Genetic basis for a rare floral mutant in an Andean species of Solanaceae.

Rachel A Coburn  
*University of Nebraska - Lincoln*

Randi H Griffin  
*University of Nebraska - Lincoln*

Stacey D Smith  
*University of Colorado Boulder*

Follow this and additional works at: [https://scholar.colorado.edu/ebio_facpapers](https://scholar.colorado.edu/ebio_facpapers)

Recommended Citation  
[https://scholar.colorado.edu/ebio_facpapers/51](https://scholar.colorado.edu/ebio_facpapers/51)
Genetic basis for a rare floral mutant in an Andean species of Solanaceae

Rachel A. Coburn\textsuperscript{2}, Randi H. Griffin\textsuperscript{2,3}, Stacey D. Smith\textsuperscript{2,4,5}

\textsuperscript{2}University of Nebraska-Lincoln, 1104 T Street, Manter Hall, Lincoln, NE 68588

\textsuperscript{3}Duke University, 130 Science Drive, Biological Sciences Bldg 108, Durham, NC 27708

\textsuperscript{4}University of Colorado at Boulder, C127 Ramaley Hall, Campus Box 334, Boulder, CO 80309
The authors thank J. Stone, D. Pilson, and members of the Smith lab for advice as well as two anonymous reviewers for helpful suggestions on the manuscript. We also thank the Ministero del Ambiente of Ecuador for providing permits to conduct research (031-IC-FLO-DBAP-MA and 05-2012-IC-FLO-DBAP-MA), and the National Herbarium of Ecuador (QCNE) and the Herbarium of the Pontificia Universidad Católica (QCA) for assistance with fieldwork. This publication was made possible by grants from the National Science Foundation (DEB 1355518), the National Center for Research Resources (5P20RR016469), and the National Institute for General Medical Science (NIGMS-INBRE 8P20GM103427), a component of the National Institutes of Health (NIH). The contents are the sole responsibility of the authors and do not represent the views of NIGMS, NIH, or NSF.

Email for correspondence: stacey.d.smith@colorado.edu
• **Premise of the Study:** White forms of typically pigmented flowers are one of the most common polymorphisms in flowering plants. Although the range of genetic changes that give rise to white phenotypes is well known from model systems, few studies have identified causative mutations in natural populations.

• **Methods:** Here we combine genetic studies, in vitro enzyme function assays and biochemical analyses to identify the mechanism underlying the loss of anthocyanin pigment production in the naturally occurring white flowered morph of *Iochroma calycinum* (Solanaceae).

• **Key Results:** Comparison of anthocyanin gene sequences revealed a putative loss-of-function mutation, an 11 amino-acid deletion in dihydroflavonol 4-reductase (DFR), in the white morph. Functional assays of *Dfr* alleles from blue and white morphs demonstrated that this deletion results in a loss of enzymatic function, indicating that the deletion could be solely responsible for the lack of pigment production. Consistent with this, quantitative PCR showed no significant differences in expression of anthocyanin genes between the morphs. Also, thin layer chromatography confirmed that the white morph continues to accumulate compounds upstream of the DFR enzyme.

• **Conclusions:** Collectively, these experiments indicate that the structural mutation at *Dfr* underlies the rare white flower morph of *I. calycinum*. This study is one of only a few examples where a flower color polymorphism is due to a loss-of-function mutation in the coding region of an anthocyanin enzyme. The rarity of such mutations in nature suggests that negative consequences prevent fixation across populations.

**Key words:** anthocyanins; delphinidin; dihydroflavonol 4-reductase; flavonol; flower color; *Iochroma*; pleiotropy; quercetin
Many species of flowering plants exhibit naturally occurring color polymorphisms (Warren and Mackenzie, 2001; Rausher, 2008). The study of these polymorphisms has provided insight into a wide range of fundamental evolutionary questions (Clegg and Durbin, 2000) from the relative importance of natural selection and genetic drift (Schemske and Bierzychudek, 2007) to how variation is maintained within species (Gigord et al., 2001; Irwin, 2003). Because of its evolutionary lability, flower color has also become an important model for investigating the predictability of evolution, and in particular, whether certain genes are hotspots for phenotypic evolution (Stern and Orgogozo, 2009; Streisfeld and Rausher, 2011).

White forms of typically purple, blue, or pink flowered species are one of the most common and well-studied classes of flower color polymorphisms (Richards, 1986; Rausher, 2008). Surveys spanning many angiosperm families indicate that these white forms lack the anthocyanin pigmentation present in the colored morphs (Warren and Mackenzie, 2001; Whittall et al., 2006; Rausher, 2008). This absence of pigmentation may be limited to the flowers (Tiffin et al., 1998; Dick et al., 2011) or characterize the entire plant (Les et al., 1989; Wu et al., 2013). Given the extensive knowledge of the flavonoid pathway that produces anthocyanins and its broad conservation across land plants (Winkel-Shirley, 2001; Grotewold, 2006; Campanella et al., 2014), it is possible to pinpoint the types of genetic changes that could result in the origin of a white-flowered morph. For example, mutations in any of the transcription factors that regulate the pathway (bHLH, MYB, and WD40) (Mol et al., 1998; Koes et al., 2005) could lead to the loss of expression of the structural genes (e.g., Chi, F3h, etc.; Fig. 1) and thus produce a white flower (de Vetten et al., 1997; Spelt et al., 2002; Morita et al., 2006). Equally, mutations in the structural genes themselves, whether in the cis-regulatory regions or in coding sequences, could cause the loss of pigmentation (e.g., Martin et al., 1985; van Houwelingen et al., 1999).
Although a wide range of mutations may potentially result in a white flowered phenotype, their pleiotropic effects on flavonoid production may limit the extent to which these mutations persist in natural populations. In addition to their function in floral pigmentation and pollinator signaling (Schemske and Bradshaw, 1999; Streisfeld and Kohn, 2007), anthocyanins and related flavonoids play important roles in plant physiology. For example, flavonols are involved in male fertility (Mo et al., 1992; Pollak et al., 1993) and auxin transport (Kuhn et al., 2011) while both flavonols and anthocyanins have been implicated in response to UV stress (Chalker-Scott, 1999; Ryan et al., 2002). Thus, mutations that disrupt flavonoid production may result in negative consequences for plant fitness beyond their effects on flower color. Mutations in upstream genes (Chs, Chi), affecting all classes of flavonoids, carry the greatest potential for negative pleiotropic effects and would be expected to be rare in nature (Coberly and Rausher, 2003). By contrast, mutations in downstream genes (Dfr and Ans, Fig. 1) could alter anthocyanin production without affecting flavones and flavonols and thus may contribute more commonly to flower color polymorphisms (Wu et al., 2013). Similarly, regulatory changes (in cis or trans) can allow floral specific loss of pigmentation with few consequences for flavonoid production in other tissues (Streisfeld and Rausher, 2009; Smith and Rausher, 2011; Streisfeld et al., 2013).

Despite this strong theoretical framework for investigating the genetic basis of white flowered morphs, we have only a handful of cases in which the causative mutations have been identified in natural populations (Ipomoea purpurea: Habu, Hisatomi, and Iida, 1998; Chang, Lu, and Rausher, 2005; Parrya nudicaulis: Dick et al., 2011; Mimulus lewisii: Wu et al., 2013). These studies identified changes in both cis-regulatory (Dick et al., 2011) and coding regions (Habu et al., 1998; Chang et al., 2005; Wu et al., 2013) affecting three different loci (Chs, Dfr and Myb). Given the large number of loci that could potentially result in a white-flowered morph
(Fig. 1), expanding these studies to additional cases may reveal a broader spectrum of causative mutations. On the other hand, if mutations affecting particular loci have strongly negative consequences, the pool of loci underlying color polymorphisms may be limited. Here we focus on *Iochroma calycinum* Benth. (Solanaceae), a blue-flowered Andean shrub that exhibits rare white morphs. Combining analysis of anthocyanin gene sequences, comparative expression studies, in vitro assays of gene function, and flavonoid biochemistry, this study seeks to identify the genetic changes responsible for the loss of pigmentation in the white *I. calycinum*.

Specifically, we aimed to distinguish between changes in gene expression and gene function as causative factors and, if possible, to determine the underlying mutation. Understanding the spectrum of changes that give rise to such polymorphisms is important not only for characterizing the number of routes to novel phenotypes, but also the extent to which the predictability of genetic evolution commonly observed above the species level (e.g., Smith and Rausher, 2011) extends to variation within populations.

**MATERIALS AND METHODS**

**Study system** - The genus *Iochroma* (Solanaceae) consists of approximately 25 species distributed throughout the Andes of South America and is best known for its diversity of flower forms and colors (Smith and Baum, 2006). *Iochroma* and allied genera fall into the Physaleae clade of the large Solanoid radiation (Olmstead et al., 2008). Like several other lineages of Physaleae, *Iochroma* includes a number of species with inflated calyxes, including *I. calycinum*. However, *I. calycinum* is the only species of *Iochroma* in which the calyx matches the deep blue color of the corolla (Fig. 2A). This species is native to northern Ecuador and Colombia and is
principally pollinated by hummingbirds (Smith et al., 2008). Like many *Iochroma* species, the population sizes are small, on the order of 25 to 50 individuals. The common name of the species is “teta de vaca,” or cow’s teat plant, because its inflated calyx swells with viscous fluid before anthesis. These fleshy fluid-filled calyces have been hypothesized to deter florivores and nectar robbers (Lagerheim, 1891; Endress, 1996).

Field surveys and herbarium studies revealed that white *I. calycinum* mutants exist, albeit in low frequencies (S. D. Smith, unpublished). While the common morph has a dark blue corolla and calyx, the white morph appears to produce no anthocyanin pigments in the vegetative or floral tissues. One white morph was sampled from the cloud forests around Chiriboga, Ecuador (0.22602°S, 78.76962°W) and another was observed in herbarium collections from that same site (QCA). No other white morphs were observed among the roughly 150 specimens examined at other herbaria (MO, F, QCNE) or among over 75 plants seen in the field, suggesting that albino flowered *I. calycinum* plants are extremely rare in nature. Blue morphs were sampled in and around the Reserva Rio Guajalito in Chiriboga, Ecuador (0.24577°S, 78.80903°W). Floral and leaf tissue of the blue morphs and the white morph were collected in silica gel for DNA sequencing and thin layer chromatography, and floral buds (ca. 1cm in length) were collected in RNAlater (Qiagen, Valencia, CA) for gene expression studies. Sampling at this bud size (roughly equivalent to *Petunia* bud stage 5) captures the expression of both upstream and downstream genes as corollas are just beginning to show color (Pollak et al., 1993; Smith and Rausher, 2011).

Herbarium vouchers for the blue (Smith 471) and white (Smith 510) morphs were deposited at MO and QCA (Appendix 1).

**Sequence analysis of anthocyanin structural genes**- Given the structure of the
flavonoid pathway (Fig. 1), there are seven enzymes required for producing the delphinidin-
derived anthocyanin pigments that are found in blue-flowered *Iochroma* species (Smith and
Rausher, 2011). White flowers could potentially arise by structural (coding sequence) mutations
at any of these loci (*Chs, Chi, F3h, F3’h, F3’5’h, Dfr*, and *Ans*). The two hydroxylating enzymes
*F3’h* and *F3’5’h* may appear to be weak candidate genes for white flowers because loss of
function at these loci should convert a blue delphinidin-producing flower to a red pelargonidin-
producing flower (Smith and Rausher, 2011). However, in all Solanaceae examined to date, *DFR*
shows some substrate specificity, usually exhibiting lower activity on DHK, the precursor of red
pelargonidin pigments, relative to other substrates (Johnson et al., 2001; Des Marais and
Rausher, 2008; Smith et al., 2013). Thus, depending on the specificity of downstream enzymes, a
mutation at *F3’h* and *F3’5’h* could result in a white flower (van Houwelingen et al., 1998;
Brugliera et al., 1999).

To identify putative loss-of-function mutations in the white morph, we amplified each of
the seven genes from one blue and one white individual using custom primers (Appendix S1; see
Supplemental Data with the online version of this article) designed from an *Iochroma cyaneum*
454 floral transcriptome (S. D. Smith, unpublished data). Although one of these genes (*Chs*) is
commonly present in multiple copies in Solanaceae, only one or a few copies are florally-
expressed (Koes et al., 1989; Van der Meer et al., 1990; Morita et al., 2012), facilitating primer
design from the *Iochroma* transcriptome. Polymerase chain reactions were carried out with
iProof High-Fidelity DNA Polymerase following manufacturer’s protocols (Bio-Rad, Hercules,
CA), and the products were purified using the QIAquick PCR Purification Kit (Qiagen, Valencia,
CA). Direct sequencing was performed using the BigDye Terminator v3.1 Sequencing Kit
(Invitrogen, Carlsbad, CA), and Sanger sequencing was conducted at Duke University’s Genome
Sequencing and Analysis Core (Durham, NC). Chromatograms were edited and analyzed using Four Peaks v.1.7.1 (Griekspoor and Groothuis, 2007) and MacClade 4.0 (Maddison and Maddison, 2005). Full length coding sequences were deposited into GenBank (Appendix 1) and the alignments and trees for each gene were deposited into TreeBase under study number S16161 (www.treebase.org).

Gene trees were estimated for each of the seven genes from both the white and blue *I. calycinum* morphs to verify orthology with known anthocyanin loci from model systems. Coding sequences from five closely related species (*Petunia hybrida, Nicotiana benthamiana, Solanum lycopersicum, Solanum tuberosum, and Iochroma cyaneum*) were retrieved from GenBank, http://solgenomics.net, or transcriptomic data in the case of *I. cyaneum* (Appendix 1). Sequences were aligned manually in MacClade 4.08a (Maddison and Maddison, 2005), and maximum likelihood trees were estimated using a heuristic search in PAUP version 4.0b10 (Swofford, 2002) with a GTR+G model and TBR branch-swapping. The resulting gene trees were rooted with *Petunia* sequences as the outgroup.

**Assay of DFR function** - Our sequence analyses identified a large deletion in the coding region of *Dfr* in the white morph. To determine if this mutation affects DFR function, we performed in vitro assays of protein function using sequences from both color morphs. The *Dfr* gene from each morph was directionally cloned into the pENTR-SD/D-TOPO vector and recombined into the pDEST14 overexpression vector using the Gateway kit (Invitrogen, Carlsbad, CA). The vectors were transformed into BL21* (DE3) *Escherichia coli* cells (Invitrogen, Carlsbad, CA). Overnight cultures of 5 mL LB-AMP100 were used to inoculate 200 mL cultures and grown at 37°C until reaching OD600 0.5. To induce expression, 500 μL of 200
mM isopropyl-beta-D-thiogalactopyranoside was added to the culture and incubated for 3 hours at 28°C. Expression of the transgene was verified with qualitative real-time PCR using 2.5 ng of cDNA template synthesized from pelleted cells from 15 mL of post-expression culture. A functional assay was performed on the crude protein extract following Des Marais and Rausher (2008), and dihydroflavonols (DHK, DHQ, and DHM) were used as the substrates. The resulting leucoanthocyanidins were heated at 95°C for 30 minutes in 95% Butanol: 5% HCl to convert the substrates into colored anthocyanidin products (delphinidin, cyanidin, and pelargonidin). These were quantified by measuring peak absorbance in a Beckman DU-800 spectrophotometer. Absorbance values were converted into μmoles by comparison with a standard dilution series. The percentage of substrate converted was calculated by dividing the μmoles of anthocyanidin produced by the μmoles of dihydroflavonol substrate provided.

**Quantification of anthocyanin gene expression** - Gene expression was measured to determine if the loss of pigment in the white morph could be due to expression level differences in pathway genes. Total RNA was extracted from RNAlater-preserved floral buds of the white morph and five blue-flowered individuals using the Spectrum kit (Sigma-Aldrich, St. Louis, MO). An on-column DNase digestion (Qiagen, Valencia, CA) was performed to remove gDNA from extracted RNA. First strand cDNA synthesis was done with 0.5 μg total RNA in a 20 μL reaction using SuperScript II Reverse Transcriptase following manufacturer’s instructions (Invitrogen Technologies, Carlsbad, CA). The resulting cDNA was diluted to 2.5 ng/μL for quantitative real-time PCR (qPCR). Each 20 μL qPCR reaction contained 2.5 ng of the cDNA template, 10 μL DyNAmo SYBR Green mix (Invitrogen Technologies, Carlsbad, CA), 0.4 μL 1x 6-carboxy-X-rhodamine passive reference dye, 0.4 μL of forward primer (10 μM), and 0.4 μL of
reverse primer (10 μM). qPCR primers for Chi, Chs, and F3h were designed as a part of this study while those for the remaining genes were taken from Smith and Rausher (2011) (Appendix S2; see Supplemental Data with the online version of this article). Elongation factor (EF1-α) was used as a reference gene. The qPCR reactions were run on an ABI 7500 Fast machine under the following conditions: 95°C for 10 minutes, followed by 40 cycles of 95°C for 20 seconds and 55°C for 1 minute, and a dissociation step to verify primer specificity. The qPCR products were additionally run on a 2% high resolution agarose gel to verify only one product was amplified. Three to six replicate reactions were run for each cDNA and primer combination. Efficiency for each gene was determined as described in Peirson et al. (2003). We calculated relative expression levels following Pfaffl (2001) to quantify the variation across loci and across blue flowered individuals and determine whether the white individual showed significantly lower expression of any of the structural genes. Differences in relative expression across individuals and morphs were analyzed using a linear mixed model with REML model fitting in JMP Pro 11 (SAS, Cary, NC), treating color as a fixed effect and individual as a nested random effect.

**Identification of anthocyanidins and related flavonoids** - Floral and vegetative flavonoids (including anthocyanidins) were identified using thin layer chromatography to establish any differences between the blue and white morph. Flavonoids were extracted from 0.025 g of dried flower or leaf tissue by soaking overnight in 2 mL 2N HCl and refluxed at 95°C for 1 hour to convert the glycosylated compounds (e.g., anthocyanins) to their corresponding aglycones (e.g., anthocyanidins; Harborne, 1998). Flavonols were separated by washing twice with 0.5 mL ethyl acetate, and anthocyanidins were extracted by washing once with 0.15 mL isosamyl alcohol. Both layers were evaporated to dryness, eluted in 50 μL MeOH.
1% HCl, and spotted onto cellulose plates for chromatography. Plates were developed for 4 to 6 hours in forestal solvent (30:3:10 acetic acid: 12N HCl: H2O), and compounds were identified by comparison of Rf values with standards. Standards for delphinidin, malvidin, petunidin, myricetin, quercetin, and kaempferol were obtained from Extrasynthese (Genay, France).

RESULTS

Sequence variation between blue and white morphs- Full coding regions were obtained for seven structural anthocyanin pathway genes (Chs, Chi, F3h, F3’h, F3’5’h, Dfr, and Ans) from both morphs of I. calycinum. Phylogenetic analyses suggest that these sequences are orthologous with the functionally characterized enzymes from model systems. At the amino acid level, the I. calycinum sequences were 77-95% similar to sequences from other Solanaceae. Moreover, the topology of the gene trees for each locus mirror the species relationships, with Iochroma more closely related to Solanum than to Nicotiana or Petunia (Appendix S3; see Supplemental Data with the online version of this article). Although, like other Solanaceae, Iochroma has multiple genomic copies of Chs (S. D. Smith, unpublished data), we isolated a single florally-expressed copy that is most similar to CHS-A, the dominant copy expressed in Petunia flowers (Koes et al., 1986). This pattern suggests that the floral-specific activity of this “A” copy may be broadly conserved across Solanaceae.

We examined sequence variation in these structural genes to identify differences that could underlie the flower color polymorphism, such as frameshifts, premature stop codons or insertion-deletion (indel) events. Comparing the sequences from the white and blue morphs of I. calycinum across all loci, we observed 14 synonymous mutations, five non-synonymous
mutations, and one deletion in Dfr (Appendix S4; see Supplemental Data with the online version of this article). These non-synonymous changes could potentially reduce enzyme function and account for the loss of pigment production, however, in each case, that amino acid variant present in the white morph allele was also present in one or more pigment-producing outgroup species (P. hybrida, N. benthamiana, S. lycopersicum, S. tuberosum, and I. cyaneum). Moreover, since none of the genes from the white morph had more than one non-synonymous difference compared to the blue morph, negative interactions among mutations at the same locus in the white morph would also not explain the phenotype. By contrast, the Dfr deletion was unique to the white morph and would be predicted to affect function given its position in the protein. The 11 amino-acid deletion spans position 216-226 in the previously elucidated grape (Vitis vinifera) crystal structure, and this area includes multiple sites important for the conformation of the active site for the enzyme (Fig. 2B; Petit et al., 2007). If this mutation were responsible for the loss of pigmentation, we would expect it to be absent in (or at least heterozygous) in all pigmented individuals. A PCR screen of five additional blue individuals confirmed that all lacked the deletion while the white individual was homozygous for the deletion (Appendix S5; see Supplemental Data with the online version of this article).

**Nonfunctional DFR in white I. calycinum**- In vitro assays comparing the function of blue and white alleles of Dfr indicate that the 11 amino-acid deletion results in a loss of enzymatic activity. A functional Dfr should convert dihydroflavonols into leucoanthocyanidins (which are converted to anthocyanidins during the assay protocol). The allele from the blue-flowered I. calycinum was able to convert all three dihydroflavonol precursors into colored anthocyanidins while the white allele was non-functional (Fig. 2C, D). The pattern of activity for
the blue morph copy was similar to other blue-flowered *Iochroma* species previously examined (Smith and Rausher, 2011; Smith et al., 2013) where activity is low on DHK (the precursor of red anthocyanins) and high on DHQ and DHM (the precursors of purple and blue anthocyanins).

By contrast, the deletion-carrying allele from the white morph of *I. calycinum* showed negligible activity on any substrate and was not statistically different from the vector without an insert (Fig. 2C). Given that the white allele of *Dfr* shows no putative loss-of-function mutations other than the deletion, these results suggest that the disruption of the active site by the deletion is sufficient to eliminate enzymatic activity.

**Expression differences between blue and white morphs** - Comparisons of gene expression levels between the white morph and five randomly selected blue morphs revealed no evidence of a significant association between gene expression and color. We observed about a 5-fold difference in expression across individuals and genes, however, in all cases the expression levels for the white sample fell within the range of the blue samples (Fig. 3). Looking within genes, levels of gene expression varied roughly 3-fold across individuals, a range similar to that observed for intraspecific expression variation in other studies (Streisfeld and Rausher, 2009; Cooley et al., 2011; Smith and Rausher, 2011) and which may reflect slight differences in the age of the buds sampled or simply developmental noise. Comparing white and blue samples, expression levels for the white individual were relatively low for *Dfr* and *Ans*, but equivalent to or higher than the blue individuals for the other five loci. Even for *Dfr* and *Ans*, the average expression was similar to at least one of the blue individuals, and results from linear mixed models showed no significant effect of color on expression (p=0.16-0.76 across genes). In addition, the differences in expression observed across the *I. calycinum* individuals are minor.
relative to the changes that affect flower color in other cases (24 to 200-fold; Streisfeld and Rausher, 2009; Des Marais and Rausher, 2010; Dick et al., 2011; Smith and Rausher, 2011; Streisfeld and Rausher, 2011). Overall, these results indicate that differences in gene expression are not likely to be responsible for the white morph of *I. calycinum*. In addition, the pattern of expression variation suggests transcription factors did not play a role in the lack of pigmentation because, in that case, we would expect to see multiple genes coordinately down regulated in white flowers (Whittall et al., 2006).

*Effects of Dfr loss of function on flavonoid production*- Thin layer chromatography (TLC) was conducted on floral and leaf tissue of blue and white *I. calycinum* to determine if anthocyanidins and related flavonoids were being produced. If the *Dfr* deletion is the only change in the pathway, we would predict that anthocyanidin production would be lost while upstream products (e.g., flavonols) would not be affected (Fig. 1). The TLC analysis confirmed the presence of the two anthocyanidins (delphinidin and malvidin, a delphinidin derivative) in blue-flowered *I. calycinum* and the lack of these compounds in the white morph (Appendix S6; see Supplemental Data with the online version of this article). These delphinidin-derived pigments are also responsible for blue and purple colors in other *Iochroma* and related species (Smith and Rausher, 2011; S. D. Smith, unpublished data). The flowers of the blue morph also produced the flavonol quercetin, and consistent with our prediction, this compound is also present in the flowers of the white morph. None of these compounds were detected in leaf tissue (Appendix S6).

**DISCUSSION**
The goal of this project was to determine the genetic basis for the lack of pigment at ion in the white morph of the normally blue-flowered *I. calycinum*. Analysis of the coding regions of anthocyanin structural genes revealed a single putative loss of function mutation (an 11-amino-acid deletion in DFR) distinguishing the blue and white morphs. Assays of DFR function using alleles cloned from each morph confirmed that this deletion causes a loss of enzyme activity, and thus blocks anthocyanin production. We did not observe any significant differences in gene expression between the blue and white morphs, suggesting that this structural mutation alone underlies the loss of flower color. Consistent with the downstream position of DFR in the pathway (Fig. 1), we found that the white morph continues to make upstream pathway products (flavonols) despite the loss of pigmentation.

*Functional evolution of DFR and flower color variation*- Like the other enzymes of the flavonoid biosynthetic pathway, DFR and its role in pigment production are deeply conserved, dating back at least to the common ancestor of land plants (Rausher, 2006; Campanella et al., 2014). This enzyme is responsible for reducing dihydroflavonols such as DHK, DHQ, and DHM to form leucoanthocyanidins, which in turn are converted into anthocyanin pigments. Although DFR has the potential to act on multiple substrates, functional analyses commonly reveal substrate specialization, where the enzyme preferentially reduces particular compounds (Johnson et al., 2001; Halbwirth et al., 2003; Xie et al., 2004; Des Marais and Rausher, 2008; Leonard et al., 2008). Previous studies demonstrated that Dfr copies from blue-flowered *Iochroma* species show higher activity on DHQ and DHM, the precursors of purple and blue anthocyanins than on DHK, the precursor of red anthocyanins (Smith et al., 2013). This study shows that the blue-
flowered *I. calycinum* shares the same pattern of substrate specificity, which has also been documented in *Petunia* (Johnson et al., 2001).

These comparative functional studies of DFR together with the recently published crystal structure from *V. vinifera* (Petit et al., 2007) have clarified the regions of the enzyme that are involved in function and specificity. Sites important for specificity fall into two clusters in the gene, one around position 133 in *V. vinifera* and another around position 227 (Petit et al., 2007). Although early studies emphasized the role of the first cluster in underlying species differences in pigmentation (Beld et al., 1989; Johnson et al., 2001), recent studies suggest that the second cluster may be equally important in functional evolution (Fischer et al., 2003; Smith et al., 2013). Both areas are part of the substrate binding site and interact with the B-ring hydroxyl groups that distinguish the classes of dihydroflavonols. The present study confirms the key role of the second cluster in DFR function. The 11 amino-acid deletion found in the *Dfr* allele from the white morph spans positions 216 to 226 in *V. vinifera* (Fig. 2B), and the effect of this mutation is to eliminate DFR activity on dihydroflavonols (Fig. 2C). Additional studies would be needed to determine if this effect is due to a loss of substrate binding specifically or the overall deformation of the active site.

**Possible consequences of loss of DFR function**- Our surveys of natural populations and herbarium collections suggest that white-flowered individuals of *I. calycinum* are rare (only found in and around the town of Chiriboga, Ecuador). The rarity of this form could relate to direct effects of the loss-of-function mutation on plant-pollinator interactions. For example, pollinators may discriminate against white flowers because of innate preferences (Launau and Maier, 1995; Weiss, 1997) or longer handling times (Waser and Price, 1983). Innate preferences
seem unlikely to play a major role in this system because the hummingbird pollinators of *I.* calycinum are known to visit a wide array of flower colors, including white, in *Iochroma* and in other Andean genera (Snow and Snow, 1980; Dziedzioch et al., 2003; Smith et al., 2008). White-flowered morphs of *Delphinium nuttalianum* have been found to incur longer handling times due to the loss of nectar guides (Waser and Price, 1983, 1985). The tubular blue flowers of *I.* calycinum are uniform in color (without apparent nectar guides) and the white form should present a strong contrast against the green background, arguing against inefficient foraging as an explanation.

An alternative explanation for the rarity of white is that the loss of DFR function has negative consequences beyond its effects on floral pigmentation. Since *Dfr* is a single copy gene in Solanaceae (Des Marais and Rausher, 2008), any mutations in the coding region will affect its function in any tissue where it is expressed. However, we observed no anthocyanins in vegetative tissue (Appendix S6), suggesting that there may be little DFR activity outside the flower. The other major class of flavonoids that require DFR function, condensed tannins (derived from leucoanthocyanidins), are not present in Solanaceae (Eich, 2008). Even though the sampled blue-flowered *I.* calycinum did not express anthocyanins in vegetative tissue, the ability to do so may still have adaptive significance. Anthocyanins are often expressed in vegetative tissues when plants are stressed, e.g., by temperature extremes or drought (Chalker-Scott, 1999), and several studies suggest that this response is adaptive (Schemske and Bierzychudek, 2001; Warren and Mackenzie, 2001; Coberly and Rausher, 2003). *Iochroma* are long-lived shrubs, likely to experience environmental stresses over their lifetimes, and the inability to produce anthocyanins might limit the potential for white morphs to survive to flowering (typically 2-3 years) and beyond.
Finally, the low frequency of the white morph may be due to neutral (non-selective) factors. For example, if the white allele arose only recently, it may still be at low frequency even if it were strongly favored by selection. Genetic drift could also play a role in shaping the frequency of color morphs, although no studies to date have found flower color variation to be neutral (reviewed in Rausher, 2008). There are multiple population genetic approaches for testing for selection against the white allele of *I. calycinum* (e.g., Slatkin, 2000) which could be used to examine the neutral hypothesis in future studies.

The genetic basis of within-species flower color polymorphisms - Despite the large number of pigmented species that exhibit white-flowered morphs (e.g., Epling and Dobzhansky, 1942; Waser and Price, 1981; Brown and Clegg, 1984; Levin and Brack, 1995), relatively little is known about the exact genetic cause for these losses of pigmentation. To date, there are four cases in which the genetic basis of white-flowered morphs has been identified, and these involve an array of mutations affecting both coding and *cis*-regulatory regions of structural genes and transcription factors. The white morph of the artic mustard *Parrya nudicaulis* is due to a *cis*-regulatory mutation in *Chs* (Dick et al., 2011), and a transposon insertion in the coding region of this gene causes white flowers in *Ipomoea purpurea* (Habu et al., 1998; Coberly and Rausher, 2003). White morphs of *I. purpurea* have also been associated with a loss-of-function allele (containing two deletions) at the *Ipmyb1* locus that encodes an R2R3-MYB transcription factor (Chang et al., 2005). Finally, a frameshift mutation causing a premature stop codon in DFR is responsible for white flowers in *Mimulus lewisii* (Wu et al., 2013). This study thus represents the second case of a loss-of-function mutation in DFR contributing to intraspecific flower color variation, although here the mutation is in-frame and likely acts by disrupting the active site.
Overall, this is the third instance in which a loss-of-function mutation in a structural gene underlies white-flowered forms within species. The wide variety of mutations involved in segregating flower color variation within populations stands in stark contrast with the pattern observed for fixed differences between species. In all cases (*Petunia axillaris*, several *Antirrhinum* species), evolutionary transitions from pigmented to white flowers have involved mutations in the R2R3-MYB transcription factors that regulate anthocyanin biosynthesis (Quattrocchio et al., 1999; Schwinn et al., 2006; Hoballah et al., 2007). This genetic convergence has been hypothesized to relate to the specificity of this class of transcription factors. Among this large gene family, one subgroup is specialized for regulating anthocyanins (Dubos et al., 2010), and different copies within this subgroup are often tissue-specific (Ramsay and Glover, 2005; Schwinn et al., 2006). Thus, mutations in the R2R3-MYBs have the potential to alter floral pigment production with minimal pleiotropic consequences for the production of other classes of flavonoids or flavonoid production in other tissues (Streisfeld and Rausher, 2011).

This contrast between the genetic basis of segregating flower color morphs and species-level differences suggests that a small and predictable subset of the mutations that give rise to white flowers within populations, namely those that affect R2R3-MYB genes, are preferentially fixed over longer evolutionary timescales (Stern and Orgogozo, 2008; Streisfeld and Rausher, 2011). Understanding the mechanism responsible for this predictability will require additional exploration of the genes underlying segregating white morphs and the selection acting on those variants. Cases where the nature of selection acting on white-flowered plants is already known (e.g., Kay, 1976; Waser and Price, 1983; Levin and Brack, 1995) would be ideal systems for these genetic studies. Given the diverse roles of flavonoids in plants (Harborne and Grayer,
1993; Winkel-Shirley, 2001), we expect that sources and strength of selection acting on mutations at different pathway genes will also be diverse. Indeed, strong purifying selection acting on some pathway loci (e.g., Chi and F3h) may explain why naturally occurring loss-of-function mutants have yet to be found. Targeting populations with exceptionally rare white morphs, as in the case of I. calycinum, may provide the best approach for discovering these “missing” mutants and characterizing the spectrum of genetic variants segregating in nature.
LITERATURE CITED


Coburn et al. 2015   Genetic basis for a rare floral mutant


APPENDIX 1. Vouchers and gene accession information. Vouchers are deposited at the University of Wisconsin-Madison Herbarium (WIS) and the Missouri Botanical Garden (MO). *Sequence from Smith and Rausher (2011) and ** from Smith, Wang and Rausher (2013).

_Taxon_; voucher; description; location; GenBank accession numbers: _Chs, Chi, F3h, F3’h, F3’5’h, Dfr, Ans._

Solanaceae _Iochroma cyaneum_ (Lindl.) M. L. Green; _Smith 265_ (WIS); Treelet, 2m, flowers blue; Cultivated at University of Wisconsin-Madison Greenhouses; KJ094348, KJ094350, KJ094352, KJ094353, KJ094354, GU595064*, KJ094351.

_I. calycinum_ Benth.; _Smith 471_ (MO); Treelet, 2m, flowers blue; Reserva Rio Guajalito, Chiriboga, Ecuador; KJ094360, KJ094355, KJ094347, KJ094344, KJ094346, JN593314**, KJ094356.

_I. calycinum_ Benth.; _Smith 510_ (MO); Treelet, 2m, flowers white; Vicinity of Chiriboga, Ecuador; KF955587, KJ094342, KJ094358, KJ094343, KJ094345, KJ094357, KJ094359.
Figure Legends

Fig. 1. Flavonoid biosynthetic pathway. Black text indicates substrates (DHK, dihydrokaempferol; DHQ, dihydroquercetin; DHM, dihydromyricetin) while gray text indicates enzymes (CHS, chalcone synthase; CHI, chalcone isomerase; F3H, flavonone 3-hydroxylase; FLS, flavonol synthase; F3’H, flavonoid 3’- hydroxylase; F3’5’H, flavonoid 3’, 5’-hydroxylase; DFR, dihydroflavonol 4-reductase; ANS, anthocyanidin synthase). Products of this pathway include flavonols, flavones, and pigmented anthocyanins. Although the pathway can produce pelargonidin (red) and cyanidin (purple) derived anthocyanins, blue-flowered Ioehroma species are pigmented with only blue delphinidin-derived anthocyanins.

Fig. 2. Relationship of Dfr structure and function to flower color. (A) Floral phenotype of blue and white morph of I. calycinum. (B) Diagram of structural mutation in Dfr associated with white morph. Residues important for substrate binding (black numbers) and active site conformation (gray numbers) are based on the crystal structure of grape (Petit et al., 2007). The deletion in the white allele is indicated by the vertical black bar. (C) Percent μmoles of the substrates DHK, DHQ, and DHM converted into pelargonidin, cyanidin, and delphinidin, respectively, by the blue morph DFR, white morph DFR, and the vector without an insert. The absorbance curves for these products are shown in (D). Peak absorbance for each product was converted to percent μmoles using standard dilution measurements. Error bars represent ± 1 SD. (D) Absorbance curves for pelargonidin, cyanidin, and delphinidin, respectively from the functional assays of the blue and white morphs of I. calycinum. The shape of these curves traces
the characteristic absorbance spectrum for each compound. Line colors follow (C). See Fig. 1 caption for abbreviations.

Fig. 3. Anthocyanin gene expression differences between white individual and five blue individuals of *I. calycinum*. Expression was measured in floral bud tissue for *Chs, Chi, F3h, F3’h, F3’5’h, Dfr,* and *Ans* (see Fig. 1 caption for gene abbreviations) and mean relative quantification values were calculated (± 1 SE). Expression values are shown relative to individual F, which was randomly chosen amongst the blue individuals. The RQ values correspond to a log$_{10}$ fold change in expression in comparison to individual F.
p-Coumaroyl CoA + 3 Malonyl-CoA

Flavonols
- Kaempferol
- Quercetin
- Myricetin

Flavones
- Flavones
- Flavones

Anthocyanins
- Anthocyanins
- Anthocyanins

**Flavonols**

**Flavones**

**Anthocyanins**
Figure 2

Click here to download Figure: Figure2_revised_120114.eps
Figure 3
Click here to download Figure: Figure3_revised_120714.eps
Appendix S1: Primers used to sequence coding regions of flavonoid pathway genes (see Fig. 1 caption for gene abbreviations). All primer sequences are shown 5’→3’. The primer numbering listed before each sequence is based off of the base pair position after the start codon of the reference sequence (last column). Negative numbers indicate primers which start before the start codon. For example, the Chs primer 1408F begins at base pair position 1408 of the GenBank sequence U47739. The Dfr primers in red were additionally used in genotyping the white and blue individuals of *I. calycinum*.

<table>
<thead>
<tr>
<th>Gene</th>
<th>Forward Primer</th>
<th>Reverse Primer</th>
<th>Ref. Seq. GB#</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chs</td>
<td>(1408F) GCT AGT TGG TAT TTT CAC ACA CAC AAA CCT</td>
<td>(3539R) GAG AAG GAG AAA TGG ACT TCC TTA CAC</td>
<td>U47739</td>
</tr>
<tr>
<td>Chi</td>
<td>(1277F) TAT CAT TAC AGG TCC CTT TGA GAA ATT CAC TCG</td>
<td>(2038R) CTG GAA TGG TAG ATT TCC TCT TCA</td>
<td>X14590</td>
</tr>
<tr>
<td>F3h</td>
<td>(-25F) AAC AAG AAA GTT AGT TTT TGA ATC GAT GGC ACC TTC</td>
<td>(1431R) CAC TTT TIG ATC CAT ATC CAC ACA TGC CTT AGT T</td>
<td>AF022142</td>
</tr>
<tr>
<td>F3 h</td>
<td>(839F) ATT GGA GAA ATG AAG GAA CAG GGG GAC</td>
<td>(8F) GTC CAA CTT TCC TCT CTA ACA TTA CAC ATC</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(1073F) CGG CTC GTT AGT GAA GAA TGG</td>
<td>(416R) ACC GTG GTC CAT AAG GTG CGA AAA</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(8F) GTC CAA CTT TCC TCT CTA ACA TTA CAC ATC</td>
<td></td>
<td></td>
</tr>
<tr>
<td>F3 5’b</td>
<td>(868F) GGT ATG GCA AAT AGG GAC AAT TTC TCT T</td>
<td>(1571R) CAA TTT GCA TAA TAA ACA TCC AAA GGT AGC CTT GG</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(898F) CTG AAG GAG AAA GGC TCA GTA CAA CC</td>
<td>(1543R) GCC TGG GAG TAA CCA TAG CTT C</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(1543R) GCC TGG GAG TAA CCA TAG CTT C</td>
<td></td>
</tr>
<tr>
<td>Dfr</td>
<td>(73F) TCT GAA AAT GGC AAG TGA AG</td>
<td>(1275R) GAA GCA AAC ATT ATC CTC AGG C</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(570F) GCT AAG AAG ATG ACA GGA TGG</td>
<td>(1016R) TGT AAA CCC CAT ATC CAT TAG C</td>
<td></td>
</tr>
<tr>
<td>Ans</td>
<td>(18F) GGG AAA AAG AAA AAG CAA GTG AGG GAA TTA G</td>
<td>(399F) ATG CTA ATG ACT TACC ACT CCT G</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(999F) TAA GCC CCT AGC TGA ATG TGT TAC CGA G</td>
<td>(1300F) CAA CAA GTT AAA CTA TTT AGA TGC TGC AGC AGC AAC</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(1350R) CAA CAA GTT AAA CTA TTT AGA TGC TGC AGC AGC AAC</td>
<td></td>
</tr>
</tbody>
</table>
Appendix S2: Primers used in qPCR of flavonoid pathway genes (see Fig. 1 caption for gene abbreviations). Primer orientation, naming, and reference sequences follow Appendix S1.

<table>
<thead>
<tr>
<th>Gene</th>
<th>Forward Primer</th>
<th>Reverse Primer</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chs</td>
<td>(1652F) GCA TGT GTG ATA AAT CAA TGA TTA AGA</td>
<td>(5707R) GCT TAG TGA GTT GGT AGT CGG</td>
<td>This study</td>
</tr>
<tr>
<td>Chi</td>
<td>(1381F) GGT CCC TTT GAG AAA TTC ACT CG</td>
<td>(1800R) TCA TCT TTA GAG AAG CTA ATC GTC AAT</td>
<td>This study</td>
</tr>
<tr>
<td>F3h</td>
<td>(590F) TAG CCA TCT ACA GGG TGA AGT</td>
<td>(1495R) GCT GTG GAC ACT TTG GGT AAA</td>
<td>This study</td>
</tr>
<tr>
<td>F3*h</td>
<td>(1111F) TTT GGA AGC CAT AGT CAA GGA AAC</td>
<td>(1318R) GGT CCC TTT GAG AAA TTC ACT CG</td>
<td>(Smith and Rausher, 2011)</td>
</tr>
<tr>
<td>F3'5'h</td>
<td>(510F) GCC ACA TGC TAA AAT CAA TGT TC</td>
<td>(664R) CCT TAA ATT CGT TGA CTT CTG CT</td>
<td>(Smith and Rausher, 2011)</td>
</tr>
<tr>
<td>Dfr</td>
<td>(373F) CAA CGA CCC TGA GAA CGA AG</td>
<td>(595R) ATA CAT CCA TCC TGT CAT CTT C</td>
<td>(Smith and Rausher, 2011)</td>
</tr>
<tr>
<td>Ans</td>
<td>(345F) CTG ACT ACA TTC CAG CAA CAA GTG</td>
<td>(678R) GTG TTC CRT GCC TCC AAC TTC</td>
<td>(Smith and Rausher, 2011)</td>
</tr>
<tr>
<td>Ef1-α</td>
<td>(qef1F) CCT GGT ATG GTT GTG ACC TTT G</td>
<td>(qef1R) AAC TTG ACA GCA ATG TGG GAA G</td>
<td>(Smith and Rausher, 2011)</td>
</tr>
</tbody>
</table>
Appendix S3: Gene trees of pathway genes (see Fig. 1 caption for gene abbreviations). Five taxa (Petunia hybrida, Nicotiana benthamiana, Solanum lycopersicum, Solanum tuberosum, and Iochroma cyaneum) were included for each locus in addition to copies from the blue and white forms of Iochroma calycinum. In Solanaceae, most of the pathway genes have single functional copies, except Chs and Chi (van Tunen et al., 1989; Koes et al., 1989; De Jong et al., 2004). For these two loci, our gene tree includes only the copy from the outgroup with the highest BLAST score (e.g., the Chi-A copy from Petunia, X14589). Each tip is labeled with either a GenBank accession number or the reference number from Sol Genomics (http://www.solgenomics.net), which begin with scaffold numbers (“Scf” or “PGSC”). All of the alignments and trees were deposited into TreeBase (Study S16161).

Online Supplemental 3
Click here to download Online Supplemental: S3_new_appendix_revised_122114.pdf
Appendix S4: Location of flavonoid pathway gene differences between the blue and the white morph of *Iochroma calycinum* (see Fig. 1 caption for gene abbreviations). Synonymous differences, non-synonymous differences, and indels position numbers are based on the *Petunia hybrida* reference gene (second column). For example, *Chs* mutation “3196” occurs at base pair position 3,196 in *P. hybrida* sequence X14591. †Denotes cases in which the amino acid variant present in the white morph allele was also present in one or more pigment-producing outgroup species, suggesting the difference between the blue and white alleles is not related to the phenotype. *Denotes a change in the white allele not found in any pigment-producing outgroups, and thus potentially related to the phenotype.

<table>
<thead>
<tr>
<th>Gene</th>
<th>Ref. Seq. GenBank #</th>
<th>Synonymous Differences</th>
<th>Non-synonymous Differences</th>
<th>Indel</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Chs</em></td>
<td>X14591</td>
<td>2690, 3196, 3328, 3589, 3703</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td><em>Chi</em></td>
<td>X14590</td>
<td>1411, 1867, 1870, 1873</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td><em>F3h</em></td>
<td>AF022142</td>
<td>952</td>
<td>1301†</td>
<td>0</td>
</tr>
<tr>
<td><em>F3'h</em></td>
<td>AF155332</td>
<td>1198</td>
<td>1584†</td>
<td>0</td>
</tr>
<tr>
<td><em>F3'5'h</em></td>
<td>D14588</td>
<td>663</td>
<td>638†</td>
<td>0</td>
</tr>
<tr>
<td><em>Dfr</em></td>
<td>KC140107</td>
<td>754</td>
<td>504†</td>
<td>755-787†</td>
</tr>
<tr>
<td><em>Ans</em></td>
<td>X70786</td>
<td>558</td>
<td>593†</td>
<td>0</td>
</tr>
</tbody>
</table>
Appendix S5: Dfr genotype of I. calycinum individuals (naming of individuals follows Fig. 2). Primers from Appendix S1 were used to amplify a 446 base-pair region of Dfr spanning the 33 base pair deletion in the white allele. Individual A was homozygous for the deletion in Dfr, while individuals B-F did not contain the deletion.
Appendix S6: Thin layer chromatography of white and blue *I. calycinum* leaf tissue (L) and floral tissue (F). Myricetin, quercetin, kaempferol, delphinidin, malvidin, and petunidin were included as standards. Flavonols are shown on the left panel and were visualized under UV light, while anthocyanidins are shown on the right panel and were visualized under natural light.