# THE GENETIC BASIS FOR THE REPEATED ORIGIN OF RED FLOWERS IN *MIMULUS AURANTIACUS*

by

AVA KIRBY

A THESIS

Presented to the Department of Biology
and the Robert D. Clark Honors College
in partial fulfillment of the requirements for the degree of
Bachelor of Science

June 2022

## An Abstract of the Thesis of

Ava Kirby for the degree of Bachelor of Science
in the Department of Biology to be taken June 2022

Title: The Genetic Basis for the Repeated Origin of Red Flowers in *Mimulus aurantiacus*

Approved: *Matthew Streisfeld, Ph.D.*
Primary Thesis Advisor

Understanding the process of speciation requires investigation of traits that can lead to divergence between populations. This thesis examines the genetic basis of variation in red pigmentation in flowers and vegetative tissues, traits that may contribute to reproductive isolation in the *Mimulus aurantiacus* (bush monkeyflower) species complex. Utilizing a series of genetic crosses that segregate for color differences, I analyzed differences among taxa in flower color patterns and vegetative anthocyanin content to investigate the genetic basis for the repeated emergence of these traits. I provide evidence that the genetic basis of red flowers differs between the two red-flowered taxa in the *M. aurantiacus* species complex: *M. parviflorus* and the red ecotype of *M. puniceus*. These results suggest parallel evolution of this trait. I also further characterize the role of the gene *MaMyb1* in influencing flower color. Finally, I add to previous evidence that suggests a correlation between vegetative and floral anthocyanins, substantiating the possibility that variation in this trait contributes to reproductive isolation. These findings add to our understanding of how reproductive isolation may have evolved in *M. aurantiacus* and motivate further investigation concerning the origins of these traits that were essential for the initial stages of speciation in this group.

## Acknowledgements

I would first like to express my sincere gratitude to my Primary Thesis Advisor, Dr. Matt Streisfeld, for guiding me through the process of researching and writing a thesis. His help and feedback were invaluable throughout the completion of this project. I would also like to thank Aidan Short for serving as second reader on my thesis committee and helping me learn how to code using R. Thank you to Dr. Corinne Bayerl for serving as my CHC representative and teaching one of the most interesting classes I have taken in college. I am grateful to Susan Belcher and the staff at the greenhouse for keeping the plants healthy throughout the duration of these experiments. I would also like to thank UO GrEBES for their support, which helped make this project possible.

I am also incredibly thankful for my wonderful friends and family, who listened to me talk about this project for over a year and provided encouragement when I needed it the most. I would especially like to thank my parents for supporting me every step of the way throughout college.

# Table of Contents

# List of Figures

# Introduction

One of the central goals in evolutionary biology is understanding how new species form. The development of distinct species is referred to as speciation, which is crucial to the evolution of the vast diversity of life on earth. Speciation occurs when populations become reproductively isolated from each other; that is, when they no longer breed with one another. When natural selection favors different traits in different populations, they can become increasingly distinct from each other. This can lead to the development of reproductive isolation, which can occur either through prezygotic isolation or postzygotic isolation. Prezygotic isolation involves the prevention of mating through mechanisms, such as the geographic separation of organisms or different mating periods. In contrast, postzygotic isolation occurs when the resulting hybrid is unfit or unable to reproduce. These mechanisms, among others, provide barriers to reproduction among populations, leading to reproductive isolation (Coyne & Orr, 2004). Understanding this process can provide insight into how the many species in the tree of life evolved, including humans.

This thesis project seeks to further understand the genetic basis of speciation through the model *Mimulus aurantiacus*, the bush monkeyflower complex, which is native to southern California. Different taxa (distinct groups of organisms) within this group vary in flower color, with most producing yellow flowers but two taxa that make red flowers. This variation in flower color may lead to differences in pollinator visitation among taxa that reduces mating, thus potentially leading to speciation. However, where the habitat of these taxa overlap, hybridization often occurs (McMinn, 1951). This makes *M. aurantiacus* a particularly interesting case study in terms of the

study of speciation, as they are not entirely distinct from each other due to the continued exchange of genetic material (Sobel & Streisfeld, 2015). This provides an opportunity to gain insight into speciation while it is in progress. Many studies focus on identifying genetic indicators of speciation long after the process occurred, by studying distinct but closely related organisms (Nosil & Schluter, 2011). However, equally important is the study of speciation as it is occurring. Studying the genetic material of diverging populations can reveal specific genes that contribute to early reproductive isolation and potentially provide insight into the overarching process of speciation (Via, 2009).

This project builds upon previous work in the Streisfeld lab focusing on the evolutionary history of *M. aurantiacus* and the genetic changes that led to variation in flower color. These experiments were conducted with the following overarching questions in mind: How do genetic changes in *M. aurantiacus* contribute to reproductive isolation, and what is the evolutionary history that led to the emergence of these genetic changes? In the following sections, I will provide further background information on the processes involved in speciation in the *M. aurantiacus* system before describing the experiments I performed to address these questions.

**Introgressive Hybridization**

Hybridization occurs when members of diverging populations interbreed and thus exchange genetic material. One of the consequences of hybridization is introgression, where the genetic information can be transferred across species boundaries due to hybridization and continued backcrossing. If this introgressed genetic variation decreases the organism's fitness in their environment, it will be selected against, and eventually removed from the population. This selection against gene flow

results in barriers to exchanging genetic material at these locations in the genome, which is necessary for speciation. However, introgression also has the potential to result in beneficial variants of genes being introduced into a new population, a process referred to as adaptive introgression. Under this model, beneficial alleles can rise in frequency if the traits they provide the organism with result in an increase in fitness (Abbott et al., 2013). Adaptive introgression represents another way that introgression can contribute to speciation, as sometimes the introduction of these beneficial variants can lead to the evolution of reproductive isolation.

Though introgression has the potential to eventually bring about speciation, it also has the potential to reduce differentiation between diverging populations (Abbott et al., 2013). The sharing of genetic material that occurs during hybridization leads to a reduction in genetic differentiation between populations as allele frequencies become more similar. Thus, divergence between populations can only persist when another force, such as natural selection, overcomes the homogenizing effects of gene flow. There are many examples of introgression leading to divergence between natural populations, one of the most famous being Darwin's finches. One study revealed the important role of hybridization in introducing genetic variation in this group, eventually leading to divergence in beak shape as different populations adapted to their particular environments (Lamichhaney et al., 2015). Therefore, introgression can be an integral mechanism leading to the origin of new species, including in *M. aurantiacus*.

**Variation in *Mimulus aurantiacus***

As previously stated, taxa within the *M. aurantiacus* species complex vary in flower color, with some growing yellow flowers and others red flowers. For this thesis project, the main taxa of interest include *M. puniceus*, *M. longiflorus*, *M. parviflorus*, and *M. aridus*. *M. puniceus*, native to the San Diego area, contains the closely related red and yellow-flowered ecotypes. These ecotypes commonly hybridize where their ranges come into contact with each other. The yellow-flowered *M. longiflorus* occurs along the coast of southern California near Los Angeles, and the yellow-flowered *M. aridus* grows to the east of San Diego in the desert of southeastern California. Finally, the red-flowered *M. parviflorus* grows on islands off the coast of southern California (McMinn, 1951) (Figure 1). A strong geographic association with flower color suggests that the divergence between the red and yellow ecotypes of subspecies *M. puniceus* is due to natural selection that is being maintained in the face of gene flow (Streisfeld & Kohn, 2005; Stankowski et al., 2015). This force of selection is likely due to adaptation to pollinator preferences, with hummingbirds preferring red flowers and hawkmoths preferring yellow flowers (Streisfeld & Kohn, 2007). This adaptation to different pollinators in the ecotypes of *M. puniceus* appears to be driving the initial stages of speciation between them.

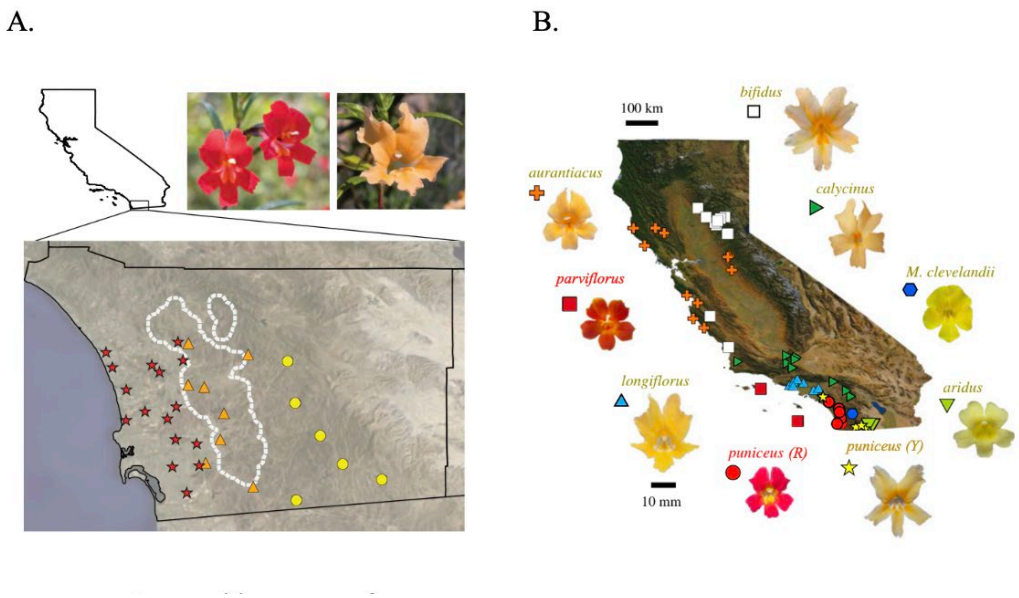


Figure 1: Geographic ranges of *M. aurantiacus* taxa

(A) The red and yellow ecotypes of *M. puniceus*. The red-flowered ecotype occurs in the west and the yellow-flowered ecotype occurs in the east. From Streisfeld et al. (2013). (B) A map of the geographic distribution of taxa in the *M. aurantiacus* species complex, including the four taxa of interest in this thesis. Adapted from Stankowski & Streisfeld (2015).

The main genetic basis of flower color differences in this species has been characterized, with a mutation in the *MaMyb2* gene regulating the amount of anthocyanin present in the plants (Streisfeld & Rausher, 2009a; Streisfeld et al., 2013). Anthocyanins are compounds that are responsible for red pigments in plants. Taxa that produce red flowers have a high concentration of anthocyanin in their petals, but the pigment is absent in yellow flowers. Previous work provides strong evidence that *MaMyb2* is under selection. *MaMyb2* allele frequencies are correlated with the geographic distribution of the red and yellow ecotypes, with fixation of the red allele in the west and the yellow allele in the east (Stankowski et al., 2015) (Figure 2). This geographic correlation with allele frequencies is not present in other areas of the

genome, indicating that selection maintains differentiation at the flower color locus, whereas selectively neutral areas of the genome remain similar due to gene flow (Streisfeld et al., 2013).

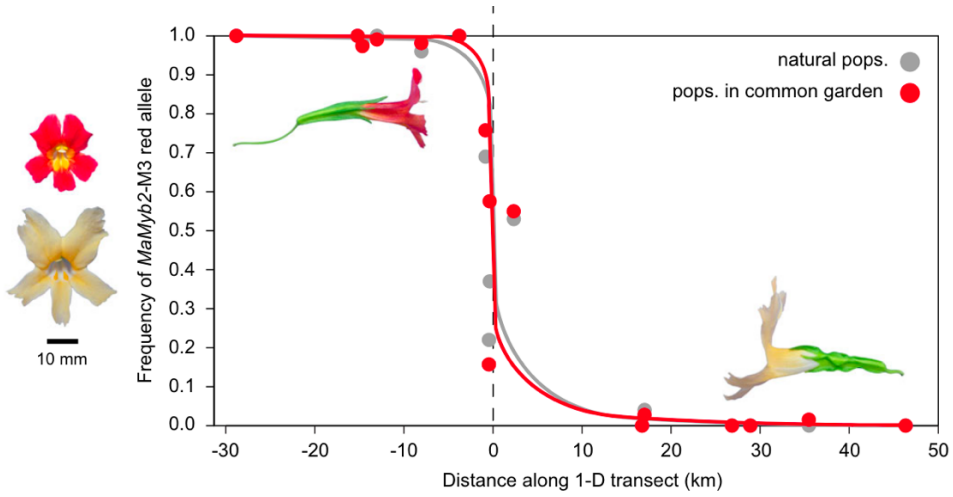


Figure 2: Geographic distribution of *MaMyb2* red allele frequencies

Frequency of the red allele at *MaMyb2*, which is strongly correlated with the red ecotype. The red allele is fixed in the red ecotype west of the hybrid zone and absent in the yellow ecotype to the east. From Stankowski et al. (2015).

Genomic analysis demonstrates that *M. puniceus* and *M. longiflorus* are relatively closely related, as they group together among other taxa in the same evolutionary clade (a group of organisms that includes a common ancestor and all descended from it). However, *M. parviflorus* and *M. aridus* are more distantly related from these taxa and are grouped in a separate clade (Chase et al., 2017). Though these evolutionary relationships have been outlined, there is much that is still unknown about the history of how the *M. aurantiacus* species complex became so diverse. Stankowski & Streisfeld (2015) showed that a variant of the *MaMyb2* appeared to have brought into the red ecotype of *M. puniceus* through introgression with another member of the species complex. The likely donor of this allele was believed to be a now extinct red-

flowered taxon similar to *M. aridus*, as the DNA sequence at *MaMyb2* in modern-day *M. aridus* is remarkably similar to that of *M. puniceus* despite these taxa being distantly related. Stankowski & Streisfeld (2015) hypothesized that the red allele was selected for in the red ecotype following this introgression event, as it was beneficial for attracting hummingbird pollinators.

The red flowers of *M. parviflorus* are also produced by a mutation in the *MaMyb2* gene, but it is unclear if this is the result of the same mutation present in *M. puniceus*. Given that *M. aridus* and *M. parviflorus* are closely related, it is possible that the mutation in *MaMyb2* arose in the common ancestor of these two taxa and was transferred to *M. puniceus* through introgression before red flowers went extinct in the ancestor of present-day *M. aridus*. Alternatively, the mutation at *MaMyb2* in *M. parviflorus* could be distinct from the one observed in *M. puniceus*, even though it occurs in the same gene. The ancestor of *M. aridus* that hybridized with the ancestor of *M. puniceus* may have instead been yellow-flowered at the time of introgression, and a mutation leading to red flowers in *MaMyb2* arose in the red ecotype of *M. puniceus* soon after this event occurred, while a separate mutation in *MaMyb2* arose in *M. parviflorus* (Stankowski & Streisfeld, 2015).

It was unknown which of these hypotheses was accurate at the time of publication (Stankowski & Streisfeld, 2015). However, when *M. parviflorus* was crossed with *M. aridus*, distinct flower color patterns were observed on the front, side, and back of the petals, something that had not been previously observed in crosses between the red and yellow ecotypes (Figure 3). This may be due to a different genetic basis for flower color between *M. aridus* and the yellow ecotype, or it could be caused

by a difference in the genetic basis for red flowers between *M. parviflorus* and the red ecotype, which would indicate that the mutation in *MaMyb2* differed between the two red-flowered taxa. The latter possibility would provide evidence for parallel evolution occurring in *M. aurantiacus*. Parallel evolution occurs when the same trait evolves independently multiple times in response to similar sources of selection. One example of parallel evolution in nature is found in the genus *Ipomoea*, where distinct mutations led to the emergence of red flowers multiple times (Streisfeld & Rausher, 2009b). In the case of *M. aurantiacus*, it is possible that the parallel evolution of red flowers occurred twice in response to pollinator selection.

Figure 3: Variation in flower color in an *M. parviflorus* x *M. aridus* cross

Flowers from a cross between *M. parviflorus* and *M. aridus* generated in Stankowski & Streisfeld (2015). In this image, the yellow in the flowers was removed in photoshop to clearly show the differences in anthocyanin pigmentation, which varies across the front, side, and back of many of the flowers.

An association has also been identified between the red allele at *MaMyb2* and the presence of anthocyanin in vegetative tissue, specifically the stem and leaves (Sobel

et al. 2019). Vegetative anthocyanins are known to provide protection to plants in stressful conditions, such as drought and UV radiation (Chalker-Scott, 1999). The presence of stressed leaf anthocyanin in *M. puniceus* varies according to geographic location and is correlated with variation in flower color between the red and yellow ecotypes (Sobel et al 2019). This association suggests that vegetative anthocyanins could be subject to selection based on differing geographic environments. Although the variation in vegetative anthocyanin between ecotypes is likely not enough to drive speciation on its own, it may act as another reproductively isolating trait that also contributes to divergence (Sobel et al., 2019).

Although the role of the gene *MaMyb2* in flower color divergence has been identified, little is known about the effects of the related gene *MaMyb1*. *MaMyb1* and *MaMyb2* are known to interact in interesting ways that influence flower color. This interaction between genes is known as epistasis. In *M. puniceus*, when plants inherit two copies of the yellow allele at *MaMyb2* (referred to as genotype *YY*), flowers will be yellow regardless of the genotype at *MaMyb1*. However, when plants inherit two copies of the red allele at *MaMyb2* (*RR*), the genotype at *MaMyb1* influences flower color variation. Specifically, if the genotype at *MaMyb1* is *RR*, then plants will have red flowers. In contrast, if the genotype at *MaMyb1* is *YY*, then plants will have yellow flowers with only a small amount of red pigment on the tube of the flower (Streisfeld unpublished data) (Figure 4).


Figure 4: Epistasis between *MaMyb1* and *MaMyb2* influences flower color

Flower color phenotypes of *M. puniceus* and their corresponding genotypes at *MaMyb1* and *MaMyb2*. From Benson (2015).

A remaining question regarding *MaMyb1* is whether dominance exists at this gene—if the genotype at *MaMyb1* is *RY* (one copy each of the red and yellow alleles), how does that influence the amount of red pigment present in the plants? Is the red allele completely dominant to the yellow allele, meaning that it masks the presence of the yellow allele entirely, or is the result somewhere in between the *RR* and *YY* genotypes? Investigating these questions may help shed light on the role that epistasis and dominance at *MaMyb1* play in anthocyanin production, and by extension its potential role in the evolution of red flowers in the *M. aurantiacus* species complex. If flowers with the genotype *RY* at *MaMyb1* have similar levels of anthocyanin pigment to those with the genotype *RR*, this would suggest that the red allele is dominant to the yellow allele at this locus, showing that at least one red allele is required to produce the highest levels of anthocyanin pigment in the flower. If this is the case, and pollinators choose to visit flowers based on their color, then genotypes *RR* and *RY* at *MaMyb1* would provide a plant with higher fitness than *YY* plants in the west (where the red

ecotype occurs). This would imply that *MaMyb1* plays an important role in the emergence of red flowers in *M. puniceus* through its interaction with *MaMyb2*. These questions were explored in this thesis.

# Methods

## *Mimulus parviflorus* Crosses

To further understand the genetics underlying variation in flower color among the taxa of the *M. aurantiacus* species complex, I grew seeds from previously made crosses that segregate for flower color. The red-flowered *M. parviflorus* was crossed with three different yellow-flowered taxa: *M. aridus*, *M. longiflorus*, and the yellow ecotype of *M. puniceus*. By using the same red-flowered parent that was crossed to different yellow-flowered parents, I could assess whether the genetic backgrounds of the yellow-flowered taxa were different. These initial crosses produced a first generation of hybrids, known as the F1 generation, which were then self-fertilized to create the F2 generation, which segregated for flower color differences. These F2 seeds were grown in a greenhouse until they began flowering.

The presence of red flower color is controlled by the genes underlying the production of the red pigment anthocyanin. To measure the presence of anthocyanin, I recorded the phenotype on the front, side, and back of the petals for each of the plants that produced flowers. These were categorized on a scale from 0 to 5, with 0 being a light yellow phenotype and 1 being an orange-yellow phenotype, both with no anthocyanin present, and 2-5 corresponding to increasingly red phenotypes. Additionally, I recorded the presence or absence of anthocyanin in the stem, leaves, and calyx (the leaf-like structure at the base of the petals) of the plants. Quantitative methods that would normally be used to measure the presence of anthocyanin were not possible for this particular experiment, as measuring the levels of anthocyanin quantitatively requires sampling the front and back of the petal simultaneously. Since

previous work revealed distinct flower color patterns on the front, side, and back of the petals in a cross between *M. parviflorus* and *M. aridus* (Figure 3), a categorical scoring system was chosen as a more effective measurement in this case.

### ***Mimulus puniceus*** **Crosses**

For comparison to the *M. parviflorus* crosses described above, additional crosses were made between the red and yellow ecotypes of the subspecies *M. puniceus*. By comparing this cross to the *M. parviflorus* x yellow ecotype cross, I could assess if the offspring of both crosses produced similar flower color patterns across the different segments of the flower. This would indicate that the genetic basis for red flowers is the same in both taxa. These crosses were also used to further understand the role of the gene *MaMyb1* in flower color variation. I compared flower colors of plants heterozygous at *MaMyb1* but homozygous for the red allele at *MaMyb2* (*RY/RR*) with those homozygous for the red allele at both genes. If both genotypes produced similar levels of anthocyanin, this would provide support for dominance of the red allele at *MaMyb1*.

Using F3 plants obtained from F2 individuals generated from a cross between red and yellow ecotype parents (Streisfeld et al., 2013), I made a series of genetic crosses. Each of the F3 plants I used was homozygous for the red or yellow allele at *MaMyb1* and *MaMyb2*. Listed in the order of *MaMyb1* and *MaMyb2* respectively, these plants had the genotypes of *RR/RR*, *YY/RR*, or *–/YY*, with the *MaMyb1* genotype unknown. I made 36 different crosses with these plants to create a variety of hybrid genotypes: *RY/RR*, *RR/RY*, *RY/RY*, *RY/YY*, and *YY/RY*. I also self-fertilized each of the plants to recreate the original homozygous genotypes for comparison. I grew the seeds

that resulted from these crosses in the same greenhouse as the *M. parviflorus* crosses and examined their flower color patterns and anthocyanin content using the same scale from 0-5 mentioned previously.

**Data Analysis**

I analyzed the flower color data for this experiment in the program RStudio (version 4.1.2). I performed a multiple correspondence analysis (MCA), a tool used to analyze categorical data and group individuals based on how similar or different they are from each other, using the R packages FactoMineR (Lê et al., 2008) and factoextra (Kassambara & Mundt, 2020). I performed two MCAs, one using the data from the *M. parviflorus* crosses and the other using the data from the *M. puniceus* crosses. If the data separated into clusters corresponding to each cross, this would suggest that the genetic background of yellow taxa had an effect on flower color, whereas a lack of clear separation would imply no effect. I also compared the two plots to each other to determine if there was a difference in flower color distribution because of the different red-flowered taxa. If the plots revealed similar patterns of flower color segregation across the different segments of the flower, it would support a shared genetic basis for flower color between *M. parviflorus* and the red ecotype of *M. puniceus*.

I also used Microsoft Excel (version 16.60) to generate graphs of vegetative anthocyanin data. I used a Fisher's exact test (Schoonjans, n.d.) to test for significant differences in the frequency of vegetative anthocyanin between different levels of floral anthocyanin. The p-value threshold for significance was 0.05. This allowed me to investigate whether a correlation existed between anthocyanin in the flower and anthocyanin in vegetative tissues in these crosses.

# Results

Of the 204 plants generated for the *M. parviflorus* crosses, 144 produced flowers and were used in my analysis: 55 from the *parviflorus* x *longiflorus* cross, 51 from the *parviflorus* x *puniceus* (yellow ecotype) cross, and 38 from the *parviflorus* x *aridus* cross. Of the 156 plants generated for the *M. puniceus* crosses, 82 produced flowers and were used in my analysis: 17 with the genotype *RR/RR*, 12 with the genotype *YY/RR*, 7 with the genotype *RY/RR*, 8 with the genotype *–/YY*, 20 derived from a cross between parents with *RR/RR* x *–/YY* genotypes, and 18 derived from a cross between *YY/RR* x *–/YY* genotypes.

**MCA of Flower Color**

The MCA of the flower color data collected on the crosses with *M. parviflorus* revealed an association between the flower color score of 0 (light yellow flowers) on the front, side, and back of the petals, as these categories clustered together. This means that individuals categorized with a score of 0 on the front were usually categorized as 0 on the side and back as well, showing that individuals with light yellow flowers had little to no variation in flower color among the different segments of the flower. The same was true for individuals with a score of 1 (orange-yellow flowers), as the categories for a score of 1 on the front, side, and back also formed a separate group. In contrast, there was almost no separation among flower color score categories if anthocyanin was present anywhere in the flowers (scores of 2-5). This reveals that individuals with floral anthocyanin in the *M. parviflorus* crosses displayed different levels of pigmentation across the segments of the flower, rather than always being

categorized as having the same flower color score on the front, side, and back. There was also little separation based on the identity of the yellow-flowered parent (Figure 5).

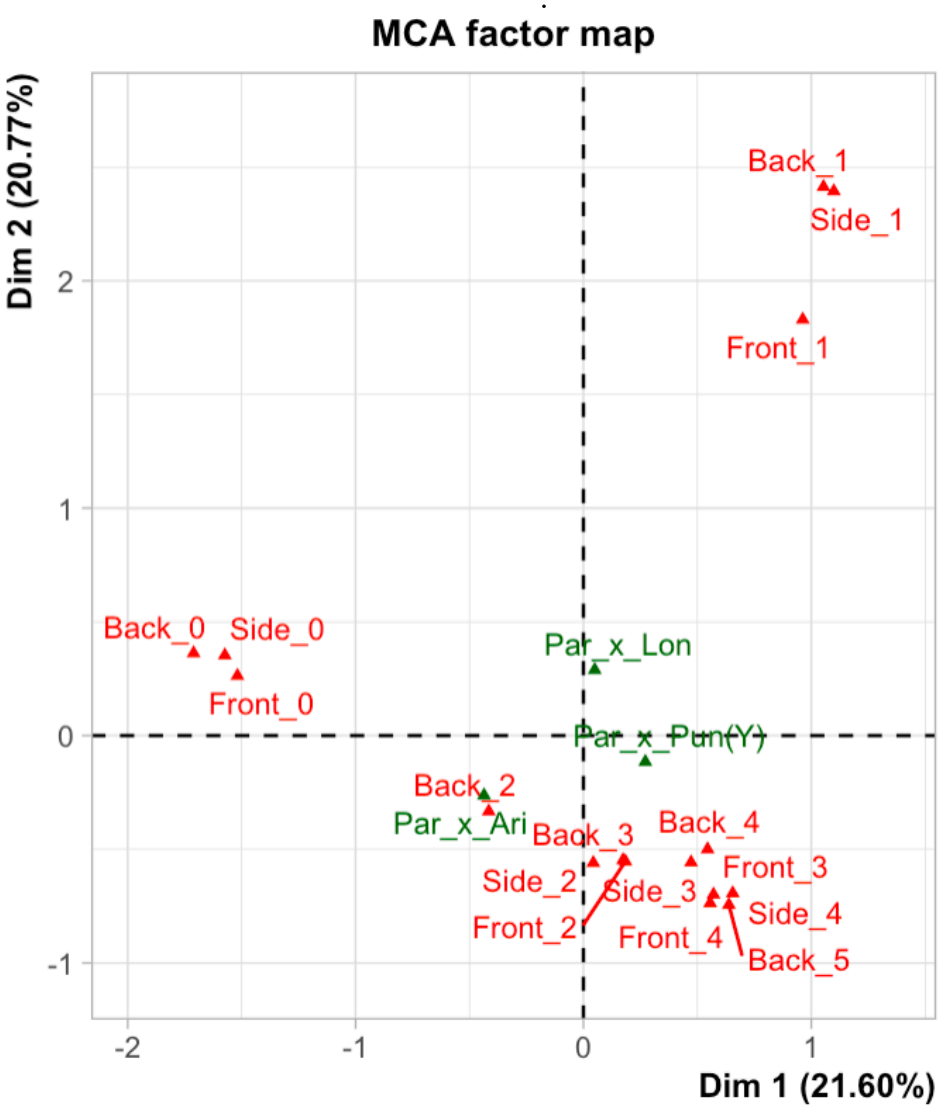


Figure 5: Flower color MCA for *M. parviflorus* crosses

An MCA of flower color data of *M. parviflorus* crosses. Flower color categories for each segment of the flower are shown in red and the positions of the crosses relative to these categories are shown in green. The labels Par_x_Lon, Par_x_Pun(Y), and Par_x_Ari represent the crosses between *M. parviflorus* and *M. longiflorus*, the yellow ecotype of *M. puniceus*, and *M. aridus*, respectively.

Similar to the *M. parviflorus* analysis, data from the crosses between the red and yellow ecotypes of *M. puniceus* also demonstrated that flowers with a color score of 1 (orange-yellow flowers) on the front, side, and back of the petals grouped together in the MCA. However, in contrast to the *M. parviflorus* crosses, the categories of medium anthocyanin pigment (a score of 3) on the front, side, and back and high anthocyanin pigment (scores of 4 or 5) on the front, side, and back also formed separate groups in the *M. puniceus* crosses, showing that individuals with floral anthocyanin displayed the same levels of pigmentation across the segments of the flower. The light shade of yellow scored as 0 and the spotty anthocyanin pigment scored as 2 were not represented in these crosses, in contrast to the *M. parviflorus* crosses.

The *MaMyb1/MaMyb2* genotypes also showed separation corresponding to the flower color scores. Plants with the *–/YY* genotype grouped with orange-yellow flowers (scored as 1), plants with the *RR/RR* and *RY/RR* genotypes grouped with red flowers (scored as 4 or 5), and those derived from the cross *RR/RR* x *–/YY* grouped with flowers with an intermediate amount of red pigment (scored as 3). Plants with the genotype *YY/RR* and those derived from the cross *YY/RR* x *–/YY* were placed intermediately between the flower color score categories of 1 and 3 (Figure 6). The common grouping of plants with the *MaMyb1* genotypes *RR* and *RY* shows that plants with these genotypes produced flowers with similar levels of floral anthocyanin.

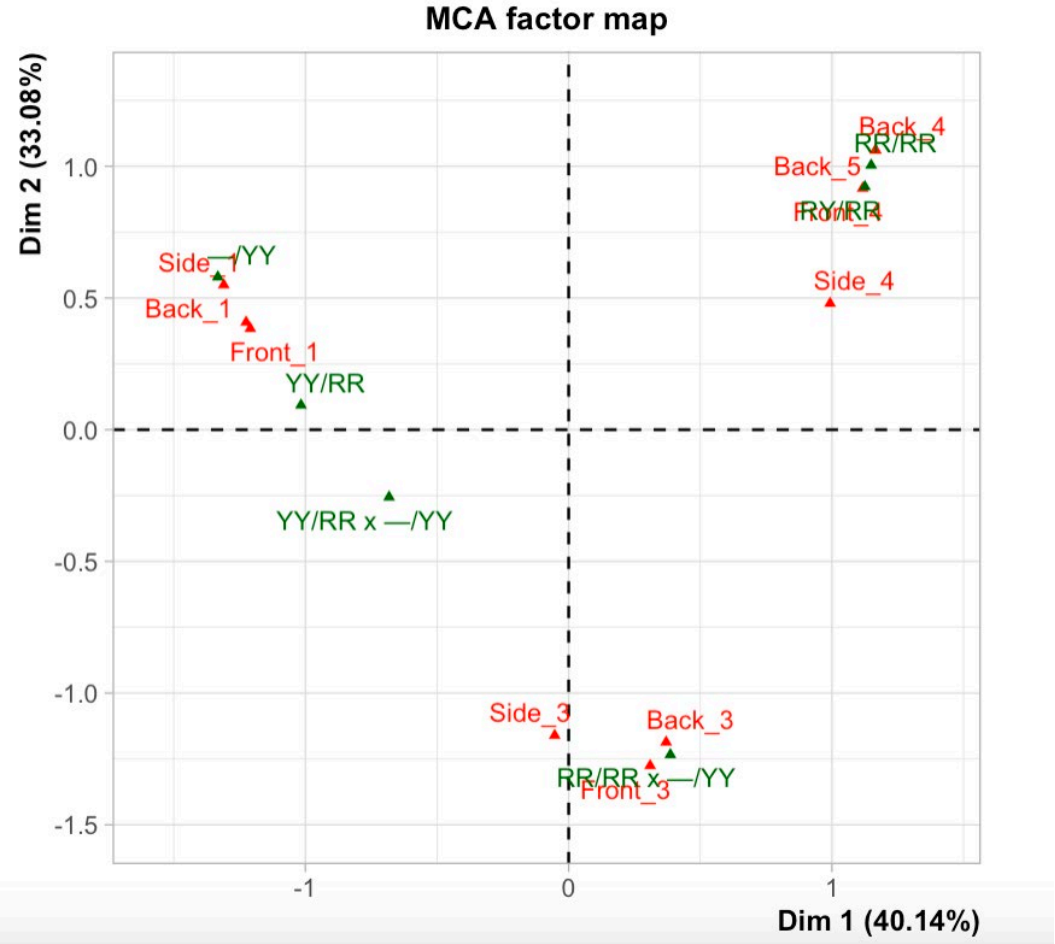


Figure 6: Flower color MCA for *M. puniceus* crosses

An MCA of flower color data from crosses between the red and yellow ecotype of *M. puniceus*. Flower color score categories for each segment of the flower are shown in red and the positions of the genotypes at *MaMyb1/MaMyb2* relative to these categories are shown in green.

## Vegetative Anthocyanin

In both sets of crosses, plants that had anthocyanin in their flowers (a flower color score of 2 or above) were more likely to also have anthocyanin in the stem and calyx in comparison to plants with no anthocyanin in their flowers (a flower color score of 0 or 1). These differences were statistically significant by a Fisher's exact test in the *M. parviflorus* crosses (stem: p=0.016, calyx: p=2.62e-6) and the *M. puniceus* crosses

(stem: p=1.63e-4, calyx: p<0.000000001). Plants that had anthocyanin in their flowers were also more likely to have anthocyanin in their leaves in the *M. parviflorus* crosses (p=0.026). However, there was no significant difference in leaf anthocyanin in the *M. puniceus* crosses (p=0.147) (Figures 7 and 8).

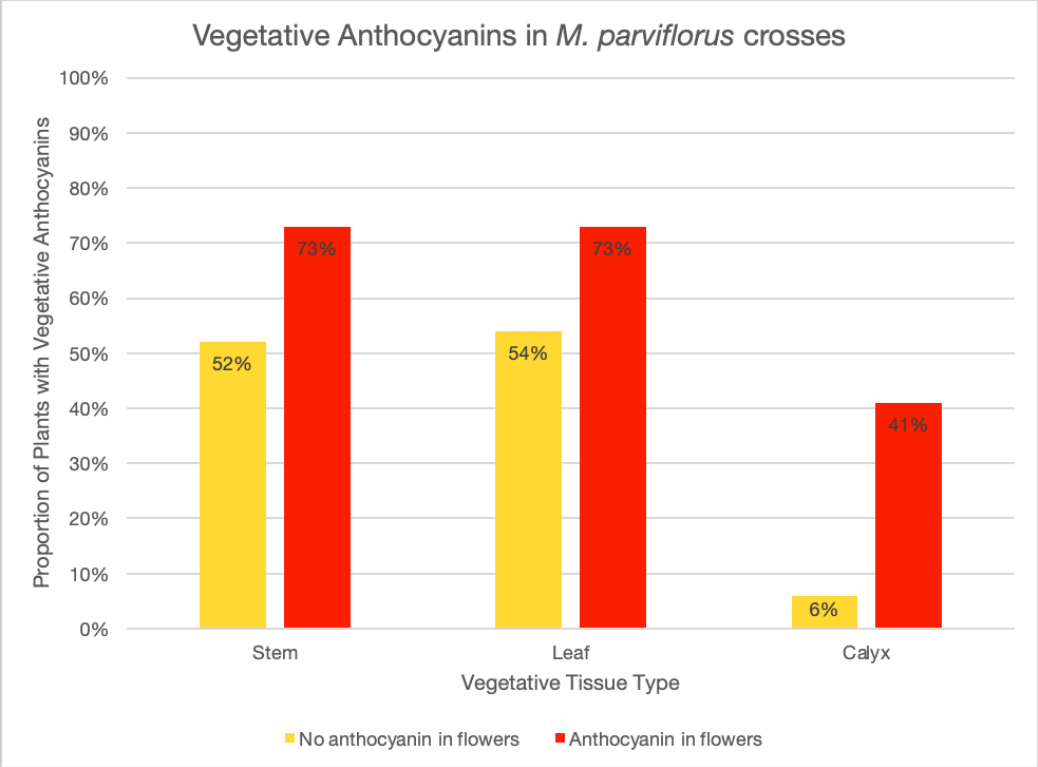


Figure 7: Vegetative anthocyanin in *M. parviflorus* crosses

A significantly higher proportion of plants with anthocyanin in their flowers (a flower color score of 2 or above) had anthocyanin in the vegetative tissue when compared to plants with no anthocyanin in their flowers (a flower color score of 0 or 1).

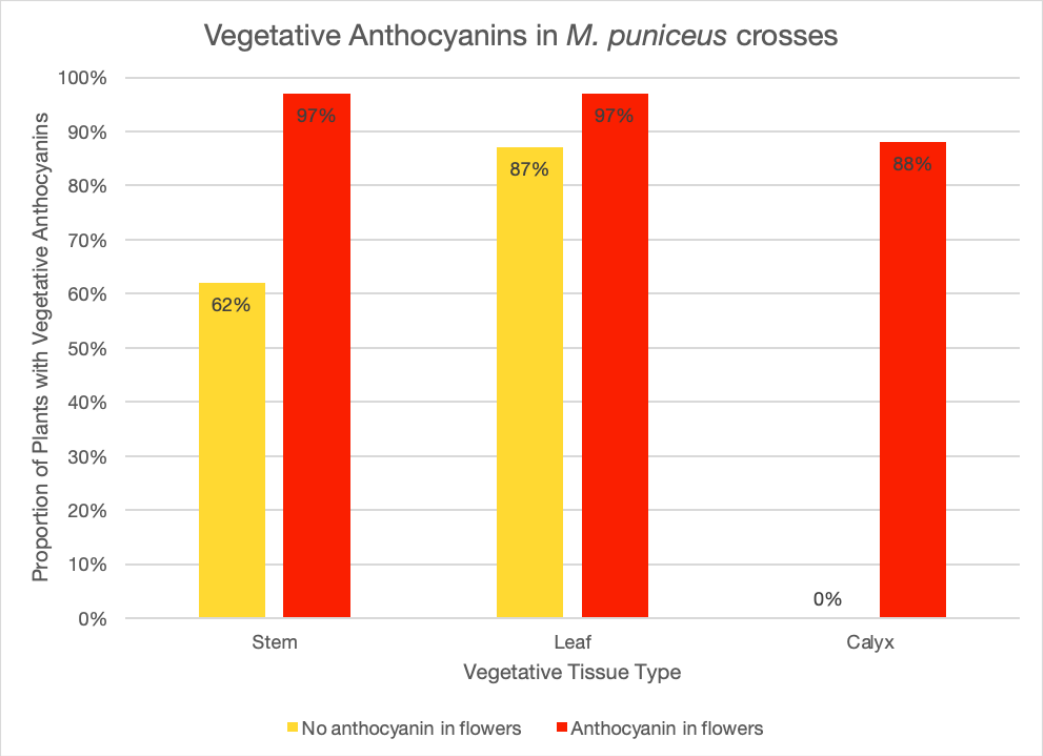


Figure 8: Vegetative anthocyanin in *M. puniceus* crosses

A significantly higher proportion of plants with anthocyanin in their flowers (a flower color score of 3 or above) had anthocyanin in the stem and calyx when compared to plants with no anthocyanin in their flowers (a flower color score of 1). However, this difference was not significant in the leaves (p=0.147).

# Discussion

To elucidate the origin of the genetic basis of red flowers in the *M. aurantiacus* complex, I compared crosses between the red-flowered *M. parviflorus* and three yellow-flowered taxa with a cross between the red and yellow ecotypes of *M. puniceus*. Although previous work showed that red flowers in both taxa were caused by a mutation at *MaMyb2* (Stankowski and Streisfeld 2015), it was unknown whether they were caused by the same mutation in both cases. In addition, I examined the *M. puniceus* crosses to test for dominance of the red allele in the gene *MaMyb1*. Finally, I examined vegetative anthocyanin content in both sets of crosses to test for previously observed correlations between vegetative anthocyanin and floral anthocyanin.

## The origin of variation in flower color

In the *M. parviflorus* crosses, there were some flowers that had no anthocyanin on any segments of the flower, while others were fully pigmented on every segment. However, many flowers in this set of crosses showed varying levels of anthocyanin across the front, side, and back of the flower, leading to a lack of clear separation of pigmented flower color score categories (Figure 5). This contrasted with the results of the cross between the red and yellow ecotypes of *M. puniceus*, where flower color scores were always similar among segments of the flower. This is reflected in the MCA for the *M. puniceus* crosses, as the flower color score categories on the front, back, and side are associated with the same pigmentation level and form separate groups (Figure 6).

The lack of similar patterns in the grouping of flower color data between the *M. parviflorus* and the *M. puniceus* red ecotype MCAs indicates that the genetic basis of

red flowers may differ between these two taxa. Based on the results, the different red taxa appear to be the main factor causing these contrasting flower color segregation patterns, as little to no separation occurred when comparing the different yellow-flowered taxa. This also suggests that the genetic basis of yellow flowers is similar in the three taxa used in this study. The differing flower color segregation patterns revealed by this experiment suggest the possibility that the mutation in *MaMyb2* arose separately in *M. parviflorus* and *M. puniceus*. This provides preliminary evidence for parallel evolution in *M. aurantiacus*, with red flowers evolving twice independently even though introgression of a distinct variant of *MaMyb2* occurred in the red ecotype of *M. puniceus*.

Another possibility to consider is that another gene besides *MaMyb2* was involved in causing these differing flower color patterns. For example, *MaMyb1* appears to affect flower color in the red and yellow ecotype, but it is currently unknown if it has an influence in the other taxa. A second gene, such as *MaMyb1*, could account for the different flower color patterns observed here rather than distinct mutations in *MaMyb2*. Experiments aimed at investigating whether other genes, including *MaMyb1*, influence flower color in *M. parviflorus* would be an important next step to help clarify whether this difference in flower color segregation patterns is truly caused by different mutations in *MaMyb2*.

Several additional limitations of this experiment call for these results to be interpreted with caution. First, the sample size was relatively low, with only 144 plants among the three *M. parviflorus* crosses producing flowers and 82 of the *M. puniceus* crosses producing flowers. This increases the risk that the differences observed could be

due to random chance rather than biological differences. However, a previous *M. parviflorus* x *M. aridus* cross where 146 flowers were scored using similar methods also revealed flower color patterns that were distinct across segments of the flower (Stankowski & Streisfeld 2015; Figure 3). The manner in which data was collected is also a potential source for error. Although categorical scoring of the level of anthocyanin pigment was chosen as a more effective method due to the varying phenotypes across the different segments of the flower, this method of collecting data is inherently less reliable than quantitative methods.

In addition, the setup of this experiment does not lend itself to direct comparison between these two groups for a variety of reasons. The red ecotype of *M. puniceus* was only crossed to one yellow taxon instead of being crossed to multiple as was *M. parviflorus*. Although the differences between the *M. parviflorus* x yellow ecotype and red ecotype x yellow ecotype cross support the conclusion that the genetic basis of red flowers differs, more compelling evidence to support this hypothesis would involve crossing the red ecotype to multiple yellow-flowered taxa. It would also be important to use the same individuals as the red and yellow-flowered parents in each cross, as this would control for variation in the genetic basis of flower color within each taxon. Furthermore, the plants generated by the *M. puniceus* crosses were F4 generation plants, whereas those generated by the crosses with *M. parviflorus* were of the F2 generation. Previous crosses between the red and yellow ecotype of *M. puniceus* led to flower color patterns in the F2 generation similar to those observed here in the F4 generation (Stankowski et al., 2015), so this is unlikely to have impacted my results. However, examining plants of the same generation would be preferable.

*M. puniceus* plants with the genotype *RY/RR* were grouped with plants with the genotype *RR/RR* in the MCA, indicating that these genotypes produce similar levels of anthocyanin in their flowers. This suggests that the red allele is dominant to the yellow allele at the *MaMyb1* gene, supporting previous evidence found in a cross between the red and yellow ecotype (Streisfeld unpublished data). The *RY/RR* and *RR/RR* genotypes were associated with the flower color score categories of 4 or 5, the highest levels of anthocyanin pigment. In contrast, plants with the genotype *YY/RR* produced very low levels of anthocyanin and grouped between the flower color scores categories of 1 and 3. Therefore, these results imply that selection would operate similarly on red ecotype plants with the *RY* and *RR* genotypes at *MaMyb1* when the genotype at *MaMyb2* is *RR*, as this produces a similar floral phenotype. In contrast, since a genotype of *YY* at *MaMyb1* leads to a reduction in floral anthocyanin, selection is likely to not favor plants with the genotype *YY/RR* in the red ecotype. However, as with the comparison between the *M. parviflorus* and *M. puniceus* crosses, these findings are only preliminary due to small sample size. This is perhaps true here to a greater extent, as only seven *RY/RR* plants were used for this analysis as a consequence of low success rate of this cross and low amounts of seeds produced within the ones that were successful.

Evidence of dominance of the red allele at *MaMyb1* suggests that this gene has an important role in shaping flower color in *M. puniceus*, and that *MaMyb2* does not fully account for variation in flower color. This is important because it shows that multiple genes can influence traits that contribute to reproductive isolation. In addition, developing a more complete understanding of the genetic basis of flower color is imperative to fully understand the divergence and potential speciation between the

different taxa of *M. aurantiacus*. Characterizing the influence of *MaMyb1* is an important step in solidifying this understanding.

**Vegetative anthocyanin is associated with increased floral anthocyanin**

I also found evidence for a correlation between anthocyanin pigment in the flowers and anthocyanin pigment in vegetative tissues, as having flowers with some anthocyanin pigment made a plant more likely to produce anthocyanin in the stem, leaves, and calyx. The observed higher frequency of vegetative anthocyanin was significantly correlated with higher levels of floral anthocyanin except in the leaves of the *M. puniceus* crosses. However, it is possible that the lack of significance here is simply due to small sample size.

In Sobel et al. (2019), vegetative anthocyanins, which can develop in response to stress, were found to be more prevalent in the red ecotype. Anthocyanin in the leaves and stems was also found to be associated with the red allele at *MaMyb2* in *M. puniceus*. This suggests that *MaMyb2* has more than one phenotypic effect, rather than just accounting for the difference in flower color. Alternatively, genetic linkage of *MaMyb2* to another gene which controls vegetative anthocyanin may be responsible for this association. My results build upon these previous conclusions by revealing that a similar correlation exists between floral and vegetative anthocyanin in *M. parviflorus*. In addition, I provide evidence for this relationship in the calyx in addition to the stem and leaves. Again, since vegetative anthocyanin correlates with geographic location in *M. puniceus*, which is likely subject to differing selection pressures (Sobel et al., 2019), it is possible that the variation in this trait represents another reproductively isolating trait that contributes to incipient speciation. This may strengthen the reproductive

isolation caused by adaptation to different pollinators by further reducing gene flow between red-flowered and yellow-flowered taxa.

**Future Directions**

Although these results provide support for additional traits and genetic changes that contribute to reproductive isolation in *M. aurantiacus*, further study is required to solidify these findings. While the differences between the *M. parviflorus* and *M. puniceus* crosses indicate that the mutations causing red flowers in these two taxa differ, suggesting parallel evolution, a more comprehensive experiment involving crossing the red ecotype and *M. parviflorus* to multiple yellow taxa would allow further testing of this hypothesis. In addition, finding a way to quantitatively measure anthocyanin content on individual segments of the flower would provide firmer evidence than the qualitative scoring method used here.

Additionally, further analysis of the effect of *MaMyb1* on the flower color phenotype is required. If dominance of the red allele exists at *MaMyb1* in the *M. puniceus*, it is likely that the frequency of *RY* and *RR* genotypes at *MaMyb1* would be relatively equal in the red ecotype, as they produce similar levels of anthocyanin, presumably leading hummingbirds (the pollinator that prefers red flowers) to visit plants with these genotypes at fairly equal rates. There would also likely be reduced frequency of the genotype *YY* at *MaMyb1* in the red ecotype, as it would lead to reduced floral anthocyanin and therefore a reduction in pollinator visitation. Genotyping of individuals from multiple populations of the red ecotype would allow for testing of this hypothesis.

Further work will also be required to determine the role of vegetative anthocyanins in reproductive isolation. Although previous work indicates that selection across different environments acts on stressed leaf anthocyanins in *M. puniceus* (Sobel et al., 2019), investigations as to whether a similar selection pressure is acting on *M. parviflorus* have not been carried out. Similarly, an association between gene expression of *MaMyb2* and vegetative anthocyanins has not been established in *M. parviflorus*. Exploring these questions would be important next steps to determine whether there is evidence that vegetative anthocyanins are a potential source of reproductive isolation in *M. parviflorus*.

By investigating a variety of different lines of evidence, I contribute to opening these new directions of research and have taken steps to further understand the evolution of reproductive isolation in the *M. aurantiacus* species complex. I provide evidence for independent origins of red flowers in the two taxa, rather than them evolving red flowers once, as was previously hypothesized. In addition, my results support the conclusion that the red allele at *MaMyb1* is dominant to the yellow allele, which contributes to furthering understanding of the genetic basis of this trait. Finally, I have provided support that vegetative and floral anthocyanins are associated with each other and with genetic variation at *MaMyb2*. These findings expand current knowledge of the history of divergence of this system and motivate further investigation, advancing understanding of the many factors that can lead to speciation.

# Bibliography


Abbott, R., Albach, D., Ansell, S., Arntzen, J. W., Baird, S. J. E., Bierne, N., Boughman, J., Brelsford, A., Buerkle, C. A., Buggs, R., Butlin, R. K., Dieckmann, U., Eroukhmanoff, F., Grill, A., Cahan, S. H., Hermansen, J. S., Hewitt, G., Hudson, A. G., Jiggins, C., … Zinner, D. (2013). Hybridization and speciation. *Journal of Evolutionary Biology*, *26*(2), 229–246. https://doi.org/10.1111/j.1420-9101.2012.02599.x

Benson, C. (2015). *The genetic architecture of a reproductively isolating trait: A widespread role of epistasis?* [Undergraduate Thesis, University of Oregon].

Chalker-Scott, L. (1999). Environmental significance of anthocyanins in plant stress responses. *Photochemistry and Photobiology*, *70*(1), 1–9. https://doi.org/10.1111/j.1751-1097.1999.tb01944.x

Coyne, J. A., & Orr, H. A. (2004). *Speciation*. Sinauer Associates.

Chase, M. A., Stankowski, S., & Streisfeld, M. A. (2017). Genomewide variation provides insight into evolutionary relationships in a monkeyflower species complex (*Mimulus* sect. *Diplacus*). *American Journal of Botany*, *104*(10), 1510–1521. https://doi.org/10.3732/ajb.1700234

Kassambara, A. and Mundt, F. (2020). factoextra: Extract and Visualize the Results of Multivariate Data Analyses. R package version 1.0.7. https://CRAN.R-project.org/package=factoextra

Lamichhaney, S., Berglund, J., Almén, M. S., Maqbool, K., Grabherr, M., Martinez-Barrio, A., Promerová, M., Rubin, C.-J., Wang, C., Zamani, N., Grant, B. R., Grant, P. R., Webster, M. T., & Andersson, L. (2015). Evolution of Darwin's finches and their beaks revealed by genome sequencing. *Nature*, *518*(7539), 371–375. https://doi.org/10.1038/nature14181

Lê, S., Josse, J. & Husson, F. (2008). FactoMineR: An R Package for Multivariate Analysis. *Journal of Statistical Software*. 25(1). pp. 1-18. 10.18637/jss.v025.i01

McMinn, H. E. (1951). Studies in the genus *Diplacus*, Scrophulariaceae. *Madroño, 11*(2), 33–128. http://www.jstor.org/stable/41422708

Nosil, P., & Schluter, D. (2011). The genes underlying the process of speciation. *Trends in Ecology & Evolution*, *26*(4), 160–167. https://doi.org/10.1016/j.tree.2011.01.001

Schoonjans, F. (n.d.). *MedCalc's Fisher exact probability calculator (version 20.109)*. MedCalc. Retrieved May 10, 2022, from https://www.medcalc.org/calc/fisher.php

Sobel, J. M., Stankowski, S., & Streisfeld, M. A. (2019). Variation in ecophysiological traits might contribute to ecogeographic isolation and divergence between parapatric ecotypes of *Mimulus aurantiacus*. *Journal of Evolutionary Biology*, *32*(6), 604–618. https://doi.org/10.1111/jeb.13442

Sobel, J. M., & Streisfeld, M. A. (2015). Strong premating reproductive isolation drives incipient speciation in *Mimulus aurantiacus*. *Evolution*, *69*(2), 447–461. https://doi.org/10.1111/evo.12589

Stankowski, S., Sobel, J. M., & Streisfeld, M. A. (2015). The geography of divergence with gene flow facilitates multitrait adaptation and the evolution of pollinator isolation in *Mimulus aurantiacus*. *Evolution*, *69*(12), 3054–3068. https://doi.org/10.1111/evo.12807

Stankowski, S., & Streisfeld, M. A. (2015). Introgressive hybridization facilitates adaptive divergence in a recent radiation of monkeyflowers. *Proceedings of the Royal Society B: Biological Sciences*, *282*(1814), 20151666. https://doi.org/10.1098/rspb.2015.1666

Streisfeld, M. A., & Kohn, J. R. (2005). Contrasting patterns of floral and molecular variation across a cline in *Mimulus aurantiacus*. *Evolution, 59*(12), 2548–2559

Streisfeld, M. A., & Kohn, J. R. (2007). Environment and pollinator-mediated selection on parapatric floral races of *Mimulus aurantiacus*. *Journal of Evolutionary Biology*, *20*(1), 122–132. https://doi.org/10.1111/j.1420-9101.2006.01216.x

Streisfeld, M. A., & Rausher, M. D. (2009a). Altered trans-regulatory control of gene expression in multiple anthocyanin genes contributes to adaptive flower color evolution in *Mimulus aurantiacus*. *Molecular Biology and Evolution*, *26*(2), 433–444. https://doi.org/10.1093/molbev/msn268

Streisfeld, M. A., & Rausher, M. D. (2009b). Genetic changes contributing to the parallel evolution of red floral pigmentation among *Ipomoea* species. *New Phytologist*, *183*(3), 751–763. https://doi.org/10.1111/j.1469-8137.2009.02929.x

Streisfeld, M. A., Young, W. N., & Sobel, J. M. (2013). Divergent selection drives genetic differentiation in an R2R3-MYB transcription factor that contributes to incipient speciation in *Mimulus aurantiacus*. *PLoS Genetics*, *9*(3), e1003385. https://doi.org/10.1371/journal.pgen.1003385

Via, S. (2009). Natural selection in action via speciation. *Proceedings of the National Academy of Sciences of the United States of America*. 106 Suppl 1. 9939-46. https://doi.org/10.1073/pnas.0901397106