Evolution of flower color genes in petunias and their wild relatives

Lucas C. Wheeler[1,†,*], Joseph F. Walker[2,3†], Julienne Ng[1], Rocío Deanna[1,4], Amy Dunbar-Wallis[1], Alice Backes[5], Pedro H. Pezzi[5], M. Virginia Palchetti[4], Andrew Monaghan[6], Loreta Brandão de Freitas[5], Gloria E. Barboza[4,7], Edwige Moyroud[2,‡], Stacey D. Smith[1,‡]

1. Department of Ecology and Evolutionary Biology, University of Colorado, 1900 Pleasant Street 334 UCB, Boulder, CO, USA, 80309-0334

2. Sainsbury Laboratory, University of Cambridge, Cambridge CB2 1LR, UK

3. Department of Biological Sciences, University of Illinois at Chicago, Chicago, IL, 60607 U.S.A

4. Instituto Multidisciplinario de Biología Vegetal (IMBIV), CONICET and Universidad Nacional de Córdoba, CC 495, CP 5000, Córdoba, Argentina

5. Laboratory of Molecular Evolution, Department of Genetics, Universidade Federal do Rio Grande do Sul, P.O. Box 15053, 91501-970, Porto Alegre, RS, Brazil

6. Research Computing, University of Colorado, 3100 Marine Street, 597 UCB Boulder, CO 80303

7. Facultad de Ciencias Químicas, Universidad Nacional de Córdoba, Haya de la Torre y Medina Allende, Córdoba, Argentina

* lwheeler9@gmail.com

† Authors contributed equally to the work

‡ Authors contributed equally to the work

Abstract

Evolutionary transitions in flower color often trace back to changes in the flavonoid biosynthetic pathway and its regulators. In angiosperms, this pathway produces a range of red, purple, and blue anthocyanin pigments. Transcription factor (TF) complexes involving members of the MYB, bHLH, and WD40 protein families control the expression of pathway enzymes. Here, we investigate flavonoid pathway evolution in the Petunieae clade of the tomato family (Solanaceae). Using transcriptomic data from 69 species of Petunieae, we estimated a new phylogeny for the clade. For the 65 species with floral transcriptomes, we retrieved transcripts encoding homologs of 18 enzymes and transcription factors to investigate patterns of evolution across genes and lineages. We found that TFs exhibit faster rates of molecular evolution than their targets, with the highly specialized MYB genes evolving fastest. Using the largest comparative dataset to date, we recovered little support for the hypothesis that upstream enzymes evolve slower than those occupying more downstream positions. However, expression levels inversely correlated with molecular evolutionary rates, while shifts in floral pigmentation were weakly related to changes affecting coding regions. Nevertheless, shifts in floral pigmentation and presence/absence patterns of MYB transcripts are strongly correlated. Intensely pigmented and patterned species express homologs of all three main MYB anthocyanin activators in petals, while pale or white species express few or none. Our findings reinforce the notion that regulators of the flavonoid pathway have a dynamic history, involving higher rates of molecular evolution than structural components, along with frequent changes in expression during color transitions.

Keywords

anthocyanins, flower color, molecular evolution, MYB, phylogenomics, Petunieae, Solanaceae, transcription factors

Introduction

The structure and function of biochemical pathways are closely tied to patterns and rates of molecular evolution. For example, enzymes positioned at early steps in these pathways have substantial control over total pathway output (flux) and often experience stronger constraints with lower overall rates of evolution (e.g. Cole and Ingvarsson 2018; Livingstone and Anderson 2009; Rausher et al. 1999, but see Alvarez-Ponce et al. 2009). As they have high flux control, upstream genes are also theoretically expected to be the targets of adaptive substitutions (Wright and Rausher 2010), a pattern found in several empirical studies (Olson-Manning et al. 2013; Passow et al. 2019). Enzymes positioned at branch points exert similarly high control (Rausher 2013; Wheeler and Smith 2019), and thus experience similar evolutionary pressures. Studies across a range of metabolic pathways indicate that, like upstream genes, enzymes at branch points exhibit elevated purifying selection (Greenberg et al. 2008; Ramsay et al. 2009) and, in some cases,

show signatures of positive selection (DallOlio et al. 2012; Flowers et al. 2007). These complex effects of the topology of biochemical pathways can manifest as systems-level relationships between molecular evolution and network properties, such as centrality and connectivity (Hahn and Kern 2005; Masalia et al. 2017; Vitkup et al. 2006).

The molecular evolution of transcription factors in relation to the metabolic pathways they regulate has received less attention, but evidence to date points to markedly different dynamics. Given their position upstream of structural gene targets, transcription factors might be expected to evolve under strong constraints, and that is indeed the case for many well-studied 'toolkit' genes (Carroll 2008). However, some core developmental genes are rapidly evolving (Purugganan and Wessler 1994; Whitfield et al. 1993), including some of the regulators that control transitions to the reproductive stage in plants (Lagercrantz and Axelsson 2000). This variation in molecular evolutionary rates across transcription factors may relate to different histories of gene duplication and different levels of functional specificity (Ascencio et al. 2017; Baum et al. 2005; Streisfeld et al. 2011), as well as differences in gene expression (Jovelin and Phillips 2011; Mukherjee et al. 2016; Yang and Gaut 2011). One challenge for understanding how the molecular evolution of transcription factors differs from the genes they regulate is the lack of studies investigating both sets of genes (Alvarez-Ponce et al. 2009). This pathway-level approach is essential for gaining insight into the potential targets of selection during macro-evolutionary transitions (e.g., Ciezarek et al. 2019; Foote et al. 2015) and assessing the consequences for other pathway genes (e.g., relaxed constraint and gene decay following trait loss, Preston et al. 2011; Springer et al. 2021).

Here we use the flavonoid pigmentation pathway in the genus *Petunia* and its wild relatives (tribe Petunieae) to investigate the relationship between macroevolutionary trait transitions and the molecular evolution of the underlying gene network. Flavonoids include the blue, purple, and red anthocyanin pigments that color many flowers and fruits, and a range of yellowish or colorless compounds (e.g., flavonols, flavones) that can act as co-pigments and create UV-absorbing patterns on flowers (Davies et al. 2012; Winkel-Shirley 2001). While our knowledge of this deeply conserved pathway builds from work in a broad range of model systems (e.g., maize, *Arabidopsis*, snapdragon), *Petunia* has served as the premier model for understanding the regulation of anthocyanin pigments and co-pigments that give rise to variation in flower color intensity, hue, and pattern (Albert et al. 2014; Berardi et al. 2021; Esfeld et al. 2018; Quattrocchio et al. 2006; Sheehan et al. 2016). Surprisingly, very little is known about the molecular basis for flower color variation in the wild relatives of petunias, which include many showy-flowered taxa of horticultural importance such as species of *Calibrachoa* (million bells), *Nierembergia* (the cupflowers), *Brunfelsia* (yesterday, today, and tomorrow), and *Fabiana* (the false heaths) (fig. 1). This wide floral variation across the ca. 182 Petunieae species provides an opportunity to test whether the mechanisms controlling flower color in model species extend to a clade-wide scale.

One emerging theme from flower color genetics is the critical role of R2R3 MYB transcription factors. These highly variable proteins have duplicated extensively in flowering plants (Gates et al. 2016; Jiang and Rao 2020) and tend to be narrowly specific in terms of their spatial and temporal expression as well as their targets (Sobel and Streisfeld 2013). Many MYB genes, acting in complex with bHLH and WD40 partners, regulate epidermal cell differentiation, contributing for example to the distribution of root hairs (Bernhardt et al. 2005) and the conical shape of petal cells (Ramsay and Glover 2005). In the context of flower color, different copies of MYBs are specialized for activating anthocyanins and co-pigment production in different petal regions, thus regulating overall color intensity and the complex pigmentation patterns such as spots (Ding et al. 2020; Martins et al. 2017) and bullseyes (Sheehan et al. 2016). The primary activators of anthocyanin

production in *Petunia* include four MYBs from subgroup 6 (AN2, AN4, DPL, and PHZ) (Albert et al. 2011; Zhang et al. 2021). Another cluster of subgroup 6 MYBs, the ASR genes, were recently described in *Petunia*, and these seem to be important early in bud development (Zhang et al. 2019). The activity of MYB repressors, including MYB27 and MYBX (Albert et al. 2011), also shape floral anthocyanin production. MYBFL, which belongs to the subgroup 7 flavonoid regulators, controls the floral expression of flavonol co-pigments (Sheehan et al. 2016). Changes in the function and expression of these diverse MYB transcription factors underlie much of the flower color variation across model species of *Petunia* (Berardi et al. 2021; Esfeld et al. 2018; Hoballah et al. 2007; Quattrocchio et al. 1999). Thus, we hypothesized that MYBs are likely to contribute to the diversification of flower color across the entire Petunieae clade, a history that would be reflected in elevated rates of molecular evolution compared to the rest of the pathway.

To test these hypotheses, we built a large and densely sampled transcriptomic dataset for Petunieae and examined the evolution of structural and regulatory genes of the flavonoid pathway across multiple flower color transitions. First, we estimated a new phylogeny for the tribe to provide a framework for identifying color transitions and testing for their molecular signatures. Next, we estimated rates of molecular evolution for structural and regulatory genes to examine how these rates vary with position and functional role. Finally, we used the repeated color transitions across the phylogeny to test whether losses of floral pigmentation are associated with changes in the selective constraint acting on coding sequences and the presence/absence of transcripts of the structural and regulatory genes. Our results uncover widely varying dynamics across the pathway and its regulators, with some loci highly conserved and others rapidly evolving. The rapid evolution of MYB genes, along with the absence of the MYB activators in the white-flowered lineages, implicates this class of transcription factors as playing a central role in flower color evolution at the clade level.

Results

Transcriptomic data resolve relationships across Petunieae

Our transcriptomic dataset spanned all genera of Petunieae, with multiple species of all non-monospecific genera, allowing us to make inferences about relationships and explore the diversity of phylogenetic signal across loci. Previous phylogenetic analyses of *Petunia* and allied genera showed moderate to strong support for the monophyly of the genera. Still, these studies were often limited in sampling and relied on a handful of markers (e.g., Ng and Smith 2016; Särkinen et al. 2013). Furthermore, relationships among the genera have been contentious (Reck-Kortmann et al. 2015) and some genera have been very difficult to resolve (Fregonezi et al. 2012). Our coalescent-based and concatenation analyses of 3672 protein-coding genes revealed congruent relationships along the tree's backbone (fig. 1, S1, S2). For example, *Petunia*, *Calibrachoa*, and *Fabiana* formed a well-supported clade with the latter two genera as sister groups (as in Reck-Kortmann et al. 2015 but contra Olmstead et al. 2008; Särkinen et al. 2013). We also recovered the small but florally diverse *Bouchetia-Hunzikeria-Plowmania* (BHP) clade and its close relationship to *Nierembergia* and *Leptoglossis* as in Särkinen et al. (2013). The large-flowered shrubby genus *Brunfelsia* is sister to this group of small herbs, a relationship also found in previous work (e.g. Filipowicz et al. 2012). In addition to these backbone relationships, our analyses highlighted the discordance across gene trees that may explain past challenges in inferring relationships among genera. The most significant conflict involves the placement of *Brunfelsia*, where we estimated that 26% of the genes conflict in

the species tree (fig. 1, S3).

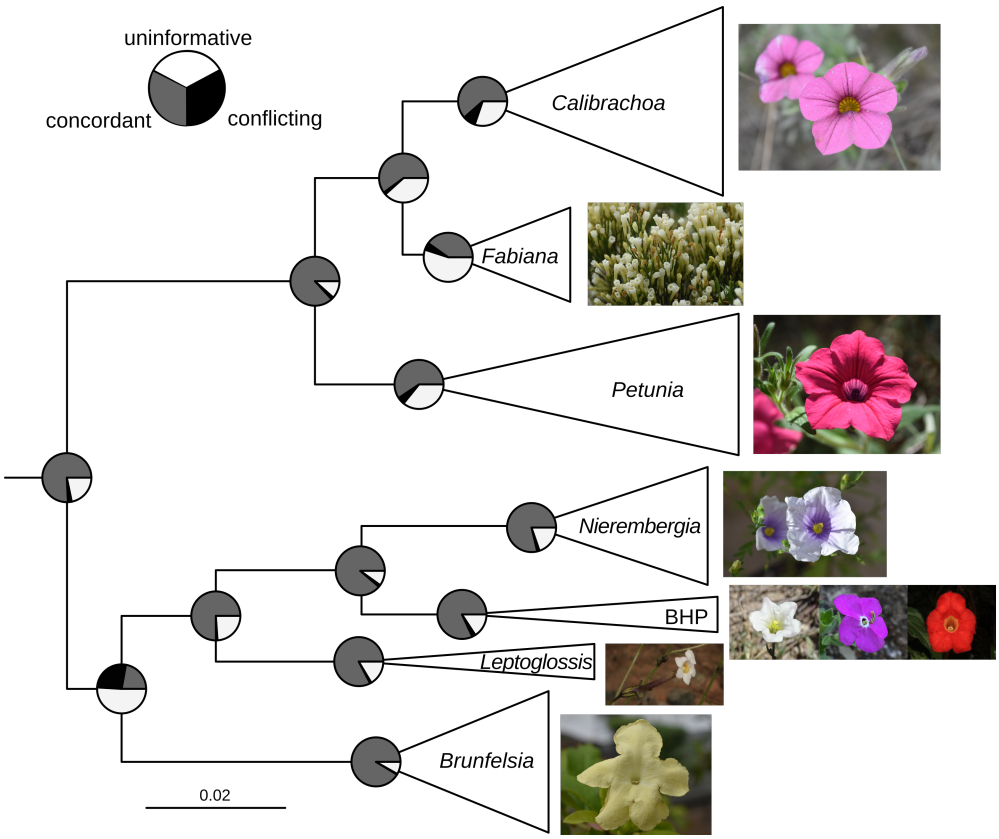

Fig. 1. Strong support across gene trees for the monophyly of Petunieae and its genera. Tree topology from the coalescent-based species tree analysis with branch lengths in substitutions per base pair of concordant genes (following Walker et al. 2021) All branches have 100% local posterior probability. The size of the triangles corresponds to the number of taxa sampled in the clade (supplementary fig. S1). Pie charts at the nodes show the level of gene tree conflict where gray, black, and white denote concordant, conflicting, and uninformative (support less than 95% UFboot or insufficient taxon sampling), respectively. Images from top to bottom (with credits): *Calibrachoa eglandulata*, *Fabiana punensis*, *Petunia reitzii*, *Nierembergia scoparia* (all by Lucas C. Wheeler), *Bouchetia erecta* (Edith Bergquist), *Hunzikeria texana* (Karla M. Benítez), *Plowmania nyctaginoides*, *Leptoglossis albiflora* (both by Rocío Deanna), and *Brunfelsia lactea* (Lucas C. Wheeler).

We also observed wide variation in patterns of concordance within the genera. Some splits are highly concordant across gene trees (e.g., the split between Antillean and South American *Brunfelsia* (Filipowicz et al. 2012) and the two subgenera of *Calibrachoa* (Fregonezi et al. 2012)) while other shallow relationships showed little agreement across gene trees (supplementary fig. S3). Discordance was particularly notable in *Calibrachoa* subg. *Stimomphis*, where previous studies have

found little support for relationships among the 25 species (Fregonezi et al. 2012, 2013). A more recent study using 10kb of plastid and nuclear sequence data recovered higher support in combined analyses (Mäder and Freitas 2019); however, few of those relationships match those we inferred (supplementary fig. S1). These disagreements are not surprising as approximately 10% of the genes follow the inferred species tree (supplementary fig. S3). The extremely short branch lengths in coalescent units (supplementary fig. S1A) are consistent with incomplete lineage sorting during a rapid radiation. However, hybridization among the highly interfertile species (Facciuto et al. 2009) could also have contributed to the discordance. Portions of the *Petunia* clade show high levels of gene tree conflict as seen in *Calibrachoa* (supplementary fig. S3).

Rates of molecular evolution vary significantly across pathway genes

We mined the high-quality floral transcriptomes (65 of 69 species) for structural and regulatory genes with well-studied roles in floral flavonoid variation in *Petunia*. We created a bioinformatic pipeline (see Methods) to retrieve ten structural genes and the homologs of 8 transcription factors (shown in fig. 2A). We recovered nearly all of the structural genes for all of the taxa and the majority of the pigment activators and repressors, including the bHLH genes AN1 and JAF13, the WD40 AN11, and five MYBs that regulate flavonoid production. We detected the homologs of An4 and the ASR genes in at most eight species; therefore, we excluded them from statistical analyses. For CHI, our pipeline recovered both the A and B copies (supplementary fig. S4). We focused on CHI-A to represent this step in the pathway as it is the predominant copy involved in flavonoid synthesis in petunia petals (van Tunen et al. 1988). However, the duplicates have similar rates of molecular evolution (results not shown). CHS also comprises a multi-gene family in *Petunia*, with CHS-A and CHS-J being closely related (Koes et al. 1989a) and the former accounting for roughly 90% of the floral expression (Koes et al. 1989b). Our pipeline retrieved a single copy across the taxa, which appears to correspond to CHS-A (supplementary fig. S5).

Using sequence alignments for these genes and their maximum likelihood trees, we estimated the ratio of non-synonymous to synonymous substitution rates (ω) as an indicator of selective constraint. We found that ω varies nearly seven-fold across loci, with the lowest value (0.09) corresponding to the most upstream structural gene in the flavonoid pathway (CHS-A) and the highest value (0.62) corresponding to MYBFL, one of the MYB transcription factors (fig. 2B; supplementary Table S3). This rate for CHS is on par with housekeeping genes, such as actin and GAPDH homologs (supplementary fig. S6A), and is consistent with strong purifying selection (Yang 2007). The genes with higher ω do not present more sites under positive selection (fig. 2B; supplementary fig. S6A), suggesting that the elevated rates instead reflect relaxed selective constraint spread across the coding regions. Pairwise comparisons among these loci supported the significant variation in molecular evolutionary rate observed across this set of genes (supplementary fig. S6B).

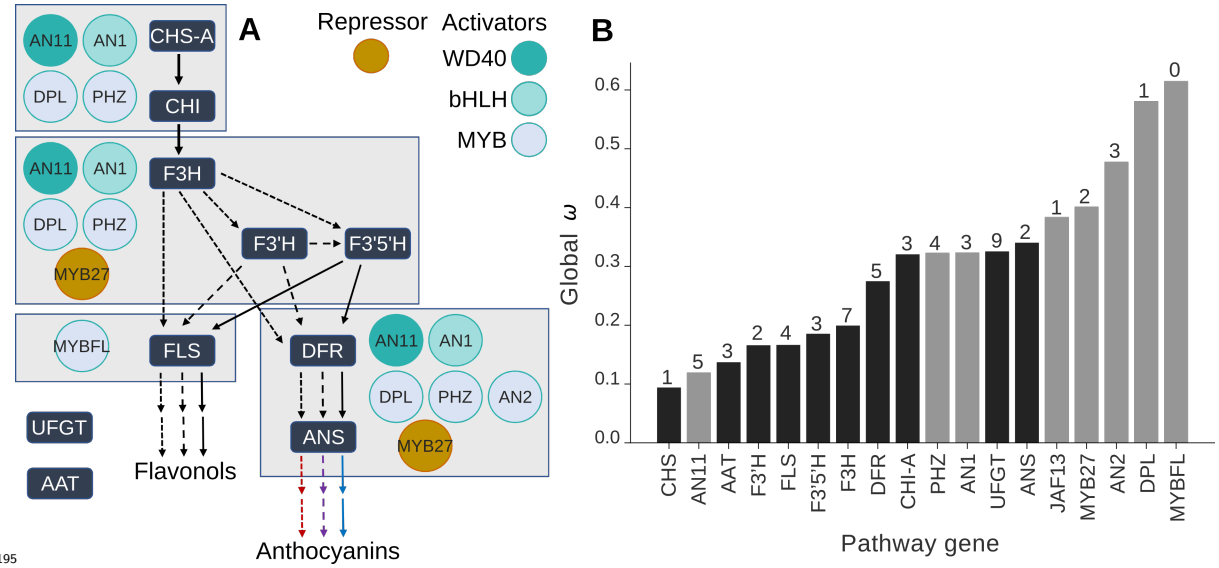

Fig. 2. Pathway genes vary seven-fold in evolutionary rate. (A) Simplified diagram of the flavonoid pathway with enzymes in the dark blue boxes. Colored circles denote regulators. The MYB genes DPL, AN2, and PHZ function as part of MYB-bHLH-WD40 (MBW) complexes, with AN1 or JAF13 (not shown) as the basic helix-loop-helix (bHLH) partner and AN11 as the WD40 partner. The dashed lines indicate the type of product (mono-, di-, or tri-hydroxylated), which correspond to anthocyanins of different hues. The two rows of arrows at the last steps indicate downstream modification and decoration of the base molecules by enzymes such as UFGT and AAT. See supplementary fig. S7D for a more detailed pathway depiction and supplementary Table S2 for gene names. (B) The global values estimated for each flavonoid pathway gene from the single ω model (structural genes in black and transcription factors in gray). Values above bars are the number of sites subject to positive selection, estimated by FUBAR (Murrell et al. 2013).

Functional category predicts evolutionary rate better than position in the network

In previous studies of the flavonoid pathway and other metabolic networks, the molecular rate of evolution and pathway position have been linked (Rausher et al. 1999, 2008); therefore, we first examined their relationship with our dataset. We repeated the classic analysis of Rausher et al. (1999) and found a similar trend of increasing evolutionary rates moving along the linear portion of the pathway across the six 'core' genes, i.e., those that comprise the shortest pathway from precursors to pigments (CHS, CHI-A, F3H, DFR, ANS, UFGT) (supplementary fig. S7A). However, this trend was not significant for ω or the non-synonymous rates (dN) alone ($R^2 = 0.72$, 0.76, $p = 0.1$, 0.08; Kendall's $T = 0.6$, 0.6, $p = 0.14$, 0.14, supplementary fig. S7A, B). We then repeated this analysis using the broader set of genes involved in the flavonoid biosynthesis. We scored position using a modified pathway pleiotropy index (Ramsay et al. 2009), where we averaged the position for genes involved in multiple reactions (see Suppl. methods). Again, we found no significant relationship with dN or ω ($R^2 = 0.22$, 0.19, $p = 0.53$, 0.60; Kendall's $T = 0.18$,

0.09, $p = 0.47$, 0.72, respectively; fig. 3A, supplementary fig. S7C). This result appears to be partly driven by the highly conserved but downstream genes FLS and AAT (fig. 3A).

To test whether gene type was associated with substitution rate, we divided pathway genes into structural and regulatory categories and compared rates between these two groups. The median ω for transcription factors is roughly double that of the structural genes (0.4 vs. 0.2, Kruskal-Wallis H-test: $H = 5.76$, $p = 0.016$, fig. 3C). The two classes of genes also differed in the range of variation in ω, with the structural genes having a compact distribution and the transcription factors spread from 0.12 for the WD40 repeat protein AN11 to 0.62 for MYB-FL (fig. 3B). Thus, structural genes from the flavonoid pathway tend to evolve more slowly and vary less in rates of evolution than transcription factors in Petunieae.

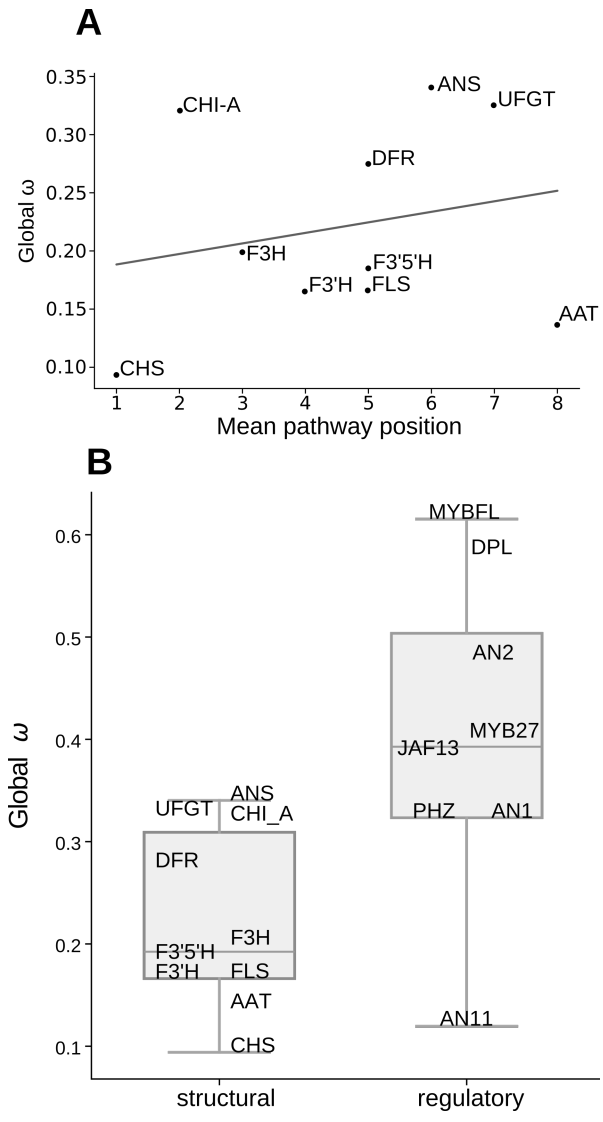

Fig. 3. Molecular evolutionary rates vary by gene type but not across pathway positions. A) Global ω estimated for structural genes ranked according to their mean position in the pathway, from most upstream (CHS) to most downstream (AAT). (B) Boxplot distributions for ω for regulatory and structural genes that code for transcription factors and enzymes, respectively. ω values of individual genes are marked by their abbreviations. The height of each box shows the interquartile range, the horizontal line shows the median, and the bars show the range of values.

Finally, we examined how these molecular evolutionary rates might vary with gene expression levels, as these two factors are often closely related (Jovelin and Phillips 2011; Mukherjee et al. 2016; Slotte et al. 2011; Yang and Gaut 2011). By mapping reads back to each assembled CDS for each species, we calculated the total reads per gene and found that this value was significantly correlated with the global ω ($R^2 = -0.63$, $p = 0.0003$; fig. 4). The average number of reads for structural genes is 14.6-fold higher than for transcription factors (Kruskal Wallis H-test: $H = 12.6$, $p = 0.0003$, supplementary fig. S8), in line with their lower ω values. We repeated the analysis with additional genes (five housekeeping genes and four florally expressed transcription factors) to determine how widely this pattern held. We recovered a very similar pattern ($R^2 = -0.63$, $p = 0.0003$, supplementary fig. S9). We also considered that sequencing error associated with variation in read counts might contribute to the observed relationship with ω. We compared our assemblies for pathway genes with those from published *Petunia* genomes. We found that the percent identity between the two was not related to read number (e.g., MYB27 and AAT were over 99% identical for *P. axillaris* sequences despite having thousands more reads for the latter; supplementary table S4). The results show that lower read counts still gave accurate assemblies, and the observed relationship with ω is not due to sequencing error.

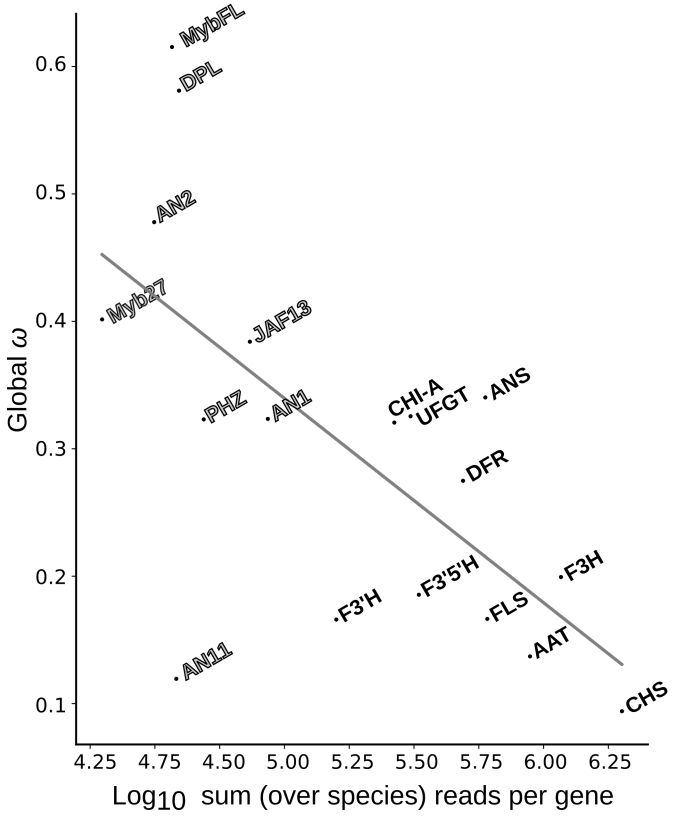

Fig. 4. Gene expression level is related to molecular evolutionary rate. Labels indicate gene names with structural genes in bold black and transcription factors in bold gray. The number of reads for each gene was computed with Salmon (Patro et al. 2017) by mapping reads from each sample to a reference file containing the assembled transcripts for the gene and summing across all samples.

Rates of molecular evolution are decoupled from flower color shifts

We next used phylogenetic comparative methods to identify shifts in the intensity of floral anthocyanin production and test the relationship of those shifts to rates of gene evolution. We hypothesized that pathway genes would experience relaxed selection in lineages that produce white flowers lacking anthocyanin pigments (Ho and Smith 2016). We also predicted that relaxed selection in white lineages would be most pronounced in the MYB transcription factors (e.g., AN2, DPL) since their roles are specific to anthocyanin production in flowers (Quattrocchio et al. 1999; Schwinn et al. 2006). Our floral biochemical profiling (see Methods) revealed wide variation in levels of anthocyanin production across the clade, providing evolutionary replication to test these hypotheses. *Petunia* and *Calibrachoa* showed the most intense pigmentation, while *Brunfelsia*, *Leptoglossis*, and *Nierembergia* experienced convergent losses (fig. 5A). Despite the many color shifts, there is a significant phylogenetic signal in floral anthocyanin concentration (Blomberg's

$K = 0.57$, $p = 0.001$) (Blomberg et al. 2003) .

We first carried out phylogenetic generalized least squares (PGLS) analysis to test for associations between variation in molecular evolutionary rates and flower color intensity. Using anthocyanin content measured with HPLC (supplementary Table S5) and tip values for rate ω (supplementary Table S6), we conducted PGLS analyses for each gene. The PGLS analyses showed two marginally significant relationships, non-significant after Bonferroni correction (supplementary Table S7). Repeating this analysis with dN, rather than ω, also revealed no significant gene-wise relationships (supplementary Table S8). Since relaxed selection may not be concentrated in particular loci but spread across the pathway, we repeated the analysis using the sum of dN across all genes for each tip; this comparison also returned no significant relationship (supplementary fig. S10). As a whole, the PGLS results indicate that shifts in the intensity of pigmentation are decoupled from rates of molecular evolution associated with flavonoid pathway genes.

Mapping anthocyanin content onto the Petunieae phylogeny indicated multiple complete losses of floral anthocyanins. Thus, we also scored pigmentation as present/absent and estimated branch models. We implemented branch models in RELAX (Wertheim et al. 2015), allowing ω to vary between background lineages producing pigments and foreground lineages without pigments while incorporating uncertainty in ancestral trait reconstruction. We estimated relaxed selective constraint for four of the 18 pathway genes examined: three regulatory genes coding for the transcription factors AN1, JAF13, AN11, and one structural gene, AAT, encoding a downstream enzyme in the flavonoid pathway (supplementary Table S9). This result is significant for AN1, the bHLH component of MBW complexes that activate anthocyanin production, where the estimated ω for lineages lacking anthocyanins is twice that of those with anthocyanins (supplementary Table S9). Combined with the PGLS analyses, these results suggest that while reductions in pigmentation and evolutionary rates across the pathway are not tightly coupled, complete losses tend to coincide with relaxed selection for some loci.

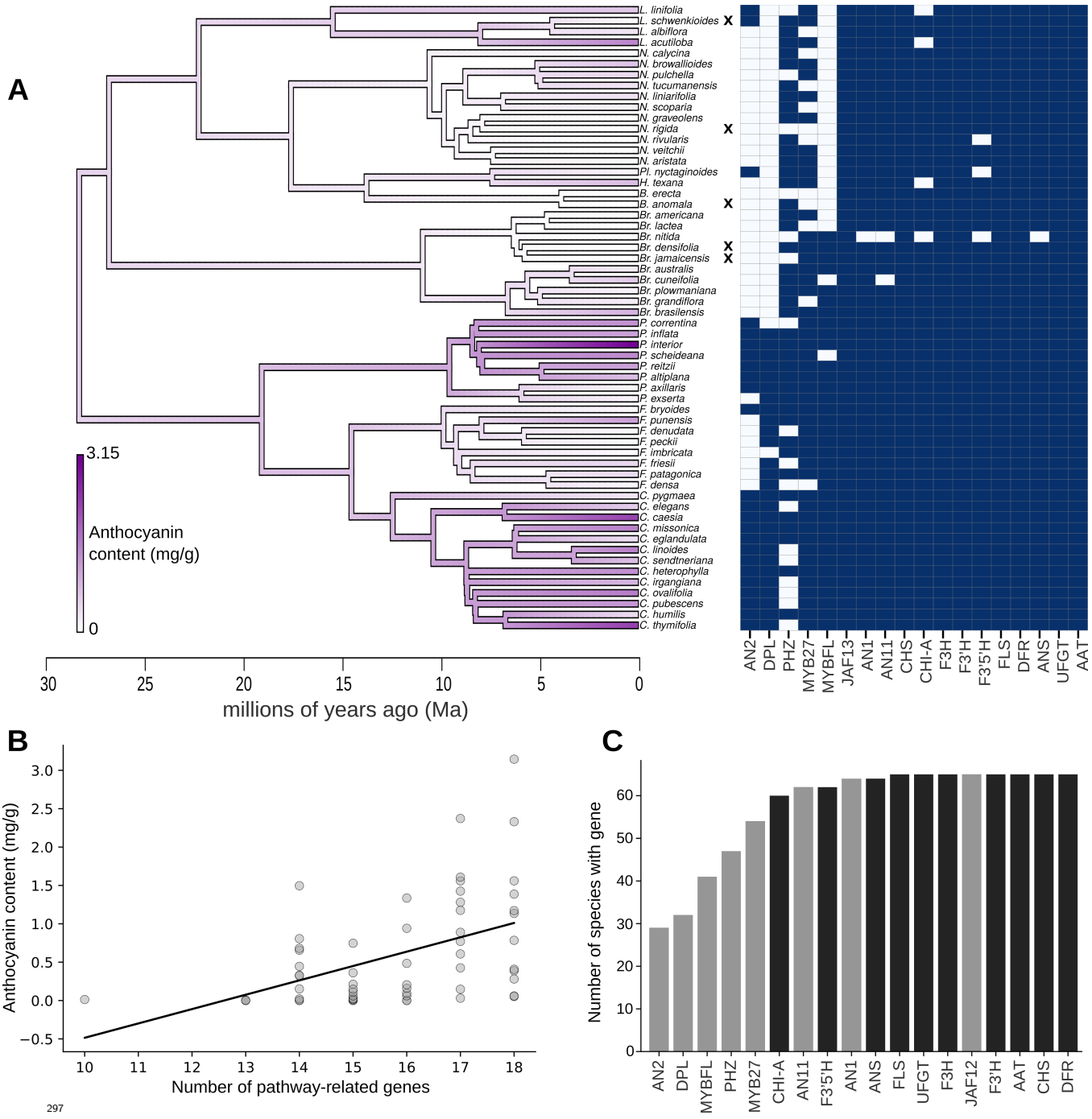

Fig. 5. Convergent losses of color are associated with fewer pathway genes expressed. (A) Maximum likelihood ancestral state reconstruction of the anthocyanin pigment concentration for the 58 species with HPLC data. Black **X** indicates a complete loss of floral

pigmentation. The heatmap shows flavonoid pathway genes detected (dark square) or not (light square) for each species based upon associated reads from the floral transcriptome. Genes are in the order MYBs, bHLH, WD40 transcription factors, then structural genes from early to late. (B) A plot of anthocyanin content against the number of pathway-related genes present for each species. (C) The number of species where we were able to detect each gene in the transcriptome. Structural genes are black, and transcription factors are gray.

Loss of floral pigmentation accompanied by lack of MYB expression

In addition to variation in rates of molecular evolution across pathway loci, our transcriptomic dataset revealed variation across species in the presence and absence of transcripts associated with the flavonoid pathway genes (supplementary fig. S4). We hypothesized that this variation in transcriptome content might be related to floral pigmentation, particularly as regulatory changes often underlie macroevolutionary color transitions (Larter et al. 2019). Treating anthocyanin concentration as a continuous trait, we again used PGLS to test for an association between pigmentation and variation in gene expression. We found a strong relationship ($R^2 = 0.21$, $p = 0.003$), with transcript detection positively correlated with anthocyanin concentration (fig. 5B). We suspected that this correlation is primarily driven by the MYB transcription factors, which are the most frequently absent among the set of pathway genes (fig. 5C), and indeed, removing these genes from the analysis eliminates the significant correlation (supplementary fig. S11). Thus, the more pigmented species express a larger number of pathway genes, and specifically, more MYB genes, in their corolla.

In contrast to the pigmented species, white-flowered species tend to be missing MYB genes while mostly retaining the structural genes. The presence of the structural targets in the absence of their activators may seem surprising, especially for the downstream genes (e.g., DFR, ANS) that are only involved in anthocyanin production (fig. 2A). Still, most of the white-flowered species express at least one activator, which may be sufficient for a low level of pathway expression. There are three white-flowered species in which no activators were detected, and in these cases, it is possible that the activators were expressed in earlier stages of development (fig. 5A). For example, the ASR genes, which were only recovered from a few pigmented species (supplementary fig. S12, S13), tend to be active only early in bud development in *Petunia* (Zhang et al. 2019).

The variation in the presence and absence of MYB transcripts across Petunieae species could be due to differences in genomic content (gene gain or loss) or gene expression in corolla tissues. To explore this possibility, we designed specific primers for AN2 and DPL, the two MYBs that were most often missing, and surveyed several taxa for the presence of these genes in the genome (see Supplemental Methods). The primers for DPL successfully amplified that specific MYB and allowed us to sequence partial copies from species of *Brunfelsia* and *Nierembergia* that lacked DPL in their transcriptomes (see Supplementary Results). For AN2, the primers designed to be copy-specific tended to amplify multiple copies, although at least one amplicon corresponded to AN2. Overall, these results suggest that DPL, and likely AN2, are present more widely in Petunieae and that their absence in the transcriptomes is due to regulatory changes.

Discussion

Drivers of evolutionary rates across the anthocyanin pathway

The topology of metabolic pathways has often been linked to rates of molecular evolution (Alvarez-Ponce et al. 2009; Montanucci et al. 2018; Vitkup et al. 2006). Early studies on the molecular evolution of the flavonoid pathway found a trend of greater constraint in the enzymes positioned at the first steps in the pathway (Rausher et al. 1999), which have high control over pathway output. Subsequent studies have recovered mixed results, with a few supporting the pattern (Lu and Rausher 2003) and others showing no association between position and rate of molecular evolution (Ho and Smith 2016; Ramos-Onsins et al. 2008; Shoeva et al. 2017). With our 65 species dataset, we observed a slight trend across the six core pathway genes (supplementary fig. S3A), but a more complete sampling of the pathway did not support this relationship (fig. 3A). The most downstream gene sampled, the anthocyanin modifier AAT, has a similar ω and an even lower dN than the first committed enzyme in the anthocyanin pathway, CHS (fig. 3C). Overall, it appears that rates of molecular evolution are only weakly related to pathway position, implicating other factors in generating the 3-fold variation in ω across the structural genes (fig. 2B). One complication in dissecting this relationship is that flux control, thought to be the underlying driver of the position effect (Rausher et al. 1999), likely evolves with the color phenotype, shifting such that the enzyme's control over the selected products is maximized (e.g., purple pigments) (Wheeler and Smith 2019; Wheeler et al. 2021). Future studies could assess the role of flux more directly by focusing on phenotypic transitions and testing whether shifting flux control alters selective regimes acting on pathway enzymes.

Although relative position within the series of biochemical steps was not correlated with the rate of molecular evolution, we found that a gene's function, as a regulator or enzyme, was predictive of selective constraint. Consistent with several studies in other pathways (e.g., Jovelin and Phillips 2011; Wu et al. 2010, but see Invergo et al. 2013), we found that transcription factors regulating the flavonoid pathway evolve 1.8 times faster on average than the structural genes they regulate (fig. 3C). The lack of positively selected sites in these genes suggests this difference is primarily due to relaxed selection (fig. 2B). The elevated evolutionary rates in transcription factors may be surprising as any coding mutations could affect the expression of multiple downstream targets (Carroll 2008; Doebley and Lukens 1998). However, this predicted pleiotropy can be reduced by redundancy and specialization (Badawi et al. 2014; Duret and Mouchiroud 2000), both of which are at play among the regulators of anthocyanin biosynthesis. Functional studies in *Petunia* suggest that the two bHLH proteins AN1 and JAF13 overlap in function, with floral pigment production being initiated by JAF13 and then reinforced with AN1 to give full coloration (Albert et al. 2014; Spelt et al. 2000). Functional roles are even more finely divided among the MYB genes regulating the flavonoid pathway, most of which are specific to particular regions of the flower (Schwinn et al. 2006) and/or branches of the pathway (Berardi et al. 2021; Sheehan et al. 2016). For example, in most *Petunia*, DPL or AN4 control vein coloration (Albert et al. 2011; Zhang et al. 2021), AN2 activates color production in the petal limb (Quattrocchio et al. 1993), and MYB-FL controls the flavonol co-pigments that give floral UV patterns (Sheehan et al. 2016). As might be predicted from these functional differences across TFs, the single copy AN11 has a slower rate of molecular evolution; the bHLH genes are intermediate; and the diverse MYBs exhibit the fastest molecular rates (fig. 2B, see also Streisfeld et al. 2011). In contrast to the functional specificity observed for

most TFs, the pathway enzymes are widely expressed in above-ground tissues, where flavonoids, including anthocyanins, are involved in many physiological functions, such as protection from UV light and mitigation of drought stress (Gould 2004).

Transcription factors may evolve faster than their enzymatic targets, not only because of their specificity, but their lower level of expression. Our results showed transcription factors have a roughly 15-fold lower level of expression on average than structural genes, and rates of molecular evolution negatively correlate with this variation (Drummond et al. 2006; Pál et al. 2001; Subramanian and Kumar 2004; Yang and Gaut 2011). The transcription factors, particularly the rapidly evolving MYB genes, had some of the lowest expression levels (fig. 4). The low level of expression of transcription factors relative to their targets has been widely observed (e.g. Czechowski et al. 2004; Ghaemmaghami et al. 2003; Vaquerizas et al. 2009) and may limit binding to the highest affinity targets (Liu and Clarke 2002). The relationship between expression levels and rates of sequence evolution may be driven by selection against misfolded proteins, which not only represent wasted energy, but can act as toxins in the cell (Drummond and Wilke 2008). Some degree of misfolding of lowly expressed proteins may not invoke high fitness costs as long as the same mutations that affect misfolding propensity do not substantially affect function. By contrast, highly expressed genes found in many tissues, like housekeepers and the flavonoid pathway enzymes, are expected to be under strong selection for robust folding, consistent with our findings. Collectively, our results show that the role of the gene in the pathway (structural or regulatory) and the level of expression associated with that function are the primary drivers of rates of molecular evolution.

The interplay of molecular evolution and floral color transitions

Macroevolutionary transitions in phenotype are often associated with suites of changes in the pathways that underlie the development of those phenotypes. While some of these changes are required to produce the new phenotype, others may accumulate after the transition, e.g., changes that stabilize the new state (Deng et al. 2010; Poon and Chao 2005; Rodríguez-Trelles et al. 2003) or changes that reflect relaxed selection on genes no longer expressed (Boakye et al. 2017; dePamphilis and Palmer 1990; Meredith et al. 2013). Our study revealed repeated cases in which lineages have transitioned to pale flowers and, in some cases, lost floral pigmentation entirely (fig. 5). We hypothesized that pigment pathway genes, particularly the floral-specific regulators, would show relaxed selective constraint in these lineages. We found no clear evidence that decreases in the color intensity relaxed the strength of selection (supplementary Table S5, S6). However, complete losses of floral pigmentation, which occurred independently in five lineages (fig. 5, supplementary Table S7), did lead to relaxed constraint for the downstream gene AAT and three transcription factors, the WD40 AN11, and the bHLH genes JAF13 and AN1, with the strongest effect in the latter (supplementary Table S6). These losses occurred within the last 10 MYA (fig. 5A, S7), a timespan over which genes with lost functions would be expected to decay (Lynch and Conery 2000; Marshall et al. 1994; Protas et al. 2007). Even though the increase was marked (e.g., from ω of 0.31 in the background to 0.62 in loss lineages for AN1), all values remain well below 1, indicating purifying selection. They may, for example, contribute to flavonoid production in other tissues and/or in different conditions (e.g., drought stress).

While the R2R3 MYB AN2 and its close relatives DPL and PHZ did not emerge from these cross-species analyses of sequence evolution, they appear linked to color transitions through the presence and absence of their associated transcripts. The five pigment-less lineages are missing DPL, and

AN2 transcripts in their corollas, as are many other pale or white-flowered species (fig. 5A). Another MYB activator, PHZ, which gives blushes to flowers and colors the vegetative tissue of petunias (Albert et al. 2011), is recovered from most of the species, suggesting it is likely responsible for the pale floral coloration of taxa like the cupflowers (fig. 1E). By contrast, the intensely pigmented and patterned species, like most *Petunia* and *Calibrachoa* (fig. 1A, B), express all three MYB activators (fig. 5A). Although this variation in transcriptome content could reflect underlying differences in genome content, our small PCR survey (see Supplemental Text) supports the regulatory hypothesis, given that DPL and AN2 appear widely spread across Petunieae genomes. These findings align well with lessons from the model petunias that the diversification of R2R3 MYB genes and their regulation has been integral to flower color evolution (Berardi et al. 2021; Esfeld et al. 2018).

Conclusions

Biochemical pathways underlie many phenotypes central to organismal function and adaptation. Patterns of molecular evolution across these pathways can provide insight into the selective forces that have historically acted on each gene and how these genes evolve in concert with phenotypic transitions. With extensive clade-level analysis of the anthocyanin pathway, our study confirms that most pathway elements are highly conserved, despite a multitude of shifts in color intensity and pattern. Evolutionary change in protein sequence is concentrated in the transcription factors of the pathway and especially the MYB genes. With their high tissue specificity and low levels of expression, coding mutations in MYBs are more often fixed than in other pathway genes. Moreover, the presence of these genes is highly variable across species, with lineages containing the most substantial dose of MYB activators producing the most intense colors. While these macroevolutionary patterns accord with our understanding of anthocyanin pathway function in model systems, extending these studies to additional clades and timescales would be valuable. For example, we expect that the correlation between the number of expressed MYBs and the intensity of color and pattern across species might extend to other groups with similarly wide flower color variation (Schwinn et al. 2006; Yuan 2019). In addition, the link between the relatively low expression of transcription factors compared to targets and the resulting relaxed constraint is likely to hold broadly.

This study also sheds light on the potential for identifying the mechanisms of phenotypic evolution at phylogenetic scales Smith et al. 2020. With an exceptionally well understood and widely conserved pathway, floral anthocyanin pigmentation is an ideal focal trait for linking genetic and developmental changes to species differences. Our work suggests that genomic scans of coding sequence variation (e.g. Muntané et al. 2018; Prudent et al. 2016) might capture some relevant genetic changes but miss important evolutionary dynamics. In the case of Petunieae flower color evolution, the MYB genes did not show shifts in coding sequence evolution during color transitions. However, their pattern of presence/absence across the transcriptomes points to a strong relationship with color variation. This result highlights the importance of developing phylogenetic genotype-to-phenotype (PhyloG2P) approaches tailored for detecting the signals of different mechanisms that can alter phenotype, from variation in coding sequences (e.g. Halabi et al. 2021) to shifts in gene content (Kiefer et al. 2019), to regulatory changes (Hu et al. 2019; Larter et al. 2018). Integrating these approaches can lead to a deeper understanding of how pathway structure and function shape phenotypic space and the potential for moving through it.

Methods

Transcriptome assembly and identification of pathway genes

We sequenced and assembled floral transcriptomes for 67 species, leaf transcriptomes for an additional two species of Petunieae for which flowering material was not available, and a floral transcriptome for the outgroup species *Browallia americana* (supplementary Table S1, Supplementary Material online, suppl. methods). Tissue sampling and RNA extraction followed (Larter et al. 2018), as described here (supplementary text, Supplementary Material online). Transcriptomes were assembled from 100bp paired-end reads following Yang et al. (2015), with the steps unified as a single Snakemake pipeline (https://osf.io/b7gcp/). Briefly, raw reads were first corrected using Rcorrector (Song and Florea 2015), and Trimmomatic (Bolger et al. 2014) was used to remove adapters. Trimmed reads were assembled with Trinity (Grabherr et al. 2011), incorporating strand-specific information. The raw Trinity assemblies were filtered for chimeric sequences using the "run_chimera_detection.py" script from Yang et al. (2015) with a custom BLAST database constructed from *Petunia*, *Solanum*, and *Arabidopsis* transcriptomes. We then used Corset (Davidson and Oshlack 2014) to collapse and cluster transcripts and finally TransDecoder (noa 2021) to predict CDS and filter predicted sequences against the same custom BLAST database, discarding CDS with no BLAST hits.

We retrieved anthocyanin pathway genes, along with a selection of housekeeping genes and other known florally-expressed transcription factors for comparison, by first collecting published sequences and using them to query BLAST databases created with makeblastdb (Madden 2013) from each raw Trinity transcriptome assembly. Our search set included the structural genes encoding CHS-A, CHI-A, F3H, FLS, F3'H, F3'5'H, DFR, ANS, UFGT, AAT; the transcription factors AN2, DPL, PHZ, AN11, AN1, JAF13, MYBFL, MYB27, AN4, ASR1, ASR2, ASR3, PH1, PH2, ODO1; and the housekeeping genes actin, tubulin, Rps18, Gapdh, Hprt (see supplementary Table S2 for full gene names). We retained matching hits (e-value cutoff = 1e-50) and then used TransDecoder to predict CDS and peptide sequences from these sets. Among these sets, we kept the single most similar sequence to the representative published sequence. Filtered sequences were blasted to the *Petunia inflata* draft genome CDS (Bombarely et al. 2016) to validate this approach. Given the close relationships among the MYB activators (AN2, AN4, DPL, PHZ, ASR1, ASR2, ASR3), we took an additional step to confirm the accuracy of our double BLAST approach, building a gene tree for the entire set to determine how the recovered sequences are related to the characterized genes from *Petunia* (see supplemental results). After this step, we excluded AN4 and the ASR genes, recovered from 8 or fewer taxa each, probably due to their low expression at the sampled bud stage (Zhang et al. 2019). For the remaining 26 loci, we inferred final alignments of the peptide sequences with MAFFT (Katoh and Standley 2013) and used these to generate codon alignments of the corresponding nucleotide CDS with pal2nal.pl (Suyama et al. 2006). We inspected all alignments for spurious sequences, and in this process, removed one truncated and unalignable CHI-A sequence from *Hunzikeria texana*. Finally, we estimated maximum likelihood gene trees from the codon alignments with a GTR+Γ model in RaxML (Stamatakis 2014) for downstream analyses. All scripts for these bioinformatic steps, along with the assembled transcriptomes, are deposited online (https://osf.io/b7gcp/).

Estimation of evolutionary rates across genes and branches

We discarded five transcriptomes from our analyses of molecular evolution. *Brunfelsia pauciflora* (BRPA), *Brunfelsia plicata* (FBRP), and *Fabiana viscosa* (FAVI) were dropped because they are derived from leaf rather than floral tissue. *Calibrachoa parviflora* (CPAR) and *Calibrachoa excellens* (CEXC) were also discarded due to the lower quality of the sequencing data (see Supplemental methods). We used the codon alignments and corresponding gene trees for the remaining 65 species to analyze patterns of molecular evolution. We used HyPhy (Kosakovsky Pond et al. 2020) to fit a single ω and a free-rates model for each gene. To test whether the global ω values were driven by a subset of sites under positive selection, we fit the FUBAR model (Murrell et al. 2013) in HyPhy. We confirmed that the genes exhibit significant variation in ω using the BUSTED model in HyPhy to conduct pairwise comparisons across a subset of genes spanning the range of observed global ω values (supplementary text, Supplementary Material online). We also compared synonymous and non-synonymous rates across genes by summing estimated branch lengths for each (dN and dS) from HyPhy. We used the non-parametric Kruskal-Wallis H-test (Kruskal and Wallis 1952) and Kendall's τ (KENDALL 1938) to compare rates across gene types and pathway positions. For testing the association between molecular evolution and losses of floral anthocyanins, we used the RELAX approach (Wertheim et al. 2015), also implemented in HyPhy. For this analysis, we used the species tree (see below) and assigned species lacking floral anthocyanins to the foreground using the phylotree.js tool (http://phylotree.hyphy.org/). We used maximum parsimony to label internal foreground branches and then fit the RELAX model to each codon alignment to test for different dN/dS rate classes between foreground (unpigmented) and background (pigmented) branches. Raw data and scripts to run these analyses are deposited (https://osf.io/b7gcp/).

Quantification of anthocyanin content

We quantified the production of anthocyanins for each sampled Petunieae species with high-performance liquid chromatography (HPLC). We sampled flowers from three individuals per species and used these to calculate the mean anthocyanin mass fraction (mg compound per g tissue) over replicates. For each individual, we collected fresh floral corolla tissue, dried the tissue with silica gel and stored the material in 2mL tubes at -80°C as in Berardi et al. (2016). For extraction of total flavonoids, 0.0005 to 0.1g of dried tissue was soaked in 1mL 2N HCL overnight. Samples were then centrifuged (3 minutes at 12,000 RPM) to pellet tissue debris and the 1mL solvent was decanted into a new 2mL tube. Samples were heated at 100-104°C for 1 hr to convert the glycosylated flavonoids into their corresponding aglycones. 400μL of ethyl acetate was then added to each tube and vortexed thoroughly to mix the solution. Samples were centrifuged at 12,000 RPM for 1 min. The ethyl acetate layer (containing flavones and flavonols) was carefully removed using a micro-pipette. This ethyl acetate extraction wash was repeated a second time. Tubes containing the remaining HCl layer was then placed open-topped in an N-EVAP nitrogen evaporator connected to an air line in a fume hood to evaporate residual ethyl acetate. 150 μL of iso-amyl alcohol was then added to the tubes, and the solution was vortexed thoroughly to mix. Samples were again centrifuged at 12,000 RPM for 1 min. The iso-amyl alcohol layer (containing anthocyanidins) was carefully removed and pipetted into new 1.5 mL tubes and this iso-amyl alcohol extraction step was repeated a second time. The combined iso-amyl alcohol layers were then dried using an N-EVAP. Each extract was eluted in 50 μL of 1% HCl in MeOH before analyses. Before injecting onto the HPLC system, we assessed the overall concentration by examining a series of dilutions (1:75, 1:50,

1:10) on an Eppendorf BioSpectrometer and ensuring that peaks between 200-680 nm were under an absorbance of 3.0 (within linear range of the spec) to avoid overloading the column. After this QC step, 10 µL at the desired dilution (none or 1:10) were injected into an Agilent 1260 HPLC system. Anthocyanidins were separated by gradient elution at 30^oC using solvents A (HPLC-grade water, 0.1% trifluoroacetic acid) and C (1-propanol, 0.1% trifluoroacetic acid) with the following program: 15% C from 0 to 4 min; linear increase to 20% C from 4 to 10 min; 20% C from 10 to 14 min; linear increase to 22.5% C from 14 to 16 min; instantaneous increase to 27.5% C; 27.5% C from 16 to 18 min; instantaneous decrease to 15% C; 15%C from 18 to 21 min. Peaks were detected at 520 and 540 nm. A blank sample was run after every three samples and between species to wash the injection needle and avoid contamination. The mobile phase was 0.5% TFA in HPLC grade water and 1% HCL in MeOH and used a 100-4.6 mm Chromalith Performance column. All results were analyzed using Agilent Chemstation software and peaks were compared to standards obtained from Extrasynthese (360nm for flavonoids and 520nm for anthocyanidins). In total, we completed HPLC analyses for 58 Petunieae species.

Species tree estimation and PGLS analyses

We used the phylotranscriptomic pipeline developed by Yang et al. (2015) to estimate species relationships. This pipeline uses a combination of BLAST searches and tree-building steps to identify homologous gene clusters and estimate ML gene trees for input into species tree estimation programs (described in the supplementary text). To obtain an ultrametric tree (with branches proportional to time) for statistical comparative analyses, we estimated branch lengths from a sample of genes present in all species and used penalized likelihood as implemented in TreePL (Smith and OMeara 2012) to carry out rate smoothing (described in the supplementary text, Supplementary Material online). We used this ultrametric tree to estimate ancestral states for anthocyanin content with the fastAnc function of the phytools package (Revell 2012) and test for associations between these shifts and multiple aspects of molecular evolution. First, we used a PGLS to test for associations between tip estimates of ω for each gene from the free-rates model and anthocyanin amount with the gls function in the NLME package (Pinheiro et al. 2021). We set the Ornstein-Uhlenbeck parameter α to be freely estimated, allowing the degree of phylogenetic structure to vary across analyses. Second, we used PGLS to examine the relationship between gene presence/absence and anthocyanin content for each gene (with non-zero missing species). Finally, we repeated the PGLS analysis scoring anthocyanins as present or absent, which is equivalent to a phylogenetic ANOVA (Rohlf 2001).

Acknowledgments

We thank members of the Smith lab for helpful discussions on the project. We thank Aléxia Gope, Drielli Canal, Sebastián Guzman Rodríguez, Louana Susa, and Ana Lúcia Cunha Dornelles of UFGRS in Porto Alegre for their assistance getting around the lab. We thank iNaturalist users Karla M. Benítez and Lindheimer, Texas Master Naturalist Edith Bergquist for contributing their photos for figure 1. We thank many who helped us acquire plant material in botanic gardens and plant research institutes: Joseph Cahill of the Ventura Botanical Gardens, Julián Greppi of INTA Castellar in Buenos Aires, Chad Husby of Fairchild Tropical Botanic Garden, Bruce Holst, Shawn McCourt, and Sally Chambers of Marie Selby Botanical Gardens, and Mike Bone of Denver Botanic

Gardens. We are also grateful to Segundo Leiva, Victor Quipuscoa, and Carmen Fernandez for assistance with field collections. We thank Sergei Pond for his advice and rapid responses regarding the use of HyPhy. This work utilized the RMACC Summit supercomputer, which is supported by the National Science Foundation (awards ACI-1532235 and ACI-1532236), the University of Colorado Boulder, and Colorado State University. The Summit supercomputer is a joint effort of the University of Colorado Boulder and Colorado State University.

Funding

This work was funded by NSF-DEB 1553114 to SDS and the Gatsby Charitable Foundation to EM. The funders had no role in study design, data collection, analysis, decision to publish, or manuscript preparation.

Availability of data and materials

The supplemental scripts and processed data files (including transcriptome assemblies) needed to conduct the analyses referenced throughout the manuscript can be found in the supplementary OSF repo (https://osf.io/b7gcp/). The raw RNA-seq data files have been uploaded to the SRA (BioProject PRJNA746328, supplementary Table S1).

Author contributions

SDS, JN, LBDF, GEB, and LCW conceived the study and outlined the experimental design. LCW, JFW, EM, and SDS developed the analyses. LCW, SDS, JN, RD, ADW, AB, PHP, and MVP conducted fieldwork to collect plant samples. ADW performed HPLC. JFW reconstructed the species phylogeny. LCW built the sequencing libraries and assembled the *de novo* transcriptomes. AM, JFW, and LCW implemented the assembly pipeline. LCW and SDS conducted the statistical analyses of the data and drafted the manuscript with revisions from JFW and EM and additional edits from GEB, JN, and MVP.

Competing interests

The authors declare that they have no competing interests.

References

TransDecoder/TransDecoder, Apr. 2021. URL `https://github.com/TransDecoder/TransDecoder`. original-date: 2015-01-25T15:02:27Z.

Albert, N. W., Lewis, D. H., Zhang, H., Schwinn, K. E., Jameson, P. E., and Davies, K. M. Members of an R2R3-MYB transcription factor family in Petunia are developmentally and environmentally regulated to control complex floral and vegetative pigmentation patterning. *The Plant Journal*, 65(5):771–784, 2011. ISSN 1365-313X. doi: https://doi.org/10.1111/j.1365-313X.2010.04465.x. URL `https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1365-313X.2010.04465.x`. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1365-313X.2010.04465.x.

Albert, N. W., Davies, K. M., Lewis, D. H., Zhang, H., Montefiori, M., Brendolise, C., Boase, M. R., Ngo, H., Jameson, P. E., and Schwinn, K. E. A Conserved Network of Transcriptional Activators and Repressors Regulates Anthocyanin Pigmentation in Eudicots. *The Plant Cell*, 26(3):962–980, Mar. 2014. ISSN 1040-4651, 1532-298X. doi: 10.1105/tpc.113.122069. URL http://www.plantcell.org/content/26/3/962. Publisher: American Society of Plant Biologists Section: Research Article.

Alvarez-Ponce, D., Aguadé, M., and Rozas, J. Network-level molecular evolutionary analysis of the insulin/TOR signal transduction pathway across 12 Drosophila genomes. *Genome Research*, 19(2):234–242, Feb. 2009. ISSN 1088-9051. doi: 10.1101/gr.084038.108. URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2652205/.

Ascencio, D., Ochoa, S., Delaye, L., and DeLuna, A. Increased rates of protein evolution and asymmetric deceleration after the whole-genome duplication in yeasts. *BMC Evolutionary Biology*, 17(1):40, Feb. 2017. ISSN 1471-2148. doi: 10.1186/s12862-017-0895-1. URL https://doi.org/10.1186/s12862-017-0895-1.

Badawi, M., Giraud, I., Vavre, F., Grève, P., and Cordaux, R. Signs of Neutralization in a Redundant Gene Involved in Homologous Recombination in Wolbachia Endosymbionts. *Genome Biology and Evolution*, 6(10):2654–2664, Oct. 2014. ISSN 1759-6653. doi: 10.1093/gbe/evu207. URL https://doi.org/10.1093/gbe/evu207.

Baum, D. A., Yoon, H.-S., and Oldham, R. L. Molecular evolution of the transcription factor LEAFY in Brassicaceae. *Molecular Phylogenetics and Evolution*, 37(1):1–14, Oct. 2005. ISSN 1055-7903. doi: 10.1016/j.ympev.2005.07.002.

Berardi, A. E., Hildreth, S. B., Helm, R. F., Winkel, B. S. J., and Smith, S. D. Evolutionary correlations in flavonoid production across flowers and leaves in the Iochrominae (Solanaceae). *Phytochemistry*, 130:119–127, Oct. 2016. ISSN 0031-9422. doi: 10.1016/j.phytochem.2016.05.007. URL https://www.sciencedirect.com/science/article/pii/S0031942216301091.

Berardi, A. E., Esfeld, K., Jäggi, L., Mandel, T., Cannarozzi, G. M., and Kuhlemeier, C. Complex evolution of novel red floral color in Petunia. *The Plant Cell*, (koab114), Apr. 2021. ISSN 1040-4651. doi: 10.1093/plcell/koab114. URL https://doi.org/10.1093/plcell/koab114.

Bernhardt, C., Zhao, M., Gonzalez, A., Lloyd, A., and Schiefelbein, J. The bHLH genes GL3 and EGL3 participate in an intercellular regulatory circuit that controls cell patterning in the Arabidopsis root epidermis. *Development*, 132(2):291–298, Jan. 2005. ISSN 0950-1991. doi: 10.1242/dev.01565. URL https://doi.org/10.1242/dev.01565.

Blomberg, S. P., Garland, T., and Ives, A. R. Testing for Phylogenetic Signal in Comparative Data: Behavioral Traits Are More Labile. *Evolution*, 57(4):717–745, 2003. ISSN 1558-5646. doi: https://doi.org/10.1111/j.0014-3820.2003.tb00285.x. URL https://onlinelibrary.wiley.com/doi/abs/10.1111/j.0014-3820.2003.tb00285.x. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.0014-3820.2003.tb00285.x.

Boakye, D. W., Jaroenlak, P., Prachumwat, A., Williams, T. A., Bateman, K. S., Itsathitphaisarn, O., Sritunyalucksana, K., Paszkiewicz, K. H., Moore, K. A., Stentiford, G. D., and Williams, B. A. P. Decay of the glycolytic pathway and adaptation to intranuclear

parasitism within Enterocytozoonidae microsporidia. *Environmental Microbiology*, 19(5): 2077–2089, 2017. ISSN 1462-2920. doi: https://doi.org/10.1111/1462-2920.13734. URL https://sfamjournals.onlinelibrary.wiley.com/doi/abs/10.1111/1462-2920.13734. _eprint: https://sfamjournals.onlinelibrary.wiley.com/doi/pdf/10.1111/1462-2920.13734.

Bolger, A. M., Lohse, M., and Usadel, B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*, 30(15):2114–2120, Aug. 2014. ISSN 1367-4803. doi: 10.1093/bioinformatics/btu170. URL https://doi.org/10.1093/bioinformatics/btu170.

Bombarely, A., Moser, M., Amrad, A., Bao, M., Bapaume, L., Barry, C. S., Bliek, M., Boersma, M. R., Borghi, L., Bruggmann, R., Bucher, M., D'Agostino, N., Davies, K., Druege, U., Dudareva, N., Egea-Cortines, M., Delledonne, M., Fernandez-Pozo, N., Franken, P., Grandont, L., Heslop-Harrison, J. S., Hintzsche, J., Johns, M., Koes, R., Lv, X., Lyons, E., Malla, D., Martinoia, E., Mattson, N. S., Morel, P., Mueller, L. A., Muhlemann, J., Nouri, E., Passeri, V., Pezzotti, M., Qi, Q., Reinhardt, D., Rich, M., Richert-Pöggeler, K. R., Robbins, T. P., Schatz, M. C., Schranz, M. E., Schuurink, R. C., Schwarzacher, T., Spelt, K., Tang, H., Urbanus, S. L., Vandenbussche, M., Vijverberg, K., Villarino, G. H., Warner, R. M., Weiss, J., Yue, Z., Zethof, J., Quattrocchio, F., Sims, T. L., and Kuhlemeier, C. Insight into the evolution of the Solanaceae from the parental genomes of Petunia hybrida. *Nature Plants*, 2(6):1–9, May 2016. ISSN 2055-0278. doi: 10.1038/nplants.2016.74. URL https://www.nature.com/articles/nplants201674. Number: 6 Publisher: Nature Publishing Group.

Carroll, S. B. Evo-Devo and an Expanding Evolutionary Synthesis: A Genetic Theory of Morphological Evolution. *Cell*, 134(1):25–36, July 2008. ISSN 0092-8674. doi: 10.1016/j.cell.2008.06.030. URL https://www.sciencedirect.com/science/article/pii/S0092867408008179.

Ciezarek, A. G., Osborne, O. G., Shipley, O. N., Brooks, E. J., Tracey, S. R., McAllister, J. D., Gardner, L. D., Sternberg, M. J. E., Block, B., and Savolainen, V. Phylotranscriptomic Insights into the Diversification of Endothermic Thunnus Tunas. *Molecular Biology and Evolution*, 36(1):84–96, Jan. 2019. ISSN 0737-4038. doi: 10.1093/molbev/msy198. URL https://doi.org/10.1093/molbev/msy198.

Cole, C. T. and Ingvarsson, P. K. Pathway position constrains the evolution of an ecologically important pathway in aspens (Populus tremula L.). *Molecular Ecology*, 27(16):3317–3330, 2018. ISSN 1365-294X. doi: https://doi.org/10.1111/mec.14785. URL http://onlinelibrary.wiley.com/doi/abs/10.1111/mec.14785. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/mec.14785.

Czechowski, T., Bari, R. P., Stitt, M., Scheible, W.-R., and Udvardi, M. K. Real-time RT-PCR profiling of over 1400 Arabidopsis transcription factors: unprecedented sensitivity reveals novel root- and shoot-specific genes. *The Plant Journal*, 38(2):366–379, 2004. ISSN 1365-313X. doi: https://doi.org/10.1111/j.1365-313X.2004.02051.x. URL https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1365-313X.2004.02051.x. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1365-313X.2004.02051.x.

DallOlio, G. M., Laayouni, H., Luisi, P., Sikora, M., Montanucci, L., and Bertranpetit, J. Distribution of events of positive selection and population differentiation in a metabolic pathway: the case of asparagine N-glycosylation. *BMC Evolutionary Biology*, 12:98, June 2012. ISSN 1471-2148. doi: 10.1186/1471-2148-12-98. URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3426484/.

Davidson, N. M. and Oshlack, A. Corset: enabling differential gene expression analysis for de novoassembled transcriptomes. *Genome Biology*, 15(7):410, July 2014. ISSN 1474-760X. doi: 10.1186/s13059-014-0410-6. URL https://doi.org/10.1186/s13059-014-0410-6.

Davies, K. M., Albert, N. W., Schwinn, K. E., Davies, K. M., Albert, N. W., and Schwinn, K. E. From landing lights to mimicry: the molecular regulation of flower colouration and mechanisms for pigmentation patterning. *Functional Plant Biology*, 39(8):619–638, Sept. 2012. ISSN 1445-4416, 1445-4416. doi: 10.1071/FP12195. URL https://www.publish.csiro.au/fp/FP12195. Publisher: CSIRO PUBLISHING.

Deng, C., Cheng, C.-H. C., Ye, H., He, X., and Chen, L. Evolution of an antifreeze protein by neofunctionalization under escape from adaptive conflict. *Proceedings of the National Academy of Sciences*, 107(50):21593–21598, Dec. 2010. ISSN 0027-8424, 1091-6490. doi: 10.1073/pnas.1007883107. URL https://www.pnas.org/content/107/50/21593. Publisher: National Academy of Sciences Section: Biological Sciences.

dePamphilis, C. W. and Palmer, J. D. Loss of photosynthetic and chlororespiratory genes from the plastid genome of a parasitic flowering plant. *Nature*, 348(6299):337–339, Nov. 1990. ISSN 1476-4687. doi: 10.1038/348337a0. URL https://www.nature.com/articles/348337a0. Number: 6299 Publisher: Nature Publishing Group.

Ding, B., Patterson, E. L., Holalu, S. V., Li, J., Johnson, G. A., Stanley, L. E., Greenlee, A. B., Peng, F., Bradshaw, H. D., Blinov, M. L., Blackman, B. K., and Yuan, Y.-W. Two MYB Proteins in a Self-Organizing Activator-Inhibitor System Produce Spotted Pigmentation Patterns. *Current Biology*, 30(5):802–814.e8, Mar. 2020. ISSN 0960-9822. doi: 10.1016/j.cub.2019.12.067. URL https://www.sciencedirect.com/science/article/pii/S0960982219317002.

Doebley, J. and Lukens, L. Transcriptional Regulators and the Evolution of Plant Form. *The Plant Cell*, 10(7):1075–1082, July 1998. ISSN 1040-4651, 1532-298X. doi: 10.1105/tpc.10.7.1075. URL http://www.plantcell.org/content/10/7/1075. Publisher: American Society of Plant Biologists Section: Review Article.

Drummond, D. A. and Wilke, C. O. Mistranslation-Induced Protein Misfolding as a Dominant Constraint on Coding-Sequence Evolution. *Cell*, 134(2):341–352, July 2008. ISSN 0092-8674, 1097-4172. doi: 10.1016/j.cell.2008.05.042. URL https://www.cell.com/cell/abstract/S0092-8674(08)00705-8. Publisher: Elsevier.

Drummond, D. A., Raval, A., and Wilke, C. O. A single determinant dominates the rate of yeast protein evolution. *Molecular Biology and Evolution*, 23(2):327–337, Feb. 2006. ISSN 0737-4038. doi: 10.1093/molbev/msj038.

Duret, L. and Mouchiroud, D. Determinants of substitution rates in mammalian genes: expression pattern affects selection intensity but not mutation rate. *Molecular Biology and Evolution*, 17 (1):68–74, Jan. 2000. ISSN 0737-4038. doi: 10.1093/oxfordjournals.molbev.a026239.

Esfeld, K., Berardi, A. E., Moser, M., Bossolini, E., Freitas, L., and Kuhlemeier, C. Pseudogenization and Resurrection of a Speciation Gene. *Current biology: CB*, 28(23):3776–3786.e7, Dec. 2018. ISSN 1879-0445. doi: 10.1016/j.cub.2018.10.019.

Facciuto, G., Pannunzio, M., Coviella, A., Bologna, P., Soto, S., Imhof, L., and Borja, M. CALIBRACHOA BREEDING ADVANCES IN ARGENTINA. *Acta Horticulturae*, (813):121–126, Mar. 2009. ISSN 0567-7572, 2406-6168. doi: 10.17660/ActaHortic.2009.813.15. URL https://www.actahort.org/books/813/813_15.htm.

Filipowicz, N., Nee, M., and Renner, S. Description and molecular diagnosis of a new species of Brunfelsia (Solanaceae) from the Bolivian and Argentinean Andes. *PhytoKeys*, 10:83–94, Mar. 2012. ISSN 1314-2003. doi: 10.3897/phytokeys.10.2558. URL https://phytokeys.pensoft.net/article/1408/. Publisher: Pensoft Publishers.

Flowers, J., Sezgin, E., Kumagai, S., Duvernell, D., Matzkin, L., Schmidt, P., and Eanes, W. Adaptive Evolution of Metabolic Pathways in Drosophila. *Molecular Biology and Evolution*, 24(6):1347–1354, June 2007. ISSN 0737-4038. doi: 10.1093/molbev/msm057. URL https://doi.org/10.1093/molbev/msm057.

Foote, A. D., Liu, Y., Thomas, G. W., Vina, T., Alföldi, J., Deng, J., Dugan, S., van Elk, C. E., Hunter, M. E., Joshi, V., Khan, Z., Kovar, C., Lee, S. L., Lindblad-Toh, K., Mancia, A., Nielsen, R., Qin, X., Qu, J., Raney, B. J., Vijay, N., Wolf, J. B. W., Hahn, M. W., Muzny, D. M., Worley, K. C., Gilbert, M. T. P., and Gibbs, R. A. Convergent evolution of the genomes of marine mammals. *Nature genetics*, 47(3):272–275, Mar. 2015. ISSN 1061-4036. doi: 10.1038/ng.3198. URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4644735/.

Fregonezi, J. N., de Freitas, L. B., Bonatto, S. L., Semir, J., and Stehmann, J. R. Infrageneric classification of Calibrachoa (Solanaceae) based on morphological and molecular evidence. *Taxon*, 61(1):120–130, 2012. ISSN 0040-0262. URL https://www.jstor.org/stable/23210319. Publisher: International Association for Plant Taxonomy (IAPT).

Fregonezi, J. N., Turchetto, C., Bonatto, S. L., and Freitas, L. B. Biogeographical history and diversification of Petunia and Calibrachoa (Solanaceae) in the Neotropical Pampas grassland. *Botanical Journal of the Linnean Society*, 171(1):140–153, Jan. 2013. ISSN 0024-4074. doi: 10.1111/j.1095-8339.2012.01292.x. URL https://doi.org/10.1111/j.1095-8339.2012.01292.x.

Gates, D. J., Strickler, S. R., Mueller, L. A., Olson, B. J. S. C., and Smith, S. D. Diversification of R2R3-MYB Transcription Factors in the Tomato Family Solanaceae. *Journal of Molecular Evolution*, 83(1-2):26–37, Aug. 2016. ISSN 1432-1432. doi: 10.1007/s00239-016-9750-z.

Ghaemmaghami, S., Huh, W.-K., Bower, K., Howson, R. W., Belle, A., Dephoure, N., O'Shea, E. K., and Weissman, J. S. Global analysis of protein expression in yeast. *Nature*, 425(6959):737–741, Oct. 2003. ISSN 1476-4687. doi: 10.1038/nature02046. URL https://www.nature.com/articles/nature02046. Number: 6959 Publisher: Nature Publishing Group.

Gould, K. S. Nature's Swiss Army Knife: The Diverse Protective Roles of Anthocyanins in Leaves. *Journal of Biomedicine and Biotechnology*, 2004(5):314–320, Dec. 2004. ISSN 1110-7243. doi: 10.1155/S1110724304406147. URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1082902/.

Grabherr, M. G., Haas, B. J., Yassour, M., Levin, J. Z., Thompson, D. A., Amit, I., Adiconis, X., Fan, L., Raychowdhury, R., Zeng, Q., Chen, Z., Mauceli, E., Hacohen, N., Gnirke,

A., Rhind, N., di Palma, F., Birren, B. W., Nusbaum, C., Lindblad-Toh, K., Friedman, N., and Regev, A. Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nature Biotechnology*, 29(7):644–652, July 2011. ISSN 1546-1696. doi: 10.1038/nbt.1883. URL https://www.nature.com/articles/nbt.1883. Number: 7 Publisher: Nature Publishing Group.

Greenberg, A. J., Stockwell, S. R., and Clark, A. G. Evolutionary Constraint and Adaptation in the Metabolic Network of Drosophila. *Molecular Biology and Evolution*, 25 (12):2537–2546, Dec. 2008. ISSN 0737-4038. doi: 10.1093/molbev/msn205. URL https://doi.org/10.1093/molbev/msn205.

Hahn, M. W. and Kern, A. D. Comparative Genomics of Centrality and Essentiality in Three Eukaryotic Protein-Interaction Networks. *Molecular Biology and Evolution*, 22(4):803–806, Apr. 2005. ISSN 0737-4038. doi: 10.1093/molbev/msi072. URL https://doi.org/10.1093/molbev/msi072.

Halabi, K., Karin, E. L., Guéguen, L., and Mayrose, I. A Codon Model for Associating Phenotypic Traits with Altered Selective Patterns of Sequence Evolution. *Systematic Biology*, 70(3):608–622, May 2021. ISSN 1063-5157. doi: 10.1093/sysbio/syaa087. URL https://doi.org/10.1093/sysbio/syaa087.

Ho, W. W. and Smith, S. D. Molecular evolution of anthocyanin pigmentation genes following losses of flower color. *BMC Evolutionary Biology*, 16(1):98, May 2016. ISSN 1471-2148. doi: 10.1186/s12862-016-0675-3. URL https://doi.org/10.1186/s12862-016-0675-3.

Hoballah, M. E., Gübitz, T., Stuurman, J., Broger, L., Barone, M., Mandel, T., Dell'Olivo, A., Arnold, M., and Kuhlemeier, C. Single GeneMediated Shift in Pollinator Attraction in Petunia. *The Plant Cell*, 19(3):779–790, Mar. 2007. ISSN 1040-4651, 1532-298X. doi: 10.1105/tpc.106.048694. URL http://www.plantcell.org/content/19/3/779. Publisher: American Society of Plant Biologists Section: Research Article.

Hu, Z., Sackton, T. B., Edwards, S. V., and Liu, J. S. Bayesian Detection of Convergent Rate Changes of Conserved Noncoding Elements on Phylogenetic Trees. *Molecular Biology and Evolution*, 36(5):1086–1100, May 2019. ISSN 0737-4038. doi: 10.1093/molbev/msz049. URL https://doi.org/10.1093/molbev/msz049.

Invergo, B. M., Montanucci, L., Laayouni, H., and Bertranpetit, J. A system-level, molecular evolutionary analysis of mammalian phototransduction. *BMC evolutionary biology*, 13:52, Feb. 2013. ISSN 1471-2148. doi: 10.1186/1471-2148-13-52.

Jiang, C.-K. and Rao, G.-Y. Insights into the Diversification and Evolution of R2R3-MYB Transcription Factors in Plants. *Plant Physiology*, 183(2):637–655, June 2020. ISSN 0032-0889, 1532-2548. doi: 10.1104/pp.19.01082. URL http://www.plantphysiol.org/content/183/2/637. Publisher: American Society of Plant Biologists Section: Research Article.

Jovelin, R. and Phillips, P. C. Expression Level Drives the Pattern of Selective Constraints along the Insulin/Tor Signal Transduction Pathway in Caenorhabditis. *Genome Biology and Evolution*, 3:715–722, Aug. 2011. ISSN 1759-6653. doi: 10.1093/gbe/evr071. URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3157841/.

Katoh, K. and Standley, D. M. MAFFT Multiple Sequence Alignment Software Version 7: Improvements in Performance and Usability. *Molecular Biology and Evolution*, 30(4):772–780, Apr. 2013. ISSN 0737-4038. doi: 10.1093/molbev/mst010. URL https://doi.org/10.1093/molbev/mst010.

KENDALL, M. G. A NEW MEASURE OF RANK CORRELATION. *Biometrika*, 30 (1-2):81–93, June 1938. ISSN 0006-3444. doi: 10.1093/biomet/30.1-2.81. URL https://doi.org/10.1093/biomet/30.1-2.81.

Kiefer, C., Willing, E.-M., Jiao, W.-B., Sun, H., Piednoël, M., Hümann, U., Hartwig, B., Koch, M. A., and Schneeberger, K. Interspecies association mapping links reduced CG to TG substitution rates to the loss of gene-body methylation. *Nature Plants*, 5(8):846–855, Aug. 2019. ISSN 2055-0278. doi: 10.1038/s41477-019-0486-9. URL https://www.nature.com/articles/s41477-019-0486-9. Number: 8 Publisher: Nature Publishing Group.

Koes, R. E., Spelt, C. E., and Mol, J. N. The chalcone synthase multigene family of Petunia hybrida (V30): differential, light-regulated expression during flower development and UV light induction. *Plant Molecular Biology*, 12(2):213–225, Feb. 1989a. ISSN 0167-4412. doi: 10.1007/BF00020506.

Koes, R. E., Spelt, C. E., and Mol, J. N. M. The chalcone synthase multigene family of Petunia hybrida (V30): differential, light-regulated expression during flower development and UV light induction. *Plant Molecular Biology*, 12(2):213–225, Feb. 1989b. ISSN 1573-5028. doi: 10.1007/BF00020506. URL https://doi.org/10.1007/BF00020506.

Kosakovsky Pond, S. L., Poon, A. F. Y., Velazquez, R., Weaver, S., Hepler, N. L., Murrell, B., Shank, S. D., Magalis, B. R., Bouvier, D., Nekrutenko, A., Wisotsky, S., Spielman, S. J., Frost, S. D. W., and Muse, S. V. HyPhy 2.5A Customizable Platform for Evolutionary Hypothesis Testing Using Phylogenies. *Molecular Biology and Evolution*, 37(1):295–299, Jan. 2020. ISSN 0737-4038. doi: 10.1093/molbev/msz197. URL https://doi.org/10.1093/molbev/msz197.

Kruskal, W. H. and Wallis, W. A. Use of Ranks in One-Criterion Variance Analysis. *Journal of the American Statistical Association*, 47(260):583–621, Dec. 1952. ISSN 0162-1459. doi: 10.1080/01621459.1952.10483441. URL https://www.tandfonline.com/doi/abs/10.1080/01621459.1952.10483441. Publisher: Taylor & Francis _eprint: https://www.tandfonline.com/doi/pdf/10.1080/01621459.1952.10483441.

Lagercrantz, U. and Axelsson, T. Rapid evolution of the family of CONSTANS LIKE genes in plants. *Molecular Biology and Evolution*, 17(10):1499–1507, Oct. 2000. ISSN 0737-4038. doi: 10.1093/oxfordjournals.molbev.a026249.

Larter, M., Dunbar-Wallis, A., Berardi, A. E., and Smith, S. D. Convergent Evolution at the Pathway Level: Predictable Regulatory Changes during Flower Color Transitions. *Molecular Biology and Evolution*, 35(9):2159–2169, Sept. 2018. ISSN 0737-4038. doi: 10.1093/molbev/msy117. URL https://doi.org/10.1093/molbev/msy117.

Larter, M., DunbarWallis, A., Berardi, A. E., and Smith, S. D. Developmental control of convergent floral pigmentation across evolutionary timescales. *Developmental Dynamics*, 248(11):1091–1100, 2019. ISSN 1097-0177. doi: https://doi.org/10.1002/dvdy.82.

URL http://anatomypubs.onlinelibrary.wiley.com/doi/abs/10.1002/dvdy.82. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/dvdy.82.

Liu, X. and Clarke, N. D. Rationalization of Gene Regulation by a Eukaryotic Transcription Factor: Calculation of Regulatory Region Occupancy from Predicted Binding Affinities. *Journal of Molecular Biology*, 323(1):1–8, Oct. 2002. ISSN 0022-2836. doi: 10.1016/S0022-2836(02)00894-X. URL https://www.sciencedirect.com/science/article/pii/S002228360200894X.

Livingstone, K. and Anderson, S. Patterns of Variation in the Evolution of Carotenoid Biosynthetic Pathway Enzymes of Higher Plants. *Journal of Heredity*, 100(6):754–761, Nov. 2009. ISSN 0022-1503. doi: 10.1093/jhered/esp026. URL https://doi.org/10.1093/jhered/esp026.

Lu, Y. and Rausher, M. D. Evolutionary Rate Variation in Anthocyanin Pathway Genes. *Molecular Biology and Evolution*, 20(11):1844–1853, Nov. 2003. ISSN 0737-4038. doi: 10.1093/molbev/msg197. URL https://doi.org/10.1093/molbev/msg197.

Lynch, M. and Conery, J. S. The Evolutionary Fate and Consequences of Duplicate Genes. *Science*, 290(5494):1151–1155, Nov. 2000. ISSN 0036-8075, 1095-9203. doi: 10.1126/science.290.5494.1151. URL https://science.sciencemag.org/content/290/5494/1151. Publisher: American Association for the Advancement of Science Section: Report.

Madden, T. *The BLAST Sequence Analysis Tool*. National Center for Biotechnology Information (US), Mar. 2013. URL https://www.ncbi.nlm.nih.gov/sites/books/NBK153387/. Publication Title: The NCBI Handbook [Internet]. 2nd edition.

Mäder, G. and Freitas, L. B. Biogeographical, ecological, and phylogenetic analyses clarifying the evolutionary history of Calibrachoa in South American grasslands. *Molecular Phylogenetics and Evolution*, 141:106614, Dec. 2019. ISSN 1055-7903. doi: 10.1016/j.ympev.2019.106614. URL https://www.sciencedirect.com/science/article/pii/S1055790319304178.

Marshall, C. R., Raff, E. C., and Raff, R. A. Dollo's law and the death and resurrection of genes. *Proceedings of the National Academy of Sciences*, 91(25):12283–12287, Dec. 1994. ISSN 0027-8424, 1091-6490. doi: 10.1073/pnas.91.25.12283. URL https://www.pnas.org/content/91/25/12283. Publisher: National Academy of Sciences Section: Research Article.

Martins, T. R., Jiang, P., and Rausher, M. D. How petals change their spots: cis-regulatory re-wiring in Clarkia (Onagraceae). *New Phytologist*, 216(2):510–518, 2017. ISSN 1469-8137. doi: https://doi.org/10.1111/nph.14163. URL https://nph.onlinelibrary.wiley.com/doi/abs/10.1111/nph.14163. _eprint: https://nph.onlinelibrary.wiley.com/doi/pdf/10.1111/nph.14163.

Masalia, R. R., Bewick, A. J., and Burke, J. M. Connectivity in gene coexpression networks negatively correlates with rates of molecular evolution in flowering plants. *PLOS ONE*, 12(7):e0182289, July 2017. ISSN 1932-6203. doi: 10.1371/journal.pone.0182289. URL https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0182289. Publisher: Public Library of Science.

Meredith, R. W., Gatesy, J., and Springer, M. S. Molecular decay of enamel matrix protein genes in turtles and other edentulous amniotes. *BMC Evolutionary Biology*, 13(1):20, Jan. 2013. ISSN 1471-2148. doi: 10.1186/1471-2148-13-20. URL https://doi.org/10.1186/1471-2148-13-20.

Montanucci, L., Laayouni, H., Dobon, B., Keys, K. L., Bertranpetit, J., and Peretó, J. Influence of pathway topology and functional class on the molecular evolution of human metabolic genes. *PLOS ONE*, 13(12):e0208782, Dec. 2018. ISSN 1932-6203. doi: 10.1371/journal.pone.0208782. URL https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0208782. Publisher: Public Library of Science.

Mukherjee, D., Mukherjee, A., and Ghosh, T. C. Evolutionary Rate Heterogeneity of Primary and Secondary Metabolic Pathway Genes in Arabidopsis thaliana. *Genome Biology and Evolution*, 8(1):17–28, Jan. 2016. ISSN 1759-6653. doi: 10.1093/gbe/evv217. URL https://doi.org/10.1093/gbe/evv217.

Muntané, G., Farré, X., Rodríguez, J. A., Pegueroles, C., Hughes, D. A., de Magalhães, J. P., Gabaldón, T., and Navarro, A. Biological Processes Modulating Longevity across Primates: A Phylogenetic Genome-Phenome Analysis. *Molecular Biology and Evolution*, 35(8):1990–2004, Aug. 2018. ISSN 1537-1719. doi: 10.1093/molbev/msy105.

Murrell, B., Moola, S., Mabona, A., Weighill, T., Sheward, D., Kosakovsky Pond, S. L., and Scheffler, K. FUBAR: A Fast, Unconstrained Bayesian AppRoximation for Inferring Selection. *Molecular Biology and Evolution*, 30(5):1196–1205, May 2013. ISSN 0737-4038. doi: 10.1093/molbev/mst030. URL https://doi.org/10.1093/molbev/mst030.

Ng, J. and Smith, S. D. Widespread flower color convergence in Solanaceae via alternate biochemical pathways. *New Phytologist*, 209(1):407–417, 2016. ISSN 1469-8137. doi: 10.1111/nph.13576. URL https://nph.onlinelibrary.wiley.com/doi/abs/10.1111/nph.13576. _eprint: https://nph.onlinelibrary.wiley.com/doi/pdf/10.1111/nph.13576.

Olmstead, R. G., Bohs, L., Migid, H. A., Santiago-Valentin, E., Garcia, V. F., and Collier, S. M. A molecular phylogeny of the Solanaceae. *TAXON*, 57(4):1159–1181, 2008. ISSN 1996-8175. doi: 10.1002/tax.574010. URL https://onlinelibrary.wiley.com/doi/abs/10.1002/tax.574010. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/tax.574010.

Olson-Manning, C. F., Lee, C.-R., Rausher, M. D., and Mitchell-Olds, T. Evolution of flux control in the glucosinolate pathway in Arabidopsis thaliana. *Molecular Biology and Evolution*, 30(1): 14–23, Jan. 2013. ISSN 1537-1719. doi: 10.1093/molbev/mss204.

Pál, C., Papp, B., and Hurst, L. D. Highly expressed genes in yeast evolve slowly. *Genetics*, 158 (2):927–931, June 2001. ISSN 0016-6731.

Passow, C. N., Bronikowski, A. M., Blackmon, H., Parsai, S., Schwartz, T. S., and McGaugh, S. E. Contrasting Patterns of Rapid Molecular Evolution within the p53 Network across Mammal and Sauropsid Lineages. *Genome Biology and Evolution*, 11(3):629–643, Mar. 2019. ISSN 1759-6653. doi: 10.1093/gbe/evy273. URL https://doi.org/10.1093/gbe/evy273.

Patro, R., Duggal, G., Love, M. I., Irizarry, R. A., and Kingsford, C. Salmon: fast and bias-aware quantification of transcript expression using dual-phase inference. *Nature methods*, 14(4):417–419, Apr. 2017. ISSN 1548-7091. doi: 10.1038/nmeth.4197. URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5600148/.

Pinheiro, J., Bates, D., DebRoy, S., Sarkar, D., and R Core Team. nlme: Linear and nonlinear mixed effects models. manual, 2021. URL https://CRAN.R-project.org/package=nlme.

Poon, A. and Chao, L. The rate of compensatory mutation in the DNA bacteriophage phiX174. *Genetics*, 170(3):989–999, July 2005. ISSN 0016-6731. doi: 10.1534/genetics.104.039438.

Preston, J. C., Martinez, C. C., and Hileman, L. C. Gradual disintegration of the floral symmetry gene network is implicated in the evolution of a wind-pollination syndrome. *Proceedings of the National Academy of Sciences*, 108(6):2343–2348, Feb. 2011. ISSN 0027-8424, 1091-6490. doi: 10.1073/pnas.1011361108. URL https://www.pnas.org/content/108/6/2343. ISBN: 9781011361106 Publisher: National Academy of Sciences Section: Biological Sciences.

Protas, M., Conrad, M., Gross, J. B., Tabin, C., and Borowsky, R. Regressive Evolution in the Mexican Cave Tetra, Astyanax mexicanus. *Current Biology*, 17 (5):452–454, Mar. 2007. ISSN 0960-9822. doi: 10.1016/j.cub.2007.01.051. URL https://www.sciencedirect.com/science/article/pii/S0960982207008901.

Prudent, X., Parra, G., Schwede, P., Roscito, J. G., and Hiller, M. Controlling for Phylogenetic Relatedness and Evolutionary Rates Improves the Discovery of Associations Between Species Phenotypic and Genomic Differences. *Molecular Biology and Evolution*, 33(8):2135–2150, Aug. 2016. ISSN 0737-4038. doi: 10.1093/molbev/msw098. URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4948712/.

Purugganan, M. D. and Wessler, S. R. Molecular evolution of the plant R regulatory gene family. *Genetics*, 138(3):849–854, Nov. 1994. ISSN 0016-6731.

Quattrocchio, F., Wing, J., Leppen, H., Mol, J., and Koes, R. Regulatory Genes Controlling Anthocyanin Pigmentation Are Functionally Conserved among Plant Species and Have Distinct Sets of Target Genes. *The Plant Cell*, 5(11):1497–1512, Nov. 1993. ISSN 1040-4651. URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC160381/.

Quattrocchio, F., Wing, J., Woude, K. v. d., Souer, E., Vetten, N. d., Mol, J., and Koes, R. Molecular Analysis of the anthocyanin2 Gene of Petunia and Its Role in the Evolution of Flower Color. *The Plant Cell*, 11(8):1433–1444, Aug. 1999. ISSN 1040-4651, 1532-298X. doi: 10.1105/tpc.11.8.1433. URL http://www.plantcell.org/content/11/8/1433. Publisher: American Society of Plant Biologists Section: Research Article.

Quattrocchio, F., Baudry, A., Lepiniec, L., and Grotewold, E. The Regulation of Flavonoid Biosynthesis. In Grotewold, E., editor, *The Science of Flavonoids*, pages 97–122. Springer, New York, NY, 2006. ISBN 978-0-387-28822-2. doi: 10.1007/978-0-387-28822-2$_4$.*URL*