Development and Evolution of Complex Reproductive Traits in the
Brassicaceae and Cleomaceae

by

Shane Francis Carey

A thesis submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

in
Systematics and Evolution

Department of Biological Sciences
University of Alberta

Abstract

Flowering plants exhibit a bewildering diversity of forms, which raises fundamental questions on how that diversity arises. A cornerstone of evolutionary developmental biology is the expansion of comparative landscapes and establishment of focal clades that enable investigation of complex and ecologically important traits. The overarching goal of this thesis is to elucidate the developmental and genetic bases of ecologically important traits in sister families Brassicaceae and Cleomaceae. I analyze developmental and anatomical data, which are key to interpreting gene expression and functional studies. RNA-seq is a powerful tool to examine differential gene expression between and within developmental stages, floral organs, and species. While gene expression data are crucial for exploring the genetics underlying morphologies across taxa, they are only correlational. As such, I established the use of virus-induced gene silencing in Cleomaceae, which provides a causal link between gene function and phenotype. Combined, these approaches are essential for assessing the genetic pathways underlying diversity of form.

Using these approaches, I address a range of questions that answer how changes in gene regulation have led to morphological diversity in Cleomaceae and Brassicaceae. First, I examine how a novel fruit morphology of segmented fruits in two species of Brassicaceae is likely due to modification of an existing fruit-patterning pathway. Second, I demonstrate that there is a high degree of conservation in the genetic basis of nectary formation and nectar secretion between *Cleome violacea* and model species *Arabidopsis thaliana*. Next, I show that a key regulator, which has been recruited across the angiosperms for the formation of monosymmetry is also necessary to impart monosymmetry in *C. violacea*. Finally, I show that differences in floral color

between two species of Cleomaceae is due to broad downregulation of the anthocyanin biosynthesis pathway, and modulation of the carotenoid biosynthesis pathway. Thus, my work examines floral and fruit traits in two families. Altogether, my thesis discusses the evolution of symmetry, pigmentation, and nectar gland development in Cleomaceae, and fruit dehiscence/disarticulation in Brassicaceae. In each case, parallel evolution explains much of the observed diversity of form, i.e., across the core eudicots similar gene pathways are responsible for the same traits, such as fruit dehiscence, nectary initiation, monosymmetry, and pigmentation. However, there is also novel gene expression underlying each explored trait, denoting the complex and beautiful nature of fruit and floral evolution.

Preface

A version of Chapter 2 was published on July 18th, 2019 (Carey et al., 2019). The experiment was designed by Dr. Jocelyn Hall, Kerrin Mendler, and me. Data collection and transcriptomic assembly was carried out by K. Mendler and me. I analyzed the data and led the writing. Dr. Hall, K. Mendler, and I edited the manuscript.

A version of Chapter 3 was published on June 14th, 2021 (Carey et al., 2021). Dr. Peter Mankowski conducted preliminary experiments and designed constructs for virus-induced gene silencing (VIGS). Dr. Mónica Higuera, Alex Rocca, Dr. Hall and I collected all presented data. Data analysis for VIGS was shared between Dr. Higuera and me. I designed experiments and wrote the initial draft. Dr. Hall, Dr. Higuera, and I edited the manuscript.

A version of Chapter 4 was published on February 9th, 2023 (Carey et al., 2023). Dr. Hall and I designed experiments and wrote the manuscript. AJ Deneka and I collected nectar and histological data. Brandi Zenchyzen conducted scanning electron microscopy. All other data was collected and analyzed by me. Dr. Hall, B. Zenchyzen and I edited the manuscript.

A version of Chapter 5 is in prep for submission to American Journal of Botany. It is a compilation of research completed by past members of the Hall lab and me. It builds off preliminary experiments completed by Dr. Mankowski and Melanie Patchell. Data collection for all presented VIGS data were a collaborative effort between Dr. Higuera, Dr. Mankowski, and me. Data analysis was completed by me, and writing was shared between Dr. Hall and myself

Chapter 6 was written by me, and all data was collected/analyzed by me. Dr. Hall, and I edited the manuscript.

"You can swim all day in the Sea of Knowledge and not get wet."

— *Norton Juster, The Phantom Tollbooth*

Acknowledgements

Thank you, Jocelyn, for the immeasurable help you have provided me during my time in your lab. You have helped to guide my life's path. Thank you, Corey, for your advice on my research, and for your advice on working life generally; I regret not going to you for guidance more often. Thank you, Brandi, for laughing with me about the numerous oddities of graduate life; you have been both a colleague and friend. I would also like to thank all the members of my candidacy committee for providing valuable advice on becoming a scientist, and my defense committee for helping me solidify my place in academia. Finally, I would like to thank NSERC for providing the funding necessary to complete my research, and the University of Alberta for providing a space to do that research.

Table of Contents

List of Tables

regulation of expression between developmental stages (≈ indicates no SDE). S1, pre-anthetic nectary; S2, anthetic nectary; S3, post-anthetic nectary.

Page 110: Table 4.2. Genes of interest with direct roles in nectar production with relative expression values from our transcriptomic dataset, putative roles, and relevant citations. SDE, Significant Differential Expression with arrows representing either up or down regulation of expression between developmental stages (≈ indicates no SDE). S1, pre-anthetic nectary; S2, anthetic nectary; S3, post-anthetic nectary.

Page 111: Table 4.3. Genes of interest with direct roles in nectary formation with relative expression values from our transcriptomic dataset, putative roles and relevant citations. SDE, Significant Differential Expression with arrows representing either up or down regulation of expression between developmental stages (≈ indicates no SDE). S1, pre-anthetic nectary; S2, anthetic nectary; S3, post-anthetic nectary.

Page 112: Table S4.1. Software versions and references for programs used in this study.

Page 112: Table S4.2. Automatically generated Trinity assembly statistics.

Page 112: Table S4.3. BUSCO analysis using version 5.1.2 with the Viridiplantae_odb10 database. Analysis was run in transcriptome mode using hmmsearch version 3.1 and metaeuk version 6.a5d39d9.

Page 113-115: Table S4.4. Highest expressed transcripts in the *Cleome violacea* nectary transcriptome (n = 51); transcripts were taken from TransDecoder filtered lists of all significantly differentially expressed transcripts and overall highest TPM transcripts.

Page 116-121: Table S4.5. (A-V) Tables generated from the KEGG automated annotation server using BLAST and bi-directional best hits with the *Arabidopsis thaliana*, *Brassica rapa* and *Tarenaya hassleriana* gene data sets. Transcripts lists for each stage had TPM expression above 10 and a coefficient of variation less than 50. Bolding indicates differences between stages that are greater than 3. Categories with less than 3 were removed. S1 = bud pre-anthesis; S2 = flower at anthesis; S3 senescent flower with ~10mm gynoecium; B. = Biosynthesis; D. Degradation. M. = Metabolism; P. = Pathway.

List of Figures

Page 29: Figure 2.6. Heatmap of contig clustered transcripts from *Erucaria erucarioides* and *Cakile lanceolata* expressed in log2 (TPM) with TMM normalization. Representative transcripts are those with the highest bitscore hits against the TAIR10 database. Bolding indicates a shared orthogroup with other transcriptome. FULa,b,c,d are copies of FUL that are present in some species across the Brassicaceae. TPM = Transcripts Per Million. TMM = Trimmed Mean of M-values. Asterisks indicate significant differential expression between proximal and joint region. (FDR-corrected α = 0.01).

Page 30: Figure S2.1. Graph of top Gene Ontology (GO) terms for *Erucaria erucarioides* (blue) and *Cakile lanceolata* (white). Sample (n) and total (N) raw counts were log2 transformed for interspecies comparison.

Page 48: Figure 3.1 (A) *Cleome violacea* whole plant. (B), ventral view of flower, and (C) lateral view of flower.

Page 49: Figure 3.2 Line drawings of *Cleome violacea*. (A) *Cleome violacea* fruit at maturity and after dehiscence. (B) Diagram of Cleomaceae and Brassicaceae fruits in the transverse plane, showing the lignification layer (ll), separation layer (sl), valve margin (vm), replum (R), valve (V), endocarp layer (en*a/b*), ovary (O), and septum (S).

Page 50: Figure 3.3 *Cleome violacea* subjected to virus-induced gene silencing (VIGS) using both heterologous and endogenous pTRV2-*CvPDS* constructs. (A, B) Leaves after silencing using *AtPDS* displaying (A) pale and (B) variegated photobleaching. (C–E) Leaflets silenced with *AtPDS* displaying (C) pale and (D,E) variegated photobleaching. (F, G) Leaves after treatment with *CvPDS* showing (F) variegated and (G) strong photobleaching. (H–J) Leaflets silenced with *CvPDS* displaying (H) variegated and (I,J) strong photobleaching. (K, L) Untreated *C. violacea* (K) leaf and (L) whole plant. (M, N) Whole plant views of *C. violacea* treated with (M) *AtPDS* and (N) *CvPDS*. Scale bars = 2 mm.

Glossary of Terms

Actinomorphic	Two or more planes of symmetry
Alternative splicing	Multiple mRNA from same pre-mRNA
Angustiseptate	Narrow septum
Dehiscent	Dry fruits that open at maturity
Disarticulation	Separation of fruit joint at maturity
Glucosinolates	Secondary metabolite of order Brassicales
Heteroarthrocarpy	Specialized type of segmented silique with two segments, each contains ovules or rudimentary ovules
Homoplasious	Similar trait with uncommon ancestor
Indehiscent	Fruit that does not open at maturity to release seeds
Monosymmetry	A single plane of symmetry
Neofunctionalization	Novel gene function of duplicated gene
Pantropical	Distributed in many tropical regions
Parallel evolution	Independent taxa evolving similar characteristics
Paralogue	Duplicated gene within same genome
Pleiotropic	Single gene with multiple functions
Polyploid	Greater than one pair of homologous chromosomes
RNA-seq	Current generation Illumina-based sequencing
Silencing-efficacy	Percentage of plants displaying VIGS phenotype
Subfunctionalization	Division of gene function
Synteny	Physical co-localization of genetic information
Valve margin	Threshold of valve and replum
Zygomorphic	One plane of symmetry

Abbreviations

HTS	High Throughput Sequencing
ORF	Open Reading Frame
GO	Gene Ontology
KEGG	Kyoto Encyclopedia of Genes and Genomes
VM	Valve Margin
PTGS	Post Transcriptional Gene Silencing
TRV	Tobacco Rattle Virus
ROS	Reactive Oxygen Species
SNPs	Single Nucleotide Polymorphisms
VIGS	Virus-induced gene silencing

Chapter 1: Introduction

Fruit and flower diversity are integral to angiosperm reproduction. Plants help support pollinator populations and maintain the biodiversity of their ecosystems (van der Kooi et al., 2021), and fruits facilitate seed dispersal (Levin et al., 2003). Flowers and fruits are under selective pressure from pollinators and dispersal agents, respectively, and are an extension of each other, i.e., fruits are the product of successful pollination and produce future generations of pollination reservoirs. Thus, fruits, flowers, and their component traits have an interlinked role in evolution and ecology. Studying such traits is synonymous with studying evolution because evolution is the change of inherited traits in a population over time, and no aspect of biology is untouched by it. Understanding the underlying developmental genetics of fruit and floral traits is key to understanding how they evolve. Herein, I explore the genetic pathways that underpin the diversity of fruits and flowers and discuss what role parallel evolution plays in that diversity.

The overarching goal of my research is to investigate the basis of variable plant traits by exploring the relationship between gene expression and phenotypic variation. Many angiosperm traits exhibit complex evolutionary patterns. I focus on a suite of traits key to plant-pollinator interactions and seed dispersal. Specifically, I examine aspects of fruit variation related to seed dispersal (fruit dehiscence and segmentation), which follows successful pollination (Ellner and Shmida, 1981; Tiffney, 1984; Willis et al., 2014). I study features essential for rewarding pollinators (nectar and nectaries) (Galliot et al., 2006; Pyke, 2016; Yan et al., 2016; Rering et al., 2018; Katzer et al., 2019; Chen et al., 2021; Edwards et al., 2022), as well as interacting and attracting them (monosymmetry and pigmentation) (Gong and Huang, 2009; Nozzolillo et al., 2010; Zhang et al., 2010; Shang et al., 2011; Lázaro and Totland, 2014). Altogether these traits play a crucial role in reproductive success (Pabón-Mora et al., 2014; Dong and Wang, 2015; Bergh et al., 2016; Pfannebecker et al., 2017; Sauquet et al., 2017; Schrager-Lavelle et al., 2017; Sengupta and Hileman, 2018; Kramer, 2019; Nikolov, 2019; Wessinger and Hileman, 2020).

1.1 RNA-seq and VIGS as tools for evo-devo

This thesis integrates a range of methodologies to expand our understanding of trait diversity in an evo-devo context: RNA-seq, histology, and functional genetics. RNA-seq provides valuable quantitative data about mRNA expression patterns (Honaas et al., 2016), and insight about genomic DNA and protein structure because RNA is central to both, i.e., information flows from DNA to RNA and from RNA to protein (Crick, 1970). Briefly, Illumina-based sequencing quantifies all mRNA from a sample at any given moment for any given tissue, like a snapshot of development. This raw data can then be assembled de novo into a transcriptome of all actively coding DNA (Honaas et al., 2016). Provided there is enough raw data (i.e., sequencing depth), a transcriptome can provide information about alternative splicing (Halperin et al., 2021), post-transcriptional modification (Furlan et al., 2021), Single Nucleotide Polymorphisms (SNPs) (Zhao et al., 2019), gene fusion (Haas et al., 2019), and gene expression (Conesa et al., 2016).

RNA-seq is valuable for comparing differences in gene expression patterns between developmental stages, organs, and or species which can be incredibly informative on how pathways may be modified to generate variable morphologies. It is particularly useful for understanding the genetic underpinnings of floral trait variation in non-model systems (because of de novo transcriptome assembly). Consequently, RNA-seq is used in four of the five data chapters in this thesis and has helped shed light on how genetic pathways produce fruit dehiscence, nectary development, monosymmetry and floral pigmentation.

While mRNA expression is an important tool for understanding trait diversity, careful developmental and anatomical characterization of those traits, combined with functional studies, are also essential to understanding how diversity arises. Developmental data informs on phenotypes whereas functional data directly assesses function of genes of interest. An excellent and frequently employed method of testing gene function is virus-induced gene silencing (VIGS). VIGS is a method of testing genes of interest, without the need for knockout lines, via the exploitation of RNA interference (Burch-Smith et al., 2006; Ramegowda et al., 2014). RNA interference is a part of the plant immune response which targets foreign genetic material (e.g., from viruses) (Wilson and Doudna, 2013). VIGS uses a modified tobacco rattle

virus (TRV) and an *Agrobacterium* vector. The TRV contains a cDNA insert corresponding to the gene of interest, which is recognized as foreign material and subsequently targeted by the RNA-induced silencing complex (RISC). Conveniently, RNA-seq in-silico data are representative of exact coding sequences which are one to one with what the RISC downregulates. This sequence similarity implies that VIGS vectors created using RNA-seq are likely to have no off-target sequences because post-transcriptional modification is already factored in (relative to genomic sequencing). VIGS is useful for quickly testing genes of interest in non-model systems where knockout lines are untested or unfeasible.

In Chapter 3, I optimize a VIGS protocol in *Cleome violacea*. As stated above, functional tests are essential to expand beyond correlation of gene expression patterns and establish causality between genes and phenotypes. This protocol is relied upon in Chapter 4 and 5 to test hypotheses regarding nectary development and floral symmetry, respectively. Chapter 4 further optimizes the protocol while functionally testing nectary development genes, which increases the overall efficacy of VIGS in *C. violacea*. Overall, RNA-seq and VIGS are integral to the study of floral evolution because they link gene expression and function.

1.2 Cleomaceae and Brassicaceae are model systems

Model systems are incredibly informative in comparative studies of the developmental and genetic basis of plant traits. They are, by definition, extensively studied taxa used to establish the framework for processes that may be shared with other taxa. *Arabidopsis thaliana* (Brassicaceae) is the premier model for many aspects of plant development (Robles and Pelaz, 2005; Irish, 2017). Other models are used to study traits not found in *Arabidopsis*. For example, *Antirrhinum majus* is a central model for organ fusion and symmetry (Robles and Pelaz, 2005; Provart et al., 2016; Irish, 2017). Maize (Poaceae) has been integral to genetics within and outside the grasses (reviewed in: Strable and Scanlon, 2009). Recently, there has been a growing interest within evo-devo to expand the comparative landscape and establish focal clades that enable investigation of complex and ecologically important traits which may not be present in model species (Damerval and Becker, 2017; Kramer et al., 2017). Because of this interest, our knowledge beyond model species is growing substantially (Roy et al., 2017;

Kramer, 2019; Stein and Granot, 2019; Slavković et al., 2021) and there are many outstanding questions remaining to be explored.

Cleomaceae and Brassicaceae are well-poised for exploring the genetic basis of plant traits because of a wealth of genetic and developmental data from *Arabidopsis* research, and because of the rich diversity of forms observed in both families (Koenig and Weigel, 2015; Bayat et al., 2018). Both are members of Brassicales, an order characterized by the presence of mustard oils as well as extensive morphological diversity (Cardinal-McTeague et al., 2016). Cleomaceae and Brassicaceae are fascinating sister lineages due to many differences between the families, although they share a floral ground plan of four sepals, four petals, six stamens and a bicarpellate gynoecium (Iltis et al., 2011). Brassicales have undergone multiple whole genome duplication (WGD) events, which are linked to species and trait diversification (Crow and Wagner, 2006; Ren et al., 2018). For example, there is a link between glucosinolates and two WGDs in the Brassicales, and it is thought that the diversification of glucosinolates, after WGD events, drove overall diversification in the order (Schranz et al., 2011; Edger et al., 2015, 2018). Brassicaceae has multiple independent WGD events, and there is evidence for at least one in the Cleomaceae (Cheng et al., 2014; del Pozo and Ramirez-Parra, 2015; Mabry et al., 2020). Thus overall, their rich phenotypical diversity, high rate of gene retention, and proximity to *Arabidopsis*, makes each family a promising group for the study of genetic diversity. Understanding the genetic underpinnings of fruit and flower development in both families will help answer fundamental questions for evolutionary biology and is the primary goal of my thesis, **i.e., do morphologically distinct fruits and flowers develop via similar pathways across the angiosperms?**

1.3 Fruit morphology in Brassicaceae

Brassicaceae exhibits great diversity in fruit morphology, which impacts seed dispersal and plant propagation (reviewed in Łangowski et al., 2016). This diversity is often homoplasious, e.g., *Arabis* and *Arabidopsis* are polyphyletic despite sharing many common traits, such as typical siliques and branched trichomes (Koch et al., 2003). Novel fruit morphologies have evolved independently multiple times in the family, e.g., heteroarthrocarpic (segmented) fruits

in the tribe Brassiceae (Hall et al., 2011), as well as angustiseptate fruits and indehiscence in many tribes (Franzke et al., 2011). In juxtaposition, most Cleomaceae fruit capsules are like *Arabidopsis* siliques, but without the septum (Iltis et al., 2011). *Arabidopsis* is an established model for fruit diversity, but there are still understudied morphologies within the family, such as segmented fruits (Hall et al., 2006, 2011).

All fruits share the function of seed dispersal. Typically, fleshy fruits accomplish seed dispersal with the aid of biological dispersal agents that consume the pericarp and discard or excrete the seed(s). Dry fruits, however, do not always rely on external agents and must disperse their seeds another way. Fruits in the Brassicaceae accomplish seed dispersal via temporal and spatial modulation of dehiscence (reviewed in Łangowski et al., 2016). In this way, they release seeds into their environment during ideal conditions. Many Brassicaceae siliques consist of three main components that enable seed dispersal: valve, replum, and the valve margin. Valve tissue is connected to the replum via the valve margin, which separates at maturity due to tensional forces. Seeds are then released into the environment, and/or remain with the maternal plant (reviewed in Ferrándiz, 2002). In Chapter 2 I investigate two subtypes of segmented fruit that are characterized by a novel transverse dehiscence zone that separates the top and bottom half of the fruits. I explore gene expression in the top, bottom, and segmented regions of both subtypes in order to compare between and within species. In addition, I establish a protocol for RNA-seq that is used in subsequent chapters. This chapter provides evidence that segmented fruits in the Brassicaceae are genetically convergent with *Arabidopsis* siliques, i.e., segmented fruits have co-opted genes that are involved in silique dehiscence.

1.4 Floral morphology in Cleomaceae

Cleomaceae is a family of approximately 270 species distributed worldwide in warmer climates, which is modest when compared to the roughly 3700 species of Brassicaceae distributed in colder climates (Cardinal-McTeague et al., 2016; Bayat et al., 2018). Unlike Brassicaceae, Cleomaceae has few economically important plants, except the minor crop plant *Gynandropsis gynandra* and the ornamental *Tarenaya hassleriana.* In addition to differences in size and

economic importance, the families also vary in terms of morphology. The Cleomaceae may have fewer species, but their flowers show greater diversity than do flowers in Brassicaceae (Bayat et al., 2018). Brassicaceae flowers are relatively uniform in morphology and polysymmetric, although some variation exists (Nikolov, 2019). Flowers of Cleomaceae are typically monosymmetric, can have stalks to reproductive organs, and vary in organ number and size (Bayat et al., 2018).

Nectaries are a prominent feature of Cleomaceae flowers and exhibit a wide range of sizes, shapes, and positions. Across the family, nectaries may be absent, adaxially positioned or annular (Tucker and Vanderpool., 2010). Altogether, they are only linked by their shared function in secreting nectar. This remains true across the angiosperms; nectaries are associated with all plant organs except for roots, and floral nectaries can be associated with any floral organ (Nicolson et al., 2007; Liao et al., 2021). In Chapter 4, I examine the key regulators underlying nectary development and nectar secretion in *Cleome violacea*. Integrating developmental, RNA-seq and VIGS data, I test whether the same regulators initiate nectary development and nectar secretion as they do in other Eudicots, and consequently demonstrate a high degree of conservation with *Arabidopsis* and *Petunia*. Interestingly, a gene known for nectary spur initiation in *Aquilegia* is also expressed in *C. violacea* nectaries. To our knowledge this is the first example of both developmental initiators being expressed simultaneously, which warrants future study. Not only do nectaries have a crucial role in pollinator reward, but they also contribute to monosymmetry (and thus pollinator attraction and orientation).

Monosymmetric flowers have a single plane of symmetry and polysymmetric flowers have two or more (Figure 1.1). Monosymmetry is often imparted by differences in top and bottom segments of the flower, and can involve one or more whorls (Endress, 1999, 2012). In Cleomaceae, monosymmetry is derived primarily by an upwards curvature of petals and reproductive organs (Patchell et al., 2011), although morphological variation within petals is also observed. Monosymmetry is widespread in the angiosperms (reviewed in Endress, 2012) and important for attracting and orienting pollinators to nectar and pollen. My research in Chapter 5 focused on corolla monosymmetry and the widely conserved role of *TCP1* in

promoting it. *Teosinte branched 1/CYCLOIDEA/Proliferating cell factor* (*TCP1*) transcription factors control floral monosymmetry across Eudicots and Monocots (reviewed in Hileman, 2014). In this chapter I show that monosymmetry may be regulated differently than in the model species *Antirrhinum majus* because of unique gene expression patterns for genes downstream of *TCP1*. Further, there appears to only be a single copy of *TCP1* in *C. violacea* when most monosymmetric species have at least two copies.

Monosymmetry is brought about in many ways, and petal color also contributes to pollinator attraction (reviewed in Miller et al., 2011). In Cleomaceae, petal colors range from white and pale pink (e.g., *Gynandropsis gynandra),* to purple (e.g., *Areocleome oxalidea*), yellow (e.g., *Arivela viscosa)*, and orange/red (e.g., *Podandrogyne decipiens*). Cleomaceae petals are also intra-variable, such as the light pink and orange stalk of *Cleome augustifolia*. The anthocyanin and carotenoid biosynthesis pathways produce most floral pigments (Tanaka et al., 2008). Both pathways contribute to the great variety of hues found amongst flowers, and this variety impacts pollinator preference (e.g., bee pollinated flowers are more often violet, yellow, or blue; Fenster et al., 2004). The genetic basis and evolutionary impact of pigmentation is well understood in some taxa but has never been characterized in Cleomaceae. My goal in Chapter 6, therefore, is to establish the genetic basis for floral pigmentation in two species of Cleomaceae, one with and one without pigmented petals. I describe the pathway activity of each species and provide foundational evidence for future research in the family. Surprisingly, carotenoid biosynthesis is active in white petals and appears to contribute to the biosynthesis of phytohormones. Both carotenoid and anthocyanin pathways are highly expressed in pigmented flowers, and there appears to be no significant difference between top and bottom petals.

In summary, I demonstrate that the genetics behind morphologically variant flowers and fruits is often similar between taxa, but never identical. Hereafter, each chapter focuses on key aspects of fruit dispersal and pollinator reward/attraction. They cover, in detail, the genetics underpinning morphological features mentioned above, adding to our general understanding of evolutionary developmental biology, and laying a foundation for future research within the Brassicaceae and Cleomaceae.

1.5 Figures
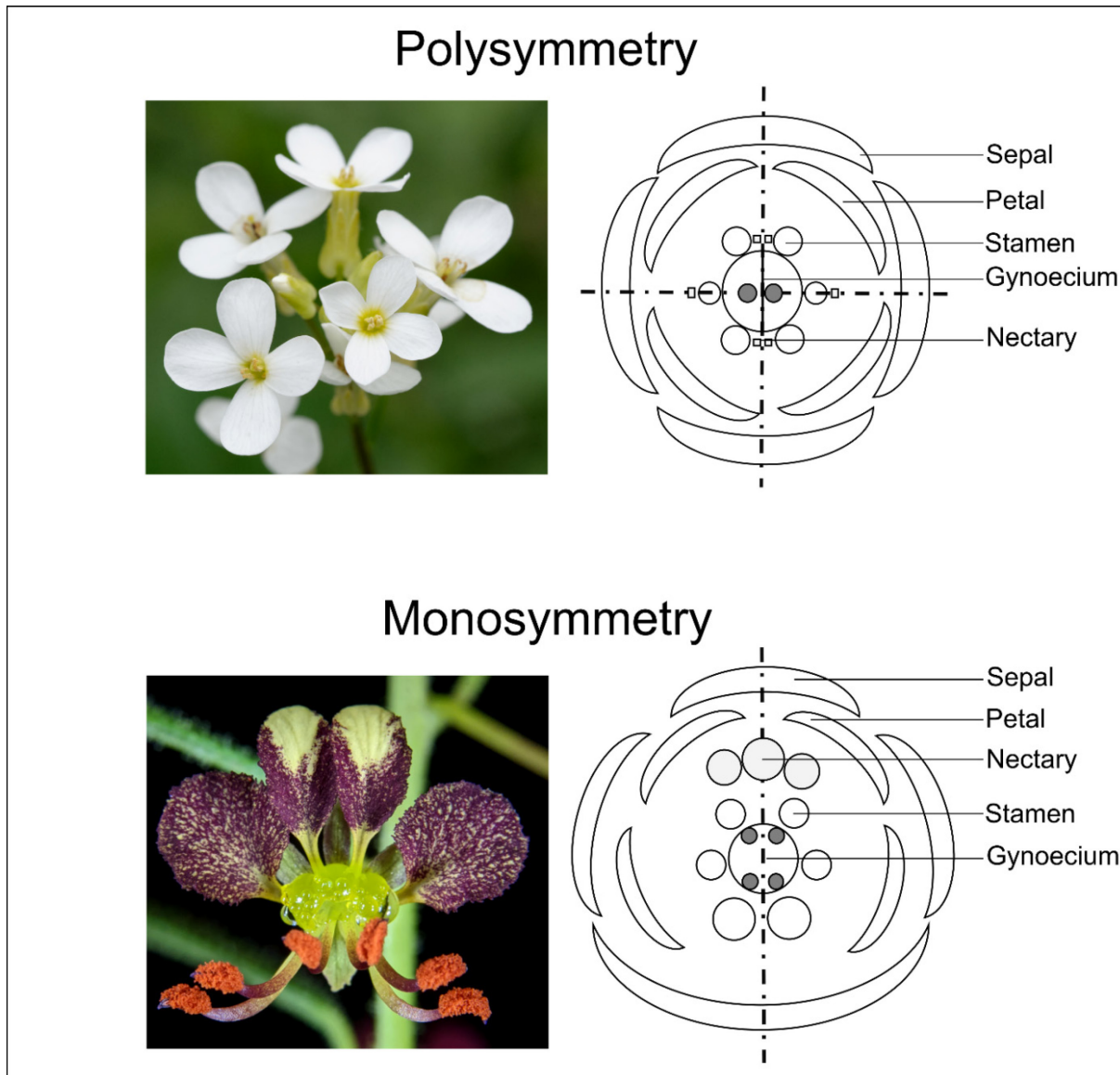
Figure 1.1. Diagram outlining polysymmetry in *Arabidopsis thaliana* and monosymmetry in *Cleome violacea*. Planes of symmetry outlined by dashed line. *Arabidopsis* photo by Marie-Lan Nguyen; Wikimedia Commons; CC-BY 2.5.

Chapter 2: How to Build a Fruit: Transcriptomics of a novel fruit type in the Brassiceae[i]

2.1 Introduction

Studying gene expression patterns across plant structures and species can elucidate how their modification may produce morphological variation (Łangowski et al., 2016; Zhang et al., 2016). Fruits are diverse and ecologically relevant plant structures to investigate because their morphological variation determines how their seeds are dispersed (Janson, 1983; Gautier-hion et al., 1985). There are many fruit morphologies in nature, and they are often categorized as fleshy or dry. Fleshy fruits are distributed primarily by animals, as the seeds are discarded before or after consuming. Dry fruits, however, may be dispersed by animals, wind, or water. Dry fruits are further classified by whether they are dehiscent, releasing seeds into the environment, or indehiscent, releasing seeds in a protected fruit wall propagule. Thus, variation in fruit morphology is directly tied to differences in dispersal capabilities.

Brassicaceae is an exemplary group to investigate the basis of fruit diversity because species in this family vary markedly in shape, structure, and size (Eldridge et al., 2016; Łangowski et al., 2016). Their variation in dehiscence is a focal point for research because it fundamentally changes fruit structure, subsequently affecting dispersal and diversification (Willis et al., 2014). A prerequisite for exploring how differences in fruit morphology are achieved across the Brassicaceae is familiarity with both the fruit structure and underlying genetic pathways of an important member of the family: *Arabidopsis thaliana.* (Dinneny et al., 2005; Avino et al., 2012). As this species is a premier model, it provides an important basis of comparison to species with variable morphology. *Arabidopsis* fruits, hereafter referred to as typical siliques, are composed of five basic elements: valve, replum, seeds, septum, and valve margins. The valve, synonymous with ovary wall in *Arabidopsis*, is the outermost tissue of the fruit that protects the developing seeds and is separated from the replum at maturity to release seeds (Figure 2.1 & 2.2). The replum is the persistent placental tissue to which the seeds are attached. The septum, which

---

[i] A version of this chapter has been published (Carey et al., 2019). Figures and tables labelled as supplemental are due to journal submission requirements and are not reflective of their contribution to the main thesis.

connects to the replum, divides the fruit into two locules or chambers. The valve and replum are separated by the valve margin, which consists of a lignification and separation layer. Thus, proper fruit formation relies on the establishment of medial (replum) and lateral (valves and valve margin) components (González-Reig et al., 2012). As the fruit dries, tension is created via the lignified layer, which facilitates the separation of the valves from the replum at the separation layer (Spence et al., 1996). This general morphology is stable across most dehiscent members of Brassicaceae.

The causal factors for dehiscence have been well characterized in *Arabidopsis* (Gu et al., 1998; Roeder et al., 1998; Girin et al., 2010, 2011), with proper formation and positioning of the valve margin being a key to this process. The valve margin pathway is essential for spatial regulation and development of valve, replum, and valve margin tissues (Ferrandiz, 2002; Dinneny and Yanofsky, 2005; Dinneny et al., 2005; Ferrándiz and Fourquin, 2014; Chávez Montes et al., 2015; Xu et al., 2015; Ballester et al., 2017). Briefly, *FRUITFULL* (*FUL*) and *REPLUMLESS* (*RPL*), as well as other upstream regulators, restrict the expression of the valve margin genes to two cell layers between the valve and replum, respectively. The valve margin genes, *SHATTERPROOF 1/2* (*SHP1/2*), *INDEHISCENT* (*IND*), *SPATULA* (*SPT*), and *ALCATRAZ* (*ALC*), are responsible for the formation of the valve margin, specifically of the separation and lignification layers that control dehiscence (Figure 2.1). Upstream regulators of *FUL* and *RPL*, e.g., *APETALA2* (*AP2*), *FILAMENTOUS FLOWER* (*FIL*), *YABBY3* (*YAB3*), and *JAGGED* (*JAG*) are also key to precise positioning of the valve margin because they tightly regulate downstream processes. In sum, replum and valve genes function in an antagonistic manner to ensure proper formation of these regions of the fruit (González-Reig et al., 2012).

A common modification in fruit morphology across Brassicaceae is indehiscence, which has been observed in at least 20 lineages, implying multiple origins of this trait (Appel and Al-Shehbaz, 2003). There are many genetic modifications that result in an indehiscent *Arabidopsis* fruit, but less is known about the basis of indehiscence observed in other species. For example, a mutation in the following genes results in indehiscent fruits in *Arabidopsis*: *SHP1/2*, *SPT*, *ALC* and *IND* (Kramer and Irish, 1999; Liljegren et al., 2000, 2004; Groszmann et al., 2011).

Overexpression of *FUL* or *NO TRANSMITTING TRACT* (*NTT*) also results in indehiscent fruits (Ferrándiz et al., 2000a; Chung et al., 2013). In studies of other species, fewer genes have been implicated in the indehiscent phenotype. One study demonstrated a deviation in expression of eight key genes between pod shatter sensitive species and shatter resistant species of *Brassica* and *Sinapis* (Zhang et al., 2016). In *Lepidium*, there has been an evolutionary shift from dehiscence to indehiscence, e.g., valve margin genes that are conserved between the dehiscent *L. campestre* and *Arabidopsis* have been lost in the indehiscent *L. apellianum* (Lenser and Theißen, 2013; Mühlhausen et al., 2013). Upregulation in upstream regulator *AP2* has been suggested as a factor in this origin of indehiscence (Mühlhausen et al., 2013).

A notable morphological adaptation is the evolution of a complex fruit type known as heteroarthrocarpy, which is only found in some members of the tribe Brassiceae (Gómez-Campo, 1999; Appel and Al-Shehbaz, 2003; Warwick and Sauder, 2005). This modified silique is defined by the presence of a variably abscising central joint, an indehiscent distal region, and a variably dehiscent proximal region (Figure 2.2). As such, this novel morphology offers an opportunity to investigate fruit variation beyond shifts from dehiscent to indehiscent. Ancestral state reconstructions reveal that typical siliques are ancestral in the tribe with multiple origins to heteroarthrocarpy. However, reconstructions vary in the precise number of times this trait has evolved. In contrast, evolutionary patterns of dehiscence and joint articulation are less clear with closely related taxa exhibiting variation in these features (Hall et al., 2011). Anatomically, heteroarthrocarpic fruits appear most like *Arabidopsis* siliques in their proximal regions, varying by a lack of a valve margin cell layer in indehiscent variants (Gómez-Campo, 1980; Hall et al., 2006, 2011). There are three described variations of heteroarthrocarpy: a non-abscising joint with a dehiscent proximal region, an abscising joint with an indehiscent proximal region, and an abscising joint with a dehiscent proximal region (Hall et al., 2006). These subtypes have evolved multiple times, perhaps as a bet hedging strategy in response to selective pressure from hostile desert environments (Hall et al., 2011; Willis et al., 2014). The evolution of the joint and subsequent heteroarthrocarpic subtypes may be developmental enablers that have facilitated changes in fruit morphology across the tribe, which would explain heteroarthrocarpy's evolutionary lability (Hall et al., 2006). Regardless of lability, all types are linked by the

mechanism in which seeds from the same fruit are released by different means. In other words, the joint is the novel and unifying feature of heteroarthrocarpy (Hall et al., 2006).

A comparison of expression patterns between heteroarthrocarpic subtypes is potentially informative for formulating hypotheses about its evolutionary origins. *Erucaria erucarioides* and *Cakile lanceolata*, hereafter referred to as *Erucaria* and *Cakile*, are two well-studied representatives for heteroarthrocarpy because of their close relation and divergent subtypes (Hall et al., 2006, 2011; Avino et al., 2012; Willis et al., 2014). In all heteroarthrocarpic subtypes the distal region remains indehiscent. *Erucaria* represents the subtype where the proximal region dehisces at maturity, releasing seeds, and the distal region remains attached to the maternal plant via the persistent replum (Figure 2.2). *Cakile* exhibits the subtype where the proximal region also remains indehiscent with joint abscission such that all seeds are dispersed as protected propagules (Figure 2.2). In previous studies it was hypothesized that the formation of the joint is the result of repositioning of the valve margin, such that the valve is only present in the proximal region of the fruit, unlike in *Arabidopsis* where it is found in the entire ovary (Hall et al., 2006). In other words, the joint is the distal portion of the valve margin. This hypothesis was partially supported by comparative gene expression data of some, but not all, genes in the valve margin pathway using a candidate gene approach (Avino et al., 2012). However, that study did not definitively determine how the pathway has been repositioned because it did not investigate upstream genes. Candidate gene approaches will, by design, overlook non-targeted genes, and a lack of in-situ hybridization does not necessarily indicate a lack of expression. Further, the basis of the joint remains unknown. No study to date has investigated transcriptional variation of heteroarthrocarpic fruits sectioned transversely into distal, joint, and proximal regions. This approach is complementary to prior research because it quantifies expression of all transcripts in discrete regions of a whole system. Expression profiles from these regions will elucidate broad patterns and potentially identify key players involved in the formation of heteroarthrocarpy. They will clarify unique and shared gene expression patterns between and within *Erucaria* and *Cakile* and will set the groundwork for future research regarding the evolution of the joint. Herein, the objective is to uncover transcript expression patterns, unique or shared, between and within, two variant heteroarthrocarpic

species. We expect gene expression to be consistent with anatomical features within fruits, and that expression of fruit patterning transcripts will be consistent with repositioning of the valve margin in heteroarthrocarpy.

2.2 Materials and Methods

2.2.1 Plant material

Seeds from *Erucaria erucarioides* (Coss. and Durieu) Müll. Berol and *Cakile lanceolata* (Willd.) O. E. Schulz were obtained from the late César Gómez-Campo's and KEW royal botanical garden's seed collections, respectively. Vouchers for *Cakile* and *Erucaria* have been deposited in the Vascular Plant Herbarium at the University of Alberta, and the Harvard University Herbaria, respectively. Seeds were germinated in 1% agar and transferred to clay pots containing a 2:1 soil (Sungro sunshine mix #4, Agawam, MA, USA) to perlite mixture. Plants were grown under a 16/8-hour light/dark schedule at 24°C with scheduled watering in the University of Alberta, Department of Biological Sciences, growth chambers.

Distal, joint, and proximal regions from 10 mm fruits (~10 days post fertilization) were collected and flash frozen in liquid nitrogen prior to storage at -80°C. Distal and proximal regions were classified as all tissue ~1 mm above or below the joint, and the joint is remaining tissue between distal and proximal regions (Figure 2.2). The 10 mm fruit size is roughly equivalent to *Arabidopsis* stage 17A fruits (Roeder and Yanofsky, 2006), which go through elongation and cell expansion before maturity. This size was chosen to capture late-stage valve margin gene expression because the valve margin is easily distinguished at this stage, and an increase in lignification is observed in key layers, e.g., enb.

2.2.2 RNA isolation and cDNA library preparation

RNA was extracted from frozen tissue using manual grinding and a Qiagen RNeasy micro kit (Hilden, Germany) with the following amendments to protocol: RNA was incubated in nuclease free water for five minutes prior to elution, and this eluate was spun through the same extraction column to maximize RNA yield. RNA concentration was verified using a Nanodrop ND-1000 spectrophotometer (Software version 3.1.2), and quality was confirmed using the

Agilent 2100 bioanalyzer (Software version B.02.09.SI720). All cDNA samples were set at the same concentration of the most dilute RNA extraction. Samples were processed using the Illumina TruSeq stranded mRNA LT sample prep kit RS-122-2101 (California, U.S.), and the procedure was followed as described in the low sample protocol. The mRNA from each sample was isolated and purified using AMPure XP magnetic beads (Agencourt; Beverly, Massachusetts) before primary and secondary strand cDNA synthesis. Unique Illumina adapters were ligated, and each sample was PCR amplified before validation. PCR was run for 15 cycles of: 98°C for 10 seconds, 60°C for 30 seconds and 72°C for 30 seconds followed by 5 minutes at 72°C and a final hold at 4°C. Samples were normalized, pooled, and sequenced by the center for applied genetics (TCAG) facilities of the Toronto Sick Kids hospital, Ontario, Canada.

2.2.3 De novo transcript assembly, differential expression, and annotation

Raw reads were trimmed and quality checked using Trim Galore! (Version 0.4.1) (Krueger, 2012) and FastQC (Version 0.11.3) (Andrews, 2010) then assembled using Trinity (Version 2.2.0) (Grabherr et al., 2011). The raw reads are available at the Sequence Read Archive (SRA) database under the BioProject ID PRJNA545186. The transcriptome shotgun assembly projects have been deposited at DDBJ/EMBL/GenBank under the accessions GHNY00000000 and GHOR00000000 for *Erucaria* and *Cakile*, respectively. The versions described in this paper are the first versions. Corset (Version 1.0.6) (Davidson and Oshlack, 2014) was used to estimate contig abundance by grouping contigs into representative gene clusters as the first step of the differential expression analysis. Contigs are defined as continuous overlapping paired-end reads. Next, edgeR (Version 3.6.2) (Robinson et al., 2009) was used to perform pairwise differential expression analysis of Trinity gene, Trinity contig, and Corset clusters between proximal, joint, and distal regions of fruits from the same species. Genes, contigs, and clusters were classified as significantly differentially expressed if log2(fold-change) was greater than 2 and the False Discovery Rate (FDR)-corrected p-value (α) was less than 0.05. The 'analyze_diff_expr.pl' script, provided with Trinity, was used to generate z-score heatmaps of all significantly differentially expressed contig clustered transcripts (α < 0.05). A z-score is used to indicate how many standard deviations a value is above or below the mean. The transcriptomes were annotated using the Basic Local Alignment Search Tool (BLAST) (Altschul et al., 1990)

algorithm on a local copy of both the National Center for Biotechnology Information (NCBI) non-redundant protein (nr) database and The *Arabidopsis* Information Resource (TAIR) 10 database (Garcia-Hernandez et al., 2002). BLASTx (E-value < 10^{-10}) was used to identify highly similar sequences, and transcripts with the highest bit-score from the TAIR database were used as representative transcripts for heatmap generation. Whole transcriptome and fruit patterning heatmaps were generated using ggplot2 (Wickham, 2009) and ggplot in R, respectively (Version 3.4.2)(R Core Team, 2013). These global gene expression patterns were compared to previously published in situ hybridization and semi-quantitative PCR of select fruit genes (Avino et al., 2012). Assembly completeness was determined using Benchmarking Universal Single Copy Orthologs (BUSCO) (Version 2.0) (Simão et al., 2015) (Table S2.1).

2.2.4 Orthologous clustering

Orthofinder (Version 1.1.8) (Emms and Kelly, 2015) was used to group orthologous transcripts from unfiltered *Erucaria* and *Cakile* transcriptomes. These groupings (orthogroups) with transcripts from both species as well as top BLAST matches for fruit patterning genes of interest were used to generate heatmaps. For Venn diagram generation, high-throughput sequencing (HTS) (Rau et al., 2013) filtered transcripts, sorted by regions, were translated to longest open reading frame (ORF) protein fasta files using TransDecoder (Version 5.0.0) (Haas et al., 2013). These files were uploaded for comparison using the Orthovenn webserver (Wang et al., 2015). HTS filtering was used to reduce file size due to the web server upload limit, and to reduce the number of insubstantial transcripts.

2.2.5 Gene ontology

Transcriptome fasta files from *Erucaria* and *Cakile* were imported to BLAST2GO (Version 2.8) (Conesa et al., 2005). Annotation files were exported and filtered, using a merged profile of all three biological replicates, to generate gene ontology (GO) terms for each region and species. These GO terms were used to produce graphs containing transcriptome hits for chosen terms. Terms were chosen based on searches for lignin, abscission, dehiscence, specific hormone keywords, and top hits. For comparison between transcriptomes, the log2 of selected GO term counts were divided over the log2 of all GO term counts (log2(n)/log2(N)).

2.3 Results

2.3.1 De novo assembly of Erucaria and Cakile transcriptome data

RNA-seq libraries were constructed from nine total replicates of triplicate distal, proximal, and joint regions. RNA samples from segmented fruits of two distinct plants were combined before sequencing to achieve optimal yield for library preparation. Sequencing from both libraries averaged 27.41 and 29.41 million paired-reads for *Erucaria* and *Cakile*, respectively. After quality trimming, read counts were reduced to 27.36 million and 28.36 million high quality reads, respectively. Inter-quartile ranges per base were minimally 33 for *Erucaria* for the first 5 base pairs, and minimally 32 in the 90th percentile; *Cakile's* inter-quartile ranges were minimally 33 for the first 5 base pairs, and minimally 29 in the 90th percentile.

The transcriptome from *Erucaria* had an average contig length of 942.83, and *Cakile's* had an average length of 877.15. The total transcript count for *Erucaria* and *Cakile* was 227,530 and 314,194 reads, respectively (Table 2.1). Corset cluster counts averaged 365,257 (*Erucaria*) and 436,177 (*Cakile*). Notably, the first replicate for *Cakile* had a read count of 269,732, which is minimally 130,000 fewer than replicate two and three. This inconsistency may have caused some issues in downstream analyses, but overall, both transcriptomes were of adequate quality and read-depth. This is supported by a BUSCO analysis, as *Erucaria* and *Cakile's* assemblies had overall completeness of 96.4% and 94.8%, respectively (Table S2.1).

2.3.2 Annotation of assembled transcripts

Both transcriptomes were compared to the nr and TAIR peptide database using a BLASTx algorithm, and all downstream analyses used the TAIR10 annotation for facilitated comparison to *Arabidopsis*. A total of 254,592 (*Cakile*) and 213,757 (*Erucaria*) transcripts with e-values ≤ 10^{-5} were matched to the TAIR10 database with multiple transcripts matched per gene. The GO analysis averaged 8,644 and 8,941 terms for *Erucaria* and *Cakile*, respectively. The top 15 GO terms consisted of 11 cellular component, three molecular function, and one biological process. Nucleus, plasma membrane, and protein binding were the top three terms, all of which are biological processes (Figure S2.1).

The majority of selected orthogroups were similar between and within species (lignin, abscission, dehiscence processes, and hormone response) (Figure 2.3). Exceptions included: cell wall modification related to abscission, general abscission, and catabolic lignification. *Cakile* has a greater ratio of cell wall modification processes, and a lower ratio of general abscission processes relative to *Erucaria*. *Erucaria* has a higher ratio of catabolic lignification processes in the joint region despite having similar ratios relative to *Cakile* in the distal and proximal regions (Figure 2.3). Overall, the GO analysis results are consistent between and within species.

Additional results from OrthoVenn showed minimal differences in orthologous clustering within species, but some differences between species (Figure 2.4). There are a greater number of shared clusters between the proximal and distal regions in *Erucaria* (2548) than *Cakile* (2306) despite *Cakile* having substantially more overall clusters than *Erucaria* (50,003 vs 32,757). Additionally, there are fewer clusters unique to the joint for *Cakile* (21) than *Erucaria* (112). In sum, there are fewer orthologous clusters in common within regions of *Cakile* fruits than within regions of *Erucaria* fruits.

2.3.3 Identification of differentially expressed transcripts in 10 mm fruits

For whole transcriptome comparison, two heatmaps of significant pairwise differentially expressed transcripts ($\alpha = 0.01$) were generated (Figure 2.5). Contig clustering was chosen for this analysis because it is a more conservative estimation of significant differential expression at the transcript level, i.e., there are a greater number of transcripts being compared with more stringent FDR correction relative to corset clustering. Values were then converted to z-score to facilitate interspecies comparison, and for visual clarity. Dendrograms highlight the differences in number of differentially expressed transcripts between both species, and show that all replicates clustered together, respectively. The joint and proximal regions of *Erucaria* are most alike in expression and are both dissimilar to the distal region (Figure 2.5). All three regions in *Cakile* have different expression patterns, and the distal region has a relatively large inter-replicate variance (Figure 2.5). There are 15,345 (*Erucaria*) and 74 (*Cakile*) significantly differentially expressed (SDE) transcripts in each transcriptome. There were no SDE *Cakile* transcripts with FDR-adjusted p-values < 0.01. The low number of SDE genes between *Cakile*

regions indicates a lack of regional distinction in terms of transcript expression. These data demonstrate a large difference in significant differential expression between the distal region relative to the joint and proximal region in *Erucaria*, and little significant variation between all three *Cakile* regions.

We compared expression profiles of 21 genes important for valve margin formation and positioning in *Arabidopsis* (Rajani and Sundaresan, 2001; Semiarti et al., 2001; Pinyopich et al., 2003; Guo et al., 2008; Girin et al., 2011; Kay et al., 2013; Romera-Branchat et al., 2013; Jaradat et al., 2014; Marsch-Martínez et al., 2014; Schiessl et al., 2014; Chávez Montes et al., 2015; Schuster et al., 2015; Simonini et al., 2016; Zhang et al., 2016; Zumajo-Cardona and Pabón-Mora, 2016) (Figure 2.6). Contig clustered transcripts were also chosen for this analysis based on matches against the TAIR database. Row dendrograms highlight the different clustering of expression patterns of VM genes between both species, e.g., *ETT*, *RPL* and *BP* cluster together in *Erucaria* but not in *Cakile* (Figure 2.6). Most fruit patterning genes for both species have no significant differences in expression across all regions, except for *FIL* and *YAB3* which were significantly upregulated in the distal region relative to the joint in *Erucaria*, and *IND* which was significantly upregulated in the joint relative to both the distal and proximal regions in *Cakile*. Upstream regulators *FIL* and *YAB3* are not expressed in late stage *Cakile* fruits, despite global expression in *Erucaria* fruits. Downstream regulator *IND* is expressed in the whole fruit in *Erucaria*, but only in the joint region of *Cakile* (Figure 2.6).

2.4 Discussion

2.4.1 Gene ontology of heteroarthrocarpic fruits

Overall, GO terms within fruits and between species are similar (Figure 2.3 and S2.1), as expected, because all sections and replicates are from developing fruit with shared components (e.g., ovary wall and septum). Additionally, GO analyses of top terms do not usually vary between closely related species (Busch et al., 2014; Sinha et al., 2015). However, despite similarities in gene ontology, the origin of heteroarthrocarpy may still be explained by deviation in expression patterns of one or more of the valve margin pathway genes (Kramer and Irish, 1999; Liljegren et al., 2000, 2004; Ferrandiz, 2002; Groszmann et al., 2011; Chung et al., 2013).

Similarities in gene ontology do not imply similarity between all expressed transcripts, so variation of just a few transcripts may be the driving factor behind heteroarthrocarpy.

2.4.2 Global transcript expression within heteroarthrocarpic fruits is consistent with anatomy

Transcript expression patterns are consistent with anatomical variances within and between fruits. The distal region of *Erucaria* has opposing transcript expression relative to both its joint and proximal regions (Figure 2.5), i.e., when transcripts are upregulated distally in *Erucaria* they are downregulated proximally. This pattern is consistent with heteroarthrocarpic fruit anatomy, as distal regions contain no valve or valve margin, and proximal regions have both (Hall et al., 2006). In contrast, all regions of *Cakile* have variable transcript expression, with the clearest distinction between the proximal and joint regions, i.e., when genes are upregulated proximally, they will be downregulated in the joint (Figure 2.5). As with *Erucaria*, expression profiles in *Cakile* vary in a manner consistent with anatomy. Superficially, one might expect the *Cakile* silique to have similar expression between all regions because the entire fruit is indehiscent, which is consistent with the pattern of significantly fewer DE genes between regions of *Cakile* (74) than *Erucaria* (15,345) (Figure 2.5). However, anatomically, the distal region of *Cakile* is more like the distal region of *Erucaria* than to its own proximal region, and its abscising joint is anatomically reminiscent of a valve margin (Hall et al., 2006). Thus, we would expect regions of the fruit to exhibit different expression patterns, which is supported by our data (Figure 2.5). Abscission zones are also found between septum and seeds, and they too share similar anatomy and expression to typical silique valve margins (Balanzà et al., 2016). Heteroarthrocarpic distal regions are unlike indehiscent non-heteroarthrocarpic siliques such as *L. appelianum*, because heteroarthrocarpic distal regions have no remnant valve margin in contrast to indehiscence observed in *Lepidium* and the proximal region of *Cakile* (Hall et al., 2006; Mühlhausen et al., 2013). Thus, we expect different expression patterns within heteroarthrocarpic fruits, as well as between heteroarthrocarpic and non-heteroarthrocarpic fruits. In summary, there is a clear difference between distal and proximal expression profiles for both *Erucaria* and *Cakile*, which is consistent with a repositioning of the valve margin, i.e., the distal region is quite distinct from the proximal region due to the lack of valve margin, or its

remnant, in the distal region. This consistency is further explored by analysis of fruit patterning transcript expression involved in valve margin formation.

2.4.3 Fruit patterning genes

Despite the substantial differences in anatomy, most valve margin genes reveal similar expression patterns across fruits in both *Erucaria* and *Cakile* (Figure 2.6). Although overall these expression patterns are consistent with a previous investigation of some, but not all, members of the VM pathway (Spence et al., 1996), they vary with regards to expression of one gene in *Erucaria*. *EeFUL1*, one of two *FUL* homologs found in *Erucaria*, was previously shown to only be expressed in the proximal region in earlier stages of carpel development (Avino et al., 2012), but all *FUL* transcripts are expressed across all regions in this study of later stage development (Figure 2.6). This discrepancy may be due to dynamic gene expression at different stages or because our methodology cannot distinguish within region differences (e.g., genes expressed in valve but not replum), so differences within regions cannot be distinguished. In contrast to *EeFUL1*, our data are consistent with a previous publication which demonstrated that other fruit patterning genes have broader expression domains than observed in *Arabidopsis* (Avino et al., 2012). *EeALC* and *EeIND* and *ClALC* were expressed in the septum of *Erucaria* and *Cakile*, respectively, which is found throughout all regions sampled in this study.

It is a compelling finding that upstream regulators *FIL/YAB3* and *JAG* have variable expression across *Erucaria* (Figure 2.6). These three genes positively regulate expression of *FUL* and valve margin genes in *Arabidopsis* such that their cooperative function has been designated together as *JAG/FIL* activity (Ferrándiz and Fourquin, 2014). Our data suggest a decoupling of this cooperation in heteroarthrocarpic fruits because these three genes do not exhibit the same expression patterns across *Erucaria* fruits (Figure 2.6). That is, no expression of *JAG* was detected in any region of either species at this stage. *FIL* and *YAB3* showed different expression patterns across fruits of *Erucaria*, but neither were detected in *Cakile*. It is important to note that plants of double mutants' *fil/yab3* in *Arabidopsis* have fruits that are remarkably reminiscent of heteroarthrocarpy, e.g., they lack valve margin in the distal region of the fruit while maintaining ovary wall identity. In contrast to heteroarthrocarpy, these mutants have

ectopic valve margin in the proximal region of their fruits (Dinneny et al., 2005). As these genes exhibit different patterns across *Cakile* and *Erucaria* and are expressed in both proximal and distal regions of *Erucaria*, heteroarthrocarpy cannot be explained by a simple lack of expression of these key regulators. Further, *FIL/YAB* are absent in the joint region of *Erucaria* (Figure 2.6), which is confounding since the joint contains small portions of both proximal and distal regions, an unavoidable consequence of segmentation during tissue collection. Nonetheless, deviation in expression patterns of these upstream regulators between *Arabidopsis* fruits and heteroarthrocarpic fruits implicates variation in their expression profiles in the origin of heteroarthrocarpy.

When exploring heteroarthrocarpy, we need to consider fruit patterning beyond the basal-apical differences that distinguish distal, joint, and proximal regions. That is, the medial (replum) patterning (Figure 2.1) is maintained in heteroarthrocarpic fruits whereas the lateral is not. This pattern is due to differences in dehiscence between proximal and distal segments; undifferentiated ovary wall is present in the distal region whereas valve or remnant valve is present in the proximal region. In other words, replum tissue is present in distal, joint, and proximal regions of heteroarthrocarpic fruits regardless of whether the ovary wall has differentiated into valve.(Roeder and Yanofsky, 2006). *FIL/YAB3* and *JAG* function antagonistically with replum promoting gene, *WUSCHEL RELATED HOMEOBOX 13* (*WOX13*), which positively regulates *RPL* in turn. This interaction is necessary for proper medial-lateral formation of *Arabidopsis* fruits. Further, *ASYMMETRIC LEAVES1* (*AS1*) and *AS2* collaborate with *JAG/FIL* function as promoters of lateral factors (González-Reig et al., 2012). The loss of both *AS1/2* and *JAG/FIL* in *Arabidopsis* results in dramatic medial-lateral differences and a substantially enlarged replum, which is more pronounced in the basal portion of the fruit (Alonso-Cantabrana et al., 2007; González-Reig et al., 2012). As *AS1/2* and *AS1* are expressed throughout *Cakile* and *Erucaria* regions, respectively, this pattern suggests that *AS1* alone is sufficient for proper replum (aka medial-lateral) formation in heteroarthrocarpic fruits. In other words, the collaboration between *JAG/FIL* function and *AS1/2* is not maintained in heteroarthrocarpic fruits. Further, *JAG/FIL* activity is non-detectable in the entire fruit, at least in *Cakile* at later stages of development. Thus, it appears that some redundancy in lateral-

medial patterning of *Arabidopsis* fruits has been lost in heteroarthrocarpic fruits, as supported by the different clustering of VM genes for each species (Figure 2.6), while simultaneously gaining apical-basal differences, e.g., dehiscence and indehiscence in the proximal and distal regions of *Erucaria*.

2.4.4 Valve margin pathway recruitment involved in abscission of the Cakile joint

The fruit of *Cakile* is distinct in that the joint abscises (disarticulates) at maturity. The joint, which represents the distal portion of the valve margin, thus represents a novel abscission zone in *Cakile*, completely separating the distal portion of the fruit. This is an unusual feature of certain heteroarthrocarpic subtypes, as there is no equivalent abscission zone in *Arabidopsis* fruits. Our data implicate the recruitment of downstream valve margin genes as responsible for joint abscission, although how that zone is positioned remains elusive.

IND is significantly upregulated in the joint region (Figure 2.6) and is primarily responsible for formation of separation and lignification layers in typical siliques (Liljegren et al., 2004; Groszmann et al., 2011), a juxtaposition of cell types also observed in the abscising joint region. Its presence in the joint may be due to a co-option of downstream valve margin pathway genes to facilitate formation of the joint abscission zone. Similar co-option is observed in seed abscission zones, although these zones typically involve *SEEDSTICK* (*STK*) in lieu of *SHP*, and the functionally similar transcription factor *HEC3* in lieu of *IND* (Balanzà et al., 2016). *SHP1/2* and *ALC* expression are both consistent with this co-option, as they are expressed in all three regions (Figure 2.6). Additionally, *SPT* expression is consistent with expression of *IND*, as expected from its downstream role in valve margin formation (Figure 2.6) (Girin et al., 2011). Further, both representative transcripts are among the 21 unique orthologous clusters in the joint of *Cakile* (Figure 2.4). This pattern is consistent with in situ hybridization data that showed *SHP2* expressed in septum and ovules of *Cakile*, and in ovules of *Erucaria* (Avino et al., 2012). Thus, the likely function of *SHP1/2* and *ALC* in the joint region would be to promote expression of *IND* (*SHP1/2*), and the formation of the separation layer (*ALC*).

What is unusual about joint abscission is that for the joint to separate, the distal and proximal regions of the replum must also separate. This expression pattern then implies that the

mechanism used to physically separate valve from replum may also be in play for replum in the joint region. Taken together with anatomical studies, our data suggest that there is a repurposing of the valve margin pathway in an otherwise indehiscent *Cakile* fruit, and that this pathway may be capable of initializing disarticulation in multiple tissue types.

2.4.5 Conclusion

Transcriptomic expression from late stage *Erucaria* and *Cakile* fruits is consistent with some conservation and some deviation of the valve margin pathway, specifically in upstream regulation, e.g., *FIL/YAB3* and *JAG*. Thus, different upstream regulators are implicated in the loss of dehiscence in Brassiceae relative to *Lepidium*, where *AP2* is likely responsible (Mühlhausen et al., 2013). Loss of expression of *FIL/YAB3* and *JAG* in *Arabidopsis* results in differing apical and basal phenotypes, which may help to explain the apical/basal differences in heteroarthrocarpic fruits (Dinneny et al., 2005). Further, heteroarthrocarpic fruits likely recruit the same mechanism used in valve and seed abscission for joint abscission (Figure 2.6). Functional tests are necessary to confirm whether redeployment of *FIL/YAB3*, *IND*, and possibly *SPT* have key roles in the origin of heteroarthrocarpy as well as joint abscission.

There have been multiple whole genome duplications in the Brassicales, which has resulted in many polyploids within the Brassicaceae family (Barker et al., 2010; Cardinal-McTeague et al., 2016; Edger et al., 2018). We considered the possibility of transcriptional differences between gene copies in distal, joint, and proximal regions that were undetected because we were unable to determine copy number in our transcriptome. For example, there are four copies of *FUL* in the Brassiceae (Brock and Hall, 2019), but each potential *FUL* copy had multiple hits from the same transcripts in both transcriptomes, so there is no definitive answer about copy number and expression (Figure 2.6). That is, we could not confirm or refute subfunctionalization of some fruit patterning genes as having a role in the origin of heteroarthrocarpy. An analysis of multiple transcripts for every fruit patterning gene showed generally similar expression for each, but further analyses are needed to determine if neo/subfunctionalization plays a role in heteroarthrocarpy.

Understanding the nature of heteroarthrocarpy, and how it relates to fruit development in *Arabidopsis*, will facilitate future studies on seed shattering in important Brassicaceous crops, and pernicious heteroarthrocarpic weeds. Lastly, these studies inform on the origin of important variation in seed packaging and dispersal capabilities.

2.5 Figures
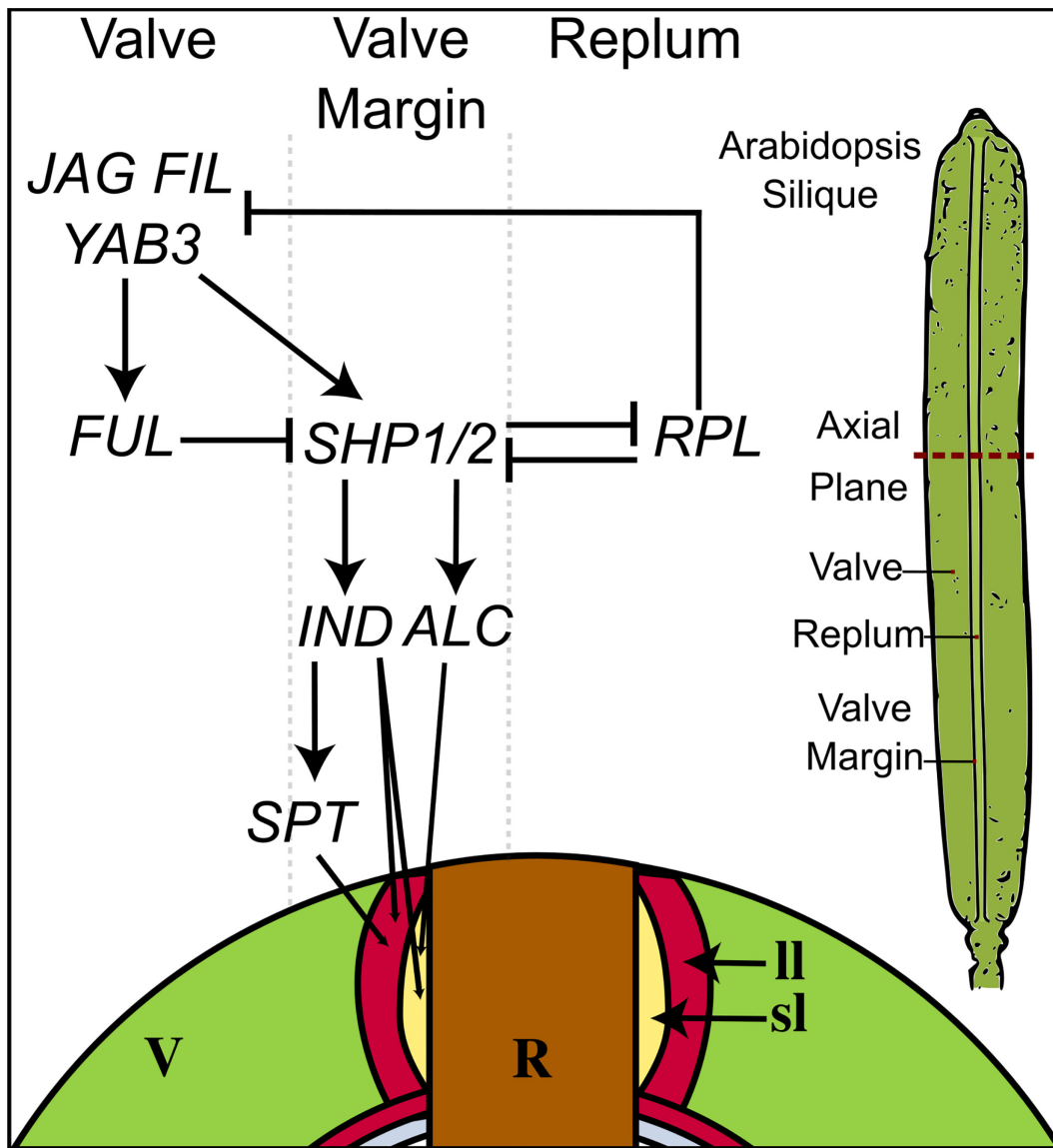
Figure 2.1. Diagram of simplified valve margin pathway for fruit dehiscence in *Arabidopsis thaliana*. V = valve R = replum. Sl = separation layer. ll = lignification layer. Valve margin = sl + ll.

Figure 2.2. Mature and young heteroarthrocarpic fruits. (A), Mature *Erucaria erucarioides* fruit in lateral view before dehiscence (left), and medial view after dehiscence (right). (B), Young *E. erucarioides* fruit representing size sampled for transcriptomics in medial view. (C), *Cakile lanceolata* fruit in lateral view before joint abscission (left), and medial view after joint abscission (right). (D), Young *C. lanceolata* fruit representing size sampled for transcriptomics in medial view. White arrows indicate joint region and blue arrows indicate replum.
Scale bars = 5 mm.
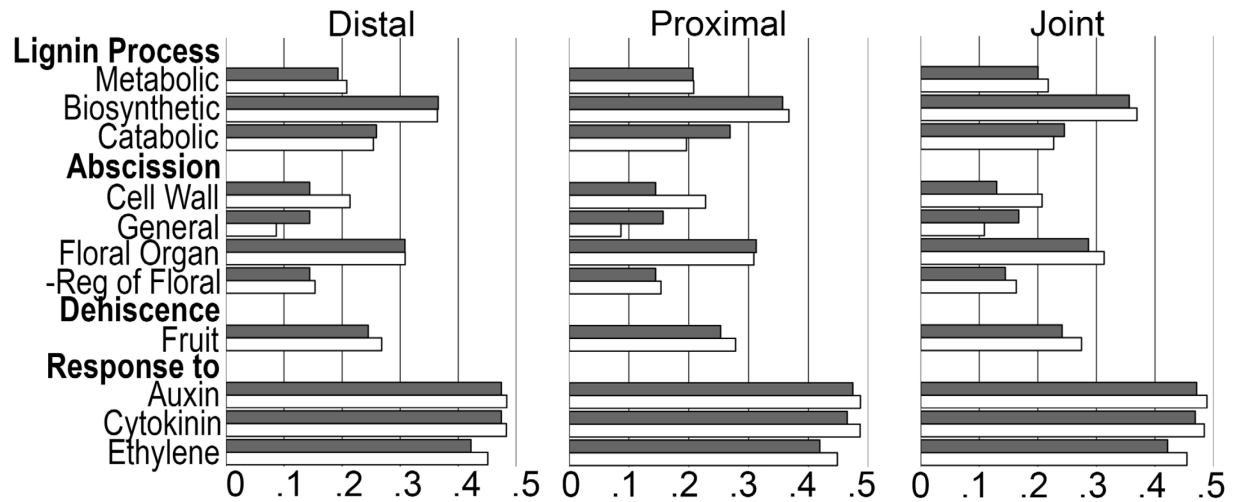
Figure 2.3. Graph of select Gene Ontology (GO) terms for *Erucaria erucarioides* (grey) and *Cakile lanceolata* (white). GO counts based on merged profile of three biological replicates per region. Sample (n) and total (N) raw counts log2 transformed for interspecies comparison. GO terms chosen based on search terms: lignin, abscission, dehiscence, and response to hormone.
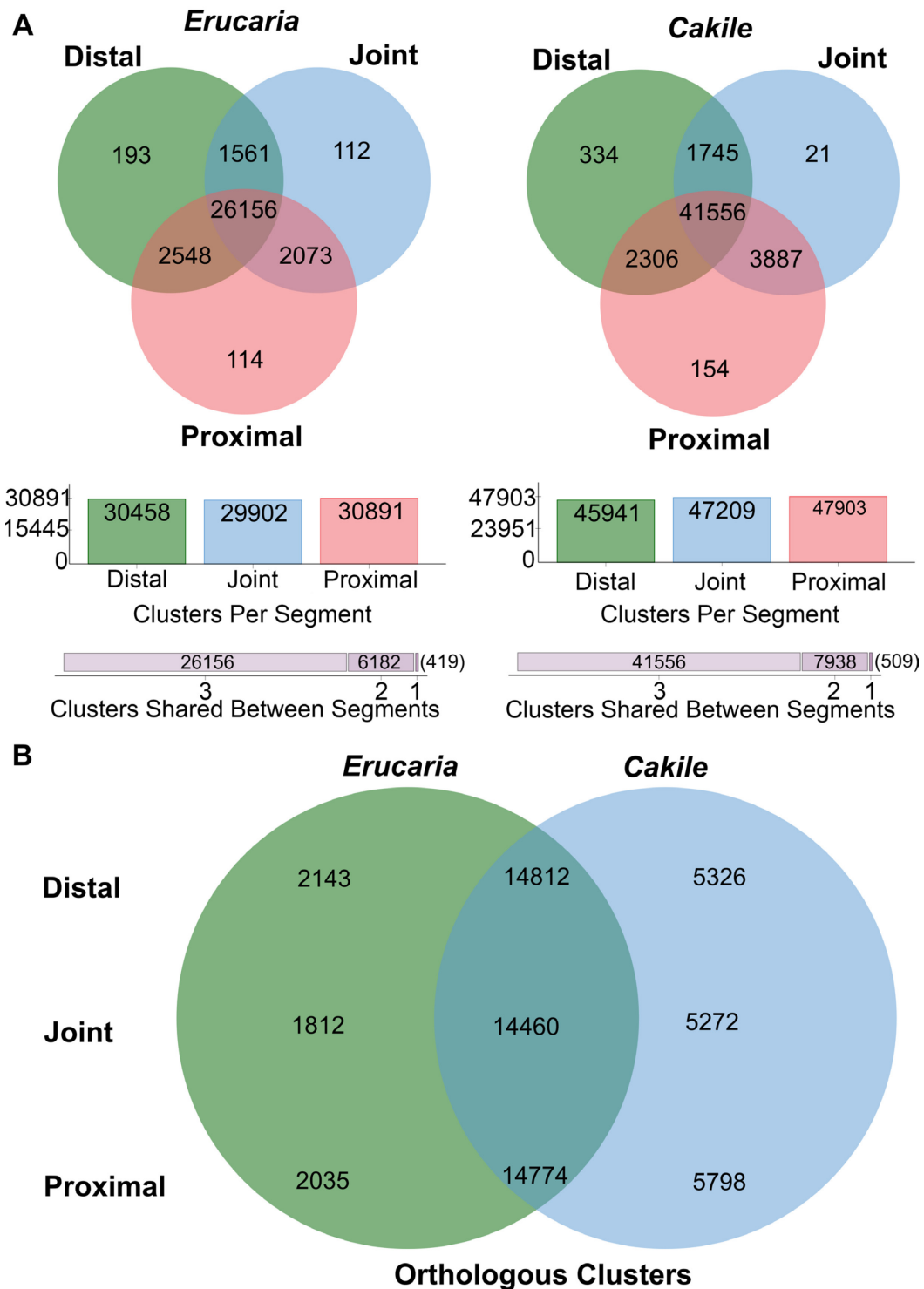
Fig 2.4. Venn diagrams of three-way and pairwise High Throughput Sequencing (HTS) filtered transcripts for *Erucaria erucarioides* and *Cakile lanceolata* transcriptomes. (A), Three-way Venn diagrams of *Erucaria* and *Cakile* orthologous clusters for distal, joint, and proximal regions. (B), Pairwise Venn diagrams of *Erucaria* and *Cakile* orthologue-clustered transcripts.
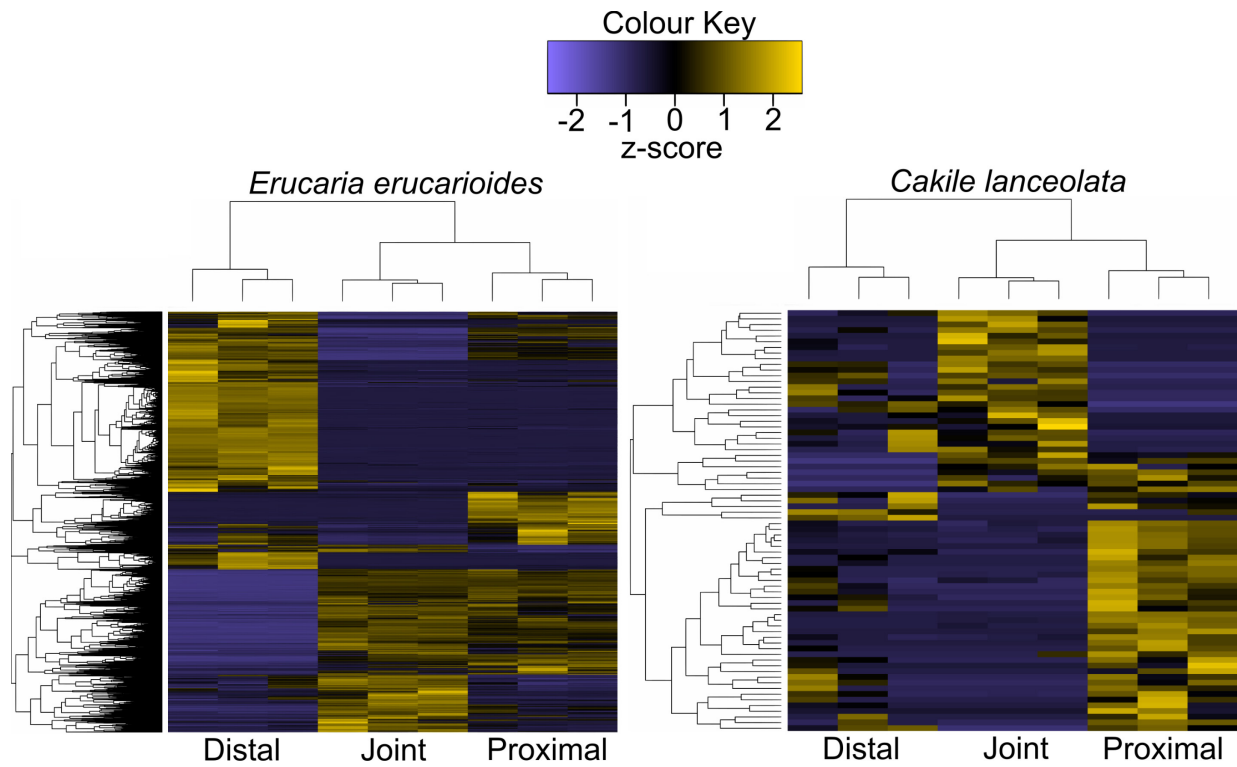
Figure 2.5. Heatmap of all significant contig clustered transcripts in the *Erucaria erucarioides* (n = 15,345) and *Cakile lanceolata* (n = 74) transcriptomes, expressed as z-scores (FDR-corrected α = 0.01).

Figure 2.6. Heatmap of contig clustered transcripts from *Erucaria erucarioides* and *Cakile lanceolata* expressed in log2 (TPM) with TMM normalization. Representative transcripts are those with the highest bitscore hits against the TAIR10 database. Bolding indicates a shared orthogroup with other transcriptome. FULa,b,c,d are copies of FUL that are present in some species across the Brassicaceae. TPM = Transcripts Per Million. TMM = Trimmed Mean of M-values. Asterisks indicate significant differential expression between proximal and joint region. (FDR-corrected α = 0.01).

Figure S2.1. Graph of top Gene Ontology (GO) terms for *Erucaria erucarioides* (blue) and *Cakile lanceolata* (white). Sample (n) and total (N) raw counts were log2 transformed for interspecies comparison.

2.6 Tables

Table 2.1. Statistics for de novo Trinity assembly of *Erucaria erucarioides* and *Cakile lanceolata* pairwise reads for all isoforms. Numbers in parentheses refer to longest isoform only.

	Erucaria	*Cakile*
N50	1544(1017)	1464(835)
Median Contig Length	578(374)	517(330)
Average Conti Length	942.83(656.94)	877.15(577.55)
Total Assembled Bases	214,521,562(92,098,767)	275,595,508(108,815,069)
Total Trinity Genes	140194	184945
Total Trinity Transcripts	227530	314194
GC%	41.89	42.05

Table S2.1. Benchmarking Universal Single Copy Ortholog (BUSCO) analysis of *Erucaria erucarioides* and *Cakile* transcriptomes.

	Erucaria	*Cakile*
Complete BUSCOs	1388 (96.4%)	1366 (94.8%)
Complete/ single-copy BUSCOs	487	532
Complete/ duplicated BUSCOs	901	834
Fragmented BUSCOs	33 (2.3%)	43(3.0%)
Missing BUSCOs	19 (1.3%)	31(2.2%)
Total BUSCO groups searched	1440	1440

Chapter 3: Virus-induced gene silencing as a tool for functional studies in *Cleome violacea*[i]

3.1 Introduction

Cleomaceae is a pantropical plant family comprising approximately 270 morphologically diverse species (Patchell et al., 2014), and despite having only a few economically important taxa, it is emerging as a model clade to investigate many areas of research (reviewed in: Bayat et al., 2018). These topics include floral morphology and development (Nozzolillo et al., 2010; Patchell et al., 2011; Endress, 1992), pollination biology (Cane, 2008; Higuera-Díaz et al., 2015), comparative genomics/transcriptomics (Schranz and Mitchell-Olds, 2006; Barker et al., 2009; Bräutigam et al., 2011, 2011; Cheng et al., 2013; Bhide et al., 2014; Kulahoglu et al., 2014), evolution of C4 photosynthesis (Marshall et al., 2007; Voznesenskaya et al., 2007; Bräutigam et al., 2011; Koteyeva et al., 2011, 2014; Williams et al., 2015), and glucosinolates (Bergh et al., 2016). An attractive feature of this family is its sister relationship to the Brassicaceae, which houses the model species *Arabidopsis thaliana* (L.) Heynh. This close phylogenetic distance facilitates transfer of knowledge from *Arabidopsis* while permitting the investigation of morphological and physiological traits not found in Brassicaceae. These traits include substantive floral diversity and C4 photosynthesis (Bayat et al., 2018). Furthermore, research of these traits is facilitated by steadily growing –omics data within Cleomaceae, including published genomes of *Tarenaya hassleriana* (Chodat) Iltis (Cheng et al., 2013) and *Cleome violacea* (L.) (Accession no. 23822) (https://genomevolution.org/coge), as well as transcriptomic data from flowers and leaves of *T. hassleriana* and *Gynandropsis gynandra* (L.) Briq.(Bräutigam et al., 2011; Bhide et al., 2014; Kulahoglu et al., 2014). In summary, Cleomaceae is a promising group to investigate important ecological and evolutionary phenomena.

Reverse genetics is key to understanding the mechanisms underlying morphological diversity in Cleomaceae because it directly tests the role(s) of candidate genes generating that diversity.

---

[i] A version of this chapter has been published (Carey et al., 2021). Figures and tables labelled as supplemental are due to journal submission requirements and are not reflective of their contribution to the main thesis.

Stable transformation techniques have been developed for both *G. gynandra* and *Tarenaya spinosa* (Jacq.) Raf. (Newell et al., 2010; Tsai et al., 2012), but they are time-consuming, in part because they require cell culture. To the best of our knowledge, only one study has used stable transformation in Cleomaceae to empirically examine mechanisms regulating C_4 photosynthesis in *G. gynandra* (Williams et al., 2015). An effective alternative to stable transformation is virus-induced gene silencing (VIGS), which is a well-established post-transcriptional gene silencing (PTGS) tool used in many model and non-model species (Burch-Smith et al., 2004, 2006; Gould and Kramer, 2007; Becker and Lange, 2010; di Stilio et al., 2010; Cheng et al., 2013). VIGS uses RNA interference to exploit plant defenses, which greatly reduces endogenous mRNA transcription without the need for stable transformation over multiple generations (Ruiz et al., 1998; Baulcombe, 1999; Burch-Smith et al., 2004). Thus, VIGS provides the opportunity to quickly downregulate individual or multiple genes and infer function via the observation of modified phenotype (Becker and Lange, 2010; Cheng et al., 2013). Notwithstanding the potential of VIGS for examining gene function in a range of species, protocols must be tailored for each new species (Senthil-Kumar et al., 2007). Tailoring of protocols can be challenging because of factors like developmental timing, susceptibility to *Agrobacterium tumefaciens* infection, and strength of PTGS response. Furthermore, not all species are amenable to reverse genetic approaches like VIGS.

Cleome violacea is a species of note for investigations within Cleomaceae (Mabry et al., 2020) and an ideal candidate for VIGS. It has a published genome and exhibits morphological traits that are important for pollinator interactions, e.g., floral monosymmetry and prominent nectaries (Figure 3.1). Furthermore, *C. violacea* displays features that facilitate its study, including small stature, high fecundity, self-fertilization, and short generation time. Here, we report on testing and optimizing the VIGS methodology in *C. violacea*. Specifically, the goals of this study are threefold: (1) establish that constructs with heterologous sequences are an informative first screen to determine efficacy of VIGS in *C. violacea*, (2) optimize the VIGS protocol for *C. violacea*, and (3) demonstrate that this protocol is effective for multiple developmental stages. We selected two genes that would maximize comparison of *C. violacea* to other VIGS models and that would adequately test different developmental processes. First,

PHYTOENE DESATURASE (*PDS*) is commonly used to determine the efficacy of VIGS in many systems due to its easy-to-score bleaching phenotype when downregulated (Gould and Kramer, 2007; Senthil-Kumar et al., 2007; Wege et al., 2007; Velasquez et al., 2009; Fujita et al., 2019). Second, the MADS-box transcription factor *FRUITFULL* (*FUL*) has an established role in fruit development in *Arabidopsis* (Gu et al., 1998; Ferrándiz et al., 2000b; Balanzà et al., 2014; Eldridge et al., 2016), a species with similar fruit morphology to *C. violacea* (Figure 3.2). Furthermore, fruit development is the last stage in reproduction and, as such, successful downregulation of *FUL* would demonstrate the efficacy of VIGS across multiple developmental stages in *C. violacea*.

3.2 Materials and Methods

3.2.1 Plant growth conditions

Inbred lines of *C. violacea* were grown from lab seed stock, and a voucher was deposited in the Vascular Plant Herbarium (ALTA) at the University of Alberta (Hall & Bolton s.n., 20 Feb 2008; #813 from *Hortus Botanicus*, Amsterdam). Seeds were sown in groups of 10 using 10-cm-wide square pots, and individual plants were transferred after treatments to 7.5-cm-diameter round pots. Growth substrate was a 2:1 mixture of sterilized (liquid cycle, 121.1°C, 20 min) Sun Gro Sunshine Mix (Sun Gro Horticulture, Agawam, Massachusetts, USA) and perlite. All plants were grown in a growth chamber at the Department of Biological Sciences at the University of Alberta, with 16 h of light and 8 h of darkness at 24°C. This temperature was chosen for *C. violacea* because it falls within the optimal range for VIGS silencing reported in *Arabidopsis* (22–24°C) (Wang et al., 2006) while maintaining fast and efficient plant growth. However, it is important to note that *C. violacea* grows well under a range of growth conditions (e.g., small to large pots, temperatures between 20–26°C, and growth medium without perlite).

3.2.2 Cloning and viral vector construct design

Viral vector construct design and inoculation were carried out in accordance with previous VIGS protocols (Gould and Kramer, 2007). Tobacco rattle virus vectors pTRV1 (donor stock no. YL192), pTRV2 (donor stock no. YL156), and pTRV2-*AtPDS* (donor stock no. YL154) were

obtained from The *Arabidopsis* Information Resource (TAIR; https://www.arabidopsis.org) using their stock center (*Arabidopsis* Biological Resource Center [ABRC], Ohio State University, Columbus, Ohio, USA; https://abrc.osu.edu). The pTRV1 vector is required for viral movement (Ziegler-Graff et al., 1991), and the pTRV2 vector contains the gene of interest inserted into the multiple cloning site (MCS) (Ratcliff et al., 2001). Three endogenous constructs were also generated using *C. violacea* mRNA: pTRV2-*CvPDS*, pTRV2-*CvFUL*, and a pTRV2-*CvPDS-CvFUL* hybrid (Table S3.1).

RNA was extracted from *C. violacea* leaves, and separately from whole flowers, using a Concert Plant RNA Reagent Kit (Invitrogen, Carlsbad, California, USA), and was then treated with DNase I (Fermentas, Hanover, Maryland, USA) and enriched using an RNeasy Mini Kit (QIAGEN, Germantown, Maryland, USA). The mRNA was isolated using a Dynabeads mRNA Purification Kit (Invitrogen). The cDNA was synthesized using SuperScript III Reverse Transcriptase (Invitrogen), poly(T) primers, and random hexamer primers, following manufacturer instructions. *Arabidopsis PDS3* degenerate primers (Table S3.1) were used to amplify a 953 bp fragment from leaf cDNA. This fragment was then cloned into a TOPO-TA plasmid vector (Invitrogen) following the manufacturer instructions, and sequenced using BigDye Terminator sequencing (Applied Biosystems, Foster City, California, USA) with vector-specific primers M13F and M13R (Applied Biosystems). From that sequence, new primers were designed to amplify a 441 bp fragment with XbaI and BamHI restriction sites on the 5' and 3' fragment ends, respectively (Table S3.1). Alternatively, *CvFUL* primers with 5' XbaI and 3' BamHI restriction enzymes were designed using an unpublished transcriptome (Table S3.1); the primers amplify from the fifth exon to the eighth exon of the coding sequence, which avoids silencing other MADS-box genes with a conserved domain sequence. A 272 bp fragment of the *CvFUL* coding sequence was amplified using those primers and floral cDNA (Table S3.1). A *CvFUL-CvPDS* construct was also made using modified primers of *CvPDS* with a 5' BamHI and 3' XhoI restriction site. This modification enabled the ligation of a *CvPDS* construct to the 3' end of *CvFUL* (Table S3.1). Restriction-modified fragments of *CvPDS*, *CvFUL*, and *CvFUL-CvPDS* were first cloned into TOPO-TA plasmid, and then excised and ligated into separately digested TRV2

vectors using XbaI, BamHI, and XhoI restriction enzymes and T4 DNA ligase (Promega Corporation, Fitchburg, Wisconsin, USA).

Viral constructs were verified for off-target silencing using siFi21 (Version 1.2.3-0008) (Lück et al., 2019). This software predicts off-targets of RNA interference (RNAi) by comparing user-defined sequences to a local database with Bowtie, a short-read aligner (Langmead et al., 2009; Langmead, 2010). Because Bowtie is purpose built for aligning to a reference, we used the *C. violacea* genome for tracking off-target silencing. One notable limitation is that Bowtie does not distinguish between introns and exons, thus any mRNA sequence that spans two exons may not be recognized. However, because the small interfering RNAs (siRNAs) tested are 21 bp in length, most targets will still be identified. For clarification, the average exon length of the closely related *Arabidopsis* is 236.8 bp (Koralewski and Krutovsky, 2011). Assuming *C. violacea* is similar, 21 bp siRNAs will produce many hits within each exon. We defined on-target sequences by using the BLAST (Altschul et al., 1990) function on https://genomevolution.org/coge. Coding sequences of *AtPDS* and *AtFUL* were matched using BLASTn against the *C. violacea* genome, and hits with the top *E*-values/bit-scores were defined as on-target. TRV2-*AtPDS*, TRV2-*CvPDS*, and TRV2-*CvFUL* were also aligned to the *C. violacea* genome as another way of verifying their similarity to *AtPDS* and *AtFUL*. The *C. violacea* genome is divided into scaffolds with sizes that range between 904 bp and 3.8 Mbp, so any sequences aligning to similar positions on the same scaffold are likely related.

TRV2 vectors were then cloned into One Shot TOP10 chemically competent *Escherichia coli* (Invitrogen) according to the manufacturer instructions. These *E*. *coli* were plated onto lysogeny broth (LB) agar plates containing 50 μg/mL kanamycin and grown overnight at 37°C. Colonies were screened via PCR with primers, 156F and 156R (Table S3.1), that span the TRV2 MCS (Gould and Kramer, 2007). Positively screened transformants were grown in LB media with the same antibiotics overnight at 37°C, and their plasmid constructs were purified using a QIAGEN Miniprep kit. Construct identity was confirmed using BigDye Terminator sequencing and 156F/ 156R primers. All sequences were obtained with an ABI-3730 DNA Analyzer (Applied Biosystems) after being cleaned with a Performa DTR V3 96-well Short Plate Kit (Edge

BioSystems, Gaithersburg, Maryland, USA). Nucleotide sequence of the *CvPDS* construct was deposited under the accession MW505002 in the National Center for Biotechnology Information (NCBI) GenBank, and a partial coding sequence of *CvFUL* is available under the accession MK584560.

3.2.3 Agrobacterium tumefaciens preparation and infiltration

Agrobacterium tumefaciens was prepared for DNA transformation as previously described (Weigel and Glazebrook, 2002). The following vectors were transformed into *A. tumefaciens* using calcium chloride heat-shock transformation: pTRV2-*AtPDS*, pTRV2-*CvPDS*, pTRV2-*CvFUL*, pTRV2-*CvPDS-CvFUL*, pTRV2-MCS, and pTRV1. For each, 100 ng of purified construct was combined with 250 μL of competent *A. tumefaciens*. Transformants were plated on LB media containing 50 μg/mL kanamycin, 50 μg/mL gentamycin, and 25 μg/mL rifampicin; all LB media used to grow *A. tumefaciens* contained these antibiotics. Transformants were then screened as before using 156F and 156R primers (Table 3.1), and glycerol stocks were made and stored at −80°C (1:1 ratio of 50% glycerol and overnight *A. tumefaciens* culture).

Vacuum infiltration was the chosen method for *C. violacea* because of its simplicity and efficiency compared with other methods (Feng et al., 2006; Becker and Lange, 2010). Alternative methods, such as agrodrench and booster inoculation (Ryu et al., 2004; Senthil-Kumar and Mysore, 2014) were briefly explored, but were abandoned due to low silencing-efficacy. The first vacuum infiltration experiments were conducted to examine mortality and silencing-efficacy in plants at different developmental stages with heterologous and endogenous constructs. First, plants were categorized based on the number of true leaves: small with 0–3 true leaves, medium with 4–6 true leaves, and large with seven or more true leaves. To prepare for infiltration, *A. tumefaciens* containing each prepared vector was streaked onto LB agar with antibiotics and incubated for 72 h at room temperature. Individual colonies were used to inoculate 5 mL of LB with antibiotics, which were grown overnight at room temperature with constant shaking of 275 rpm to an OD600 of approximately 1.0. Sequential inoculations using the same media and conditions were then used to scale up growth to a final working volume of 500 mL. That is, 5 mL cultures were used to inoculate 50 mL of LB, which

were used as starter cultures for 500 mL. Starter cultures one-tenth of the total volume produce sufficient growth in ~24 h. This process can be scaled up or down depending on the desired number of treated seedlings, and we recommend 500 mL for treatment groups with 50 to 100 seedlings. The final cultures contained: antibiotics, 1 mM MES buffer, and 0.02 mM acetosyringone. For each pTRV2-containing *A. tumefaciens* culture, one pTRV2-containing *A. tumefaciens* culture must be grown because a 1:1 mixture is required for infiltration. Each culture was then aliquoted into 50 mL conical centrifuge tubes and centrifuged at 3200 *g* for 20 min at 4°C. The supernatant was decanted from each tube, and cells were resuspended in 20 mL of infiltration buffer (10 mM MES, 10 mM $MgCl_2$, and 0.2 mM acetosyringone). Resuspended colonies were recombined in a sterile beaker, and infiltration buffer was added to an OD600 of 2.0 ± 0.1. Colonies were then left at room temperature for a minimum of 4 h to acclimatize. It should be noted that it is also viable to leave cultures overnight if conditions are sterile.

The pTRV2 *A. tumefaciens* resuspensions were combined 1:1 with pTRV1 *A. tumefaciens* resuspensions. Silwet L-77 surfactant was added to each mixture at 100 µL/L. Groups of seedlings were extracted from the soil, rinsed in distilled water, and briefly air-dried before being submerged in the resuspensions and placed in a vacuum chamber. The vacuum chamber was evacuated for 10 sec, or until the mixture bubbled, and then the vacuum was turned off and seedlings were left for 2 min before pressure was quickly released. The resuspension mixtures were as follows: pTRV1 + pTRV2-*CvPDS*, pTRV1+ pTRV2-*AtPDS*, pTRV1 + pTRV2-*CvFUL*, pTRV1 + pTRV2-*CvFUL-CvPDS*, and pTRV1 + pTRV2-MCS as a negative control. Ten seedlings were infiltrated each time, and the same mixtures were reused for repeated infiltrations (i.e., 10 more seedlings up to 100 seedlings per 500 mL). Additionally, a group of seedlings were extracted from the soil, rinsed, and air-dried, and then replanted as an untreated control.

Phenotypic scoring varied depending on the construct. Phenotypes for both *CvPDS-* and *AtPDS-* downregulated plants became apparent between three and four weeks post-inoculation. Plants treated with pTRV2-*CvPDS* and pTRV2-*AtPDS* had their leaves scored using visual approximation of bleaching and were binned into the following categories: pale (<50% of total bleaching),

variegated (50–80%), and strong (>80%). Plants treated with pTRV2-*CvFUL* and pTRV2-*CvFUL-CvPDS* were scored approximately 30 days after vacuum infiltration. Phenotypes were initially binned into three categories based on the severity of fruit curling and the indentation of valve tissue around the ovules: (1) indentation only, (2) mild curling with indentation, and (3) heavy curling with indentation. These data were used to compare the effects of *CvFUL* silencing to *FUL* knockout in *Arabidopsis* (Gu et al., 1998), although less severe phenotypes were expected because of incomplete mRNA silencing inherent to VIGS.

Leaf and fruit tissue from treated and control plants were excised, flash-frozen in liquid nitrogen, and stored at −80°C. Following described protocols (Patchell et al., 2011), scanning electron microscopy (SEM) was used to examine both untreated *C. violacea* and pTRV2-*CvFUL-CvPDS*-treated fruits. Leaf and fruit phenotypes were imaged using a Nikon SMZ 1500 dissecting microscope (Nikon, Tokyo, Japan) with a QImaging Retiga 4000R camera (Teledyne DALSA, Waterloo, Ontario, Canada), and handheld digital Canon DS126181 camera (Canon, Tokyo, Japan). Images were standardized, scaled, color balanced, and assembled into figures using Inkscape version 0.92.5 (https://inkscape.org) and GIMP version 2.10.18 (https://www.gimp.org).

3.2.4 RT-qPCR expression analysis

RNA was extracted from approximately 100 mg of treated and control leaf/fruit tissue using an RNeasy Plant Mini Kit (ID:74904) (QIAGEN). The following amendments were made to the Quick-Start protocol: Step 1a, tissue was disrupted with a miniature pestle directly in a 1.5 mL Eppendorf tube. Step 1b was skipped. In Step 2, RLC buffer was used in lieu of RLT buffer for fruit tissue, and tubes were left to incubate at room temperature for 6 min with 1 min of vortexing at 2500 rpm. Step 6 was replaced with steps 4 and 5 from the QIAGEN Micro Kit Quick-Start Protocol to incorporate the column-based DNase I treatment. Step 10 was incubated for 5 min, and eluate was run through the same column twice. All steps with 15 sec centrifuge directions were increased to 30 sec due to centrifuge timer limitations. RNA was quantified and checked for integrity using the NanoDrop ND-1000 (version 3.1; Thermo Fisher Scientific, Waltham, Massachusetts, USA) and Bioanalyzer 2100 (version B.02.09.SI720; Agilent,

Santa Clara, California, USA) from the Molecular Biology Service Unit, University of Alberta, Canada. The cDNA for each sample was generated as previously described using 1000 ng of RNA.

Quantitative reverse transcriptase PCR (RT-qPCR) was used to compare relative transcript levels between treated and untreated controls using the delta-delta Ct method. All experiments used *ACTIN* as a reference gene. Primers were designed using Primer Express 3.0 (Applied Biosystems), ensuring at least one primer was outside of the construct (Table S3.1). Primers were also tested for optimal efficiency using a dilution series. RT-qPCR was run using 10 μL reaction volumes containing 5.0 μL SYBR Green master mix (0.25× SYBR Green, 0.1× ROX, 0.3 units Platinum *Taq* Polymerase [Invitrogen], and 0.2 mM dNTPs), 2.5 μL of 3.4 μM forward and reverse primer, and 20 ng of cDNA. Assays were run using both the 384-well QuantStudio 6 Real-Time PCR System and the 7500 Fast Real-Time PCR System (both from Applied Biosystems). Three technical replicates per sample were analyzed with a minimum of three biological replicates (range of 3 to 5). One individual plant is equivalent to one biological replicate. VIGS-treated plants will often display varying phenotypes on one plant, and in such cases only one phenotype was used for RT-qPCR.

3.3 Results

3.3.1 No off-target silencing was detected for endogenous pTRV2 constructs

The endogenous constructs had no off-target silencing and were on-target for our genes of interest (Table S3.1). The coding sequence of *AtPDS* had top hits to scaffold 38 (650,987 bp), which was similar to both endogenous and heterologous constructs. The coding sequence of *AtFUL* had a top hit with scaffold 1565 (19,090 bp), similar to pTRV2-*CvFUL*, which only had a single hit. No off-targets were found for pTRV2-*CvPDS*, pTRV2-*CvFUL*, or pTRV2-*CvPDS-CvFUL*, and only minimal off-target silencing was detected for pTRV2-*AtPDS* (Figure S3.1). pTRV2-*AtPDS* had 84 total and 36 efficient hits to scaffold 38. pTRV2-*AtPDS* also had six off-target hits to scaffold 283 (212,661 bp), although none were efficient. The efficient hits are matched on a more stringent set of parameters and can be considered more representative of true off-targets

(Lück et al., 2019). In summary, there were no major off-target sites for any of the constructs, and all constructs appear to target our genes of interest, *PDS* and *FUL*.

*3.3.2 pTRV2-*AtPDS *and pTRV2-*CvPDS *infection induces silencing of gene targets in* Cleome violacea

An effective VIGS system entails low mortality and high silencing-efficacy of transient mutant phenotypes. We explored mortality and silencing-efficacy in *C. violacea* using binary vectors pTRV1 + pTRV2-*CvPDS* and pTRV1 + pTRV2-*AtPDS*. Initial experiments tested both heterologous (pTRV2-*AtPDS*) and endogenous (pTRV2-*CvPDS*) constructs to determine how susceptible *C. violacea* is to VIGS (Figure 3.3). Treatment with pTRV2-*AtPDS* revealed that *C. violacea* was amenable to VIGS, as most surviving plants exhibited photobleaching (Table 3.1, Figure. 3.3A-E). Of the 198 plants treated with pTRV2-*AtPDS*, 119 survived, with roughly half displaying photobleaching. Relative to pTRV2-*AtPDS*, plants treated with pTRV2-*CvPDS* had the highest silencing-efficacy (81%) despite a lower survival rate (Table 3.1, Figure.3.3F-J). The greatest silencing-efficacy (74%) was in medium plants (Table 3.1). That is, 37 of the 58 surviving plants that showed altered morphology were medium-sized seedlings (Table 3.1). Ultimately, medium plants were chosen for all subsequent experiments because of their stronger response to infiltration, i.e., stronger photobleaching. We also tested a species of Brassicaceae, *Erucaria erucarioides* Müll. Berol., with heterologous pTRV2-*AtPDS*, but no further experiments with *E. erucarioides* were completed due to low silencing-efficacy (Table S3.2). We mention these preliminary data to show that phylogenetic distance, and similarity of VIGS construct to the target gene, does not necessarily correlate with efficacy of VIGS (Figure S3.2).

Follow-up experiments tested whether the observed phenotypes were due to downregulation of *CvPDS*. Here, we only used TRV2-*CvPDS* because it had the highest overall silencing-efficacy from earlier trials. The pTRV2-*CvPDS* treatment group had 36 of 86 surviving plants display photobleaching, most with strong photobleaching (Table 3.2). On average, we observed photobleaching from three to four weeks after inoculation, with the altered phenotypes lasting for several weeks. Leaves showing 50–80% and >80% photobleaching had significant downregulation of *CvPDS* relative to controls and treated plants with unaltered phenotype

(Figure 3.4A). Although we scored and phenotyped leaves with pale bleaching (1–50%) (Figure S3.3A), they were not included in the RT-qPCR analysis because we are primarily interested in strong VIGS phenotypes, i.e., our goal was to demonstrate that *CvPDS* is downregulated when we see obvious photobleaching. Leaves with complete photobleaching were also scored (Figure S3.3A); however, they were binned with strong photobleaching (>80%) because of their infrequency. These results demonstrate that VIGS functions well in *C. violacea*, is experimentally replicable, and produces strong silencing with endogenous and heterologous constructs.

3.3.3 Agrobacterium-mediated pTRV2-CvFUL *infection disrupts fruit morphology in* Cleome violacea

Another factor in an effective VIGS system is the ability to study genes across developmental stages. As such, we investigated pTRV2-*CvFUL* treatment on the final stage of reproduction in flowering plants, fruit development. Fruits of *C. violacea* are elongate, bicarpellate capsules, which are characteristic for Cleomaceae. Like Brassicaceae, the fruits have valves, which are the portion of the ovary walls that separate at maturity (Figure 3.2A). Also, like Brassicaceae, fruits of *C. violacea* have a prominent persistent placenta called the replum. The primary difference between Cleomaceae capsule fruits and Brassicaceae siliques is that Cleomaceae fruits lack the false septum separating the two locules (Figure 3.2B) (Iltis et al., 2011). Like with *PDS*, initial experiments with pTRV2-*CvFUL-CvPDS* were completed to determine effectiveness in *C. violacea*. The *PDS* marker was used to facilitate phenotyping, as we were unsure of the silencing effects of *CvFUL* alone. Large seedlings were chosen for the initial experiments to account for the developmental timing of fruiting. A range of altered fruit morphologies were observed four to six weeks post-inoculation. Of the 60 large seedlings treated (≥7 true leaves), there was a 95% survival rate, and 36% of survivors displayed some degree of curling and/or indentation (Table 3.1, Figure 3.5A-D).

Later experiments used pTRV2-*CvFUL* alone, as initial experiments were relatively easy to score. Like with pTRV2-*CvFUL-CvPDS*, pTRV2-*CvFUL*-treated plants began displaying altered fruit morphologies by four to six weeks post-inoculation (Table 3.2, Figure 3.5E-H). Most

prominently, fruit curling and ovule indentation were observed, i.e., the valve was appressed around the developing seeds (Figure 3.5E-H, and S3.3H, I). In non-control treated plants of *C. violacea*, fruit length was decreased when fruit curling was moderate to heavy (≥90°); however, those with light curling (<90°) and/or indentation did not have a significant difference relative to fruits with non-altered phenotype (Figure S3.4). Future studies with *CvFUL* knockouts would be more informative because variability of PTGS response is likely a major factor in fruit length with VIGS. For seedlings treated with pTRV2-*CvFUL*, 37 of the 85 surviving plants showed some altered phenotype, mostly mild curvature (Table 3.2). Incomplete and broad variability of *CvFUL* silencing made it difficult to accurately bin phenotypes by severity (Figure 3.5E-H), despite initial efforts (Figure S3.3B-E). Because our primary goal was to assess the general viability of *CvFUL* silencing using VIGS, and thus the potential for other fruit-patterning genes, we pooled data from all curled fruits for RT-qPCR regardless of severity. This way, ambiguities from incomplete silencing and unknown effects would be reduced while still demonstrating that *CvFUL* had been successfully downregulated. *CvFUL* was significantly downregulated relative to untreated plants, pTRV2-MCS-treated plants, and normal-phenotype treated plants (Figure 3.4B). It should be noted that pTRV2-MCS had a greater than expected effect on fruit morphology (Figure S3.3C, G); this is due to viral vector response and not downregulation of *CvFUL* (Figure 3.4B). It seems likely that *C. violacea* fruits are more sensitive to tobacco rattle virus relative to other structures, which may account for the mild bulging in some fruits (Figure S3.3G). It would be informative for future studies to observe the effects of viral response at varying concentrations of *A. tumefaciens* inoculation.

The curling, indentation, and shortening of fruits in pTRV2-*CvFUL-CvPDS*-treated plants are accompanied by differences in both cell size and shape in the valves and repla of those fruits (Figure 3.6). Cells in the repla of pTRV2-*CvFUL-CvPDS* treated plants are less elongated and less uniformly dispersed than their untreated counterparts (Figure 3.6B,F). Furthermore, their ends are blunter and less tapered than in untreated plants. Cells in the valve regions also show marked differences between untreated and treated plants (Figure 3.6C,D,G,H). Specifically, the smaller valve cells of pTRV2-*CvFUL* treated plants are uneven and irregularly spaced (Figure 3.6H). Smaller valve cells in untreated plants are more elongated, with even edges, and are

somewhat regularly spaced (Figure 3.6D). Both treated and untreated fruits have larger bulbous cells in their valve regions (Figure 3.6C, G), although those from *CvFUL-CvPDS*-treated plants appear to be smaller, rounder, and less raised. Both untreated and treated fruits have glandular and multicellular trichomes, although there seem to be more on treated fruits (Figure 3.6C, G and Figure 3.3B,D). The apparent increase seems likely to be a secondary effect of VIGS, as trichomes are known to be involved in pathogen response (Wang et al., 2011). Overall, these results demonstrate that silencing of *CvFUL* in *C. violacea* produces modified phenotypes in both valve and replum tissue.

3.4 Discussion

VIGS is a well-established method for gene knockdown in angiosperms and has provided insight in determining the mechanisms underlying differences across diverse lineages such as Ranunculales, Asterids, Caryophyllales, and Rosids. It remains a powerful tool to address questions about organ identity, sex determination, induction of flowering, and compound leaf development (Hidalgo et al., 2012; Hsieh et al., 2013; Fujita et al., 2019). As such, the major goal of this study was to establish VIGS in Cleomaceae, which is garnering interest as a model family (Bayat et al., 2018). We established that *C. violacea* has high silencing efficacy and moderate mortality with VIGS treatment. Furthermore, *C. violacea* is amenable to treatment with readily available *Arabidopsis* constructs, which implies a potential to explore this method for additional members of Cleomaceae. Finally, we demonstrate that both leaves and fruits are susceptible to treatment, suggesting that other developmental stages and organs could also be studied with this system.

3.4.1 Arabidopsis *constructs are valuable for assessing susceptibility to VIGS in Cleomaceae*

Initial screens carried out on small, medium, and large seedlings with a heterologous *PDS* construct (pTRV2-*AtPDS*) were informative about overall susceptibility of *C. violacea* to VIGS (Table 3.1, Figure 3.3A-E). Readily available heterologous constructs are a useful first pass for determining whether select species from Cleomaceae are amenable to VIGS. Survival and silencing-efficacy was similar in *C. violacea* between both heterologous and endogenous *PDS*

constructs (Tables 3.1, 3.2). Overall, more plants survived after treatment with the *Arabidopsis* construct than the endogenous construct, and both constructs had a similar percentage of photobleaching in medium plants. The increased mortality in pTRV2-*CvPDS*-treated plants was likely due to stronger photobleaching because photobleaching reduces efficiency of photosystem II (Wang et al., 2009). The use of endogenous constructs resulted in significant downregulation of *CvPDS* in *C. violacea* (Figure 3.4A). Downregulation of *PDS* is also observed in *Nicotiana benthamiana* Domin using *PDS* constructs from other species of Solanaceae (Senthil-Kumar et al., 2007); even constructs from *Taxus baccata* (Pinophyta) were successfully used to downregulate *PDS* in *N. benthamiana* (Hosseini Tafreshi et al., 2012). Close phylogenetic distance between *C. violacea* and *Arabidopsis* was not indicative of higher susceptibility to VIGS, as another member of Brassicaceae had lower silencing-efficacy despite greater *PDS* sequence similarity (Figure S3.2, Table S3.3). Different species displaying variable susceptibility to VIGS is unsurprising given that different cultivars of the same species, e.g., *Gerbera* (Asteraceae), also show different susceptibility to the same *PDS* vectors (Deng et al., 2012). That is, it appears that susceptibility to a specific vector, and to VIGS itself, is not necessarily predictable by phylogenetic distance.

There are many other factors in determining susceptibility, e.g., the type of viral vector, the length of the insert, and plant growth temperature. Although it varies by study, the effective lower limit for *PDS* silencing using TRV is as low as ~190 bp in *N. benthamiana* (Liu and Page, 2008), and even a size of 23 bp can produce silencing effects when using a potato virus vector (Thomas et al., 2001). Additionally, VIGS appears to be more efficient at lower temperatures, and inhibited at high temperatures, in some species. For example, *Lycopersicon esculentum* Mill. (tomato) (Fu et al., 2006), *Catharanthus roseus* (L.) G. Don (periwinkle) (Sung et al., 2014), and *Petunia hybrida* (Broderick and Jones, 2014); nonetheless, temperatures between 22–24°C function well for *C. violacea* and *Arabidopsis* (Wang et al., 2006). Our data, in congruence with studies in *N. benthamiana*, suggest that a construct sharing at least partial identity to the gene of interest is capable of knockdown, regardless of species origin. In summary, heterologous constructs are an informative first screen for VIGS in *C. violacea* and should be considered when exploring other members of Cleomaceae.

3.4.2 VIGS is a practical tool for studying fruit development in Cleomaceae

Significant downregulation of the key fruit pattern gene *CvFUL* resulted in altered fruit morphology of *C. violacea* (Figure 3.4B, 3.5). In preliminary treatments, *CvFUL* was downregulated in combination with a *CvPDS* marker (Table 3.1, Figure 3.5A-D). We acknowledge the usefulness of marker genes for identifying difficult-to-spot mutant phenotypes and for quick identification when overall silencing-efficacy is low. However, there are limitations to marker genes. For example, in studies of floral symmetry, photobleaching would obscure pigmentation differences between adaxial and abaxial petals of *C. violacea* (Figure 3.1). Additionally, there are potentially confounding effects of decreased plant vitality from photobleaching. For these reasons, and because *CvFUL*-silencing phenotypes are relatively easy to score, albeit difficult to bin, additional VIGS *CvFUL* treatments were completed without a marker (Figure 3.5E-H).

FUL has a critical role in ensuring that the valve margin genetic pathway is restricted to a thin layer of cells between the valve and replum in *Arabidopsis* (Dinneny et al., 2005). Specifically, *FUL* is expressed in the valves and functions in positioning the valve margin in combination with *REPLUMLESS* (*REP*), which has a similar role in the replum. Given the similarities in overall fruit morphology between Brassicaceae and Cleomaceae (Figure 3.2B), we predicted a conserved role in the genetic basis of fruit patterning via *FUL*. In *Arabidopsis*, *ful-1* mutants have fruits with valve cells that have not properly developed and are much rounder than wild-type plants (Figures 3, 7 from: Gu et al., 1998). Mutant *ful Arabidopsis* fruits and, to a lesser extent, other fruit-patterning mutants are shorter in length and have alterations in their valve cellular structure and uniformity (Figure 4 from:Liljegren et al., 2004). Like in *Arabidopsis*, *CvFUL* influences the uniformity of cell expansion in the valve (Figure 3.6D,H). Superficially, treated *C. violacea* fruits (Figure S3.3D) appear more like *Arabidopsis ful-1 35S::AGL8* rescue mutants (Figure 7 from;Gu et al., 1998), which are less stunted than *ful-1* mutants. Incomplete downregulation of *CvFUL* likely explains the somewhat normal fruit length observed during our trials.

Replum cells in fruits of *C. violacea* (Figure 3.6F) are irregularly shaped in *CvFUL-CvPDS*-silenced fruits, a phenotype that is also observed in fruits of *ful-1 Arabidopsis* mutants (Figure 3E from (Gu et al., 1998). We propose that the alteration in replum growth is likely due to shearing forces from the offset growth of valve tissue on either side, which is likely due to incomplete silencing of *CvFUL*. When *FUL* is downregulated, it prevents the formation of a dehiscence zone between the replum and the valve (reviewed in: Ferrándiz and Fourquin, 2014). Thus, both sides of the valve variably push and pull on the replum, and because it cannot separate at the dehiscence zone, it warps the replum and causes fruit curling. A complete knockout of *CvFUL* should not curl because the replum valve boundary would be disrupted evenly along the length of the fruit, i.e., cells in the valves would exert even pressure on the replum. Altogether, these data exhibit how VIGS can be used to target late developmental stages prior to the initiation of those stages.

Cleomaceae is quickly emerging as a model family to address a range of evolutionary and developmental questions (Bayat et al., 2018). Additional functional tools are invaluable for examining the morphological novelties present in the family. We have demonstrated that mild modifications to established VIGS protocols (Gould and Kramer, 2007; Kramer et al., 2007) are sufficient for high silencing-efficacy of *PDS* and *FUL* in *C. violacea* (Tables 3.1, 3.2). We propose that the use of *Arabidopsis* constructs is a viable first step in future investigations of additional species in the family because of the effectiveness of pTRV2-*AtPDS* constructs in this study. Furthermore, the published draft genome of *C. violacea* makes this species a prime candidate for further research on a range of traits.

3.5 Figures

Figure 3.1. (A) *Cleome violacea* whole plant. (B), ventral view of flower, and (C) lateral view of flower.
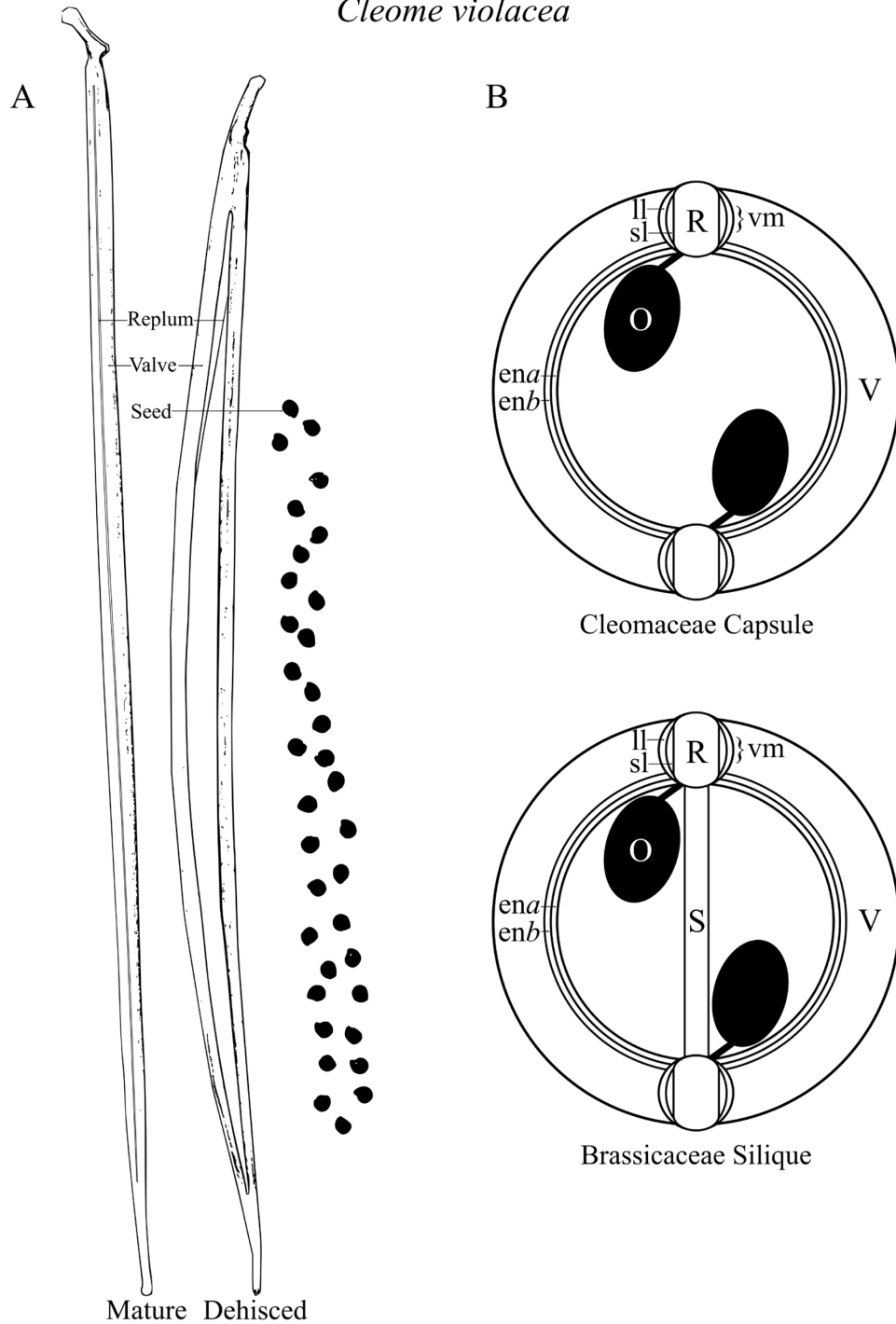
Figure 3.2 Line drawings of *Cleome violacea*. (A) *Cleome violacea* fruit at maturity and after dehiscence. (B) Diagram of Cleomaceae and Brassicaceae fruits in the transverse plane, showing the lignification layer (ll), separation layer (sl), valve margin (vm), replum (R), valve (V), endocarp layer (en*a/b*), ovary (O), and septum (S).

Figure 3.3 *Cleome violacea* subjected to virus-induced gene silencing (VIGS) using both heterologous and endogenous pTRV2-*CvPDS* constructs. (A, B) Leaves after silencing using *AtPDS* displaying (A) pale and (B) variegated photobleaching. (C–E) Leaflets silenced with *AtPDS* displaying (C) pale and (D,E) variegated photobleaching. (F, G) Leaves after treatment with *CvPDS* showing (F) variegated and (G) strong photobleaching. (H–J) Leaflets silenced with *CvPDS* displaying (H) variegated and (I,J) strong photobleaching. (K, L) Untreated *C. violacea* (K) leaf and (L) whole plant. (M, N) Whole plant views of *C. violacea* treated with (M) *AtPDS* and (N) *CvPDS*. Scale bars = 2 mm.

Figure 3.4 Relative quantification (RQ) vs. control of (A) *PHYTOENE DESATURASE* (*CvPDS*)–silenced leaves and (B) *FRUITFULL* (*CvFUL*)–silenced fruits from *Cleome violacea* after virus-induced gene silencing. (A) Untreated (*n* = 4), pTRV2-MCS (*n* = 3), normal phenotype (*n* = 5), >80% photobleaching (*n* = 3), and 50–80% photobleaching (*n* = 3). (B) Untreated (*n* = 3), pTRV2-MCS (*n* = 3), normal phenotype (*n* = 3), and curled (*n* = 6). Error bars indicate RQ maximum and minimum values calculated from standard error. *ACTIN* was used as an endogenous control. Statistical significance was determined using a Welch's *t*-test on delta CT means (α = 0.05), indicated by an asterisk.

Figure 3.5 (A-D) *Cleome violacea* subjected to virus-induced gene silencing using pTRV2-*CvFUL*-*CvPDS* and (E-H) pTRV2-*CvFUL* constructs. (A–D) Fruits displaying photobleaching and atypical development after treatment with pTRV2-*CvPDS-CvFUL*. (E–H) Fruits displaying atypical development after treatment with pTRV2-*CvFUL*. pTRV2-*CvFUL-CvPDS*-treated *C. violacea* with (I) normal development and (J) untreated fruit. Scale bar = 2 mm.
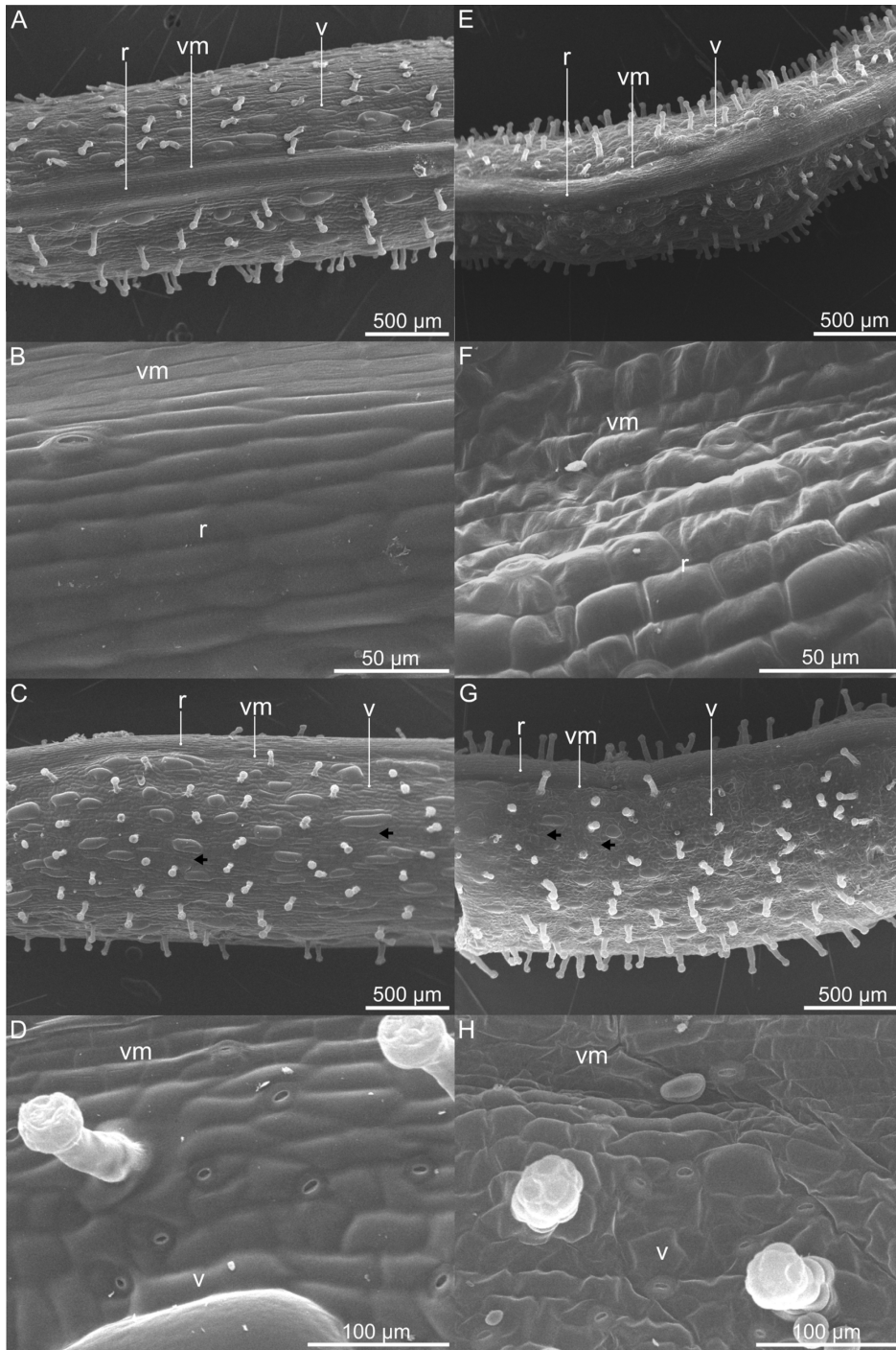
Figure 3.6 Scanning electron micrographs of (A-D) untreated and (E-H) pTRV2-*CvFUL-CvPDS*–
treated *Cleome violacea* fruits. Medial view of (A) untreated and (E) treated fruit; medial view
of (B) untreated and (F) treated replum and valve margin; lateral view of (C) untreated and (G)
treated fruit; lateral view of (D) untreated and (H) treated valve and valve margin. Bulbous
valve cells are indicated by black arrows. r = replum, v = valve, vm = valve margin.
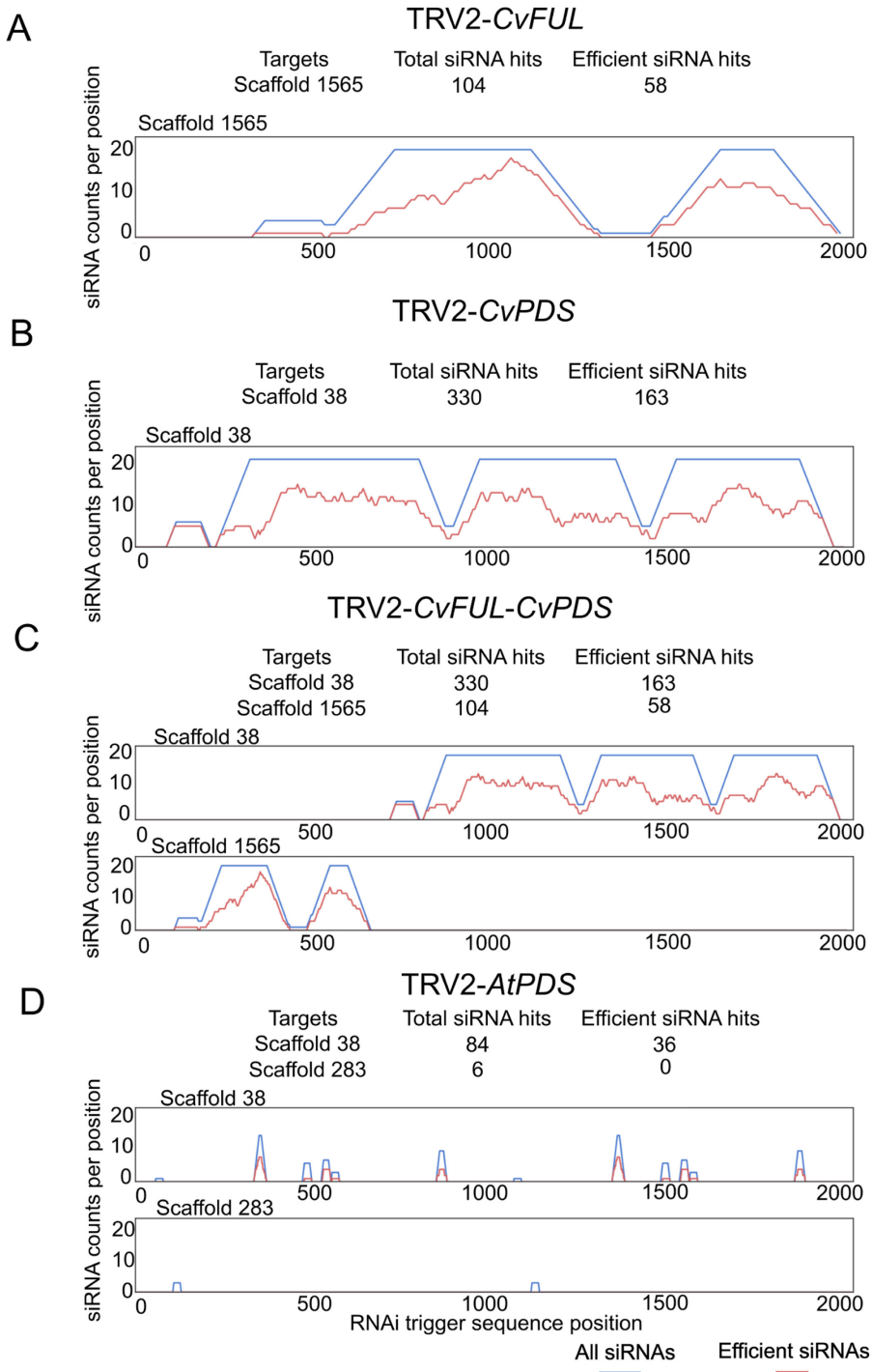
Figure S3.1. SiFi21 results for pTRV2 constructs used in this study. (A) pTRV2-*CvFUL*, (B) pTRV2-*CvPDS*, (C) pTRV2-*CvFUL-CvPDS*, (D) pTRV2-*AtPDS*. Results are compared to the draft genome of *Cleome violacea* (https://genomevolution.org/coge; accession no. 23822) and organized by scaffolds.
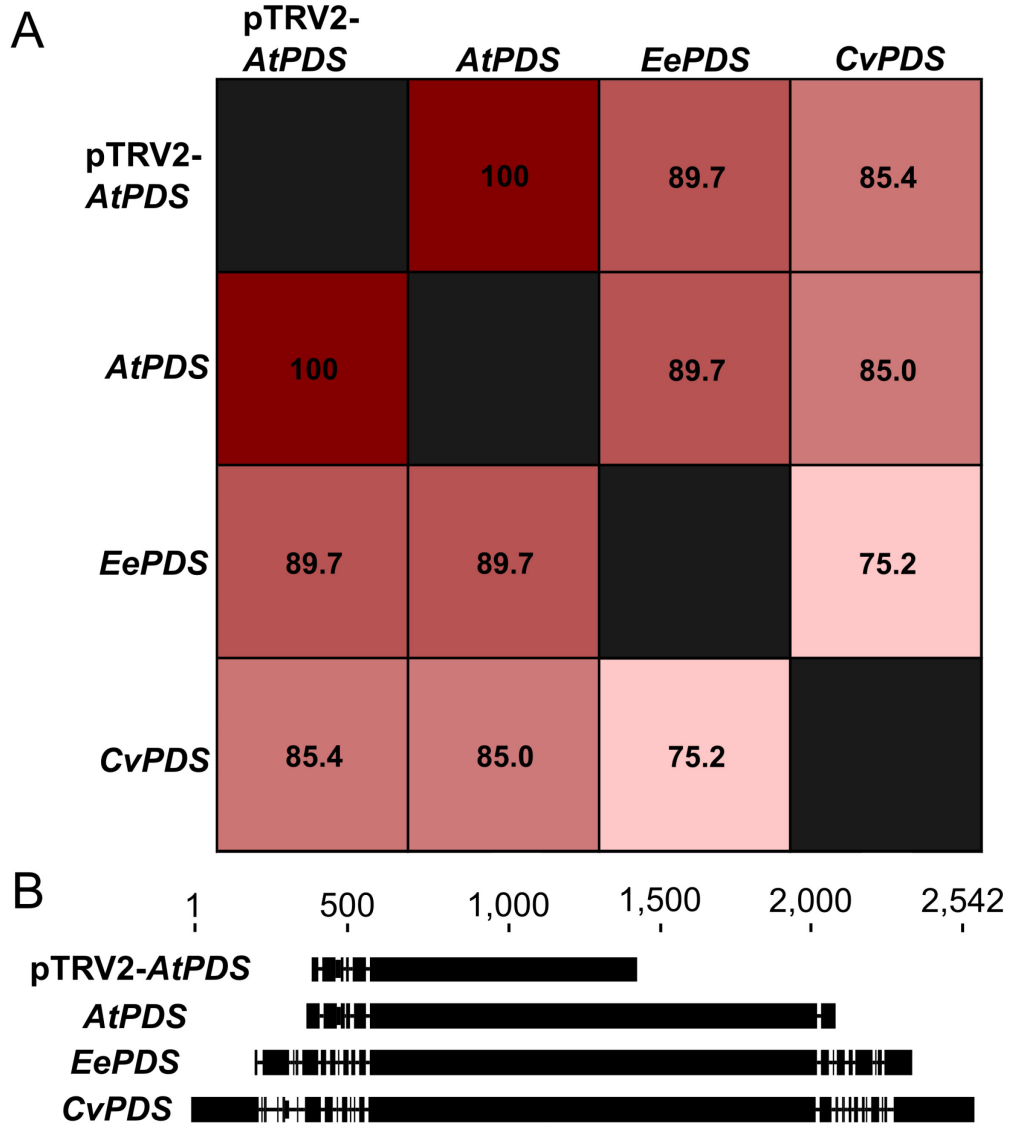
Figure S3.2. (A) Percent identity of the pTRV2-*AtPDS* construct relative to *PDS* from *Arabidopsis thaliana*, *Erucaria erucarioides*, and *Cleome violacea*. (B) Graph of the alignment using Geneious alignment software with 93% similarity cost matrix and default settings.
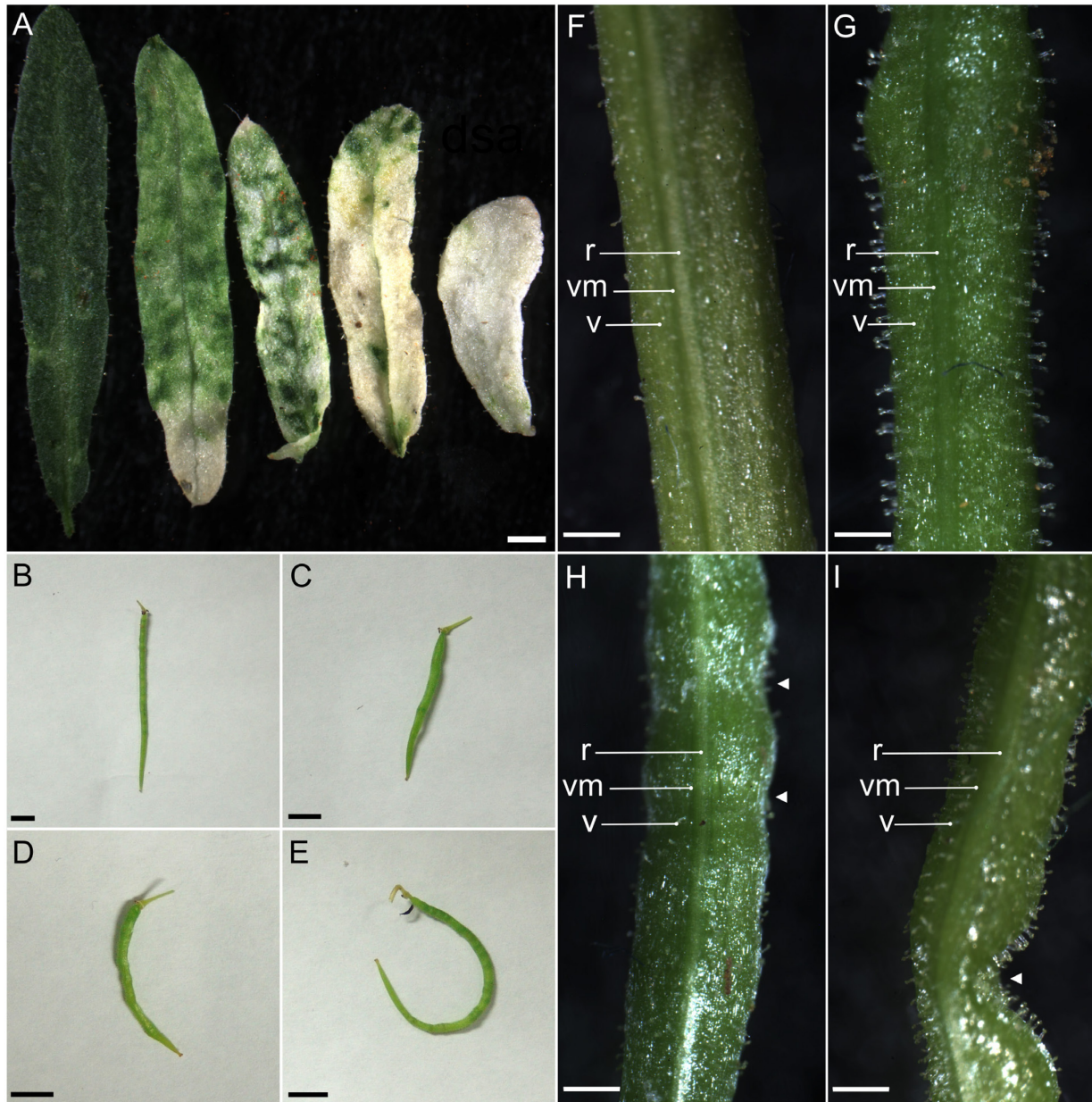
Figure S3.3. (A) *Cleome violacea* leaflets, taken from whole leaves with identical phenotypes, after treatment with pTRV2-*CvPDS*. From left to right: untreated, pale bleaching (0–50%); variegated bleaching (50–80%); strong bleaching (>80%); and complete bleaching. Scale bar = 2 mm. (B–I) *Cleome violacea* fruits after treatment with TRV2-*CvFUL*. (B, F) untreated control (C,G) pTRV2-MCS treated (D,H) fruits with mild curvature with indentation, (E,I) fruits with moderate curvature and indentation. r = replum, v = valve; vm = valve margin; arrows indicate indentation. Scale bars = 0.5 cm (B–E) and 500 µm (F–I).
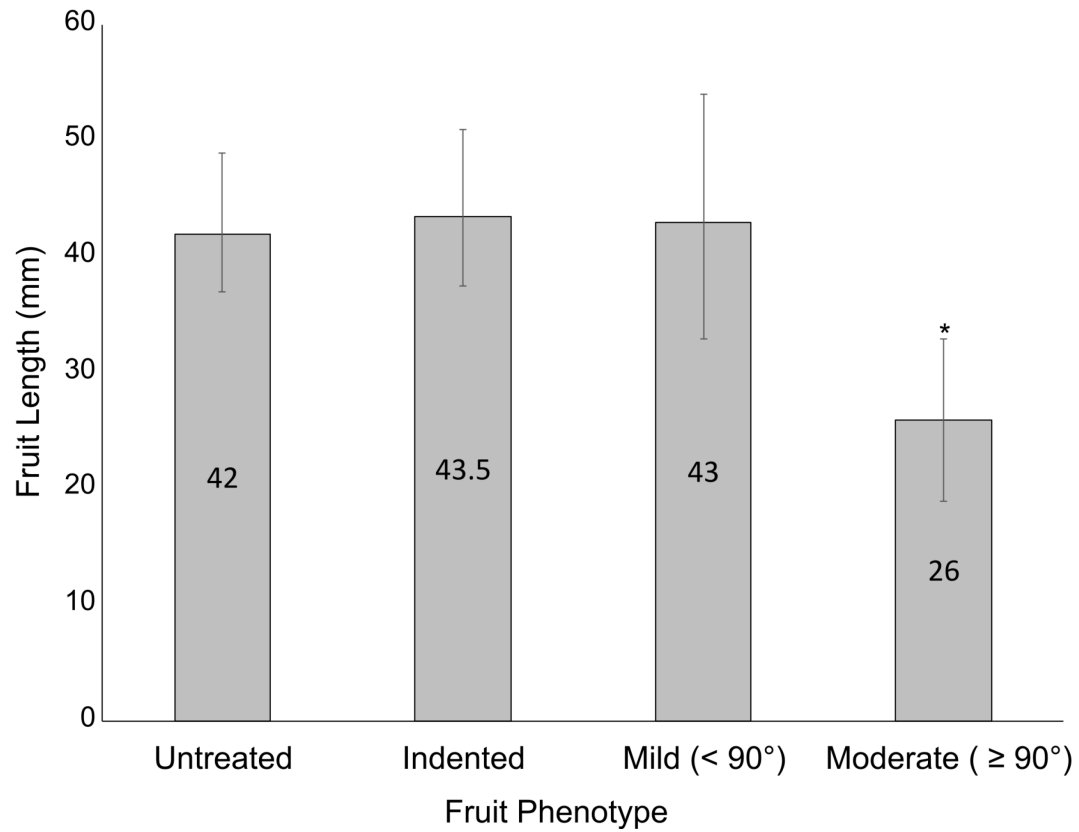
Figure S3.4. Length of *Cleome violacea* fruits treated with pTRV2-*CvFUL*. Normal phenotype (n = 24), indentation (n = 8), mild and moderate curvature with indentation (n = 5). Significance was determined using Welch's t-test (α = 0.01), indicated by an asterisk.

3.6 Tables

Table 3.1. Survival and silencing-efficacy data combined from multiple *Cleome violacea* virus-induced gene silencing trials. Mortality was recorded up to 30 days post-inoculation. Plants were categorized based on the number of true leaves: small (s) with 0–3 true leaves, medium (m) with 4–6 true leaves, and large (l) with ≥7 true leaves.

Treatment	N				Surviving			Displaying Phenotype		
	s	m	l	Total	s	m	l	s	m	l
Untreated	23	30	18	71	20	24	17	n/a	n/a	n/a
pTRV2-MCS	49	85	21	155	5	32	18	n/a	n/a	n/a
pTRV2-*AtPDS*	51	68	79	198	9	53	57	2	39	29
pTRV2-*CvPDS*	25	121	77	223	0	25	33	0	20	17
pTRV2-*CvFUL*-*CvPDS*	0	2	60	62	0	1	57	0	0	21

Table 3.2. Survival and silencing-efficacy data, representing two virus-induced gene silencing trials of medium plants (4–6 true leaves) of *Cleome violacea* grouped by treatment.

Treatment	N	Surviving	Displaying Phenotype
Untreated	50	43	n/a
Vacuum + surfactant	50	45	n/a
pTRV2-MCS	100	80	n/a
CvPDS	100	86	
0% bleached			38
<50% bleached			5
50–80% bleaching			6
>80% bleaching			25
CvFUL	100	85	
Indentation			8
Mild Curvature			28
Moderate curvature			8
Heavy curvature			1

Table S3.1. List of primers used for construct design and RT-qPCR in this study. Brackets indicate restriction sites for incorporation into the pTRV2 vector. Extra base pairs upstream of restriction sites are to improve digestion efficiency. EcoRI (G^AATTC), XbaI (T^CTAGA), BamHI (G^GATCC), and XhoI (C^TCGAG).

Primer	Forward (5' - 3')	Reverse (5' - 3')
AtPDS	CGC[GAATTC]TGCGGCGAAT TTGCCTTATCAAAACG	CGC[TCTAGA]AACTCTTAACCGTGC CATCGTCATTGAG
CvPDS	GG[TCTAGA]TAGTAGATTTGA TTTCCCAGAT	AA[GGATCC]TAGAATTTAGTCGTAC TTCCCC
EePDS	GG[TCTAGA]TAGTAGATTTGA TTTCCCAGAT	AA[GGATCC]TTGAGTTAAGTCGTAC TTCCCC
CvFUL	CGG[TCTAGA]ACCAAGTCAT GTTCGAATCCATAGC	GCT[GGATCC]AAGCAGGAAGAAGA GAAT
CvFUL- CvPDS*	CG[GGATCC]TAGTAGATTTGA TTTCCCAGAT	AA[CTCGAG]TAGAATTTAGTCGTAC TTCCCC
PDS Brassicales Degenerate	TGGAAGGARCACTCMATGAT W TTYGCHATG	ACRACATGRTACTTSAVDATTTTWG CYTT
156	TTACTCAAGGAAGCACGATG AGC	GAACCGTAGTTTAATGTCTTCGGG
FUL qPCR	AGACCTCGATTCTTTGACCTT GA	TGGATTCGAACATGACTTGGTT
PDS qPCR	CGATCGATGCTGGAATTGGT	GACATCATAGAAGCAACAATGAAG GA

*Primers for *CvFUL-CvPDS* introduced modified restriction sites to append *CvPDS* downstream of *CvFUL.*

Table S3.2. Hits to the *Cleome violacea* genome (https://genomevolution.org/coge; accession no. 23822) using *Arabidopsis thaliana* coding sequences of *FUL* and *PDS*, TRV2-*AtPDS*, TRV2-*CvPDS*, and TRV2-*CvFUL* constructs. Default parameters were used for CoGe BLASTn.

Query	Scaffold	Position	HSP no.	Length	Percent ID	*E*-value	Bit-score
FUL AT5G60910	1565	8713	2	197	90.8	8.00E-73	277
FUL AT5G60910	14	859,465	11	146	86.9	6.00E-41	171
FUL AT5G60910	6	979,237	13	189	82	8.00E-40	167
FUL AT5G60910	23	142,350	18	146	83.5	3.00E-32	142
FUL AT5G60910	2	1,080,064	20	147	82.9	4.00E-31	139
PDS AT4G14210	38	217,286	4	217	85.2	3.00E-59	233
PDS AT4G14210	38	215,791	6	156	91.6	6.00E-57	225
PDS AT4G14210	38	218,154	8	187	85	8.00E-49	198
PDS AT4G14210	38	218,531	12	144	86.8	1.00E-39	167
PDS AT4G14210	38	216,676	15	124	89.5	2.00E-38	164
PDS AT4G14210	38	216,161	17	151	83.4	3.00E-33	146
PDS AT4G14210	38	217,590	19	108	88.8	7.00E-31	139
TRV2-*ATPDS*	38	215,791	5	156	91.6	3.00E-57	225
TRV2-*ATPDS*	38	217,286	9	161	87.5	5.00E-48	194
TRV2-*ATPDS*	38	216,676	14	124	89.5	9.00E-39	164
TRV2-*ATPDS*	38	216,161	16	151	83.4	1.00E-33	146
TRV2-*CvFUL*	1565	11,560	10	103	99	3.00E-45	183
TRV2-*CvPDS*	38	216,161	1	151	99.3	1.00E-75	285
TRV2-*CvPDS*	38	216,676	3	124	100	1.00E-61	239
TRV2-*CvPDS*	38	217,286	7	116	100	4.00E-57	223

Table S3.3. Survival and silencing-efficacy data combined from multiple *Erucaria erucarioides* virus-induced gene silencing trials. Mortality was recorded up to 30 days post-inoculation. Plants were categorized based on the number of true leaves: small (s) with 0–3 true leaves, medium (m) with 4–6 true leaves, and large (l) with ≥7 true leaves.

Treatment	N				Surviving			Displaying Phenotype		
	s	m	l	Total	s	m	l	s	m	l
Untreated	19	16	8	43	14	14	7	n/a	n/a	n/a
pTRV2-MCS	37	46	0	83	28	38	0	n/a	n/a	n/a
pTRV2-*AtPDS*	59	111	5	175	50	97	5	8	0	0
pTRV2-*EePDS*	67	40	5	112	46	34	4	5	4	0

Chapter 4: Nectary Development in *Cleome violacea*[i]

4.1 Introduction

Flowers exhibit tremendous diversity of form, much of which is driven by plant-pollinator interactions. Responses to similar pollinator environments have resulted in repeated evolution of floral forms across angiosperms (reviewed in: Endress, 2011; Sauquet et al., 2017; Wessinger and Hileman, 2020). Such traits include, but are not limited to, monosymmetry (zygomorphy), organ fusion, spurs, and heterostyly (Specht and Howarth, 2015; Kramer, 2019; Phillips et al., 2020; Wessinger and Hileman, 2020). This repeated evolution raises fundamental questions about the developmental and genetic bases of their evolutionary shifts (Sobel and Streisfeld, 2013; Specht and Howarth, 2015; Kramer, 2019; Wessinger and Hileman, 2020). Among these questions is whether the same genetic pathways have been recruited in independent origins of these traits (Specht and Howarth, 2015; Wessinger and Hileman, 2020). Remarkable and repeated recruitment of the same genetic pathway is clear with certain traits, notably monosymmetry (reviewed in: Preston et al., 2009; Preston et al., 2011; Hileman, 2014a; Hileman, 2014b; Wessinger and Hileman, 2020), but whether the genetic basis of other features is conserved remains unclear.

Nectaries, and the nectar they secrete, are integral to plant-animal interactions and, as such, warrant detailed investigation across taxa (Liao et al., 2021). They have evolved multiple times across angiosperms and are remarkably variable in position, structure, and morphology (Bernardello, 2007; Nepi, 2007; Nepi et al., 2018; Liao et al., 2021; Slavkovic et al., 2021). Despite this variation, nectaries are united by the simple function of producing nectar, a complex sugar-rich solution that contains a wide range of metabolites and microbes (Nepi, 2007; Heil, 2011; Nepi et al., 2018; Slavkovic et al., 2021; Liao et al., 2021). As a critical reward to insects, and potential attractor, nectaries and their nectar drive many macroevolutionary patterns via relationships with pollinators and other animals (Parachnowitsch et al., 2019; Liao et al., 2021). Nectaries are associated with all plant organs except for roots, and floral nectaries

---

[i] A version of this chapter has been published (Carey et al., 2023). Figures and tables labelled as supplemental are due to journal submission requirements and are not reflective of their contribution to the main thesis.

can be associated with any floral organ (Nepi, 2007; Liao et al., 2021). Nectary morphology can be structured (i.e., distinct morphology with identifiable cell types) or unstructured (i.e., no specialized morphology) (Nepi, 2007; Slavkovic et al., 2021). Nectar secretion ranges from modified stomata (nectarostomata), to specialized trichomes, and even cell rupture (reviewed in: Nepi, 2007; Slavkovic et al., 2021). This diversity in morphology and secretion mechanisms differs across families and within genera (Bernardello, 2007). Also, variable nectar composition impacts pollinator interactions (Nepi et al., 2018; Parachnowitsch et al., 2019). This extensive diversity calls into question whether nectary development is underpinned by similar or different developmental programs in taxa with variable nectaries.

A genetic breakthrough in nectary research was the establishment of *CRABS CLAW* (*CRC*), a *YABBY* family transcription factor, as essential for nectary initiation (Alvarez and Smyth, 1998; Bowman and Smyth, 1999; Lee et al., 2005a). In Arabidopsis, *CRC* knockouts do not develop nectaries (Bowman and Smyth, 1999). *CRC* has since been shown as essential for nectary formation across the core eudicots (Lee et al., 2005b; Fourquin et al., 2014), and is expressed in extrafloral nectaries (Lee et al., 2005b). *CRC* protein dimerizes with other *YABBY* transcription factors and has an important role in *Arabidopsis* carpel development (Alvarez and Smyth, 1998; Alvarez and Smyth, 1999; Alvarez and Smyth, 2002; Lee et al., 2005a) that is widely conserved (Orashakova et al., 2009; Fourquin et al., 2014; Pfannebecker et al., 2017). The expression of *CRC* across core eudicot nectaries, regardless of morphology or position, suggests that *CRC* regulation of nectary development is consistent across the clade (Lee et al., 2005b). To the best of our knowledge, functional studies of nectaries have only been conducted in four core Eudicot taxa: *Petunia* (Solonales) (Lee et al., 2005b; Morel et al., 2018), *Gossypium* (Malvales) (Pei et al., 2021), *Pisum* (Fabales) (Fourquin et al., 2014), and *Arabidopsis* (Brassicales) (Bowman and Smyth, 1999). *CRC* is shown as essential for nectary development in the aforementioned taxa, except for *Gossypium* where the gene *GoNe* is required for both floral and extra floral nectaries (Pei et al., 2021). Thus, investigations of additional taxa are needed to uncover the extent of this potential conserved role of *CRC*.

The role of *CRC* as essential for nectary development does not extend beyond the core-eudicots. For example, all petals of *Aquilegia* have elongated spurs, which bear nectaries in their distal tips. In this taxa, three *STYLISH* (*STY*) homologs, a member of the *SHORT INTERNODES* (*SHI*) gene family, are redundantly necessary for the formation of nectaries in spurs as well as style development (Min et al., 2019). Thus, both *CRC* and *STY* are involved in nectary and gynoecial development (Pfannebecker et al., 2017), which raises questions about developmental pathways shared between nectaries and carpels.

Upstream regulators of *CRC* are also shared between *Petunia* and *Arabidopsis* (Morel et al., 2018). *CRC* is insufficient for ectopic nectary formation (Baum et al., 2001), which reveals a necessity for upstream regulators. In *Arabidopsis*, these regulators include ABC(E) class genes *APETALA2/3* (*AP2/3*), *PISTILLATA* (*PI*), *AGAMOUS* (*AG*), and *SEPALLATA1/2/3* (*SEP1/2/3*), as well as MADS-box gene *SHATTERPROOF 1/2* (*SHP1/2*) (Reviewed in Slavkovic et al., 2021). In sum, *SHP1/2* and *AG* act redundantly to promote *CRC*, such that knockouts of each one alone does not prevent nectary formation, although combined they do (Lee et al., 2005a). The floral meristem identity genes *LEAFY* (*LFY*)and *UNUSUAL FLORAL ORGANS* (*UFO*) are also upstream of *CRC* and function to restrict *CRC* expression to nectaries and carpels (Bowman and Smyth, 1999; Slavkovic et al., 2021). Loss of *SEP1/2/3* also prevents nectaries from developing (Lee et al., 2005a). Individual knockouts of any aforementioned gene do not prevent nectary formation, although they can impact shape and size (e.g., *lfy, ufo, pi, ag*) (Baum et al., 2001). Double and triple knockouts however cause a loss of nectaries (e.g., *lfy* & *ufo*, *sep1/2/3*) (reviewed in: Slavkovic et al., 2021). Nectary inhibition may be indirect because meristem identity genes act upstream of ABC class genes, i.e., nectary formation may be halted because their associated organs fail to form. In *Petunia*, nectary formation is also dependent on C class genes, i.e., *euAG* and *PLEN* are essential for nectary formation (Morel et al., 2018). This redundancy of MADS-box genes implies that the entire regulatory pathway was established prior to the Rosid/Asterid split (Morel et al., 2018; Slavkovic et al., 2021).

Beyond nectary formation, genes have been identified that are important for nectary size and growth. In *Petunia*, two *euAP2* genes, *BLIND ENHANCER* (*BEN*) and *REPRESSOR OF B FUNCTION*

(*ROB*), impact floral nectary size such that *rob1 rob2 rob3* triple mutants have flowers with larger nectaries than wildtype (Morel et al., 2018). This phenotype is enhanced when *BEN* is also knocked out, such that much of the carpel is converted to nectary tissue (Morel et al., 2018). Whereas in Arabidopsis, *BLADE ON PETIOLE 1/2* (*BOP1/2*) are essential for nectary growth independent of *CRC* (Mckim et al., 2008). Knockouts of *BOP1/2* result in nectaries that are small and not fully differentiated into parenchyma and secretory tissue (Mckim et al., 2008).

Phytohormones also play an important role in nectary development, composition, and secretion. *AUXIN RESPONSE FACTOR 6/8* (*ARF6/8*) promote and coordinate nectary formation in *Arabidopsis* (Reeves et al., 2012) and *Aquilegia* (Zhang et al., 2020). Thus, while these taxa differ in which key regulator promotes nectary formation, they have a shared response to hormone signalling, which reflects the central role of plant hormones in floral evolution (Wessinger and Hileman, 2020). Auxin plays an additional role in nectar secretion via *PIN FORMED 6* (*PIN6*) expression, which is positively correlated with nectar production (Bender et al., 2013). Also correlated with an increase in nectar production is jasmonic acid (JA), which peaks in concentration just prior to nectar secretion in *Brassica napus* (Radhika et al., 2010). Further, both auxin and JA are regulated by gibberellic acid (GA) (Reeves et al., 2012), which speaks to the complex interplay between auxin, JA, and GA.

Investigations of additional taxa are critical for assessing not only the extent of the conserved role of *CRC*, but also how it is regulated, and the potential pathway deviations across taxa with different nectary shapes and positions. Towards addressing these outstanding questions, Cleomaceae is an excellent model for investigating floral development. Cleomaceae is a small, cosmopolitan family of circa 270 species placed in 25 genera (Bayat et al., 2018). This family houses floral variation in traits likely associated with pollinator interactions, including petal color, petal size, and gynophores/androgynophores (Iltis et al., 2011; Higuera-Diaz et al., 2015; Bayat et al.,2018). Importantly, members of the family exhibit a wide range of nectary size, shape, and position. Across the family, nectaries may be absent, adaxially positioned, or annular (Tucker and Vanderpool, 2010). Cleomaceae is sister to Brassicaceae and the phylogenetic framework within the family is established (Patchell et al., 2014; Barrett et al.,

2017; Bayat et al., 2018). While some floral developmental patterns are described (Erbar and Leins, 1997; Patchell et al., 2011), most information regarding nectaries is based on floristic work (e.g., Tucker and Vanderpool, 2010). There is also limited empirical information on pollinators, which has revealed generalist and specialist systems across the family (Cane, 2008; Fleming et al., 2009; Higuera-Diaz et al., 2015; Raju and Rani, 2016). Of note, functional approaches have been established for *Cleome violacea* (Carey et al., 2021). This species is amenable to investigations of nectaries as their flowers have prominent, 3-lobed nectaries adaxially positioned between petals and stamen (Figure 4.1).

The overarching goal of this study was to determine the genetic basis of floral nectaries of *Cleome violacea*. Towards this end, we first characterized nectary anatomy and nectar volume. Second, we conducted a detailed transcriptomic analysis of nectaries from pre-anthetic, anthetic, and post-anthetic (post-fertilization) flowers to document gene expression patterns during nectary development and assess possible convergences in underlying genetic pathways. Finally, we conducted functional studies on key genes to test their direct role and putative interactions in nectary development and nectar production.

4.2 Materials and methods

4.2.1 Plant growth conditions

Inbred lines of *C. violacea* were grown from lab seed stock. A voucher was deposited in the vascular plant herbarium at the University of Alberta (ALTA; Hall & Bolton s.n., 20th February 2008; #813 from Hortus Botanicus, Amsterdam). Seeds were sown individually in 7.5 cm diameter pots containing sterilized (20 min, liquid, 121.1°C) Sun Gro Sunshine Mix (Agawam, Massachusetts, USA). All plants were grown in a growth chamber at the University of Alberta, Department of Biological Sciences with 16 h of full spectrum LED light at 22°C and 8 h of darkness at 18°C.

4.2.2 Histology and Scanning Electron Microscopy (SEM)

Inflorescence tips, small buds (<1 mm wide), medium buds (1-1.5 mm wide), large buds (2.5-3 mm wide), flowers, and post-anthesis flowers were collected and fixed in FAA solution (50%

EtOH, 5% glacial acetic acid, 10% formalin, 35% MilliQ water) and vacuum infiltrated as outlined previously (Hall et al., 2006; Patchell et al., 2011). Plant samples were then dehydrated in an ethanol series (50% to 100%). All ethanol solutions were kept at 4°C and samples were incubated for 2 hours. The 100% ethanol solutions were left overnight. Samples were infiltrated with CitriSolv (Decon Labs, USA) by changing to a 1:1 ethanol:CitriSolv solution, then changing to 100% CitriSolv. Each CitriSolv solution was incubated for two hours at room temperature with shaking. Samples were infiltrated with Tissue-Prep paraplast (Leica Biosystems, Canada) with 2-3 changes daily for five days then embedded in paraplast. Samples were sectioned to 8 μm using a Microm HM 325 (GMI, Inc., Ramsey, MN, USA) microtome prior to mounting. Slides were cleared with CitriSolv and dehydrated in isopropanol before staining with 0.025% Alcian blue and 0.01% Safranin O in 0.1M acetate buffer for two hours. Slides were examined using a Nikon (Tokyo, Japan) Eclipse 801 microscope with a Nikon DS-Ri1 photo system.

Samples used for scanning electron microscopy (SEM) were fixed and dehydrated as indicated above. After dehydration, samples were critical point dried with carbon dioxide using a CPD 030 critical point dryer (Bal-Tec AG, Liechtenstein, Germany). Specimens were then dissected and mounted on scanning electron microscopy stubs with conductive carbon tabs and sputter coated with gold using a Hummer 6.2 sputter coater (Anatech USA, Sparks, Nevada, USA). Finally, specimens were imaged using a ZEISS EVO 10 scanning electron microscope (Carl Zeiss AG, Oberkochen, Germany). Contrast and brightness of micrographs were adjusted using GIMP version 2.10.18 (https://gimp.org).

4.2.3 Nectar volume

Nectar volume of *C. violacea* was measured by pooling nectar from all the open flowers on each plant (2-7 flowers) in a capillary tube (Morrant et al., 2009). This measurement was taken for 20 plants and repeated at the same time each day for four consecutive days. Only flowers with visible nectar were measured. Individual flowers typically senesce three days after anthesis and stop producing nectar. Average nectar volume was calculated for each day and then graphed. A student's t-test was run for binary comparisons between day 1-4.

4.2.4 RNA isolation and cDNA library preparation

Nectaries were collected from *C. violacea* flowers at three stages of development: pre-anthetic (buds 2.5-3 mm wide), anthetic (first day of anthesis) and post-anthetic (fertilized flowers with fruits at approximately 10 mm in length). RNA from these three developmental stages of four biological replicates were extracted to provide 1) an overview of gene expression at the end of nectary development and 2) insight into how gene expression changes before, at, and after anthesis. Nectary tissue was excised, flash frozen in liquid nitrogen, and stored at -80°C. Following: Carey et al., (2019), RNA was extracted from manually-ground frozen tissue using a Qiagen RNeasy micro kit (Hilden, Germany) and cDNA was generated using the Illumina TruSeq stranded mRNA LT sample prep kit RS-122-2101 (California, U.S.). In this case, mRNA for each sample was isolated using nucleomag beads (Macherey-Nagel, Düren, Germany). Samples were sent to The Center for Applied Genetics (TCAG) at the Toronto Sick Kids Hospital in Ontario, Canada where they were normalized, pooled, and sequenced on a HiSeq 2500.

4.2.5 De novo transcript assembly, differential expression, and annotation

Raw reads were downloaded from the TCAG webserver and processed as in (Carey et al., 2019) using updated software (Table S4.1). The raw reads are available at the Sequence Read Archive (SRA) database (BioProject: PRJNA912718). After differential expression analysis with edgeR (Robinson et al., 2009), transcripts were classified as significantly differentially expressed if they had a log2 (fold-change) greater than four and a False Discovery Rate (FDR)-corrected p-value (α) less than 0.001. The 'analyze_diff_expr.pl' script, provided with Trinity (Haas et al., 2013), was used to generate a matrix of all significantly differentially expressed contig clustered transcripts, which was then used to generate a z-score heatmap in R (R Core Team, 2013).

We performed an additional z-score analysis with Trinity transcripts filtered using TransDecoder.LongOrfs and TransDecoder.Predict to remove potential misassemblies. In total, 81,151 of 143,919 transcripts remained. A list of the 81,151 transcripts was used to extract significant transcripts from the original matrix file produced by 'analyze_diff_expr.pl',as well as from a list of all transcripts with expression greater than 100 TPM. Additionally, any transcripts

with one or more biological groupings below 10 TPM, or with a coefficient of variation greater than or equal to 50, were removed.

All transcripts from the larger (original) Trinity fasta file were annotated using BLASTx (Altschul et al., 1990) with default parameters and a local copy of the Araport11 protein database. Transcripts with the highest bit-score from the TAIR database were used as representative transcripts. Gene specific heatmaps were generated using ggplot2 and ggplot in R (R Core Team, 2013), respectively. Assembly completeness was determined using Benchmarking Universal Single Copy Orthologs (BUSCO) (Simao et al., 2015), and an ExN90 profile.

Transcripts unique to each stage were uploaded to the KEGG automatic annotation server (KAAS) using the bi-directional best hit against the following organism databases: *Arabidopsis thaliana* (Brassicaceae), *Brassica napa* (Brassicaceae), and *Tarenaya hassleriana* (Cleomaceae). Transcripts were considered unique if their expression was ≥ 10 TPM with a coefficient of variation < 50. A list of all KEGG entries was compiled, excluding most human diseases and other mammalian-exclusive categories. Of note, some categories were kept because they are convergent with pathways in plants and/or yeast.

4.2.6 Virus-induced gene silencing (VIGS)

Viral vector constructs were designed following (Carey et al., 2021). Tobacco rattle virus vectors pTRV1 (donor stock no. YL192) and pTRV2 (donor stock no. YL156) were obtained from The *Arabidopsis* Information Resource (TAIR; https://www. arabidopsis.org) using their stock center (*Arabidopsis* Biological Resource Center [ABRC], Ohio State University, Columbus, Ohio, USA; https://abrc.osu.edu). The pTRV2 vector is used for downregulating genes of interest (Ratcliff et al., 2001) and the pTRV1 vector assists with viral movement (Ziegler-Graff et al., 1991). Six new endogenous constructs were generated for this study using *C. violacea* mRNA: pTRV2-*CvANS*, pTRV2-*CvAG*, pTRV2-*CvAG-CvSHP*, pTRV2-*CvCRC-CvANS*, pTRV2-*CvSWEET9-CvANS*, and pTRV2-DN802_c0_g1_i4-*CvANS*.The *CvANS* construct was used as a marker gene and positive control. Treatment with pTRV2-*CvSHP* was explored in a preliminary round of VIGS but produced no remarkable floral phenotype and, as such, was abandoned in future trials.

All constructs were generated as follows. All cDNA was synthesized following manufacturer instructions using SuperScript III Reverse Transcriptase (Invitrogen), poly(T) primers, and random hexamer primers. All primers were designed using the transcriptomic data from this study. All amplification was done using Invitrogen recombinant Taq DNA Polymerase (Waltham, Massachusetts, USA) using the manufacturer protocol, 50 µL reaction volumes, and 35 cycles. All amplicons were verified using agarose gel electrophoresis, and colonies were screened via PCR with primers, 156F and 156R that span the TRV2 multiple cloning site (Gould and Kramer, 2007). Manufacturer protocols were used for each step unless otherwise noted. First, a 533 bp insert of *CvANS* was amplified using forward and reverse primers with added BAMHI [G^GATCC] and XHOI [C^TCGAG] restriction sites, respectively. Amplicons were purified using a QIAquick PCR purification kit and digested alongside empty TRV2 vector with NEB BAMHI and XHOI restriction enzymes (Ipswich, Massachusetts, USA). Digests were purified using the Quantum Prep Freeze 'N Squeeze DNA Gel Extraction protocol with 200 µL pipette tips and 2 mL tubes in lieu of spin columns. Eluate was further purified using ethanol precipitation (https://projects.iq.harvard.edu/hlalab/resources-0). Digests were ligated together using NEB T4 DNA ligase and immediately transformed using One Shot™ TOP10 Chemically Competent *E. coli*. *Escherichia coli* was incubated for 24 h at 37°C in Miller LB broth (Sigma-Aldrich, Burlington, Massachusetts, USA) containing 50 µg/mL kanamycin. The pTRV2-*CvANS* construct was verified using colony PCR, and colonies containing the appropriately sized plasmids were extracted using a GeneJET Plasmid Miniprep Kit (Thermo Fisher, Waltham, Massachusetts, USA), verified using agarose gel electrophoresis, and transformed into chemically competent *Agrobacterium* GV3101; cells were prepared and transformed according to protocol (Luo et al., 2008). All media used to grow *Agrobacterium* contained 50 µg/mL kanamycin, 50 µg/mL gentamycin, and 25 µg/mL rifampicin. Plasmids from transformed *Agrobacterium* were verified via restriction digestion and agarose gel electrophoresis, and finally sanger sequencing. *Agrobacterium* containing the appropriate pTRV2-CvANS vector were grown for 48 h at 28°C and mixed 1:1 with sterile 50% glycerol prior to storage at -80°C.

All other vectors were made following the same protocol. Amplicons from *CvCRC*, *CvSWEET9*, and DN802_c0_g1_i4 were ligated to pTRV2-*CvANS* vectors using XBAI [T^CTAGA] and BAMHI

restriction sites; *ANS* acting both as a positive control and marker gene for facilitated phenotyping. The pTRV2-*CvAG* construct was generated using BAMHI and XHOI restriction sites, as with pTRV2-*CvANS*.The *CvSHP* amplicon was then ligated to the pTRV2-*CvAG* vector using XBAI and BAMHI restriction sites. No *CvANS* marker was used for the *CvAG* or *CvAG-CvSHP* constructs because downregulation of *CvAG* is distinct. Viral constructs were verified for off-target silencing using siFi21 (Lück et al., 2019).

Agrobacterium tumefaciens was prepared for DNA transformation as previously described (Carey et al., 2021). All vectors were transformed into *A. tumefaciens* using calcium chloride heat-shock transformation. For each, 100 ng of purified construct was combined with 250 µL of competent *A. tumefaciens*. Transformants were plated on LB media containing the aforementioned antibiotics. Transformants were then screened as before using 156F and 156R primers, and glycerol stocks were made and stored at −80°C (1:1 ratio of 50%glycerol and overnight *A. tumefaciens* culture).

The vacuum infiltration protocol, which has been shown to be an effective infiltration method with *C. violacea*, was modified from Carey et al. (2021). For each vector, *A. tumefaciens* cultures were serially inoculated up to 1000 mL cultures containing antibiotics, 1mM MES buffer and 0.02 mM acetosynringone. A 1:1 ratio of pTRV1 cultures were also serially inoculated up to 1000 mL. The final cultures were grown until they reached an OD600 between 0.8-1, and then immediately centrifuged and resuspended in infiltration buffer (10 mM MES, 10 mM MgCl2 and 0.2 mM acetosyringone) to an OD600 of 4.0 ± 0.1 and left for four hours to acclimatize. *Agrobacterium* containing pTRV1 should be inoculated 1-2 hours prior to pTRV2 cultures because they have a slower growth rate. An OD600 of four was chosen because it has been reported to achieve greater yields, and when pTRV2 and pTRV1 are combined their OD600 values half to an optimal OD600 of 2.0 (Wang et al., 2006). The serial inoculation was halted at OD600 0.8 to capture log-phase growth. The pTRV2 and pTRV1 suspensions were combined prior to infiltration at a 1:1 ratio. Silwet L-77 surfactant was added to each mixture at 100 µL/L. Groups of seedlings were extracted from the soil, rinsed in reverse osmosis water, briefly air-dried, submerged in *Agrobacterium*, and placed in a vacuum chamber. The chamber was

evacuated to -20 inHg and held for 2 minutes. Vacuum pressure was then quickly released, and plants were rinsed and planted in fresh soil. Finally, plants were grown at 22°C for 16 h and 18°C for 8 h because it was found that lower temperatures consistently resulted in better VIGS efficacy in *Petunia* (Broderick and Jones, 2014).

All treated plants began showing phenotypes five weeks post-inoculation, and phenotypes lasted until plant senescence. Phenotypes for construct pTRV2-*CvANS* were scored based on reduction in maroon pigmentation in petals, which is hereafter referred to as yellowing, i.e., a reduction in anthocyanins resulted in petals with increased yellow pigmentation. A flower was scored as having a moderate phenotype when at least two petals displayed obvious yellowing. Flowers were scored as having a strong phenotype if all four petals displayed obvious yellowing. Flowers with less yellowing than moderate flowers, but which were distinct from untreated flowers, were scored as having a mild phenotype. There were no observed instances of only a single petal yellowing.

The yellowing phenotype assisted with scoring of *CvCRC*, *CvSWEET9*, and DN802_c0_g1_i4 constructs. Phenotypes for pTRV2-*CvCRC-CvANS* were scored based on complete or partial absence of nectary. Yellowed flowers with complete nectaries, and non-yellow flowers without nectaries were also recorded because it is possible for only a single gene to be silenced even with multiple gene constructs. Phenotypes for pTRV2-*CvSWEET9-CvANS* were scored based on visual inspection of nectary gland for presence of nectary droplets using a dissection microscope. Phenotypes for pTRV2-DN802_c0_g1_i4-*CvANS* were indistinguishable from pTRV2-*CvANS*.

Phenotypes for pTRV2-*CvAG* and pTRV2-*CvAG-CvSHP* were scored based on *AG* silenced phenotypes in *Arabidopsis* because of conservation of ABC gene function (Mizukami and Ma 1997). Silencing efficacy was based on the extent of repetition of perianth whorls and the absence of reproductive whorls. For both constructs, presence/absence of nectaries, absence of reproductive whorls, and repetition of perianth whorls were noted. Plant tissue from treated and control plants was excised, flash-frozen in liquid nitrogen, and stored at -80°C. Phenotypes were imaged using a Nikon SMZ 1500 dissecting microscope (Nikon, Tokyo, Japan) and a

handheld digital Canon DS126181 camera (Canon, Tokyo, Japan). Images were standardized, scaled, color balanced, and assembled into figures using Inkscape version 0.92.5 (https://inkscape.org) and GIMP.

4.3 Results

4.3.1 Morphology and nectar production in Cleome violacea

Anthetic nectaries of *C. violacea* are adaxially positioned between petals and stamen. These nectaries are prominent due to their relatively large size (i.e., roughly half the size of an abaxial petal). Nectaries are tri-lobulate with two larger lateral lobes and a smaller central lobe at anthesis (Figure 4.1B–E). Following the terminology of Nepi (2007), these structured nectaries have prominent epidermis, and mostly consist of specialized parenchyma with vascular tissue interspersed throughout (Figure 4.2). Nectar is secreted via nectarostomata, which are present prior to anthesis (Figure 4.3). Floral nectaries are first visible late in development when developing stamens reach the same length as petals. At this stage, sepals and petals are growing, stamens have differentiated into filaments and anthers, and the gynoecium is formed with papillate stigma. Nectaries mature concordantly with stamens and reach maturity just prior to anthesis (Figure 4.1B). Nectary primordia are visible in small buds (Figures 4.2) when sepals are maturing.

In small buds (<1.0 mm wide), nectaries are oblong and marginally lobed. At this stage, cells appear parenchymal with no differentiation of epidermis or vascular tissue, although the cuticle is apparent (Figure 4.2A–C). Medium buds (1-1.5 mm wide) have more pronounced lobes with differentiated epidermis (Figure 4.2D–F). Large buds (2.5-3.0 mm wide) have larger lobes comprised of parenchymal cells which make up the bulk of the nectary (Figure 4.2G–I). Epidermal cells are 1-2 layers thick on medial and lateral nectary lobes (Figure 4.2Q). In large buds, vascular tissue is distributed throughout the specialized parenchyma and likely connects with other vasculature near the receptacle base (Figure 4.2P) and with nectarostomata on the nectary surface (Figure 4.3). These nectarostomata are present on large buds (Figure 4.3A) and anthetic flowers (Figure 4.3B). Nectaries produce a low volume of nectar that decreases in volume after anthesis (Figure 4.4). Anthetic nectaries produce an average of 0.17 ± 0.07 µL) of

nectar (Figure 4.4). Nectar volume decreases after day 1 but remains stable over three consecutive days of sampling at ~ 0.11 ± 0.04 μL) (Figure 4.4).

4.3.2 Expression profiles show distinct gene expression patterns pre to post-anthesis

The transcriptome is of suitable quality and completeness for downstream analyses. The transcriptomic read depth averaged 19.9 million reads across 12 biological replicates, totaling just over 239M paired end trimmed reads. Median Phred scores are between 34 and 39 for each base pair of all 143,919 Trinity transcripts (Table S4.2), which indicates a base call accuracy between 99.7% and 99.99% (data not shown). The E90N50 value of the assembled transcriptome is 2227, and peaks at 2259 for Ex93 and Ex94 (Figure S4.1). A peak around Ex90 generally indicates a high level of transcriptome completeness. Further, the Benchmarking Universal Single Copy Orthologues (BUSCO) analysis of Viridiplantae orthologues (Simao et al., 2015) revealed that the transcriptome was 99.6% complete with 2 fragmented BUSCOs (Table S4.3).

Heatmap patterns of gene expression of pre-anthetic, anthetic and post-anthetic nectaries are consistent across two distinct thresholds. We compared all 4521 significantly differentially expressed transcripts, as well as the 1214 transcripts above 100 TPM (with a coefficient of variation < 50 in one or more biological groupings) (Figure 4.5). Pre and post-anthetic nectaries have opposing expression profiles, such that transcripts upregulated in pre-anthetic nectaries are generally downregulated in post-anthetic nectaries. Anthetic nectaries have no large clusters of up or downregulated transcripts and appear to be partially transitional, although they have a few unique clusters of differential expression (Figure 4.5). Expression patterns of transcripts filtered using TransDecoder were similar to the unfiltered list for significantly differentially expressed transcripts, as well as those above 100 TPM (Figure S4.5). In sum, each of the three developmental stages is genetically distinct.

To provide additional insight into gene regulatory networks governing nectar secretion and nectary development, we also assembled the highest expressed transcripts across all stages from the TransDecoder-filtered dataset. This list included 20 of the highest expressed transcripts from pre-anthetic, anthetic and post-anthetic stages that were significantly

differentially expressed in pairwise comparisons, and the top 20 highest overall expressed transcripts that were not differentially expressed. Due to overlap, there were 56 transcripts in total (40 among the differentially expressed and 16 non-differentially highly expressed transcripts). A few transcripts matched to the same genes leaving 51 unique accessions. Out of the 51, eight had no obvious role specific to nectary function (e.g., ubiquitous cellular process; Table S4.4). Five were related to photosynthesis, 14 to water transport and sugar production, 16 to stress response and six to cell growth. *YABBY5*, which can dimerize with *CRC* (Gross et al., 2018) was also among the highest expressed (Table S4.4). Putative gene function was estimated using gene description information from TAIR (www.arabidopsis.org), and a brief review of the literature.

4.3.3 Energy metabolism and hormonal regulation across nectary development

Genetic networks were assessed to determine which categories were active at each sampled stage of nectary development. KEGG analyses revealed several categories that had different relative counts (i.e., putative orthologs) in one or more stages: energy metabolism, biosynthesis of secondary metabolites, translation, replication and repair, environmental adaptation, and cell growth and death. We interpret greater counts as greater biological activity. A difference of three or less was disregarded to account for any potential noise in the data, e.g., invalid isoforms created during the assembly process. In pre-anthetic nectaries, oxidative phosphorylation (35 relative to 25 in the other two stages), and thermogenesis (40 relative to 32 in the other two stages) are the only subcategories with a greater number of putative orthologs (Table S4.5). Mitochondrial oxidative phosphorylation in plants is known to provide ATP for cellular functions (e.g., sucrose metabolism) and is tightly linked to photosynthesis (Braun, 2020). However, photosynthetic processes are similar between all stages (Table S4.5). Anthetic nectaries have no categories with greater hits than the other two stages (Table S4.5). Post-anthetic nectaries have increased biological activity in two categories, replication and repair, and cell growth and death; three of the five subcategories for cell growth and death are directly related to yeast (Table S4.5), i.e., nectary yeast are likely contributing to ortholog abundance in this category. Overall, most categories have a similar number of putative orthologues across all stages.

Although plant hormone signalling is important in nectary development and nectar secretion, there was no indication of differences between hormone signalling related orthologs between stages, based on the KEGG analysis (Table S4.5). However, expression analyses indicated significant differential expression in genes related to these pathways (Figure 4.6). Auxin, JA, and GA are known to regulate transcriptional expression related to nectar secretion (Slavkovic et al., 2021), and ethylene interacts synergistically with auxin (Muday et al., 2012), although to our knowledge has no direct link to nectaries. Examples of highly expressed transcripts were *AUXIN RESPONSE FACTOR 6* (*ARF6*) and *JASMONATE ZIM-DOMAIN PROTEIN 1* (*JAZ1*) (Figure 4.6A, C). Three *GIBBERELLIN 2-OXIDASE* genes were expressed in pre-anthetic, anthetic and post-anthetic nectaries, respectively (Figure 4.6B). We also found significant upregulation of multiple ethylene related transcripts in pre-anthetic nectaries (e.g., *ETHYLENE RESPONSE FACTOR 1* (*ERF1*) and *ETHYLENE FORMING ENZYME* (*EFE1*). (Figure 4.6D). These data suggest that a combination of auxin, JA, and GA influence nectary development and nectar secretion in *C. violacea*.

4.3.4 Yeast and bacteria are present on Cleome violacea *nectaries*

Transcriptomic and SEM data provide evidence that yeast and bacteria colonize *C. violacea* nectaries. There are a total of 46 and 44 hits (e-value < 1e-50) to bacterial and yeast-related rRNA in the *C. violacea* transcriptome, respectively (Figure 4.7). Ribosomal rRNA can still be present in poly(A)-enriched libraries in appreciable percentages (Kim et al., 2019), which is valuable for finding non-plant related expression. Generally, expression of fungal and bacterial rRNA was inconsistent across biological replicates and stages; suggesting that colonization may be replicate specific. However, there are a few instances where expression is consistent across replicates and stages, which may indicate an established biological interaction (Figure 4.7). These data are further supported by the obstructions surrounding and within the nectarostomata (e.g., what appears to be budding yeast cells) (Figure 4.3F–H) and are consistent with the abundance of yeast related KEGG terms (Table S4.5). For example, there are nearly twofold more KEGG terms related to the yeast cell cycle in post-anthetic nectaries (43) than pre-anthetic (22) or anthetic (20) nectaries (Table S4.5). Carotenoid related genes *FLAVONOL SYNTHASE 1*(*FLS1*), *PHYTOENE SYNTHASE* (*PSY*), and *CHALCONE SYNTHASE* (*CHS*) are

also highly expressed at various developmental stages (Table 4.1 and Figure S4.2). All three genes have purported roles in combating biotic stress (Dao et al., 2011; Havaux, 2014; Naparlo et al., 2019).

4.3.5 Dynamic expression patterns of genes involved in nectar and nectary formation.

After establishing global patterns and active biological networks, we examined expression patterns of 17 genes of interest with uncertain roles, nine genes known to be involved in nectar production, and ten genes with direct roles in nectary formation (e.g., expression in *Arabidopsis* nectaries) (Tables 4.1-4.3; Figure S4.2). These analyses reveal dynamic expression patterns from pre to post-anthetic nectaries. Seven genes linked to nectar production are significantly upregulated in either pre-anthetic and/or anthetic nectaries: *BAM1*, *PIN6*, *MYB21*, *SWEET9*, *JAZ*, *G20X* and *JMT* (Table 4.2). This pattern mirrors the onset of nectar production. Interestingly, expression profiles for nectary development genes (Table 4.3) are generally opposite to nectar production (Table 4.2). That is, of the genes examined with established roles in nectary development, most transcripts are downregulated in pre-anthetic nectaries, and three of ten genes explored are evenly expressed across all stages (Table 4.3). Downregulated genes include *PI*, *AG*, *ARF6*, *ARF8*, and *STY* (Table 4.3). Of the 17 genes with uncertain roles, six are evenly expressed across all three developmental stages: *SEP1/4*, *TOPLESS* (*TPL*), *SUPRESSOR OF OVEREXPRESSION OF CO 1* (*SOC1*), *GIGANTEA* (*GI*), and *FLOWERING LOCUS D LIKE* (*FLD-like*). Of the remaining, no clear pattern emerges (Table 4.1).

4.3.6 VIGS demonstrates plausible roles for CvAG*,* CvSHP*, and* CvSWEET9 *in nectary development*

VIGS experiments tested the putative function of five genes. Four of the five genes targeted for downregulation were highly expressed and have established roles in nectary development and nectar production: *CvAG*, *CvSHP*, *CvCRC*, and *CvSWEET9* (Tables 4.2, 4.3 and Figure S4.2). We also downregulated an uncharacterized transcript (DN802_c0_g1_i4) because it was among the highest expressed in the transcriptome, has a similar profile to *CvSWEET9*, and has no significant match to either the TAIR11 database or the nr database, despite an ORF of 375 bp (Figure S4.3). It partially matches *AT412520.1* (e-value = 8.19e-4) from the TAIR11 database and

MW419336 (e-value = 6.33e-7) from the nr database. These hits were not considered further because of their large e-values.

ANS was used both as a positive control vector (pTRV2-*CvANS*) and as a marker gene to facilitate scoring of phenotypes for *CvCRC* (pTRV2-*CvCRC-CvANS*), *CvSWEET9* (pTRV2-*CvSWEET9-CvANS*), and DN802_c0_g1_i4 (pTRV2-DN802_c0_g1_i4-*CvANS*). Untreated *C. violacea* and plants treated with pTRV2-MCS constructs were also used as controls (Figure 4.8A–D). Treatment with pTRV2-*CvANS* produced flowers with primarily yellow pigmentation on adaxial and abaxial petals (Figure 4.8E, F). Yellowing is purported to be from a disruption of the anthocyanin production pathway and was the visual marker used for other constructs because it is not expected to alter the form or function of nectaries. *CvANS* is only moderately expressed in post-anthetic nectaries (Figure S4.2).

With exception of pTRV2-DN802_c0_g1_i4 (Figure S4.4), all other treatment groups had marked phenotypes related to nectary and nectar formation. DN802_c0_g1_i4 was highly expressed in pre-anthetic and anthetic nectaries, and relatively downregulated in post-anthetic nectaries (Figure S4.2). Treatment with pTRV2-DN802_c0_g1_i4-*CvANS* resulted in flowers that were phenotypically indistinguishable from the *CvANS* control (Figure S4.4; Figure 4.8E, F), despite a high efficacy and mortality among treated flowers relative to control (Table S4.6). It is plausible that this transcript is related to water and/or nutrient transport because of its unusually high mortality during silencing (Table S4.6), absence of discernible silencing phenotype (Figure S4.4), and similar expression profile to *CvSWEET9* (Figure S4.2), but further research is required.

Functional studies suggest the role of *CvCRC* and *CvSWEET9* in nectary and nectar formation, respectively. *CvCRC* is expressed across all three developmental stages investigated, but with no significant difference between stages (Table 4.3). In contrast, *CvSWEET9* was downregulated in post-anthetic nectaries as compared to pre-anthetic and anthetic nectaries (Table 4.2). Treatment with pTRV2-*CvCRC-CvANS* resulted in either partial or total loss of nectaries in all flowers with yellowing phenotype (Figure 4.9; Table S4.6). Treatment with pTRV2-*CvSWEET9-*

CvANS produced flowers with a visible reduction in nectar on their nectaries (Figure 4.10), and no detectable sugar using a refractometer (n = 10; data not shown).

Individual and combined constructs of *CvAG* and *CvSHP* demonstrate these genes are functionally redundant in regulation of nectary formation. Treatment with pTRV2-*CvAG* and pTRV2-*CvAG-CvSHP* produced flowers without reproductive whorls and repeating perianth (Figures 4.11, 4.12). Flowers treated with pTRV2-*CvAG* still produced nectaries, but their position and structure were altered relative to untreated flowers (Figure 4.11). This misplacement is likely due to a loss of reproductive whorls and repeated morphology. Flowers treated with pTRV2-*CvAG-CvSHP* generally produced no nectaries, although occasionally they were present and reduced in size (Figure 4.12A, F).

4.4 Discussion

4.4.1 Cleome violacea *have structured nectaries that produce nectar secreted via nectarostomata*

The nectary of *C. violacea* is striking in that it is a prominent feature of the flower, due to its large size and location (Figure 4.1). The nectary is adaxially positioned between stamens and adaxial petals, contributing to monosymmetry of the flower in addition to petal color patterning and reproductive organ curvature. Nectaries appear late in development, well after initiation of stamens and gynoecium. Once formed, mature nectaries are a large 3-lobulate structure (Figures 4.1–4.3). *Cleome violacea* nectaries are characteristic of many other structured nectaries (Nepi, 2007): (1) the nectary epidermis has thick cuticle, (2) the nectary parenchyma is made up of small, dense cells, and (3) the vasculature is interspersed throughout the nectary and likely connects with vascular bundles in the receptacle (Figure 4.2P). Presence of nectarostomata on nectaries of *C. violacea* has been described previously (Erbar and Leins, 1997b). In the annular nectary of *Cleomella sparsifolia* (= *Cleome sparsifolia*), nectarostomata appear on abaxial tips (Lee et al., 2005b), suggesting nectarostomata may be common in Cleomaceae. Nectarostomata are modified stomata that secrete carbohydrate rich solutions for pollinator reward; a similar genetic pathway regulates both nectarostomata and unmodified stomata (Pillitteri et al., 2008; Baylis et al., 2013).

Nectar is known to be secreted via a few different methods, and most commonly via nectarostomata in a granulocrine or eccrine based manner (Nepi, 2007). Outlined by Roy et al., (2017), eccrine based secretion begins with the breakdown of starch and subsequent synthesis of sucrose, which is then transported out of the cell and hydrolyzed before secretion out of nectarostomata in droplets of nectar. Key to the export of sucrose is *SWEET9*, which is essential for sugar transport in nectaries of *Arabidopsis*, *Brassica rapa*, and *Nicotiana* (Lin et al., 2014). Another relevant gene, *CWINV4* is important for nectar formation in *Arabidopsis*, specifically cleaving sucrose in the extracellular space which has the effect of moving water towards sugar, forming nectary droplets (Ruhlmann et al., 2010). In *C. violacea*, multiple lines of evidence support eccrine-based nectar secretion. First, there are nectarostomata on the nectary surface (Figure 4.3C–H) which likely connect with the vasculature present throughout parenchymal tissue (Figure 4.2P). Second, both *CvSWEET9* and *CvCWINV4* are highly expressed in pre-anthetic and anthetic nectaries (Table 4.2 and Figure S4.2). Additionally, we identified 14 highly expressed transcripts that are related to sugar production or water transport, e.g., five of which are related to aquaporins found in *Aquilegia* (Singh et al., 2020) (Figure S4.5 and Table S4.4). Finally, nectar secretion is lessened when *CvSWEET9* is completely downregulated (Figure 4.10). In sum, nectary secretion in *C. violacea* is dependent on *CvSWEET9*, as demonstrated for *Arabidopsis*, *Brassica* and *Nicotiana*, which supports its key role in sucrose export across the core eudicots (Lin et al., 2014).

Nectar is secreted at anthesis, accumulates on *C. violacea* nectary lobes (Figure 4.1E). and has a low average secretion volume (0.17 ± 0.07 µL) (Figure 4.4). This volume is lower than averages of wild populations of other species of Cleomaceae: *Cleomella serrulata* (0.85 ± 0.96 µL) and *Polanisia dodecandra* (0.63 ± 0.32 µL) (Higuera-Diaz et al., 2015). However, it is similar to the average volume produced by one species of Brassicaceae: *Erysimum mediohispanicum* (0.136 ± 0.010 µL). The differences in nectar volume may be in part explained by flower size as *C. violacea* has much smaller flowers than *C. serrulata* and *P. dodecandra*. It may also reflect different pollinator environments; flowers of *C. serrulata* and *P. dodecandra* have a wide range of visitors and somewhat overlap in geography in some areas of North America (Higuera-Diaz et al., 2015). It is also unclear if there is a reduction of nectar in lab-grown inbred lines of *C.*

violacea relative to wild populations. No empirical pollination study has been conducted on *C. violacea* to date, which is native to Spain (GBIF.org), so there is no information on which pollinators would be attracted to and rewarded by its nectar.

4.4.2 CvCRC*,* CvSHP*, and* CvAG*, exhibit conserved roles with other core Eudicots in nectary formation*

CRC is essential for nectary formation in *Arabidopsis*, *Petunia*, *Pisum*, and *Medicago* in addition to having an important role in carpel formation (Baum et al., 2001; Lee et al., 2005b; Fourquin et al., 2014). As with these taxa and other core Eudicots (Lee et al., 2005b; Slavkovic et al., 2021), *CvCRC* is expressed in developing nectaries without any significant difference in gene expression patterns from late-stage buds to post-anthetic flowers (Table 4.3 and Figure S4.2). Like with other species, loss of *CvCRC* resulted in an absence of nectaries (Figure 4.9), which demonstrates that *CvCRC* is essential for nectary formation in *C. violacea*. While strong *CsCRC* expression in nectaries of *C. sparsifolia* implied the conserved role of *CRC* (Lee et al., 2005b), this study provides the first functional evidence of the direct contribution of *CRC* to nectary formation in Cleomaceae. Since the nectaries of *C. sparsifolia* are annular, forming a ring around the stamen base, as compared to the adaxial position of *C. violacea* nectaries (Figure 4.1), these data indicate that upstream regulators of *CRC* are likely important for nectary position and morphology within Cleomaceae flowers. In *Medicago* and *Pisum*, inconspicuous nectaries form at the base of the staminal tube. Like with Fabaceae, nectaries of *Arabidopsis* are found at the base of stamens, although in this instance forming six glands on the abaxial side. Nectaries in *Petunia* form a ring at the base of the gynoecium (Morel et al., 2018). Interestingly, unlike knockout or knockdowns of *CRC* in *Arabidopsis* (Alvarez and Smyth, 1999; Alvarez and Smyth,2002), Fabaceae (Ferrandiz and Fourquin, 2014), and poppy (Orashakova et al., 2009), we did not observe many notable changes to gynoecium or fruit formation in *CvCRC* knockdowns (but see Figure 4.9C). Additional studies are necessary to explore the extent of *CRC*'s role in gynoecium development and its conservation in *C. violacea*. The highly expressed *YABBY5* (Table 4.1, Figure S4.2) should also be explored due to its ability to dimerize with *CRC* (Gross et al., 2018), i.e., it may share a role with *CRC* in *C. violacea*. *CRC* homologs have variable importance in carpel formation across core Eudicots (Morel et al., 2018), which also warrants

further examination in *C. violacea*. As shown in Arabidopsis (Pinyopich et al., 2003), we predict a high redundancy of gene function for gynoecial formation in *C. violacea* given its importance to plant fitness.

MADS-box genes *AG* and *SHP* act redundantly upstream of *CRC* in both *Arabidopsis* and *Petunia* to initiate nectary development (Morel et al., 2018). The regulatory roles of these genes appear to be conserved in nectary formation of *C. violacea*. Both *CvAG* and *CvSHP* are strongly expressed across all stages of development (Table 4.3 and Figure S4.2). Treatment with pTRV2-*CvSHP* or pTRV2-*CvAG* alone is insufficient to prevent nectaries from forming (Figure 4.11 & Table S4.6). Treatment with pTRV2-*CvSHP* alone has no effect on floral phenotype (Table S4.6) but treatment with pTRV2-*CvAG* disrupts the formation of whorls 3 and 4 (Figure 4.11). These phenotypes in *C. violacea* are like *Arabidopsis ag-1* mutants (Baum et al., 2001). Only doubly silenced flowers do not produce nectaries, although they are otherwise like flowers treated with pTRV2-*CvAG* (Figure 4.12). Our data is consistent with the model from Wollmann et al. (2010) which shows a balance between *AP2* and *AG* activities, i.e., in pTRV2-*CvAG* treated flowers, stamens occasionally appear petaloid (Figure 4.11F). Thus, there is the possibility that the overlapping of whorls may be the condition which contributes to nectary formation because all the ABC genes are expressed in nectaries to some degree (Table 4.3 and Figure S4.2). When flowers are treated with pTRV2-*CvAG*, and petals form haphazardly, nectary tissue surrounds each petal at the base of the flower and the lobe-like structure is lost. Perhaps this is because nectary tissue here has no boundary due to the absence of reproductive whorls (Figure 4.11B). These results are consistent with those observed in *Arabidopsis* and *Petunia* in that *CRC* expression is dependent on both *AG* and *SHP* lineages (Morel et al., 2018). It is striking that the upstream regulators are likely shared between these three taxa. However, like Morel et al. (2018) our data cannot distinguish whether this shared regulation is due to a single evolutionary origin of nectaries or due to the conservation of *CRC* in carpel development.

Intriguingly, when *CvTCP1* is downregulated in *C. violacea*, nectaries are altered with phenotypes ranging from reduced lobes to complete absence (Chapter 5). Like *CRC*, the regulatory pathway upstream of *TCP1* is unclear, although the key contribution of TCP

homologs towards many types of floral monosymmetry has been demonstrated across angiosperms (Preston and Hileman, 2009; Hileman, 2014b; Wessinger and Hileman, 2020). Given that expression domains of *AG* and *SHP* are much broader across the flowers, other genetic factors are required for restriction of nectaries to a single whorl (Morel et al., 2018). In *C. violacea*, *CvTCP1* may be involved, at least indirectly. As noted above, functional data for floral nectaries to date has been conducted on flowers whose nectaries are distributed evenly around floral organs (e.g., circular around *Petunia* gynoecium and at the base of all stamens in *Arabidopsis*), unlike the adaxial positioning of the nectary in *C. violacea*. Thus, adaxial floral identity may be required for nectary formation in *C. violacea*, although it is unclear if *TCP1* has a direct role in nectary initiation. Functional studies of Cleomaceae with annular nectaries, such as *Tarenaya hassleriana*, would inform on decoupling nectary position and identity in the family.

Less is known about nectary size than initiation. In *Petunia*, *BEN* and *ROB* are important for nectary size (Morel et al., 2018), whereas *BOP1/2* impacts nectary size in *Arabidopsis* (Mckim et al., 2008). It is perhaps unsurprising that no *BEN* or *ROB* homologs were expressed in nectaries of *C. violacea*, but *CvBOP2* is expressed throughout all stages examined (Table 4.3 and Figure S4.2). Unlike *BEN* and *ROB*, the interactions between *BOP1/2* and other floral homeotic genes, with regards to nectary formation and size, are not as well understood (Slavkovic et al., 2021). Further experiments are needed to determine whether *CvBOP2* contributes to nectary size in *C. violacea*. In our analysis of highly expressed transcripts (Figure S4.5 and Table S4.4), six transcripts are potentially linked to cell growth in nectaries, although they have only been characterized in leaves (e.g., *EXL2*) and roots (e.g., *PRX44*) (Schröder et al., 2009; Marzol et al., 2022). Future studies should explore genes similar to those identified in this study, as well as earlier stages of nectary development. Altogether, these expression patterns suggest that pathways determining nectary size are not conserved across the core Eudicots.

Gene expression data suggests additional conservation as well as deviation in the genetic pathway of nectary development between *Arabidopsis* and *C. violacea*. Transcriptomic data shows many genes important for nectary formation are conserved across *Arabidopsis* and

Cleome, including ABC genes *AG*, *AP2*, *AP3*, *PI,* and MADS-box gene *SHP* (Table 4.3 and Figure S4.2). Notably, *AqSTY* has been shown as essential for nectary formation in *Aquilegia* (Min et al., 2019) and *CvSTY* is expressed in nectaries of *C. violacea*. While expression is low, it is significantly differentially expressed and down regulated in pre-anthetic flowers (Table 4.3 and Figure S4.2). This co-expression presents a tantalizing hypothesis that *CvSTY* and *CvCRC* are not mutually exclusive pathways in *C. violacea* nectary development. *STY* likely interacts with *CRC* in developing carpels of *Arabidopsis* (Kuusk et al., 2002) such that interactions in other floral structures are feasible. In addition, *STY* is also linked to auxin biosynthesis (Baylis et al., 2013), which is important to nectary development.

4.4.3 Nectar of Cleome violacea *is complex, as is its secretion method*

Nectar is a multifaceted sugar solution that changes in composition over time and includes microorganisms as well as secondary metabolites made by both plant and microbes (Alvarez-Perez et al., 2012; Chappell and Fukami, 2018; Parachnowitsch et al., 2019; Liao et al., 2021; Jacquemyn et al., 2021). Our data are consistent with bacteria and fungi colonization of *C. violacea* nectaries (Figure 4.7) and reflect complexities in these interactions. Many of the identified microorganisms from this study are commonplace in soil and/or have been previously isolated from nectar (e.g., *Sphingomonas*, *Pseudomonas*, and *Erythrobasidium*) (Figure 4.7) (Alvarez-Perez et al., 2012; Jacquemyn et al., 2013). However, the exact nature of the relationship (i.e., mutualism, commensalism, or parasitism) cannot be determined with gene expression data alone, especially since there was variation across replicates. Nonetheless, we found six transcripts that potentially play a role in combating biotic stress from our analysis of highly expressed transcripts (Figure S4.5 and Table S4.4), e.g., *LIPID TRANSFER PROTEIN 2* (*LTP2*) and *β-GLUCOSIDASE 19* (*BGLU19*) (Molina and García-Olmedo, 1997; Li et al., 2019) KEGG counts also showed enriched plant-pathogen interactions (26, 25, and 26 in pre-anthetic, anthetic, and post-anthetic nectaries, respectively.) (Table S4.5). Further, compounds typically produced by nectar-associated microbial communities (e.g., alcohols, isoprenoids, and ketones) (Rering et al., 2018) are difficult to distinguish with transcriptomics because many of these metabolites are also produced by the plant. Additionally, yeasts are known to chemically alter metabolites already present in nectar (Vannette and Fukami, 2016).

In all stages of developing nectaries of *C. violacea*, we found roughly even KEGG counts of carotenoid and flavonoid biosynthesis (~14 and ~8 across stages, respectively) (Table S4.5). Flavonoids have antioxidant activity, which reduces reactive oxygen species (ROS) in *Arabidopsis* (Vannette and Fukami, 2016), and they have also been linked to the reduction of *E. coli* fimbria, which may reduce biofilm formation (Lee et al., 2011). *FLS1* is significantly upregulated in anthetic and post-anthetic nectaries (Table 4.1 and Figure S4.2). Additionally, accumulation of flavanols have also been shown to increase survival of yeast by reducing oxidative stress (Naparlo et al., 2019). Thus, flavonoid accumulation may be a way to inhibit bacterial biofilms while simultaneously supporting symbiotic yeast. *CHS*, which is highly expressed in our transcriptome (Table 4.1, S4.4 and Figure S4.2), is also linked to resistance against biotic and abiotic stress such as UV, temperature, wounding, and bacteria (Dao et al., 2011). However, even though their role in the reduction of ROS can potentially impact biotic stress, carotenoids are more commonly linked to abiotic stress (Havaux, 2014), pollinator attraction (Cazzonelli, 2011), and photoprotection (Demmig-Adams, 1990), so further research is required.

Phytohormone expression in *C. violacea* is complex with evidence supporting convergence to other Eudicots. In Brassicales, phytohormones play an important role in gland development and nectar secretion (Slavkovic et al., 2021). In both *Aquilegia* and *Arabidopsis*, auxin is linked to nectary initiation via *ARF6* and *ARF8* (Nagpal et al., 2005; Reeves et al., 2012) to nectar production via *PIN6* (Bender et al., 2013). *Cleome violacea* nectaries express multiple ARFs and PINs across development (Figure 4.6A), although not all expression is identical to that in *Arabidopsis* (Tables 4.2, 4.3). All PINs serve to promote the flow of auxin between cells (Křeček et al., 2009), so there is likely conservation of function between Cleomaceae and Brassicaceae. Auxin however does not function alone, and gland development is complicated by phytohormone interactions. For example, JA is positively and negatively regulated by auxin and GA, respectively. Both GA (Wiesen et al., 2016) and JA (Radhika et al., 2010) are linked to nectar secretion (Slavkovic et al., 2021), e.g., JA is positively correlated with nectar production. For other biological processes (e.g., seedling development) there is substantial crosstalk between auxin and ethylene (Muday et al., 2012). To our knowledge there have been no studies to date

that characterize ethylene function in nectaries, although few do show ethylene-related genes present in nectary tissues (Tang et al., 1994) or ethylene production with *CRC* promoters (Switzenberg et al., 2015). Our transcriptome has multiple ethylene-related genes that are expressed across all developmental stages, *ETHYLENE INSENSITIVE 4* (*EIN4*) and *ETHYLENE INSENSITIVE-LIKE 3* (*EIL3*) (Figure 6C). Further, *CHITINASE LIKE 1* (*CLK1*), which modulates ethylene biosynthesis in root development, is among the highest expressed transcripts (Table S4.4). However, it is yet unclear what role these genes play, and whether they have a function unique to nectaries.

4.4.4 Conclusions

As no explicit ancestral reconstruction states of nectaries have been performed in Brassicales or core Eudicots, it remains unknown whether nectaries in Cleomaceae and Brassicaceae represent a single or independent origin of nectaries. The data presented in this study demonstrate a high degree of conservation between Cleomaceae and Brassicaceae, which would be consistent with a single origin of nectaries in these sister lineages. *CvCRC* functions as it does in *Arabidopsis* and is regulated redundantly by MADS-box genes *AG* and *SHP*. *Cleome violacea* nectaries are eccrine-based and appear to regulate their own energy production. Given multiple origins of other traits (e.g., monosymmetry), we cannot exclude the possibility of independent recruitment in the roles of *CRC*, *AG*, *SHP* and *SWEET9* for nectary development and nectar secretion, respectively. Research on the evolution and development of nectaries and on nectar biology is ripe for interdisciplinary research (Liao et al., 2021). Here we show that *Cleome violacea* is a promising model for nectary development in the Cleomaceae that will pave the way forward for future nectary research on other key factors such as morphology and pollination.

4.5 Figures

Figure 4.1. *Cleome violacea* flowers at various stages of development. (A) Large undissected floral bud. (B) Large dissected floral bud showing nectary. (C) Newly anthetic flower. (D) Post-anthetic flower with developing fruit. (E) Magnified view of anthetic nectary. Scale bars = 1 mm.
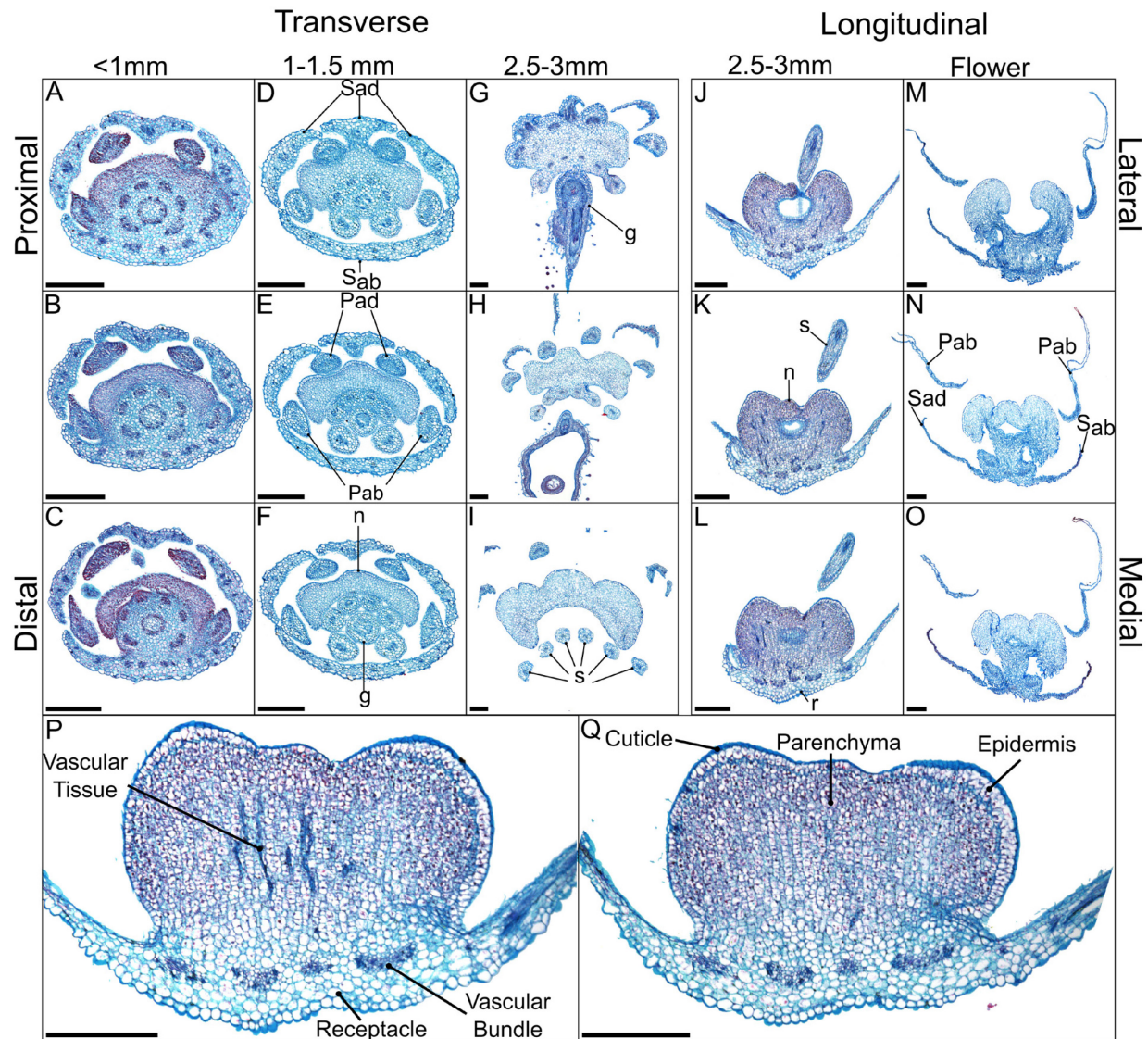
Figure 4.2. Alcian blue/safranin O-stained sections of *Cleome violacea* nectaries at pre-anthetic, anthetic and post-anthetic stages. From left to right: (A-C) small, (D-F) medium, and (G-I) large buds in transverse view with proximal-distal indicating relative distance to receptacle. (J-L) Large bud and (M-O) flowers in longitudinal view with lateral-medial indicating relative distance from center. (P-Q) Longitudinal view of 8 μm sections of the same large floral bud with and without vascular tissue, respectively. Scale bars = 250 μm. Sad = adaxial sepal; Sab = abaxial sepal; Pad = adaxial petal; Pab = abaxial petal; s = stamen; g = gynoecium, r = receptacle.
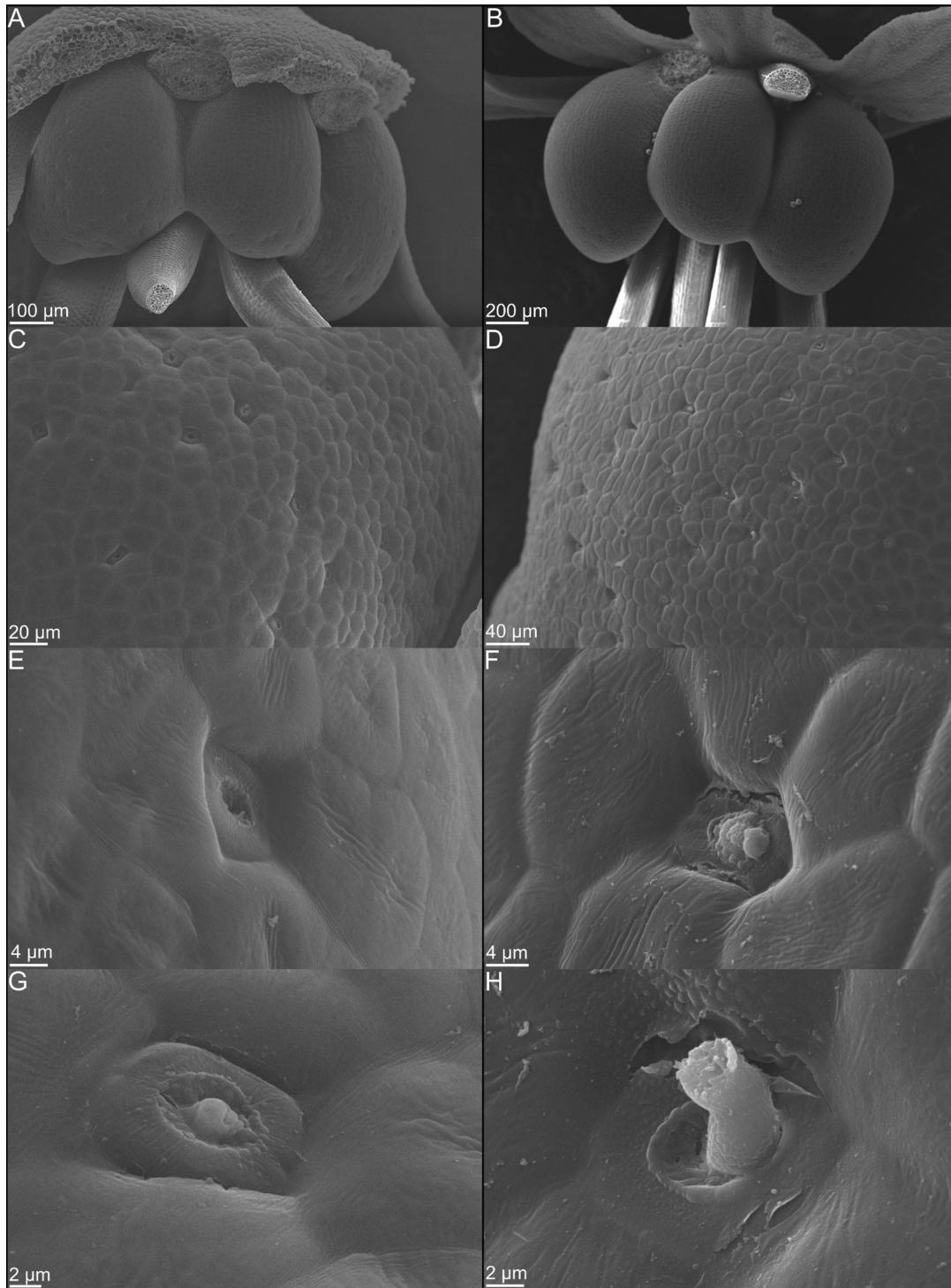
Figure 4.3. Scanning electron micrographs of whole nectaries from *Cleome violacea* at (A) pre-anthetic and (B) anthetic stages. (C) Distribution of nectarostomata on pre-anthetic nectary lobe and (D) anthetic nectary lobe. Examples of nectarostomata from (E-G) bud and (F-H) anthetic flowers.
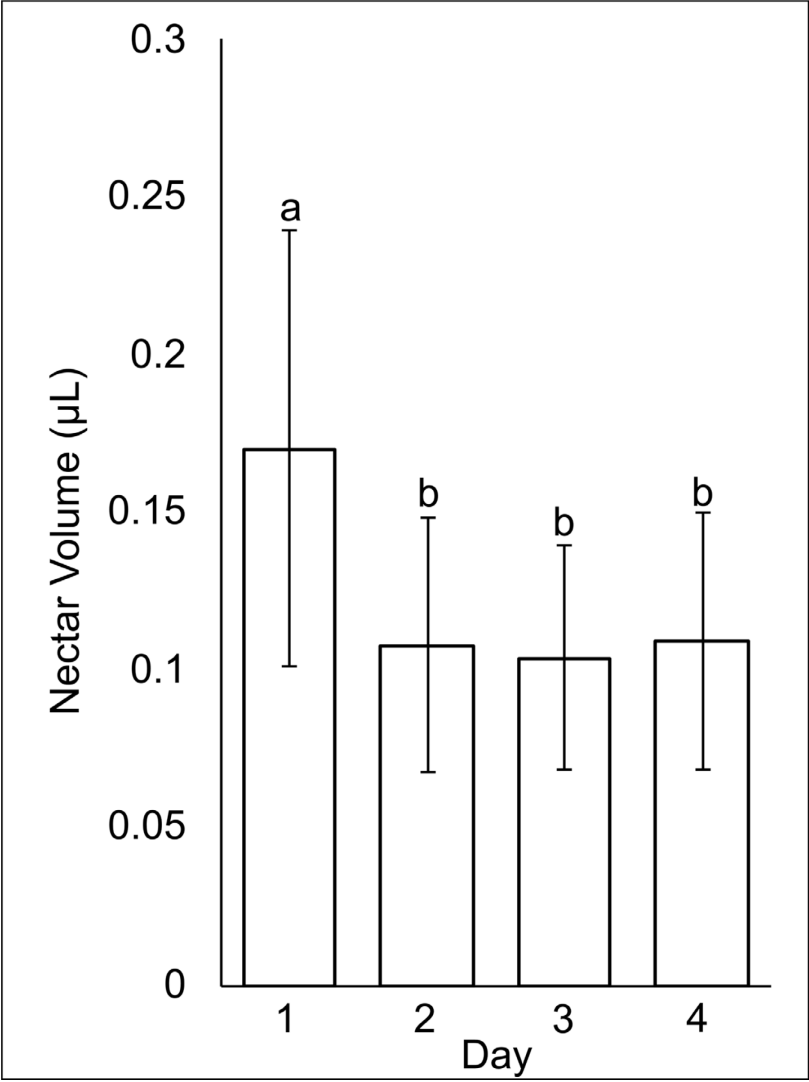
Figure 4.4. Nectar volume from *Cleome violacea* flowers taken on first day of anthesis (Day 1) and three days post-anthesis (Days 2, 3, and 4). Averaged value of flowers from 20 plants. Significance measured using paired, one-tailed, student's t-tests (a < 0.01).
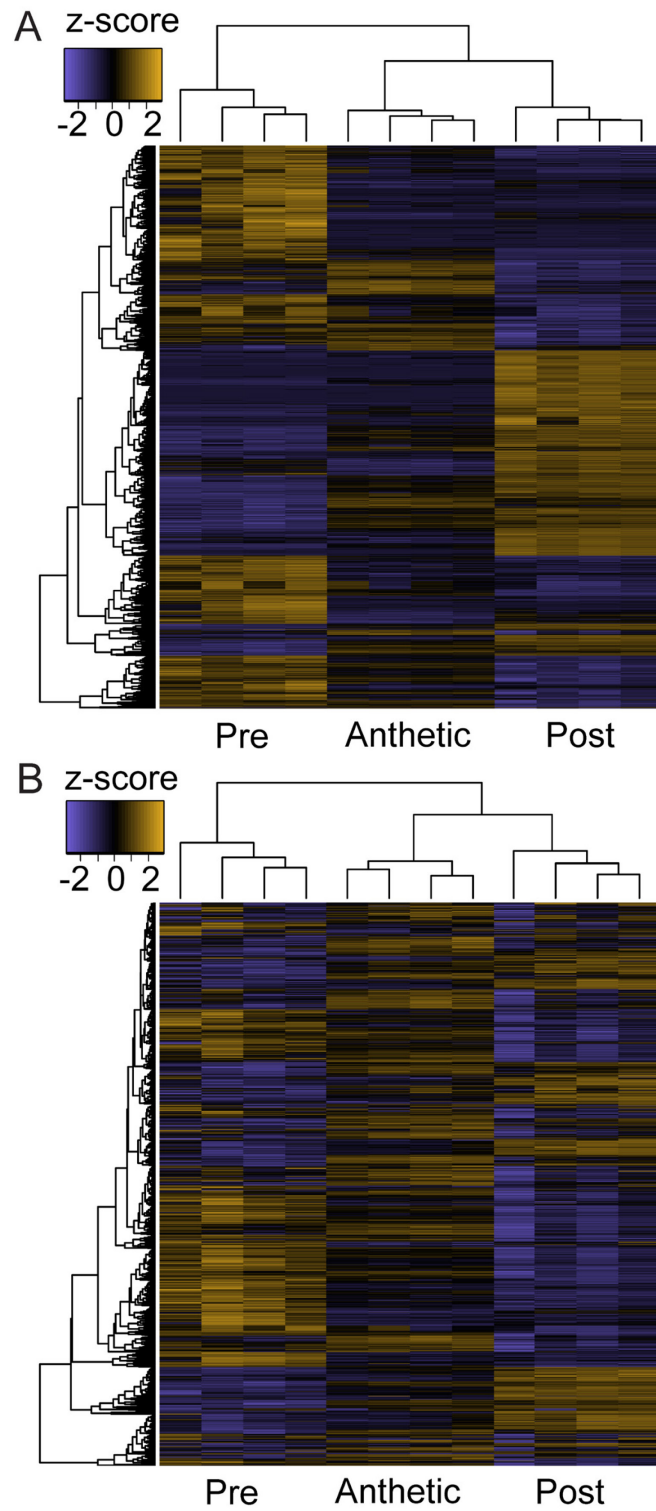
Figure 4.5. Z-score heatmaps of (A) all differentially expressed transcripts and (B) transcripts with TPM > 100 from *Cleome violacea* pre-anthetic, anthetic, and post-anthetic nectaries. TPM = Transcripts per million.
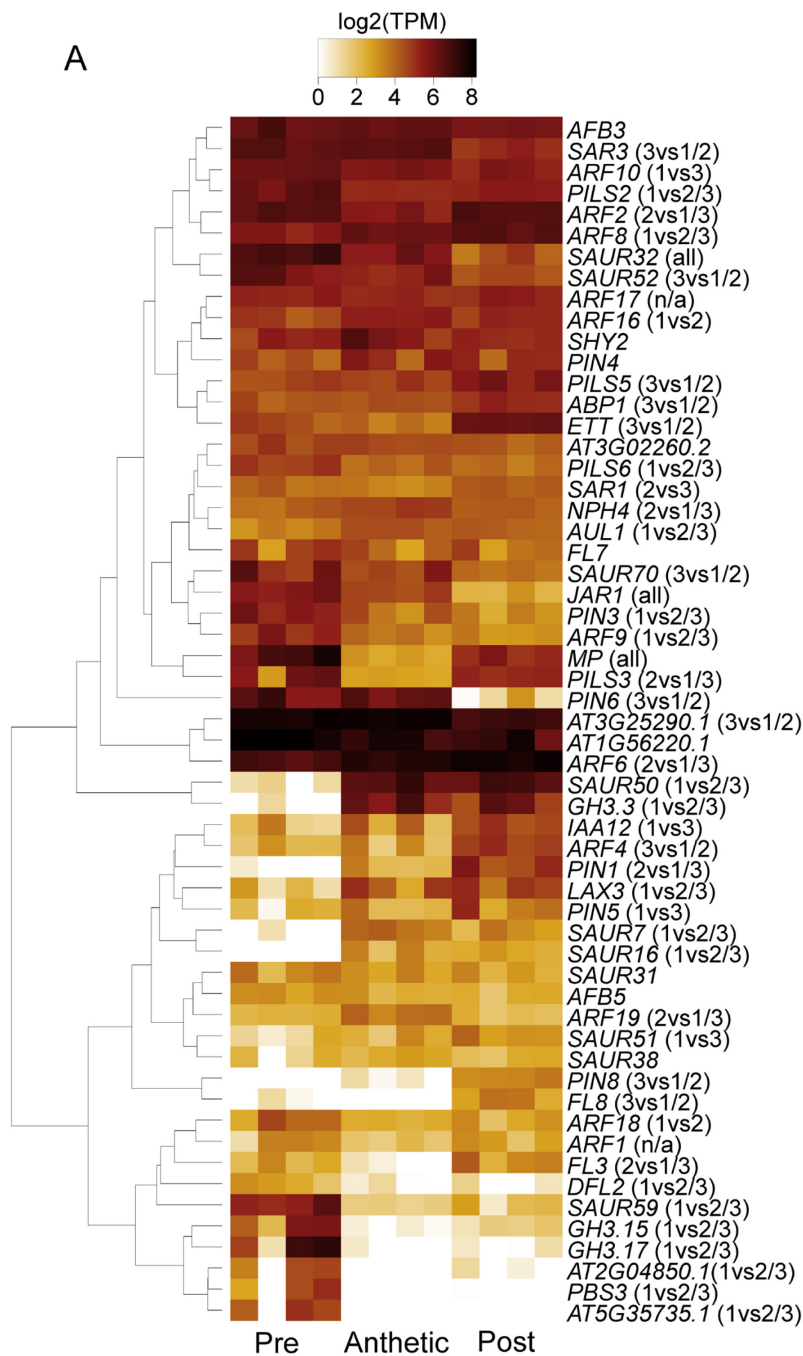
Figure 4.6A. A heatmap of phytohormone-related transcripts expressed in pre-anthetic, anthetic, and post-anthetic nectaries of *Cleome violacea* displayed in log2(TPM). Representative transcripts of genes related to (A) auxin, (B) gibberellic acid (C) jasmonic acid, and (D) ethylene. Significant 1 to 1 differences displayed in brackets. 1 = pre-anthetic; 2 = anthetic; 3 = post-anthetic.
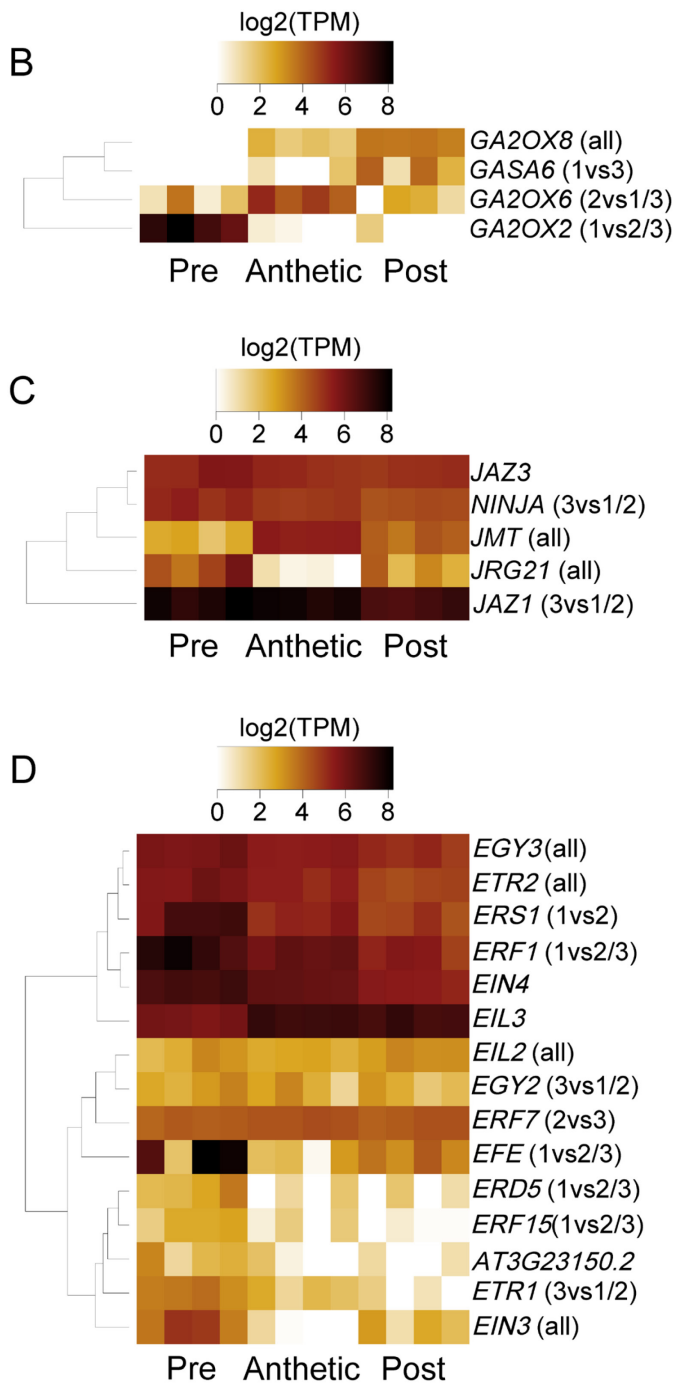
Figure 4.6B-D. A heatmap of phytohormone-related transcripts expressed in pre-anthetic, anthetic, and post-anthetic nectaries of *Cleome violacea* displayed in log2(TPM). Representative transcripts of genes related to (A) auxin, (B) gibberellic acid (C) jasmonic acid, and (D) ethylene. Significant 1 to 1 differences displayed in brackets. 1 = pre-anthetic; 2 = anthetic; 3 = post-anthetic.
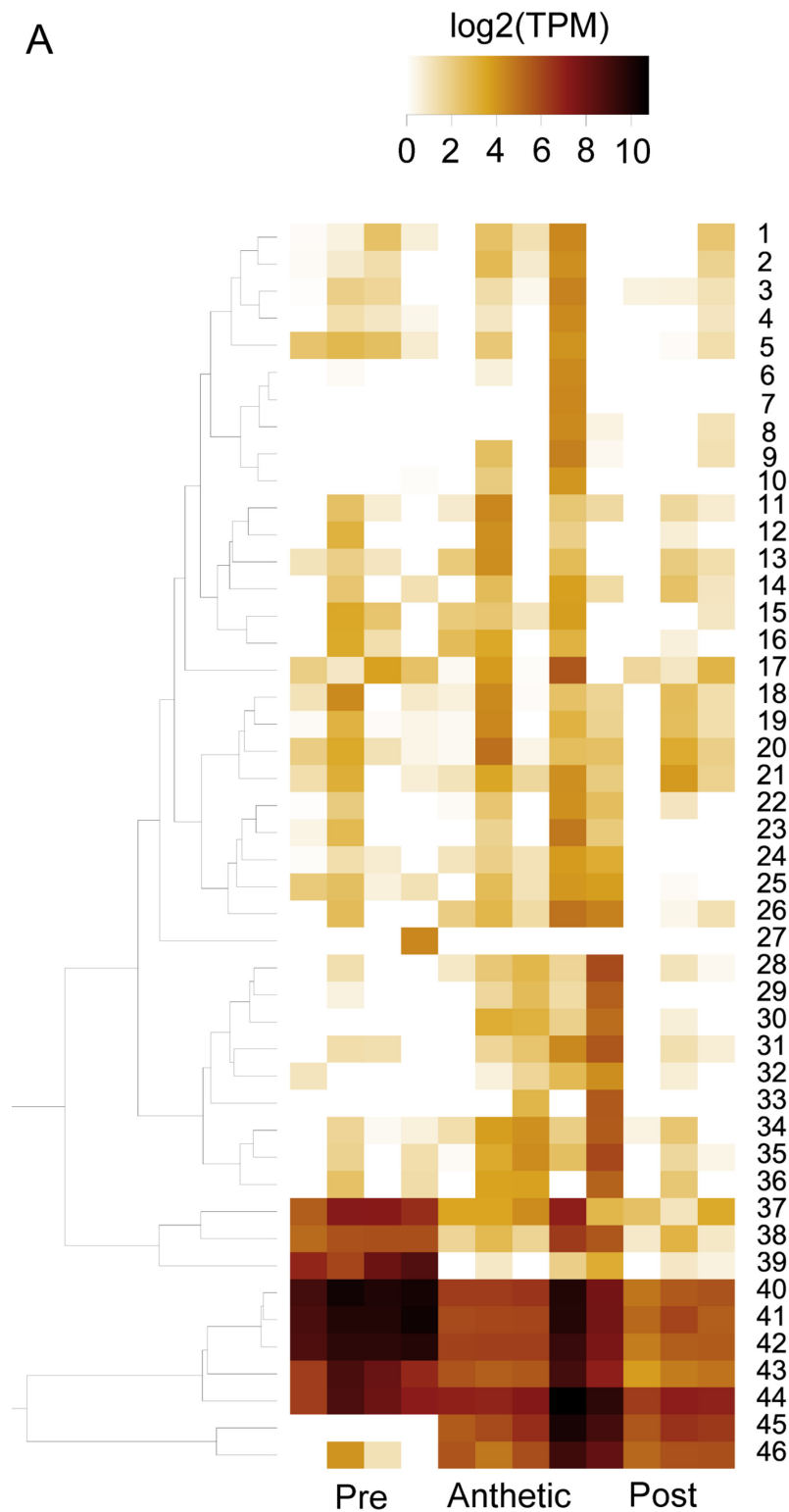
Figure 4.7A. Heatmap of (A) 16s bacterial rRNA and (C) 18s fungal rRNA related transcripts in pre-anthetic, anthetic, and post-anthetic nectaries of *Cleome violacea* displayed in log2(TPM). Genera and NCBI accession of respective transcripts for (B) bacteria and (D) fungi.

B

#	Genus	Accession ID
2;18	*Acinetobacter*	nr_102814.1/nr_148847.1
24;27		nr_115299.1/nr_026208.1
38;44	*Corynebacterium*	nr_026380.1/nr_029292.1
33	*Curvibacter*	nr_024702.1
8;13;20	*Cutibacterium*	nr_040847.1
28;35	*Cylindrospermum*	nr_102462.1
15	*Dokdonella*	nr_042397.1
22	*Domibacillus*	nr_108861.1
32	*Ehrlichia*	nr_121714.1
25;37	*Exiguobacterium*	nr_043478.1
5	*Insolitispirillum*	nr_044314.1
45	*Kribbella*	nr_025735.1
23	*Microbacterium*	nr_044934.1
43	*Microcoleus*	nr_172606.1
10	*Neochroococcus*	nr_172603.1
3;17	*Paracoccus*	nr_024658.1/nr_025857.1
21;34;40		nr_025714.1/nr_157668.1
16	*Prolinoborus*	nr_104948.1
26	*Propionibacterium*	nr_151943.1
4	*Propionivibrio*	nr_025455.1
1;6;11	*Pseudomonas*	nr_041715/116651/134795
14;29;42		nr_026395/024704/041715
9;36	*Schlegelella*	nr_043802.1
46	*Sphingomonas*	nr_108991.1
12	*Stella*	nr_025582.1
7;19;30	*Tepidimonas*	nr_042418.1
31;39;41		nr_109514.1

Figure 4.7B. Heatmap of (A) 16s bacterial rRNA and (C) 18s fungal rRNA related transcripts in pre-anthetic, anthetic, and post-anthetic nectaries of *Cleome violacea* displayed in log2(TPM). Genera and NCBI accession of respective transcripts for (B) bacteria and (D) fungi.
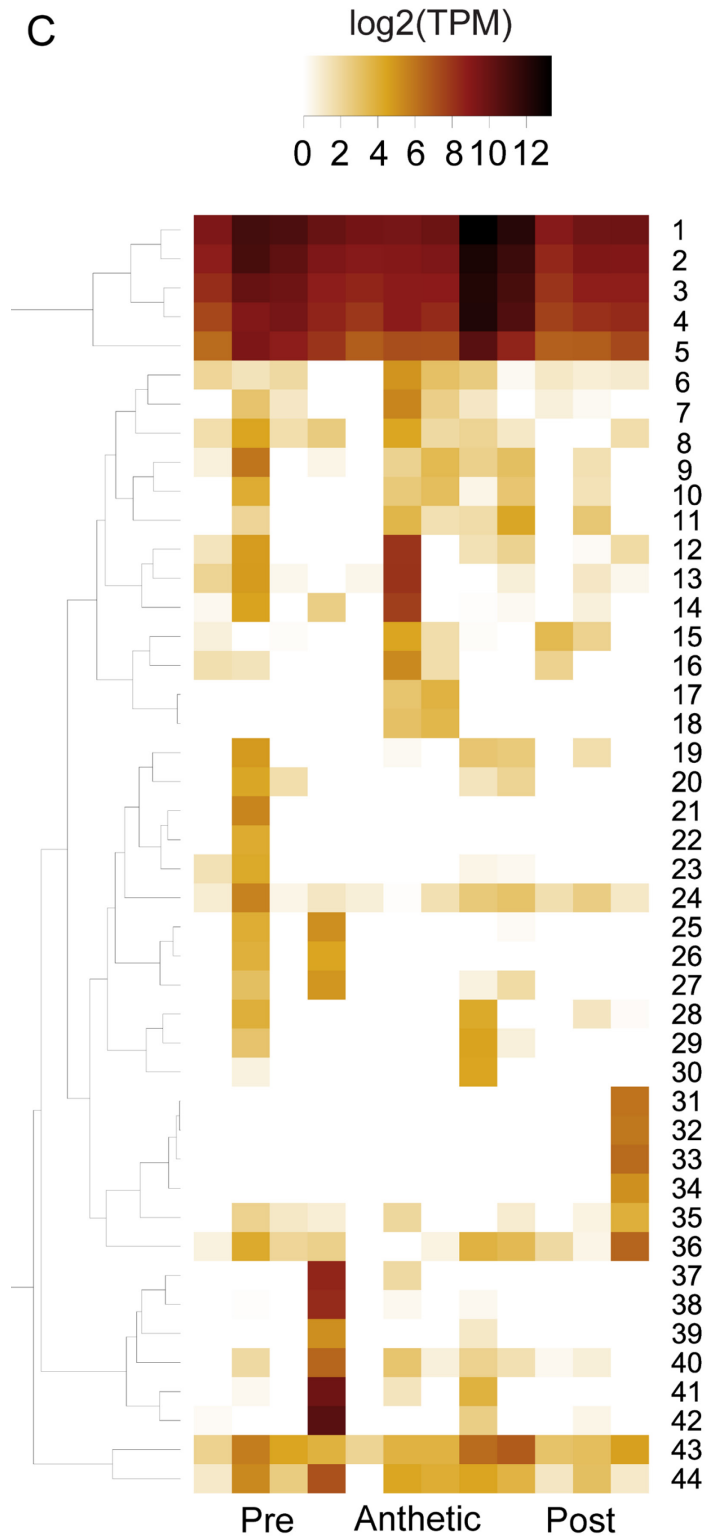
Figure 4.7C. Heatmap of (A) 16s bacterial rRNA and (C) 18s fungal rRNA related transcripts in pre-anthetic, anthetic, and post-anthetic nectaries of *Cleome violacea* displayed in log2(TPM). Genera and NCBI accession of respective transcripts for (B) bacteria and (D) fungi.

D

#	Genus	Accession ID
19	*Acremonium*	ng_062813.1
9;34	*Amycosphaerella*	ng_061084.1
41	*Ascochyta*	ng_061064.1
7;15;17;39;42	*Aspergillus*	ng_063229.1/ng_074950.1
8;14;23;30;40;43	*Brunneoclavispora*	ng_061284.1
13	*Candida*	ng_063392.1
1	*Chytridium*	ng_061075.1
37	*Cutaneotrichosporon*	ng_062024.1
5	*Erythrobasidium*	ng_063520.1
36	*Hyaloraphidium*	ng_017172.1
18	*Hyphozyma*	ng_070854.1
31	*Infundichalara*	ng_061112.1
27	*Leiothecium*	ng_061073.1
38	*Lentithecium*	ng_016507.1
10;28	*Loramyces*	ng_062693.1
21	*Malassezia*	ng_062732.1
3;12;20;44	*Mortierella*	ng_070277/070287/070251
6;25;35	*Olpidium*	ng_017176.1
29	*Penicillium*	ng_074879.1
26	*Phaeotremella*	ng_063469.1
22	*Ramicandelaber*	ng_070281.1
2;4	*Rhexocercosporidium*	ng_061015.1
16	*Rozella*	ng_017174.1
33	*Symbiotaphrina*	ng_070853.1
11	*Teratosphaeria*	ng_065572.1
24	*Westerdykella*	ng_016502.1
32	*Xenoacremonium*	ng_061110.1

Figure 4.7D. Heatmap of (A) 16s bacterial rRNA and (C) 18s fungal rRNA related transcripts in pre-anthetic, anthetic, and post-anthetic nectaries of *Cleome violacea* displayed in log2(TPM). Genera and NCBI accession of respective transcripts for (B) bacteria and (D) fungi..

Figure 4.8. *Cleome violacea* flowers from untreated and treatment control groups. (A) Untreated newly anthetic flower and (B) maturing flower. pTRV2-MCS treated flower displaying (C) moderate and (D) mild viral phenotype. (E, F) pTRV2-*CvANS* treated flowers displaying moderate yellowing petal phenotypes. Scale bars = 1 mm.

Figure 4.9 Flowers of *Cleome violacea* treated with pTRV2-*CvCRC-CvANS* constructs. (A) Flower with strong yellowing phenotype and no nectary. (B) Flower with moderate yellowing phenotype and no nectary. (C) Flower with moderate yellowing phenotype, no nectary, and enlarged gynoecium. (D) Flower with moderate yellowing phenotype and no nectary. (E) Flower with half normal and half yellowing petals with partially absent nectary. (F) Flower with strong yellowing phenotype and reduced lateral nectary lobes. Scale bars = 1 mm.

Figure 4.10. Flowers of *Cleome violacea* treated with pTRV2-*CvSWEET9-CvANS* constructs. (A) Flower with moderate yellowing and nectary with reduced nectar accumulation. (B) Magnified view of nectary in A. (C) Flower with partial yellowing and partial normal phenotype. (D) Flower with near-normal pigmentation and reduced nectar production. (E) Magnified nectary from C displaying decreased nectar accumulation correlating with yellowing phenotype.

Scale bars = 1 mm.

Figure 4.11. Flowers of *Cleome violacea* treated with pTRV2-*CvAG* constructs. (A) Flower with repeating perianth whorls. (B) Nectary from flower similar to A with petals removed. (C) Flower with normal adaxial petals, repeating perianth whorls and adaxial nectary. (D) Flower with repeating perianth whorls and distally positioned nectary. (E) Flower with petaloid stamens and adaxial nectary. (F) Flower with repeating perianth whorls and adaxial nectary. White arrowheads indicate nectary position. Scale bars = 1 mm.

Figure 4.12. Flowers of *Cleome violacea* treated with pTRV2-*CvAG-CvSHP* constructs. (A) Flower with partial nectary. (B-E) Flowers with repeating perianth whorls and no nectary. (F) Flower with repeating perianth whorls and partial nectary. Black and white arrowheads represent reduced and absent nectary, respectively. Scale bars = 1 mm.
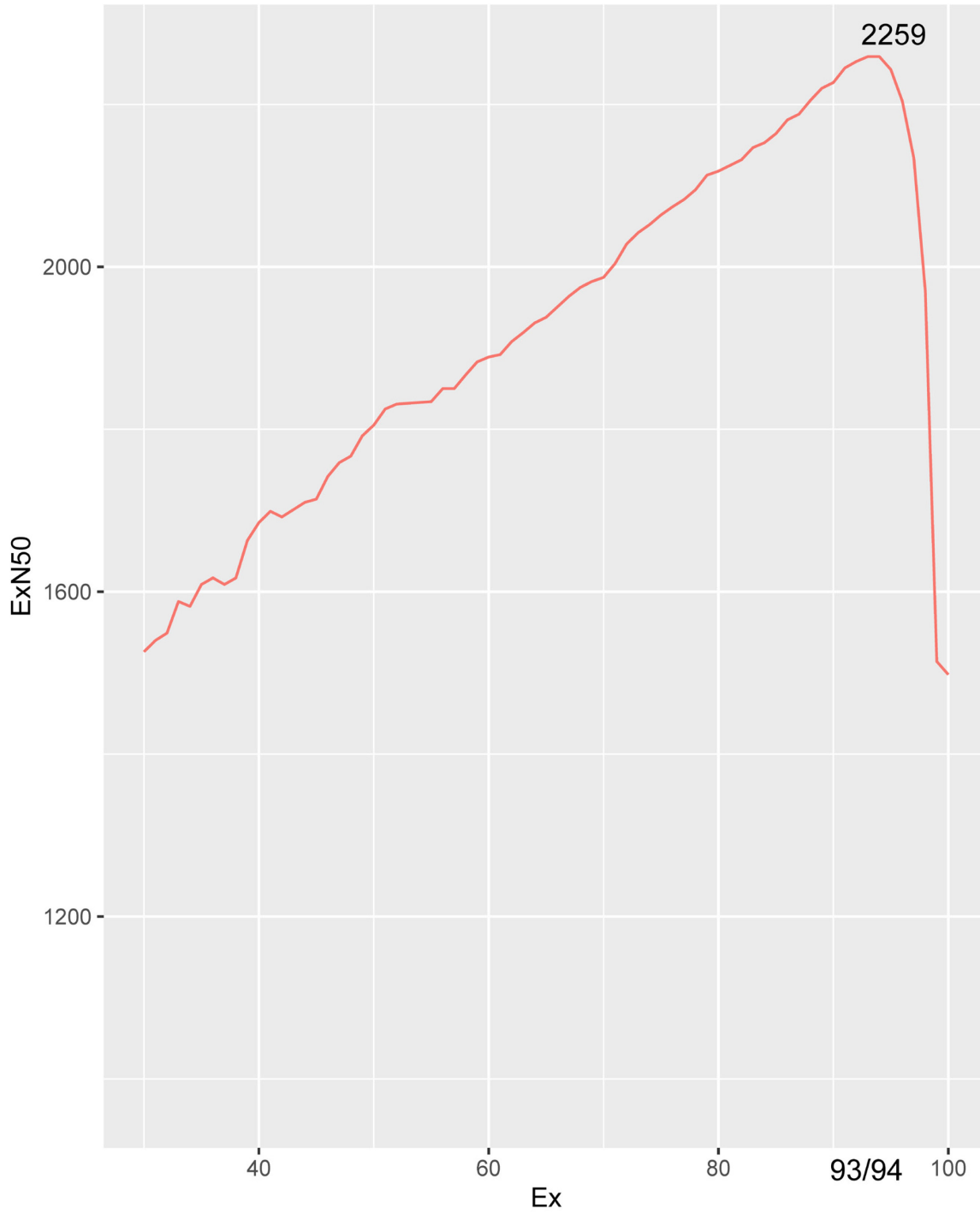
Figure S4.1. ExN50 graph generated by the 'contig_ExN50_statistic.pl' and 'plot_ExN50_statistic.Rscript' scripts provided with Trinity.
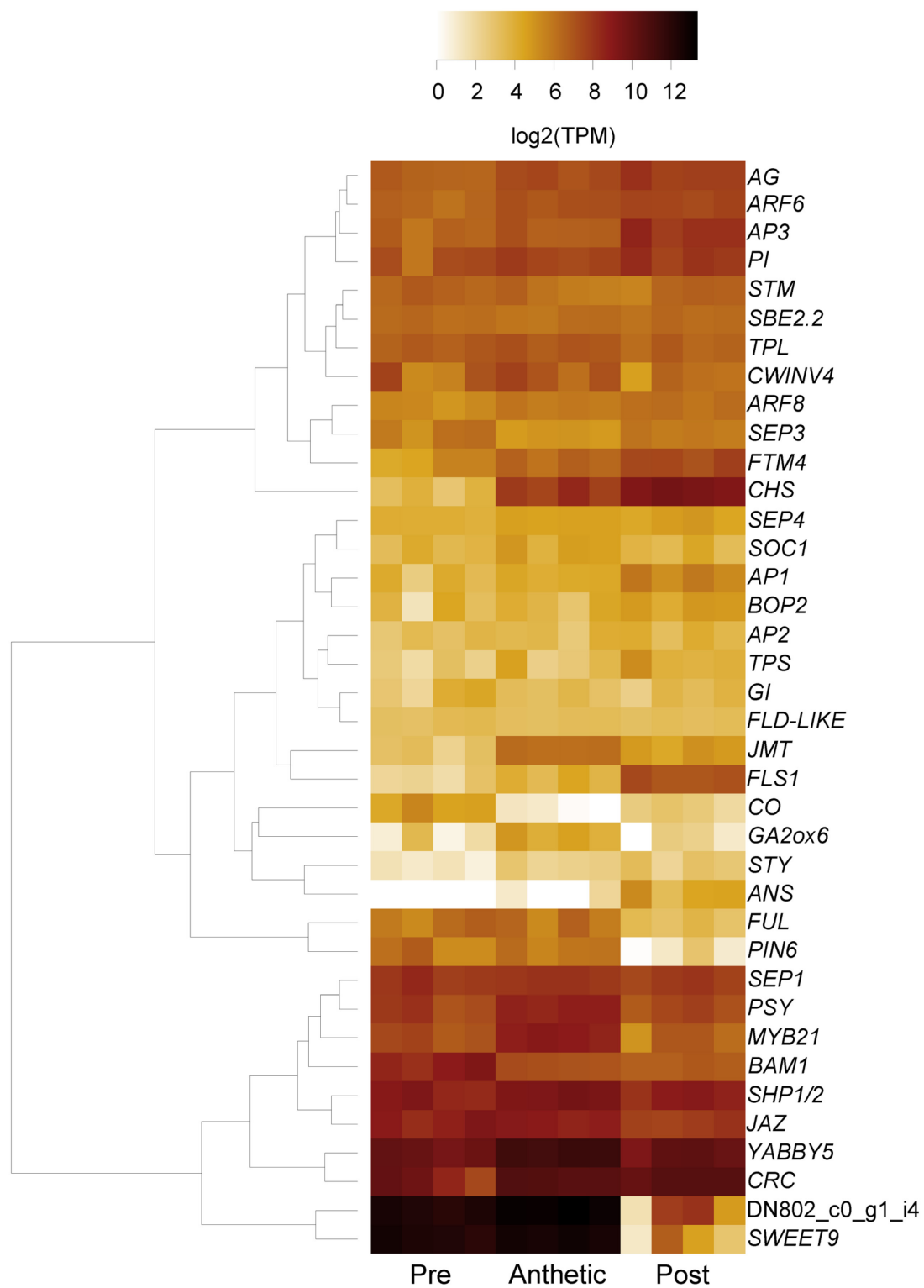
Figure S4.2. A heatmap of nectary-related genes displayed as log2(TPM).

```
         1              10                  20                  30                  40
         AATAAGAAGGCTTTGGTGTCACCAGTCCCAAAGCAGAGGGTTCGGCGAT
50                 60                  70                  80                  90
CTGATTAACTGAGCTACGTCGTCGTTTAGTCAAGCCTCTCGGTGTATAAA
100                110                 120                 130                 140
TACCCACAAACTCATAATACTTATATCATCACTCACAAACATACATTGTT
150                160                 170                 180                 190
ATTTGCATTCTATAGAAGAGAAGAGATCAACAATGGCCTCAAAGAGCTCG
200                210                 220                 230                 240
AACACCACAATGCCTCTTCTCATTGCACTCAACATCCTTGTCCTCGGCCT
250                260                 270                 280                 290
CATCAGCGAAGGCGTCCTCGCCTGGTCGCCTGCGCCCCAGCCGACCCCGC
300                310                 320                 330                 340
CCCCGTCCTAACCGTTGGGACATGCCCCAAGAATACCCTCCAACTATTG
350                360                 370                 380                 390
CCATGCATCATTCTCCGCCTCGTCAGAATTACGGAGAACGACATGCGGGC
400                410                 420                 430                 440
GAGGCGGTGTTGTTCCGTCCTCAATGGTTTGCCTGATAACCAAGCAGCCG
450                460                 470                 480                 490
GTTGCCTCTGTACCGCGTTTAGACGCGGCGCTGCCAGAAACTCCATCACT
500                510                 520                 530                 540
GGAATCCTTAATACTTGCGGTCGGAAAATCACTCCACCGGGCTTCCAATG
550                560                 570                 580                 590
CCCATAAACAACTTCACTTCTTGCCCTAAAACCCTATCTTATATATGAAG
600                610                 620                 630                 640
AAACAAAATAAACACAAACTCACGTATAGTATACGATTGAGGGCGTTGTG
650                660                 670                 680                 690
TAATCTTGCATCTTTGTGAAAGTAATCATGAATAAGAATGAGTATTGGCT
700                710                 720                 730          740      744
TTTTAAAAAAAAAAAAAAAAAAAACGGAAGAGGGGCGGGGAGGGAA
```

Figure S4.3. Sequence and ORF of uncharacterized Trinity transcript DN802_c0_g1_i4.The ORF is highlighted and begins at bp 182.

Figure S4.4. Flowers of *Cleome violacea* treated with pTRV2-DN802_c0_g1_i4-*CvANS* constructs. Phenotypes were indiscernible from pTRV2-*CvANS* control. (A) Flower with mild yellowing and underdeveloped stamens. (B) Flower with strong yellowing and no nectar production. (C) Flower with mild yellowing and underdeveloped stamens. (D) Flower with moderate yellowing. (E) Flower with moderate yellowing and underdeveloped stamen. (F) Flower with moderate yellowing. Scale bars = 1 mm.
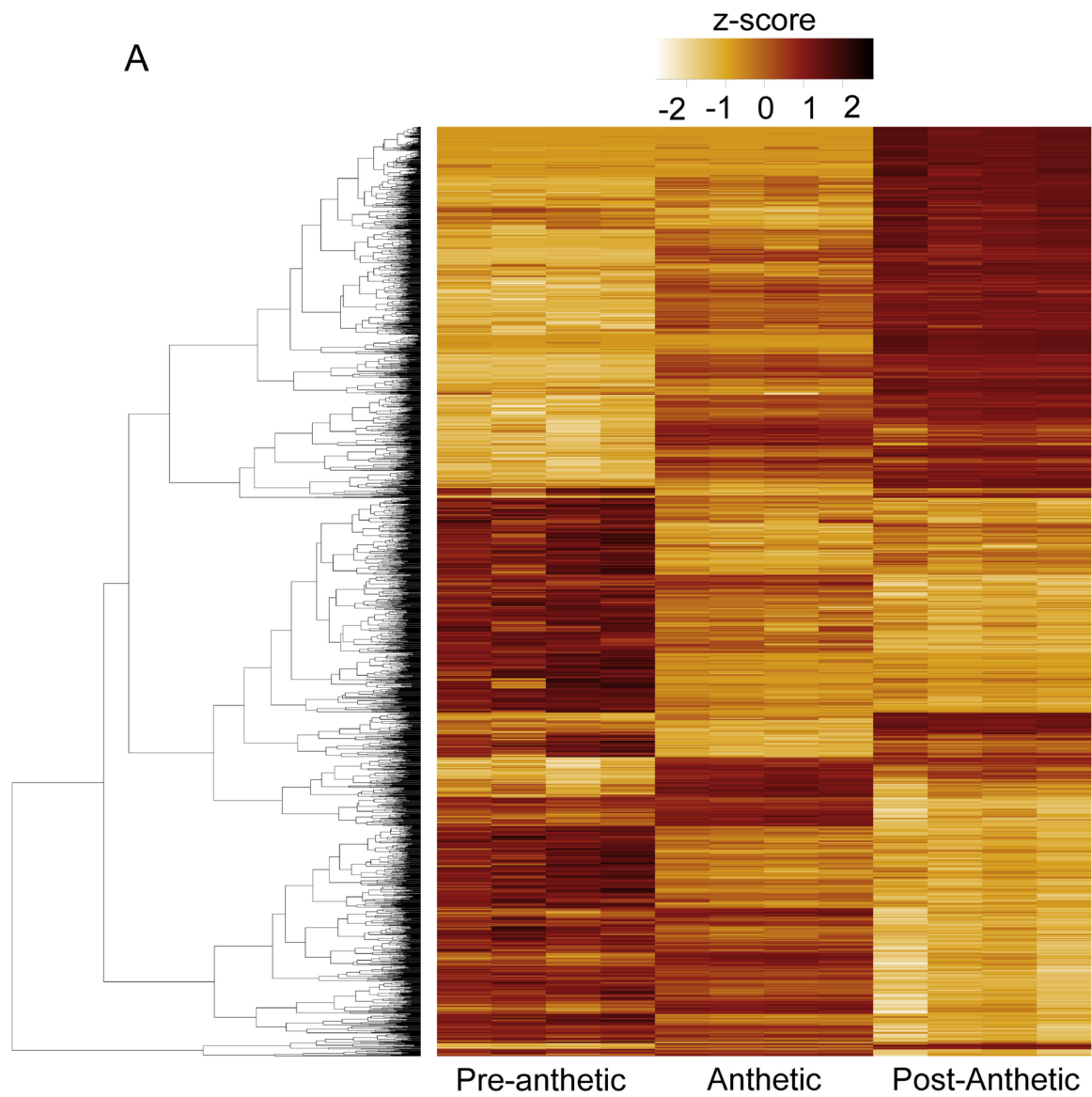
S4.5A. Z-score heatmaps of TransDecoder filtered transcripts. (A) All differentially expressed transcripts and (B) transcripts with TPM > 100 from *Cleome violacea* pre-anthetic, anthetic, and post-anthetic nectaries.
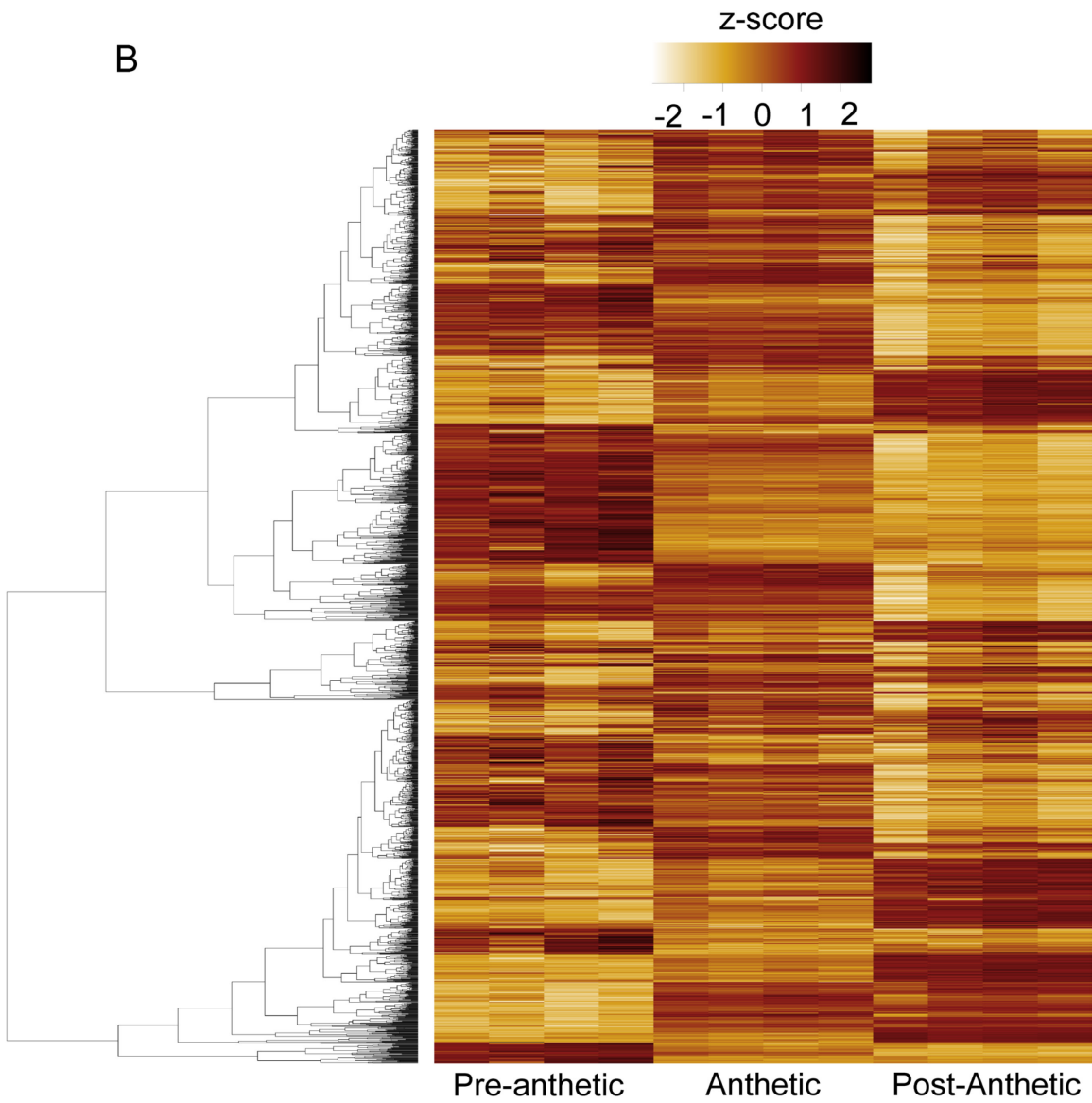
Figure S4.5B. Z-score heatmaps of TransDecoder filtered transcripts. (A) All differentially expressed transcripts and (B) transcripts with TPM > 100 from *Cleome violacea* pre-anthetic, anthetic, and post-anthetic nectaries.

4.6 Tables

Table 4.1. Genes of interest not directly implicated in nectary development with relative expression values from our transcriptomic dataset, putative roles, and relevant citations. SDE, Significant Differential Expression with arrows representing either up or down regulation of expression between developmental stages (≈ indicates no SDE). S1, pre-anthetic nectary; S2, anthetic nectary; S3, post-anthetic nectary.

Gene	SDE	Putative Role	Citation
TPL	≈	Co-repressor of *AG* which is recruited by *AP2* in floral organ identity	Krogan et al., 2012
STM	↓S2 vs S1/S3	Controls carpel development and requires the function of *AG*	Scofield et al., 2007
SEP3	↓S2 vs S1/S3	Functions in combination with B and C class genes to activate *CRC*	Lee et al., 2005
FTM4	↑S3 vs S2/S3	Encodes an intracellular LRR protein that interacts with *AG*	Torti et al., 2012
SEP4	≈	Functions in combination with B and C class genes to activate *CRC*	Lee et al., 2005
SOC1	≈	Functions together with *FUL* to promote development of inflorescence	Preston et al., 2011
AP1	↑S3 vs S1/S2	Regulates fatty acid biosynthesis with *CRC* in *Arabidopsis*	Han et al., 2012
GI	≈	Regulates miR172, which in turn regulates *BEN* and *ROB* in *Petunia*	Jung et al., 2007
FLD-LIKE	≈	Required for systemic acquired resistance in *Arabidopsis*	Sing et al., 2013
FUL	↓S3 vs S1/S2	Expressed in *Arabidopsis* nectaries from stage 9 to stage 14	Baum et al., 2001
CO	↑S1 vs S2/S3	Promotes *SOC1* and *FT*	Jung et al., 2007
SEP1	≈	Functions in combination with B and C class genes to activate *CRC*	Lee et al., 2005
YABBY5	↑S2 vs S1/S3	Can dimerize with *CRC* via the *YABBY* domain	Gross et al., 2018
TPS	↑S3 vs S1/S2	Levels of trehalose change in parallel with sucrose; regulates stomatal conductance and water use.	Lunn et al., 2014
FLS1	↑S3 vs S1/S2	Accumulation of flavanols may increase survival of yeast by reducing oxidative stress	Naparlo et al., 2019
PSY	↑S2 vs S1/S3	Carotenoid oxidation products function as a plant stress signal	Havaux 2014
CHS	↑S3 vs S1/S2	Linked to resistance of biotic stress	Dao et al.,2011

Table 4.2. Genes of interest with direct roles in nectar production with relative expression values from our transcriptomic dataset, putative roles, and relevant citations. SDE, Significant Differential Expression with arrows representing either up or down regulation of expression between developmental stages (≈ indicates no SDE). S1, pre-anthetic nectary; S2, anthetic nectary; S3, post-anthetic nectary.

Gene	SDE	Putative Role	Citation
CWINV4	≈	Hydrolyzes sucrose into fructose and glucose; knockouts do not produce nectar	Ruhlman et al., 2010
SBE2.2	≈	Involved in starch synthesis; upregulated early in development in ornamental tobacco.	Ren et al., 2007
BAM1	↑S1 vs S2/S3	Starch breakdown; upregulation at early stages in *C. pepo*	Solhaug et al., 2019
PIN6	↓S3 vs S1/S2	Expression level is positively correlated with nectar production in *Arabidopsis*	Bender et al.,2013
MYB21	↑S2 vs S1/S3	Induces negative feedback loop on jasmonate biosynthesis	Reeves et al., 2012
SWEET9	↓S3 vs S1/S2	Required for nectar secretion in *Arabidopsis*	Lin et al., 2014
JAZ	↓S3 vs S1/S2	Represses jasmonic acid (JA) signalling in a negative feedback loop	Chico 2008
G20X	↑S2 vs S1/S3	Inactivates gibberellic acid (GA), which increases expression of genes involved in nectar production	Wiesen et al., 2015
JMT	↑S2 vs S1/S3	Forms Methyl Jasmonate from JA; JA conjugates linked to increased nectar production.	Radhika et al., 2010

Table 4.3. Genes of interest with direct roles in nectary formation with relative expression values from our transcriptomic dataset, putative roles and relevant citations. SDE, Significant Differential Expression with arrows representing either up or down regulation of expression between developmental stages (≈ indicates no SDE). S1, pre-anthetic nectary; S2, anthetic nectary; S3, post-anthetic nectary.

Gene	SDE	Putative Role	Citation
AP3	↑S3 vs S1/S2	Downregulation disrupts nectary placement and nectar secretion in *Arabidopsis*	Baum et al., 2001
PI	↑S3 vs S1	Downregulation disrupts nectary placement and nectar secretion in *Arabidopsis*	Baum et al., 2001
AP2	≈	Downregulation disrupts nectar secretion in *Arabidopsis*	Baum et al., 2001
AG	↓S1 vs S2/S3	Redundantly activates *CRC* with *SHP1/2*	Morel et al., 2018
ARF8	↓S1 vs S2/S3	Affects nectary size and gene expression redundantly with *AUXIN RESPONSE FACTOR 6 (ARF6)* in *Arabidopsis*	Reeves et al., 2012
ARF6	↓S1 vs S2/S3	Affects nectary size and gene expression redundantly with *ARF8* in *Arabidopsis*	Reeves et al., 2012
BOP2	≈	Promote formation of nectary glands independent of *CRC*	McKim et al., 2008
SHP1/2	↑S2 vs S3	Redundantly regulates *CRC* with *AG*	Morel et al., 2018
STY	↓S1 vs S2/S3	Controls nectary development in *Aquilegia* independent of *CRC*	Min et al., 2019
CRC	≈	Essential but not sufficient for nectary formation in the core-eudicots.	Lee et al., 2005

Table S4.1. Software versions and references for programs used in this study.

Software	Version	References
Trim Galore	0.6.6	Krueger, 2012
FastQC	0.11.9	Andrews, 2010
Trinity	2.12	Haas et al., 2013
Corset	1.09	Davidson and Oshlack, 2014
edgeR	3.32.1	Robinson et al., 2009
R	4.0.5	R Core Team, 2013
BLAST+	2.2.31	Altschul et al., 1990
Geneious	11.09	Geneious 11.09 (https://www.geneious.com)
SiFi21	1.2.3	Lück et al., 2019
BUSCO	5.1.2	Simão et al., 2015
Transdecoder	5.0.1	https://github.com/TransDecoder

Table S4.2. Automatically generated Trinity assembly statistics.

	All Isoforms	Longest Isoform
Contig N10	5247	4641
Contig N20	4101	3584
Contig N30	3424	2945
Contig N40	2936	2472
Contig N50	2541	2032
Median contig length	991	368
Average contig	1478.87	865.39
Total assembled bases	212837382	62520918
Total trinity 'genes':	72246	
Total trinity transcripts:	143919	
Percent GC	43.71	

Table S4.3. BUSCO analysis using version 5.1.2 with the Viridiplantae_odb10 database. Analysis was run in transcriptome mode using hmmsearch version 3.1 and metaeuk version 6.a5d39d9.

Complete BUSCOs (C)	423
Complete and single-copy BUSCOs (S)	44
Complete and duplicated BUSCOs (D)	379
Fragmented BUSCOs (F)	2
Missing BUSCOs (M)	0
Total BUSCO groups Searched (n)	425
C:99.6%[S:10.4%,D:89.2%],F:0.4%,M:0%,n:425	

Table S4.4. Highest expressed transcripts in the *Cleome violacea* nectary transcriptome (n = 51); transcripts were taken from TransDecoder filtered lists of all significantly differentially expressed transcripts and overall highest TPM transcripts.

Name	Accession	Related Function of Homologues	Potential Role in *C. violacea* Nectaries	Citation
BGLU19	AT3G21370	Disease resistance in rice	Biotic stress	Li et al., 2019
APX1	AT1G07890	H2O2 scavenging	Biotic stress	Pnueli et al., 2003
LTP2	AT2G38530	Enhanced tolerance to *Pseudomonas*	Biotic stress	Molina and Olmedo 1997
SAG21	AT4G02380	Root development and biotic stress	Biotic stress	Salleh et al., 2012
GAMMA-VPE	AT4G32940	Resistance to *Botrytis*	Biotic stress	Baarlen et al., 2007
AT4G30170	AT4G30170	Oxidative stress response in sulfate-transporter mutant	Abiotic stress	Maruyama-Nakashita et al., 2003
CHS	AT5G13930	Accumulation of flavonoids and abiotic/biotic resistance	Abiotic/biotic stress	Dao et al.,2011
AT5G38100	AT5G38100	General response to water stress	Water stress	Chen et al., 2003
FER1	AT5G01600	Protection against oxidative stress	Abiotic stress	Ravet et al., 2009
ASMT	AT4G35160	Drought tolerance	Water stress	Zuo et al., 2014
ERD14	AT1G76180	Dehydration stress	Water stress	Kiyosue et al., 1994
HTA12	AT5G02560	Stomatal closure	Nectar secretion via nectarostomata	Xu et al., 2016
SPDS1	AT1G23820	Water stress	Water stress	Alcazar et al., 2006
AGT3	AT2G38400	Salt/ABA related water stress	Water stress	Bray 2004
AT1G62480	AT1G62480	Signal transduction related to NaCl stress	Water stress	Jiang et al., 2007
AT1G11910.2	AT1G11910	Draught tolerance	Water stress	Fernando et al., 2020

Table S4.4. Continued

AT5G10770	AT5G10770	Draught and salinity stress	Water stress	Shariatipour and Heidari et al., 2018
PSBR	AT1G79040	Photosystem II complex	Energy metabolism	Allahverdiyeva et al., 2007
ELIP1	AT3G22840	Photoprotection	Energy metabolism	Casazza et al., 2005
LHB1B1	AT2G34430	Photosystem II binding protein	Energy metabolism	Cheminant et al., 2011
LHCB2.1	AT2G05100	Photosystem II binding protein	Energy metabolism	Cheminant et al., 2011
FBN1B	AT4G22240	Thylakoid maintenance	Energy metabolism	Gámez-Arjona et al., 2014
YAB5	AT2G26580	Can dimerize with CRC	Interacts with CRC	Gross et al., 2018
GAPC1	AT3G04120	Glycolysis/carbon flux/mitochondrial function	Energy metabolism	Rius et al., 2008
AT5G54940.3	AT5G54940	Expressed in root hair	Unclear	Li and Lan 2015
AILP1	AT5G19140	Expressed in shoot meristem/aluminium ion response	Unclear	Yadav et al., 2014
AT1G78040.3	AT1G78040	Functionally uncharacterized extension family member	Unclear	Luo et al., 2012
UBQ3	AT5G03240	Protein degradation	Unclear	Sun and Callis 1997
TUA2	AT1G50010	Ubiquitous / many roles	Unclear	Abe et al., 2004
TUA4	AT1G04820	Ubiquitous / many roles	Unclear	Abe et al., 2004
AT2G20870	AT2G20870	Cell wall protein downregulated in *ft* mutants	Unclear	Cai et al., 2007
VAT1	AT5G16290	BCAA synthesis / accumulation	Nectar composition	Chen et al., 2010
TP2	AT1G07340	High-affinity monosaccharide transporter	Nectar composition	Schneidereit et al., 2003
CWINV1	AT3G13790	CWINV4 is required for nectar production	Nectar production	Ruhlmann et al., 2010
SUSY1	AT5G20830	Highly expressed in *Nicotiana* floral nectaries	Nectar production	Ren et al., 2007
PIP2;5	AT3G54820	Aquaporin in *Aquilegia*	Nectar production	Singh et al., 2020
PIP1A	AT3G61430	Aquaporin in *Aquilegia*	Nectar production	Singh et al., 2020
GAMMA-TIP	AT2G36830	Aquaporin in *Aquilegia*	Nectar production	Singh et al., 2020
PIP2;2	AT2G37170	Aquaporin in *Aquilegia*	Nectar production	Singh et al., 2020

Table S4.4. Continued

DELTA-TIP	AT3G16240	Aquaporin in *Aquilegia*	Nectar production	Singh et al., 2020
BGLU20	AT1G75940	Anther Specific Glucosidase	Glucose production	Rubinelli et al., 1998
SWEET9	AT2G39060	Nectar Production	Nectar production	Lin et al., 2014
ACP2	AT1G54580	Increased expression in presence of sucrose	Nectar composition	Bonaventure and Ohlrogge 2002
CASPL1E2	AT4G15620	Bind with *GAI*	Unclear	Barro-Trastoy et al., 2022
SIP3	AT4G30960	Aquaporin	Nectar production	Quigley et al., 2001
ATBBE26	AT5G44400	Cell wall related in lateral root development	Growth and development	Xun et al., 2020
XTH24	AT4G30270	Turgor-driven polar cell elongation	Growth and development	Lee et al., 2017
EXL2	AT5G64260	Suppresses brassinosteroid-dependent growth and controls C allocation	Growth and development	Schröder et al., 2012
SAMDC	AT3G02470	Production of intermediates in the polyamine biosynthetic pathway	Growth and development	Ge et al., 2006
CLK1	AT1G05850	Modulates ethylene biosynthesis to regulate root development	Growth and development	Gu et al., 2019
PRX44	AT4G26010	Root hair growth	Growth and development	Marzol et al., 2022

Table S4.5. (A-V) Tables generated from the KEGG automated annotation server using BLAST and bi-directional best hits with the *Arabidopsis thaliana*, *Brassica rapa* and *Tarenaya hassleriana* gene data sets. Transcripts lists for each stage had TPM expression above 10 and a coefficient of variation less than 50. Bolding indicates differences between stages that are greater than 3. Categories with less than 3 were removed. S1 = bud pre-anthesis; S2 = flower at anthesis; S3 senescent flower with ~10mm gynoecium; B. = Biosynthesis; D. Degradation. M. = Metabolism; P. = Pathway.

(A) Carbohydrate Metabolism	ID	S1	S2	S3
Glycolysis	10	31	34	33
Amino/Nucl sugar M.	520	32	32	34
Pyruvate M.	620	30	32	32
Starch/sucrose M.	500	30	30	31
Inositol phosphate M.	562	25	26	26
Glyoxylate M.	630	26	25	25
Citrate cycle	20	19	19	19
Propanoate M.	640	18	19	19
Ascorbate/aldarate M.	53	17	19	17
Fructose/mannose M.	51	16	16	17
Pentose phosphate P.	30	16	16	16
Galactose M.	52	13	14	14
Pentose interconversions	40	13	12	12
Butanoate M.	650	12	12	12
C5-dibasic acid M.	660	5	5	5

(B) Amino Acid Metabolism	ID	S1	S2	S3
Cysteine/methionine M.	270	38	40	41
Glycine, serine/threonine M.	260	29	30	29
Alanine, aspartate/glutamate M.	250	27	27	27
Valine, leucine/isoleucine D.	280	22	23	22
Arginine/proline M.	330	22	23	22
Phenylalanine, tyrosine/tryptophan B.	400	22	22	23
Arginine B.	220	20	20	20
Tryptophan M.	380	14	17	16
Tyrosine M.	350	10	13	14
Phenylalanine M.	360	9	13	13
Histidine M.	340	11	12	11
Lysine D.	310	8	12	11
Valine, leucine/isoleucine B.	290	10	10	10
Lysine B.	300	7	7	8

(C) Nucleotide Metabolism	ID	S1	S2	S3
Purine M.	230	40	40	43
Pyrimidine M.	240	23	24	27

(D) Metabolism of Other Amino Acids	ID	S1	S2	S3
beta-Alanine M.	410	15	17	16
Glutathione M.	480	16	15	17
Selenocompound M.	450	9	9	9
Cyanoamino acid M.	460	8	7	9
Phosphonate/phosphinate M.	440	3	3	3

(E) Glycan Biosynthesis and Metabolism	ID	S1	S2	S3
N-Glycan B.	510	24	26	26
Various types of N-glycan B.	513	19	20	21
GPI-anchor B.	563	12	13	13
Other glycan D.	511	8	9	9
O-Antigen nucleotide sugar B.	541	7	7	7
Other types of O-glycan B.	514	6	6	6
Glycosaminoglycan D.	531	5	5	5
Glycosphingolipid B.	603	3	3	3
Lipopolysaccharide B.	540	3	3	3

(F) Lipid Metabolism	ID	S1	S2	S3
Glycerophospholipid M.	564	31	32	33
Glycerolipid M.	561	26	30	28
Fatty acid B.	61	16	16	16
Steroid B.	100	16	16	16
alpha-Linolenic acid M.	592	13	14	14
Sphingolipid M.	600	10	12	12
Fatty acid D.	71	10	11	11
B. of unsaturated fatty acids	1040	7	7	7
Ether lipid M.	565	6	6	6
Arachidonic acid M.	590	5	5	6
Fatty acid elongation	62	5	5	5
Cutin, suberine/wax B.	73	4	5	5
B./D. of ketone bodies	72	3	3	3
Linoleic acid M.	591	3	3	3

(G) Metabolism of Cofactors and Vitamins	ID	S1	S2	S3
Porphyrin/chlorophyll M.	860	29	29	31
Ubiquinone B.	130	13	17	21
Pantothenate/CoA B.	770	14	15	16
Folate B.	790	14	13	14
Nicotinate/nicotinamide M.	760	12	12	12
Thiamine M.	730	9	10	10
One carbon pool by folate	670	9	10	10
Riboflavin M.	740	7	8	9
Vitamin B6 M.	750	7	7	7
Biotin M.	780	6	7	8
Retinol M.	830	5	4	5

(H) Xenobiotics Biodegradation and Metabolism	ID	S1	S2	S3
Drug M. - other enzymes	983	12	13	15
M. of xenobiotics	980	4	3	4
Benzoate D.	362	3	3	3
Styrene D.	643	3	3	3
Chloroalkane/ene D.	625	3	2	3
Drug M. - cytochrome P450	982	3	2	3

(I) Energy Metabolism	ID	S1	S2	S3
Oxidative phosphorylation	190	35	25	25
Carbon fixation	710	23	23	23
Methane M.	680	17	16	17
Sulfur M.	920	13	14	14
Carbon fixation in prokaryotes	720	13	13	13
Photosynthesis antenna proteins	196	12	12	12
Nitrogen M.	910	10	8	9
Photosynthesis	195	8	6	8

(J) Metabolism of Terpenoids and Polyketides	ID	S1	S2	S3
Terpenoid backbone B.	900	27	28	27
Carotenoid B.	906	15	14	14
Brassinosteroid B.	905	4	3	4
Diterpenoid B.	904	3	3	3
Zeatin B.	908	4	3	2
Sesquiterpenoid/triterpenoid B.	909	2	3	3

(K) Biosynthesis of other Secondary Metabolites	ID	S1	S2	S3
Phenylpropanoid B.	940	18	20	21
Flavonoid B.	941	7	8	8
Tropane B.	960	7	8	8
Isoquinoline alkaloid B.	950	5	7	7
Monobactam B.	261	5	5	6
Stilbenoid B.	945	4	4	4
Glucosinolate B.	966	3	4	4
Streptomycin B.	521	3	4	4
Prodigiosin B.	333	3	3	3

(L) Transcription	ID	S1	S2	S3
Spliceosome	3040	70	70	69
Basal transcription	3022	22	21	22
RNA polymerase	3020	12	12	12

(M) Translation	ID	S1	S2	S3
Nucleocytoplasmic transport	3013	81	78	84
mRNA surveillance	3015	46	46	46
Ribosome biogenesis	3008	41	39	43
Ribosome	3010	32	31	32
Aminoacyl-tRNA B.	970	25	26	26

(N) Replication and Repair	ID	S1	S2	S3
Nucleotide excision repair	3420	25	22	30
DNA replication	3030	12	9	27
Base excision repair	3410	13	13	20
Homologous recombination	3440	12	10	19
Mismatch repair	3430	9	6	13
Non-homologous end-joining	3450	5	5	7

(O) Environmental Adaptation	ID	S1	S2	S3
Thermogenesis	4714	40	32	32
Plant-pathogen interaction	4626	26	25	26
Circadian rhythm - plant	4712	17	17	18

(P) Cellular Community - Prokaryotes/Eukaryotes	ID	S1	S2	S3
Quorum sensing	2024	12	11	12
Biofilm formation - *E. coli*	2026	3	3	3
Tight junction	4530	12	13	13
Focal adhesion	4510	6	7	8
Adherens junction	4520	5	6	6
Gap junction	4540	4	4	5

(Q) Signal Transduction	ID	S1	S2	S3
Plant hormone signaling	4075	33	34	34
MAPK signaling P. - plant	4016	28	32	30
PI3K-Akt signaling P.	4151	22	23	24
mTOR signaling P.	4150	23	22	24
AMPK signaling P.	4152	20	21	21
Phosphatidylinositol signaling	4070	18	18	19
HIF-1 signaling P.	4066	15	16	17
Sphingolipid signaling P.	4071	12	15	16
FoxO signaling P.	4068	13	13	14
Wnt signaling P.	4310	11	13	13
Apelin signaling P.	4371	12	12	13
MAPK signaling P. - yeast	4011	12	11	12
MAPK signaling P.	4010	8	10	11
Two-component system	2020	10	9	9
Phospholipase D signaling P.	4072	9	9	10
TGF-beta signaling P.	4350	8	8	8
Calcium signaling P.	4020	7	8	8
Notch signaling P.	4330	7	7	7
Hippo signaling P.	4390	7	7	7
cAMP signaling P.	4024	6	7	8
Ras signaling P.	4014	5	6	7
cGMP-PKG signaling P.	4022	5	6	7
Hedgehog signaling P.	4340	5	5	5

(R) Membrane Transport	ID	S1	S2	S3
ABC transporters	2010	5	5	6
Bacterial secretion system	3070	5	5	5

(S) Transport and Catabolism	ID	S1	S2	S3
Autophagy - yeast	4138	39	38	38
Endocytosis	4144	36	36	36
Peroxisome	4146	32	32	32
Lysosome	4142	26	27	26
Autophagy - other	4136	19	19	19
Phagosome	4145	15	16	16
Mitophagy - yeast	4139	12	12	12

(T) Folding, Sorting and Degradation	ID	S1	S2	S3
Protein processing in ER	4141	63	63	62
Ubiquitin proteolysis	4120	34	34	37
RNA D.	3018	33	32	34
Proteasome	3050	24	24	24
Protein export	3060	14	14	14
SNARE interactions	4130	7	7	7
Sulfur relay system	4122	6	6	6

(U) Cell Growth and Death	ID	S1	S2	S3
Cell cycle	4110	27	28	46
Cell cycle - yeast	4111	22	20	43
Meiosis - yeast	4113	18	15	33
Oocyte meiosis	4114	16	17	22
Cellular senescence	4218	14	16	20
Necroptosis	4217	11	11	11
Apoptosis	4210	9	11	12
p53 signaling P.	4115	9	9	9
Ferroptosis	4216	7	8	8
Cell cycle - Caulobacter	4112	5	5	5

(V) Cell Motility	ID	S1	S2	S3
Regulation of actin	4810	9	10	11

Table S4.6. Phenotyping data for all VIGS treatment groups used in this study. * The *CvSHP* treatment group was from a preliminary study and used different controls.

Construct	Phenotype	Total Plants	Total Plant mortality	Total plants displaying phenotype	Total Altered Flowers	Altered flowers per plant (avg.)
CRC +*ANS*	Reduced or Absent Nectary + Yellow Petals	100	27	30	102	3
AG +*SHP*	Reduced or Absent Nectary + Increased Whorls + No stamen or carpel	60	12	14	71	5
UNCH +*ANS*	N/A	20	10	9	17	2
SWEET9 +*ANS*	No nectar production	60	16	13	33	3
*SHP**	N/A	46	11	0	0	0
AG	Increased Whorls + No stamen or carpel	20	5	7	69	10
ANS	Yellow Petals	20	6	7	32	5
MCS	N/A	10	1	n/a	n/a	n/a
Untreated	N/A	10	0	n/a	n/a	n/a

Chapter 5: *TCP1* imparts monosymmetry in *Cleome violacea* (Cleomaceae), a case-study in continued exploration of a remarkable parallelism[i]

5.1 Introduction

Much of floral evolution can be attributed to evolutionary pressure from pollinators, which can result in convergence of traits (i.e., synchronized evolution of plant and pollinator) (Wessinger and Hileman, 2020). The repeated evolution of monosymmetric flowers represents a valuable and well-studied example of convergent evolution of a complex trait (reviewed in Hileman, 2014; Specht and Howarth, 2015; Wessinger and Hileman, 2020). Across angiosperms, floral monosymmetry (i.e., zygomorphy or bilateral symmetry) has evolved independently over 100 times from polysymmetry (i.e., actinomorphy or radial symmetry) (Reyes et al., 2016). Monosymmetry is most often reflected by differences in adaxial (dorsal) and abaxial (ventral) floral morphology, which may be imparted in either single or multiple whorls, and may include organ fusion (Endress, 1999, 2012). This phenomenon ranges from subtle displays to highly elaborate forms (Endress, 2012). The transition to, and elaboration of, monosymmetry is often correlated with increased speciation (Sargent, 2004; O'Meara et al., 2016) and a shift from general to specialized pollinators (Citerne et al., 2010).

Notwithstanding the phylogenetic distribution and incredible variation exhibited across monosymmetric flowers, there is substantial evidence of parallel recruitment of the same genetic program in investigated transitions from polysymmetry to monosymmetry (reviewed in: Hileman, 2014; Kramer, 2019; Wessinger and Hileman, 2020). Homologues *of Teosinte branched 1/CYCLOIDEA/Proliferating cell factor* (*TCP1*) transcription factors control floral monosymmetry across Eudicots and Monocots. Their role has been demonstrated by functional studies in Asteraceae (Juntheikki-Palovaara et al., 2014), Brassicaceae (Busch and Zachgo, 2007), Fabaceae (Feng et al., 2006), Gesneriaceae (Liu et al., 2021), Plantaginaceae (Reardon et al., 2009), and others (Jabbour et al., 2009; Hileman, 2014). In these, and other monosymmetric

---

taxa, *TCP1* exhibits polarized expression, and is most often restricted to the adaxial portion of the flower (Busch and Zachgo, 2007; Preston and Hileman, 2009; Martín-Trillo and Cubas, 2010; Zhang et al., 2010; Specht and Howarth, 2015). In addition to this adaxial expression pattern, monosymmetry is often, but not always, correlated with multiple copies of *TCP1* (Rosin and Kramer, 2009)*.* As such, late developmental expression of two *TCP1* paralogues is believed to be the primary driver behind adaxial and abaxial petal differentiation in flowering plants (Martín-Trillo and Cubas, 2010; Hileman, 2014; Damerval and Becker, 2017; Kramer, 2019). However, a single copy of *TCP1* is sufficient to produce monosymmetric flowers in some species such as *Iberis amara* (Brassicaceae) (Busch and Zachgo, 2007).

While the putative role and expression of *TCP1* orthologs are well investigated in many taxa (reviewed in Hileman, 2014; Wessinger and Hileman, 2020), the broader genetic pathway involving *TCP1* is most studied in the Lamiales. Snapdragon has been the primary genetic model for monosymmetry; this Asterid exhibits a fused bilabiate corolla and many abaxial/adaxial differences (reviewed in Hileman, 2014; Kramer, 2019). Briefly, two *TCP1* paralogs, *CYCLOIDEA* (*CYC*) and *DICHOTOMA (DICH)*, are expressed in the adaxial region of the flower and promote adaxial identity (Almeida et al., 1997). *CYC*/*DICH* positively regulate *RADIALIS* (*RAD*), a MYB family transcription factor, whose expression is also only found in the adaxial region of the flower (Corley et al., 2005). In turn, a paralog of *RAD*, *DIVARICATA* (*DIV*), is a key regulator of abaxial identity. While *DIV* has potential for broad expression across the flower, *RAD* excludes *DIV* from the adaxial region by competing protein interactions with *DRIF1/2* (*DIV* and *RAD* Interacting Factors 1 and 2), which are required for *DIV* function (Raimundo et al., 2013). These gene and protein interactions of *CYC*, *RAD*, and *DIV* are conserved across Lamiales (Citerne et al., 2000; Zhong and Kellogg, 2015; Su et al., 2017; Sengupta and Hileman, 2022) as well as Solanales (Sengupta and Hileman, 2022). In fact, it has been recently proposed that the *CYC-RAD-DIV-DRIF* interactions have been co-opted as an entire genetic pathway for floral monosymmetry from female organ development in Lamiales (Sengupta and Hileman, 2022). The interactions of the genetic pathway in its entirety have not been investigated thoroughly beyond these Asterid clades.

Continued dissection of monosymmetric floral diversity is warranted to understand the extent and intricacies of the widespread parallel recruitment of *TCP1* and the potential involvement of other regulatory genes. Additional data on different taxa broadens the comparative landscape and may provide insight on outstanding questions. For example, it is unclear how *TCP1* orthologs were repeatedly recruited in monosymmetry (Sengupta and Hileman, 2018; Wessinger and Hileman, 2020). TCP genes are known to play a general role in promoting or repressing cell proliferation and cell growth processes (Preston and Hileman, 2009), but their specific function for promoting monosymmetry is not established (Hileman, 2014). Further, *TCP1* is autoregulated in *Antirrhinum* and across Lamiales (Yang et al., 2012; Sengupta and Hileman, 2018, 2022). Whether *TCP1* auto-regulation occurs outside of Lamiales has not been addressed in detail, although TCP binding sequences have been described in other orders (Sengupta and Hileman, 2018).

Towards this end, Cleomaceae (Brassicales) is an excellent family to study because it is an exception to some of the evolutionary trends associated with monosymmetry (Bayat et al., 2018). For example, monosymmetry is often touted as a key innovation that leads to lineage diversification, e.g., Fabaceae and Orchidaceae (O'Meara et al., 2016). However, Cleomaceae comprises only a small number of species (~270) in comparison to its larger and predominantly dissymmetric sister family, Brassicaceae (~3700) (Bayat et al., 2018). This trend is particularly intriguing as the two families share in their general floral ground plans: four sepals, four petals, six stamens and bicarpellate gynoecium. However, Cleomaceae exhibit greater floral diversity in monosymmetry, organ number and organ elaboration (Bayat et al., 2018). Further, monosymmetry in the Cleomaceae is not correlated with specialized pollination, as is observed in other families (Armbruster et al., 2014; Cardoso et al., 2018; Reiter et al., 2018). In fact, limited data reveal a roughly equal proportion of specialized and generalized pollination syndromes in the cosmopolitan family (Cane, 2008; Fleming et al., 2009; Higuera-Díaz et al., 2015; Bayat et al., 2018).

In Cleomaceae, monosymmetry is typically imparted by the curvature of petals and reproductive whorls. Some species exhibit further elaboration of pigmentation, size, and shape

between adaxial and abaxial petals as well as the presence of variably shaped adaxial nectary glands (Patchell et al., 2011; Bayat et al., 2018). Petals are oriented upwards (i.e., adaxially) whereas reproductive whorls curve downwards before curving up again, i.e., sigmoidal curvature (Patchell et al., 2011). There are two distinct developmental pathways that result in similar mature flowers across species, early monosymmetry and early dissymmetry (Patchell et al., 2011). Briefly, the early monosymmetric pattern is characterized by an enlarged abaxial sepal primordium during development, while early dissymmetry has four equally sized sepal primordia. Early monosymmetry is typically accompanied by dimorphic petals varying in size and/or color. This developmental pattern is also seen in species of Brassicaceae, such as *Arabidopsis thaliana* (L.)*.* Early dissymmetry typically results in flowers with petals of similar size and color. Regardless of early developmental patterns, most aspects of monosymmetry exhibited by members of Cleomaceae occur late in development (Patchell et al., 2011), which is consistent with the predicted role of *TCP1*.

The overarching goal of this chapter is to examine whether, and to what extent, *TCP1* is responsible for floral monosymmetry in *Cleome violacea*. This taxon is a representative of early monosymmetry that exhibits prominent differences in adaxial and abaxial petals as well as a large adaxial nectary (Figure 5.1). We use RNA-seq and RT-qPCR to evaluate broad gene expression patterns and those of candidate genes. We also compare expression patterns of *CvTCP1* to two copies of *Tarenaya hassleriana ThTCP1.* Inclusion of *T. hassleriana* provides a comparison to the early dissymmetric developmental pathway to assess if there are potential genetic differences underlying the two developmental trajectories towards monosymmetry in Cleomaceae. We then perform functional tests using virus-induced gene silencing (VIGS) to assess the contribution of *CvTCP1* to monosymmetry in *C. violacea*. We also conduct in silico analyses to examine if *CvTCP1* can potentially autoregulate. This study represents the first investigation of the genetic basis of monosymmetric flowers in Cleomaceae. Our results demonstrate that *CvTCP1* contributes to adaxial identity in *C. violacea* and provide preliminary evidence of deviations in CYC/DICH-RAD-DIV pathway interactions as compared to Lamiales.

5.2 Materials and Methods

5.2.1 Plant Growth

Lines of *C. violacea* and *T. hassleriana* were grown from seed in a 2:1 soil to perlite mixture (Sungro sunshine mix #4, Agawam, MA, USA) with 16 hours of light and 8 hours of darkness at 24 °C in the University of Alberta, Department of Biological Sciences, growth chambers. These plants are inbred lines originally from seed (*C. violacea*: Horticus Botanicus, #813 and *T. hassleriana*: B&T seeds 27019), and have vouchers deposited in the University of Alberta Vascular Plant Herbarium (ALTA).

5.2.2 RNA Extractions

RNA extractions were performed using a RNeasy Micro Kit (Qiagen, Hilden, Germany) for quantitative reverse transcriptase PCR (RT-qPCR), VIGS, and RNA-seq. The following amendments were applied to the manufacturers protocol: Ground tissue in RLT buffer was incubated for six minutes to maximize lysis. RNA-loaded columns were left to sit for five minutes before centrifugal elution at the final step, and the same eluate was loaded again in the same RNA column to ensure maximum yield. RNA was stored at -80°C prior to downstream analyses. Elutions were quantified using a NanoDrop 1000 (V. 3.1; Thermo Fisher Scientific, Waltham, Massachusetts, USA) and qualified using an Agilent 2100 Bioanalyzer (V. B.02.09.SI720; Agilent, Santa Clara, California, USA) prior to all downstream preparations. For RNA-seq, RNA extractions for each biological replicate were diluted to the lowest sample concentration before cDNA library preparation.

5.2.3 Identification of CvTCP1

TCP1 orthologues were originally identified in (Brock, 2014) from genomic and cDNA samples of *C. violacea*. Total RNA was extracted from multiple inflorescence stages using a Concert Plant RNA Reagent kit (Invitrogen, Carlsbad, CA, USA). Messenger RNA was purified using a Dynabeads mRNA Direct Kit (Invitrogen), and then converted to cDNA using Superscript III First Strand Synthesis System (Invitrogen). A 1000-1100bp fragment of *CvTCP1* was consistently amplified from both cDNA and genomic templates using forward and reverse primers which were originally used to amplify a *TCP1* homologue in *I. amara* (Table 5.1). PCR conditions for

amplification included an initial denaturation of 10 minutes at 94°C, followed by 36 cycles of denaturation at 94°C for 0.5 minutes, annealing for 55°C for 1 min, and extension at 72°C for 2 minutes followed by a final extension at 72°C. The *CvTCP1* fragment was excised and purified using a QIAquick gel extraction kit (Qiagen), cloned using pGEM T-easy plasmids (Promega) and propagated in JM109 *E. coli* (Promega). Colonies were screened using M13 forward and reverse primers (Table 5.1) to visualize variation in cloned fragment size. Plasmids were extracted from JM109 cells using miniprep kits (Qiagen, Hilden, Germany) and purified using Performa DTR V3 96-well short plates (Edge Biosystems, Gaithersburg, MD). Purified plasmids were sequenced by cycle sequencing reactions using M13 forward and reverse primers (ABI Big Dye v.3.1, Applied Biosystems, Foster City, CA).

5.2.4 RT-qPCR

In RT-qPCR experiments, we examined *TCP1* expression in different floral organs in both *C. violacea* and *T. hassleriana*. *ThTCP1* orthologues identified from genomic analyses of *T. hassleriana* are designated Th21666 and Th24587 (Cheng et al., 2013). Primers were ordered from Integrated Device Technologies (IDT, Coralville, Iowa, USA), and tested using IDT gBlocks Gene Fragments; a gBlock is a ~1kb synthetic oligonucleotide with primer binding sites for all primers (Table 5.1). After successful testing, cDNA was synthesized using a Thermo Scientific RevertAid H Minus First Strand cDNA Synthesis Kit (Thermo Fisher Scientific, Waltham, Massachusetts, USA). The synthesis protocol was modified by adding 1 µL of 100 µM polyT and 1 µL of 100 µM random hexamer primer in lieu of just one primer type (Kramer et al., 1998). Before the cDNA was used for RT-qPCR, it was diluted 1:5 with IDT Nuclease-free Water, and stored at -20°C. All RT-qPCR analyses used the QuantStudio 6 Flex Real-Time PCR System (Life Technologies, Carlsbad, California, USA). Parameters of operation for RT-qPCR were: 2-ΔΔ-Ct, 384 well plates, and 10 µl reaction volumes. Each 10 µL reaction contained a SYBR Green master mix (0.25xSYBR Green, 0.1xROX, 0.3U Invitrogen Platinum Taq Polymerase and 0.2mM dNTPs), 300 mM of forward and reverse primer, and variable concentrations of DNA. All reactions were run in technical triplicates, and ACTIN was the endogenous control (Table 5.1). For *C. violacea*, a single biological replicate was defined as four plants grown in one 4-inch square pot because flowers from a single plant did not yield sufficient RNA for analyses. For *T.*

hassleriana, a single plant was defined as a biological replicate because flowers from one plant yielded sufficient RNA. Three biological replicates were analyzed for both species. Floral buds less than 3.0 mm were defined as small, and buds greater than or equal to 3.0 mm were defined as large. Sepals, adaxial petals, abaxial petals, glands, stamens, and gynoecia were dissected, flash frozen, and stored at -80 °C for downstream RNA extraction.

5.2.5 Transcriptome Assembly and Analysis

We also generated RNA-seq expression profiles for adaxial and abaxial petals in *C. violacea* using newly anthetic flowers (N = 3). Petal tissue was flash frozen and stored at -80°C prior to RNA isolation. All steps of RNA isolation, library preparation, assembly and analysis were followed as described in Chapter 2 using updated software (Table 5.2). The updated Araport11 database (Cheng et al., 2017) was used for transcript annotation in lieu of The *Arabidopsis* Information Resource (TAIR) 10 database. Raw reads are available at the Sequence Read Archive (SRA) database, and the Transcriptome Assembly is deposited at DDBJ/EMBL/GenBank. RNA-seq petal data was used to validate *CvTCP1* gene expression from prior RT-qPCR experiments as well as identify other genes of interest. Analysis parameters for the gene ontology (GO) and Kyoto encyclopedia of genomes and genes (KEGG) analysis were altered considerably from prior studies and are outlined here in greater detail.

5.2.6 Gene Ontology

Two orthology-related analyses were run to compare pathway expression underlying adaxial and abaxial petals. First, we ran a GO analysis which compiles biological, cellular, and molecular processes into a tangible view of functionality, e.g., pigmentation. Second, we ran a KEGG analysis which shows the activity of biological pathways as determined by gene expression data, so we can better understand petal development from a viewpoint of which pathways are active, e.g., flavonoid biosynthesis. These analyses are limited in that they do not consider differential expression, i.e., a presence/absence threshold must be selected. Despite these constraints, both analyses are effective for understanding biological significance within and between petals.

For these analyses we compiled the following fasta files: all transcripts above or equal to five TPM, all transcripts greater than or equal to 100 TPM, and all significantly differentially expressed transcripts (FDR < 0.05 and fold change ≥ 4). The lists were split into separate adaxial and abaxial matrix files based on averaged expression across biological replicates. Geneious (Table 5.2) was used to extract fasta formatted transcripts from the Trinity assembly fasta file, and OrfPredictor (Table 5.2) was used for longest open reading frame (ORF) prediction and translation. For the KEGG analyses, adaxial and abaxial fasta files were uploaded to the KAAS server (https://www.genome.jp/kegg/kaas) using bidirectional hits and the *Arabidopsis, Brassica rapa* (Brassicaceae)*, and T. hassleriana* organism lists. For the GO analyses the translated amino acid sequences were uploaded to the OrthoVenn2 web server (https://orthovenn2.bioinfotoolkits.net/home). Typically, OrthoVenn2 is used to determine orthology between different species, but in principle can be used to determine differences between floral organs of the same species (Landis et al., 2016) (See Chapter 1).

5.2.7 VIGS

We followed the previously described VIGS protocol for *C. violacea* (See Chapter 3). Four constructs were used for this study. A construct-free vector (pTRV2-MCS) served as a negative control, and a vector containing *PHYTOENE DESATURASE* (*PDS*) (pTRV2-*CvPDS*) served as a positive control. Constructs containing class II *CvTCP1* (pTRV2-*CvTCP1*) and class I *CvTCP14* (pTRV2-*CvTCP14*) were generated using *C. violacea* genomic DNA. We used SiFi21 (Table 5.2) to verify construct functionality in-silico. RNAi trigger sequences (i.e., virtual 21bp siRNA related to pTRV2 constructs) were used to predict hits to the *C. violacea* transcriptome and genome. An on-target hit was defined as any expected transcript or genomic hit with a low ratio of alternative hits (≤ 3 hits). An off-target hit was defined as any unexpected hit, or a siRNA with many alternative hits (>3 hits).

All vectors were transformed into *Agrobacterium tumefaciens,* and glycerol stocks were prepared and stored at -80°C. Vacuum infiltration was used for all vectors. VIGS phenotypes were scored between 20 and 40 days after inoculation and collected for RT-qPCR verification.

5.2.8 Autoregulation

We searched for class II TCP binding sites upstream of genomic *CvTCP1* to assess potential autoregulation. The first binding sequence searched for was 5' -GTGGNCCC- 3' because it is the preferential site for class 2 TCPs (Kosugi and Ohashi, 2002; Yang et al., 2012). We also searched for the core TCP binding site (5' –GGNCCC- 3'). To find binding sites upstream of *CvTCP1* we first used the online CoGe BLAST and aligned the following sequences to the draft genome: *AtTCP1* (AT1G67260.1) and both copies of *T. hassleriana ThTCP1*, *Th21666* and *Th24587* (XM_010530156.2 & XM_010530157.2). Scaffolds 169 and 59 were the top hits for all sequences and, as such, both were downloaded from the CoGe webserver. After downloading, we used Geneious (Table 5.2) to search for TCP binding motifs on each scaffold and reported those within 10kb upstream of the *CvTCP1* start codon.

5.3 Results

5.3.1 Abaxial and adaxial petals differ in morphology and gene expression profiles

Cleome violacea is a small herb that grows up to 30 cm in height and produces bracteate racemes bearing small flowers approximately 8 mm wide. Flowers of this species exhibit monosymmetry in all whorls. The abaxial sepal is wider than the lateral and adaxial sepals. The adaxial and abaxial petal pairs differ in shape, size, and color. Adaxial petals are maroon with a yellow spot whereas abaxial petals are uniformly maroon with yellow undertones. Both abaxial and adaxial petals are clawed and similar in length, measuring 3-4 mm. Adaxial petals are narrower at 1.5-2.0 mm relative to abaxial petals that are 2.5- 3.5 mm wide. Reproductive whorls curve upwards at anthesis, and a prominent 3-lobed nectary is positioned between adaxial petals and stamens (Figure 5.1).

We generated RNA-seq data for both adaxial and abaxial petals that reveal distinct expression profiles in both. In total, 517 of 110,332 transcripts are significantly differentially expressed between adaxial and abaxial petals (Figure 5.2A). Of these, 247 are upregulated in abaxial petals and 270 are upregulated in adaxial petals. Generally, transcripts above 5 TPM are more abundant in abaxial petals, whereas those between 1 and 5 TPM are more abundant in adaxial petals (Figure S5.1).

Orthology-related analyses also reveal adaxial and abaxial differences. A GO analysis of all transcriptomic expression greater than 5 TPM resulted in 13293 shared, 32 unique-adaxial, and 19 unique-abaxial GO groupings (Figure 5.2B). Among the top 50 shared terms across petal types the majority are biological processes related to growth (Table 5.3). Of note, the metabolic processing of aromatic compounds (GO:0006725) has a count of 696, which suggests the presence of floral volatiles (Farré-Armengol et al., 2020) .The most abundant cluster of eight transcripts unique to abaxial petals is undefined, and two are involved in regulation of transcription (GO:0006355; Table 5.4). There are several clusters unique to adaxial petals that appear to be related to nectar (e.g., starch metabolic process, starch biosynthetic process, and sucrose transport (Table 5.4). Many of these terms are only strongly expressed in replicate 2 (Figure 5.4) and may be due to the close association of nectary tissue and adaxial petals.

KEGG analyses show the activity of biological pathways as determined by gene expression data and are similar between abaxial and adaxial petals. The KEGG analysis of all transcriptomic expression greater than 5 TPM resulted in 3599 and 3591 terms for abaxial and adaxial petals, respectively. Like the GO analysis, many terms are related to growth and development, e.g., plant hormone signal transduction is 15th with a count of 38 abaxially and 36 adaxially. Surprisingly, yeast autophagy is the 10th top term with a count of 47 abaxially and 46 adaxially (Table 5.5).

To further investigate gene expression patterns between petal types we also ran GO and KEGG analyses for significantly differentially expressed transcripts and highly expressed transcripts. All significantly differentially expressed transcripts are represented by 62 KEGG pathways (Table 5.6), excluding pathways related to mammals that were captured due to conservation between kingdoms. Among the significantly differentially expressed transcripts, photosynthesis had the greatest variance between abaxial and adaxial petals at 17:1, followed by antenna proteins at 10:1 (Table 5.6). Similarly for the GO analysis, photosynthesis had a 2:0 ratio of terms in abaxial and adaxial petals, respectively. The top five categories amongst highly expressed transcripts were oxidative phosphorylation, photosynthesis, ER protein processing, glycolysis, and carbon fixation in photosynthetic organisms (Table 5.6). The top three GO categories for significant and

highly expressed transcripts were biological process, metabolic process, and response to stimulus (Table 5.7). These data indicate that abaxial petals are enriched with energy metabolism related transcripts relative to adaxial petals.

5.3.2 There are no class 2 TCP consensus binding sequences upstream of a single copy of CvTCP1

Scaffold 169 was consistently the top hit In a BLAST alignment of multiple *TCP1* related sequences to the draft *C. violacea* genome, and among them DN10350_c1_g3_i2 (*CvTCP1*) had the highest overall bitscore (Table 5.8). There appears to be an incomplete copy of *CvTCP1* on scaffold 59 which was the second top hit for all aligned sequences but had a maximum matching length of 201. In contrast, the matching length of *CvTCP1* on scaffold 169 is 1079 (Table 5.8). We estimated that there is only a single copy of *CvTCP1* because all sequences aligned to a similar position on scaffold 169 and no full size ORFs were found on any other scaffold. Generally, *TCP1* has an ORF of greater than 1000 in most species, e.g., 1080 in *Arabidopsis* and 1233 in *C. violacea.*

After obtaining a plausible copy number for *CvTCP1,* we investigated potential autoregulation by searching for binding sequences around its loci on the draft genome. Specifically, we searched for sequences withing 10kb upstream of the *CvTCP1* start codon on scaffold 169 and 59. The core binding sequence for all TCP genes, type I and II, is 5'-GGNCCC-3', and the consensus sequence for class I TCPs and class II *TCPs* are 5'-GGNCCCAC-3' and 5'-G(T/C)GGNCCC-3', respectively (Kosugi and Ohashi, 2002). There are two core binding sequences within 9.5kb upstream of the predicted *CvTCP1* start codon on scaffold 169, but neither resemble the class II consensus sequence (Table 5.9).

5.3.3 TCP1 exhibits polarized expression in petals of both Cleome violacea and Tarenaya hassleriana

CvTCP1 exhibits significant upregulation in adaxial petals of both *C. violacea* and *T. hassleriana* as well as widespread expression across other floral organs, as demonstrated in both RNA-seq and RT-qPCR data (Figure 5.3). RT-qPCR data reveal that *TCP1* is expressed in sepals, nectary gland, stamens, and gynoecium of anthetic flowers, but not abaxial petals (Figure 5.3). This

polarized expression pattern of *TCP1* across petals is consistent between *C. violacea* and *T. hassleriana* (Figure 5.3B, C), despite each species exhibiting different morphological developmental pathways and mature floral morphologies. Morphologically, the mature flowers of *T. hassleriana* are like *C. violacea* in petal and organ curvature. However, the anthetic adaxial and abaxial petals of *T. hassleriana* are uniform in color, size, and shape, and it has two copies of *TCP1* designated *TH21666* and *Th24587* (Cheng et al., 2013). RT-qPCR data demonstrate that both copies are expressed in developing buds, sepals, nectary, stamens, and carpels of anthetic flowers (Figure 5.3C). However, only *Th24587* is expressed in adaxial petals at anthesis, despite the morphological uniformity of mature petals.

Examination of abaxial/adaxial expression patterns of other TCP genes was also explored with the *C. violacea* RNA-seq dataset. These analyses reveal that a single copy of *CvTCP14* is significantly upregulated in abaxial petals, an intriguing pattern that is opposite of *CvTCP1* expression (Figure 5.3A). In contrast, other TCP genes do not exhibit significant differential expression between petal pairs (Figure 5.3). The identity of *CvTCP14* was determined by BLAST alignment to the Araport11 database and draft genome. DN44436_c0_g1_i1 BLAST aligned to *AtTCP14*(AT3G47620.1) and is located at position 504,596 on scaffold 28 of the genome; It matches with 100% identity to 463 bp (Table 5.8). DN44436_c0_g1_i1 yielded no other strong hits to the draft genome (Table 5.8) and is suspected to be the only copy present in *C. violacea*.

Aside from TCP related expression, transcripts corresponding with downstream floral symmetry and floral organ identity genes were also expressed. It has been suggested that B-Class genes *APETALA 3* (*AP3*) and *PISTILLATA* (*PI*) may be involved with monosymmetry in *T. hassleriana* flowers (Cheng et al., 2013). In *C. violacea,* transcripts related to B-class genes are highly expressed in abaxial/adaxial petals but display no differential expression. This pattern is also true for transcripts related to A-class genes *APETALA 1/2* (*AP1/2*). Transcripts related to C-class gene, *AGAMOUS* (*AG*), and floral symmetry gene, *RADIALIS*-like (*RAD*), are inconsistently expressed across replicates, limiting interpretation. There is, however, a significantly upregulated transcript (DN14369_c0_g1_i3) related to a *T. hassleriana DIVARICATA-like* (*DIV-like)* gene in adaxial petals (Figure 5.3A). The top two NCBI megablast alignments for this

transcript are for predicted *T. hassleriana DIV-like* transcript variants XM_010550842.2 and XM_010550840.2. It is intriguing that a *DIV-like* transcript is expressed in opposition to *TCP1*, which suggests a role in abaxial petal identity, although functional studies are required.

5.3.4 VIGS demonstrates that CvTCP1 *promotes adaxial identity but* CvTCP14 *function is less clear*

Plants treated with TRV2-*CvTCP1* displayed anthetic adaxial petals that appeared morphologically like non-treated abaxial petals (abaxialization), and nectaries were occasionally altered or absent (Figure 5.4). Complete abaxialization of adaxial petals was observed in 26 of 124 treated plants and 16 plants displayed some alteration of adaxial petal morphology (Table 5.10). *CvTCP1* treatment variance was presumably due to inconsistent viral silencing of *CvTCP1* (Figure 5.4). Pigmentation was altered in knockdowns such that yellow spots were reduced (Figure 5.4E, F) or entirely absent in adaxial petals (Figure 5.4 C, D; Table 5.10). Modification of nectaries also ranged from a mild decrease in lobe size (Figure 5.4D) to semi and complete absence of nectary lobes (Figure 5.4E, C). Curvature and number of organs in reproductive whorls, which also contribute to overall monosymmetry of the flower, were not altered in any of the knockdown flowers (Figure 5.4).

Although *CvTCP14* expression was upregulated in abaxial petals (Figure 5.3A), pTRV2-*CvTCP14* treatments exhibited alterations to abaxial/adaxial petals that were challenging to interpret. Treatment with pTRV2-*CvTCP14* resulted in modified floral phenotypes in 40 of 100 total plants (Table 5.10). Of note, many flowers displayed two or more binned phenotypes. The two most common phenotypes were a reduction/absence of one or both adaxial petals (39 of 100 flowers) (Figure 5.5A, B, D, E) followed by a reduction/absence in abaxial petals (20 of 100 flowers) (Table 5.10 and Figure 5.5 C, E, F). Like pTRV2-*CvTCP1* treatments, the curvature of reproductive whorls was similar in both treated and non-treated flowers such that these organs continued to curve upwards. However, unlike phenotypes from *CvTCP1* treatments, the morphology and size of nectaries in pTRV2-*CvTCP14* treated flowers were similar to non-treated. Overall, treatment with pTRV2-*CvTCP14* appears to affect petal size (Figure 5.5 A-C) and medial/lateral sides of the flower, but not abaxial/adaxial identity.

5.4 Discussion

Research on the genetic basis of monosymmetry has uncovered stunning parallelisms across angiosperms via the repeated recruitment of *TCP1* (reviewed in Hileman, 2014; Juntheikki-Palovaara et al., 2014; Wessinger and Hileman, 2020). Here, we establish that an atypical monosymmetric taxon represents another valuable case study for *TCP1's* role in producing monosymmetric flowers by promoting adaxial identity. In *C. violacea*, monosymmetry manifests from differences in size and color between petal types, position and shape of nectary, and curvature of reproductive whorls (Figure 5.1). Although *TCP1* homologs are upregulated in adaxial petals of anthetic flowers, they are also expressed in other floral organs of both *C. violacea* and *T. hassleriana* (Figure 5.3B, C). Functional analyses demonstrate that *CvTCP1* contributes to adaxial petal identity and nectary development but is unlikely to contribute to curvature of reproductive whorls (Figure 5.4). Unlike *CvTCP1,* the role of *CvTCP14* is unclear regarding its contribution to floral monosymmetry (Figure 5.5), despite being upregulated in anthetic abaxial petals (Figure 5.3). When compared to other Brassicales, these data suggest conserved expression and function of TCP homologs across the order that exhibit some deviations from roles established in the Asterids. These deviations include widespread expression of *TCP1* as well as co-expression of other members of the CYC pathway.

5.4.1 Gene expression patterns correlate with morphological differences in adaxial and abaxial petals

RNA-seq analyses reflect morphological differences between adaxial and abaxial petals in *C. violacea*. Adaxial petals, which are smaller and have yellow spots, have similar expression profiles to maroon abaxial petals (Figure S5.1), although a subset of transcript expression is distinct. However similar, expression profiles are not identical. Expression patterns vary from those observed in the closely related *Iberis amara* (Brassicaceae), which has monosymmetric flowers because of size differences between adaxial and abaxial petals. In *I. amara*, adaxial petals have more overall gene upregulation, and are smaller than abaxial petals due to unequal cell proliferation (Busch and Zachgo, 2007; Busch et al., 2014). However, *I. amara* flowers exhibit none of the differences in color and curvature observed in *C. violacea* abaxial/adaxial petals. RNA-Seq data from *I. amara* indicated that photosynthetic genes were significantly

upregulated in adaxial petals, which the authors attributed to size differences in the petal pairs relative to green claws (Busch et al., 2014). Petal claws in *C. violacea* are yellow and appear to be proportionally larger on adaxial petals (Figure 5.1). Despite the similarity in proportions to *I. amara* petals, it is *C. violacea* abaxial petals that exhibit an increase in photosynthesis related KEGG and GO terms (Table 5.6, 5.7). The difference may be attributable to biological function in abaxial petals rather than size differences, but further analysis of separate petal regions is required. Overall, between-species comparisons suggest it may not be global gene expression patterns that contribute to morphological differences in petal pairs, but expression of key transcriptional regulators.

5.4.2 CvTCP1 *imparts adaxial identity in Cleome violacea while the role of* CvTCP14 *is less clear*

Of all TCP and floral identity genes examined, only a few exhibited significant differential expression (Figure 5.3A). Of particular interest are *TCP1* and *DIV* because of their established role in monosymmetry (reviewed in Hileman, 2014; Kramer, 2019; Wessinger and Hileman, 2020), and *CvTCP14* because of its significant differential expression in abaxial petals (FDR < 0.016) (Figure 5.3A). Upregulation of *CvTCP1* in adaxial petals (FDR < 2.2E-7) was expected given the extent to which *TCP1/CYC* exhibits polarized expression in other monosymmetric angiosperms (reviewed in Hileman, 2014; Kramer, 2019; Wessinger and Hileman, 2020). These patterns suggest that *CvTCP1* is responsible for adaxial petal identity, a conserved role in monosymmetric taxa, whereas *CvTCP14* may function in abaxial identity or play another yet uncharacterized role.

Functional studies of *CvTCP1* confirm its role in promoting adaxial identity of flowers in *C. violacea* (Figure 5.4). When plants were treated with pTRV2-*CvTCP1*, adaxial petals became abaxialized, such that the petals were larger and more uniformly maroon in coloration. Thus, *CvTCP1* promotes adaxial identity by reducing size and altering coloration of petals (Figure 5.4). TCP genes widely promote and/or repress tissue growth (Danisman, 2016; Martín-Trillo and Cubas, 2010). As such, *C. violacea* is added to the expanding list of taxa in which *TCP1* is recruited to impart monosymmetry via promoting adaxial identity (reviewed in Hileman, 2014).

As observed in *Antirrhinum* but not *Iberis*, down-regulation of *TCP1* impacts more than the corolla. When flowers are treated with pTRV2-*CvTCP1*, flowers may also exhibit reduced and/or missing adaxial nectary glands (Figure 5.4E, C). In *Arabidopsis,* nectary gland initiation and development are regulated by a combination of genes: *AG*, *LEAFY*, and *SHATTERPROOF1/2* (*SHP1/2*), which activate key regulator *CRABSCLAW* (*CRC*) (reviewed in Slavković et al., 2021). *CRC* has been shown to be recruited across core Eudicots for nectary development but not other Eudicots (reviewed in Wessinger and Hileman, 2020; Liao et al., 2021). A loss of nectaries in pTRV2-*CvTCP1* treated flowers suggests that there are other regulators upstream of *CRC* that are connected to *TCP1*. However, absent nectaries may be due to loss of adaxial identity, thus *CvTCP1* would only be indirectly upstream of nectary formation. Regardless, future studies should investigate the mechanistic link between nectary formation and adaxial identity.

Curiously, reproductive whorls were not altered during *CvTCP1* down-regulation despite widespread expression in those whorls at anthesis (Figure 5.3B). Such patterns of phenotype restriction were found in *Aquilegia* after downregulating *AUXIN RESPONSE FACTOR 6* and *8*, i.e., downregulation of *ARF6/8* did not modify phenotypes in all organs where the genes are typically expressed (Zhang et al., 2020). In *C. violacea*, this may be due to pathway redundancy for correct positioning of stamens and stigma, given their importance for ensuring successful pollination. Alternatively, some organs, e.g., petals and nectaries, may be more sensitive to dosing of focal genes (Zhang et al., 2020).

How *TCP1* expression is restricted to the adaxial portion of flowers is not well understood outside of *Antirrhinum* (Wessinger and Hileman, 2020). One possible mechanism for polarized expression is autoregulation, whereby a gene product enhances or represses its own transcription. Autoregulation was proposed to be key to the evolution of monosymmetry in Lamiales (Sengupta and Hileman, 2018, 2022) where researchers looked for the core TCP binding sequence (5' –GGNCCC- 3') 3 kb upstream of the *CYC* transcription start site. In *C. violacea* there do not appear to be any TCP binding sequences within 3 kb of the *CvTCP1* transcription start site on scaffold 169, and only one binding sequence 681 bp upstream of the partial *CvTCP1* ORF on scaffold 59. We expanded the search for binding sequences up to 10kb

from the initiation site because distant regions of DNA can become proximal due to folding (Hernandez-Garcia and Finer, 2015). We found two core binding sequences upstream on scaffold 169, and two on scaffold 59. Neither resemble the class II consensus sequence for *TCP1* (5'-G(T/C)GGNCCC-3'), however, flanking sequences outside the core binding sequence are not critical for *TCP1* binding (Gao et al., 2015). It seems possible that *CvTCP1* autoregulates, although other autoregulating orthologues have more upstream binding sites, e.g., *Antirrhinum majus TCP1* has eight binding sites within 3 kb of its start site.

Because *DIV* shares a role with *TCP1* in floral monosymmetry, and is expressed in abaxial petals, we thought that it may play a role in *TCP1* regulation. We found one binding sequence within 3 kb of *CvTCP1* on scaffold 169 as well as one on scaffold 59. In total, there were five *DIV* related binding sequences within 10 kb upstream of *CvTCP1* on scaffold 169. The *TCP1* and *DIV* binding sequences appear at a frequency of 1 every 3209 bp and 1 every 1516 bp on scaffold 169, thus by random chance we would expect to find roughly 3 and 6 binding sites upstream of *CvTCP1*, respectively. Altogether it appears that *CvTCP1* does not rely on autoregulation or regulation from *DIV.* It is unclear whether other Cleomaceae *TCP1* orthologues also lack upstream TCP binding sites. Future studies are required to demonstrate protein-DNA binding of *CvTCP1*, or another transcription factor, upstream of *TCP1.*

In contrast to *CvTCP1*, it is unlikely that *CvTCP14* plays a direct role in symmetry by influencing abaxial identity. *CvTCP14* has an intriguing expression pattern of upregulation in abaxial petals (Figure 5.3A), which suggests a role in abaxial petal identity in *C. violacea*. When plants are treated with pTRV2-*CvTCP14*, modification of floral phenotype is observed in both adaxial and abaxial petals (Figure 5.5). In *Arabidopsis*, *AtTCP14* regulates internode length and leaf shape, and *tcp14* knockouts cause severe defects in carpel, sepal and petal morphology, including a reduction in organ size. *AtTCP14* maintains cell density in leaves, and cell density increases in its absence (Kieffer et al., 2011). As such, the upregulation of *CvTCP14* in abaxial petals at anthesis may be explained by the larger size of abaxial petals, i.e., relatively rare transcripts may be diluted out during library preparation. Alternatively, abaxial petals may take longer to develop because of their larger size, and *CvTCP14* may be equally expressed in adaxial petals at earlier

developmental stages. It is interesting nonetheless to consider *CvTCP14* silencing phenotypes, as it seems to affect either the left or right portion of the flower (aka medial-lateral gradient) rather than adaxial-abaxial gradient. Like the role of *AtTCP14* (Kieffer et al., 2011), it is possible this phenotype is due to reduction in petal size and/or disruption of cell organization. Thus, *CvTCP14* may play a role in floral development through the regulation of cell density, but likely does not have a direct role in monosymmetry.

5.4.4 Within and beyond Cleomaceae: TCP exhibits dynamic gene expression across Brassicales

Comparison of candidate gene expression between two genera of Cleomaceae points towards a conserved role of *TCP1* in monosymmetry that is independent of early developmental pathway or whole genome duplications. Unlike *C. violacea*, *T. hassleriana* exhibits an early dissymmetry pathway. Further, *T. hassleriana* has undergone a whole genome triplication that is not shared with *C. violacea* (Cheng et al., 2013; Mabry et al., 2020) and, as such, retains two copies of *TCP1*. Both copies exhibit widespread expression in developing bud and, at anthesis, in the sepals, gland, stamens, gynoecium and leaves (Figure 5.3C). *Th24587* is upregulated in adaxial petals compared to abaxial. This pattern is also observed in *CvTCP1*: It is widespread across developing buds and, at anthesis, in sepals, adaxial petals, gland, stamens and gynoecium (Figure 5.3B). The similarities observed between these species suggest that *TCP1's* contribution to monosymmetry in Cleomaceae is conserved such that the gene functions to promote adaxial identity, especially in flowers. This pattern is also consistent with features that we associate with contributing to monosymmetry occurring late in development (Patchell et al., 2011). Interestingly, the widespread expression of *TCP1* at anthesis in multiple organs is also shared between the two genera, a phenomenon that warrants examination in additional species. Many Cleomaceae species have curved reproductive organs at anthesis. Finally, these data indicate that only a single copy of *TCP1* is needed to impart monosymmetry in Cleomaceae (Table 5.8), which is akin to Brassicaceae (Busch and Zachgo, 2007). It may be that *TCP1* is widely duplicated in monosymmetric taxa, but only one copy plays a direct role in adaxial/abaxial differentiation whereas the other copy has a broader role more typical to the rest of the TCP gene family (Danisman, 2016).

Expression of *TCP1* across many floral whorls is exhibited in other species of Brassicales. In most monosymmetric taxa, *TCP1* exhibits polarized expression, either adaxial or abaxial, throughout development (Hileman, 2014; Wessinger and Hileman, 2020). In contrast, some species of monosymmetric Brassicaceae exhibit widespread expression of *TCP1* homologs across the floral meristem, which later becomes restricted to adaxial petals (Busch et al., 2012). *AtTCP1* is transiently expressed in adaxial regions of *Arabidopsis* before petal primordia are formed (Cubas et al., 2001) even though flowers are dissymmetric at maturity. This intriguing pattern led to the proposal that early adaxial expression was ancestral in the family and that a heterochronic shift in *TCP1* homolog expression resulted in monosymmetric taxa (Busch and Zachgo, 2009). It is important to point out that these seminal papers on Brassicaceae explore *IaTCP1* expression in inflorescences, shoots, leaves, and adaxial/abaxial petals (Busch and Zachgo, 2007; Busch et al., 2012, 2014). However, beyond in situ data of early developing buds, it is not known if *IaTCP1* is expressed in sepals, stamens, or gynoecia of anthetic flowers, with notable exception of *TCP1* being expressed in petals of *Iberis amara* (Busch and Zachgo, 2007). Recent work on *Tropaeolum longifolium*, a member of non-core Brassicales provides additional insight on TCP expression patterns in another monosymmetric member of the order (Martínez-Salazar et al., 2021). Flowers of *Tropaeolum* impart monosymmetry via development of an adaxial spur that is both sepal and petal derived (Martínez-Salazar et al., 2021). In *T. longifolium* many TCP homologs were examined, revealing dynamic and varied expression. Of note, *TiTCP1*, is upregulated in the adaxial petal lobes, stamens and carpels relative to sepals (Figure 10 in (Martínez-Salazar et al., 2021)). In combination, these data suggest that while differential expression of *TCP1* homologs in petals contributes to monosymmetry, expression of *TCP1* is likely widespread in other organs (e.g., sepals, stamens, carpels) that may or may not contributes towards floral monosymmetry. These findings are consistent with our hypothesis that only a single copy of *TCP1* functions in floral monosymmetry of Brassicales, even in taxa with multiple copies. They are also consistent with a recent hypothesis that the entire TCP genetic program also has function in carpel development in Lamiales (Sengupta and Hileman, 2022), suggesting that widespread expression of TCP in other taxa might not be related to monosymmetry. Regardless, this pattern underscores the need to examine *TCP1* homologs

across many whorls in both monosymmetric and polysymmetric flowers of Brassicales to detangle possible functions of these genes beyond promoting adaxial identity.

In addition to the role of *TCP1*, expression data hints at potential involvement of MYB family transcription, *DIV*, in some monotypic Brassicales. Like *CvTCP1*, *CvDIV* is also significantly upregulated in adaxial petals (FDR < 0.014) (Figure 5.3A). Another monosymmetric member of the Brassicales, *Tropaeolum longifolium* (Tropaeolaceae), also exhibits restriction of *DIV-like* homologs (*TlDVL2*) to the adaxial petal later in development despite a broad expression pattern earlier (Martínez-Salazar et al., 2021). *RAD* orthologs also exhibit broad expression domains in *T. longifolium* although appear to be absent in abaxial petal lobes and, in contrast, restricted to the adaxial petal (Martínez-Salazar et al., 2021). To the best of our knowledge, expression patterns of *DIV* orthologs have not been reported for *Iberis*. Data from *Cleome* and *Tropaeolum* suggest that *DIV* and possibly *RAD* orthologs may have a divergent role in Brassicales relative to the Asterids, where *DIV* functions to regulate adaxial identity (Galego and Almeida, 2002). Unlike in *Antirrhinum*, *CvDIV* is not restricted to the abaxial portion of the flower. The restriction of this gene to adaxial portion of flowers in some Brassicales warrants further investigation to determine if there is a novel interaction between *DIV*, *RAD* and *TCP1* in monosymmetric taxa of this order.

5.4.5 Conclusions

Here we validate that continued examination of well-established parallelisms uncovers novel insights. First, we confirm in yet another monosymmetric taxon in which *TCP1* has been recruited to promote adaxial identity and, as a result, imparts floral monosymmetry. However, the polarized expression of *TCP1* is only observed in petal whorls of Cleomaceae. The widespread expression in all other organs at anthesis is unusual and suggests additional roles in contributing towards monosymmetry of Cleomaceae, perhaps related to curvature of reproductive organs. The loss or reduction of adaxial nectaries in knockdown *CvTCP1* flowers represents a novel phenotype. The lack of phenotypes in other floral whorls when *CvTCP1* is downregulated implies functional redundancy with other regulatory genes to ensure critical orientation of stamens and carpels. Finally, the intriguing co-expression of *CvTCP1* and *CvDIV* in

adaxial petals hints that interactions among the *CYC/DICH-RAD-DIV-DRIF* genetic program may not be conserved across Brassicales and Lamiales. Further investigation of the entire program is needed to detangle both conservation and deviation from this program.

5.5 Figures

Figure 5.1. *Cleome violacea* inflorescence (A), front-plane view of flower (B), and median-plane view of flower (C).
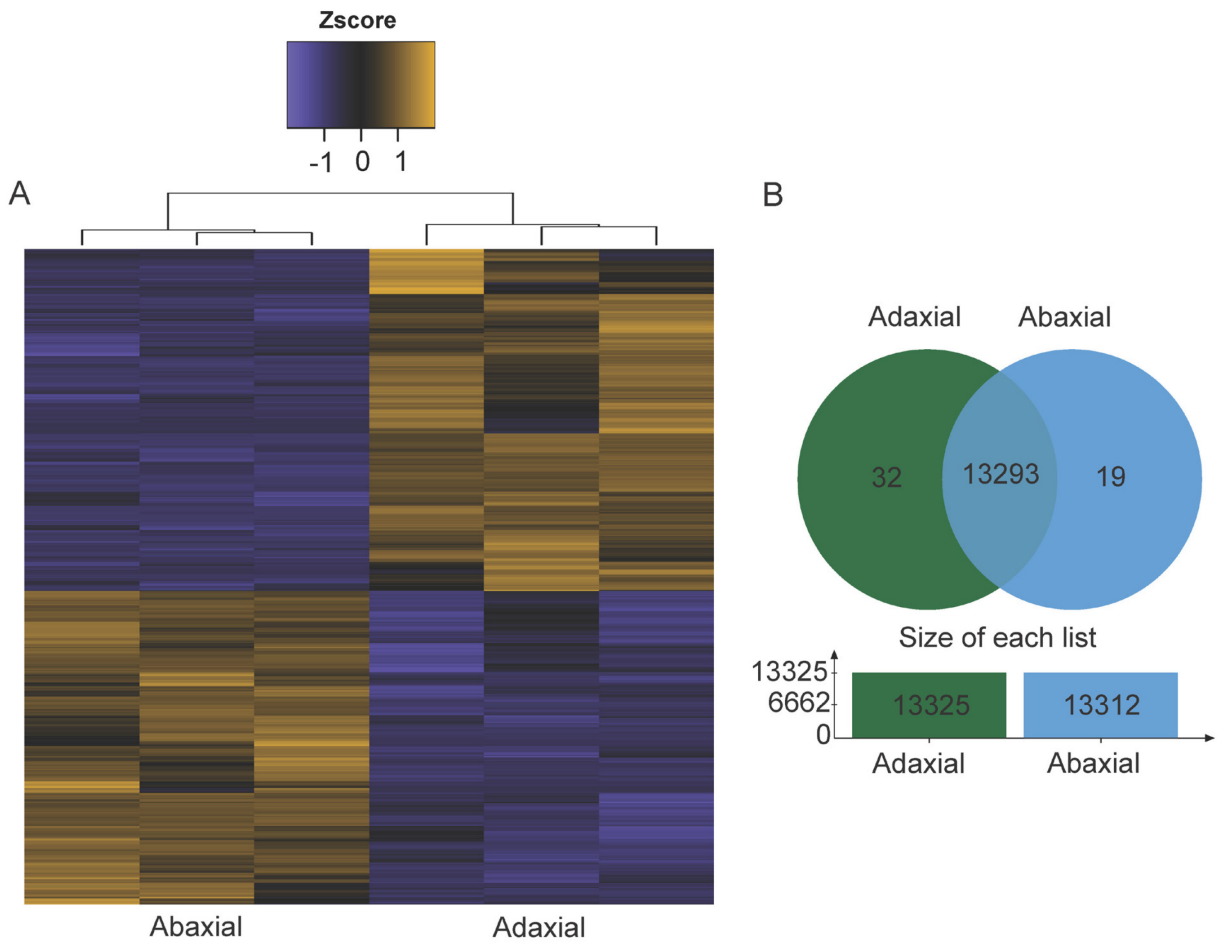
Figure 5.2. Heatmap of all significantly differentially expressed contig clustered transcripts from the *Cleome violacea* transcriptome. FDR α = 0.05; n = 517 (A). Venn diagram of all GO terms for abaxial and adaxial transcripts greater than or equal to 5 TPM (B).
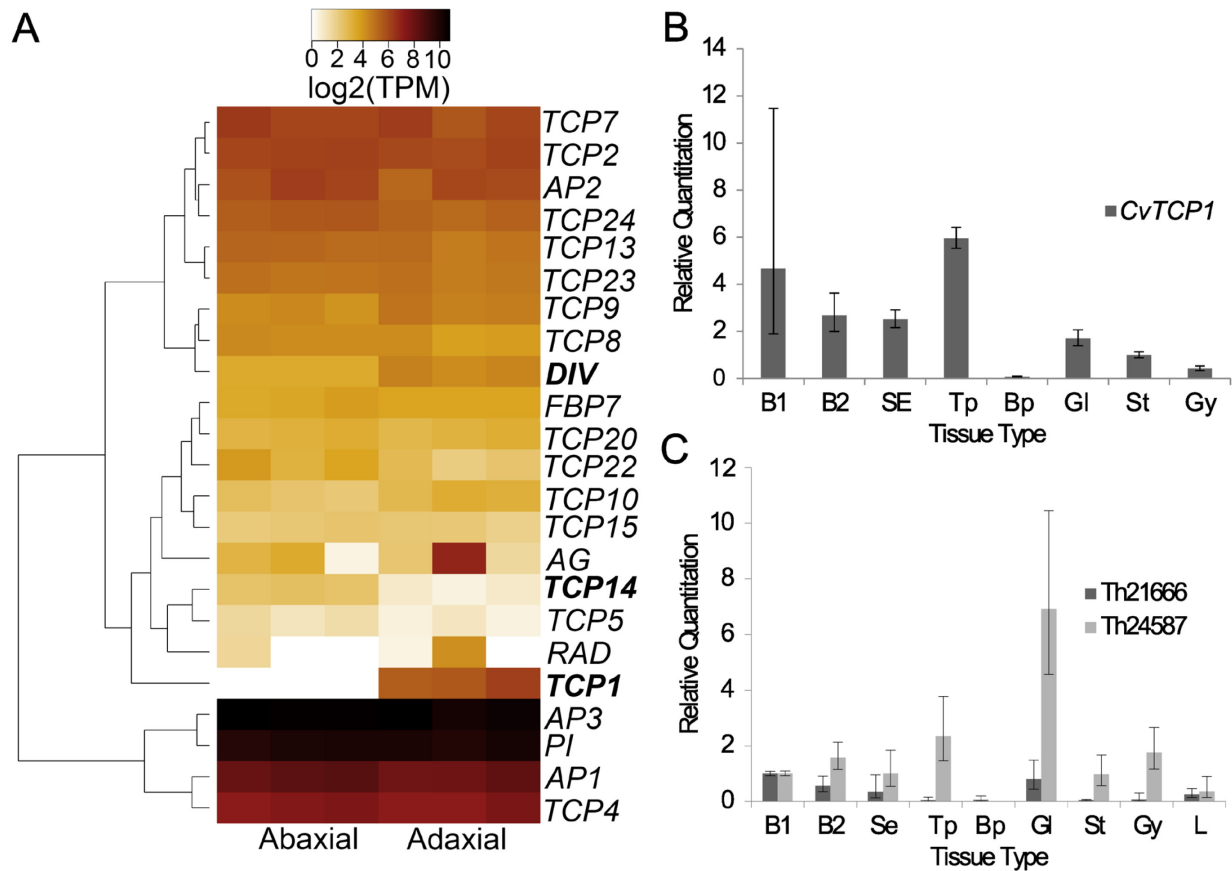
Figure 5.3. (A) Gene expression of TCP and ABC model genes from *Cleome violacea* transcriptome. (B) *CvTCP1* expression validation using qPCR across multiple tissue types in *C. violacea*. (C) *CvTCP1* homologue expression in *Tarenaya hassleriana* across multiple tissue types. B1 = young buds; B2 = mature buds; SE = sepals; TP = adaxial petals; BP = Abaxial petals; GL = nectary gland; ST = stamens; GY = gynoecium; L = leaves. Bolding indicates significant differential expression (α < 0.05).

Figure 5.4. Observed silencing phenotypes from virus induced gene silencing on *Cleome violacea* using the *pTRV2-CvTCP1* vector. (A) Untreated control flower; (B) Empty vector control; (C) Maroon flower without nectary; (D) Maroon Flower with nectary; (E) Partial maroon flower with partial nectary; (F) Reduced adaxial petal. Scale bars = 1mm.

Figure 5.5. Observed silencing phenotypes from virus induced gene silencing on *Cleome violacea* using the *pTRV2-CvTCP14* vector. Treated flowers with no secondary viral effects (A-C) and with potential secondary effects (D-E). Reduced adaxial and absent abaxial (A); Reduced and altered adaxial with reduced abaxial (B); Reduced adaxial and absent abaxial (C); absent adaxial and opposite reduced abaxial (D). severely reduced adaxial and opposite absent abaxial (E). Yellowed adaxial with absent abaxial petals (F); Untreated flower (G); Severely altered TRV2-MCS flower (H). Scale bars = 1mm
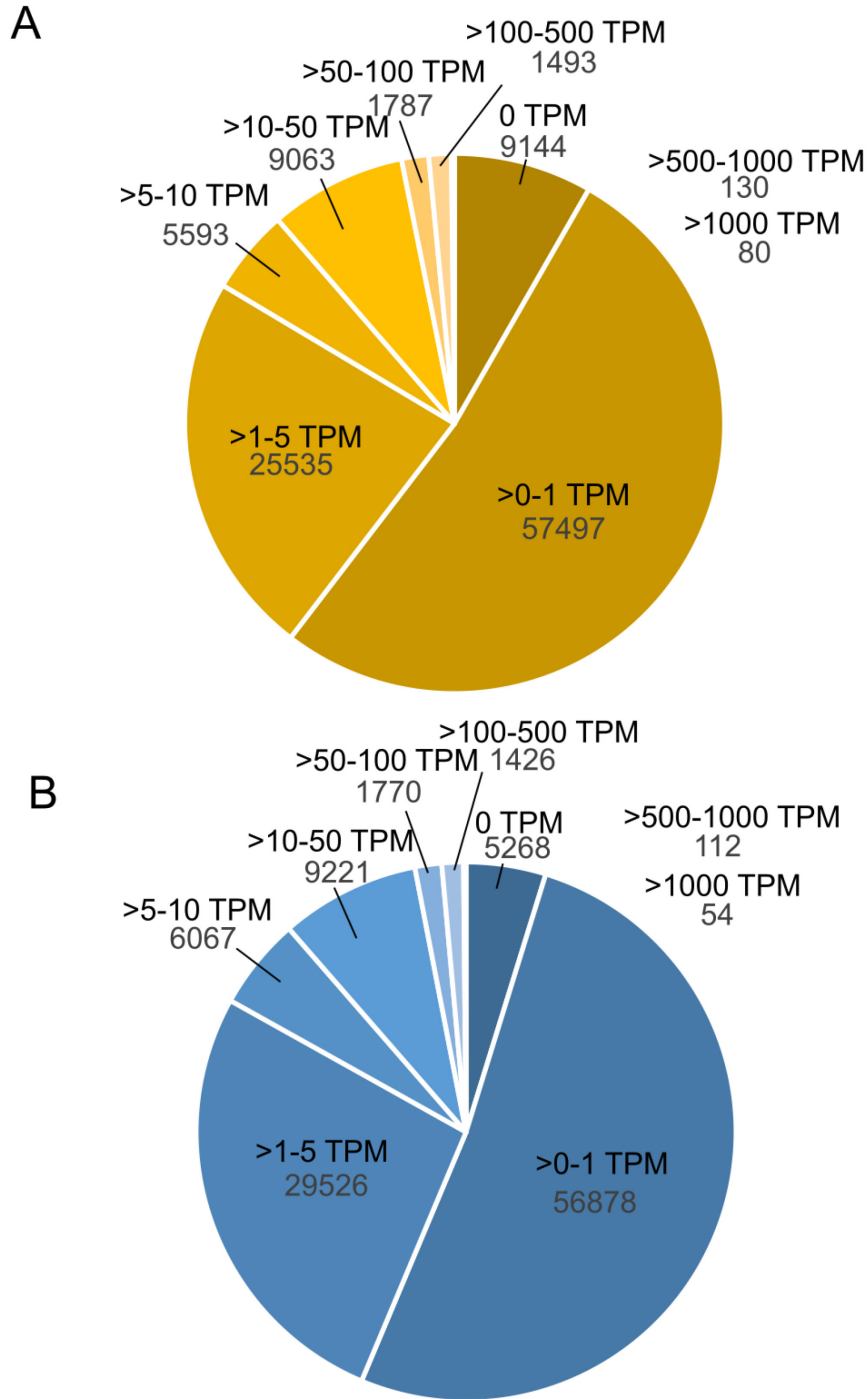
Figure S5.1. Distribution of 110,332 transcripts binned by TPM for abaxial (A) and adaxial (B) petals.

A

log2(TPM)

0 1 2 3 4 5

B

#	Gene	Accession
30;32	*CYP71B13*	AT5G25140.1
2	*ISTL9*	AT1G52315.1
23;24;25	*NSP1*	AT3G16400.1
13;20	*PERK9*	AT1G68690.1
36	*RFL2*	AT1G12300.1
7	*SAG101*	AT5G14930.2
45		AT1G19397.1
35	*SESA2*	AT4G27150.1
33	*SESA3*	AT4G27160.1
9	*STL8*	AT1G79910.1
8	*SPL9*	AT2G42200.1
10	*UMAMIT21*	AT5G64700.1
27	*UMAMIT22*	AT1G43650.1
26;29		AT2G37440.1
34		AT4G27160.1
1		AT5G05240.1
15		AT3G04360.1
4;5;6		AT2G43180.1
16		AT3G61210.1
12		AT1G55450.2
19		AT5G14930.2
18;31		AT1G28590.1
14		AT2G13350.1
17		AT5G66430.1
11		AT4G23160.2
22;28		AT2G16790.1
21		AT2G40630.1
43		AT1G63150.1
37		AT1G63080.1
46		AT1G19397.1
38;39;40		AT1G19397.1
41;42;44		AT1G19397.1
3		AT3G57910.1

Figure S5.2A-B. Heatmap of log2(TPM) transcripts from GO terms from abaxial (A-B) and adaxial petals (C-D).

C

D

#	Gene	Accession
16	*ALATS*	AT1G50200.1
3	*ALATS*	AT1G50200.2
25;27	*bHLH071*	AT5G46690.1
42	*CYP705A19*	AT3G20100.1
41	*CYP705A23*	AT3G20140.1
38;67	*DSEL*	AT4G18550.1
28	*ECH2*	AT1G76150.1
17	*ECH2*	AT1G76150.1
51;59	*ELT6*	AT3G02030.1
37;38;49;50	*ELT6*	AT3G02030.2
14;15	*EXT4*	AT1G76930.2
22;23	*GASA6*	AT1G74670.1
55	*GGL13*	AT1G75900.1
13;46	*GSTU27*	AT3G43800.1
46	*GSTU27*	AT3G43800.1
34;47	*KCS15*	AT3G52160.1
44;52	*KCS7*	AT1G71160.1
72;73	*LHB1B1*	AT2G34430.1
1;6	*MTERF7*	AT5G07900.1
18;19;24	*MTP11*	AT2G39450.1
63	*OTU1*	AT2G28120.1
58	*PIC30*	AT2G39210.1
70;71	*PPME1*	AT1G69940.1
45	*PUM7*	AT1G78160.1
10;26	*SBE2.2*	AT5G03650.1
53	*SEOA*	AT3G01670.1
43	*SEOR1*	AT3G01680.1
4	*SS5*	AT5G65685.4
7	*SS5*	AT5G65685.11
8	*SS5*	AT5G65685.12
64	*STP10*	AT3G19940.1
56;62	*STP14*	AT1G77210.1
21	*STP4*	AT3G19930.1
11;69	*SVB3*	AT5G46230.1
54;60	*SWEET3*	AT5G53190.1
40;57	*TT4*	AT5G13930.1
31;35;65;66		AT1G20120.1
39		AT5G47635.1
2		AT1G05310.1
20		AT5G19730.1
5;9		AT2G39440.1
29;61		AT1G18960.1
12		AT5G47635.1
32;68		AT1G06250.1
33;36		AT5G17340.1
30		AT1G35730.2

Figure S5.2C-D. Heatmap of log2(TPM) transcripts from GO terms from abaxial (A-B) and adaxial petals (C-D).

5.6 Tables

Table 5.1. List of primers used for VIGS construct design and RT-qPCR in this study. Brackets indicate restriction sites for incorporation into the pTRV2 vector. Extra base pairs upstream of restriction sites are binding sequences.

Primer	Forward (5' – 3')'	Reverse (5' – 3')	Probe (5' – 3')
TCP1 qPCR	GCGGCTAGGAGGAAGATCAA	TGCTAGAACTCTGATCCTTCATCATAA	56FAM/CATCAGCAC/ZEN/CGTTGCTTCCGAGC/3IABkFQ/
TCP1 cDNA	GAGTCTGGTGAA CGGTGGAT	AGTCCTCATCAAAGGGTGCA	
TCP1 VIGS R/E	CTCT[GGATCC]GAGTCTGGTGAACGG TGGAT	CTCT[TCTAGA]AGTCCTCATCAAAGG GTGCA	
TCP1 Iberis amara	ACAATGGAGTGTACCCTCTCTCTCTTTACC	TTATAGTTGCTGCTAGAACTCTGSTCTAC	
TCP14 VIGS R/E	CTCT[TCTAG]AGCCTCTCTCTTCTCCCTCCA	CTCT[CTCGAG]CACTGTGGCTCCTCCTGAAG	
ACTIN qPCR control	TGACTCGGGTGATGGTGTGT	AAGATCAAGACGAAGGATAGCATGT	
156	TTACTCAAGGAAGCACGATGAGC	GAACCGTAGTTTAATGTCTTCGGG	56FAM/CACTGTTCC/ZEN/AATTTACGAGGGTTATGCCC/3IABkFQ/
M13	d[GTAAAACGACGGCCAG]	d[CAGGAAACAGCTATGAC]	

Table 5.2. Software versions and references for all programs used in this study.

Software	Version	References
Trim Galore	0.6.6	Krueger 2012
FastQC	0.11.9	Andrews 2010
Trinity	2.12	Manfred et al., 2013
Corset	1.09	Nadia & Oshlack 2014
edgeR	3.32.1	Robinson et al., 2009
R	4.0.5	R Core Team (2021)
Transdecoder	5.5.0	https://github.com/TransDecoder
Orthofinder	2.5.2	Emms & Kelly 2015
BLAST+	2.2.31	Altschul et al., 1990
Geneious	11.09	"Geneious 11.09 (https://www.geneious.com)"
OrfPredictor	2.3	X. J. Min et al., 2005
SiFi21	1.2.3	Lück et al., 2019

Table 5.3. Top 50 shared terms between adaxial and abaxial petals in an OrthVenn GO analysis of Trinity transcripts with TPM expression greater than or equal to 5.

GO #	Name	#
GO:0008150	biological process	4148
GO:0008152	metabolic process	2427
GO:0009987	cellular process	1853
GO:0044237	cellular metabolic process	1766
GO:0043170	macromolecule metabolic process	1455
GO:0050896	response to stimulus	1397
GO:0065007	biological regulation	1060
GO:0044238	primary metabolic process	1055
GO:0006807	nitrogen compound metabolic process	963
GO:0016070	RNA metabolic process	822
GO:0006725	cellular aromatic compound metabolic process	696
GO:0046483	heterocycle metabolic process	694
GO:0032502	developmental process	626
GO:0006139	nucleobase compound metabolic process	572
GO:0032501	multicellular organismal process	524
GO:0051234	establishment of localization	475
GO:0006810	transport	463
GO:0006464	cellular protein modification process	431
GO:0016043	cellular component organization	386
GO:0006396	RNA processing	366
GO:0019538	protein metabolic process	363
GO:0006082	organic acid metabolic process	342
GO:0005975	carbohydrate metabolic process	328
GO:0006793	phosphorus metabolic process	320

Table S5.3 Continued

GO:0006412	translation	278
GO:0051179	localization	265
GO:0006508	proteolysis	252
GO:0006629	lipid metabolic process	251
GO:0000003	reproduction	243
GO:0007154	cell communication	240
GO:0015031	protein transport	217
GO:0044255	cellular lipid metabolic process	217
GO:0051641	cellular localization	207
GO:0016192	vesicle-mediated transport	191
GO:0043412	macromolecule modification	189
GO:0051704	multi-organism process	171
GO:0005976	polysaccharide metabolic process	148
GO:0051186	cofactor metabolic process	142
GO:0006996	organelle organization	138
GO:0040007	growth	132
GO:0006518	peptide metabolic process	128
GO:0006811	ion transport	127
GO:0042254	ribosome biogenesis	126
GO:0043603	cellular amide metabolic process	124
GO:0032989	cellular component morphogenesis	111
GO:0034622	cellular macromolecular complex assembly	93
GO:0009117	nucleotide metabolic process	84
GO:0015979	photosynthesis	84
GO:0007049	cell cycle	79
GO:0016049	cell growth	77

Table 5.4. GO analysis cluster list of unique abaxial and adaxial terms from a list of transcripts greater than or equal to 5TPM.

Cluster #	Transcripts (#)	SwissProt ID	GO #	GO Annotation
Abaxial				
cluster16	8	N/A	N/A	N/A
cluster953	3	O49326	GO:0080028	P:nitrile biosynthetic process
cluster960	3	Q0QLE4	GO:0051187	P:cofactor catabolic process
cluster1025	3	Q0WKV3	GO:0090615	P:mitochondrial mRNA processing
cluster1092	3	P17333	GO:0045735	F:nutrient reservoir activity
cluster1496	2	Q9FXJ2	GO:0016042	P:lipid catabolic process
cluster1497	2	N/A	N/A	N/A
cluster1498	2	Q9SX33	GO:0045332	P:phospholipid translocation
cluster1499	2	Q6NMB7	GO:0022857	F:transmembrane transporter activity
cluster1500	2	Q6XMI3	GO:0009611	P:response to wounding
cluster1501	2	Q700W2	GO:0006355	P:regulation of transcription, DNA-templated
cluster1502	2	Q55EX9	GO:0008168	F:methyltransferase activity
cluster1503	2	Q0WT19	GO:0046856	P:phosphatidylinositol dephosphorylation
cluster1504	2	P58050	GO:0016709	F:oxidoreductase activity
cluster1505	2	N/A	N/A	N/A
cluster1506	2	Q4F883	GO:1900057	P:positive regulation of leaf senescence
cluster1507	2	N/A	N/A	N/A
cluster1508	2	N/A	N/A	N/A
cluster11872	2	Q07488	GO:0006979	P:response to oxidative stress

Table 5.4 Continued

Adaxial				
cluster127	6	Q9LW26	GO:0019432	P:triglyceride biosynthetic process
cluster205	5	P0DKJ6	GO:0016042	P:lipid catabolic process
cluster1243	3	O80632	GO:0010042	P:response to manganese ion
cluster1245	3	Q0WVX5	GO:0005982	P:starch metabolic process
cluster9197	2	Q84WM7	GO:0009860	P:pollen tube growth
cluster9198	2	Q93XX2	GO:0010088	P:phloem development
cluster9199	2	O49523	GO:0010187	P:negative regulation of seed germination
cluster9200	2	Q9SUY9	GO:0006633	P:fatty acid biosynthetic process
cluster9201	2	P36428	GO:0006400	P:tRNA modification
cluster9202	2	P30924	GO:0019252	P:starch biosynthetic process
cluster9203	2	F4J2K4	GO:0007018	P:microtubule-based movement
cluster9204	2	Q9LNC2	GO:0016042	P:lipid catabolic process
cluster9205	2	N/A	N/A	N/A
cluster9206	2	Q9LSI7	GO:0055046	P:microgametogenesis
cluster9207	2	Q56XR0	GO:0006355	P:regulation of transcription, DNA-templated
cluster9208	2	Q8GW61	GO:0005351	F:carbohydrate:proton symporter activity
cluster9209	2	Q9C992	GO:0006633	P:fatty acid biosynthetic process
cluster9210	2	P24825	GO:0009813	P:flavonoid biosynthetic process
cluster9211	2	N/A	N/A	N/A
cluster9212	2	Q9LZG7	GO:0009407	P:toxin catabolic process
cluster9213	2	P13851	GO:0018298	P:protein-chromophore linkage
cluster9214	2	O23038	GO:0045490	P:pectin catabolic process
cluster9215	2	Q39228	GO:0015770	P:sucrose transport
cluster9216	2	N/A	N/A	N/A
cluster9217	2	N/A	N/A	N/A
cluster9218	2	N/A	N/A	N/A
cluster9219	2	Q6NQN5	GO:0051119	F:sugar transmembrane transporter activity
cluster9220	2	N/A	N/A	N/A
cluster9221	2	N/A	N/A	N/A
cluster9222	2	Q9C9R6	GO:0006417	P:regulation of translation
cluster11870	2	Q9LJY5	GO:0009958	P:positive gravitropism
cluster11871	2	N/A	N/A	N/A

Table 5.5. Top 50 shared terms between adaxial and abaxial petals in a KEGG analysis of Trinity transcripts with TPM expression greater than or equal to 5.

KEGG #	Name	Abaxial	Adaxial
3010	Ribosome	110	109
3040	Spliceosome	90	90
4141	Protein processing in endoplasmic reticulum	73	73
3013	Nucleocytoplasmic transport	64	66
4714	Thermogenesis	59	59
190	Oxidative phosphorylation	55	54
4144	Endocytosis	54	54
4120	Ubiquitin mediated proteolysis	52	51
3015	mRNA surveillance pathway	48	50
4138	Autophagy - yeast	47	46
270	Cysteine and methionine metabolism	43	43
3008	Ribosome biogenesis in eukaryotes	42	44
230	Purine metabolism	41	42
3018	RNA degradation	41	41
4075	Plant hormone signal transduction	38	36
520	Amino sugar and nucleotide sugar metabolism	37	37
564	Glycerophospholipid metabolism	37	37
3050	Proteasome	37	37
4146	Peroxisome	36	35
10	Glycolysis / Gluconeogenesis	35	35
4142	Lysosome	34	35
620	Pyruvate metabolism	33	33
860	Porphyrin metabolism	33	33
4016	MAPK signaling pathway - plant	33	33
260	Glycine, serine and threonine metabolism	32	32

Table 5.5 Continued.

4110	Cell cycle	31	30
500	Starch and sucrose metabolism	30	30
510	N-Glycan biosynthesis	30	30
561	Glycerolipid metabolism	30	29
630	Glyoxylate and dicarboxylate metabolism	28	29
3420	Nucleotide excision repair	28	27
240	Pyrimidine metabolism	27	28
900	Terpenoid backbone biosynthesis	27	27
970	Aminoacyl-tRNA biosynthesis	27	27
3022	Basal transcription factors	27	27
250	Alanine,aspartate and glutamate metabolism	26	27
562	Inositol phosphate metabolism	26	26
4111	Cell cycle - yeast	26	25
4150	mTOR signaling pathway	26	26
4626	Plant-pathogen interaction	26	26
195	Photosynthesis	25	24
710	Carbon fixation in photosynthetic organisms	25	25
3020	RNA polymerase	25	26
4145	Phagosome	24	24
4114	Oocyte meiosis	23	22
4152	AMPK signaling pathway	23	23
280	Valine, leucine and isoleucine degradation	22	22
330	Arginine and proline metabolism	22	22
400	Phenylalanine,tyrosine and tryp biosynthesis	22	23
513	Various types of N-glycan biosynthesis	22	23

Table 5.6. All KEGG terms for Significantly differentially expressed transcripts and transcripts greater than 100 TPM sorted by petal type.

KEGG #	Description	KO Count			
		SDE		100TPM	
		Ab	Ad	Ab	Ad
195	Photosynthesis	17	1	21	20
196	Photosynthesis - antenna proteins	10	1	9	9
630	Glyoxylate and dicarboxylate metabolism	6	0	13	12
4075	Plant hormone signal transduction	5	4	9	8
500	Starch and sucrose metabolism	4	2	10	7
710	Carbon fixation in photosynthetic organisms	4	0	18	16
20	Citrate cycle	3	0	13	11
52	Galactose metabolism	3	0	4	4
190	Oxidative phosphorylation	3	0	50	50
380	Tryptophan metabolism	3	0	7	5
592	alpha-Linolenic acid metabolism	3	2	6	5
620	Pyruvate metabolism	3	0	16	13
940	Phenylpropanoid biosynthesis	3	1	13	13
4146	Peroxisome	3	1	8	7
10	Glycolysis / Gluconeogenesis	2	0	18	17
40	Pentose and glucuronate interconversions	2	2	4	6
53	Ascorbate and aldarate metabolism	2	0	11	10
260	Glycine, serine and threonine metabolism	2	2	11	11
270	Cysteine and methionine metabolism	2	1	16	14
350	Tyrosine metabolism	2	0	3	2
460	Cyanoamino acid metabolism	2	1	3	3
520	Amino sugar and nucleotide sugar metabolism	2	1	16	13
561	Glycerolipid metabolism	2	1	5	5
591	Linoleic acid metabolism	2	1	2	2
680	Methane metabolism	2	0	7	7
860	Porphyrin and chlorophyll metabolism	2	0	10	10
900	Terpenoid backbone biosynthesis	2	0	11	11
941	Flavonoid biosynthesis	2	0	10	10
30	Pentose phosphate pathway	1	0	9	8
71	Fatty acid degradation	1	0	6	6

Table 5.6 Continued.

250	Alanine, aspartate and glutamate metabolism	1	0	7	5
480	Glutathione metabolism	1	0	10	10
564	Glycerophospholipid metabolism	1	2	3	4
625	Chloroalkane and chloroalkene degradation	1	0	2	2
626	Naphthalene degradation	1	0	1	1
643	Styrene degradation	1	0	1	0
670	One carbon pool by folate	1	0	3	3
720	Carbon fixation pathways in prokaryotes	1	0	4	4
740	Riboflavin metabolism	1	0	1	0
830	Retinol metabolism	1	0	1	1
910	Nitrogen metabolism	1	0	3	2
944	Flavone and flavonol biosynthesis	1	0	2	2
945	Stilbenoid, diarylheptanoid and gingerol biosynthesis	1	0	4	4
4016	MAPK signaling pathway - plant	1	0	3	3
4111	Cell cycle - yeast	1	0	4	4
4120	Ubiquitin mediated proteolysis	1	0	10	10
4138	Autophagy - yeast	1	0	5	5
4139	Mitophagy - yeast	1	0	1	1
4141	Protein processing in endoplasmic reticulum	1	0	20	20
4142	Lysosome	1	0	12	12
4150	mTOR signaling pathway	1	0	10	10
4210	Apoptosis	1	0	5	5
230	Purine metabolism	0	1	7	7
280	Valine, leucine and isoleucine degradation	0	2	7	8
290	Valine, leucine and isoleucine biosynthesis	0	1	5	6
310	Lysine degradation	0	1	3	2
565	Ether lipid metabolism	0	2	1	2
590	Arachidonic acid metabolism	0	1	1	1
640	Propanoate metabolism	0	1	6	5
770	Pantothenate and CoA biosynthesis	0	1	6	6
906	Carotenoid biosynthesis	0	1	7	7
4144	Endocytosis	0	1	14	13

Table 5.7. All GO terms for significantly differentially expressed transcripts and transcripts greater than 100 TPM sorted by petal type.

GO#	Description	SDE			100TPM		
		Ab	Ad	Both	Ab	Ad	Both
8150	biological process	18	6	9	2	2	566
8152	metabolic process	7	1	5	1	1	299
50896	response to stimulus	9	2	4	1	1	235
43170	macromolecule metabolic process	3	1	5	0	1	210
44237	cellular metabolic process	5	1	4	1	1	210
9987	cellular process	5	3	6	1	1	194
44238	primary metabolic process	2	1	3	0	0	124
6412	translation	0	0	0	0	0	102
6807	nitrogen compound metabolic process	2	1	2	1	0	91
65007	biological regulation	5	2	2	0	0	79
6082	organic acid metabolic process	1	0	0	0	0	64
32502	developmental process	2	2	1	0	0	64
19538	protein metabolic process	1	0	1	0	0	63
6725	cellular aromatic compound metabolic process	1	1	2	0	0	57
46483	heterocycle metabolic process	1	1	2	1	0	55
51234	establishment of localization	2	3	2	0	0	54
6810	transport	1	3	2	0	0	52
32501	multicellular organismal process	2	0	0	0	0	52
6629	lipid metabolic process	2	0	0	0	0	46
44255	cellular lipid metabolic process	1	0	0	0	0	44
6139	nucleobase-containing compound metabolic process	1	1	2	0	0	41
16043	cellular component organization	0	0	0	0	0	39
6508	proteolysis	1	0	1	0	0	36
16070	RNA metabolic process	1	1	2	0	0	36
5975	carbohydrate metabolic process	1	1	2	0	1	35
6793	phosphorus metabolic process	0	0	1	1	0	34
15979	photosynthesis	2	0	0	0	0	29

Table 5.7 Continued.

51179	localization	2	0	1	0	0	26
6464	cellular protein modication process	2	0	1	0	0	25
6811	ion transport	1	0	0	0	0	23
51704	multi-organism process	3	1	0	0	1	22
6091	generation of precursor metabolites and energy	1	0	0	0	0	21
3	reproduction	0	0	0	0	0	21
43603	cellular amide metabolic process	0	0	0	0	0	21
6518	peptide metabolic process	0	0	0	0	0	20
7154	cell communication	1	0	1	0	0	19
51641	cellular localization	0	0	1	0	0	19
45333	cellular respiration	0	0	0	0	0	19
15031	protein transport	0	1	1	0	0	18
51186	cofactor metabolic process	0	0	0	1	0	18
5976	polysaccharide metabolic process	1	0	2	0	1	17
40007	growth	0	1	0	0	0	17
19748	secondary metabolic process	1	0	0	0	0	15
6996	organnelle organization	0	0	1	0	0	14
9117	nucleotide mtabolic process	0	0	0	1	0	12
32989	cellullar component morphogenesis	0	0	1	0	0	12
16049	cell growth	0	1	0	0	0	11
71555	cell wall organization	0	1	0	0	1	10
45229	external encapsulating structure organization	0	1	0	0	0	8
2376	immune system process	0	1	0	1	1	6
7049	cell cycle	0	0	1	0	0	6
6081	cellular aldehyde metabolic process	0	0	0	1	0	5
6869	lipid transport	0	0	0	0	1	3
15893	drug transport	1	0	0	0	0	1
15849	organic acid transport	0	0	1	0	0	1
51276	chromosome organization	0	0	1	0	0	0
7059	chromosome segregation	0	0	1	0	0	0
6865	amino acid transport	0	0	1	0	0	0

Table 5.8. CoGe BLAST of *CvTCP1* and other *TCP1* homologues to the draft genome.

HSP No.	Query	Organism	Scaffold	Position	Length	% ID	E-value	Bit-Score
1	*CvTCP1*	*Cleome violacea* (ID 23822)	169	290877	1079	99.6	0	2052
2	*CvTCP1*	*Cleome violacea* (ID 23822)	59	353400	201	89	7.00E-62	241
3	*CvTCP1*	*Cleome violacea* (ID 23822)	59	353088	109	80.7	3.00E-13	79.5
4	*CvTCP1*	*Cleome violacea* (ID 23822)	3	134182	63	84.1	0.00000001	64.1
5	*CvTCP1*	*Cleome violacea* (ID 23822)	205	48009	48	91.6	0.0000002	60.3
6	*CvTCP1*	*Cleome violacea* (ID 23822)	6	1226392	97	77.3	0.0000002	60.3
7	*CvTCP1*	*Cleome violacea* (ID 23822)	197	120996	143	73.4	0.000003	56.4
8	*CvTCP1*	*Cleome violacea* (ID 23822)	105	104017	35	94.2	0.000003	56.4
9	*CvTCP1*	*Cleome violacea* (ID 23822)	46	449906	56	83.9	0.000003	56.4
10	*CvTCP1*	*Cleome violacea* (ID 23822)	17	96151	38	92.1	0.000003	56.4
1	*Th21666*	*Cleome violacea* (ID 23822)	169	290633	273	89	3.00E-92	342
2	*Th21666*	*Cleome violacea* (ID 23822)	59	353380	184	84.7	2.00E-38	164
3	*Th21666*	*Cleome violacea* (ID 23822)	169	290923	84	86.9	8.00E-19	98.7
4	*Th21666*	*Cleome violacea* (ID 23822)	169	289832	149	77.8	2.00E-15	87.2
5	*Th21666*	*Cleome violacea* (ID 23822)	6	1226354	69	85.5	7.00E-12	75.7
6	*Th21666*	*Cleome violacea* (ID 23822)	197	120996	143	74.1	0.000000008	62.2

Table 5.8 Continued.

7	*Th21666*	*Cleome violacea* (ID 23822)	169	290698	52	88.4	0.000004	56.4
8	*Th21666*	*Cleome violacea* (ID 23822)	169	290785	65	81.5	0.000004	56.4
1	*Th24587*	*Cleome violacea* (ID 23822)	169	290698	306	85.9	5.00E-83	312
2	*Th24587*	*Cleome violacea* (ID 23822)	59	353353	153	85.6	1.00E-36	158
3	*Th24587*	*Cleome violacea* (ID 23822)	169	289797	103	89.3	9.00E-30	135
4	*Th24587*	*Cleome violacea* (ID 23822)	169	290923	87	90.8	1.00E-25	121
5	*Th24587*	Cleome violacea (ID 23822)	6	1226392	109	78.8	2.00E-12	77.6
6	*Th24587*	*Cleome violacea* (ID 23822)	169	290285	80	82.5	3.00E-11	73.7
7	*Th24587*	*Cleome violacea* (ID 23822)	59	353200	83	81.9	3.00E-11	73.7
8	*Th24587*	*Cleome violacea* (ID 23822)	59	353088	56	87.5	2E-09	68
9	*Th24587*	*Cleome violacea* (ID 23822)	59	353462	56	85.7	9E-08	62.2
10	*Th24587*	*Cleome violacea* (ID 23822)	197	120996	143	74.1	9E-08	62.2
1	*Th_X1*	*Cleome violacea* (ID 23822)	169	290633	541	80.5	7.00E-100	367
2	*Th_X1*	*Cleome violacea* (ID 23822)	169	290923	581	78.3	2.00E-92	342
3	*Th_X1*	*Cleome violacea* (ID 23822)	59	353380	184	84.7	2.00E-38	164
4	*Th_X1*	*Cleome violacea* (ID 23822)	59	353195	106	82	1.00E-17	94.9
5	*Th_X1*	*Cleome violacea* (ID 23822)	169	289832	149	77.8	2.00E-15	87.2
6	*Th_X1*	*Cleome violacea* (ID 23822)	6	1226354	69	85.5	7.00E-12	75.7
7	*Th_X1*	*Cleome violacea* (ID 23822)	197	120996	143	74.1	8E-08	62.2
8	*Th_X1*	*Cleome violacea* (ID 23822)	53	142066	56	85.7	8E-08	62.2
9	*Th_X1*	*Cleome violacea* (ID 23822)	59	353088	58	84.4	3E-07	60.3
10	*Th_X1*	*Cleome violacea* (ID 23822)	3	134186	45	88.8	0.000001	58.4
1	*Th_x2*	*Cleome violacea* (ID 23822)	169	290633	541	80.5	7.00E-100	367
2	*Th_x2*	*Cleome violacea* (ID 23822)	169	290923	581	78.3	2.00E-92	342
3	*Th_x2*	*Cleome violacea* (ID 23822)	59	353380	184	84.7	2.00E-38	164
4	*Th_x2*	*Cleome violacea* (ID 23822)	59	353195	106	82	1.00E-17	94.9

Table 5.8 Continued.

5	*Th_x2*	*Cleome violacea* (ID 23822)	169	289832	135	78.5	3.00E-14	83.4
6	*Th_x2*	*Cleome violacea* (ID 23822)	6	1226354	69	85.5	7.00E-12	75.7
7	*Th_x2*	*Cleome violacea* (ID 23822)	197	120996	143	74.1	8E-08	62.2
8	*Th_x2*	*Cleome violacea* (ID 23822)	53	142066	56	85.7	8E-08	62.2
9	*Th_x2*	*Cleome violacea* (ID 23822)	59	353088	58	84.4	3E-07	60.3
10	*Th_x2*	*Cleome violacea* (ID 23822)	3	134186	45	88.8	0.000001	58.4
1	*AtTCP1*	*Cleome violacea* (ID 23822)	169	290582	191	84.8	1.00E-49	200
2	*AtTCP1*	*Cleome violacea* (ID 23822)	59	353330	128	84.3	5.00E-26	121
3	*AtTCP1*	*Cleome violacea* (ID 23822)	6	1226423	107	76.6	5E-08	62.2
1	*TCP14*	*Cleome violacea* (ID 23822)	28	504596	463	100	0	890
2	*TCP14*	*Cleome violacea* (ID 23822)	112	153127	222	83.7	7.00E-56	219
3	*TCP14*	*Cleome violacea* (ID 23822)	6	646164	236	82.6	3.00E-55	217
4	*TCP14*	*Cleome violacea* (ID 23822)	18	621801	192	84.3	6.00E-49	196
5	*TCP14*	*Cleome violacea* (ID 23822)	772	3792	216	81	1.00E-43	179
6	*TCP14*	*Cleome violacea* (ID 23822)	0	2053143	195	77.9	4.00E-28	127
7	*TCP14*	*Cleome violacea* (ID 23822)	66	407024	187	78	6.00E-27	123
8	*TCP14*	*Cleome violacea* (ID 23822)	105	257602	192	77.6	2.00E-26	121
9	*TCP14*	*Cleome violacea* (ID 23822)	2	1943809	165	78.1	7.00E-23	110
10	*TCP14*	*Cleome violacea* (ID 23822)	120	228985	177	76.8	4.00E-21	104

Table 5.9. Putative *TCP1* and *DIVARICATA* binding sequences upstream of the start site on scaffold 169 and 59. Bolding indicates conserved sequences. Brackets indicate the consensus binding sequence for all *TCPs*.

Putative Gene	Scaffold	Sequence	Strand	BP from start site
TCP1	169	AA(**GGTCCC**)TT	Sense	7,325
		AA(**GGTCCC**)AT	Sense	9,145
	59	AT(**GGTCCC**)AC	Sense	681
		GA(**GGTCCC**)GA	Sense	9,351
		TG(**GGCCC**)AA	Sense	9,142
DIV	169	A**GATA**AGA	Antisense	5,852
		G**GATA**ACC	Antisense	1,834
		G**GATA**CGG	Sense	6,802
		G**GATA**CCG	Antisense	8,538
		C**GATA**AGA	Antisense	7,295
	59	A**GATA**AGG	Antisense	9,258
		A**GATA**CCA	Sense	3,690
		G**GATA**AGG	Antisense	7,967
		G**GATA**ACA	Antisense	9,195
		G**GATA**ACA	Antisense	1,546
		G**GATA**CCA	Antisense	6,122

Table 5.10. Percent silencing-efficacy and all observed phenotypes for virus induced gene silencing in *Cleome violacea* treated with pTRV2-*CvTCP1* and pTRV2-*CvTCP14.*

Silencing-efficacy Data	*TCP1*	*TCP14*
Total Plants	124	100
Plants displaying phenotype (#)	42	40
Floral phenotypes		
Maroon	26	n/a
Reduced Adaxial	16	
Reduction or Absence of 1 or 2 adaxial petals	n/a	39
Reduction or Absence of 1 or 2 abaxial petals		20
Reduction or Enlargement of 1 or 2 adaxial spots		6
Increased ratio of yellowing in adaxial petals		6
Increased ratio of yellowing in abaxial petals		2
Increased ratio of yellowing in whole flower		10
Undistinguishable from control pTRV2 group		23

Table 5.11. Assembly statistics for *Cleome violacea* adaxial and abaxial transcriptome.

Contig	All Transcripts	Longest Isoform
N10	5035	4459
N20	3917	3443
N30	3248	2836
N40	2766	2408
N50	2398	2029
Median Length	1095	404
Average Length	1472.38	945.49
Total Bases	162435466	52655541
Total Genes	55691	
Total Transcripts	110322	
GC %	44.18	

Table 5.12. Benchmarking Universal Single Copy Orthologs results indicating overall completeness of the *Cleome violacea* transcriptome.

BUSCO Summary	*Cleome violacea*
Complete BUSCOs (C)	2046 (96.4%)
Complete and Single Copy BUSCOs (S)	862
Complete and Duplicated BUSCOs (D)	1184
Fragmented BUSCOs (F)	42 (2.0%)
Missing BUSCOs (M)	33 (1.6%)
Total BUSCO Groups Searched	2121

Chapter 6: Floral Pigmentation in *Cleome violacea* and *Gynandropsis gynandra*.

6.1 Introduction

Pigmentation is one of many facets of floral morphology that directly affects pollination and subsequent reproduction (Galliot et al., 2006). It does this through variation of colors and patterns in petals/sepals (e.g., Hydrangea) (Miller et al., 2011) and reproductive whorls (e.g., Orchids) (Yu and Goh, 2001). Flowers can be categorized into separate pollination syndromes based on shared features and the types of pollinators they commonly attract (Galliot et al., 2006). Although many exceptions exist, pollination syndromes are an excellent beginning for understanding patterns of attraction because they are an adaptive response to pollinators (Macior, 1971). There are eight classified pollination syndromes (Fenster et al., 2004),e.g., bird adapted flowers are red because most birds are tetrachromats (Hart and Hunt, 2007) and thus sensitive to red flowers which offer a strong contrast against green leaves. There is positive selective pressure for genes that promote specific pollinator and floral rewards, although pollinators are not always the primary agent (Souto-Vilarós et al., 2018). Notwithstanding, pollination is directly linked to at least some of the vast array of pigmentation seen in nature (Miller et al., 2011).

Most of the pigment diversity observed in nature is brought about by combinations of anthocyanins (i.e., a class of flavonoids) and carotenoids (Grotewold, 2006). Betalains are a third type of pigment, but only found within the Caryophyllales, in lieu of anthocyanins (Tanaka et al., 2008). There are over 9,000 flavonoid compounds found in all land plants except Hornworts (Wang et al., 2011; Davies et al., 2020), but only a few subclasses contribute to floral pigmentation in angiosperms. The primary subclasses are the anthocyanins and anthocyanidins, which present hues of yellow, red, blue and purple (Grotewold, 2006). In total there are six major types of anthocyanidins: cyanidin (red), pelargonidin (orange), delphinidin (blue) and their O-methylated derivatives, peonidin, malvidin, and petunidin, respectively (Khoo et al., 2017). These pigments are often aqueous and found in vacuoles within petal tissue (Grotewold, 2006). Anthocyanins are like anthocyanidins, but with an additional sugar moiety that increases

their stability (Yonekura-Sakakibara et al., 2019). Anthocyanins and anthocyanidins are sensitive to pH (Khoo et al., 2017), and their hues can change with varying acidity from red to purple and blue to blue-green; they also become colorless in strongly alkaline conditions (Khoo et al., 2017).

Post-translational regulation via changes in pH is one of a few ways pigment diversity is created (Hichri et al., 2011). This regulation can also be measured indirectly via mRNA quantification of intracellular pH transport pathways (Pittman, 2012). Flavanols, another subclass, are found in some pale-yellow flowers, alongside other flavonoids (Dudek et al., 2016); they are thought to have a photoprotective quality, as observed in geographic variants of poppy (Yonekura-Sakakibara et al., 2019; Dudek et al., 2020). Altogether, more than 30 anthocyanidins and 400 anthocyanins have been identified (Kong et al., 2003).

Phenotypically, petal colour is altered via changes to pigment intensity, hue, or patterning (e.g., petal spots). Genotypically, mutations occur either directly to anthocyanin/carotenoid pathway enzymes, to their regulatory transcription factors (TFs), or to the cis-regulatory elements (CREs) bound by TFs (Streisfeld and Rausher, 2011). Other outcomes, such as direct loss-of-function, are less common (i.e., less likely to be fixed in a population) because they have greater deleterious effects in the organism overall (Sobel and Streisfeld, 2013). A major goal in evolutionary developmental biology is understanding how these changes in genotype affect diversity of phenotype. One way of measuring this is via mRNA quantification because mutations to either TFs or CREs change the transcription rates of pathway enzymes. To understand the genetic underpinnings of diversity within and between white and pigmented flowers, we characterized expression of relevant biosynthetic pathways.

Much of the anthocyanin biosynthesis pathway (ABP) is modulated, positively and negatively, by the R2R3-MYB transcription factor family (Sobel and Streisfeld, 2013; Hsu et al., 2015; Sagawa et al., 2016; Kramer et al., 2017; Su et al., 2017; Wang et al., 2019; Xi et al., 2019; Duncan and Rausher, 2020; Zhong et al., 2020). For example, transgenic overexpression of *MYB6* and *MYB10* results in a significant increase of anthocyanin content in *Populus tomentosa* (Salicaceae) (Wang et al., 2019), *Prunus armeniaca* (Rosaceae) (Xi et al., 2019), and

Phaelenopsis spp. (Orchidaceae) (Hsu et al., 2015)*.* Contrastingly, overexpression of *MYB1* suppresses anthocyanin content in *Gerbera hybrida* (Asteraceae) (Zhong et al., 2020), and is known to activate spot formation in *Clarkia* (Onagraceae) (Martins et al., 2017). In addition to regulation of anthocyanins, there is also evidence that R2R3-MYB regulation extends to carotenoids. *CAROTENOID PIGMENTATION 1* (*RCP1*) is known to increase carotenoid expression while simultaneously downregulating anthocyanin production in *Mimulus lewisii* (Phrymaceae) (Sagawa et al., 2016).

Carotenoids, specifically carotenes and xanthophylls, are the other major contributors to floral pigmentation diversity (Zhu et al., 2010). They are found in all plants and algae, as well as some bacteria and fungi (Nisar et al., 2015), and more than 1,100 compounds are classified (Yabuzaki, 2017). They are known to provide photoprotective qualities (Demmig-Adams, 1990), among other functions, and their biosynthesis is here described: In the presence of light, geranylgeranyl pyrophosphate (GGPS) is converted in multiple stages to all-trans-lycopene, a bright-red carotene. The pathway splits after the production of lycopene, and it is converted into either α-carotene or β-carotene (orange), which are further converted to lutein (yellow) or one of several xanthophylls, respectively (Tanaka et al., 2008; Zhu et al., 2010). In addition to their role in pigmentation, xanthophylls are precursors in the biosynthesis of abscisic acid (ABA), which is a key plant hormone involved in dormancy, organ size and stomatal closure (Chen et al., 2020).

Despite the thorough classification of carotenoids, there is a gap in knowledge about their upstream regulation, particularly in flowers. Current evidence is related to carotenoid accumulation in knockout experiments, e.g., knockouts of chromatin structure regulators, *HIGH PIGMENT 1* and *2* (*HP1/2*), accumulate high quantities of carotenoids. Similarly, *ELONGATED HYPOCOTYL 5* (*HY5*) and *CONSTITUTIVE PHOTOMORPHOGENIC 1* (*COP1*) result in accumulation of carotenoids in tomato (Kilambi et al., 2013). Our understanding is further obscured by post-translational regulation, e.g., there is no practical difference between pathway expression in yellow and white petals of chrysanthemums (Ohmiya et al., 2006) because of *CAROTENOID CLEAVAGE DIOXEGENASE* (*CCD*). *CCD*, which is expressed in white chrysanthemums, reduces

the effective level of carotenoid pigment, masking its transcriptional expression; the suppression of *CCD* reverted white petals to yellow (Ohmiya et al., 2006). Similar function is observed with the *Orange* (Or) gene in *Brassica oleracea* (Osorio, 2019). The benefit of post-translational cleavage, in lieu of transcriptional repression, is not currently understood, although it may be important for maintaining basal quantities required for hormone production. Nonetheless, we expect *CCD* mRNA to be present in higher concentrations in white petals.

Cleomaceae is well suited to study expression related to the ABP and CBP, because it exhibits a broad range of floral colours. Some of its 25 recognized genera are grouped by unifying floral colours, including the yellow flowers of *Arivela,* the orange/red infloresences of *Podandrogyne*, and purple petals with yellow markings commonly found in *Sieruela* (Bayat et al., 2018). In contrast, other taxa have white petals: *Tarenaya spinosa*, *Gynandropsis gynandra*, and *Polanisia dodecandra* (Bayat et al., 2018)*.* Whereas all petals of some species are the same color, other species have distinct color patterns between adaxial and abaxial petals. For example, *Cleome violacea* has adaxial petals with yellow spots. The combination of both yellows and reds in the same petals suggests that anthocyanins and carotenoids both contribute to pigmentation between and within species.

This study compares a pigmented and white flowered species of Cleomaceae to elucidate the differences between ABP and CBP contributions to pigmentation between and within species. *Cleome violacea* has maroon petals with yellow undertones and each adaxial petal has a yellow spot. In contrast, all petals of *G. gynandra* are white. The two primary goals of this chapter are: (1) examine the transcriptional changes between top and bottom petals within species, and (2) understand how carotenoid and flavonoid pathways vary between white and pigmented flowers. I hypothesize that only subtle changes in gene expression are required to modulate pigmentation between top and bottom petals of *C. violacea*, i.e., modification of pigmentation will be related to expression of upstream regulators. Further, because white petals often still express carotenoid and anthocyanin pathway genes, other factors such as *CCD* should be more abundant in *G. gynandra*.

6.2 Materials and Methods:

6.2.1 Plant growth conditions

Cleome violacea and *G. gynandra* were grown from lab seed stock (original source #813 from Hortus Botanicus and accession TOT8917 from Malawi (Sogbohossou et al., 2020), respectively). Vouchers were deposited in the Vascular Plant Herbarium, University of Alberta (ALTA). Seeds were sown in sterilized soil (SunGro sunshine mix #4) and grown in a University of Alberta, Department of Biological Sciences, growth chamber (16h day) at 24 °C for *C. violacea* and 12h day/12h night at 28 °C/18 °C for *G. gynandra.* Top and bottom petals were collected from newly opened flowers and flash frozen in liquid nitrogen prior to storage at -80 °C. For *C. violacea* three biological replicates each of adaxial and abaxial petals were collected, and for *G. gynandra,* four biological replicates of adaxial and abaxial petals were collected. For both species a biological replicate is defined as 5-10 flowers taken from a single plant. On average, more flowers were collected for *C. violacea* to collect an adequate quantity of RNA.

6.2.2 RNA extraction and transcriptomic library preparation

RNA was extracted from frozen tissue following the Qiagen RNeasy micro kit protocol (Hilden, Germany). RNA concentration was measured using a Nanodrop ND-1000 spectrophotometer (Version 3.1.2) and quality was confirmed using the Agilent 2100 bioanalyzer (version B.02.09S1720). The cDNA was prepared using the Illumina TruSeq stranded mRNA LT sample prep kit (California, U.S.) following the low sample protocol. The mRNA from 14 samples was isolated and purified using nucleomag beads (Macherey-Nagel, Düren, Germany) prior to cDNA synthesis. Unique Illumina adapters were ligated to each sample and PCR amplified prior to cDNA validation. The PCR cycle was run with the following settings: 15x 98°C for 10 seconds, 60°C for 30 seconds and 72°C for 30 seconds followed by 1x 72°C for five minutes and a final hold at 4°C. Samples were normalized, pooled, and sequenced by the center for applied genetics (TCAG) facilities of the Toronto Sick Kids hospital, Ontario, Canada.

6.2.3 De novo transcript assembly and analysis

Trimming, quality checking, assembly and differential expression analysis were performed as previously described using updated software (Table S6.1) (See Chapter 2). Briefly, raw reads

were trimmed using Trim Galore and quality checked with FastQC prior to de novo assembly with Trinity. Significant Differential expression (SDE) was measured using edgeR with a false discovery rate (FDR) corrected significance threshold of $\alpha = 0.05$. Both transcriptomes were annotated using BLASTn with default parameters against a local copy of the Araportt11 database (Cheng et al., 2017).

6.2.4 Orthologous clustering

For both species, a list of all transcripts with expression values greater than or equal to five transcripts per million (TPM) were extracted from the Trinity master fasta file using 'faSomeRecords.py' (https://github.com). A TPM value of five was chosen because it roughly equates to one transcript per cell, assuming 200,000 mRNAs per cell (Shapiro et al., 2013). The list was created using the sort function in excel. Next, the trimmed FASTA file was converted to a peptide list of longest open reading frames (ORFs) using 'TransDecoder.LongOrfs' (Table S6.1), and the included perl script, 'get_longest_ORF_per_transcript.pl'. These scripts compile a list of putative protein-coding transcripts, and the longest ORF peptide translations were used as a best estimate of coding genes for both transcriptomes. Four lists in total, one for each petal and species type, were uploaded to the Orthovenn2 webserver for one-to-one comparison between each grouping (https://orthovenn2.bioinfotoolkits.net). The Venn diagram generated shows a count of gene clusters unique and shared between top and bottom petals of *C. violacea* and *G. gynandra*.

6.2.5 KEGG analysis

The KEGG (Kyoto Encyclopedia of Genes and Genomes) analysis was completed using the KAAS (KEGG Automatic Annotation Server: https://www.genome.jp/kegg/kaas/). KEGG is a compendium of known gene pathways which are used to annotate novel genomic/transcriptomic data. The analysis is limited in that it does not incorporate expression level, i.e., any transcripts with a TPM above five are considered present regardless of magnitude. However, it does provide preliminary evidence about presence/absence of pathway expression in tissues of interest. For both species, the BLAST search algorithm was used with the bi-directional best hit assignment method. Initially, all SDE transcripts ($\alpha = 0.05$) from

adaxial and abaxial petals of both species were separately analyzed to explore which pathways were completely or partially active. Because there were relatively few SDE transcripts for either species, a second analysis was run using all transcripts with expression values above five TPM.

6.2.6 Heatmap generation

First, whole transcriptome z-score matrices were generated using the 'analyze_diff_expr.pl' script provided with Trinity, using default parameters. These matrices contain the z-score modified values of expression data for all significant transcripts across both transcriptomes. Z-score measures the relationship to the mean of all values from a sample. Z-score heatmaps were generated using ggplot2 in R (Table S6.1) To identify potential genes of interest, expression profiles from the carotenoid, anthocyanin/flavonoid pathways were generated. Representative genes were determined using the highest bit-score alignment from the Araport11 annotation. The R package ggplot was used to generate the heatmap files.

6.2.7 Transcriptome quality and completeness

Benchmarking Universal Single-Copy Orthologs (BUSCO) analysis was used to determine the overall completeness of each transcriptome (Simão et al., 2015). This tool evaluates the overall gene content of respective transcriptome/genome based on a Viridiplantae dataset. In conjunction with BUSCO, the 'contig_ExN50_statistic.pl' script provided with Trinity was used to calculate ExN50, i.e., the highest expressed transcripts that represent N% of all normalized expression data. For example, at Ex90, the given N50 would be representative of 90% of normalized expression data.

6.3 Results and Discussion

6.3.1 Statistical analysis of transcriptome indicates a high-quality assembly.

Multiple lines of evidence indicate that both datasets are of excellent quality. An average of 30.3 and 34.8 million high quality reads were assembled for the *C. violacea* and *G. gynandra* transcriptomes, respectively. Although not universal, typical sequencing depth of 30 million or greater is considered suitable for non-model de novo studies (Francis et al., 2013). For each 150 base pair read of both transcriptomes, Phred scores were no less than 32 for base pairs 1-5 and

145-150 (Data not shown). Other base pairs had scores of 38 or higher, i.e., the probability of an incorrect base call was around 1.6 in 10,000. While there is no consensus on an ideal Phred score, values greater than 20 (1 in 100 base call error) are usually the minimum cutoff (Mbandi et al., 2014). Our values indicate that any one transcript is unlikely to have more than 1 incorrectly called base pair.

Transcriptome completeness was measured using BUSCO and Ex90N50 values. BUSCO analyses showed the *C. violacea* transcriptome at 96.4% complete with an Ex90N50 value of 2143, and the *G. gynandra* transcriptome at 92.60% complete with an Ex90N50 of 2028 (Figure 6.1; Table 6.1). There is no universally accepted percentage for BUSCO, but values near to 100% are likely best. We suspect that the *G. gynandra* transcriptome is less complete because it shares the *Th-α* whole genome duplication (WGD) event with *Tarenaya hassleriana* (Mabry et al., 2020), and thus has similar synonymous substitution rates (Ks) (Figure 2 in (van den Bergh et al., 2014)). Such events may impact gene retention (Brunet et al., 2006) and so may affect completeness. Partial incompleteness may also be due to tissue type and stage of development chosen for each transcriptome (Veeckman et al., 2016), i.e., not all genes are actively expressed at all stages of development or in all organs. Generally, a higher Ex90N50 value indicates longer contig lengths and a higher quality transcriptome. In a study comparing different quality statistics, Trinity scored the fourth highest average Ex90N50 value relative to other assemblers for *Arabidopsis thaliana* (L.) (1451) (Hölzer and Marz, 2019), and the second highest BUSCO statistic in *Arabidopsis* for completeness (77.6%). Because of the relatively high Ex90N50, BUSCO, read-depth, and FastQC scores, both transcriptomes were considered suitable for downstream analyses.

6.3.2 General expression patterns within and between C. violacea and G. gynandra

Both *C. violacea* and *G. gynandra* have distinct expression patterns across abaxial versus adaxial petal pairs (Figure 6.2). However, *C. violacea* petals have more transcriptional activity than *G. gynandra* as demonstrated by a roughly nine-fold increase of SDE transcripts in the *C. violacea* dataset (317 vs 35) (Figure 6.2). This difference likely reflects that the petal pairs of *G. gynandra* are similar in both morphology and pigmentation whereas *C. violacea* petals differ in both. Both

species have more gene upregulation in adaxial petals (Figure 6.2) as compared to abaxial, which is consistent with gene expression patterns in monosymmetric flowers with adaxial and abaxial differences (Chapter 5). There are 7554 of 9820 shared orthologs between both species, and 1211 and 1055 unique to *C. violacea* and *G. gynandra*, respectively (Figure 6.3). Altogether, pigmented petals of *C. violacea* have more active transcription than white petals of *G. gynandra*, and more of those transcripts are upregulated in adaxial petals regardless of pigmentation.

6.3.3 MYB expression is consistent with pigmentation in C. violacea *and* G. gynandra

We expected to see an enrichment of R2R3-MYB family transcription factors in *C. violacea* flowers because they are linked to the regulation of the ABP in many pigmented taxonomic groupings (Sobel and Streisfeld, 2013; Hsu et al., 2015; Sagawa et al., 2016; Kramer et al., 2017; Su et al., 2017; Wang et al., 2019; Xi et al., 2019; Duncan and Rausher, 2020; Zhong et al., 2020), and because floral pigmentation is biased towards cis-regulatory control (Sobel and Streisfeld, 2013). Likewise, we expected to see fewer R2R3-MYB transcription factors in *G. gynandra* because its flowers lack pigmentation. There are a plethora of MYBs associated with numerous functions in plant development (Wu et al., 2022), but we focused on the few MYBs directly linked to flavonoid synthesis (Table 6.2). Of these, nine are expressed in *C. violacea* and one is expressed in *G. gynandra* (Table 6.2). Overall, it appears that MYB expression is not responsible for differences in adaxial/abaxial pigmentation because there is no significant differential expression between petals for any of its regulators or pathway enzymes. Interestingly, *MYB1*, the only ABP-linked MYB expressed in *G. gynandra* flowers, is known to upregulate early ABP genes in *Gerbera* (*CHS* and *FLS*) (Zhong et al., 2020) and promotes petal spot formation in *Clarkia* (Martins et al., 2017). Paradoxically, *MYB1* expression is also positively correlated with a decrease in anthocyanin accumulation (Zhong et al., 2020). The decrease in anthocyanins may be in part due to the upregulation of *FLS*, the first step in flavonol biosynthesis, which diverts anthocyanin precursor away from *DFR*. Because no ABP genes are expressed in *G. gynandra*, we believe *MYB1* may act instead as a negative regulator. This could be possible via modifications directly to *MYB1*, its CRE binding site, or post-translational modification from another *MYB*. For example, *MYBL2* and *MYB27* modify the MBW and switch

its function from activator to repressor of late-stage ABP (Dubos et al., 2008; Boase et al., 2014), although no homologues are expressed in *G. gynandra*. There are, however, many yet unlinked MYBs expressed in *G. gynandra* petals (Data not shown). Future work should focus on characterizing *GgMYB1*, its binding sequence, and associated enzymes.

In *C. violacea*, all ABP genes are expressed except for O-methyltransferase (*OMT*), which adds a methyl group to anthocyanins, inducing a color shift in some species (Du et al., 2015). Its absence may be due to a lower concentration of anthocyanins relative to anthocyanidins, which is correlated with moderate expression of *UFGT* (Figure 6.5), although HPLC analysis is required for confirmation. All other late-stage ABP genes are highly expressed (>1000 TPM), which is correlated with an accumulation of anthocyanins, and explains the maroon pigmentation on adaxial and abaxial petals. Although not significant, adaxial petals have reduced expression of all ABP genes. This reduction is likely due to their smaller size and yellow carotenoid spot. It appears that the *C. violacea* ABP is regulated by an MBW (MYB-bHLH-WD40) complex, which likely consists of *MYB90* or *WER/GL1, TT8*, and *TTG1* respectively (Hichri et al., 2011). It is possible for other genes to be a part of the MBW complex, but only the aforementioned are expressed (Figure 6.5). The MBW complex upregulates late-stage ABP genes to promote accumulation of anthocyanins (Xu et al., 2021). There is active expression of both positive and negative regulatory factors in *C. violacea* (Figure 6.5), which suggests precise regulation of pigment concentrations. For example, *CAPRICE* (*CPC*) disrupts the formation of the MBW complex by competing with *PAP1/2* (*MYB75/90*) (Zhu et al., 2009), thus reducing anthocyanin accumulation.

Contrary to this, *MYB21* downregulates the production of jasmonic acid, which in turn reduces the expression of *JAZ* proteins through a complex negative feedback loop (Reeves et al., 2012). Similar to *CPC*, *JAZ* proteins interfere with the MBW complex by binding directly with *MYBs,* and also bHLH components (Shan et al., 2009). Lastly, *MYB3* interacts with MBW factors *TT8* and *EGL3* to repress anthocyanin accumulation in specific conditions, e.g., salt stress in *Arabidopsis* (Kim et al., 2022). Altogether, MYB and ABP gene expression is correlated with other taxonomic groupings for both tested species (He et al., 2010; Hichri et al., 2011; del Valle et al., 2019;

Mekapogu et al., 2020; Zhong et al., 2020). However, *MYB10* expression is notably absent in *C. violacea*, which is correlated with anthocyanin accumulation in *G. hybrida* (Laitinen et al., 2008). This absence may be due to phylogenetic distance and the pleiotropic nature of *MYB* transcription factors, e.g., *Arabidopsis MYB10* is typically expressed in roots under iron-limited conditions (Palmer et al., 2013) and not known to be expressed in flowers. R2R3-MYBs don't need a high degree of sequence similarity to function transgenically (Ramsay and Glover, 2005), thus it can be inferred that another MYB serves a similar function as *MYB10* in *C. violacea*. It is likely that other components of MYB-bound protein complexes (e.g., WD40) also contribute to functionality, i.e., MYB expression is only one part of the story.

6.3.4 Carotenoid synthesis

Unlike the ABP, the CBP is expressed in both pigmented and white flowers (Figure 6.6). Generally, patterns between adaxial and abaxial petals are similar between both species (Figure 6.5), which implies that carotenoid biosynthesis is roughly equal regardless of petal type. This result is surprising considering the yellow petal spot on *C. violacea* adaxial petals, although it is consistent with photoprotective qualities of carotenoids and/or synthesis of ABA (Nisar et al., 2015). That is, a baseline concentration of carotenoid pigment can be expected in all floral cells in order to protect DNA, and ABA biosynthesis may occur equally in abaxial and adaxial petals. Both *C. violacea* and *G. gynandra* petals express transcripts related to carotenoid biosynthesis including the synthesis of lycopene and astaxanthins, e.g., zeasxanthin, biolaxanthin and xanthoxin. However, it appears that transcripts related to lutein synthesis are not expressed during later petal development in *G. gynandra* (Figure 6.6). Both lycopene-epsilon cyclase 1 & 2 are below five TPM in *G. gynandra* petals, which suggests that no lutein is present (in addition to white phenotype). CBP heatmap expression data is consistent within and between petals of both species, with a few exceptions (Figure 6.5). Generally, carotenoid related genes in *G. gynandra* are weakly expressed relative to *C. violacea* (Figure 6.5), i.e., *G. gynandra* transcripts reach a maximum of ~64 TPM while *C. violacea* reach a maximum of ~1024 TPM, a 16-fold difference in expression. While some differences between transcriptomes are to be expected, it is likely that some of the variation is biologically relevant, as both transcriptomes were produced using identical procedures (therefore limiting technical variance). It seems likely that

carotenoid production in *G. gynandra* may be strictly for the synthesis of ABA (Finkelstein, 2013). It is unlikely that carotenoids in *G. gynandra* petals offer any photoprotection because they are not present in high enough concentrations to absorb any wavelength of light (i.e., white petals). It is also plausible that CBP expression is leaky because complete silencing would result in deleterious phenotypes in other parts of the plant (i.e., absence of ABA).

It was hypothesized that *CCD1* would be upregulated in *G. gynandra* relative to *C. violacea* because *CCD1* protein degrades carotenoids, and its disruption converted yellow *Brassica* (Zhang et al., 2015) and Chrysanthemum petals (Ohmiya et al., 2006) to white. However, it is less expressed by more than a factor of six (~60TPM vs ~400 TPM) in *G. gynandra* than *C. violacea*, respectively. This is consistent with species of *Ipomoea*, where *CCD1* expression is not correlated to carotenoid concentrations (Yamamizo et al., 2010). What may be happening is that *CCD1* fine-tunes pigmentation. This fine tuning is consistent with the similar magnitude of *CCD1* expression in both species, i.e., *CCD1* TPM is generally the same order of magnitude of TPM as other carotenoid genes. Contrastingly, *aldehyde oxidase 3* (*AAO3*) and *9-cis-epoxycarotenoid dioxygenase 3* (*NCED3*) are highly expressed in *G. gynandra* and only minimally expressed in *C. violacea* (Figure 6.5). High expression of end-stage pathway genes may help drive the pathway forward to ABA synthesis and avoid accumulation of precursors.

6.3.5 Conclusion

In this preliminary study of gene regulation of pigmentation, we demonstrated that the ABP is only expressed in pigmented *C. violacea* flowers, and the CBP is expressed in both focal species, albeit at lower TPM in white-petaled *G. gynandra*. Our hypothesis that *CCD1* would be upregulated in white petals was unsupported because it is more highly expressed in *C. violacea*. However, *CCD1* is expressed at roughly the same magnitude relative to other CBP genes in both species. We suggest that it is used in the Cleomaceae to tightly regulate levels of carotenoids in pigmented and white flowers, although studies in other species are required.

There is a high level of conservation seen in regulatory *MYBs* controlling the ABP pathway in *C. violacea* relative to other angiosperms. As expected, the ABP is not active in *G. gynandra* and anthocyanins appear to serve no secondary function in Cleomaceae flowers. The absence of

MYB10 in *C. violacea* is interesting but consistent with the highly pleiotropic nature of *R2R3-MYBs.* It is likely that *MYB10* function is replaced by another *MYB*, we suggest MYB111 performs this function in *C. violacea* because it is known to regulate early ABP genes, although functional studies are required (Medina-Puche et al., 2014). Future studies should focus on targeted silencing of highly expressed *MYBs* outlined in this study, and/or overexpression lines in white flowered species.

Many unanswered questions about floral pigmentation remain, such as what is controls spot formation in adaxial *C. violacea* petals? This question could be addressed with dissection and sequencing of yellow and maroon sections of adaxial petals. One limitation in this study is that we cannot determine the exact types and concentrations of anthocyanins and carotenoids present in either species using RNA-seq data alone. Next steps should include a complete chemical analysis using HPLC, which should confirm that anthocyanins are completely absent in *G. gynandra* (Figure 6.4)*,* and that lutein is only present in *C. violacea* because of relevant pathway expression (Figure 6.5 & 6.6). Following relevant gene expression, carotenoid precursors in *G. gynandra* should be only those that are directly linked to ABA synthesis, e.g., β-carotene, zeaxanthin and violaxanthin (Figure 6.5 & 6.6). This chapter is the first of its kind in Cleomaceae and has laid the foundation for future pigmentation research in the family.

6.4 Figures
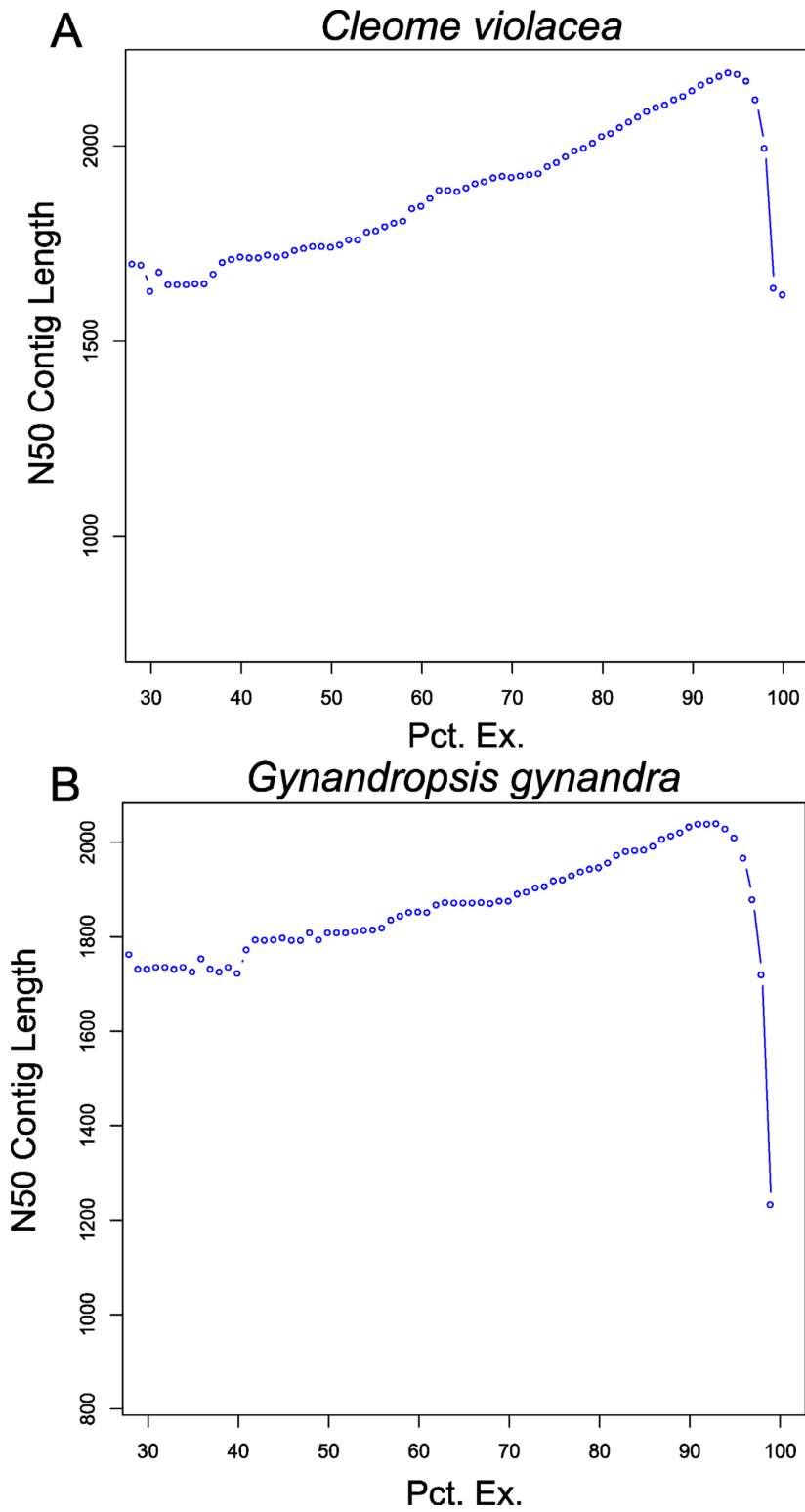
Figure 6.1 ExN50 data-plot for the *Cleome violacea* (A) and *Gynandropsis gynandra* (B) petal transcriptomes generated by the 'TrinityStats.pl' script and 'plot_ExN50_statistic.Rscript'.
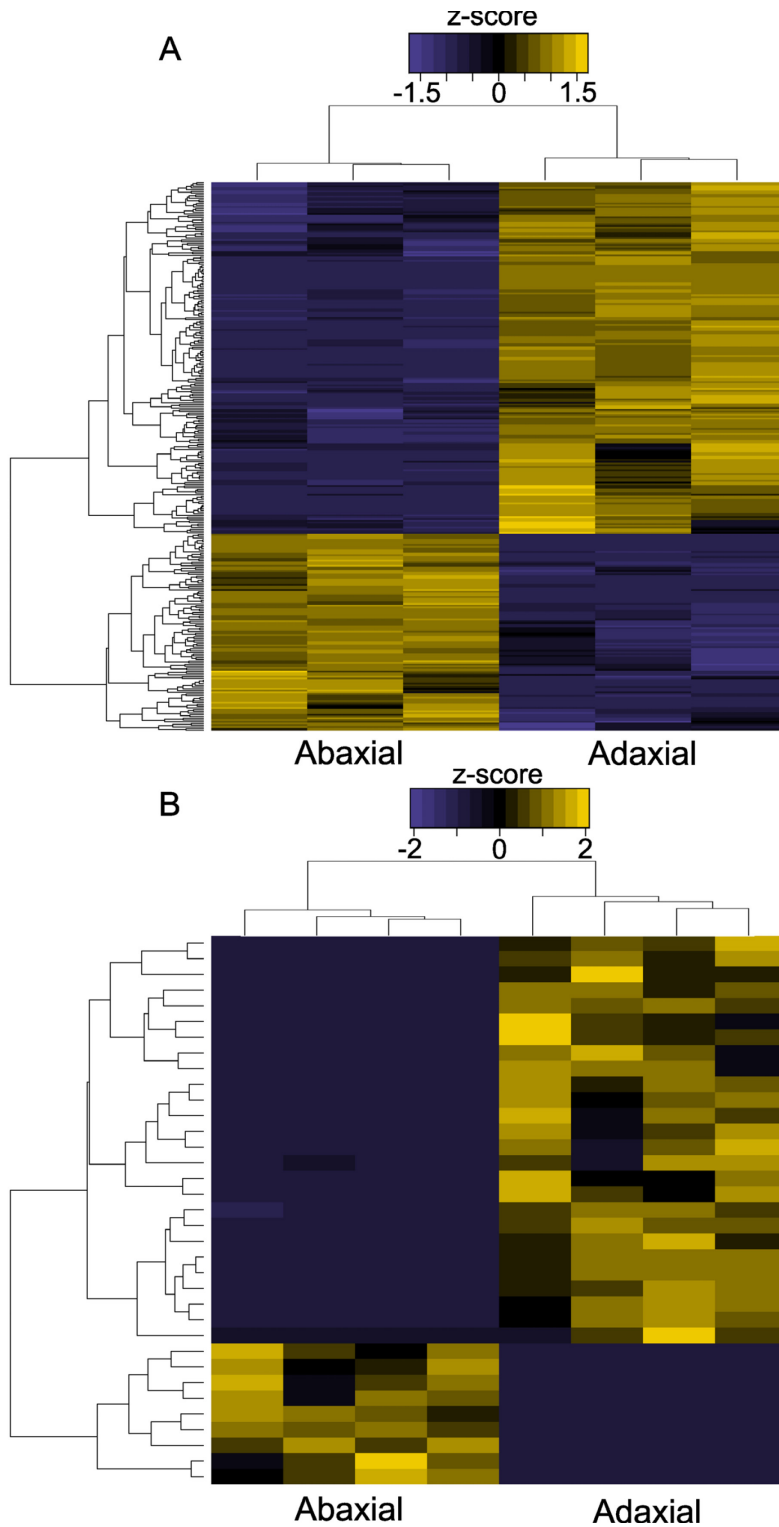
Figure 6.2 Heatmap of all significant edgeR contig clustered transcripts in the *Cleome violacea* (n = 317) and *Gynandropsis gynandra* (n = 35) transcriptomes, expressed as z-scores (FDR-corrected α = 0.01).
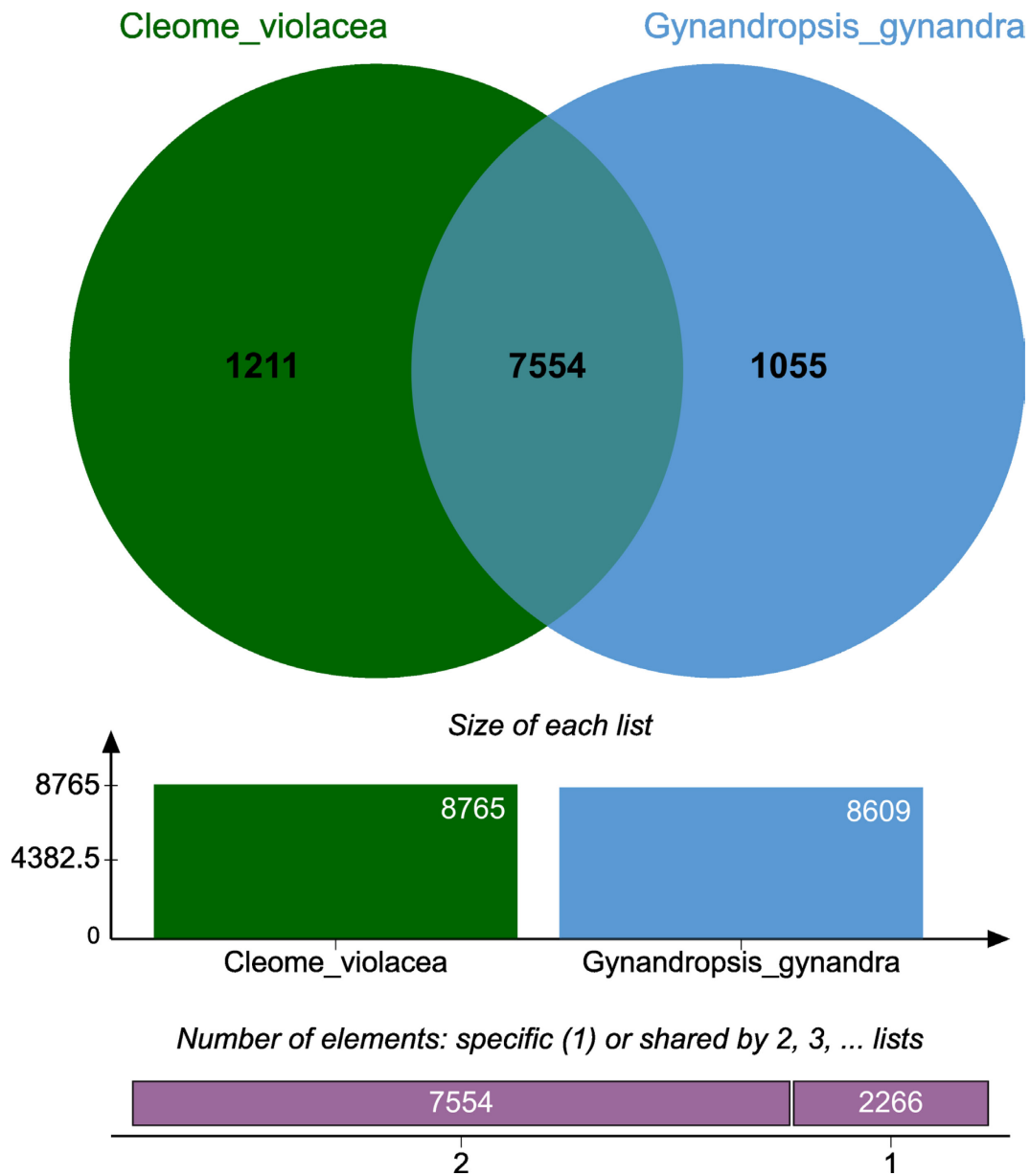
Figure 6.3 Graph of Gene Ontology (GO) terms for *Cleome violacea* (green) and *Gynandropsis gynandra* (blue). GO counts based on transcript list with replicate average greater than or equal to five TPM.
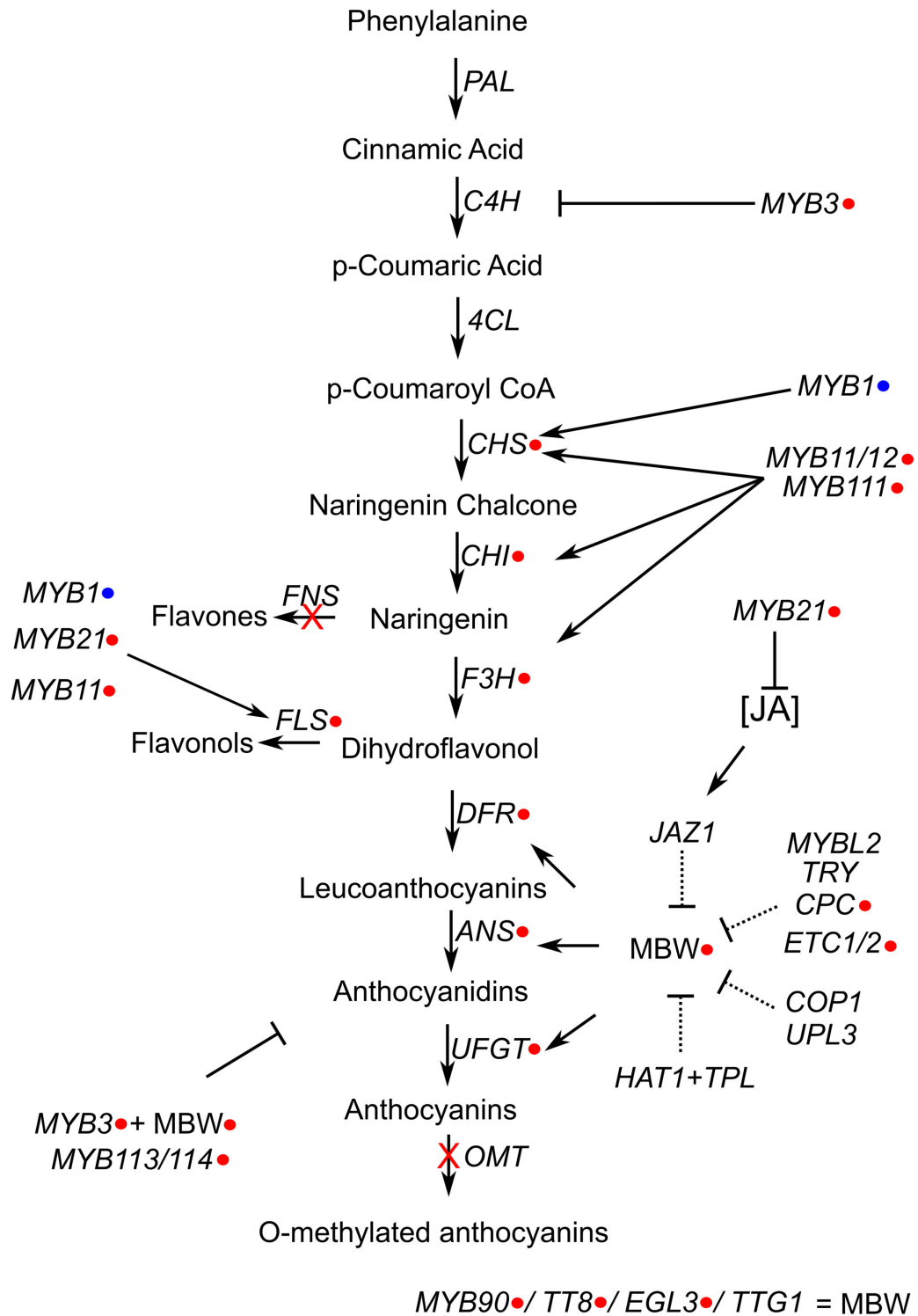
Figure 6.4. Simplified Anthocyanin biosynthetic pathway. Red and blue markers indicate genes that are uniquely expressed in *C. violacea* and *G. gynandra* only, respectively. Crossed out arrows indicate that no gene expression is observed for either species.
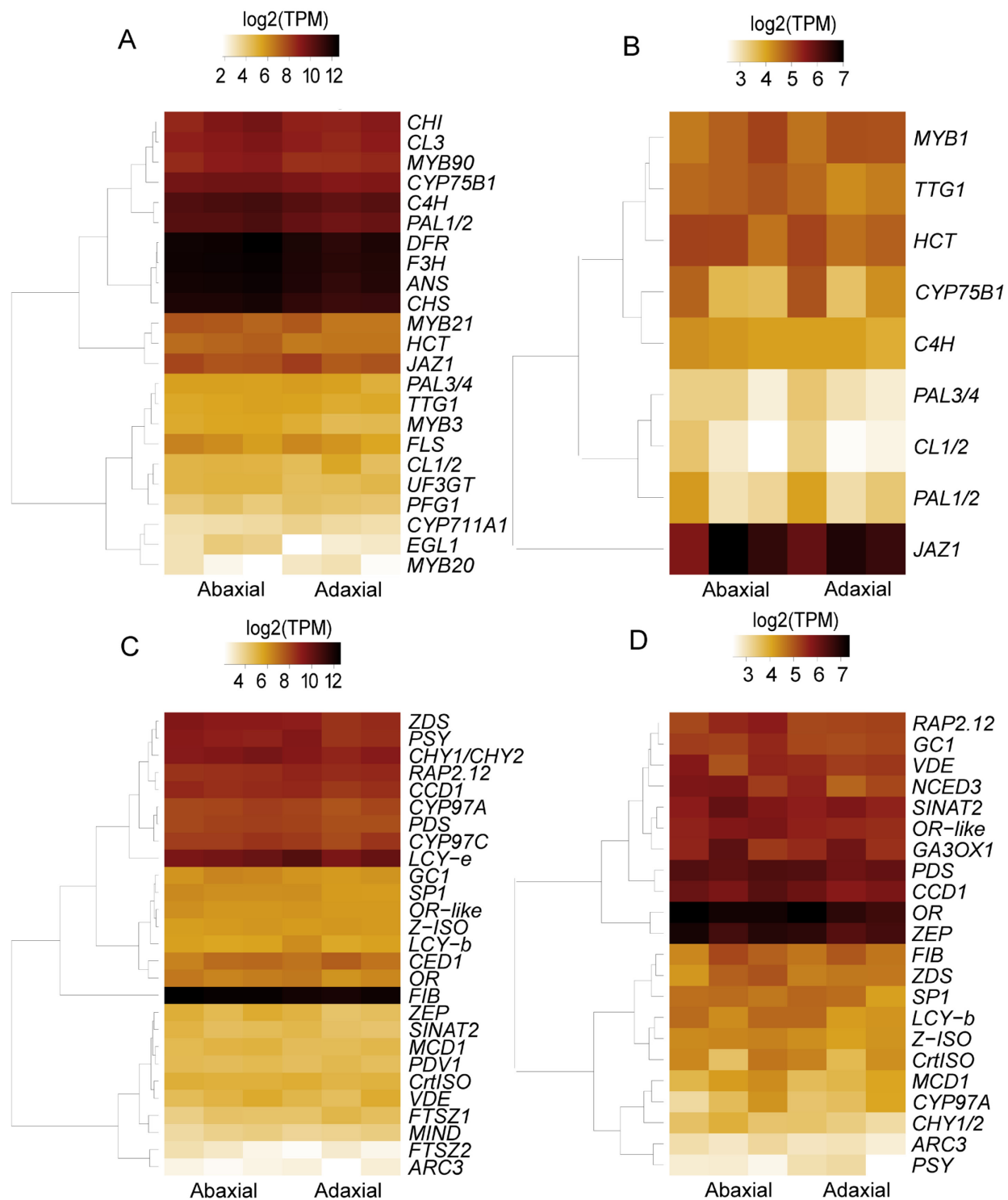
Figure 6.5. Expression heatmap in log2(TPM) of anthocyanin (A-B) and carotenoid (C-D) biosynthesis pathways for *Cleome violacea* (A-C) and *Gynandropsis gynandra* (B-D).
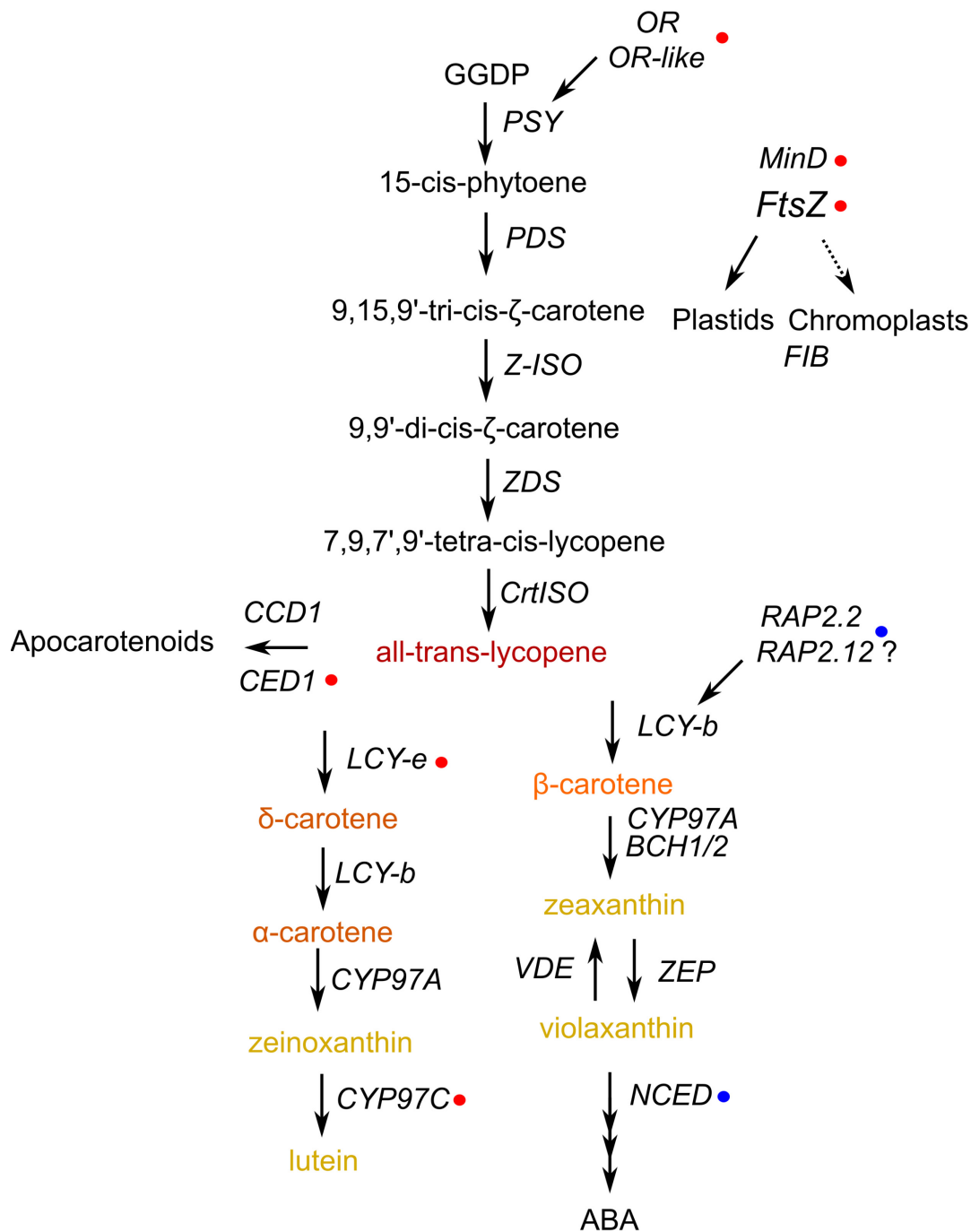
Figure 6.6. Simplified carotenoid biosynthetic pathway. Red and pink markers indicate genes that are uniquely expressed in *C. violacea* and *G. gynandra* only, respectively. Crossed out arrows indicate that no gene expression is observed for either species.
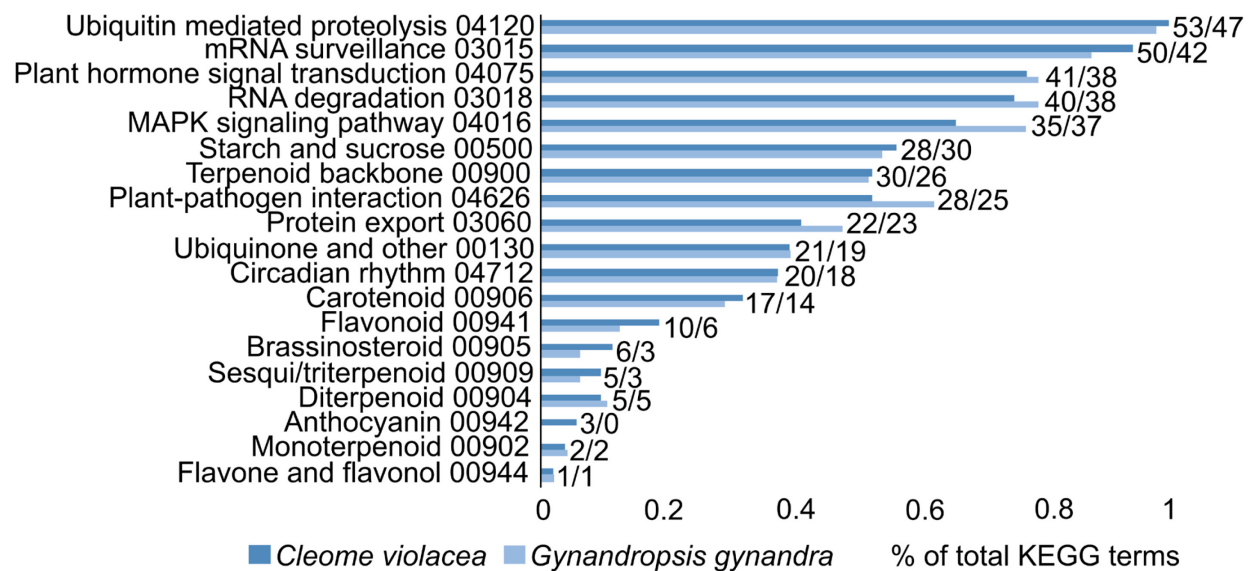
Figure S6.1. Top 19 KEGG counts for all transcripts greater than or equal to 5 TPM.

6.5 Tables

Table 6.1. Benchmarking Universal Single-Copy Orthologs (BUSCO) (A) and contig analysis (B) of *Cleome violacea* and *Gynandropsis gynandra* transcriptomes. Longest isoform values in brackets.

BUSCO	*Cleome violacea*	*Gynandropsis gynandra*
Complete BUSCOs (C)	2046	1963
Complete and single-copy BUSCOs (S)	862	697
Complete and duplicated BUSCOs (D)	1184	1266
Fragmented BUSCOs (F)	42	90
Missing BUSCOs (M)	33	68
Total BUSCO groups searched	2121	2121
Completeness	96.40%	92.60%
Assembly Statistics	*Cleome violacea*	*Gynandropsis gynandra*
Contig N10	5035 (4459)	4792 (4126)
Contig N20	3917 (3443)	3736 (3145)
Contig N30	3248 (2836)	3133 (2512)
Contig N40	2766 (2408)	2657 (2052)
Contig N50	2398 (2029)	2272 (1587)
Median contig length	1095 (404)	835 (382)
Average contig length	1472.38 (945.49)	1323.25 (788.85)
Total assembled bases	162435466 (52655541)	233538229 (69116152)
Total trinity genes	55691	87616
Total trinity transcripts	110322	176488
Percent GC:	44.18	41.75

Table 6.2. Functional and regulatory genes related to the anthocyanin biosynthesis pathway (ABP) for *Cleome violacea* (Cv) and *Gynandropsis gynandra* (Gg). - = expression below 5 TPM, + = expression <100 TPM and ++ = Expression > 100 TPM.

Gene	Putative Role	Citation	*Cv*	Gg
PAL1/2	Phenylpropanoid Biosynthesis	Vogt 2010	+++	+
C4H			+++	+
4CL3			++	+
CHS	Early Flavonoid Pathway Genes	Yonekura-Sakakibara et al., 2019	+++	-
CHI			++	-
F3H			+++	-
FLS			+	-
FNS			-	-
DFR	Late Flavonoid Pathway Genes		+++	-
ANS			+++	-
UF3GT			+	-
MYB3	Represses *C4H* with *LNK1/2* co-repressors	Zhou et al., 2017	+	-
MYB113/114	Upregulates *F3H*, *DFR*,*ANS*,*UFGT*	Gonzalaz et al., 2008	++	-
MYB1	Upregulates *CHS* and *FLS*	Zhong et al., 2020	-	+
MYB11/12/111	Upregulates *CHS*, *CHI*, *F3H* and *FLS*	Stracke et al., 2007	+	-
MYB21	Promotes *FLS1* and decreases JA levels	Zhang et al., 2021; Reeves et al., 2012	++	-
MYB90	Potential constituent of MBW Complex; promotes Late Flavanoid structural genes	Xu et al., 2015	++	-
TT8			++	-
EGL3			+	-
TTG1			+	+
TT2			-	-
GL1			++	-
WER			++	-
JAZ1	Destabilizes MBW complex	LaFountain & Yuan 2021	++	+
TRY	Inhibits formation of function MBW complexes	Ishida et al., 2007	-	-
CPC			+	-
ETC1/2	Enhancers of *TRY* and *CPC* in trichoem and root hair formation	Kirik et al., 2004	+	-
COP1	Degrades components of MBW complex	Maier & Hoecker 2005	+	+
UPL3	Degrades *EGL3/GL1*	Patra et al., 2013	+	+
HAT1	Inhbits MBW and recruits *TPL*	Zheng et al., 2019	+	+
TPL	Recruited by *HAT1* as co-repressor.		+	+

Table 6.3 Functional and regulatory genes related to the carotenoid biosynthesis pathway (CBP) for *Cleome violacea* (Cv) and *Gynandropsis gynandra* (Gg). - = expression below 5 TPM, + = expression <100 TPM and ++ = Expression > 100 TPM.

Gene	Putative Role	Citation	*Cv*	Gg
PSY	Lycopene synthesis	zhu et al., 2010	++	+
PDS			++	++
Z-ISO			+	+
ZDS			++	+
CrtISO			+	+
LCY-e	Lutein synthesis		++	-
LCY-b	Zeanxthin/violaxanthin synthesis		+	+
BCH1/2			++	+
CYP97C			++	-
CYP97A			++	+
VDE			+	+
ZEP			+	++
NCED3	ABA synthesis; first step from violaxanthin		-	+
OR	Regulation of *PSY*	Osorio 2019	++	++
OR-like			+	+
RAP2.12	Regulation of *LCY-b*	Dalal et al., 2010	++	+
RAP2			-	+
CED1	Apocarotenoid synthesis	Simkin et al., 2004	++	-
FTSZ1	Plastid/Chromoplast division; correlated with pigmented flowers	Moehs et al., 2001	+	-
FTSZ2			+	-
CCD1	Apocarotenoid synthesis;	Simkin et al., 2004	++	++
MIND	Plastid/Chromoplast division; correlated with pigmented flowers	Moehs et al., 2001	+	-
CHRB/FIB	Chromoplast specific protein; Storage in certain chromoplast-types is achieved by carotenoid esterification which allows their association with specialized proteins known as fibrillins or plastid lipid associated proteins.	Smirra et al., 1993; Vishnevetsky et al., 1999	+++	+
GA3OX1	ABA synthesis	Finklestein 2002	-	+
MCD1	Associated with MIND	Nakanishi 2009	+	+
ARC3	Chimera of *FtsZ*; Involved with chromoplasts?	Sun et al., 2020	+	+
GC1	Involved with chloroplast division; Involved with chromoplasts?		+	++
SINAT2	Interacts with *RAP2*	Welsch et al., 2007	+	+
PDV1	Overexpression increases carotenoids and chromoplast size	Sun et al., 2020	+	-
SP1	Promotes chloroplast to chromoplast transition in Tomato	Ling et al., 2021	+	+

Table S6.1. Software versions used in this study.

Software	Version	Citation
Trim Galore	0.6.2	(Krueger, 2012)
FastQC	0.11.5	(Andrews, 2010)
Trinity	2.8.5	(Grabherr et al., 2011)
Corset	1.09	(Davidson and Oshlack, 2014)
edgeR	3.10	(Robinson et al., 2009)
R	3.6.3	(R Core Team, 2013)
Transdecoder	5.5.0	(Haas et al., 2013)
Orthofinder	2.5.2	(Emms and Kelly, 2015)
BLAST+	2.11.0	(Altschul et al., 1990)

Chapter 7: Conclusions and future of research in Cleomaceae

How morphological variation has arisen is a fundamental question in evolutionary biology that requires an interdisciplinary approach to address. Although all taxonomic groupings are morphologically intriguing, Brassicales (specifically sister families Brassicaceae and Cleomaceae) are well suited to address fundamental questions about the genetic basis of morphological diversity. Brassicales has tremendous diversity of fruits and flowers, and is home to the premier model species, *Arabidopsis thaliana*. This diversity ties in well with the overarching goal of my thesis, which is to examine the evolutionary and developmental basis of fruit and floral features in Brassicaceae and Cleomaceae, respectively. Central questions include: (1) how are genetic programs modified to produce novel fruit structures in Brassicaceae? (2) Are the same or different genetic pathways recruited in the evolution of floral features in Cleomaceae as compared to other species? (3) What is the genetic basis of multiple traits important for pollinator attractions and rewards in Cleomaceae?

RNA-seq, combined with functional genetics, are useful in answering my central questions and exploring the underlying genetics of fruit and floral phenotypes. In Chapter 2, we examine fruit patterning genes between species and within fruits of segmented members of Brassicaceae. There is a great deal of conservation in gene expression patterns of *Cakile* and *Erucaria* relative to *Arabidopsis*, but also deviations from the valve margin pathway that are consistent only with segmented fruit morphology, e.g., indehiscence in distal segments of both species and a novel abscission zone in *Cakile*. We conclude that there is a co-option of valve margin pathway genes to achieve these unique morphologies. This study shows how modifications of existing pathways correlate with the incredible homoplasy/lability in fruit features of the family. Patterns established in this chapter have parallels to research in *Zahora* (Brassicaceae), Rubiaceae (Gentianales), and *Cucumis melo* (Cucurbitaceae) (Koch and Lemmel, 2019; Salazar-Duque et al., 2021; Du et al., 2022).

Chapter 3 establishes *Cleome violacea* as a viable model for functional studies by optimizing a VIGS protocol for two genes expressed in early and late floral development (*PDS* and *FUL*). This chapter is fundamental to subsequent chapters on nectary development (Chapter 4) and

monosymmetry (Chapter 5) because it directly test the roles of genes of interest to their phenotype. In these chapters, I document additional instances where the same genetic pathways are recruited in separate origins of floral traits in *Cleome violacea*. As with many other monosymmetric taxa, *CvTCP1* is responsible for floral monosymmetry in *C. violacea* (Chapter 5). Similarly, *CRC* and its regulators *SHP/AG* are necessary for nectary formation in Cleomaceae (Chapter 4), and other core eudicots (Luo et al., 1996; Almeida et al., 1997).

While adding additional case-studies is essential for documenting the extent of parallelisms, the nuances uncovered between species bring additional light to these processes. For example, *STY*, which is essential for nectary spur initiation in *Aquilegia* (Min et al., 2019) is significantly differentially expressed in *C. violacea* petals. Expression of both *CRC* and *STY* regulators hints at a possible interaction and/or a quasi-subfunctionalization of roles. Further, I establish that there is likely only one copy of *TCP1* in *C. violacea* and that downstream genes *RAD/DIV* have unexpected expression patterns. Thus, *C. violacea* unlike most other monosymmetric taxa, except for *I. amara* (Busch and Zachgo, 2007; Busch et al., 2014), has a single copy of *TCP1* contributing to floral monosymmetry. Functional data suggest a link between the TCP gene family and nectary initiation, i.e., when *TCP1* is downregulated nectaries occasionally fail to form. These results reveal an exciting example of the interplay between different developmental systems. In every investigated aspect of floral biology there were parallels and variations in regulation relative to other taxa, and even between floral organs. In summary, underlying genetic pathways often share similarities with distant taxa and many genes are pleiotropic. Undoubtedly there is more to be uncovered, and this research is only a beginning for floral development in the Cleomaceae. In the following paragraphs I provide an overview of potential next steps for expanding knowledge in the family.

7.1 In-silico studies in Cleomaceae

De novo transcriptome assembly is an invaluable tool for studying gene expression, especially in non-model taxa with limited genetic data, e.g., *Lilium ledebourii* (Sheikh-assadi et al., 2022) and *Lantana camara* (Shah et al., 2020), and was no exception for data generation in this thesis.

What then is the next step for *C. violacea,* and how can useful data be generated to aid in research on other non-model Cleomaceae?

A high-quality genome of *C. violacea* would be valuable for genetic comparisons within the family. Currently, *Arabidopsis* is the nearest taxon with such a genome, but there are potential issues when making comparisons. *Arabidopsis* flowers are bisymmetric, unpigmented and without conspicuous nectaries. Thus, gene retention for important flowering genes may be different than in Cleomaceae. In Cleomaceae, there are currently three species with published genomes, *C. violacea, T. hassleriana,* and *G. gynandra*. However, all versions are drafts and incompletely assembled (i.e., they contain hundreds of scaffolds). Draft genomes offer only limited information about synteny, which informs on functional relationships between syntenic genes that can be evolutionarily advantageous or share regulatory mechanisms, e.g., circadian clock genes (Lai et al., 2020). Additionally, a complete picture of cis regulatory modules (CRMs) may be missing; CRMs are non-random clusters of cis regulatory elements (CREs) that range in distance from promoters from Kb to Mb (Long and Miano, 2007). A high-quality genome provides an exhaustive list of CRMs upstream and downstream of all gene initiation sites, and thus provides the opportunity to make predictions about whole pathway expression when paired with RNA-seq data.

In chapters 4 and 5, we outlined in detail the downstream effects that regulatory genes had on nectaries and monosymmetry, respectively, although it remains to be determined what the downstream effector genes are for *CRC* and *TCP1*. There is, occasionally, an evolutionary advantage for related genes to share a greater degree of synteny, e.g., in mammalian hox genes (Lee et al., 2006), and in Rhizobiales (Guerrero et al., 2005). Thus, genes directly downstream of either *CRC* or *TCP1* are more likely to be functionally related. Moreover, closely related taxa tend to have greater synteny, e.g., Poaceae (Moore et al., 1995), so a complete genome in *C. violacea* would be informative for other Cleomaceae. Lastly, we could definitively confirm that there is a single functional copy of *TCP1* in the genome. This would be equally illuminating for all other regulatory genes investigated in this thesis, e.g., *CRC* and/or *MYB1*.

7.2 Functional studies in Cleomaceae

VIGS has been widely used for functionally testing genes in a multitude of species and is a well-established methodology (Becker and Lange, 2010) optimal for taxa without genomic information, and/or no established gene knockout protocol, as is the case for many Cleomaceae. Despite the availability of permanent gene editing techniques such as CRISPR/Cas9, VIGS still has a role in modern evo-devo. Unlike CRISPR/Cas9, VIGS works in shorter timespans (single generation), and is not restricted by embryo-lethal mutations, although phenotypes are impermanent. Silencing is effective even when the gene construct is from a different taxon (Hosseini Tafreshi et al., 2012), e.g., *Arabidopsis* constructs, which are available to order, can be used as a first pass for Cleomaceae species of interest. VIGS is also useful for preliminary experiments prior to floral dip and CRISPR/Cas9 because all systems use Agrobacterium as a transformation vector, i.e., the infiltration protocol is similar in all three methodologies (Zhang et al., 2006; Wang et al., 2009; Hooghvorst et al., 2019). CRISPR/Cas9 would be a plausible next step for functional studies in *C. violacea* because a protocol could be easily adapted from chapter 3.

Chapters 4 and 5 of this thesis combine RNA-seq and VIGS to investigate genes of interest and then functionally test those genes, but it would be more informative to reverse the order of investigation, i.e., sequence silenced tissue to determine downstream genetic effects. In theory, silenced tissue could be collected for sequencing with RNA-seq, which would allow genetic observation of downstream genes, e.g., those downstream from *CvCRC* and *CvTCP1*. To accomplish this, Agrobacterium and TRV2 contaminates would need to be filtered in silico prior to assembly, but such filtering is already possible (Hart et al., 2020); otherwise, this method would be no different than sequencing untreated plants. This methodology has been proposed for lines of foxtail millet (Tang, 2014), but information is limited. It would also be interesting to incorporate gblocks (custom DNA fragments up to 3000bp in length) into VIGS construct design. This way, multiple restriction sites and/or multiple genes of interest could be incorporated into a single DNA fragment, eliminating many practical difficulties of construct design. Additionally, TRV2 inserts can be anywhere from ~200bp to ~1300bp (Becker and Lange, 2010), so up to six gene fragments could be included in the same vector. It is unknown whether the molecular

machinery responsible for RNAi would be a limiting factor, but this would increase the potential of VIGS.

7.3 Future research

Many questions remain about downstream genes in nectary development and monosymmetry in *C. violacea*, and it is unclear exactly what effects *MYB* transcription factors have on carotenoid and anthocyanin pigmentation. For example, we now know that *CvTCP1* governs floral symmetry in much the same way as other core eudicots, but we do not yet understand the role of *STY*, or the interactions *DICH, RAD*, and *DIV* have in Cleomaceae, if any. *CvDIV* was expressed at a higher TPM in adaxial petals, which may in this case have a function opposite than in *Antirrhinum.* Further, there is no known homologue in *C. violacea* for *DICH*, and *CvRAD* is only expressed in one abaxial replicate. These results suggest that downstream of *CvTCP1* the pathway functions differently than distantly related taxa. Silencing of *CvDIV* would be the most informative next step; if *CvDIV* shares responsibility for adaxialization of petals with *CvTCP1* then its absence should result in a novel petal morphology that is similar to pTRV2-*CvTCP1* treated plants. Combinations of these genes should be silenced in *C. violacea* to further flesh out differences and interactions. Another key step would be to constitutively express *CvTCP1*; I predict that it would have two effects, 1) adaxialization of abaxial petals and 2) ectopic nectary formation around the adaxialized petals. This would expand *CvTCP1*'s role in floral symmetry, confirm if it is upstream of *CvCRC*, and provide evidence that the nectary is petal-associated in *C. violacea.*

It is imperative that the next step for *C. violacea* as a model system be the establishment of a protocol for stable transformation. Because *C. violacea* is already amenable to VIGS (Chapter 3-5), elements of this new protocol could be taken from already established VIGS methodology. I propose that *CvCRC* be the first target for stable transformation because 1) it is an easily detectable phenotype, 2) it does not appear to affect gynoecium development in pTRV2-*CvCRC* treated plants (i.e., *crc* genotypes can be propagated)*,* and 3) it will facilitate transcriptomic experiments that test the expression of genes downstream of *CRC*. Once a protocol has been established the next step will be to target upstream regulators that are purported to have a

broad effect on floral phenotypes (e.g., *TCP1*). Downstream genes (e.g., *SWEET9*) are better suited for VIGS because they have limited functions and overall, the process is less laborious.

Nectary initiation in *C. violacea* is highly paralleled with other core eudicots, but we do not yet know what factors influence nectary size in *C. violacea* as it expresses no homologues to *BEN* and *ROB* (which affect nectary size in *Petunia* (Morel et al., 2018)). I hypothesize that size may be influenced by *CRC* mRNA concentration. VIGS treatment of *CvCRC* occasionally resulted in a reduced, but not absent nectary. This result suggests potential involvement of mRNA dosage, i.e., If *CvCRC* were only responsible for nectary initiation then we would expect a binary result and not partial formation. Dosage of *CRC* mRNA could be tested using recombinant *CvCRC* DNA with an attached heat-shock promoter that is sensitive to external temperatures; if mRNA concentration were a factor, changes in temperature would affect nectar size. Alternatively, ethylene may influence *C. violacea* nectary size, based on the expression of ethylene-related genes at all stages of nectary development. *AG, SHP,* and *CRC* all play roles in carpel development, so it is plausible that ethylene, which has a role in fruit ripening, could have been co-opted for nectary development. Evidence to support these hypotheses is limited, but nonetheless intriguing.

In Chapter 6, I presented a solid foundation of gene expression in the ABP and CBP in *C. violacea* and *G. gynandra*. Unlike symmetry and nectary development, we have a solid understanding of pathway genes but a limited understanding of upstream regulators specific to Cleomaceae. Explicitly, there are many MYB transcription factors linked to the ABP and CBP, but it is unclear the effects each has during development. The next step in understanding the genetic regulators of pigmentation would be to functionally test multiple *MYBs*, *CCD1*, and *OR* using VIGS. Phenotypes would be simple to score, and HPLC could be used to measure the effects on pigment concentration post silencing. Further to this, we do not yet understand the causal factor for the yellow spot on *C. violacea* adaxial petals. It is likely carotenoid pigment but there is no evidence of what genes regulate the threshold between yellow carotenoid and maroon anthocyanin. To answer this question, a transcriptome of top and bottom sections of adaxial petals should be sequenced in order to determine expressional differences.

Ultimately, this thesis has greatly expanded the collective knowledge on the genetic underpinnings of heteroarthrocarpic fruits in the Brassicaceae and nectary development, monosymmetry, and pigmentation in the Cleomaceae. Additionally, my thesis has generated many more questions about the incompletely characterized pathways underlying those processes (an unavoidable consequence of any scientific inquiry). I hope that others will endeavor to build on this research, and completely delineate all aspects of floral development in an order I have come to care so much about.

References

Almeida, J., Rocheta, M., and Galego, L. (1997). Genetic control of flower shape in *Antirrhinum majus*. *Development* 124, 1387–1392.

Alonso-Cantabrana, H., Ripoll, J. J., Ochando, I., Vera, A., Ferrándiz, C., and Martínez-Laborda, A. (2007). Common regulatory networks in leaf and fruit patterning revealed by mutations in the *Arabidopsis* ASYMMETRIC LEAVES1 gene. *Development* 134, 2663–2671.

Altschul, S., Gish, W., Miller, W., Myers, E., and Lipman, D. (1990). Basic local alignment search tool. *J. Mol. Biol* 215, 403–410.

Alvarez, J., and Smyth, D. R. (1998). Genetic pathways controlling carpel development in *Arabidopsis* thaliana. *J Plant Res* 111, 295–298.

Alvarez, J., and Smyth, D. R. (1999). CRABS CLAW and SPATULA, two *Arabidopsis* genes that control carpel development in parallel with AGAMOUS. *Development* 126, 2377–2386.

Alvarez, J., & Smyth, D. R. (2002). CRABS CLAW and SPATULA genes regulate growth and pattern formation during gynoecium development in *Arabidopsis thaliana*. *International journal of plant sciences*, *163*, 17-41.

Álvarez-Pérez, S., Herrera, C. M., and de Vega, C. (2012). Zooming-in on floral nectar: A first exploration of nectar-associated bacteria in wild plant communities. *FEMS Microbiol Ecol* 80, 591–602.

Andrews, S. (2010). FastQC: a quality control tool for high throughput sequence data. Available at: http://www.bioinformatics.babraham.ac.uk/projects/fastqc.

Appel, O., and Al-Shehbaz, I. A. (2003). Cruciferae. In K.kubitzki and C.Bayer[eds.],The families and genera of vascular plants. in *Springer-Verlag, Berlin, Germany*, 75–174.

Armbruster, W. S., Shi, X. Q., and Huang, S. Q. (2014). Do specialized flowers promote reproductive isolation? Realized pollination accuracy of three sympatric Pedicularis species. *Ann Bot* 113, 331–340.

Avino, M., Kramer, E. M., Donohue, K., Hammel, A. J., & Hall, J. C. (2012). Understanding the basis of a novel fruit type in Brassicaceae: conservation and deviation in expression patterns of six genes. *EvoDevo 3*, 1-12.

Balanzà, V., Martínez-Fernández, I., and Ferrándiz, C. (2014). Sequential action of FRUITFULL as a modulator of the activity of the floral regulators SVP and SOC1. *J Exp Bot* 65, 1193-1203.

Balanzà, V., Martínez-Fernández, I., & Ferrándiz, C. (2014). Sequential action of FRUITFULL as a modulator of the activity of the floral regulators SVP and SOC1. *Journal of experimental botany*, *65*, 1193-1203

Ballester, P., & Ferrándiz, C. (2017). Shattering fruits: variations on a dehiscent theme. *Current Opinion in Plant Biology 35*, 68-75.

Barker, M. S., Vogel, H., & Schranz, M. E. (2009). Paleopolyploidy in the Brassicales: analyses of the Cleome transcriptome elucidate the history of genome duplications in *Arabidopsis* and other Brassicales. *Genome biology and evolution 1*, 391-399.

Barrett, R. L., Roalson, E. H., Ottewell, K., Byrne, M., Govindwar, S. P., Yadav, S. R., et al. (2017). Resolving Generic Boundaries in Indian-Australasian Cleomaceae: Circumscription of Areocleome, Arivela, and Corynandra as Distinct Genera. *Syst Bot 42*, 694–708.

Baulcombe, D. C. (1999). Fast forward genetics based on virus-induced gene silencing. *Current opinion in plant biology 2*, 109-113.

Baum, S. F., Eshed, Y., and Bowman, J. L. (2001). The *Arabidopsis* nectary is an ABC-independent floral structure. *Development 128*, 4657–4667.

Bayat, S., Schranz, M. E., Roalson, E. H., and Hall, J. C. (2018). Lessons from Cleomaceae, the Sister of Crucifers. *Trends Plant Sci 23*, 808–821.

Baylis, T., Cierlik, I., Sundberg, E., and Mattsson, J. (2013). SHORT INTERNODES/STYLISH genes, regulators of auxin biosynthesis, are involved in leaf vein development in *Arabidopsis thaliana*. *New Phytologist 197*, 737–750.

Becker, A., and Lange, M. (2010). VIGS - genomics goes functional. *Trends Plant Sci 15*, 1–4.

Bender, R. L., Fekete, M. L., Klinkenberg, P. M., Hampton, M., Bauer, B., Malecha, M., et al. (2013). PIN6 is required for nectary auxin response and short stamen development. *Plant Journal 74*, 893–904.

Bergh, E. van den, Hofberger, J. A., and Schranz, M. E. (2016). Flower power and the mustard bomb: Comparative analysis of gene and genome duplications in glucosinolate biosynthetic pathway evolution in cleomaceae and brassicaceae. *Am J Bot 103*, 1212–1222.

Bernardello, G. (2007). A systematic survey of floral nectaries. In *Nectaries and nectar* (pp. 19-128). Springer, Dordrecht.

Bhide, A., Schliesky, S., Reich, M., Weber, A. P. M., and Becker, A. (2014). Analysis of the floral transcriptome of Tarenaya hassleriana (Cleomaceae), a member of the sister group to the Brassicaceae: towards understanding the base of morphological diversity in Brassicales. *BMC Genomics* 15, 140.

Boase, M. R., Ngo, H., Jameson, P. E., and Schwinn, K. E. (2014). A Conserved Network of Transcriptional Activators and Repressors Regulates Anthocyanin Pigmentation in Eudicots. *The Plant Cell 26*, 962-980.

Bowman, J. L., and Smyth, D. R. (1999). CRABS CLAW, a gene that regulates carpel and nectary development in *Arabidopsis*, encodes a novel protein with zinc finger and helix-loop-helix domains. *Development* 126, 2387–2396.

Braun, H. P. (2020). The Oxidative Phosphorylation system of the mitochondria in plants. *Mitochondrion* 53, 66–75.

Bräutigam, A., Kajala, K., Wullenweber, J., Sommer, M., Gagneul, D., Weber, K. L., et al. (2011). An mRNA blueprint for C4 photosynthesis derived from comparative transcriptomics of closely related C3 and C4 species. *Plant Physiol* 155, 142–156.

Bräutigam, A., Mullick, T., Schliesky, S., and Weber, A. P. M. (2011). Critical assessment of assembly strategies for non-model species mRNA-Seq data and application of next-generation sequencing to the comparison of C3 and C4 species. *J Exp Bot* 62, 3093–3102.

Brock, K. C. (2014). Tracking the Evolutionary History of Development Genes: Implications for the Diversification of Fruits and Flowers in the Brassicaceae and Cleomaceae. *Master's thesis*. University of Alberta, Edmonton, Alberta, Canada.

Brock, K. C., and Hall, J. C. (2019). Multiple lineages of fruitfull exhibit dynamic patterns of gene evolution after genome triplication in the Brassiceae tribe (Brassicaceae). *Botany* 97, 293–310.

Broderick, S. R., and Jones, M. L. (2014). An Optimized Protocol to Increase Virus-Induced Gene Silencing Efficiency and Minimize Viral Symptoms in *Petunia*. *Plant Mol Biol Report* 32, 219–233.

Brunet, F. G., Crollius, H. R., Paris, M., Aury, J. M., Gibert, P., Jaillon, O., et al. (2006). Gene loss and evolutionary rates following whole-genome duplication in teleost fishes. *Mol Biol Evol* 23, 1808–1816.

Burch-Smith, T. M., Anderson, J. C., Martin, G. B., and Dinesh-Kumar, S. P. (2004). Applications and advantages of virus-induced gene silencing for gene function studies in plants. *Plant Journal* 39, 734–746.

Burch-Smith, T. M., Schiff, M., Liu, Y., and Dinesh-Kumar, S. P. (2006). Efficient virus-induced gene silencing in *Arabidopsis*. *Plant Physiol* 142, 21–27.

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.

Busch, A., Horn, S., & Zachgo, S. (2014). Differential transcriptome analysis reveals insight into monosymmetric corolla development of the crucifer Iberis amara. *BMC plant biology 14*, 1-17

Busch, A., and Zachgo, S. (2007). Control of corolla monosymmetry in the Brassicaceae Iberis amara. *Proceedings of the National Academy of Sciences* 104, 16714–16719.

Busch, A., and Zachgo, S. (2009). Flower symmetry evolution: Towards understanding the abominable mystery of angiosperm radiation. *BioEssays* 31, 1181–1190.

Cane, J. H. (2008). Breeding biologies, seed production and species-rich bee guilds of *Cleome lutea* and *Cleome serrulata* (Cleomaceae). *Plant Species Biol* 23, 152–158.

Cardinal-McTeague, W. M., Sytsma, K. J., and Hall, J. C. (2016). Biogeography and diversification of Brassicales: A 103 million year tale. *Mol Phylogenet Evol* 99, 204–224.

Cardoso, J. C. F., de Deus, F. F., Silva, S. C. S., and Oliveira, P. E. (2018). Bow to the middle: reproductive system and style behaviour of *Tococa guianensis*, a widespread Melastomataceae from the Neotropics. *Plant Systematics and Evolution* 304, 259–267.

Carey, S., Zenchyzen, B., Deneka, A. J., & Hall, J. C. (2023). Nectary development in *Cleome violacea*. *Frontiers in Plant Science* 13, 1085900.

Carey, S., Higuera-Díaz, M., Mankowski, P., Rocca, A., and Hall, J. C. (2021). Virus-induced gene silencing as a tool for functional studies in *Cleome violacea*. *Appl Plant Sci* 9, 1–13.

Carey, S., Mendler, K., and Hall, J. C. (2019). How to build a fruit: Transcriptomics of a novel fruit type in the Brassiceae. *PLoS One* 14, e0209535

Cazzonelli C. I. (2011). Carotenoids in nature: insights from plants and beyond. *Funct. Plant Biol.* 38, 833–847.

Chappell, C. R., and Fukami, T. (2018). Nectar yeasts: a natural microcosm for ecology. *Yeast* 35, 417–423.

Chávez Montes, R. A. C., Herrera-Ubaldo, H., Serwatowska, J., & de Folter, S. (2015). Towards a comprehensive and dynamic gynoecium gene regulatory network. *Current Plant Biology 3*, 3-12.

Chen, K., Li, G. J., Bressan, R. A., Song, C. P., Zhu, J. K., and Zhao, Y. (2020). Abscisic acid dynamics, signaling, and functions in plants. *J Integr Plant Biol* 62, 25–54.

Chen, T., Zhou, Y., Zhang, J., Peng, Y., Yang, X., Hao, Z., et al. (2021). Integrative analysis of transcriptome and proteome revealed nectary and nectar traits in the plant-pollinator interaction of *Nitraria tangutorum* Bobrov. *BMC Plant Biol* 21, 1–13.

Cheng, C. Y., Krishnakumar, V., Chan, A. P., Thibaud-Nissen, F., Schobel, S., and Town, C. D. (2017). Araport11: a complete reannotation of the *Arabidopsis thaliana* reference genome. *Plant Journal* 89, 789–804.

Cheng, F., Wu, J., Wang, X., Warwick, S., Francis, A., Al-Shehbaz, I., et al. (2014). Genome triplication drove the diversification of *Brassica* plants. *Hortic Res* 1, 14024.

Cheng, S., van den Bergh, E., Zeng, P., Zhong, X., Xu, J., Liu, X., et al. (2013). The Tarenaya hassleriana genome provides insight into reproductive trait and genome evolution of crucifers. *Plant Cell* 25, 2813–2830.

Chico J. M., Chini A., Fonseca S., And Solano R. (2008). *JAZ* repressors set the rhythm in jasmonate signaling. *Curr. Opin. Plant Biol.* 11, 486–494.

Chung, K. S., Lee, J. H., Lee, J. S., & Ahn, J. H. (2013). Fruit indehiscence caused by enhanced expression of NO TRANSMITTING TRACT in *Arabidopsis thaliana*. *Molecules and cells 35*, 519-525.

Citerne, H., Jabbour, F., Nadot, S., & Damerval, C. (2010). The evolution of floral symmetry. In *Advances in botanical research* 54, 85-137.

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.

Coen, E. S. (1996). Floral symmetry. *The EMBO Journal* 15, 6777-6788.

Conesa, A., Gotz, S., Garcia-Gomez, J., Terol, J., Talon, M., and Robles, M. (2005). Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics* 21, 3674–3676.

Conesa, A., Madrigal, P., Tarazona, S., Gomez-Cabrero, D., Cervera, A., McPherson, A., et al. (2016). A survey of best practices for RNA-seq data analysis. *Genome Biol* 17, 1–19.

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.

Crick, F. (1970). Central dogma of molecular biology. *Nature 227*, 561-563.

Crow, K. D., and Wagner, G. P. (2006). What is the role of genome duplication in the evolution of complexity and diversity? *Mol Biol Evol* 23, 887–892.

Cubas, P., Coen, E., and Zapater, J. M. M. (2001). Ancient asymmetries in the evolution of flowers. *Current Biology* 11, 1050–1052.

Damerval, C., and Becker, A. (2017). Genetics of flower development in Ranunculales – a new, basal eudicot model order for studying flower evolution. *New Phytologist* 216, 361–366.

Danisman, S. (2016). TCP Transcription Factors at the Interface between Environmental Challenges and the Plant's Growth Responses. *Front Plant Sci* 7, 1–13.

Dao, T. T. H., Linthorst, H. J. M., and Verpoorte, R. (2011). Chalcone synthase and its functions in plant resistance. *Phytochemistry Reviews* 10, 397–412.

Davidson, N. M., and Oshlack, A. (2014). Corset: Enabling differential gene expression analysis for de novo assembled transcriptomes. *Genome Biol* 15, 1–14.

Davies, K. M., Jibran, R., Zhou, Y., Albert, N. W., Brummell, D. A., Jordan, B. R., et al. (2020). The Evolution of Flavonoid Biosynthesis: A Bryophyte Perspective. *Front Plant Sci* 11, 1–21.

del Pozo, J. C., and Ramirez-Parra, E. (2015). Whole genome duplications in plants: An overview from *Arabidopsis*. *J Exp Bot* 66, 6991–7003.

del Valle, J. C., Alcalde-Eon, C., Escribano-Bailón, M. T., Buide, M. L., Whittall, J. B., and Narbona, E. (2019). Stability of petal color polymorphism: The significance of anthocyanin accumulation in photosynthetic tissues. *BMC Plant Biol* 19, 1–13.

Demmig-Adams, B. (1990). Carotenoids and photoprotection in plants: A role for the xanthophyll zeaxanthin. *Biochimica et Biophysica Acta (BBA) - Bioenergetics* 1020, 1–24.

Deng, X., Elomaa, P., Nguyen, C. X., Hytönen, T., Valkonen, J. P. T., and Teeri, T. H. (2012). Virus-induced gene silencing for Asteraceae-a reverse genetics approach for functional genomics in Gerbera hybrida. *Plant Biotechnol J* 10, 970–978.

Di Stilio, V. S., Kumar, R. A., Oddone, A. M., Tolkin, T. R., Salles, P., & McCarty, K. (2010). Virus-induced gene silencing as a tool for comparative functional studies in Thalictrum. *PLoS One 5*, e12064

Dinneny, J. R., Weigel, D., and Yanofsky, M. F. (2005). A genetic framework for fruit patterning in *Arabidopsis thaliana*. *Development* 132, 4687–4696.

Dinneny, J. R., and Yanofsky, M. F. (2005). Drawing lines and borders: How the dehiscent fruit of *Arabidopsis* is patterned. *BioEssays* 27, 42–49.

Dong, Y., and Wang, Y.-Z. (2015). Seed shattering: from models to crops. *Front Plant Sci* 6, 476.

Du, H., Wu, J., Ji, K. X., Zeng, Q. Y., Bhuiya, M. W., Su, S., et al. (2015). Methylation mediated by an anthocyanin, O-methyltransferase, is involved in purple flower coloration in Paeonia. *J Exp Bot* 66, 6563–6577.

Du, X., Liu, H., Zhu, Z., Liu, S., Song, Z., Xia, L., et al. (2022). Identification of Candidate Chromosome Region Related to Melon (Cucumis melo L.) Fruit Surface Groove Trait Through Biparental Genetic Mapping and Genome-Wide Association Study. *Front Plant Sci* 13, 828287.

Dubos, C., Le Gourrierec, J., Baudry, A., Huep, G., Lanet, E., Debeaujon, I., ... & Lepiniec, L. (2008). MYBL2 is a new regulator of flavonoid biosynthesis in *Arabidopsis thaliana*. *The Plant Journal 55*, 940-953.

Dudek, B., Schneider, B., Hilger, H. H., Stavenga, D. G., and Martínez-Harms, J. (2020). Highly different flavonol content explains geographic variations in the UV reflecting properties of flowers of the corn poppy, *Papaver rhoeas* (Papaveraceae). *Phytochemistry* 178, 112457.

Dudek, B., Warskulat, A. C., and Schneider, B. (2016). The occurrence of flavonoids and related compounds in flower sections of *Papaver nudicaule*. *Plants* 5, 93–126.

Duncan, T. M., and Rausher, M. D. (2020). Selection favors loss of floral pigmentation in a highly selfing morning glory. *PLoS One* 15, 1–18.

Edger, P. P., Hall, J. C., Harkess, A., Tang, M., Coombs, J., Mohammadin, S., et al. (2018). Brassicales phylogeny inferred from 72 plastid genes: A reanalysis of the phylogenetic localization of two paleopolyploid events and origin of novel chemical defenses. *Am J Bot* 105, 463–469.

Edger, P. P., Heidel-Fischer, H. M., Bekaert, M., Rota, J., Glöckner, G., Platts, A. E., et al. (2015). The butterfly plant arms-race escalated by gene and genome duplications. *Proceedings of the National Academy of Sciences 112*, 8362-8366.

Edwards, M. B., Ballerini, E. S., and Kramer, E. M. (2022). Complex developmental and transcriptional dynamics underlie pollinator-driven evolutionary transitions in nectar spur morphology in Aquilegia (columbine). *Am J Bot 109*, 1360-1381.

Eldridge, T., Ł, Ł., Stacey, N., Jantzen, F., Moubayidin, L., Sicard, A., et al. (2016). Fruit shape diversity in the Brassicaceae is generated by varying patterns of anisotropy. *Development* 143, 3394-3406

Ellner, S., & Shmida, A. (1981). Why are adaptations for long-range seed dispersal rare in desert plants?. *Oecologia 51*, 133-144.

Emms, D. M., & Kelly, S. (2015). OrthoFinder: solving fundamental biases in whole genome comparisons dramatically improves orthogroup inference accuracy. *Genome biology 16*, 1-14.

Endress, P. K. (1992). Evolution and floral diversity: the phylogenetic surroundings of *Arabidopsis* and *Antirrhinum*. *International Journal of Plant Sciences 153*, 106-122.

Endress, P. K. (1999). Symmetry in flowers: Diversity and evolution. *Int J Plant Sci* 160, 3-23

Endress, P. K. (2011). Evolutionary diversification of the flowers in angiosperms. *Am J Bot* 98, 370–396.

Endress, P. K. (2012). The Immense Diversity of Floral Monosymmetry and Asymmetry Across Angiosperms. *Botanical Review* 78, 345–397.

Erbar C., Leins P. (1997. a). Different patterns of floral development in whorled flowers, exemplified by apiaceae and brassicaceae. *Int. J. Plant Sci* 158, 49–64.

Erbar, C., and Leins, P. (1997b). Studies on the early floral development in Cleomoideae (Capparaceae) with emphasis on the androecial development. *Plant Systematics and Evolution* 206, 119–132.

Farré-Armengol, G., Fernández-Martínez, M., Filella, I., Junker, R. R., and Peñuelas, J. (2020). Deciphering the Biotic and Climatic Factors That Influence Floral Scents: A Systematic Review of Floral Volatile Emissions. *Front Plant Sci* 11, 1–16.

Feng, X., Zhao, Z., Tian, Z., Xu, S., Luo, Y., Cai, Z., et al. (2006). Control of petal shape and floral zygomorphy in Lotus japonicus. *Proceedings of the National Academy of Sciences* 103, 4970–4975.

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.

Ferrandiz, C. (2002). Regulation of fruit dehiscence in *Arabidopsis*. *J Exp Bot* 53, 2031–2038.

Ferrándiz, C., and Fourquin, C. (2014). Role of the FUL-SHP network in the evolution of fruit morphology and function. *J Exp Bot 65*, 4505-4513.

Ferrándiz, C. (2002). Regulation of fruit dehiscence in *Arabidopsis*. *Journal of Experimental Botany*, 53, 2031-2038.

Ferrándiz, C., Gu, Q., Martienssen, R., and Yanofsky, M. F. (2000a). Redundant regulation of meristem identity and plant architecture by FRUITFULL, APETALA1 and CAULIFLOWER. *Development* 127, 725–734.

Ferrándiz, C., Liljegren, S. J., and Yanofsky, M. F. (2000b). Negative regulation of the SHATTERPROOF genes by FRUITFULL during *Arabidopsis* fruit development. *Science* 289, 436–438.

Finkelstein, R. (2013). Abscisic acid synthesis and response. *The Arabidopsis book/American society of plant biologists 11*, e0166.

Fleming, T. H., Geiselman, C., and Kress, W. J. (2009). The evolution of bat pollination: A phylogenetic perspective. *Ann Bot* 104, 1017–1043.

Fourquin, C., and Ferrándiz, C. (2014). The essential role of NGATHA genes in style and stigma specification is widely conserved across eudicots. *New Phytologist* 202, 1001-1013

Fourquin, C., Primo, A., Martínez-Fernández, I., Huet-Trujillo, E., & Ferrándiz, C. (2014). The CRC orthologue from *Pisum sativum* shows conserved functions in carpel morphogenesis and vascular development. *Annals of botany 114*, 1535-1544

Francis, W. R., Christianson, L. M., Kiko, R., Powers, M. L., Shaner, N. C., and D Haddock, S. H. (2013). A comparison across non-model animals suggests an optimal sequencing depth for de novo transcriptome assembly. *BMC Genomics* 14, 1–12.

Franzke, A., Lysak, M. A., Al-Shehbaz, I. A., Koch, M. A., and Mummenhoff, K. (2011). Cabbage family affairs: The evolutionary history of Brassicaceae. *Trends Plant Sci* 16, 108–116.

Fu, D. Q., Zhu, B. Z., Zhu, H. L., Zhang, H. X., Xie, Y. H., Jiang, W. B., et al. (2006). Enhancement of virus-induced gene silencing in tomato by low temperature and low humidity. *Mol Cells* 21, 153–160.

Fujita, N., Kazama, Y., Yamagishi, N., Watanabe, K., Ando, S., Tsuji, H., et al. (2019). Development of the VIGS system in the dioecious plant *Silene latifolia*. *Int J Mol Sci* 20, 1–12.

Furlan, M., de Pretis, S., and Pelizzola, M. (2021). Dynamics of transcriptional and post-transcriptional regulation. *Brief Bioinform* 22, 1–13.

Galego, L., and Almeida, J. (2002). Role of DIVARICATA in the control of dorsoventral asymmetry in *Antirrhinum* flowers. *Genes Dev* 16, 880–891.

Galliot, C., Hoballah, M. E., Kuhlemeier, C., and Stuurman, J. (2006). Genetics of flower size and nectar volume in *Petunia* pollination syndromes. *Planta* 225, 203–212.

Galliot, C., Stuurman, J., and Kuhlemeier, C. (2006). The genetic dissection of floral pollination syndromes. *Curr Opin Plant Biol* 9, 78–82.

Gao, Y., Zhang, D., & Li, J. (2015). *TCP1* modulates DWF4 expression via directly interacting with the GGNCCC motifs in the promoter region of DWF4 in *Arabidopsis thaliana*. Journal of Genetics and Genomics 42, 383-392.

Garcia-Hernandez, M., Berardini, T., Chen, G., Crist, D., and Doyle, A. (2002). TAIR: a resource for intergrated *Arabidopsis* data. *Funct Integr Genomics* 2, 239–253.

Gautier-Hion, A., Duplantier, J., Quris, R., Feer, F., Sourd, C., Decoux, J., Dubost, G., Emmons, L., Erard, C., Hecketsweiler, P., Moungazi, A., Roussilhon, C., Thiollay, J. (1985). Fruit characters as a basis of fruit choice and seed dispersal in a tropical forest vertebrate community. *Oecologia 65*, 324-337.

Girin, T., Paicu, T., Stephenson, P., Fuentes, S., Korner, E., O'Brien, M., et al. (2011). INDEHISCENT and SPATULA Interact to Specify Carpel and Valve Margin Tissue and Thus Promote Seed Dispersal in *Arabidopsis*. *Plant Cell* 23, 3641–3653.

Girin, T., Stephenson, P., Goldsack, C. M. P., Kempin, S. A., Perez, A., Pires, N., et al. (2010). Brassicaceae INDEHISCENT genes specify valve margin cell fate and repress replum formation. *Plant Journal* 63, 329–338.

Gómez-Campo, C. (1980). Morphology and morpho-taxonomy of the tribe Brassiceae. *Brassica crops and wild allies 1,* 3-31.

Gómez-Campo, C. (1999). Seedless and seeded beaks in the tribe Brassiceae. *Cruciferae Newsletter*, 11-12.

Gong, Y. B., and Huang, S. Q. (2009). Floral symmetry: Pollinator-mediated stabilizing selection on flower size in bilateral species. *Proceedings of the Royal Society B: Biological Sciences* 276, 4013–4020.

González-Reig, S., Ripoll, J. J., Vera, A., Yanofsky, M. F., and Martínez-Laborda, A. (2012). Antagonistic Gene Activities Determine the Formation of Pattern Elements along the Mediolateral Axis of the *Arabidopsis* Fruit. *PLoS Genet 8*, e1003020.

Gould, B., and Kramer, E. M. (2007). Virus-induced gene silencing as a tool for functional analyses in the emerging model plant Aquilegia (columbine, Ranunculaceae). *Plant Methods* 3, 1–12.

Grabherr, M. G., Haas, B. J., Yassour, M., Levin, J. Z., Thompson, D. A., Amit, I., et al. (2011). Trinity: reconstructing a full-length transcriptome without a genome from RNA-Seq data HHS Public Access. *Nat Biotechnol. Nat Biotechnol* 29, 644–652.

Gross T., Broholm S., And Becker A. (2018). CRABS CLAW acts as a bifunctional transcription factor in flower development. *Front. Plant Sci* 9.

Groszmann, M., Paicu, T., Alvarez, J. P., Swain, S. M., and Smyth, D. R. (2011). SPATULA and ALCATRAZ, are partially redundant, functionally diverging bHLH genes required for *Arabidopsis* gynoecium and fruit development. *Plant Journal* 68, 816–829.

Grotewold, E. (2006). The genetics and biochemistry of floral pigments. *Annu Rev Plant Biol* 57, 761–780.

Gu, Q., Ferrandiz, C., Yanofsky, M. F., and Martienssen, R. (1998). The FRUITFULL MADS-box gene mediates cell differentiation during *Arabidopsis* fruit development. *Development* 125, 1509–1517.

Guerrero, G., Peralta, H., Aguilar, A., Dí, R., Villalobos, M. A., Medrano-Soto, A., Mora J. (2005). Evolutionary, structural and functional relationships revealed by comparative analysis of syntenic genes in Rhizobiales. *BMC Evol Biol* 5, 1-19.

Guo, M., Thomas, J., Collins, G., and Timmermans, M. C. P. (2008). Direct Repression of KNOX Loci by the ASYMMETRIC LEAVES1 Complex of *Arabidopsis*. *the Plant Cell* 20, 48–58.

Haas, B. J., Dobin, A., Li, B., Stransky, N., Pochet, N., and Regev, A. (2019). Accuracy assessment of fusion transcript detection via read-mapping and de novo fusion transcript assembly-based methods. *Genome Biol* 20, 1–16.

Haas, B. J., Papanicolaou, A., Yassour, M., Grabherr, M., Blood, P. D., Bowden, J., et al. (2013). De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. *Nat Protoc* 8, 1494.

Hall, J. C., Tisdale, T. E., Donohue, K., & Kramer, E. M. (2006). Developmental basis of an anatomical novelty: heteroarthrocarpy in *Cakile lanceolata* and *Erucaria erucarioides* (Brassicaceae). *International Journal of Plant Sciences 167*, 771-789

Hall, J. C., Tisdale, T. E., Donohue, K., Wheeler, A., Al-Yahya, M. A., and Kramer, E. M. (2011). Convergent evolution of a complex fruit structure in the tribe Brassiceae (Brassicaceae). *Am J Bot* 98, 1989–2003.

Halperin, R. F., Hegde, A., Lang, J. D., Raupach, E. A., Narayanan, V., Huentelman, M., et al. (2021). Improved methods for RNAseq-based alternative splicing analysis. *Sci Rep* 11, 1–15.

Han X., Yin L., And Xue H. (2012). Co-Expression analysis identifies CRC and AP1 the regulator of arabidopsis fatty acid biosynthesis. *J. Integr. Plant Biol* 54, 486–499.

Hart, A. J., Ginzburg, S., Xu, M., Fisher, C. R., Rahmatpour, N., Mitton, J. B., et al. (2020). EnTAP: Bringing faster and smarter functional annotation to non-model eukaryotic transcriptomes. *Mol Ecol Resour* 20, 591–604.

Hart, N. S., and Hunt, D. M. (2007). Avian visual pigments: Characteristics, spectral tuning, and evolution. *American Naturalist* 169, 7-26.

Havaux, M. (2014). Carotenoid oxidation products as stress signals in plants. *Plant Journal* 79, 597–606.

He, F., Mu, L., Yan, G. L., Liang, N. N., Pan, Q. H., Wang, J., et al. (2010). Biosynthesis of anthocyanins and their regulation in colored grapes. *Molecules* 15, 9057–9091.

Heil, M. (2011). Nectar: Generation, regulation and ecological functions. *Trends Plant Sci* 16, 191–200.

Hernandez-Garcia, C. M., Finer, J. J. (2014). Identification and validation of promoters and cis-acting regulatory elements. *Plant Science 217*, 109-119.

Hichri, I., Barrieu, F., Bogs, J., Kappel, C., Delrot, S., and Lauvergeat, V. (2011). Recent advances in the transcriptional regulation of the flavonoid biosynthetic pathway. *J Exp Bot* 62, 2465–2483.

Hidalgo, O., Bartholmes, C., and Gleissberg, S. (2012). Virus-induced gene silencing (VIGS) in *Cysticapnos vesicaria*, a zygomorphic-flowered Papaveraceae (Ranunculales, basal eudicots). *Ann Bot* 109, 911–920.

Higuera-Díaz, M., Manson, J. S., and Hall, J. C. (2015). Pollination biology of *Cleomella serrulata* and *Polanisia dodecandra* in a protected natural prairie in southern Alberta, Canada. *Botany* 93, 745–757.

Hileman, L. C. (2014a). Bilateral flower symmetry - how, when and why? *Curr Opin Plant Biol* 17, 146–152.

Hileman, L. C. (2014b). Trends in flower symmetry evolution revealed through phylogenetic and developmental genetic advances. *Philosophical Transactions of the Royal Society B: Biological Sciences* 369, 20130348.

Hölzer, M., and Marz, M. (2019). De novo transcriptome assembly: A comprehensive cross-species comparison of short-read RNA-Seq assemblers. *Gigascience* 8, 1–16.

Honaas, L. A., Wafula, E. K., Wickett, N. J., Der, J. P., Zhang, Y., Edger, P. P., et al. (2016). Selecting superior de novo transcriptome assemblies: Lessons learned by leveraging the best plant genome. *PLoS One* 11, e0146062.

Hooghvorst, I., López-Cristoffanini, C., and Nogués, S. (2019). Efficient knockout of phytoene desaturase gene using CRISPR/Cas9 in melon. *Sci Rep* 9, 1–7.

Hosseini Tafreshi, S. A., Shariati, M., Mofid, M. R., Nekui, M. K., and Esmaeili, A. (2012). Heterologous virus-induced gene silencing as a promising approach in plant functional genomics. *Mol Biol Rep* 39, 2169–2178.

Hsieh, M. H., Lu, H. C., Pan, Z. J., Yeh, H. H., Wang, S. S., Chen, W. H., et al. (2013). Optimizing virus-induced gene silencing efficiency with Cymbidium mosaic virus in *Phalaenopsis* flower. *Plant Science* 201–202, 25–41.

Hsu, C. C., Chen, Y. Y., Tsai, W. C., Chen, W. H., and Chen, H. H. (2015). Three R2R3-MYB transcription factors regulate distinct floral pigmentation patterning in Phalaenopsis spp. *Plant Physiol* 168, 175–191.

Iltis, H. H., Hall, J. C., Cochrane, T. S., & Sytsma, K. J. (2011). Studies in the Cleomaceae I. On the Separate Recognition of Capparaceae, Cleomaceae, and Brassicaceae1. *Annals of the Missouri Botanical Garden 98*, 28-36.

Irish, V. (2017). The ABC model of floral development. *Current Biology* 27, 887–890.

Jabbour, F., Nadot, S., and Damerval, C. (2009). Evolution of floral symmetry: a state of the art. *C R Biol* 332, 219–231.

Jacquemyn, H., Lenaerts, M., Tyteca, D., and Lievens, B. (2013). Microbial diversity in the floral nectar of seven Epipactis (Orchidaceae) species. *Microbiologyopen* 2, 644–658.

Jacquemyn, H., Pozo, M. I., Álvarez-Pérez, S., Lievens, B., and Fukami, T. (2021). Yeast–nectar interactions: metacommunities and effects on pollinators. *Curr Opin Insect Sci* 44, 35–40.

Janson, C. H. (1983). Adaptation of Fruit Morphology to Dispersal Agents in a Neotropical Forest Published by : American Association for the Advancement of Science Stable URL : http://www.jstor.org/stable/1690124 Adaptation of Fruit Morphology to Dispersal Agents in a Neotropi. *Adv Sci* 219, 187–189.

Jaradat, M. R., Ruegger, M., Bowling, A., Butler, H., & Cutler, A. J. (2014). A comprehensive transcriptome analysis of silique development and dehiscence in *Arabidopsis* and *Brassica* integrating genotypic, interspecies and developmental comparisons. *GM crops & food 5*, 302-320.

Jung J. H., Seo Y. H., Pil J. S., Reyes J. L., Yun J., Chua N. H., et al. . (2007). The *GIGANTEA*-regulated microRNA172 mediates photoperiodic flowering independent of *CONSTANS* in *Arabidopsis*. *Plant Cell* 19, 2736–2748.

Juntheikki-Palovaara, I., Tähtiharju, S., Lan, T., Broholm, S. K., Rijpkema, A. S., Ruonala, R., and Elomaa, P. (2014). Functional diversification of duplicated CYC 2 clade genes in regulation of inflorescence development in *Gerbera hybrida* (Asteraceae). *The Plant Journal 79*, 783-796

Katzer, A. M., Wessinger, C. A., and Hileman, L. C. (2019). Nectary size is a pollination syndrome trait in *Penstemon*. *New Phytologist* 223, 377–384.

Kay, P., Groszmann, M., Ross, J. J., Parish, R. W., and Swain, S. M. (2013). Modifications of a conserved regulatory network involving INDEHISCENT controls multiple aspects of reproductive tissue development in *Arabidopsis*. *New Phytologist* 197, 73-87.

Khoo, H. E., Azlan, A., Tang, S. T., and Lim, S. M. (2017). Anthocyanidins and anthocyanins: Colored pigments as food, pharmaceutical ingredients, and the potential health benefits. *Food Nutr Res* 61, 1361779.

Kieffer, M., Master, V., Waites, R., and Davies, B. (2011). *TCP14* and *TCP15* affect internode length and leaf shape in *Arabidopsis*. *Plant Journal* 68, 147–158.

Kilambi, H. V., Kumar, R., Sharma, R., and Sreelakshmi, Y. (2013). Chromoplast-specific carotenoid-associated protein appears to be important for enhanced accumulation of carotenoids in hp1 tomato fruits. *Plant Physiol* 161, 2085–2101.

Kim, D., Jeong, S., Samantha, J., Sung, Y., Park, C., Soo, H., et al. (2022). MYB3 plays an important role in lignin and anthocyanin biosynthesis under salt stress condition in *Arabidopsis*. *Plant Cell Rep 41*, 1549-1560.

Kim, I. v., Ross, E. J., Dietrich, S., Döring, K., Sánchez Alvarado, A., and Kuhn, C. D. (2019). Efficient depletion of ribosomal RNA for RNA sequencing in planarians. *BMC Genomics* 20, 1–12.

Koch, M. A., and Lemmel, C. (2019). *Zahora*, a new monotypic genus from tribe brassiceae (brassicaceae) endemic to the moroccan sahara. *PhytoKeys* 135, 119.

Koch, M., Al-Shehbaz, I. A., & Mummenhoff, K. (2003). Molecular systematics, evolution, and population biology in the mustard family (Brassicaceae). *Annals of the Missouri Botanical Garden* 90, 151-171

Koenig, D., and Weigel, D. (2015). Beyond the thale: comparative genomics and genetics of *Arabidopsis* relatives. *Nat Rev Genet* 16, 285.

Kong, J.-M., Chia, L.-S., Goh, N.-K., Chia, T.-F., and Brouillard, R. (2003). Analysis and biological activities of anthocyanins. *Phytochemistry* 64, 923–933.

Koralewski, T. E., and Krutovsky, K. v. (2011). Evolution of exon-intron structure and alternative splicing. *PLoS One* 6, e18055.

Kosugi, S., & Ohashi, Y. (2002). DNA binding and dimerization specificity and potential targets for the TCP protein family. *The Plant Journal 30*, 337-348.

Koteyeva, N. K., Voznesenskaya, E. v., Cousins, A. B., and Edwards, G. E. (2014). Differentiation of C4 photosynthesis along a leaf developmental gradient in two Cleome species having different forms of Kranz anatomy. *J Exp Bot* 65, 3525–3541.

Koteyeva, N. K., Voznesenskaya, E. v., Roalson, E. H., and Edwards, G. E. (2011). Diversity in forms of C4 in the genus *Cleome* (Cleomaceae). *Ann Bot* 107, 269–283.

Kramer, E. M., & Irish, V. F. (1999). Evolution of genetic mechanisms controlling petal development. *Nature 399*, 144-148

Kramer, E. M. (2019). Plus ça change, plus c'est la même chose: The developmental evolution of flowers. *Current topics in developmental biology 131*, 211-238.

Kramer, E. M., Dorit, R. L., and Irish, V. F. (1998). Molecular evolution of genes controlling petal and stamen development: Duplication and divergence within the APETALA3 and PISTILLATA MADS-box gene lineages. *Genetics* 149, 765–783.

Kramer, E. M., Holappa, L., Gould, B., Jaramillo, M. A., Setnikov, D., and Santiago, P. M. (2007). Elaboration of B gene function to include the identity of novel floral organs in the lower eudicot Aquilegia. *Plant Cell* 19, 750–766.

Kramer, E. M., Kong, H., and Rausher, M. D. (2017). Plant evolutionary developmental biology. Introduction to a special issue. *New Phytologist* 216, 335–336.

Křeček, P., Skůpa, P., Libus, J., Naramoto, S., Tejos, R., Friml, J., & Zažímalová, E. (2009). The PIN-FORMED (PIN) protein family of auxin transporters. *Genome biology 10*, 1-11.

Krogan N. T., Hogan K., And Long J. A. (2012). *APETALA2* negatively regulates multiple floral organ identity genes in arabidopsis by recruiting the co-repressor *TOPLESS* and the histone deacetylase HDA19. *Development* 139, 4180–4190.

Krueger, F. (2012). Trim Galore!. A wrapper tool around Cutadapt and FastQC to consistently apply quality and adapter trimming to FastQ files. *Babraham Bioinformatics 516*.

Külahoglu, C., Denton, A. K., Sommer, M., Maß, J., Schliesky, S., Wrobel, T. J., ... and Weber, A. P. (2014). Comparative transcriptome atlases reveal altered gene expression modules between two Cleomaceae C3 and C4 plant species. *The Plant Cell 26*, 3243-3260.

Kuusk S., Sohlberg J. J., Long J. A., Fridborg I., Sundberg E. (2002). *STY1* and *STY2* promote the formation of apical tissues during Arabidopsis gynoecium development. *Development* 129, 4707–4717.

Lai, X., Bendix, C., Yan, L., Zhang, Y., Schnable, J. C., and Harmon, F. G. (2020). Interspecific analysis of diurnal gene regulation in panicoid grasses identifies known and novel regulatory motifs. *BMC Genomics* 21, 1-17.

Laitinen, R. A. E., Ainasoja, M., Broholm, S. K., Teeri, T. H., and Elomaa, P. (2008). Identification of target genes for a MYB-type anthocyanin regulator in Gerbera hybrida. *J Exp Bot* 59, 3691–3703.

Landis, J. B., O'Toole, R. D., Ventura, K. L., Gitzendanner, M. A., Oppenheimer, D. G., Soltis, D. E., & Soltis, P. S. (2016). The phenotypic and genetic underpinnings of flower size in Polemoniaceae. *Frontiers in Plant Science 6*, 1144

Langmead, B. (2010). Aligning short sequencing reads with Bowtie. *Current protocols in bioinformatics 32*, 11-7.

Langmead, B., Trapnell, C., Pop, M., & Salzberg, S. L. (2009). Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome biology 10*, 1-10.

Łangowski, Ł., Stacey, N., and Østergaard, L. (2016). Diversification of fruit shape in the Brassicaceae family. *Plant Reprod* 29, 149–163.

Lázaro, A., and Totland, O. (2014). The influence of floral symmetry, dependence on pollinators and pollination generalization on flower size variation. *Ann Bot* 114, 157–165.

Lee, A. P., Koh, E. G., Tay, A., Brenner, S., & Venkatesh, B. (2006). Highly conserved syntenic blocks at the vertebrate Hox loci and conserved regulatory elements within and outside Hox gene clusters. *Proceedings of the National Academy of Sciences 103*, 6994-6999.

Lee, J. H., Regmi, S. C., Kim, J. A., Cho, M. H., Yun, H., Lee, C. S., et al. (2011). Apple flavonoid phloretin inhibits Escherichia coli O157:H7 biofilm formation and ameliorates colon inflammation in rats. *Infect Immun* 79, 4819–4827.

Lee, J. Y., Baum, S. F., Alvarez, J., Patel, A., Chitwood, D. H., and Bowman, J. L. (2005a). Activation of CRABS CLAW in the nectaries and carpels of *Arabidopsis*. *Plant Cell* 17, 25–36.

Lee, J. Y., Baum, S. F., Oh, S. H., Jiang, C. Z., Chen, J. C., and Bowman, J. L. (2005b). Recruitment of CRABS CLAW to promote nectary development within the eudicot clade. *Development* 132, 5021–5032.

Lenser, T., & Theißen, G. (2013). Conservation of fruit dehiscence pathways between *Lepidium campestre* and *Arabidopsis thaliana* sheds light on the regulation of INDEHISCENT. *The Plant Journal 76*, 545-556

Levin, S. A., Muller-Landau, H. C., Nathan, R., and Chave, J. (2003). The Ecology And Evolution Of Seed Dispersal: A Theoretical Perspective. *Annu. Rev. Ecol. Evol. Syst* 34, 575–604.

Liao, I. T., Hileman, L. C., and Roy, R. (2021). On the horizon for nectar-related research. *Am J Bot* 108, 2326–2330.

Liljegren, S. J., Ditta, G. S., Eshed, Y., Savidge, B., Bowman, J. L., and Yanofsky, M. F. (2000). *SHATTERPROOF* MADS-box genes control seed dispersal in *Arabidopsis*. *Nature* 404, 766–770.

Liljegren, S. J., Roeder, A. H., Kempin, S. A., Gremski, K., Østergaard, L., Guimil, S., ... and Yanofsky, M. F. (2004). Control of fruit patterning in *Arabidopsis* by INDEHISCENT. *Cell 116*, 843-853

Li B. B., Liu Y. G., Tao W. U., Wang J. P., Xie G. R., Chu Z. H., et al. . (2019). *OsBGLU19* and *OsBGLU23* regulate disease resistance to bacterial leaf streak in rice. *J. Integr. Agric* 18, 1199–1210.

Lin, I. W., Sosso, D., Chen, L. Q., Gase, K., Kim, S. G., Kessler, D., ... & Frommer, W. B. (2014). Nectar secretion requires sucrose phosphate synthases and the sugar transporter SWEET9. *Nature 508*, 546-549.

Liu, E., and Page, J. E. (2008). Optimized cDNA libraries for virus-induced gene silencing (VIGS) using tobacco rattle virus. *Plant Methods* 4, 1–13.

Liu, J., Wu, J., Yang, X., and Wang, Y. Z. (2021). Regulatory pathways of CYC-like genes in patterning floral zygomorphy exemplified in *Chirita pumila*. *J Syst Evol* 59, 567–580.

Long, X., and Miano, J. M. (2007). Remote control of gene expression. *Journal of Biological Chemistry* 282, 15941–15945.

Lück, S., Kreszies, T., Strickert, M., Schweizer, P., Kuhlmann, M., and Douchkov, D. (2019). siRNA-Finder (si-Fi) Software for RNAi-Target Design and Off-Target Prediction. *Front Plant Sci* 10, 1–12.

Lunn J. E., Delorge I., Figueroa C. M., Van Dijck P., And Stitt M. (2014). Trehalose metabolism in plants. *Plant J* 79, 544–567.

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

Luo K. M., Harding S. A., Tsai C. J. (2008). A modified T-vector for simplified assembly of hairpin RNAi constructs. *Biotechnol. Lett* 30, 1271–1274.

Mabry, M. E., Brose, J. M., Blischak, P. D., Sutherland, B., Dismukes, W. T., Bottoms, C. A.,... and Pires, J. C. (2020). Phylogeny and multiple independent whole-genome duplication events in the Brassicales. *Am J Bot* 107, 1148–1164.

Macior, L. W. (1971). Co-Evolution of Plants and Animals—Systematic Insights From Plant-Insect Interactions. *Taxon*, *20*, 17-28.

Marsch-Martínez, N., Zúñiga-Mayo, V. M., Herrera-Ubaldo, H., Ouwerkerk, P. B., Pablo-Villa, J., Lozano-Sotomayor, P., ... and de Folter, S. (2014). The NTT transcription factor promotes replum development in *Arabidopsis* fruits. *The Plant Journal 80*, 69-81.

Marshall, D. M., Muhaidat, R., Brown, N. J., Liu, Z., Stanley, S., Griffiths, H., et al. (2007). Cleome, a genus closely related to *Arabidopsis*, contains species spanning a developmental progression from C3 to C4 photosynthesis. *Plant Journal* 51, 886–896.

Martínez-Salazar, S., González, F., Alzate, J. F., and Pabón-Mora, N. (2021). Molecular framework underlying floral bilateral symmetry and nectar spur development in *Tropaeolum*, an atypical member of the Brassicales. *Am J Bot* 108, 1315–1330.

Martins, T. R., Jiang, P., and Rausher, M. D. (2017). How petals change their spots: cis-regulatory re-wiring in Clarkia (Onagraceae). *New Phytologist* 216, 510–518.

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

Marzol E., Borassi C., Sardoy M. C., Ranocha P., Aptekmann A. A., Bringas M., et al. . (2022). Class III peroxidases PRX01, PRX44, and PRX73 control root hair growth in arabidopsis thaliana. *Int. J. Mol.* Sci. 23, 5375.

Mbandi, S. K., Hesse, U., Rees, D. J. G., and Christoffels, A. (2014). A glance at quality score: Implication for de novo transcriptome reconstruction of Illumina reads. *Front Genet* 5, 1–5.

McCarthy, D. Chen, Y., and Smyth, K.(2012) Differential expression analysis of multifactor RNA-Seq experiments with respect to biological variation. *Nucleic Acids Res* 40, 4288–4297.

McKim, S. M., Stenvik, G. E., Butenko, M. A., Kristiansen, W., Cho, S. K., Hepworth, S. R., et al. (2008). The BLADE-ON-PETIOLE genes are essential for abscission zone formation in *Arabidopsis*. *Development* 135, 1537–1546.

Medina-Puche, L., Cumplido-Laso, G., Amil-Ruiz, F., Hoffmann, T., Ring, L., Rodríguez-Franco, A., et al. (2014). MYB10 plays a major role in the regulation of flavonoid/phenylpropanoid metabolism during ripening of *Fragaria × Ananassa* fruits. *J Exp Bot* 65, 401–417.

Mekapogu, M., Vasamsetti, B. M. K., Kwon, O. K., Ahn, M. S., Lim, S. H., and Jung, J. A. (2020). Anthocyanins in floral colors: Biosynthesis and regulation in *chrysanthemum* flowers. *Int J Mol Sci* 21, 1–24.

Miller, R., Owens, S. J., and Rørslett, B. (2011). Plants and colour: Flowers and pollination. *Opt Laser Technol* 43, 282–294.

Min, Y., Bunn, J. I., and Kramer, E. M. (2019). Homologs of the *STYLISH* gene family control nectary development in *Aquilegia*. *New Phytologist* 221, 1090–1100.

Mizukami, Y., & Ma, H. (1997). Determination of Arabidopsis floral meristem identity by *AGAMOUS*. *The Plant Cell* 9, 393-408.

Molina A., García-Olmedo F. (1997). Enhanced tolerance to bacterial pathogens caused by the transgenic expression of barley lipid transfer protein *LTP2*. *Plant J* 12, 669–675.

Moore, G., Devos, K. M., Wang, Z., & Gale, M. D. (1995). Cereal genome evolution: grasses, line up and form a circle. *Current biology 5*, 737-739

Morrant D. S., Schumann R., And Petit S. (2009). Field methods for sampling and storing nectar from flowers with low nectar volumes. *Ann. Bot.* 103, 533–542.

Morel, P., Heijmans, K., Ament, K., Chopy, M., Trehin, C., Chambrier, P., ... and Vandenbussche, M. (2018). The floral C-lineage genes trigger nectary development in *Petunia* and *Arabidopsis*. *The Plant Cell 30*, 2020-2037.

Muday, G. K., Rahman, A., and Binder, B. M. (2012). Auxin and ethylene: collaborators or competitors? *Trends Plant Sci* 17, 181–195.

Mühlhausen, A., Lenser, T., Mummenhoff, K., & Theißen, G. (2013). Evidence that an evolutionary transition from dehiscent to indehiscent fruits in *Lepidium* (Brassicaceae) was caused by a change in the control of valve margin identity genes. *The Plant Journal 73*, 824-835.

Nagpal, P., Ellis, C. M., Weber, H., Ploense, S. E., Barkawi, L. S., Guilfoyle, T. J., et al. (2005). Auxin response factors *ARF6* and *ARF8* promote jasmonic acid production and flower maturation. *Development* 132, 4107–4118.

Naparlo, K., Zyracka, E., Bartosz, G., and Sadowska-Bartosz, I. (2019). Flavanols protect the yeast *Saccharomyces cerevisiae* against heating and freezing/thawing injury. *J Appl Microbiol* 126, 872–880.

Nepi, M. (2007). Nectary structure and ultrastructure. In *Nectaries and nectar* (Springer), 129–166.

Nepi, M., Grasso, D. A., and Mancuso, S. (2018). Nectar in Plant–Insect Mutualistic Relationships: From Food Reward to Partner Manipulation. *Front Plant Sci* 9, 1–14.

Newell, C. A., Brown, N. J., Liu, Z., Pflug, A., Gowik, U., Westhoff, P., et al. (2010). *Agrobacterium tumefaciens*-mediated transformation of *Cleome gynandra* L., a C4 dicotyledon that is closely related to *Arabidopsis thaliana*. *J Exp Bot* 61, 1311–1319.

Nicolson, S. W., Nepi, M., & Pacini, E. (Eds.). (2007). *Nectaries and nectar* (Vol. 4). Dordrecht: Springer.

Nikolov, L. A. (2019). Brassicaceae flowers: diversity amid uniformity. *J Exp Bot* 70, 2623–2635.

Nisar, N., Li, L., Lu, S., Khin, N. C., and Pogson, B. J. (2015). Carotenoid metabolism in plants. *Mol Plant* 8, 68–82.

Nozzolillo, C., Amiguet, V. T., Bily, A. C., Harris, C. S., Saleem, A., Andersen, Ø. M., & Jordheim, M. (2010). Novel aspects of the flowers and floral pigmentation of two *Cleome* species (Cleomaceae), *C. hassleriana* and *C. serrulata*. *Biochemical Systematics and Ecology 38*, 361-369.

Ohmiya, A., Kishimoto, S., Aida, R., Yoshioka, S., and Sumitomo, K. (2006). *Carotenoid cleavage dioxygenase* (*CmCCD4a*) contributes to white color formation in *Chrysanthemum* petals. *Plant Physiol* 142, 1193–1201.

O'Meara, B. C., Smith, S. D., Armbruster, W. S., Harder, L. D., Hardy, C. R., Hileman, L. C., ... and Diggle, P. K. (2016). Non-equilibrium dynamics and floral trait interactions shape extant angiosperm diversity. *Proceedings of the Royal Society B: Biological Sciences 283*, 20152304.

Orashakova, S., Lange, M., Lange, S., Wege, S., and Becker, A. (2009). The CRABS CLAW ortholog from California poppy (Eschscholzia californica, Papaveraceae), EcCRC, is involved in floral meristem termination, gynoecium differentiation and ovule initiation. *The Plant Journal* 58, 682–693.

Osorio, C. E. (2019). The Role of Orange Gene in Carotenoid Accumulation: Manipulating Chromoplasts Toward a Colored Future. *Front Plant Sci* 10, 1235.

Pabón-Mora, N., Wong, G. K. S., & Ambrose, B. A. (2014). Evolution of fruit development genes in flowering plants. *Frontiers in plant science 5*, 300.

Palmer, C. M., Hindt, M. N., Schmidt, H., Clemens, S., & Guerinot, M. L. (2013). *MYB10* and *MYB72* are required for growth under iron-limiting conditions. *PLoS genetics 9*, e1003953.

Parachnowitsch, A. L., Manson, J. S., and Sletvold, N. (2019). Evolutionary ecology of nectar. *Ann Bot* 123, 247–261.

Patchell, M. J., Bolton, M. C., Mankowski, P., and Hall, J. C. (2011). Comparative floral development in Cleomaceae reveals two distinct pathways leading to monosymmetry. *Int J Plant Sci* 172, 352–365.

Patchell, M. J., Roalson, E. H., & Hall, J. C. (2014). Resolved phylogeny of Cleomaceae based on all three genomes. *Taxon 63*, 315-328.

Pei, Y., Zhang, J., Wu, P., Ye, L., Yang, D., Chen, J., et al. (2021). GoNe encoding a class VIIIb AP2/ERF is required for both extrafloral and floral nectary development in *Gossypium*. *Plant Journal* 106, 1116–1127.

Pelaz, S., Ditta, G. S., Baumann, E., Wisman, E., and Yanofsky, M. F. (2000). B and C floral organ identity functions require *SEPALLATTA* MADS-box genes. *Nature* 405, 200–203.

Pfannebecker, K. C., Lange, M., Rupp, O., & Becker, A. (2017). An evolutionary framework for carpel developmental control genes. *Molecular biology and evolution 34*, 330-348.

Phillips H. R., Landis J. B., Specht C. D. (2020). Revisiting floral fusion: the evolution and molecular basis of a developmental innovation. *J. Exp. Bot* 71, 3390–3404.

Pillitteri, L. J., and Bogenschutz, N. L. (2008). The bHLH Protein , *MUTE*, Controls Differentiation of Stomata and the Hydathode Pore in *Arabidopsis*. 49, 934–943.

Pinyopich, A., Ditta, G. S., Savidge, B., Liljegren, S. J., Baumann, E., Wisman, E., & Yanofsky, M. F. (2003). Assessing the redundancy of MADS-box genes during carpel and ovule development. *Nature 424*, 85-88.

Pittman, J. K. (2012). Multiple transport pathways for mediating intracellular pH homeostasis: The contribution of H+/ion exchangers. *Front Plant Sci* 3, 1–8.

Preston, J. C., and Hileman, L. C. (2009). Developmental genetics of floral symmetry evolution. *Trends Plant Sci* 14, 147–154.

Preston, J. C., Hileman, L. C., and Cubas, P. (2011). Reduce, reuse, and recycle: Developmental evolution of trait diversification. *Am J Bot* 98, 397–403.

Provart, N. J., Alonso, J., Assmann, S. M., Bergmann, D., Brady, S. M., Brkljacic, J., et al. (2016). 50 years of *Arabidopsis* research: Highlights and future directions. *New Phytologist* 209, 921–944.

Pyke, G. H. (2016). Floral Nectar: Pollinator Attraction or Manipulation? *Trends Ecol Evol* 31, 339–341.

R Core Team (2013). R: A language and environment for statistical computing. *R Foundation for statistical Computing, Vienna, Austria*.

Radhika, V., Kost, C., Boland, W., and Heil, M. (2010). The role of jasmonates in floral nectar secretion. *PLoS One* 5, 1–6.

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 Journal* 75, 527–538.

Rajani, S., and Sundaresan, V. (2001). The *Arabidopsis* myc/bHLH gene alcatraz enables cell separation in fruit dehiscence. *Current Biology* 11, 1914–1922.

Raju, A. J. S., and Rani, D. S. (2016). Reproductive ecology of *Cleome gynandra* and *Cleome viscosa* (Capparaceae). *Phytol Balc* 22, 15–28.

Ramegowda, V., Mysore, K. S., and Senthil-Kumar, M. (2014). Virus-induced gene silencing is a versatile tool for unraveling the functional relevance of multiple abiotic-stress-responsive genes in crop plants. *Front Plant Sci* 5, 323.

Ramsay, N. A., and Glover, B. J. (2005). MYB-bHLH-WD40 protein complex and the evolution of cellular diversity. *Trends Plant Sci* 10, 63–70.

Ratcliff, F., Montserrat Martin-Hernandez, A., and Baulcombe, D. C. (2001). Tobacco rattle virus as a vector for analysis of gene function by silencing. *The Plant Journal* 25, 237-245

Rau, A., Gallopin, M., Celeux, G., and Jaffrézic, F. (2013). Data-based filtering for replicated high-throughput transcriptome sequencing experiments. *Bioinformatics* 29, 2146–2152.

Reardon, W., Fitzpatrick, D. A., Fares, M. A., and Nugent, J. M. (2009). Evolution of flower shape in Plantago lanceolata. *Plant Mol Biol* 71, 241–250.

Reeves, P. H., Ellis, C. M., Ploense, S. E., Wu, M. F., Yadav, V., Tholl, D., et al. (2012). A regulatory network for coordinated flower maturation. *PLoS Genet* 8, e1002506.

Reiter, N., Lawrie, A. C., and Linde, C. C. (2018). Matching symbiotic associations of an endangered orchid to habitat to improve conservation outcomes. *Ann Bot* 122, 947–959.

Ren, R., Wang, H., Guo, C., Zhang, N., Zeng, L., Chen, Y., et al. (2018). Widespread Whole Genome Duplications Contribute to Genome Complexity and Species Diversity in Angiosperms. *Mol Plant* 11, 414–428.

Ren G., Healy R. A., Horner H. T., James M. G., And Thornburg R. W. (2007). Expression of starch metabolic genes in the developing nectaries of ornamental tobacco plants. *Plant Sci* 173, 621–637.

Rering, C. C., Beck, J. J., Hall, G. W., McCartney, M. M., and Vannette, R. L. (2018). Nectar-inhabiting microorganisms influence nectar volatile composition and attractiveness to a generalist pollinator. *New Phytologist* 220, 750–759.

Reyes, E., Sauquet, H., and Nadot, S. (2016). *Perianth symmetry changed at least 199 times in angiosperm evolution. Taxon 65*, 945-964.

Robinson, M., McCarthy, D., and Smyth, G. (2009). edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* 26, 139–140.

Robles, P., and Pelaz, S. (2005). Flower and fruit development in *Arabidopsis thaliana*. *International Journal of Developmental Biology* 49, 633–643.

Roeder, A., Ferrándiz, C., and MF, Y. (1998). The Role of the *REPLUMLESS* Homeodomain Protein in Patterning the *Arabidopsis* Fruit. *Cell Press* 13, 1630–1635.

Roeder, A., and Yanofsky, M. (2006). Fruit Development in *Arabidopsis*. In *The Arabidopsis Book*, 1–50.

Romera-Branchat, M., Ripoll, J. J., Yanofsky, M. F., & Pelaz, S. (2013). The WOX 13 homeobox gene promotes replum formation in the *Arabidopsis thaliana* fruit. *The Plant Journal 73*, 37-49.

Rosin, F. M., and Kramer, E. M. (2009). Old dogs, new tricks: Regulatory evolution in conserved genetic modules leads to novel morphologies in plants. *Dev Biol* 332, 25–35.

Roy, R., Schmitt, A. J., Thomas, J. B., and Carter, C. J. (2017). Review: Nectar biology: From molecules to ecosystems. *Plant Science* 262, 148–164.

Ruhlmann, J. M., Kram, B. W., and Carter, C. J. (2010). *Cell wall invertase 4* is required for nectar production in *Arabidopsis*. *J Exp Bot* 61, 395–404.

Ruiz, M. T., Voinnet, O., & Baulcombe, D. C. (1998). Initiation and maintenance of virus-induced gene silencing. *The Plant Cell*, *10*, 937-946.

Ryu, C. M., Anand, A., Kang, L., and Mysore, K. S. (2004). Agrodrench: A novel and effective agroinoculation method for virus-induced gene silencing in roots and diverse Solanaceous species. *Plant Journal* 40, 322–331.

Sagawa, J. M., Stanley, L. E., Lafountain, A. M., Frank, H. A., Liu, C., and Yuan, Y. W. (2016). An R2R3-MYB transcription factor regulates carotenoid pigmentation in *Mimulus lewisii* flowers. *New Phytologist* 209, 1049–1057.

Salazar-Duque, H., Alzate, J. F., Urrea Trujillo, A., Ferrándiz, C., and Pabón-Mora, N. (2021). Comparative anatomy and genetic bases of fruit development in selected Rubiaceae (Gentianales). *Am J Bot* 108, 1838–1860.

Sargent, R. D. (2004). Floral symmetry affects speciation rates in angiosperms. *Proceedings of the Royal Society B: Biological Sciences* 271, 603–608.

Sauquet, H., Von Balthazar, M., Magallón, S., Doyle, J. A., Endress, P. K., Bailes, E. J., ... and Schönenberger, J. (2017). The ancestral flower of angiosperms and its early diversification. *Nature communications 8*, 1-10.

Schiessl, K., Muiño, J. M., & Sablowski, R. (2014). *Arabidopsis JAGGED* links floral organ patterning to tissue growth by repressing Kip-related cell cycle inhibitors. *Proceedings of the National Academy of Sciences 111*, 2830-2835.

Schrager-Lavelle, A., Klein, H., Fisher, A., and Bartlett, M. (2017). Grass flowers: An untapped resource for floral evo-devo. *J Syst Evol* 55, 525–541.

Schranz, M. E., Edger, P. P., Pires, J. C., van Dam, N. M., and Wheat, C. W. (2011). Comparative genomics in the Brassicales: ancient genome duplications, glucosinolate diversification and pierinae herbivore radiation. *Genetics, genomics and breeding of oilseed brassicas 206*, 206-217.

Schranz, M. E., and Mitchell-Olds, T. (2006). Independent ancient polyploidy events in the sister families Brassicaceae and Cleomaceae. *Plant Cell* 18, 1152–1165.

Schröder F., Lisso J., Lange P., Müssig C. (2009). The extracellular *EXO* protein mediates cell expansion in arabidopsis leaves. *BMC Plant Biol* 9, 1–12.

Schuster, C., Gaillochet, C., & Lohmann, J. U. (2015). *Arabidopsis* HECATE genes function in phytohormone control during gynoecium development. *Development 142*, 3343-3350.

Scofield S., Dewitte W., And Murray J. A. H. (2007). The KNOX gene *SHOOT MERISTEMLESS* is required for the development of reproductive meristematic tissues in *Arabidopsis*. *Plant J* 50, 767–781.

Simao F. A., Waterhouse R. M., Ioannidis P., Kriventseva E. V., Zdobnov E. M. (2015). BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics* 31, 3210–3212.

Singh S., Bhatt V., Kumar V., Kumawat S., Khatri P., Singla P., ... and Sonah, H. (2020). Evolutionary understanding of aquaporin transport system in the basal eudicot model species aquilegia coerulea . Plants 9, 799.

Singh V., Roy S., Giri M. K., Chaturvedi R., Chowdhury Z., Shah J., Nandi, A. K. (2013). Arabidopsis thaliana *FLOWERING LOCUS d* is required for systemic acquired resistance. *Mol. Plant-Microbe Interact* 26, 1079–1088.

Semiarti, E., Ueno, Y., Tsukaya, H., Iwakawa, H., Machida, C., and Machida, Y. (2001). The *ASYMMETRIC LEAVES2* gene of *Arabidopsis thaliana* regulates formation of a symmetric

lamina, establishment of venation and repression of meristem-related homeobox genes in leaves. *Development* 128, 1771–1783.

Sengupta, A., and Hileman, L. C. (2018). Novel traits, flower symmetry, and transcriptional autoregulation: New hypotheses from bioinformatic and experimental data. *Front Plant Sci* 871, 1–15.

Sengupta, A., and Hileman, L. C. (2022). A *CYC–RAD–DIV–DRIF* interaction likely pre-dates the origin of floral monosymmetry in Lamiales. *Evodevo* 13, 1–19.

Senthil-Kumar, M., Hema, R., Anand, A., Kang, L., Udayakumar, M., and Mysore, K. S. (2007). A systematic study to determine the extent of gene silencing in *Nicotiana benthamiana* and other Solanaceae species when heterologous gene sequences are used for virus-induced gene silencing. *New Phytologist* 176, 782–791.

Senthil-Kumar, M., and Mysore, K. S. (2014). Tobacco rattle virus-based virus-induced gene silencing in Nicotiana benthamiana. *Nat Protoc* 9, 1549–1562.

Shah, M., Alharby, H. F., Hakeem, K. R., Ali, N., Rahman, I. U., Munawar, M., et al. (2020). De novo transcriptome analysis of Lantana camara L. revealed candidate genes involved in phenylpropanoid biosynthesis pathway. *Sci Rep* 10, 1–14.

Shan, X., Zhang, Y., Peng, W., Wang, Z., and Xie, D. (2009). Molecular mechanism for jasmonate-induction of anthocyanin accumulation in *Arabidopsis*. *J Exp Bot* 60, 3849–3860.

Shang, Y., Venail, J., Mackay, S., Bailey, P. C., Schwinn, K. E., Jameson, P. E., ... and Davies, K. M. (2011). The molecular basis for venation patterning of pigmentation and its effect on pollinator attraction in flowers of *Antirrhinum*. *New Phytologist 189*, 602-615.

Shapiro, E., Biezuner, T., and Linnarsson, S. (2013). Single-cell sequencing-based technologies will revolutionize whole-organism science. *Nat Rev Genet* 14, 618–630.

Sheikh-Assadi, M., Naderi, R., Salami, S. A., Kafi, M., Fatahi, R., Shariati, V., ... and Claros, M. G. (2022). Normalized Workflow to Optimize Hybrid De Novo Transcriptome Assembly for Non-Model Species: A Case Study in Lilium ledebourii (Baker) Boiss. *Plants 11*, 2365.

Simão, F. A., Waterhouse, R. M., Ioannidis, P., Kriventseva, E. v., and Zdobnov, E. M. (2015). BUSCO: Assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics* 31, 3210–3212.

Simonini, S., Deb, J., Moubayidin, L., Stephenson, P., Valluru, M., Freire-Rios, A., ... and Østergaard, L. (2016). A noncanonical auxin-sensing mechanism is required for organ morphogenesis in *Arabidopsis*. *Genes & development 30*, 2286-2296.

Sinha, S., Raxwal, V. K., Joshi, B., Jagannath, A., Katiyar-Agarwal, S., Goel, S., et al. (2015). De novo transcriptome profiling of cold-stressed siliques during pod filling stages in Indian mustard (*Brassica juncea* L.). *Front Plant Sci* 6, 932.

Slavković, F., Dogimont, C., Morin, H., Boualem, A., and Bendahmane, A. (2021). The Genetic Control of Nectary Development. *Trends Plant Sci* 26, 260–271.

Sobel, J. M., and Streisfeld, M. A. (2013). Flower color as a model system for studies of plant evo-devo. *Front Plant Sci* 4, 1–17.

Sogbohossou, E. D., Achigan-Dako, E. G., Mumm, R., de Vos, R. C., & Schranz, M. E. (2020). Natural variation in specialised metabolites production in the leafy vegetable spider plant (*Gynandropsis gynandra* L.(Briq.)) in Africa and Asia. *Phytochemistry 178*, 112468

Solhaug E. M., Roy R., Chatt E. C., Klinkenberg P. M., Mohd-Fadzil N., Hampton M., ... and Carter, C. J. (2019). An integrated transcriptomics and metabolomics analysis of the Cucurbita pepo nectary implicates key modules of primary metabolism involved in nectar synthesis and secretion. *Plant Direct* 3, e00120.

Souto-Vilarós, D., Vuleta, A., Jovanović, S. M., Budečević, S., Wang, H., Sapir, Y., et al. (2018). Are pollinators the agents of selection on flower colour and size in irises? *Oikos* 127, 834–846.

Specht, C. D., and Howarth, D. G. (2015). Adaptation in flower form: A comparative evodevo approach. *New Phytologist* 206, 74–90.

Spence, J., Vercher, Y., Gates, P., & Harris, N. (1996). 'Pod shatter'in *Arabidopsis thaliana Brassica napus* and B. juncea. *Journal of Microscopy 181*, 195-203.

Stein, O., and Granot, D. (2019). An overview of sucrose synthases in plants. *Front Plant Sci* 10, 1–14.

Strable, J., and Scanlon, M. J. (2009). Maize (Zea mays): A model organism for basic and applied research in plant biology. *Cold Spring Harb Protoc* .

Streisfeld, M. A., and Rausher, M. D. (2011). Population genetics, pleiotropy, and the preferential fixation of mutations during adaptive evolution. *Evolution* 65, 629–642.

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 Phytologist* 215, 1582–1593.

Sung, Y. C., Lin, C. P., and Chen, J. C. (2014). Optimization of virus-induced gene silencing in Catharanthus roseus. *Plant Pathol* 63, 1159–1167.

Switzenberg, J. A., Beaudry, R. M., and Grumet, R. (2015). Effect of *CRC*::etr1-1 transgene expression on ethylene production, sex expression, fruit set and fruit ripening in transgenic melon (*Cucumis melo* L.). *Transgenic Res* 24, 497–507.

Tanaka, Y., Sasaki, N., and Ohmiya, A. (2008). Biosynthesis of plant pigments: Anthocyanins, betalains and carotenoids. *Plant Journal* 54, 733–749.

Tang, X., Gomes, A. M., Bhatia, A., & Woodson, W. R. (1994). Pistil-specific and ethylene-regulated expression of 1-aminocyclopropane-1-carboxylate oxidase genes in *petunia* flowers. *The Plant Cell*, *6*, 1227-1239.

Tang, Y. (2014). RNA-seq analysis of virus-induced gene silenced (VIGS) lines in foxtail millet. *USDOE Joint Genome Institute*.

Thomas, C. L., Jones, L., Baulcombe, D. C., and Maule, A. J. (2001). Size constraints for targeting post-transcriptional gene silencing and for RNA-directed methylation in *Nicotiana benthamiana* using a potato virus X vector. *Plant Journal* 25, 417–425.

Tiffney, B. H. (1984). Seed Size, Dispersal Syndromes , and the Rise of the Angiosperms: Evidence and Hypothesis Author. *Annals of the Missouri Botanical Garden* 71, 551–576.

Torti S., Fornara F., Vincent C., Andrés F., Nordström K., Göbel U., … and Coupland, G. (2012). Analysis of the *Arabidopsis* shoot meristem transcriptome during floral transition identifies distinct regulatory patterns and a leucine-rich repeat protein that promotes flowering. *Plant Cell* 24, 444–462.

Tsai, Y. T., Chen, P. Y., and To, K. Y. (2012). Plant regeneration and stable transformation in the floricultural plant Cleome spinosa, a C 3 plant closely related to the C 4 plant C. gynandra. *Plant Cell Rep* 31, 1189–1198.

Tucker, G. C., and Vanderpool, S. S. (2010). Cleomaceae. In Flora of north Q30 America north of Mexico: Magnoliophyta: Salicaceae to brassicaceae (New York:Oxford University Press).

van den Bergh, E., Külahoglu, C., Bräutigam, A., Hibberd, J. M., Weber, A. P., Zhu, X. G., & Schranz, M. E. (2014). Gene and genome duplications and the origin of C4 photosynthesis: birth of a trait in the Cleomaceae. *Current Plant Biology 1*, 2-9.

van der Kooi, C. J., Vallejo-Marín, M., and Leonhardt, S. D. (2021). Mutualisms and (A)symmetry in Plant–Pollinator Interactions. *Current Biology* 31, 91–99.

Vannette, R. L., and Fukami, T. (2016). Nectar microbes can reduce secondary metabolites in nectar and alter effects on nectar consumption by pollinators. *Ecology* 97, 1410–1419.

Veeckman, E., Ruttink, T., and Vandepoele, K. (2016). Are we there yet? Reliably estimating the completeness of plant genome sequences. *Plant Cell* 28, 1759–1768.

Velasquez, A., Chakravarthy, S., and Martin, G. B. (2009). Virus-induced Gene Silencing (VIGS) in *Nicotiana benthamiana* and Tomato. *Journal of Visualized Experiments*, e1292.

Voznesenskaya, E. V., Koteyeva, N. K., Chuong, S. D., Ivanova, A. N., Barroca, J., Craven, L. A., and Edwards, G. E. (2007). Physiological, anatomical and biochemical characterisation of photosynthetic types in genus Cleome (Cleomaceae). *Functional Plant Biology 34*, 247-267.

Wang, C., Cai, X., Wang, X., and Zheng, Z. (2006). Optimization of tobacco rattle virus-induced gene silencing in *Arabidopsis*. *Functional Plant Biology* 33, 347–355.

Wang, L., Lu, W., Ran, L., Dou, L., Yao, S., Hu, J., et al. (2019). R2R3-MYB transcription factor MYB6 promotes anthocyanin and proanthocyanidin biosynthesis but inhibits secondary cell wall formation in *Populus tomentosa*. *Plant Journal* 99, 733–751.

Wang, M., Wang, G., Ji, J., and Wang, J. (2009). The effect of *pds* gene silencing on chloroplast pigment composition, thylakoid membrane structure and photosynthesis efficiency in tobacco plants. *Plant Science* 177, 222–226.

Wang, Y., Chen, S., and Yu, O. (2011). Metabolic engineering of flavonoids in plants and microorganisms. *Appl Microbiol Biotechnol* 91, 949–956.

Wang, Y., Coleman-Derr, D., Chen, G., and Gu, Y. Q. (2015). OrthoVenn: A web server for genome wide comparison and annotation of orthologous clusters across multiple species. *Nucleic Acids Res* 43, 78–84.

Warwick, S. I., and Sauder, C. A. (2005). Phylogeny of tribe Brassiceae (Brassicaceae) based on chloroplast restriction site polymorphisms and nuclear ribosomal internal transcribed spacer and chloroplast *trn* L intron sequences. *Canadian Journal of Botany* 83, 467–483.

Wege, S., Scholz, A., Gleissberg, S., and Becker, A. (2007). Highly efficient virus-induced gene silencing (VIGS) in California poppy (Eschscholzia californica): An evaluation of VIGS as a strategy to obtain functional data from non-model plants. *Ann Bot* 100, 641–649.

Weigel, D., and Glazebrook, J. (2002). *Arabidopsis: A Laboratory Manual*. Cold Spring Harbor, N.Y: Cold Spring Harbor Laboratory Press.

Weigel, D., & Meyerowitz, E. M. (1994). The ABCs of floral homeotic genes. *Cell 78*, 203-209.

Wessinger, C. A., and Hileman, L. C. (2020). Parallelism in Flower Evolution and Development. *Annu Rev Ecol Evol Syst* 51, 387–408.

Wickham, H., & Wickham, H. (2009). Getting started with qplot. ggplot2: *elegant graphics for data analysis,* 9-26.

Wiesen, L. B., Bender, R. L., Paradis, T., Larson, A., Perera, M. A. D. N., Nikolau, B. J.,… and Carter, C. J. (2016). A Role for *GIBBERELLIN 2-OXIDASE6* and Gibberellins in Regulating Nectar Production. *Mol Plant* 9, 753–756.

Williams, B. P., Burgess, S. J., Reyna-Llorens, I., Knerova, J., Aubry, S., Stanley, S., and Hibberd, J. M. (2015). An untranslated cis-element regulates the accumulation of multiple C4 enzymes in *Gynandropsis gynandra* mesophyll cells. *Plant Cell* 28, 454–465.

Willis, C. G., Hall, J. C., Rubio De Casas, R., Wang, T. Y., and Donohue, K. (2014). Diversification and the evolution of dispersal ability in the tribe Brassiceae (Brassicaceae). *Ann Bot* 114, 1675–1686.

Wilson, R. C., and Doudna, J. A. (2013). Molecular mechanisms of RNA interference. *Annu Rev Biophys* 42, 217–239.

Wollmann, H., Mica, E., Todesco, M., Long, J. A., and Weigel, D. (2010). On reconciling the interactions between *APETALA2*, miR172 and *AGAMOUS* with the ABC model of flower development. *Development* 137, 3633–3642.

Wu, Y., Wen, J., Xia, Y., Zhang, L., and Du, H. (2022). Evolution and functional diversification of R2R3-MYB transcription factors in plants. *Hortic Res* 9.

Xi, W., Feng, J., Liu, Y., Zhang, S., and Zhao, G. (2019). The R2R3-MYB transcription factor *PaMYB10* is involved in anthocyanin biosynthesis in apricots and determines red blushed skin. *BMC Plant Biol* 19, 1–14.

Xu, L., Palatnik, J., Springer, P., Mao, L., Hepworth, S. R., Ca, S., et al. (2015). Beyond the Divide: Boundaries for Patterning and Stem Cell Regulation in Plants. *Front. Plant Sci* 6, 1052.

Xu, P., Wu, L., Cao, M., Ma, C., Xiao, K., and Li, Y. (2021). Identification of MBW Complex Components Implicated in the Biosynthesis of Flavonoids in Woodland Strawberry *Frontiers in Plant Science 12*, 774943.

Yabuzaki, J. (2017). Carotenoids Database: Structures, chemical fingerprints and distribution among organisms. *Database* 1–11.

Yamamizo, C., Kishimoto, S., and Ohmiya, A. (2010). Carotenoid composition and carotenogenic gene expression during *Ipomoea* petal development. *J Exp Bot* 61, 709–719.

Yan, J., Wang, G., Sui, Y., Wang, M., and Zhang, L. (2016). Pollinator responses to floral colour change, nectar, and scent promote reproductive fitness in Quisqualis indica (Combretaceae). *Sci Rep* 6, 1–10.

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. *Plant Cell* 24, 1834–1847.

Yonekura-Sakakibara, K., Higashi, Y., and Nakabayashi, R. (2019). The Origin and Evolution of Plant Flavonoid Metabolism. *Front Plant Sci* 10, 1–16.

Yu, H., and Goh, C. J. (2001). Molecular genetics of reproductive biology in orchids. *Plant Physiol* 127, 1390–1393.

Zhang, B., Liu, C., Wang, Y., Yao, X., Wang, F., Wu, J., ... & Liu, K. (2015). Disruption of a *CAROTENOID CLEAVAGE DIOXYGENASE 4* gene converts flower colour from white to yellow in *Brassica* species. *New Phytologist 206*, 1513-1526.

Zhang, R., Min, Y., Holappa, L. D., Walcher-Chevillet, C. L., Duan, X., Donaldson, E., et al. (2020). A role for the Auxin Response Factors ARF6 and ARF8 homologs in petal spur elongation and nectary maturation in Aquilegia. *New Phytologist* 227, 1392–1405.

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. *Proceedings of the National Academy of Sciences* 107, 6388–6393.

Zhang, X., Henriques, R., Lin, S. S., Niu, Q. W., and Chua, N. H. (2006). Agrobacterium-mediated transformation of *Arabidopsis thaliana* using the floral dip method. *Nat Protoc* 1, 641–646.

Zhang, Y., Shen, Y. Y., Wu, X. M., and Wang, J. B. (2016). The basis of pod dehiscence: anatomical traits of the dehiscence zone and expression of eight pod shatter-related genes in four species of Brassicaceae. *Biol Plant* 60, 343–354.

Zhao, Y., Wang, K., Wang, W. L., Yin, T. T., Dong, W. Q., and Xu, C. J. (2019). A high-throughput SNP discovery strategy for RNA-seq data. *BMC Genomics* 20, 1–10.

Zhong, C., Tang, Y., Pang, B., Li, X., Yang, Y., Deng, J., et al. (2020). The R2R3-MYB transcription factor GhMYB1a regulates flavonol and anthocyanin accumulation in *Gerbera hybrida*. *Hortic Res* 7.

Zhong, J., and Kellogg, E. A. (2015). Duplication and expression of CYC2-like genes in the origin and maintenance of corolla zygomorphy in Lamiales. *New Phytologist* 205, 852–868.

Zhu, C., Bai, C., Sanahuja, G., Yuan, D., Farré, G., Naqvi, S., et al. (2010). The regulation of carotenoid pigmentation in flowers. *Arch Biochem Biophys* 504, 132–141.

Zhu, H. F., Fitzsimmons, K., Khandelwal, A., and Kranz, R. G. (2009). CPC, a single-repeat R3 MYB, is a negative regulator of anthocyanin biosynthesis in *Arabidopsis*. *Mol Plant* 2, 790–802.

Ziegler-Graff, V., Guilford, P. J., and Baulcombe, D. C. (1991). Tobacco rattle virus RNA-1 29K gene product potentiates viral movement and also affects symptom induction in tobacco. *Virology* 182, 145–155.

Zumajo-Cardona, C., and Pabón-Mora, N. (2016). Evolution of the APETALA2 Gene Lineage in Seed Plants. *Mol Biol Evol* 33, 1818-1832