Evolution of novel genetic programs: insights from flower and fruit development

By

Aniket Sengupta
M.Sc., University of Delhi, 2011
B.Sc., University of Calcutta, 2009

Submitted to the graduate degree program in Ecology and Evolutionary Biology and the Graduate Faculty of the University of Kansas in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

Chair: Lena C. Hileman

Paulyn Cartwright

Mark Mort

Wm Leo Smith

Matthew Buechner

Date Defended: 02-July-2019

The dissertation committee for Aniket Sengupta certifies that this is the approved version of the following dissertation:

Evolution of novel genetic programs: insights from flower and fruit development

Chair: Lena Hileman

Date Approved: 02-July-2019

Abstract

An outstanding question in evolutionary biology is how genetic programs (interaction between multiple genes or their products) that define novel phenotypes evolve. There are two major ways such genetic programs that define novel phenotypes can evolve. First, by a *de novo* assembly of previously non-interacting genes during the origin of the novel phenotype. Secondly, by co-option (re-deployment, re-recruitment) of existing programs from other functions towards defining the novel phenotype.

Monosymmetry of flowers is a novel phenotype that has evolved at least 130 times from polysymmetric flowers during the diversification of flowering plants. Flower monosymmetry in the order Lamiales is defined by an interaction of CYCLOIDEA, RADIALIS, DIVARICATA, and DIVARICATA and RADIALIS interacting factors (CYC–RAD–DIV–DRIF). This interaction is best understood in the Lamiales species *Antirrhinum majus*. The evolutionary history of the CYC–RAD–DIV–DRIF genetic program that defines Lamiales flower monosymmetry is unclear. It is an open question whether this genetic program was assembled *de novo* near the base of Lamiales during the evolution of flower monosymmetry or it was co-opted from a different function.

We find evidence that the CYC–RAD–DIV–DRIF genetic program, which is crucial for defining the novel phenotype of flower monosymmetry in Lamiales, was likely co-opted from a different function, possibly fruit and ovary development. We come to this conclusion through a comparative analysis between representative taxa from the Lamiales and its close relative the Solanales. This suggests that the evolution of flower monosymmetry in Lamiales may have been facilitated by the availability of the CYC–RAD–DIV–DRIF genetic program. This genetic program was likely co-opted in defining flower monosymmetry in Lamiales for the following

reasons. This program was ancestrally involved in regulating cell size: and this function makes it a likely candidate for defining a novel phenotype that has variously shaped petals and aborted stamens. Also, the CYC–RAD–DIV–DRIF genetic program functions by a competitive interaction between RAD and DIV proteins—the two competing proteins could easily be co-opted in defining the two morphologically distinct regions of the monosymmetric flowers.

Little is known about the regulators that affect the transcription of *CYC* in Lamiales. We find predictive, bioinformatics-based evidence for a Lamiales-specific transcriptional autoregulation of *CYC*, suggesting that the evolution of flower monosymmetry in this lineage was associated with an autoregulation-mediated sustained, stable, and high transcription of *CYC*. Our data elucidates the evolutionary origin of the CYC–RAD–DIV–DRIF genetic program that defines flower monosymmetry in Lamiales. This interaction was likely co-opted *en bloc* from fruit and ovary development, with additional lineage-specific changes in gene regulation leading to transcriptional autoregulation of *CYC*.

Entangled in this question is the fact that ovaries and fruits of many species in the tribe Antirrhineae (order Lamiales), including *Antirrhinum majus*, are monosymmetric—with unequal dorsal and ventral locules. We determine the genetic and micromorphological basis of this phenotype and also estimate its evolutionary history. We identify at least five evolutionary transitions from polysymmetric to monosymmetric ovaries, and at least seven reversals to polysymmetry across the tribe. Ovary monosymmetry in *Antirrhinum* and its closest relatives is likely controlled non-cell autonomously by a CYC–RAD interaction. CYC upregulates *RAD* expression in dorsal petals and stamens in early stages of development, and RAD protein from these tissues migrates to the dorsal locule to increases cell proliferation in dorsal ovary wall, causing the ovary to be monosymmetric.

Acknowledgments

This work stems from many years of effort, which would have been demanding, if not for the academic and social support of many others. I want to thank my thesis adviser Dr. Lena Hileman for shaping me into the scientist I am today, for cherishing diversity in her lab, and for working towards ensuring that I get a salary during the summer months. This work would not have been possible without her advice and efforts. I also want to acknowledge other members of my advisory committee Dr. Paulyn Cartwright, Dr. Mark Mort, Dr. Wm. Leo Smith, and Dr. Matthew Buechner for helping me keep track of my progress, and for their help in brainstorming my thesis chapters when experiments failed.

I acknowledge the help and encouragement Aagje Ashe, Graduate Coordinator, and Dorothy Johanning, Program Assistant, have provided to me over the years. Other than keeping the department running behind the scenes, they were always happy to explain the departmental policies without the usual American-terminology to an international student. In the same line, I want to acknowledge Katie Sadler, Greenhouse Manager (University of Kansas) for her unparalleled efforts in keeping plants healthy and growing, irrespective of weather, and of the availability of resources.

Science is a labor of love, but it benefits immensely from a stable salary. This research would have been impossible without financial support towards personal expenses in the form of research fellowships and scholarships from the following: my adviser Dr. Lena Hileman, The Botany Endowment (University of Kansas), Dr. Joy Ward (University of Kansas). The following sources helped me fund travel to conferences: The Botany Endowment (University of Kansas), The Botany Conference (Botanical Society of America), and the departmental Graduate Student Organization (EEB-GSO, University of Kansas). The Botany Endowment, the Office of

Graduate Studies, and the EEB-GSO at the University of Kansas generously supported my travel to a conference to present a historical research unrelated to my thesis.

I worked as a teaching assistant with Dr. Chris Haufler and Dr. Bob Timm for the course Biology 413: History and Diversity of Organisms. They encouraged me to experiment with my pedagogic methods and helped me become a better instructor. Teaching this course was one of the most enjoyable experiences in graduate school.

My co-workers in the Hileman Lab have been a constant source of support. I want to thank Amanda Katzer and Vibhuti Singh for motivating me. No less important are the undergraduate researchers I had the privilege to work with—Nick Nolte, Jesse Kaighin, and Sukhindervir Sandhu. Other graduate students and post-doctoral researchers in the department were a steady source of scientific and social advice. Aleah Henderson, James Fischer, Hannah Kinmonth-Schultz (and family), Keely Brown, Lucas Hemmer, and Ahmed Alnazi were always there when I needed them.

There are many in Lawrence who helped me feel at home several thousand miles away from the place I was raised. The list is long and in no particular order: the Sandal family, the Lawson family, the Peters family, the Clark family, the Nolte Family, Don and family, Indranuja and family, Sacipriya and family, Cory and Tate Betswick, The Jacob House, Tyler Flory and family, Natalya Lowther, and Ryan Whitesell. They were a source of comfort when I needed it, and opened their homes and hearts to an international student during weekends and holidays. My friends at the Aimee's Coffeehouse, Lawrence—that provided both respite after a long week of work, and a relaxing place to write parts of this thesis—deserve my gratitude. I thank patrons and baristas at the coffeehouse—Paige, Sydney, Bert Haverkate-Ens, Arielle Jacobs, Hunter, and Brandon. Masanori Mac Shiomi has been a constant, reliable friend for all these years and has

found words of encouragement when I needed them the most. Nate Morsches, I thank, for his mentorship, and for his affectionate friendship.

I am grateful to my friends who are located outside the US. They never let the distance, the time difference, or the cost of internet in the developing world keep them from reaching out. Sreeja Bhattacharya, my childhood friend in India, I thank for going through the trials and triumphs of graduate school with me, even if it was over the phone. Krati Vikram Singh, a friend from my Masters program in India, for giving me well thought advice and encouragement. I thank my friends in Argentina (Mauricio Vinassa and Yesica Mungo) and in Afghanistan (the Vedi Family) for their advice on, well, everything.

I have had several major medical concerns throughout my time as a doctoral researcher. My medical care providers at the Student Health Services, University of Kansas, supported me through this strenuous period—beyond what was required by their professions. They are as follows: Dr. Robert Brown (primary care provider), Dr. Pavika Saripalli (primary care provider), Dr. Keith Floyd (psychologist), and Evan Riordan (pharmacist).

I have always been botanically inclined. My first graffiti on my parents' wall was a boy and a flowering plant (*Hibiscus*?). My earliest memories are of being filled with awe by mosses and algae. I learnt by heart the outline of the Linnaean Hierarchy of the Vegetable Kingdom as a teenager. I was a passionate amateur. The professors at the Department of Botany, Presidency College, Kolkata, transformed me into botanist—Dr. Tarit Kumar Sadhu, Dr. Sumita Ghosh, and Dr. Samarendra Nath Ghosh. My professors in my Masters program at the University of Delhi streamlined me towards research in botanical diversity—Dr. R. Geeta, Dr. P. Uniyal, Dr. A. K. Pandey. My interest in evo-devo, and curiosity about research in the West were stoked, and supported by Dr. R. Geeta, my Masters Thesis adviser. My teachers in grade school encouraged

me to be an orator, and helped me become a better public speaker despite my speech concerns. I still reap the benefits of their encouragement when I present science at meetings. It might have been decades ago, but I still hold the same affection and respect for them—Mrs. Suchitra Sengupta and Mrs. Smriti Choudhury.

Last, but never the least, my want to express my deep gratitude towards my family, especially my mother. Without her support and sacrifices, I wouldn't have made it this far. My appreciation for her cannot be fully expressed in language. To her the extent of my indebtedness is immense, and heart-felt.

Table of contents

Introduction

Novel phenotypes are a recurring feature across the tree of life. Interestingly, novel phenotypes do not evolve by utilizing new genes, but evolve by co-opting existing genes (for example, Bharathan et al., 2002; Busch and Zachgo, 2007; Citerne et al., 2013; Hay and Tsiantis, 2010; Panganiban et al., 1997; Spitz et al., 2001, 2; Stern, 2013; Werner et al., 2010; Jiggins et al., 2017). However, genes do not usually function in isolation but interact with other genes or gene products to affect phenotype. Hence, it is likely that the genes co-opted towards defining novel phenotypes have been a part of genetic regulatory program or interaction in the ancestral species. Whether, during co-option of existing genes to novel phenotypes, these genes were co-opted in isolation or along with the entire genetic program is an open question is biology. There have been few studies addressing this question, and mostly in animal systems (for example, Hay and Tsiantis, 2010; Jiggins et al., 2017; Panganiban et al., 1997).

Monosymmetry of flowers is a novel phenotype that has evolved at least a 130 times from polysymmetric flowers during the diversification of flower plants (Reyes et al., 2016). All independently acquired monosymmetric flowers, at least the ones that have been tested at a molecular level, are defined by or have been shown to be associated with expression of *CYC* genes. Genetic basis of the flower monosymmetry is best understood in the order Lamiales that includes the species *Antirrhinum majus* (snapdragon). In Lamiales, flower monosymmetry is defined by a genetic program CYC–RAD–DIV–DRIF (**Figure 1**). Whether CYC was co-opted in isolation to this novel phenotype or whether the entire CYC–RAD–DIV–DRIF program was co-opted towards defining Lamiales flower monosymmetry is an open question. Interestingly, a RAD–DIV–DRIF program is known from *Solanum lycopersicum* (tomato, order Solanales),

where RAD suppresses cell expansion in fruit wall (Machemer et al., 2011). Solanales are close relatives of Lamiales (Stull et al., 2015), and tomato has putatively ancestrally polysymmetric flowers. This makes tomato an ideal outgroup to test the origin of the CYC–RAD–DIV–DRIF that defines flower monosymmetry in Lamiales. It is not known whether the RAD–DIV–DRIF interaction in tomato fruits involves CYC as well (like it does in snapdragon flower monosymmetry). Also, it is not known whether CYC–RAD–DIV–DRIF interaction or at least a RAD–DIV–DRIF interaction is present in Lamiales. The predictions under a hypothesis that a CYC–RAD–DIV–DRIF genetic program was present in fruits of the common ancestor of Lamiales and Solanales and was later co-opted *en bloc* towards defining flower monosymmetry in Lamiales are summarized in **Figure 1**.

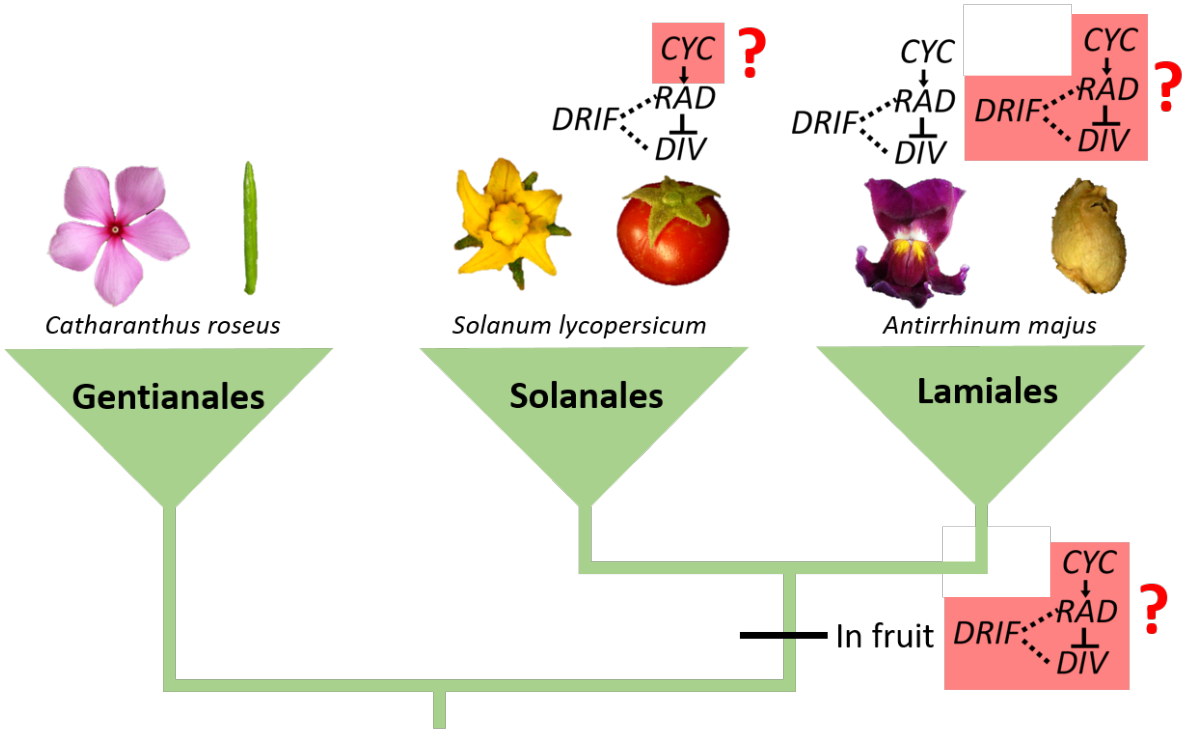

Figure 1. A model explaining the co-option of *CYC*, *RAD*, *DIV*, and *DRIF* genes in defining flower monosymmetry in Lamiales. Molecular interactions in red boxes are predictions under the hypothesis that a CYC–RAD–DIV–DRIF genetic program was present in fruits of the common ancestor of Lamiales and Solanales and was later co-opted *en bloc* towards defining flower monosymmetry in Lamiales. Molecular interactions in white background have been previously reported. *Catharanthus roseus* images by Arria Belli and SAplants (Wikimedia Commons).

Changes in gene regulation can be associated with the origin of novel phenotypes (for example, Espley et al., 2009; Tao et al., 2012). Very little is known about whether changes in regulation of *CYC* is associated with the evolution of flower monosymmetry in Lamiales. We expand and bioinformatically test the hypothesis (Yang et al., 2012) that the evolution of flower monosymmetry in Lamiales is associated with the origin of transcriptional autoregulation of *CYC*. Transcription factors that define novel and crucial phenotypes can often regulate their own transcription, and such autoregulation can be mediated by the presence of multiple autoregulatory sites in the *cis*-regulatory regions of these genes to which the protein products can bind (reviewed in Sengupta and Hileman, 2018). We bioinformatically test whether the origin of flower monosymmetry in Lamiales is likely associated with such an enrichment of predicted autoregulatory sites in the putative *cis*-regulatory regions of Lamiales *CYC* genes to which the CYC transcription factors may bind and regulate their own transcription.

Embedded in the question whether a CYC–RAD–DIV–DRIF program us present in *Antirrhinum* fruits (**Figure 1**) is the fact the fruits of this species are monosymmetric. We estimate the dynamic phylogenetic history of this phenotype in the tribe Antirrhineae. We determine the genetic and micromorphological basis of this phenotype.

Chapter 1: Testing the hypothesis that a flower symmetry developmental module evolved via recruitment from an ancestral role in fruit development

Summary

The petal, stamen, and ovary whorls of *Antirrhinum majus* (snapdragon, order Lamiales) flowers are monosymmetric. Previous work demonstrated that the following gene products control petal and stamen whorl monosymmetry—CYCLOIDEA (CYC), DICHOTOMA (DICH), RADIALIS (RAD), DIVARICATA (DIV), and DIV and RAD Interacting Factor (DRIF) 1 & 2. However, a direct link between these genes and ovary symmetry has not been made. Through characterization of existing mutants, we show that petal/stamen whorl symmetry genes control monosymmetry of the snapdragon ovary. This likely occurs through non-cell autonomous RAD function that leads to a wider pericarp on the dorsal side of the ovary relative to the pericarp on the ventral side. This difference is mediated by promotion of cell proliferation in the dorsal pericarp at early stages of ovary differentiation. Evolutionary changes to non-cell autonomous RAD function may underlie the dynamic history of ovary symmetry in the tribe Antirrhineae, to which snapdragon belongs. We identify at least five evolutionary transitions from polysymmetric to monosymmetric ovaries, and at least seven reversals to polysymmetry across the tribe, all in the conserved context of petal and stamen whorl monosymmetry. In addition, we identified a novel peak in *RAD* expression late in *A.majus* ovary development. This late expression is independent of positive regulators of *RAD* in the petal and stamen whorls. We find this peak of *RAD* expression late in ovary development in an early and a late diverging member of the tribe Antirrhineae, suggesting that this pattern is conserved across the tribe. Interestingly, an ortholog of RAD in *Solanum lycopersicum* (tomatoes, order Solanales) similarly suppresses fruit wall cell

expansion. Integrating our results with previous findings in *S. lycopersicum*, we hypothesize that RAD function, and a RAD-DIV-DRIF interaction in ovary development, pre-dates recruitment of this module to a role in petal and stamen whorl monosymmetry.

Introduction

Antirrhinum majus (snapdragon; hereafter *Antirrhinum*) flowers are highly specialized for bee pollination, and this specialization includes flower monosymmetry (bilateral symmetry; zygomorphy; **Figure 2**, **Figure 3**). Each flower has three distinct morphological regions that define monosymmetry across all four flower whorls: first whorl sepals, second whorl petals, third whorl stamens and inner whorl carpels (ovary). These three distinct regions are the dorsal (top; adaxial), ventral (bottom; abaxial), and lateral sides of the flower (**Figure 2**, **Figure 3).**

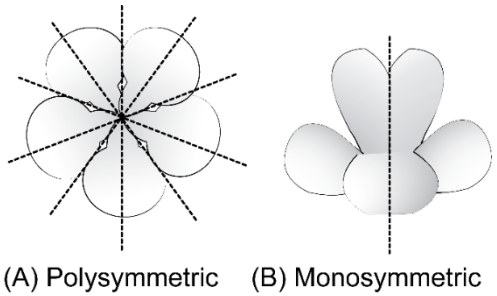

Figure 2. Major types of flower symmetry shown with hypothetical flowers. **A.** Polysymmetric (radially symmetric, actinomorphic) flower. **B.** Monosymmetric (bilaterally symmetric, zygomorphic) flower.

Monosymmetry of snapdragon flowers along the dorso-ventral axis is defined by a competitive interaction involving TCP (TEOSINTE BRANCHED1, CYCLOIDEA, and PROLIFERATING CELL FACTORS) and MYB (first described from an avian myeloblastosis virus) transcription factors. TCP and MYB proteins are found as large gene families in flowering plants (Martín-Trillo and Cubas, 2010; Yanhui et al., 2006) and play diverse roles beyond flower symmetry patterning, including aspects of vegetative and reproductive development (Ambawat et al., 2013; Martín-Trillo and Cubas, 2010; Parapunova et al., 2014).

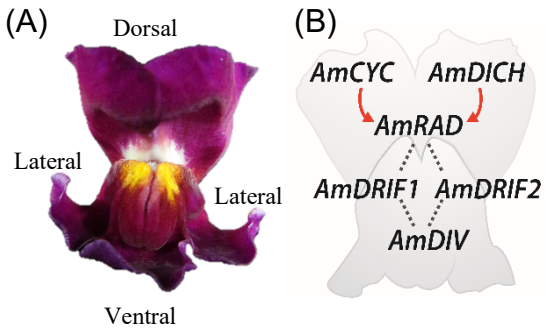

Figure 3. *Antirrhinum* flower monosymmetry. **A.** *Antirrhinum* flower in face-view showing morphological differentiation along the dorso-ventral axis (line JI-7 from John Innes Centre, UK). **B.** Regulatory mechanisms involved in *Antirrhinum* flower symmetry development. Positive regulation (arrows), and protein-protein interactions (dotted lines) are shown.

The dorsal side of *Antirrhinum* flowers consists of the dorsal sepal, dorsal halves of the lateral sepals, the dorsal petals, the dorsal halves of the lateral petals, and the dorsal sterile stamen (staminodium). The identity of dorsal organs in the petal and stamen whorls is defined by the combined action of two recently duplicated TCP paralogs, CYCLOIDEA (*Am*CYC) and DICHOTOMA (*Am*DICH) (Corley et al., 2005; Hileman and Baum, 2003; Luo et al., 1996, 1999). These two transcription factors define dorsal flower morphology partly by activating the transcription of a downstream MYB gene, *RADIALIS* (*AmRAD*; **Figure 3**) (Corley et al., 2005). *Am*RAD protein competes with another MYB protein, DIVARICATA (*Am*DIV) that defines ventral petal and stamen whorl morphology. Through this negative interaction, *Am*RAD excludes the ventral flower identity specified by *Am*DIV from the dorsal side of the developing snapdragon flower (**Figure 3**). Specifically, *Am*RAD and *Am*DIV compete for interaction with two other MYB-family protein partners called DIV and RAD Interacting Factors 1 and 2 (*Am*DRIF1 and *Am*DRIF2) (Almeida et al., 1997; Corley et al., 2005; Galego and Almeida, 2002a; Raimundo et al., 2013)(**Figure 3**). *Am*DIV requires protein-protein interaction with *Am*DRIF1&2 to function as a transcription factor to regulate downstream targets (**Figure 3**) (Perez-Rodriguez et al., 2005; Raimundo et al., 2013). In the dorsal flower domain, *Am*RAD outcompetes *Am*DIV for interaction with *Am*DRIF1&2, *Am*DIV function (Raimundo et al., 2013). Interestingly, *AmRAD* mutants also display an additional mutant phenotype in the lateral petals, though *AmRAD* mRNA is not detectable there (Corley et al., 2005). This suggests that

*Am*RAD protein is involved in non-cell autonomous control of lateral petal identity, in which the small *Am*RAD protein migrates from the dorsal organs to the neighboring lateral petals (Corley et al., 2005).

Evidence strongly supports the hypothesis that *CYC, RAD, DIV*, and *DRIF* genes and protein interactions are conserved in specifying monosymmetric flower development from a common ancestor early in the diversification of Lamiales (the order to which *Antirrhinum* belongs)(Citerne et al., 2000, 2000; Corley et al., 2005; Gao et al., 2008, 2008; Liu et al., 2013; Luo et al., 1996, 1999; Preston et al., 2011, 2014; Raimundo et al., 2013; Su et al., 2017; Yang et al., 2010, 2012; Zhong and Kellogg, 2015b, 2015a; Zhou et al., 2008). This is not surprising. Flower monosymmetry is homologous across Lamiales, derived from a monosymmetric ancestor early in Lamiales diversification although there have been multiple transitions back to polysymmetry in derived Lamiales lineages (Reeves and Olmstead, 2003; Reyes et al., 2016). Therefore, a complex set of regulatory interactions appears to have evolved early in Lamiales diversification to specify monosymmetry flower development.

An outstanding question in the evolution of development is how genetic networks, similar to the *CYC*-dependent program controlling *Antirrhinum* flower symmetry*,* evolve to shape complex, derived phenotypes. Existing genes are often recruited (re-deployed; co-opted) to define new phenotypes (Stern, 2013; True and Carroll, 2002). Co-option of single genes in defining novel phenotypes has been tested in a diversity of organisms, including re-deployment of *CYC* orthologs for independent transitions to flower monosymmetry across flowering plants (Bharathan et al., 2002; Busch and Zachgo, 2007; Citerne et al., 2013; Spitz et al., 2001, 2; Stern, 2013; Werner et al., 2010). However, the question whether such genes are re-deployed in isolation or the entire corresponding networks are re-deployed collectively has received limited

attention, and mostly in animal systems (Bharathan et al., 2002; Brakefield et al., 1996; Sordino et al., 1995; True and Carroll, 2002). Whether the *CYC*-dependent program was assembled *de novo* at the base of Lamiales or was recruited to a role in flower monosymmetry as a pre-assembled unit remains unknown. If the *CYC*-based network was recruited as a pre-assembled unit, this would constitute evidence that transitions to bilateral flower symmetry are facilitated by the presence of an ancestral genetic network that can be re-deployed *en bloc* to a novel role in flower development.

Solanales are the sister order to Lamiales + Vahliaceae (Stull et al., 2015) and primarily develop polysymmetric flowers (**Figure 1**). Compelling data from studies in the Solanales model species, *Solanum lycopersicum* (tomato), suggest that a RAD-DIV-DRIF module plays a role in tomato fruit development by modulating cell size (Machemer et al., 2011). The RAD component, *Sl*RADlike4 (or fruit SANT/MYB-like 1, FSM1), is an ortholog of *Am*RAD (Gao et al., 2017; Sengupta and Hileman, 2018). *SlRADlike4* suppresses cell expansion in the developing tomato fruit walls (pericarp). The DIV component, *Sl*DIVlike5 (or MYBI) is not an ortholog, but a paralog of *Am*DIV (Gao et al., 2017; Sengupta and Hileman, 2018). Similarly, the DRIF component, Fruit SANT/MYB Binding protein1 (FSB1) is also not an ortholog, but a paralog of *Am*DRIF1&2 (Raimundo et al., 2013). The surprising similarity of this three-component regulatory interaction raises the possibility that the common ancestor of Lamiales and Solanales utilized a RAD-DIV-DRIF module to regulate ovary/fruit development and that this module was re-deployed *en bloc* to a role patterning flower monosymmetry during Lamiales diversification.

In this study, we test whether expression of *Antirrhinum* flower symmetry genes are associated with later stages of ovary/fruit similar to tomato. In doing so, we identified a novel peak of symmetrical *AmRAD* expression late in *Antirrhinum* ovary development. This expression

overlaps with broad patterns of *AmDIV/DIVlike* and *AmDRIF* expression, but not *AmCYC* or *AmDICH* expression, providing support for the hypothesis that the RAD-DIV-DRIF module is conserved in ovary/fruit development.

Entangled in this question is the fact that *Antirrhinum* ovaries are monosymmetric. *Antirrhinum* (and all species in the tribe Antirrhineae to which *Antirrhinum* belongs) develops bicarpellate ovaries and fruit with one dorsally positioned and one ventrally positioned locule. The morphology of the dorsal and ventral locules in *Antirrhinum* are unequal (Sutton, 1988). However, the role of flower symmetry genes in establishing ovary monosymmetry in *Antirrhinum* has not been explicitly tested. Therefore, we additionally determined whether elements of the CYC-dependent program are required for *Antirrhinum* ovary monosymmetry in addition to petal and stamen whorl monosymmetry. We found that dorsal flower identity gene products *Am*CYC and *Am*RAD are required for proper development of ovary monosymmetry. We place this finding into a broader context of dynamic ovary symmetry evolution in the tribe to which *Antirrhinum* belongs.

Results

AmCYC and *AmRAD* function in *Antirrhinum* ovary monosymmetry

Monosymmetry of the wildtype *Antirrhinum* ovary is clearly visible from early stages of flower development (**Figure 5)** through fruit ripening (**Figure 4)**. The ventral locule at maturity is expanded near the base, its ventral surface makes an angle of 38.2 ± 4.2 degrees (at anthesis) and dehisces before the dorsal locule by two large pores that usually do not merge. On the other hand, the dorsal locule is not expanded near the base, making an angle of 84.6 ± 3.5 degrees (at

anthesis), overarches the apex of the ventral locule, and dehisces after the ventral locule by two small pores that merge late in fruit development.

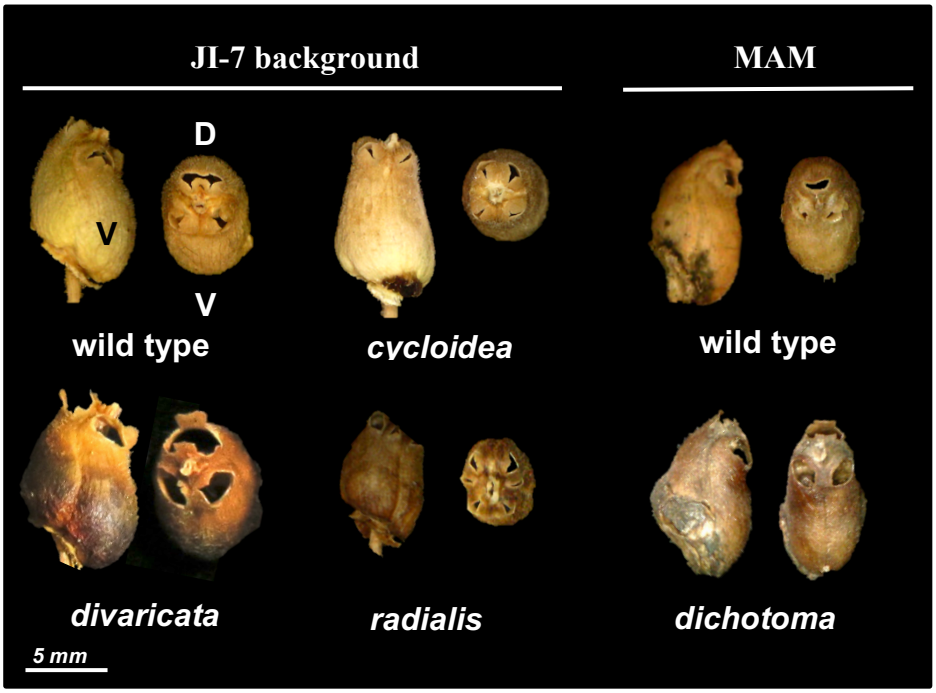

Figure 4. *Antirrhinum* wild type and mutant fruits at dehiscence. D: dorsal, V: ventral. All fruit oriented as in J1-7 wild type.

Figure 5. *Antirrhinum* (WT JI-7) reproductive tissues and comparison of WT to *Amrad* ovaries**. A.** Inflorescence. **B.** Flower bud, stage-11. **C.** Ovary and style, stage-13. **D.** Ovary and style, stage-14 (anthesis). **E.** Ovary seven days after anthesis. A-E scale bar: 5 mm. **F.** Wild type and *Amrad* ovaries, stage-14. Top panel, intact ovaries; bottom panel, longitudinal section. For each carpel, the dorsal locule is on the left. Scale bar: 2 mm.

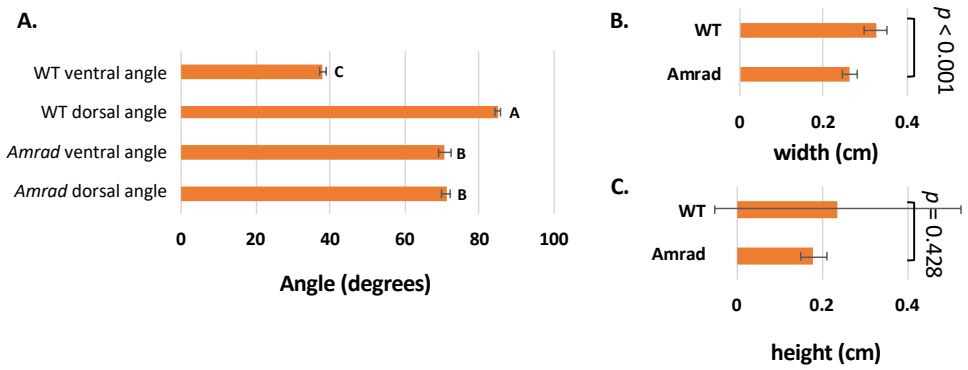

Figure 6. Wild type and *Amrad* ovary morphology at anthesis. A. Dorsal and ventral locule angles (as in Fig. 10); letters at the tips of bars are groupings based on Tukey's post-hoc comparisons; $p < 0.001$. B. Ovary width and C. ovary height; p-values are from T-test performed on bracketed tissues.

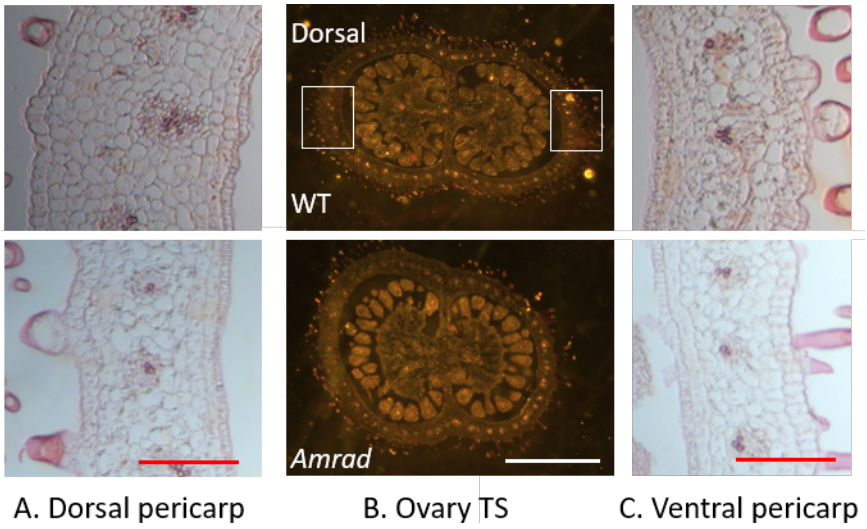

Figure 7. Wild type and *Amrad* ovary micromorphology at anthesis. The ovary in transverse section is displayed in B; boxed areas are magnified in A and C. Orientation same as in WT. Scale bar, red: 0.1 mm, white 1 mm.

Antirrhinum ovaries and fruits in *Amcyc* and *Amrad* mutant backgrounds, but not in *Amdich* or *Amdiv* mutant backgrounds, have ventralized dorsal locules (**Figure 4**). Ventralization includes equalization of the angle at dorsal and ventral surfaces (**Figure 6**), loss of curvature at the apex of the dorsal locule, and development of two pores in the dorsal locule that do not merge **(Figure 4)**. In the *Amrad* mutant background, additional pores occasionally appear in the lateral regions, with the total number of locules either remaining two, or becoming three, irrespective of the total number of pores.

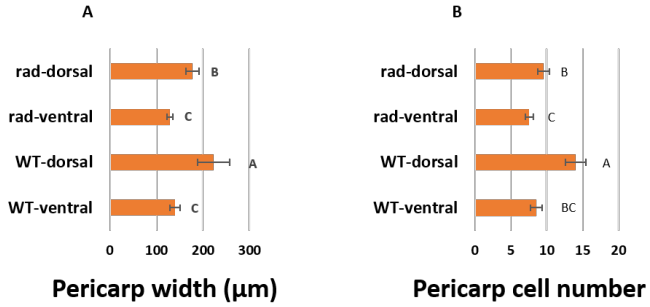

Figure 8. Wild type and *Amrad* pericarp micromorphology at anthesis. Letters at the tips of bars are groupings based on Games-Howell's or Tukey's post-hoc comparisons. A. Pericarp width ($p = 0.000$). B. Pericarp cell number ($p = 0.000$).

Ventralization is in part the result of changes to patterns of cell proliferation, specifically, in the pericarp (ovary wall) on the dorsal side. The dorsal pericarp in the wildtype ovaries at anthesis has more cells than the ventral pericarp, and hence it is wider (**Figure 7**, **Figure 8**). In *Amrad* mutants the dorsal pericarp is ventralized—it is narrower and has fewer cells than the wildtype (**Figure 7**, **Figure 8**). The width of the ventral pericarp remains unaffected in *Amrad* mutants (**Figure 7**, **Figure 8**), consistent with the hypothesized dorsal-specific cell-autonomous function of *AmRAD* in ovaries.

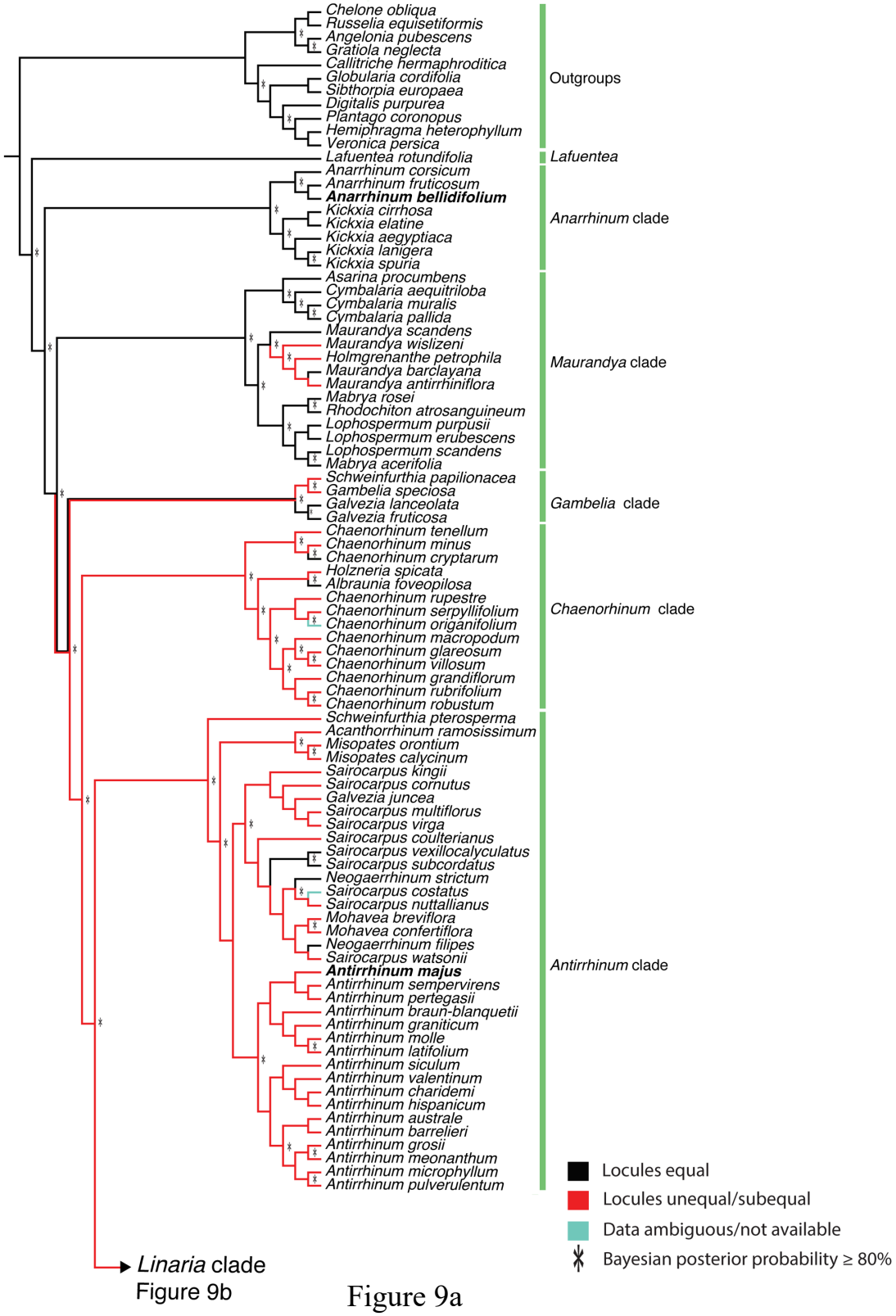
Chelone obliqua
Russelia equisetiformis
Angelonia pubescens
Gratiola neglecta
Callitriche hermaphroditica
Globularia cordifolia
Sibthorpia europaea
Digitalis purpurea
Plantago coronopus
Hemiphragma heterophyllum
Veronica persica
Outgroups
Lafuentea rotundifolia
Lafuentea
Anarrhinum corsicum
Anarrhinum fruticosum
Anarrhinum bellidifolium
Kickxia cirrhosa
Kickxia elatine
Kickxia aegyptiaca
Kickxia lanigera
Kickxia spuria
Anarrhinum clade
Asarina procumbens
Cymbalaria aequitriloba
Cymbalaria muralis
Cymbalaria pallida
Maurandya scandens
Maurandya wislizeni
Holmgrenanthe petrophila
Maurandya barclayana
Maurandya antirrhiniflora
Mabrya rosei
Rhodochiton atrosanguineum
Lophospermum purpusii
Lophospermum erubescens
Lophospermum scandens
Mabrya acerifolia
Maurandya clade
Schweinfurthia papilionacea
Gambelia speciosa
Galvezia lanceolata
Galvezia fruticosa
Gambelia clade
Chaenorhinum tenellum
Chaenorhinum minus
Chaenorhinum cryptarum
Holzneria spicata
Albraunia foveopilosa
Chaenorhinum rupestre
Chaenorhinum serpyllifolium
Chaenorhinum origanifolium
Chaenorhinum macropodum
Chaenorhinum glareosum
Chaenorhinum villosum
Chaenorhinum grandiflorum
Chaenorhinum rubrifolium
Chaenorhinum robustum
Chaenorhinum clade
Schweinfurthia pterosperma
Acanthorrhinum ramosissimum
Misopates orontium
Misopates calycinum
Sairocarpus kingii
Sairocarpus cornutus
Galvezia juncea
Sairocarpus multiflorus
Sairocarpus virga
Sairocarpus coulterianus
Sairocarpus vexillocalyculatus
Sairocarpus subcordatus
Neogaerrhinum strictum
Sairocarpus costatus
Sairocarpus nuttallianus
Mohavea breviflora
Mohavea confertiflora
Neogaerrhinum filipes
Sairocarpus watsonii
Antirrhinum majus
Antirrhinum sempervirens
Antirrhinum pertegasii
Antirrhinum braun-blanquetii
Antirrhinum graniticum
Antirrhinum molle
Antirrhinum latifolium
Antirrhinum siculum
Antirrhinum valentinum
Antirrhinum charidemi
Antirrhinum hispanicum
Antirrhinum australe
Antirrhinum barrelieri
Antirrhinum grosii
Antirrhinum meonanthum
Antirrhinum microphyllum
Antirrhinum pulverulentum
Antirrhinum clade
Linaria clade
Figure 9b
Locules equal
Locules unequal/subequal
Data ambiguous/not available
Bayesian posterior probability ≥ 80%

Figure 9a

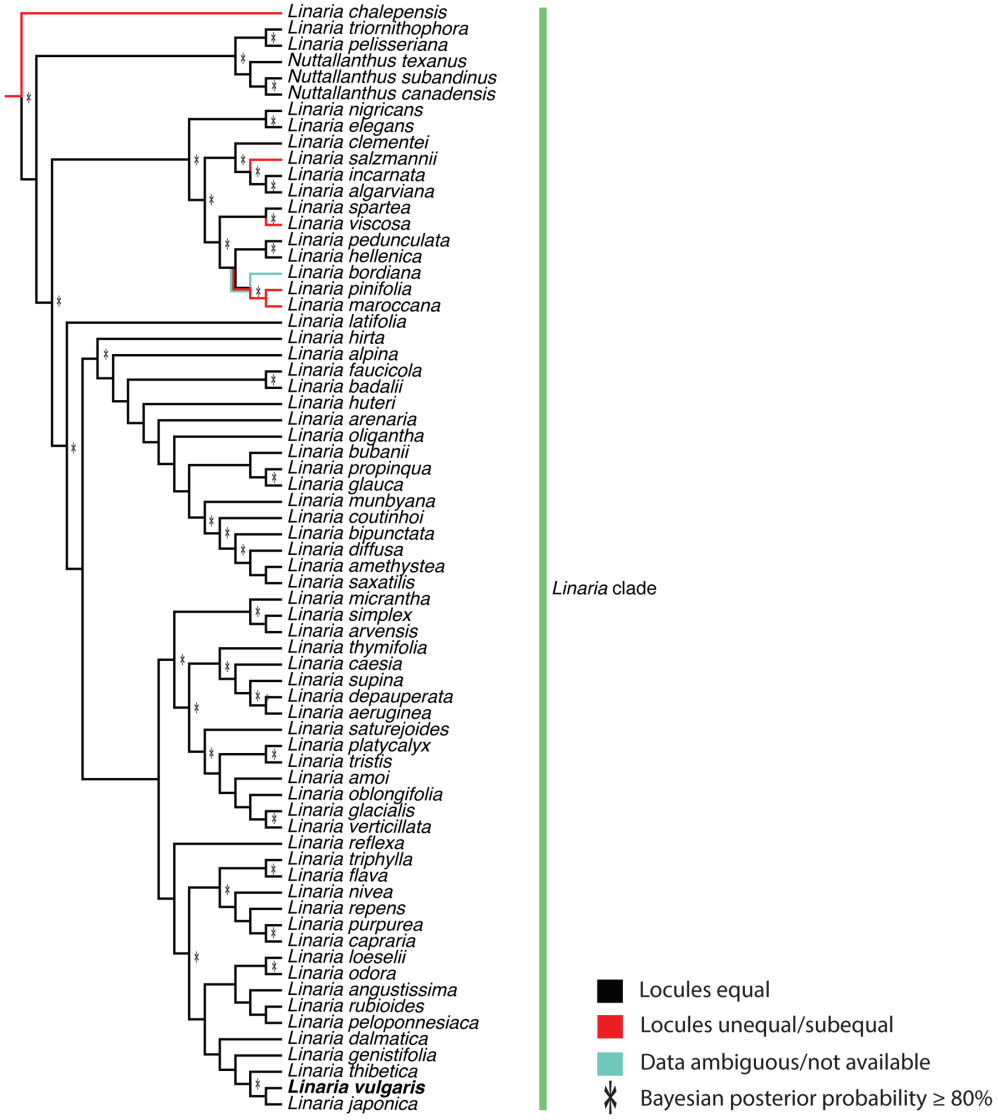

Figure 9b

Figure 9. Evolutionary history of ovary symmetry in the tribe Antirrhineae inferred by Mesquite employing a parsimony-based ancestral state reconstruction. Species studied in this work are in bold. The backbone Bayesian tree is from Ogutcen and Vamosi, 2016.

Dynamic evolution in ovary symmetry during Antirrhineae diversification

The tribe Antirrhineae ancestrally had two, equivalent (one dorsal, one ventral) locules (**Figure 9**) and this feature is retained in the early diverging lineages *Lafuentea rotundifolia* and the sub-tribe *Anarrhinum* clade. Inequality or sub-equality of the locules (and corresponding fruit monosymmetry) has evolved at least five times, and has been lost at least seven times (the branch leading to the *Gambelia* clade has an equivocal state). Inequality or sub-equality of locules is ancestral to the *Antirrhinum* clade.

This state of locule inequality/sub-equality represented in the *Antirrhinum* clade evolved either in the common ancestor of the *Chaenorhinum* + *Antirrhinum* + *Linaria* lineage, or earlier, including the common ancestor of these lineages plus the *Gambelia* lineage (ancestral state is equivocal in the *Gambelia* clade). This puts the age of locule inequality at *ca.* 22.9 mya (not including *Gambelia* clade) to *ca.* 24.4 mya (including *Gambelia* clade)(dates are from Ogutcen and Vamosi, 2016). Locule inequality/subequality was subsequently lost in the *Linaria* clade after the divergence of *Linaria chalepensis*.

Expression of *AmRAD*, *AmDIV*/*DIV-like1* and *AmDRIF1&2* are consistent with a function in ovary development independent of dorso-ventral identity

We used quantitative real-time PCR (qRT-PCR) to determine relative expression of *Antirrhinum* symmetry genes across stages of ovary development with the objective of assessing evidence for RAD-DIV-DRIF function during ovary/fruit development similar to that found in tomato (Machemer et al., 2011). *AmCYC*, *AmDICH*, *AmRAD*, *AmDIV*, *AmDRIF1*, and *AmDRIF2* are involved in defining flower monosymmetry in *Antirrhinum. AmDIV-like1*, a close

paralog of *AmDIV*, has not been implicated in the control of flower symmetry, but it is important for understanding the ancestral expression and function of its sister gene, *AmDIV*.

We found that upstream regulators of dorsal flower identity, *AmCYC* and *AmDICH* (**Figure 3**), have relatively high expression in tissues with petals and stamens—inflorescences and entire flower buds (**Figure 10**). This is consistent with their singular role in establishing dorsal petal and stamen identity (Luo et al., 1996). We found *AmCYC* and *AmDICH* expression to be sparingly low to undetectable in isolated ovary tissue of any stage (**Figure 10**).

We found that the dorsal flower identity gene, *AmRAD* (**Figure 3**), is expressed in tissues with petals and stamens—inflorescences and entire flower buds (**Figure 10**), consistent with its previously identified role in establishing dorsal petal and stamen identity (Corley et al., 2005). In addition, we found a striking pattern whereby *AmRAD* expression peaks in late stages of ovary development, stage-14 (anthesis) flowers. We sequenced the qRT-PCR amplicon from stage-14 carpels and confirmed that the primers were amplifying the correct template. The late high expression of *RAD* is apparently conserved in the tribe Antirrhineae. The *AmRAD* orthologs in an early diverging member (*Anarrhinum bellidifolium*, *AbRAD*) and a late diverging member (*Linaria vulgaris*, *LvRAD*) have a peak of expression in ovaries at anthesis (**Figure 11**).

Given the surprisingly high relative expression of *AmRAD* in isolated ovary tissue, we performed further qRT-PCR on sub-parts of the carpels from stage-14 flowers to determine possible localization of *AmRAD* transcripts. We found that expression of *AmRAD* is significantly higher in ovary walls than in the internal ovules+septa of the ovary, but that ovary expression of *AmRAD* is not asymmetric across the dorsal-ventral locules of the ovary (**Figure 10**). Therefore, *AmRAD* expression during later stages of ovary differentiation and maturation is independent of dorso-ventral positional information within the flower. Further, the late, high expression of

AmRAD in ovaries is independent of *AmCYC*. The expression of *AmRAD* is not significantly different between the stage-13 ovaries of wildtype and *Amcyc* background **(Figure 12)**. This lack of significant difference is not due to low sensitivity of the testing methods—downregulation of

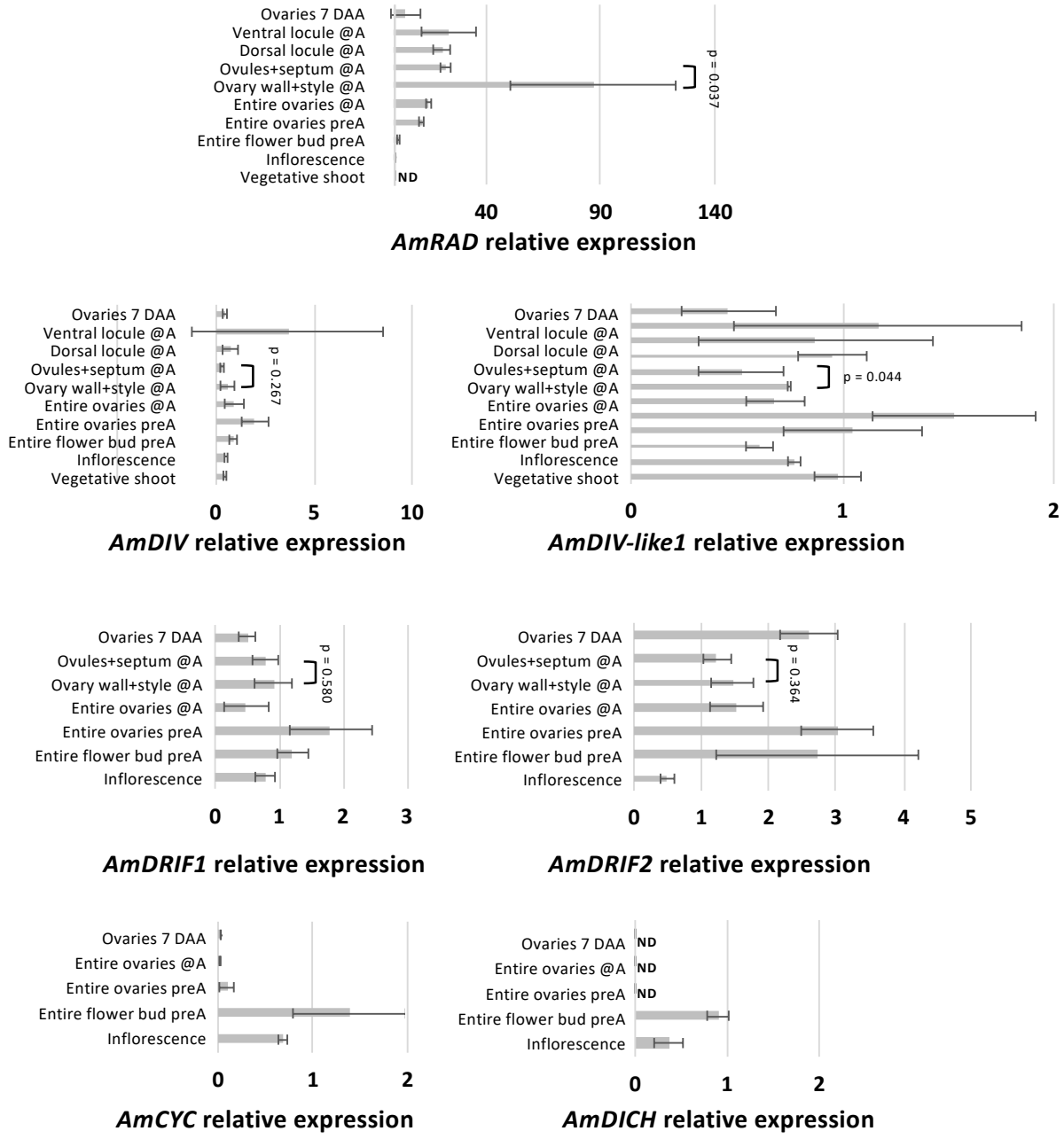

Figure 10. Relative expression of genes involved in petal and stamen symmetry development across reproductive wild type *Antirrhinum* tissues in the JI-7 background. Error bars are standard deviations of samples. ND: expression not determinable; p-values from T-tests performed on the bracketed tissues. 7 DAA, 7 days after anthesis; @A, at anthesis (stage-14); preA, pre-anthesis; entire ovary preA is from stage-13.

AmRAD in *Amcyc* inflorescence can be captured by qRT-PCR **(Figure 12)** even though *AmRAD* is expressed in inflorescence tissues at a much lower level than ovaries.

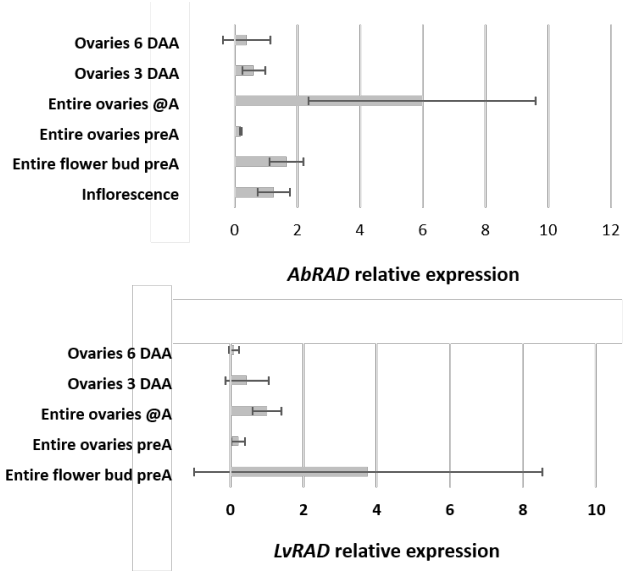

Figure 11. Expression of *AmRAD* orthologs in *Anarrhinum bellidifolium* (top) and *Linaria vulgaris* (bottom). DAA, days after anthesis; @A, at anthesis; preA, pre-anthesis.

AmDRIF1&2 and *AmDIV* function coordinate with *AmRAD* to restrict *Am*DIV activity to developing ventral petals and stamens (Raimundo et al., 2013). We found that along with *AmRAD, AmDRIF1&2*, *AmDIV* and *AmDIV-like1* are expressed in anthesis stage ovaries at levels comparable to their expression in inflorescence and entire flower bud tissues (**Figure 10)**. This supports the hypothesis that their gene products may interact with AmRAD during ovary differentiation and maturation.

Conflicting evidence that putative *AmDIV* and *AmDIV-like1* autoregulation is negatively affected by AmRAD

Extensive biochemical analyses support the hypothesis that *Am*RAD suppresses *Am*DIV function through competitive interaction for binding to *Am*DRIF1 and *Am*DRIF2 (Raimundo et

al., 2013). There is evidence that *AmDIV*, *AmCYC*, and *AmDICH* affect the transcription of *AmDIV* (Galego and Almeida, 2002a). In stage-10 flowers, *AmDIV* mRNA is expressed in all petals but is concentrated in the inner epidermis of the ventral petal. However, in div-35 mutants (frameshift mutation), the accumulation in the inner epidermis of the ventral petal is removed and expression becomes uniform across the ventral petal (in addition to the expression in all other petals) (Galego and Almeida, 2002a). In *Amcyc-Amdich* double mutants, the accumulation in ventral petal spreads to all other petals—every petal has higher *AmDIV* mRNA levels in their inner epidermis (Galego and Almeida, 2002a). This suggests that the transcription of *Am*DIV is under the negative regulation of *Am*CYC and *Am*DICH. However, *Am*CYC and *Am*DICH are not known to be downregulators of transcription, and possibly regulate the transcription of *AmDIV* by disrupting its autoregulation. This disruption may be mediated by *Am*RAD or another factor.

Testing for competitive inhibition of *Am*DIV or *Am*DIV-like1 by *Am*RAD during *Antirrhinum* ovary development is beyond the scope of this study. However, there are tentative data suggesting that that competitive inhibition of *Am*DIV or *Am*DIV-like1 by *Am*RAD may disrupt a putative *AmDIV* positive autoregulation. Specifically, *Am*DIV is known to bind to the conserved DNA sequence 5′-[AGC]GATA[AC][GC][GAC]-3′, and within 3 kb upstream of the *AmDIV* transcriptional start site are two conserved DIV DNA binding motifs (Raimundo et al., 2013; Sengupta and Hileman, 2018). The possibility that *AmDIV* expression is altered by the abundance of *Am*RAD protein allows us to indirectly test negative regulatory interactions by determining expression of *AmDIV* and *AmDIV-like1* in wild type compared to *Amrad* backgrounds with the expectation that *AmDIV* and/or *AmDIV-like1* expression should increase in the absence of functional *Am*RAD protein.

Similar to *AmDIV*, within 3 kb of the transcriptional start site of *AmDIV-like1* we identified three consensus DIV DNA binding sites, 5′-[AGC]GATA[AC][GC][GAC]-3′ (**Table 1**). The MYB DNA binding domain of *Am*DIV-like1 shares 100% amino acid identity with the corresponding region of *Am*DIV, suggesting similar capacity for interaction with DIV DNA binding sites. The MYB protein-protein interaction domain of *Am*DIV-like1 has 80% amino acid identity with the corresponding region of *Am*DIV, suggesting similar capacity for interactions with *Am*DRIF1&2. In addition, we demonstrated above that *AmDIV-like1* expression is significantly higher in ovule/septum tissue of the ovary compared to ovary wall/style tissue, a pattern negatively correlated with *AmRAD* expression (**Figure 10**). These findings support the possibility that both *Am*DIV and *Am*DIV-like1 may function to auto and/or cross-regulate their own expression in a manner that can be detected in the *Amrad* mutant background.

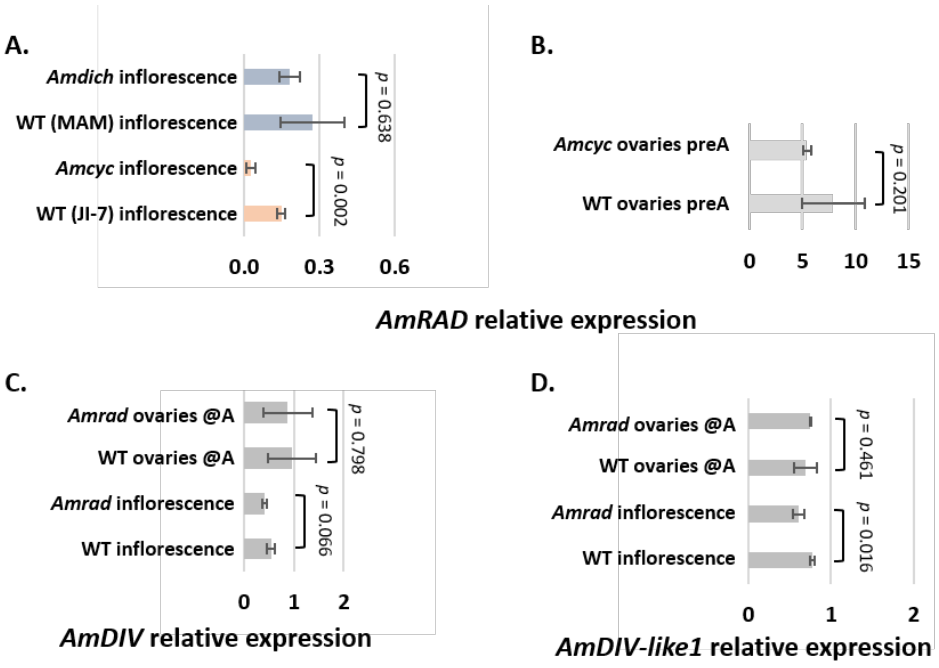

Figure 12. Quantitative RT-PCR expression of *AmRAD, AmDIV* and *AmDIV-like1* in mutant *Antirrhinum* backgrounds compared to wild type**. A**. *AmRAD* expression the inflorescences in *Amcyc* and *Amdich* mutant backgrounds; *Amdich* is in the MAM background and *Amcyc* is in the JI-7 background, **B**. *AmRAD* expression the preanthesis ovaries in *Amcyc* and wildtype backgrounds. **C.** *AmDIV* expression in *Amrad* mutant background. **D.** *AmDIV-like1* expression in *Amrad* mutant background. Error bars are standard deviations of samples. The p-values are from T-test performed on the bracketed tissues.

However, we find that neither *AmDIV* nor *AmDIV-like1* expression is significantly upregulated in the *Amrad* mutant background compared to wild type (**Figure 12**). This is the case for both inflorescence tissues as well as ovary tissues where *AmRAD* is expressed at relatively high levels in wild type (**Figure 12**). In fact, *AmDIV-like1* expression is significantly lower in inflorescences of *Amrad* mutants relative to wild type (**Figure 12**). To ensure that we are able to confidently detect regulatory interactions in mutant backgrounds we confirmed that *AmRAD* is significantly downregulated in the *Amcyc* background inflorescences (**Figure 12**). Our gene expression results suggest that neither *AmDIV* nor *AmDIV-like1* are directly or indirectly positively regulated, or the regulatory interaction is of small effect and not detectable by our methods. Therefore, we are unable to use *AmDIV/DIV-like1* expression as a test for *Am*RAD-*Am*DIV/DIV-like1 competitive interaction during *Antirrhinum* ovary development.

Table 1. Predicted consensus DIV-binding sites upstream of *AmDIV-like1* (5′ to 3′)

Sequence	DNA strand	bp upstream of transcription start
AGATAACG	sense	1373–1366
AGATAAGA	anti-sense	283–276
AGATAAGG	anti-sense	1628–1621

Discussion

*Am*RAD may function non-cell autonomously to specify *Antirrhinum* ovary monosymmetry

Loss-of-function mutations in *Antirrhinum AmRAD* and its positive regulator, *AmCYC*, not only transform dorsal petals and stamens to lateral identity (Luo et al., 1996; Corley et al., 2005), but they radialize the central ovary (and subsequent fruit) by converting the dorsal locule to ventral identity. This is consistent with a recent finding in *Misopates orontium*, a close relative of *Antirrhinum*, in which loss of CYC function similarly results in ventralization of the dorsal

locule (Lönnig et al., 2018). However, this finding is difficult to reconcile with the observation that mRNA of neither *AmCYC* nor *AmRAD* is detectable in early stages of *Antirrhinum* ovary development (Corley et al., 2005). Early stages of ovary development are crucial for expression of genes that may affect symmetry because the inequality of the two locules is patterned early after the ovule primordium initiates. A clear function for *AmCYC* and *AmRAD* in establishing dorsal locule identity, but a lack of corresponding mRNA in the stages when patterning occurs, suggests that *AmCYC* functions non-cell autonomously via *Am*RAD protein for differentiation of the dorsal locule.

Previous work points to non-cell autonomous function of *AmRAD* during petal whorl development. *AmRAD* transcripts are localized to the developing dorsal petal primordia at early stages of *Antirrhinum* flower development and are not detected in lateral petal primordia. Yet, in the *Amrad* mutant background the dorsal side of each lateral petal has altered development. This observation led Corley et al., (2005) to hypothesize that *AmRAD* is transcribed and translated in the dorsal floral organs followed by *Am*RAD protein moving into the lateral petals to non-cell autonomously affect morphology those organs. Our data suggest a similar pattern of *AmRAD* non-cell autonomous action. We hypothesize that *Am*CYC upregulates *AmRAD* transcription in the dorsal petals and stamens early in development and translated *Am*RAD protein migrates to the developing dorsal locule where it establishes dorsal identity and restricts ventral locule identity.

We attempted to use potential autoregulation of *AmDIV* and/or *AmDIV-like1* to test for competitive inhibition of *Am*DIV or *Am*DIV-like1 by *Am*RAD in developing ovaries, but our data suggest that neither of these transcription factors is actually under direct or indirect positive

autoregulation. Still, it is likely that *Am*RAD in the developing ovary functions by outcompeting *Am*DIV or *Am*DIV-like1, or another MYB transcription factor.

A non-cell autonomous function of *CYC* in fruit symmetry seems to be consistent with the expansion of expression of *AmCYC* and *AmDICH* orthologs to lateral organs in *Mohavea confertiflora* (nested within the genus *Antirrhinum*). This expansion is correlated with dorsalization of lateral stamens (Hileman et al., 2003). Unlike *Antirrhinum majus*, the locules in *Mohavea* open by one pore each, (reviewed in Sutton, 1988), which is similar to the dorsal locule of Antirrhinum. The orthologs of *AmCYC* and *AmDICH* in *Mohavea* are not expressed in carpels (Hileman et al., 2003), but it is possible that along with the expansion of the autonomous function to the lateral organs, the non-cell autonomous function of *Mohavea RAD* expanded to include both the dorsal and ventral locule, resulting in dorsal-like identity of *Mohavea* ventral locules.

AmDICH is likely not involved in defining fruit symmetry, as apparent from the wild type fruit symmetry in the *Amdich* background (**Figure 4**). This is in-line with our observation that *AmRAD* expression is not significantly downregulated in the *Amdich* mutant background and *AmDICH* is not detected in ovaries at any developmental stage (**Figure 10**). *AmDIV* may also not be involved in defining fruit symmetry, as apparent from the wild type morphology in the *Amdiv* background (**Figure 4**). However, it is possible that *Amdiv* has subtle effects on ventral locule morphology or that *AmDIV* function in ventral locule morphology can only be detected in an *Amdiv,Amdiv-like1* double mutant background and no *AmDIV-like1* mutant lines exist. Notably, both *AmDIV* and *AmDIV-like1* are transcribed in early stages of carpel development (Galego and Almeida, 2002a). Therefore, it is possible that AmDIV-like1 by itself, or in co-operation with its

paralog AmDIV, interacts with *Am*RAD to establish *Antirrhinum* fruit monosymmetry along the dorso-ventral axis.

A possible role for *RAD* during evolution of ovary symmetry across Antirrhineae

Flowers in the tribe Antirrhineae have ancestrally monosymmetric petal and stamen whorls, but their ovaries (and fruits) ancestrally lacked monosymmetry with identical dorsal and ventral locules. However, monosymmetry in fruits has evolved at least five times, and has been lost at least seven times independently in the tribe (**Figure 9**). This result demonstrates that symmetry in fruits can experience evolutionary changes without affecting the symmetry in other floral whorls. For example, the majority of the genus *Linaria* has experienced a reversal from unequal to equal loculed condition (where the dorsal and ventral locules lose their morphological differences), but the corolla remains monosymmetric (the dorsal petals are distinct from the ventral ones).

Our results point to a non-cell autonomous action of *AmRAD* in defining fruit monosymmetry in *Antirrhinum*. This non-autonomous function is likely shared with close relatives *Misopates orontium* and *Mohavea confertiflora*, as described above. One mechanism by which fruit symmetry may evolve without pleiotropic effects on conserved petal and stamen whorl monosymmetry is through changes in the non-cell autonomous action of *RAD* homologs. Under this scenario, the ancestral condition in the tribe was likely cell autonomous action of *RAD* with non-autonomous function evolving in the ancestor of *Chaenorhinum* + *Antirrhinum* + *Linaria*, and possibly independently in *Maurandya*, *Gambelia* and a few *Linaria* lineages. We hypothesize this was followed by reversal to the cell autonomous function of *RAD* in the ancestor of *Linaria*. Tests of our hypotheses concerning dynamic evolution of *RAD* cell

autonomous function await biochemical (protein antibody) assays for RAD across the Antirrhineae.

Expression of *AmRAD*, *AmDIV/DIV-like1*, and *AmDRIF1&2* are consistent with a function in ovary development independent of dorso-ventral identity

We identified a novel peak in *AmRAD* expression late in ovary development. The absence of *AmCYC* mRNA late in ovary development, and the distribution of *AmRAD* transcripts across both dorsal and ventral locules, suggests that *AmRAD* is upregulated by a factor other that *Am*CYC in later stage ovaries. These results also indicate an important developmental function in later stages of ovary/fruit development, especially in the ovary wall where *AmRAD* expression is highest. This function is likely independent of fruit symmetry. We were not able to find any phenotype under the control of this late *AmRAD* expression in our survey of *Antirrhinum* ovary micromorphology in transverse sections. It is possible that *AmRAD* controls a phenotype that we did not test, possibly along the longitudinal plane. It is likely that this function involves *Am*RAD competitively excluding *Am*DIV/DIV-like1 from interacting with *Am*DRIF1/2. This hypothesis is based on the fact that high expression of *AmRAD* in ovaries coincides with expression of *AmDIV/DIV-like1* and *AmDRIF1&2* in those tissues, and the only known biochemical interactions known for *Am*RAD homologues involve homologs of *Am*DIV/DIV-like1 and *Am*DRIF1/2.

A conserved ancestral function of RAD–DIV–DRIF in fruits likely pre-dates Lamiales flower monosymmetry.

In Lamiales, *Am*RAD is known to function in defining floral monosymmetry along the dorso-ventral axis, and monosymmetry evolved in Lamiales after its divergence from close relative, the Solanales. Tomatoes are a model species in the order Solanales. A RAD–DIV–DRIF like interaction has been reported from tomato fruits, where the RAD component suppresses cell expansion in the pericarp tissue. Pericarp, or the fruit wall, is the ovule wall after fertilization. We provide suggestive evidence that *Am*RAD has a function in late ovary/fruit development, and that this function involves *AmDIV*, *AmDIV-like1*, and *AmDRIF1&2* in that expression of these gene overlaps with expression of *AmRAD* in later stages of ovary development. Hence, we hypothesize an ancestral function of *RAD*-like genes is in controlling micromorphology during ovary wall development. This conclusion is supported by the fact that a high expression of *RAD* in ovaries is conserved across Antirrhineae. This also suggests that the RAD–DIV–DRIF interaction, which is crucial in defining Lamiales flower symmetry, did not evolve during the origin of flower monosymmetry in Lamiales but was co-opted from a different function, likely fruit/carpel development, to define the dorso-ventral monosymmetry in Lamiales flowers.

Methods

Plant material

We acquired *Antirrhinum majus* seeds from The John Innes Centre (JIC), Norwich, Norfolk, England, UK, and The Leibniz Institute of Plant Genetics and Crop Plant Research (IPK), Gatersleben, Germany. The seeds were imported under the permit number P37-16-01034 granted to Lena Hileman (application number P587-160901-023) by United States Department

of Agriculture (USDA), Animal and Plant Health Inspection Service (APHIS), Riverdale, MD, USA. Seeds for mutant *Amdich* (MAM 95) and its corresponding wildtype (MAM 428) are from IPK. Only one *Amdich* line is available from IPK, and hence is likely the one described previously (Luo et al., 1999). The rest of the lines are from JIC and are as follows: wildtype JI-7, *Amcyc* JI-608 (Luo et al., 1996), *Amcyc-Amdich* JI-720, *Amrad* JI-654 (Corley et al., 2005), and *Amdiv* JI-13 (the original mutant div-35 described by Almeida et al., 1997, is lost, but has been reconstituted from crossing Gatersleben pal-car ; del mutant to JI stock 15 (pal-tub stabiliser); personal communication, Lucy Copsey, Research Assistant, JIC). We broadcast the seeds on sowing medium, lightly covered them with vermiculite, and germinated them under a short night (i.e., "long day") condition at 20–26° C. We transferred the germinated seedlings to larger individual pots, maintained them in the same light and temperature regimen, treated them for arthropod pests and fertilized them with chemical fertilizer once a week.

Tissue sampling

We collected the following tissues for qRT-PCR analyses (**Figure 5**): pre- floral induction vegetative shoot (includes leaves, stem, and shoot apex), inflorescence (*ca*. 8.0 mm long, *ca*. 3.5 mm wide), stage-11 flower bud (flower bud ca. 4.0 mm in length, corolla equal in length to calyx, petal tips white in wildtype; Vincent and Coen, 2004), carpels from stage-13 flower buds (flower bud ca. 1.0 cm in length; Vincent and Coen, 2004), carpels from stage-14 (anthesis) flowers (corolla mouth open; anthers bright yellow, turgid, but yet to dehisce; Vincent and Coen, 2004), dorsal and ventral ovary locules from flowers at anthesis, carpel wall (consisting ovary wall and style) and the ovules (consisting ovules and the septum) from flowers at anthesis, and fruits (only the post-fertillization ovary) seven days after anthesis (DAA) that

were derived from flowers pollinated manually at anthesis. We did not test all tissues for all genes; if there was evidence suggesting that the expression in a tissue was likely to be absent or irrelevant, we did not test its expression by qRT-PCR. For example, given the sparingly small expression of *AmCYC* in carpels of anthesis-stage flowers, we did not test for *AmCYC* levels in dorsal versus ventral ovary locules. We did not test earlier developmental stages because gene expression has been determined in those stages by previous workers.

Isolating RAD orthologs from *Anarrhinum bellidifolium* and *Linaria vulgaris*

RAD orthologs were isolated by PCR (Bullseye Taq DNA polymerase, Midwest Scientific, St. Louis, MO, USA) using the degenerate primers RAD-70-F (GCATTGGCGGTTTACGAYMAAG) and RAD-240-R (ACYRGTGGTCCTRTAGTTRGG) (Preston et al., 2011). The PCR products were cloned/sequenced in pGEM-T vector system (Promega Corporation, Madison, WI, USA). Putative RAD orthology was assessed using phylogenetics (data not shown).

Quantitative RT-PCR assays

We extracted Total RNA from three biological replicates of each tissue type using RNeasy plant minikit (Qiagen, Germantown, MD, USA) or TRI Reagent (Thermo Fisher Scientific, Waltham, MA, USA), followed by DNase treatment (TURBO DNase, Thermo Fisher Scientific), and cDNA synthesis (iScript cDNA Synthesis Kit, Bio-Rad, Hercules, CA, USA). We performed quantitative RT-PCR in a StepOnePlusTM Real-Time PCR System (Thermo Fisher Scientific) using Bullseye EvaGreen qPCR Mastermix (Midwest Scientific, St. Louis, MO, USA) for *AmDRIF1* (GenBank ID JX966358.1), *AmDRIF2* (GenBank ID JX966359.1), and

AmDIV-like1(GenBank ID), or SYBR Select Master Mix (Thermo Fisher Scientific) for *AmCYC* (GenBank ID Y16313.1)*, AmDICH* (GenBank ID AF199465.1)*, AmRAD* (GenBank ID AY954971) and *AmDIV* (GenBank ID AY077453.1). Expression of *AbRAD* and *LvRAD* was measured using PowerUp SYBR Green Master Mix (Thermo Fisher Scientific). We normalized expression of target genes against a constitutively expressed gene *AmUBIQUITIN5* (*AmUBQ5*), or its homologs in *Anarrhinum* and *Linaria*. This gene has been reported to have little transcriptional variation across tissue types and developmental stages (Preston and Hileman, 2010). The qRT-PCR primer pairs are as listed in **Table 2**; we determined PCR primer efficiencies using DART (Peirson et al., 2003). We analyzed expression employing ΔΔCt method using Microsoft Excel. For ΔΔCt calculation, we normalized expression against a single stage-11 flower bud biological replicate. We compared expression between pairs of tissues using two sample T-test assuming equal variances.

Table 2. Antirrhineae qRT-PCR primers (5′ to 3′)

Gene	Forward primer	Reverse primer
AmUBQ5	GCGCAAGAAGAAGACCTACAC	CTTCCTGAGCCTCTGCACTT
AmCYC	CATCCTCCCTTCACTCTCGC	TGAACAAAGCGGTGGACTCA
AmDICH	TGAGTGGAACCCCTCAGTTC	CCCAAACATTGAAGGGTGGT
AmRAD	GGACGAACACCGGAAGAAGT	GTTGCCCCGACCATAGCTTA
AmDIV	GGGGACTGGAGGAACATCTC	CGATGGAGTTTGGTTGTCGC
AmDIV-like1	GATCACGGGTTTTGGCAGT	ATCGACCCTGCAGTCCAAC
AmDRIF1	GCCTTGGATCAAATTTCGGC	AGGAAGAATGGAGCTGGCAA
AmDRIF2	AATGGTCATGGAGAGTGGGG	TATAGCTTGCTCCTCTGGGG
AbRAD	TTGGACCAACGTGGCGAG	AGGGCACTTTACCACTCTCA
LvRAD	GCTAATGTGGCTAGGGCTGT	GGCACTTTCCCGCTCTCAAT

For *AmRAD* qRT-PCR, we sequenced the qRT-PCR amplicon from stage-14 carpels confirm that the primers were amplifying the correct template. We isolated the amplicon by QIAquick Gel Extraction Kit (Qiagen), then polished it with Bullseye Taq DNA polymerase

(Midwest Scientific) to generate A-tails, ligated the polished fragments into pGEM-T vector (Promega Corporation, Madison, WI, USA) by TA-cloning, transformed the vector into electrocompetent *E. coli* DH5α (NEB, Ipswich, MA, USA), and finally plated the bacteria on blue-white selection plates with carbenicillin. We picked white colonies from the plates, grew them in liquid media with carbenicillin, and then we alkaline lysed them with QIAprep Spin Miniprep Kit (Qiagen) to isolate plasmid DNA. We sequenced plasmids from two colonies with Retrogen, San Diego, CA, USA, using M13 primers (forward 5′-CGCCAGGGTTTTCCCAGTCACGAC-3′ and reverse 5′-TCACACAGGAAACAGCTATGAC-3′).

Consensus DIV-binding site predictions

We downloaded three kilobases of non-coding sequence upstream of *AmDIV-like1* transcription start site from a published genome (Li et al., 2019, version3 available at http://bioinfo.sibs.ac.cn/Am). We defined the AmDIV-binding site as 5′-[AGC]GATA[AC][GC][GAC]-3′ (Raimundo et al., 2013). We performed the alignments and predictions in Geneious 10.2.3 (Kearse et al., 2012).

Macromophological analysis of *Antirrhinum* ovaries

We photographed stage-14 (anthesis) ovaries from 23 wild type (JI-7) and 17 *Amrad* (JI-654) flowers. We took measurements using the software ImageJ2 (http://imagej.net/). We measured angle of ovary curvature as depicted in

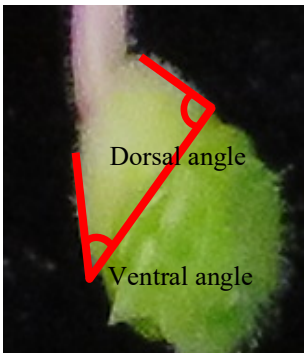

Figure 13. Measuring angles at the dorsal and ventral faces of Antirrhinum ovary at anthesis.

Figure 13. We treated the proximal boundary of the ovary parallel to boundary of the nectary as the base for measuring angles. Ovary width was measured at the base and parallel to the necary; ovary height was measured along the septum from the base of the ovary to the ovary-style junction. The ovary surface is covered in trichomes, and we ensured that we measured the angle made by the surface of the ovary and not the tips of the trichomes. We performed statistical analyses Minitab 18. We first performed Levene's test to test for equality of variances (not significantly different for any of the comparisons). For the angular measurements, we then performed classic one-way ANOVA, and Tukey's post-hoc pairwise comparisons. We performed two sample T-tests assuming equal variances to compare ovary width and height between wild type and *Amrad*.

Micromorphological analysis of *Antirrhinum* ovaries

We fixed seven ovaries each from wildtype (JI-7) and *Amrad* (JI-654) from stage-14 (anthesis) flowers in FAA (formaldehyde, acetic acid, and ethanol). We kept the dorsal sepal attached to the base of the carpel to provide information on dorso-ventral orientation. Fixed carpels were dehydrated through a series of ethanol dilutions, stained with eosine orange, cleared with citrisolv (Thermo Fisher Scientific), and embedded in paraplast plus (Thermo Fisher Scientific). The carpels were transversely sectioned at 10 μm thickness in a microtome (AJ Griner, Kansas City, MS) with blades from Feather Safety Razor Co. Ltd (Japan). We adhered the sections on probe plus sides (Thermo Fisher Scientific), removed paraplast plus with citrisolv, rehydrated the sections through dilutions of ethanol, ending in a final rehydration in PBS at pH 7. We stained the rehydrated sections with safranin, and mounted them in water glass solution (glycerin, 37% sodium silicate, in ratio 1:2).

We took measurements in the mid–upper, trachomatous region of the ovary and avoided including any nectary tissue in our analysis. We measured the following variables at the dorsal-most and ventral-most regions of the sections: width of the pericarp along the radial axis of the ovary (including outer and inner epidermis) and the number of cells in the pericarp along the radial axis of the ovary (including outer and inner epidermis). We took measurements on a Leica DM500 B microscope at 10x objective lens calibrated using a stage micrometer. Statistical analysis was performed in Minitab version 18. We first tested homogeneity of variances with Levene's test. For datasets with homogenous variances, we performed classic Fisher's ANOVA followed by Tukey's post-hoc pairwise comparison. When the assumption of equality of variances was violated; hence we performed Welch's ANOVA followed by Games-Howell post-hoc pairwise comparison.

Phylogenetic history of tribe Antirrhineae fruit symmetry

We used a previously published species-level Antirrhineae phylogeny as the backbone for inference of ovule/fruit symmetry evolution (Ogutcen and Vamosi, 2016). This tree includes 157 species, representing close to half the species diversity of the tribe. We used Mesquite 3.40 (Maddison and Maddison, 2018) to estimate the phylogenetic history of fruit symmetry (dorso-ventral locule equality/inequality) using a parsimony based approach with categorical character states.

We scored the species for the following states: equal locules, unequal (or sub-equal) locules, and ambiguous state (or data unavailable). We scored the species for these states primarily from a monograph of the tribe (Sutton, 1988) with additional information from the following sources: *Angelonia pubescens* (Barringer, 1981), *Callitriche hermaphroditica* (images

in Watson and Dallwitz, 2018), *Chelone obliqua* (Ghebrehiwet et al., 2000), *Digitalis purpurea* (Juan, 2000; image in Lucid Central, Digitalis purpurea; image in The Plant List version 1.1, Digitalis purpurea), *Globularia cordifolia* (inferred from elliptical shape described in Bojnanský and Fargašová, 2007), *Gratiola neglecta* (Juan, 2000; inferred from shape described in The Jepson Herbarium, Gratiola neglecta), *Hemiphragma heterophyllum* (image in Flowers of India, Hemiphragma heterophyllum; image in University of British Columbia Botanical Garden, Hemiphragma heterophyllum), *Plantago coronopus* (image in Go Botany, Plantago coronopus; image number seite 690 in Sturm and Strum, 1796), *Russelia equisetiformis* (images in Ahmed et al., 2016), *Sibthorpia europaea* (image in delta-intkey, Sibthorpia europaea; Juan, 2000), *Veronica persica* (Juan, 2000), *Lafuentea rotundifolia* (image in Ivorra, 2014; image number TAB.CXVI in Willkomm, 1881), *Lophospermum purpusii* (image number Tab 8697 in Smith, 1917), *Lophospermum erubescens* (Ixitixel, 2008; inferred from spherical or oblong shape described in Walters, 2000), *Sairocarpus kingii* (The Jepson Herbarium, Antirrhinum filipes), *Sairocarpus cornutus* (The Jepson Herbarium, Antirrhinum filipes), *Neogaerrhinum filipes* (CalPhotos, Antirrhinum filipes; The Jepson Herbarium, Antirrhinum filipes), and *Sairocarpus watsonii* (Consortium of Intermountain Herbaria Detailed Collection Record Information, Sairocarpus watsonii).

When species name used in the backbone phylogeny (Ogutcen and Vamosi, 2016) did not match any name in the monograph (Sutton, 1988), we determined synonymy from IPNI (International Plant Names Index) and the Plant List (The Plant List version 1.1). The list of synonyms determined are as follows: *Maurandya antirrhiniflora* (*Maurandella antirrhiniflora*), *Maurandya wislizeni* (*Epixiphium wislizeni*), *Neogaerrhinum strictum* (*Antirrhinum kelloggii*), *Sairocarpus costatus* (*Antirrhinum costatum*), and *Neogaerrhinum filipes* (*Antirrhinum filipes*).

Chapter 2: Testing the hypothesis that a conserved CYC–RAD module was co-opted to flower monosymmetry in Lamiales

Summary

A CYC–RAD–DIV–DRIF interaction defines the novel phenotype of flower monosymmetry in Lamiales. Solanales are sister to Lamaiales + Vahliaceae. Tomato (*Solanum lycopersicum*) is a model species in the order Solanales and has a putatively ancestral state of flower polysymmetry. In tomato, a RAD–DIV–DRIF interaction has been reported in the fruits, suggesting that a RAD–DIV–DRIF interaction was co-opted to the novel phenotype of flower monosymmetry. However, whether CYC was a part of this ancestral interaction and was co-opted with the rest of the program has been an open question. Here, we report a CYC–RAD regulatory interaction in tomato. Our data suggest that a CYC–RAD–DIV–DRIF interaction was likely present before the divergence of Solanales and Lamiales, and that the likely function of this genetic program was ovary/fruit development. The evolution of flower monosymmetry was likely facilitated by the co-option of the entire CYC–RAD–DIV–DRIF interaction.

Introduction

An outstanding question in evolutionary biology is how genetic interactions (programs, networks, pathways) that define novel phenotypes evolve. There are two ways genetic interactions that define novel phenotypes can evolve. First, by a *de novo* assembly during the origin of the novel phenotype, where previously non-interacting genes are assembled into a new interaction. Second, by co-option of an existing network to a new function of defining the novel phenotype.

Monosymmetry of flowers is a novel phenotype, and it has evolved at least 130 times during the diversification of flowering plants (Reyes et al., 2016), including in the order Lamiales. Monosymmetric flowers evolved at or near the base of the order Lamiales (Reeves and Olmstead, 2003; Reyes et al., 2016). In this order, a CYC–RAD–DIV–DRIF interaction defines flower monosymmetry (Almeida et al., 1997; Corley et al., 2005; Galego and Almeida, 2002a; Hileman and Baum, 2003; Luo et al., 1996, 1999; Raimundo et al., 2013). In the previous chapter, we provide evidence that the RAD–DIV–DRIF interaction, which is crucial in defining Lamiales flower symmetry, was likely co-opted from a different function. This function was likely in ovary/fruit development. However, whether CYC is a part of this interaction that pre-dates the origin of monosymmetric flowers, is an open question. Crucial to testing this hypothesis is determining whether a CYC–RAD interaction is present in lineages sister to the Lamiales.

Figure 14. A tomato flower

The order Solanales is sister to Lamiales + Vahliaceae (Stull et al., 2015). Given the sister relationship with Lamiales + Vahliaceae, Solanales is an ideal outgroup for testing whether a CYC–RAD interaction is ancestral to Lamiales + Solanales. Tomato, or *Solanum lycopersicum* (syn. *Lycopersicum esculentum*), is a model species in the order Solanales, and has polysymmetric flowers (**Figure 14**). This state of polysymmetry is putatively ancestral, and not derived from an ancestor with flower monosymmetry.

A RAD–DIV–DRIF interaction has been reported from tomato fruit development (Machemer et al., 2011). The RAD component in this interaction is an ortholog of *Am*RAD (Gao et al., 2017; Sengupta and Hileman, 2018), but the DIV component (Gao et al., 2017;

Sengupta and Hileman, 2018) and the DRIF component (Raimundo et al., 2013) are paralogs of *Am*DIV and *Am*DRIF1/2, respectively. The RAD component in this interaction is called *Sl*RADlike4 (or fruit SANT/MYB-like 1, *Sl*FSM1). *Sl*RADlike4 protein suppresses cell elongation in developing tomato fruit walls (Machemer et al., 2011). The DIV component is called *Sl*DIVlike5 (or *Sl*MYBI), and the DRIF component is called *Solanum lycopersicum* Fruit SANT/MYB Binding protein1 (*Sl*FSB1). The surprising similarity of this three-component interaction to that identified in the development of *Antirrhinum* flower monosymmetry raises the possibility that the ancestral role of RAD, DIV, and DRIF genes, prior to the divergence of Lamiales and Solanales, was likely in ovary/fruit development.

However, it is yet untested whether CYC is a part of the RAD–DIV–DRIF interaction in tomato. Key to understanding this question is to test whether there is a CYC–RAD regulatory interaction in tomato. If CYC orthologs in tomatoes transcriptionally regulate *RAD* orthologs, it would suggest that the CYC–RAD–DIV–DRIF interaction that defines flower monosymmetry in Lamiales, did not evolve *de novo* during the origin of flower monosymmetry, but it was co-opted *en bloc* from a different function.

We undertook virus-induced gene silencing to test for a CYC–*RAD* interaction in tomato, and find that a CYC–*RAD* interaction is present in tomatoes. This suggests that the CYC–RAD–DIV–DRIF interaction, that defines flower monosymmetry in Lamiales, was co-opted from a different ancestral function, possibly ovary and fruit development. The evolution of monosymmetric flowers in Lamiales, was thus likely facilitated by co-option of an existing genetic interaction. This adds to the evidence that novel phenotypes can evolve by co-opting existing genetic interactions.

Results

Expression of *SlTCP7*, *SlTCP26*, *SlRADlike4*, *SlDIVlike5*, and *SlDIVlike6* suggests potential interaction

We used quantitative real-time PCR to determine relative expression of the homologs of *Antirrhinum* flower symmetry genes in tomato (**Table 3**). We find that all of these genes, with the exception of *SlRADlike1*, are broadly expressed across tomato vegetative and reproductive tissues (**Figure 15**). Overlapping expression is an important criterion for genes to interact with each other. Interestingly, the expression of these genes overlap in ovaries and fruits, and is often high in those tissues. This suggests that these genes have a key role in ovary and fruit development. This is consistent with the previously described interaction of *Sl*RADlike4 and *Sl*DIVlike5 in tomato fruits where these two proteins compete for *Sl*FSB1 (Machemer et al., 2011), which is a paralog of *Am*DRIF1&2 (Raimundo et al., 2013).

Table 3. List of tomato genes tested with qPCR in this study and their homology to snapdragon flower symmetry genes

Snapdragon gene	Tomato ortholog	Tomato paralog
AmCYC, AmDICH	*SlTCP7, SlTCP26*	
AmRAD	*SlRADlike1, SlRADlike4*	
AmDIV, AmDIV-like1	*SlDIVlike6*	*SlDIVlike5*

A CYC–*RAD* regulatory interaction is present in tomato

There are two *AmCYC* orthologs in tomato—*SlTCP7* and *SlTCP26*. There are two *AmRAD* orthologs in tomato—*SlRADlike1* and *SlRADlike4*. We selected *SlTCP26* and *SlRADlike4* to test for a *CYC–RAD* interaction. We did not select *SlTCP7* because its expression is low in whole stage-20 flowers relative to other tissues (**Figure 15**) and downregulation was

difficult to assess in VIGS experiments (data not shown). We selected *SlRADlike4*, because *SlRADlike4* has an expression pattern similar to that of *SlTCP26* (unlike *SlRADlike1*) (**Figure 15**), making *SlRADlike4* a potential candidate for transcriptional regulation by *Sl*TCP26. Also,

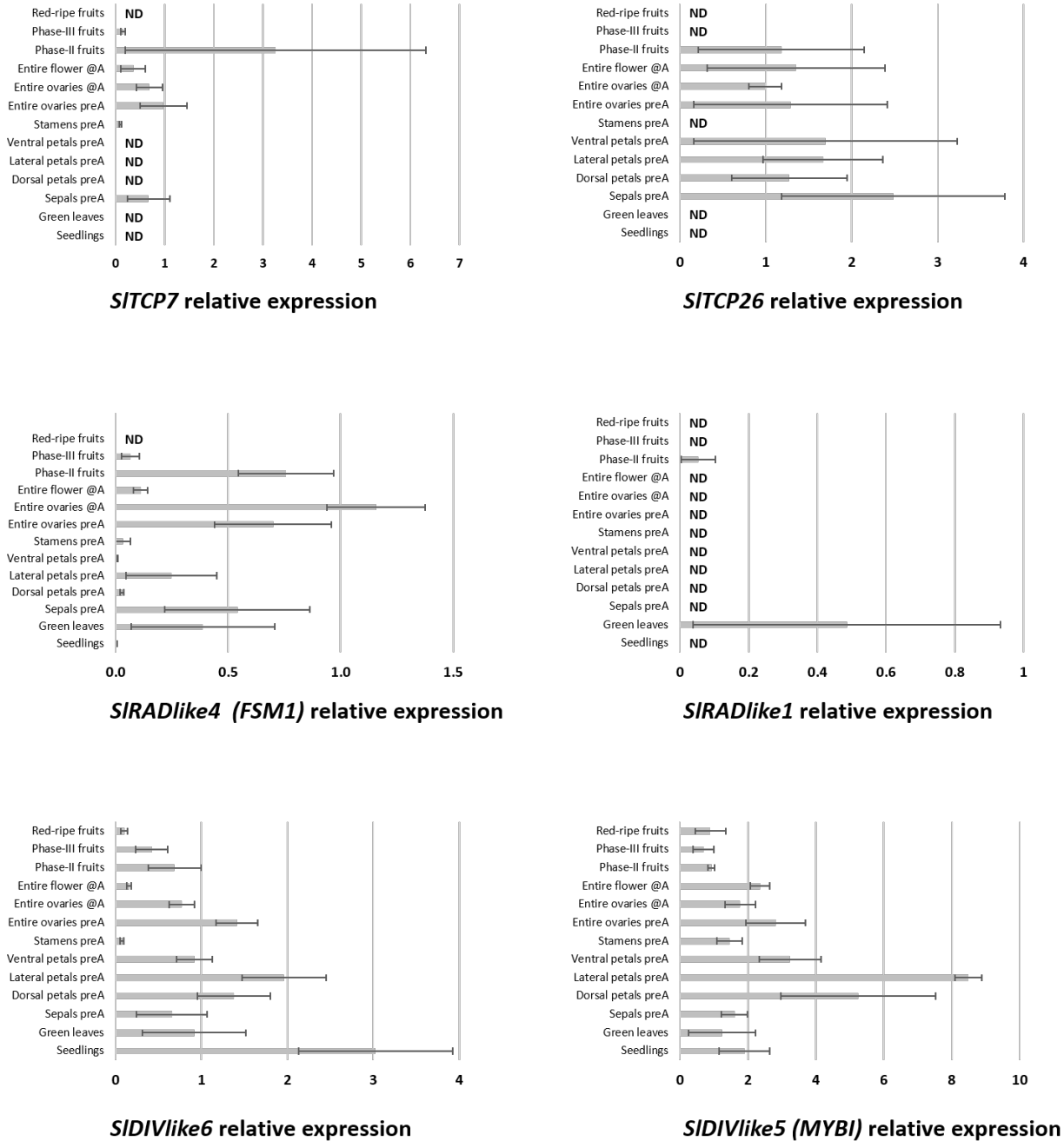

Figure 15. Relative expression of CYC, RAD and DIV orthologs and one DIV paralog in tomato across reproductive wildtype background. Error bars are standard deviations of samples. ND: expression not determinable; @A, at anthesis (stage-20); preA, pre-anthesis (stage-16).

SlRADlike4 has five predicted TCP-binding sites within the first 3000 kb upstream of its transcription start site, unlike *SlRADlike1*, which has none (**Table 4**). Our previous work has demonstrated that *RAD* genes that are known or predicted to be under the transcriptional control of CYC proteins are enriched in predicted TCP-binding sites in the first 3000 kb upstream of their transcription start sites (Sengupta and Hileman, 2018).

Table 4. Predicted TCP-binding sites upstream of *SlRADlike4* (5′ to 3′)

Sequence	DNA strand	bp upstream of transcription start
GGG**CCC**	anti-sense	300–295
GGG**CCC**	sense	300–295
GGTC**CC**	sense	761–756
GGAC**CC**	anti-sense	762–757
GGAC**CC**	anti-sense	1136–1131

We downregulated *SlTCP26* expression in tomato by employing virus-induced gene silencing (VIGS) (**Figure 16**). There was a concomitant decline in *SlRADlike4* expression in the same tissues (**Figure 16**). This provides strong evidence that *SlRADlike4* is under the transcriptional control of *Sl*TCP26. We predict this transcriptional control to be direct—*Sl*TCP26 likely binds to the predicted TCP-binding sites present upstream the transcription start site of *SlRADlike4* (**Table 4**). This provides evidence of a CYC–*RAD* regulatory interaction in tomato.

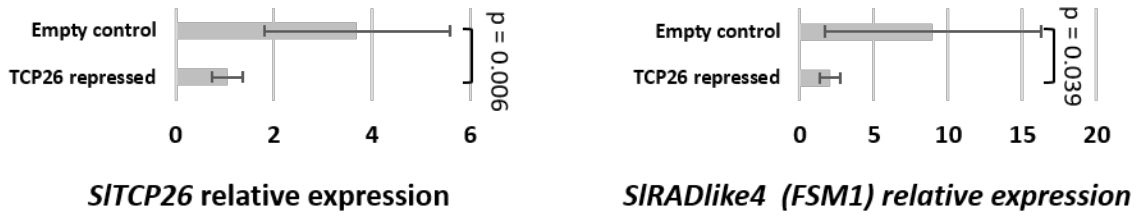

Figure 16. Downregulation of *SlTCP26* and its effect on *SlRADlike4.* Error bars are standard deviations of samples. The p-values are from T-tests performed on the bracketed tissues.

Discussion

SlTCP26 transcriptionally regulates *SlRADlike4* in tomato

Downregulating *SlTCP26* by virus-induced gene silencing leads to a corresponding decrease in *SlRADlike4* expression. This provides strong evidence for transcriptional control of *SlRADlike4* by *Sl*TCP26. However, our data do not provide evidence as to whether this interaction is direct (*Sl*TCP26 protein binding to the 5′ *cis*-regulatory sequence of *SlRADlike4*) or indirect (downstream targets of *SlTCP26* binding to the 5′ *cis*-regulatory sequence of *SlRADlike4*).

TCP proteins (similar to *Sl*TCP26) are known or predicted to be transcription factors that bind to the consensus TCP-binding site 5′–GGNCCC-3′ (Costa et al., 2005; Kosugi and Ohashi, 2002; Yang et al., 2012). *RAD* orthologs that are known or predicted to be under the direct transcriptional regulation by CYC orthologs are likely to be enriched in predicted TCP-binding sites in the first 3000 kb upstream their transcription start site (Sengupta and Hileman, 2018). *SlRADlike4* has five such predicted TCP-binding sites within the first 3000 kb upstream of its transcription start site. Together, the data from bioinformatics analysis and gene silencing experiments suggest that *Sl*TCP26 directly upregulates the transcription of *SlRADlike4*.

CYC-RAD-DIV-DRIF interaction was co-opted to flower monosymmetry from other functions

A CYC–RAD–DIV–DRIF interaction defines flower monosymmetry in Lamiales. A part of this interaction, RAD–DIV–DRIF interaction, is present in Solanales, and affects fruit development in tomato (Machemer et al., 2011). We provide evidence in Chapter 1 that the

RAD-DIV-DRIF interaction is conserved across Lamiales+Solanales ovary/fruit development. Here we report a CYC-*RAD* interaction in tomato, where *Sl*TCP26 transcriptionally upregulates *SlRADlike4* (**Figure 16**). This suggests that the entire CYC–RAD–DIV–DRIF interaction is likely ancestral to Lamiales+Solanales, and was co-opted *en bloc* to define the novel phenotype of flower monosymmetry in Lamiales. The likely ancestral function of this interaction was in ovary/fruit development.

The DIV and DRIF components involved in Lamiales flower monosymmetry and Solanales fruit development are not orthologous (Gao et al., 2017; Raimundo et al., 2013; Sengupta and Hileman, 2018). We hypothesize that the DIV and DRIF genes involved in the ancestral CYC–RAD–DIV–DRIF interaction underwent lineage-specific replacements by other homologs. Both DIV and DRIF genes are a part of multigene families in plants, making other homologs readily available (Gao et al., 2017; Raimundo et al., 2013; Sengupta and Hileman, 2018).

Existing genes are often recruited to define novel phenotypes (Stern, 2013; True and Carroll, 2002); this process is called co-option (or re-deployment). Co-option of single genes in defining novel phenotypes has been reported from a wide variety of organisms, including the co-option of CYC to define flower monosymmetry (Bharathan et al., 2002; Busch and Zachgo, 2007; Citerne et al., 2003; Spitz et al., 2001; Stern, 2013; Werner et al., 2010). However, genes usually do not function in isolation, but they function as a part of a larger genetic program where they interact with other genes to affect a phenotype. An immediate question arises: whether co-option involves single genes, or whether entire genetic programs are co-opted collectively. This question has received limited attention and mostly from work in animal systems (Bharathan et al., 2002; Brakefield et al., 1996; Sordino et al., 1995; True and Carroll, 2002). Our results add to

the evidence that evolution of novel phenotypes can be associated with or facilitated by the co-option of entire genetic programs.

Methods

Plant material

Tomato seeds of variety ‘micro-tom’ were used for this study. Seeds were broadcast on wet soil, covered with a transparent lid, and kept in short night conditions at 23–24° Celsius. We transferred the germinated seedlings to larger individual pots, maintained them in the same light and temperature regimen, treated them for arthropod pests and fertilized them with chemical fertilizer once a week.

Tissue sampling

We collected the following tissues for qRT-PCR analyses: seedlings (cotyledons fully expanded, first two true leaves are 0.5 cm long , green leaves (leaf 3 mm long, three leaflets visible, green, with no dark pigmentation on the underside), stage-16 sepals, stage-16 dorsal petals, stage-16 lateral petals, stage-16 ventral petals, stage-16 stamens, stage-16 carpels, stage-20 carpels, stage-20 whole flowers, phase-II fruits, phase-III fruits, and red-ripe fruits. The flower stages (i.e., stage-16 and stage-20) were determined based on descriptions in a previous study (Brukhin et al., 2003). Stage-20 corresponds to flowers at anthesis (recently opened). Solanaceae flowers are slightly rotated relative to the main axis (Knapp, 2004), and, further, the carpels are oblique relative to the median plane of the flower (Craene, 2010; Murray, 1945), making the dorso-ventral axis difficult to identify. The dorso-ventral axis in flowers were determined based on previously published illustrations (Knapp, 2004). Tomatoes also have a

sympodial inflorescence architecture, making it challenging to determine the dorso-ventral axis in the second flower onwards. The dorso-ventral axes in such flowers were determined based on previously published illustrations (Hake, 2008). The fruit stages (i.e., phase-II, phase-III, and red-ripe) were determined from a previous study (Gillaspy et al., 1993). Phase-I 'fruits' (Gillaspy et al., 1993) correspond to the carpels of stage-20 flowers.

Identifying orthologs

The following genes were studied in this work: *SlTCP7* (Solyc02g089830.1.1), *SlTCP26* (Solyc03g045030.1), *SlRADlike1* (Solyc01g109670.2.1), *SlRADlike4* (Solyc10g052470.1), *SlDIVlike6* (Solyc06g076770.2.1), and *SlDIVlike5* (Solyc05g055240.2.1). Orthology was identified from supplementary figures 1 and 2. The methods for estimating the phylogenetic history of these genes are described in chapter 3 (and in Sengupta and Hileman, 2018).

Quantitative RT-PCR assays

We sampled three biological replicates for each tissue type. We extracted Total RNA using RNeasy plant minikit (Qiagen, Germantown, MD, USA) or TRI Reagent (Thermo Fisher Scientific, Waltham, MA, USA), followed by DNase treatment (TURBO DNase, Thermo Fisher Scientific), and cDNA synthesis (iScript cDNA Synthesis Kit, Bio-Rad, Hercules, CA, USA). We performed qRT-PCR in a StepOnePlusTM Real-Time PCR System (Thermo Fisher Scientific) using Bullseye EvaGreen qPCR Mastermix (Midwest Scientific, St. Louis, MO, USA. We normalized expression of target genes against a constitutively expressed gene *Elongation factor 1-alpha* (*EF1a*). We analyzed expression employing ΔΔCt method. We determined the

efficiency of PCR using DART (Peirson et al., 2003). The primers used in qRT-PCR are as listed in **Table 5**.

Table 5. Tomato qRT-PCR primers (5′ to 3′)

Gene	Forward primer	Reverse primer
SlEF1a	TACTGGTGGTTTTGAAGCTG	AACTTCCTTCACGATTTCATCATA
SlTCP7	GGAACCCTATTCTTCACTCCTC	TTGAGTGACCATTTGCGGCT
SlTCP26	TCTTGGTTTCACTGGCAACC	TCTTCATGGGGAACGACCTT
SlDIVlike5	AGACGAACCACCTGCAATCA	TCGATCTTTAAGCTCCTGGATTCA
SlDIVlike6	ACGGTCTCTCTTGTCTTGTGA	ACAGGTTTCGGTTCGTTCCT
SlFSM1	ATTTGCCTTGGAACCTGCCT	GGACATGACGAGTACAAGAGCA
SlRADlike1	GGTTCCAACAAAGCAATGAGGG	GCTTCCTTCATATAGTACATCCATGA

Virus-induced gene silencing

We used a pRTV1 and pTRV2 based system to downregulate *SlTCP26* (Dinesh-Kumar et al., 2003; Liu et al., 2002; Padmanabhan and Dinesh-Kumar, 2009). This gene has two alternative transcripts (HM921069.1 and XM_010319513.2). We designed a construct that would silence both transcripts. We amplified a 416 bp fragment of the *SlTCP26* transcript from cDNA using the primers 5′–TCTCTAGAAGGCCTCCATGGTGAAACTAGCCACAAATC–3′ and 5′–TCTTCGGGACATGCCCGGGCTTAGATTGAAGAAGATGACG–3′ and cloned it into pTRV2 using NEBuilder HiFi DNA Assembly Master Mix (New England Biolabs. Ipswich, Massachusetts, USA). The underlined portions of the primers are overhangs that facilitate cloning into pTRV2. The fragment encompasses both coding and non-coding parts of the transcript. We used *Agrobacterium tumefaciens* GV3101 to introduce the pTRV1/2 into tomato seedlings (as described in Dinesh-Kumar et al., 2003). As a control, we infiltrated some plants with the empty pTRV2 vector (without the insert) along with the pTRV1. We sampled flowers at

anthesis (stage-20) to test for downregulation (using extraction and qRT-PCR methods discussed above). Six pTRV2-insert flowers and eight control flowers (from different plants) were used for testing downregulation of *SlTCP26* and *SlRADlike4*. We compared the mean *SlTCP26* expression of these genes in the control and experimental sets by T-test.

Chapter 3: Novel traits, flower symmetry, and transcriptional autoregulation: new hypotheses from bioinformatic and experimental data

Summary

A common feature in developmental networks is the autoregulation of transcription factors which, in turn, positively or negatively regulate additional genes critical for developmental patterning. When a transcription factor regulates its own expression by binding to *cis*-regulatory sites in its gene, the regulation is direct transcriptional autoregulation (DTA). Indirect transcriptional autoregulation (ITA) involves regulation by proteins expressed downstream of the target transcription factor. We review evidence for a hypothesized role of DTA in the evolution and development of novel flowering plant phenotypes. We additionally provide new bioinformatic and experimental analyses that support a role for transcriptional autoregulation in the evolution of flower symmetry. We find that 5' upstream non-coding regions are significantly enriched for predicted autoregulatory sites in Lamiales *CYCLOIDEA* genes—an upstream regulator of flower monosymmetry. This suggests a possible correlation between autoregulation of *CYCLOIDEA* and the origin of monosymmetric flowers near the base of Lamiales, a pattern that may be correlated with independently derived monosymmetry across eudicot lineages. We find additional evidence for transcriptional autoregulation in the flower symmetry program, and report that *Antirrhinum DRIF2* may undergo ITA. In light of existing data and new data presented here, we hypothesize how *cis*-acting autoregulatory sites originate, and find evidence that such sites (and DTA) can arise subsequent to the evolution of a novel phenotype.

Introduction

A common feature in developmental networks is the autoregulation of transcription factors which, in turn, positively or negatively regulate additional genes critical for developmental patterning. A *trans*-acting protein is considered transcriptionally autoregulated when the protein itself, or downstream factors, modulate its expression. Transcriptional autoregulation can be either direct, or indirect. In direct transcriptional autoregulation (DTA), a protein binds to *cis*-regulatory sites in its gene and modulates expression. Indirect transcriptional autoregulation (ITA) involves regulation by proteins expressed downstream of the target transcription factor (**Figure 17**). Both DTA and ITA have the potential to enter run-away positive feedback processes. Expression of such genes is likely reduced or stabilized by additional regulatory factors. Transcription factor autoregulation is widespread. For example, at least 40% of transcription factors in *Escherichia coli* are autoregulated (Rosenfeld et al., 2002), and similar direct and indirect autoregulation has been reported across the tree of life—in viruses, prokaryotes, and eukaryotes (for example, Hochschild, 2002; Martínez-Antonio and Collado-Vides, 2003; Holloway et al., 2011; Tao et al., 2012; Gallo-Ebert et al., 2013; and reviewed in Bateman, 1998; Crews and Pearson, 2009), including those with complex development (for example, Cripps et al., 2004; Holloway et al., 2011; Ye et al., 2016). DTA has been demonstrated in processes as diverse, and crucial as the origin of certain cancers (Pasqualucci et al., 2003), and the onset of flowering (Tao et al., 2012).

The widespread occurrence of transcription factor autoregulation suggests a beneficial role in the function and evolution of genetic programs. Here we provide a review of evidence for

DTA in key flowering plant developmental programs. We provide new data supporting the hypothesis that DTA facilitated the evolution of flower monosymmetry in Lamiales. Together these data provide compelling evidence for the hypothesis that DTA plays a role in facilitating

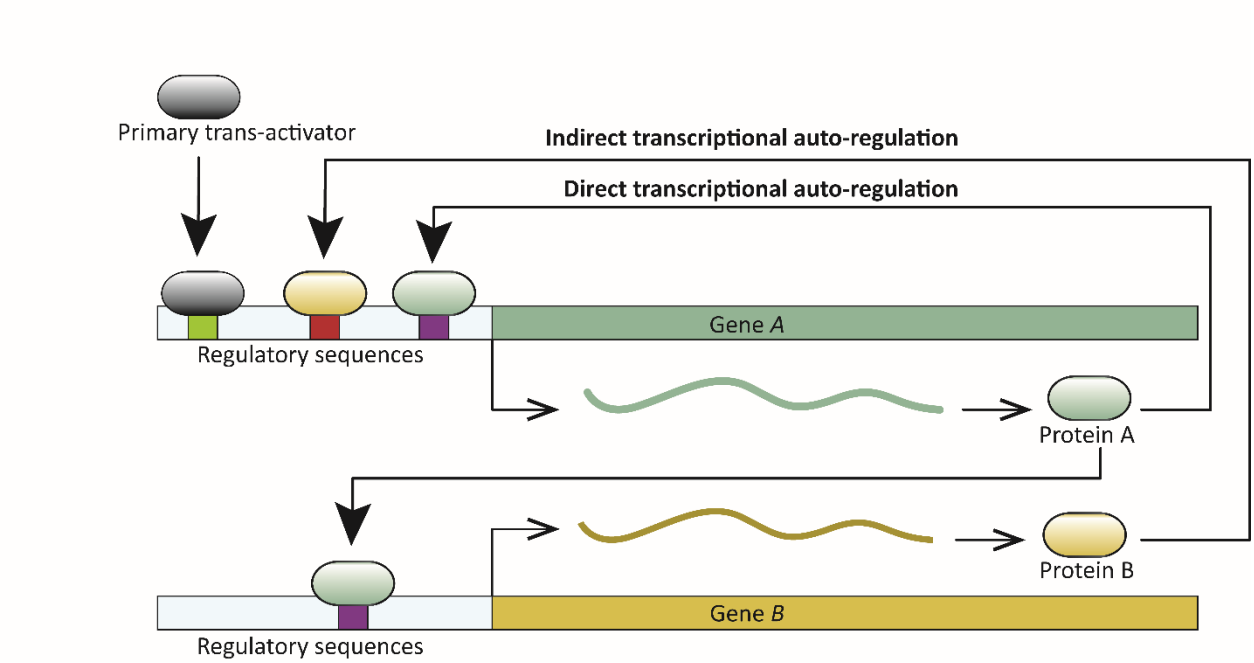

Figure 17. Schematic representation of direct and indirect transcriptional autoregulation involving two transcription factors. Gene A undergoes DTA and also regulates transcription of gene B. In turn, Gene A undergoes ITA when B regulates its transcription.

the evolution of novelty.

Advantages of autoregulation

Several models suggest that autoregulation, especially DTA, can maintain a steady level of expression independent of other factors. If so, genes that are more likely to be autoregulated should be those that experience fleeting regulatory signals, or are positioned upstream in genetic regulatory networks with crucial developmental functions (Crews and Pearson, 2009; Singh and Hespanha, 2009). For example, several transcription factors involved in antibiotic resistance are reported to be autoregulated, resistance being a crucial phenotype (Hoot et al., 2010; Toth Hervay et al., 2011). Similarly, entering or exiting lytic and lysogenic stages is a key

developmental decision in lambda bacteriophages, and this decision is partly controlled by the autoregulation of a transcription factor, CI (Hochschild, 2002). The prediction that transcription factors upstream in regulatory networks are more likely to undergo autoregulation has been tested in the model eukaryote yeast, *Saccharomyces cerevisiae*. In yeasts, where all possible transcription factor interactions have either been tested or predicted, master regulatory genes are significantly more likely to experience autoregulation than are other regulators (Odom et al., 2006). Similarly, five out of six master regulatory genes in human hepatocytes bind to their own promoters, *i.e.*, undergo DTA (Odom et al., 2006).

How regulatory networks define stable phenotypes is an important question in evolution and development. Simulations of developmental network evolution suggest that autoregulated genes are more robust when faced with random mutations and environmental perturbation (Pinho et al., 2014). The model that DTA stabilizes expression by reducing system noise has been tested in the gene *hunchback* in *Drosophila melanogaster*. Models where the HOX transcription factor Hunchback binds to the *hunchback* promoter (*i.e., hunchback* undergoes DTA) predict less promoter binding-unbinding noise, making the system more robust (Holloway et al., 2011). Experimental work in *hunchback* mutants whose protein cannot bind to DNA (hence, cannot undergo DTA) supports this prediction (Holloway et al., 2011).

In addition to enhancing system robustness, autoregulation provides a mechanism for maintaining expression through key stages of development (reviewed below) that are potentially critical for patterning phenotype. However, the developmental role of DTA has only been tested by mutational studies in a handful of cases. To determine the role of transcription factor DTA, the direct binding between the protein product of a gene and that gene's *cis*-regulatory DNA can be either intensified or weakened through direct DNA manipulation. For example, addition or

deletion of *cis*-regulatory self-binding sites can be used to test for the specific developmental role of DTA within a given species (Espley et al., 2009; Tao et al., 2012; Gallo-Ebert et al., 2013). A complementary, but more difficult approach is to alter transcription factor peptide sequence by mutagenesis in order to modify affinity towards the self-binding sites, *e.g.,* in the *hunchback* exampled discussed above (Holloway et al., 2011). In some model systems, it is possible to repress activity of a transcription factor by overexpressing a dominant chimeric version of the peptide with a repressor domain added to the carboxy-terminus. The chimeric protein can repress the function of the native transcription factor by competitive inhibition (for example, Hiratsu et al., 2003; Koyama et al., 2010). Recent advances in CRISPR/Cas9 gene-editing technologies (Ma et al., 2016, 9) will certainly facilitate exploration of DTA function, at least in model species.

Review of DTA in flowering plant developmental evolution

Once an initial signal for activation of gene expression has been received, a transcription factor capable of DTA can contribute to swift developmental decisions. A clear example comes from work on the developmental transition to flowering (**Figure 18**). Flowering time is a key life-history transition in plant development, intimately tied to environmental cues and aging in order to ensure reproductive success (reviewed in Ó'Maoiléidigh et al., 2014). In *Arabidopsis thaliana*, the transition from vegetative to reproductive development is regulated in part by a MADS-box transcription factor, SUPPRESSOR OF OVEREXPRESSION OF CONSTANS 1 (SOC1). SOC1 undergoes DTA through the binding of SOC1 protein to four *cis*-regulatory *CArG-box* self-binding sites close to the SOC1 transcription start site (Tao et al., 2012). The flowering transition is significantly delayed in the insertional mutant *soc1-2* which carries a loss-

of-function mutation in the coding sequence of *SOC1*. The delayed flowering phenotype is largely rescued when *soc1-2* lines are transformed with a wild type *SOC1* allele (including the wild type promoter). This mutant-rescue system with known self-binding sites in the *SOC1* promoter creates an elegant system for testing the specific role of *SOC1* DTA in establishing tight control of the flowering time phenotype. In heterozygous rescue lines where the self-binding sites in the transgenic allele have been mutated by substituting nucleotides at the first two and last two positions of the *CArG-box* binding site, flowering is delayed (Tao et al., 2012).

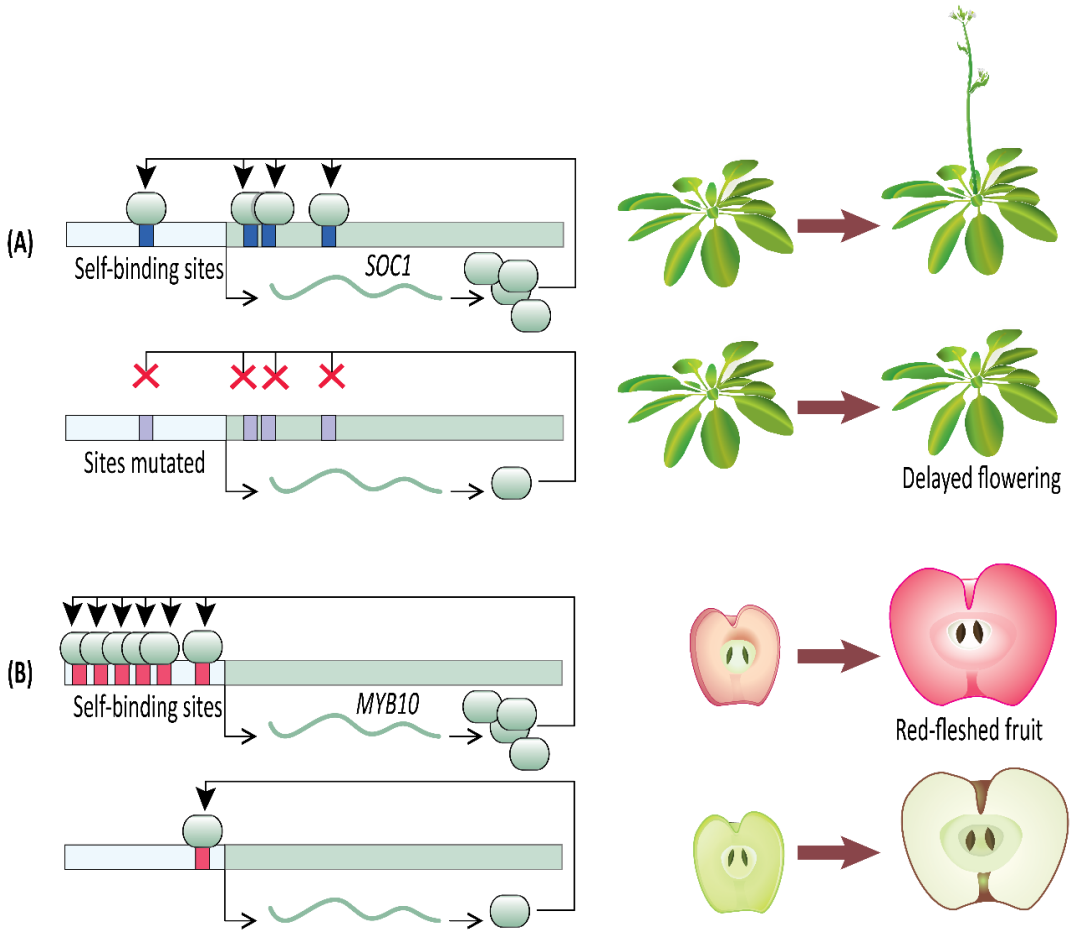

Figure 18. DTA in novel plant phenotypes. (A) Disrupting DTA by removal of *cis*-acting autoregulatory sites in *Arabidopsis SOC1* delays onset of flowering. (B) The number of autoregulatory sites in apple *MYB10 cis*-regulatory sequence is correlated with fruit flesh color.

This suggests that the DTA of SOC1 has a key role in transition to flowering. Tao et al. (2012) provide further evidence of SOC1 autoregulation using an estradiol-inducible expression system. Estradiol-induction allows tight control over transgenic protein entering the nucleus and functioning as a transcription factor. Within two hours of estradiol-induction of transgenic SOC1, expression of endogenous *SOC1* tripled in comparison to a control. This rapid increase in *SOC1* expression after releasing transgenic SOC1 protein to the nucleus suggests SOC1 plays a direct role in its own upregulation. Together, these SOC1 experiments in *Arabidopsis* provide clear evidence that once induced, a transcription factor undergoing DTA can rapidly increase its expression level to swiftly respond to a signal and affect developmental outcomes.

Sustained, stable, and high expression is likely key to defining complex phenotypes. Other than increasing the expression level at a certain point during development (as described in SOC1 above), DTA would provide selective advantage if it could sustain the expression for an extended time through consecutive developmental events. A way to test this would be to determine how expression changes when homologous autoregulatory and non-autoregulatory sites between a pair of recently diverged paralogs are swapped. *Arabidopsis* APETALA1 (*At*AP1) and CAULIFLOWER (*At*CAL) are two recently duplicated paralogs (Wang et al., 2012) and this system was employed by Ye et al. (2016) to test the role of DTA for sustaining expression in developmental patterning. *At*AP1 defines sepal development, and Ye et al. (2016) found that strong expression of *AtAP1* is initiated in floral meristems, and that the expression continues to near-mature flower stages (stage-12). *At*AP1 also undergoes DTA wherein it binds to a *CArG-box* located in its *cis*-regulatory region and activates *AtAP1* transcription. On the other hand, *AtCAL* does not undergo DTA, is expressed at a low level in early stage flowers, with the expression vanishing soon after stage-4 (Ye et al., 2016). In an elegant system, Ye et al. (2016)

generated β-glucuronidase (GUS) reporter-constructs driven by *AtAP1* and *AtCAL* promoter regions. When the *CArG-box* in the *GUS* reporter construct with the *AtAP1* promoter was replaced with the homologous non-autoregulatory nucleotides from the *AtCAL* promoter, two changes occurred. First, the overall expression level of *GUS* dropped, and second, the expression duration was shortened, approximating that of *AtCAL* in wild type plants. On the contrary, when *GUS* was placed under the control of an *AtCAL* promoter whose non-autoregulatory nucleotides had been replaced with the homologous *CArG-box* from the *AtAP1* promoter, *GUS* expression level increased and extended to near-mature stage flowers. This suggests that DTA of *AtAP1* not only has a role in maintaining high expression levels compared to the non-autoregulated paralog, but has a critical role in sustaining the expression for an extended period. This study did not directly test the role of *AtAP1* DTA, its loss or acquisition, in defining phenotype. However, direct evidence for acquisition or loss of DTA on the evolution of a novel phenotype comes from domesticated apples.

Malus domestica (domesticated apple) provides compelling evidence for the importance of DTA on phenotypic outcomes (**Figure 18**). The color of fruit flesh in many domesticated apple varieties ranges from white to red. Variation in fruit color is regulated by the transcription factor MYB10, which upregulates anthocyanin expression, especially cyanidin-3-galactoside (Espley et al., 2007, 2009, 2013). Anthocyanin-regulating MYBs have been reported from a wide variety of angiosperm species (reviewed in Lin-Wang et al., 2010), including *Malus* (Espley et al., 2009), *Prunus* (Starkevič et al., 2015), *Myrica* (Niu et al., 2010), *Arabidopsis* (Gonzalez et al., 2008), and *Ipomoea* (Mano et al., 2007). *Malus domestica* has two alleles of *MYB10* that are identical in their coding sequences but differ in their promoter sequences. Allele *R1* promoter contains one MBY10 autoregulatory binding site, whereas allele *R6* promoter

contains six repeats of the autoregulatory site (Espley et al., 2009). The white-fleshed domestic apple varieties are homozygous for the one-repeat *R1* allele, whereas the red-fleshed varieties are *R1*/*R6* heterozygotes or *R6/R6* homozygotes, which leads to increased anthocyanin production via DTA (Espley et al., 2009). It is not clear which allele is ancestral in domesticated apples. Of the four *Malus* species that contributed to the domesticated apple genome (Cornille et al., 2014), *M. sieversii* can be either *R6*/*R6* (Espley et al., 2009; Lin-Wang et al., 2010, 3) or *R1*/*R6* (Espley et al., 2009; Nocker et al., 2012), and *M. baccata* is *R1*/*R1* (Nocker et al., 2012). Of the other species in the genus *Malus* tested for *MYB10* promoter sequence, all but one have the *R1*/*R1* genotype (Nocker et al., 2012).

Though it is not clear whether *R1*/*R1* (white flesh) or *R6*/*R6* (red flesh) is ancestral in the genus *Malus*, it is clear from studies in domesticated apple that changes to fruit flesh color are regulated by addition or loss of autoregulatory sites in the *MYB10* promoter. The evidence from flesh coloration in apples suggests an interesting possibility. Self-activating loops of DTA can serve as easy modules for evolving elevated or reduced gene expression levels. Such evolutionary shifts in gene expression have potentially adaptive developmental consequences accompanied by minimal pleiotropy. Genes, including transcription factors, are often regulated by *trans*-activators that bind to the *cis*-acting elements in the regulatory region of the target gene. Theoretically, these target genes can be upregulated in three ways: adding more *cis*-regulatory sites recognized by either the existing or novel *trans*-activators, upregulating the expression of the existing *trans*-activators, or acquiring new (or additional) self-binding sites in the promoter region. Addition of *cis*-regulatory sites recognized by *trans*-activators can be ineffective if the expression level of the *trans*-activator is limiting. Additionally, increasing the expression level of the *trans*-activator can have pleiotropic consequences. However, acquiring new (or additional)

cis-regulatory self-binding sites can lead to increased expression of the target gene while bypassing the limitations associated with *trans*-activation. Similarly, reduced expression levels can evolve with minimal pleiotropic consequences through the loss of existing autoregulatory sites.

The evidence from SOC1, *At*AP1, and MYB10 provide insight into why genes involved in defining novel phenotypes are likely to undergo DTA. Autoregulatory loops can serve as a quick developmental switch that can rapidly respond to an inbound signal, they can provide high expression levels, and extend that expression through consecutive developmental events. Lastly, DTA can act as a module that can be used to evolve increased or decreased expression with minimal pleiotropic effect, allowing the evolution of novel phenotypes that require such directional changes in protein levels. Quick evolutionary shifts in developmental function of paralogs and divergent alleles can therefore occur through gain or loss of DTA, most likely through gain or amplification of self-binding sites in *cis*-regulatory sequences of focal genes.

Evidence for DTA in flower symmetry evolution

An emerging system for studying the role of DTA in both development and evolution is flower symmetry. DTA has been implicated in the control of monosymmetry (bilateral symmetry; zygomorphy) (Yang et al., 2012), and may represent a critical step for the evolution of this floral novelty. Monosymmetric flowers are considered a key innovation defining flower form in many species-rich flowering plant lineages including Lamiales, asterids, legumes, and orchids (Sargent, 2004; Vamosi and Vamosi, 2010). Therefore, assessing the role of DTA in the development of flower monosymmetry may provide critical insights into patterns of gene network modification that facilitate novel trait evolution. Below, we review the genetic control

of monosymmetry in Lamiales alongside the evidence for DTA. We test for previously unreported regulatory interactions in the *Antirrhinum majus* flower symmetry program, as well as the potential for DTA in a major radiation of taxa with primarily monosymmetric flowers, the Lamiales. Lastly, we comment on possible wide-spread DTA in repeated origins of monosymmetry across flowering plants.

Flowering plants are ancestrally polysymmetric (radially symmetric; actinomorphic; **Figure 2**) (Sauquet et al., 2017). Evolutionary shifts away from polysymmetry include asymmetry (no axis of flower symmetry) and disymmetry (two non-equivalent axes of flower symmetry), but monosymmetry (a single axis of flower mirror-image symmetry; **Figure 2**) is by far the most common form of non-radial symmetry in flowering plants. Monosymmetric flowers have evolved at least 130 times independently during flowering plant diversification (Reyes et al., 2016). The role of floral symmetry in pollination was recognized as early as 1793 by Sprengel in his monumental German work *Das entdeckte Geheimniss der Natur im Bau und in der Befruchtung der Blumen* (reviewed in the following: Endress, 1999; Fenster et al., 2004, 2009; Neal et al., 1998). Monosymmetric flowers are often associated with specialized pollination by animals (Kampny, 1995; reviewed in Neal et al., 1998), rarely in wind pollinated species (rarely in wind pollinated species, as shown by Yuan et al., 2009), and transitions to monosymmetry are strongly associated with increased speciation rates (O'Meara et al., 2016; Sargent, 2004).

The genetics of monosymmetry is best understood in the model species *A. majus* (snapdragon, Lamiales). The flowers of *A. majus* have two distinct morphological regions—the dorsal (top; adaxial) side, and the ventral (bottom; abaxial) side (**Figure 19**). Monosymmetry of *A. majus* flowers along the dorso-ventral axis is defined by a competitive interaction involving

TCP and MYB transcription factors. TCP (TEOSINTE BRANCHED1, CYCLOIDEA, and PROLIFERATING CELL FACTORS) and MYB (first described from avian myeloblastosis virus) proteins are found as large gene families in flowering plants (Yanhui et al., 2006; Martín-Trillo and Cubas, 2010) and play diverse roles in aspects of vegetative and reproductive developmental patterning (Ambawat et al., 2013; Martín-Trillo and Cubas, 2010; Parapunova et al., 2014). The dorsal side of *A. majus* flowers is defined by the combined action of two recently duplicated TCP paralogs, CYCLOIDEA (*Am*CYC) and DICHOTOMA (*Am*DICH) (Corley et al., 2005; Hileman and Baum, 2003; Luo et al., 1996, 1999). These two transcription factors define dorsal flower morphology partly by activating the transcription of a downstream MYB protein, RADIALIS (*Am*RAD; **Figure 19**) (Corley et al., 2005). *Am*RAD post-translationally negatively regulates another MYB protein, DIVARICATA (*Am*DIV), which defines ventral flower morphology. Through this negative interaction, *Am*RAD excludes the ventral flower identity specified by *Am*DIV from the dorsal side of the developing *A. majus* flower (**Figure 19**). Specifically, *Am*RAD and *Am*DIV compete for interaction with two MYB-family protein partners called DIV and RAD Interacting Factors 1 and 2 (*Am*DRIF1 and *Am*DRIF2) (Almeida et al., 1997; Corley et al., 2005; Galego and Almeida, 2002b; Raimundo et al., 2013). *Am*DIV requires protein-protein interaction with *Am*DRIF1 or 2 to function as a transcription factor and upregulate its own transcription, as well as to regulate downstream targets (**Figure 19**) (Perez-Rodriguez et al., 2005; Raimundo et al., 2013). In the dorsal flower domain, *Am*RAD outcompetes *Am*DIV for interaction with *Am*DRIF1/2, preventing accumulation of *Am*DIV protein (Raimundo et al., 2013).

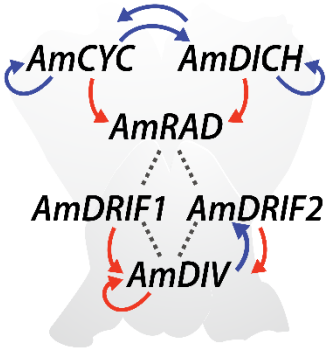

Figure 19. Regulatory mechanisms involved in *A. majus* flower symmetry. Previously reported transcriptional regulation (red arrows), transcriptional regulation predicted in this study (blue arrows), and previously reported protein-protein interactions (dotted lines) are shown.

Because flower monosymmetry has evolved multiple times, a considerable amount of effort has gone into testing whether elements of the *A. majus* symmetry program function to specify dorso-ventral differentiation in other flowering plant lineages. Interestingly, all monosymmetric species tested at a molecular level so far show evidence that a TCP-based regulatory network is likely involved in differentiation along the dorso-ventral flower axis. These studies span eudicot and monocot lineages and primarily, but not exclusively, show a pattern of dorsal-specific floral expression of *TCP* homologs (for example, Bartlett and Specht, 2011; Busch and Zachgo, 2007; Chapman et al., 2012; Citerne et al., 2003, 2010; Damerval et al., 2013; Howarth et al., 2011; Preston and Hileman, 2012; Wang et al., 2008; Yuan et al., 2009) (and reviewed in Hileman, 2014). In core eudicots, there are three lineages of *CYCLOIDEA* (*CYC*)-like TCP genes resulting from two rounds of duplication near the origin of core eudicots: the *CYC1*-, *CYC2*-, and *CYC3*-lineages (Citerne et al., 2013; Howarth and Donoghue, 2006). *AmCYC* and *AmDICH* belong to the *CYC2*-lineage, and in an interesting pattern, all TCP genes implicated in floral monosymmetry in core eudicots belong to the same *CYC2*-lineage (Citerne et al., 2010; and reviewed in Hileman, 2014). How these orthologous genes were recruited

convergently during the multiple evolutionary origins of floral monosymmetry, from an as yet unclear function in species with ancestral polysymmetry, remains an open question.

Detailed developmental studies in *A. majus* have provided key insights into the regulatory interactions that shape flower monosymmetry, and *A. majus* as a model represents a species-rich lineage of flowering plants, Lamiales. Monosymmetry evolved early in Lamiales diversification (Reyes et al., 2016; Zhong and Kellogg, 2015b), and developmental genetic studies in additional Lamiales species provide further insight into the regulatory network that shapes bilateral flower symmetry across the entire lineage. Notably, detailed expression and functional studies of *CYC*, *RAD* and *DIV* orthologs in Gesneriaceae, a sister lineage to the bulk of Lamiales species diversity, have contributed to a fuller understanding of regulatory interactions that shape Lamiales flower monosymmetry (Citerne et al., 2000; Gao et al., 2008; Liu et al., 2014; Smith et al., 2004; Yang et al., 2010, 2012; Zhou et al., 2008, 2). From studies in *A. majus* (Plantaginaceae) and *Primulina heterotricha* (syn. *Chirita heterotricha;* Gesneriaceae), there is strong evidence that at least two components of the flower symmetry network undergo DTA–*DIV* and *CYC* (**Figure 19**).

As mentioned above, AmDIV forms heterodimers with *Am*DRIF1 and 2 to specify ventral flower identity in *A. majus* (Raimundo et al., 2013). *Am*DIV-*Am*DRIF dimers bind to a consensus sequence that includes the conserved *I-box* motif, 5′-GATAAG-3′ located 2596 bp upstream of the *AmDIV* transcription start site (Raimundo et al., 2013), providing compelling evidence that AmDIV is involved in an autoregulatory loop. Autoregulation of *DIV* orthologs has not been tested outside of *A. majus*. In *P. heterotricha,* peloric (radialized) forms due to flower ventralization have reduced expression levels of *CYC* orthologs, *PhCYC1C* and *PhCYC1D* (Yang et al., 2012), presenting strong evidence that these two genes define dorsal identity of

monosymmetric *P. heterotricha* flowers. Experimental evidence suggests that *Ph*CYC1 and *Ph*CYC2 undergo DTA; PhCYC1 and PhCYC2 proteins bind to the consensus TCP-binding sequence 5′-GGNCCC-3′ in the putative promoter regions of both *PhCYC1* and *PhCYC2* (Yang et al., 2010, 2012). Autoregulation of *CYC* orthologs has not been tested outside of *P. heterotricha*.

These initial insights from *A. majus* and *P. heterotricha* lead to a set of important evolutionary questions. Is autoregulation of *CYC* orthologs conserved across Lamiales? And has a pattern of autoregulation repeatedly evolved in *CYC2*-lineage orthologs from lineages with independently derived monosymmetric flowers? This second question is especially compelling given that *CYC2*-lineage ortholog expression is expected to persist from early through later stages of flower development in order to specify asymmetric morphological differentiation along the dorso-ventral floral axis in lineages with flower monosymmetry.

Results

Predicted TCP- and DIV-binding sites in *A. majus* are consistent with known and hypothesized transcriptional regulation

In *A. majus,* we found consensus TCP-binding sites in four of the six genes known to be involved in *A. majus* flower symmetry (**Figure 19**,

Table 6). *AmCYC* and *AmDICH* had eight and four predicted TCP-binding sites in their upstream non-coding sequences, respectively, and likely regulate their own and each other's expression. Notably, *AmCYC* DTA has been hypothesized previously (Costa et al., 2005), and the presence of predicted autoregulatory sites in *AmCYC* and *AmDICH* is consistent with the putative auto and cross-regulation of *P. heterotricha PhCYC1C* and *PhCYC1D* (Yang et al., 2012). *AmRAD*,

known to be positively regulated by *Am*CYC and *Am*DICH (Corley et al., 2005; Costa et al., 2005), had two predicted consensus TCP-binding sites in its upstream non-coding sequence. *AmDIV* and *AmDRIF2* did not have predicted TCP-binding sites in their upstream non-coding sequences, consistent with evidence that they are unlikely to be under direct transcriptional regulation by *Am*CYC, *Am*DICH, or any other more distantly related TCP transcription factors.

Table 6. Predicted consensus TCP-binding sites in the upstream non-coding sequences of *A. majus* flower symmetry genes. Bases in bold indicate conservation in the consensus binding site. *AmDIV and AmDRIF2* lack consensus TCP-binding sites in their upstream non-coding sequences. Costa et al., 2005 reported TCP-binding sites for *AmRAD* and suggested the presence of autoregulatory sites in the non-coding sequence upstream of *AmCYC*.

Gene	Sequence	DNA strand	bp upstream of transcription start
AmCYC	**GGGCCC**	sense	2454-2457
	GGGCCC	sense	544-549
	GGGCCC	anti-sense	544-549
	GGGCCC	anti-sense	2452-2457
	GGCCCC	sense	2451-2456
	GGCCCC	sense	2292-2297
	GGCCCC	sense	543-548
	GGCCCC	anti-sense	2453-2458
AmDICH	**GGGCCC**	sense	1170-1175
	GGGCCC	anti-sense	1170-1175
	GGCCCC	sense	965-970
	GGCCCC	anti-sense	1171-1176
AmRAD Costa et al., 2005	**GGCCCC**	sense	1521-1526
	GGCCCC	sense	1489-1494
AmDRIF1	**GGTCCC**	anti-sense	2394-2399

Consensus TCP-binding sites (plus 100 bp flanking sequence from either side) initially identified in the upstream non-coding sequences of *AmCYC* and *AmDICH* were used to search for similar sites elsewhere in the *A.majus* genome. These searches resulted in only self-hits to *AmCYC* and *AmDICH* upstream non-coding sequences or cross-paralog matches between *AmCYC* and *AmDICH*. This result suggests that these sites evolved *de novo* and not through

translocation of existing sites from elsewhere in the genome. Similarly, our search for consensus TCP-binding sites from *M. lewisii CYC2*-lineage genes in the *M. lewisii* genome resulted in only self-hits.

We identified two consensus DIV-binding sites in the *AmDIV* upstream non-coding sequence (**Table 7**), one of which was previously implicated by Raimundo *et al.* (2013) in *AmDIV* DTA. *AmCYC*, *AmRAD*, *AmDRIF1* and *AmDRIF2,* but not *AmDICH,* also had predicted DIV-binding sites in their upstream non-coding sequences (**Table 7**). It is unlikely that the predicted DIV-binding sites in the upstream non-coding sequences of *AmCYC* or *AmRAD* function for *Am*DIV binding. This is because *Am*DIV function is impaired in the presence of *Am*RAD proteins through competitive inhibition.

Table 7. Predicted consensus DIV-binding sites in the upstream non-coding sequences of *A. majus* flower symmetry genes. Bases in bold indicate conservation in the consensus binding site. *AmDICH* lacks consensus DIV-binding sites in upstream non-coding sequences. One of two consensus DIV-binding sites in *AmDIV* was reported by Raimundo et al. (2013)

Gene	Sequence	DNA strand	bp upstream of transcription start
AmCYC	A**GATA**AGG	anti-sense	329-336
AmRAD	A**GATA**ACA	anti-sense	798-805
	G**GATA**ACG	anti-sense	1051-1058
	C**GATA**AGA	anti-sense	2843-2850
AmDIV Raimundo et al., 2013	A**GATA**AGG	sense	2595-2602
	C**GATA**CCC	sense	1557-1564
AmDRIF1	G**GATA**CGG	sense	711-718
	A**GATA**AGG	sense	242-249
	A**GATA**AGC	anti-sense	505-512
AmDRIF2	A**GATA**ACC	anti-sense	1892-1899

Expression analyses suggest additional autoregulation of DIV in A. majus

Given the presence of predicted DIV-binding sites in *AmDRIF1* and *AmDRIF2* upstream non-coding sequences (**Table 7**), we tested whether *AmDRIF1* and/or *AmDRIF2* expression is

significantly altered in the *A. majus div* mutant background compared to wild type. We found that *AmDRIF1*, despite having multiple *DIV* consensus binding sites in its upstream region, was not under either direct or indirect regulation by *Am*DIV (p=0.453; **Figure 20**). *AmDRIF1* may be regulated by a non-DIV MYB transcription factor(s) that binds to the consensus DIV-binding motif. On the other hand, we found significantly lower levels of *AmDRIF2* expression in *div* mutant flower buds compared to wild type (p=0.031; **Figure 20**). This suggests that *AmDRIF2* is either directly or indirectly positively regulated by *Am*DIV. In turn, *AmDIV* is positively regulated by *Am*DRIF2-*Am*DIV heterodimers (Raimundo et al., 2013). Therefore, *AmDIV* appears to experience both direct and indirect transcriptional autoregulation through interaction of *AmDIV cis*-regulatory sequences with *Am*DRIF2-*Am*DIV heterodimers.

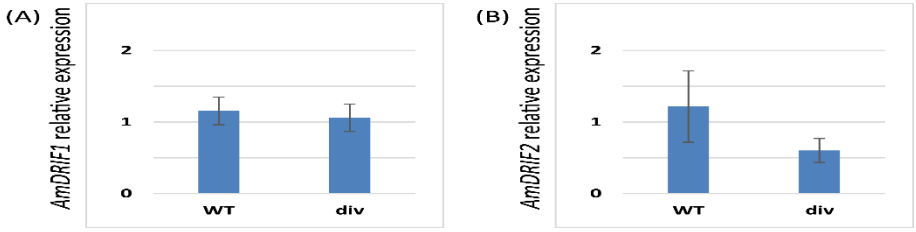

Figure 20. Relative expression of *AmDRIF1* (A) and *AmDRIF2* (B) in wild type and *divaricata* mutant lines. The expression level of *AmDRIF2* is significantly lower in the div background suggesting that AmDIV positively regulates *AmDRIF2* transcription. The values are mean ± standard deviation

Putative TCP-binding sites are enriched in upstream non-coding sequences of Lamiales *CYC2*-lineage genes

While no *CYC2*-lineage gene outside *P. heterotricha* has been experimentally tested for DTA, it is possible to infer the potential for DTA by screening for the consensus TCP-binding site, 5′-GGNCCC-3′ (Gao et al., 2015; Kosugi and Ohashi, 2002; Yang et al., 2012), in putative *cis*-regulatory regions of Lamiales *CYC2*-lineage genes. Given that flower monosymmetry is

homologous in *P. heterotricha* and *A. majus*, evolving early in the diversification of Lamiales (Reyes et al., 2016; Zhong and Kellogg, 2015b), a straight-forward hypothesis is that *CYC2*-lineage DTA evolved early in Lamiales and has been retained in Lamiales lineages with monosymmetric flowers. Under this hypothesis, Lamiales with flower monosymmetry will retain consensus TCP-binding site(s) in putative *CYC2*-lineage *cis*-regulatory sequences. The availability of multiple Lamiales genomes (Supplementary Table 1) allowed us to begin testing the hypothesis that autoregulation is potentially conserved across Lamiales *CYC* orthologs.

We identified orthologs of *AmCYC/AmDICH* (*CYC2*-lineage genes) from genome-sequenced Lamiales plus representative core eudicots (Supplementary Table 1, Supplementary Fig. 1). We identified orthologs of *AmRAD* and *AmDIV* from genome-sequenced Lamiales plus representative orthologs from sister lineages to Lamiales, Gentianales and Solanales (Supplementary Table 2, Supplementary Fig. 2). As with *P. heterotricha* and *A. majus*, recent duplication events lead to paralog complexity for *CYC2*-lineage genes (Supplementary Figure 1). We found that at least one *CYC2*-lineage gene from each core eudicot species had consensus TCP-binding sites(s) in the upstream non-coding sequence (Supplementary Tables 3 and 4), with two exceptions. The only *CYC2*-lineage genes in *Vitis vinifera* (Vitales), *CYCLOIDEA-like 2a*, and *Gossypium raimondii* (Malvales), *TCP1*, had no consensus TCP-binding sites in their upstream non-coding sequences.

We found consensus TCP-binding sites in the upstream non-coding sequences of *CYC2*-lineage genes in a wide variety of core eudicots with flowers with mono-, poly-, and dissymmetry (Supplementary Tables 3 and 4). However, *prima facia*, the *CYC2*-lineage orthologs from Lamiales appeared to be enriched for consensus TCP-binding sites. We tested for enrichment of consensus TCP-binding sites in the non-coding sequences upstream of Lamiales

CYC2-lineage genes. Additionally, we tested the upstream non-coding sequences of non-Lamiales core-eudicot *CYC2*-lineage genes, and Lamiales *RAD* and *DIV* orthologs for enrichment in consensus TCP-binding sites. We predict that *RAD* orthologs may show enrichment of the consensus TCP-binding site due to conserved regulation of *RAD* by CYC-like transcription factors across Lamiales, but that Lamiales *DIV* orthologs are not likely to be enriched for the consensus TCP-binding site given that there is no previous data indicating regulation of *DIV* orthologs by CYC-like transcription factors or other TCP proteins.

Table 8. Results from Analysis of Motif Enrichment (AME) tests for consensus TCP-binding sites in the upstream non-coding sequences of symmetry gene orthologs.

Test sequences (putative *cis*-regulatory regions)	**Control sequences**	***p*-value**	**Genes surveyed**	**Species surveyed**
Lamiales *DIV* orthologs	Shuffled test sequences	0.517	15	9
Lamiales *RAD* orthologs	Shuffled test sequences	**0.0406**	33	9
Lamiales *CYC2* orthologs	Shuffled test sequences	**0.0169**	20	9
Non-Lamiales core eudicot *CYC2* orthologs	Shuffled test sequences	0.352	39	17

As expected, we found that the upstream non-coding sequences of Lamiales *DIV* orthologs were not significantly enriched for the consensus TCP-binding sites (p=0.517; **Table 8**), and that the upstream non-coding sequences of Lamiales *RAD* orthologs were significantly enriched for the consensus TCP-binding site (p=0.0406; **Table 8**). This result is consistent with CYC-like transcription factors acting as regulators of *RAD,* but not *DIV* across Lamiales. Strikingly, we found that the upstream non-coding sequences of *CYC2*-lineage genes in Lamiales were significantly enriched in consensus TCP-binding sites (p=0.0169; **Table 8**) in-line with the hypothesis that *CYC* autoregulation evolved early in Lamiales, coincident with the evolution of monosymmetric flower, and has been maintained during Lamiales diversification. Notably, this pattern of enrichment appears specific to Lamiales. We tested for similar enrichment of the

consensus TCP-binding site in non-Lamiales core eudicot *CYC2*-lineage orthologs and found no evidence for a similar pattern of binding site enrichment (p=0.352; **Table 8**).

Discussion

Binding site enrichment supports the hypothesis that DTA of *CYC* is associated with the origin of flower monosymmetry in Lamiales

Positive regulation of *RAD* by CYC2-lineage genes for specifying flower monosymmetry is conserved across much of Lamiales (Corley et al., 2005; Su et al., 2017; Zhou et al., 2008). That we find significant enrichment of consensus TCP-binding sites in Lamiales *RAD* upstream non-coding sequences is in-line with conservation of this *CYC*-*RAD* regulatory module. Strikingly, our data demonstrate that Lamiales *CYC2*-lineage genes are also significantly enriched for consensus TCP-binding sites in upstream non-coding sequences. This supports the hypothesis that the origin of Lamiales flower monosymmetry coincides with the evolution of *CYC2*-lineage DTA. Further empirical studies in emerging Lamiales models (for example, Liu et al., 2014; Su et al., 2017, 2) will allow this hypothesis to be tested, as well as the alternative, that *CYC2*-lineage genes undergo transcriptional regulation by other TCP family proteins. As additional eudicot genomes become available, tests for TCP-binding site enrichment can be carried out in other lineages with bilaterally symmetrical flowers for which a role of *CYC2*-lineage genes has been implicated, for example, Fabaceae (Wang et al., 2008; Xu et al., 2013) and Malpighiaceae (Zhang et al., 2010).

Evaluating the pan-eudicot model for monosymmetry involving DTA of *CYC2*-lineage genes

A model hypothesizing the role of DTA for the parallel origin of monosymmetric flowers across eudicots was put forward by Yang et al. (2012; Fig 6) based on two primary lines of evidence. First, the observed differences in duration of flower specific expression of *CYC2*-lineage genes between species with monosymmetric vs. non-monosymmetric flowers. Second, the reported absence of consensus TCP-binding sites in the upstream non-coding sequences of *CYC2*-lineage genes from non-monosymmetric flowers. Specifically, *Arabidopsis thaliana*, *Brassica rapa*, *Vitis vinifera*, and *Solanum lycopersicum* do not have monosymmetric flowers and were reported to lack consensus TCP-binding sites in their *CYC2*-lineage genes compared to *Glycine max*, *Medicago trunculata*, *Mimulus guttatus*, *Primulina heterotricha*, *Oryza sativa*, and *Zea mays* (representing three independent origins of monosymmetry) that have consensus TCP-binding sites (Yang et al., 2012).

This model relies heavily on observations from *Arabidopsis* flowers where the expression of the sole *CYC2*-lineage gene (*AtTCP1*) is transiently dorsal-specific and the flowers are non-monosymmetric (Cubas et al., 2001). It is clear that *At*TCP1 does not play a critical role in floral organ differentiation in *Arabidopsis*, given no floral-specific DTA or other means by which expression can persist to later stages of flower differentiation. However, the pattern in *Arabidopsis* may not be universal for non-monosymmetric flowers. Closely related monosymmetric and non-monosymmetric Brassicaceae flowers do not exhibit a consistent pattern of early dorsal-specific expression (Busch et al., 2012). Evidence from Brassicaceae suggests that *Arabidopsis*-like dorsal-restricted expression early in flower development is not a pre-requisite for the evolution of flower monosymmetry via DTA. Beyond Brassicaceae, there

are examples of ancestrally non-monosymmetric flowers in core-eudicots where expression of *CYC2*-lineage genes is not localized spatially and/or restricted to an early developmental stage. These examples include *Bergia texana* (Elatinaceae) (Zhang et al., 2010), *Viburnum plicatum* (Adoxaceae) (Howarth et al., 2011), and *Solanum lycoperscicum* (Solanaceae, ancestral state ambiguous) (Parapunova et al., 2014), as well as an early-diverging eudicot, *Eschscholzia californica* (Papaveraceae) (Kölsch and Gleissberg, 2006).

Yang et al. (2012) reported a correlation between flower monosymmetry vs. non-monosymmetry and the presence vs. absence of consensus TCP-binding sites in corresponding upstream non-coding sequences of *CYC2*-lineage genes. This contributed to the model for the origin of flower monosymmetry facilitated by the evolution of *CYC2*-lineage DTA. In our expanded sampling we find that consensus TCP-binding sites are present in the upstream non-coding sequences of many *CYC2*-lineage genes across eudicots irrespective of flower symmetry. Yet, in an interesting pattern, all species with independently derived monosymmetric flowers that we investigated (Fabales, Lamiales, Brassicales, Asterales) have at least one *CYC2*-lineage ortholog with a consensus TCP-binding sequence in the upstream non-coding sequences (Supplementary Table 3 and 4). On the other hand, many species with non-monosymmetric flowers also have at least one *CYC2*-lineage ortholog with a consensus TCP-binding sequence in their upstream non-coding sequences (Supplementary Table 3 and 4). Notably, we find that the sole *CYC2*-lineage gene in *Arabidopsis* (*AtTCP1*), and a second *CYC2*-lineage gene in tomato that was not included in Yang et al. (2012), *Solanum lycopersicum TCP26* (Solyc03g045030.1), have consensus TCP-binding sites in their upstream non-coding sequences (Supplementary Table 4).

*At*TCP1 binds to all combinations of the consensus sequence 5′-GGNCCC-3′ *in vitro*, and flanking regions have limited significance in this interaction (Gao et al., 2015). *In vivo*, *At*TCP1 can directly bind to the two TCP-binding sites located in the regulatory region of the a downstream gene *DWARF4* (Gao et al., 2015). This suggests that the *Arabidopsis* TCP1 transcription factor can likely bind to the predicted TCP-binding site in its own upstream non-coding sequence, and hence possibly undergoes DTA. *AtTCP1* is expressed and is functional across the shoot organs throughout development, from seedlings to inflorescences (Koyama et al., 2010). This persistent expression is consistent with it having a predicted autoregulatory site. Expression surveys employing *in situ* mRNA hybridization (Cubas et al., 2001) and *AtTCP1* promoter fused to a β-glucuronidase (GUS) construct (Koyama et al., 2010) did not detect *AtTCP1* expression in later stages of flower development. It is interesting that the expression of a gene that is widely expressed in and controls development of many different organs is specifically downregulated in flowers. It is possible that *AtTCP1* is negatively regulated during late stages of *Arabidopsis* flower development, or continues to be expressed in flowers but a level that can only be detected by more sensitive methods, like quantitative rt-PCR.

Predicted *CYC2*-lineage autoregulatory sites are strongly associated with monosymmetry supporting the potential importance for DTA in establishing high and continuous asymmetric expression through later stages of flower organ differentiation (**Figure 21**). However, this pattern is not exclusive: *CYC2*-lineage orthologs from many species lacking monosymmetry also have predicted TCP-binding sites. This may be autoregulation for alternative developmental pathways, or regulation of *CYC2*-lineage genes by upstream TCP activators. At this point, experimental tests of TCP gene autoregulation are too sparse to draw solid conclusions regarding the role of DTA in independent origins of flower monosymmetry across core eudicots.

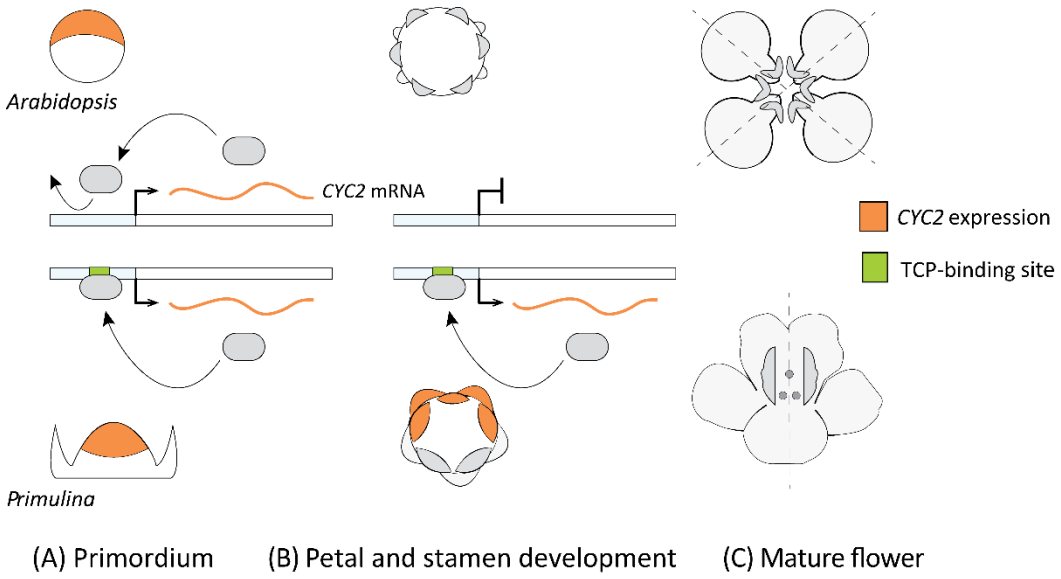

Figure 21. A previously proposed model explaining flower symmetry in *P. heterotricha* and *Arabidopsis*. (A) *CYC2*-lineage genes are expressed in early stage flower primordia of both *Arabidopsis* (dorsal part) and *P. heterotricha* (apical part). (B) Expression in *P. heterotricha* continues by DTA to later stages crucial for defining flower monosymmetry, this is not the case in *Arabidopsis.* (C) At anthesis, *P. heterotricha* is monosymmetric, *Arabidopsis* is not.

Origin and evolution of autoregulatory sites in DTA

Any *cis*-regulatory site can evolve by two primary processes, *de novo* by mutation and/or recombination in ancestral non-regulatory sequences, or by duplication of existing regulatory sites from a different location in the genome. Both have been reported in the origin of *cis*-regulatory sites involved in DTA. For example, the *CArG-box* sites involved in *Arabidopsis* AP1 autoregulation discussed earlier evolved by substitutions in the ancestral sequence that likely had a weak affinity for AP1 (Ye et al., 2016). Once evolved, these sites can undergo duplications, as reported in the apple MYB10 gene that controls fruit flesh color (Espley et al., 2009; Nocker et al., 2012).

How did the predicted autoregulatory sites in *CYC2*-lineage genes originate? We did not detect consensus TCP-binding sites with accompanying flanking sequences elsewhere in the *A.*

majus or *M. lewisii* genomes. This suggests that these predicted autoregulatory sites evolved *in situ* and are not a result of duplication from a different part of the genome, *i.e.,* similar to the origin of the autoregulatory sites in *Arabidopsis AP1* (Ye et al., 2016). However, multiple consensus TCP-binding sites are present within single *A. majus* and *M. lewisii CYC2*-lineage genes. To further test whether these multiple TCP-binding sites within a single putative regulatory region evolved by local, intra-genic duplication, as in the case of *MYB10* promoter in apples (Espley et al., 2009; Nocker et al., 2012), we aligned all *A. majus* and *M. lewisii* consensus TCP-binding sites, along with 100 bp flanking on either side, from within single upstream non-coding regions. We found no evidence that any of the predicted TCP-binding sites are derived from tandem duplication within *CYC* regulatory regions, again suggesting that multiple binding sites evolved *de novo.*

Chicken or egg: novel function or DTA first?

We have discussed potential roles of DTA in development, but how does DTA itself evolve? Autoregulation is common among genes positioned upstream in genetic regulatory networks with crucial developmental functions (discussed in Crews and Pearson, 2009; Hoot et al., 2010; specifically tested in yeasts and hepatocytes by Odom et al., 2006; Pasqualucci et al., 2003; Tao et al., 2012; Toth Hervay et al., 2011). This observed pattern leads to an interesting chicken or egg conundrum. Which evolves first in genes recruited to new developmental functions: the novel function, or the autoregulation? Two scenarios can explain the observed pattern that crucial genes are often autoregulated. 1) DTA evolves first, and such genes are recruited for new functions that require extended stable expression. Or, 2) New function evolves

first, and such genes, under selective pressure to provide extended stable expression, evolve DTA.

Evidence supporting scenario 2 is found in the *Arabidopsis AtAP1* example. This A-class floral homeotic gene in Brassicaceae underwent a duplication that generated the paralogs *AP1* and *CAL* gene lineages (Wang et al., 2012). AtAP1 defines sepals in *Arabidopsis thaliana*, but this function has not been reported elsewhere, and is likely an innovation in the genus *Arabidopsis* (Huijser et al., 1992; Litt, 2007; Lowman and Purugganan, 1999; Ruokolainen et al., 2010; Shepard and Purugganan, 2002). Except for the *AP1* paralog in *Arabidopsis* species, no Brassicaceae *AP1/CAL* gene tested to date undergoes DTA (Ye et al., 2016). And, as described above, DTA is an integral component of AtAP1 A-class function in flower development. Further, while the *AP1* orthologs of two *Arabidopsis* species have *CArG-box* in their *cis*-regulatory region that allows them to undergo DTA, other Brassicaceae species have *CArG-box*-like sequences with mismatches in the homologous gene region. In one such homolog, *Capsella rubella AP1*, the binding affinity of the mismatched *CArG-box*-like sequence was tested and can only weakly bind to AP1 protein. Hence, *Capsella rubella AP1* is likely not autoregulated (Ye et al., 2016). This suggests that the autoregulation of *Arabidopsis AP1* evolved either after or during, but not before, its recruitment to A-class function.

A major unanswered question that will clarify the origin of DTA in *Arabidopsis AP1* is whether its orthologs have similar functions in other Brassicaceae species. It is challenging to identify the ancestral state of autoregulation for any gene primarily for two reasons: there has been little functional work outside the model species, and predictive surveys are limited because genomes sequencing has been biased towards lineages with those model species. As plant sciences expands away from models systems (Poaceae, Brassicaceae, and Solanaceae), a wider

phylogenetic sampling will facilitate reconstruction of ancestral molecular interactions.

Conclusions

The origins and evolution of autoregulation will likely remain elusive until extensive experimental evidence emerges from multiple plant (and animal) lineages that inform ancestral and derived roles for autoregulation in development. It is, however, not surprising that a large number of transcription factors involved in defining crucial or novel phenotypes undergo direct transcriptional autoregulation, as this form of regulation is expected to both enhance and stabilize gene expression patterns critical for developmental patterning. We find evidence for enrichment of self-binding sites in Lamiales *CYC2*-lineages genes. This enrichment may reflect evolution of a novel pattern of direct transcriptional autoregulation early in Lamiales diversification, coincident with the origin of a key morphological innovation, floral monosymmetry. It is likely that the putative autoregulatory binding sites associated with Lamiales *CYC2*-lineages genes evolved via *de novo* mutations. Whether direct transcriptional autoregulation is conserved across Lamiales awaits further experimental evidence, as does the hypothesis that independent origins of flower monosymmetry may be associated with the evolution of positive transcriptional autoregulation.

Methods

Homolog predictions and phylogenetic analyses

AmCYC, AmDICH, *AmRAD*, and *AmDIV* orthologs were identified from published sources and online databases by tBLASTx (Altschul *et al.,* 1990). The gene names/identifiers and sources are listed in Supplementary Tables 1 and 2. Gene identifiers are also included with terminal genes on the phylogenies (Supplementary Figs. 1 and 2). A subset of included genes were available as full-length coding sequences from public databases. A subset of included genes were available as partial coding sequences from public databases. For partial coding sequences from species with available genome data, we predicted the full-length coding sequences either manually by aligning to previously reported homologs, or by prediction with AUGUSTUS ((either the web portal or the option in Geneious; Kearse et al., 2012; Stanke et al., 2004). A subset of included genes were identified by BLAST (Altschul *et al.,* 1990) from annotated genomes. We predicted the coding sequences either manually or with AUGUSTUS when our BLAST searches hit a region in a genome where no or partial genes were predicted. For *Mimulus lewisii DIV* and *RAD* homologs, we first BLAST searched the available transcriptome and subsequently mapped the hits to the genome. Two sets of sequences used here were not publicly available, the genes from *Ipomoea lacunosa* whose genome sequence was generously shared by Dr. Mark Rausher (Duke University), and *Mimulus guttatus RADlike1*, which was shared by Dr. Jinshun Zhong (University of Vermont; (the sequence has been used in the following work: Zhong et al., 2017).

We translationally aligned the coding sequences (omitting the stop codon) of *CYC*-like genes using MAFFT v7.388 (Katoh et al., 2002) in Geneious 10.2.3 (Kearse et al., 2012) with the following parameters: algorithm–auto, scoring matrix–BLOSUM62, gap opening penalty–1.1, offset value–0.124. The entire alignment was used for downstream phylogenetic analyses.

The *CYC*-like gene tree was estimated using a Bayesian approach (Metropolis-coupled Markov chain Monte Carlo) in MrBayes 3.2.6 (Ronquist et al., 2012) with uninformative priors for 10 million generations on the online CIPRES portal at https://www.phylo.org (Miller et al., 2010). The core-eudicot *CYC*-like tree was rooted with Rananculales *CYC*-like genes in FigTree (http://tree.bio.ed.ac.uk/software/figtree/).

DIV- and *RAD*-like genes were translationally aligned using an approach similar to *CYC*-like genes except for the following: gap opening penalty–1.53, and offset value–0.123. We removed the columns with 70% or more gaps from the alignment, and from the subsequent file used only the conserved first *MYBI* domain and nucleotides immediately 3′ to this domain. *DIV*- and *RAD*-like gene trees were estimated using the same approach as for *CYC*-like genes. Resulting *DIV*- and *RAD*-like trees were mid-point rooted in FigTree (http://tree.bio.ed.ac.uk/software/figtree/). For all sequences included in our phylogenetic analyses, nexus format nucleotide alignment along with the Bayesian parameter block, and the unaligned coding sequences in fasta format available from the Dryad Digital Repository: https://doi.org/10.5061/dryad.tv54037.

Consensus TCP and DIV-binding site predictions

We downloaded up to 3 kb non-coding sequence upstream of the transcription start sites of target Lamiales *CYC, RAD* and *DIV* homologs from corresponding genomes. We downloaded up to 3 kb non-coding sequence upstream of the transcription start sites of representative core eudicot *CYC* homologs from corresponding genomes. All genomic sources are listed in Supplementary Table 1. Within these sequences, we searched for the consensus TCP-binding site 5′-GGNCCC-3′ (Costa et al., 2005; Gao et al., 2015; Kosugi and Ohashi, 2002; Yang et al.,

2012) on both strands using Geneious 10.2.3 (Kearse et al., 2012). In *A. majus* only, we searched for the consensus DIV-binding site, 5′-[AGC]GATA[AC][GC][GAC]-3′ (Raimundo et al., 2013) in 3 kb upstream non-coding sequences of the six genes known to be involved in *A. majus* flower symmetry (**Figure 19**) using Geneious 10.2.3 (Kearse et al., 2012). To determine whether the consensus TCP-binding sites found in the *A. majus* and *M. lewisii* upstream *CYC* homolog sequences were derived from other genomic locations, we used the predicted TCP-binding sites, plus 100 bp on either side, as BLAST queries against the available genomes in Geneious 10.2.3 (Kearse et al., 2012).

Analysis of Motif Enrichment

We tested for consensus TCP-binding site enrichment using *Analysis of Motif Enrichment* (AME, http://meme-suite.org/tools/ame, McLeay and Bailey, 2010). AME can identify known or user-provided motifs that are relatively enriched in a given set of sequences compared with shuffled versions of those sequences or with user-provided control sequences. *AME* does not discriminate among motifs based on their locations within the sequences. The following options were selected: sequence scoring method—average odds score, motif enrichment test—rank sum test, and background model—uniform model. We defined the consensus TCP-binding site as 5′-GGNCCC-3′ (Costa et al., 2005; Gao et al., 2015; Kosugi and Ohashi, 2002; Yang et al., 2012), and query sequences as 3 kb upstream of transcription start sites of focal genes, and used shuffled sequences as the control. The upstream non-coding sequences are available in fasta format from the Dryad Digital Repository: https://doi.org/10.5061/dryad.tv54037.

Quantitative reverse-transcriptase PCR (rt-PCR)

Antirrhinum majus wild type (genotype JI 7) and *divaricata* mutants (genotype JI 13) were acquired from John Innes Centre, UK, under USDA permit number P37-16-01034. Five flower buds of the same developmental stage (stage 11, flower bud ca. 4.0 mm in length, corolla equal in length to calyx, petal tips white in wild type; Vincent and Coen, 2004) were sampled from each genotype. RNA was extracted using RNeasy plant minikit (Qiagen, Germantown, MD, USA), followed by DNase treatment (TURBO™ DNase, ThermoFisher Scientific, Waltham, MA, USA), and cDNA synthesis (iScript cDNA Synthesis Kit, Bio-Rad, Hercules, CA, USA). Quantitative rt-PCR was performed on a StepOnePlus™ Real-Time PCR System (ThermoFisher Scientific) using SYBR™ Select Master Mix (ThermoFisher Scientific). Quantitative rt-PCR was carried out for three technical replicates for each of five biological replicates per genotype. Expression was normalized against *UBIQUITIN5*. This gene has been reported to have little transcriptional variation across tissue types and developmental stages (Preston and Hileman, 2010). Expression was analysed by the ΔΔCt method. Significant differences in relative expression between genotypes were determined using two sample t-test assuming equal variances in Minitab. The quantitative rt-PCR primers were as follows: AmDRIF1_RT_F4: GCCTTGGATCAAATTTCGGC; AmDRIF1_RT_R4: AGGAAGAATGGAGCTGGCAA; AmDRIF2_RT_F1a: AATGGTCATGGAGAGTGGGG; AmDRIF2_RT_R1:TATAGCTTGCTCCTCTGGGG; AmUBQ5_qPCR_F1: GCGCAAGAAGAAGACCTACAC; AmUBQ5_qPCR_R1: CTTCCTGAGCCTCTGCACTT. Efficiency of PCR was determined using DART (Peirson et al., 2003).

Supplementary files

Supplementary tables

Supplementary Table 1. Names and sources of CYC homologs used in this study.

Supplementary Table 2. Names and sources of DIV and RAD homologs used in this study.

Supplementary Table 3. Predicted TCP-binding sites in the upstream non-coding sequences of non-Lamiales core eudicot CYC2-lineage genes.

Supplementary Table 4. Predicted TCP-binding sites in the upstream non-coding sequences of Lamiales CYC2-lineage, RAD- and DIV-orthologs.

Supplementary figures

Supplementary Fig. 1. Bayesian majority-rule consensus tree of CYC-like genes in eudicots. Bayesian posterior probabilities at nodes.

Supplementary Fig. 2. Bayesian majority-rule consensus tree of DIV- and RAD-like genes in Lamiales, Solanales, and Gentianales. Bayesian posterior probabilities at nodes.

References

Ahmed, E. M., Desoukey, S. Y., Fouad, M. A., and Kamel, M. S. (2016). A Pharmacognostical Study of Russelia equisetiformis Sch. & Cham. 8, 19.

Almeida, J., Rocheta, M., and Galego, L. (1997). Genetic control of flower shape in *Antirrhinum majus*. *Development* 124, 1387–1392. Available at: http://dev.biologists.org/content/124/7/1387 [Accessed March 18, 2014].

Ambawat, S., Sharma, P., Yadav, N. R., and Yadav, R. C. (2013). MYB transcription factor genes as regulators for plant responses: an overview. *Physiol. Mol. Biol. Plants* 19, 307–321. doi:10.1007/s12298-013-0179-1.

Barringer, K. A. (1981). *A taxonomic revision of* Angelonia *(Scrophulariaceae)*. University of Connecticut Available at: https://opencommons.uconn.edu/dissertations/AAI8205161.

Bartlett, M. E., and Specht, C. D. (2011). Changes in expression pattern of the *TEOSINTE BRANCHED1*-like genes in the Zingiberales provide a mechanism for evolutionary shifts in symmetry across the order. *Am. J. Bot.* 98, 227–243. doi:10.3732/ajb.1000246.

Bateman, E. (1998). Autoregulation of Eukaryotic Transcription Factors. *Prog. Nucleic Acid Res. Mol. Biol.* 60, 133–168. doi:10.1016/S0079-6603(08)60892-2.

Bharathan, G., Goliber, T. E., Moore, C., Kessler, S., Pham, T., and Sinha, N. R. (2002). Homologies in leaf form inferred from *KNOXI* gene expression during development. *Science* 296, 1858–1860. doi:10.1126/science.1070343.

Bojnanský, V., and Fargašová, A. (2007). *Atlas of Seeds and Fruits of Central and East-European Flora: The Carpathian Mountains Region*. Springer Science & Business Media.

Brakefield, P. M., Gates, J., Keys, D., Kesbeke, F., Wijngaarden, P. J., Montelro, A., et al. (1996). Development, plasticity and evolution of butterfly eyespot patterns. *Nature* 384, 236–242. doi:10.1038/384236a0.

Brukhin, V., Hernould, M., Gonzalez, N., Chevalier, C., and Mouras, A. (2003). Flower development schedule in tomato Lycopersicon esculentum cv. sweet cherry. *Sex. Plant Reprod.* 15, 311–320. doi:10.1007/s00497-003-0167-7.

Busch, A., Horn, S., Mühlhausen, A., Mummenhoff, K., and Zachgo, S. (2012). Corolla Monosymmetry: Evolution of a Morphological Novelty in the Brassicaceae Family. *Mol. Biol. Evol.* 29, 1241–1254. doi:10.1093/molbev/msr297.

Busch, A., and Zachgo, S. (2007). Control of corolla monosymmetry in the Brassicaceae *Iberis amara*. *Proc. Natl. Acad. Sci. U. S. A.* 104, 16714–16719. doi:10.1073/pnas.0705338104.

CalPhotos, Antirrhinum filipes *CalPhotos Antirrhinum Filipes Twining Snapdragon Berkeley Nat. Hist. Mus. Univ. Calif. Berkeley*. Available at: https://calphotos.berkeley.edu/cgi/img_query?enlarge=0000+0000+0209+0168 [Accessed March 6, 2019].

Chapman, M. A., Tang, S., Draeger, D., Nambeesan, S., Shaffer, H., Barb, J. G., et al. (2012). Genetic analysis of floral symmetry in Van Gogh's sunflowers reveals independent recruitment of *CYCLOIDEA* genes in the Asteraceae. *Public Libr. Sci. Genet.* 8, e1002628. doi:10.1371/journal.pgen.1002628.

Citerne, H., Jabbour, F., Nadot, S., and Damerval, C. (2010). "The evolution of floral symmetry," in Advances in Botanical Research (Academic Press), 85–137. Available at: http://www.sciencedirect.com/science/article/pii/S0065229610540035 [Accessed March 18, 2014].

Citerne, H. L., Le Guilloux, M., Sannier, J., Nadot, S., and Damerval, C. (2013). Combining Phylogenetic and Syntenic Analyses for Understanding the Evolution of TCP ECE Genes in Eudicots. *Public Libr. Sci. ONE* 8, e74803. doi:10.1371/journal.pone.0074803.

Citerne, H. L., Luo, D., Pennington, R. T., Coen, E., and Cronk, Q. C. B. (2003). A phylogenomic investigation of *CYCLOIDEA-Like TCP* genes in the Leguminosae. *Plant Physiol.* 131, 1042–1053. doi:10.1104/pp.102.016311.

Citerne, H. L., Möller, M., and Cronk, Q. C. B. (2000). Diversity of cycloidea -like Genes in Gesneriaceae in Relation to Floral Symmetry. *Ann. Bot.* 86, 167–176. doi:10.1006/anbo.2000.1178.

Consortium of Intermountain Herbaria Detailed Collection Record Information, Sairocarpus watsonii Available at: http://intermountainbiota.org/portal/collections/individual/index.php?occid=13398926 [Accessed March 6, 2019].

Corley, S. B., Carpenter, R., Copsey, L., and Coen, E. (2005). Floral asymmetry involves an interplay between TCP and MYB transcription factors in *Antirrhinum*. *Proc. Natl. Acad. Sci. U. S. A.* 102, 5068–5073. doi:10.1073/pnas.0501340102.

Cornille, A., Giraud, T., Smulders, M. J. M., Roldán-Ruiz, I., and Gladieux, P. (2014). The domestication and evolutionary ecology of apples. *Trends Genet.* 30, 57–65. doi:10.1016/j.tig.2013.10.002.

Costa, M. M. R., Fox, S., Hanna, A. I., Baxter, C., and Coen, E. (2005). Evolution of regulatory interactions controlling floral asymmetry. *Development* 132, 5093–5101. doi:10.1242/dev.02085.

Craene, L. P. R. D. (2010). *Floral Diagrams: An Aid to Understanding Flower Morphology and Evolution*. Cambridge University Press.

Crews, S. T., and Pearson, J. C. (2009). Transcriptional autoregulation in development. *Curr. Biol. CB* 19, R241–R246. doi:10.1016/j.cub.2009.01.015.

Cripps, R. M., Lovato, T. L., and Olson, E. N. (2004). Positive autoregulation of the Myocyte enhancer factor-2 myogenic control gene during somatic muscle development in Drosophila. *Dev. Biol.* 267, 536–547. doi:10.1016/j.ydbio.2003.12.004.

Cubas, P., Coen, E., and Zapater, J. M. M. (2001). Ancient asymmetries in the evolution of flowers. *Curr. Biol.* 11, 1050–1052. doi:10.1016/S0960-9822(01)00295-0.

Damerval, C., Citerne, H., Guilloux, M. L., Domenichini, S., Dutheil, J., Craene, L. R. de, et al. (2013). Asymmetric morphogenetic cues along the transverse plane: Shift from disymmetry to zygomorphy in the flower of Fumarioideae. *Am. J. Bot.* 100, 391–402. doi:10.3732/ajb.1200376.

delta-intkey, Sibthorpia europaea *Descr. Lang. Taxon.* Available at: http://delta-intkey.com/angio/images/ebot9691.jpg [Accessed March 6, 2019].

Dinesh-Kumar, S. P., Anandalakshmi, R., Marathe, R., Schiff, M., and Liu, Y. (2003). "Virus-Induced Gene Silencing," in *Plant Functional Genomics* Methods in Molecular Biology™., ed. E. Grotewold (Humana Press), 287–293. Available at: http://link.springer.com/protocol/10.1385/1-59259-413-1%3A287 [Accessed November 22, 2014].

Endress, P. K. (1999). Symmetry in Flowers: Diversity and Evolution. *Int. J. Plant Sci.* 160, S3. Available at: http://search.ebscohost.com/login.aspx?direct=true&db=a9h&AN=5100860&site=ehost-live [Accessed November 8, 2013].

Espley, R. V., Bovy, A., Bava, C., Jaeger, S. R., Tomes, S., Norling, C., et al. (2013). Analysis of genetically modified red-fleshed apples reveals effects on growth and consumer attributes. *Plant Biotechnol. J.* 11, 408–419. doi:10.1111/pbi.12017.

Espley, R. V., Brendolise, C., Chagné, D., Kutty-Amma, S., Green, S., Volz, R., et al. (2009). Multiple Repeats of a Promoter Segment Causes Transcription Factor Autoregulation in Red Apples. *Plant Cell* 21, 168–183. doi:10.1105/tpc.108.059329.

Espley, R. V., Hellens, R. P., Putterill, J., Stevenson, D. E., Kutty-Amma, S., and Allan, A. C. (2007). Red colouration in apple fruit is due to the activity of the MYB transcription factor, MdMYB10. *Plant J.* 49, 414–427. doi:10.1111/j.1365-313X.2006.02964.x.

Fenster, C. B., Armbruster, W. S., and Dudash, M. R. (2009). Specialization of flowers: is floral orientation an overlooked first step? *New Phytol.* 183, 502–506. doi:10.1111/j.1469-8137.2009.02852.x.

Fenster, C. B., Armbruster, W. S., Wilson, P., Dudash, M. R., and Thomson, J. D. (2004). Pollination Syndromes and Floral Specialization. *Annu. Rev. Ecol. Evol. Syst.* 35, 375–403. doi:10.1146/annurev.ecolsys.34.011802.132347.

Flowers of India, Hemiphragma heterophyllum Available at: https://www.flowersofindia.net/catalog/slides/Nash%20Jhaar.html [Accessed March 6, 2019].

Galego, L., and Almeida, J. (2002a). Role of *DIVARICATA* in the control of dorsoventral asymmetry in *Antirrhinum* flowers. *Genes Dev.* 16, 880–891. doi:10.1101/gad.221002.

Galego, L., and Almeida, J. (2002b). Role of DIVARICATA in the control of dorsoventral asymmetry in *Antirrhinum* flowers. *Genes Dev.* 16, 880–891. doi:10.1101/gad.221002.

Gallo-Ebert, C., Donigan, M., Liu, H.-Y., Pascual, F., Manners, M., Pandya, D., et al. (2013). The Yeast Anaerobic Response Element AR1b Regulates Aerobic Antifungal Drug-dependent Sterol Gene Expression. *J. Biol. Chem.* 288, 35466–35477. doi:10.1074/jbc.M113.526087.

Gao, A., Zhang, J., and Zhang, W. (2017). Evolution of RAD- and DIV-Like Genes in Plants. *Int. J. Mol. Sci.* 18, 1961. doi:10.3390/ijms18091961.

Gao, Q., Tao, J.-H., Yan, D., Wang, Y.-Z., and Li, Z.-Y. (2008). Expression differentiation of CYC-like floral symmetry genes correlated with their protein sequence divergence in Chirita heterotricha (Gesneriaceae). *Dev. Genes Evol.* 218, 341–351. doi:10.1007/s00427-008-0227-y.

Gao, Y., Zhang, D., and Li, J. (2015). TCP1 Modulates DWF4 Expression via Directly Interacting with the GGNCCC Motifs in the Promoter Region of DWF4 in Arabidopsis thaliana. *J. Genet. Genomics* 42, 383–392. doi:10.1016/j.jgg.2015.04.009.

Ghebrehiwet, M., Bremer, B., and Thulin, M. (2000). Phylogeny of the tribe Antirrhineae (Scrophulariaceae) based on morphological and*ndh*F sequence data. *Plant Syst. Evol.* 220, 223–239. doi:10.1007/BF00985047.

Go Botany, Plantago coronopus Available at: https://gobotany.newenglandwild.org/species/plantago/coronopus/ [Accessed March 6, 2019].

Gonzalez, A., Zhao, M., Leavitt, J. M., and Lloyd, A. M. (2008). Regulation of the anthocyanin biosynthetic pathway by the TTG1/bHLH/Myb transcriptional complex in Arabidopsis seedlings. *Plant J. Cell Mol. Biol.* 53, 814–827. doi:10.1111/j.1365-313X.2007.03373.x.

Hake, S. (2008). Inflorescence Architecture: The Transition from Branches to Flowers. *Curr. Biol.* 18, R1106–R1108. doi:10.1016/j.cub.2008.10.024.

Hay, A., and Tsiantis, M. (2010). KNOX genes: versatile regulators of plant development and diversity. *Development* 137, 3153–3165. doi:10.1242/dev.030049.

Hileman, L. C. (2014). Bilateral flower symmetry — how, when and why? *Curr. Opin. Plant Biol.* 17, 146–152. doi:10.1016/j.pbi.2013.12.002.

Hileman, L. C., and Baum, D. A. (2003). Why do paralogs persist? Molecular evolution of *CYCLOIDEA* and related floral symmetry genes in Antirrhineae (Veronicaceae). *Mol. Biol. Evol.* 20, 591–600. doi:10.1093/molbev/msg063.

Hileman, L. C., Kramer, E. M., and Baum, D. A. (2003). Differential regulation of symmetry genes and the evolution of floral morphologies. *Proc. Natl. Acad. Sci. U. S. A.* 100, 12814–12819. doi:10.1073/pnas.1835725100.

Hiratsu, K., Matsui, K., Koyama, T., and Ohme-Takagi, M. (2003). Dominant repression of target genes by chimeric repressors that include the EAR motif, a repression domain, in Arabidopsis. *Plant J.* 34, 733–739. doi:10.1046/j.1365-313X.2003.01759.x.

Hochschild, A. (2002). The λ Switch: cI Closes the Gap in Autoregulation. *Curr. Biol.* 12, R87–R89. doi:10.1016/S0960-9822(02)00667-X.

Holloway, D. M., Lopes, F. J. P., Costa, L. da F., Travençolo, B. A. N., Golyandina, N., Usevich, K., et al. (2011). Gene Expression Noise in Spatial Patterning: hunchback Promoter Structure Affects Noise Amplitude and Distribution in Drosophila Segmentation. *PLOS Comput. Biol.* 7, e1001069. doi:10.1371/journal.pcbi.1001069.

Hoot, S. J., Brown, R. P., Oliver, B. G., and White, T. C. (2010). The UPC2 Promoter in Candida albicans Contains Two cis-Acting Elements That Bind Directly to Upc2p, Resulting in Transcriptional Autoregulation. *Eukaryot. Cell* 9, 1354–1362. doi:10.1128/EC.00130-10.

Howarth, D. G., and Donoghue, M. J. (2006). Phylogenetic analysis of the "ECE" (CYC/TB1) clade reveals duplications predating the core eudicots. *Proc. Natl. Acad. Sci.* 103, 9101–9106. doi:10.1073/pnas.0602827103.

Howarth, D. G., Martins, T., Chimney, E., and Donoghue, M. J. (2011). Diversification of *CYCLOIDEA* expression in the evolution of bilateral flower symmetry in Caprifoliaceae and *Lonicera* (Dipsacales). *Ann. Bot.* 107, 1521–1532. doi:10.1093/aob/mcr049.

Huijser, P., Klein, J., Lönnig, W. E., Meijer, H., Saedler, H., and Sommer, H. (1992). Bracteomania, an inflorescence anomaly, is caused by the loss of function of the MADS-box gene squamosa in Antirrhinum majus. *EMBO J.* 11, 1239–1249. Available at: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC556572/ [Accessed August 23, 2017].

International Plant Names Index *IPNI*. Available at: http://www.ipni.org/ipni/plantnamesearchpage.do [Accessed March 6, 2019].

Ivorra, A. (2014). Joyas botánicas de Almería, Lafuentea rotundifolia. *Joyas Botánicas Almería*. Available at: http://www.almerinatura.com/joyas/lafuentea-rotundifolia.html [Accessed March 6, 2019].

Ixitixel (2008). Kletter-Gloxinie (Lophospermum erubescens) - Fruit. Available at: https://commons.wikimedia.org/wiki/File:Lophospermum_erubescens_(Fruit).jpg [Accessed March 6, 2019].

Jiggins, C. D., Wallbank, R. W. R., and Hanly, J. J. (2017). Waiting in the wings: what can we learn about gene co-option from the diversification of butterfly wing patterns? *Philos. Trans. R. Soc. B Biol. Sci.* 372. doi:10.1098/rstb.2015.0485.

Juan, R. (2000). SEM and Light Microscope Observations on Fruit and Seeds in Scrophulariaceae from Southwest Spain and their Systematic Significance. *Ann. Bot.* 86, 323–338. doi:10.1006/anbo.2000.1188.

Kampny, C. M. (1995). Pollination and Flower Diversity in Scrophulariaceae. *Bot. Rev.* 61, 350–366. Available at: http://www.jstor.org.www2.lib.ku.edu/stable/4354259 [Accessed July 23, 2017].

Katoh, K., Misawa, K., Kuma, K., and Miyata, T. (2002). MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Res.* 30, 3059–3066. doi:10.1093/nar/gkf436.

Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., et al. (2012). Geneious Basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* 28, 1647–1649. doi:10.1093/bioinformatics/bts199.

Knapp, S. (2004). "Floral diversity and evolution in Solanaceae," in *Developmental Genetics and Plant Evolution*, eds. Q. C. B. Cronk, R. M. Bateman, and J. A. Hawkins (CRC Press).

Kölsch, A., and Gleissberg, S. (2006). Diversification of *CYCLOIDEA-like TCP* Genes in the Basal Eudicot Families Fumariaceae and Papaveraceae s.str. *Plant Biol.* 8, 680–687. doi:10.1055/s-2006-924286.

Kosugi, S., and Ohashi, Y. (2002). DNA binding and dimerization specificity and potential targets for the TCP protein family. *Plant J.* 30, 337–348. doi:10.1046/j.1365-313X.2002.01294.x.

Koyama, T., Sato, F., and Ohme-Takagi, M. (2010). A role of TCP1 in the longitudinal elongation of leaves in Arabidopsis. *Biosci. Biotechnol. Biochem.* 74, 2145–2147. doi:10.1271/bbb.100442.

Li, M., Zhang, D., Gao, Q., Luo, Y., Zhang, H., Ma, B., et al. (2019). Genome structure and evolution of <i>Antirrhinum majus>/i> L. *Nat. Plants* 5, 174. doi:10.1038/s41477-018-0349-9.

Lin-Wang, K., Bolitho, K., Grafton, K., Kortstee, A., Karunairetnam, S., McGhie, T. K., et al. (2010). An R2R3 MYB transcription factor associated with regulation of the anthocyanin biosynthetic pathway in Rosaceae. *BMC Plant Biol.* 10, 50. doi:10.1186/1471-2229-10-50.

Litt, A. (2007). An Evaluation of A-Function: Evidence from the APETALA1 and APETALA2 Gene Lineages. *Int. J. Plant Sci.* 168, 73–91. doi:10.1086/509662.

Liu, B.-L., Pang, H.-B., Yang, X., and Wang, Y.-Z. (2014). Functional and evolutionary analyses of Primulina heterotrichaCYC1C gene in tobacco and Arabidopsis transformation systems. *J. Syst. Evol.* 52, 112–123. doi:10.1111/jse.12067.

Liu, G., Liu, Y., Chen, Z., and Liu, Z. (2013). [Establishment of transformation system of Lonicera macranthoides mediated by Agrobacterium tumefaciens]. *Zhong Yao Cai Zhongyaocai J. Chin. Med. Mater.* 36, 1904–1907.

Liu, Y., Schiff, M., and Dinesh-Kumar, S. P. (2002). Virus-induced gene silencing in tomato. *Plant J.* 31, 777–786. doi:10.1046/j.1365-313X.2002.01394.x.

Lönnig, W.-E., Stüber, K., Saedler, H., and H. Kim, J. (2018). Biodiversity and Dollo's Law: to what extent can the phenotypic differences between *Misopates orontium* and *Antirrhinum majus* be bridged by mutagenesis? *Bioremediation Biodivers. Bioavailab. V1 1-30 2007*.

Lowman, A. C., and Purugganan, M. D. (1999). Duplication of the Brassica oleracea APETALA1 floral homeotic gene and the evolution of domesticated cauliflower. *J. Hered.* 90, 514–520.

Lucid Central, Digitalis purpurea *Lucid Cent. Weeds Aust. Biosecurity Qld. Ed. Qld. Goverment*. Available at:

https://keyserver.lucidcentral.org/weeds/data/media/Html/digitalis_purpurea.htm [Accessed March 6, 2019].

Luo, D., Carpenter, R., Copsey, L., Vincent, C., Clark, J., and Coen, E. (1999). Control of organ asymmetry in flowers of *Antirrhinum*. *Cell* 99, 367–376. doi:10.1016/S0092-8674(00)81523-8.

Luo, D., Carpenter, R., Vincent, C., Copsey, L., and Coen, E. (1996). Origin of floral asymmetry in *Antirrhinum*. *Nature* 383, 794–799. doi:10.1038/383794a0.

Ma, X., Zhu, Q., Chen, Y., and Liu, Y.-G. (2016). CRISPR/Cas9 Platforms for Genome Editing in Plants: Developments and Applications. *Mol. Plant* 9, 961–974. doi:10.1016/j.molp.2016.04.009.

Machemer, K., Shaiman, O., Salts, Y., Shabtai, S., Sobolev, I., Belausov, E., et al. (2011). Interplay of MYB factors in differential cell expansion, and consequences for tomato fruit development. *Plant J.* 68, 337–350. doi:10.1111/j.1365-313X.2011.04690.x.

Maddison, W. P., and Maddison, D. R. (2018). *Mesquite: a modular system for evolutionary analysis*. Available at: http://www.mesquiteproject.org.

Mano, H., Ogasawara, F., Sato, K., Higo, H., and Minobe, Y. (2007). Isolation of a Regulatory Gene of Anthocyanin Biosynthesis in Tuberous Roots of Purple-Fleshed Sweet Potato. *Plant Physiol.* 143, 1252–1268. doi:10.1104/pp.106.094425.

Martín-Trillo, M., and Cubas, P. (2010). TCP genes: a family snapshot ten years later. *Trends Plant Sci.* 15, 31–39. doi:10.1016/j.tplants.2009.11.003.

Martínez-Antonio, A., and Collado-Vides, J. (2003). Identifying global regulators in transcriptional regulatory networks in bacteria. *Curr. Opin. Microbiol.* 6, 482–489. doi:10.1016/j.mib.2003.09.002.

McLeay, R. C., and Bailey, T. L. (2010). Motif Enrichment Analysis: a unified framework and an evaluation on ChIP data. *BMC Bioinformatics* 11, 165. doi:10.1186/1471-2105-11-165.

Miller, M. A., Pfeiffer, W., and Schwartz, T. (2010). Creating the CIPRES Science Gateway for inference of large phylogenetic trees. in *2010 Gateway Computing Environments Workshop (GCE)*, 1–8. doi:10.1109/GCE.2010.5676129.

Murray, M. A. (1945). Carpellary and Placental Structure in the Solanaceae. *Bot. Gaz.* 107, 243–260. doi:10.1086/335344.

Neal, P. R., Dafni, A., and Giurfa, M. (1998). Floral symmetry and its role in plant-pollinator systems: terminology, distribution, and hypotheses. *Annu. Rev. Ecol. Syst.* 29, 345–373. doi:10.1146/annurev.ecolsys.29.1.345.

Niu, S.-S., Xu, C.-J., Zhang, W.-S., Zhang, B., Li, X., Lin-Wang, K., et al. (2010). Coordinated regulation of anthocyanin biosynthesis in Chinese bayberry (Myrica rubra) fruit by a R2R3 MYB transcription factor. *Planta* 231, 887–899. doi:10.1007/s00425-009-1095-z.

Nocker, S. van, Berry, G., Najdowski, J., Michelutti, R., Luffman, M., Forsline, P., et al. (2012). Genetic diversity of red-fleshed apples (Malus). *Euphytica* 185, 281–293. doi:10.1007/s10681-011-0579-7.

Odom, D. T., Dowell, R. D., Jacobsen, E. S., Nekludova, L., Rolfe, P. A., Danford, T. W., et al. (2006). Core transcriptional regulatory circuitry in human hepatocytes. *Mol. Syst. Biol.* 2, 2006.0017. doi:10.1038/msb4100059.

Ogutcen, E., and Vamosi, J. C. (2016). A phylogenetic study of the tribe Antirrhineae: Genome duplications and long-distance dispersals from the Old World to the New World. *Am. J. Bot.* 103, 1071–1081. doi:10.3732/ajb.1500464.

Ó'Maoiléidigh, D. S., Graciet, E., and Wellmer, F. (2014). Gene networks controlling Arabidopsis thaliana flower development. *New Phytol.* 201, 16–30. doi:10.1111/nph.12444.

O'Meara, B. C., Smith, S. D., Armbruster, W. S., Harder, L. D., Hardy, C. R., Hileman, L. C., et al. (2016). Non-equilibrium dynamics and floral trait interactions shape extant angiosperm diversity. *Proc. R. Soc. B Biol. Sci.* 283. doi:10.1098/rspb.2015.2304.

Padmanabhan, M., and Dinesh-Kumar, S. P. (2009). Virus-Induced Gene Silencing as a Tool for Delivery of dsRNA into Plants. *Cold Spring Harb. Protoc.* 2009, pdb.prot5139. doi:10.1101/pdb.prot5139.

Panganiban, G., Irvine, S. M., Lowe, C., Roehl, H., Corley, L. S., Sherbon, B., et al. (1997). The origin and evolution of animal appendages. *Proc. Natl. Acad. Sci.* 94, 5162–5166. doi:10.1073/pnas.94.10.5162.

Parapunova, V., Busscher, M., Busscher-Lange, J., Lammers, M., Karlova, R., Bovy, A. G., et al. (2014). Identification, cloning and characterization of the tomato TCP transcription factor family. *BMC Plant Biol.* 14, 157. doi:10.1186/1471-2229-14-157.

Pasqualucci, L., Migliazza, A., Basso, K., Houldsworth, J., Chaganti, R. S. K., and Dalla-Favera, R. (2003). Mutations of the BCL6 proto-oncogene disrupt its negative autoregulation in diffuse large B-cell lymphoma. *Blood* 101, 2914–2923. doi:10.1182/blood-2002-11-3387.

Peirson, S. N., Butler, J. N., and Foster, R. G. (2003). Experimental validation of novel and conventional approaches to quantitative real-time PCR data analysis. *Nucleic Acids Res.* 31, e73–e73. doi:10.1093/nar/gng073.

Perez-Rodriguez, M., Jaffe, F. W., Butelli, E., Glover, B. J., and Martin, C. (2005). Development of three different cell types is associated with the activity of a specific MYB transcription factor in the ventral petal of *Antirrhinum majus* flowers. *Development* 132, 359–370. doi:10.1242/dev.01584.

Pinho, R., Garcia, V., Irimia, M., and Feldman, M. W. (2014). Stability Depends on Positive Autoregulation in Boolean Gene Regulatory Networks. *PLOS Comput. Biol.* 10, e1003916. doi:10.1371/journal.pcbi.1003916.

Preston, J. C., Barnett, L. L., Kost, M. A., Oborny, N. J., and Hileman, L. C. (2014). Optimization of virus-induced gene silencing to facilitate evo-devo studies in the emerging model species *Mimulus guttatus* (Phrymaceae). *Ann. Mo. Bot. Gard.* 99, 301–312. doi:10.3417/2010120.

Preston, J. C., and Hileman, L. C. (2010). SQUAMOSA-PROMOTER BINDING PROTEIN 1 initiates flowering in *Antirrhinum majus* through the activation of meristem identity genes. *Plant J. Cell Mol. Biol.* 62, 704–712. doi:10.1111/j.1365-313X.2010.04184.x.

Preston, J. C., and Hileman, L. C. (2012). Parallel evolution of TCP and B-class genes in Commelinaceae flower bilateral symmetry. *EvoDevo* 3, 6. doi:10.1186/2041-9139-3-6.

Preston, J. C., Martinez, C. C., and Hileman, L. C. (2011). Gradual disintegration of the floral symmetry gene network is implicated in the evolution of a wind-pollination syndrome. *Proc. Natl. Acad. Sci. U. S. A.* 108, 2343–2348. doi:10.1073/pnas.1011361108.

Raimundo, J., Sobral, R., Bailey, P., Azevedo, H., Galego, L., Almeida, J., et al. (2013). A subcellular tug of war involving three MYB-like proteins underlies a molecular antagonism in *Antirrhinum* flower asymmetry. *Plant J.* 75, 527–538. doi:10.1111/tpj.12225.

Reeves, P. A., and Olmstead, R. G. (2003). Evolution of the TCP Gene Family in Asteridae: Cladistic and Network Approaches to Understanding Regulatory Gene Family

Diversification and Its Impact on Morphological Evolution. *Mol. Biol. Evol.* 20, 1997–2009. doi:10.1093/molbev/msg211.

Reyes, E., Sauquet, H., and Nadot, S. (2016). Perianth symmetry changed at least 199 times in angiosperm evolution. *Taxon* 65, 945–964. doi:10.12705/655.1.

Ronquist, F., Teslenko, M., van der Mark, P., Ayres, D. L., Darling, A., Höhna, S., et al. (2012). MrBayes 3.2: Efficient Bayesian Phylogenetic Inference and Model Choice Across a Large Model Space. *Syst. Biol.* 61, 539–542. doi:10.1093/sysbio/sys029.

Rosenfeld, N., Elowitz, M. B., and Alon, U. (2002). Negative Autoregulation Speeds the Response Times of Transcription Networks. *J. Mol. Biol.* 323, 785–793. doi:10.1016/S0022-2836(02)00994-4.

Ruokolainen, S., Ng, Y. P., Broholm, S. K., Albert, V. A., Elomaa, P., and Teeri, T. H. (2010). Characterization of SQUAMOSA-like genes in Gerbera hybrida, including one involved in reproductive transition. *BMC Plant Biol.* 10, 128. doi:10.1186/1471-2229-10-128.

Sargent, R. D. (2004). Floral symmetry affects speciation rates in angiosperms. *Proc. R. Soc. Lond. B Biol. Sci.* 271, 603–608. doi:10.1098/rspb.2003.2644.

Sauquet, H., Balthazar, M. von, Magallón, S., Doyle, J. A., Endress, P. K., Bailes, E. J., et al. (2017). The ancestral flower of angiosperms and its early diversification. *Nat. Commun.* 8, 16047. doi:10.1038/ncomms16047.

Sengupta, A., and Hileman, L. C. (2018). Novel Traits, Flower Symmetry, and Transcriptional Autoregulation: New Hypotheses From Bioinformatic and Experimental Data. *Front. Plant Sci.* 9. doi:10.3389/fpls.2018.01561.

Shepard, K. A., and Purugganan, M. D. (2002). The genetics of plant morphological evolution. *Curr. Opin. Plant Biol.* 5, 49–55. doi:10.1016/S1369-5266(01)00227-8.

Singh, A., and Hespanha, J. P. (2009). Evolution of gene auto-regulation in the presence of noise. *IET Syst. Biol.* 3, 368–378. doi:10.1049/iet-syb.2009.0002.

Smith, J. F., Hileman, L. C., Powell, M. P., and Baum, D. A. (2004). Evolution of GCYC, a Gesneriaceae homolog of CYCLOIDEA, within Gesnerioideae (Gesneriaceae). *Mol. Phylogenet. Evol.* 31, 765–779. doi:10.1016/j.ympev.2003.09.012.

Smith, M. (1917). Curtis's Botanical Magazine, Maurandia purpusii, Matilda Smith del., John Nugent Fitch lith.

Sordino, P., van der Hoeven, F., and Duboule, D. (1995). *Hox* gene expression in teleost fins and the origin of vertebrate digits. *Nature* 375, 678–681. doi:10.1038/375678a0.

Spitz, F., Gonzalez, F., Peichel, C., Vogt, T. F., Duboule, D., and Zákány, J. (2001). Large scale transgenic and cluster deletion analysis of the *HoxD* complex separate an ancestral regulatory module from evolutionary innovations. *Genes Dev.* 15, 2209–2214. doi:10.1101/gad.205701.

Stanke, M., Steinkamp, R., Waack, S., and Morgenstern, B. (2004). AUGUSTUS: a web server for gene finding in eukaryotes. *Nucleic Acids Res.* 32, W309–W312. doi:10.1093/nar/gkh379.

Starkevič, P., Paukštytė, J., Kazanavičiūtė, V., Denkovskienė, E., Stanys, V., Bendokas, V., et al. (2015). Expression and Anthocyanin Biosynthesis-Modulating Potential of Sweet Cherry (Prunus avium L.) MYB10 and bHLH Genes. *PloS One* 10, e0126991. doi:10.1371/journal.pone.0126991.

Stern, D. L. (2013). The genetic causes of convergent evolution. *Nat. Rev. Genet.* 14, 751–764. doi:10.1038/nrg3483.

Stull, G. W., Stefano, R. D. de, Soltis, D. E., and Soltis, P. S. (2015). Resolving basal lamiid phylogeny and the circumscription of Icacinaceae with a plastome-scale data set. *Am. J. Bot.* 102, 1794–1813. doi:10.3732/ajb.1500298.

Sturm, J., and Strum, J. G. (1796). *Deutschlands Flora in Abbildungen (nur Tafeln)*. Stuttgart Available at: http://caliban.mpipz.mpg.de/sturm/flora/sturm.pdf.

Su, S., Xiao, W., Guo, W., Yao, X., Xiao, J., Ye, Z., et al. (2017). The CYCLOIDEA–RADIALIS module regulates petal shape and pigmentation, leading to bilateral corolla symmetry in Torenia fournieri (Linderniaceae). *New Phytol.* 215, 1582–1593. doi:10.1111/nph.14673.

Sutton, D. A. (1988). *A revision of the tribe Antirrhineae*. British Museum (Natural History).

Tao, Z., Shen, L., Liu, C., Liu, L., Yan, Y., and Yu, H. (2012). Genome-wide identification of SOC1 and SVP targets during the floral transition in Arabidopsis. *Plant J.* 70, 549–561. doi:10.1111/j.1365-313X.2012.04919.x.

The Jepson Herbarium, Antirrhinum filipes *Jepson Herb. Univ. Calif. Berkeley*. Available at: http://ucjeps.berkeley.edu/eflora/eflora_display.php?tid=13566 [Accessed March 6, 2019].

The Jepson Herbarium, Gratiola neglecta *Jepson Herb. Univ. Calif. Berkeley*. Available at: http://ucjeps.berkeley.edu/eflora/eflora_display.php?tid=27298 [Accessed March 6, 2019].

The Plant List version 1.1 Available at: http://www.theplantlist.org/ [Accessed March 6, 2019].

The Plant List version 1.1, Digitalis purpurea *Plants World Online Kew Sci. World Online*. Available at: http://powo.science.kew.org/taxon/urn:lsid:ipni.org:names:802077-1 [Accessed March 6, 2019].

Toth Hervay, N., Hodurova, Z., Balazfyova, Z., and Gbelska, Y. (2011). Autoactivated KlPDR1 gene in the control of multidrug resistance in Kluyveromyces lactis. *Can. J. Microbiol.* 57, 844–849. doi:10.1139/w11-071.

True, J. R., and Carroll, S. B. (2002). Gene co-option in physiological and morphological evolution. *Annu. Rev. Cell Dev. Biol.* 18, 53–80. doi:10.1146/annurev.cellbio.18.020402.140619.

University of British Columbia Botanical Garden, Hemiphragma heterophyllum *Bot. Photo Day*. Available at: https://botanyphoto.botanicalgarden.ubc.ca/2008/04/hemiphragma_heterophyllum_1/ [Accessed March 6, 2019].

Vamosi, J. C., and Vamosi, S. M. (2010). Key innovations within a geographical context in flowering plants: towards resolving Darwin's abominable mystery. *Ecol. Lett.* 13, 1270–1279. doi:10.1111/j.1461-0248.2010.01521.x.

Vincent, C. A., and Coen, E. S. (2004). A temporal and morphological framework for flower development in Antirrhinum majus. *Can. J. Bot.* 82, 681–690. doi:10.1139/b04-042.

Walters, S. M. (2000). *The European Garden Flora: A Manual for the Identification of Plants Cultivated in Europe, Both Out-of-doors and Under Glass*. Cambridge University Press.

Wang, B., Zhang, N., Guo, C.-C., Xu, G.-X., Kong, H.-Z., and Shan, H.-Y. (2012). Evolutionary divergence of the APETALA1 and CAULIFLOWER proteins. *J. Syst. Evol.* 50, 502–511. doi:10.1111/j.1759-6831.2012.00211.x.

Wang, Z., Luo, Y., Li, X., Wang, L., Xu, S., Yang, J., et al. (2008). Genetic control of floral zygomorphy in pea (*Pisum sativum* L.). *Proc. Natl. Acad. Sci.* 105, 10414–10419. doi:10.1073/pnas.0803291105.

Watson, L., and Dallwitz, M. J. (2018). The families of flowering plants: descriptions, illustrations, identification, and information retrieval. *Fam. Flower. Plants Descr. Illus. Identif. Inf. Retr.* Available at: http://delta-intkey.com/angio/ [Accessed March 16, 2018].

Werner, T., Koshikawa, S., Williams, T. M., and Carroll, S. B. (2010). Generation of a novel wing colour pattern by the Wingless morphogen. *Nature* 464, 1143–1148. doi:10.1038/nature08896.

Willkomm, H. M. (1881). *Illustrationes florae Hispaniae insularumque Balearium*. Stuttgart Available at: http://www.plantillustrations.org/illustration.php?id_illustration=239337&SID=0&mobile=0&code_category_taxon=9&size=1.

Xu, S., Luo, Y., Cai, Z., Cao, X., Hu, X., Yang, J., et al. (2013). Functional Diversity of CYCLOIDEA-like TCP Genes in the Control of Zygomorphic Flower Development in Lotus japonicus. *J. Integr. Plant Biol.* 55, 221–231. doi:10.1111/j.1744-7909.2012.01169.x.

Yang, X., Cui, H., Yuan, Z.-L., and Wang, Y.-Z. (2010). Significance of consensus CYC-binding sites found in the promoters of both *ChCYC* and *ChRAD* genes in *Chirita heterotricha* (Gesneriaceae). *J. Syst. Evol.* 48, 249–256. doi:10.1111/j.1759-6831.2010.00086.x.

Yang, X., Pang, H.-B., Liu, B.-L., Qiu, Z.-J., Gao, Q., Wei, L., et al. (2012). Evolution of Double Positive Autoregulatory Feedback Loops in CYCLOIDEA2 Clade Genes Is Associated with the Origin of Floral Zygomorphy[W]. *Plant Cell* 24, 1834–1847. doi:10.1105/tpc.112.099457.

Yanhui, C., Xiaoyuan, Y., Kun, H., Meihua, L., Jigang, L., Zhaofeng, G., et al. (2006). The MYB Transcription Factor Superfamily of Arabidopsis: Expression Analysis and Phylogenetic Comparison with the Rice MYB Family. *Plant Mol. Biol.* 60, 107–124. doi:10.1007/s11103-005-2910-y.

Ye, L., Wang, B., Zhang, W., Shan, H., and Kong, H. (2016). Gains and Losses of Cis-regulatory Elements Led to Divergence of the Arabidopsis APETALA1 and CAULIFLOWER Duplicate Genes in the Time, Space, and Level of Expression and Regulation of One Paralog by the Other. *Plant Physiol.* 171, 1055–1069. doi:10.1104/pp.16.00320.

Yuan, Z., Gao, S., Xue, D.-W., Luo, D., Li, L.-T., Ding, S.-Y., et al. (2009). *RETARDED PALEA1* controls palea development and floral zygomorphy in rice. *Plant Physiol.* 149, 235–244. doi:10.1104/pp.108.128231.

Zhang, W., Kramer, E. M., and Davis, C. C. (2010). Floral symmetry genes and the origin and maintenance of zygomorphy in a plant-pollinator mutualism. *Proc. Natl. Acad. Sci.* 107, 6388–6393. doi:10.1073/pnas.0910155107.

Zhong, J., and Kellogg, E. A. (2015a). Duplication and expression of CYC2-like genes in the origin and maintenance of corolla zygomorphy in Lamiales. *New Phytol.* 205, 852–868. doi:10.1111/nph.13104.

Zhong, J., and Kellogg, E. A. (2015b). Stepwise evolution of corolla symmetry in *CYCLOIDEA2*-like and *RADIALIS*-like gene expression patterns in Lamiales. *Am. J. Bot.* 102, 1260–1267. doi:10.3732/ajb.1500191.

Zhong, J., Preston, J. C., Hileman, L. C., and Kellogg, E. A. (2017). Repeated and diverse losses of corolla bilateral symmetry in the Lamiaceae. *Ann. Bot.* doi:10.1093/aob/mcx012.

Zhou, X.-R., Wang, Y.-Z., Smith, J. F., and Chen, R. (2008). Altered expression patterns of TCP and MYB genes relating to the floral developmental transition from initial zygomorphy to

actinomorphy in *Bournea* (Gesneriaceae). *New Phytol.* 178, 532–543. doi:10.1111/j.1469-8137.2008.02384.x.