Molecular Evolution of Anthocyanin Biosynthesis in Morning Glories

by

David Lee Des Marais

Department of Biology
Duke University

Date:_______________________
Approved:

____
Mark D. Rausher, Supervisor

____
Erich Grotewold

____
Francois Lutzoni

____
Gregory A. Wray

____
John H. Willis

Dissertation submitted in partial fulfillment of
the requirements for the degree of Doctor
of Philosophy in the Department of
Biology in the Graduate School
of Duke University

2008

ABSTRACT

Molecular Evolution of Anthocyanin Biosynthesis in Morning Glories

by

David Lee Des Marais

Department of Biology
Duke University

Date:_______________________

Approved:

____
Mark D. Rausher, Supervisor

____
Erich Grotewold

____
Francois Lutzoni

____
Gregory A. Wray

____
John H. Willis

An abstract of a dissertation submitted in partial
fulfillment of the requirements for the degree
of Doctor of Philosophy in the Department of
Biology in the Graduate School
of Duke University

2008

Abstract

Determining the genetic basis of adaptation has become a central focus of evolutionary biology, and the incorporation of increasingly sophisticated analytical tools from molecular biology has made identifying causal genes a practical reality. The work presented herein addresses the effects of pleiotropic constraint on evolutionary change at the level of individual genes and genetic networks. In the first chapter, I combine molecular phylogenetic analyses and direct assays of enzymatic function to determine the evolutionary processes following a gene duplication in the anthocyanin pathway. My results show that, prior to duplication, the DFR gene was constrained from functional improvement by its multiple enzymatic roles. Following duplication, this constraint was released and adaptive evolution proceeded along both paralog lineages. In the second chapter, I determine the molecular genetic basis of a flower color transition in morning glories that is associated with change in pollinator attraction. A regulatory change in a branching enzyme F3'H in the flavonoid biosynthetic pathway reduced flux down the cyanidin-producing branch, conferring nearly exclusive production of red pelargonidin pigment in flowers. I further demonstrate that this regulatory change was restricted to floral tissue, and that ancestral pathway flux predominates in vegetative tissues. I propose that deleterious pleiotropic effects prevented evolutionary change via enzymatic changes in the pathway due to the numerous essential products downstream of this branching point. Together, these two results show that evolutionary change may be constrained by the molecular genetic context in which prospective adaptive mutations occur.

Contents

List of Tables

List of Figures

Acknowledgements

I give my deepest and most heartfelt thanks to my advisor, Mark Rausher, for intellectual, financial and, though he may not have realized, emotional, support and for teaching me clarity of thought. For reminding me to always step back and ask what The Question was or should have been. Mark welcomed a wayward systematist into his lab, seduced me with the gospel of evolutionary genetics and let me explore the field in my own way. I cannot imagine a better mentor for a young scientist.

My greatest failing in grad school has been in the infrequency with which I have utilized my excellent committee. John Willis provided energy and excitement but also carefully reasoned intellectual and career advice when called for, and has always made me feel like my ideas were interesting and relevant. Erich Grotewold patiently responded to countless rambling emails from a molecular biology neophyte and helped keep me grounded with his perspective from outside the evolutionist's bubble. Greg Wray was my first laboratory host at Duke and I credit him and his excellent lab group with beginning my proper education in molecular evolution. Finally, Francois Lutzoni was a mentor and friend before I ever arrived at Duke and played a significant role in my decision to join the student body here. Francois first taught me how to test evolutionary models using phylogenetics and his excitement about my thesis gave me confidence in the usefulness of my work to the broader biological community. I am indebted to each of these fine scientists and mentors.

I have overlapped with an indecent number of biograds in my time here and would thank them all given the time. In particular, Stuart McDaniel, Matt Rockman and Andrea Sweigart were inspirational to me early in my graduate career. Matt Hahn and Becky Zufall encouraged me to join the Rausher Lab and made it a comfortable yet intellectually challenging home once I signed on. Over the years, the other members of the lab have endured my unsolicited opinions, advice and mentoring with grace and taught me much about how to do science and how to conduct oneself in the laboratory and scientific community.

I have also been lucky to count many young biologists as my closest friends. Jason Stajich, Dan Speiser, David Lowry, Arielle Cooley, Amy Bouck, Elizabeth Derryberry, Meleah Hickman, Rich Phillips, Eric Schuettpelz, Tom Chappel and Liz Scordato are all amazing scientists and irreplaceable friends. I hope to share in your discoveries in the world and in science long after our time here.

For many practical and technical considerations, I thank Beverly Calhoun, Lisa Bukovnic, Jim Tunney, Marcy Speer, and Clint King. Anne Lacey, Durham mother to all of us biograd orphans, should be canonized. I'll say no more.

Over the years, I have been funded by the National Science Foundation, the Society for Molecular Biology and Evolution, the Duke University Program in Genetics (and Genomics; someone please explain to me how the new name is not redundant) and the Department of Biology. Chapter One of this thesis could not have happened without this funding.

My family, from 3000 miles away, has always supported my long efforts at over education. My father has gracefully served as both a caring parent and, increasingly over the years, a scientific mentor. My mother has likewise provided a loving and calm ear to my long-distance phone calls. Her commitment to education has inspired me to become a better teacher. My sister has inspired me by living the life I might have had Darwin not bitten me in a college classroom 12 years ago.

Finally, I give my greatest thanks and affection to Robin Hopkins. Robin has been with me through my hardest and most exciting time here at Duke and has weathered any number of fits of disappointment, excitement, depression and relief along the way. Her skill and commitment as a scientist will never cease to amaze me and her generosity and kindness as a friend are unparalleled.

Introduction

The spectacular diversity of the natural world is even more remarkable when we consider the diversity of genes and gene functions that shape living systems. One fascinating result of the genomic revolution has been quantifying the seemingly small number of genes in eukaryotic genomes as compared to the number of gene functions these genomes are responsible for carrying out. The so-called G-value paradox, reflecting the observation that gene number does not necessarily correspond to organismal complexity, can be easily understood in the context of multiple-function genes. However, the multifunctionality of genes creates a dilemma for adaptation: how can natural selection act on some aspects of a gene's function without interfering with other essential functions of that gene? We can begin to address this question by studying how homologous genes from closely related species differ in function, especially when there is a clear, adaptive, phenotypic difference between the species. In this thesis, I exploit past work on flavonoid biosynthesis in the morning glory family, Convolvulaceae, to address two different methods that plants have used to escape pleiotropic constraint: gene duplication and evolution by regulatory mutation.

Chapter 1 describes the evolutionary changes following duplication of the flavonoid biosynthetic gene dihydroflavonol 4-reductase. While many studies have focused on the evolutionary fate and consequence of gene duplicates as a subfield in itself, paralogous genes continue to provide us with a wealth of information on broader mechanisms of molecular evolution. Some of the first estimates of the neutral rate of nucleotide substitution made use of degenerated gene duplicates, or pseudogenes (Li et al. 1981). Likewise, one of the clearest cases of adaptive evolution at the level of the gene is one involving gene duplicates (Zhang et al. 2002). My work on the evolution of DFR in the Convolvulaceae indirectly demonstrates intragenic pleiotropy, or an adaptive conflict within in a gene, and directly shows the effects of adaptive substitution on enzyme function. Of particular interest to the broader field of molecular evolution is my result showing a strong improvement in enzymatic function in the absence of a conclusive sequence-based test of adaptive change. While such phylogenetic tests for changing selection pressures have long been understood to be conservative, my results demonstrate this analytical problem directly. More generally, the work presented in chapter 1 shows that organisms can more broadly exploit a multi-functional enzyme via the simple mechanism of gene duplication.

Chapter 2 addresses the genetic basis of a conspicuous phenotypic change associated with an adaptation for hummingbird pollination. An outstanding question in molecular evolution concerns the relative role of regulatory versus structural gene changes during adaptation. Due to the many compounds produced in the same pathway as the floral pigmenting anthocyanins, some level of evolutionary constraint may be expected to act during bouts of adaptation in this pathway. For example, enzymes that are essential for producing the blue floral pigment cyanidin are also essential for producing many compounds involved in plant defense. So, while natural selection may favor mutations conferring changed anthocyanin biosynthesis in flower petals, these same mutations might have deleterious consequences in stem or leaf tissue. In chapter 2, I show that a red-flowered clade of morning glories lost expression of the key branching enzyme flavonol 3’ hydroxylase in the floral tissue due to a cis-regulatory mutation but retains F3’H expression in other tissues. These plants were thereby able to evolve a novel phenotype associated with pollinator attraction without reducing their ability to defend against ecological stresses. In this case, the multifunctionality of the gene had to do with spatial expression; by changing only the module of expression conferring floral function the plant could undergo adaptive evolution without pleiotropic costs.

1. Escape from Adaptive Conflict After Duplication in an Anthocyanin Pathway Gene[1]

1.1 Introduction

Gene duplications have been recognized as an important source for evolutionary innovation and adaptation since at least Haldane (1933) and their varying fates may partly explain the vast disparity in observed genome sizes (Hahn et al. 2005). The expected fates of most gene duplications involve primarily non-adaptive substitutions leading to either non-functionalization of one duplicate copy or subfunctionalization (Lynch and Conery 2000), neither of which yields novel function. A significant evolutionary problem is thus elucidating the mechanisms of adaptive evolutionary change leading to evolutionary novelty. Presently, the most widely recognized adaptive process involving gene duplication is neofunctionalization, in which one copy undergoes directional selection to perform a novel function following duplication (Ohno 1970). An alternative, but understudied, adaptive fate that has been proposed is escape from adaptive conflict (EAC), in which a single-copy gene is selected to perform a novel function while maintaining its ancestral function (Hughes 1994; Piatigorsky and Wistow 1990). This gene is constrained from improving either novel or ancestral function because of detrimental pleiotropic effects on the other function. Following duplication, one copy is free to improve novel function, while the other is selected to improve

[1] This chapter appeared previously as Des Marais, D. L. and M. D. Rausher. 2008. Escape from adaptive conflict after duplication in an anthocyanin pathway gene. Nature 454: 762-765.

ancestral function. I first present two criteria that can be used to distinguish neofunctionalization from EAC. Using both tests for positive selection and assays of enzyme function, I then demonstrate that adaptive evolutionary change in a duplicated gene of the anthocyanin biosynthetic pathway in morning glories (*Ipomoea*) is best interpreted as EAC. Finally, I argue that this phenomenon likely occurs more often than has been previously believed and may thus represent an important mechanism for the generation of evolutionary novelty.

Novel gene function associated with gene duplication can arise in either of two ways. According to the neofunctionalization model, novel function arises after gene duplication, with one copy maintaining ancestral function, while the second copy is selected to perform a new function. By contrast, under EAC, novel function arises first in the single-copy ancestral gene, which results in a reduction in that copy's ability to perform the original function. Following duplication, each copy is free to specialize on either the original or the novel function and improve those functions. While both of these processes have been recognized for at least two decades as alternative possibilities (Ohno 1970; Piatigorsky and Wistow 1990), I am aware of only one previous attempt to distinguish between them (Hittinger and Carroll 2007).

Two criteria may be used to distinguish between neofunctionalization and EAC. First, under EAC adaptive change occurs in both duplicate copies, whereas under neofunctionalization only one copy undergoes adaptive change because purifying selection acts to maintain ancestral function in the other copy. Second, under EAC ancestral function is improved, whereas under neofunctionalization it is not. Here I apply both of these criteria to distinguish between neofunctionalization and EAC as explanations for adaptive change in duplicated copies of the anthocyanin biosynthetic pathway gene dihydroflavonol-4-reductase (DFR).

Plant dihydroflavonol-4-reductases (EC number 1.1.1.219) function most conspicuously in the reduction of several flavonoid precursors of anthocyanin pigments and their related phytoalexins, though the full scope of DFR function is unknown. The widely distributed 3-hydroxyanthocyanidins pelargonidin, cyanidin and delphinidin are primarily responsible for red, purple and blue flowers in angiosperms and are the downstream products of DFR activity on dihydrokaempferol (DHK), dihydroquercetin (DHQ) and dihydromyricetin (DHM), respectively (Schwartz-Sommer et al. 1987). DFR from several taxa is able to reduce naringenin and eriodyctiol to produce the flavan-4-ols which are themselves precursors of phlobaphenes and of two rare anthocyanidins, apigeninidin and luteolininidin (Fischer et al. 2003; Halbwirth et al. 2003). In the common morning glory, *Ipomoea purpurea,* DFR is present as a small, tandemly arrayed 3-gene family spanning ca. 17kb (Inagaki et al. 1999). All three copies are expressed in *I. purpurea* (although *DFR-A* and *DFR-C* are expressed in overlapping but fewer tissues relative to *DFR-B*) and have the conserved intron/exon structure observed across eudicots.

1.2 Results and Discussion

Phylogenetic analysis of *DFR* in the *Ipomoea* gene family and closely related single-copy taxa shows that each *DFR* copy identified in *I. purpurea*, A, B and C, forms a clade (Figure 1). Each clade of orthologs, the entire *DFR* gene family, *DFR* from all Convolvulaceae and *DFR* from all Solanales constitute well-supported clades. Sister to the three-copy species clade is a large clade of morning glories characterized by a single copy of *DFR* and represented here by *Evolvulus glomeratus*; most Solanaceae species are also single-copy and my results suggest that there was a lineage specific duplication in *Petunia*. The gene tree also reveals two separate *DFR* duplication events in the Convolvulaceae: the first gave rise to *DFR-B* and another lineage which experienced a second duplication event creating the *DFR-A* and *DFR-C* copies.

Codon-based models of sequence evolution (Yang 1998; Zhang et al. 2005) indicate that single-copy *DFR*s have historically experienced purifying selection (dN/dS=0.13; Fig. 1 and Tables 1-7), as did the base of the clade containing the Convolvulaceae *DFR* gene family (dN/dS =0.104). Following the first gene duplication, the lineage subtending *DFR-B* has a ratio of replacement to synonymous substitutions statistically indistinguishable from the single copy *DFR*s (dN/dS =0.183). By contrast, this ratio has increased on the lineage subtending *DFR-A/C* clade (dN/dS = infinite). A likelihood model infers that 18 replacement and no synonymous substitutions occurred at the base of the A/C clade. The probability of this pattern occurring under neutrality is P=0.004, indicating the action of repeated positive selection along this branch after the first duplication. Non-synonymous substitution was also significantly elevated above background on the branch subtending the *DFR-A* clade (dN/dS = 0.842, p=0.009) and marginally so on the branch subtending the DFR-C clade (dN/dS = 0.29, p=0.06), although a branch-sites test fails to detect positive selection on either of these (p=0.21 and p>0.9, respectively). Changes immediately following the second duplication thus hint at positive selection, but do not definitively reveal it.

Subsequent non-synonymous substitution on copies in the B and the A/C clades varies slightly from background rates but still bears the signature of purifying selection. A clade-sites test for the *DFR-A/C* clade as a whole as compared to *DFR-B* as a whole reveals that the majority of codon sites in both clades remain under strong purifying selection (Figure 1). It thus appears that there was a burst of adaptation in the A and C copies very soon after the first, and possibly after the second, duplication. Following this burst, selection once again became primarily purifying.

I chose *Solanum lycopersicon* and *Evolvulus glomeratus* as representative "pre-duplication" species and *Ipomoea purpurea* and *Convolvulus arvensis* as representative "post-duplication" species to determine whether enzyme function changed following duplication. Protein products of each *DFR* copy from these four taxa were assayed for enzymatic activity on three common DFR substrates (dihydrokaempferol, dihydroquercetin and dihydromyricetin) and two substrates less commonly reduced by plant DFR enzymes (naringenin and eriodyctiol).

Three major patterns arise from this analysis. First, the pre-duplication copies exhibit moderate activity on DHK, DHQ and DHM (Fig. 2C, 2D, 2E), very low activity on eriodyctiol (Fig. 2B), and minimal activity on naringenin (Fig. 2A). Second, the post-duplication A and C copies exhibit essentially no activity on any of the five substrates (Fig. 2). Moreover, DFR-A and C activity on eriodyctiol, DHK, DHQ and DHM are significantly reduced compared to activity in the pre-duplication copies (Tables 9-11). Third, activity on all five substrates is higher for the post-duplication DFR-B copies than for the pre-duplication copies, all but DHQ significantly so even after a Bonferroni correction for multiple comparisons. The increased DFR-B activity on the molecularly smaller naringenin (4.6-fold), eriodyctiol (6.1-fold) and DHK (7.9-fold) are substantially greater than the increases on DHQ (1.4-fold) and DHM (1.4-fold).

These results indicate that adaptive change in DFR immediately subsequent to the first duplication is best explained as escape from adaptive conflict. Adaptive change apparently occurred in both copies after the first duplication, consistent with expectations under EAC. Adaptive change in the A/C copy is indicated by the signature of repeated positive selection in the lineage subtending the A/C clade. Although such a signature of positive selection was not detected on the B copy, enzyme functional analysis indicates that adaptive substitution likely occurred. In particular, following duplication, enzyme activity on ancestral flavonoid substrates increased substantially. Because most mutations are believed to have detrimental effects on enzyme stability or activity (DePristo et al. 2005), I believe that neutral evolution in the B-lineage can be rejected in favor of positive selection that is undetected at the sequence level, possibly because it involved few substitutions.

Further evidence for adaptive change along the B-lineage is provided by an analysis of the substitutions that occurred at its base. The crystal structure of DFR from *Vitis vinifera* shows that Asn133 lines the substrate binding pocket and forms a hydrogen bond with the 4’ hydroxyl (and, if present, the 3’ hydroxyl) of the substrate B-ring (Petit et al. 2007). DFR from maize (Halbwirth et al. 2003) and *Pyrus* (Fischer et al. 2003) has activity on naringenin, eriodyctiol, DHK and DHQ (DHM activity was not assayed in either of these studies). Both of these DFR variants have asparagine in the 133 position; in fact, nearly all known DFR genes from angiosperms have Asn133. My phylogenetic reconstruction reveals that an Asn133Asp substitution occurred prior to the divergence of the Gentianales and Solanales (Fig. 3); this aspartate is retained in my two assayed single-copy species and correlates with relatively poorer flavonoid function. DFR-A copies from different species have different non-Asn amino acids at this site, and all DFR-C copies have the hydrophobic isoleucine; substituting the hydrophobic leucine at residue 133 has been shown to change substrate specificity in *Gerbera* DFR (Johnson et al. 2001). Following gene duplication, the Asn133 variant re-evolved in the DFR-B lineage, coinciding with the improvement of DFR-B activity on the five substrates examined. Given this pattern, it is tempting to suggest that the original substitution of Asp for Asn at the base of the Gentianales and Solanales reflects the acquisition by the single-copy DFR of the unknown second function, and that the reverse transition in the DFR-B lineage reflects loss of this function in this lineage. Moreover, the failure to regain Asn in DFR-A and DFR-C is consistent with the loss of the ability of these copies to metabolize dihydroflavonols.

The improvement of ancestral function in the DFR-B lineage satisfies the second criterion for the operation of EAC. Apparently, the ability of DFR to reduce dihydroflavonols was constrained in the ancestral single-copy gene and this constraint was released when the first duplication occurred.

Ideally, if EAC explains evolutionary change in duplicated copies of DFR, one should be able to demonstrate that the different copies perform different functions and that both of these functions were performed by the ancestral single-copy DFR, albeit at reduced efficiency. Unfortunately, I am unable to confirm this expectation because I have not yet identified what function(s) the DFR A and C copies perform. Although it is clear that these copies have lost the ability to perform one ancestral function of DFR (reduction of dihydroflavonols), I do not have direct evidence that they perform an ancestral rather than a novel function. Nevertheless, the improvement of flavonoid function in the DFR-B copies points to this function being constrained in the ancestral single-copy gene. Escape from adaptive conflict is thus clearly a feature of DFR evolution.

A potential alternate explanation for my results is that the activity spectrum of DFR-B reflects the ancestral state of DFR, and that the loss of ability to metabolize naringenin and eriodyctiol, as well as reduced function on DHK, DHQ and DHM, in *E. glomeratus* and *S. lycopersicon* reflect independent losses of activity. However, an ancestral state reconstruction of substrate specificity does not support this interpretation: an analysis using all species for which activity on all five flavonoid substrates were tested indicates that while ability to utilize all five substrates is deeply ancestral in the angiosperms, the ability to use naringenin and eriodyctiol was significantly reduced on the lineage leading to *Solanum*, *Evolvulus*, and *Ipomoea* (Fig. 4). Loss of function on these two substrates occurred prior to the *DFR* duplication in *Ipomoea*. This pattern is consistent with the EAC hypothesis: the reconstruction demonstrates first a reduction of ability to metabolize dihydroflavonols, naringenin and eriodyctiol, corresponding to when the single-copy DFR gained an additional (currently unknown) function; and then an increase in the ability to metabolize these compounds after duplication, corresponding to release from adaptive constraint.

A possibility not explored in the current study is that DFR-A and DFR-C have not lost flavonoid activity, rather that other changes occurred in these enzymes such that my assay conditions did not accurately sample their optimal catalytic activity. Changes in their protein sequences may have led to poor protein folding in the *E. coli* cell line, or perhaps led to a change in optimal pH conditions. While such changes may not constitute a change in "function," these types of changes would suggest that some significant evolutionary changes occurred following duplication. And the absence of definitive change in function stands apart from the clear inference of a strong signature of positive selection early in the history of DFR-A and C.

My analysis illustrates the larger issue that many published examples of neofunctionalization may actually represent EAC. There are some cases of neofunctionalization that by the two criteria used in my analysis are better explained by neofunctionalization than by EAC (e.g. Benderoth et al. 2006; Zhang et al. 2002). However, there are also many purported cases of neofunctionalization that have not characterized ancestral function prior to duplication, that have not determined whether adaptive evolution occurred in one or in both duplicate copies subsequent to duplication, and that have not determined whether either copy exhibits improvement of known ancestral function (e.g. the origin of vertebrate MHC, Spagnuolo et al. 1997, and plant R-gene alleles, Michelmore and Meyers 1998). In these cases, EAC cannot be ruled out. Consequently, it remains unclear whether the absence of reported cases of EAC indicates a true rarity of its occurrence, or just that much of what is considered neofunctionalization is actually EAC.

Unfortunately, the theoretical literature is silent on the expected relative frequencies of these two processes. Although some models predict that the relative importance of neofunctionalization and subfunctionalization depends on factors such as population size (Lynch and Force 2000; Lynch et al. 2001), these results cannot necessarily be extrapolated to either a comparison of neofunctionalization and EAC or of EAC and subfunctionalization. While EAC resembles subfunctionalization in that it results in the partitioning of ancestral functions, subfunctionalization is envisioned as involving only neutral or slightly deleterious mutations (Force et al. 1999), while EAC involves primarily adaptive substitutions. Similarly, while neofunctionalization involves adaptive mutations in one copy of a duplicate, EAC involves adaptive mutations in both copies. Neither of these features of EAC is incorporated into any existing models of the fate of duplicate genes. Intuitively, we might expect that the probability of preservation of duplicate copies is higher for EAC than for neofunctionalization, given that the duplicated gene has already acquired multiple functions, because mutations in either copy can start the EAC process. However, the overall occurrence of EAC vs. neofunctionalization will also depend upon the fraction of duplicated genes that have acquired new function before duplication, a quantity for which no information exists.

That the *DFR-B* copy also evolved the capability of reducing naringenin and eriodyctiol is surprising because the presumed downstream products of these reactions are rare in angiosperms and unreported from the Solanales (Iwashina 2000). It is therefore doubtful that this new capacity is actually utilized in post-duplication *Ipomoea* species. Nevertheless, 3-deoxyanthocyanins and their polymers have been reported in some crops to have insecticidal and anti-fungal properties (Nicholson et al. 1987) which suggests that the ability of DFR-B to synthesize their precursors may constitute an exaptation for the evolution of such resistance. Under this interpretation, improvement of DFR-B activity on the more common substrates DHK, DHQ and DHM may have been the adaptive change while improvement in flavonol reduction was a byproduct. Interestingly, this result indicates that the evolution of novel function may sometimes be a byproduct of the evolution in enhanced ancestral activity.

1.3 Methods

1.3.1 Plant Materials and Extractions

I used the phylogeny of Stefanovic et al. (2002) to target my sampling to a set of species that would help determine the timing of evolutionary transitions from a single-copy *DFR* ancestor to the three-copy family observed in *I. purpurea* and *I. nil* (Inagaki et al. 1999). I focused on 6 species representing the phylogenetic diversity of the Convolvulaceae: *Ipomoea purpurea*, *Convolvulus arvensis*, *Evolvulus glomeratus*, *Merremia dissecta*, and *Merremia umbellata* and *Solanum lycopersicon* from the sister family Solanaceae. All plants were grown from seed. DNA and RNA were extracted from freshly harvested buds, leaves and stem tissue using the DNeasy Plant Mini kit and RNeasy kit (Qiagen, Germantown, MD), respectively, according to the manufacturer's protocol.

1.3.2 Cloning and Sequencing of DFR copies

To isolate *DFR* copies from previously unstudied taxa, I designed primers based on a consensus of known DFR copies from the Convolvulaceae. Initially, forward primer DFR59F: 5'-GCGTCACCGGAGCTGCTGG-3' was paired with DFR405R: 5'-GAGGAAGTGAAAACCAGC-3' or the GeneRacer oligo dT in 3'RACE reactions using either genomic DNA or cDNA synthesized from total RNA as template. Gel-extracted PCR products were cloned into TOPO-TA 2.1 vectors (Invitrogen, Carlsbad, CA). Individual clones were sequenced using BigDye version 3.0 (Applied Biosystems, Foster

City, CA) and visualized on ABI 3730 sequencers. To amplify the variable 5' ends of the *DFR* copies, I designed nested gene-specific primers and carried out 5' RACE reactions using the 5' RACE System (Invitrogen) according to the manufacturer's specifications. To confirm all copies from a given taxon had been identified, I also used multiple pairs of degenerate primers in PCRs using independently synthesized cDNA pools as template (primer sequences available on request). Gene- specific primers were then used to amplify the full-length cDNAs of each paralog, which were cloned into TA vectors and transformed into Top10 cells (Invitrogen) for storage at -80°C.

1.3.3 Phylogenetic Analysis of Gene Tree Topology, Changes in Selection and Timing of Duplication

I combined new sequences collected in the present study with *Ipomoea* sequences available in GenBank and several outgroups to make a single alignment. All new sequences reported herein have been deposited in GenBank (accession numbers EU189072-189082). Because the N- and C-terminal regions were highly variable, I was unable to discern positional homology with reasonable certainty. These regions (consensus nucleotide positions 13-66 and 1054-1320) were excluded from all subsequent analyses. Alignments are available on request.

The optimal tree and estimates of branch-support were determined using a Bayesian MCMC approach in the program MrBayes 3.1.2 (Ronquist and Huelsenbeck 2003). In Bayesian analyses, I used two models: in the first model, all sites were included in one partition assumed to evolve according to the GTR+G model. In the second model, one partition contained all 1^{st} and 2^{nd} codon positions and a second partition contained all 3^{rd} positions, treating each partition independently with its own GTR+G parameters. MCMC analyses were started with flat priors and run for 2 million generations, a burn-in was discarded and consensus trees were determined using MrBayes. At the conclusion of each MCMC run, I estimated that the chains had converged upon a stable set of parameters by calculating the potential scale reduction factor using MrBayes.

In addition to Bayesian phylogenetic analyses, maximum parsimony tree searches were conducted using PAUP 4.0b10 (Swofford 2000) with equally weighted substitutions. Tree searches using the Maximum Likelihood criterion were implemented in GARLI (Zwickl 2006) with the GTR+G model, as selected using a top-down hierarchical likelihood ratio test. These trees had the identical in-group topology as the MCMC-estimated consensus and are therefore not presented here.

I determined if changes in selection followed the origin of duplicate copies of DFR using the codeml program in the PAML 3.15 package (Yang 1997) to perform lineage and clade-specific analyses of Ka/Ks ratios (ω). I first tested a series of nested models that allowed particular branches to have ω independent of background ratios: the lineage subtending all duplicated *DFR* copies, the lineage subtending the *DFR-B* clade, the lineage subtending the *DFR-A/C* clade, the lineage subtending the *DFR-A* clade, and the lineage subtending the *DFR-C* clade. I also tested each of these branches for a signature of positive selection using the branch-sites test of Zhang et al. (2005), and each of the clades arising from these branches using the clade-sites test of Yang et al. (2005). I applied a sequential Bonferroni correction (Holm 1979) to account for the multiple comparisons made in these analyses and report only results that were significant following this correction.

The lineage-based tests for change in selection pressure described above revealed an absence of synonymous substitutions among the 277 synonymous sites on the lineage subtending the *DFR-A/C* clade. codeml infers 18 nonsynonymous changes among 779 nonsynonymous sites on that branch. Given the ratio of nonsynonymous to synonymous sites across the whole gene (779/277) the probability of observing 18 nonsynonymous and zero synonymous substitutions given that there were 18 total substitutions is the binomial probability $(779/1056)^{18}$ which equals 0.004.

I also fit a series of models that allow particular codon positions to switch between selection classes (negative, positive, neutral) on any branch on the tree (Guindon et al. 2004). These tests build on the sites models of Yang et al. and do not require branches of interest to be specified *a priori*. Specifically, I implemented an extension of the M2a model of Yang et al. (2005) where $\omega_1 < 1$, $\omega_2 = 1$ and ω_3 is freely estimated from the data. Using the terminology of Guindon et al. (2004), I fit three models, M2a, M2a + S1 (which allows for equal switching probabilities between negative, neutral and positive selection) and M2a + S2 (which allows the class switching to be estimated freely). These models are nested and can be compared using a likelihood ratio test with one degree of freedom between each (M2a < M2a +S1 < M2a + S2). Ts/Tv values were estimated from the data using the F3x4 codon model of Goldman and Yang (1994) as implemented in FitModel version 0.5. Results are presented as Table 9 and show codon positions do switch between selection regimes, but not in a biased fashion. Further, no sites were inferred to be under repeated positive selection, which is unsurprising given my hypothesis that directional selection was restricted to a very small subset of lineages (i.e. those following duplication events) which I tested directly using Yang's branch-sites models.

1.3.4 *In vitro* Assays of Enzyme Function

Full-length cDNA transcripts of each *DFR* copy from *I. purpurea*, *C. arvensis*, *E. glomeratus* and *S. lycopersicon* were amplified with Invitrogen Pfx polymerase and cloned directly into pENTR/D-TOPO vectors and then recombined into pDEST17 vectors using Clonase 2.1 (Invitrogen) according to manufacturer specifications. Plasmids containing inserts in the correct orientation and coding frame were transformed into BL21 Star (DE3) *E. coli* for overexpression. After overnight growth in 5mL LB-Amp_{100}, I inoculated 200 mL of each culture and grew the cells at 37°C and 200rpm until mid-log phase (OD600 = 0.4). To induce expression, IPTG was added to a final concentration of 0.5mM and the cultures were grown for 3 hours at 37°C. Cultures were then spun down and the pellet resuspended in Lysis Buffer (50mM HEPES, 50mM NaCl, 1mM EDTA, 1mM PMSF, 100uM TPCK, 100uM TLCK), lysed for 30 minutes on ice with 1 mg/mL lysozyme and pelleted again leaving the protein-containing supernatant.

Substrates and standards for the enzyme assays were obtained from commercial manufacturers: dihydrokaempferol (TransMit; Marburg, Germany), dihydroquercetin

(Alexis Biochemical; Lausen, Switzerland), dihydromyricetin (Apin Chemicals; Abingdon, UK), eriodyctiol (Indofine; Belle Mead, NJ), naringenin (Sigma), pelargonidin (Indofine; Hillsborough, NJ), delphinidin (Polyphenols Laboratories AS; Sandnes, Norway), cyanidin (Indofine), apigeninidin (Apin) and luteolinidin (Apin). 250 µl of enzyme extract was added to reaction mixtures for *in vitro* functional enzyme assays as described previously (Stafford and Lester 1982). For each gene construct, two replicates each of 30 minute and 3 hour time points were taken. The immediate product of these assays are unstable leucoanthocyanidins (DHK, DHQ and DHM as substrates) or flavan-4-ols (naringenin and eriodyctiol as substrates). The former were converted to stable anthocyanidins as described (Stafford and Lester 1982) and the latter as described (Stafford 1965). Products and their appropriate standards were identified spectrophotometrically and by TLC and then quantified spectrophotometrically on a Shimadzu UV-2401PC. Two negative control assays were also run: one assayed overexpressed *I. purpurea* chalcone synthase enzyme and the other substituted lysis buffer for enzyme.

I used a two-way ANOVA to partition variance of enzyme activity by species, time point and their interaction. To determine whether the pre-duplication copies differed from the post-duplication copies, I conducted a standard contrast between the means of the pre-duplication copies in *Evolvulus glomeratus* DFR and *Solanum lycopersicon* DFR and the post-duplication copies from *Ipomoea purpurea* and *Convolvulus arvensis.* All analyses were performed in SAS. Significance was assessed at P=0.05 after employing a sequential Bonferroni correction (Holm 1979). Only values significant after this correction are marked with an asterisk in Table 9.

1.3.5 Reconstructing Ancestral States of Enzyme Specificity

I collected information from the literature reporting DFR enzyme activity on naringenin, eriodyctiol, DHK, DHQ and DHM and scored these as binary characters (presence/absence of activity; Halbwirth et al. 2003, Petit et al. 2007, Fischer et al. 2003, Martens et al. 2002, Gerats et al. 1982, Punyasiri et al. 2004). To address whether the single-copy DFR enzyme that was ancestral to the gene family under study here possessed activity on all five substrates, I reconstructed the most parsimonious (Maddison and Maddison 2000) and most likely (Pagel 1999) ancestral states in Mesquite version 1.1 (Maddison and Maddison 2007). In the likelihood model, I used branchlengths determined as in the ML topology search, above, and modeled character evolution assuming equal probability of gains and losses. These analyses demonstrate that although activity on naringenin and eriodyctiol is anciently ancestral in the angiosperms, lack of such activity is three times more likely than its presence in the lineage immediately ancestral to the focal taxa examined in this study (Fig. 4).

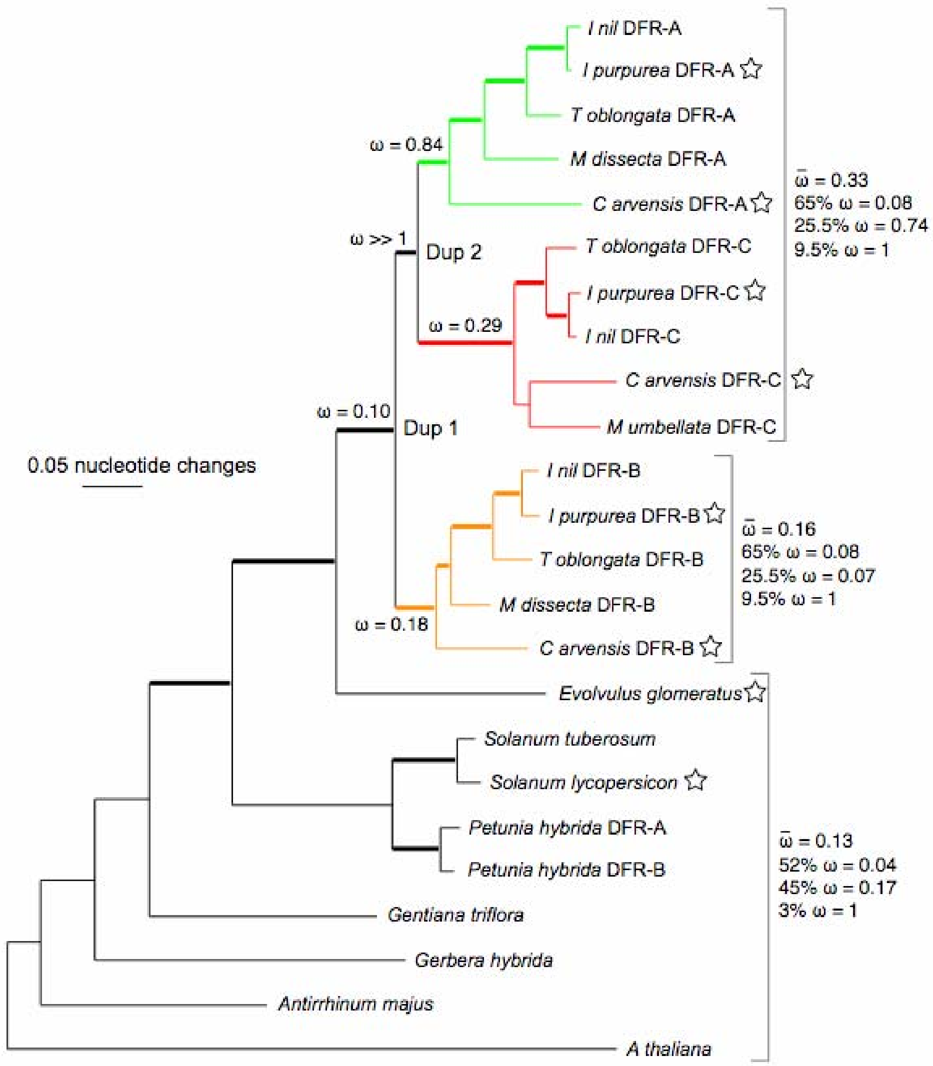

Figure 1: Gene tree of sampled DFR copies from Convolvulaceae and outgroup taxa. Thickened lines indicate Bayesian posterior probabilities greater than 0.95. Starred genes are those included in enzymatic studies. Results of selection tests are reported for tested branches, as are those for clades.

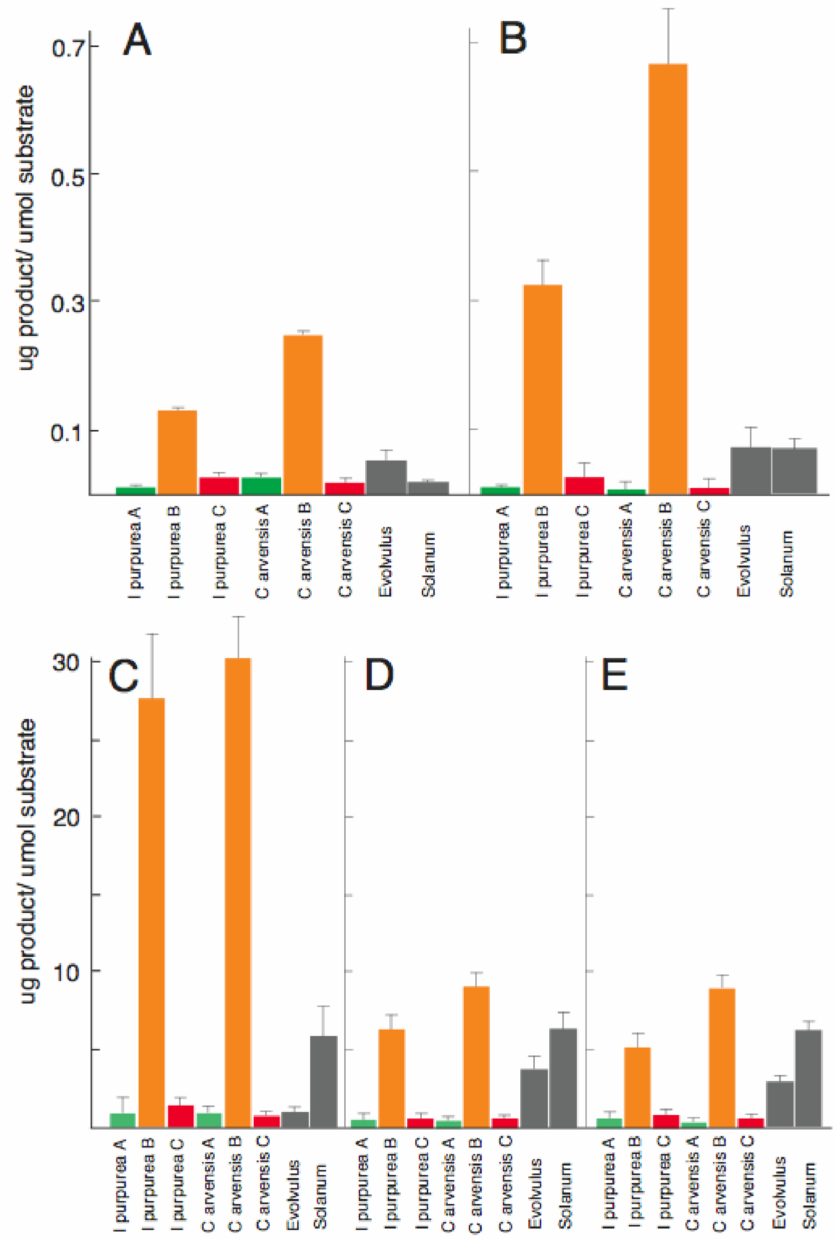

Figure 2: Results of enzyme activity assays. Substrates are Naringenin (A), Eriodyctiol (B), Dihydrokaempferol (C), Dihydroquercetin (D), and Dihydromyricetin (E). Colors reflect colored ortholog clades as in Figure 1-1. Bars are one standard error.

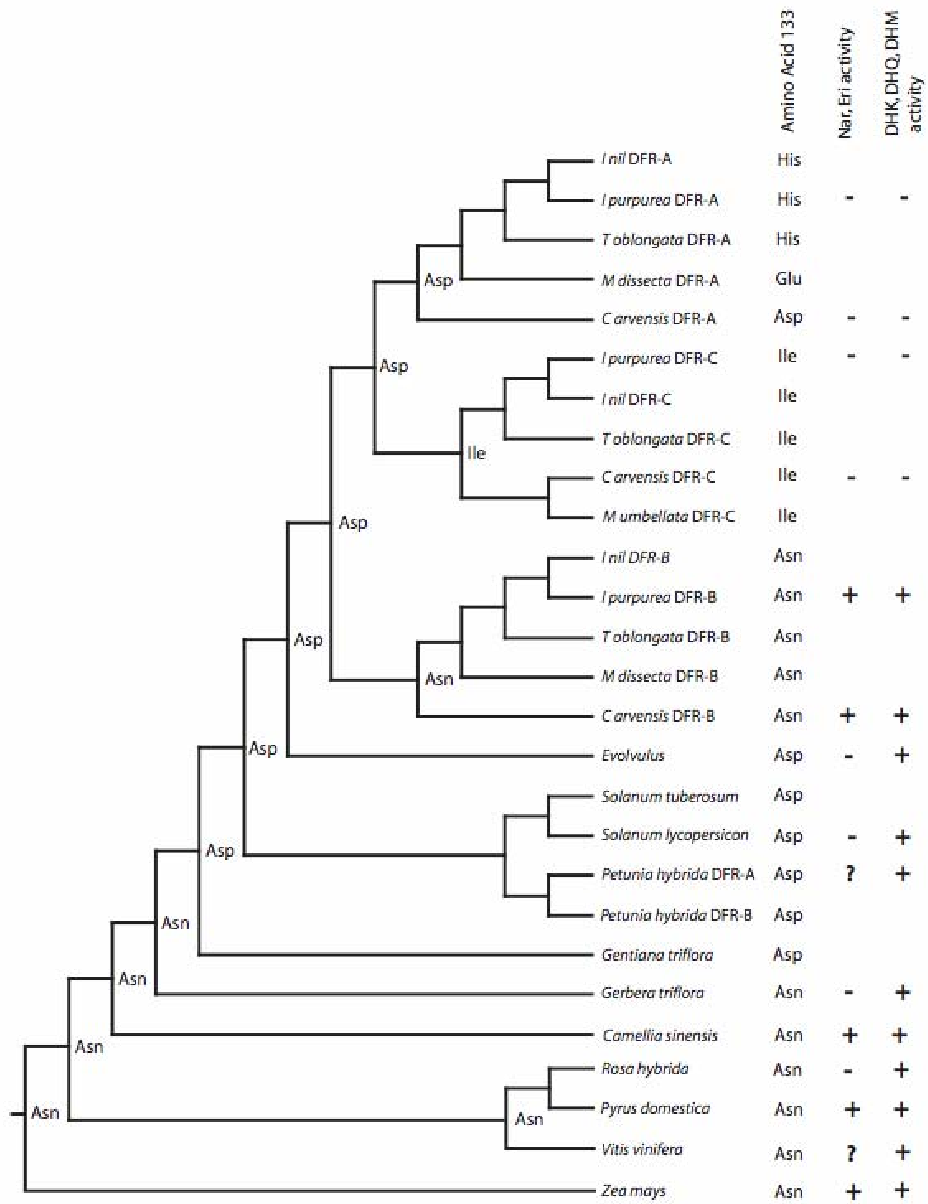

Figure 3: Gene tree with DFR function from across the angiosperms.

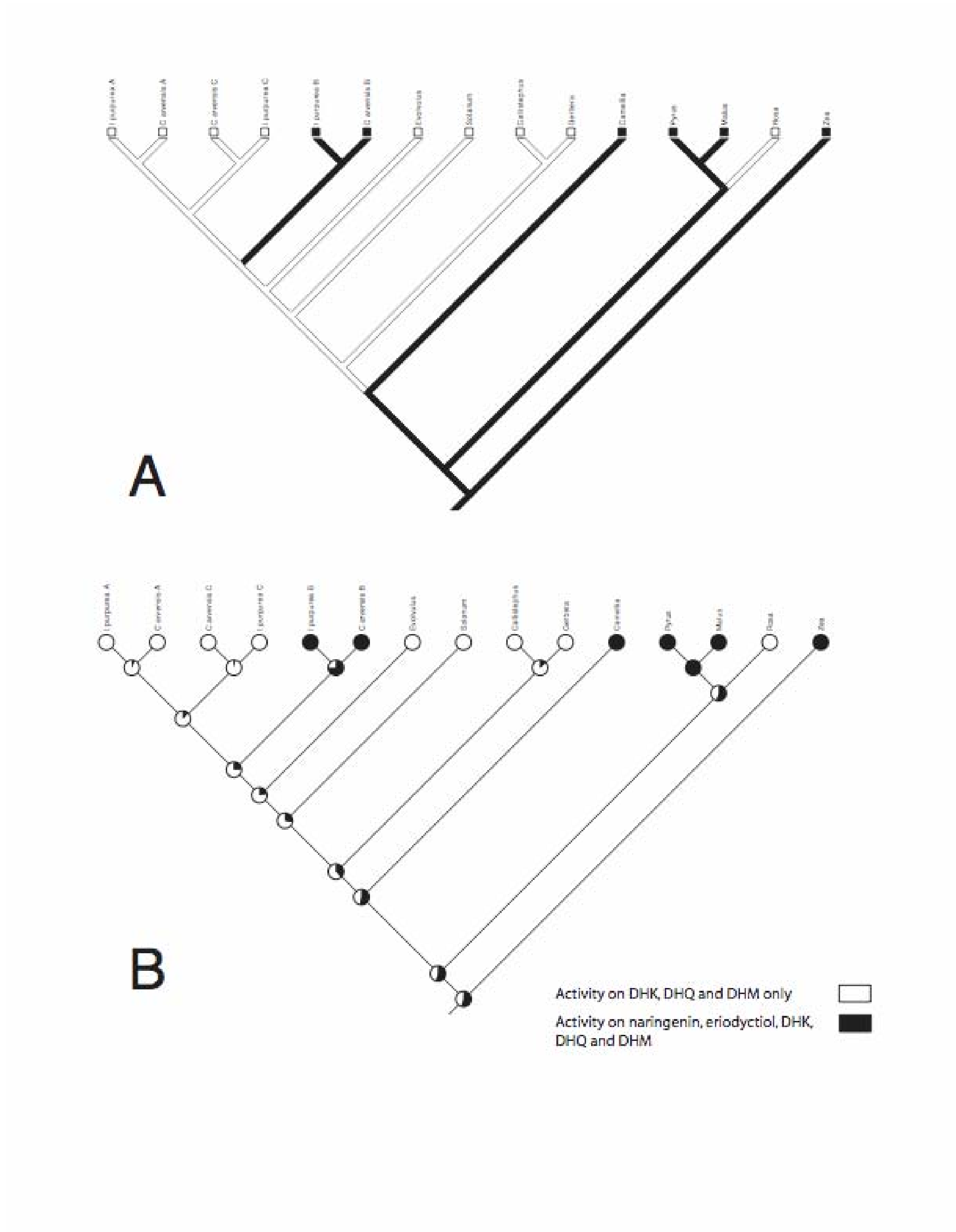

Figure 4: Ancestral state reconstruction of DFR enzyme activity. (A) Maximum parsimony reconstruction. (B) Maximum likelihood reconstruction. For ML analyses, area of pies indicate proportional support for different ancestral states.

Table 1: Lineage specific models of selection

Model	- lnL	Significant	Back-ground	Before Dup	Base AC	Base B	Base A	Base C
L0	9859.5913	-	ω = 0.1778	-	-	-	-	-
L1	9859.1018	No	ω = 0.1798	ω = 0.1110	-	-	-	-
L2	9851.0763	Yes (vs L0)	ω = 0.1753	ω = 0.1041	ω = inf.	ω = 0.1833	-	-
L3	9847.7389	Yes (vs L2)	ω = 0.1720	ω = 0.1043	ω = inf	ω = 0.1759	ω = 0.8426	-
L4	9849.3473	No (vs L2)	ω = 0.1717	ω = 0.1072	ω = inf	ω = 0.1785	-	ω = 0.2968

Table 2: Branch-sites models for A/C lineage

Model	-lnL	Significant	Class 0	Class 1	Class 2a	Class 2b
Null 1 (M1a)	8716.9331	-	$p = 0.886$ $\omega = 0.1257$	$p = 0.114$ $\omega = 1$	-	-
Null 2 (neutral)	8710.1584	Yes (vs Null 1)	$p = 0$ $\omega_{background} = 0.1288$ $\omega_{AC} = 0.1222$	$p = 0$ $\omega_{background} = 1$ $\omega_{AC} = 1$	$p = 0.886$ $\omega_{background} = 0.1222$ $\omega_{AC} = 1$	$p = 0.114$ $\omega_{background} = 1$ $\omega_{AC} = 1$
Positive Selection	8709.1045	No (vs Null 2)	$p = 0$ $\omega_{background} = 0.1220$ $\omega_{AC} = 0.1294$	$p = 0$ $\omega_{background} = 1$ $\omega_{AC} = 1$	$p = 0.886$ $\omega_{background} = 0.1222$ ω_{AC} = inf.	$p = 0.114$ $\omega_{background} = 1$ ω_{AC} = inf.

Table 3: Clade-sites models comparing single-copy to DFR family

Model	-lnL	Significant at 0.05?	Class 0	Class 1	Class 2
Null (M1a)	8716.9331	-	p = 0.886 ω = 0.1257	p = 0.114 ω = 1	
Model C	8634.2404	Yes	p = 0.517 $\omega_{background}$ = 0.036 $\omega_{post\text{-}dup}$ = 0.036	p = 0.030 $\omega_{background}$ = 1 $\omega_{post\text{-}dup}$ = 1	p = 0.453 $\omega_{background}$ = 0.176 $\omega_{post\text{-}dup}$ = 0.449

Table 4: Clade-sites models comparing DFR-B to DFR-A/C

Model	-lnL	Significant at 0.05?	Class 0	Class 1	Class 2
Null (M1a)	4872.1013	-	p = 0.803 ω = 0.134	p = 0.197 ω = 1	
Model C	4854.9880	Yes	p = 0.647 ω_B = 0.078 $\omega_{A/C}$ = 0.078	p = 0.095 ω_B = 1 $\omega_{A/C}$ = 1	p = 0.258 ω_B = 0.067 $\omega_{A/C}$ = 0.736

Table 5: Clade-sites models comparing DFR-A to DFR-C

Model	-lnL	Significant at 0.05?	Class 0	Class 1	Class 2
Null (M1a)	3736.3838	-	*p* = 0.743 ω = 0.102	*p* = 0.257 ω = 1	
Model C	3734.4619	No	*p* = 0.746 ω_A = 0.102 ω_C = 0.102	*p* = 0.208 ω_A = 1 ω_C = 1	*p* = 0.046 ω_A = 0.173 ω_C = 1.810

Table 6: Branch-sites models for base of the DFR-A clade

Model	-lnL	Significant	Class 0	Class 1	Class 2a	Class 2b
Null 1 (M1a)	8716.9331	-	*p* = 0.886 ω = 0.1257	*p* = 0.114 ω = 1	-	-
Null 2 (neutral)	8715.9470	No (vs Null 1)	*p* = 0.663 $\omega_{background}$ = 0.1245 ω_A = 0.1245	*p* = 0.084 $\omega_{background}$ = 1 ω_A = 1	*p* = 0.224 $\omega_{background}$ = 0.1245 ω_A = 1	*p* = 0.029 $\omega_{background}$ = 1 ω_A = 1
Positive Selection	8715.1600	No (vs Null 2)	*p* = 0.8486 $\omega_{background}$ = 0.1245 ω_A = 0.1245	*p* = 0.106 $\omega_{background}$ = 1 ω_A = 1	*p* = 0.040 $\omega_{background}$ = 0.1245 ω_A = 19.47	*p* = 0.005 $\omega_{background}$ = 1 ω_A = 19.47

Table 7: Branch-sites models for base of DFR-C clade

Model	-lnL	Significant	Class 0	Class 1	Class 2a	Class 2b
Null 1 (M1a)	8716.9331	-	p = 0.886 ω = 0.1257	p = 0.114 ω = 1	-	-
Null 2 (neutral)	8714.4092	Yes (vs Null 1)	p = 0.745 $\omega_{background}$ = 0.1224 ω_C = 0.1224	p = 0.095 $\omega_{background}$ = 1 ω_C = 1	p = 0.142 $\omega_{background}$ = 0.1224 ω_C = 1	p = 0.018 $\omega_{background}$ = 1 ω_C = 1
Positive Selection	8714.4092	No (vs Null 2)	p = 0.745 $\omega_{background}$ = 0.1224 ω_C = 0.1224	p = 0.095 $\omega_{background}$ = 1 ω_C = 1	p = 0.142 $\omega_{background}$ = 0.1224 ω_C = 1	p = 0.018 $\omega_{background}$ = 1 ω_C = 1

Table 8: Model parameters for dN/dS ratio switching

	M2a	M2a + S1	M2a + S2
lnL	-8875.763712	-8831.749794	-8836.182688
$\omega_1\omega_2\omega_3$	0.044, 1, 0.293	0.318, 1, 0.003	0, 1, 0.668
p_1 p_2 p_3	0.544, 0.030, 0.425	0.271, 0.093, 0.636	0.738, 0.013, 0.248

Table 9: F-statistics and significance levels for means contrasts of enzyme assays

DFR-A	Apigeninidin		Luteolinidin		Pelargonidin		Cyanidin		Delphinidin	
Factor	df	*F*	df	*F*	df	*F*	df	*F*	df	*F*
Species	3	1.73	3	5.2	3	30.7*	3	140.48*	3	49.07*
Time	1	0.77	1	1.44	1	17.24*	1	37.18*	1	2.8
Time x Spp	3	0.06	3	1.66	3	15.24*	3	12.73*	3	1.83
Contrast	1	1.68	1	15.57*	1	33.22*	1	12.73*	1	102.02*

DFR-B	Apigeninidin		Luteolinidin		Pelargonidin		Cyanidin		Delphinidin	
Factor	df	*F*	df	*F*	df	*F*	df	*F*	df	*F*
Species	3	39.86*	3	15.79*	3	26.51*	3	5.57	3	24.62*
Time	1	1.76	1	0.83	1	5.1	1	9.85	1	17.81*
Time x Spp	3	0.03	3	0.02	3	0.65	3	0.65	3	0.34
Contrast	1	90.04*	1	36.08*	1	85.39*	1	5.93	1	18.9*

DFR-C	Apigeninidin		Luteolinidin		Pelargonidin		Cyanidin		Delphinidin	
Factor	df	*F*	df	*F*	df	*F*	df	*F*	df	*F*
Species	3	1.26	3	2.96	3	28.28*	3	147.3*	3	67.98*
Time	1	1.82	1	3.93	1	13.42	1	40.54*	1	4.36
Time x Spp	3	0.03	3	0.44	3	16.01*	3	13.18*	3	2.18
Contrast	1	0.89	1	8.24	1	26.97*	1	319.94*	1	143.59*

2. A Mutation in a Single Gene Controls the Flower Color Change Conferring Hummingbird Pollination in *Ipomoea quamoclit*

2.1 Introduction

Determining which genes are responsible for adaptive change, and understanding the rules that govern the kinds of mutations that are allowed in these genes is now a central goal of evolutionary genetics (Orr 2005). Theoretical and empirical results from diverse fields predict that pleiotropy may constrain the number, size and kinds of mutations available during adaptive episodes, and in which genes these mutations occur. Whether evolutionary change is accomplished by mutations in coding or non-coding sequences has been of particular interest in recent years (Carroll et al. 2001; Hoekstra and Coyne 2007; Stern 2000; Wray 2007). Identifying the causal gene or genes underlying adaptive change is a first step towards understanding these mechanisms that constrain, generate and maintain biological diversity.

Flowers are conspicuous and highly specialized adaptations for facilitating pollination. A pollination syndrome is a suite of floral characters associated with the attraction of a particular species or guild of pollinators (Faegri and Van Der Pijl 1966; Fenster et al. 2004). One commonly observed syndrome is that associated with bird pollination, which has evolved numerous times from an ancestral bee-pollination syndrome (Cronk and Ojeda 2008; Thomson and Wilson 2008). Bird pollinated flowers are typically red or orange in color, have long, slender floral tubes, reproductive parts that extend beyond the face of the flower, little or no scent, and produce relatively large amounts of dilute nectar. Some general rules have been proposed for the order of assembly of this suite of characters, for example whether characters that attract pollinators precede those which enhance effective pollination by particular species (Thomson and Wilson 2008), though these rules remain untested.

In most flowering plants, red, blue and purple pigments are anthocyanins, which are produced in the flavonoid biosynthetic pathway (Grotewold 2006). Flavonoids constitute a large and functionally diverse group of secondary metabolic compounds that are present in all land plants and whose biochemistry and genetic control is exceptionally well-characterized (Harborne 1967; Forkman and Heller 1999; Quattrocchio et al. 2006). Most flowers are pigmented with elaborations of the 3-hydroxyanthocyanidins pelargonidin, cyanidin and delphinidin, which are produced through successive branching points in the flavonoid pathway (Fig. 5). The downstream enzymes dihydroflavonol-4-reductase (DFR) and anthocyanidin synthase (ANS) act independent of the presence of extra hydroxyl groups on the substrate B-ring which are added by the branching enzymes flavonol 3' hydroxylase (F3'H) or flavonol 3' 5' hydroxylase (F3'5'H). Changing flux down one of these branches, then, can change the relative amounts of the pigments produced and thereby change flower color.

However, changing flux along branches of the flavonoid pathway may incur pleiotropic costs for a plant. In addition to their role in floral pigmentation, various classes of flavonoids are expressed throughout developing and adult plants where they serve diverse functions including signalling, enhancing pollen viability and defending against herbivores, UV damage and other environmental stresses (Gould 2004; Shirley 1996). While many of these compounds arise from early branch points in the flavonoid pathway, condensed tannins and some flavonols are derived from the same immediate precursors as is cyanidin, and so their synthesis also requires activity of the branching enzyme F3'H (Figure 5). Modification of cyanidin itself through addition of various sugar moieties leads to compounds with important defense roles in vegetative tissues (Gould 2004). How the diverse vegetative roles of flavonoids influence evolutionary transitions between flower hues remains an open question.

In the morning glories, genus *Ipomoea* (Convolvulaceae), the production of cyanidin giving rise to purple flowers conferring bee pollination is the ancestral condition with numerous transitions to white, yellow or red flowers (McDonald 1991). The species of one clade of red-flowered morning glories, *Ipomoea* subgenus *Quamoclit* section *Mina*, produce abundant pelargonidin pigment in their flowers and show several other floral characters associated with bird pollination. The *Mina* species *I. quamoclit* and *I. hederifolia* are visited most often by hummingbirds and sulfur butterflies, and very rarely by bees, whereas birds rarely visit the blue flowered species outside the *Mina* clade, such as *I. hederacea* and *I. trichocarpa* (Wolfe and Sowell 2006). Prior genetic work comparing the red-flowered *Mina* species *I. quamoclit* with the distantly related purple-flowered *I. purpurea* inferred three independently sufficient changes for the transition to red flowers in the *Mina* clade: downregulation of the branching enzyme F3'H, reduction in the enzymatic activity of F3'H, and reduction in the activity of DFR on the cyanidin precursor dihydroquercetin (Zufall and Rausher 2004). These authors proposed that the redundant losses of cyanidin-producing functions constituted preliminary evidence for the degeneration of a biochemical pathway following a relaxation of purifying selection. However, considering the role played by vegetative flavonoids, some of these losses might incur considerable pleiotropic costs as changes in biosynthetic enzymes could not be restricted to floral tissues. The identities and synthetic control of flavonoids in the vegetative tissues of red-flowered morning glory species is currently unknown.

Here I describe the genetic changes in the flavonoid biosynthetic pathway that gave rise to red flowers in the *Ipomoea Mina* clade. Directly identifying how, if at all, these plants regulate differential production of flavonoids in flowers and vegetative tissues will inform us how evolution tinkers with branching biochemical pathways in the generation of evolutionary novelty, and if the evolution of novelty has deleterious effects on the prospects for future evolutionary change. Additionally, identifying the gene responsible for one component of a floral syndrome establishes a framework for future developmental and ecological work to determine the order of assembly of this complex adaptation.

2.2 Results and Discussion

2.2.1 The genetic basis of red flowers in *Ipomoea* section *Mina*

Phylogenetic analysis (Miller et al. 2004) provides a framework for tracing the evolutionary history of flavonoid production in the *Mina* clade (Fig. 6). The 16 species (Austin and Huaman 1996) in this monophyletic group all have red or orange flowers and are predominantly pollinated by birds and butterflies. Sister to the *Mina* clade is a small monophyletic clade of purple-flowered species. In the present study, I compare one of these purple species, *I. ternifolia*, to several red flowered species. For functional assays, I focus on the red-flowered *I. quamoclit* which is sister to the remaining *Mina* species.

HPLC analysis of the floral anthocyanidin content of *I. ternifolia* reveals that the predominant pigment in this purple-flowered species is cyanidin, with trace amounts of pelargonidin (Table 10). This finding is consistent with previous studies which have identified cyanidin as the pigment underlying purple flowers in many other *Ipomoea* species (Zufall 2003); delphinidin production has not been observed in *Ipomoea* species. The red flowers of *I. quamoclit* and *I. coccinea* produce almost exclusively pelargonidin (Table 10). Combining these results with prior work on the distribution of anthocyanidins in *Ipomoea* species suggests a single loss of cyanidin production and gain of pelargonidin production in flowers in the *mina* clade (Fig. 6).

To stimulate production of defense-related flavonoids in vegetative tissue, I subjected *I. ternifolia*, *I. quamoclit* and *I. coccinea* plants to intense heat and light stress over a five day period, with normal watering. HPLC analysis of the anthocyanidin content of the vegetative tissue of these plants reveals that all three species produce predominantly cyanidin in response to heat and light stress (Table 10). My process of extracting and purifying anthocyanidins resulted in the loss of higher-order structures which may have represented the actual compounds deployed in leaves. In *I. batatas*, a cultivated congener of the morning glories studied here, a large number of acylated cyanidin and peonidin derivatives have been identified (Islam et al. 2002), and it is probable that similar cyanidin derivatives are found in the *Mina* species. Condensed tanins are polymerized derivatives of cyanidin and its precursor leucocyanidin and, while these compounds have not been studied in morning glories, the cyanidin detected in my stressed tissues may also be from depolymerized condensed tannins. My HPLC results indicate that red-flowered *Mina* species must retain a functional cyanidin producing branch in their flavonoid pathway, suggesting that the enzymatic function of the branching enzyme F3’H and downstream enzymes DFR and ANS have not changed appreciably in these taxa.

Because DFR and ANS enzymes are shared between the pathway branches that give rise to pelargonidin, cyanidin, and delphinidin, changes in expression of these genes would affect flux down all three branches; expression changes in these genes are therefore not candidates for the transition from cyanidin to pelargonidin production in flowers. Gene duplication and subsequent pathway specialization by enzymes is one route to avoid such a constraint and, indeed, *DFR* is present as a three-gene family in morning glories. However, I have shown previously that only one of these copies (DFR-

B) has activity on the dihydroflavonols (Des Marais and Rausher 2008), suggesting that specific copies have not been recruited for different branches of 3-hydroxyanthocyanidin production.

Considering the apparent conservation of cyanidin-producing enzymatic function in red-flowered species, I turn now to understanding the expression of the branching enzyme F3'H. Previous results from Northern blots indicated that *F3'H* is downregulated in red *I. quamoclit* flowers as compared to the distantly related purple-flowered *I. purpurea* (Zufall and Rausher 2004). Quantitative PCR (Q-PCR) analysis of *F3'H* expression demonstrates that, relative to purple-flowered *I. ternifolia, I. quamoclit* and *I. coccinea* have greater than 10-fold fewer *F3'H* transcripts in their floral tissue (Fig. 7). A one-way ANOVA by species confirms this difference ($F = 30.7885$, $P=0.0015$). Comparison of means with the Tukey-Kramer test at $\alpha = 0.05$ shows that the expression differences between species group by flower color: *I. ternifolia* (mean = Log -0.1425) ≠ *I. quamoclit* (mean = Log -1.337) = *I. coccinea* (mean = Log -1.540). This downregulation of floral *F3'H* expression is correlated with the observed loss of floral cyanidin production in the studied *Mina* species (Fig. 6).

F3'H expression in the cyanidin-producing vegetative tissue of the heat-stressed *I. quamoclit* and *I. coccinea* plants is statistically indistinguishable from the vegetative tissue of *I. ternifolia* (Fig. 7). Although a one-way ANOVA by species shows a slightly significant effect ($F=10.2572$, $P=0.0456$), a Tukey test shows that *I. quamoclit* (mean = Log 0.115) = *I. ternifolia* (mean = Log 0.052), *I. ternifolia* = *I. coccinea* (mean = Log -

0.34), *I. coccinea* ≠ *I. quamoclit*. The observed difference in means, then, is between the two red species, but the variance of each overlaps with that of *I. ternifolia*. It appears that *F3'H* expression is under independent control in vegetative and floral tissues, and that the evolutionary change at the base of the *Mina* clade was restricted to expression in floral tissue.

To directly test the effect of the change in *F3'H* floral expression on the production of floral cyanidin in *Ipomoea quamoclit*, I constructed a transgenic morning glory. I placed the endogenous *I. quamoclit F3'H* gene under the control of the CaMV 35S promoter, which confers strong, constitutive expression throughout the plant. This construct was stably transformed into wild-type *I. quamoclit* callus tissue and adult plants were grown from these calli. I recovered three lines representing independent transformation events and grew the plants under normal conditions in the greenhouse. Figure 8 shows a flower from a T2 individual from one of these lines compared to wild-type *I. quamoclit*, indicating the rescue of the ancestral, purple-flowered, phenotype. All three lines produced individuals displaying purple flowers. HPLC analysis of the anthocyanidins extracted from transgenic flowers reveal a high ratio of cyanidin: pelargonidin production (Table 10). The lower ratio of cyanidin: pelargonidin in transgenic lines as compared to wild-type purple species is due to continued production of some pelargonidin in the transgenic plants. These results clearly demonstrate that restoring expression of *F3'H* in floral tissues is sufficient to activate the cyanidin branch of the anthocyanin pathway.

The observed expression change in floral tissue could be due to changes either in *cis* or *trans* relative to the *F3'H* gene. I used transient transformation of *I. purpurea* flowers via particle bombardment to test the relative strengths of *F3'H* promoters from purple and red flowered species. Figure 9 shows the relative activities of a *F3'H* promoter driving the reporter gene GUS (beta-galacturonidase) as a function of constitutively-expressed luciferase to control for transformation efficiency. A two-tailed t-test confirms that the promoter from cyanidin-producing *I. purpurea F3'H* drives greater GUS production than does the promoter from cyanidin-lacking *I. quamoclit* F3'H (Student's t-test: $t=5.11$, d.f. $= 20$, $p<0.0001$). The differential response of the promoters from red and blue species indicates that the low observed expression of *F3'H* in *I. quamoclit* flowers is due to factors acting in *cis*. Thus, the evolution of pelargonidin-producing flowers in the *Ipomoea Mina* clade from cyanidin-producing ancestors was achieved through a change in the promoter sequence that previously conferred *F3'H* expression in flower tissue.

Alternative explanations for the difference in *F3'H* expression driven by promoters of purple and red flowers involve the transcription factors that control *F3'H*. Two experiments are underway to address this alternative. First, the reciprocal transient expression assay will be conducted by bombarding GUS driven by F3'H promoters from *I. quamoclit* and *I. purpurea* into *I. quamoclit* floral tissue. This experiment will confirm that the sampled promoter fragment from *I. quamoclit* included homologous elements to the sequence used from *I. purpurea* and will confirm that there has not been coevolutionary divergence of transcription factor- promoter sequence interactions between the two species. The definitive test of whether expression changes are due to elements in cis would be to assay the phenotype of a stably transgenic *I. quamoclit* line with *I. ternifolia F3'H* promoter driving *I. quamoclit F3'H*.

In the absence of results from these two planned experiments, there is some evidence from the literature that transcription factor changes may not be responsible for the observed difference in *Mina F3'H* floral expression. In similar transient assays to those employed here, the *Petunia* transcription factors *jaf13* (a bHLH protein) and *an2* (a MYB protein), which regulate their endogenous downstream biosynthetic genes *dfr* and *ans* but not the upstream gene *chsA*, are able to activate the promoter of the upstream *c2* (which codes for CHS) gene from the distantly-related maize genome; in maize, a single set of transcription factors regulate all six core structural genes of the flavonoid pathway (Quattrocchio et al. 1998, 1993). This finding implies highly conserved recognition of promoter motifs by *Petunia* transcription factors which faithfully replicate the activity of their maize host regulators. A second alternative explanation for the downregulation of *F3'H* is that the transcription factor(s) necessary for *F3'H* activation are themselves not expressed in *I. quamoclit* floral tissue. I feel that this explanation is unlikely. In *Ipomoea*, *F3'H* appears to be closely co-regulated with *DFR* and *ANS* and, to a lesser extent, with *CHS*, *CHI* and *F3H* (Morita et al. 2006; Park et al. 2007; Tiffin et al. 1998). If the *I. quamoclit* transcription factors are not expressed in floral tissue, my transgenic *I. quamoclit F3'H* overexpression lines would not have produced cyanidin in their flowers because DFR and ANS proteins would not have been present to convert DHQ into cyanidin. I therefore conclude that a mutation in the promoter region of *F3'H*, specific to the sequence(s) which activated expression in flower tissue, converted a purple-flowered ancestor into a plant producing red flowers.

Mounting evidence suggests that adaptations involving change in plant pigments can be attributed to a small number of genes of apparently large effect. Using a QTL approach, Bradshaw and Schemske (Bradshaw et al. 1995) showed that a single locus in red-flowered *Mimulus cardinalis* conferred 88% of the variation associated with red-hued carotenoid deposition in a mapping cross with the pink *M. lewisii* and that, when introgressed into *M. lewisii,* this single change conferred a 68-fold increase in the number of hummingbird visits (Bradshaw and Schemske 2003). The gene underlying this difference in pigmentation has not been identified. Quattrocchio and colleagues (Quattrocchio et al. 1999) identified several independent loss of function mutations in a single transcription factor which confers white flowers in *Petunia axillaris*. The absence of pigmentation in *P. axillaris* distinguishes it from a sympatric congener, the red-flowered *P. integrifolia*, though the authors propose that the loss of pigmentation occurred after the separation of these lineages and was not causal in their divergence. Two loci coding for transcription factors, *ROSEA* and *ELUTA*, are correlated with a flower color polymorphism observed in two subspecies of *Antirrhinum majus* across a cline and the *ROSEA* locus shows the effects of selection acting to maintain the difference across a hybrid zone (Schwinn et al. 2006, Whibley et al. 2006). These studies, combined with my work directly demonstrating a role for *F3'H* expression in the transition from blue to red flowers in morning glories, demonstrate the important part played by regulatory changes in the evolution of flower colors.

2.2.2 Possible Pathway Degeneration

A long-standing hypothesis in biology is that evolutionary specialization may constrain the potential directions of subsequent evolutionary change (Gould 1977; Gould 2002; Simpson 1953). In the genus *Ipomoea*, red flowers have evolved at least four times independently (Zufall 2003), though no phylogenetic transitions back to the ancestral, bee-pollinated, condition have been observed. This directionality from bee to bird pollinated flowers (and blue/purple to orange/red pigments) appears to be common in several groups of plants (Bruneau 1997; Kay et al. 2005; Whittall and Hodges 2007; Wilson et al. 2007), a finding which suggests that there may be constraints on back transitions to bee pollination. As described earlier, the topology of the anthocyanin biosynthetic pathway suggests that the transition from purple to red floral pigments involves a loss-of-function mutation; I have shown here that a single regulatory change causes this transition. Prior work suggested that genetic constraint in the anthocyanin pathway, resulting from pathway degeneration following the transition to exclusive pelargonidin production, might explain the lack of evolutionary transitions from red to purple flowers (Zufall and Rausher 2004). Specifically, loss in enzymatic activity of F3'H on the substrate DHK, DFR activity on the substrate DHQ, or ANS activity on the substrate leucocyanidin, might occur if cyanidin production were to become vestigial due to pollinator shift. If this degeneration took place, then red-flowered lineages experiencing selective pressure to change flower color back to purple would need to fix multiple independent mutations in order to produce purple flowers. This condition would be a formidable obstacle to the evolution of bee-pollinated flowers from red bird-pollinated ancestors. However, my results showing abundant cyanidin accumulation in the vegetative tissues of *I. quamoclit* and *I. coccinea* argue strongly against any enzyme degeneration having occurred, and that change in a single gene might be sufficient to re-evolve blue flowers.

Two additional lines of evidence suggest that the cyanidin-branch enzymes have not degenerated in the *Mina* clade. Functional F3’H enzyme is indicated by the results of analysis of sequence evolution. In the absence of conserved function, the ratio of nonsynonymous to synonymous substitutions in a gene is expected to increase, reflecting the accumulation of mutations in a neutral fashion (Li 1997). However, a phylogenetic test for change in dN/dS ratios following the origin of red flowers reveals no departure from background rates of purifying selection (lnL of a single ratio for the tree = -5294.2377, two-ratio model lnL = -5294.1547; likelihood ratio test $\lambda = 0.166$, $P>0.9$).

Results of direct assays of DFR enzyme function suggest that DFR retains ancestral activity in red species as well. Previous work (Zufall and Rausher 2004) suggested that DFR from *I. quamoclit* had reduced activity on the cyanidin precursor dihydroquercetin (DHQ) relative to the purple species *I. purpurea*. Using a more sensitive in vitro assay of DFR substrate specificity, I show here that heterologously expressed DFR from *I. coccinea* and *I. quamoclit* retain activity on both DHQ and the pelargonidin precursor dihydrokaempferol (DHK; Fig. 10). Using the same assay conditions, Des Marais and Rausher (2008) showed that DFR-B from morning glories have an approximately 4:1 ratio of activity on DHK:DHQ; similar ratios were found for the species sampled here (Fig. 10). A one-way ANOVA shows that the DHK activities of red-flowered *I. coccinea* and *I. quamoclit* are not statistically different from purple-flowered *I. ternifolia* $F = 0.1617$, $P = 0.8576$.

I therefore find no evidence of cyanidin pathway degeneration in the red-flowered species *I. coccinea* or *I. quamoclit*. What, then, explains the paucity of evolutionary back-transitions to blue bee-pollinated flowers from red hummingbird pollinated flowers? In addition to red pigment, the flowers of *Mina* species have several other features of bird-pollination including long, narrow corollas and exserted anthers (Zufall 2003). Increasing evidence suggests that red flowers and the bird syndrome are best interpreted as adaptations against bee visitation. Despite abundant evidence that birds primarily visit red flowers, many species will also forage on blue flowers, suggesting that they can effectively discriminate either color against a leafy green background (reviewed in Chittka and Waser 1997). By contrast, hymenopteran (including bee) visual pigments do not contrast red against a green background well, especially under some lighting conditions (Chittka and Waser 1997). Optimal foraging theory suggests that bees may not invest the effort to find red flowers in the presence of suitable blue ones and, therefore, that birds may preferentially visit red flowers simply to avoid resource competition with bees (Rodriguez-Girones and Santamaria 2004).

My results suggest that the evolution of purple flowers from red ones may simply involve the re-acquisition of a promoter motif in *F3'H*. How might bees respond if such a mutation were to arise in a species with otherwise bird-adapted flowers (e.g. with long, tubular corolla, small face and exserted stamens and pistil)? Bees generally land on large flower petals to collect pollen as a food source, whereas pollination by hummingbirds is airborne and incidental and therefore net pollen removal from floral displays by birds is minimal; accordingly, bird pollinated species are generally expected to produce fewer pollen grains (Thomson et al. 2000). In *Penstemon*, experimentally introducing bird-type floral characters (other than color) into a bee-type background reduces the effectiveness of bee pollination, suggesting that if bees were attracted to purple flowers that are otherwise bird-type, they would serve as inefficient pollinators (Castellanos et al. 2004). Additionally, there may be selection against purple flowered bird-adapted species, because attracting bees may leave flowers vulnerable to nectar robbing. Under this scenario, bees attracted to purple, bird-type flowers would be inefficient pollinators at best, and deterrents to effective pollinators at worst. Behavioral and ecological factors, rather than genetic constraint, may therefore explain the observed unidirectionality of pollinator syndrome evolution.

2.3 Methods

2.3.1 Plant Materials and Extraction of Nucleic Acids

Plant material was collected fresh from living specimens of *I. ternifolia, I. lobata, I. lutea*, *I. quamoclit* and *I. coccinea* growing in our greenhouse. RNA was extracted from buds and stem tissue using the Spectrum Plant Total RNA extraction kit (Sigma-Aldrich Corp., St, Louis, MO) according to the manufacturer's protocol. DNA was extracted from leaf tissue using a modified cTAB protocol.

2.3.2 Pigment Extraction and Identification

Anthocyanins were extracted from fresh floral and vegetative tissue by soaking 250 mg of tissue overnight at room temperature in 2M HCl and then boiling for 30 min. Following centrifugation, the supernatant was cleaned three times with ethyl acetate, leaving anthocyanidins in the aqueous phase. Anthocyanidins were subsequently extracted into isoamyl alcohol, which was then evaporated under vacuum. Dried purified anthocyanidins were resuspended in methanol-HCl (99:1 v/v) and stored at -20°C until processing.

Anthocyanidins were separated via high performance liquid chromatography (HPLC) on a Shimadzu 10A system (Shimadzu Corp., Columbia, MD). Separation in a 30°C oven was achieved with a propanol gradient from 15-27.5% in H_2O. Trifluoroacetic acid was added to all HPLC solvents to a final concentration of 0.1%. Retention times were calibrated against commercially synthesized standards: pelargonidin (Indofine;

Hillsborough, NJ), delphinidin (Polyphenols Laboratories AS; Sandnes, Norway), and cyanidin (Indofine).

2.3.3 *F3'H* Gene Tree and Selection Analyses

F3'H coding sequence was amplified using cDNA from *I. ternifolia*, *I. coccinea*, *I. lutea* and *I. lobata* as templates. Primers D1-1F (5'-GCAGTACTAGTATGGCTACCTTAACCTTAATT-3') and D1-2R (5'-GCGTGGATCCTT AATTGAGAGTATAGAGATG-3') were used to amplify the complete coding sequence of *I. ternifolia* and *I. coccinea*; nested PCR using the *F3'H* primers reported by Whittall et al. (2006) were used to isolate 854 b.p. fragments from *I. lutea* and *I. lobata*. All novel sequences reported herein have been deposited in GenBank (FJ004840-FJ004843).

F3'H sequences were aligned with published sequences available in GenBank. The alignment is available by request. For selection analyses, I enforced the species topology of Miller et al. (2004). I determined if changes in selection on *F3'H* coding sequence followed the origin of red flowers using the codeml program in the PAML 3.15 package (Yang 1997). I tested two models: the null model assigned a single dN/dS ratio to all branches of the tree. In the alternate model, the branches of the *Mina* clade were assigned a separate ratio. I tested whether the two-ratio model was a better fit of the data using a likelihood ratio test with one degree of freedom between the models.

2.3.4 *F3'H* Expression Analysis

cDNA was synthesized from floral RNA at concentrations of 75 ng/μL from total RNA from three individuals each of *I. ternifolia*, *I. quamoclit* and *I. coccinea* and two individuals for vegetative tissue. 20 μL quantitative PCR (Q-PCR) assays were run with Finnzymes SYBR-Green PCR mix (Espoo, Finland) and 10 ng cDNA template using the following primers whose sequence was conserved in all three taxa: forward primer F3'H 137F: 5'-GGCCGGTGTTCGGAAACCTTCC-3' and reverse primer F3'H 384R: 5'-CCTAAGCATGCGCCACCGC-3'. The elongation factor 1-alpha gene (*EF1α*) was used as a control gene as amplified by forward primer EF1-F: 5'-TCCACTTCAGGATGTGTACAAGA-3' and reverse primer EF1-R: 5'-AGCAACATTCTTCACATTGAACC-3' kindly provided by M. Streisfeld. For both genes, reactions were run on an ABI 7000 (Applied Biosystems, Foster City, CA) for 40 cycles with an annealing temperature of 60°C.

I determined the relative expression of *F3'H* between species by assigning the level of expression of one *I. ternifolia* individual as zero, and calculating the log-change of all other individuals relative to that sample. A one-way analysis of variance (ANOVA) partitioned the response by species. To test which, if any, of the means of species differed from each other, I performed Tukey-Kramer HSD test at alpha=0.05. All statistical analyses were performed in JMP ver. 7.01 (SAS, Cary, NC).

2.3.5 DFR Enzyme Substrate Specificity Assays

The ability of DFR-B from *I. purpurea*, *I. quamoclit* and *I. coccinea* to reduce dihydrokaempferol and dihydroquercetin was assayed *in vitro* and the results analyzed statistically as described previously (Des Marais and Rausher 2008). Two replicates of each assay were run for 3 hours. A one-way analysis of variance (ANOVA) partitioned the response by flower color.

2.3.6 Transgenics

To generate *I. quamoclit* callus tissue for transformation, anthers were plated on agar-solidified MS minimal organics medium (Murashige and Skoog 1962) supplemented with 3% sucrose, 15 mg/L folic acid (FA) and 0.5 mg/L 6-benzylaminopurine (BAP). The tissue was sub-cultured every 3 weeks. Approximately 4 weeks after initiation, callus formed and 4 weeks after callus formation, somatic embryo developed from the callus. Embryos were separated from developing callus and transferred to fresh medium every 4 weeks. Embryogenic callus cultures were maintained at 26°C in the dark.

The full-length transcript of *I. quamoclit F3'H* was subcloned between the CaMV 35S promoter and nopaline-synthase terminal sequence and then subcloned into a Cambia 1301 binary vector. This construct (D5) was transformed into *Agrobacterium tumefasciens* strain LBA4404. Pro-embryo masses were co-cultured with *Agrobacterium* on MS medium supplemented with 15 mg/L FA, 0.5 mg/L BAP and 200 μM acetosyringone. After co-cultivation, tissue was transferred to MS minimal organics medium supplemented with 15 mg/L FA, 0.5 mg/L BAP, 400 mg/L carbenicillin, and 50 mg/L hygromycin sulfate. Callus was subcultured every 3 weeks to fresh medium until hygromycin resistant embryos formed.

Whole plants were regenerated from embryogenic callus by transferring embryos to MS minimal organics medium containing either 0.5 mg/L Zeatin or 2.0 mg/L BAP and 0.2 mg/L IAA. Elongating shoots were transferred to 0.5X MS minimal organics medium for rooting. Once roots developed, these plants were transferred to MetroMix 200 seedling mix (SungGro Horticulture, Bellevue, Washington) and placed in a growth chamber under 16 hr days at 22°C. Plants were then transferred to a greenhouse. Three lines representing independent transformation events were recovered. These plants were selfed, and their offspring were scored for presence of the transgene via PCR, anthocyanidin content via HPLC as described above, and *F3'H* expression via Q-PCR, as described above.

Proximal promoters for *I. quamoclit* and *I. purpurea F3'H* were isolated by genome walking with the SeeGene kit (SeeGene, Seoul, Korea) following the manufacturer's protocol. A modified pUC19 plasmid (kindly provided by F. Quattrocchio) was engineered with the following cassette: 1299 b.p. of *I. purpurea F3'H* promoter and 5' UTR and the first 27 b.p. of the F3'H coding sequence spliced as a translational fusion with the β-glucuronidase (GUS) gene and nopaline-synthase (nos).

An equivalent construct was made with 1342 b.p. of *I. quamoclit* F3'H promoter and 5' UTR.

These test promoter GUS constructs were co-bombarded with a pUC19-based plasmid carrying 35S::Luciferase (kindly provided by R. Zentella) into 1.5 day preanthesis *I. purpurea* flowers deficient in CHS activity (and, hence, lacking endogenous anthocyanin pigment which allowed ready detection of reporter protein activity). Dissected petal limbs were spread, adaxial side up, onto MS plates containing 2.5% sucrose. DNA was precipitated onto 0.6μm gold particles as described (Sambrook and Russell 2001). Biolistic delivery was achieved under vacuum using 1350 psi rupture disks in a DuPont Helios 1000 particle bombardment system (DuPont, Wilmington, Deleware). Bombarded tissue was incubated overnight at room temperature and then ground in protein lysis buffer (100 mM sodium phosphate buffer, pH 7.0, 10 mM dithiothreitol, 20 μg/mL leupeptin, 20% glycerol glycerol in H_2O) in chilled mortars and pestles. After centrifugation, for the luciferase reporter, 75 μL of cleared protein lysate was added to 100 μl of 2X luciferase assay buffer (LAB: 60 mM Tris-SO_4, 20 mM $MgCl_2$, 20 mM DTT, 2 mM EDTA in H_2O) containing 0.1M ATP. For the GUS reporter, 50μL of protein lysate was added to 200 μL MUG assay buffer (2.5 mM 4-methylumbelliferae beta-D-glucuronide trihydrate, 50 mM sodium phosphate, 10 mM EDTA, 10 mM DTT, 20 μg/mL leupeptin, 20% v/v methanol in H_2O) and incubated overnight at 37ºC. For visualization, 200 μL of 0.1M Na_2HCO_3 was added tp 50 μL of the MUG reaction. Both luciferase and GUS expression was quantified in solution on a Perkin Elmer Fusion α–HP FT plate spectrophotometer.

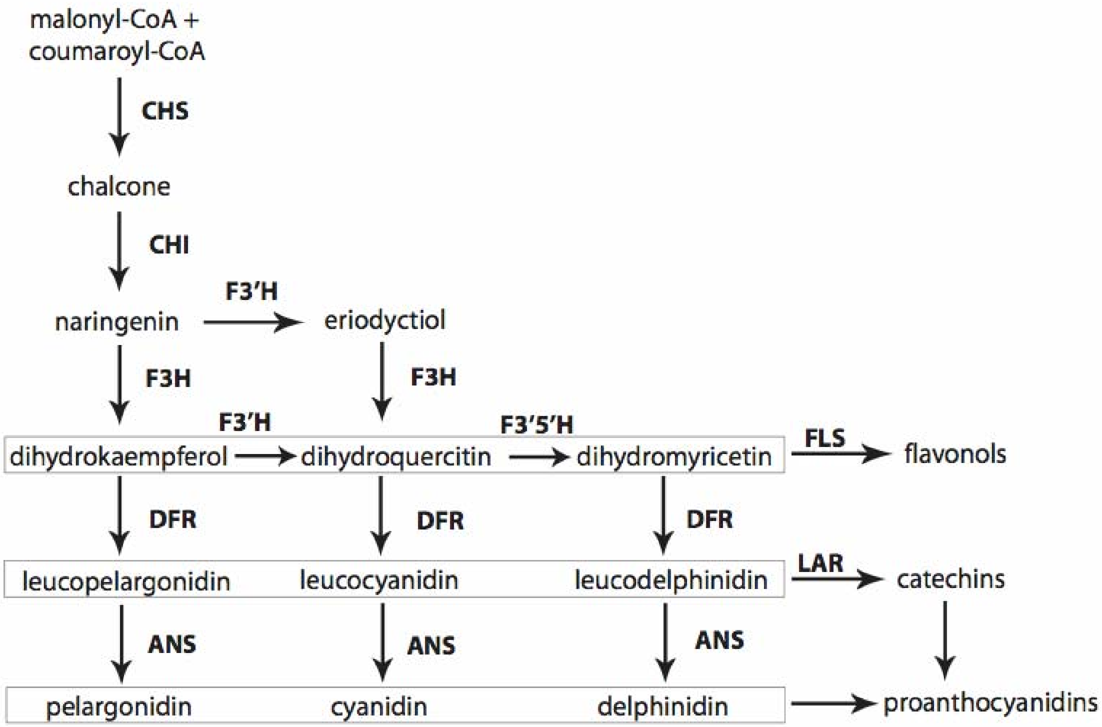

Figure 5: Generalized topology of the flavonoid biosynthesis pathway

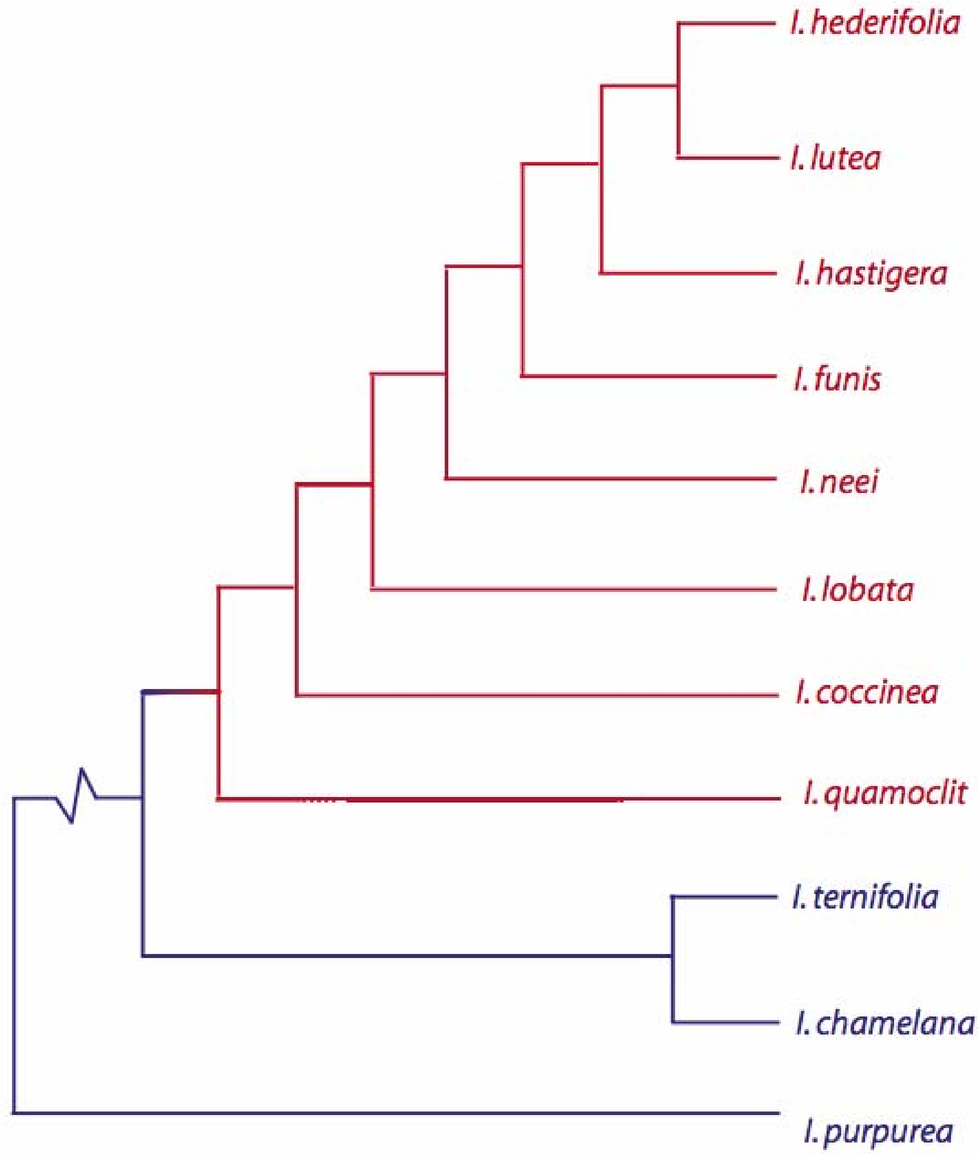

Figure 6: Phylogenetic relationships among *Ipomoea* section *Mina* species and close relatives

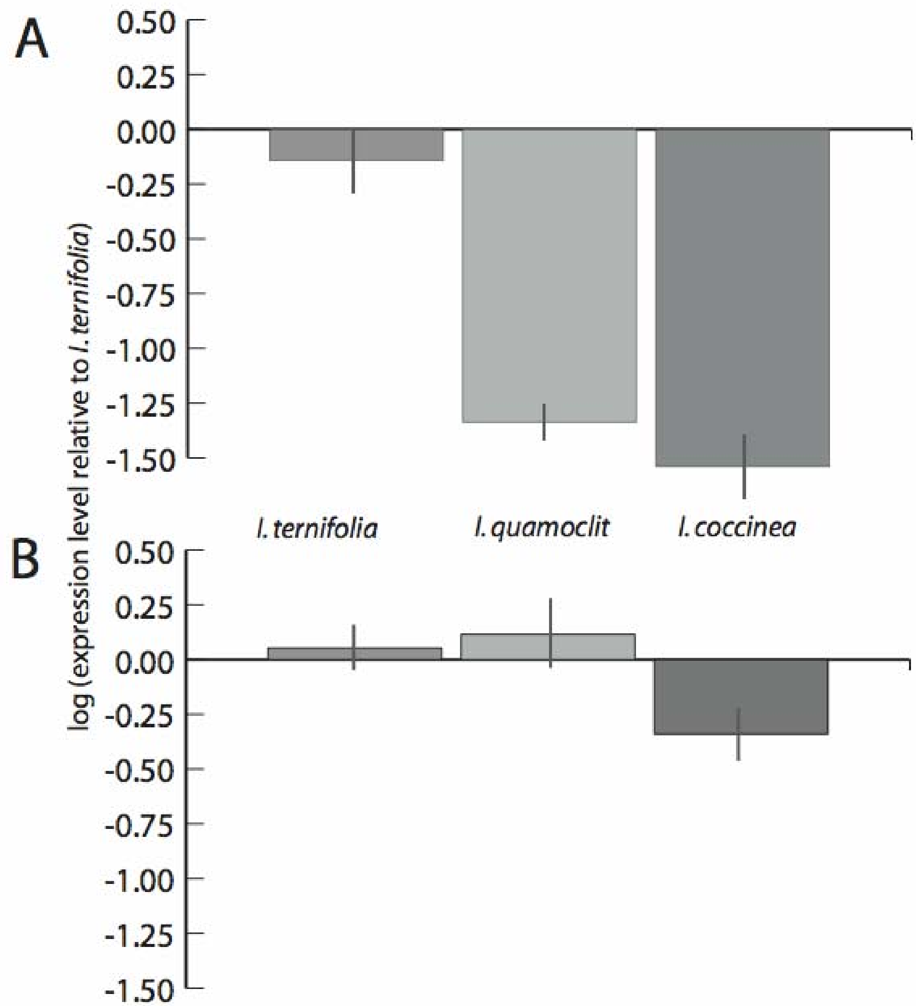

Figure 7: Relative *F3'H* expression levels among species as determined by Q-PCR for (A) Floral tissue and (B) Vegetative tissue. Bars are one standard error.

Figure 8: Wild type *Ipomoea quamoclit* (left) and transgenic line 71030-052 overexpressing *F3'H* (right)

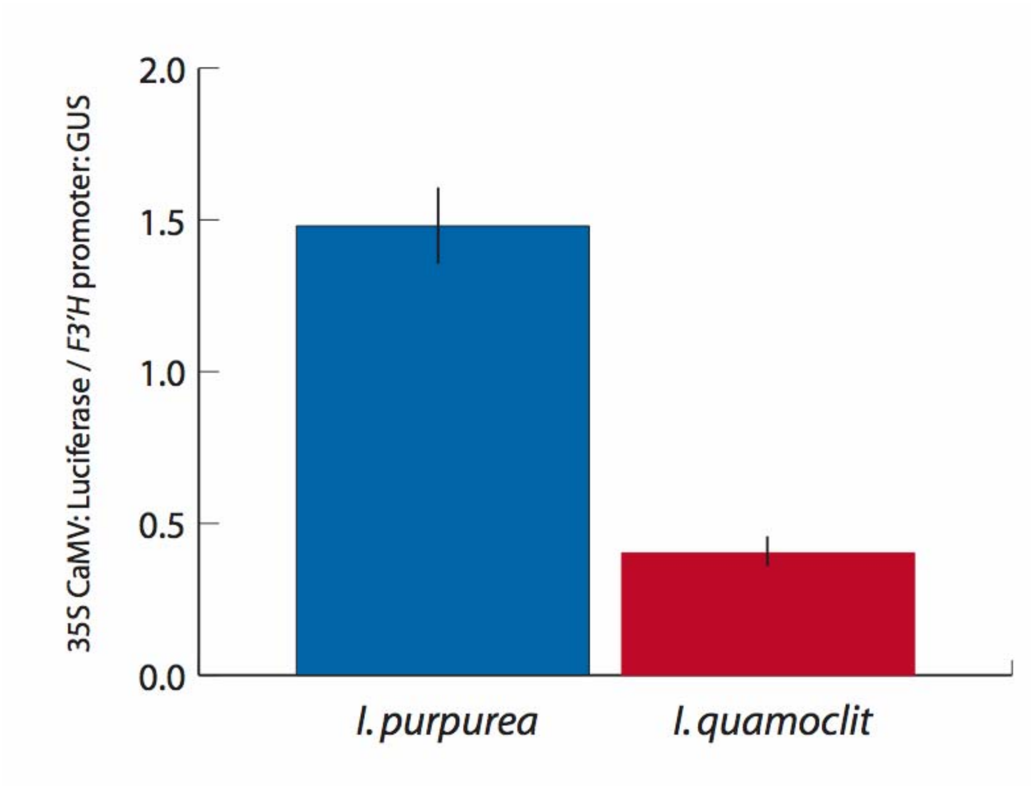

Figure 9: Relative strength of *F3'H* promoters in transient transgenic experiments. Bars are one standard error.

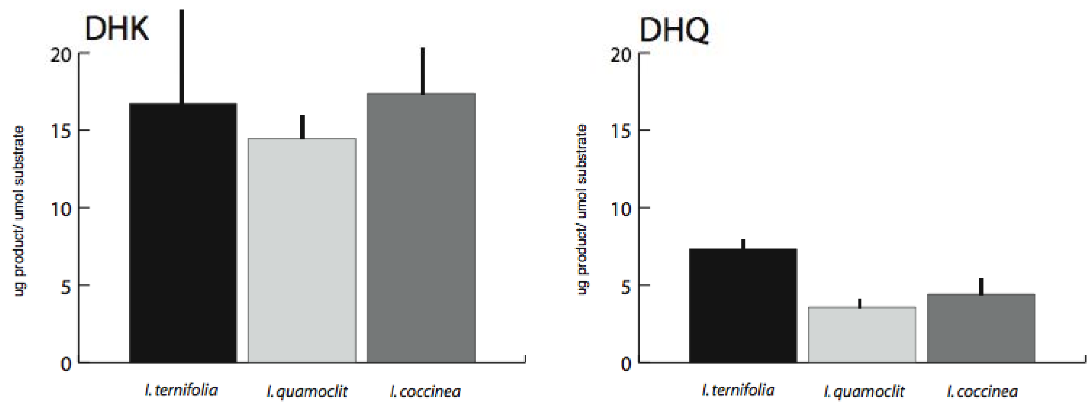

Figure 10: DFR enzymatic activity. Bars are one standard error

Table 10: Ratios of production of cyanidin to pelargonidin in the indicated tissues as determined by HPLC

Species	Cyanidin/ Pelargonidin	
	Floral	Vegetative
I. ternifolia	11.04	8.4
I. quamoclit	0.112	19.5
I. coccinea	0.03	14.9
Transgenic *I. quamoclit* Line 71030-052	4.57	-

Appendix 1 Protocol for *in vitro* assay of dihydroflavonol 4-reductase enzyme activity

This protocol was developed from the assay described by Stafford and Lester (1982). While their protocol made use of native proteins extracted from fresh flower tissue, the method presented here uses enzymes overexpressed in *E. coli*. We have successfully made DFR enzymes using Invitrogen's pDEST17 and pDEST14 vectors, though other workers have had success with a variety of bacterial and yeast overexpression vectors; we use the *E. coli* line BL-21 Star (DE3) from Invitrogen. The assay works well for the DFR substrates naringenin, eriodyctiol, dihydrokaempferol, dihydroquercetin and dihydromyricetin, though these substrates may have different optimal assay conditions that we have not yet determined. The assay protocol presented below is for 1 mL reactions; due to the cost of the reagents, we generally do 500 μL reactions. Adjust both assays contents and subsequent extraction steps accordingly.

Bacterial Overexpression

1. Start 5 mL overnight cultures from single colony picks. Grow in shaking incubator at 37°C in LB+Amp_{100}.

2. Autoclave 200 mL LB and a 250mL Erhlenmeyer flask for each clone. Add ampicillin to LB at a concentration of 100 mg/L. Leave the LB in the non-shaking incubator overnight to get the broth warm.

3. The next morning, draw 2 μL of overnight culture and run a PCR to check for the presence of the insert. Use T7 primer and an internal reverse GSP. (Alternatively, do this at the end of the 3-hour grow-out)

4. Dilute overnight into 200mL LB+Amp_{100} and grow at 37°C in shaking incubator at 200rpm until OD_{600} reaches ~0.5 (2-3 hrs from room temp LB).

5. Add 500 µL 200 mM IPTG solution to the 200 mL culture.

6. After three hours, aliquot samples into labeled 50 mL tubes and spin down in swing-rotor at 4°C for 20 minutes.

7. Pour off LB and store at -80°C.

Preparing Bulk Protein for Enzyme Assays

1. The DFR assay works well working off protein extracted from the pellets of 100 mL of culture grown for 3 hours.

2. Thaw pelleted cells on ice. Add 500 µL Lysis buffer (to make 10 mL of buffer, add to 7.85 mL H_20: 500 µL 1M HEPES, 87 mg NaCl, 1mL 200 mM PMSF, 100 µL 100 mM TPCK, 50 µL 200 mM TLCK) and 10 µL Lysozyme (stock: 100 mg/mL) to one tube of cells from 50 mL of culture. Flick to resuspend pellet.

3. Set swing-rotor temperature to 4°C.

4. Incubate lysis reactions on ice 30 minutes with occasional mixing.

5. Spin down 20 minutes at 4°C and transfer supernatant to pre-chilled Eppendorf. DFR is a soluble protein and will remain in the supernatant.

6. Yield is generally ~1.5 mL crude bacterial protein extract from two 50 mL tubes of cells.

DFR Assay

In 1 mL reaction:

1 µmol substrate (= 0.304 mg DHQ, 0.288 mg DHK, 0.32 mg DHM, 0.286 mg eriodyctiol: dissolve in >95% MeOH. For Sigma naringenin, which is racemic, use 2 µmol, 0.54 mg)

1µmol NADPH (= 1 vial ~ 1 mg; I generally dissolve this in the Tris)

6μmol glucose-6-P (= 1.7 mg)
1 U glucose-6-P dehydrogenase (= 0.25 mg)
200-400 μg enzyme (I use 250 μL of bacterial extract)
0.1M Tris (pH 7.4) to 1 mL

1. Mix and incubate 3 hr at 30ºC

2. Add 15 μL 6N HCl to stop the reaction.

3. Immediately extract three times in 1 mL ethyl acetate (I shake these vigorously and then spin down ca. 30 sec in the centrifuge. EtAc is the top layer and contains the leucoanthocyanidins, which you want).

4. Wash twice with 200 μL H_2O (H_2O forms a small bubble at the bottom- try to get as much as possible out).

5. Evaporate EtAc under vacuum at 35-45ºC into ~ 100 μL aqueous phase.

6a. For leucoanthocyanidins (DHK, DHQ, DHM as substrate), add 1 mL butanol-HCl reagent (95:5 v/v) and heat at 95ºC for 30 min (in the past, we have used n-butanol but t-butanol is preferable to reduce ether formation at the substrate 3-carbon).

6b. For flavan-4-ols (naringenin, eriodyctiol as substrate), add 1mL 2N HCl and heat at 95ºC for 30 min, then extract into isoamyl to clean, if desired.

7. After formation of anthocyanidins, the products can be visualized using light spectroscopy and compared to purified standards for quantification.

Appendix 2. Transient Transformation of *Ipomoea purpurea* flower petals via particle bombardment.

In the absence of a reliable, efficient and timely stable genetic transformation system, transient transformation allows researchers to directly test genetic hypotheses *in vivo*. We developed the following protocol to assay promoter function though, in principal, the protocol should be effective for functionally evaluating protein coding sequences as well. This protocol makes use of two independent reporter genes: in the presence of the substrate 4-methylumbelliferae beta-D-glucuronide trihydrate, beta-galacturonidase (GUS) emits visible light when stimulated by nm wavelengths. We used GUS as the reporter for our experimental promoter sequences. The luciferase gene was isolated from jellyfish; when presented with its substrate luciferin, luciferase emits detectable light. We used luciferase as the reporter gene for a constitutively active promoter, CaMV 35S, which provided an estimate of transformation efficiency in each particle bombardment replicate. Both of these reporter constructs were delivered to the host tissue in derivatives of the widely available pUC18 plasmid which is easily stored and isolated from *E. coli*.

The following protocol is the amalgamation of several pre-existing protocols, reported previously. In practice, nearly all host species and tissue types will have their own optimal experimental parameters.

Preparing the gold for shooting leaf or flower tissue

1. Weigh 60 mg of gold, pour into a 1.5 mL eppie.
2. Add 1 mL of 70% EtOH to the tube.
3. Vortex 5 min, let stand for 15 min.
4. Spin 5 sec, remove the supernatant.
5. Add 1 mL of dH_20 to the tube, vortex 10 sec, let stand 1 min.
6. Spin 5 sec, remove supernatant.
7. Repeat step 5-6 three more times.
8. Add 1 mL 50% glycerol to the eppendorf tube, store at -20°C.

Preparing DNA and Gold mixture

1. Allow gold mixture to warm slightly, then vortex. Add 50 μL to a 1.5 mL eppie.
2. Sequentially add, while vortexing at medium speed
   10 μL of DNA mixture (1 μg/μL) (5 μL each of control and experimental vector)
   50 μL of 2.5 M $CaCl_2$
   20 μL of 0.1M spermidine (free base; comes as 6.3M)
3. Vortex an additional 3 min.
4. Let sit on the benchtop for 1 min.
5. Spin down at high speed for 2 sec.
6. Discard supernatant.
7. Add 140 μL 70% EtOH, flick gently (don't disturb the pellet), remove supernatant.

8. Add 140 μL 100% EtOH, flick gently (don't disturb the pellet), remove supernatant.

9. Add 50 μL 100% EtOH. This ethanol needs to be very pure, as any water will cause clumping during the next step.

Coating the Microcarrier

1. Submerge microcarriers in a small dish of 100% EtOH (I typically do this before starting the gold preparation and let it sit).

2. Using forceps, remove microcarriers from EtOH and let dry on a Kimwipe.

3. Load microcarriers into the disks.

4. Vortex gold mixture and quickly pipette off 10 μL.

5. Add 10 μl gold-particle mix to center of microcarrier.

6. Repeat for as many microcarriers as needed (NOTE: re-vortex after every two).

7. Let particle mix dry on microcarrier.

Bombarding

1. Turn on gas cylinder. Set regulator to 1500 psi.

2. Turn on vacuum pump; turn on power to chamber (left red switch).

3. Place rupture disk into center of screw-on cylinder; screw cylinder on; tighten it with torque tool (until the bar just touches the black sleeve).

4. Place stopping screen and microcarrier into holder and screw on lid.

5. Place tissue on shelf (may adjust shelf height) and close door.

6. Turn on vacuum switch (middle red switch, up position); make sure vacuum takes and door is held shut.

7. When vacuum guage reads 25, it's ready to fire.

8. Turn vacuum switch to "hold" (down position).

9. Press "fire" switch (right red switch) and hold until it fires.

10. Turn vacuum switch to vacuum to vent.

11. Repeat steps 3 – 10 for as many shots as you want to make.

12. When done, turn off everything: turn off gas tank, hold the fire switch on until pressure drops to 0, turn off chamber, turn off vacuum pump.

14. Incubate bombarded tissue overnight at room temperature.

Quantitative assay (luciferase)

1. Prepare grinding buffer [keep at -20^{o}C]:

100 mM Na Phosphate pH 7.0	1ml of 1M/10 ml
10 mM DTT	100 µl of 1M/10 ml
20 µg/ml Leupeptin	20 µl 10mg/ml stock/10 ml
20% (v/v) glycerol	4 ml 50% stock/10 ml
H_20	4.88 ml/ 10 ml

2. Prepare 2X Luciferase Assay Buffer (LAB) [keep at -20^{o}C]

60 mM Tris-SO_4 pH 7.7*	600µL 1M/ 10 mL
20 mM $MgCl_2$	200µL 1M/ 10 mL
20 mM DTT	200 µL 1M/ 10 mL
2 mM EDTA	40µL 0.5M/ 10 mL
H_20	8.95mL/ 10 mL

* Tris adjusted to pH w/ H_2S0_4

3. Prepare 100 µl/sample of 1X LAB + 1 mM luciferin (Luciferin stock- dissolve 3.18mg in 100µL H_20 for a 100 mM solution. Store at -20C). Load these into white-bottom 96-well plate.

4. Prepare 100 µl/sample of 2X LAB + 2µl of 0.1 M ATP (Assay Buffer).

5. Grind tissue in 300 µL grinding buffer with mortar and pestle.

6. Spin extract in microfuge for 10 min. at 4°C, and use it for assays (the extracts can be kept at -20°C).

7. Remove 75 µL supernatant to a tube containing 100 µl of 2X LAB+ATP.

8. Add complete sample to 100 µl of 1X LAB with Luciferin (in 96-well plates) as prepared above. Pipette up and down.

9. Visualize samples by measuring the luminescence of each for 5 seconds in a spectrophotometer.

Quantitative assay (GUS)

1. Prepare GUS assay buffer [keep at -20°C]

2.5 mM MUG	250 µL 0.1M/ 10 mL
50 mM NaPO4, pH 7.0	500 µL 1M/ 10 mL
10 mM EDTA	200 µL 0.5M/ 10 mL
10mM DTT	100 µl 1M/ 10 mL
20 µg/ml Leupeptin	20 µL stock (10 mg/mL)/ 10 mL
20% (v/v) MeOH	2 mL/ 10 mL
H_20	6.93 mL/ 10 mL

2. Add 200 µL/ sample to clear-bottom 96-well plate

3. Add 50 µL of sample to well

4. Seal plate with parafilm and incubate overnight (~20 hr) at 37°C

5. (The next day) add 50 µL of overnight reaction to 200µL 0.2M Na_2CO_3

6. Visualize samples by exciting samples at 365 nm and measuring the light emitted at 455 nm.

Work Cited

Benderoth, M., S. Textor, A. J. Windsor, T. Mitchell-Olds, J. Gershenzon, and J. Kroymann. 2006. Positive selection driving diversification in plant secondary metabolism. Proc. Natl. Acad. Sci. U. S. A. 103:9118-23.

Austin, D. F., and Z. Huaman. 1996. A Synopsis of *Ipomoea* (Convolvulaceae) in the Americas. Taxon 45:3-38.

Bradshaw, H. D., and D. W. Schemske. 2003. Allele Substitution at a Flower Colour Locus Produces a Pollinator Shift in Monkeyflowers. Nature 426:176-178.

Bradshaw, H. D., S. M. Wilbert, K. G. Otto, and D. W. Schemske. 1995. Genetic Mapping of Floral Traits Associated with Reproductive Isolation in Monkeyflowers. Nature 376:762-765.

Bruneau, A. 1997. Evolution and Homology of Bird Pollination Syndromes in *Erythrina* (Leguminosae). Amer. J. Bot. 84:54-71.

Carroll, S. B., J. K. Grenier, and S. D. Weatherbee. 2001. From DNA to Diversity: Molecular Genetics and the Evolution of Animal Design. Blackwell, Malden, Massachusettes.

Castellanos, M. C., P. Wilson, and J. D. Thomson. 2004. "Anti-Bee" and "Pro-Bird" Changes during the Evolution of Hummingbird Pollination in *Penstemon* Flowers. J. Evol. Biol. 17:876-885.

Chittka, L., and N. Waser. 1997. Why Red Flowers are not Invisible to Bees. Israel J. Plant Sci. 45:169-183.

Cronk, Q., and I. Ojeda. 2008. Bird-Pollinated Flowers in an Evolutionary and Molecular Context. J. Exper. Bot. 59:715-727.

Des Marais, D. L., and M. D. Rausher. 2008. Escape from Adaptive Conflict after Duplication in an Anthocyanin Pathway Gene. Nature 454:762-765.

DePristo, M. A., D. M. Weinreich, and D. L. Hartl. 2005. Missense meanderings in sequence space: a biophysical view of protein evolution. Nat. Rev. Genet. 6:678-87.

Faegri, K., and L. Van Der Pijl. 1966. The Principals of Pollination Ecology. Pergamon Press, New York.

Fenster, C. B., W. S. Armbruster, P. Wilson, M. R. Dudash, and J. D. Thomson. 2004. Pollination Syndromes and Floral Specialization. Annual Review of Ecology and Systematics 35:375-403.

Fischer, T. C., H. Halbwirth, B. Meisel, K. Stich, and G. Forkman. 2003. Molecular Cloning, Substrate Specificity of the Functionally Expressed Dihydroflavonol 4-Reductases from Malus domestica and Pyrus communis Cultivars and the Consequences for Flavonoid Metabolism. Arch. Biochem. and Biophys. 412:223-230.

Force, A., M. Lynch, F. B. Pickett, A. Amores, Y.-l. Yan, and J. Postlethwait. 1999. Preservation of Duplicate Genes by Complementary, Degenerative Mutations. Genetics 151:1531-1545.

Forkman, G., and W. Heller. 1999. Biosynthesis of Flavonoids. Pp. 713-748 *in* D. Barton, K. Nakanishi and O. Meth-Cohn, eds. Comprehensive Natural Products Chemistry. Elsevier, Amsterdam.

Gerats, A. G. M., P. Vlaming, M. Doodeman, B. Al, and A. W. Schram. 1982. Genetic Control of the Conversion of Dihydroflavonols into flavonols and anthocyanins in flowers of *Petunia hybrida*. Planta 155:364-368.

Goldman, N., and Z. Yang. 1994. A Codon-based Model of Nucleotide Substitution for Protein-coding Sequences. Mol. Biol. and Evol. 11:725-736.

Gould, K. S. 2004. Nature's Swiss Army Knife: The Diverse Protective Roles of Anthocyanins in Leaves. J. Biomed. Biotech. 5:314-320.

Gould, S. J. 1977. Ontogeny and Phylogeny. Belknap Press, Cambridge.

Gould, S. J. 2002. The Structure of Evolutionary Theory. Harvard University Press, Cambridge, MA.

Grotewold, E. 2006. The Genetics and Biochemistrt of Floral Pigments. Annual Review of Plant Biology 57:761-780.

Guindon, S., A. G. Rodrigo, K. A. Dyer, and J. P. Huelsenbeck. 2004. Modelling the Site-Specific Variation of Selection Patterns Along Lineages. Proc. Natl. Acad. Sci. U. S. A.

Hahn, M. W., T. De Bie, J. E. Stajich, C. Nguyen, and N. Cristiani. 2005. Estimating the Tempo and Mode of Gene Family Evolution from Comparative Genomic Data. Genome Res. 15:1153-1160.

Halbwirth, H., S. Martens, U. Wienand, G. Forkmann, and K. Stich. 2003. Biochemical Formation of Anthocyanins in Silk Tissue of *Zea mays*. Plant Sci. 164:489-495.

Haldane, J. B. S. 1933. The Part Played by Recurrent Mutation in Evolution. Am. Nat. 67:5-19.

Harborne, J. B. 1967. Comparative Biochemistry of the Flavonoids. Academic Press, New York.

Hittinger, C. T., and S. B. Carroll. 2007. Gene Duplication and the Adaptive Evolution of a Classic Genetic Switch. Nature 449:677-681.

Hoekstra, H. E., and J. A. Coyne. 2007. The Locus of Evolution: Evo Devo and the Genetics of Adaptation. Evolution 61:995-1016.

Holm, S. 1979. A Simple Sequentially Rejective Multiple Test Procedure. Scand J. Statistics 6:65-70.

Hughes, A. L. 1994. The Evolution of Functionally Novel Proteins After Gene Duplication. Proc. Royal Soc. London, Ser. B 256:119-124.

Inagaki, Y., Y. Johzuka-Hisatomi, T. Mori, S. Takahashi, Y. Hayakawa, S. Peyachoknagul, Y. Ozeki, and S. Iida. 1999. Genomic Organization of the Genes Encoding Dihydroflavonol 4-reductase for Flower Pigmentation in the Japanese and Common Morning Glories. Gene 226:181-188.

Islam, M. S., M. Yoshimoto, N. Terahara, and O. Yamakawa. 2002. Anthocyanin Compositions in Sweetpotato (*Ipomoea batatas* L.) Leaves. Bioscience, Biotechnology and Biochemistry 66:2483-2486.

Iwashina, T. 2000. The Structure and Distribution of the Flavonoids in Plants. J. Plant Res. 113:287-299.

Johnson, E. T., S. Ryu, H. Yi, B. Shin, H. Cheong, and G. Choi. 2001. Alteration of a Single Amino Acid Changes the Substrate Specificity of Dihydroflavonol 4-Reductase. Plant J. 25: 325-333.

Kay, K. M., P. A. Reeves, R. G. Olmstead, and D. W. Schemske. 2005. Rapid Speciation and the Evolution of Hummingbird Pollination in Neotropical *Costus* subgenus *costus* (Costaceae): Evidence from nrDNA ITS and ETS sequences. Amer. J. Bot. 92:1899-1910.

Li, W.-Y. 1997. Molecular Evolution. Sinauer, Sunderland, Massachusetts

Li, W.-Y., T. Gojobori and M. Nei. Pseudogenes as a Paradigm of Neutral Evolution. Nature 292: 237-239.

Lynch, M., and J. S. Conery. 2000. The Evolutionary Fate and Consequences of Duplicate Genes. Science 290:1151-1155.

Lynch, M., and A. Force. 2000. The Probability of Duplicate Gene Preservation by Subfunctionalization. Genetics 154:459-473.

Lynch, M., M. O'Hely, B. Walsh, and A. Force. 2001. The Probability of Preservation of a Newly Arisen Gene Duplicate. Genetics 159:1789-1804.

Maddison, W. P., and D. R. Maddison. 2000. MacClade 4. Sinauer, Sunderland, Massachusetts.

Maddison, W. P., and D. R. Maddison. 2007. Mesquite: A Modular System for Evolutionary Anlaysis.

Martens, S., T. Teeri, and G. Forkmann. 2002. Heterologous expression of dihydroflavonol 4-reductases from various plants. FEBS Lett 531:453-458.

McDonald, A. 1991. Origin and Diversity of Mexican Convolvulaceae. Anales del Instituto de Biologia, Universidad Nacional Autonoma de Mexico, Serie Botanica 62:65-82.

Michelmore, R. W., and B. C. Meyers. 1998. Clusters of Resistance in Plants Evolve by Divergent Selection and a Birth-and-Death Process. Genome Research 8:1113-1130.

Miller, R. E., J. A. McDonald, and P. S. Manos. 2004. Systematics of Ipomoea subgenus Quamoclit (Convolvulaceae) based on ITS Sequence Data and a Bayesian Phylogenetic Analysis. Amer. J. Bot. 91:1208-1218.

Morita, Y., M. Saitoh, A. Hoshino, E. Nitasaka, and S. Iida. 2006. Isolation of cDNAs for R2R3-MYB, bHLH and WDR Transcriptional Regulators and Identification of *c* an *ca* Mutations Conferring White Flowers in the Japanese Morning Glory. Plant Cell Physiol. 47:457-470.

Murashige, T., and F. Skoog. 1962. A Revised Medium for Rapid Growth and Bioassays with Tobacco Tissue Cultures. Physiol. Plant 15:473-497.

Nicholson, R. L., S. S. Kollipara, J. R. Vincent, P. C. Lyons, and G. Cadena-Gomez. 1987. Phytoalexin Synthesis by the Sorghum Mesocotyl in Response to Infection by Pathogenic and Nonpathogenic Fungi. Proc. Natl. Acad. Sci. U. S. A. 84:5520-5524.

Ohno, S. 1970. Evolution by Gene Duplication. Springer-Verlag, Berlin.

Orr, H. A. 2005. The Genetic Theory of Adaptation: A Brief History. Nature Reviews Genetics 6:119-127.

Pagel, M. 1999. The Maximum Likelihood Approach to Reconstructing Ancestral Character States of Discrete Characters on Phylogenies. Syst. Biol. 48:612-622.

Park, K.-I., N. Ishikawa, Y. Morita, J.-D. Choi, A. Hoshino, and S. Iida. 2007. A bHLH Regulatory Gene in the Common Morning Glory, Ipomoea purpurea, Controls Anthocyanin Biosynthesis in Flowers, Proanthocyanidin and Phytomelanin Pigmentation in Seeds, and Seed Trichome Formation. Plant J. 49:641-654.

Petit, P., T. Granier, B. L. d'Estaintot, C. Manigand, K. Bathany, J.-M. Schmitter, V. Lauvergeat, S. Hamdi, and B. Gallois. 2007. Crystal Structure of Grape Dihydroflavonol 4-Reductase, A Key Enzyme in Flavonoid Biosynthesis. J. Mol. Biol. 368:1345-1357.

Piatigorsky, J., and G. Wistow. 1990. The Recruitment of Crystallins: New Functions Precede Gene Duplication. Science 252:1078-1079.

Punyasiri, P. A. N., I. S. B. Abeysinghe, V. Kumar, D. Treutter, D. Duy, C. Gosch, S. Martens, G. Forkmann, and T. C. Fischer. 2004. Flavonoid Biosynthesis in the Tea Plant Camellia sinensis: Properties of Enzymes of the Prominent Epicatechin and Catechin Pathways. Arch. Biochem. and Biophys. 431:22-30.

Quattrocchio, F., A. Baudry, L. Lepiniec, and E. Grotewold. 2006. The Regulation of Flavonoids. Pp. 97-122 *in* E. Grotewold, ed. The Science of Flavonoids. Springer, New York.

Quattrocchio, F., J. Wing, K. van der Woude, J. N. M. Mol, and R. Koes. 1998. Analysis of bHLH and MYB Domain Proteins: Species-Specific Regulatory Differences are Caused by Divergent Evolution of Target Anthocyanin Genes. Plant J. 16:475-488.

Quattrocchio, F., J. Wing, K. van der Woude, E. Souer, N. de Vetten, J. Mol, and R. Koes. 1999. Molecular Analysis of the *anthocyanin2* Gene of Petunia and Its Role in the Evolution of Flower Color. Plant Cell 11:1433-1444.

Quattrocchio, F., J. F. Wing, H. T. C. Leppen, J. N. M. Mol, and R. E. Koes. 1993. Regulatory Genes Controlling Anthocyanin Pigmentation Are Functionally Conserved amon Plant Species and Have Distinct Sets of Target Genes. Plant Cell 5:1497-1512.

Rodriguez-Girones, M. A., and L. Santamaria. 2004. Why Are So Many Bird Pollinated Flowers Red? PLoS Biol. 2:1515-1519.

Ronquist, F., and J. Huelsenbeck. 2003. MrBayes 3: Bayesian Phylogenetic Inference Under Mixed Models. Bioinformatics 19:1572-1574.

Sambrook, J., and D. W. Russell. 2001. Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Schwartz-Sommer, Z., N. Shepherd, E. Tacke, A. Gierl, W. Rohde, L. Leclerq, M. Mattes, R. Berndtgen, P. Peterson, and H. Saedler. 1987. Influence of Transposable Elements on the Structure and Function of the *A1* Gene of *Zea mays*. EMBO J. 6:287-294.

Schwinn, K., J. Venail, Y. Shang, S. Mackay, V. Alm, E. Butelli, R. Oyama, P. Bailey, K. Davies, and C. Martin. 2006. A small family of MYB-regulatory genes controls floral pigmentation intensity and patterning in the genus Antirrhinum. Plant Cell 18:831-51.

Shirley, B. W. 1996. Flavonoid Biosynthesis: 'New' Functions for an 'Old' Pathway. Trends Plant Sci. 1:377-382.

Simpson, G. G. 1953. The Major Features of Evolution. Simon and Schuster, New York.

Stafford, H. A. 1965. Flavonoids and Related Phenolic Compounds Produced in the First Internode of *Sorghum vulgare* Pers. in Darkness and in Light. Plant Physiol. 40:130-138.

Stafford, H. A., and H. H. Lester. 1982. Enzymatic and Nonenzymatic Reduction of (+)-dihydroquercetin to its 3.4,-diol. Plant Physiol. 70:695-698.

Stefanovic, S., L. Krueger, and R. G. Olmstead. 2002. Monophyly of the Convolvulaceae and Circumscription of their Major Lineages Based on DNA Sequences of Multiple Chloroplast Loci. Amer. J. Bot. 89:1510-1522.

Stern, D. L. 2000. Perspective: Evolutionary Developmental Biology and the Problem of Variation. Evolution 54:1079-1091.

Swofford, D. L. 2000. PAUP*. Phylogenetic Analysis Using Parsimony (*and Other Methods). Sinauer Associates, Sunderland, Massachusettes.

Thomson, J. D., and P. Wilson. 2008. Explaining Evolutionary Shifts Between Bee and Hummingbird Pollination: Convergence, Divergence and Directionality. International J. Plant Sci. 169:23-38.

Thomson, J. D., P. Wilson, M. Valenzuela, and M. Malzone. 2000. Pollen Presentation and Pollination Syndromes, with Special Reference to *Penstemon*. Plant Species Biol. 15:11-29.

Tiffin, P., R. E. Miller, and M. D. Rausher. 1998. Control of Expression Patterns of Anthocyanin Structural Genes by Two Loci in the Common Morning Glory. Genes and Genetic Systems 73:105-110.

Whibley, A. C., N. B. Langlade, C. Andalo, A. I. Hanna, A. Bangham, C. Thebaud, and E. Coen. 2006. Evolutionary Paths Underlying Flower Color Variation in *Antirrhinum*. Science 313:963-967.

Whittall, J., C. Voelckel, and S. A. Hodges. 2006. Convergence, constraint and the role of gene expression during adaptive radiation: Floral anthocyanins in *Aquilegia*. Mol. Ecol. 15:4645-1657.

Whittall, J. B., and S. A. Hodges. 2007. Pollinator Shifts Drive Increasingly Long Nectar Spurs in Columbine Flowers. Nature 444:706-710.

Wilson, P., A. D. Wolfe, W. S. Armbruster, and J. D. Thomson. 2007. Constrained Lability in Floral Evolution: Counting Convergent Origins of Humminbird Pollination in *Penstemon* and *Keckiella*. New Phytologist 176:883-890.

Wolfe, L. M., and D. R. Sowell. 2006. Do Pollination Syndromes Partition the Pollinator Community? A Test Using Four Sympatric Morning Glory Species. International J. Plant Sci. 167:1169-1175.

Wray, G. A. 2007. The Evolutionary Significance of cis-Regulatory Mutatons. Nature Reviews Genetics 8:206-216.

Yang, Z. 1997. PAML, A Program for Phylogenetic Analysis by Maximum Likelihood. CABIOS 13:555-556.

Yang, Z. 1998. Likelihood Ratio Tests for Detecting Positive Selection and Application to Primate Lysozyme Evolution. Mol. Biol. Evol. 15:568-573.

Yang, Z., W. S. Wong, and R. Nielsen. 2005. Bayes empirical bayes inference of amino acid sites under positive selection. Mol. Biol. Evol. 22:1107-18.

Zhang, J., R. Nielsen, and Z. Yang. 2005. Evaluation of an Improved Branch-Site Likelihood Method for Detecting Positive Selection at the Molecular Level. Mol. Biol. Evol. 22:2472-2479.

Zhang, J., Y.-P. Zhang, and H. F. Rosenberg. 2002. Adaptive Evolution of a Duplicated Panreatic Ribonuclease Gene in a Leaf-Eating Monkey. Nat. Genet. 30:411-415.

Zufall, R. 2003. Evolution of Red Flowers in *Ipomoea*. Biology. Duke University, Durham, North Carolina.

Zufall, R. A., and M. D. Rausher. 2004. Genetic Changes Associated with Floral Adaptation Restrict Future Evolutionary Potential. Nature 428:847-850.

Zwickl, D. J. 2006. Genetic Algorithm Approaches for the Phylogenetic Analysis of Large Biological Sequence Datasets Under the Maximum Likelihood Criterion. The University of Texas at Austin.

Biography

David Lee Des Marais

Born March 22, 1978 in Stanford, California

Education

Ph. D., Biology, 2008: Duke University, Durham, North Carolina

B. A., Integrative Biology, 1996: University of California at Berkeley

Publications

Des Marais, David L. and Mark D. Rausher. 2008. Escape from Adaptive Conflict after Duplication in an Anthocyanin Pathway Gene. Nature 454: 762-765.

Sweeney, Alison, David L. Des Marais, Yih-En Andrew Ban and Sonke Johnsen. 2007. Evolution of a Graded Refractive Index in Squid Lens. Journal of the Royal Society Interface 4: 685-698.

Des Marais, David L., Alan R. Smith Donald M. Britton and Kathleen M. Pryer. 2003. Phylogenetic Relationships and Evolution of Extant Horsetails, *Equisetum*, based on Chloroplast DNA Sequence Data (*trnL-F* and *rbcL*). International Journal of Plant Science 164 (5): 737-751.

Grants, Fellowships and Academic Honors

2007 Walter Fitch Prize, Society for Molecular Biology and Evolution.

2004 National Science Foundation Doctoral Dissertation Improvement Grant

2002 NSF Deep Gene Student Summer Grant

2002 NSF Graduate Research Fellowship Honorable Mention

2001 Duke University – University Genetics Fellowship

2001 Botanical Society of America Young Botanist Award

2001 UC Berkeley Integrative Biology Joseph Le Conte Award

2000 California Native Plant Society Myrtle Wolf Award