Article - Discoveries

Title: Convergent evolution at the pathway level: predictable regulatory changes during flower color transitions

Maximilian Larter[1], Amy Dunbar-Wallis[1], Andrea E. Berardi[12], and Stacey D. Smith[1].

[1]Department of Ecology and Evolutionary Biology, University of Colorado-Boulder, CO, USA
[2]Institute of Plant Sciences, University of Bern, Switzerland

Corresponding author: Maximilian Larter - maximilian.larter@colorado.edu

Abstract

The predictability of evolution, or whether lineages repeatedly follow the same evolutionary trajectories during phenotypic convergence remains an open question of evolutionary biology. In this study, we investigate evolutionary convergence at the biochemical pathway level and test the predictability of evolution using floral anthocyanin pigmentation, a trait with a well-understood genetic and regulatory basis. We reconstructed the evolution of floral anthocyanin content across 28 species of the Andean clade Iochrominae (Solanaceae) and investigated how shifts in pigmentation are related to changes in expression of 7 key anthocyanin pathway genes. We used phylogenetic multivariate analysis of gene expression to test for phenotypic and developmental convergence at a macroevolutionary scale. Our results show that the four independent losses of the ancestral pigment delphinidin involved convergent losses of expression of the three late pathway genes (*F3'5'h*, *Dfr* and *Ans*). Transitions between pigment types affecting floral hue (e.g. blue to red) involve changes to the expression of branching genes *F3'h* and *F3'5'h*, while the expression levels of early steps of the pathway are strongly conserved in all species. These patterns support the idea that the macroevolution of floral pigmentation follows predictable evolutionary trajectories to reach convergent phenotype space, repeatedly involving regulatory changes. This is likely driven by constraints at the pathway level, such as pleiotropy and regulatory structure.

Introduction

The prevalence of convergent evolution, i.e. the emergence of similar forms in different lineages, suggests that evolution may often be repeatable, with similar outcomes resulting from "replaying the tape of life" (Gould 1990; Orgogozo 2015). Nonetheless, the extent to which these convergent phenotypes arise via similar underlying mechanisms remains an open question (Stern and Orgogozo 2009; Rosenblum et al. 2014). In some cases, convergent phenotypes evolve through distinct genetic or developmental changes (Steiner et al. 2009; Nery et al. 2016), while in others, similar mechanisms are involved (Chan et al. 2010; Martin and Orgogozo 2013). The predictable recruitment of the same mechanisms may be explained by genetic constraints, for example if only one or a few loci have the potential to influence the phenotype, e.g. in tetrodotoxin resistance (Brodie III and Brodie Jr. 2015) and albinism of cavefish (Protas et al. 2006). By contrast, when multiple paths can lead to the same phenotype, the repeated involvement of particular loci or types of changes suggest biases, either in mutation rates or fixation probability (Stern and Orgogozo 2008; Gompel and Prud'homme 2009; Christin et al. 2010; Streisfeld and Rausher 2011; Conte et al. 2012). Still, the ability to assess such biases is limited for many traits where the underlying developmental pathway and the potential routes to convergent phenotypes are largely unknown.

Floral anthocyanin pigmentation provides a powerful system for studying phenotypic convergence and the predictability of evolution at the molecular and developmental levels. The anthocyanin branch of the large flavonoid biosynthetic pathway (hereafter the anthocyanin pathway) produces three classes of red, purple and blue pigments (fig. 1) which comprise the most common pigments determining color in flowers and fruits (Grotewold 2006). Small changes in the anthocyanin pathway (i.e. mutations at a single locus) can radically alter floral hue and/or color intensity (Koes et al. 1994; Mol et al. 1998; Rausher 2006; Rausher 2008), with important ecological consequences, e.g. pollinator attraction (Schemske and Bradshaw 1999; Hoballah et al. 2007; Sheehan et al. 2016). Upstream steps of this pathway (enzymes CHS, CHI, F3H; fig. 1) are shared across multiple flavonoid compounds (e.g., flavones and flavonols). The subsequent steps involve the so-called branching enzymes F3'H and F3'5'H, which determine

the type of anthocyanin produced, and the downstream enzymes DFR and ANS, which convert the colorless precursors into pigments. Although the entire flavonoid pathway contains dozens of genes, this group of core enzymes (fig. 1) are largely responsible for controlling metabolic flux toward anthocyanin pigments, and together determine the amount and type of anthocyanins produced (Grotewold 2006). This pathway is regulated at the level of transcription, and the regulatory genes are well known (Quattrocchio et al. 2006). In most flowering plants, the upstream steps are jointly regulated by MYB transcription factors while the downstream steps are activated by a complex of MYB, bHLH and WD40 (MBW) transcription factors (Mol et al. 1998; Stracke et al. 2007). The regulation of the branching genes varies across species and is generally not as well understood as for the structural genes (De Vetten et al. 1999; Schwinn et al. 2006; Albert et al. 2014).

In the context of the anthocyanin biosynthetic pathway, evolutionary transitions in flower color appear to occur via a small and predictable subset of possible mechanisms. For example, although losses of pigmentation could occur by inactivation along any branch of the pathway, convergent evolutionary transitions to white flowers have been exclusively linked to changes in the MYB genes that control the downstream elements of the pathway (reviewed in Sobel and Streisfeld 2013). The fixation of these regulatory mutations preserves the potential for expressing flavonoids in other tissues, where they function, for example, in protection from UV light (Ryan et al. 2001), signaling (Watkins et al. 2017) and pollen fertility (Mo et al. 1992). Similarly, repeatable patterns of genetic evolution have been described for the evolution of floral hue, where transitions from blue to red flowers frequently often involve genetic and developmental changes inferred to incur minimal pleiotropic effects (Wessinger and Rausher 2012). Nonetheless, these patterns of apparent predictability in the mechanisms of flower color evolution rely on a relatively small number of studies (e.g., five in the case of fixed gains of floral anthocyanins; Sobel and Streisfeld 2013).

Here, we use the Iochrominae clade (Solanaceae) to trace the evolution of anthocyanin gene expression during repeated shifts among flower colors and quantify the extent of convergence in pathway regulation. This group of around 40 species, native to South America, displays a remarkable variety of flower colors, largely due to variation in anthocyanin pigmentation. Previous studies have shown that the ancestral flower color in the clade is blue with mainly

delphinidin-based anthocyanins (Smith and Baum 2006; Ho and Smith 2016). Multiple lineages have lost these pigments to shift to white or yellow flowers (Ho and Smith 2016), and two lineages have evolved orange to red flowers with cyanidin or pelargonidin pigments (Berardi et al. 2016). In two cases, the genetic basis for these color transitions has been identified. In red-flowered *Iochroma gesnerioides*, the shift to pelargonidin pigments occurred via the trans-regulatory loss of *F3'h* expression, and the inactivation of *F3'5'h* through gene deletion (Smith and Rausher 2011). In the white flowered *I. loxense*, the loss of pigment production involved an R3 MYB repressor that blocks expression of multiple downstream genes (Gates et al. 2017).

Although the genetic basis is unknown for the remaining color transitions in Iochrominae, several lines of evidence point to an important role of pathway regulation. First, a study of molecular evolution of three core pathway genes (the upstream *Chi* and *F3h* and the anthocyanin specific *Ans*), revealing consistent purifying selection across the clade, including lineages that have lost floral anthocyanins (Ho and Smith 2016). Second, flavonoids, including anthocyanins, are abundant in vegetative tissue (Berardi et al. 2016) as well as fruits (Smith SD, unpublished data), which is consistent with the broad conservation of pathway function. Finally, phylogenetic analyses point to at least one regain of floral anthocyanins following loss, suggesting functionality of the structural pathway genes was maintained, allowing renewed floral expression of the pathway (Ho and Smith 2016).

In order to test the prediction that convergent flower color shifts in Iochrominae have repeatedly involved similar regulatory changes, we combined a suite of biochemical, developmental and phylogenetic approaches. First, we profiled anthocyanin pigment production across 28 species, and estimated the evolutionary history of anthocyanin pigmentation in this clade. Next, we quantified the expression of seven pathway genes and used statistical comparative methods to investigate the relationship between gene expression and pigment phenotypes. Given the likely pleiotropic effects of changes to the early genes of the pathway on other flavonoids, we predicted that the shifts from blue to white flowers involve losses of expression of downstream elements of the pathway. By contrast, we expected that transitions among the blue, purple and red pigments involve differential regulation of the branching enzymes. We further tested these predictions using multivariate approaches to cluster the species in gene expression space. Together our

analyses reveal a tight relationship between floral expression of pathway genes, anthocyanin pigment production, and flower color evolution in Iochrominae.

Results:

Evolutionary shifts in pigment production

Consistent with previous work in this system, species within Iochrominae produce all three classes of anthocyanins, and the types present are closely associated with the flower color phenotype. The blue-flowered species produce delphinidin-derived anthocyanins (fig. 2). While some blue-flowered Iochrominae make these pigments exclusively, most also synthesize small amounts (<15%) of cyanidin-derived anthocyanins (supplementary figure S1). The two red-flowered sister species, *I. gesnerioides* and *I. fuchsioides* produce the red anthocyanidin pelargonidin, and the orange-red flowered *I. edule* is the only species which produces exclusively cyanidin (supplementary figure S1). As predicted, six white or yellow flowered species produce no detectable anthocyanidins (supplementary figure S1).

Analyses of trait evolution with stochastic mapping supported strong directionality in transitions among pigment types (fig. 2), similar to previous work in this and other clades (reviewed in Rausher 2008). The most common transitions were inferred to be convergent losses of ancestral delphinidin production, with stochastic mapping estimating an average of five losses across the clade (fig. 2; supplementary table S3). The persistence of just four unpigmented lineages suggests that one of these transitions was reversed. The earliest loss likely occurred c. 3 million years ago on the branch leading to the "A" clade containing the unpigmented species *I. ellipticum*, *I. salpoanum*, *I. confertiflorum*, *Acnistus arborescens* (fig. 2A). Additional more recent shifts involved *I. lehmannii*, *I. loxense* and *D. solanacea*. Stochastic mapping estimated that transitions from unpigmented to producing cyanidin occurred at least twice in the clade's history (fig. 2B), although only one extant species, *I. edule*, produces this pigment exclusively. Its position, nested within the unpigmented "A" clade, suggests that it has recently regained

floral anthocyanin expression. Shifts to floral pelargonidin production are inferred to be rare, likely occurring only once in the common ancestor of the two red-flowered species. Several transitions (e.g., pelargonidin to cyanidin, delphinidin to cyanidin) were never observed among the stochastic maps (fig. 2; supplementary table S3), indicating they are unlikely to have occurred during the history of Iochrominae.

Evolution of anthocyanin pathway gene expression

Our comparative analysis of gene expression revealed distinct evolutionary patterns across pathway genes. Expression levels of the upstream genes (e.g. *Chs*, *Chi*, and *F3h*) were similar across Iochrominae regardless of pigment composition, with log fold-change values mostly between -1 and +1, i.e. transcript concentrations within an order of magnitude of the reference blue-flowered *I. cyaneum* (fig. 3A). There were a few exceptions; for example, *Chs* had low expression (log fold-change of -3) in the white-flowered *I. ellipticum* and *F3h* showed moderate down-regulation (log fold-changes of around -1.5) in several species which lack floral anthocyanins (*I. ellipticum*, *I. confertiflorum* and *D. solanacea*; fig. 3A). The phylogenetic ANOVA suggested that despite these cases, species without floral anthocyanins did not have significantly lower expression for *Chs*, *Chi*, or *F3h* than their pigmented relatives (fig. 3B). This pattern of conservation of upstream gene expression in lineages that have lost floral anthocyanins is consistent with their continued production of other floral flavonoids, in particular flavonols (supplementary figure S1). These compounds are typically as abundant in flower tissues as anthocyanins and remain abundant in most white- and yellow-flowered species (Berardi et al. 2016; Larter et al., unpubl. data).

In contrast with the evolutionary conservation of expression observed for the upstream genes, the other elements of the pathway showed strong color-specific patterns of expression. In the two red-flowered species, which produce only pelargonidin, we observed extremely low expression of the branching genes *F3'h* and *F3'5'h* (with log fold change of -3 or lower, or c. 1000 times fewer transcripts than in the reference blue species), while the remaining steps of the pathway had comparable expression levels to blue-flowered species (fig. 3B). In the unpigmented lineages, the two downstream genes *Dfr* and *Ans* showed similarly strong down-regulation (log fold change of around -3), along with the branching gene *F3'5'h* (fig. 3B). The retention of *F3'h*

expression in these lineages is likely tied to the prevalence of quercetin, a flavonol that requires F3'H activity (fig. 1), in their floral tissue (supplementary fig. S1). The cyanidin-producing species *I. edule*, which is nested within the unpigmented A clade, has recovered expression of *F3'5'h*, *Dfr* and *Ans*, although we were unable to assess statistical significance with only a single tip (fig. 3A, B).

Co-expression patterns of the anthocyanin pathway genes

Given that anthocyanin pathway genes are known to be jointly regulated by MBW complexes, we used multivariate approaches to quantify co-expression and relate these patterns to floral pigmentation variation. The phylogenetic PCA (pPCA) of expression levels grouped species by primary pigment types, and thus flower color (fig. 4A). The first two PC axes explained around 90% of the variance (fig. 4A). The first PC axis divides the pigmented and unpigmented species (fig. 4A), with high loadings of expression levels of *F3'5'h*, *Dfr* and *Ans* (fig.4B). The second PC axis separates the red pelargonidin-producing species from the rest of Iochrominae, with high loadings of expression of branching genes *F3'h* and *F3'5'h*. The two centermost points, slightly detached from the blue-flowered cluster, are the red-orange *I. edule* and purple-spotted *S. punctata*. The separation of *I. edule* might be expected as it is the only species producing primarily cyanidin and has high *F3'h* expression but relatively low *F3'5'h* expression. *S. punctata* produces primarily delphinidin, but expression levels for pathway genes are generally lower than the other blue and purple taxa, perhaps due to the spotted phenotype.

The pattern of grouping in gene expression space by pigment type was confirmed by k-means clustering. With k = 3, we recovered the same three clusters: the species that 1) make no anthocyanins, 2) those that make pelargonidin, and 3) those that make either delphinidin- or cyanidin-based anthocyanins. This latter cluster is broken when k = 4, but not into groups clearly reflecting different pigmentation phenotypes. Both sub-clusters contain blue-flowered species predominantly producing delphinidin, however they are slightly differentiated along PC1 of the pPCA of gene expression (supplementary figure S3).

In addition to identifying phenotypic clusters of species, the pPCA highlights the high levels of correlated expression among groups of pathway genes (fig.4C). The upstream genes *Chs* and *Chi* are significantly co-expressed with *F3h* across the phylogeny ($r^2 = 0.55$ and 0.36; $p < 0.001$ and

0.02, respectively; supplementary table S4), and these genes show no significant correlations with the remaining genes. A similar pattern is exhibited by *F3'5'h* and the downstream genes *Dfr* and *Ans* ($r^2 = 0.49 – 0.87$, $p < 0.001$). Intriguingly, the expression of the branching gene *F3'h* shows no significant correlations with the expression of any of the other pathway genes. This result aligns with previous crossing studies indicating an independent (unidentified) trans-regulatory factor for this branch of the pathway in *Iochroma* (Smith and Rausher 2011).

Discussion:

Regulatory mechanisms drive evolution of floral pigmentation intensity

Our results show that convergent losses of floral anthocyanin production are driven at the pathway level by convergence in the expression of pathway genes in Iochrominae. Ancestral delphinidin production has been lost independently in four white- and yellow-flowered lineages, and in each case these transitions are associated with 1000-fold decrease in expression of the branching gene *F3'5'h* and the downstream genes *Dfr* and *Ans*. These repeatable, coordinated expression shifts at the macroevolutionary level are striking given the wide range of mechanisms known to generate unpigmented flowers in natural populations and horticultural varieties (Streisfeld and Rausher 2011). For example, both regulatory and structural mutations at the *Chs* loci have been found to underlie white morphs of some normally pigmented species (Habu et al. 1998; Coberly and Rausher 2003; Dick et al. 2011). This contrast between the range of mechanisms that can give rise to the derived phenotype and those that underlie species level transitions suggests bias towards particular genetic paths, possibly due to mutational target size or pleiotropy (Gompel and Prud'homme 2009). Given that flavonoids play important roles in physiological processes throughout the plant (Winkel-Shirley 2001), pleiotropic effects of mutations that change flower color are particularly likely to affect their probability of fixation over short and long evolutionary timescales (Coberly and Rausher 2008).

Although we have not identified the exact genetic changes that caused the coordinated down-regulation of *F3'5'h*, *Dfr* and *Ans* in the unpigmented lineages, several lines of evidence point to a trans-regulatory mechanism. First, the tight co-expression of these three genes across the phylogeny suggests that they belong to a single regulatory module. Indeed, *Dfr* and *Ans* are

known to be jointly regulated by an MBW transcription factor complex in a wide range of taxa (Holton and Cornish 1995; Mol et al. 1998). It appears that *F3'5'h* belongs to this regulatory module in Iochrominae (Gates et al. 2017) as it does many but not all eudicots (De Vetten et al. 1999; Butelli et al. 2008; Albert et al. 2014). Furthermore, losses of pigmentation in other lineages also tend to be driven by trans-regulatory mutations (reviewed in Wessinger and Rausher 2012; Sobel and Streisfeld 2013). These often involve the MBW complex (e.g., Schwinn et al. 2006; Hoballah et al. 2007) and lead to loss of expression in the downstream genes (Whittall et al. 2006; Cooley et al. 2011).

Among the trans-regulators of the downstream genes, the MYB genes have emerged as the dominant players in flower pigment evolution. Indeed, changes in these transcription factors have been found to underlie every case of evolutionary gain or loss of anthocyanins in which the genetic basis has been dissected (reviewed in Sobel and Streisfeld 2013). This pattern may relate to the fact that MYBs are often highly tissue-specific, and determine the transcriptional targets for the MBW complex (Ramsay and Glover 2005). Thus either cis- or coding mutations at these loci will have limited pleiotropic effects, relative to mutations in other elements of the complex or in the pathway itself (Streisfeld and Rausher 2011). In addition to this activating role as part of the MBW complex, MYB proteins can act at transcriptional repressors of the anthocyanin pathway (Aharoni et al. 2001; Albert et al. 2014; Liu et al. 2015). Most transitions to white or yellow flowers involve loss-of-function mutations in the MYB activators (Sobel and Streisfeld 2013) although in *Iochroma loxense*, a gain-of-function in an MYB repressor was responsible (Gates et al. 2017). Given the similarity we observe in expression profiles, with strongest downregulation in *F3'5'h*, *Dfr*, and *Ans*, changes at this same MYB repressor may underlie the losses of floral pigmentation in other Iochrominae lineages. Investigating this possibility will require a multi-faceted approach as the rapid evolution of MYB genes and their immense copy number (over 100 in *Iochroma*; Gates et al. 2016) complicates standard candidate gene experiments.

Understanding the mechanisms responsible for the repeated losses of floral anthocyanins in Iochrominae has important implications for potential regains of pigmentation. Given that the structure of the pathway appears conserved (Ho and Smith 2016) and remains active in non-floral tissues (Berardi et al. 2016; S.D. Smith, unpubl. data), an evolutionary gain of floral

anthocyanins likely requires only re-activation of the pathway in floral tissue. This could occur via downregulation of a repressor (Yuan et al 2013) or upregulation of an activator (Streisfeld et al. 2013). Indeed, this may be the case with *I. edule*, which has regained floral anthocyanins following a loss roughly five million years ago. Compared to its white and yellow relatives in the A clade, this species has elevated *F3'5'h*, *Dfr* and *Ans* expression, a change that could be achieved by their shared regulator. Nonetheless, *I. edule* did not regain the ancestral blue delphinidin pigmentation but instead produces purple-pink cyanidin. This unique feature may represent incomplete reversal, given that the expression of *F3'5'h* (along with *Dfr* and *Ans*) is still lower than all of the blue-flowered species (fig. 3B). Alternately, additional changes in the pathway following the loss may have altered the consequences for re-activation of the *F3'5'h-Dfr-Ans* module. For example, DFR may have shifted its specificity to the less hydroxylated DHQ as cyanidin is more common in vegetative tissue (Wessinger and Rausher 2012). The functionality of *F3'5'h* may have also degraded since its products (myricetin and delphinidin) do not appear in vegetative tissues of A-clade species (Berardi et al. 2016; supplementary figure S1).

Genetic targets of shifts in floral hue

In contrast to shifts in the intensity of pigmentation, shifts in floral hue due to transitions between anthocyanin pigments are driven by changes to the branching genes *F3'h* and *F3'5'h*. The relative activity of these enzymes determines the abundance of the pigment precursors (the three dihydroflavonols, fig. 1) and in turn, the flux toward the three classes of pigments. Thus, shifts in *F3'h* and *F3'5'h* expression can alter which branch of the anthocyanin pathway is most active thereby changing the main floral anthocyanin pigment (Rausher 2008). Previous work in *Iochroma* demonstrated that the loss of both F3'H and F3'5'H activity is required for the complete transition from the ancestral delphinidin production to the derived pelargonidin production (Smith and Rausher 2011). In the red-flowered *I. gesnerioides*, this was achieved by flower-specific loss of *F3'h* expression due to a trans-regulatory change along with degeneration of the *F3'5'h* gene. Here we find that the downregulation of *F3'h* is also found in the red-flowered sister species *I. fuchsioides*. Both species produce trace amounts of cyanidin in flowers and leaves and substantial amounts of quercetin in leaves, indicating that F3'H remains functional (Berardi et al. 2016). Additional studies will be required to determine whether *I.*

fuchsioides, like *I. gesnerioides*, lacks a functional *F3'5h* gene. Even if *I. fuchsioides* has retained *F3'5'h*, its expression is essentially undetectable in flowers (fig. 3A). In this case, the initial loss of expression of *F3'5'h* in the common ancestor of *I. fuchsioides* and *I. gesnerioides* may have set the stage for the later degeneration of the coding sequence in one species (*I. gesnerioides*).

Although the two red-flowered *Iochroma* are the result of a single evolutionary transition, the patterns we observe mirror those found in shifts in hue in other taxa. For example, shifts from delphinidin to cyanidin or pelargonidin production consistently involve coding and/or regulatory changes at *F3'5'h* (Nakatsuka et al. 2006; Takahashi et al. 2010; Hopkins and Rausher 2011; Wessinger and Rausher 2014) and exclusively regulatory changes at *F3'h* (Streisfeld and Rausher 2009; Des Marais and Rausher 2010). Given these patterns, Wessinger and Rausher (2012; 2014) posited that *F3'5'h* is a "disposable" gene while *F3'h* is constrained due to its functions beyond flower color. Indeed, the flavonol quercetin, which requires F3'H activity has higher antioxidant and antiradical activity (Burda and Oleszek 2001; Hirano et al. 2001) and provides better UV protection (Ryan et al. 2001; Ryan et al. 2002) than myricetin, which requires F3'5'H activity. Thus, the pleiotropic cost of losing *F3'h* function is likely higher than that for *F3'5'h* (Streisfeld and Rausher 2011).

Macroevolutionary signatures of regulatory structure

Given the complex patterns of pigment production across tissues and in response to environmental cues, the anthocyanin pathway has long served as a model for studying the spatial and temporal regulation of metabolic pathways. Early studies of anthocyanin regulation revealed that steps in the pathway are often coordinately controlled by shared transcription factors, but that these relationships can differ across taxa. For instance, the anthocyanin biosynthetic genes are divided into two co-expressed clusters ('early' and 'late' genes) in Eudicots, but which genes belong to those clusters varies across taxa (e.g., Martin et al. 1991; Quattrocchio 1993; Pelletier and Winkel-Shirley 1996). The discovery of these regulatory relationships has and continues to rely on detailed studies of transcription factor and structural gene expression in experimental model systems (e.g., Quattrocchio 1993; Butelli et al. 2008).

Our analyses suggest that comparative studies of gene expression at the pathway level may provide a complementary approach to detecting co-regulated modules. The phylogenetic PCA recovered the strong co-expression of *F3'5'h*, *Dfr* and *Ans*, a group that corresponds to 'late' genes in the model system *Petunia* (Mol et al. 1998; Albert et al. 2014). The traditional 'early' genes in *Petunia* (*Chs*, *Chi*, *F3h*; Quattrocchio 1993) emerge here as a loose cluster, likely reflecting the redundant regulation of these steps by multiple MYB transcription factors (Quattrocchio et al. 2006). Finally, the expression of *F3'h*, which is controlled by a separate trans-regulator (Smith and Rausher 2011), is decoupled from other parts of the pathway. The close correspondence between the correlation structure of expression variation in these genes and the regulatory structure known from plant model systems likely reflects the predominance of trans-regulatory mutations in flower color evolution. If indeed the fluctuation in pigment intensity across the clade arises from changes in shared MYB transcription factors, multiple genes will show highly correlated variation at a phylogenetic scale. Equally, gene that experience independent control will lack such correlated variation.

Conclusions and future directions

Evolutionary constraints at the pathway level restrict the available trajectories between floral pigmentation phenotypes, leading to strong predictability of the mechanisms involved in phenotypic transitions. The predictable involvement of regulatory changes in branching and downstream steps of the pathway likely arises from selection for maintaining the production of fitness-related flavonoid compounds in other tissues. At the scale of an entire clade, our study supports key features of flower color evolution that have emerged from a growing number of studies in model systems. In particular, we show that evolutionary losses of anthocyanins predictably involve reversible regulatory shut-down of the late pathway genes. Moreover, changes of floral hue require reduced activity of the branching genes, although the underlying mechanism varies due to pleiotropic and regulatory constraints. One explanation for the repeatability of evolution is that while many changes can produce the same phenotypic outcome (e.g. a coding change to an early pathway gene inducing loss of floral pigments), some are filtered out through selection within populations, due to their negative pleiotropic effects (Gompel and Prud'homme 2009). Rare unpigmented morphs are found in multiple blue-flowered Iochrominae species, with previous research showing that in one case, this intraspecific variation

is due to a loss-of-function mutation to the late gene *Dfr* (Coburn et al. 2015). Further work in several polymorphic species could shed some light on the fixation advantage of minimally pleiotropic mutations at the population level explaining the bias in mechanisms visible at longer evolutionary timescales (Stern and Orgogozo 2008; Gompel and Prud'homme 2009).

Methods:

Study system

Iochrominae are distributed across Western South America, from Columbia and Peru to Brazil and northern Argentina. Flower buds were collected from natural populations for 19 of the species. The remaining nine species were grown from seed at the UC-Boulder EBIO Greenhouses. Greenhouse plants were placed outside two to three weeks prior to flower bud collection to provide natural light (without UV filtering) to maximize the stimulation of the anthocyanin pathway. We obtained samples from at least 3 individuals per species, with the exception of *I. ellipticum* from the Galapagos Islands, for which we only have one individual, and *Vassobia breviflora* for which we only have 2 individuals. Vouchers and accession numbers are available in supplementary table S1.

Anthocyanin identification and quantification

We separated and quantified anthocyanidins (aglycones of anthocyanins) and closely related flavonoids (flavones, flavonols) from floral tissue using High Performance Liquid Chromatography (HPLC). Following Berardi et al. (2016), we extracted anthocyanidins and other flavonoids from silica-dried floral tissue using the acid hydrolysis procedure of Harborne (1998) with minor modifications. Dried extracts were re-suspended in methanol with 1% HCl (v/v), and 10 μL were injected and analyzed using an Agilent Infinity 1260 Infinity Series HPLC system with a 100 x 4.6 mm Chromolith Performance RP-18 endcapped column (EMD Chemicals, Darmstadt, Germany) and a photodiode array detector (PDA; scanning from 200 to 600 nm at a step of 2 nm). Separation of the compounds followed the protocols of Berardi et al. (2016). Peaks were detected at 520 nm for anthocyanidins and 365 nm for flavones and flavonols and identified by comparison of retention times and UV-VIS absorbance spectra with

commercial standards (Extrasynthese, Genay, France). Chromatograms were visualized using Agilent ChemStation for LC 3D Systems (Rev. B0403, Waldbronn, Germany). Concentrations were scaled by tissue weight and were estimated by comparing the area of each chromatogram peak to those from dilution series of the corresponding standards. Two of the anthocyanidins (cyanidin and delphinidin) have methylated derivatives, and these were summed for each to obtain the total amount of cyanidin- and delphinidin-derived anthocyanins, i.e. matching each branch of the pathway (fig. 1). Relative amounts of the three classes (pelargonidin-, cyanidin-, and delphinidin-derived anthocyanins) were calculated to identify the principal pigment and minor pigments in each species. HPLC data for nine species in this dataset was previously published in Berardi et al. (2016).

Quantification of anthocyanin pathway gene expression

Floral buds were collected at a stage roughly equivalent to *Petunia* bud Stage 5 (Pollak et al. 1993), at which time both the early and late parts of the pathway are expressed (Smith and Rausher 2011). Tissue was preserved in RNAlater (Qiagen, Valencia, CA), or flash frozen if collected from a greenhouse plant. The protocol for measuring gene expression followed previous work (Coburn et al. 2015). Briefly, total RNA was extracted using the Spectrum kit (Sigma-Aldrich, St. Louis, Missouri, USA), and on-column DNase digestion (Qiagen) was used to remove genomic DNA. Following reverse transcription with the SuperScriptII Reverse Transcription kit (Life Technologies, CA), newly synthesized cDNA was used as a template for amplification of the gene fragments using previously designed primers for seven pathway genes *Chs*, *Chi*, *F3h*, *F3'h*, *F3'5'h*, *Dfr* and *Ans* and a reference "housekeeping" gene, the elongation factor EF1-α. The qPCR mix contained 10µL of SYBR green dye (Finnzymes, Espoo, Finland), 0.8 µL primers, 0.4 µL reference dye, and 0.8 µL cDNA (2 ng), and 8 µL H_2O. The qPCR program proceeded with 10 minutes at 95°C, then 40-50 cycles of 95°C for 20 seconds followed by 55°C for 60 seconds, ending with a dissociation stage, at 72°C for 10 minutes. Three replicate reactions were run for each individual, and the resulting amplification curves were imported into R for analysis (R Core Team 2017). From the log-linear phase of these amplification curves, we obtained the parameters of the relative expression ratio model (Pfaffl 2001), i.e. the crossing points and amplification efficiency (Peirson et al. 2003). We normalized expression of each of the pathway genes by the expression of the reference gene, *ef1*-α (following Pfaffl 2001) and

calculated $\log_{10}$ fold-change in gene expression relative to a single control individual of the blue-flowered *I. cyaneum*. This blue color is similar to other blue-flowered *Iochroma* species and is thought to be the ancestral state in the clade, and we used the same individual of *I. cyaneum* as a reference in previous gene expression studies (Smith and Rausher 2011). In the comparison of any individual's expression to this control individual, a fold-change of 0 means equal expression in sample and control, whereas positive or negative fold-change values indicate higher and lower expression relative to the control, respectively. Each unit of fold-change represents an order-of-magnitude change in the initial concentration of gene transcripts in the sample, again relative to the control.

Stochastic mapping of pigment evolution

In order to understand evolutionary shifts in floral pigmentation, we reconstructed the evolution of pigment production in Iochrominae using maximum likelihood with *ace* {ape} (Paradis et al. 2004) and stochastic mapping (SM) using *make.simmap* {*phytools*} (Revell 2012) in R version 3.4.2 (R Core Team 2017). SM is a Bayesian approach that simulates possible histories ("maps") of discrete character evolution on a phylogeny based on a model of character evolution (Huelsenbeck et al. 2003). Here we employed an all-rates different model, where transition rates among all states are allowed to vary independently. The simulations can be summarized to give the posterior probability of each state at each node, as well as posterior distributions of the frequency of events, i.e. the number of transitions to and from each character state across the tree. Using HPLC pigment profiling, we scored taxa as based on their primary pigment type (delphinidin, cyanidin, pelargonidin, or none), giving us four possible states. Based on our maximum likelihood reconstruction (supplementary figure S2) in addition to biochemical and phylogenetic analyses at the family level (Ng and Smith 2016), we fixed the root state to delphinidin for SM and completed 1000 simulations of character evolution. For these and subsequent analyses, we used the previously published, time-calibrated, maximum clade credibility tree for Iochrominae (Smith and Goldberg 2015).

Comparative analyses of gene expression

We used phylogenetic comparative methods to test the relationship between levels of gene expression and pigment type. We removed *I. ellipticum* from the dataset because we only have one individual. For each of the seven pathway genes, we implemented a phylogenetic ANOVA using *gls* {*nlme*} (Pinheiro et al. 2014). For this analysis, we grouped together all species based on which pigments they make (delphinidin, pelargonidin, and none) in order to test how gene expression varies by pigment type. With only one species (*I. edule*) producing cyanidin as the primary pigment, we excluded this species. In order to account for phylogenetic non-independence, we incorporated the correlation structure based on the phylogeny, with the strength of signal modeled with Pagel's λ (Pagel 1999) using *corPagel* {ape}. This correlation structure, with λ estimated from the data, produced a better fit than Brownian motion or the Ornstein-Uhlenbeck model (results not shown).

Next, we examined patterns of co-expression across pathway genes in the context of phenotypes using multivariate comparative methods. We implemented phylogenetic PCA (pPCA) because standard PCA in comparative settings can mislead inferences (Uyeda et al. 2015). We used the function *phyl.pca* {phytools} (Revell 2009), again with Pagel's λ to account for the strength of phylogenetic signal. To estimate pairwise gene expression relationships, we extracted the phylogenetic correlation matrix using code from the *phyl.pca* function. These pairwise correlations are equivalent to phylogenetic generalized least squares (or PGLS) using Pagel's λ. We then computed p-values with the Bonferonni correction for multiple testing, using *corr.p* {psych}. Given that species with the same pigment type might be expected to group in gene expression space, we detected clusters using k-means clustering (supplementary figure S3), assuming 2 to 4 clusters with the {stats} package in R (R Core Team 2017). This method aims to minimize within-cluster variance while maximizing differences among clusters for a given number of clusters.

Acknowledgments

This work was supported by the National Science Foundation (Grant DEB1355518 to S.D. Smith). We thank the Servicio Nacional Forestal y de Fauna Silvestre (SERFOR) of Peru for permission to carry out investigation under permits Nº 0259-2016 and N° 041-2017 and the

Servicio Nacional de Áreas Naturales Protegidas (SERNANP) under permit N° 004-2013. We are also grateful to the Direccion Provincial del Ambiente Pichincha, Ecuador for permission to collect under permit Nº 05-2012-IC-FAU-DPAP-MA. We thank members of the Smith lab for constructive feedback and helping to improve the manuscript.

References

Aharoni A, De Vos CHR, Wein M, Sun Z, Greco R, Kroon A, Mol JNM, O'Connell AP. 2001. The strawberry FaMYB1 transcription factor suppresses anthocyanin and flavonol accumulation in transgenic tobacco. Plant J. 28:319–332.

Albert NW, Davies KM, Lewis DH, Zhang HB, Montefiori M, Brendolise C, Boase MR, Ngo H, Jameson PE, Schwinn KE. 2014. A Conserved Network of Transcriptional Activators and Repressors Regulates Anthocyanin Pigmentation in Eudicots. Plant Cell 26:962–980.

Berardi AE, Hildreth SB, Helm RF, Winkel BSJ, Smith SD. 2016. Evolutionary correlations in flavonoid production across flowers and leaves in the Iochrominae (Solanaceae). Phytochemistry 130:119–127.

Brodie III ED, Brodie Jr. ED. 2015. Predictably convergent evolution of sodium channels in the arms race between predators and prey. Brain. Behav. Evol. 86:48–57.

Burda S, Oleszek W. 2001. Antioxidant and Antiradical Activities of Flavonoids. J. Agric. Food Chem. 49:2774–2779.

Butelli E, Titta L, Giorgio M, Mock H-P, Matros A, Peterek S, Schijlen EGWM, Hall RD, Bovy AG, Luo J, et al. 2008. Enrichment of tomato fruit with health-promoting anthocyanins by expression of select transcription factors. Nat. Biotechnol. 26:1301–1308.

Chan YF, Marks ME, Jones FC, Jr GV, Shapiro MD, Brady SD, Southwick AM, Absher DM, Grimwood J, Schmutz J, et al. 2010. Adaptive evolution of pelvic reduction in sticklebacks by recurrent deletion of a *Pitx1* enhancer. Science. 327(5963):302–305.

Christin PA, Weinreich DM, Besnard G. 2010. Causes and evolutionary significance of genetic convergence. Trends Genet. 26:400–405.

Coberly LC, Rausher MD. 2003. Analysis of a chalcone synthase mutant in *Ipomoea purpurea* reveals a novel function for flavonoids: Amelioration of heat stress. Mol. Ecol. 12:1113–1124.

Coberly LC, Rausher MD. 2008. Pleiotropic effects of an allele producing white flowers in Ipomoea purpurea. Evolution. 62:1076–1085.

Coburn RA, Griffin RH, Smith SD. 2015. Genetic basis for a rare floral mutant in an andean species of solanaceae. Am. J. Bot. 102:264–272.

Conte GL, Arnegard ME, Peichel CL, Schluter D. 2012. The probability of genetic parallelism and convergence in natural populations. Proc. R. Soc. B Biol. Sci. 279:5039–5047.

Cooley AM, Modliszewski JL, Rommel ML, Willis JH. 2011. Gene duplication in mimulus underlies parallel floral evolution via independent trans-regulatory changes. Curr. Biol. 21:700–704.

Dick CA, Buenrostro J, Butler T, Carlson ML, Kliebenstein DJ, Whittall JB. 2011. Arctic mustard flower color polymorphism controlled by petal-specific downregulation at the threshold of the anthocyanin biosynthetic pathway. PLoS One 6.

Gates DJ, Olson BJSC, Clemente TE, Smith SD. 2017. A novel R3 MYB transcriptional repressor associated with the loss of floral pigmentation in Iochroma. New Phytol. 217: 1346-1356.

Gates DJ, Strickler SR, Mueller LA, Olson BJSC, Smith SD. 2016. Diversification of R2R3-MYB Transcription Factors in the Tomato Family Solanaceae. J. Mol. Evol. 83:26–37.

Gompel N, Prud'homme B. 2009. The causes of repeated genetic evolution. Dev. Biol. 332:36–47.

Gould SJ. 1990. Wonderful life: the Burgess Shale and the nature of history. New York, NY: WW Norton & Company

Grotewold E. 2006. The Science of Flavonoids. (Grotewold E, editor.). New York, NY: Springer Science & Business Media

Habu Y, Hisatomi Y, Iida S. 1998. Molecular characterization of the mutableflakedallele for flower variegation in the common morning glory. Plant J. 16:371–376.

Harborne JB. 1998. Phytochemical methods: A guide to modern techniques of plant analysis. 3rd ed. London, UK: Chapman and Hall

Hirano R, Sasamoto W, Matsumoto A, Itakura H, Igarashi O, Konda K. 2001. Antioxidant Ability of Various Flavonoids against DPPH Radicals and LDL Oxidation. J. Nutr. Sci. Vitaminol. (Tokyo). 47:357–362.

Ho WW, Smith SD. 2016. Molecular evolution of anthocyanin pigmentation genes following losses of flower color. BMC Evol. Biol. 16:98.

Hoballah ME, Gubitz T, Stuurman J, Broger L, Barone M, Mandel T, Dell'Olivo a, Arnold M, Kuhlemeier C. 2007. Single gene-mediated shift in pollinator attraction in Petunia. Plant Cell 19:779–790.

Holton TA, Cornish EC. 1995. Genetics and Biochemistry of Anthocyanin Biosynthesis. Plant Cell 7:1071.

Hopkins R, Rausher MD. 2011. Identification of two genes causing reinforcement in the Texas wildflower Phlox drummondii. Nature 469:411–414.

Huelsenbeck JP, Nielsen R, Bollback JP. 2003. Stochastic mapping of morphological characters. Syst. Biol. 52:131–158.

Koes RE, Quattrocchio F, Mol JNM. 1994. The Flavonoid Biosynthetic Pathway in Plants: Function and Evolution. BioEssays 16:123–132.

Liu J, Osbourn A, Ma P. 2015. MYB transcription factors as regulators of phenylpropanoid metabolism in plants. Mol. Plant 8:689–708.

Des Marais DL, Rausher MD. 2010. Parallel evolution at multiple levels in the origin of hummingbird pollinated flowers in Ipomoea. Evolution. 64:2044–2054.

Martin A, Orgogozo V. 2013. The loci of repeated evolution: A catalog of genetic hotspots of phenotypic variation. Evolution. 67:1235–1250.

Martin C, Prescott A, Mackay S, Bartlett J, Vrijlandt E. 1991. Control of Anthocyanin

Biosynthesis in Flowers of Antirrhinum-Majus. Plant J. 1:37–49.

Mo Y, Nagel C, Taylor LP. 1992. Biochemical complementation of chalcone synthase mutants defines a role for flavonols in functional pollen. Proc Natl Acad Sci USA 89:7213–7217.

Mol J, Grotewold E, Koes R. 1998. How genes paint flowers and seeds. Trends Plant Sci. 3:212–217.

Nakatsuka T, Nishihara M, Mishiba K, Hirano H, Yamamura S. 2006. Two different transposable elements inserted in flavonoid 3′,5′-hydroxylase gene contribute to pink flower coloration in Gentiana scabra. Mol. Genet. Genomics 275:231–241.

Nery MF, Borges B, Dragalzew AC, Kohlsdorf T. 2016. Selection on different genes with equivalent functions: the convergence story told by Hox genes along the evolution of aquatic mammalian lineages. BMC Evol. Biol. 16:113.

Ng J, Smith SD. 2016. Widespread flower color convergence in Solanaceae via alternate biochemical pathways. New Phytol. 209:407–417.

Orgogozo V. 2015. Replaying the tape of life in the twenty-first century. Interface Focus 5:20150057.

Pagel M. 1999. Inferring the historical patterns of biological evolution. Nature 401:877–884.

Paradis E, Claude J, Strimmer K. 2004. APE: Analyses of Phylogenetics and Evolution in R language. Bioinformatics 20:289–290.

Peirson SN, Butler JN, Foster RG. 2003. Experimental validation of novel and conventional approaches to quantitative real-time PCR data analysis. Nucleic Acids Res. 31:e73.

Pelletier MK, Winkel-Shirley B. 1996. Analysis of flavanone 3-hydroxylase in Arabidopsis seedlings. Plant Physiol. 111:339–345.

Pfaffl MW. 2001. A new mathematical model for relative quantification in real-time RT-PCR. Nucleic Acids Res. 29:16–21.

Pinheiro J, Bates D, DebRoy S, Sarkar D. 2014. R Core Team (2014) nlme: linear and nonlinear mixed effects models. R package version 3.1-117. Available at http://CRAN. R-project.org/package=nlme.

Pollak PE, Vogt T, Mo Y, Taylor LP. 1993. Chalcone Synthase and Flavonol Accumulation in Stigmas and Anthers of Petunia hybrida. Plant Physiol. [Internet] 102:925–932. Available from:http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=158865&tool=pmcentrez&rendertype=abstract

Protas ME, Hersey C, Kochanek D, Zhou Y, Wilkens H, Jeffery WR, Zon LI, Borowsky R, Tabin CJ. 2006. Genetic analysis of cavefish reveals molecular convergence in the evolution of albinism. 38:107–111.

Quattrocchio F. 1993. Regulatory Genes Controlling Anthocyanin Pigmentation Are Functionally Conserved among Plant Species and Have Distinct Sets of Target Genes. Plant Cell Online 5:1497–1512.

Quattrocchio F, Baudry A, Lepiniec L, Grotewold E. 2006. The regulation of flavonoid biosynthesis. In: Grotewold E, editor. The science of flavonoids. New York: Springer New York. p. 97–122.

R Core Team. 2017. R: A Language and Environment for Statistical Computing. Available from: https://www.r-project.org/

Ramsay NA, Glover BJ. 2005. MYB-bHLH-WD40 protein complex and the evolution of cellular diversity. Trends Plant Sci. 10:63–70.

Rausher MD. 2006. The evolution of flavonoids and their genes. In: Grotewold E, editor. The science of flavonoids. New York: Springer New York. p. 97–122.

Rausher MD. 2008. Evolutionary Transitions in Floral Color. Int. J. Plant Sci. 169:7–21.

Revell LJ. 2009. Size-correction and principal components for interspecific comparative studies. Evolution. 63:3258–3268.

Revell LJ. 2012. phytools: An R package for phylogenetic comparative biology (and other things). Methods Ecol. Evol. 3:217–223.

Rosenblum EB, Parent CE, Brandt EE. 2014. The Molecular Basis of Phenotypic Convergence. Annu. Rev. Ecol. Evol. Syst. 45:203–226.

Ryan KG, Swinny EE, Markham KR, Winefield C. 2002. Flavonoid gene expression and UV photoprotection in transgenic and mutant Petunia leaves. Phytochemistry 59:23–32.

Ryan KG, Swinny EE, Winefield C, Markham KR. 2001. Flavonoids and UV photoprotection in Arabidopsis mutants. Zeitschrift für Naturforsch. C 56:745–754.

Schemske DW, Bradshaw HD. 1999. Pollinator preference and the evolution of floral traits in monkeyflowers (Mimulus). Proc. Natl. Acad. Sci. 96:11910–11915.

Schwinn K, Venail J, Shang Y, Mackay S, Alm V, Butelli E, Oyama R, Bailey P, Davies K, Martin C. 2006. A small family of MYB-regulatory genes controls floral pigmentation intensity and patterning in the genus Antirrhinum. Plant Cell 18:831–851.

Sheehan H, Moser M, Klahre U, Esfeld K, Dell'olivo A, Mandel T, Metzger S, Vandenbussche M, Freitas L, Kuhlemeier C. 2016. MYB-FL controls gain and loss of floral UV absorbance, a key trait affecting pollinator preference and reproductive isolation. Nat. Genet. 48:159–166.

Smith SD, Baum DA. 2006. Phylogenetics of the florally diverse Andean clade Iochrominae (Solanaceae). Am. J. Bot. 93:1140–1153.

Smith SD, Goldberg EE. 2015. Tempo and mode of flower color evolution. Am. J. Bot. 102:1014–1025.

Smith SD, Rausher MD. 2011. Gene loss and parallel evolution contribute to species difference in flower color. Mol. Biol. Evol. 28:2799–2810.

Sobel JM, Streisfeld MA. 2013. Flower color as a model system for studies of plant evo-devo. Front. Plant Sci. 4:1–17.

Steiner CC, Römpler H, Boettger LM, Schöneberg T, Hoekstra HE. 2009. The genetic basis of phenotypic convergence in beach mice: Similar pigment patterns but different genes. Mol. Biol. Evol. 26:35–45.

Stern DL, Orgogozo V. 2008. The loci of evolution: How predictable is genetic evolution? Evolution. 62:2155–2177.

Stern DL, Orgogozo V. 2009. Is genetic evolution predictable? Science. 323(5915):746–751.

Stracke R, Ishihara H, Huep G, Barsch A, Mehrtens F, Niehaus K, Weisshaar B. 2007. Differential regulation of closely related R2R3-MYB transcription factors controls flavonol accumulation in different parts of the Arabidopsis thaliana seedling. Plant J. 50:660–677.

Streisfeld MA, Rausher MD. 2009. Genetic changes contributing to the parallel evolution of red floral pigmentation among Ipomoea species. New Phytol. 183:751–763.

Streisfeld MA, Rausher MD. 2011. Population genetics, pleiotropy, and the preferential fixation of mutations during adaptive evolution. Evolution. 65:629–642.

Streisfeld MA, Young WN, Sobel JM. 2013. Divergent Selection Drives Genetic Differentiation in an R2R3-MYB Transcription Factor That Contributes to Incipient Speciation in Mimulus aurantiacus. PLoS Genet. 9.

Takahashi R, Dubouzet JG, Matsumura H, Yasuda K, Iwashina T. 2010. A new allele of flower color gene W1 encoding flavonoid 3’5’-hydroxylase is responsible for light purple flowers in wild soybean Glycine soja. BMC Plant Biol. 10:155.

De Vetten N, Ter Horst J, Van Schaik HPP, de Boer A, Mol J, Koes R. 1999. A cytochrome b5 is required for full activity of flavonoid 3’, 5’-hydroxylase, a cytochrome P450 involved in the formation of blue flower colors. Proc. Natl. Acad. Sci. USA. 96:778–783.

Watkins J, Chapman JM, Muday GK. 2017. Abscisic acid-induced reactive oxygen species are modulated by flavonols to control stomata aperture. Plant Physiol. 175:pp.01010.2017.

Wessinger CA, Rausher MD. 2012. Lessons from flower colour evolution on targets of selection.

J. Exp. Bot. 63:695–709.

Wessinger CA, Rausher MD. 2014. Ecological transition predictably associated with gene degeneration. Mol. Biol. Evol. 32:347–354.

Whittall JB, Voelckel C, Kliebenstein DJ, Hodges SA. 2006. Convergence, constraint and the role of gene expression during adaptive radiation: Floral anthocyanins in Aquilegia. Mol. Ecol. 15:4645–4657.

Winkel-Shirley B. 2001. Flavonoid biosynthesis. A colorful model for genetics, biochemistry, cell biology, and biotechnology. Plant Physiol. 126:485–493.

Figures and Tables

Figure 1. Simplified view of the flavonoid biosynthesis pathway adapted from Rausher (2006). Arrows indicate biochemical reactions catalyzed by enzymes (white boxes). Intermediate metabolites are enclosed in grey boxes. Abbreviations as follows: LCP = leucopelargonidin; LCC = leucocyanidin; LCD = leudodelphinidin; CHS = chalcone synthase; CHI = chalcone isomerase; F3H = flavanone 3-hydroxylase; F3'H = flavonoid 3'-hydroxylase; F3'5H = flavonoid 3',5'-hydroxylase; DFR = dihydroflavonol 4-reductase; ANS = anthocyanidin synthase. The methylated cyanidin derivative peonidin and methylated delphinidin derivatives malvidin and petunidin not shown. The dashed arrow indicates that F3’5’H does not have 3’ hydroxylation activity in Iochrominae (Smith and Rausher 2011).

Figure 2. Evolution of main floral anthocyanidin pigments in Iochrominae. A. Phylogeny of Iochrominae, with ancestral state reconstruction of anthocyanidins from stochastic mapping, with root state fixed at delphinidin. Pies at nodes represent frequencies of node states across 1000 simulations of character evolution. Disks at tips are colored by species main anthocyanidin pigment (supplementary fig. S1): red for pelargonidin, pink for cyanidin, and blue for delphinidin. B. Transitions between character states, summarized over 1000 stochastic histories, with arrow width proportional to average number of transitions (see supplementary table S2). Crosses indicate nodes with posterior probability < 0.95.

Figure 3. Anthocyanin pathway gene expression in Iochrominae. A. Gene expression for each species, with bars corresponding to each tip of the phylogeny. Colors of bars and dots at tips represent main floral anthocyanidins: red for pelargonidin, pink for cyanidin, and blue for delphinidin (supplementary fig. S2). Bracket represents the “A clade” sensu Smith and Baum (2006). B. Log-fold-change gene expression aggregated by floral anthocyanidin production. Pink circles represent *I. edule*, that makes cyanidin. The letters represent significant differences under a pairwise Tukey comparison test (alpha = 0.05).

Figure 4. Phylogenetic PCA of anthocyanin pathway gene expression. A. Species scores on PC1 and PC2, with symbols detailed in box below. Symbols are filled based on pigment production as follows: red for pelargonidin, pink for cyanidin, white for delphinidin, and yellow for no anthocyanidins. Symbol outline colors indicate k-means clustering with k=3 (as in text): black, red and blue outlines. B. Arrows show loadings for each variable on the PC axes, scale shown on bottom and right axes. C. Pairwise gene expression correlation table. Circle size and color scale represent magnitude of the correlation coefficients, and stars within circles show significance level ($p < 0.05$, 0.01 and 0.001 for 1, 2 and 3 stars respectively).