Full Title: Coding-sequence evolution does not explain divergence in petal anthocyanin pigmentation between *Mimulus luteus* var. *luteus* and *M. l. variegatus*

Short Title: Monkeyflower pigment expansion not due to coding-sequence evolution

Walker E. Orr[1]
Ji Yang Kim[1]
Tejas Raj[2]
Ellen K. Hom[1]
Ashley E. Person[1]
Anne Vonada[1]
John A. Stratton[2]
Arielle M. Cooley[1,*]

[1]Whitman College Biology Department, Walla Walla WA, 99362, U.S.A.
[2]Whitman College Computer Science Department, Walla Walla WA, 99362, U.S.A.
*Corresponding author

Abstract

Phenotypic transitions in related taxa often share a common genetic basis, which suggests that there are constraints that shape the process of evolution at the genetic level. For example, noncoding changes in a gene might be favored relative to coding changes due to being less constrained by pleiotropic effects. Here we evaluate the importance of coding-sequence changes to the recent evolution of a novel anthocyanin pigmentation trait in the monkeyflower genus *Mimulus*. The magenta-flowered *Mimulus luteus* var. *variegatus* recently gained petal

lobe anthocyanin pigmentation via a single-locus Mendelian difference from its sister taxon, the yellow-flowered *M. l. luteus*. Previous work showed that the differentially expressed transcription factor gene *MYB5a/NEGAN* is the single causal gene. However, it was not clear whether *MYB5a* coding-sequence evolution (in addition to the observed patterns of differential expression) might also have contributed to increased anthocyanin production in *M. l. variegatus*. Quantitative image analysis of tobacco leaves, transfected with highly expressed *MYB5a* coding sequence from each taxon, revealed robust anthocyanin production driven by both alleles compared to a negative control. Counter to expectations, significantly higher anthocyanin production was driven by the coding sequence from the low-anthocyanin taxon *M. l. luteus*. Together with previously-published expression studies, this supports the hypothesis that petal pigment in *M. l. variegatus* was not gained by protein-coding changes, but instead via non-coding *cis*-regulatory evolution. Finally, while constructing the transgenes needed for this experiment, we unexpectedly discovered two sites in *MYB5a* that appear to be post-transcriptionally edited – a phenomenon that has been rarely reported, and even less often explored, for nuclear-encoded plant mRNAs.

260 words

Keywords: anthocyanin regulation; floral color patterning; gene expression; *cis*-regulatory evolution; R2R3 MYBs; *Mimulus*; petal lobe pigmentation; post-transcriptional mRNA editing; A-to-I editing; transient transformation; digital image analysis

Introduction

To what extent are the molecular mechanisms of evolutionary diversification predictable? Biologists have long been interested in understanding constraints on the evolutionary process, which can increase our ability to predict the molecular mechanisms that underlie a specific trait.

53 For example, coding and noncoding mutations have been hypothesized to differ in their
54 contributions to evolutionary change (Hoekstra and Coyne, 2007; Stern and Orgogozo, 2008).
55 Noncoding "*cis*-regulatory" regions integrate upstream signals to determine the conditions under
56 which a protein will be expressed. These regions tend to be modular, so mutations can easily
57 alter the gene expression driven by one *cis*-regulatory region without altering expression patterns
58 driven by neighboring regions (Wray, 2007). By contrast, a change in a gene's coding sequence
59 tends to be more pleiotropic since it will likely alter the gene product's function in all contexts
60 under which that gene is expressed. Stern and Orgogozo (2008) proposed a model predicting
61 that the contribution of coding changes to phenotypic transitions will decline with evolutionary
62 distance between the diverged taxa, due to lower average fitness values for coding changes
63 compared to noncoding changes. The pattern has been documented empirically by Wittkopp et
64 al. (2008) in *Drosophila*.
65
66 In plants, transitions in anthocyanin pigmentation are especially well suited to investigating the
67 mechanisms of molecular evolution (Davies et al., 2012; Sobel and Streisfeld, 2013; LaFountain
68 and Yuan, 2021; Li et al., 2022). Anthocyanins are a class of flavonoids that are responsible for
69 purple, pink, and red colors in diverse angiosperm tissues, including leaves, seed coats, and
70 flowers (Durbin et al., 2003). In the model plant *Arabidopsis thaliana* and many other species, the
71 anthocyanin pathway consists of "early" genes *CHS, CHI,* and *F3'H*, followed by the "late" genes
72 *DFR, ANS,* and *UF3GT* (Dubos et al., 2010). The "late" genes are coordinately activated by MYB-
73 bHLH-WD40 (MBW) transcription factor complexes (Grotewold, 2006). Much of the target
74 specificity of an MBW complex is conferred by the MYB partner in the complex: different tissues
75 make use of different *MYB* genes to stimulate anthocyanin production, but may use the same
76 *bHLH* gene (Quattrocchio et al., 2006). *MYB* genes thus serve as "input-output" integrators that
77 regulate gene batteries under highly specific contexts. Consistent with their position in gene
78 networks and the hypothesized importance of "input-output" genes in repeated evolution events,

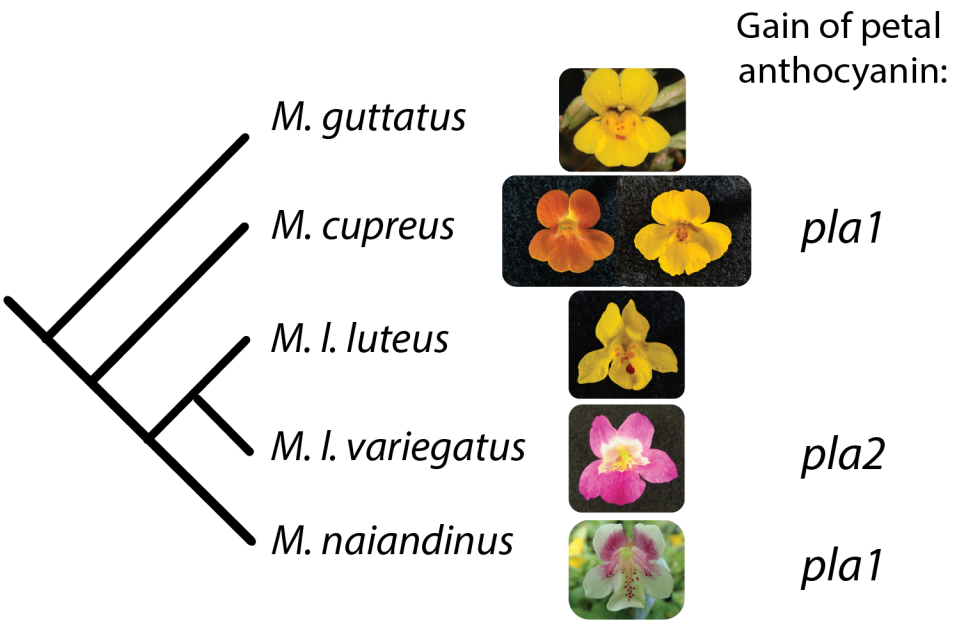
MYB genes demonstrate remarkable reuse in floral pigment transitions across a variety of species (Streisfeld and Rausher, 2011; Yuan et al., 2013).

The independent evolution of petal lobe anthocyanin (PLA) pigmentation in three lineages in the *luteus* group of the monkeyflower genus *Mimulus* (synonym *Erythranthe*; see Barker (2012) and Lowry et al. (2019) for a debate within the monkeyflower research community over which genus name to use) provides an opportunity to test the relative importance of coding and noncoding changes at different evolutionary time scales. *Mimulus* lends itself to studies in plant evo-devo, thanks to a diversity of species with short generation times, high fecundity, amenability to greenhouse cultivation, and a large range of environmental adaptation (Wu et al., 2008; Yuan 2018). Sequenced genomes have been published for several species including *M. guttatus* (Puzey et al., 2017) and *M. l. luteus* (Edger et al., 2017).

The *luteus* group of *Mimulus* has an ancestral phenotype of yellow, carotenoid-pigmented flowers with red, anthocyanin-pigmented spots in the nectar guide region (Fig. 1). The overall color of the anthocyanin-pigmented petal tissue ranges from orange in *M. cupreus*, to red in *M. l. luteus*, to magenta in *M. l. variegatus* and *M. naiandinus*, depending on the relative intensity of carotenoid and anthocyanin pigmentation (Fig. 1).

In three members of the *luteus* group—the magenta-flowered *M. l. variegatus* and *M. naiandinus,* and the orange-flowered *M. cupreus*—anthocyanin pigment has expanded into the petal lobes. The gain of PLA is a derived, single-locus Mendelian trait in all three taxa. In *M. cupreus* and *M. naiandinus*, this is controlled by the *pla1* locus, which contains candidate anthocyanin-activating transcription factor genes *MYB2* and *MYB3a* (Cooley et al., 2011). A rare yellow-flowered morph of *M. cupreus* does not bear petal lobe anthocyanins (PLA). Since it is found in a single population in Chile, intermingled with orange morphs of *M. cupreus* (Cooley et al., 2008), it likely represents

105 a secondary loss of the PLA trait, and also segregates as a single-locus trait mapping to *pla1*. In
106 *M. l. variegatus*, in contrast, the gain of PLA is conferred by an unlinked second locus named
107 *pla2*. The *pla2* locus contains candidate gene *MYB5a/NEGAN*. All of the candidate genes at *pla1*
108 and *pla2* belong to subgroup 6 of the R2R3 MYB gene family (Cooley et al., 2011), which regulate
109 the "late" anthocyanin biosynthetic genes via participation in the MBW complex (reviewed in Lloyd
110 et al. (2017); see Stracke et al. (2001) for a description of the subgroup 6 genes).
111



112
113 **Fig. 1.** Petal lobe anthocyanin has been gained repeatedly in the *luteus* group of *Mimulus*. Closely related
114 species outside the *luteus* group, like *M. guttatus*, are typically yellow-flowered, with red anthocyanin
115 pigmentation restricted to the nectar guide region of the corolla. *Mimulus cupreus* and *M. naiandinus* have
116 each gained petal lobe anthocyanin via a single-locus change at genomic region *pla1*, while the magenta-
117 petaled *M. luteus* var. *variegatus* gained petal lobe anthocyanin via a change at *pla2*. A rare yellow-flowered
118 morph of *M. cupreus*, found in a single population in Chile, has lost petal lobe anthocyanin via a change at
119 *pla1*. Figure modified from Zheng et al. (2021).
120
121 While transitions in floral pigment traits have been extensively studied in plants, pigment losses
122 have been investigated more often than pigment gains (Rausher, 2008). Studying the repeated

123 gain of PLA in the *luteus* group presents an opportunity to rectify this. PLA transitions in the group
124 also enable a test of the Stern and Orgogozo (2008) prediction that protein-coding changes are
125 relatively more important within populations, while noncoding changes become increasingly
126 abundant as evolutionary divergence increases. Consistent with this prediction, the rare yellow
127 morph of *M. cupreus* - a within-population polymorphism linked to *pla1* (Cooley and Willis 2009;
128 Cooley et al. 2011) - appears to be associated with a deletion or other major mutation in exon 3
129 or 4 of *MYB2* (Supplemental Figures S1-S4). Based on the Stern and Orgogozo (2008) model,
130 we expect that the fixed gains of PLA in *M. l. variegatus*, orange *M. cupreus*, and *M. naiandinus*,
131 are more likely to be caused by *cis*-regulatory evolution.
132
133 Of the three taxa that recently gained PLA, *M. l. variegatus* is the most thoroughly characterized.
134 In *M. l. variegatus,* a combination of genetic mapping, *MYB5a* RNAi and overexpression, and
135 transcriptomic studies of wild-type versus *MYB5a* RNAi lines, shows that *MYB5a* is both
136 necessary and sufficient for the gain of PLA (Cooley et al. 2011; Zheng et al., 2021), and points
137 to one particular splice variant of the gene as participating in anthocyanin activation.
138
139 Unlike most *MYB* genes, which have a conserved three-exon structure (Dubos et al., 2010),
140 *MYB5a* in *M. l. luteus* and *M. l. variegatus* has four exons. The exon 1-2-4 splice variant, but not
141 the 1-2-3 variant, was found to be abundant specifically in anthocyanin-pigmented petal tissue:
142 the petal lobes of *M. l. variegatus* and the nectar guide regions of both taxa (Zheng et al. 2021).
143 Knockdown of the 1-2-4 splice variant resulted in the loss of petal pigmentation from *M. l.*
144 *variegatus* (Zheng et al., 2021). Finally, the exon 1-2-4 splice variant of *M. l. variegatus* - but not
145 the 1-2-3 splice variant - contains the subgroup 6 motif found in all known R2R3 MYB anthocyanin
146 activators (Stracke et al., 2001). Thus, the exon 1-2-4 splice variant of *MYB5a* appears to be
147 responsible for the gain of petal lobe anthocyanins in *M. l. variegatus*. While its pattern of spatially
148 specific expression suggests a mechanism of *cis*-regulatory evolution, it is unknown whether

149 divergence in the protein-coding sequence of *MYB5a* additionally contributed to the evolution of
150 increased anthocyanin production in *M. l. variegatus*.
151
152 In previous studies (Zheng et al., 2021), we relied on the stable transformation procedure
153 published by Yuan et. al. (2013) to test hypotheses about *MYB* gene function. While the method
154 is capable of producing stably transformed offspring, it has the disadvantage of having a low
155 transformation efficiency (about one seed per thousand in *M. l. variegatus*) and a large time cost
156 of about five months between infiltration of plants and flowering of transformant offspring.
157 Transient transformation is an attractive alternative for more rapid tests of gene function,
158 particularly for genes - such as pigment activators - that are expected to produce an easily visible
159 phenotype. In transient expression, the transgene is delivered to plant cells and transcribed by
160 the plant's transcriptional machinery without necessarily being incorporated into the plant's
161 genome, and transient transformation is regularly used in the tobacco genus *Nicotiana* (Kapila et
162 al., 1997; Schöb et al., 1997; Yang et al. 2000; Sparkes 2006). *Nicotiana* is relatively closely
163 related to *Mimulus*, as the two genera belong to the sister orders of Solanales and Lamiales,
164 respectively, and *Nicotiana* is a highly tractable system for transgenic experimentation. It is
165 routinely used for heterologous genetic experiments *in planta*, including tests of flower color genes
166 from both rosids and asterids (Montefiori et al., 2015; Tian et al., 2017). Ding and Yuan (2016)
167 adapted methods from *Nicotiana* for use in *Mimulus lewisii*. In our hands, however, transient
168 transformation caused substantial leaf tissue death in both *M. lewisii* and *M. l. luteus*. We therefore
169 returned to *Nicotiana*, using *N. tabacum* as the host for transient tests of *MYB5a* gene function.
170
171 If the gain of petal lobe anthocyanin (PLA) in the magenta-flowered *M. l. variegatus* is caused
172 solely by a *cis*-regulatory-driven spatial expansion of *MYB5a* function, then we predict that the
173 coding sequences of *MYB5a* from *M. l. variegatus* and the yellow-flowered *M. l. luteus* will be
174 equally capable of stimulating anthocyanin production. We tested this hypothesis by transiently

175 expressing each taxon's *MYB5a* exon 1-2-4 sequence in leaves of *N. tabacum*. We used a custom
176 image analysis pipeline to rapidly generate quantitative estimates of pigment production in the
177 transformed leaves.
178
179 Somewhat surprisingly, the *luteus* allele of *MYB5a* drove significantly stronger anthocyanin
180 pigmentation than did the *variegatus* allele. This indicates that *M. l. variegatus* did not evolve its
181 greater anthocyanin pigmentation via coding-sequence changes at *MYB5a*, and points to *cis*-
182 regulatory evolution at the causal *MYB5a* gene as a more likely driver of the gain of petal lobe
183 anthocyanin pigmentation in *M. l. variegatus*.
184
185 **Methods**
186
187 *Plant materials and growth conditions*
188
189 *Nicotiana tabacum* cv. Petit Havana SR1 seeds were obtained from Lehle Seeds (Round Rock,
190 TX, USA). *Mimulus lewisii*, line LF10HT1 (two generations inbred) seeds were a gift from the
191 Yao-Wu Yuan lab at the University of Connecticut (Storrs, CT, USA). *Mimulus luteus* var. *luteus*
192 and *M. l. variegatus* were originally collected in Chile from the El Yeso and Río Cipreses
193 populations, respectively (Cooley et al. 2008), and self-fertilized repeatedly with single-seed
194 descent to generate highly inbred lines. In this work, we utilized the 12-generations inbred line
195 *M. l. luteus* EY7 and the 11-generations inbred line *M. l. variegatus* RC6.
196
197 Seeds were surface-planted on wet soil and grown at Whitman College (Walla Walla, WA, USA)
198 in a greenhouse with 16-hour day lengths and temperatures ranging from 15ºC to 30ºC. Plants
199 were misted daily and fertilized three times per week with Open Sesame flowering fertilizer (Fox
200 Farm, Samoa, CA, USA).

201

202 *Determining the genomically encoded sequence of* MYB5a *from* M. l. variegatus

203

204 While investigating the *MYB5a* protein-coding regions of *M. l. luteus* and *M. l. variegatus*, using
205 the cloning and sequencing methods described in Zheng et al. (2021), we discovered an
206 unexpected new sequence in the fourth exon of *M. l. variegatus Myb5a*. In two different
207 sequencing reactions, originating from two distinct cDNA syntheses from a single mRNA
208 extraction from young bud *M. l. variegatus* tissue, a "GG" variant was found in which adenines
209 at positions 582 and 684 of the exon 1-2-4 splice variant were replaced with guanines (Figure
210 2). A third sequencing reaction from the same mRNA extraction produced the expected "AA"
211 sequence that was previously reported in Zheng et al. (2021). Because the same two variants
212 were observed, twice each, in the same two samples, we discounted the likelihood of four
213 random sequencing errors. Instead, we hypothesized that the "GG" variant might be encoded in
214 the genome of *M. l. variegatus*, perhaps representing an alternate allele of *MYB5a* or a closely
215 related gene duplicate, or that it might be the result of post-transcriptional editing.

216

217 To determine whether the "GG" variant represented the genomic sequence of an unknown,
218 related *MYB* gene, we cloned and sequenced a fragment of genomic *MYB5a* from an F1 hybrid
219 of *M. l. luteus* x *M. l. variegatus* using primers Myb5_64F and Myb5_57R (Supplemental Table
220 S1), which encompass the first of the two sites in question; these primers were selected
221 because they reliably amplified *MYB5a* from both *M. l. luteus* and *M. l. variegatus*. We reasoned
222 that, if the "GG" variant were from a paralogous *MYB* gene, then we should be able to recover
223 both variants from the F1 hybrid. The "AA" variant would originate from the *M. l. variegatus*
224 *MYB5a* and the "GG" variant would originate from the other, unknown gene; we would also
225 expect to recover the *M. l. luteus* allele of *MYB5a*. If the "GG" variant instead represented
226 residual heterozygosity in *M. l. variegatus*, or post-transcriptional editing of the mRNA, then any

227 single F1 hybrid would contain only one of the two *M. l. variegatus* variants, along with the *M. l.*
228 *luteus* allele of *MYB5a*.
229
230 To determine whether the “GG” variant represented genomically encoded, residual
231 heterozygosity in our highly inbred *M. l. variegatus*, we similarly cloned and sequenced genomic
232 *MYB5a* from the *M. l. variegatus* inbred line. If the “GG” variant were the result of post-
233 transcriptional modification, it should be absent from all genomic DNA samples (both the F1
234 hybrid and the *M. l. variegatus*).
235
236 Finally, to determine whether the “GG” variant could be repeatably isolated from cDNA, we
237 performed new mRNA extractions, and cloned and sequenced *MYB5a* cDNA, from two types of
238 *M. l. variegatus* and *M. l. luteus* floral tissues: the nectar guide region, which is anthocyanin-
239 pigmented in both taxa, and the petal lobe region, which is anthocyanin-pigmented only in *M. l.*
240 *variegatus.* Because exons 3 and 4 are partial duplicates of each other, primers Myb5_64F and
241 Myb5_57R (Supplemental Table S1) amplified both splice variants (exon 1-2-3 and 1-2-4) from
242 *M. l. luteus*, though only the exon 1-2-4 variant from *M. l. variegatus*. See Supplemental Figure
243 S6 for an illustration of primer binding sites for both taxa.

Figure 2. Putative A-to-I editing sites in the exon 1-2-4 splice variant of *M. I. variegatus MYB5a*. A. Dark gray bars show the exon structure from start codon to stop codon. Light gray bars show the DNA-binding R2 and R3 domains common to all members of the R2R3 MYB gene family (Stracke et al. 2001). "Subgroup 6" is a sequence motif that is conserved across all R2R3 MYB genes that encode activators of anthocyanin biosynthesis (Stracke et al. 2001). The two putative A-to-I editing sites are each marked as "edited site". B. Chromatograms from *MYB5a* Variant AA and Variant GG. The two polymorphic sites are both located in the fourth exon of *MYB5*, 584 and 686 nucleotides downstream of the translation start site. Nucleotide and amino acid differences are highlighted. Sequences were obtained using Sanger sequencing and were visualized using Geneious R10.

Nucleic acid extraction, PCR, and cloning

257 Genomic DNA was extracted from young leaves and floral buds using the Zyppy DNeasy
258 Extraction kit (Zymo Research, CA, USA) according to the manufacturers' protocol. RNA was
259 extracted from buds using the E.Z.N.A. Plant RNA Kit (Omega Bio-Tek, GA, USA) with the
260 DNAse I digestion protocol added to it. cDNA was synthesized using the qScriptTM cDNA
261 Synthesis Kit (Quanta BioSciences, Inc., MD, USA). Quality and concentration of DNA and RNA
262 were quantified using a nanodrop.

263 Fragments of *MYB5a* spanning one or both adenine/guanine polymorphic sites were PCR
264 amplified using the primers listed in Supplemental Table S1 with G-Biosciences Taq polymerase
265 (St. Louis, MO, USA). Reactions were run with 10μM forward and reverse primers, G-
266 Biosciences 10x buffer, and 2.5μM dNTPs. Annealing temperatures were set to 3℃ below the
267 primer's lowest melting temperature and the number of PCR cycles ranged from 30-32.

268 PCR products were purified and cloned into pGEM vectors in *E. coli* as described in Zheng et al.
269 (2021). Colonies were PCR-screened for inserts of the correct size using primers M13F(-20)
270 and M13R(-24). Sanger sequencing was performed by Eton Biosciences (San Diego, CA, USA)
271 and sequences were visualized using Geneious R9 and R10.
272
273 *Strategy for testing for functional equivalence of two coding sequences*
274
275 Once the "AA" allele had been identified as the only genomically encoded *MYB5a* sequence
276 present in the magenta-flowered *M. l. variegatus*, transgenes were constructed to test whether
277 the exon 1-2-4 splice variant was functionally equivalent to the corresponding allele from the
278 yellow-flowered *M. l. luteus* (which lacks petal lobe anthocyanins), as described below. Each
279 transgene, as well as a negative control, was transfected into leaves of *Nicotiana tabacum*, and
280 the area of the spot of anthocyanin pigment produced following each infiltration was quantified.

281 The heterologous *N. tabacum* system was selected because our pilot studies in *M. l. luteus* and
282 another monkeyflower species, *M. lewisii*, failed to produce visible anthocyanin pigment.
283
284 *Bacterial culturing for transgene construction*
285
286 *Escherichia coli* cultures were grown at 37ºC in Luria-Bertani (LB) broth: 10g/L tryptone, 5g/L
287 yeast extract, 10g/L NaCl in demineralized water, sterilized by autoclaving. *Agrobacterium*
288 *tumefaciens* cultures were grown at 28 ºC in LB broth with the NaCl concentration reduced from
289 10g/L to 5g/L.
290
291 Cells containing plasmids with a Kanamycin-resistance gene were grown in media containing 50
292 µg/mL kanamycin. The *A. tumefaciens* strain, GV3101, used in these studies contains gentamicin-
293 and rifampicin-resistance genes; these cultures were grown in media additionally containing 50
294 µg/mL gentamicin and 25 µg/mL rifampicin. Liquid cultures were grown at the appropriate
295 temperature with shaking at 200 rpm. To isolate individual colonies, cells were grown on plates
296 containing LB media with appropriate selective antibiotics and 15 g/L agar.
297
298 *Construction of Gateway® Entry Vectors*
299
300 The exon 1-2-4 splice variant of *MYB5a* was amplified from *M. l. variegatus* petal cDNA, using
301 primers cacc10F and Myb5_69R (Supplemental Table S1) and New England Biolabs® Phusion®
302 High-Fidelity DNA Polymerase. Amplicons were transformed into the pEARLEYGATE101
303 Gateway vector (Earley et al. 2006). From there, the coding sequence without a stop codon was
304 amplified from plasmid DNA containing *M. l. variegatus MYB5a* CDS using the same primers and
305 polymerase.
306

307 The pENTR-D/TOPO Cloning Kit was used to produce directionally-cloned Gateway® entry
308 clones carrying *M. l. variegatus MYB5a* CDS. The reaction mixture was transformed into TOP10™
309 *E. coli* cells, and colonies were screened for the presence of the insert via PCR, using the
310 Myb5_10F, M13R(-24) primer pair. Three colonies that gave a band at ~1-kb were selected for
311 sequencing. The M13F(-20), M13R(-24) primer pair was used to amplify and sequence the insert.
312 Sanger sequencing was performed by Eton BioScience® (San Diego, CA, USA) and checked for
313 errors against reference sequence in Geneious® version 9.1.8 (https://www.geneious.com).
314
315 For *M. l. luteus*, amplification of the exon 1-2-4 splice variant was attempted without success using
316 cDNA extracted from young bud tissue of *M. l. luteus* lines EY1 and EY7. *M. l. luteus* EY7 is
317 known to express *MYB5a* in the anthocyanin-spotted nectar guide region of the flower bud, but
318 expression levels are low (Zheng et al. 2021). Instead, this protein-coding region was synthesized
319 by GENEWIZ (Plainfield, NJ, USA) based on the published genomic sequence of *M. l. luteus*
320 (Edger et al., 2017). The sequence was delivered in a pUC57 vector, but with attL1 and attL2
321 homology sites added to the 5'- and 3'- ends of the gene, respectively, to facilitate Gateway
322 recombination. Upon receipt, the plasmid was transformed into TOP10™ chemically competent
323 *E. coli* cells, and colonies were screened for the presence of the insert via PCR, using Myb5_12F
324 internal forward primer and M13R(-24) reverse primer (primer table). Sanger sequencing was
325 performed by Eton BioScience, and sequence was checked for errors against the reference
326 sequence in Geneious.
327
328 *Construction of Gateway® Plant Expression Vectors*
329
330 The LR Clonase II kit (ThermoFisher Scientific, Waltham, MA, USA) was used to transfer each
331 insert from an entry vector to destination vector pEARLEYGATE101 (Earley et al. 2006). Entry

332 vectors used in this study were pENTR with *M.l.variegatus MYB5a* CDS and pUC57 with
333 *M.l.luteus MYB5a* CDS.
334
335 Reactions were transformed into TOP10™ chemically competent *E. coli* and screened for the
336 presence of the insert via PCR and sequencing, using an insert-specific forward primer
337 (Myb5_10F for *M. l. variegatus* and Myb5_12F for *M. l. luteus*) and a reverse primer, att-R2, that
338 binds to the recombination site at the 3'-end of the insert in recombined pEARLEYGATE vectors.
339
340 To exclude colonies with unrecombined entry vector, restriction endonuclease digests were
341 performed as an additional diagnostic on plasmid purified from those colonies. Because it cuts
342 both within the destination vector and within the insert, the HinDIII enzyme was used (Promega
343 Corp., Madison, WI). Reactions were incubated for 60 minutes at 37ºC, then heat-inactivated for
344 15 minutes at 65ºC. Colonies that gave the expected digest pattern for recombined destination
345 construct as well as the correct insert DNA sequence were chosen to proceed with this project.
346
347 *Transformation into* Agrobacterium
348
349 GV3101 electrocompetent *A. tumefaciens* cells, mixed with 1 µL isolated plasmid DNA (25-350
350 ng/µL) from each construct, were briefly exposed to a 2.5 kV, 200 ohm, 25 µF pulse using a
351 BioRad® MicroPulser Electroporator (BioRad Laboratories, Hercules, CA, USA). The mixture was
352 then immediately combined with 1 mL room-temperature LB media without antibiotics, incubated
353 at 28ºC for 2-3 hours with shaking at 200 rpm., then plated on selective media. Putative
354 transformants were tested for transgene insertion using a PCR screen with primers pEG-35S-
355 attB1_F and att-R2 (Supplemental Table S1).
356

357 *pGFP plasmid for negative control*

358

359 To screen for non-specific effects of transgene infiltration, the *A. tumefaciens*-compatible GFP
360 expression plasmid pGFPGUSPlus was used as a negative control. pGFPGUSPlus was a gift
361 from Claudia Vickers (Addgene plasmid # 64401; http://n2t.net/addgene:64401;
362 RRID:Addgene_64401) (Vickers et al., 2007). Using the protocol in the previous section, the
363 plasmid was transformed into *A. tumefaciens*.

364

365 We initially planned to use GFP fluorescence as a read-out of transgene activity, if anthocyanin
366 pigment failed to be produced by either the *luteus* or *variegatus* coding sequence. However, we
367 found that the focal trait of anthocyanin pigmentation was indeed reliably produced. GFP
368 fluorescence, beyond that produced by autofluorescence, was not evident in a pilot study of
369 infiltrated *N. tabacum* leaves (Supplemental Figure S5), and thus was not examined further.

370

371 *Transient transformation*

372

373 A video documenting our transformation methodology, adapted from Ding and Yuan (2016), is
374 available upon request.

375

376 *Agrobacterium* colonies were PCR-screened to verify that they contained the desired transgene
377 (pGFPGUSPlus, or *MYB5a* CDS from *M. l. luteus* or *M. l. variegatus*), and were then grown 16-
378 24 hours in 5 mL LB (*Agrobacterium* recipe) at 28°C with appropriate antibiotics. The small culture
379 was brought up to 50 mL with LB plus antibiotics and grown 12-16 hours at 28°C.

380

381 The 50-mL cultures were centrifuged in a Beckman-Coulter Allegra 25R temperature-controlled
382 benchtop centrifuge (Beckman Coulter Inc, Brea, CA, USA) at 4ºC and 6,000 RCF for 15 minutes.
383 The pellet was resuspended in a volume of 5% sucrose (w/v) solution equivalent to between one
384 half and one time the original culture's volume. Resuspensions were adjusted, through dilution
385 with the sucrose solution, to have the same optical density at 600 nm (OD_{600}) across all three
386 transgene types within each trial, with an OD_{600} range of 1.6-1.9 across trials. Departing from Ding
387 and Yuan (2016), acetosyringone and Silwet L-77 were not included in the resuspension solution;
388 omitting these reagents was proposed as a possible solution for leaf tissue damage previously
389 observed in infiltrated *Mimulus lewisii* leaves (B. Ding, personal communication).
390
391 A B-D 1-mL slip tip disposable SubQ syringe (Becton, Dickinson and Company, Franklin Lakes,
392 NJ, USA) with the needle removed was used to deliver *A. tumefaciens* cells to the leaves of young
393 (1-3 months) *N. tabacum*. Using a gloved finger, the top of the leaf was held firmly while the
394 underside of the leaf was injected with the syringe until the liquid had visibly spread past the site
395 of injection. A volume of 100-200µL resuspended cells was injected per spot. The *M. l. luteus*
396 *MYB5a* transgene and the *M. l. variegatus MYB5a* transgene were infiltrated in pairs, alternating
397 with each leaf which transgene was injected into the left versus the right side of the leaf. A smaller
398 number of negative controls was performed, with pGFPGUSPlus infiltrations approximately
399 evenly distributed between the two sides of other leaves on the same plants.
400
401 *Image acquisition and preparation*
402
403 Leaves were imaged starting three days after infiltration, since Li et. al. (2009) reported that
404 maximum expression of *Agrobacterium*-delivered transgenes occurs 3 days after infiltration.
405 Because accumulation of visible gene product (anthocyanin) continued for several days after
406 infiltration in some samples, leaves were imaged until a maximum of twelve days after infiltration.

407 Digital photographs of infected leaves were then taken in a dark room with a Nikon D3500 DSLR
408 camera with 18-55mm lens. The camera was fixed on a stand and the leaf was illuminated by a
409 Sylvania Ceramic Metal Halide bulb, which exceeds 15,000 lumens, as the light source. All
410 exposures produced both RAW and jpeg images, with RAW images used for the analysis and
411 jpeg images used for interoperability with image annotation software. VGG Image Annotator was
412 used to demarcate the regions of interest covering the extent of the infiltrated leaf tissue and the
413 center of each injection site.
414
415 *Analysis of images and data*
416
417 S_{green}, which is the strength of the green channel relative to the total of all three color channels,
418 was previously found to correlate highly (R^2 >= 0.63, p<.001) with anthocyanin concentration
419 across a range of taxa and plant tissues (del Valle et. al. 2018). This is because anthocyanin
420 absorbs light in the green region of the visible spectrum. The index, S_{green}, is given by:
421
422 $S_{green} = N_{green} / (N_{green} + N_{blue} + N_{red})$
423
424 where N is the intensity value for the green, blue, or red color channel (del Valle et al. 2018).
425
426 A custom Python program imported the RAW image files for processing. The Python program
427 used a modified version of the MacDuff color chart detection algorithm
428 (https://github.com/mathandy/python-macduff-colorchecker-detector) to automatically detect the
429 panels of known broad-spectrum reflectance values on a reference color chart. Image pixel values
430 were converted into normalized reflectance values based on a linear fit of the red, green, and blue
431 signal strengths in those panels of known reflectance. The program then averages the relative
432 greenness (S_{green}) value over all pixels of the annotated region of interest minus a circular region

433 20 pixels in radius at the injection site, which typically exhibited tissue damage from the injection
434 syringe. This yielded a single S_{green} value for each sample.
435
436 To simultaneously evaluate S_{green} values across all three treatments (the negative controls, the
437 *luteus* transgenes, and the *variegatus* transgenes), a one-way ANOVA was performed in R 4.1.2,
438 followed by a Tukey's post hoc test. To further compare effects of only the *luteus* and *variegatus*
439 transgenes to each other, taking advantage of the paired design of infiltrations for these two
440 treatments, a two-tailed paired t-test was used.
441
442 A number of leaves infiltrated with *MYB5a* did not produce visible quantities of red pigmentation.
443 A *luteus-variegatus* treatment pair was categorized as "no visible pigment" if the S_{green} value for
444 either member of the pair showed equal or lower amounts of red pigmentation than the average
445 of all the negative controls. This categorization corresponded well with a by-eye assessment.
446 Data were analyzed both with and without these apparently-unsuccessful infiltration pairs. Overall
447 rates of success (defined as production of visible pigment) were compared between the *luteus*
448 and *variegatus* transgenes using a χ^2 contingency test in R 4.1.2.
449
450 *Data availability*
451
452 A detailed transient-transformation methods video is available upon request. Images of all
453 leaves analyzed are linked as Supplemental Figure S9.A, S9.B, and S9.C in the Supplemental
454 Data. Full S_{green} (pigment intensity) data are available as Supplemental Table S2. *MYB5a*
455 sequences from the exon 1-2-4 splice variants of *M. l. luteus* and *M. l. variegatus* have been
456 previously published: https://www.ncbi.nlm.nih.gov/nuccore/MT361119.1 (*M. l. luteus*, also listed
457 as *E. l. lutea*) and https://www.ncbi.nlm.nih.gov/nuccore/2019733960 *(M. l. variegatus*, also

listed as *E. l. variegata*). Code for image analysis is available at https://github.com/WhitmanOptiLab/PigmentSpotting.

Results

Two sites in MYB5a *are occasionally sequenced as guanine rather than adenine*

The locations of two putatively edited sites, in exon 4 of *M. l. variegatus MYB5a*, are shown in Figure 2, along with chromatograms from the “AA” variant versus the “GG” variant from our initial, fortuitous discovery of the sequence difference. The novel “GG” variant was found in two different sequencing reactions, originating from two cDNA synthesis reactions from a single mRNA extraction of *M. l. variegatus* young bud outer-petal tissue.

The “GG” allele of MYB5a *from* M. l. variegatus *is not genomically encoded*

Cloning *MYB5a* gDNA from a *variegatus x luteus* F1 hybrid yielded 29 colonies containing a *MYB* sequence. Of these, 13 contained the “AA” variant of the *M. l. variegatus* allele. The remaining 16 contained the *M. l. luteus* allele. The “GG” variant was not discovered in these genomic DNA samples.

Cloning *MYB5a* gDNA from a highly inbred line of *M. l. variegatus* yielded 37 colonies containing a *MYB* sequence. All of these were the “AA” variant of *M. l. variegatus MYB5a*. The “GG” variant was not discovered in these genomic DNA samples. Sample PCR colony screens, from both F1 hybrid gDNA and *M. l. variegatus* gDNA, are shown in Supplemental Figure S7.

483 After collecting new floral bud tissues, and cloning and sequencing *MYB5a* cDNA from them, we
484 identified one additional colony containing a G at the edited site encompassed by our primers
485 (Table 1). As before, the variant was obtained from the petal lobes ("outer petal") of *M. l.*
486 *variegatus*. The other 29 *M. l. variegatus* colonies contained the "AA" variant. In *M. l. variegatus,*
487 only the exon 1-2-4 splice variant was recovered, as expected based on the utilization of
488 primers Myb5_64F and Myb5_57R (Supplemental Figure S6). These same primers were,
489 however, competent to amplify both 1-2-3 and 1-2-4 splice variants from *M. l. luteus*, and they
490 did. We found 28 colonies containing the exon 1-2-4 splice variant of *M. l. luteus MYB5a*, and
491 13 containing the exon 1-2-3 splice variant, with both splice variants appearing in both inner and
492 outer petal tissue (Table 1).
493
494 Recovering *M. l. luteus MYB5a* sequence from outer petal tissue was unexpected, and may
495 reflect imprecise separation of the two tissue types during floral bud dissection. The finding is
496 consistent with RT-PCR of *MYB5a* from our four cDNA samples, which indicated the presence
497 of some *MYB5a* transcript in the *M. l. luteus* outer petal sample (Supplemental Figure S8).
498
499
500 **Table 1. Sequencing *MYB5a* cDNA from *M. l. luteus* and *M. l. variegatus* developing flower bud**
501 **tissue identifies splice variants in *M. l. luteus,* and sequence variants in *M. l. variegatus.*** "Inner
502 petal" corresponds to the nectar-guide-spotted throat region of the flower; "outer petal" corresponds to the
503 petal lobes.

MYB5a variant	*luteus* inner petal	*luteus* outer petal	*variegatus* inner petal	*variegatus* outer petal
Exons 1-2-4 with "A"	17	11	10	19

MYB5a variant	*luteus* inner petal	*luteus* outer petal	*variegatus* inner petal	*variegatus* outer petal
Exons 1-2-4 with “G”	-	-	-	1
Exons 1-2-3	10	3	-	-

504

505 *Infiltration of transgenes into* N. tabacum *leaves*

506 Of the 124 leaves that received paired infiltrations of the 1-2-4 splice variant of *MYB5a* (*M. l.*
507 *variegatus* on one side of the central vein, and *M. l. luteus* on the other), 26 were eliminated due
508 to tearing or inadvertent marking over the pigmented area. The remaining 98 were scored for
509 the presence of visible anthocyanin pigmentation, and also quantitatively analyzed for pigment
510 abundance. Of the 25 leaves infiltrated on each side of the midvein with pGFPGusPlus as a
511 negative control, five were eliminated due to tearing or marking errors and the remaining 20
512 leaves (40 infiltrations) were quantitatively analyzed for anthocyanin pigment abundance
513 (Supplemental Table S2, Supplemental Figure S9).

514

515 MYB5a *from both* M. l. luteus *and* M. l. variegatus *drives strong anthocyanin production*

516

517 Both *MYB5a* transgenes resulted in significantly redder leaf tissue, as indicated by lower S_{green}
518 values, than did the negative control (Figure 3 and 4; $F(2, 233) = 12.43$; $p<0.0001$**)**. Surprisingly,
519 the allele of *MYB5a* from the yellow-flowered *M. l. luteus* drove significantly greater anthocyanin
520 production than the corresponding allele from the magenta-flowered *M. l. variegatus* (Tukey’s
521 post hoc test: $p=0.0102$). After excluding the negative control to compare only the *luteus* and
522 *variegatus* transgenes via a paired t-test, the difference was even more strongly significant ($t = -$
523 6.3656, $df = 97$, $p = 6.461*10^{-9}$; mean difference = $6.456*10^{-3}$ S_{green} units).

524

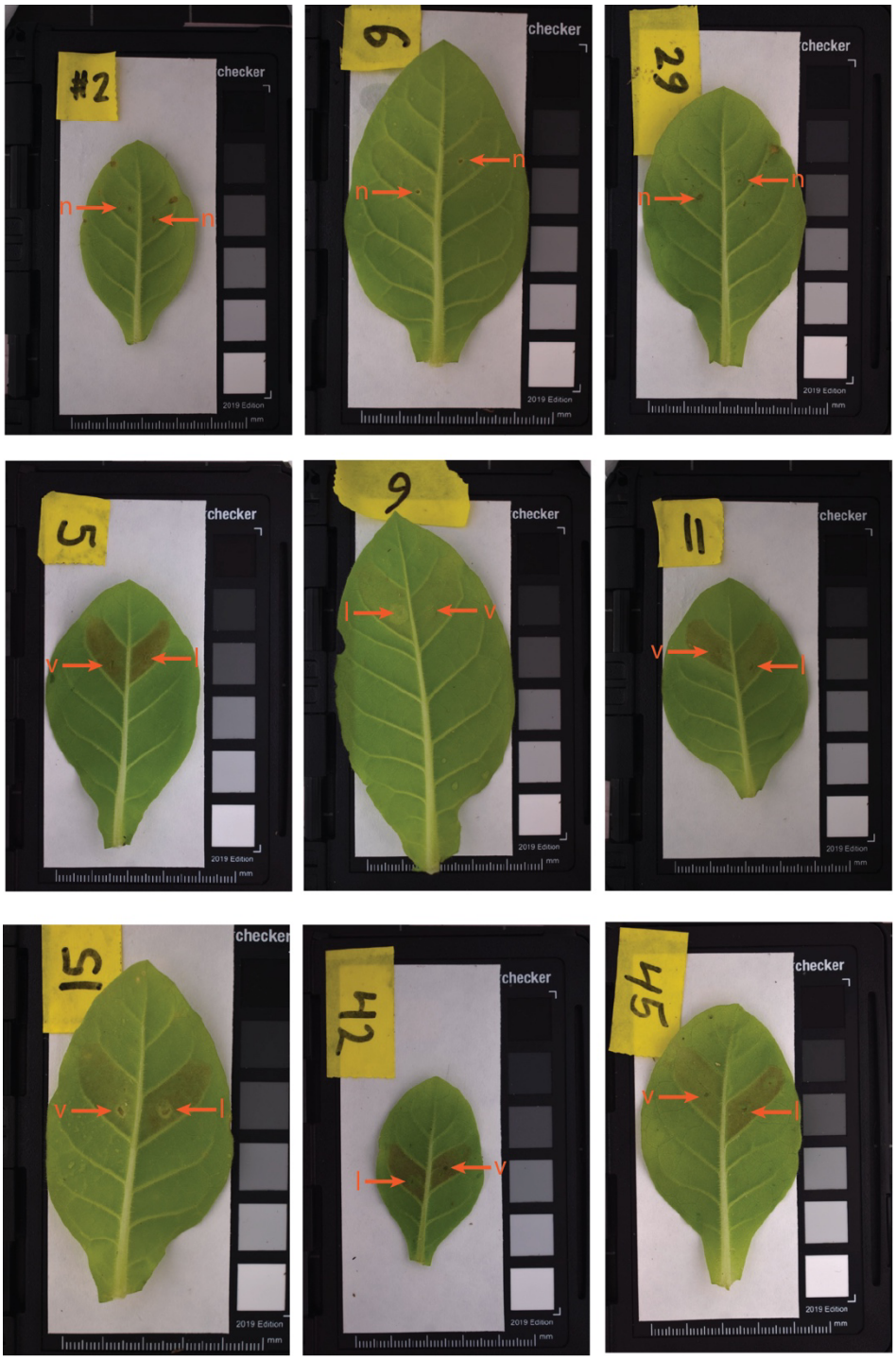
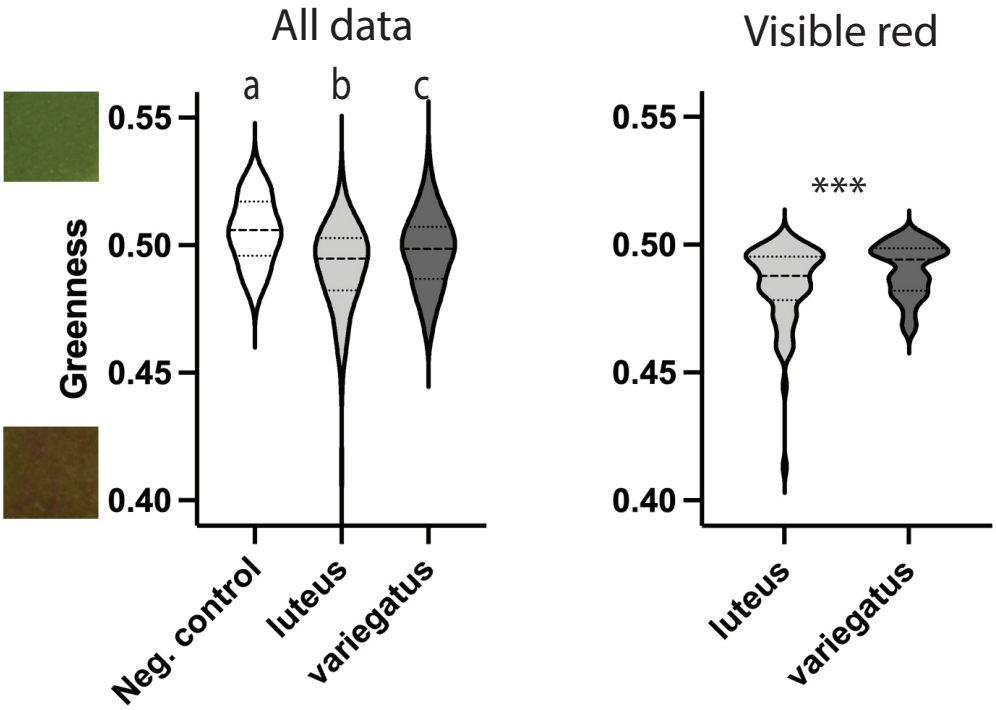


Figure 3. *MYB5a* transgenes yielded a range of anthocyanin biosynthesis levels in *N. tabacum* leaves. In each image, the injection site is indicated with an arrow, labeled as n (negative control), l (*luteus* allele of *MYB5a*), or v (*variegatus* allele of *MYB5a*). Leaf photos for this figure were uniformly

529 brightened by 30% in Powerpoint to make anthocyanin pigmentation and injection sites easier to see,
530 although analyses were performed on un-brightened images.
531



532

533 **Figure 4. *MYB5a* alleles from both the yellow-petaled *M. l. luteus* and the magenta-petaled *M. l.***
534 ***variegatus* drive production of red anthocyanin pigmentation in *N. tabacum* leaf tissue.** Left, data
535 from all samples; *N* = 40 (negative control), 98 (*luteus* coding sequence), 98 (*variegatus* coding
536 sequence). Right, data only from matched *luteus*-*variegatus* pairs for which both members of the pair
537 produced visible pigmentation; *N* = 66 pairs. Each graph is a violin plot with the median shown as a
538 dashed line and the first and third quartiles as dotted lines. On the Y-axis, low values correspond to
539 redder leaf tissue and high values correspond to greener tissue. Color swatches on the Y-axis were taken
540 from injection sites with S_{green} values of 0.53 and 0.41 respectively. Swatches were taken from leaf photos
541 that had been uniformly brightened by 30% in Powerpoint, to maintain consistency with the previous
542 figure, although analyses were performed on un-brightened images. Letters a, b, and c are significance
543 groupings at $p<0.05$ based on a one-way Anova with Tukey's post hoc test; *** indicates $p<0.001$ based
544 on a two-tailed paired t-test.
545

546 We wondered whether this result might have been influenced by different rates of infiltration
547 success, rather than by different quantities of pigment produced in successful infiltration events.
548 Overall, 79/98 *luteus* infiltrations, but only 67/98 *variegatus* infiltrations, resulted in the
549 accumulation of visible anthocyanin pigmentation in a tobacco leaf (Table 2). The difference in
550 success rate between the two alleles approached but did not reach statistical significance (χ^2 =
551 3.249, df = 1, p = 0.0715).
552
553 **Table 2. The *M. l. luteus* allele of *MYB5a* more often produced visible anthocyanin pigmentation in**
554 **tobacco leaves than did the *M. l. variegatus* allele, although the difference is not statistically**
555 **significant.** χ^2 = 3.249, df = 1, p = 0.0715.

Transgene	Leaves with visible anthocyanin	Leaves without visible anthocyanin
luteus MYB5a	79	19
variegatus MYB5a	67	31

556
557 To exclude possible effects of varying infiltration success, we analyzed pigment production only
558 from the pairs of infiltrations for which both members of the pair produced visible amounts of
559 leaf anthocyanin pigment (Figure 4 right panel). In these pairs, a slightly smaller but still highly
560 significant difference was found, again with the *luteus* coding sequences corresponding to lower
561 S_{green} values and thus more-intense red pigmentation (t = -4.4855, df = 65, p = $3.032*10^{-5}$; mean
562 difference = $6.170*10^{-3}$ S_{green} units).
563
564 **Discussion**
565

566 When closely related taxa show phenotypic divergence, is the molecular mechanism a
567 mutational difference in the protein-coding region of a gene, or in the noncoding, *cis*-regulatory
568 region? We used transient transgenic assays to investigate this question for the MYB5a
569 anthocyanin-activating transcription factor from two varieties of monkeyflower: the yellow-
570 flowered *Mimulus luteus* var. *luteus*, which lacks anthocyanin pigment in its petal lobes, and the
571 magenta-flowered *M. l. variegatus*, which recently evolved petal lobe anthocyanin via an
572 unknown change within the *MYB5a* gene. Using quantitative image-analysis based methods, we
573 report that there is a significant functional difference between the two protein-coding regions
574 when expressed in *N. tabacum* leaves, but in the opposite direction of what would be expected
575 if coding sequence evolution were responsible for the increased pigmentation observed in *M. l.*
576 *variegatus*. Together with a previous finding that *MYB5a* is more strongly expressed in *M. l.*
577 *variegatus* petal lobes than in its conspecific (Zheng et al. 2021), this result strongly indicates
578 that *cis*-regulatory evolution is responsible for the recent expansion of pigmentation in *M. l.*
579 *variegatus*.
580
581 *Improved tools for rapid transgenic assays in Mimulus*
582
583 The success of the two *Mimulus* transgenes at activating anthocyanin production in *N. tabacum*
584 is encouraging for future functional studies in *Mimulus*, though not unprecedented. One factor
585 that can limit the ability of a MYB to function in a heterologous system is the availability of a
586 functional bHLH co-factor. In some cases, co-expression of the focal MYB gene's native bHLH
587 partner has been necessary for successful anthocyanin activation (Espley et al., 2009; Lin-
588 Wang et al., 2010), but in other cases, anthocyanins have been induced in *N. tabacum* without
589 also expressing bHLH from the same system (Fraser et. al., 2013). In one such study,
590 Montefiori et. al. (2015) expressed *AcMyb110* from Kiwifruit (*Actinidia* sp., order Ericales) in *N.*
591 *tabacum* leaves and successfully stimulated anthocyanin production. They identified two

592 endogenous bHLH transcription factors in *N. tabacum*, NtJAF13 and NtAN1, that associated
593 with AcMYB110 to stimulate expression of anthocyanin biosynthetic pathway genes, and that
594 may also have interacted with the *Mimulus* MYB5a protein in our experiment. While stable
595 transgenics are more useful for many applications, the ease and rapidity of transient
596 transformation approaches makes it a valuable tool when studying a readily quantifiable
597 phenotype such as pigment production.
598
599 One caveat to the quantification of anthocyanin activation by the *luteus* and *variegatus* transgenes
600 is that high levels of transgene expression could potentially alter patterns of functionality that only
601 appear at lower concentrations (Koes et. al., 2005). Developing a more closely related species
602 as a platform for functional tests would also be beneficial. In our hands, preliminary tests with both
603 *M. luteus* and congener *M. lewisii* resulted in high levels of tissue death and damage, but the
604 latter appears to have promise as a host for transient transgenic assays (Ding and Yuan 2016).
605
606 Using *N. tabacum* for direct side-by-side comparison of *Agrobacterium*-delivered transgenes was
607 first reported by Van der Hoorn et. al. (2000). This strategy takes advantage of leaf symmetry and
608 the clearly delineated leaf sectors in *N. tabacum* to compare two genes side-by-side in an identical
609 biological background. Coupled with a nondestructive way to quantify the resulting phenotype, we
610 believe this remains an underutilized strategy for functional comparisons between genes. Finally,
611 we note that the significant difference in pigment production between the *luteus* and *variegatus*
612 transgenes would almost certainly not have been detected using a qualitative assessment of
613 pigmentation, highlighting the benefits of even relatively simple computer-vision based programs
614 to quantify color traits in plants.
615
616 *Possible A-to-I editing of mRNA*
617

618 In the process of building *MYB5a* overexpression transgenes, we discovered what appears to
619 be the first documented case, to our knowledge, of post-transcriptional editing in an
620 anthocyanin-related gene. Two sites within the *M. l. variegatus* allele are encoded as adenine in
621 the genome, yet occasionally produce mRNA sequences that read as a guanine in Sanger
622 sequencing.
623
624 Inosine is a guanine analog, most often created in cells by the deamination of adenine
625 (Srinivasan et al., 2021), that is reported as guanine in Sanger sequencing (Cattenoz et al.,
626 2013). Adenine-to-inosine (A-to-I) editing was first discovered in *Xenopus laevis* mRNA by Bass
627 and Weintraub (1988), and is abundant in metazoans (Cattenoz et al., 2013), with one study
628 predicting over 36,000 A-to-I editing sites in the human genome (Li et al., 2009).
629
630 Although A-to-I editing of mRNA transcripts does not yet appear to have been directly
631 investigated in plants, we hypothesize that the two new bases are in fact inosine, given that
632 inosine is reported as guanine by Sanger sequencing (Cattenoz et al., 2013). Nuclear A-to-I
633 post-transcriptional editing has been reported in plant tRNA (Delannoy et al., 2009; Karcher and
634 Bock 2009; Zhou et al., 2014), and the deaminase enzymes required for A-to-I editing have
635 been putatively discovered in *Arabidopsis thaliana* (Zhou et al., 2014). Though mRNA A-to-I
636 editing has not been described in plants, it is widespread across the domains of life, including
637 fungi (reviewed in Teichert 2018), animals (reviewed in Knoop 2011), and bacteria (first reported
638 by Bar-Yaacov et al., 2017).
639
640 In contrast, reports of A-to-G editing in plants appear to be based solely on sequencing-based
641 approaches that would report inosines incorrectly as guanines (Pan et al., 2022), with A-to-I
642 editing apparently first proposed by Meng et al. (2010) on the basis of A-to-"G" mRNA editing
643 discovered by sequencing plant transcriptomes. True A-to-G editing does not appear to be a

644 verified biological phenomenon in any taxon. We therefore consider A-to-G editing to be less
645 likely than A-to-I editing in our study.
646
647 A variety of methods exist for confirming A-to-I editing, including chromatographic approaches
648 (Wolf et al., 2002; Chan et al., 2010) and “inosine chemical erasing” (ICE)-Seq (Sakurai et al.,
649 2010). However, A-to-I editing is commonly detected and quantified by the simple method used
650 here, in which reverse transcription and sequencing of mRNA reveals unexpected “guanines” in
651 some proportion of transcripts (e.g. Gu et al., 2012).
652
653 When inosine is present in tRNA, it can pair promiscuously with A, C, or U. In mRNA transcripts,
654 in contrast, it is translated as though it were guanine (Srinivasan et al. 2021). Regardless of
655 whether the edited bases result in guanine or inosine, then, they are likely to be interpreted by
656 the translational machinery of the cell as guanine. In the *M. l. variegatus* allele of *MYB5a*, both
657 edited sites would result in an amino acid change: from asparagine to serine at nucleotide
658 position 584, and from glutamic acid to glycine at position 686.
659
660 How conservative are these changes? One metric is Grantham’s Distance, based on
661 composition, polarity and molecular volume (Grantham 1974). Using this metric, amino acid
662 pairs have similarity scores ranging from 5 for the highly similar leucine-isoleucine pair to a
663 maximum of 215 for cysteine-tryptophan. The first putative editing site reported here
664 (asparagine-serine, both of which have polar side chains) has a modest Grantham’s Distance of
665 46. The putative glutamic acid to glycine substitution - replacing a negatively charged side chain
666 with a single hydrogen - has a larger Grantham’s Distance of 98. The implications for MYB5a
667 protein folding and function are, however, unknown. Overall, the mechanisms for A-to-I (or A-to-
668 G) editing in plant mRNAs, and their functional impacts, comprise a barely-explored area within
669 plant molecular biology, which seems likely to yield new discoveries upon further investigation.

670 Nanopore native RNA sequencing methods, recently used by Nguyen et. al. (2022) to globally
671 identify inosine in human, mouse, and *Xenopus* transcriptomes, might be applied fruitfully to
672 plant transcriptomes with the same aim.
673
674 **Conclusions**
675
676 *Mimulus luteus* var. *luteus* and *M. l. variegatus* differ strikingly in floral phenotype, thanks to a
677 derived loss of yellow carotenoid pigment and gain of magenta cyanidin pigment in the latter.
678 The expansion of cyanidin to the petal lobes of *M. l. variegatus* has previously been tracked to
679 the *MYB5a* transcription factor gene, for which the patterns of petal expression correlate well
680 with the presence versus absence of cyanidin pigment. Here, we use transient transgenics in *N.*
681 *tabacum*, coupled with quantitative digital image analysis, to show that the protein-coding region
682 of *MYB5a* from the high-anthocyanin *M. l. variegatus* does not show any increased anthocyanin-
683 activation ability (in fact, the reverse) compared to the orthologous sequence from the low-
684 anthocyanin *M. l. luteus*. This finding adds further support to the hypothesis that evolution in *cis*
685 to *MYB5a* is the molecular mechanism for the gain of this novel anthocyanin trait in *M. l.*
686 *variegatus.*
687
688 We additionally report the discovery of what appears to be post-transcriptional mRNA editing.
689 The edits are reported as A-to-G by Sanger sequencing, but we argue that A-to-I editing is more
690 likely based on what is known about RNA editing in plants and other organisms. Overall, our
691 work highlights the utility of floral diversification for identifying the molecular mechanisms of
692 evolution, as well as the scope for continued new discoveries in the realm of plant molecular
693 genetics.
694
695 **Acknowledgements and Funding**

The authors thank N. Forsthoefel, B. Ding, and Y.-W. Yuan for assistance with and advice on the transgenic techniques. We thank D. Vernon, B. Moss, J. Puzey, and their students for discussion of some of the results presented here. AMC was supported by NSF-DEB-1655311, NSF-DEB-1754075, and NSF-IOS-2031272. AEP was additionally supported by a Whitman College Abshire Award for undergraduate research.

References

Barker, W. R., G. L. Nesom, P. M. Beardsley, and N. S. Fraga. 2012. A taxonomic conspectus of Phrymaceae: a narrowed circumscription for Mimulus, new and resurrected genera, and new names and combinations. Phytoneuron 39:1–60.

Bar-Yaacov, D., E. Mordret, R. Towers, T. Biniashvili, C. Soyris, S. Schwartz, O. Dahan, and Y. Pilpel. 2017. RNA editing in bacteria recodes multiple proteins and regulates an evolutionarily conserved toxin-antitoxin system. Genome Res. 27:1696–1703.

Bass, B. L., and H. Weintraub. 1988. An unwinding activity that covalently modifies its double-stranded RNA substrate. Cell 55:1089–1098.

Beardsley, P. M., and R. G. Olmstead. 2002. Redefining Phrymaceae: the placement of *Mimulus* , tribe Mimuleae, and *Phryma*. Am. J. Bot. 89:1093–1102.

Cattenoz, P. B., R. J. Taft, E. Westhof, and J. S. Mattick. 2013. Transcriptome-wide identification of A > I RNA editing sites by inosine specific cleavage. RNA 19:257–270.

Chan, C. T. Y., M. Dyavaiah, M. S. DeMott, K. Taghizadeh, P. C. Dedon, and T. J. Begley. 2010. A Quantitative Systems Approach Reveals Dynamic Control of tRNA Modifications during Cellular Stress. PLoS Genet 6:e1001247.

720 Cooley, A. M., G. Carvallo, and J. H. Willis. 2008. Is Floral Diversification Associated with
721 Pollinator Divergence? Flower Shape, Flower Colour and Pollinator Preference in
722 Chilean Mimulus. Annals of Botany 101:641–650.
723 Cooley, A. M., J. L. Modliszewski, M. L. Rommel, and J. H. Willis. 2011. Gene Duplication
724 in Mimulus Underlies Parallel Floral Evolution via Independent trans-Regulatory
725 Changes. Current Biology 21:700–704.
726 Cooley, A. M., and J. H. Willis. 2009. Genetic divergence causes parallel evolution of flower
727 color in Chilean *Mimulus*. New Phytologist 183:729–739.
728 Davies, K. M., N. W. Albert, and K. E. Schwinn. 2012. From landing lights to mimicry: the
729 molecular regulation of flower colouration and mechanisms for pigmentation patterning.
730 Functional Plant Biol. 39:619.
731 del Valle, J. C., A. Gallardo-López, M. L. Buide, J. B. Whittall, and E. Narbona. 2018. Digital
732 photography provides a fast, reliable, and noninvasive method to estimate anthocyanin
733 pigment concentration in reproductive and vegetative plant tissues. Ecol Evol 8:3064–
734 3076.
735 Delannoy, E., M. Le Ret, E. Faivre-Nitschke, G. M. Estavillo, M. Bergdoll, N. L. Taylor, B. J.
736 Pogson, I. Small, P. Imbault, and J. M. Gualberto. 2009. *Arabidopsis* tRNA Adenosine
737 Deaminase Arginine Edits the Wobble Nucleotide of Chloroplast tRNAArg(ACG) and Is
738 Essential for Efficient Chloroplast Translation. The Plant Cell 21:2058–2071.
739 Ding, B., and Y.-W. Yuan. 2016. Testing the utility of fluorescent proteins in Mimulus lewisii
740 by an Agrobacterium-mediated transient assay. Plant Cell Rep 35:771–777.
741 Dubos, C., R. Stracke, E. Grotewold, B. Weisshaar, C. Martin, and L. Lepiniec. 2010. MYB
742 transcription factors in Arabidopsis. Trends in Plant Science 15:573–581.
743 Durbin, M. 2003. Genes that determine flower color: the role of regulatory changes in the
744 evolution of phenotypic adaptations. Molecular Phylogenetics and Evolution 29:507–
745 518.

746 Edger, P. P., R. Smith, M. R. McKain, A. M. Cooley, M. Vallejo-Marin, Y. Yuan, A. J.
747 Bewick, L. Ji, A. E. Platts, M. J. Bowman, K. L. Childs, J. D. Washburn, R. J. Schmitz,
748 G. D. Smith, J. C. Pires, and J. R. Puzey. 2017. Subgenome Dominance in an
749 Interspecific Hybrid, Synthetic Allopolyploid, and a 140-Year-Old Naturally Established
750 Neo-Allopolyploid Monkeyflower. Plant Cell 29:2150–2167.
751 Espley, R. V., C. Brendolise, D. Chagné, S. Kutty-Amma, S. Green, R. Volz, J. Putterill, H.
752 J. Schouten, S. E. Gardiner, R. P. Hellens, and A. C. Allan. 2009. Multiple Repeats of a
753 Promoter Segment Causes Transcription Factor Autoregulation in Red Apples. The
754 Plant Cell 21:168–183.
755 Feller, A., K. Machemer, E. L. Braun, and E. Grotewold. 2011. Evolutionary and
756 comparative analysis of MYB and bHLH plant transcription factors: Plant MYB and
757 bHLH factors. The Plant Journal 66:94–116.
758 Fraser, L. G., A. G. Seal, M. Montefiori, T. K. McGhie, G. K. Tsang, P. M. Datson, E.
759 Hilario, H. E. Marsh, J. K. Dunn, R. P. Hellens, K. M. Davies, M. A. McNeilage, H. N. D.
760 Silva, and A. C. Allan. 2013. An R2R3 MYB transcription factor determines red petal
761 colour in an Actinidia (kiwifruit) hybrid population. BMC Genomics 14:28.
762 Grantham, R. 1974. Amino Acid Difference Formula to Help Explain Protein Evolution.
763 Science 185:862–864.
764 Grotewold, E. 2006. The genetics and biochemistry of floral pigments. Annu. Rev. Plant
765 Biol. 57:761–780.
766 Gu, T., F. W. Buaas, A. K. Simons, C. L. Ackert-Bicknell, R. E. Braun, and M. A. Hibbs.
767 2012. Canonical A-to-I and C-to-U RNA Editing Is Enriched at 3′UTRs and microRNA
768 Target Sites in Multiple Mouse Tissues. PLoS ONE 7:e33720.
769 Hoekstra, H. E., and J. A. Coyne. 2007. The Locus of Evolution: Evo Devo and the
770 Genetics of Adaptation. Evolution 61:995–1016.

771 Kapila, J., R. De Rycke, M. Van Montagu, and G. Angenon. 1997. An Agrobacterium-
772 mediated transient gene expression system for intact leaves. Plant Science 122:101–
773 108.
774 Karcher, D., and R. Bock. 2009. Identification of the chloroplast adenosine-to-inosine tRNA
775 editing enzyme. RNA 15:1251–1257.
776 Knoop, V. 2011. When you can't trust the DNA: RNA editing changes transcript sequences.
777 Cell. Mol. Life Sci. 68:567–586.
778 Koes, R., W. Verweij, and F. Quattrocchio. 2005. Flavonoids: a colorful model for the
779 regulation and evolution of biochemical pathways. Trends in Plant Science 10:236–
780 242.
781 LaFountain, A. M., and Y. Yuan. 2021. Repressors of anthocyanin biosynthesis. New
782 Phytol 231:933–949.
783 Li, C., W. Yu, J. Xu, X. Lu, and Y. Liu. 2022. Anthocyanin Biosynthesis Induced by MYB
784 Transcription Factors in Plants. IJMS 23:11701.
785 Li, J. B., E. Y. Levanon, J.-K. Yoon, J. Aach, B. Xie, E. LeProust, K. Zhang, Y. Gao, and G.
786 M. Church. 2009a. Genome-Wide Identification of Human RNA Editing Sites by
787 Parallel DNA Capturing and Sequencing. Science 324:1210–1213.
788 Li, J.-F., E. Park, A. G. von Arnim, and A. Nebenführ. 2009b. The FAST technique: a
789 simplified Agrobacterium-based transformation method for transient gene expression
790 analysis in seedlings of Arabidopsis and other plant species. Plant Methods 5:6.
791 Lin-Wang, K., K. Bolitho, K. Grafton, A. Kortstee, S. Karunairetnam, T. K. McGhie, R. V.
792 Espley, R. P. Hellens, and A. C. Allan. 2010. An R2R3 MYB transcription factor
793 associated with regulation of the anthocyanin biosynthetic pathway in Rosaceae. BMC
794 Plant Biol 10:50.
795 Lloyd, A., A. Brockman, L. Aguirre, A. Campbell, A. Bean, A. Cantero, and A. Gonzalez.
796 "Advances in the MYB–BHLH–WD Repeat (MBW) pigment regulatory model: Addition

797 of a WRKY factor and co-pption of an anthocyanin MYB for betalain regulation." Plant
798 and Cell Physiology 58: 1431–41.
799 Lowry, D. B., J. M. Sobel, A. L. Angert, T. Ashman, R. L. Baker, B. K. Blackman, Y.
800 Brandvain, K. J. R. P. Byers, A. M. Cooley, J. M. Coughlan, M. R. Dudash, C. B.
801 Fenster, K. G. Ferris, L. Fishman, J. Friedman, D. L. Grossenbacher, L. M. Holeski, C.
802 T. Ivey, K. M. Kay, V. A. Koelling, N. J. Kooyers, C. J. Murren, C. D. Muir, T. C. Nelson,
803 M. L. Peterson, J. R. Puzey, M. C. Rotter, J. R. Seemann, J. P. Sexton, S. N. Sheth,
804 M. A. Streisfeld, A. L. Sweigart, A. D. Twyford, M. Vallejo-Marín, J. H. Willis, K. M.
805 Wright, C. A. Wu, and Y. Yuan. 2019. The case for the continued use of the genus
806 name *Mimulus* for all monkeyflowers. Taxon 68:617–623.
807 Meng, Y., D. Chen, Y. Jin, C. Mao, P. Wu, and M. Chen. 2010. RNA editing of nuclear
808 transcripts in Arabidopsis thaliana. BMC Genomics 11:S12.
809 Montefiori, M., C. Brendolise, A. P. Dare, K. Lin-Wang, K. M. Davies, R. P. Hellens, and A.
810 C. Allan. 2015. In the Solanaceae, a hierarchy of bHLHs confer distinct target
811 specificity to the anthocyanin regulatory complex. Journal of Experimental Botany
812 66:1427–1436.
813 Nguyen, T. A., J. W. J. Heng, P. Kaewsapsak, E. P. L. Kok, D. Stanojević, H. Liu, A.
814 Cardilla, A. Praditya, Z. Yi, M. Lin, J. G. A. Aw, Y. Y. Ho, K. L. E. Peh, Y. Wang, Q.
815 Zhong, J. Heraud-Farlow, S. Xue, B. Reversade, C. Walkley, Y. S. Ho, M. Šikić, Y.
816 Wan, and M. H. Tan. 2022. Direct identification of A-to-I editing sites with nanopore
817 native RNA sequencing. Nat Methods 19:833–844.
818 Pan, Y., M. Li, J. Huang, W. Pan, T. Shi, Q. Guo, G. Yang, and X. Nie. 2022. Genome-
819 Wide Identification and Characterization of RNA/DNA Differences Associated with
820 Drought Response in Wheat. IJMS 23:1405.
821 Puzey, J. R., J. H. Willis, and J. K. Kelly. 2017. Population structure and local selection
822 yield high genomic variation in *Mimulus guttatus*. Mol Ecol 26:519–535.

823 Quattrocchio, F., W. Verweij, A. Kroon, C. Spelt, J. Mol, and R. Koes. 2006. PH4 of Petunia
824 Is an R2R3 MYB Protein That Activates Vacuolar Acidification through Interactions with
825 Basic-Helix-Loop-Helix Transcription Factors of the Anthocyanin Pathway. The Plant
826 Cell 18:1274–1291.
827 Sakurai, M., T. Yano, H. Kawabata, H. Ueda, and T. Suzuki. 2010. Inosine cyanoethylation
828 identifies A-to-I RNA editing sites in the human transcriptome. Nat Chem Biol 6:733–
829 740.
830 Schöb, H., C. Kunz, and F. Meins. 1997. Silencing of transgenes introduced into leaves by
831 agroinfiltration: a simple, rapid method for investigating sequence requirements for
832 gene silencing. Mol Gen Genet 256:581–585.
833 Sobel, J. M., and M. A. Streisfeld. 2013. Flower color as a model system for studies of plant
834 evo-devo. Front. Plant Sci. 4.
835 Sparkes, I. A., J. Runions, A. Kearns, and C. Hawes. 2006. Rapid, transient expression of
836 fluorescent fusion proteins in tobacco plants and generation of stably transformed
837 plants. Nat Protoc 1:2019–2025.
838 Srinivasan, S., A. G. Torres, and L. Ribas de Pouplana. 2021. Inosine in Biology and
839 Disease. Genes 12:600.
840 Stern, D. L., and V. Orgogozo. 2008. The Loci of Evolution: How Predictable Is Genetic
841 Evolution? Evolution 62:2155–2177.
842 Stracke, R., M. Werber, and B. Weisshaar. 2001. The R2R3-MYB gene family in
843 Arabidopsis thaliana. Current Opinion in Plant Biology 4:447–456.
844 Streisfeld, M. A., and M. D. Rausher. 2011. Population genetics, pleiotropy, and the
845 preferential fixation of mutations during adaptive evolution. Evolution 65:629–642.
846 Teichert, I. 2018. Adenosine to inosine mRNA editing in fungi and how it may relate to
847 fungal pathogenesis. PLoS Pathog 14:e1007231.

848 Tian, J., J. Zhang, Z. Han, T. Song, J. Li, Y. Wang, and Y. Yao. 2017. McMYB12
849 Transcription Factors Co-regulate Proanthocyanidin and Anthocyanin Biosynthesis in
850 Malus Crabapple. Sci Rep 7:43715.
851 Wittkopp, P. J., B. K. Haerum, and A. G. Clark. 2008. Regulatory changes underlying
852 expression differences within and between Drosophila species. Nat Genet 40:346–350.
853 Wolf, J. 2002. tadA, an essential tRNA-specific adenosine deaminase from Escherichia coli.
854 The EMBO Journal 21:3841–3851.
855 Wray, G. A. 2007. The evolutionary significance of cis-regulatory mutations. Nat Rev Genet
856 8:206–216.
857 Wu, C. A., D. B. Lowry, A. M. Cooley, K. M. Wright, Y. W. Lee, and J. H. Willis. 2008.
858 Mimulus is an emerging model system for the integration of ecological and genomic
859 studies. Heredity 100:220–230.
860 Yang, Y., R. Li, and M. Qi. 2000. In vivo analysis of plant promoters and transcription
861 factors by agroinfiltration of tobacco leaves. Plant J 22:543–551.
862 Yuan, Y. 2019. Monkeyflowers (*Mimulus*): new model for plant developmental genetics
863 and evo-devo. New Phytol 222:694–700.
864 Yuan, Y.-W., K. J. Byers, and H. Bradshaw. 2013a. The genetic control of flower–pollinator
865 specificity. Current Opinion in Plant Biology 16:422–428.
866 Yuan, Y.-W., J. M. Sagawa, R. C. Young, B. J. Christensen, and H. D. Bradshaw. 2013b.
867 Genetic Dissection of a Major Anthocyanin QTL Contributing to Pollinator-Mediated
868 Reproductive Isolation Between Sister Species of *Mimulus*. Genetics 194:255–263.
869 Zheng, X., K. Om, K. A. Stanton, D. Thomas, P. A. Cheng, A. Eggert, E. Simmons, Y.-W.
870 Yuan, G. D. Conradi Smith, J. R. Puzey, and A. M. Cooley. 2021. The regulatory
871 network for petal anthocyanin pigmentation is shaped by the MYB5a/NEGAN
872 transcription factor in *Mimulus*. Genetics 217:iyaa036.

873 Zhou, W., D. Karcher, and R. Bock. 2014. Identification of Enzymes for Adenosine-to-
874 Inosine Editing and Discovery of Cytidine-to-Uridine Editing in Nucleus-Encoded
875 Transfer RNAs of Arabidopsis. Plant Physiol. 166:1985–1997.
876
877
878
879 **Supplemental tables**
880
881 **Supplemental Table S1.** Primers used. F and R in primer names indicate forward and reverse
882 primer directions with respect to the direction of transcription. *MYB5a* primers bind to both *M. l.*
883 *variegatus* and *M. l. luteus* alleles unless otherwise noted.

Primer ID	Sequence (5' to 3')	Target	Notes
att-R2	CACCACTTTGTACAAGAAAGCTG	pEARLEYGATE vector	
cacc_10F	CACCTTGCAGAGCATGGAAAACAC	MYB5a, end of 5'UTR through beginning of coding sequence	"cacc" tag enables directional cloning into entry vector
GAPDH_1F	TTGAAGGGAATCTTGGGCTA	GAPDH	used as a positive control
GAPDH_2R	CATTTGACGTACCATAAACGAGT	GAPDH	used as a positive control
M13F(-20)	GTAAAACGACGGCCAGT	pENTR entry vector	
M13R(-24)	AACAGCTATGACCATG	pENTR entry vector	

Myb5_10F	TTGCAGAGCATGGAAAACAC	MYB5a, end of 5'UTR through beginning of coding sequence	
Myb5_12F	TGTAGGTGTAAGAAAAGGTGCAT	Beginning of MYB5a Exon 1, luteus allele only	
Myb5_53R	TTAATTAGGCCCCAGTAGGC	End of MYB5a Exon 4, variegatus allele only	includes stop codon
Myb5_57R	CCATCTTCTGTCGTCGTAGTTTC	MYB5a Exon 4	can bind exon 3, though with 2 mismatches
Myb5_64F	GACGGCGGCGAAATTACT	MYB5a Exon 4	also binds exon 3, in luteus only
Myb5_69R	ATTAGGCCCCAGTAGGC	End of MYB5a Exon 4, variegatus allele only	does not include stop codon
pEG-35S-attB1_F	ACGCTCGAGATCACAAGTTT	pEARLEYGATE vector	

Supplemental Table S2. S_{green} pigment intensity data for all analyzed samples. Damaged leaves were excluded from the dataset prior to analysis. A lower S_{green} value corresponds to a greater intensity of red pigmentation. Each pair of *luteus* and *variegatus* transgenes that were infiltrated into the same leaf also share an ID number. In the “Vis_red” column, 1 indicates pigment intensity higher than the average for the negative controls.

ID	Treatment	Vis_red	Sgreen
19	luteus	1	0.45896124

20	luteus	1	0.47148911
21	luteus	1	0.46284725
22	luteus	1	0.47377697
23	luteus	1	0.44480534
24	luteus	1	0.47310128
25	luteus	1	0.46713436
28	luteus	1	0.46502909
30	luteus	1	0.46001506
32	luteus	1	0.47255904
33	luteus	1	0.41304113
34	luteus	1	0.4789886
36	luteus	1	0.45663214
38	luteus	1	0.46054921
39	luteus	1	0.49623173
42	luteus	0	0.50579161
43	luteus	1	0.47908797
44	luteus	1	0.49753298
45	luteus	1	0.48242528
46	luteus	1	0.47112646
48	luteus	0	0.50685166
49	luteus	1	0.48953625
51	luteus	1	0.4946138
52	luteus	1	0.50356411
53	luteus	0	0.51186847
54	luteus	1	0.48408803
55	luteus	1	0.48813383
56	luteus	1	0.49909275
57	luteus	1	0.48070741
58	luteus	1	0.48181555
59	luteus	1	0.48455214
60	luteus	1	0.49736191
62	luteus	1	0.49150095
63	luteus	1	0.4850712
64	luteus	1	0.48059246
65	luteus	1	0.47644334
68	luteus	1	0.48498292
70	luteus	1	0.49468007
71	luteus	1	0.5003172
73	luteus	1	0.48761652
94	luteus	1	0.50385092

95	luteus	1	0.49936115
96	luteus	1	0.48075892
98	luteus	1	0.48728567
99	luteus	1	0.5016936
101	luteus	0	0.50752067
102	luteus	1	0.4857078
104	luteus	1	0.49520219
107	luteus	1	0.48361114
108	luteus	1	0.46805803
128	luteus	0	0.50970645
129	luteus	1	0.49818394
130	luteus	1	0.49095988
131	luteus	1	0.49361628
132	luteus	1	0.49463075
133	luteus	1	0.50097837
134	luteus	0	0.50725935
135	luteus	1	0.50149842
136	luteus	1	0.5012832
137	luteus	1	0.49518309
138	luteus	1	0.502672
139	luteus	0	0.51151598
140	luteus	0	0.509741
141	luteus	0	0.50641157
142	luteus	0	0.51103531
143	luteus	1	0.49785915
144	luteus	1	0.50101639
145	luteus	0	0.51817363
146	luteus	1	0.49673595
147	luteus	1	0.49658392
148	luteus	1	0.50435194
149	luteus	1	0.49478991
150	luteus	1	0.50428831
151	luteus	1	0.49645666
152	luteus	0	0.52145751
153	luteus	0	0.50715806
154	luteus	1	0.4989888
157	luteus	0	0.5279056
158	luteus	0	0.51197026
161	luteus	1	0.48309196
162	luteus	1	0.49193773

167	luteus	1	0.48526353
168	luteus	1	0.49288641
169	luteus	1	0.48946587
170	luteus	1	0.49558514
171	luteus	1	0.48158542
172	luteus	1	0.48120953
173	luteus	1	0.49439967
174	luteus	1	0.50384465
175	luteus	0	0.5114833
176	luteus	1	0.49654562
177	luteus	1	0.48933943
178	luteus	1	0.49135485
179	luteus	1	0.49939164
180	luteus	0	0.51514498
181	luteus	0	0.51083798
182	luteus	0	0.50657476
183	luteus	1	0.50250531
19	variegatus	1	0.47568877
20	variegatus	1	0.48868988
21	variegatus	1	0.46880163
22	variegatus	1	0.476834
23	variegatus	1	0.47909708
24	variegatus	1	0.48677177
25	variegatus	1	0.46918763
28	variegatus	1	0.48686645
30	variegatus	1	0.48905031
32	variegatus	1	0.47155967
33	variegatus	1	0.46714941
34	variegatus	1	0.4838239
36	variegatus	1	0.46573637
38	variegatus	1	0.47379907
39	variegatus	1	0.48784078
42	variegatus	0	0.51741385
43	variegatus	1	0.4794843
44	variegatus	1	0.48321804
45	variegatus	1	0.48173758
46	variegatus	1	0.50040993
48	variegatus	0	0.52520581
49	variegatus	1	0.50306358
51	variegatus	1	0.49606229

52	variegatus	1	0.49970266
53	variegatus	0	0.50902986
54	variegatus	1	0.48608338
55	variegatus	1	0.49465192
56	variegatus	1	0.50304332
57	variegatus	0	0.50605713
58	variegatus	1	0.47967438
59	variegatus	1	0.50309228
60	variegatus	1	0.50266188
62	variegatus	1	0.49723691
63	variegatus	1	0.49567132
64	variegatus	1	0.48869598
65	variegatus	1	0.50188501
68	variegatus	0	0.50639038
70	variegatus	1	0.47943648
71	variegatus	1	0.48588973
73	variegatus	1	0.49556374
94	variegatus	0	0.51197658
95	variegatus	1	0.50227522
96	variegatus	1	0.48536213
98	variegatus	1	0.49801555
99	variegatus	0	0.50660931
101	variegatus	1	0.50401541
102	variegatus	1	0.49641449
104	variegatus	1	0.49945548
107	variegatus	1	0.48027251
108	variegatus	1	0.48198389
128	variegatus	0	0.52135956
129	variegatus	0	0.50577999
130	variegatus	1	0.49626116
131	variegatus	1	0.49693573
132	variegatus	1	0.49175623
133	variegatus	0	0.52033851
134	variegatus	0	0.5189482
135	variegatus	1	0.49854457
136	variegatus	0	0.51242068
137	variegatus	1	0.49919095
138	variegatus	0	0.50810671
139	variegatus	0	0.51677281
140	variegatus	0	0.50841716

141	variegatus	0	0.50649654
142	variegatus	0	0.51604924
143	variegatus	1	0.49630413
144	variegatus	0	0.51275691
145	variegatus	0	0.52673709
146	variegatus	1	0.49852838
147	variegatus	1	0.49625944
148	variegatus	0	0.50718913
149	variegatus	1	0.49693789
150	variegatus	0	0.50736216
151	variegatus	0	0.50968741
152	variegatus	0	0.52494794
153	variegatus	0	0.50955283
154	variegatus	1	0.50310402
157	variegatus	0	0.53515949
158	variegatus	0	0.50600882
161	variegatus	1	0.48170408
162	variegatus	1	0.49886199
167	variegatus	1	0.48870643
168	variegatus	1	0.50240691
169	variegatus	1	0.4864144
170	variegatus	1	0.50069586
171	variegatus	1	0.49361669
172	variegatus	1	0.4820093
173	variegatus	1	0.49556454
174	variegatus	0	0.51279883
175	variegatus	0	0.51206144
176	variegatus	1	0.49898868
177	variegatus	1	0.49354642
178	variegatus	1	0.49645237
179	variegatus	1	0.49709808
180	variegatus	0	0.51553433
181	variegatus	0	0.51572583
182	variegatus	0	0.51128091
183	variegatus	1	0.50510844
184	pGFP		0.52134394
185	pGFP		0.49493349
186	pGFP		0.50720189
187	pGFP		0.48427841
188	pGFP		0.50338806

189	pGFP	0.52251484
190	pGFP	0.50606369
191	pGFP	0.5069713
192	pGFP	0.51953843
193	pGFP	0.50197984
194	pGFP	0.52441718
195	pGFP	0.48660946
196	pGFP	0.50257154
197	pGFP	0.5067341
198	pGFP	0.48867401
199	pGFP	0.50075806
200	pGFP	0.49162335
201	pGFP	0.51551854
202	pGFP	0.50892671
203	pGFP	0.52311187
204	pGFP	0.52580258
205	pGFP	0.48691814
206	pGFP	0.50820532
207	pGFP	0.4929026
208	pGFP	0.50398024
209	pGFP	0.51763381
210	pGFP	0.50977392
211	pGFP	0.50713293
212	pGFP	0.52212015
213	pGFP	0.50461793
214	pGFP	0.51932389
215	pGFP	0.48918478
216	pGFP	0.49894963
217	pGFP	0.49885496
218	pGFP	0.48199319
219	pGFP	0.49959746
220	pGFP	0.49140784
221	pGFP	0.51215037
222	pGFP	0.50583205
223	pGFP	0.52408216

Supplemental figures

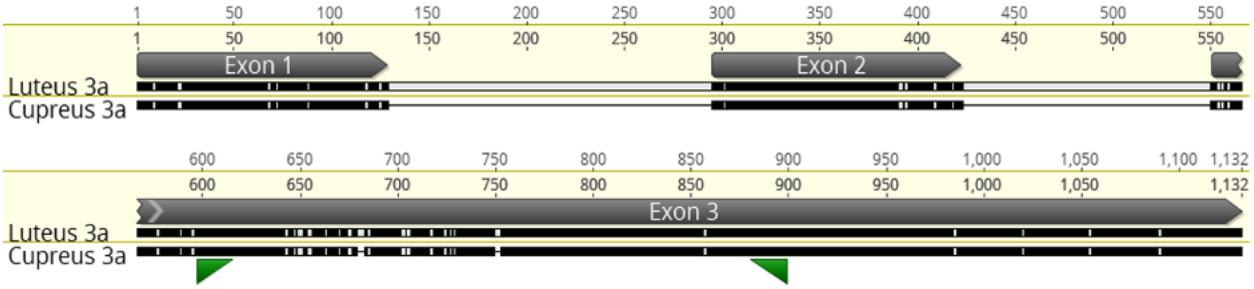
894

895

896

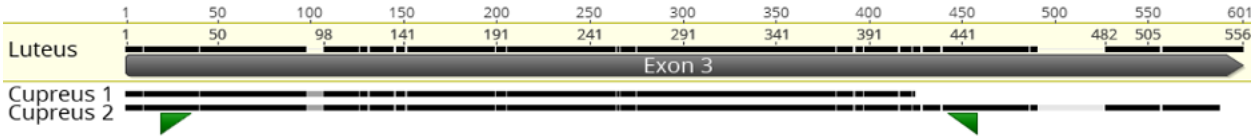
897

898 A



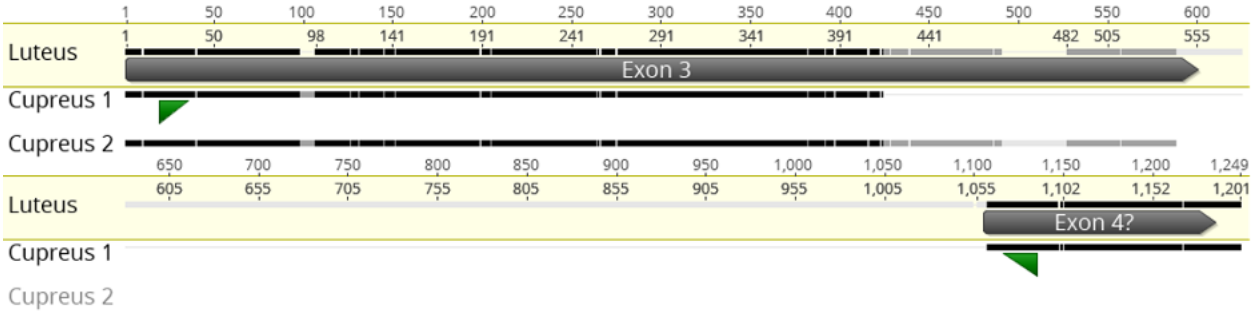
899

900 B



901

902 C



903

904 **Supplemental Figure S1.** *Mimulus cupreus MYB3a* and *MYB2* transcripts aligned to *M. luteus*

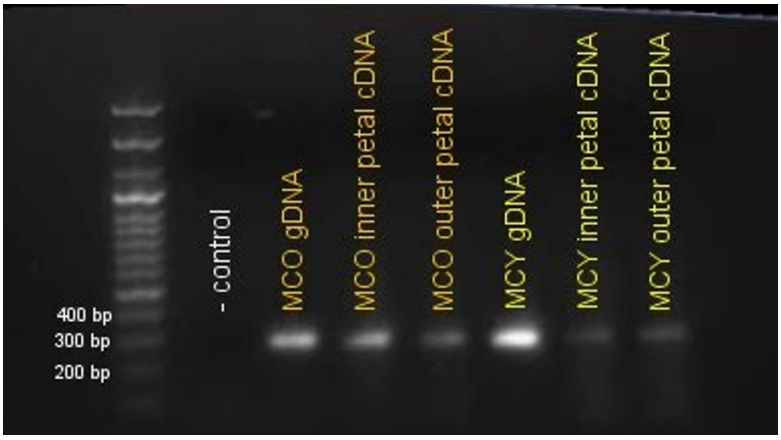
905 genomic sequence. A. *MYB3a*. Primers Myb2/3_1F and Myb3a_2R are shown in green. B.

906 *MYB2*.

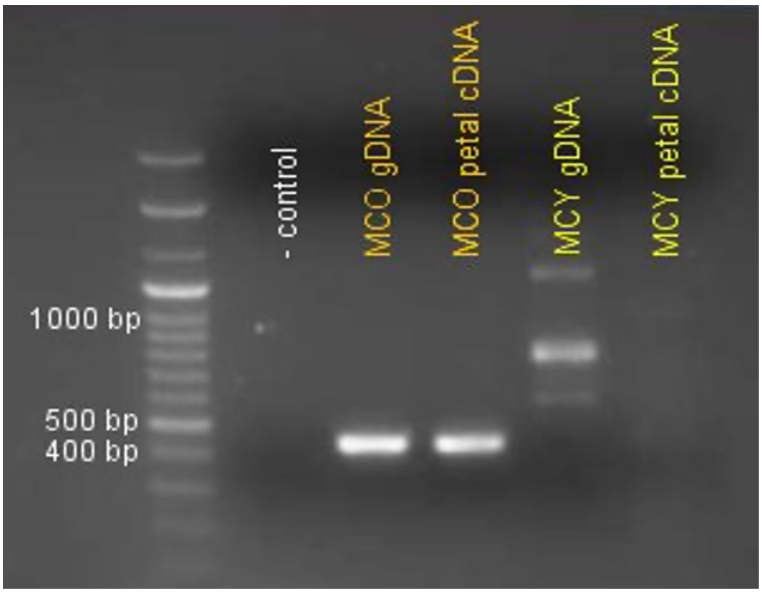
907 Two *M. cupreus* transcripts were recovered from transcriptome sequence; the region in which

908 they diverge is labeled as “Cupreus 1” and “Cupreus 2,” possibly corresponding to alternative

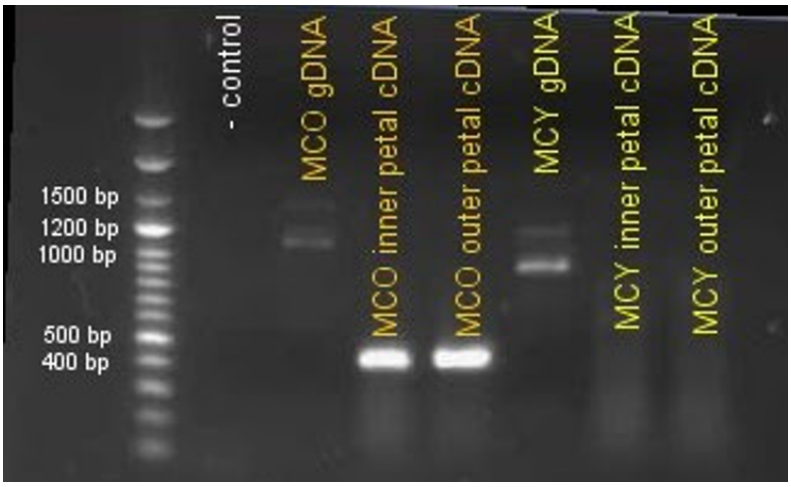
909 splice variants. Primers Myb 2_1F and Myb 2b_5R shown in green. These primers are expected
910 to amplify only transcript 2. C. The same *MYB2* transcripts with primers Myb 2_1F and Myb
911 2b_7R shown in green.
912



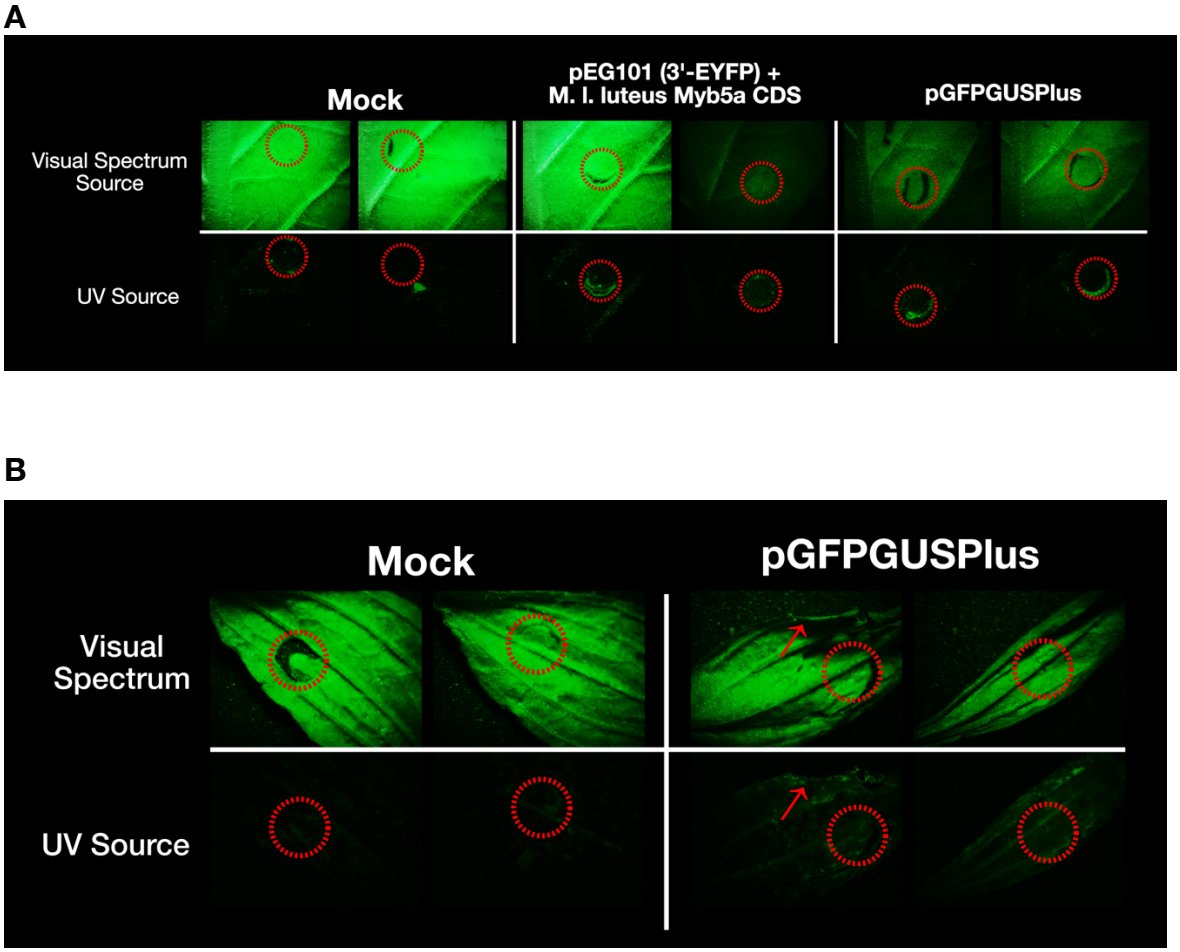
913
914 **Supplemental Figure S2.** Primers Myb 2/3_1F and Myb 3a_2R were used to amplify a portion
915 of *Myb3a* exon 3, encoding a transcription factor gene that is one of two candidates for the gain
916 of petal anthocyanin pigmentation in the orange-flowered morph of *M. cupreus* (Cooley et al.,
917 2011). The recent loss of pigmentation in the rare yellow morph of *M. cupreus* also maps to the
918 same region. A product of the expected length (300 bp) was amplified out of gDNA for both
919 orange and yellow-flowered *M. cupreus.* The product was also amplified out of cDNA for both
920 inner and outer petal tissue from both morphs, indicating that Myb 3a **is** expressed in all of
921 these tissues.
922



Supplemental Figure S3. Primers Myb 2_1F and Myb 2b_5R were used to amplify a portion of *Myb2* transcript 2, exon 3, encoding a transcription factor gene that is the second of two candidates for the derived gain of petal anthocyanin pigmentation in the orange-flowered morph of *M. cupreus*, and the even more recent loss of petal anthocyanin in the rare yellow morph. A product of expected length (around 450 bp) was amplified out of orange-flowered *M. cupreus* gDNA. The product was also amplified out of cDNA from both inner and outer petal from orange-flowered *M. cupreus,* indicating that it is expressed in the orange morph. However, no product of the expected length was amplified out of yellow-flowered *M. cupreus* gDNA. Several longer products were amplified less brightly, and are likely due to nonspecific annealing. No product was seen in the yellow-flowered petal cDNA either.



Supplemental Figure S4. Primers Myb 2_1F and Myb 2b_7R were used to amplify a region of *Myb2* extending from the beginning of the third exon through the beginning of the predicted fourth exon. The expected length of the product spanning the third intron of Myb2 was estimated to be approximately 1050 bp based on the *M. luteus* genome. A PCR amplification done using standard Taq polymerase showed bands at 1050 bp and 1500 bp in orange-flowered *M. cupreus* gDNA and at 950 bp and 1300 bp in yellow-flowered *M. cupreus* gDNA. Amplification using long-amp Taq polymerase revealed that a 950 bp product was also present in gDNA from orange-flowered *M. cupreus*. Sequencing of these products showed that the 1050 bp fragment from orange *M. cupreus* corresponds to *Myb2*, while the two bands seen in yellow *M. cupreus* contain sequences with no resemblance to any anthocyanin-related Myb gene. In orange-flowered *M. cupreus* cDNA from both inner and outer petal, a product of approximately 430 bp was amplified. This indicates that the transcript is spliced as expected, and is expressed in inner and outer petal tissue of orange-flowered *M. cupreus.* No product was observed in the amplification out of yellow-flowered *M. cupreus* petal cDNA, indicating that the Myb 2b transcript is not expressed in yellow *M. cupreus* petals.



954 **Supplemental Figure S5.** Screen for GFP fluorescence in transfected *N. tabacum* and *M.*
955 *lewisii* leaves. Fluorescence microscopy was performed on a Jenco GL7-3L zoom stereo
956 microscope (Jenco International, Portland, OR, USA) equipped with a NIGHTSEA green filter
957 (NIGHTSEA LLC, Lexington, MA, USA). Light was provided by both a full-spectrum light source
958 supplied with the microscope as well as an external NIGHTSEA ultraviolet light (UV) light
959 source, which was used to excite fluorescence in tissue infiltrated with GFP-expressing
960 plasmids. Leaves infiltrated with pGFPGUSPlus were used as a positive control; leaves
961 infiltrated with a mock control and YFP-expressing pEARLEYGATE101 with *M. l. luteus Myb5a*
962 CDS were used as negative controls. A) Fluorescence microscopy with *Nicotiana tabacum*
963 leaves. Left, two replicates treated with 5% sucrose; middle, two replicates treated with
964 pEARLEYGATE101 with *M. l. luteus Myb5a* CDS in *Agrobacterium tumefaciens*; right, two

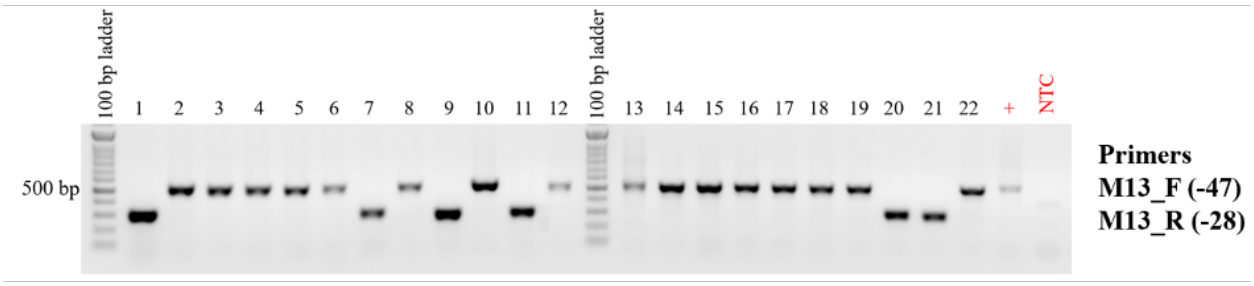
965 replicates treated with pGFPGUSPlus in *A. tumefaciens*. Top row, visual spectrum light source;
966 bottom row, UV source. Some leaves imaged with visual spectrum light source appear dimmer;
967 this is an artifact caused by the inadvertent use of a different light setting. Red circles
968 correspond to injection sites. B) Fluorescence microscopy with *Mimulus lewisii* leaves. Left, two
969 replicates treated with 5% sucrose; right, two replicates treated with pGFPGUSPlus in *A.*
970 *tumefaciens*. Top row, visual spectrum light source; bottom row, UV source. Red circles
971 correspond to injection sites. Note apparent tissue damage at edge of leaves treated with
972 pGFPGUSPlus (red arrows).
973

Myb5-10F
Myb5_12F
Exon1
1. luteus
Frame 2
2. variegatus
Frame 2
Exon 1
Exon2
Exon3
Exon 2
Exon 3
64F
Subgroup 6
2 mismatches to 57R
5
Exon4
Exon 4
53R
57R

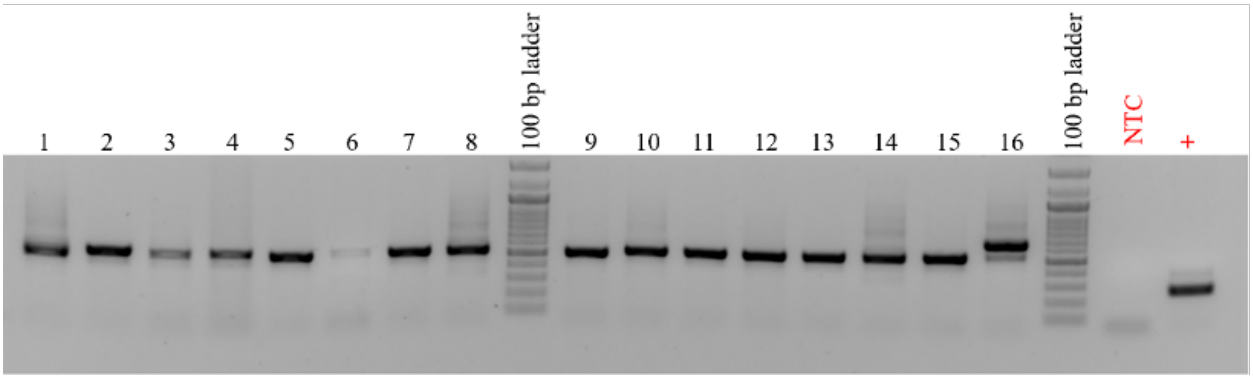
974

Supplemental Figure S6. Alleles of *MYB5a* from the yellow-petaled *M. l. luteus* (top sequence on each line) and the magenta-petaled (anthocyanin-pigmented) *M. l. variegatus* (bottom sequence on each line). Primers mentioned in the Methods are shown in dark green (forward primers) and light green (reverse primers). Exons are marked in gray. Two putative A-to-I editing sites are marked in yellow. Note that Subgroup 6, which is a hallmark of all known anthocyanin-activating R2R3 MYBs, is present in both taxa in exon 4 as the amino acid sequence RPRPRTF. In exon 3, however, a frameshift mutation in *M. l. variegatus* eliminates this feature, making it unlikely that the exon 1-2-3 splice variant of *M. l. variegatus* would be capable of anthocyanin activation.

A

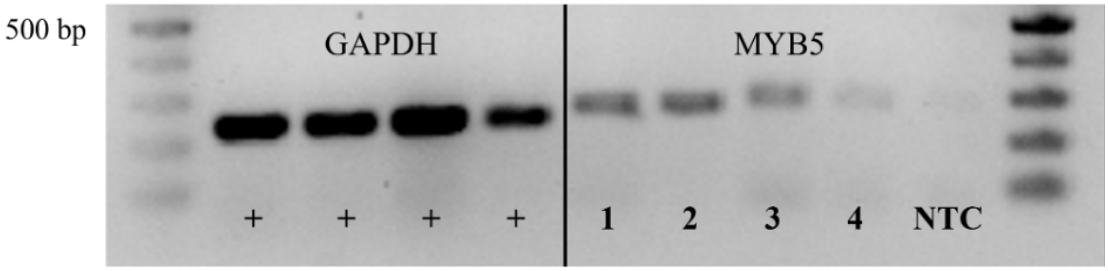


B



991 **Supplemental Figure S7.** pGEM Cloning of an exon 4 fragment of *MYB5a*. A fragment
992 containing the first edited site was PCR-amplified using primers 64F-57R and cloned, and
993 colonies were screened for fragment insertion. Lanes with band size ~500 bp are colonies that
994 contain the *MYB5a* exon 4 insert. Lanes with bands ~300 denote empty vectors. (A) Primers
995 M13_F (-47) and M13_R (-28) were used to screen white colonies from a pGEM cloning attempt
996 with *M. l. variegatus* x *M. l. luteus* F1 hybrid gDNA. 16 colonies appear to have taken up the
997 MYB5 exon 4 insert. (B) Primers M13_F (-47) and M13_R (-28) were used to screen white
998 colonies from a pGEM cloning attempt with Mlv gDNA. 15 colonies appear to have successfully
999 taken up the insert. The band at ~650bp most likely denotes a vector with a longer, incorrect
1000 insert.

1001



1002

1003 **Supplemental Figure S8.** Relative expression of *MYB5a* in *M. l. variegatus* and *M. l. luteus*.
1004 Left: positive control using GAPDH primers 1F-2R, on each of the samples shown on the right.
1005 Right: An exon 4 fragment of *MYB5a* amplified using Myb5_64F and Myb5_57R from cDNA
1006 from *M. l. variegatus* inner petal (1) and outer petal (2); the red-spotted *M. l. luteus* inner petal
1007 (3); and the yellow *M. l. luteus* outer petal (4). NTC, No Template Control. All band sizes were
1008 as expected.

1009

1010

1011 **Supplemental Figure S9.** Leaf photos utilized in quantitative comparisons of anthocyanin
1012 production. Photos excluded due to leaf damage are not included. A. pGFP negative controls,
1013 with two transgene injections per leaf. B. Leaves for which the *M. l. variegatus MYB5a* exon 1-2-
1014 4 transgene was injected on the left side of the leaf and the *M. l. luteus* allele on the right. C.
1015 Leaves for which the *M. l. variegatus MYB5a* exon 1-2-4 transgene was injected on the right
1016 side of the leaf and the *M. l. luteus* allele on the left. Files uploaded separately.
1017
1018
1019