doi:10.1038/nature09641

Identification of two genes causing reinforcement in the Texas wildflower *Phlox drummondii*

Robin Hopkins[1] & Mark D. Rausher[1]

Species formation generates biological diversity and occurs when traits evolve that prevent gene flow between populations. Discerning the number and distribution of genes underlying these traits and, in a few cases, identifying the genes involved, has greatly enhanced our understanding over the past 15 years of species formation (reviewed by Noor and Feder[1] and Wolf *et al.*[2]). However, this work has almost exclusively focused on traits that restrict gene flow between populations that have evolved as a by-product of genetic divergence between geographically isolated populations. By contrast, little is known about the characteristics of genes associated with reinforcement, the process by which natural selection directly favours restricted gene flow during the formation of species. Here we identify changes in two genes that appear to cause a flower colour change in *Phlox drummondii*, which previous work has shown contributes to reinforcement. Both changes involve *cis*-regulatory mutations to genes in the anthocyanin biosynthetic pathway (ABP). Because one change is recessive whereas the other is dominant, hybrid offspring produce an intermediate flower colour that is visited less by pollinators, and is presumably maladaptive. Thus genetic change selected to increase prezygotic isolation also appears to result in increased postzygotic isolation.

Natural selection can directly favour species formation through a process termed reinforcement. If two incipient species experience secondary contact and produce maladaptive hybrids, selection favours decreased gene flow and increased reproductive isolation between them[3–5]. Reinforcement can be recognized by a resulting pattern known as reproductive character displacement: reproductive isolation is greater in sympatry than in allopatry[4]. Although the occurrence of reinforcement was historically controversial[6,7], empirical studies have documented reinforcing selection and reproductive character displacement in birds, insects, amphibians, plants and mammals, suggesting that it is a common step in the process of species formation (reviewed by Ortiz-Barrientos *et al.*[8] and Pfennig and Pfennig[9]). Despite this work, little is known about genetic changes associated with reinforcement (but see studies of quantitative trait loci by Ortiz-Barrientos *et al.*[10] and Saether *et al.*[11]).

Divergence of floral colour in *P. drummondii* constitutes one of the best-documented cases of reinforcement in plants[12] and exhibits the classic pattern of reproductive character displacement. *P. drummondii* and the closely related *P. cuspidata* produce similar light-blue flowers throughout the allopatric parts of their ranges. However, in the area of sympatry, *P. drummondii* has dark-red flowers, representing the only natural evolution of red flowers in the *Phlox* clade[12]. Both species of *Phlox* and colours of *P. drummondii* are pollinated by the same array of species of Lepidoptera[12]. Hybrids between these two species are formed at rates as high as 11% in the area of sympatry[12]. The hybrids are vigorous but have high, although not complete, male and female sterility[13–15]. Experimental crosses indicate as many as 40% of hand-pollinated hybrid flowers will mature at least one seed and as many as 72% of crosses sired by hybrid pollen will set one seed[15]. Additionally, allozyme data show low levels of gene flow between these species of *Phlox*[14]. Although other traits may contribute to prezygotic isolation (including possible reinforcement traits such as self-compatibility[12]), Levin demonstrated that the shift from light-blue to dark-red flowers in *P. drummondii* decreases interspecific hybridization by 66%, indicating that the change in flower colour substantially increases prezygotic reproductive isolation[12]. Given the above estimates, the hybridization rate before the evolution of dark-red flower colour could have been as high as 28%, which, with low hybrid fitness, would presumably create strong selection to decrease hybridization.

We determined that the evolutionary transition from light-blue to dark-red flower colour in *P. drummondii* results from changes of large effect at two loci. F_2 populations derived from crosses between the allopatric colour variant (light blue) and the sympatric colour variant (dark red) segregate four discrete flower colours: dark blue, light blue, dark red and light red (Fig. 1a). Quantification of the spectral reflectance of 200 F_2 flowers, transformed into two-dimensional Commission internationale de l'éclairage (CIE) 1976 colour space[16], followed by discriminant analysis verified our discrete classifications (Supplementary Fig. 1 and Supplementary Table 4a, b). The ratios of counts within these categories are very close to the 9:3:3:1 ratios expected from two loci with complete dominance ($\chi^2_{(3,\, N=618)} = 0.92$, $P = 0.8206$) (Supplementary Table 3). One locus, *H*, determines flower hue, with blue allele dominant to red, whereas the second locus, *I*, determines colour intensity, with the dark allele dominant to the light.

These two loci appear to determine the types and amounts of anthocyanin floral pigment produced in *P. drummondii*. Anthocyanin pigments, the final products of the well-characterized and highly conserved ABP, are derived from six common types of anthocyanidin by the addition of sugar and/or methyl moieties[17]. Less-hydroxylated anthocyanidins give rise to redder pigments whereas more-hydroxylated anthocyanidins give rise to bluer pigments[17]. Correspondingly, blue-flowered *P. drummondii* (*H-*) produce anthocyanins derived from both the less-hydroxylated cyanidin and peonidin pigments, as well as the more-hydroxylated malvidin pigment, whereas red flowers (*hh*) produce exclusively the less-hydroxylated pigments (Fig. 1b). The change in floral hue thus results from redirecting flux from the malvidin branch of the anthocyanin pathway to the cyanidin/peonidin branch (Fig. 1c). Individuals with increased colour intensity (*I-*) produce more pigment than *ii* individuals, without an effect on pigment composition (Fig. 1b and Supplementary Table 5a, b). These biochemical patterns, coupled with the structure of the ABP (Fig. 1c), suggest candidate genes for the hue and intensity loci.

The structure of the ABP (Fig. 1c) suggests that the loss of flux down the malvidin branch of the ABP might result from changes in one of three candidate genes: (1) loss of function or reduced expression of the gene coding for the branching enzyme flavanoid 3′5′-hydroxylase (F3′5′H); (2) alteration of the substrate specificity owing to a coding mutation in the gene for dihydroflavonol 4-reductase (DFR) making the enzyme unable to metabolize the malvidin precursor dihydromyricitin; and (3) a similar alteration of the substrate specificity of anthocyanidin synthase (ANS). To examine these possibilities, we analysed multiple F_2 populations for co-segregation of floral hue with single nucleotide polymorphism (SNP) markers in the candidate genes. In a total of 100

[1]Department of Biology, Box 90338, Duke University, Durham, North Carolina 27708, USA.

©2011 Macmillan Publishers Limited. All rights reserved

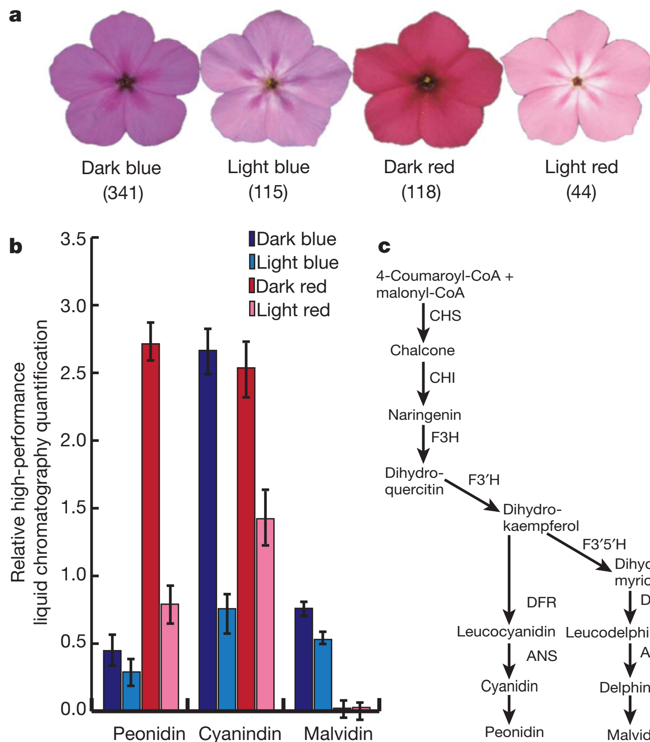

Figure 1 | Flower colour phenotypes in F_2 individuals. a, Representative pictures of the four flower colours in F_2 populations: dark blue, light blue, dark red, light red (from left to right). Total counts in F_2 populations indicated under each flower. Ratios of counts are similar to 9:3:3:1 ($\chi^2_{(3,\ N=618)} = 0.92$, $P = 0.8206$). **b**, Relative anthocyanidin pigment production results in variation in F_2 flower colour. Production of all three pigments is significantly different between hue classes. The light- and dark-red flowers produce no malvidin ($F_{(1,25)} = 134.03$, $P < 0.0001$), and significantly more peonidin ($F_{(1,25)} = 58.88$, $P < 0.0001$) and cyanidin ($F_{(1,25)} = 21.57$, $P = 0.0002$). The amount of pigment production is significantly different between classes of flower intensity, with high-intensity flowers producing more of each flower colour pigment ($F_{\mathrm{malvidin}\ (1,25)} = 8.64$, $P = 0.0092$; $F_{\mathrm{peonidin}\ (1,25)} = 32.11$, $P < 0.0001$; $F_{\mathrm{cyanidin}\ (1,25)} = 66.74$, $P < 0.0001$). Standard errors are shown. See Supplementary Table 5a–c for full data and multivariate analysis of variance (MANOVA) results. **c**, A simplified schematic of the ABP, showing core enzymes to the right of the arrows, with the substrates and products indicated at the ends of the arrows. The two branches of the pathway active in *P. drummundii* produce two 3′-hydroxylated red pigments (cyanidin and peonidin) and one 5′-hydroxylated blue pigment (malvidin).

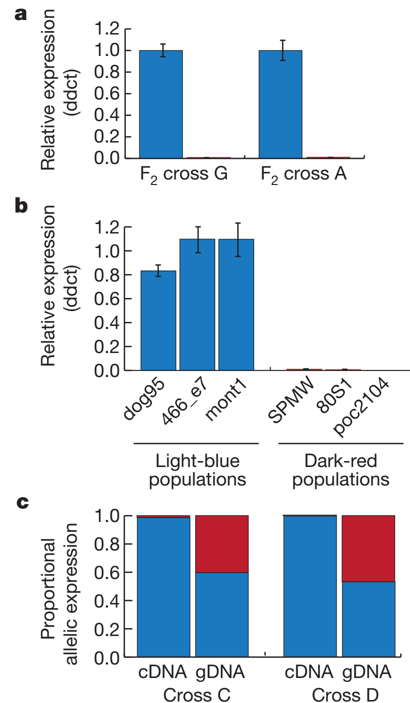

Figure 2 | Results of expression experiments on the hue locus (*F3′5′h*). Quantitative PCR results showing relative expression of *F3′5′h* in blue and red individuals in two F_2 populations (**a**) and six field populations (**b**). Transcripts of *F3′5′h* in red floral tissue (shown in red) are significantly lower than levels detected in blue tissue (shown in blue) in both F_2 populations ($F_{(1,26)} = 250.89$, $P < 0.0001$) and field-collected individuals ($F_{(1,36)} = 30.21$, $P = 0.0053$). Delta–delta cycle-threshold (ddct) indicates relative transcript levels, and bars indicate two standard error units. **c**, Allele-specific expression of *F3′5′h* in multiple heterozygous individuals (*Hh*) from two crosses. Each bar represents the relative contribution of the red allele (shown in red) and the blue allele (shown in blue) to the total expression detected in heterozygous individuals from F_2 families C and D. Relative allelic representation in complementary DNA (cDNA) from heterozygous individuals is significantly different from the relative allelic representation in genomic DNA (gDNA) ($F_{(1,32)} = 375.93$, $P < 0.0001$), indicating a *cis*-regulatory change controlling expression of the red allele. The genomic samples from heterozygous individuals show nearly equal allelic representation (0.5). See Supplementary Fig. 3 for experimental control data.

individuals there was perfect co-segregation between *F3′5′h* and floral hue (Supplementary Table 7). Moreover, genotype at *F3′5′h* explains 77% of the variation in flower hue.

This genetic association corresponds to a downregulation of *F3′5′h* in red flowers. Among F_2 individuals, there is a nearly 100-fold decrease in *F3′5′h* transcripts in red-flowered compared with blue-flowered individuals (Fig. 2a). In addition, red- and blue-flowered individuals collected from multiple populations throughout the range of *P. drummondii* have a comparable difference in *F3′5′h* expression (Fig. 2b) These results demonstrate three features of the genetic change associated with the shift to red flowers: (1) variation in hue is associated with transcript level of *F3′5′h*; (2) genotype at *F3′5′h* predicts transcript level; and (3) the expression difference evident between naturally occurring flower colour variants segregates as a single locus in F_2 individuals.

These patterns imply that expression variation at *F3′5′h* is caused by variation in a *cis*-regulatory element. Allele-specific expression assays confirm this inference: in heterozygous (*Hh*) individuals, the red allele is almost completely downregulated (Fig. 2c). This allelic imbalance indicates a *cis*-regulatory change, whereas equal expression of the two alleles would indicate *trans*-acting regulation of expression[18].

Based on the structure of the ABP, we identified three candidate genetic changes that could explain the increased pigment production in dark (*I-*) flowers: (1) a *cis*-regulatory change that increases production of an ABP enzyme with control over pathway flux; (2) increased enzymatic efficiency (through coding-sequence mutations) of a rate-controlling enzyme in the ABP; and (3) increased expression or function of an ABP transcription factor coordinately regulating the expression of several enzymes. To evaluate these possibilities, we cloned and sequenced genes coding for the core enzymes of the pathway (Fig. 1c), as well as an R2R3-Myb transcription factor orthologous to those known to regulate ABP enzymes in other angiosperms. Of these genes, the marker in *R2R3-Myb* co-segregates perfectly with flower colour intensity in 100 F_2 individuals (Supplementary Table 8) and genotype explains 71% of the intensity variation. These observations suggest *Myb* corresponds to the intensity locus.

This correspondence is further supported by quantification of *Myb* transcript in F_2 individuals. *Myb* expression is significantly higher in dark individuals than light individuals (Fig. 3a), demonstrating that both genotype and expression of this transcription factor are associated with variation in pigment production. Furthermore, field-collected individuals show the same association between expression of *Myb* and flower colour intensity (Fig. 3b). Finally, analysis of allele-specific expression in heterozygous (*Ii*) individuals show that the 'dark' allele is

©2011 Macmillan Publishers Limited. All rights reserved

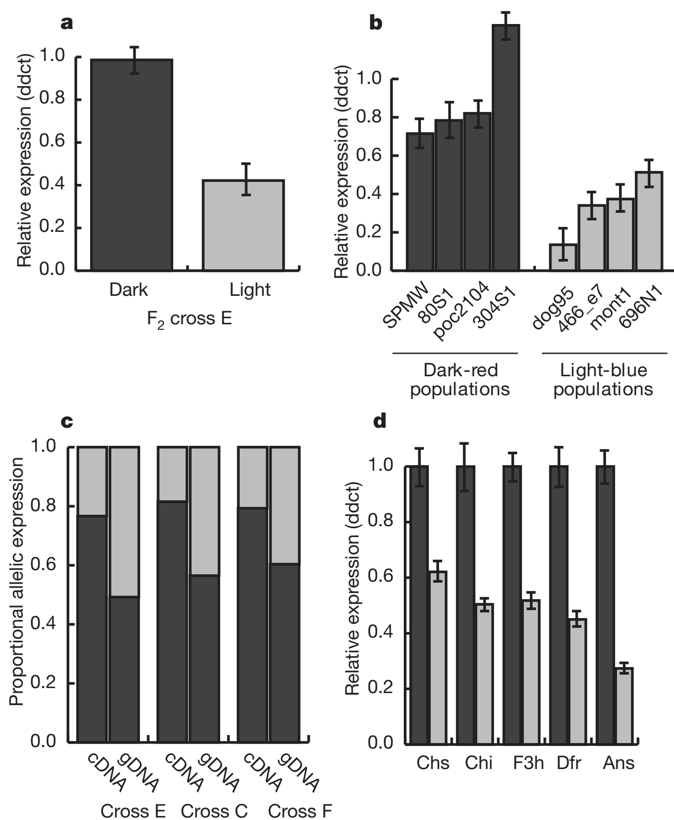

Figure 3 | Results of expression experiments of the intensity locus (*R2R3-Myb*). Relative expression of *R2R3-Myb* in dark and light F_2 individuals (**a**) and natural field populations (**b**). There is a significant upregulation associated with dark individuals relative to light both in F_2 individuals ($F_{(1,12)} = 8.11$, $P = 0.0173$) and in field-collected individuals ($F_{(1,50} = 14.91$, $P = 0.0023$). **c**, Allele-specific expression indicates significantly different allelic representation in cDNA relative to gDNA in heterozygous individuals (*Ii*) from three F_2 families ($F_{(1,12)} = 116.74$, $P < 0.0001$). The over-representation of the dark allele (dark grey) relative to the light allele (light grey) indicates a *cis*-regulatory change. See Supplementary Fig. 4 for complete control data. **d**, There is significant upregulation of transcript levels of all core ABP enzymes in field-collected dark individuals (dark grey) relative to light individuals (light grey). $F_{\text{chs (1,36)}} = 5.19$, $P = 0.0305$; $F_{\text{chi (1,36)}} = 5.07$, $P = 0.032$; $F_{\text{f3h (1,36)}} = 15.81$, $P = 0.0004$; $F_{\text{dfr (1,36)}} = 13.45$, $P = 0.001$; $F_{\text{ans (1,36)}} = 36.7$, $P < 0.0001$. For full MANOVA results see Supplementary Table 10. Standard errors are indicated on graphs.

upregulated relative to the 'light' allele (Fig. 3c), a pattern indicative of a *cis*-regulatory change at the *Myb* locus.

In all species that have been examined, this *Myb* coordinately activates several, if not all, of the ABP enzyme-coding genes[19]. Thus we expect that if changes in expression of *Myb* influence pigment intensity, there should be correlated expression changes in the ABP enzyme-coding genes. This expectation was realized: all five core-enzyme genes exhibited significant upregulation in dark flowers relative to light flowers (Fig. 3d). These results indicate that a *cis*-regulatory mutation at the *R2R3-MYB* transcription factor causes the increased colour intensity of *P. drummondii* flowers in sympatric populations.

Our investigation has provided the first identification of genetic changes causing reinforcement of species boundaries. Ideally, we would like to have confirmed the identity of the hue and intensity loci by either fine mapping or transformation, but this is not currently feasible in non-model systems like *P. drummondii*. Nevertheless, we are confident that we have identified the correct genes. Not only do *F3′5′h* and *Myb* co-segregate perfectly with the hue and intensity loci, respectively, but both exhibit changes in expression levels in the directions that are expected to produce the observed phenotypic changes. Moreover, the alternative possibility that the changes are due to linked transcription factors is ruled out by differences in allele-specific expression.

We have shown that reinforcement may involve changes in a few genes, each change having a large phenotypic effect. Our results expand upon two previous analyses of the genetic architecture of reinforcement, which report the involvement of a small number of quantitative trait loci[10,11]. This simple genetic architecture is consistent with theoretical expectations, which indicate that selection for reinforcement is most likely to result in increased reproductive isolation when the phenotypic effect of the assortative mating allele is large and selection for reproductive isolation is strong[7,20].

Ortiz-Barrientos *et al.*[10] suggest that reinforcing reproductive isolation should be inherited as a dominant trait because of Haldane's sieve (the greater probability of a new dominant adaptive mutation reaching fixation than a recessive mutation[21]). The intensity locus fits this expectation, with the derived dark allele dominant to the light allele. In contrast, the hue locus shows the reverse pattern, with the derived red allele recessive to the ancestral blue allele. Although it is easier for selection to increase the frequency of a novel dominant allele, Haldane's sieve can be overcome with strong selection[21]. Additionally, probability of fixation has been found to be independent of dominance when adaptations are not new mutations but are standing genetic variation[22]. Either of these possibilities could explain the fixation of the red allele in sympatry.

Much work investigating the process of speciation focuses on categorizing traits as a particular type of reproductive isolation mechanism (that is, prezygotic or postzygotic) (reviewed by Lowry *et al.*[23] and Nosil *et al.*[24]). Recently it has become clear that a single trait can affect multiple types of reproductive isolation[25,26], but it remains unclear how commonly this occurs. Although the most obvious effect of floral-colour evolution in *P. drummondii* is increased premating isolation with *P. cuspidata*, the associated genetic changes may also have led to increased postzygotic isolation. Hybrids between the dark-red-flowered *P. drummondii* and the light-blue-flowered *P. cuspidata* have dark-blue flowers, which differ from the two parental species. Levin has shown that pollinators discriminate against this intermediate hybrid flower colour[27]. Although further experimental tests are required, this discrimination probably reduces male outcross success and possibly seed set, causing reduced fitness of the hybrids beyond that associated with intrinsic postzygotic isolation. Our work shows how the pattern of genetic dominance can influence whether a trait can contribute to multiple types of isolation. This hypothesized extrinsic postzygotic isolation arises only because the novel allele at the intensity locus is dominant, whereas the novel allele at the hue locus is recessive. Had the novel alleles at both loci exhibited the same dominance, hybrids would have had the same flower colour as one of the parents.

Finally, both of the genetic changes contributing to reinforcement in *P. drummondii* involve *cis*-regulatory mutations, rather than functional (coding-sequence) mutations in pathway enzymes. The types of mutation described here are similar to those described for other cases of flower colour evolution in which reinforcement is not known to be involved[28–30]. These similarities suggest that the genetic basis of reinforcement may be comparable to that of adaptations not associated with reinforcement. However, only by indentifying the genetic basis of reinforcement in other systems can we understand what aspects of the molecular genetic basis of these traits are generally important for reinforcement and which are unique to a particular system.

METHODS SUMMARY

F_2 populations were created from crossing parents collected from allopatry with those collected from sympatry (Supplementary Table 1). F_2 individuals were phenotyped for flower colour and sampled for DNA and RNA extractions. Anthocyanidins were extracted from floral tissue using standard protocol and identified and quantified using high-performance liquid chromatography[30]. Candidate genes were amplified from parental individuals to identify SNPs, and F_2 individuals were genotyped. The Roche Universal ProbeLibrary system was used to quantify expression of all candidate

©2011 Macmillan Publishers Limited. All rights reserved

genes from F_2 and field-collected populations. Pyrosequencing technology was used to quantify allele-specific expression in individuals heterozygous at flower colour loci[18].

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 23 August; accepted 2 November 2010.
Published online 9 January 2011.

1. Noor, M. A. F. & Feder, J. L. Speciation genetics: evolving approaches. *Nature Rev. Genet.* **7,** 851–861 (2006).
2. Wolf, J. B. W., Lindell, J. & Backstrom, N. Speciation genetics: current status and evolving approaches. *Phil. Trans. R. Soc. B* **365,** 1717–1733 (2010).
3. Dobzhansky, T. *Genetics and the Origin of Species* (Columbia University Press, 1937).
4. Howard, D. J. in *Hybrid Zones and the Evolutionary Process* (ed. Harrison, R. G.) 46–69 (Oxford University Press, 1993).
5. Servedio, M. R. & Noor, M. A. F. The role of reinforcement in speciation: theory and data. *Annu. Rev. Ecol. Evol. Syst.* **34,** 339–364 (2003).
6. Butlin, R. Speciation by reinforcement. *Trends Ecol. Evol.* **2,** 8–13 (1987).
7. Felsenstein, J. Skepticism towards Santa Rosalia, or why are there so few kinds of animals. *Evolution* **35,** 124–138 (1981).
8. Ortiz-Barrientos, D., Grealy, A. & Nosil, P. The genetics and ecology of reinforcement implications for the evolution of prezygotic isolation in sympatry and beyond. *Year Evol. Biol.* **2009,** 156–182 (2009).
9. Pfennig, K. S. & Pfennig, D. W. Character displacement: ecological and reproductive responses to a common evolutionary problem. *Q. Rev. Biol.* **84,** 253–276 (2009).
10. Ortiz-Barrientos, D., Counterman, B. A. & Noor, M. A. F. The genetics of speciation by reinforcement. *PLoS Biol.* **2,** 2256–2263 (2004).
11. Saether, S. A. *et al.* Sex chromosome-linked species recognition and evolution of reproductive isolation in flycatchers. *Science* **318,** 95–97 (2007).
12. Levin, D. A. Reproductive character displacement in *Phlox*. *Evolution* **39,** 1275–1281 (1985).
13. Levin, D. A. Hybridization between annual species of *Phlox* – population structure. *Am. J. Bot.* 54, 1122-& (1967).
14. Levin, D. A. Interspecific hybridization, heterozygosity and gene exchange in *Phlox*. *Evolution* **29,** 37–51 (1975).
15. Ruane, L. G. & Donohue, K. Pollen competition and environmental effects on hybridization dynamics between *Phlox drummondii* and *Phlox cuspidata*. *Evol. Ecol.* **22,** 229–241 (2008).
16. Gonnet, J. F. CIELab measurement, a precise communication in flower color: an example with carnation (*Dianthus Caryophyllus*) cultivars. *J. Hortic. Sci.* **68,** 499–510 (1993).
17. Holton, T. A. & Cornish, E. C. Genetics and biochemistry of anthocyanin biosynthesis. *Plant Cell* **7,** 1071–1083 (1995).
18. Wittkopp, P. J., Haerum, B. K. & Clark, A. G. Evolutionary changes in *cis* and *trans* gene regulation. *Nature* **430,** 85–88 (2004).
19. Koes, R., Verweij, W. & Quattrocchio, F. Flavonoids: a colorful model for the regulation and evolution of biochemical pathways. *Trends Plant Sci.* **10,** 236–242 (2005).
20. Kirkpatrick, M. Reinforcement and divergence under assortative mating. *Proc. R. Soc. Lond. B* **267,** 1649–1655 (2000).
21. Haldane, J. B. S. A mathematical theory of natural and artificial selection, part 1. *Trans. Cambr. Phil. Soc.* **23,** 19–41 (1924).
22. Orr, H. A. & Betancourt, A. J. Haldane's sieve and adaptation from the standing genetic variation. *Genetics* **157,** 875–884 (2001).
23. Lowry, D. B., Modliszewski, J. L., Wright, K. M., Wu, C. A. & Willis, J. H. The strength and genetic basis of reproductive isolating barriers in flowering plants. *Phil. Trans. R. Soc. B* **363,** 3009–3021 (2008).
24. Nosil, P., Vines, T. H. & Funk, D. J. Perspective: reproductive isolation caused by natural selection against immigrants from divergent habitats. *Evolution* **59,** 705–719 (2005).
25. Dambroski, H. R. *et al.* The genetic basis for fruit odor discrimination in *Rhagoletis* flies and its significance for sympatric host shifts. *Evolution* **59,** 1953–1964 (2005).
26. Lowry, D. B. & Willis, J. H. A widespread chromosomal inversion polymorphism contributes to a major life-history transition, local adaptation, and reproductive isolation. *PLoS Biol.* **8,** e1000500 (2010).
27. Levin, D. A. The exploitation of pollinators by species and hybrids of *Phlox*. *Evolution* **24,** 367–377 (1970).
28. Hoballah, M. E. *et al.* Single gene-mediated shift in pollinator attraction in *Petunia*. *Plant Cell* **19,** 779–790 (2007).
29. Schwinn, K. *et al.* A small family of MYB-regulatory genes controls floral pigmentation intensity and patterning in the genus *Antirrhinum*. *Plant Cell* **18,** 831–851 (2006).
30. Des Marais, D. L. & Rausher, M. D. Parallel evolution at multiple levels in the origin of hummingbird pollinated flowers in *Ipomoea*. *Evolution* **64,** 2044–2054 (2010).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank D. Des Marais, J. Tung, S. Johnsen and T. Juenger for technical advice, D. Levin for assistance in locating natural populations, and M. Noor for comments on the manuscript. This work was supported by a National Science Foundation grant to M.D.R. and a National Science Foundation Doctoral Dissertation Research Improvement Grant to R.H. R.H. was supported in part by the National Science Foundation Graduate Research Fellowship Program.

Author Contributions R.H. and M.D.R. designed the project; R.H. performed the experiments and the analyses; R.H. and M.D.R. wrote the paper.

Author Information The DNA sequences reported here are deposited in GenBank under accession numbers HQ127319–HQ127344 and HQ323688–HQ323691. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to R.H. (robin.hopkins@duke.edu) or M.D.R. (mrausher@duke.edu).

©2011 Macmillan Publishers Limited. All rights reserved

METHODS

F_2 crosses. Seeds were collected from plants in six populations (Supplementary Table 1) in the spring of 2006, and germinated and grown at the Duke University greenhouses. Seeds were soaked for 48 h in 500 p.p.m. giberellic acid, planted in Metro-Mix 360 (Sun Gro Horticulture), and stratified for 6 days at 4 °C. Plants were transplanted at the four-true-leaf stage into a 50:50 mix of Fafard 4P (Conrad Fafard) and PermaTill One Time (Carolina Stalite Company). Light-blue plants were crossed with dark-red plants to form F_1 seeds. F_1 seeds were grown as above and self-fertilized to create F_2 populations. Self-incompatibility was overcome using bud pollination.

Flower colour phenotyping. All F_2 individuals were categorically phenotyped for flower colour, and a χ^2 test was used to determine if the ratio of categorical flower colour counts differed from 9:3:3:1 (Supplementary Table 3). A subset of individuals were quantitatively phenotyped with a StellarNet EPP2000C spectrometer with a SL1 visible light source and a SL5 deuterium +halogen ultraviolet–visible light source (StellarNet). Raw reflectance was transformed into two-dimensional CIE 1976 LUV colour space[16] (Supplementary Fig. 1). This transformation creates two axes of variation: the x axis corresponds to u' and the y axis corresponds to v'. The white point is located at ($v' = 0.4683305$, $u' = 0.197833$). The v' coordinate represents a measure of hue, and distance from the white point is a quantitative measure of intensity. We used canonical discriminant analysis to confirm that flower colour grouped into four discrete categories (Supplementary Table 4a, b). Anthocyanindins were extracted from a total of 26 individuals (12 from cross A, 14 from cross B) using standard methods of dissolving petal tissue in HCl followed by isoamyl alcohol extractions[30]. Pigments were identified using high-performance liquid chromatography as described in Des Marais and Rausher[30]. Pigments were quantified by calculating the area under the high-performance liquid chromatography peak and scaled to the corresponding standards. We used a MANOVA to determine if pigment amount differed by intensity, hue, family and all interactions of main effects. We used subsequent individual ANOVA models to determine significance of each of the above effects on amount of individual pigments (Supplementary Table 5a, b). All analyses used SAS software version 9.1 (SAS Institute).

Genetic association. Leaf tissue was collected from F_2 and parental individuals. A modified cetyltrimethylammonium bromide (CTAB) extraction was used to isolate the DNA (as in Kelly and Willis[31] but with an additional 2% Triton X-100 added to the CTAB solution and a 3 M sodium acetate wash after ethanol precipitation). Genes in the ABP were amplified, first using degenerate primers designed from orthologous genes in closely related species and then with species-specific primers in subsequent amplifications (Supplementary Table 6). Parental sequences for each gene were used to identify SNPs segregating in the F_2 populations. A subset of F_2 individuals were genotyped at each ABP gene. For those genes that showed an association, subsequent F_2 individuals were genotyped. χ^2 tests were used to confirm associations between genotype and phenotype (Supplementary Tables 7 and 8). A mixed-model ANOVA with F_2 family as a random effect was used to determine how much quantitative flower colour variation was explained by *F3′5′h* and the *R2R3-Myb* genotype. For this analysis we used flower colour reflectance transformed into CIE colour space to determine quantitative measures of hue and intensity (see above). The y axis, corresponding to the value of v', is the measure of hue and the distance each flower colour point is from the white point ($v' = 0.4683305$, $u' = 0.197833$) is the measure of intensity. All analyses used SAS software version 9.1 (SAS Institute).

RNA expression analyses. All expression analyses were performed on flower-bud tissue collected approximately 2 days before opening. RNA was extracted from individuals using the SpectrumTM Plant Total RNA Kit (Sigma-Aldrich). The Roche Applied Science Universal ProbeLibrary (Roche Diagnostics) was used to quantify expression of each gene in the ABP[32]. Probes sites and primers were designed using the online design centre (http://www.roche-applied-science.com/sis/rtpcr/upl/ezhome.html) (Supplementary Table 9). A Thermo Scientific Verso 1-Step RT-qPCR kit was used to amplify according to the manufacturer's instructions (Fisher Scientific). Twenty-six F_2 individuals were used in the *F3′5′h* expression assay. Twelve F_2 individuals were used in the *Myb* expression assay. Two replicate reactions were performed for each sample and the average cycle-threshold value was used in all analyses. Ef1-α was amplified in each sample to control for total amount of RNA in each extraction. Raw expression data were analysed as in Rieu and Powers[33]. We used a mixed-model ANOVA to detect a significant difference in expression in candidate genes between colour groups. F_2 family was used as a random effect in the model.

We collected seeds from natural populations of *P. drummondii* in both allopatry and sympatry (Supplementary Table 2), grew them in the Duke University greenhouse and extracted RNA from bud tissue as described above. Thirty-six individuals were used from field-collected population expression assays, with an additional 14 individuals from two additional populations for the *Myb* assay. Transcript levels of all genes in the ABP were quantified as above. We used a MANOVA to determine if flower colour intensity has an effect on expression of non-causal ABP genes (*Chs*, *Chi*, *F3h*, *Dfr*, *Ans*). Subsequent ANOVAs were used to determine if each individual gene shows a significant effect of intensity on expression (Supplementary Table 10). All analyses were performed using SAS software version 9.1 (SAS Institute).

Allele-specific expression. RNAs from F_1 individuals and heterozygous F_2 individuals were used to quantify allele-specific expression at both the *F3′5′h* and the *Myb* genes. A schematic of the assay design (Supplementary Fig. 2) shows the SNP identity, as well as amplification and sequencing primer sequences. RNA was extracted as described above from 28 individuals for the *Myb* assays (11 heterozygotes and seven homozygous controls) and from eight individuals for the *F3′5′h* assay (six heterozygotes and two homozygous controls). A reverse transcription reaction was performed using InvitrogenTM SuperScript II (Life Technologies) according to the manufacturer's instructions. DNA was extracted from leaf tissue as described above. For each individual, two replicate DNA and four replicate cDNA PCRs were run. No-template controls and no-sequencing-primer controls were performed as well. Pyrosequencing reactions were run on all samples using PyroMARKTM Q96ID (Qiagen)[18,34]. Experiments on both genes were independently replicated.

31. Kelly, A. J. & Willis, J. H. Polymorphic microsatellite loci in *Mimulus guttatus* and related species. *Mol. Ecol.* **7,** 769–774 (1998).
32. Mouritzen, P. *et al.* The ProbeLibrary™-expression profiling 99% of all human genes using only 90 dual-labeled real-time PCR probes. *Biotechniques* **37,** 492–495 (2004).
33. Rieu, I. & Powers, S. J. Real-time quantitative RT-PCR: design, calculations, and statistics. *Plant Cell* **21,** 1031–1033 (2009).
34. Ahmadian, A. *et al.* Single-nucleotide polymorphism analysis by pyrosequencing. *Anal. Biochem.* **280,** 103–110 (2000).

©2011 Macmillan Publishers Limited. All rights reserved