
- 847 48. Broman KW, Wu H, Sen S, Churchill GA. R/qtl: QTL mapping in experimental crosses.
848 *Bioinformatics.* 2003;19(7):889-90. doi: 10.1093/bioinformatics/btg112. PubMed PMID:
849 ISI:000182634800014.
- 850 49. Rastas P, Calboli FCF, Guo BC, Shikano T, Merila J. Construction of Ultradense Linkage
851 Maps with Lep-MAP2: Stickleback F-2 Recombinant Crosses as an Example. *Genome Biol*
852 *Evol.* 2016;8(1):78-93. doi: 10.1093/gbe/evv250. PubMed PMID: ISI:000370971600007.
- 853 50. Lawrence M, Huber W, Pages H, Aboyoun P, Carlson M, Gentleman R, et al. Software
854 for Computing and Annotating Genomic Ranges. *Plos Comput Biol.* 2013;9(8). doi: ARTN
855 e1003118
856 10.1371/journal.pcbi.1003118. PubMed PMID: ISI:000323885400001.
- 857 51. Endler JA. *Geographic Variation, Speciation, and Clines*: Princeton University Press;
858 1977.
- 859 52. Hewitt GM. Hybrid Zones - Natural Laboratories for Evolutionary Studies. *Trends Ecol*
860 *Evol.* 1988;3(7):158-67. doi: Doi 10.1016/0169-5347(88)90033-X. PubMed PMID:
861 ISI:A1988N945300004.
- 862 53. Momigliano P, Florin AB, Merila J. Biases in Demographic Modeling Affect Our
863 Understanding of Recent Divergence. *Mol Biol Evol.* 2021;38(7):2967-85. doi:
864 10.1093/molbev/msab047. PubMed PMID: ISI:000671060500022.
- 865 54. Stankowski S, Streisfeld MA. Introgressive hybridization facilitates adaptive divergence
866 in a recent radiation of monkeyflowers. *P Roy Soc B-Biol Sci.* 2015;282(1814):154-62. doi: Artn
867 20151666
868 10.1098/Rspb.2015.1666. PubMed PMID: ISI:000363361800019.
- 869 55. Polechova J, Barton N. Genetic Drift Widens the Expected Cline but Narrows the
870 Expected Cline Width. *Genetics.* 2011;189(1):227-U905. doi: 10.1534/genetics.111.129817.
871 PubMed PMID: ISI:000294721600018.
- 872 56. Jofre GI, Rosenthal GG. A narrow window for geographic cline analysis using genomic
873 data: Effects of age, drift, and migration on error rates. *Mol Ecol Resour.* 2021;21(7):2278-87.
874 doi: 10.1111/1755-0998.13428. PubMed PMID: ISI:000656092200001.
- 875 57. Rockman MV. The Qtn Program and the Alleles That Matter for Evolution: All That's
876 Gold Does Not Glitter. *Evolution.* 2012;66(1):1-17. doi: 10.1111/j.1558-5646.2011.01486.x.
877 PubMed PMID: ISI:000298739100001.
- 878 58. Griswold CK. Gene flow's effect on the genetic architecture of a local adaptation and its
879 consequences for QTL analyses. *Heredity.* 2006;96(6):445-53. doi: 10.1038/sj.hdy.6800822.
880 PubMed PMID: ISI:000237751700006.
- 881 59. Lenormand T. Gene flow and the limits to natural selection. *Trends Ecol Evol.*
882 2002;17(4):183-9. doi: Doi 10.1016/S0169-5347(02)02497-7. PubMed PMID:
883 ISI:000174446400011.
- 884 60. Fishman L, Stathos A, Beardsley PM, Williams CF, Hill JP. Chromosomal
885 Rearrangements and the Genetics of Reproductive Barriers in *Mimulus* (Monkey Flowers).
886 *Evolution.* 2013;67(9):2547-60. doi: 10.1111/evo.12154. PubMed PMID: ISI:000323828500007.
- 887 61. Todesco M, Owens GL, Bercovich N, Legare JS, Soudi S, Burge DO, et al. Massive
888 haplotypes underlie ecotypic differentiation in sunflowers. *Nature.* 2020;584(7822):602-+. doi:
889 10.1038/s41586-020-2467-6. PubMed PMID: ISI:000546555800006.

- 890 62. Lowry DB, Willis JH. A Widespread Chromosomal Inversion Polymorphism Contributes
891 to a Major Life-History Transition, Local Adaptation, and Reproductive Isolation. *Plos Biol.*
892 2010;8(9). doi: ARTN e1000500
893 10.1371/journal.pbio.1000500. PubMed PMID: ISI:000282279200019.
- 894 63. Jones FC, Grabherr MG, Chan YF, Russell P, Mauceli E, Johnson J, et al. The genomic
895 basis of adaptive evolution in threespine sticklebacks. *Nature.* 2012;484(7392):55-61. doi:
896 10.1038/nature10944. PubMed PMID: ISI:000302343400033.
- 897 64. Hager ER, Harringmeyer OS, Wooldridge TB, Theingi S, Gable JT, McFadden S, et al. A
898 chromosomal inversion drives evolution of multiple adaptive traits in deer mice. *bioRxiv.* 2021.
899 doi: <https://doi.org/10.1101/2021.01.21.427490>.
- 900 65. Yeaman S, Whitlock MC. The Genetic Architecture of Adaptation under Migration-
901 Selection Balance. *Evolution.* 2011;65(7):1897-911. doi: 10.1111/j.1558-5646.2011.01269.x.
902 PubMed PMID: ISI:000292447000006.
- 903 66. Barton NH. What role does natural selection play in speciation? *Philos T R Soc B.*
904 2010;365(1547):1825-40. doi: 10.1098/rstb.2010.0001. PubMed PMID: ISI:000277208600009.
- 905 67. Barton NH. Does hybridization influence speciation? *J Evolution Biol.* 2013;26(2):267-9.
906 doi: 10.1111/jeb.12015. PubMed PMID: ISI:000313747600007.
- 907 68. Lee CR, Wang BS, Mojica JP, Mandakova T, Prasad KYSK, Goicoechea JL, et al.
908 Young inversion with multiple linked QTLs under selection in a hybrid zone. *Nat Ecol Evol.*
909 2017;1(5). doi: ArtN 0119
910 10.1038/S41559-017-0119. PubMed PMID: ISI:000417173100009.
- 911 69. Sobel JM, Chen GF, Watt LR, Schemske DW. The Biology of Speciation. *Evolution.*
912 2010;64(2):295-315. doi: 10.1111/j.1558-5646.2009.00877.x. PubMed PMID:
913 ISI:000273821700001.
- 914 70. Sobel JM, Stankowski S, Streisfeld MA. Variation in ecophysiological traits might
915 contribute to ecogeographic isolation and divergence between parapatric ecotypes of *Mimulus*
916 *aurantiacus*. *J Evolution Biol.* 2019;32(6):604-18. doi: 10.1111/jeb.13442. PubMed PMID:
917 ISI:000472662300008.
- 918 71. Kulmuni J, Westram AM. Intrinsic incompatibilities evolving as a by-product of
919 divergent ecological selection: Considering them in empirical studies on divergence with gene
920 flow. *Molecular ecology.* 2017;26(12):3093-103. doi: 10.1111/mec.14147. PubMed PMID:
921 ISI:000402836700003.
- 922 72. Stelkens RB, Schmid C, Seehausen O. Hybrid Breakdown in Cichlid Fish. *Plos One.*
923 2015;10(5). doi: ARTN e0127207
924 10.1371/journal.pone.0127207. PubMed PMID: ISI:0003564444000050.
- 925 73. Thompson K, Peichel CH, Rennison DK, McGee MR, Albert AYK, Vines T, et al.
926 Analysis of ancestry heterozygosity suggests that hybrid incompatibilities in threespine
927 stickleback are environment dependent. *Plos Biol.* 2022;20(1). doi: ARTN e3001469
928 10.1371/journal.pbio.3001469. PubMed PMID: ISI:000740930300001.
- 929 74. Streisfeld MA, Rausher MD. Altered trans-Regulatory Control of Gene Expression in
930 Multiple Anthocyanin Genes Contributes to Adaptive Flower Color Evolution in *Mimulus*
931 *aurantiacus*. *Mol Biol Evol.* 2009;26(2):433-44. doi: 10.1093/molbev/msn268. PubMed PMID:
932 ISI:000262525700018.
- 933
934
935

936 **Supplementary tables and figure captions**

937 **Table S1. Location information and samples sizes for the sample sites used in the study.**

938 Population codes correspond with those in Figure 1.

939

940 **Table S2. Estimated parameters from each demographic model.** N_a , size of the ancestral
941 population; N_r , size of the red population, N_y , size of the yellow population; T_s , duration of the
942 split; m_{ry} , migration from yellow into red; m_{yr} , migration from red into yellow; T_{am} , duration of
943 ancient migration; T_{SC} , duration of secondary contact.

944

945 **Figure S1. Cartoons of each model tested in the demographic analysis.** (i) SI, strict isolation;
946 (ii) AM, ancient migration; (iii) IM, isolation with migration; (iv) SC, Secondary contact; (v)
947 PSC, Periods of secondary contact. The remaining four models—AM2*m*, IM2*m*, SC2*m*,
948 PSC2*m*—are the same as models *ii-v*, except that migration rates are inferred for two groups of
949 loci to simulate the effect of a porous barrier to gene flow. The model parameters are as follows:
950 N_a , size of the ancestral population; N_r , size of the red population, N_y , size of the yellow
951 population; T_s , duration of the split; m_{ry} , migration from yellow into red; m_{yr} , migration from red
952 into yellow; m_{ery} , effective migration from yellow into red; m_{eyr} , effective migration from red into
953 yellow; T_{am} , duration of ancient migration; T_{SC} , duration of secondary contact.

954

955 **Figure S2. Information on the size and content of the 100 SNP windows.** Distributions of
956 window size (kb) and the number of RAD tags sequenced (post filtering) within each window.

957

958 **Figure S3. Floral traits measured for the QTL analysis.** In addition to the size-related traits,
959 we also measured anthocyanin and carotenoid pigment levels. Anthocyanins were extracted in
960 1% acidic methanol from a single disc collected from one of the top petals the first day each
961 flower opened. Absorbance of extracts was measured with a spectrophotometer at 520 nm, as
962 described previously [74]. Carotenoids were extracted in hexane following a similar protocol,
963 and absorbance was measured at 450 nm.

964

965 **Figure S4. Fits of the different demographic models to the observed data.** The observed
966 unfolded joint site-frequency spectrum (JSFS) calculated from the RADseq data set is shown on
967 the left. The first column shows the modeled JSFS for the best fit of each model (See Fig. S1 for
968 a cartoon of each model). The second column shows the residuals of the best model fit to the
969 observed JSFS. The third column shows the residuals plotted as a histogram.

970

971 **Figure S5. The results of an *Admixture* analysis conducted on the full SNP dataset.** Each bar
972 represents an individual and shows its probability of membership (Q score) to the two different
973 clusters. Individuals are grouped by whether they come from the distribution of the red ecotype,
974 yellow ecotype, or from within the hybrid zone.

975

976 **Figure S6. Cline fits for each chromosome.** Black points are the mean ancestry scores plotted
977 along the one dimensional transect. The dashed line is the ML sigmoid model.

978

979 **Figure S7. Relationship between the cline similarity score (cs) and F_{CT} .** The line through the
980 plot is the least squares regression ($r^2 = 0.38$).

981

982 **Figure S8. Trait variation in greenhouse-raised red ecotype, yellow ecotype, F_1 , and F_2**
983 **individuals.** The histogram in each plot shows the distribution for each trait in the F_2 population.
984 The top, middle and lower box plots show the distributions for the yellow ecotype, red ecotype
985 and F_1 , respectively. ΔSD indicates the number of standard deviation that the red and yellow
986 ecotypes differ by.

987

988 **Figure S9. LOD scores plotted across the genome for each trait in the QTL analysis.** The
989 horizontal red line shows the LOD score cutoff for the detection of a significant QTL. Note the
990 different y-axis scales among traits.

991

992 **Figure S10. QTL effect plots for each of the 26 QTL discovered in the analysis.** Each plot
993 shows the mean (circle) and standard deviation (crosses) for each trait plotted against the three
994 genotypes. RR, both alleles inherited from a red grandparent; YY, both alleles inherited from a
995 yellow grandparent; RY, Heterozygous. Red lines indicate that the effect occurs in the opposite
996 direction of that inferred by the trait differences measured in the red and yellow grandparents
997 (see Fig. S8).

998

999 **Figure 11. Distribution of effect sizes (percent variation explained among F_2 plants) for the**
1000 **26 QTL detected in the study.**

1001

1002 **Figure S12. Results from the permutation test for the significant colocalization of QTL.** The
1003 dashed red line shows the average number of observed overlaps per QTL, while the histogram
1004 shows the observed null distribution, generated from 9,999 random permutations. The p -value
1005 for the test is shown at the top of the plot.

1006

1007 **Figure S13. Correlation matrix showing the Pearson correlation between each pair of traits**
1008 **in the F_2 population.** The color and shape of each ellipse indicates the strength and direction of
1009 each correlation.

1010

1011 **Figure S14. Results from the permutation tests for significant enrichment of candidate**
1012 **barrier loci within the QTL Bayes credible intervals.** The test is performed with two different
1013 arbitrary cutoffs for defining candidate barrier loci (top 1% and top 5% of the distribution) for
1014 two different statistics (the cline similarity score and F_{CT}). The red line shows the observed
1015 number of candidate barrier loci falling within the QTL intervals, while the histogram shows the

1016 observed null distribution, generated from 9,999 random permutations. The p -value for the test is
1017 shown at the top of the plot.

1018

1019 **Figure 1. Clinal variation across a bush monkeyflower hybrid zone. (top)** Typical flower
1020 phenotypes of the red and yellow ecotypes, and a map of the 25 sampling locations in San Diego
1021 County. The size of the circles shows variation in the sample sizes, which range from 4 to 18
1022 individuals, totaling 292 individuals. The dashed line indicates the center of the hybrid zone,
1023 previously inferred from spatial variation in the frequency of alternative alleles at the *MaMyb2*
1024 locus. **(bottom)** Clines in allele frequency at the *MaMyb2* locus (red circles) and the mean floral
1025 trait PC1 score (blue squares) across the one-dimensional transect. The solid and dashed lines are
1026 the ML sigmoid cline models for *MaMyb2* allele frequency and trait PC1 score, respectively. The
1027 gray shaded rectangle represents the width of the hybrid zone.

1028

1029 **Figure 2. Demographic modeling reveals a history of gene flow following isolation. (a) -**
1030 Δ AIC scores for the 9 demographic models fitted to the observed JSFS using $\partial a \partial i$. The base
1031 models (left of the dashed line) include a single migration parameter (m) for all loci, whereas the
1032 $2m$ models include separate migration parameters for neutral loci (m) and loci affected by a
1033 barrier to gene flow (m_e). The best model (SC2 m) has a $-\Delta$ AIC of 0, with more negative values
1034 indicating that models were a poorer fit. **(b)** A graphical depiction of the SC2 m model. The width
1035 of the columns is proportional to the population size estimates for the ancestral (N_a), red (N_R),
1036 and yellow (N_Y) populations. the height of the red and yellow bars is proportional to the total
1037 time in generations (T_S) that has passed since the split. The blue bar shows the period during
1038 which secondary gene flow (T_{SC}) occurred. The difference in arrow size is proportional to the
1039 difference in the bi-directional migration rate, m . The rates of effective migration (m_e) are too
1040 small to show graphically.

1041

1042 **Figure 3. Clines in ancestry scores at different scales of genomic organization. (a)** Genome-
1043 wide cline, inferred from the mean ancestry scores in each population along the 1-D transect.
1044 Position 0 on the horizontal axis corresponds to the cline center estimated from *MaMyb2* allele
1045 frequencies (see Fig 1b). The vertical bars show the standard deviation in ancestry scores for
1046 each population. The dashed line is the ML cline model, and the gray band is the two-unit
1047 support envelope. Three parameters of interest, including the cline center (c), width (w), and total
1048 change in ancestry across the cline (ΔQ), are indicated on the plot. **(b)** Ancestry clines estimated
1049 separately for each chromosome. Only the ML curves are shown for clarity (but see Fig S6). The
1050 dashed line is the mean cline, estimated by taking the average of the ML parameters for all
1051 chromosomes. **(c)** Ancestry clines estimated for 2,173, 100-SNP windows. The dashed cyan line
1052 shows the cline shape for the genome-wide cline (as shown in panel a), while the dashed orange
1053 line is the mean cline shape, estimated by taking the average of the ML parameters obtained for
1054 all windows. Each solid line is the ML sigmoid curve for one of the genomic windows. The
1055 curves are colored according to the value of the cline similarity score (cs), which indicates how

1056 similar the shape and position of each cline is to the genome-wide cline. Redder clines are more
1057 similar to the genome-wide cline and bluer clines are less similar (See main text for more
1058 details).

1059

1060 **Figure 4. Cline-based genome scan and locations of QTL for floral traits. (top)** The scaled
1061 cline similarity (*cs* score) score in each 100 SNP window plotted against the physical position of
1062 the window in the bush monkeyflower genome. The points are colored as in Figure 3c, with
1063 redder points containing windows with *cs* scores that are more similar to the genome-wide
1064 pattern and bluer points are less similar (See main text for more details). The orange asterisk
1065 denotes the average barrier score among all windows. The position of the *MaMyb2* gene that
1066 controls differences in flower color is shown. **(bottom)** The positions of the QTL for the 13
1067 measured floral traits plotted along the physical position of the genome. The red vertical line
1068 corresponds to the best estimate of the QTL peak, and the width of the rectangles denotes the
1069 95% Bayes credible intervals of the estimated QTL position.

1070

Table S1. Location information and samples sizes for the sample sites used in the study.
Population codes correspond with those in Figure 1.

Pop name	Latitude	Longitude	1-D distance (in meters)	<i>N</i>
CRS	33.13037	-117.30717	-26461	11
UCSD	32.8894	-117.23618	-25369	12
SDP	32.9981	-117.23538	-23857	12
MT	32.82095	-117.06175	-19135	12
JMC	32.73732	-116.9541	-13303	18
ELF	33.08595	-117.1453	-12893	11
LH	32.80582	-116.9867	-12730	12
FLP	32.89422	-117.08982	-12557	12
ELT	33.06088	-117.11877	-12168	12
PMD	32.93787	-117.05913	-8097	12
WM	32.82133	-116.90228	-4298	17
BS	33.0148	-117.01643	-3710	18
DLR	33.16818	-117.05237	-2153	9
MW	33.00718	-116.95978	-827	17
OAK	32.91407	-116.88932	-363	12
LKW	33.16372	-117.0161	1177	15
DLZ	32.6525	-116.78597	2891	4
WCR	32.9805	-116.826517	8277	12
AND	32.868233	-116.7461	9643	4
BC	33.12262	-116.80468	10311	13
POTR	32.6038	-116.63392	18064	11
INJ	33.09785	-116.66432	22509	11
BCRD	32.94958	-116.63795	22615	6
PCT	32.73258	-116.46983	32097	7
LO	32.6767	-116.33123	45224	12

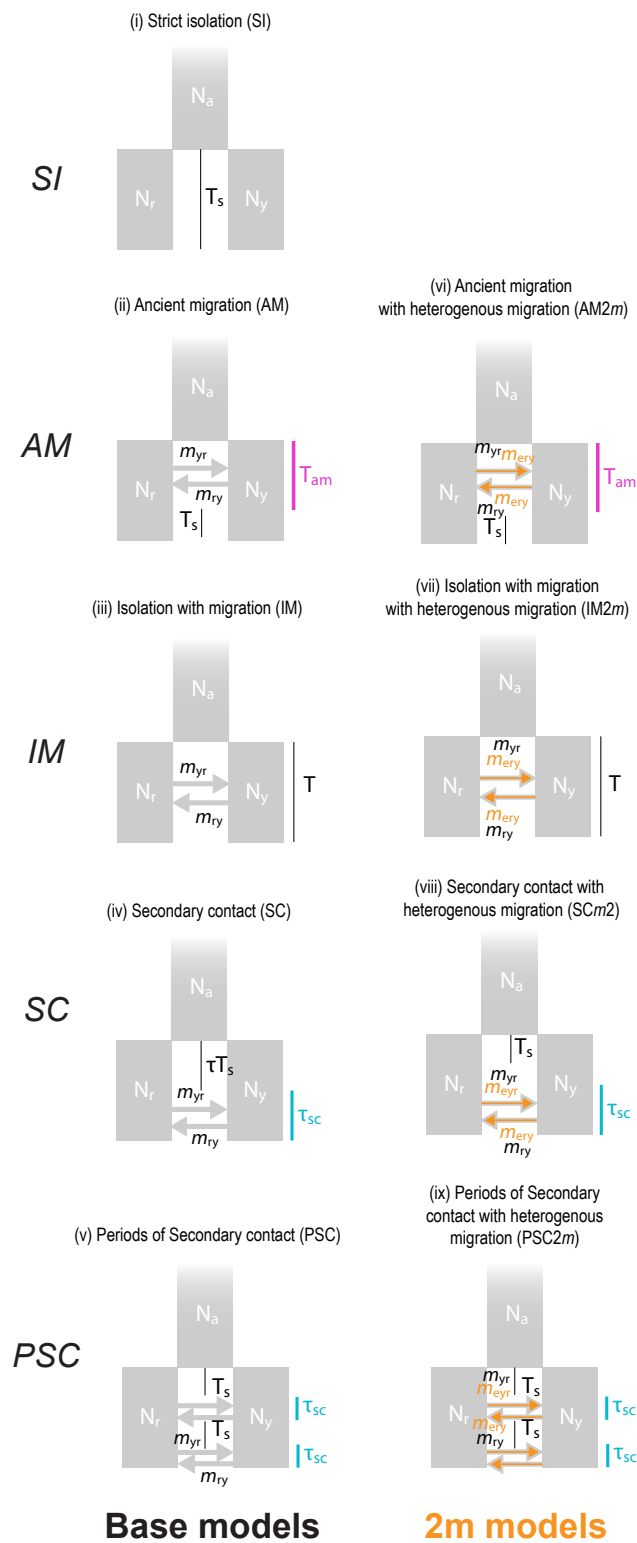


Figure S1. Cartoons of each model tested in the demographic analysis.

(i) SI, strict isolation; (ii) AM, ancient migration; (iii) IM, isolation with migration; (iv) SC, Secondary contact; (v) PSC, Periods of secondary contact. The remaining four models—AM2m, IM2m, SC2m, PSC2m—are the same as models *ii-v*, except that migration rates are inferred for two groups of loci to simulate the effect of a porous barrier to gene flow. The model parameters are as follows: N_a , size of the ancestral population; N_r , size of the red population, N_y , size of the yellow population; T_s , duration of the split; m_{yr} , migration from yellow into red; m_{ry} , migration from red into yellow; m_{ery} , effective migration from yellow into red; m_{eyr} , effective migration from red into yellow; T_{am} , duration of ancient migration; T_{sc} , duration of secondary contact.

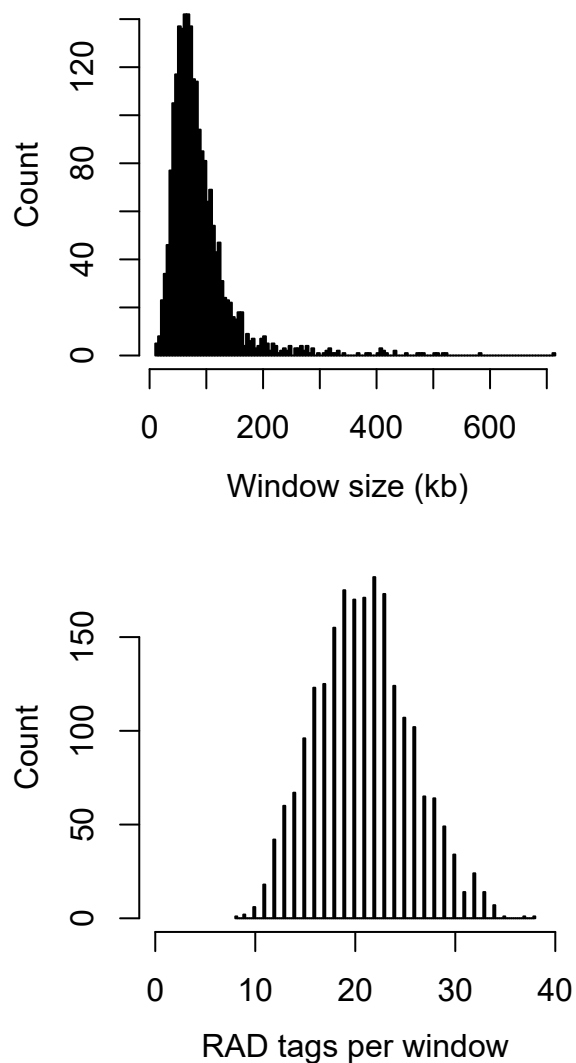


Figure S2. Information on the size and content of the 100 SNP windows. Distributions of window size (kb) and the number of RAD tags sequenced (post filtering) within each window.

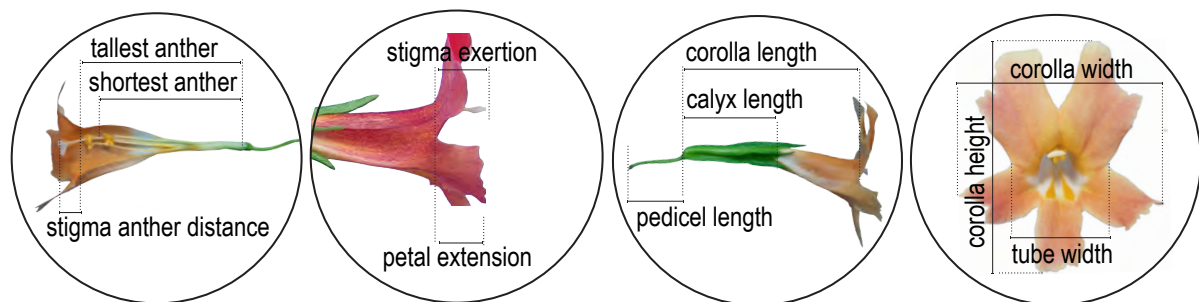


Figure S3. Floral traits measured for the QTL analysis. In addition to the size-related traits, we also measured anthocyanin and carotenoid pigment levels. Anthocyanins were extracted in 1% acidic methanol from a single disc collected from one of the top petals the first day each flower opened. Absorbance of extracts was measured with a spectrophotometer at 520 nm, as described previously (Streisfeld and Rausher 2009). Carotenoids were extracted in hexane following a similar protocol, and absorbance was measured at 450 nm.

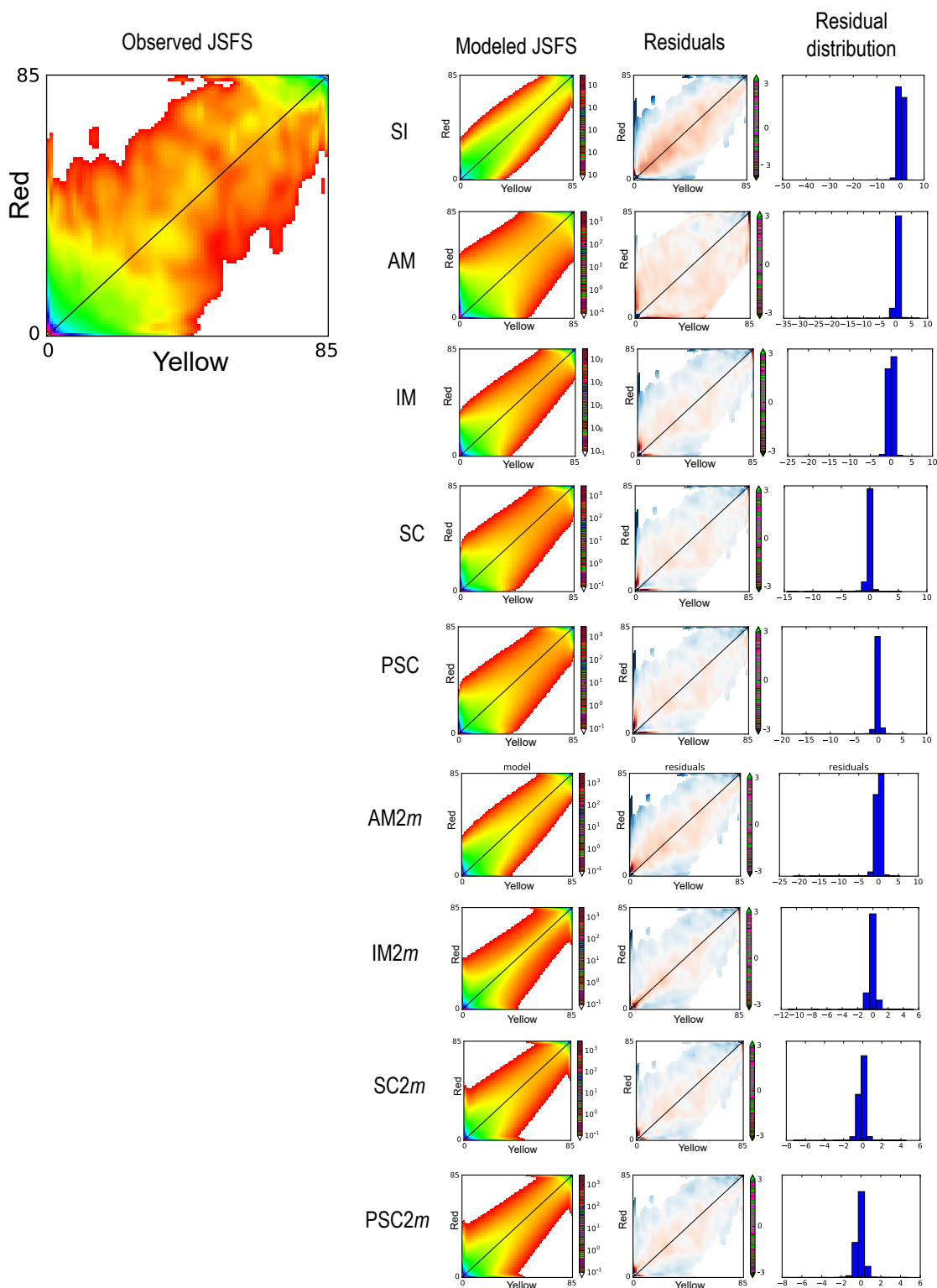


Figure S4. Fits of the different demographic models to the observed data. The observed unfolded joint site-frequency spectrum (JSFS) calculated from the RADseq data set is shown on the left. The first column shows the modeled JSFS for the best fit of each model (See Fig. S1 for a cartoon of each model). The second column shows the residuals of the best model fit to the observed JSFS. The third column shows the residuals plotted as a histogram.

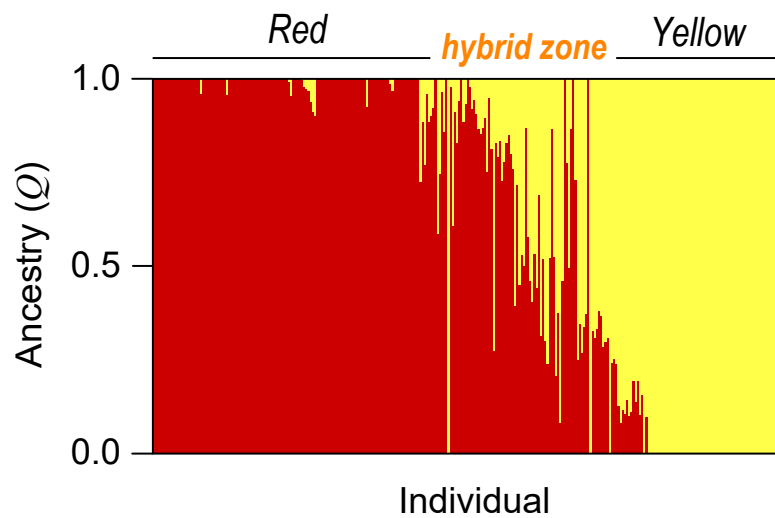


Figure S5. The results of an *Admixture* analysis conducted on the full SNP dataset. Each bar represents an individual and shows its probability of membership (Q score) to the two different clusters. Individuals are grouped by whether they come from the distribution of the red ecotype, yellow ecotype, or from within the hybrid zone.

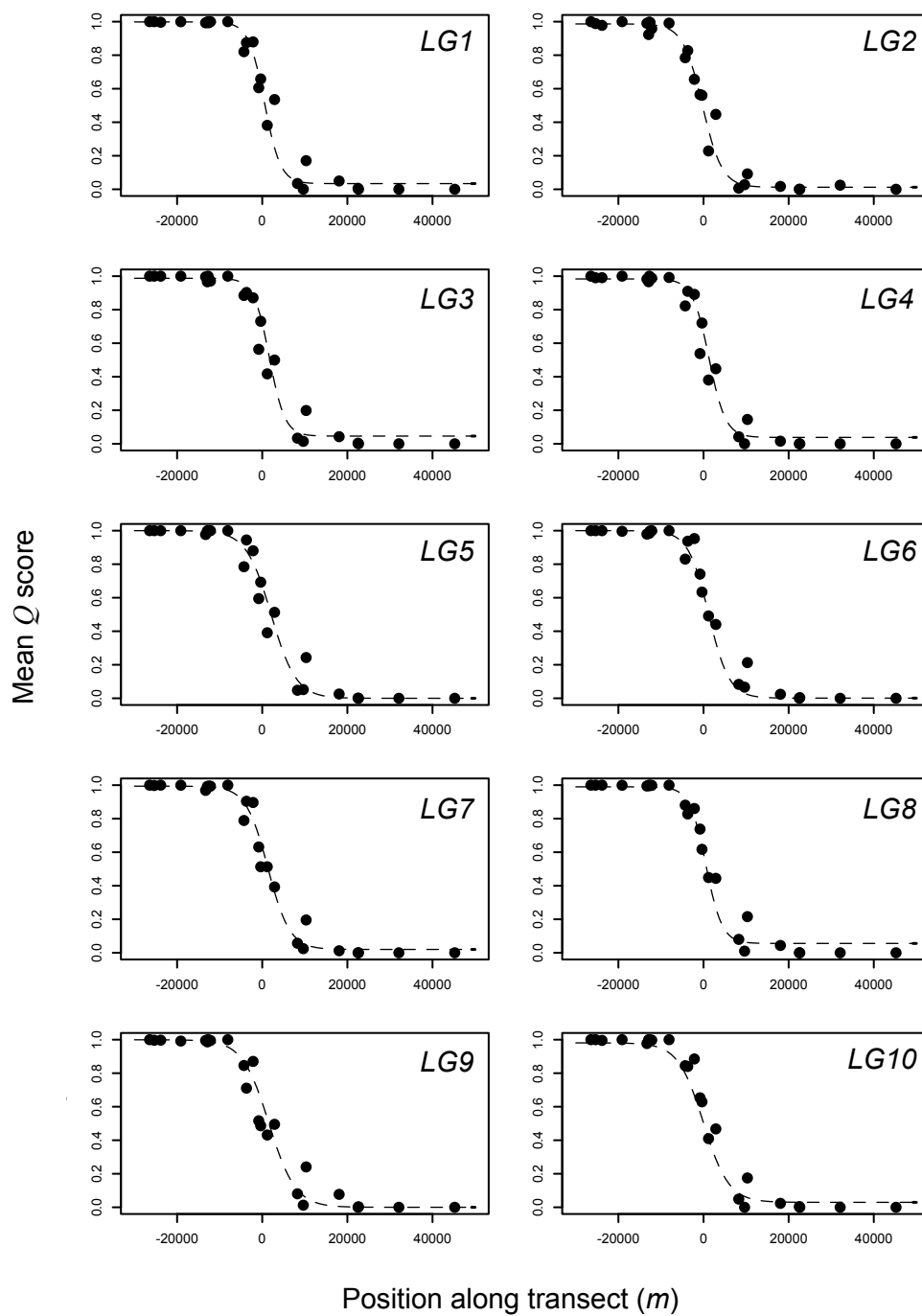


Figure S6. Cline fits for each chromosome. Black points are the mean ancestry scores plotted along the one dimensional transect. The dashed line is the ML sigmoid model.

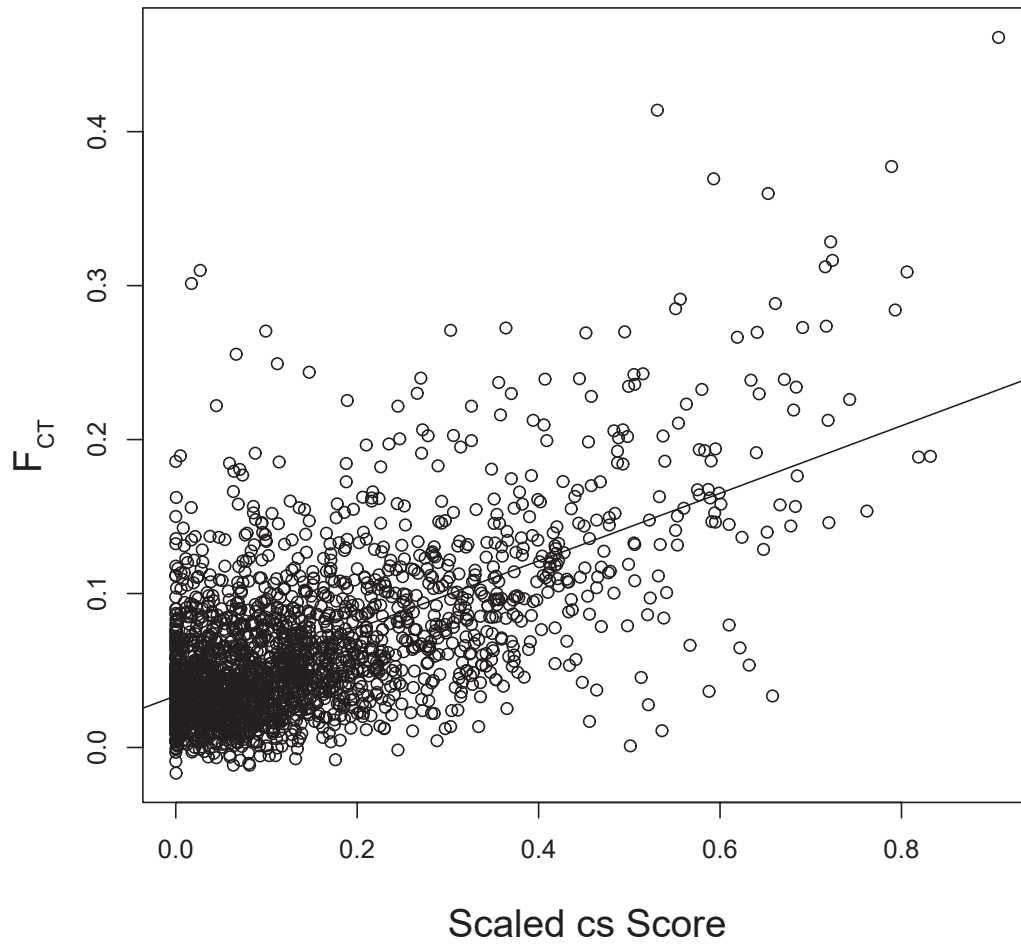


Figure S7. Relationship between the cline similarity score (cs) and F_{CT} . The line through the plot is the least squares regression ($r^2 = 0.38$).

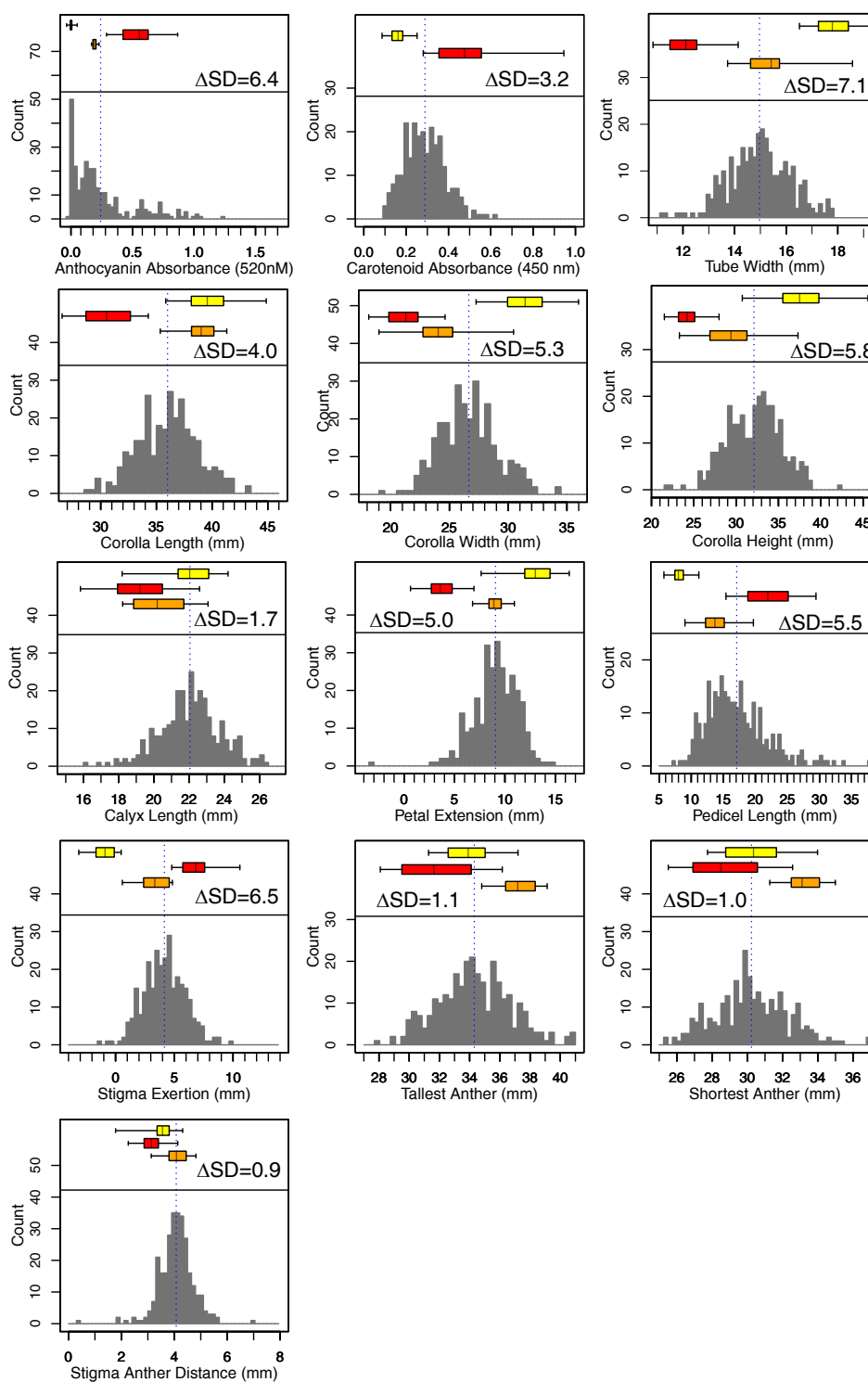


Figure S8. Trait variation in greenhouse-raised red ecotype, yellow ecotype, F₁, and F₂ individuals. The histogram in each plot shows the distribution for each trait in the F₂ population. The top, middle and lower box plots show the distributions for the yellow ecotype, red ecotype and F₁, respectively. Δ SD indicates the number of standard deviation that the red and yellow ecotypes differ by.

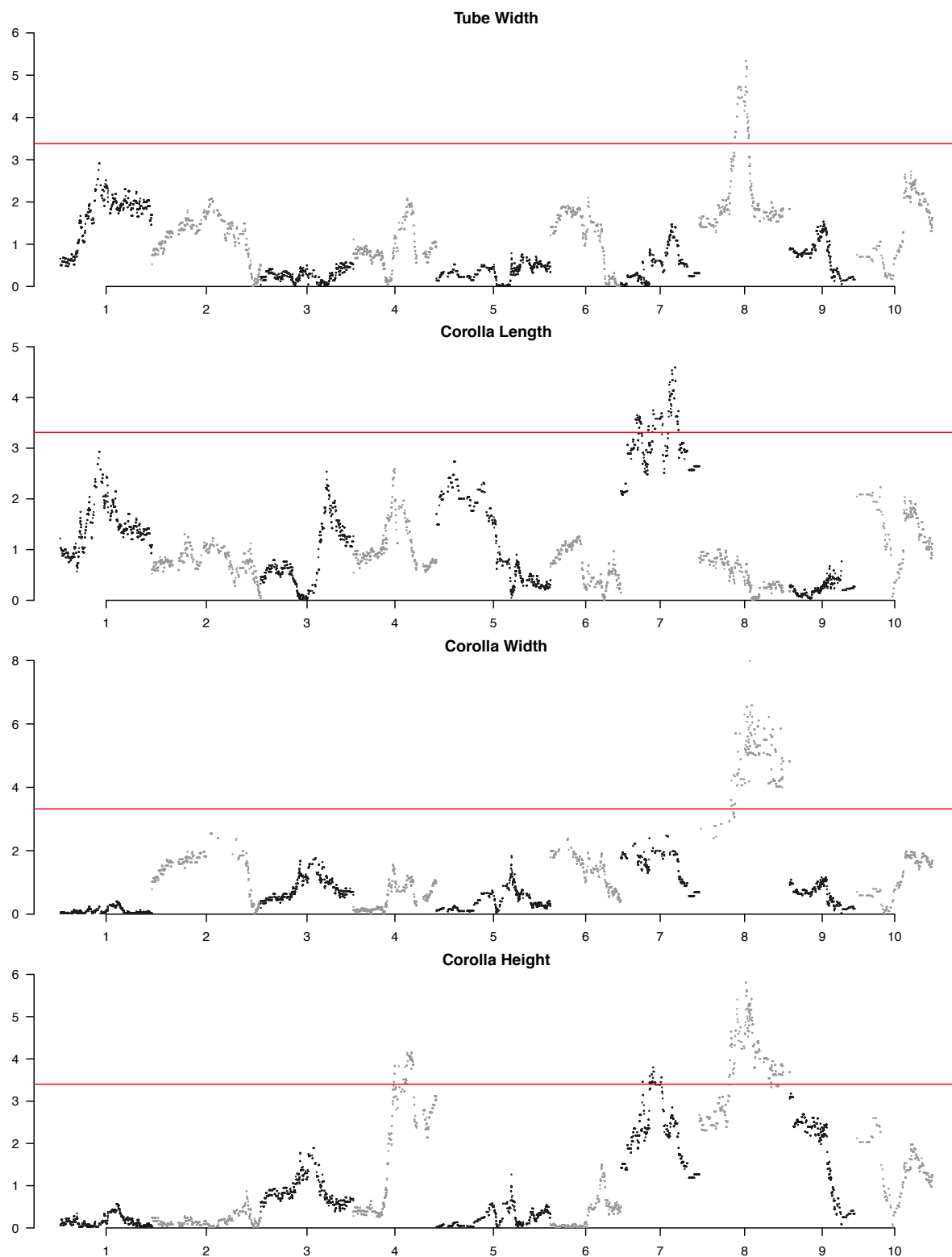


Figure S9. LOD scores plotted across the genome for each trait in the QTL analysis. The horizontal red line shows the LOD score cutoff for the detection of a significant QTL. Note the different y-axis scales among traits.

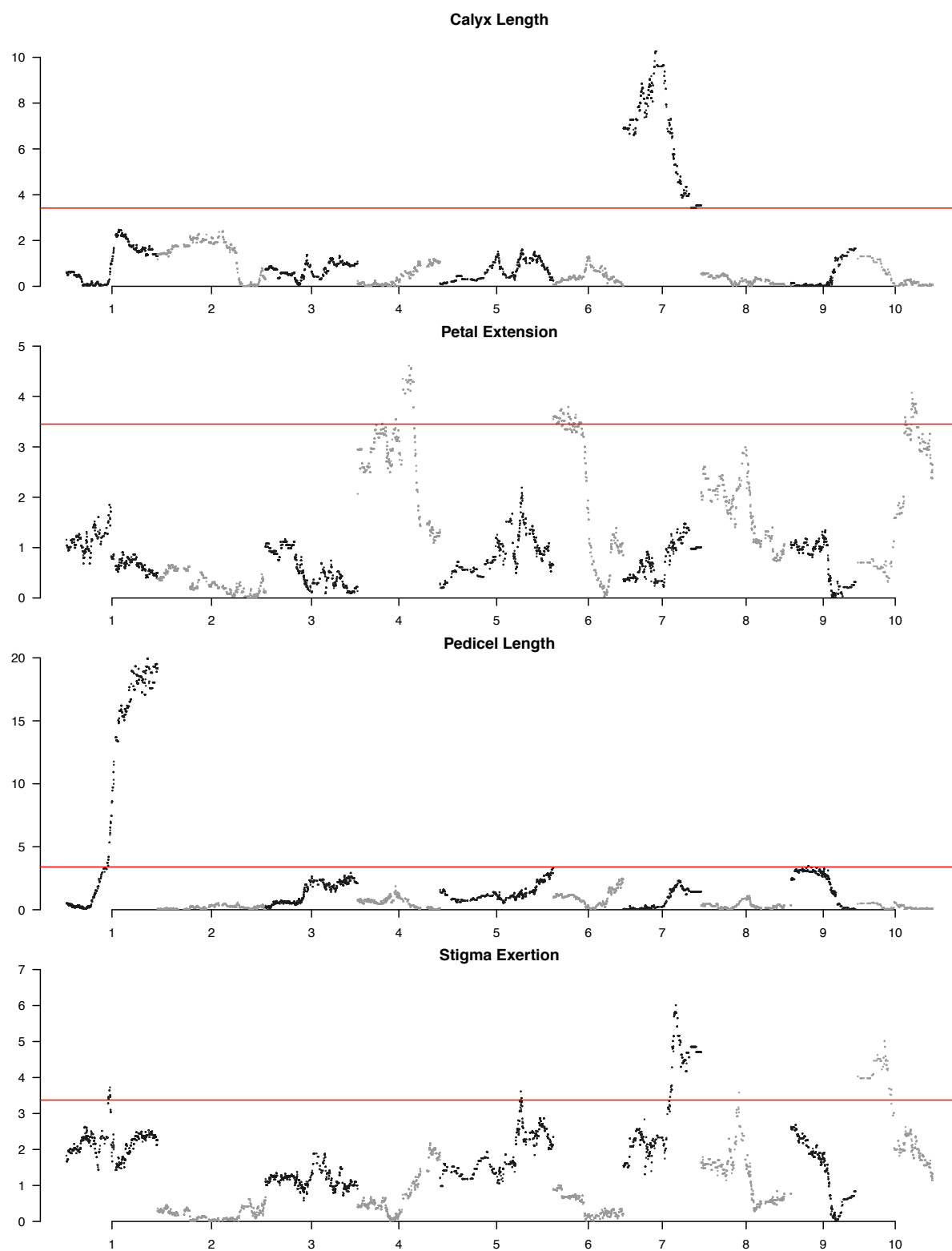


Figure S9 continued.

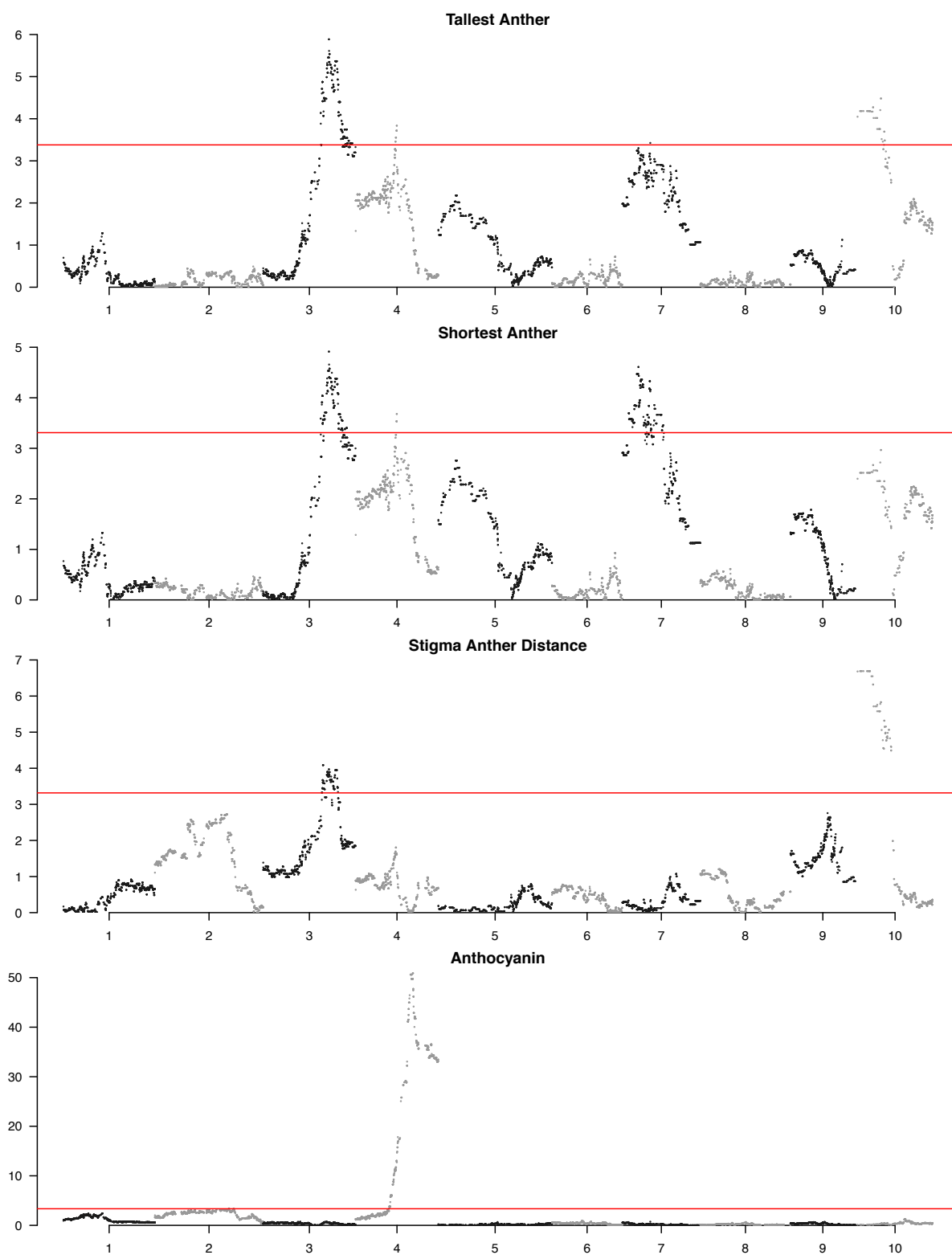


Figure S9 continued.

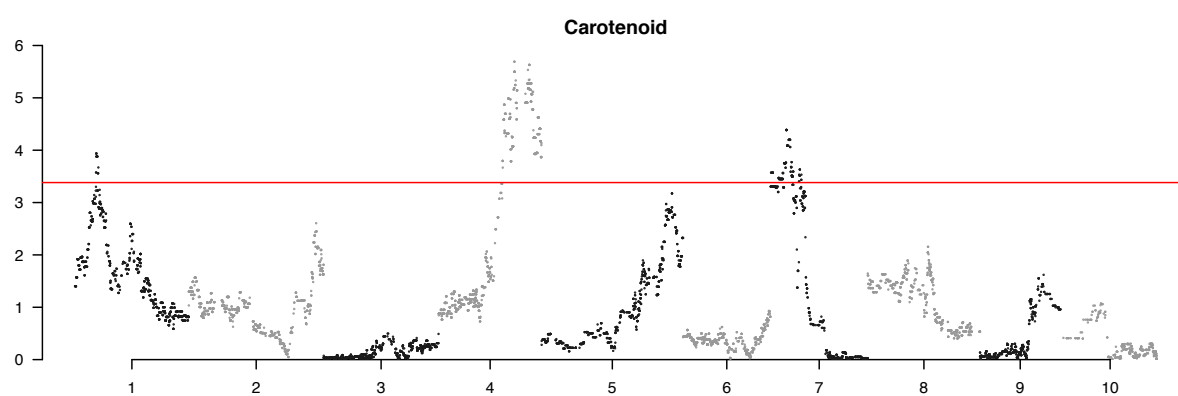


Figure S9 continued.

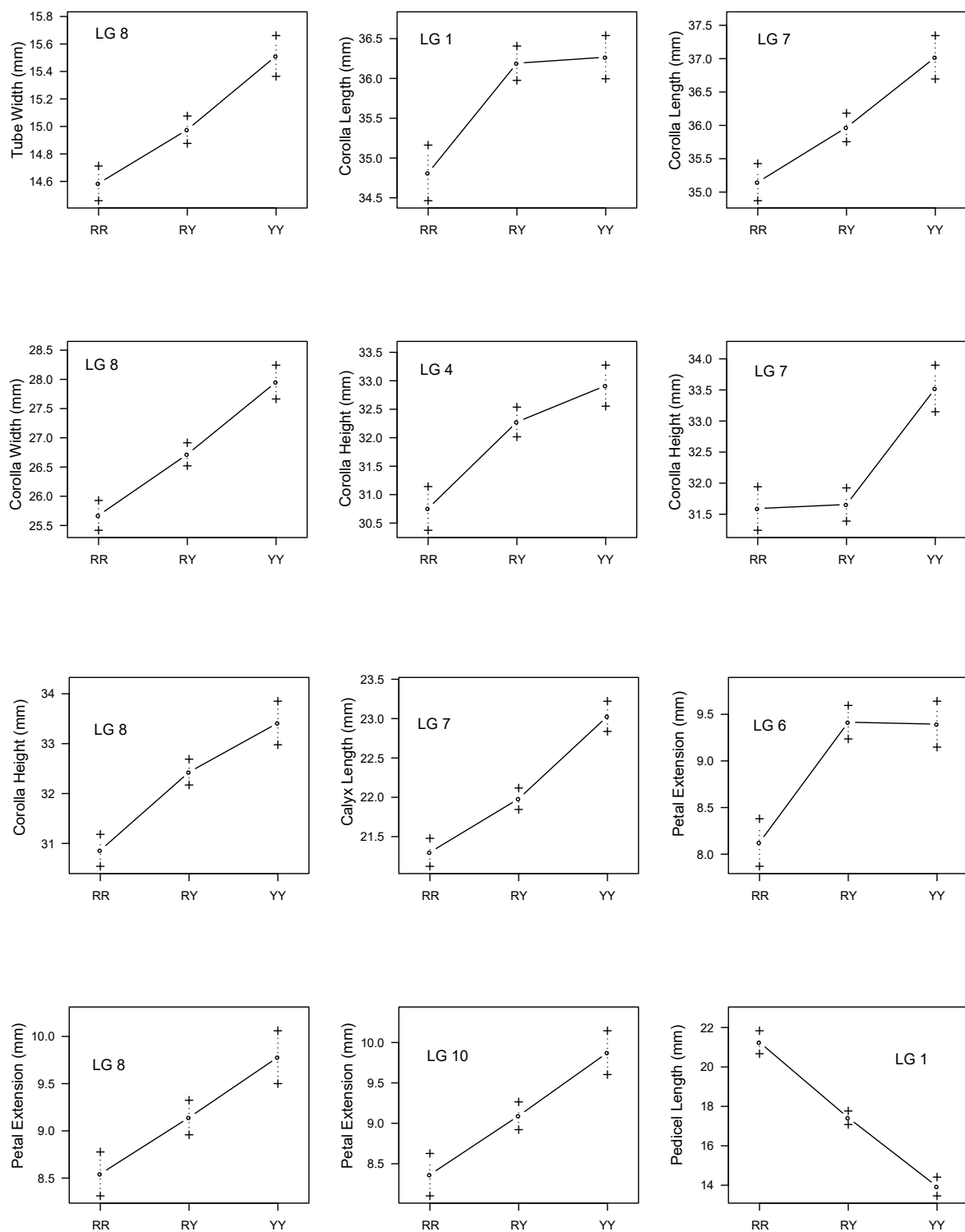


Figure S10. QTL effect plots for each of the 26 QTL discovered in the analysis. Each plot shows the mean (circle) and standard deviation (crosses) for each trait plotted against the three genotypes. RR, both alleles inherited from a red grandparent; YY, both alleles inherited from a yellow grandparent; RY, Heterozygous. Red lines indicate that the effect occurs in the opposite direction of that inferred by the trait differences measured in the red and yellow grandparents (see Fig. S8).

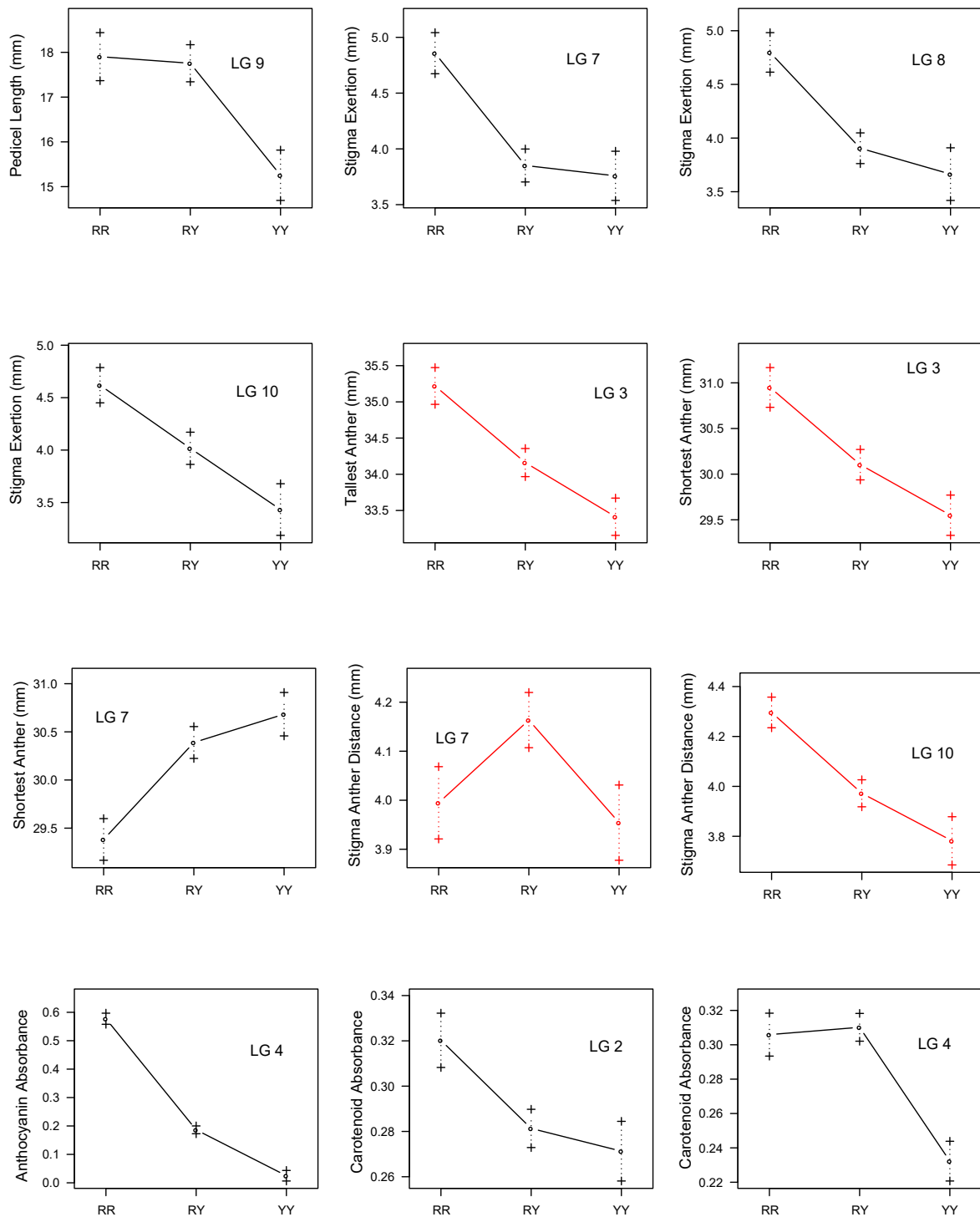


Figure S10 continued.

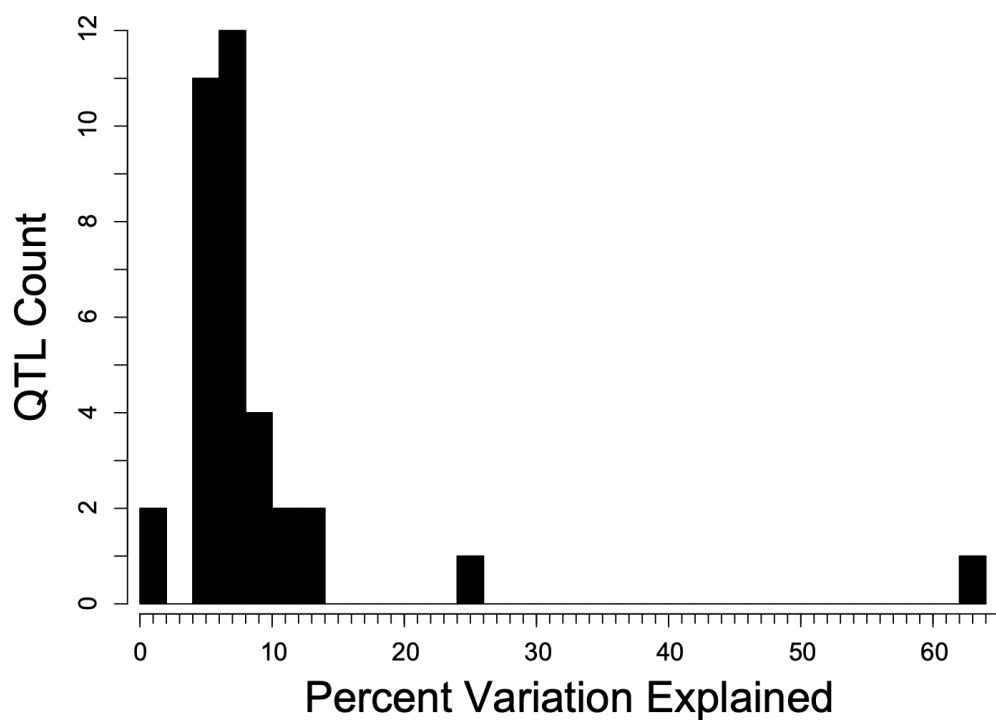


Figure 11. Distribution of effect sizes (percent variation explained among F₂ plants) for the 26 QTL detected in the study.

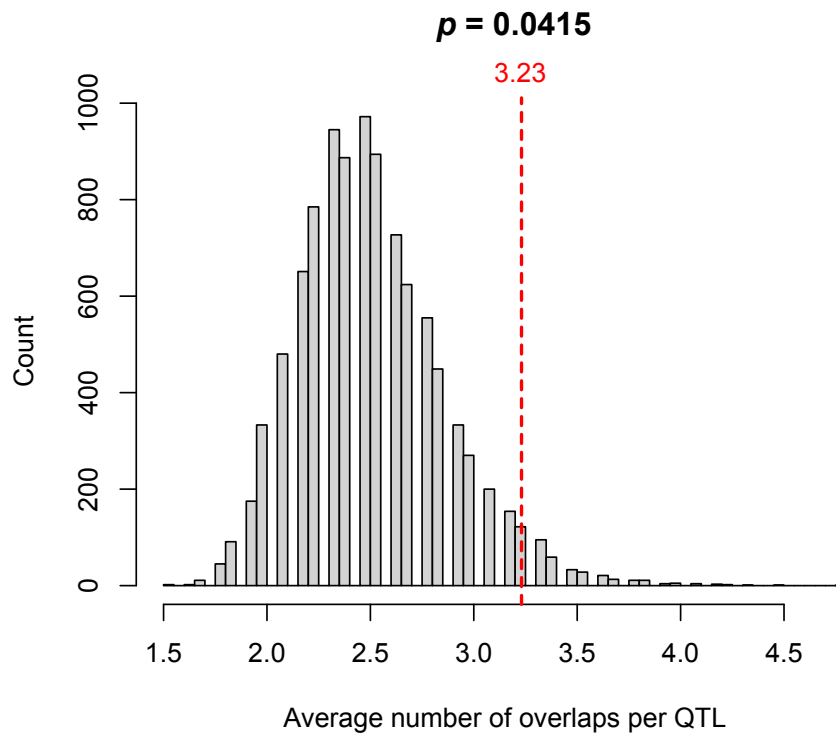


Figure S12. Results from the permutation test for the significant colocalization of QTL. The dashed red line shows the average number of observed overlaps per QTL, while the histogram shows the observed null distribution, generated from 9,999 random permutations. The p -value for the test is shown at the top of the plot.

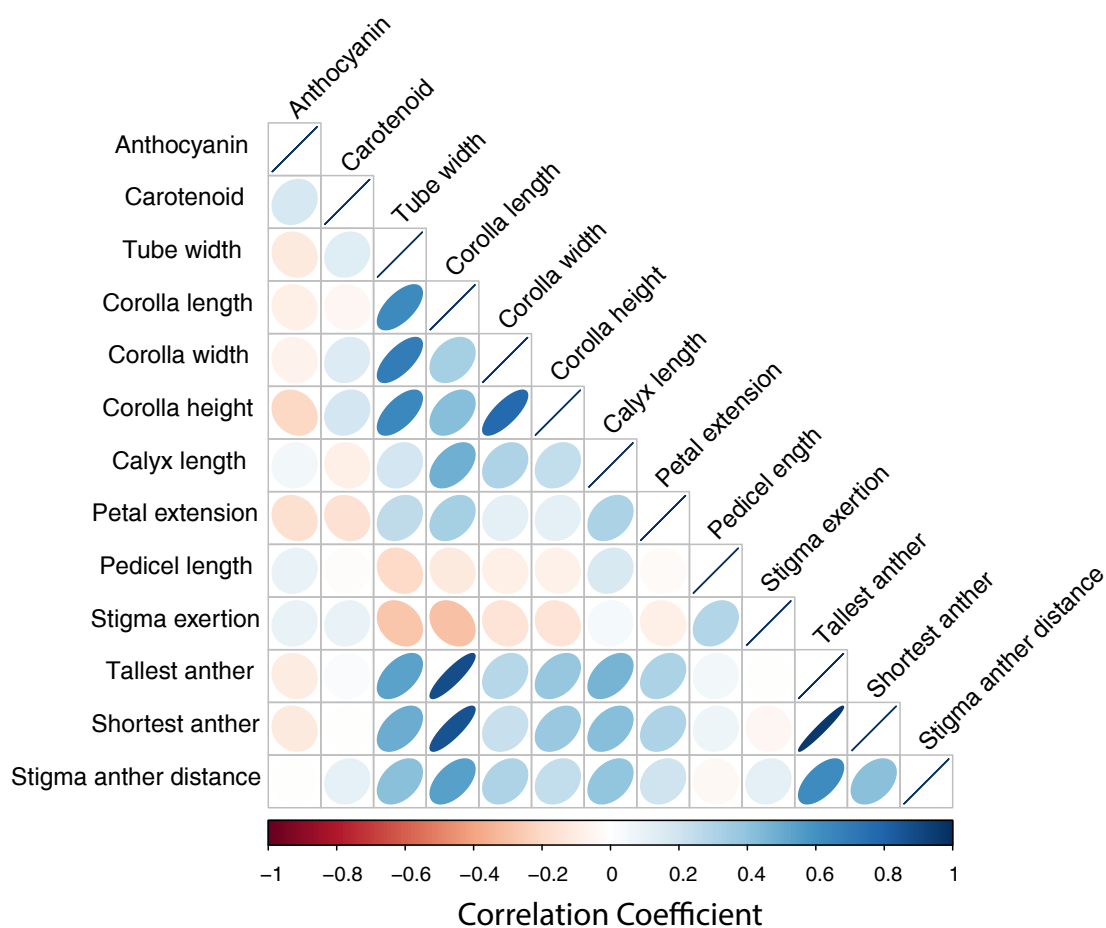


Figure S13. Correlation matrix showing the Pearson correlation between each pair of traits in the F₂ population. The color and shape of each ellipse indicates the strength and direction of each correlation.

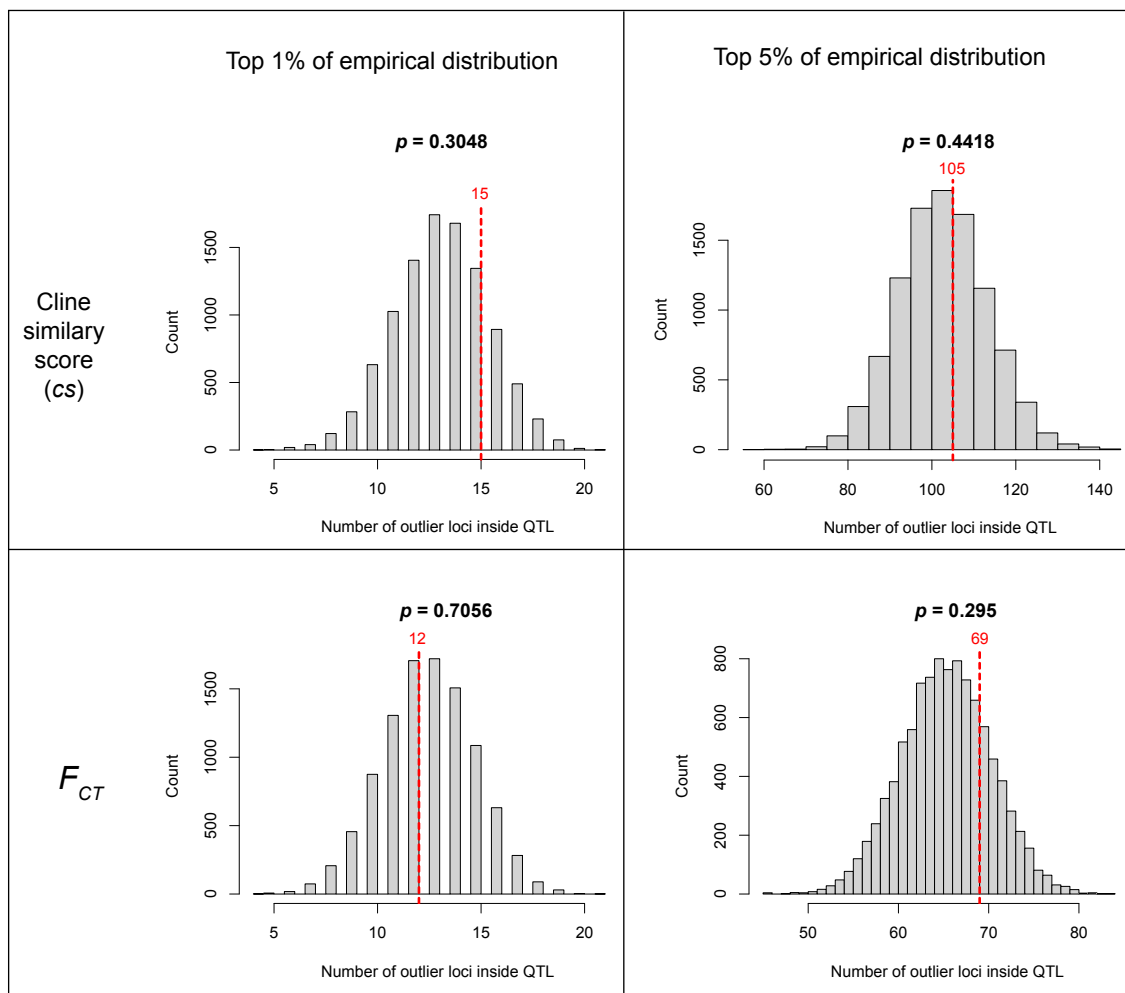


Figure S14. Results from the permutation tests for significant enrichment of candidate barrier loci within the QTL Bayes credible intervals. The test is performed with two different arbitrary cutoffs for defining candidate barrier loci (top 1% and top 5% of the distribution) for two different statistics (the cline similarity score and F_{CT}). The red line shows the observed number of candidate barrier loci falling within the QTL intervals, while the histogram shows the observed null distribution, generated from 9,999 random permutations. The p -value for the test is shown at the top of the plot.

Supplement 1:

Calculation of ‘cline similarity’ score

Rationale

This supplement outlines the rationale behind the calculation of the cline similarity score, which describes the relative shape and position of a cline, relative to some other cline of interest. In our case, we wanted to be able to view the variation in multiple parameters that describe different features of a geographic cline in an integrated way, so we could do things, like plot clinal variation along a genome.

Consider a simple, sigmoid cline, like the one shown in Figure A. This cline model, which is the most commonly used in empirical studies, is described by 4 fitted parameters: the cline centre (c), cline width (w), the mean trait value on the ‘high’ side of the cline (Q_{\max}), and the mean trait value on the ‘low’ side of the cline (Q_{\min}). The last 2 parameters can be reduced to a single parameter, $\Delta Q = Q_{\max} - Q_{\min}$, which quantifies the total change in a trait across the transect. It is worth noting that we use the notation ΔQ to denote change in ancestry scores across the cline, rather than Δz , which is typically used for a quantitative trait (or Δp for allele frequency).

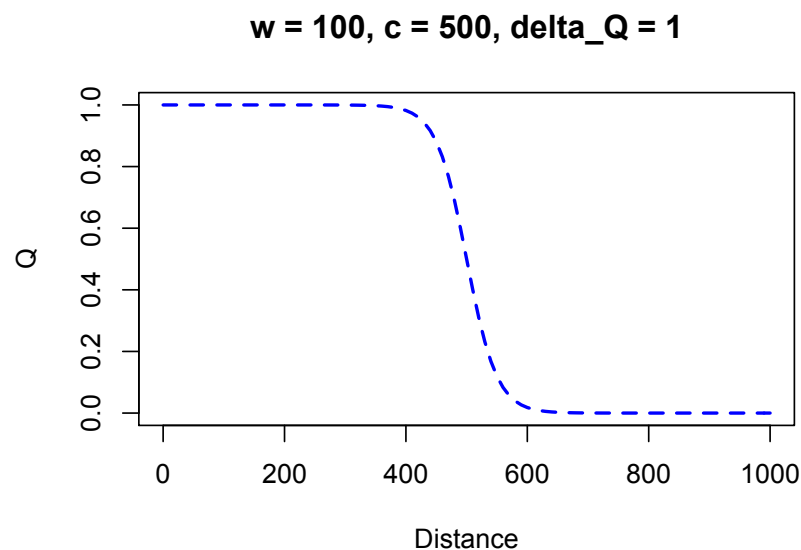


Fig. A. Simple sigmoid cline model described by 3 parameters, ΔQ , w , and c .

Each parameter tells us something about the feature of the cline and is relevant to inferences that we might make about selection in a hybrid zone (provided that we are happy to make some assumptions). For example, the width of the cline is inversely proportional to the strength of selection acting on the trait; the centre tells us about the spatial pattern of selection across the hybrid zone; and the change in the trait value across the cline tells us about the strength of the difference in the trait across the cline. In a situation where selected and neutral loci are at equilibrium, the above cline would indicate that variation in the trait is maintained by selection.

The problem is that inferences cannot easily be made from a single cline parameter. This is highlighted in Figure B, where several clines are shown, with certain parameters fixed and others varying. The left plot shows three clines with the same values of w and c , but substantially different values of Δz ; The middle plot shows multiple clines with the same values of Δz and c , but with different values of w ; the third shows clines with identical shapes (i.e., same Δz and w), but different values of c .

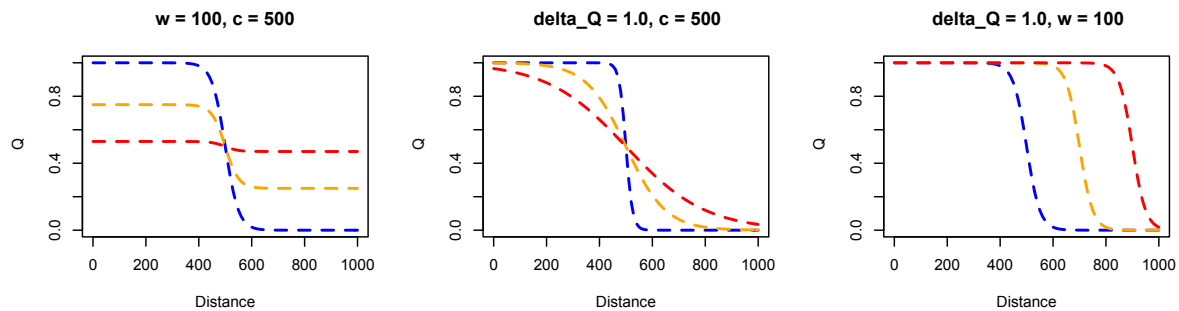


Fig. B. Examples of clines where a single parameter varies but the other two parameters are fixed.

In the above plots, we can easily evaluate each cline in terms of its fit to some hypothesis by visually inspecting them. However, when we have many clines, this is impractical and highly subjective. But these plots clearly show that we cannot infer much by looking at the distribution of a single parameter, especially in real data where all parameters will vary simultaneously.

The ‘cline similarity’ score

To help simplify the process of identifying clines of interest, we created an *ad-hoc* statistic that we call the ‘cline similarity’ or ‘ cs ’ score. The cs score summarises cline shape and relative position with a single number that can be calculated from the ML cline parameters. cs is calculated as:

$$\text{Cline similarity, } cs = \left(\frac{\Delta Q}{w+l} \right) e^{-(l/c/l)^2} \quad (\text{eq. 1}).$$

Where ΔQ , w and c , are defined as above and l is a scaling factor that will be explained further below. Eq. 1 can be broken down into two terms that determine the cs score in different ways. The first term is a shape score, that describes the shape of a cline, independent of the location of the cline centre:

$$\text{Shape score} = \left(\frac{\Delta Q}{w+l} \right) \quad (\text{eq. 2}).$$

The second term is a centre penalisation score, which downgrades the shape score based on the value of the fitted cline centre relative to some point of reference (explained further below).

$$\text{Centre penalisation score} = e^{-(l/l)^2} \quad (\text{eq. 3.}]$$

These formulae and the decisions made to arrive at them will next be explained by walking through its calculation for the *Mimulus* dataset.

Application to the *Mimulus* dataset

For the *Mimulus* dataset, we fit clines to ancestry scores in 2,173 non-overlapping windows, each containing 100 SNPs. The details of how the fits were done are described in the main text, but the clines can be fitted in any way. We only need the parameters for the fitted model to calculate cs .

Starting with the shape score, we wanted a quantity that described variation in the shape of the cline. Ultimately, we wanted clines with a small w and large ΔQ to have high shape scores, and clines with very low ΔQ and high w to have a low shape scores. Figure C shows the joint distribution of ΔQ and w for all of the windows. The numbers in the plot correspond to the cline fits for 10 selected windows, which are shown on the right side of the plot. The orange star (cline 2), is not for a specific window, but rather is the fit to mean Q score for each location, averaging across all windows (i.e., it is a cline fitted to the mean of means for each location). The green star shows the location of the genome-wide cline in this space (i.e., the one shown in figure 3a of the main text, and figure D (ii)).

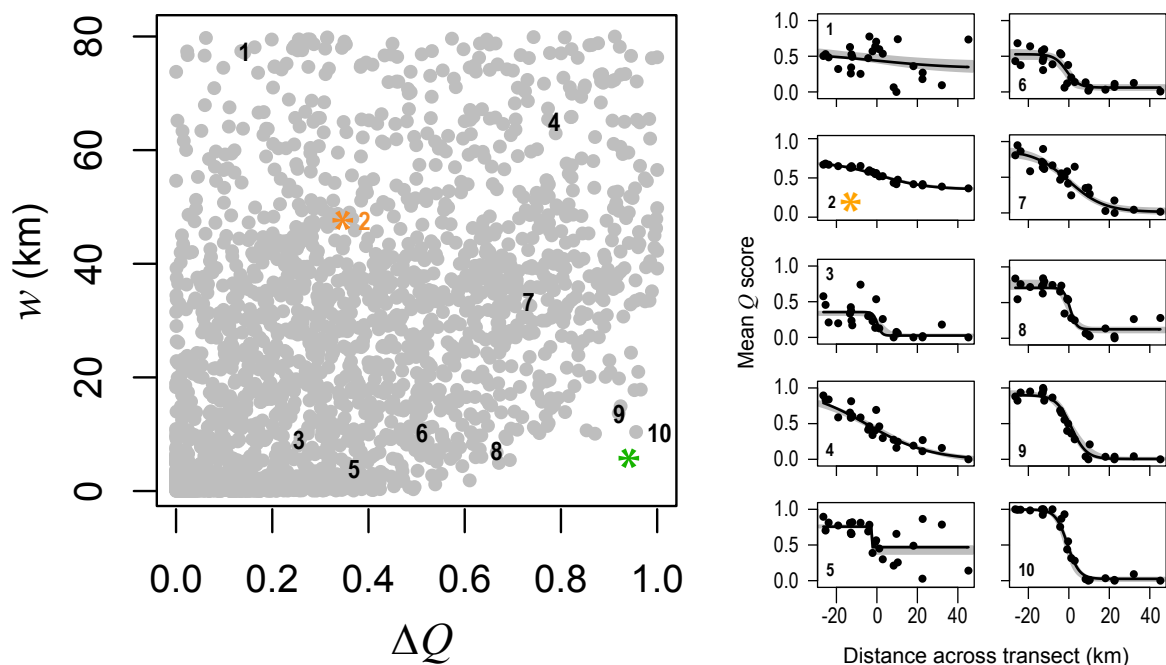


Fig. C. Joint distribution of w and Δz for the *Mimulus* dataset. The clines on the right show fits that correspond to the numbers in the bivariate plot.

The bottom right corner of Figure C contains the clines that are most interesting in terms of our hypothesis regarding barrier loci. The ones in the top left and bottom left are of little

interest, because ΔQ is always very small. The ones in the top right are more interesting, because they show high ΔQ but are very wide; these clines could be generated by processes like isolation by distance, or the collapse of secondary clines when barriers to gene flow are porous.

Figure D(i) shows the same plot, but with the points coloured by the shape score (10 coloured bins, but the values are continuous), as given in Eq. 2. Here we define l , the scaling parameter, as $0.5t$, where t is the length of the transect. l determines precisely how variation in the shape scores is spread out across the joint distribution of ΔQ and w . In our case, $0.5t$, gives a distribution shape scores that suits our purposes, but other values might be more suitable for other study systems.

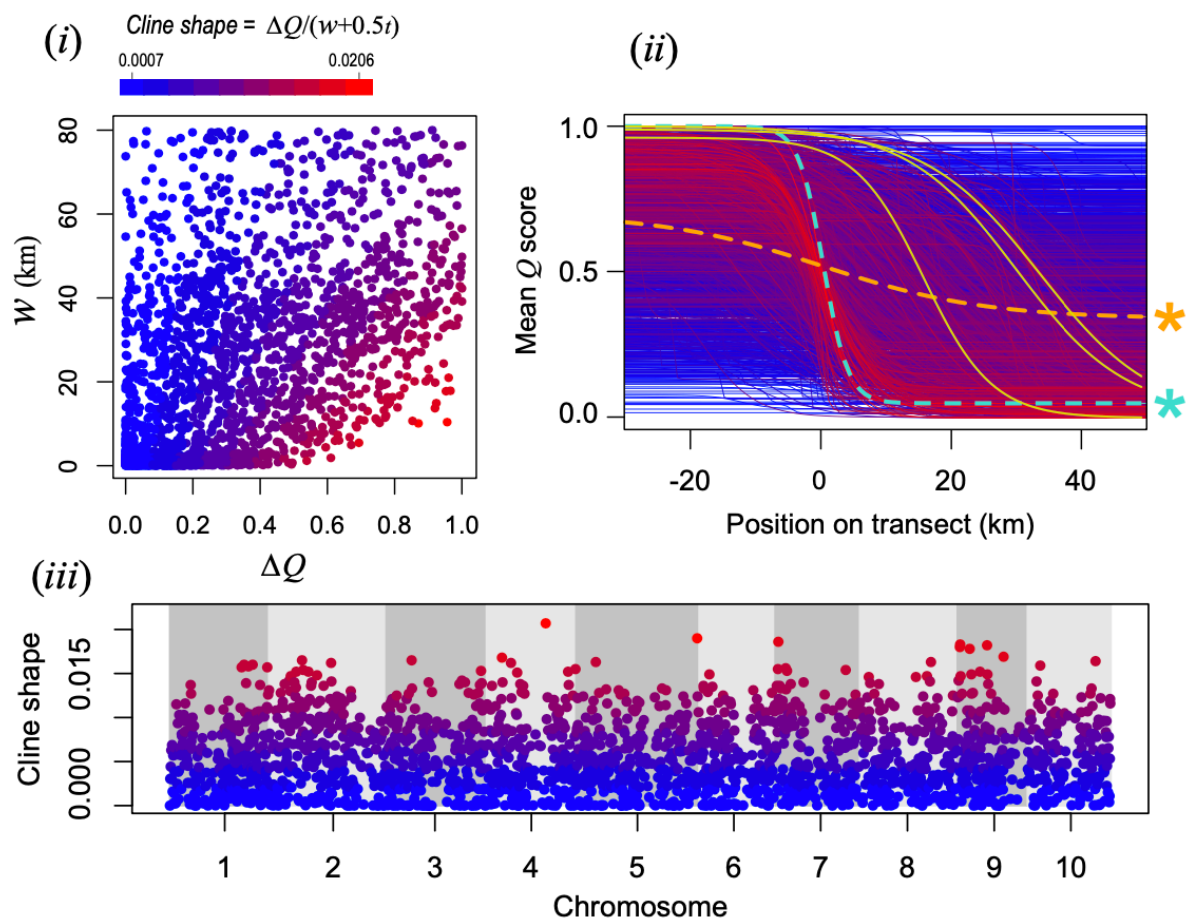


Fig. D. Shape scores for the 2,173 windows. (i) Joint distribution of w and ΔQ , coloured by the shape score. (ii) Cline models coloured by the shape score. (iii) The shape scores plotted across the genome. See text for more details.

By colouring the cline models for each window by their shape scores, as shown in Figure D(ii), we can see that clines with ‘hotter’ colours do have shapes more similar to the genome wide cline (the dashed cyan line) than ‘cooler’ colours. However, it is also clear that some clines with high shape scores have centres that are displaced a very long way from the genome wide cline. Some of these have been coloured solid yellow to help highlight them.

Although these clines are interesting for a variety of reasons, they are not spatially associated with the centre of the hybrid zone that also coincides with the transition in floral trait differences. Ideally, we would like our summary statistic to also reflect where the cline is positioned in relation to this location. To do this, we penalise the shape score based on the position of the cline centre, relative to a position of interest, which needs to be given position 0. This could be a feature of the environment or a cline in a focal marker or trait. In our case, the position of interest is the centre of the genome-wide ancestry cline.

Figure E(i) shows how the value of the inferred centre informs the centre penalisation function outlined in Eq. 3. If the fitted cline centre coincides exactly with the point of interest ($c = 0$), the cs score for that cline is simply the shape score (i.e., $cs = \text{shape score}$). However, as the difference between fitted centre and point of interests increases (i.e., $|0 - c|$), then the shape score is downgraded according to the centre penalization function, resulting in a cs score lower than the shape score.

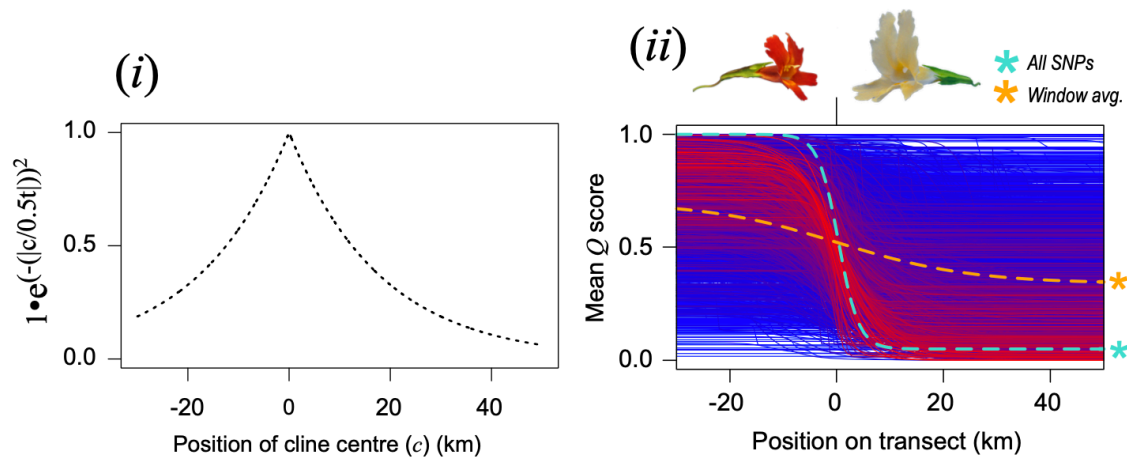


Figure E. Penalisation of the shape score to give the final cs score. (i) The effect of the cline centre penalisation function on the shape score. Here, an arbitrary shape score of 1 is used for illustrative purposes. (ii) The cline models for all windows after the cline penalisation score, coloured by the overall cs score.

The results of this final transformation step can be seen in figure 3c of the main text, which is reproduced in Figure E(ii) for convenience. In that figure, the cs scores are scaled between 0 and 1, where 0 is the cline with the lowest cs score, and 1 is the cs score for the genome-wide cline.

Some afterthoughts (if others attempt something similar)

This was done 5 years ago and if I was doing this again from scratch, I would probably think about it a bit differently. First, I think it might be useful to try and incorporate the variance explained by the cline model into the cs score. i.e., if the data around the cline are extremely noisy (common for clines with a small ΔQ , e.g., cline 1 in Figure C), we might conclude that these are less likely to be of interest. The variance explained was used primarily in Westram et al. 2017 and seemed to be good at identifying non-neutral clines in simulations.