Pollination biology and transcriptomics of Cleomaceae species exhibiting different pollination systems

by

Monica Paola Higuera

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

Doctor in Philosophy

in

Systematics & Evolution

Department of Biological Sciences

University of Alberta

Abstract

Pollinator mediated-selection is responsible for much of the phenotypic diversity exhibited by flowers. Flowering plants that attract a diverse range of pollinators represent a generalist pollination system in which selection is exerted by many different pollinators. However, generalist pollination systems are difficult to study. They are frequently overlooked, because directional selection may be limited and diffuse selection exerted by pollinators in floral traits is hard to test. Thus, specialist flowering plants attract specific pollinator guilds, which mediates directional or stabilizing selection on floral traits. Selection by a guild of pollinators, such as bees, birds or bats, has led to the convergence of floral phenotypes, known as pollination syndromes. Flowering plants associated to pollination syndromes belong to a specialist pollination system. While, flowering plants that attract a diverse range of pollinators represent a generalist pollination system in which selection is exerted by many different pollinators. However, generalist pollination systems are difficult to study. They are frequently overlooked, because directional selection may be limited and diffuse selection exerted by pollinators in floral traits is hard to asses. One way to understand the evolution of floral traits and pollination systems involves exploring the genes encoding floral traits subject to pollinator-mediated selection. One good family for exploration of these topics is the Cleomaceae, which exhibits variation in floral traits, has species displaying generalist and specialist pollination systems and is an emerging model system for studying floral evolution. In this thesis, I used natural history of Cleomaceae species and molecular methods to compare the pollination biology and genes encoding floral traits from plants exhibiting generalized or specialized floral traits. I observed plants with floral traits that could be subject to pollinator-mediated selection such as color, and nectar glands,

and then I used transcriptomics to identify candidate genes that may be responsible for differences in these floral traits.

My thesis begins by proposing a comprehensive framework that incorporates both ecological and genetic pollination studies, which I have defined as integrative pollination studies. I reviewed genetic techniques used to explore genes underlying floral traits and discuss the use of next generation sequencing to study pollination. Second, I studied the pollination biology of two Cleomaceae species present in Alberta, which differ in main floral traits: *Cleomella serrulata* (Pursh) E.H. Roalson and J.C. Hall and *Polanisia dodecandra* (L.) DC to determine their pollination systems. I determined that both species have a generalist pollination system, but there were differences in the richness and composition of pollinators between them, due to main differences in floral traits and population locations. To examine differences in the floral evolution of species exhibiting generalized and specialized floral traits, I identified putative candidate genes responsible for pollinator attraction, then I studied the floral transcriptome of these two generalist species, along with the specialist *Melidiscus giganteus* (L.) Raf. Despite phylogenetic distance, the generalist species *C. serrulata* and *P. dodecandra* shared more transcripts when compare with *M. giganteus.* I then examined putative candidate genes for key floral traits and found differences in genes coding for color and nectary traits between generalist and specialist species that could be subjected to selection by pollinators. However, the genes coding for floral traits in specialist and generalist species were similar overall. Finally, I aimed to down-regulate candidate genes important for pollinator attraction to obtain plants with modified phenotypes that could be used in pollinators preference tests. I tested the feasibility of the virus-inducing gene silencing (VIGS) technique in *P. dodecandra*. Unfortunately, this species was not amenable to this technique, but it was

feasible in *Cleome violacea* L., which may be a good candidate species for further pollination studies.

Overall, my results indicate that species exhibiting generalized and specialized floral traits were similar in terms of gene repertoires despite being pollinated by different pollinators and despite phylogenetic distance. The generalist species were pollinated by different pollinator assemblages although the populations were located nearby. Though the floral transcriptomes of the generalist species were more similar to each other than to that of the specialist, generalists and specialist were alike in terms of gene repertoires. Taken together, this thesis advances our understanding of the pollination biology and genetics of the floral traits in Cleomaceae. I recommend that Cleomaceae be used as a comprehensive model to study the transitions between generalist and specialist pollination systems, and to study the evolution of floral traits important for pollinator attraction.

Preface

Chapter 3 of this thesis has been published as "Higuera-Díaz M, Manson JS, Hall JC. 2015. Pollination biology of *Cleomella serrulata* and *Polanisia dodecandra* in a protected natural prairie in southern Alberta, Canada", at *Botany* **93**: 745-757. I collected the data, analyzed and wrote the manuscript. JS Manson and JC Hall helped with experimental design, manuscript editing, and revisions.

Dedication

To my husband who always hold my hand, my beautiful daughter Suanti and my family

Acknowledgements

I would like to thank my supervisors Dr. Jocelyn C. Hall and Dr. Jessamyn Manson for their guidance and their support during my PhD studies, but also, I am very grateful for the support that they provided me, when I endured the most sad and difficult moments in my life. Without your help and support, I would not have been able to continue my doctoral studies. You were the best supervisors for me, many thanks for your mentoring, support and patience during the last six years. I also want to thank to Dr. James Cahill for being in my committee, many thanks for your guidance and support on my project, also many thanks for made me think broadly and for challenge me, and of course, I would like to thank you for introducing me to your students, and for adopting me in your lab meetings and celebrations, it was a great experience for me. I would like to thank Dr. Mike Deyholos for being in my committee for 2 years and for his support on my project.

I want to thank Dr. Felix Sperling, Janet Sperling, José Díaz and Lucy Díaz for their support during my arrival to Edmonton, many thanks for your hospitality and for making me feel at home. I am also very grateful to my lab mates and also good friends Kelsey Bernard, Melanie Patchell, Shane Carey, Peter Mankowski, Mariano Avino, Ashton Sturm, Angela Phung, Monica Khöler, and Michael Foisy. Many thanks for your support and your friendship during this process. I also want to thank my very good friends who were my support in happy and sad moments: Gisela Stötz, Gregory Peck, Tan Bao, Jared Bernard, William Torres, Orión Reyes, Isabel Barrios, Guillermo Bueno, Alec Carrigy, Gian Carlo Gianoli, Fred Noddin, Caroline Whitehouse, Sandra Durán, Charlotte Brown, Margarete Dettlaff, Megan Ljubotina, Leonardo Galindo, Angélica Quesada, Javier Luque, Kecia Kerr, Claudia Ayala, Jaime Pinzón, Jenny Fonseca, and Nicolas Abarca. Many thanks for all the happy moments and for taking care of me and my daughter when she was born.

My research would not have been possible without the people that help me in the field, many thanks to Jared Bernard Jenny Kleininger for their hard work during the field seasons. Thanks to Drew Taylor, range biologist at Suffield CFB, for his logistical help during my project. Many thanks to Cory Sheffield, Ashton Sturm, Felix Sperling, Gary Anweiler, Matthias Buck and Tyler Cobb for help with insect identification. Many thanks to Dorothy Fabijan for assistance with plant voucher preparation. I would also like to thank

Corey Davis and Troy Locke for their advice during my project, and many, many thanks to Kerrin Mendler, Leonardo Galindo and Dr. Mike Deyholos for their bioinformatics help.

Last but not least, I want to thank my life partner Giovanny Fagua for all his support, love and care, my cute baby Suanti for her smile and happiness that encourage me every day, to my family Stella, Francisco, Andrea, Gustavo, Vero, and Diego for their encouragement and support. Finally, many, many thanks to my mom and my mother in law for taking care of my baby, while, I was finishing my thesis.

My PhD studies had the financial support of the Administrative Department of Science, Technology and Innovation of Colombia (COLCIENCIAS, "Generación del Bicentenario" scholarship). My research was funded for a ACA Grant in Biodiversity (supported by the Alberta Conservation Association), and a NSERC Discovery Grant held by Dr. Jocelyn C. Hall.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Chapter 1: Introduction

Pollination is a process of vital biological importance for plant reproduction, in which pollen is transferred to the stigma to facilitate fertilization (Proctor *et al.* 1996; Willmer 2011). This transfer can be performed by abiotic and biotic vectors. Abiotic vectors include water and wind, while biotic vectors include, birds, mammals, reptiles and insects. Most angiosperms have biotic vectors that transfer pollen, which are known as pollinators, and insects are the most diverse group of pollinators. Plants exhibit a huge variation in floral traits associated with pollinator attraction, which promotes great diversity in flower size, shape, colors and glands (Herrera and Pellmyr 2002b; Proctor *et al.* 1996; Willmer 2011). The astonishing diversity in floral traits is the result of selection exerted by pollinators (Proctor *et al.* 1996; Willmer 2011), and pollinator-mediated selection on floral traits has been frequently studied (Baucom *et al.* 2011; Bolstad *et al.* 2010; Campbell *et al.* 1996; Cronk and Ojeda 2008; Galen 1999; Johnston 1991; Medel *et al.* 2003; Sandring and Agren 2009; Sobral *et al.* 2015; Willmer 2011).

Pollinator-mediated selection in floral traits can cause strong directional selection and/or stabilizing selection on specific traits, which could promote floral trait diversity (Fenster *et al.* 2004; Schemske and Bradshaw 1999; Sletvold *et al.* 2016; Zhao and Wang 2015). In order to study evolution and pollinator-mediated selection on floral traits, it is crucial to understand which genes encode floral traits, because genes and their underlying variation play a major role in the floral phenotypes exhibited by plants (Sapir 2009b; Sedeek *et al.* 2013). Additionally, changes in pollinator behavior can in turn lead to variation in these floral traits (Devaux *et al.* 2014; Fenster *et al.* 2004; Kay and Sargent 2009; Sapir 2009b). There are two major types of pollination systems: generalist systems, in which flowers exhibit simple floral traits, the rewards are exposed to pollinators, and they are pollinated by a wide taxonomic range of pollinators, but not all the generalist species are pollinated by the same kind of pollinators and variation exist (Herrera 2005; Herrera and Pellmyr 2002b; Proctor *et al.* 1996; Willmer 2011), and specialist systems, in which flowers exhibit complex morphologies, the rewards are frequently concealed, and they are pollinated by a specific pollinator guild (Herrera and Pellmyr 2002b; Proctor *et al.* 1996; Willmer 2011). Flowering plant species belonged to a specialist pollination system attract specific pollinator guilds, which mediates strong selection on floral traits. Thus,

selection by a guild of pollinators, such as bees, birds or bats, has led to the convergence of floral phenotypes, known as pollination syndromes, in which traits as morphology, color, size, scent and rewards are correlated with pollinator traits (Proctor *et al.* 1996; Rosas‐Guerrero *et al.* 2014; Willmer 2011). Most of the research exploring the genetic basis of pollination systems have been done in species exhibiting specialist pollination systems, specifically species exhibiting different pollination syndromes, because these species have distinct traits that correlate with specific pollinators and the traits therefore are likely tied to pollinator-mediated selection, and in some cases, the genes involved in this process have been identified (Cronk and Ojeda 2008; Epperson and Clegg 1992; Gübitz *et al.* 2009; Hoballah *et al.* 2007; Klahre *et al.* 2011; Schemske and Bradshaw 1999; Verdonk *et al.* 2005; Xu *et al.* 2012). Although, in nature, generalist pollination systems are more frequent than the specialist ones, the genetic basis of generalist pollination systems is frequently underexplored, because the traits are under diffuse selection exerted by different pollinators (Kessler and Baldwin 2011; Sapir 2009b).

In my thesis research, I used Cleomaceae, a group that exhibits both types of pollination systems as a model to explore the pollination biology of two species exhibiting generalized floral traits and to elucidate the genetic variation underlying generalized and specialized floral traits. This allowed me to identify putative candidate genes encoding major floral traits that are likely subject to diffuse pollinator-mediated selection in generalists and directional pollinator-mediated selection in specialist pollination systems. My research advances the pollination biology knowledge of species that have generalized traits, and in the understanding of genetic differences and similarities between Cleomaceae species exhibiting generalized and specialized floral traits. It is also one of the few studies that has used a huge amount of data derived from Next Generation Sequencing (NGS) to compare gene repertoires in both pollination systems.

1.1. Genetic studies of pollination systems

Pollination studies that involve ecology and genetics have great potential to understand complex interactions such as pollination systems and their evolution (Clare *et al.* 2013; Kay and Sargent 2009; Kessler and Baldwin 2011; Kessler *et al.* 2008; Mayer *et al.* 2011; Sapir 2009b). The numbers of studies involving the genetic basis of pollination

systems have increased in recent years (Clare *et al.* 2013). Both forward and reverse genetic approaches (reviewed in Chapter 2), have been used to better understand the genetic control and regulation of major floral traits, pollinator shifts between closely related species, and floral traits under pollinator-mediated selection (Kay and Sargent 2009; Kessler and Baldwin 2011; Kessler *et al.* 2012; Owen and Bradshaw 2011; Sapir 2009a, b; Sapir and Armbruster 2010; Sedeek *et al.* 2013; Sharma and Kramer 2013; Xu *et al.* 2012; Yuan *et al.* 2013*a*). These approaches and their potential applications for addressing the evolution of pollination systems and the putative genes encoding major floral traits for pollinators attraction are detailed in Chapter 2.

To study the evolution of plant-pollinator interactions and the putative genes encoding floral traits, next generation sequencing (NGS) is a powerful technique that generates vast amounts of genomics and transcriptomics data in model and non-model species. Whole genome sequencing is often conducted in model organisms and provides valuable information about insertions/deletions, mutations, single nucleotide variants, and differences in genes copy number (Clare *et al.* 2013; Hawkins *et al.* 2010). Alternatively, RNA sequencing is frequently used in model and non-model organisms and allows the identification of all transcripts present in the sample. Transcriptome data sets provide information on nucleotide variation, identification of rare transcripts, and transcript expression levels across different tissues types, time periods, and across species. In addition, we can perform transcriptome comparisons among species and in natural populations. Transcriptomics have been used to study which floral traits and genes are involved for pollinator attraction in *Ophrys*, a sexually deceptive orchid representing a specialist pollination system (Sedeek *et al.* 2013). Moreover, transcriptomics have been used to find candidate self-incompatibility genes in self and cross pollinated flowers of *Erigeron breviscapinus* (Vaniot) Hand.-Mazz.(Zhang *et al.* 2015b). NGS techniques and their application in pollination research are reviewed in Chapter 2 and these techniques are applied in Chapter 4. In Chapter 5, two of the methods proposed Chapter 2, virus-inducing gene silencing (VIGS) and chemical mutagenesis, are explored as possible tools for modifying floral traits in Cleomaceae species.

1.2. Pollination biology studies in Cleomaceae

Cleomaceae belongs to the order Brassicales and is the sister family to Brassicaceae, which includes many economically important crops and the model plant for evolutionary studies *Arabidopsis thaliana* (L) Heynh. Cleomaceae is a relatively small plant family that is comprised of approximately 270 spp., it has a cosmopolitan distribution, though the major diversity is found in warm temperate and tropical regions (Patchell *et al.* 2014; van den Bergh *et al.* 2016a). The family is compelling for studying pollination biology because its flowers exhibit variation in morphological traits traditionally associated with pollinator attraction. These include different floral symmetries, pigment color, petal shape, size and curvature, stamen number and length, gynophore (stalk structure that supports the gynoecium) elongation and curvature, nectar guides and nectar production, and nectar gland color, size and location (Cane 2008a; Higuera-Díaz *et al.* 2015; Machado *et al.* 2006; Patchell *et al.* 2011; Raju and Rani 2016). Because of their floral morphology, most Cleomaceae species presumably exhibit a generalist pollination system and studies suggest they are visited by a wide range of pollinator guilds (Cane 2008a; Higuera-Díaz *et al.* 2015; Martins and Johnson 2013; Raju and Rani 2016). However, some species exhibit a specialist pollination system, with pollinators including bats and hawkmoths (Fleming *et al.* 2009; Machado *et al.* 2006). Despite the variation exhibited in floral traits, pollination studies are scarce in the family (Cane 2008a; Higuera-Díaz *et al.* 2015; Machado *et al.* 2006; Martins and Johnson 2013; Raju and Rani 2016). Pollination biology has been studied in only 10 species, of which five exhibit a generalist and five a specialist pollination system (Table 1.1). In Chapter 3, I investigated the pollination biology of two native Cleomaceae species populations in southern Alberta, Canada, *C. serrulata* and *P. dodecandra*. I focused on these species because they were present in the province and their pollination biology is understudied. I collected empirical pollination data (Chapter 3), which I related with their floral transcriptome (Chapter 4).

1.3. Genomics and transcriptomics studies in Cleomaceae

Cleomaceae is a well explored model to address major evolutionary questions beyond pollination biology. These questions include understanding the presence and consequences of whole genome duplications in Cleomaceae and Brassicaceae (Barker *et al.*

2009; Mohammadin *et al.* 2015; Schranz and Mitchell-Olds 2006a; van den Bergh *et al.* 2016a; van den Bergh *et al.* 2014). In addition, research in Cleomaceae involves systematic and phylogenetic inquiries (Hall 2008; Inda *et al.* 2008; Patchell *et al.* 2014; Riser *et al.* 2013); origins, evolution and regulation of C_3 and C_4 carbon fixation pathways (Brautigam *et al.* 2011a; Feodorova *et al.* 2010; Külahoglu *et al.* 2014; van den Bergh *et al.* 2014); evolution of glucosinolates biosynthetic pathway and their role in bat pollinators attraction (van den Bergh *et al.* 2016a); arm race between Brassicales and their Pieridae butterflies herbivores (Edger *et al.* 2015); evolution and morphological diversity in Cleomaceae (Bhide *et al.* 2014; Cheng *et al.* 2013; Edger *et al.* 2015); and the study of long non-coding RNAs and their conservation in Brassicaceae and Cleomaceae (Mohammadin *et al.* 2015).

Many genomics and transcriptomics data sets have been generated to address the above questions. For example, the transcriptomics data generated to understand the evolution of C_4 carbon fixation pathways has been used to examine glucosinolate diversity and their potential link with the bat pollination syndrome (Külahoglu *et al.* 2014). Moreover, researchers have identified putative genes involved in floral pigment production responsible for flower color, which is a key trait for pollinator attraction (Bhide *et al.* 2014; van den Bergh 2017). Currently, there are full genomes available for two species of Cleomaceae *Tarenaya hassleriana* (Chodat) Iltis (Cheng et al. 2013) and *Cleome violacea* L. (https://genome.jgi.doe.gov/ClevioStandDraft/). In addition, there is RNA sequencing data for different tissues and developmental stages in five Cleomaceae species and one hybrid (Table 1.2). However, the above aforementioned did not explore the genetic variation underlying different pollination systems in Cleomaceae, and few studies address the genes encoding major floral traits and the role of the genes for pollinator attraction (Bhide *et al.* 2014; Külahoglu *et al.* 2014; van den Bergh *et al.* 2016a). Furthermore, empirical pollination studies and genetic basis of floral traits are overlooked in this family. In Chapter 4, I performed a transcriptome analysis in three Cleomaceae species to elucidate gene repertoire variations in species exhibiting generalized and specialized floral traits. I identified genes putatively associated with floral traits important for pollinator attraction. Further, I performed comparisons between species with generalist and specialist traits, which is important to find putative candidate genes coding for floral traits that could be subject to selection pressures by pollinators.

1.4. Objectives and outline

This thesis explores the pollination biology and the genes encoding for floral traits in three Cleomaceae species, which exhibit generalized and specialized traits. In Chapter 2, I perform a literature review into the molecular techniques and new technologies used to study floral traits and how these can be incorporated into ecological pollination studies through an integrative pollination studies approach; in which, I connect empirical pollination data and the genes encoding for traits important for pollinator attraction. In Chapter 3, I investigate the pollination biology of wild populations of *Cleomella serrulata* and *Polanisia dodecandra* in southern Alberta, Canada. My aimed was to identify the type of pollination system in these species, describing and comparing the floral visitor for both species and their effect on plant fitness. In Chapter 4, I perform a transcriptome analysis in three Cleomaceae species exhibiting different pollination systems to make qualitative comparisons between the genes found in species with generalized and specialized floral traits and pollinators and to find putative candidate genes encoding floral traits important for pollinator attraction. In Chapter 5, I determine the efficacy of VIGS and mutagenesis methodologies to obtain altered phenotypes in *P. dodecandra*, which could be used in future studies on pollinator preference tests. Finally, in Chapter 6, I end my thesis with a synthesis of general findings of my research and future directions.

1.5. References

Barker MS, Vogel H, Schranz ME. 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.

Baucom RS, Chang SM, Kniskern JM, Rausher MD, Stinchcombe JR. 2011. Morning Glory as a powerful model in ecological genomics: Tracing adaptation through both natural and artificial selection. *Heredity (Edinb)* **107**: 377-385.

Bhide A, Schliesky S, Reich M, Weber A, 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.

Bolstad GH, Armbruster WS, Pélabon C, Pérez-Barrales R, Hansen TF. 2010. Direct selection at the blossom level on floral reward by pollinators in a natural population of *Dalechampia schottii*: Full-disclosure honesty? *New Phytologist* **188**: 370-384.

Brautigam A, Kajala K, Wullenweber J, Sommer M, Gagneul D, Weber KL, Carr KM, Gowik U, Mass J, Lercher MJ, Westhoff P, Hibberd JM, Weber AP. 2011a. An mRNA blueprint for C4 photosynthesis derived from comparative transcriptomics of closely related C3 and C4 species. *Plant Physiology* **155**: 142-156.

Campbell DR, Waser NM, Price MV. 1996. Mechanisms of hummingbird-mediated selection for flower width in *Ipomopsis aggregata*. *Ecology* **77**: 1463-1472.

Cane JH. 2008a. Breeding biologies, seed production and species-rich bee guilds of *Cleome lutea* and *Cleome serrulata* (Cleomaceae). *Plant Species Biology* **23**: 152-158.

Cheng S, van den Bergh E, Zeng P, Zhong X, Xu J, Liu X, Hofberger J, de Bruijn S, Bhide AS, Kuelahoglu C, Bian C, Chen J, Fan G, Kaufmann K, Hall JC, Becker A, Brautigam A, Weber AP, Shi C, Zheng Z, Li W, Lv M, Tao Y, Wang J, Zou H, Quan Z, Hibberd JM, Zhang G, Zhu XG, Xu X, Schranz ME. 2013. The *Tarenaya hassleriana* genome provides insight into reproductive trait and genome evolution of crucifers. *Plant Cell* **25**: 2813-2830.

Clare EL, Schiestl FP, Leitch AR, Chittka L. 2013. The promise of genomics in the study of plant-pollinator interactions. *Genome Biology* **14**: 207.

Cronk Q, Ojeda I. 2008. Bird-pollinated flowers in an evolutionary and molecular context. *Journal of Experimental Botany* **59**: 715-727.

Devaux C, Lepers C, Porcher E. 2014. Constraints imposed by pollinator behaviour on the ecology and evolution of plant mating systems. *Journal of Evolutionary Biology* **27**: 1413-1430.

Edger PP, Heidel-Fischer HM, Bekaert M, Rota J, Glöckner G, Platts AE, Heckel DG, Der JP, Wafula EK, Tang M, Hofberger JA, Smithson A, Hall JC, Blanchette M, Bureau TE, Wright SI, dePamphilis CW, Eric Schranz M, Barker MS, Conant GC, Wahlberg N, Vogel H, Pires JC, Wheat CW. 2015. The butterfly plant arms-race escalated by gene and genome duplications. *Proceedings of the National Academy of Sciences* **112**: 8362-8366.

Epperson BK, Clegg MT. 1992. Unstable white flower color genes and their derivatives in Morning Glory. *Journal of Heredity* **83**: 405-409.

Fenster CB, Armbruster WS, Wilson P, Dudash MR, Thomson JD. 2004. Pollination syndromes and floral specialization. *Annual Review of Ecology Evolution and Systematics* **35**: 375-403.

Feodorova TA, Voznesenskaya EV, Edwards GE, Roalson EH. 2010. Biogeographic patterns of diversification and the origins of C4 in *Cleome* (Cleomaceae). *Systematic Botany* **35**: 811-826.

Fleming TH, Geiselman C, Kress WJ. 2009. The evolution of bat pollination: A phylogenetic perspective. *Annals of Botany* **104**: 1017-1043.

Galen C. 1999. Why do flowers vary? The functional ecology of variation in flower size and form within natural plant populations. *BioScience* **49**: 631-640.

Gübitz T, Hoballah ME, Dell'Olivo A, Kuhlemeier C. 2009. *Petunia* as a model system for the genetics and evolution of pollination syndromes. In: Gerats T, Strommer J, eds. *Petunia*. New York: Springer 29-49.

Hall JC. 2008. Systematics of Capparaceae and Cleomaceae: An evaluation of the generic delimitations of *Capparis* and *Cleome* using plastid DNA sequence data. *Botany-Botanique* **86**: 682-696.

Hawkins RD, Hon GC, Ren B. 2010. Next-Generation genomics: An integrative approach. *Nature Reviews. Genetics* **11**: 476-486.

Herrera CM. 2005. Plant generalization on pollinators: Species property or local phenomenon? *American Journal of Botany* **92**: 13-20.

Herrera CM, Pellmyr O. 2002b. *Plant-animal interactions : An evolutionary approach.* Oxford: Blackwell Science.

Higuera-Díaz M, Manson JS, Hall JC. 2015. Pollination biology of *Cleomella serrulata* and *Polanisia dodecandra* in a protected natural prairie in southern Alberta, Canada. *Botany* **93**: 745-757.

Hoballah ME, Gubitz T, Stuurman J, Broger L, Barone M, Mandel T, Dell'Olivo A, Arnold M, Kuhlemeier C. 2007. Single gene-mediated shift in pollinator attraction in *Petunia*. *Plant Cell* **19**: 779-790.

Inda LA, Torrecilla P, Catalán P, Ruiz-Zapata T. 2008. Phylogeny of *Cleome* L. and its close relatives *Podandrogyne* Ducke and *Polanisia* Raf. (Cleomoideae, Cleomaceae) based on analysis of nuclear ITS sequences and morphology. *Plant Systematics and Evolution* **274**: 111-126.

Johnston MO. 1991. Natural selection on floral traits in two species of *Lobelia* with different pollinators. *Evolution* **45**: 1468-1479.

Kay KM, Sargent RD. 2009. The role of animal pollination in plant speciation: Integrating ecology, geography, and genetics. *Annual Review of Ecology Evolution and Systematics* **40**: 637-656.

Kessler D, Baldwin IT. 2011. Back to the past for pollination biology. *Current Opinion in Plant Biology* **14**: 429-434.

Kessler D, Bhattacharya S, Diezel C, Rothe E, Gase K, Schottner M, Baldwin IT. 2012. Unpredictability of nectar nicotine promotes outcrossing by hummingbirds in *Nicotiana attenuata*. *Plant Journal* **71**: 529-538.

Kessler D, Gase K, Baldwin IT. 2008. Field experiments with transformed plants reveal the sense of floral scents. *Science* **321**: 1200-1202.

Klahre U, Gurba A, Hermann K, Saxenhofer M, Bossolini E, Guerin PM, Kuhlemeier C. 2011. Pollinator choice in *Petunia* depends on two major genetic loci for floral scent production. *Current Biology* **21**: 730-739.

Külahoglu C, Denton AK, Sommer M, Mass J, Schliesky S, Wrobel TJ, Berckmans B, Gongora-Castillo E, Buell CR, Simon R, De Veylder L, Brautigam A, Weber AP. 2014. Comparative transcriptome atlases reveal altered gene expression modules between two Cleomaceae C3 and C4 plant species. *Plant Cell* **26**: 3243-3260.

Machado IC, Lopes AV, Leite AV, de Brito Neves C. 2006. *Cleome spinosa* (Capparaceae): Polygamodioecy and pollination by bats in urban and Caatinga areas, northeastern Brazil. *Botanische Jahrbücher* **127**: 69-82.

Martins DJ, Johnson SD. 2013. Interactions between hawkmoths and flowering plants in East Africa: Polyphagy and evolutionary specialization in an ecological context. *Biological Journal of the Linnean Society* **110**: 199-213.

Mayer C, Adler LS, Armbruster WS, Dafni A, Eardley C, Huang SQ, Kevan PG, Ollerton J, Packer L, Ssymank A, Stout JC, Potts SG. 2011. Pollination ecology in the 21st Century: Key questions for future research. *Journal of Pollination Ecology* **3**: 8-23.

Medel R, Botto-Mahan C, Kalin-Arroyo M. 2003. Pollinator-mediated selection on the nectar guide phenotype in the Andean Monkey flower, *Mimulus luteus*. *Ecology* **84**: 1721-1732.

Mohammadin S, Edger PP, Pires JC, Schranz ME. 2015. Positionally-conserved but sequence-diverged: Identification of long non-coding RNAs in the Brassicaceae and Cleomaceae. *BMC Plant Biology* **15**: 217.

Owen CR, Bradshaw HD. 2011. Induced mutations affecting pollinator choice in *Mimulus lewisii* (Phrymaceae). *Arthropod-Plant Interactions* **5**: 235-244.

Patchell MJ, Bolton MC, Mankowski P, Hall JC. 2011. Comparative floral development in Cleomaceae reveals two distinct pathways leading to monosymmetry. *International Journal of Plant Sciences* **172**: 352-365.

Patchell MJ, Roalson EH, Hall JC. 2014. Resolved phylogeny of Cleomaceae based on all three genomes. *Taxon* **63**: 315-328.

Proctor MCF, Yeo P, Lack A. 1996. *The natural history of pollination.* Portland: Timber Press.

Raju AJS, Rani DS. 2016. Reproductive ecology of *Cleome gynandra* and *Cleome viscosa* (Capparaceae). *Phytologia Balcanica* **22**: 15-28.

Riser JP, Cardinal-McTeague WM, Hall JC, Hahn WJ, Sytsma KJ, Roalson EH. 2013. Phylogenetic relationships among the North American cleomoids (Cleomaceae): A test of Iltis's reduction series. *American Journal of Botany* **100**: 2102-2111.

Rosas‐Guerrero V, Aguilar R, Martén‐Rodríguez S, Ashworth L, Lopezaraiza‐Mikel M, Bastida JM, Quesada M. 2014. A quantitative review of pollination syndromes: Do floral traits predict effective pollinators? *Ecology Letters* **17**: 388-400.

Sandring S, Agren J. 2009. Pollinator-mediated selection on floral display and flowering time in the perennial herb *Arabidopsis lyrata*. *Evolution* **63**: 1292-1300.

Sapir Y. 2009a. Effects of floral traits and plant genetic composition on pollinator behavior. *Arthropod-Plant Interactions* **3**: 115-129.

Sapir Y. 2009b. Pollination genetics: Using molecular genetic underlying floral traits to study pollination ecology in an evolutionary context. *Israel Journal of Plant Sciences* **57**: 141-149.

Sapir Y, Armbruster WS. 2010. Pollinator-mediated selection and floral evolution: From pollination ecology to macroevolution. *New Phytologist* **188**: 303-306.

Schemske DW, Bradshaw HD. 1999. Pollinator preference and the evolution of floral traits in Monkey flowers (*Mimulus*). *Proceedings of the National Academy of Sciences of the United States of America* **96**: 11910-11915.

Schranz ME, Mitchell-Olds T. 2006a. Independent ancient polyploidy events in the sister families Brassicaceae and Cleomaceae. *The Plant Cell Online* **18**: 1152-1165.

Sedeek KE, Qi W, Schauer MA, Gupta AK, Poveda L, Xu S, Liu ZJ, Grossniklaus U, Schiestl FP, Schluter PM. 2013. Transcriptome and proteome data reveal candidate genes for pollinator attraction in sexually deceptive orchids. *PLoS ONE* **8**: e64621.

Sharma B, Kramer EM. 2013. Virus-induced gene silencing in the rapid cycling columbine *Aquilegia coerulea* "Origami". In: Becker A, ed. *Virus-Induced Gene Silencing: Methods and Protocols*. Totowa, NJ: Humana Press, 71-81.

Sletvold N, Trunschke J, Smit M, Verbeek J, Ågren J. 2016. Strong pollinator-mediated selection for increased flower brightness and contrast in a deceptive orchid. *Evolution* **70**: 716-724.

Sobral M, Veiga T, Domínguez P, Guitián JA, Guitián P, Guitián JM. 2015. Selective pressures explain differences in flower color among *Gentiana lutea* populations. *PLoS ONE* **10**: e0132522.

van den Bergh E. 2017. *Comparative genomics and trait evolution in Cleomaceae, a model for ancient polyploidy,* PhD, Wageningen University

van den Bergh E, Hofberger JA, Schranz ME. 2016a. Flower power and the mustard bomb: Comparative analysis of gene and genome duplications in glucosinolate biosynthetic pathway evolution in Cleomaceae and Brassicaceae. *American Journal of Botany* **103**: 1212-1222.

van den Bergh E, Külahoglu C, Bräutigam A, Hibberd JM, Weber AP, Zhu X-G, Schranz 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.

Verdonk JC, Haring MA, van Tunen AJ, Schuurink RC. 2005. ODORANT1 regulates fragrance biosynthesis in *Petunia* flowers. *Plant Cell* **17**: 1612-1624.

Willmer P. 2011. *Pollination and floral ecology.* Princeton: Princeton University Press.

Xu SQ, Schluter PM, Grossniklaus U, Schiestl FP. 2012. The genetic basis of pollinator adaptation in a sexually deceptive orchid. *Plos Genetics* **8**.

Yuan YW, Byers KJRP, Bradshaw HD. 2013*a*. The genetic control of flower-pollinator specificity. *Current Opinion in Plant Biology* **16**: 422-428.

Zhang W, Wei X, Meng H-L, Ma C-H, Jiang N-H, Zhang G-H, Yang S-C. 2015b. Transcriptomic comparison of the self-pollinated and cross-pollinated flowers of *Erigeron breviscapus* to analyze candidate self-incompatibility-associated genes. *BMC Plant Biology* **15**: 248.

Zhao Z-G, Wang Y-K. 2015. Selection by pollinators on floral traits in generalized *Trollius ranunculoides* (Ranunculaceae) along Altitudinal Gradients. *PLoS ONE* **10**: e0118299.

Table 1.1. Different pollination systems exhibited by Cleomaceae species explored to date.

Pollination system	Species	Major pollinators	References
Generalist	*Arivela viscosa* (L.) Raf.	Bees, butterflies, flies	Raju and Rani 2016
	Cleomella lutea (Hook) E.H. Roalson and J.C. Hall	Bees, butterflies, flies	Cane 2008a
	Cleomella serrulata (Pursh) E.H. Roalson and J.C. Hall	Bees, butterflies, flies	Cane 2008a
	Gynandropsis gynandra (L.) Briq*	Bees, butterflies, flies, moths	Raju and Rani 2016, Martins and Johnson 2013
Specialist	*Cleome anomala* Kunth	Bats	Fleming *et al*. 2009
	Cleomella arborea (Nutt) E.H. Roalson and J.C. Hall	Bats	Fleming *et al*. 2009
	Cleome moritziana Klotzsch ex Eichler	Bats	Fleming *et al*. 2009
	Melidiscus giganteus (L.) Raf.	Bats	Fleming *et al*. 2009
	Tarenaya spinosa (Jacq.) Raf.	Bats, moths	Machado *et al.* 2006

* African populations appear to be specialists (Martins and Johnson 2013), while Asian populations appear to be generalists (Raju and Rani (2016).

Table 1.2 Available RNA sequencing data of different tissues in Cleomaceae species.

Species	Tissue	References
Cleome droserifolia (Forssk) Delile	Flowers, roots, young leaves	Mohammadin *et al*. 2015
Cleome violacea (L.) Raf.	Buds, open flowers, fruits, roots, seedlings, stem	Edger *et al.* 2015
Gynandropsis gynandra (L.) Briq.	Flowers, leaves, roots, seedlings	Brautigam *et al.* 2011a, Külahoglu *et al.* 2014
Tarenaya spinosa (Jacq.) Raf.	Leaves	Barker *et al.* 2009, Brautigam *et al.* 2011a
Tarenaya hassleriana (Chodat) Iltis	Apical meristem, buds, leaves, open flowers, roots, seedlings	Bhide *et al.* 2014, Külahoglu *et al.* 2014

Chapter 2: Bridging pollination ecology and floral developmental evolution through molecular techniques: A review

2.1. Introduction

Animal pollination is a critical process for reproduction of most angiosperms species. Through specific cues and rewards, flowers attract pollinators such as insects, birds, lizards, bats, and small mammals to vector and deposit pollen between conspecific plants (Proctor *et al.* 1996; Willmer 2011). Due to its importance, plant-pollinator interactions have been widely studied from an evolutionary perspective, focusing on generalized *vs.* specialized pollination systems (Armbruster 2012; Johnson and Steiner 2000; Waser *et al.* 1996), pollination syndromes (Fenster *et al.* 2004; Fernandez-Mazuecos and Glover 2017; Martin *et al.* 2008; Ollerton *et al.* 2009; Ollerton *et al.* 2015), the influence of pollinators on the evolution of plant mating systems (Devaux *et al.* 2014; Karron *et al.* 2012; Mayer *et al.* 2011), the evolution of floral traits (Fernandez-Mazuecos and Glover 2017; Schiestl and Johnson 2013; Smith *et al.* 2008; Specht *et al.* 2012), pollinator-mediated selection (Schiestl and Johnson 2013; Smith 2010; Smith *et al.* 2008; van der Niet *et al.* 2014), and how variation in floral traits can influence pollinator behavior, leading to changes in pollinator communities (Chittka and Raine 2006; Dyer *et al.* 2007; Mayer *et al.* 2011). There are three common approaches to studying pollination: (1) ecological studies, which focus on behavioral responses of pollinators to different floral phenotypes; (2) evolutionary studies that focus on the genetics underlying floral traits presumably involved in pollination; and (3) phylogenetic studies that look at changes in morphology in plants over time (Figure 2.1). However, the link between the genetic basis of floral traits and their effects on pollinator attraction is still poorly understood (Fernandez-Mazuecos and Glover 2017; Sapir 2009b; Sapir and Armbruster 2010; Yuan *et al.* 2013*a*). It is important to establish this link, because changes in genes encoding main floral traits may affect the phenotype, which could affect the pollinator behavior and in turn plant fitness. Many studies have determined the genetic basis of floral traits in a range of lineages (Anderson *et al.* 2011; Stellari *et al.* 2004; Yuan *et al.* 2013*a;* Zhang *et al.* 2010); yet, few investigations directly test how genetic changes in floral traits affect pollinator attraction and plant fitness (but see Kessler *et al.* 2008; Owen and Bradshaw 2011; Sheehan

et al. 2012; Yuan *et al.* 2013*b*). Studies that integrate both the genetics underlying a floral trait and pollinator response to variation in the same trait, hereafter referred to as integrative pollination studies, provide a powerful tool to link phenotypic variation in flowers and pollinator behavior to pollinator-mediated selection and the diversification of flowering plants (Figure 2.1).

Studying pollination using genetic tools is an example of a new field called ecogenomics, which study how ecological interactions are affected by different factors including development, genotypic, and subsequent phenotypic variation (Ouborg and Vriezen 2007). We are now poised to make great strides in linking the genetic basis of floral traits and the impact of pollinators in shaping these traits. Data on the genetic basis of traits potentially linked to pollinator interactions is now available from a range of plant species. Moreover, these studies are complemented by an unprecedented amount of accumulated genomic data in model and non-model organisms obtained using Next Generation Sequencing (NGS) technologies (Almeida *et al.* 1997; Bender *et al.* 2012; Bhide *et al.* 2014; Hoballah *et al.* 2007; Lulin *et al.* 2012; Sedeek *et al.* 2013; Sheehan *et al.* 2012; Xu *et al.* 2012). The 'omics' (genomics, transcriptomics, proteomics and metabolomics) derived data from NGS and other technologies can be used to perform broad scale genomic comparison studies in plants that differ in floral morphology, reproductive strategy, life history traits and pollinator guilds, which can in turn be used to directly test function of trait variation on pollinator behavior and relate them to plant fitness. Here we review 1) techniques used to address the genetic basis of floral traits important for animal-mediated pollination; 2) potential applications and pitfalls of NGS with an emphasis on identifying which candidate genes and metabolic pathways can be modified for integrated pollination studies; and 3) examples of emerging plant systems that are good candidates for these studies. We also highlight how this approach can be used to study the evolution of specialist, generalist and deceptive pollination systems. Finally, we identify questions and make suggestions for future research, emphasizing the significance of NGS for studying pollination using an integrative approach.

2.2. Beyond the classic morphological manipulation techniques of floral traits

2.2.1. Pollinator preference studies

The first step to investigating the relationship between floral traits and pollination is often a study of pollinator preferences and these studies have primarily focused on flower color (Chittka and Kevan 2005; Gumbert 2000; Hansen *et al.* 2012), flower symmetry (Horridge 1996; Møller 1995), inflorescence display sizes (Bayo-Canha *et al.* 2007; Schmid-Hempel and Speiser 1988), floral scents (Gaskett 2011; Kessler *et al.* 2008), and nectar production (Erhardt and Rusterholz 1998; Hansen *et al.* 2012). Most studies have been performed with honeybees and bumblebees, but other taxa like beetles, butterflies, flies, birds and bats have also been examined (Baker *et al.* 1998; Dornhaus and Chittka 1999; Erhardt and Rusterholz 1998; Goldblatt and Manning 2000; Harder and Johnson 2005; Horridge 1996; Lewis 1986; Menzel and Muller 1996; Møller 1995; Møller 2000; Neal *et al.* 1998). Furthermore, many studies have addressed how certain combinations of floral traits attract pollinator guilds and how pollination syndromes are associated with floral diversification and speciation (Fenster *et al.* 2004; Fenster *et al.* 2009; Hodges and Kramer 2007; Whittall and Hodges 2007; Willmer 2011; Wilson *et al.* 2004).

Researchers have historically studied pollinator preference using techniques such as artificial flowers, clipping and painting petals, and artificial scents. While these techniques are valuable, they are not ideal; artificial flowers are simple models that mimic floral shape but they are often far from the essence of a real flower. Furthermore, some of these experiments physically damage the flowers and the observed changes in pollinator behavior may be, at least in part, a response of confounding factors such as alterations in floral size and shape, and chemical signals (Neal *et al.* 1998). Some floral traits such as floral scents, flowering date, nectar composition, and pollen production prove very difficult to manipulate (Neal *et al.* 1998; Raguso 2008). The development of three-dimensional printed flowers has resolved morphological issues, allowing for the modification of color, shape, size, even scent (synthetized scents are applied to the artificial flowers), and other physical traits in artificial flowers (Policha *et al.* 2016). However, the plastic material used, the colors, and their texture are still only facsimiles of real flowers.

2.2.2. Molecular approaches to testing pollinator preferences

The most effective approach to modifying floral traits while limiting confounding

factors may be through genetic manipulation. We can genetically modify plants to alter floral phenotypes, which in turn can be used to evaluate pollinator response and subsequent effects on plant fitness. A key prerequisite for these experiments is appropriate knowledge of the genetic basis of these traits. Prior to the current –omics era, most evo-devo studies compared gene expression patterns and, ideally, functional tests to identify candidate genes for manipulation (Table 2.1). In general, candidate genes are selected based on known gene function primarily in model species, which allow us to make inferences about conservation of gene function in species other than model systems (Cronk *et al.* 2002; Gübitz *et al.* 2009; Pflieger *et al.* 2001). The comparative gene expression approach has led to substantial understanding of genes underlying floral traits, including floral color (e.g., Clegg and Durbin 2000; Dyer *et al.* 2007; Epperson and Clegg 1992; Hoballah *et al.* 2007; Quattrocchio *et al.* 1999; Rausher and Fry 1993; Whibley *et al.* 2006; Xu *et al.* 2012), scent (e.g., Dudareva *et al.* 1996; Dudareva *et al.* 1998; Kessler *et al.* 2010; Kessler *et al.* 2008; Sedeek *et al.* 2013; Spitzer-Rimon *et al.* 2012; Verdonk *et al.* 2005; Xu *et al.* 2012), and symmetry (e.g., Altieri 1999; Bey *et al.* 2004; Hileman and Cubas 2009; Kramer 2009*a;* Pashley *et al.* 2006; Zhang *et al.* 2010).

Genetic mapping of lab crosses using quantitative trait loci (QTLs) or near isogenic lines (NILs) is another approach to identify the genetic bases of key plant traits (reviewed Anderson and Mitchell-Olds 2011; Cronk *et al.* 2002; Rifkin 2012). QTLs are a statistical method used to link morphological and genetic data where the main goal is to identify chromosomal regions and linked genes underlying the trait of interest. Historically, the genes in QTLs were largely unknown, but more recent studies are pairing next generation sequencing with QTLs analysis to concurrently identify genes (Celik *et al.* 2017; Luo *et al.* 2012; Pootakham *et al.* 2015; Zou *et al.* 2013). The QTL approach (Table 2.1) has been used to study complex floral traits in *Helianthus* (Sapir 2009a)*, Iris* (Martin *et al.* 2008)*, Mimulus* (Schemske and Bradshaw 1999; Yuan *et al.* 2013*b*)*, Penstemon* (Wessinger *et al.* 2014*)*, and *Petunia* (Gübitz *et al.* 2009; Klahre *et al.* 2011). NILs, a similar method, are used as starting material to get introgression lines through recombination of subsequent crosses. Usually cross-parental lines are obtained in which one individual exhibits a trait of interest previously mapped to a specific locus and the other individual lacks this trait. This technique can be used between species to make hybrids (Kooke *et al.* 2012). NILs were

used successfully to evaluate pollinator preferences in *Mimulus* (Bradshaw and Schemske 2003; Table 1).

Reverse genetic approaches may be a good approach for integrated pollination studies for two major reasons. First, these experiments explicitly test whether candidate genes are responsible for specific floral traits (Kessler and Baldwin 2011; Neal *et al.* 1998; Sapir 2009b; Sapir and Armbruster 2010; Yuan *et al.* 2013*a*). Second, they generate modified floral phenotypes, which can be used in pollinator preference experiments. These approaches can result in plants with stable or transient phenotypes. Virus-induced gene silencing (VIGS), a transient post-transcriptional method for down-regulating genes of interest, is being applied to a range of model and non-model species (Gould and Kramer 2007; Lange *et al.* 2013; Senthil-Kumar and Mysore 2014). VIGS uses a viral vector carrying a fragment of one or more target genes but it has variable penetrance (Gould and Kramer 2007; Senthil-Kumar and Mysore 2011; Stratmann and Hind 2011). Genetic transformation mediated by *Agrobacterium tumefaciens* (Smith and Townsend) generates stable mutants (Weigel and Glazebrook 2002). This technique was used successfully in *Nicotiana attenuata* using RNAi constructs to test the effect of Chalcone synthase, a gene underlying floral scent, in pollinator attraction (Kessler *et al.* 2012; Table 1; Kessler *et al.* 2008). However, transformation success is limited in many non-model species (Becker and Lange 2010; Collier *et al.* 2005; Wege *et al.* 2007). Alternatively, CRISPR/cas9 (clustered regularly interspaced short palindromic repeats/cas associated protein 9) is a promising, newly developed and stable genome editing technique, in which a target DNA sequence is transformed by adding, removing or replacing DNA nucleotides using CRISPR associated proteins (Bortesi and Fischer 2015; Kumar and Jain 2015).

Forward genetic approaches may also generate altered floral phenotypes for pollinator preferences studies, although the genes that are modified are unknown without additional genotyping. In chemical mutagenesis, seeds are immersed in a mutagenizing agent such as ethyl methanesulfonate (EMS). Plants grown from the treated seeds are self-fertilized to obtain a second generation, which are then screened for variation in morphological traits (Leyser 2000; Maple and Moller 2007; Swarbreck *et al.* 2008; Weigel and Glazebrook 2002). To date, one study has used chemical mutagenesis in plants to address questions related to floral morphology and pollinator behavior in *Mimulus lewisii*

(Owen and Bradshaw 2011; Table 1).

While only transformation and mutagenesis have been used to study pollination, there are a wealth of approaches that can be applied. Moreover, the availability of multiple techniques provides alternative approaches should certain methods fail in a particular species. In particularly, techniques such as VIGS and CRISPR are promising tools for future studies.

2.3. Next generation sequencing, transcriptomics and pollination ecology

Next generation sequencing can be used to discover genes responsible for floral traits, especially in non-model species (Brautigam and Gowik 2010; Metzker 2010). NGS approaches have several advantages over the more traditional Sanger sequencing method. First, NGS techniques produce a huge amount of data quickly and less expensive than Sanger sequencing or per base pair (Henson *et al.* 2012; van Dijk *et al.* 2014). NGS can also obtain sequence information from a single DNA molecule, and instead of using individual techniques such as molecular markers and microarrays, we can get most the information all at once (Wang *et al.* 2009). Here we focus on using RNA-seq approaches, as they are likely to substantially contribute to our understanding of genetic basis of floral traits. We focus on RNA-seq because this technique provides differential gene expression data not only in terms of presence/absence, but also quantitative data such as transcripts abundance, which allow us to perform differential expression comparisons of genes across related species exhibiting differences in floral traits.

RNA-seq is a powerful NGS technique to characterize and quantify transcriptomes, which comprise the analysis of mRNA, small RNAs, and non-coding RNAs transcripts (Lemmon and Lemmon 2013; Wang *et al.* 2009). RNA-seq has several advantages, including the detection of all the transcripts present in the sample, identification of rare transcripts, and precise quantification of differential transcript expression. The main advantage of RNA-seq, however, is the possibility of performing *de novo* assembly, making it a powerful technique in non-model species to identify differentially expressed transcripts and novel genes without a reference genome (Brautigam and Gowik 2010; Garber *et al.* 2011; Martin *et al.* 2013; Strickler *et al.* 2012; Wang *et al.* 2010a; Wang *et al.* 2009; Ward *et al.* 2012). However, the technique requires fresh and high quality RNA as starting material (but see Gallego Romero *et al.* (2014)), high sequencing depth to get

rarely expressed genes, and RNA obtained from the same types of tissue from each individuals at the same developmental stage to perform quantitative comparisons (Brautigam and Gowik 2010; Martin *et al.* 2013; Strickler *et al.* 2012; Ward *et al.* 2012). In addition, RNA-seq has challenges associated with reconstructing transcripts and sequence assembly such as short read lengths, alternative splicing which results in different isoforms, reads comprised of exon-exon junctions, and differential expression of transcript abundance (Honaas *et al.* 2016; Ozsolak and Milos 2011; Wang *et al.* 2009). Furthermore, analyses of the large amount of sequence data is challenging due to data storage, processing time, and data interpretation, but this depends on the methodology, genome size, replicates number among others (Garber *et al.* 2011; Honaas *et al.* 2016; Lemmon and Lemmon 2013; Ozsolak and Milos 2011; Wang *et al.* 2010a; Wang *et al.* 2009). Moreover, the choice of pipeline can be challenging and factors such as proportion of mapping reads, recovery of widely expressed genes, N50 statistics (length at which the assembly bp reaches 50% of the total assembly length), and number of unigenes or unique represented genes should be considered when choosing an assembler (Honaas *et al.* 2016).

Taking advantage of decreased sequencing cost and the huge amount of data produced from NGS, three ambitious macro projects are performing RNA-seq in a broad group of plants (1) the 1000 plants or 1KP, http://onekp.com/project.html, (2) Brassicales map alignment project or BMAP, http://www.brassica.info/resource/sequencing/bmap.php, and (3) the medicinal plant genomics resource or MPGR, http://medicinalplantgenomics.msu.edu/index.shtml. These projects include research areas such as agriculture, angiosperm diversification, biochemistry, extremophytes, and invasive species, among others. The data derived from these projects may provide resources for transcriptome comparisons of floral traits associated with pollination in different plant species; however, caution must be taken because the tissue used for RNA extraction comes mostly from leaves. These sites took the first steps towards making transcriptome data for angiosperms broadly accessible and are a starting point for identifying transcripts specific to floral traits.

In addition to identifying and comparing particular genes, transcriptomes can be used to examine gene regulatory networks (GRN) and metabolic pathways through enrichment analysis of differentially expressed genes (Wang and Cairns 2013), which is

important because many floral traits are encoded and regulated by multiple genes. For example, studies examining GRN have found that morphological changes in leaves are due to modulation or change in peripheral transcription factors, not in key factors as expected (Ichihashi and Tsukaya 2015). In contrast, studies involving GRN in flowers have been focused on genes controlling the initiation of flower development in *A. thaliana* L., in which key transcription factors are responsible for flower development and floral whorls identity (Ó'Maoiléidigh *et al.* 2014; Wellmer and Riechmann 2010).

Despite the increasing number of available plant transcriptomes, this approach has rarely been used to link floral traits to pollinator responses. Sedeek *et al.* (2013), studying the transcriptome of three closely related deceptive orchids, identified candidate genes responsible for pollinator attraction and reproductive isolation like genes regulating alkane and alkene production (flower scent) and anthocyanin biosynthesis (flower color). This comparative transcriptome analysis is an example of finding putative candidate genes in related species for floral traits associated with pollination attraction and reproductive isolation. Broader comparative analyses that utilize the information generated by large-scale 'omics' projects are an opportunity to establish the function of the genes in key ecological interactions such as pollination (Ungerer *et al.* 2007).

2.4. Linking plant genetics and pollinator behavior

2.4.1. 'omics' potential to study floral rewards and pleiotropy

One major advantage of 'omics' data is the opportunity to study traits that are difficult to manipulate in traditional ways, such as nectar and pollen production. Despite the importance of nectar and pollen for pollinators, the underlying genes of these traits are poorly known, though some candidate genes have been proposed to encode for the development of floral nectaries, the production of nectar, and the rates of nectar secretion (Bender *et al.* 2012; Lin *et al.* 2014). Researchers are working towards identifying the genetic components underlying nectar composition, function, and structure in *Arabidopsis thaliana* and other species of Brassicaceae (http://www.nectarygenomics.org), using mutant lines, gene expression profiling, transcriptome analysis, metabolomics data, and gene knockout studies (Bender *et al.* 2012). In addition, Yant *et al.* (2015) studied the transcriptome of petal nectar spurs in *Aquilegia* and discovered a new set of candidate

genes involved in the development of nectar spurs that had not been considered before. On the other hand, the genetic basis that defines the nutritional composition of pollen, which is likely a key factor in driving pollinator response, is still a black box. Novel molecular methods and resources provide a unique opportunity to assess how nectar and pollen mediate plant-pollinator interactions.

'Omics' data can also address floral traits that are expressed or regulated by many genes and can conversely allow us to examine genes that express or regulate many traits. For example, it can be used to identify and understand pleiotropy in floral traits, in which a single gene could influences several floral traits (Smith 2016). Generally, researchers assess whether the pleotropic effects on the traits promote or constraint adaptive evolution. Selection on floral pleotropic traits have been described for several angiosperm species (Rausher 2008) and genes with pleiotropic effects on floral traits have been identified in mutant lines of model species such as *A. thaliana* (Smith 2016). Since pleotropic genes affect multiple traits, we should be cautious when manipulating candidate genes to study the relevance of a single trait for pollinator behavior, because the candidate genes could have cascading and unpredicted effects on other floral or plant traits. On the other hand, using candidate genes that exhibit pleiotropic effects could be used to examine how variations in multiple and correlated floral traits affect pollination and plant fitness.

2.4.2. 'Omics' as a tool for integrative pollination studies

Integrative pollination studies have great potential to improve our understanding of the evolution of plant-pollinator interactions by allowing us to discover which genes are encoding floral traits that are under selection and observing pollinator response to changes on those traits. The integrated approach can validate the importance of specific traits for pollinator attraction but can also lead to unexpected responses by pollinators or other plant visitors. Several studies have used genetic manipulation to examine pollinator attraction and shifts in pollination guilds due to changes in flower color (e.g., Bradshaw and Schemske 2003; Dyer *et al.* 2007; Hoballah *et al.* 2007; Schemske and Bradshaw 1999). In contrast, few studies look beyond pollinator response to fitness consequences for the plants. Kessler *et al.* (2008, 2010), used RNAi constructs, where nectar and scent properties were manipulated to test pollinator visitation. Further, Kessler *et al.* (2008) took into consideration how changes in transformed plants may have influenced other visitors such as

nectar robbers and herbivores. For example, silencing the gene for nicotine production in nectar led to increased florivory and nectar robbing, while changes in repellent and attractive floral volatiles had consequences for pollinator visit number and plant fitness (Kessler *et al.* 2008). This example illustrates the power of an experimental approach that integrates genetic manipulation with pollinator response and plant fitness, in which new information regarding the behavior of pollinators, herbivores and nectar robbers was obtained.

'Omics' is an excellent approach for finding candidate genes that reflect pollinator-mediated selection (Figure 2.2a). Frequently these candidate genes are those that reveal differential expression patterns, which are down- or up-regulated. An integrative pollination study often starts by identifying a trait (or traits) that exhibits phenotypic, and potentially genotypic, variation and is predicted to influence plant-pollinator interactions (Figure 2.2a). Using RNA-seq, we can then identify possible candidate gene(s) underlying this focal trait. Next, the genes are modified through VIGS, stable transformation or CRISPR to try to generate phenotype variation of the focal trait. These modified plants are in turn exposed to pollinators to evaluate their response and subsequent effects on plant fitness (Figure 2.2a). Pollinator response can be measured in number of visits, visitation length, or pollen transfer rates, but changes in male and female plant fitness are the ultimate indicator of pollinator-mediated selection on the focal trait. These studies can be performed in the lab under controlled conditions using naïve pollinators, which allows for unbiased evaluation of pollinator preference. Alternatively, by placing genetically modified plants into the field, we can evaluate interactions with a broader range of animals including nectar and pollen robbers and herbivores, leading to estimates of how a floral trait affects plant fitness under natural conditions. Field experiments can indicate the importance of selection on floral traits by antagonists and may also reveal pleiotropic traits. However, experiments in natural settings also have drawbacks: pollinator response in field-based experiments may be influenced by previous foraging experience (Clare *et al.* 2013) and there is a risk of introducing genetically modified seeds into natural plant populations, depending on the method of modification used.

'Omics' investigations on particular species may be expanded across promising clades that house species with variable traits and different pollinators. In addition, 'omics'

derived data can be used to compare gene sequences in a group of taxa that have multiple traits gain, losses or modifications (Chanderbali *et al.* 2016). We can also make comparisons across distant taxa to study evolutionary adaptive changes in the trait among species (Figure 2.2b). For example, having comparative transcriptome data across a clade or clades might allow us to infer whether the same genes and/or pathways are being recruited and if those changes are linked to particular floral traits. An 'omics' approach can generate the initial experimental baseline for future studies; processes associated with the evolution of traits detected by 'omics' can then be validated through manipulating plant genotype and evaluating pollinator response. To date, these kinds of approaches have been used in orchids and the availability of 'omics' data is rapidly increasing in this group (Mondragón-Palomino 2013).

2.5. Using emerging model and non-model plants to address the genetic basis of pollination systems

The emerging genomic resources on plants with divergent evolutionary histories, allow us to choose new models to explore plant-pollinators interactions. A good plant model should be fast and easy to grow, have a short generation time, a wide range of floral traits across species and clades, be pollinated by different vectors, have genomic resources, and be amenable to experimental genetic manipulation (Baucom *et al.* 2011; Kramer 2009*b*). Frequently, pollination studies focus only on specialist pollination systems since there is often a strong link between variation in phenotype, pollinator preference and plant fitness. Generalist pollination systems are overlooked because there may be limited directional selection on floral traits, which can make them difficult to study (Fernández *et al.* 2009; Johnson and Steiner 2000). However, many plants belonging to different families exhibit generalist pollination systems and the interaction between these plants and their pollinators is frequently undetermined. On the other hand, the evolution of deceptive pollination systems (e.g., sexual or food deception) allow us to study how the deception is maintained in this parasitic plant-pollinator relationship. Here, we present here three promising taxa, Cleomaceae, *Aquilegia* (Ranunculaceae), and *Ophrys* (Orchidaceae), that serve as emerging plant models to study the genetic basis of generalist, specialist and sexually deceptive pollination systems, respectively.

2.5.1. Generalist pollination system

Most Cleomaceae species represent a generalist pollination system, where flowers can be effectively pollinated by a wide range of pollinators (Cane 2008a; Higuera-Díaz *et al.* 2015; Martins and Johnson 2013; Raju and Rani 2016). However, some species such as *Cleome anomala* Kunth*, Cleome isomeris* Green*, Cleome moritziana* Klotzsch ex Eichler*, M. giganteus* (L.) Raf., and *Tarenaya spinosa* (Jacq.) Raf., exhibit a specialist pollination system, being pollinated by bats and hawkmoths (Fleming *et al.* 2009; Machado *et al.* 2006), but these species are exceptions to the rule. In general, Cleomaceae is an excellent model to explore the genetic basis of generalist pollination systems because of their great variation in floral traits associated with advertisement, including flower symmetry, flower sizes, petal colour, petal dimorphism, ultraviolet light reflection, stamen length, gynophore elongation, nectar guides, and floral fragrances (Cane 2008a; Erbar and Leins 1997; Iltis *et al.* 2011b; Nozzolillo *et al.* 2010; Patchell *et al.* 2011). In addition, its flowers also present variation in traits related to pollinator rewards such as nectar glands, nectar production, nectar sugar concentration, pollen colour, and pollen production (Cane 2008a; Erbar and Leins 1997; Higuera-Díaz *et al.* 2015; Raju and Rani 2016). This variation leads to the attraction of different animal pollinators such as bats, birds, flies, bees, wasps, and butterflies (Cane 2008a; Higuera-Díaz *et al.* 2015; Raju and Rani 2016; van den Bergh *et al.* 2016a).

Cleomaceae has a growing number of genomic resources including the genome and the transcriptome of leaves and flower tissues of *Cleome violacea* L., the transcriptome of *Gynandropsis gynandra* (L.) Briq., and *T. spinosa*, the genome of *Tarenaya hassleriana* (Chodat) Iltis, and the floral transcriptome of a hybrid *T. hassleriana* (Barker *et al.* 2009; Bhide *et al.* 2014; Brautigam *et al.* 2011a; Braütigam *et al.* 2011b; Cheng *et al.* 2013; Edger *et al.* 2015; Külahoglu *et al.* 2014). The floral transcriptome of *T. hassleriana* led to the identification of the genes putatively controlling the anthocyanin pathway, floral phenology, organ development, and male sterility, all traits associated with pollination (Bhide *et al.* 2014). Moreover, when compared with the leaf transcriptome, a representative proportion of transcription factor genes were unique for the floral transcriptome (Bhide *et al.* 2014). Cleomaceae is therefore well poised to conduct integrated pollination studies because the combination of different pollination systems, differences in main floral traits,

and increasing availability of genomic resources. Furthermore, Cleomaceae species are amenable to stable transformation (Newell *et al.* 2010) and VIGS techniques (Chapter 5), which favors its use in functional studies.

2.5.2. Specialist pollination system

The genus *Aquilegia* exhibits flowers with diverse morphology that have evolved in response to pollinator-mediated selection by different types of animals (Whittall and Hodges 2007). Evolutionary studies have demonstrated that variation in the length of the nectar spurs is due to pollinator-mediated directional selection, and that shifts in pollinator communities have occurred as a result of rapid changes in nectar spur lengths, leading to evolutionary switches from bees that select for short nectar spurs to hummingbirds and from hummingbirds to hawk moths, which select for longer nectar spurs (Whittall and Hodges 2007; Yant *et al.* 2015). Flowers also differ in perianth shape, colour and nectar traits, exhibiting classic examples of pollination syndromes (Hodges and Kramer 2007; Kramer 2009*a;* Whittall and Hodges 2007).

The genetics of flower development and petal color in *Aquilegia* have been well studied, and the genes involved in anthocyanin biosynthesis have been described (Hodges and Kramer 2007; Whittall *et al.* 2006). Most genetic research in *Aquilegia* involves the study of the ABC model of flower development to describe homologous and paralogous genes, with particular interest in genes involved in petal identity and staminodium development such as *APETALA* paralogs and *PISTILLATA* (Kramer 2009*b;* Kramer *et al.* 2007b). The genomic resources for *Aquilegia* are comprised of a big EST collection, QTLs for the main floral traits, and whole genome sequences for three species (Kramer 2009*a;* Whittall *et al.* 2006; Yant *et al.* 2015). *Aquilegia* species are also amenable to transient transformation using VIGS (Gould and Kramer 2007; Sharma and Kramer 2013). Because *Aquilegia* represents a highly specialized pollination system, studies can focus on the genetics of traits that restrict attraction to different pollinator guilds among lineages. This system may allow us to examine whether groups of traits associated with syndromes have additive or non-additive effects on pollination using a full-factorial approach to manipulate genotypes and phenotypes.

2.5.3. Sexually deceptive pollination system

Orchids (Orchidaceae) have highly specialized pollination systems (Schiestl and Schlüter 2009; Scopece *et al.* 2010). The astounding morphological diversity of floral traits in these plants is a frequent subject of study (Gaskett 2011; Jersáková *et al.* 2006; Schiestl 2005). The genus *Ophrys* represents a sexually deceptive pollination system, in which flowers attract bees by mimicking the female mating signals of their pollinators and pollination is the result of pseudocopulation by male bees (Sedeek *et al.* 2013; Xu *et al.* 2012). *Ophrys* is the only model system in which transcriptome and proteome data have been used as part of an integrative approach to studying pollination by looking at genes for different floral bouquets in closely related species. These bouquets attract different pollinators, leading to reproductive isolation (Sedeek *et al.* 2013; Xu *et al.* 2012). Through genetic manipulation and subsequent pollinator preference studies, Xu et al (2012) determined that *SADs* (STEAROYL-ACP DESATURASES) genes control the alkene double bonds that affect the carbon length chains of compounds produced in floral scent, leading to unique scent profiles that attract different pollinators (Xu *et al.* 2012). Thus, SADs genes are likely under strong pollinator-mediated selection. The highly specialized sexually deceptive pollination systems allow us to study which genes are involved in the attraction of specialized pollinators and explore which floral traits are more likely to experience strong directional selection.

2.6. Conclusions and future perspectives

An integrated approach that combines genomic tools with ecological and evolutionary experiments can greatly expand our understanding of plant-pollinator interactions. This approach provides new opportunities to study traits that are difficult to manipulate with classical methods, such as floral rewards and olfactory attractants. Using ‘omics’, we can ask more complex questions by focusing on genes controlling whole suites of traits or pollination systems in non-model species. We can also compare genes encoding for floral traits in pollination systems across a wide taxonomic range of plants, allowing us to identify new regulatory pathways and networks controlling key floral traits for pollinator attraction.

Despite its advantages, some challenges remain. For example, functional gene

studies are time-consuming and not always feasible. Furthermore, many plants are difficult to transform, although techniques like CRISPR may help resolve this issue. Genomic information is still limited for non-model plants, although the availability of plant genomes and transcriptomes is growing every day. Data analysis from the huge amount of DNA sequences generated from NGS approaches takes time and there is a steep learning curve associated with processing and interpreting data. Indeed, using 'omics' and genetic manipulation to identify, modify and validate the role of floral traits in plant-pollinator interactions does not mean that classical methods should be discounted. Rather, these approaches can be highly complementary.

Understanding the evolution of plant-pollinator interactions requires an interdisciplinary approach and is therefore an opportunity for collaborative studies. An integrative approach would take advantage of the experience from pollination biologists who can perform specific pollinator behavioral tests, eco-evo-devo biologists can study suite of traits and their evolution, and bioinformaticians who can analyze and synthesize the data derived from NGS. A multidisciplinary approach will allow us to better understand the eco-evolutionary processes underlying plant-pollinator interactions. Furthermore, integrative studies that involve genetic manipulations, NGS, and transcriptome analysis, have the potential to help us discover which genes are involved in the evolution of different pollination systems and syndromes. Finally, an integrative approach to studying pollination goes beyond describing patterns in floral trait variation to understanding the mechanisms that drive the underlying genetic basis of floral traits under pollinator-mediated selection. This approach, therefore allows us to explore how floral traits have evolved in response to specific pollinator taxa or to a community of pollinators, which will help to reveal how floral diversity is generated and maintained in natural populations.

2.7. References

Almeida J, Rocheta M, Galego L. 1997. Genetic control of flower shape in *Antirrhinum majus*. *Development* **124**: 1387-1392.

Altieri MA. 1999. The ecological role of biodiversity in agroecosystems. *Agriculture, Ecosystems & Environment* **74**: 19-31.

Anderson JT, Mitchell-Olds T. 2011. Ecological genetics and genomics of plant defenses: Evidence and approaches. *Functional Ecology* **25**: 312-324.

Anderson JT, Willis JH, Mitchell-Olds T. 2011. Evolutionary genetics of plant adaptation. *Trends in Genetics* **27**: 258-266.

Armbruster WS. 2012. Evolution and ecological implications of "specialized" pollinator rewards. In: Patiny S, ed. *Evolution of plant-pollinator relationships*. Cambridge; New York: Cambridge University Press, 44-67.

Baker HG, Baker I, Hodges SA. 1998. Sugar composition of nectars and fruits consumed by birds and bats in the tropics and subtropics. *Biotropica* **30**: 559-586.

Barker MS, Vogel H, Schranz ME. 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.

Baucom RS, Chang SM, Kniskern JM, Rausher MD, Stinchcombe JR. 2011. Morning Glory as a powerful model in ecological genomics: Tracing adaptation through both natural and artificial selection. *Heredity (Edinb)* **107**: 377-385.

Bayo-Canha A, Delgado-Benarroch L, Weiss J, Egea-Cortines M. 2007. Artificial decrease of leaf area affects inflorescence quality but not floral size in *Antirrhinum majus*. *Scientia Horticulturae* **113**: 383-386.

Becker A, Lange M. 2010. VIGS - genomics goes functional. *Trends in Plant Science* **15**: 1-4.

Bender R, Klinkenberg P, Jiang Z, Bauer B, Karypis G, Nguyen N, Perera MADN, Nikolau BJ, Carter CJ. 2012. Functional genomics of nectar production in the Brassicaceae. *Flora - Morphology, Distribution, Functional Ecology of Plants* **207**: 491-496.

Bey M, Stuber K, Fellenberg K, Schwarz-Sommer Z, Sommer H, Saedler H, Zachgo S. 2004. Characterization of *Antirrhinum* petal development and identification of target genes of the class B MADS box gene DEFICIENS. *Plant Cell* **16**: 3197-3215.

Bhide A, Schliesky S, Reich M, Weber A, 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.

Bortesi L, Fischer R. 2015. The CRISPR/Cas9 system for plant genome editing and beyond. *Biotechnology Advances* **33**: 41-52.

Bradshaw HD, Schemske DW. 2003. Allele substitution at a flower colour locus produces a pollinator shift in Monkeyflowers. *Nature* **426**: 176-178.

Brautigam A, Gowik U. 2010. What can next generation sequencing do for you? Next generation sequencing as a valuable tool in plant research. *Plant Biology* **12**: 831-841.

Brautigam A, Kajala K, Wullenweber J, Sommer M, Gagneul D, Weber KL, Carr KM, Gowik U, Mass J, Lercher MJ, Westhoff P, Hibberd JM, Weber AP. 2011a. An mRNA blueprint for C4 photosynthesis derived from comparative transcriptomics of closely related C3 and C4 species. *Plant Physiology* **155**: 142-156.

Braütigam A, Mullick T, Schliesky S, Weber AP. 2011b. 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. *Journal of Experimental Botany* **62**: 3093-3102.

Cane JH. 2008a. Breeding biologies, seed production and species-rich bee guilds of *Cleome lutea* and *Cleome serrulata* (Cleomaceae). *Plant Species Biology* **23**: 152-158.

Celik I, Gurbuz N, Uncu AT, Frary A, Doganlar S. 2017. Genome-wide SNP discovery and QTL mapping for fruit quality traits in inbred backcross lines (IBLs) of *Solanum pimpinellifolium* using genotyping by sequencing. *BMC Genomics* **18**: 1.

Chanderbali AS, Berger BA, Howarth DG, Soltis PS, Soltis DE. 2016. Evolving ideas on the origin and evolution of flowers: New perspectives in the genomic era. *Genetics* **202**: 1255-1265.

Cheng S, van den Bergh E, Zeng P, Zhong X, Xu J, Liu X, Hofberger J, de Bruijn S, Bhide AS, Kuelahoglu C, Bian C, Chen J, Fan G, Kaufmann K, Hall JC, Becker A, Brautigam A, Weber AP, Shi C, Zheng Z, Li W, Lv M, Tao Y, Wang J, Zou H, Quan Z, Hibberd JM, Zhang G, Zhu XG, Xu X, Schranz ME. 2013. The *Tarenaya hassleriana* genome provides insight into reproductive trait and genome evolution of crucifers. *Plant Cell* **25**: 2813-2830.

Chittka L, Kevan PG. 2005. Flower colour as advertisement. In: Dafni A, Kevan PG, Husband BC, eds. *Practical pollination biology*. Ontario, Canada: Enviroquest Ltd., 157 - 196.

Chittka L, Raine NE. 2006. Recognition of flowers by pollinators. *Current Opinion in Plant Biology* **9**: 428-435.

Clare EL, Schiestl FP, Leitch AR, Chittka L. 2013. The promise of genomics in the study of plant-pollinator interactions. *Genome Biology* **14**: 207.

Clegg MT, Durbin ML. 2000. Flower color variation: A model for the experimental study of evolution. *Proceedings of the National Academy of Sciences* **97**: 7016-7023.

Collier R, Fuchs B, Walter N, Kevin Lutke W, Taylor CG. 2005. *Ex vitro* composite plants: An inexpensive, rapid method for root biology. *The Plant Journal* **43**: 449-457.

Cronk QCB, Bateman RM, Hawkins JA. 2002. *Developmental genetics and plant evolution.* New York: Taylor & Francis.

Devaux C, Lepers C, Porcher E. 2014. Constraints imposed by pollinator behaviour on the ecology and evolution of plant mating systems. *Journal of Evolutionary Biology* **27**: 1413-1430.

Dornhaus A, Chittka L. 1999. Insect behaviour: Evolutionary origins of bee dances. *Nature* **401**: 38-38.

Dudareva N, Cseke L, Blanc VM, Pichersky E. 1996. Evolution of floral scent in *Clarkia*: Novel patterns of S-linalool synthase gene expression in the *C. breweri* flower. *The Plant Cell Online* **8**: 1137-1148.

Dudareva N, Raguso RA, Wang JH, Ross EJ, Pichersky E. 1998. Floral scent production in *Clarkia breweri* - III. Enzymatic synthesis and emission of benzenoid esters. *Plant Physiology* **116**: 599-604.

Dyer AG, Whitney HM, Arnold SEJ, Glover BJ, Chittka L. 2007. Mutations perturbing petal cell shape and anthocyanin synthesis influence bumblebee perception of *Antirrhinum majus* flower colour. *Arthropod-Plant Interactions* **1**: 45-55.

Edger PP, Heidel-Fischer HM, Bekaert M, Rota J, Glöckner G, Platts AE, Heckel DG, Der JP, Wafula EK, Tang M, Hofberger JA, Smithson A, Hall JC, Blanchette M, Bureau TE, Wright SI, dePamphilis CW, Eric Schranz M, Barker MS, Conant GC, Wahlberg N, Vogel H, Pires JC, Wheat CW. 2015. The butterfly plant arms-race

escalated by gene and genome duplications. *Proceedings of the National Academy of Sciences* **112**: 8362-8366.

Epperson BK, Clegg MT. 1992. Unstable white flower color genes and their derivatives in Morning Glory. *Journal of Heredity* **83**: 405-409.

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

Erhardt A, Rusterholz H-P. 1998. Do Peacock butterflies (*Inachis io* L.) detect and prefer nectar amino acids and other nitrogenous compounds? *Oecologia* **117**: 536-542.

Fenster CB, Armbruster WS, Wilson P, Dudash MR, Thomson JD. 2004. Pollination syndromes and floral specialization. *Annual Review of Ecology Evolution and Systematics* **35**: 375-403.

Fenster CB, Martén-Rodriguez S, Schemske DW. 2009. Pollination syndromes and the evolution of floral diversity in *Iochroma* (Solanaceae). *Evolution* **63**: 2758-2762.

Fernández VA, Galetto L, Astegiano J. 2009. Influence of flower functionality and pollination system on the pollen size-pistil length relationship. *Organisms Diversity & Evolution* **9**: 75-82.

Fernandez-Mazuecos M, Glover BJ. 2017. The evo-devo of plant speciation. *Nature Ecology & Evolution* **1**: 110.

Fleming TH, Geiselman C, Kress WJ. 2009. The evolution of bat pollination: A phylogenetic perspective. *Annals of Botany* **104**: 1017-1043.

Gallego Romero I, Pai AA, Tung J, Gilad Y. 2014. RNA-seq: impact of RNA degradation on transcript quantification. *BMC Biology* **12**: 42.

Garber M, Grabherr MG, Guttman M, Trapnell C. 2011. Computational methods for transcriptome annotation and quantification using RNA-seq. *Nature Methods* **8**: 469-477.

Gaskett AC. 2011. Orchid pollination by sexual deception: Pollinator perspectives. *Biological Reviews* **86**: 33-75.

Goldblatt P, Manning JC. 2000. The long-proboscid fly pollination system in Southern Africa. *Annals of the Missouri Botanical Garden* **87**: 146-170.

Gould B, Kramer EM. 2007. Virus-induced gene silencing as a tool for functional analyses in the emerging model plant *Aquilegia* (columbine, Ranunculaceae). *Plant Methods* **3**: 6.

Gübitz T, Hoballah ME, Dell'Olivo A, Kuhlemeier C. 2009. *Petunia* as a model system for the genetics and evolution of pollination syndromes. In: Gerats T, Strommer J, eds. *Petunia*. New York: Springer 29-49.

Gumbert A. 2000. Color choices by Bumble bees (*Bombus terrestris*): Innate preferences and generalization after learning. *Behavioral Ecology and Sociobiology* **48**: 36-43.

Hansen DM, Van der Niet T, Johnson SD. 2012. Floral signposts: Testing the significance of visual 'nectar guides' for pollinator behaviour and plant fitness. *Proceedings Biological Sciences* **279**: 634-639.

Harder LD, Johnson SD. 2005. Adaptive plasticity of floral display size in animal-pollinated plants. *Proceedings of the Royal Society B: Biological Sciences* **272**: 2651-2657.

Henson J, Tischler G, Ning Z. 2012. Next-generation sequencing and large genome assemblies. *Pharmacogenomics* **13**: 901-915.

Higuera-Díaz M, Manson JS, Hall JC. 2015. Pollination biology of *Cleomella serrulata* and *Polanisia dodecandra* in a protected natural prairie in southern Alberta, Canada. *Botany* **93**: 745-757.

Hileman L, Cubas P. 2009. An expanded evolutionary role for flower symmetry genes. *Journal of Biology* **8**: 90.

Hoballah ME, Gubitz T, Stuurman J, Broger L, Barone M, Mandel T, Dell'Olivo A, Arnold M, Kuhlemeier C. 2007. Single gene-mediated shift in pollinator attraction in *Petunia*. *Plant Cell* **19**: 779-790.

Hodges SA, Kramer EM. 2007. Columbines. *Current Biology* **17**: R992-R994.

Honaas LA, Wafula EK, Wickett NJ, Der JP, Zhang Y, Edger PP, Altman NS, Pires JC, Leebens-Mack JH, dePamphilis CW. 2016. Selecting superior *De Novo* transcriptome assemblies: Lessons learned by leveraging the best plant genome. *PLoS ONE* **11**: e0146062.

Horridge GA. 1996. The honeybee (*Apis mellifera*) detects bilateral symmetry and discriminates its axis. *Journal of Insect Physiology* **42**: 755-764.

Ichihashi Y, Tsukaya H. 2015. Behavior of leaf meristems and their modification. *Frontiers in Plant Science* **6**: 1060.

Iltis HH, Hall JC, Cochrane TS, Sytsma KJ. 2011b. Studies in the Cleomaceae I. On the separate recognition of Capparaceae, Cleomaceae, and Brassicaceae. *Annals of the Missouri Botanical Garden* **98**: 28-36.

Jersáková J, Johnson SD, Kindlmann P. 2006. Mechanisms and evolution of deceptive pollination in orchids. *Biological Reviews* **81**: 219-235.

Johnson SD, Steiner KE. 2000. Generalization versus specialization in plant pollination systems. *Trends in Ecology & Evolution* **15**: 140-143.

Karron JD, Ivey CT, Mitchell RJ, Whitehead MR, Peakall R, Case AL. 2012. New perspectives on the evolution of plant mating systems. *Annals of Botany* **109**: 493-503.

Kessler D, Baldwin IT. 2011. Back to the past for pollination biology. *Current Opinion in Plant Biology* **14**: 429-434.

Kessler D, Bhattacharya S, Diezel C, Rothe E, Gase K, Schottner M, Baldwin IT. 2012. Unpredictability of nectar nicotine promotes outcrossing by hummingbirds in *Nicotiana attenuata*. *Plant Journal* **71**: 529-538.

Kessler D, Diezel C, Baldwin IT. 2010. Changing pollinators as a means of escaping herbivores. *Current Biology* **20**: 237-242.

Kessler D, Gase K, Baldwin IT. 2008. Field experiments with transformed plants reveal the sense of floral scents. *Science* **321**: 1200-1202.

Klahre U, Gurba A, Hermann K, Saxenhofer M, Bossolini E, Guerin PM, Kuhlemeier C. 2011. Pollinator choice in *Petunia* depends on two major genetic loci for floral scent production. *Current Biology* **21**: 730-739.

Kooke R, Wijnker E, Keurentjes JJ. 2012. Backcross populations and near isogenic Lines. In: Rifkin SA, ed. *Quantitative Trait Loci (QTL): Methods and protocols*. New York: Springer, 3-16.

Kramer EM. 2009*a*. *Aquilegia*: A new model for plant development, ecology, and evolution. *Annual Review of Plant Biology* **60**: 261-277.

Kramer EM. 2009*b*. New model systems for the study of developmental evolution in plants. In: Jeffery WR, ed. *Current Topics in Developmental Biology*. Burlington: Academic Press, 67-105.

Kramer EM, Holappa L, Gould B, Jaramillo MA, Setnikov D, Santiago PM. 2007b. Elaboration of B gene function to include the identity of novel floral organs in the lower eudicot *Aquilegia*. *Plant Cell* **19**: 750-766.

Külahoglu C, Denton AK, Sommer M, Mass J, Schliesky S, Wrobel TJ, Berckmans B, Gongora-Castillo E, Buell CR, Simon R, De Veylder L, Brautigam A, Weber AP. 2014. Comparative transcriptome atlases reveal altered gene expression modules between two Cleomaceae C3 and C4 plant species. *Plant Cell* **26**: 3243-3260.

Kumar V, Jain M. 2015. The CRISPR–Cas system for plant genome editing: Advances and opportunities. *Journal of Experimental Botany* **66**: 47-57.

Lange M, Yellina AL, Orashakova S, Becker A. 2013. Virus-induced gene silencing (VIGS) in plants: An overview of target species and the virus-derived vector systems. In: Becker A, ed. *Virus-induced gene silencing: Methods and protocols*. Totowa, NJ: Humana Press, 1-14.

Lemmon E, Lemmon A. 2013. High-throughput genomic data in systematics and phylogenetics. *The Annual Review of Ecology, Evolution, and Systematics* **44**: 1-23.

Lewis AC. 1986. Memory constraints and flower choice in *Pieris rapae*. *Science* **232**: 863-865.

Leyser O. 2000. Mutagenesis. *Methods in Molecular Biology* **141**: 133-144.

Lin IW, Sosso D, Chen LQ, Gase K, Kim SG, Kessler D, Klinkenberg PM, Gorder MK, Hou BH, Qu XQ, Carter CJ, Baldwin IT, Frommer WB. 2014. Nectar secretion requires sucrose phosphate synthases and the sugar transporter SWEET9. *Nature* **508**: 546-549.

Lulin H, Xiao Y, Pei S, Wen T, Shangqin H. 2012. The first Illumina-based de novo transcriptome sequencing and analysis of Safflower flowers. *PLoS One* **7**: e38653.

Luo L, Zhu Y, Xiong M. 2012. Quantitative Trait Locus (QTL) analysis for Next-Generation Sequencing with the functional linear models. *Journal of medical genetics* **49**: 513-524.

Machado IC, Lopes AV, Leite AV, de Brito Neves C. 2006. *Cleome spinosa* (Capparaceae): Polygamodioecy and pollination by bats in urban and Caatinga areas, northeastern Brazil. *Botanische Jahrbücher* **127**: 69-82.

Maple J, Moller SG. 2007. Mutagenesis in *Arabidopsis*. *Methods in Molecular Biology* **362**: 197-206.

Martin LB, Fei Z, Giovannoni JJ, Rose JK. 2013. Catalyzing plant science research with RNA-seq. *Frontiers in Plant Science* **4**: 1-10.

Martin NH, Sapir Y, Arnold ML. 2008. The genetic architecture of reproductive isolation in Louisiana irises: Pollination syndromes and pollinator preferences. *Evolution* **62**: 740-752.

Martins DJ, Johnson SD. 2013. Interactions between hawkmoths and flowering plants in East Africa: Polyphagy and evolutionary specialization in an ecological context. *Biological Journal of the Linnean Society* **110**: 199-213.

Mayer C, Adler LS, Armbruster WS, Dafni A, Eardley C, Huang SQ, Kevan PG, Ollerton J, Packer L, Ssymank A, Stout JC, Potts SG. 2011. Pollination ecology in the 21st Century: Key questions for future research. *Journal of Pollination Ecology* **3**: 8-23.

Menzel R, Muller U. 1996. Learning and memory in Honeybees: From behavior to neural substrates. *Annual Review of Neuroscience* **19**: 379-404.

Metzker ML. 2010. Applications of Next-Generation Sequencing technologies - the next generation. *Nature Reviews Genetics* **11**: 31-46.

Møller AP. 1995. Bumblebee preference for symmetrical flowers. *Proceedings of the National Academy of Sciences* **92**: 2288-2292.

Møller AP. 2000. Developmental stability and pollination. *Oecologia* **123**: 149-157.

Mondragón-Palomino M. 2013. Perspectives on MADS-box expression during orchid flower evolution and development. *Frontiers in Plant Science* **4**: 377.

Neal PR, Dafni A, Giurfa M. 1998. Floral symmetry and its role in plant-pollinator systems: Terminology, distribution, and hypotheses. *Annual Review of Ecology and Systematics* **29**: 345-373.

Newell CA, Brown NJ, Liu Z, Pflug A, Gowik U, Westhoff P, Hibberd JM. 2010. *Agrobacterium tumefaciens*-mediated transformation of *Cleome gynandra* L., a C(4)

dicotyledon that is closely related to *Arabidopsis thaliana*. *Journal of Experimental Botany* **61**: 1311-1319.

Nozzolillo C, Amiguet VT, Bily AC, Harris CS, Saleem A, Andersen OM, 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.

Ó'Maoiléidigh DS, Graciet E, Wellmer F. 2014. Gene networks controlling *Arabidopsis thaliana* flower development. *New Phytologist* **201**: 16-30.

Ollerton J, Alarcón R, Waser NM, Price MV, Watts S, Cranmer L, Hingston A, Peter CI, Rotenberry J. 2009. A global test of the pollination syndrome hypothesis. *Annals of Botany* **103**: 1471-1480.

Ollerton J, Rech AR, Waser NM, Price MV. 2015. Using the literature to test pollination syndromes — Some methodological cautions. *Journal of Pollination Ecology* **16**: 119-125.

Ouborg NJ, Vriezen WH. 2007. An ecologist's guide to ecogenomics. *Journal of Ecology* **95**: 8-16.

Owen CR, Bradshaw HD. 2011. Induced mutations affecting pollinator choice in *Mimulus lewisii* (Phrymaceae). *Arthropod-Plant Interactions* **5**: 235-244.

Ozsolak F, Milos PM. 2011. RNA sequencing: Advances, challenges and opportunities. *Nature Reviews Genetics* **12**: 87-98.

Pashley CH, Ellis JR, McCauley DE, Burke JM. 2006. EST databases as a source for molecular markers: Lessons from *Helianthus*. *Journal of Heredity* **97**: 381-388.

Patchell MJ, Bolton MC, Mankowski P, Hall JC. 2011. Comparative floral development in Cleomaceae reveals two distinct pathways leading to monosymmetry. *International Journal of Plant Sciences* **172**: 352-365.

Pflieger S, Lefebvre V, Causse M. 2001. The candidate gene approach in plant genetics: A review. *Molecular Breeding* **7**: 275-291.

Policha T, Davis A, Barnadas M, Dentinger BT, Raguso RA, Roy BA. 2016. Disentangling visual and olfactory signals in mushroom-mimicking *Dracula* orchids using realistic three-dimensional printed flowers. *New Phytologist* **210**: 1058-1071.

Pootakham W, Jomchai N, Ruang-areerate P, Shearman JR, Sonthirod C, Sangsrakru D, Tragoonrung S, Tangphatsornruang S. 2015. Genome-wide SNP discovery and identification of QTL associated with agronomic traits in oil palm using genotyping-by-sequencing (GBS). *Genomics* **105**: 288-295.

Proctor MCF, Yeo P, Lack A. 1996. *The natural history of pollination.* Portland: Timber Press.

Quattrocchio F, Wing J, van der Woude K, Souer E, de Vetten N, Mol J, Koes R. 1999. Molecular analysis of the anthocyanin2 gene of *Petunia* and its role in the evolution of flower color. *Plant Cell* **11**: 1433-1444.

Raguso RA. 2008. Start making scents: The challenge of integrating chemistry into pollination ecology. *Entomologia Experimentalis et Applicata* **128**: 196-207.

Raju AJS, Rani DS. 2016. Reproductive ecology of *Cleome gynandra* and *Cleome viscosa* (Capparaceae). *Phytologia Balcanica* **22**: 15-28.

Rausher MD. 2008. Evolutionary transitions in floral color. *International Journal of Plant Sciences* **169**: 7-21.

Rausher MD, Fry JD. 1993. Effects of a locus affecting floral pigmentation in *Ipomoea purpurea* on female fitness components. *Genetics* **134**: 1237-1247.

Rifkin SA. 2012. *Quantitative trait loce (QTL): Methods and protocols.* New York: Springer.

Sapir Y. 2009a. Effects of floral traits and plant genetic composition on pollinator behavior. *Arthropod-Plant Interactions* **3**: 115-129.

Sapir Y. 2009b. Pollination genetics: Using molecular genetic underlying floral traits to study pollination ecology in an evolutionary context. *Israel Journal of Plant Sciences* **57**: 141-149.

Sapir Y, Armbruster WS. 2010. Pollinator-mediated selection and floral evolution: From pollination ecology to macroevolution. *New Phytologist* **188**: 303-306.

Schemske DW, Bradshaw HD. 1999. Pollinator preference and the evolution of floral traits in Monkey flowers (*Mimulus*). *Proceedings of the National Academy of Sciences of the United States of America* **96**: 11910-11915.

Schiestl FP. 2005. On the success of a swindle: Pollination by deception in orchids. *Naturwissenschaften* **92**: 255-264.

Schiestl FP, Johnson SD. 2013. Pollinator-mediated evolution of floral signals. *Trends in Ecology & Evolution* **28**: 307-315.

Schiestl FP, Schlüter PM. 2009. Floral isolation, specialized pollination, and pollinator behavior in orchids. *Annual Review of Entomology* **54**: 425-446.

Schmid-Hempel P, Speiser B. 1988. Effects of Inflorescence Size on Pollination in *Epilobium angustifolium*. *Oikos* **53**: 98-104.

Scopece G, Cozzolino S, Johnson Steven D, Schiestl Florian P. 2010. Pollination efficiency and the evolution of specialized deceptive pollination systems. *The American Naturalist* **175**: 98-105.

Sedeek KE, Qi W, Schauer MA, Gupta AK, Poveda L, Xu S, Liu ZJ, Grossniklaus U, Schiestl FP, Schluter PM. 2013. Transcriptome and proteome data reveal candidate genes for pollinator attraction in sexually deceptive orchids. *PLoS ONE* **8**: e64621.

Senthil-Kumar M, Mysore KS. 2011. New dimensions for VIGS in plant functional genomics. *Trends in Plant Science* **16**: 656-665.

Senthil-Kumar M, Mysore KS. 2014. Tobacco rattle virus–based virus-induced gene silencing in *Nicotiana benthamiana*. *Nature Protocols* **9**: 1549-1562.

Sharma B, Kramer EM. 2013. Virus-induced gene silencing in the rapid cycling columbine *Aquilegia coerulea* "Origami". In: Becker A, ed. *Virus-Induced Gene Silencing: Methods and Protocols*. Totowa, NJ: Humana Press, 71-81.

Sheehan H, Hermann K, Kuhlemeier C. 2012. Color and scent: How single genes influence pollinator attraction. *Cold Spring Harbor Symposia on Quantitative Biology* **77**: 117-133.

Smith SD. 2010. Using phylogenetics to detect pollinator-mediated floral evolution. *New Phytologist* **188**: 354-363.

Smith SD. 2016. Pleiotropy and the evolution of floral integration. *New Phytologist* **209**: 80-85.

Smith SD, Ane C, Baum DA. 2008. The role of pollinator shifts in the floral diversification of *Iochroma* (Solanaceae). *Evolution* **62**: 793-806.

Specht C, Yockteng R, Almeida A, Kirchoff B, Kress WJ. 2012. Homoplasy, pollination, and emerging complexity during the evolution of floral development in the tropical Gingers (Zingiberales). *The Botanical Review* **78**: 440-462.

Spitzer-Rimon B, Farhi M, Albo B, Cna'ani A, Ben Zvi MM, Masci T, Edelbaum O, Yu Y, Shklarman E, Ovadis M, Vainstein A. 2012. The R2R3-MYB-like regulatory factor EOBI, acting downstream of EOBII, regulates scent production by activating ODO1 and structural scent-related genes in *Petunia*. *Plant Cell* **24**: 5089-5105.

Stellari GM, Jaramillo MA, Kramer EM. 2004. Evolution of the APETALA3 and PISTILLATA lineages of MADS-box-containing genes in the basal angiosperms. *Molecular Biology and Evolution* **21**: 506-519.

Stratmann JW, Hind SR. 2011. Gene silencing goes viral and uncovers the private life of plants. *Entomologia Experimentalis Et Applicata* **140**: 91-102.

Strickler SR, Bombarely A, Mueller LA. 2012. Designing a transcriptome next-generation sequencing project for a nonmodel plant species. *American Journal of Botany* **99**: 257-266.

Swarbreck D, Wilks C, Lamesch P, Berardini TZ, Garcia-Hernandez M, Foerster H, Li D, Meyer T, Muller R, Ploetz L, Radenbaugh A, Singh S, Swing V, Tissier C, Zhang P, Huala E. 2008. The *Arabidopsis* Information Resource (TAIR): Gene structure and function annotation. *Nucleic Acids Research* **36**: D1009-D1014.

Ungerer MC, Johnson LC, Herman MA. 2007. Ecological genomics: Understanding gene and genome function in the natural environment. *Heredity* **100**: 178-183.

van den Bergh E, Hofberger JA, Schranz ME. 2016a. Flower power and the mustard bomb: Comparative analysis of gene and genome duplications in glucosinolate biosynthetic pathway evolution in Cleomaceae and Brassicaceae. *American Journal of Botany* **103**: 1212-1222.

van der Niet T, Peakall R, Johnson SD. 2014. Pollinator-driven ecological speciation in plants: New evidence and future perspectives. *Annals of Botany* **113**: 199-212.

van Dijk EL, Auger H, Jaszczyszyn Y, Thermes C. 2014. Ten years of next-generation sequencing technology. *Trends in Genetics* **30**: 418-426.

Verdonk JC, Haring MA, van Tunen AJ, Schuurink RC. 2005. ODORANT1 regulates fragrance biosynthesis in *Petunia* flowers. *Plant Cell* **17**: 1612-1624.

Wang L, Li PH, Brutnell TP. 2010a. Exploring plant transcriptomes using ultra high-throughput sequencing. *Briefings in Functional Genomics* **9**: 118-128.

Wang X, Cairns MJ. 2013. Gene set enrichment analysis of RNA-Seq data: Integrating differential expression and splicing. *BMC Bioinformatics* **14**: 1-10.

Wang Z, Gerstein M, Snyder M. 2009. RNA-Seq: a revolutionary tool for transcriptomics. *Nature Reviews Genetics* **10**: 57-63.

Ward JA, Ponnala L, Weber CA. 2012. Strategies for transcriptome analysis in nonmodel plants. *American Journal of Botany* **99**: 267-276.

Waser NM, Chittka L, Price MV, Williams NM, Ollerton J. 1996. Generalization in pollination systems, and why it matters. *Ecology* **77**: 1043-1060.

Wege S, Scholz A, Gleissberg S, 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. *Annals of Botany* **100**: 641-649.

Weigel D, Glazebrook J. 2002. *Arabidopsis: A laboratory manual.* New York: Cold Spring Harbor Laboratoty Press.

Wellmer F, Riechmann JL. 2010. Gene networks controlling the initiation of flower development. *Trends in Genetics* **26**: 519-527.

Wessinger CA, Hileman LC, Rausher MD. 2014. Identification of major quantitative trait loci underlying floral pollination syndrome divergence in *Penstemon*. *Philosophical Transactions of the Royal Society B: Biological Sciences* **369**.

Whibley AC, Langlade NB, Andalo C, Hanna AI, Bangham A, Thébaud C, Coen E. 2006. Evolutionary paths underlying flower color variation in *Antirrhinum*. *Science* **313**: 963-966.

Whittall JB, Hodges SA. 2007. Pollinator shifts drive increasingly long nectar spurs in Columbine flowers. *Nature* **447**: 706-709.

Whittall JB, Voelckel C, Kliebenstein DJ, Hodges SA. 2006. Convergence, constraint and the role of gene expression during adaptive radiation: Floral anthocyanins in *Aquilegia*. *Molecular Ecology* **15**: 4645-4657.

Willmer P. 2011. *Pollination and floral ecology.* Princeton: Princeton University Press.

Wilson P, Castellanos MC, Hogue JN, Thomson JD, Armbruster WS. 2004. A multivariate search for pollination syndromes among Penstemons. *Oikos* **104**: 345-361.

Xu SQ, Schluter PM, Grossniklaus U, Schiestl FP. 2012. The genetic basis of pollinator adaptation in a sexually deceptive orchid. *Plos Genetics* **8**.

Yant L, Collani S, Puzey J, Levy C, Kramer EM. 2015. Molecular basis for three-dimensional elaboration of the *Aquilegia* petal spur. *Proceedings of the Royal Society of London B: Biological Sciences* **282**.

Yuan YW, Byers KJRP, Bradshaw HD. 2013*a*. The genetic control of flower-pollinator specificity. *Current Opinion in Plant Biology* **16**: 422-428.

Yuan YW, Saqawa JM, Young RC, Christensen BJ, Bradshaw HDJ. 2013*b*. Genetic dissection of a major anthocyanin QTL contributing to pollinator-mediated reproductive isolation between sister species of *Mimulus*. *Genetics* **194**: 255-263.

Zhang W, Kramer EM, Davis CC. 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.

Zou Z, Ishida M, Li F, Kakizaki T, Suzuki S, Kitashiba H, Nishio T. 2013. QTL analysis using SNP markers developed by Next-Generation Sequencing for identification of candidate genes controlling 4-Methylthio-3-Butenyl Glucosinolate contents in roots of radish, *Raphanus sativus* L. *PLoS ONE* **8**: e53541.

Table 2.1. Plant models in alphabetic order used to address the genetic basics of floral traits and their influence in pollinator behaviour.

Taxon	Floral traits	Method	Identified genes	'Omics' resources	Major references
Antirrhinum majus (Snapdragon)	- Petal colour - Petal epidermal cells shape	- Candidate gene - Recombinant inbred lines (RILs)	- NIVEA - MIXTA	- ESTs	Almeida *et al.* 1997; Bey *e*t *al*. 2004; Comba êt al. 2000; Dyer *et al.* 2007; Whibley *et at.* 2006
Clarkia	- Floral scents	- Artificial hybrids	- LIS	- ESTs	Dudareva *et al.* 1996, 1998
Helianthus annuus (Common Sunflower)	- Flowering phenology - Floral disk area - Height of the primary flowering head	- RILs	- Disk area, flowering date, and height QTLs	- ESTs - 80% sequenced genome -Floral/root/stem transcriptome	Pashley *et al.* 2006; Sapir 2009a
Ipomoea purpurea (Morning Glory)	- Petal colour	- Candidate gene	- F3'5'H	- ESTs	Baucom *et al.* 2011, Epperson and Clegg 1992; Rausher and Fry 1993
Iris brevicaulis and	- Anther extension	- Artificial	- Flower QTLs	- ESTs	Martin *et al.*

Iris fulva (Zigzag Iris and Copper Iris)	- Flowering phenology - Flower stalk height - Nectar guide area - Sepal blade chroma, brightness and hue - Sepal length - Stylar branch length	hybrids	morphology/color	- Floral/leaf transcriptome	2008
Mimulus cardinalis and *M. lewisii* (Scarlet Monkey Flower and Lewis' Monkey Flower)	- Nectar guides - Nectar volume - Petal anthocyanin and carotenoids concentrations - Petal colour patterns - Petal reduction - Trichomes	- Chemical mutagenesis - Genetic transformation - Inbreeding lines - Near isogenic lines (NILs)	- ANS QTLs - ROI1 - R2R3 - YUP	- ESTs - Corolla transcriptome - Whole genome sequence	Bradshaw and Schemske 2003; Owen and Bradshaw 2011, Yuan *et al.* 2013b
Nicotiana attenuata (Wild Tobacco)	- Floral scents	- Candidate gene - Genetic transformation	- NaCHAL1/2 - NaPMT1/2	- ESTs	Kessler *et al.* 2008, 2010, 2012

Penstemon	- Petal colour	- Artificial hybrids	- DFR - F3'H - F3'5'H - Flower morphology/color nectar volume/concentration QTLs	- Whole genome shotgun sequence	Wessinger *et al.* 2014
Petunia	- Floral scents - Petal colour	- Candidate gene - Genetic transformation - Near isogenic lines - Recombinant inbreeding lines	- AN2 - EOBI/II - ODO1	- ESTs - Flower/seedling transcriptome	Hoballah *et al.* 2007; Klahre *et al.* 2011; Quattrocchio *et al.* 1999; Spitzer-Rimon *et al.* 2012; Verdonk *et al.* 2005

Figure 2.1. Ecology and evo-devo pollination studies and how they can be used and linked to perform integrative pollination studies, in which we can study evolution and variation of floral traits, the genes encoding floral traits, and their effect in pollinator attraction.

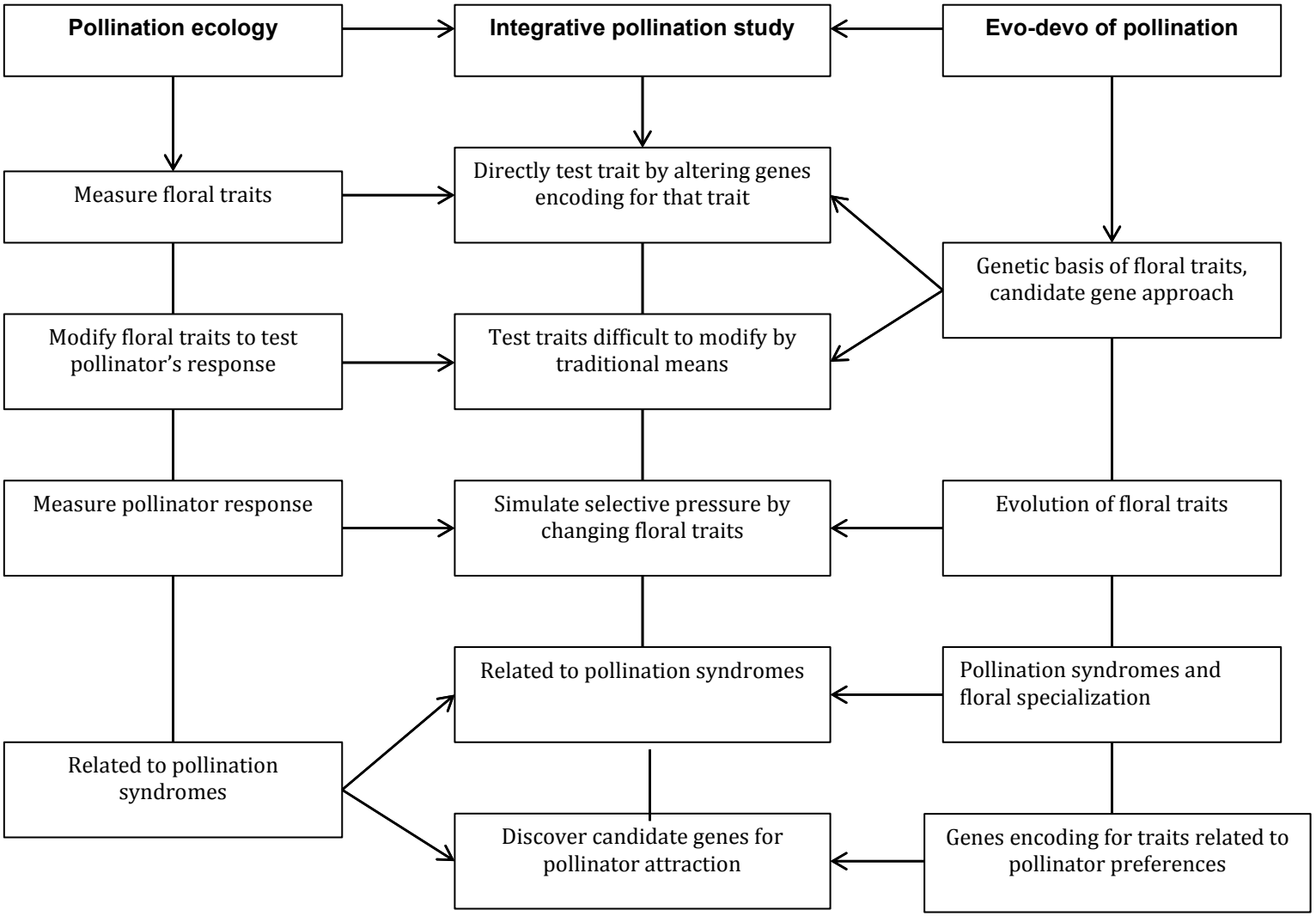
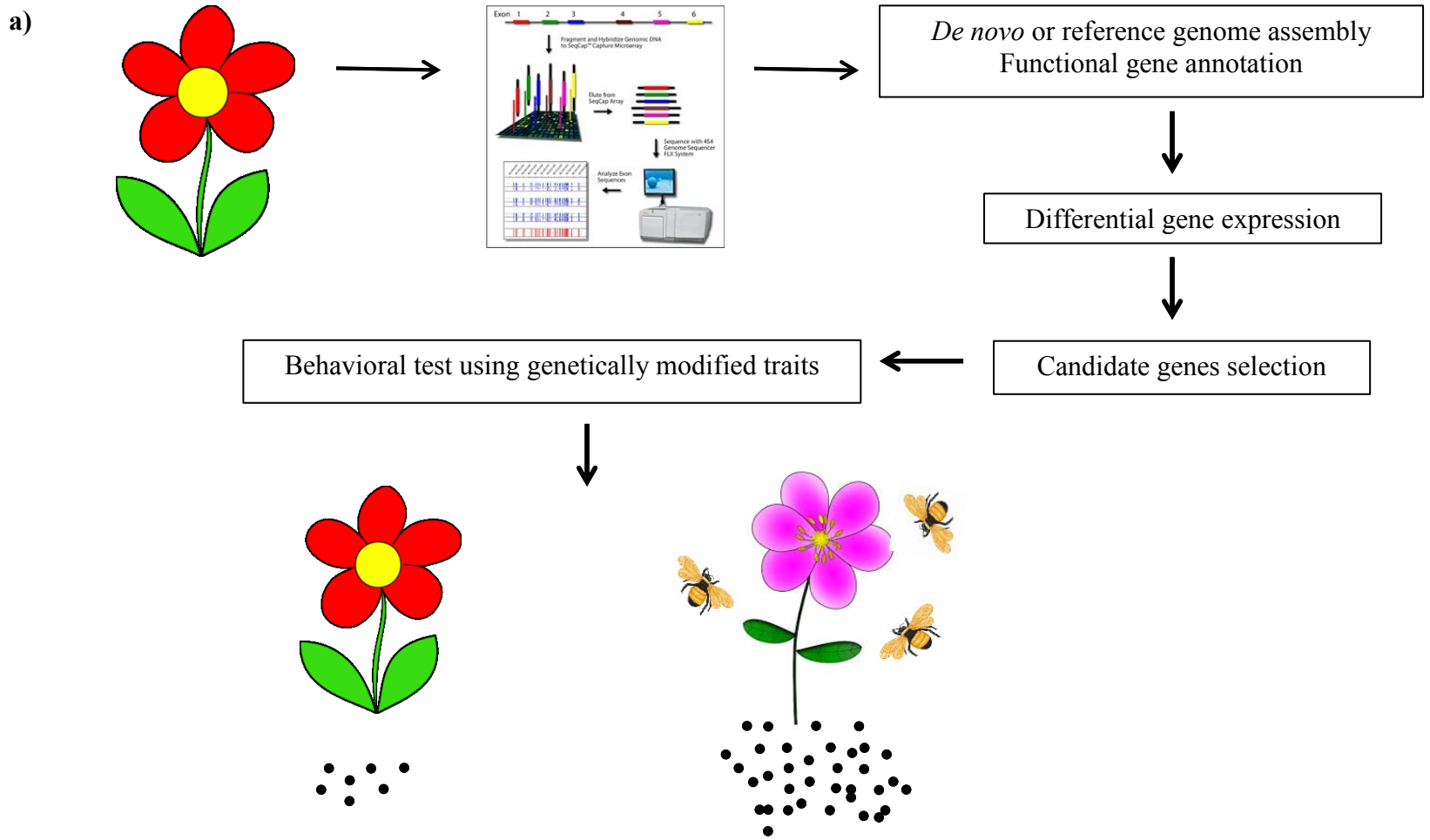
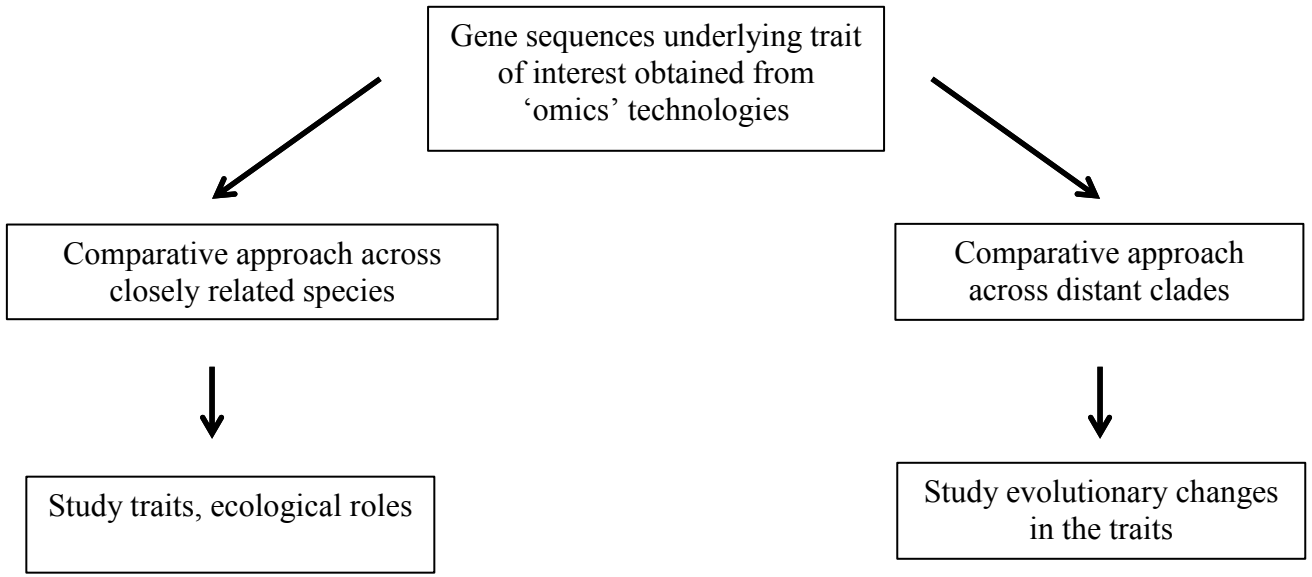


Figure 2.2. a). A pipeline for modifying floral traits using next generation sequencing, candidate gene approach, and forward or reverse genetics techniques to modify the floral trait to perform pollinator preference tests. **b).** Using 'omics' to study ecological role of traits in closely related taxa and to study evolutionary changes in traits in distant clades.



b)


Gene sequences underlying trait of interest obtained from 'omics' technologies
Comparative approach across closely related species
Comparative approach across distant clades
Study traits, ecological roles
Study evolutionary changes in the traits

Chapter 3: Pollination biology of *Cleomella serrulata* and *Polanisia dodecandra* in a protected natural prairie in Southern Alberta, Canada

3.1. Introduction

Animal-mediated pollination is a process of great biological importance for the reproduction of plants and is an essential provisioning ecosystem service in which pollinators increase seed set of crops and native plants (Klein *et al.* 2007; Proctor *et al.* 1996; Willmer 2011). The availability of floral resources is critical to the survival of pollinators, which in turn increases the chances of cross-pollination in plants (Free 1993). Native plants play a crucial role in contributing to the biodiversity of natural ecosystems and also provide food, forage resources and nesting sites for different populations of pollinators. Thus, maintaining sustainable native plant populations and their associated pollinator communities benefits wild and managed ecosystems (Garibaldi *et al.* 2014). Unfortunately, many native plants and their pollinators are facing threats including reduction and loss of habitat and the introduction of invasive species. This is particularly true of plants found in native prairies, which are one of the most threatened ecosystems in North America (Fuhlendorf *et al.* 2002; Samson and Knopf 1994; Sheffield *et al.* 2014).

In prairie ecosystems, native flowering forbs and shrubs provide diverse foraging resources for pollinators. Although dominant native plant species in prairies are wind pollinated (Kevan 1999; Moldenke 1976), shrub and forb species frequently exhibit open and non-specialized flowers that are visited by a broad spectrum of pollinators (Hegland and Totland 2005; Kevan 1999). This event is known as a generalist pollination system (Hegland and Totland 2005; Herrera 1996; Waser *et al.* 1996). Annual plant species tend to have generalized pollination systems because relying on specialist pollinators, which may experience annual fluctuations in abundance, is risky (Waser *et al.* 1996). Studying native prairie plants with generalized pollination systems provides valuable information about the availability of floral resources to pollinators in an ecosystem where the scarcity of flowers could limit pollinator populations. In Alberta, two rare native species of Cleomaceae persist in the southeast part of the province and exhibit flowers consistent with generalist syndrome (Cane 2008a, b; Smith *et al.* 2010).

Cleomaceae is a cosmopolitan plant family comprising 18 genera and 150 to 200 species (Patchell *et al.* 2014). Cleomaceae species have been used to explore variation in petal pigments (Nozzolillo *et al.* 2010), flower and fruit morphology (Iltis 1957; Iltis 1958), and the evolution of floral symmetry (Patchell *et al.* 2011). The flowers exhibit great variation in morphological traits associated with attractants and rewards for pollinators including flower size and symmetry, petal colour, ultraviolet light reflection, floral fragrances, nectar glands, and pollen production (Cane 2008a; Erbar and Leins 1997; Iltis *et al.* 2011b; Nozzolillo *et al.* 2010). However, the pollination biology of most species belonging to this family is poorly understood.

Cleomella serrulata (Pursh) E.H. Roalson and J.C. Hall (Beeweed, previously known as *Cleome serrulata* – hereafter referred to as *Cleomella*) and the rare in Alberta *Polanisia dodecandra* (L.) DC. (Clammyweed – hereafter referred to as *Polanisia*) are found in Alberta prairies. Although *Cleomella* is currently listed as a secure species in Alberta, its distribution appears to have substantially decreased over the years based on herbarium records. In contrast, *Polanisia* is listed as a rare species in Alberta and is vulnerable to local extirpation (Wallis 2001). Both species exhibit a mixed-mating system, in which plants can reproduce through self or cross-fertilization (Cane 2008a; Cruden and Lyon 1985; Wiens 1984). *Cleomella* was reported to attract pollinators from a wide variety of wild bee guilds and is also used extensively by managed pollinators, including *Megachile rotunda* F. and *Apis mellifera* L. in the United States (Cane 2008a). In addition, *Cleomella* seeds are sold and planted in the peripheries of agriculture areas, orchards, gardens and seed crops to promote pollinator populations. Despite the importance of *Cleomella* and *Polanisia* as forage for a number of bee species, information on the pollination biology of these plants is restricted to breeding biology of *Cleomella* (Cane 2008a) and morphological description for *Polanisia* (Iltis 1958). Although both species exhibit the same floral symmetry pattern, they differ in key morphological traits for pollinator attraction such as petal colour, inflorescence size and nectar production. Thus, we might expect variation in the communities of insects visiting both species and differences in the frequency of visits, because these flowers do not exhibit the same floral traits. Their floral morphology suggests that they may be pollinated by insects, probably bees and butterflies, but it is uncertain which species of insects are effective pollinators.

The major goal of this study is to determine the pollination biology for wild populations of *Cleomella* and *Polanisia* in southern Alberta. To achieve this, we (i) described population densities and distributions of both plant species; (ii) determined their floral phenology; (iii) established time of stigma receptivity and anthers dehiscence; and (iv) identified main flower visitors and their activity periods.

3.2. Materials and Methods

3.2.1. Study site

This research was conducted during the summers of 2013 and 2014 at Suffield National Wildlife Area (Suffield), which comprises the largest remaining tract of native prairie in Alberta (458 km^2). We found two populations of *Cleomella* in the Casa Berardi area (50°14'02.3" N 110°39'27.8" W, 668 m) and the Komati area (50°26'09.9" N 110°31'36.6" W, 707 m). Due to logistical issues, the Komati population was only sampled during 2013. We found one population of *Polanisia* in the Fish Creek area (50°14'36.9" N 110°39'42.5" W, 682 m). The plants of both populations of *Cleomella* were growing in dry-mixed grassland on the roadside, while *Polanisia* plants were growing on sandy South Saskatchewan River bluffs (Appendix 3.1).

3.2.2. Description of focal taxa

Cleomella serrulata is an annual herb. Leaves are alternate and trifoliate with slender elliptic leaflets (Moss and Packer 1983). Bracteate racemes contain numerous flowers and are located in the main stem and lateral branches (Figure 3.1). The inflorescences may have two kinds of hermaphroditic flowers with completely or incompletely developed pistils. At the beginning of the flowering season new flowers fail to fully develop pistils, but late in season inflorescences present both kinds of flowers (Cane 2008a). Flowers are monosymmetric; they have four green-reddish lanceolate sepals, four purple or soft pink petals rounded at the tip, with adaxial petals shorter than the abaxial ones (Figure 3.2). Flowers have six stamens longer than the petals, and a long bicarpellate gynoecium on a gynophore. An adaxial nectary is located between the petal and androecium whorls (Figure 3.1d); it is green, elongated and laminar with a three-toothed apex, and produces a drop of viscous nectar (Figure 3.1c). Fruits are cylindrical-oblong capsules with wrinkled ovoid blackish seeds (Moss and Packer 1983). In Alberta, the plants

are distributed in mixed grasslands, pastures, roadsides, and stabilized sand dunes (Moss and Packer 1983).

Polanisia dodecandra is also an annual herb. The stems vary from unbranched to basally branched, and are pubescent and densely covered by stipitate glands that produce a strong disagreeable smell (Iltis 1958; Moss and Packer 1983). Leaves are alternate and trifoliate with ovolanceolated leaflets. Bracteate racemes contain numerous flowers and are located in the main stem and lateral branches (Figure 3.3). Inflorescences have hermaphroditic flowers (Figure 3.3). Flowers are monosymmetric; they have four purple-greenish sepals and four white or white-pinkish petals notched at the tip, with adaxial petals longer than the abaxial ones (Figure 3.4). Flowers have 8 to 12 stamens variable in length, and a long bicarpellate gynoecium on a short gynophore (Figure 3.4). An adaxial nectary is located between the petal and androecium whorls (Figure 3.3d); it is obliquely truncated and concave in the apex, orange or reddish, and produces a drop of highly viscous nectar (Figure 3.3c). Fruits are erect, oblong capsules containing subspherical brownish seeds (Moss and Packer 1983; Tucker and Vanderpool 2010). In Alberta, the plants are distributed in stream and river banks, dunes, and roadsides, generally growing in dry and sandy soils (Moss and Packer 1983).

3.2.3. Flowering phenology

Cleomella starts flowering in mid-late June and continues to flower until late August. The study performed by Cane (2008a) in Utah, found the flowers exhibited nocturnal anthesis and they opened 1 – 3 h after sunset. Flowers presented complete anther dehiscence and receptive stigmas 8 h after anthesis.

There are no published studies of *Polanisia* flowering phenology. To describe the floral phenology of *P. dodecandra*, we recorded time and duration of anthesis, time of stigma receptivity, and anther dehiscence for 20 randomly selected plants. We evaluated stigma receptivity in a greenhouse experiment, examining 2 stigmas per plant (n=20). Plants were grown at 24 °C in long day conditions corresponding to 16 hours light and 8 hours dark (Department of Biological Sciences, University of Alberta). We tested stigma receptivity from excised pistils each day via a peroxidase test (Dafni 1992; Kearns and Inouye 1993). In the field, we recorded number of inflorescences per plant, number of flowers per inflorescences, fruit production per plant, and the length of flowering period.

Although the mating system of *Cleomella* (Cane (2008a) and *Polanisia* (Wiens (1984) were previously described, we attempted pollination limitation experiments using exclusion bags in 2013. Because the flowers are hermaphroditic we had to antherectomize flowers before anthesis to avoid self-pollination. However, most of the flowers died in the field due to mechanical damage. These experiments were not pursued further.

3.2.4. Nectar volume and sugar concentration

To determine the nectar volume available for pollinators and nectar sugar concentration, we performed floral visitor exclusion experiments in 2014. To exclude floral visitors, we covered inflorescences with mesh bags for 24 h prior to sampling. For each of the two species, we randomly selected 50 plants with large buds and covered the terminal inflorescences. We sampled nectar production during two consecutive sunny days and nectar samples were taken after 24 hours of nectar accumulation. For *Cleomella*, nectar volume and sugar concentration were sampled between 08:00 and 16:00 h. After 16:00 h most of the nectar had evaporated due to high temperatures, so no more sampling was conducted. For *Polanisia*, nectar volume and sugar concentration were measured from 08:00 h to 22:00 h in two-hour intervals to determine if there are variations in nectar production rate and sugar concentration throughout the day. Floral nectar was extracted with a 10 μl microcapillary tube. For each species, we calculated average nectar volume for 10 flowers per plant on 5 plants. Once the nectar volume was depleted the flowers did not produce more nectar. Thus, we used different flowers at every two-hour interval. We took nectar measurements from flowers closest to the inflorescence meristem following a clockwise direction and always used flowers at the same developmental stage (flowers open for one day after anthesis). Due to the small volume produced by each flower it was not possible to measure sugar concentration per flower, thus, the nectar from 10 flowers was pooled to measure sugar concentration. Nectar volume was calculated following the method described by (Dafni 1992), and Kearns and Inouye (1993). Nectar volume was then diluted 1:3 in distilled water due to the high nectar viscosity. After dilution, the sugar concentration was measured by a 0 – 30% refractometer (Extech Instruments). The values obtained from the refractometer were multiplied by the dilution factor to calculate the correct nectar concentration.

3.2.5. Pollinator observations and visitation rates

In 2013, we performed a pilot study to collect and identify pollinator taxa visiting flowers. After experts helped verify taxonomic identifications, we made a visual guide with images of all taxa and used this to identify subsequent visitors. Every time that we collected a new taxon it was added to the visual guide. Male and female Hymenoptera often differ in their morphology and we included examples of each sex when possible to ensure that individuals were not identified as separate taxa.

To determine the insect visitation rates and their activity periods, we selected 20 groups of five plants for both *Cleomella* and *Polanisia*. The focal area of observed plants was approximately 2 m^2. Two people, who were familiar with insect taxonomy, performed the pollinator observations using the visual guide. Each person sat next to an adjacent plant group and recorded all insect visitors observed in 10-minute intervals and then we switched to another group. We observed the visitors for 20 hours from 06:00 to 02:00 h. To perform the nocturnal observations, we wore red headlights to illuminate the flowers. We did not perform observations between 02:00 to 06:00 h because of cool morning temperatures and scarcity of insect visitors. We recorded the total number of visits and the number of inflorescences visited. In addition, we described whether insects were foraging for nectar or pollen and the number of flowers and inflorescences visited. We collected representative plant vouchers and insect specimens for taxonomic identification, the latter of which included at least one specimen per observed taxon (241 specimens/2219 observed). Vouchers specimens were deposited in the University of Alberta Vascular Plant Herbarium (ALTA) and E. H Strickland Entomological Museum, respectively.

3.2.6. Data analysis

We compared morphological characteristics of *Cleomella* and *Polanisia* using R (R Development Core Team 2013) to determine if observed differences were significant. Specifically, we compared plant height, inflorescences per plant and flowers per inflorescences, fruit set, nectar volume, and nectar sugar concentration between the two species. For normally distributed variables, we performed a Student's t-test to compare trait differences between species in different years. For data that was not normally distributed, we performed a Mann-Whitney U test. We also performed the tests between years to determine if there was significant annual variation in the measured traits.

We calculated pollinator richness as the total number of taxa collected during the sampling period. To compare the difference in sampling effort between taxa and between years according to the number of observations obtained, we created a rarefaction curve for each plant species using Biodiversity Pro version 2 (McAleece *et al.* 1997).

To compare the pollinator community composition between plant species and sampling years, we performed a qualitative similarity analysis in Biodiversity-Pro version 2 (McAleece *et al.* 1997). We made a presence-absence matrix of arthropod visitors per plant population and sampling year and we performed a Jaccard coefficient analysis. The obtained matrix was grouped using UPGMA (Unweighted Pair Group Method with Arithmetic Mean).

3.3. Results

3.3.1. Description of *Cleomella* and *Polanisia* populations

Both populations of *Cleomella* were in native mixed grassland growing along roadsides in densely aggregate clusters. The Komati population comprised ~600 plants, while in Casa Berardi we found ~500 plants in 2013 and ~300 plants in 2014; thus, the population size in Casa Berardi decreased by 40% from 2013 to 2014. Plant height ranged from 16 cm to 80 cm; inflorescences per plant ranged from 1 to 64; flower number per inflorescence ranged from 2 to 301; and fruit set per plant ranged from 0 to 164. In 2014, the Casa Berardi population had taller plants and more inflorescences per plant compared to the Casa Berardi and Komati 2013 populations, while the 2013 Casa Berardi population produced the highest fruit set relative to Casa Berardi 2014 and Komati populations. The Komati plant population had the most flowers per inflorescence across sites (Table 3.1). Fruit set was significantly different between Casa Berardi 2013 and Komati populations (Appendix 3.2), although other morphological variables, including flower number, did not differ. However, plant height, inflorescences per plant, flowers per inflorescence and fruit set were significantly different between Casa Berardi 2013 and 2014 populations, which could be due to weather changes in the sampling periods (Appendix 3.2 – 3.3).

Polanisia plants were in dispersed clusters along river bluffs. The population consisted of ~2500 plants in 2013, but only ~600 plants in 2014, representing a remarkable reduction in population size of approximately 76%. Plant height ranged from 4 cm to 58

cm; inflorescences number per plant ranged from 1 to 27; flower number per inflorescence ranged from 1 to 51; and fruit set per plant ranged from 1 to 103. In 2013, the population had, on average, taller plants, more inflorescences per plant, more flowers per inflorescence and a higher fruit set when compared to 2014 (Table 3.1). Plant height, inflorescences per plant, flower per inflorescence and fruit set were significantly different in Fish Creek in 2013 compared to 2014 (Appendix 3.2 – 3.3).

When comparing between species within years, all the variables differed significantly except the fruit set between Komati *Cleomella* population and *Polanisia* 2013 (Appendix 3.2 – 3.3).

3.3.2. Flowering phenology

Cleomella started flowering in early or late July. For both populations, flowering persisted until late August, and individual flowers were open for 5 days on average (Figure 3.1a–b). New flowers exhibited nocturnal anthesis, opening between 23:00 – 02:00 h; anthers released the pollen and stigmas became receptive 6 – 8 hours after anthesis. At the beginning of the flowering season, most flowers had a short and presumably non-functional pistil. Therefore, these flowers did not produce any fruits. As the flowering season advanced, plants started to produce two kinds of flowers: (i) flowers with short and non-functional pistils reaching 1/5 of the stamens length and (ii) flowers with long and fully-functional pistils that reach the stamens length (Figure 3.1d). When the plants had a high number of fruits, they again started producing primarily flowers with non-functional pistils.

Polanisia started flowering in late June or early July and the length of the flowering period was two months. Inflorescences persisted until late August and individual flowers lasted 7 days on average (Figure 3.3a–b). Flowers started to open shortly after the sunset between 23:00 – 01:00 h; however, the stigma became receptive 4 days on average before anthesis. Exerted pistils presented in floral buds were receptive after two days of bud development, with continuous receptivity during anthesis. The anthers started gradually releasing pollen after anthesis and the longest stamens were first to exhibit anther dehiscence, while anthers of shorter stamens were last to release pollen. This plant also produced flowers with short and non-functional pistils at the beginning of the flowering season, but, in contrast with *Cleomella* later in season, plants produced only flowers with long and completely functional pistils. Interestingly, some plants produced two kinds of

completely distorted fruits: (i) fruits twisted over the longitudinal axis, and (ii) fruits with three or four sides instead of just two sides. These abnormal fruits did not produce viable seeds and in all of the cases seeds were aborted.

3.3.3. Nectar volume and sugar concentration

Cleomella flowers have a green nectar gland that started producing a viscous nectar drop two days before anthesis (Figure 3.1c–d). The nectar drops persisted approximately 15 h after anthesis from 01:00 – 16:00 h, but usually evaporated due to high daily temperatures or was removed by insects. Once the nectar evaporated or was removed, flowers did not secrete more nectar; however, during the flowering season nectar was always available for pollinators because new flowers opened daily (Figures 3.1c–d). Covered flowers produced the highest nectar volume of 0.85 ± 0.96 µl between 08:00 – 10:00 h; however, variation in nectar volume production between flowers was high. After 10:00 h, nectar production decreased 68% on average, with nectar volume of 0.28 ± 0.08 µl and the variation between flowers was low (Figure 3.5a). Sugar concentration was relatively steady across time, but it was higher between 10:00 – 12:00 h (Figure 3.5b). In general, the average of total sugar contents in floral nectar was 40.8 ± 7.68 %.

Polanisia flowers have an orange nectar gland (Figure 3.3c–d) that produced a highly viscous nectar drop, which was available to pollinators during anthesis (Figure 3.3d). The nectar drops persisted for approximately 48 h after anthesis. Covered flowers produced the highest nectar volume, 0.63 ± 0.32 µl between 10:00 – 12:00 h; however, variation in nectar volume production between flowers was high. After 14:00 h nectar production decreased 60% on average, the nectar volume was 0.25 ± 0.06 µl and the variation between flowers was low (Figure 3.5c). The sugar concentration was steady between 08:00 – 14:00 h and it decreased between 14:00 – 16:00 h; interestingly the sugar concentration increased to a maximum peak between 20:00 – 22:00 h corresponding to 43.32 ± 4.74 % (Figure 3.5d). The sugar concentration differed significantly between species, while the nectar volume did not (Appendix 3.2).

3.3.4. Pollinator observations and visitation rates

We observed 2219 insect visitors corresponding to 8 orders, 48 families and 150 taxa visiting the two-plant species for the total survey period (Appendix 4) Supplementary Table S3). For the *Cleomella* population in Komati, we recorded 85 taxa and 442 visits,

while in Casa Berardi we recorded 76 taxa and 276 visits in 2013, and 96 taxa and 1327 visits in 2014. Thus, visitor richness and insect visits increased 20.8% and 79.2% respectively in 2014 (Appendix 3.4 – 3.5). For *Polanisia*, we recorded 44 taxa and 81 visits in 2013, while we recorded 32 taxa and 93 visits in 2014. Furthermore, richness of visitors was 27.3% higher in 2013, while the insect visits were 12.9% higher in 2014 (Appendix 3.4 – 3.5). The rarefaction analysis indicated that we did not get the number expected of pollinator species for either plant population based on the obtained rarefaction curves, which were far from the asymptote except for *Cleomella* Casa Berardi 2014 population (Appendix 3.6).

The average visitation rate in *Cleomella* populations was 4.52 insects per plant per hour (Appendix 3.4 – 3.5). We found no significant differences in the number of pollinator visits between Casa Berardi and Komati populations in 2013 (Appendix 3.2). However, pollinator visitation number was significantly different between Casa Berardi in 2013 and 2014 (Appendix 3.2). Hymenoptera had the highest richness in both sampling periods followed by Diptera and Lepidoptera (Figure 3.6a–b). Within Hymenoptera, bee families presented the highest richness and Formicidae the highest number of visits. The most frequent visitors were *Ammophila azteca* (sphecid wasp), *Bombus borealis*, and *B. ternarius* (Appendix 3.4). Most Apidae visited the plants for pollen but occasionally collected nectar too, while other hymenopterans visited the plants for nectar and transferred pollen passively between flowers. *Bombus borealis*, one of the most frequent pollinators, represented 32.29% of overall visits. *Bombus borealis* started visiting the flowers at 07:00 h, and they exhibited higher visitation rates at 09:00 h and 14:00h. After 15:00 h the bees did not visit the plants at all (Appendix 3.7). In contrast, *B. ternarius*, another common visitor, started visiting the flowers at 06:00 h and represented 38.61% of overall visits (Appendix 3.4). Bumblebees had three high frequency visitation periods: 08:00 h, 10:00 h and 21:00 h, and the frequency of visits decreased rapidly between 12:00 h and 19:00 h (Appendix 3.7). Both bumblebee species were visiting plants for nectar and pollen, touched anthers and pistils and flew between plants, acting as potential pollen vectors. In addition, we observed the first nocturnal visitors for *Cleomella*, which includes four Noctuidae taxa, a crambid and an erebid moth (Appendix 3.4). We collected *Synnervus plagiatus* (sand wasp), which has not previously been recorded in Alberta, and *Bombus griseocolis*, only the

second individual collected in Alberta (Appendix 3.4), and only recently reported in the province by Sheffield *et al.* (2014). We also observed that Formicidae taxa feeding on nectar and pollen as well as residing in plants so they could occasionally act as pollination vectors. Furthermore, we found *Phymata americana* (ambush bug), and Thomisidae spiders (Appendix 3.4). These taxa, which were not likely to be effective pollinators, actively chased away other flower visitors and preyed upon syrphid flies, sphecid and pompilid wasps, small bees, and butterflies.

The average visitation rate for *Polanisia* was 1.41 per plant per hour (Appendix 3.4 – 3.5). We did not find significant differences in pollinator visit number between sampling periods (Appendix 3.2). Hymenoptera had the highest richness and number of visits in 2013 followed by Diptera, Lepidoptera and Coleoptera (Figure 3.7a–b). In contrast, Diptera and Hymenoptera had the highest richness in 2014, while Diptera had the highest number of visits (Figure 3.7a–b). Within Hymenoptera, bee families presented the highest richness. The most frequent taxa were *Microbembex monodonta* (sand wasp), *Perdita* sp. 1, and Syrphidae sp. 4 (Appendix 3.5). *Microbembex monodonta* represented 8.33% of overall visits and started visiting the plants at 08:00 h, with highest activity at 09:00 h, but their frequency of visits decreased after 13:00h (Appendix 3.5). Syrphidae sp. 4, another common pollinator, represented 18.75% of overall visits. Hoverflies visited flowers from 09:00 h, with peak visitation rates at 10:00 h and 13:00 h; after 14:00 h the visit activity ceased completely (Appendix 3.8). Sand wasps and hoverflies were feeding on nectar and pollen from different plants, acting as potential vectors for cross-pollination. We also found ants feeding on nectar and pollen, although their frequency was lower when compared with ants on *Cleomella* (Appendix 3.4).

When we compared the pollinator visitor composition between species across sampling periods, the Jaccard similarity analysis showed that populations clustered according to species (Figure 3.8). Jaccard index values range from 0 to 100, where 0 indicated no taxa shared between communities and 100 represent identical communities. The similarity index of *Polanisia* populations was 26.67%, while, *Cleomella* Casa Berardi and Komati populations clustered together and the similarity index was 47.27%. The similarity values indicated that *Polanisia* populations had fewer taxa in common when compared with *Cleomella* populations, which shared more taxa.

3.4. Discussion

3.4.1. Distribution of focal species

Native prairies are one of the most endangered ecosystems in North America due to anthropogenic activity (Fuhlendorf *et al.* 2002; Mlot 1990; Samson and Knopf 1994). Much of the natural vegetation has been gradually replaced by tame pastures or by crop production, leading to a loss in the biodiversity of native flora and fauna (Fuhlendorf *et al.* 2002; Kevan 1999; Samson and Knopf 1994). More than 70% of Canada native prairies were transformed into rangelands or crops before the 1990s (Sheffield *et al.* 2014). In Alberta, Suffield NWA is the largest remnant of native prairie, but grasses are the dominant vegetation type, while forbs are scarce. Native flowering plants such as *Cleomella* and *Polanisia* may therefore represent an important source of forage for pollinators. Despite many potential habitats for *Cleomella* in Alberta, it is restricted to remnant grasslands, disturbed sites and roadsides (Moss and Packer 1983). The plants form a densely aggregated cluster, which is visited by a broad spectrum of pollinators. On the other hand, *Polanisia* is located on the western and northern limit of its distribution and although the species also forms densely aggregated clusters, the rate of visits by pollinators was 92% lower when compared with *Cleomella*. The study population of *Polanisia* was growing on a steep slope in an unstable sandy area, and was subjected to strong winds, which could deter accessibility of these plants by insects. Fluctuations in population size could be the result of annual variation in environmental factors, including the availability of foraging resources for pollinators.

Both *Cleomella* and *Polanisia* grow in disturbed areas or in areas where colonization by other plant species is difficult. These species have been used for restoration projects on overgrazed lands, for soil stabilization in disturbed areas, and to rehabilitate rangelands after fires in western regions of North America (Cane 2008b; Smith *et al.* 2010). But once the soil has been restored, these plants are naturally replaced by more competitive forbs (Cane 2008a, b; Smith *et al.* 2010). According to herbarium records, years ago both plants had a wider distribution in Alberta, but the populations were small and localized. However, after visiting the locations where plants had been previously reported, the populations were no longer there. Further ecological studies are needed to determine if

competitive exclusion or habitat losses are a determining factor in the scarcity of *Cleomella* and *Polanisia* populations in Alberta.

3.4.2. Flowering phenology and mating systems of focal species

Both *Cleomella* and *Polanisia* have a mixed-mating system, exhibiting floral characteristics to attract pollinators and promote cross-pollination, as well as strategies to avoid self-pollination (Cane 2008a; Cruden and Lyon 1985; Wiens 1984). For example, *Polanisia* stigma is receptive days prior to the release of pollen from anthers. The low rate of insect visits and the location of plants could be favoring self-pollination in this population. However, selfing in *Polanisia* may result in a high rate of inbreeding in the population, as suggested by several plants that produced fruits with abnormal morphology and unviable seeds. Poor seed production may be contributing to *Polanisia*'s relative rarity compared to *Cleomella*.

Both species produce staminate flowers early in the flowering period. The production of staminate flowers is a mechanism to differentially allocate resources to male and female functions, which allows plants to control the timing of fruit initiation and production (Lloyd and Bawa 1984; Miller and Diggle 2007). In *Cleomella*, once the plants produce many fruits, they start to produce staminate flowers again, while *Polanisia* plants switch to producing hermaphrodite flowers later in the season. *Polanisia* plants produced fewer inflorescences as well as fewer and fruits per inflorescences, and fruits when compared with *Cleomella*. In addition, the flowers received a low visitation rate by pollinators. Thus, later in season the plants likely produce more fruits to ensure their reproductive success.

3.4.3. Focal species produce high nectar volume and moderately concentrated nectar sugars

Even though the nectar gland of *Cleomella* is inconspicuous, it produces a substantial volume of nectar for pollinators. The highest nectar volume was produced between 08:00 and 10:00 h, which is consistent with the high frequency of pollinators. Most visits were conducted from 07:00 until 13:00 h, when frequency of visits decreased due to high temperatures, although plants continued receiving insect visits at lower frequencies until late at night. Sugar content in nectar was steady in *Cleomella*. The moderate sugar concentrations produced by the flowers attract a wide spectrum of

pollinators and Hymenoptera and Diptera taxa were the most common visitors for this plant.

Despite the conspicuous nectar gland and the UV reflectance patterns in the nectar, *Polanisia* received few insect visits during the sampling period. The highest nectar volume was produced between 10:00 – 12:00, which is directly related with the number of insect visits performed at this time. Surprisingly, the highest sugar concentration was found between 20:00 – 22:00, but no crepuscular or nocturnal insects were observed visiting the plants at this time. This may be due to the plant's distribution; they are growing on the Western and northern boundary of the species' range, which may be inconsistent with the optimal pollinator community.

3.4.4. Pollinator communities are diverse and differ between species

Cleomella and *Polanisia* attracted a wide spectrum of visitors, the most abundant and diverse being from Hymenoptera, Diptera, and Lepidoptera. The focal species shared most of the visitor taxa including bees, wasps, butterflies and moths. Although six taxa visited exclusively *Polanisia,* overall *Cleomella* attracted more insect species than *Polanisia.* This difference may be due primarily to the pink petal color, the higher number of open flowers, and the greater number of flowers forming inflorescences of *Cleomella*, which exhibit a bigger display than *Polanisia*. Species exhibiting a bigger floral display can attract a greater number of visitors when compared with similar species with smaller floral displays (Galloway *et al.* 2002; Klinkhamer and de Jong 1990; Schmid-Hempel and Speiser 1988). Most of the taxa visiting both plants are generalist pollinators including polylectic bees such as halictid bees and bumblebees. However, we also observed oligolectic bees such as *Andrena* and *Perdita* species. *Perdita* species have been reported as oligolectic bees for *Cleomella* (Cane 2008a) and these species collected pollen actively in both *Cleomella* and *Polanisia*, but we did not observe them in the surrounding plants. There are 298 bees species reported for Alberta prairies (Sheffield *et al.* 2014) and bees visiting *Cleomella* and *Polanisia* account for 8.38% of the bee fauna reported for Alberta. Our results therefore suggest that these plants contribute to supporting a large and variable community of pollinators.

Perhaps unsurprising based on morphology and distribution, the focal species differ in key floral traits for pollinator attraction. *Cleomella* petal colour, which varies from pink

to purple, might make them more visible to pollinators. Surprisingly, although the nectar of *Polanisia* flowers produces UV reflection and the nectar gland is vividly colored, the visitation rate was lower than *Cleomella* flowers. Nectar volume and sugar content were higher in *Cleomella* than in *Polanisia*, which correlates with the higher visitation frequencies observed in *Cleomella.* However, the location of the populations might also have an effect in visitation frequencies. *Cleomella* plants were located in sites that gave pollinators easy access, while *Polanisia* plants were located in a river bluff subject to strong winds that make pollinator access difficult. In addition, differences between the visitation frequencies of both species could be due to other factors that we did not consider here such as pollen limitation, floral fragrances, flower constancy, and the presence of other floral species that could be competing or facilitating pollinator visits.

Two pollinators that were predicted to visit *Cleomella* and *Polanisia* were conspicuously absent. *Bombus occidentalis* has been recorded in the area in past studies (Hobbs 1968) and was at one time very common. However, *B. occidentalis* populations are declining in Western North America, likely due to pathogen infections and habitat fragmentation (Cameron *et al.* 2011), and the species has recently been listed as endangered in Canada. We also recorded only two visits of *Apis mellifera* to the plants. The low frequency of visits performed by *A. mellifera* is likely due to the distance from where the managed hives are kept by the beekeepers. Overall, the bee fauna visiting *Cleomella* in Alberta is slightly less rich than that reported for populations in Utah, United States (Cane 2008a, b). Our results suggest that *Cleomella* and *Polanisia* provide important food resources that help support a rich community of pollinators including different bee guilds, flies, butterflies and other insects. Although, flowers of both species are anthetic at night, we did not observe a high visitation frequency of nocturnal pollinators. Nocturnal anthesis is also presented in *Cleomella lutea* (Hook) E.H. Roalson and J.C. Hall (Cane 2008a) and *Cleome spinosa* (Dafni *et al.* 1987). Interestingly, crepuscular and nocturnal moths visited *Cleomella*, but their frequency was lower when compared with diurnal pollinators and they were active only between 11:00 – 01:00. After 01:00 the temperature decreased and no floral visitors were observed. Cane (2008a), suggested that nocturnal anthesis could be an ancestral trait in Cleomaceae species, but further evolutionary studies are needed to test this

interesting assumption. Finally, it is possible that the observed differences in pollinator assemblages were due to the different habitats where the species were located.

3.4.5. Implications for agriculture and conservation

Generalist native plants that attract a diverse range of pollinators are essential for the conservation of wild pollinators, particularly in habitats where flowers are a limited resource. In addition, native plants provide much needed dietary diversity for managed pollinators and are an important alternative food resource for honey bees in crop margins (Garibaldi *et al.* 2014). *Cleomella* and *Polanisia* may be good candidate species to increase floral resources in prairie agroecosystems because they grow in disturbed areas and could be planted in field margins or along tracks and roads next to crop areas to support managed and native pollinator communities. Furthermore, *Cleomella* and *Polanisia* could be used in restoration projects after fires or for industrial use in Alberta prairies and grasslands, since they are ruderal plants, and grow well in disturbed areas (Cane 2008a, b; Smith *et al.* 2010). Therefore, *Cleomella* and *Polanisia* could be used to promote and maintain the diversity of wild pollinators not only in range and croplands, but also in parklands, orchards and gardens, thereby providing an essential ecosystem service.

Pollinators are threatened by habitat transformation in many ecosystems, but prairie pollinators in North America are particularly at risk because anthropogenic disturbance is high (Kevan 1999). Promoting the conservation of native plants in natural grasslands will support biodiversity of pollinators and ensure they continue to provide pollination services to natural and agricultural ecosystems (Altieri 1999; Duelli *et al.* 1999; Kevan 1999).

3.5. References

Altieri MA. 1999. The ecological role of biodiversity in agroecosystems. *Agriculture, Ecosystems & Environment* **74**: 19-31.

Cameron SA, Lozier JD, Strange JP, Koch JB, Cordes N, Solter LF, Griswold TL. 2011. Patterns of widespread decline in North American Bumble bees. *Proceedings of the National Academy of Sciences* **108**: 662-667.

Cane JH. 2008a. Breeding biologies, seed production and species-rich bee guilds of *Cleome lutea* and *Cleome serrulata* (Cleomaceae). *Plant Species Biology* **23**: 152-158.

Cane JH. 2008b. Pollinating bees crutial to farming wildflower seed for U.S. habitat restoration. In: James RR, Pitts-Singer TL, eds. *Bee pollination in agricultural ecosystems*. New York: Oxford University Press, 48 - 64.

Cruden RW, Lyon DL. 1985. Patterns of biomass allocation to male and female functions in plants with different mating systems. *Oecologia* **66**: 299-306.

Dafni A. 1992. *Pollination ecology: a practical approach.* Oxford: IRL Press at Oxford University Press.

Dafni A, Eisikowitch D, Ivri Y. 1987. Nectar flow and pollinators' efficiency in two co-occurring species of *Capparis* (Capparaceae) in Israel. *Plant Systematics and Evolution* **157**: 181-186.

Duelli P, Obrist MK, Schmatz DR. 1999. Biodiversity evaluation in agricultural landscapes: Above-ground insects. *Agriculture, Ecosystems & Environment* **74**: 33-64.

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

Free JB. 1993. *Insect pollination of crops.* London: Academic Press.

Fuhlendorf SD, Zhang H, Tunnell TR, Engle DM, Cross AF. 2002. Effects of grazing on restoration of southern mixed Prairie soils. *Restoration Ecology* **10**: 401-407.

Galloway LF, Cirigliano T, Gremski K. 2002. The contribution of display size and dichogamy to potential geitonogamy in *Campanula americana*. *International Journal of Plant Sciences* **163**: 133-139.

Garibaldi LA, Carvalheiro LG, Leonhardt SD, Aizen MA, Blaauw BR, Isaacs R, Kuhlmann M, Kleijn D, Klein AM, Kremen C, Morandin L, Scheper J, Winfree R. 2014. From research to action: Enhancing crop yield through wild pollinators. *Frontiers in Ecology and the Environment* **12**: 439-447.

Hegland SJ, Totland Ø. 2005. Relationships between species' floral traits and pollinator visitation in a temperate grassland. *Oecologia* **145**: 586-594.

Herrera CM. 1996. Floral traits and plant adaptation to insect pollinators: A devil's advocate approach. In: Lloyd DG, Barrett SCH, eds. *Floral biology*. Boston, MA: Springer, 65-87.

Hobbs GA. 1968. Ecology of species of *Bombus* (Hymenoptera: Apidae) in southern Alberta: VII. Subgenus bombus. *The Canadian Entomologist* **100**: 156-164.

Iltis HH. 1957. Studies in the Capparidaceae. III. Evolution and phylogeny of the western North American Cleomoideae. *Annals of the Missouri Botanical Garden* **44**: 77-119.

Iltis HH. 1958. Studies in the Capparidaceae—IV. *Polanisia* Raf. *Brittonia* **10**: 33-58.

Iltis HH, Hall JC, Cochrane TS, Sytsma KJ. 2011b. Studies in the Cleomaceae I. On the separate recognition of Capparaceae, Cleomaceae, and Brassicaceae. *Annals of the Missouri Botanical Garden* **98**: 28-36.

Kearns CA, Inouye DW. 1993. *Techniques for pollination biologists.* Colorado: University Press of Colorado.

Kevan PG. 1999. Pollinators as bioindicators of the state of the environment: Species, activity and diversity. *Agriculture, Ecosystems & Environment* **74**: 373-393.

Klein A-M, Vaissière BE, Cane JH, Steffan-Dewenter I, Cunningham SA, Kremen C, Tscharntke T. 2007. Importance of pollinators in changing landscapes for world crops. *Proceedings of the Royal Society B: Biological Sciences* **274**: 303-313.

Klinkhamer PG, de Jong TJ. 1990. Effects of plant size, plant density and sex differential nectar reward on pollinator visitation in the protandrous *Echium vulgare* (Boraginaceae). *Oikos*: 399-405.

Lloyd DG, Bawa KS. 1984. Modification of the gender of seed plants in varying conditions. *Evolutionary Biology* **17**: 255-338.

McAleece N, Gage JDG, Lambshead PJD, Paterson GLJ. 1997. BioDiversity Professional statistics analysis software.

Miller JS, Diggle PK. 2007. Correlated evolution of fruit size and sexual expression in andromonoecious *Solanum* sections *Acanthophora* and *Lasiocarpa* (Solanaceae). *American journal of botany* **94**: 1706-1715.

Mlot C. 1990. Restoring the prairie. *BioScience* **40**: 804-809.

Moldenke AR. 1976. California pollination ecology and vegetation types. *Phytologia* **34**.

Moss EH, Packer JG. 1983. *Flora of Alberta: A manual of flowering plants, conifers, ferns, and fern allies found growing without cultivation in the Province of Alberta, Canada.* Toronto; Buffalo: University of Toronto Press.

Nozzolillo C, Amiguet VT, Bily AC, Harris CS, Saleem A, Andersen OM, 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.

Patchell MJ, Bolton MC, Mankowski P, Hall JC. 2011. Comparative floral development in Cleomaceae reveals two distinct pathways leading to monosymmetry. *International Journal of Plant Sciences* **172**: 352-365.

Patchell MJ, Roalson EH, Hall JC. 2014. Resolved phylogeny of Cleomaceae based on all three genomes. *Taxon* **63**: 315-328.

Proctor MCF, Yeo P, Lack A. 1996. *The natural history of pollination.* Portland: Timber Press.

R Development Core Team. 2013. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing.

Samson F, Knopf F. 1994. Prairie conservation in North America. *BioScience* **44**: 418-421.

Schmid-Hempel P, Speiser B. 1988. Effects of Inflorescence Size on Pollination in *Epilobium angustifolium*. *Oikos* **53**: 98-104.

Sheffield CS, Frier SD, Dumesh S. 2014. The bees (Hymenoptera: Apoidea, Apiformes) of the Prairies Ecozone, with comparisons to other Grasslands of Canada. In: Giberson DJ, Carcamo HA, eds. *Arthropods of Canadian Grasslands*. Canada: Biological Survey of Canada, 427-467.

Smith FS, Maywald PD, Lloyd-Reilley J, Maher SD, Pawelek KA, Scott Jr AW, Garza J. 2010. Zapata germplasm rio grande Clammyweed. *Native Plants Journal* **11**: 269-273.

Tucker GC, Vanderpool SS. 2010. Cleomaceae. In: Committee FNAE, ed. *Flora of North America North of Mexico*. New York: Oxford University Press, 199-223.

Wallis C. 2001. *Proposed protocols for inventories of rare plants of the Grassland Natural Region.* Edmonton: Alberta Sustainable Resource Development, Fish and Wildlife Service.

Waser NM, Chittka L, Price MV, Williams NM, Ollerton J. 1996. Generalization in pollination systems, and why it matters. *Ecology* **77**: 1043-1060.

Wiens D. 1984. Ovule survivorship, brood size, life history, breeding systems,and reproductive success in plants. *Oecologia* **64**: 47-53.

Willmer P. 2011. *Pollination and floral ecology.* Princeton: Princeton University Press.

Table 3.1. Means and standard deviations of plant height, flowering, and fruit set measurements for *Cleomella serrulata* and *Polanisia dodecandra* at Casa Berardi, Komati, and Fish Creek area. Data collected in 2013 and 2014 (N=100 per site).

	Plant height (cm)	**Inflorescences per plant**	**Flowers per inflorescence**	**Fruit set (per plant)**
C. serrulata Casa Berardi 2013	43.85 ± 10.7	9.07 ± 5.58	6.65 ± 5.39	50.78 ± 40.56
C. serrulata Casa Berardi 2014	47.29 ± 13.85	13.35 ± 11.96	20.08 ± 12.61	27.52 ± 32.69
C. serrulata Komati 2013	41.2 ± 19.2	9.15 ± 10.8	22.18 ± 14.08	28.37 ± 32.64
P. dodecandra Fish Creek 2013	35.3 ± 8.51	3.98 ± 3.09	14.44 ± 6.73	22.84 ± 13.61
P. dodecandra Fish Creek 2014	17.1 ± 7.96	2.23 ± 1.95	11.28 ± 5.38	14.52 ± 14.6

Figure 3.1. Photos of *Cleomella serrulata* (A) Terminal raceme inflorescences bearing abundant small flowers. (B) *Bombus ternarius* downside position and *Formica* sp. visiting flowers. (C) Nectar droplet. (D) Flowers with petals removed to expose the nectary at the base. Scale bars: 1 mm.

Figure 3.2. *Cleomella serrulata* flowers. (A) Frontal view. (B) Lateral view. Scale bar: 1 cm.

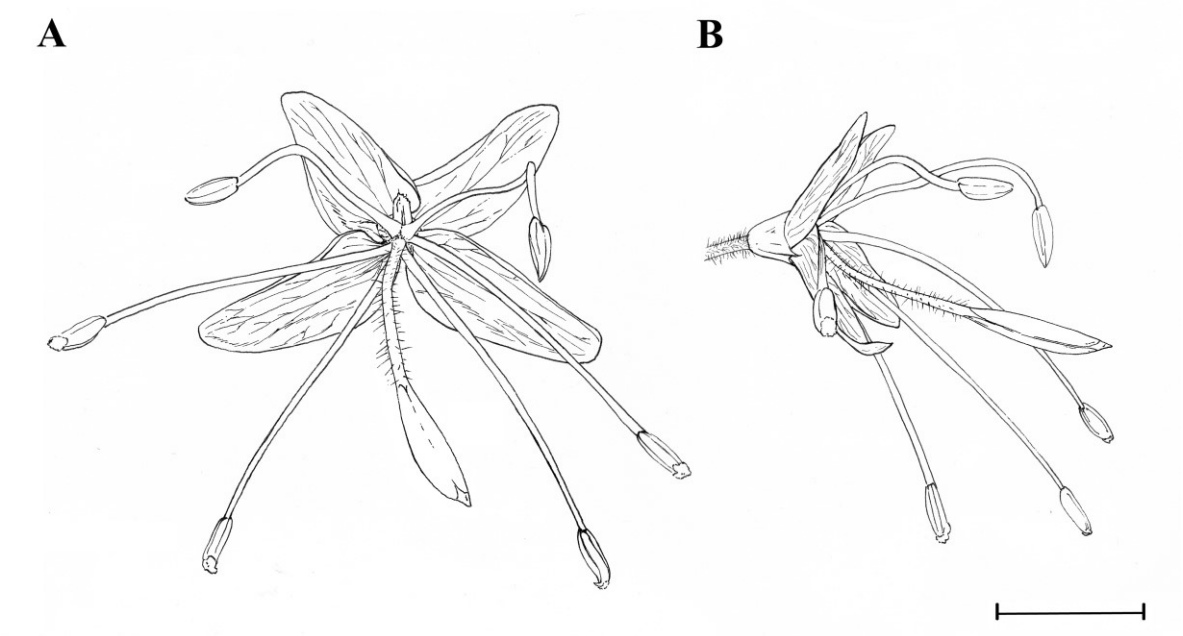


Figure 3.3. Photos of *Polanisia dodecandra*. (A) Terminal raceme inflorescences bearing abundant small flowers. (B) *Lasioglossum* sp. collecting pollen from the anther. (C) Nectar droplet. (D) Flowers with petals removed to expose the nectary at the base. Scale bars: 1 mm.

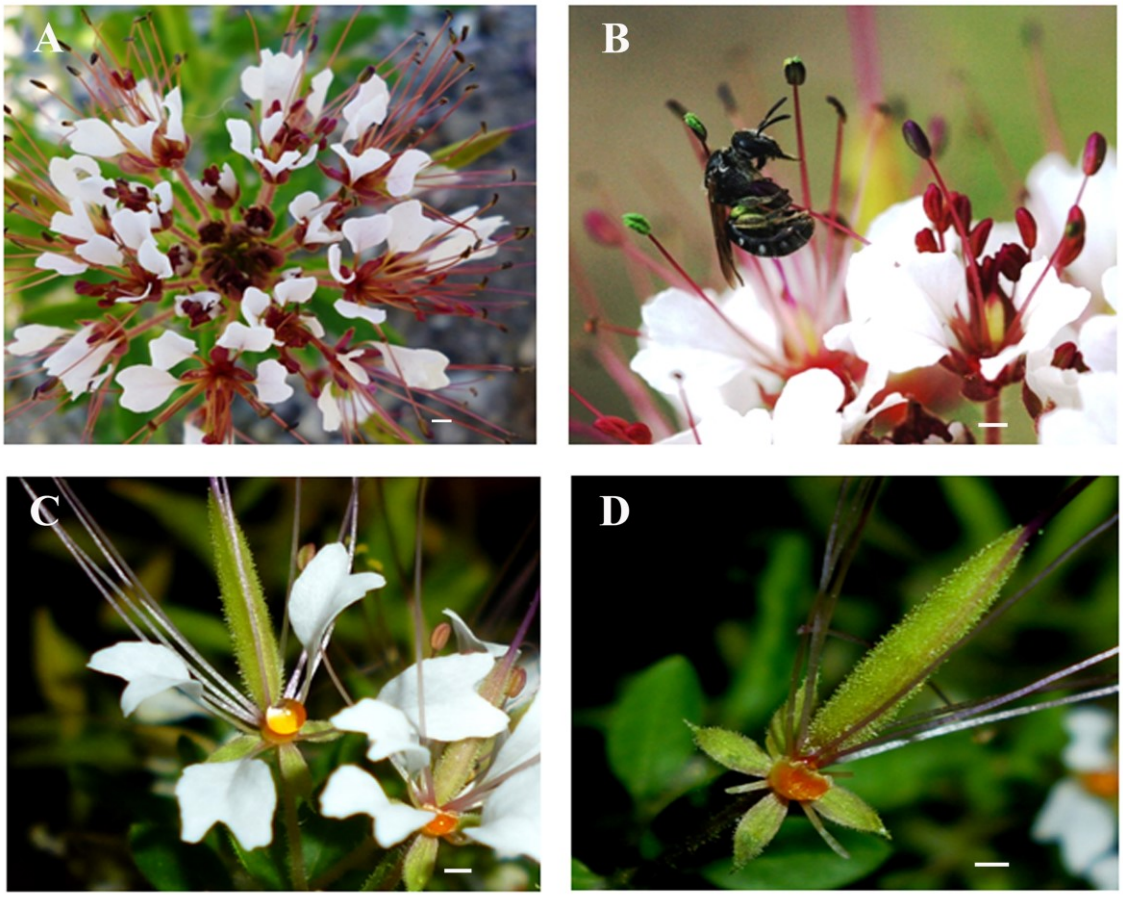


Figure 3.4. *Polanisia dodecandra* flowers. (A) Frontal view. (B) Lateral view. Scale bar: 1 cm.

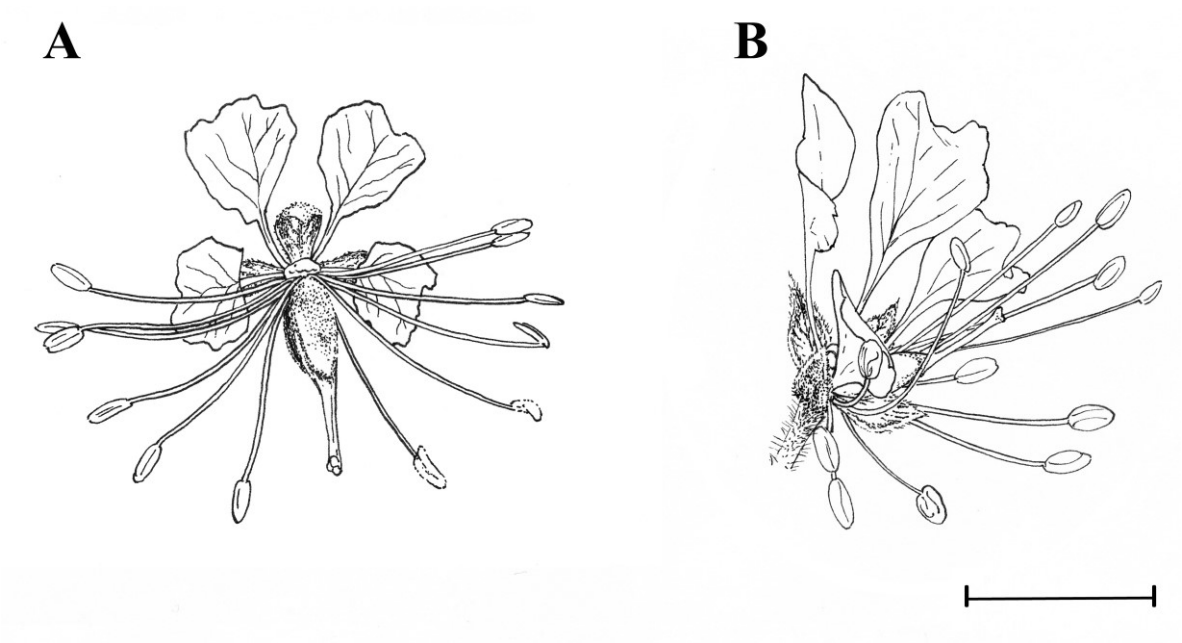


Figure 3.5. Nectar production (A) *Cleomella serrulata*, (B) *Polanisia dodecandra*, and sugar concentration (C) *Cleomella serrulata*, (D) *Polanisia dodecandra* at different time intervals. Error bars show the standard deviation (N=50).

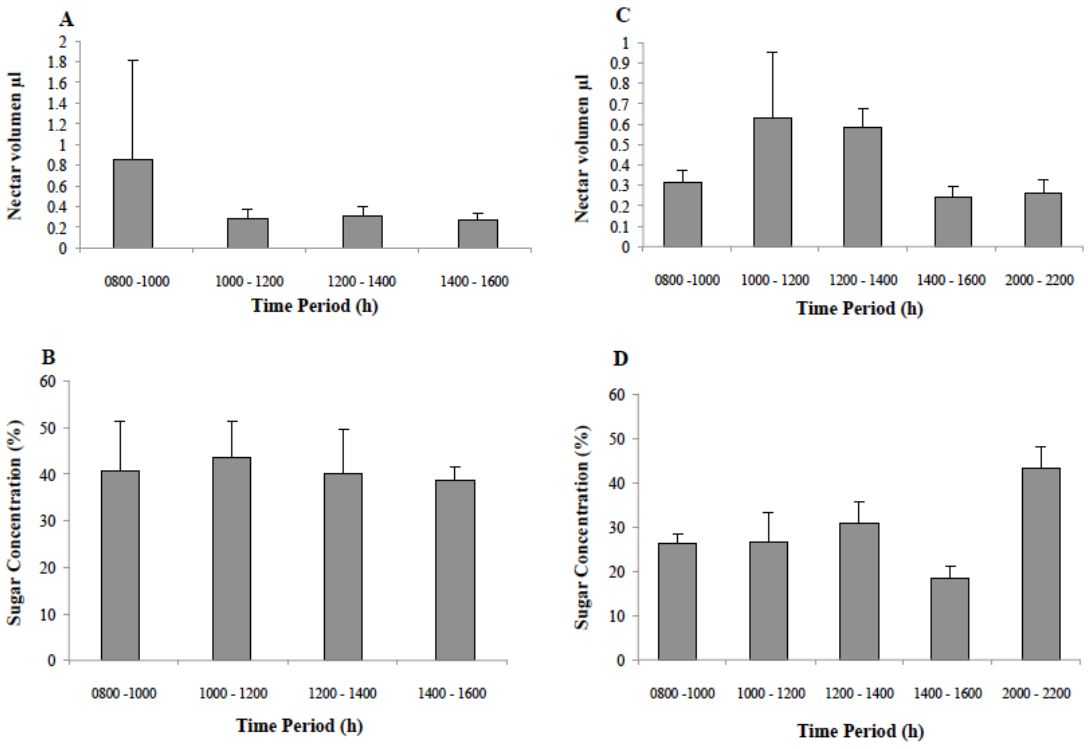


Figure 3.6. Richness and number of insect visits for *Cleomella serrulata* flowers in 2013 and 2014 at Komati (KM) and Casa Berardi (CB) areas. (A) Species richness per insect order. (B) Number of visits per insect order.

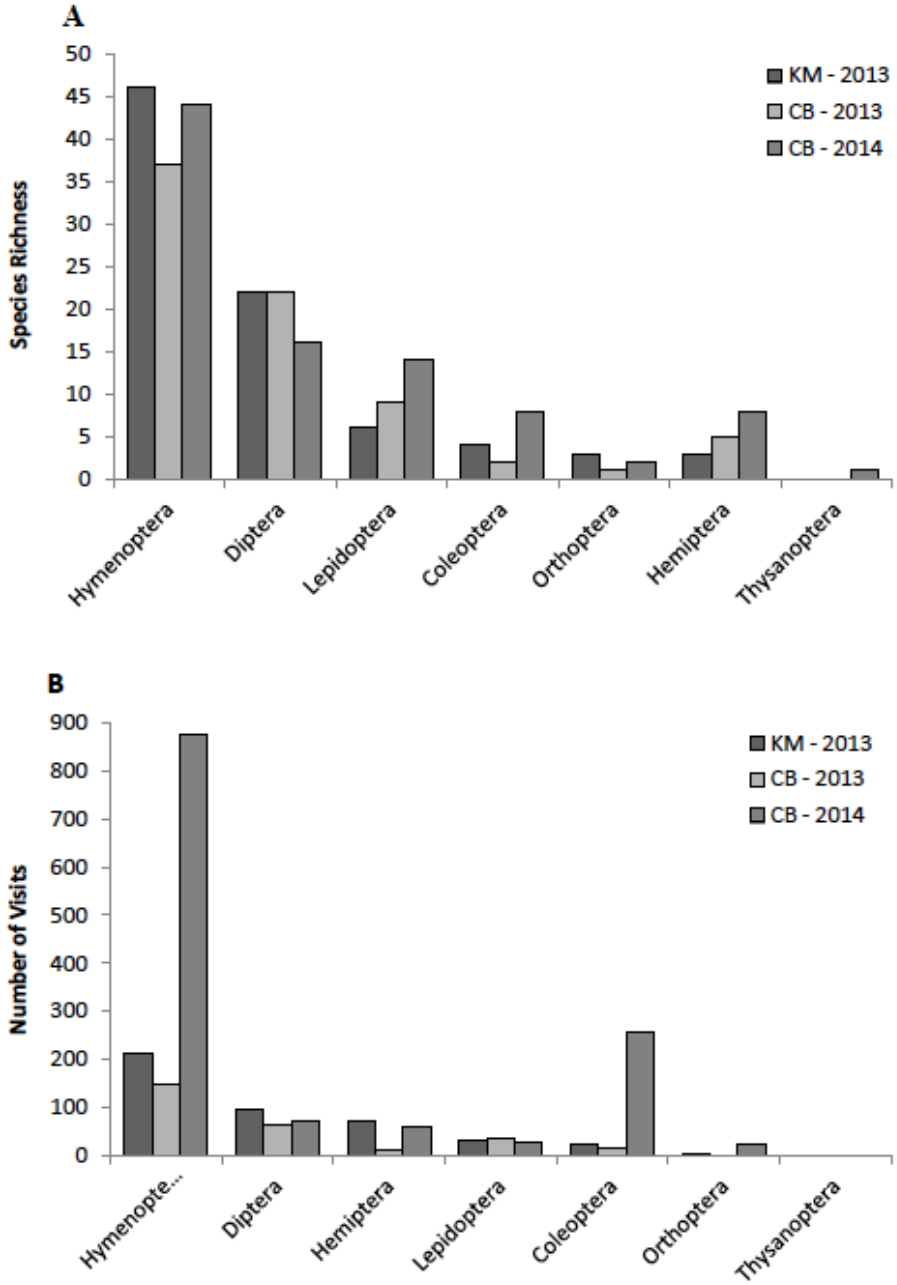


Figure 3.7. Richness and number of insect visits for *Polanisia dodecandra* flowers in 2013 and 2014 at Fish Creek (FC) area. (A) Species richness per insect order. (B) Number of visits per insect order.

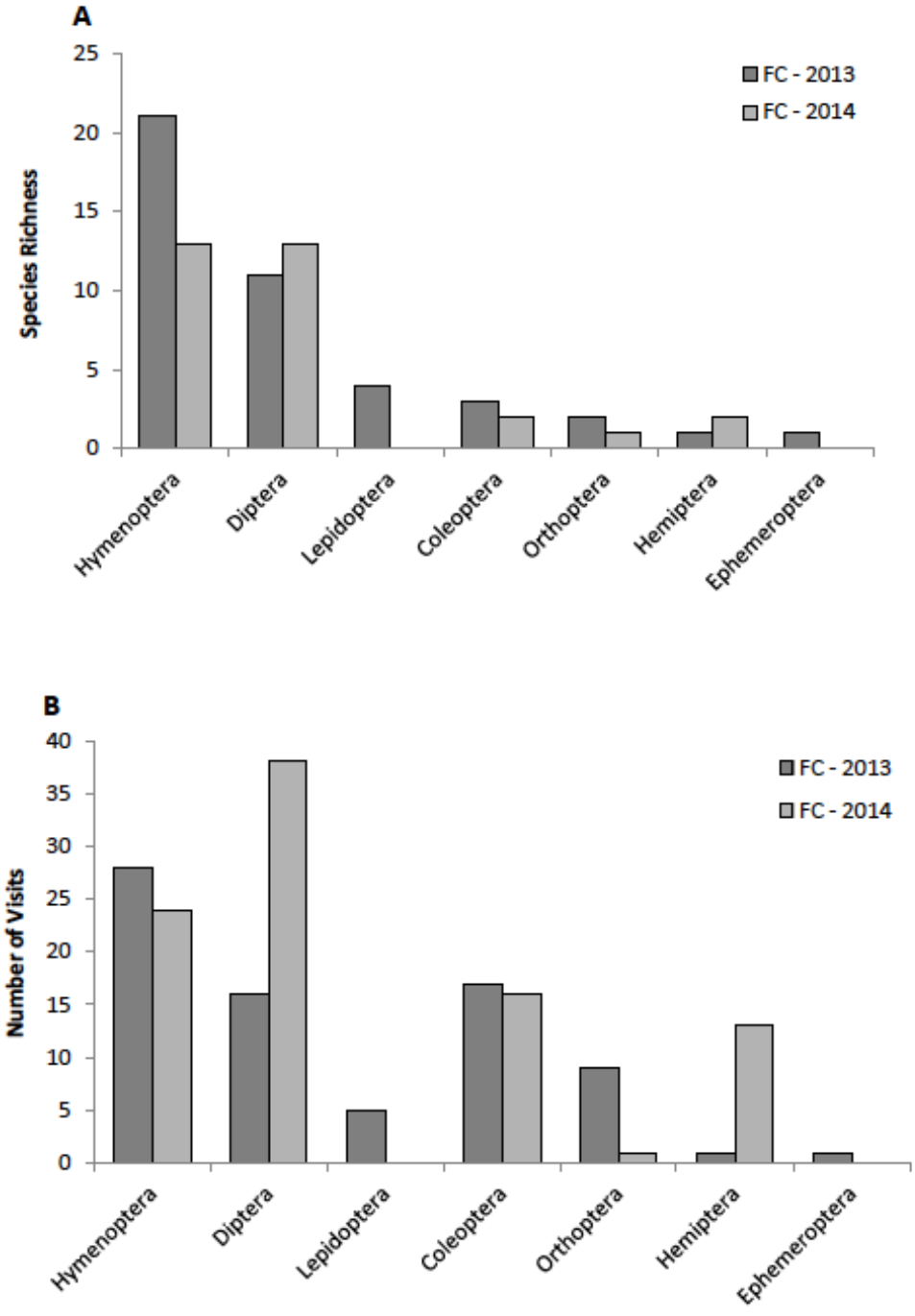
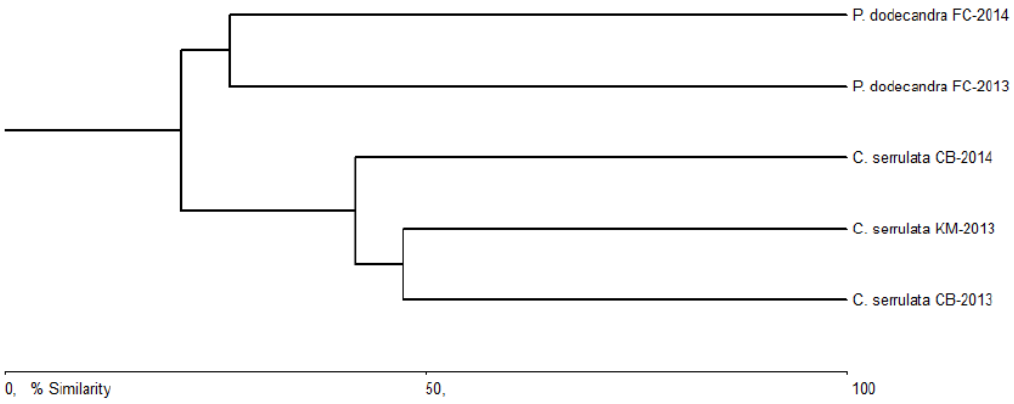


Figure 3.8. Jaccard similarity index dendrogram. CB=Casa Berardi, FC=Fish Creek, KM=Komati.



Chapter 4: Comparing floral transcriptomes of three Cleomaceae species reveals little variation between different pollination systems

4.1. Introduction

For centuries, plant-pollinator interactions have fascinated researchers who have dedicated their efforts to understanding the evolutionary processes of these relationships. Plant-pollinator interactions focus on how pollinator-mediated selection on specific floral traits has led to the stunning morphological diversity of flowers, and how the pollination by the same kind of animal pollinators promotes the convergence of floral phenotypes, which leads to pollination syndromes (Fenster *et al.* 2004; Johnson and Steiner 2000; Mitchell *et al.* 2009; Ollerton *et al.* 2009; Schiestl and Johnson 2013; van der Niet *et al.* 2014). Studies on pollinator-mediated selection on floral traits usually focus on the relationship between pollinators and floral phenotypes, in which the floral traits are correlated with particular groups of pollinators, but the link between the genes underlying these floral traits is frequently unknown (Kessler and Baldwin 2011; Kessler *et al.* 2008; Sapir 2009b; Yuan *et al.* 2013*a*). Floral phenotypes are largely controlled by gene identity and gene expression (Bey *et al.* 2004; Chanderbali *et al.* 2016; Hoballah *et al.* 2007; Ma 2001). To understand the evolution and pollinator-mediated selection on floral traits, it is critical to know which genes encode floral traits, because these genes play a main role in the determination of floral phenotypes (Sapir 2009b; Sedeek *et al.* 2013).

One of the most studied aspects of plant-pollinator interactions is the evolution of specialist pollination systems, where flower have complex morphologies, concealed rewards, and are pollinated by a specific pollinator guild (Herrera 2005; Herrera and Pellmyr 2002a; Proctor *et al.* 1996; Willmer 2011). Specialization in pollination systems promotes floral diversification and speciation (Alexandre *et al.* 2015; Mayer *et al.* 2011; Schiestl and Schlüter 2009). To date, most of the studies connecting genes, floral traits and pollinator preference have focused on specialized pollination systems, where pollinators exert directional and stabilizing selection on specific floral traits, and the genes encoding for these traits and their variations can be identified (Clare *et al.* 2013; Kessler and Baldwin 2011; Kessler *et al.* 2012; Owen and Bradshaw 2011; Sapir 2009b; Sedeek *et al.* 2013). However, generalist pollination systems, where plants are pollinated

by multiple pollinator guilds, are more common in nature (Waser *et al.* 1996). In fact, plants exhibit a generalization-specialization gradient in terms of the diversity of their pollinator communities and the type (diffuse vs. directional) or strength of selection on floral traits likely depends on where flowers fall along this gradient (Herrera 1996; Herrera 2005; Johnson and Steiner 2000; Waser *et al.* 1996). However, to explore transitions between pollination systems we need to know the ancestral state and determine the direction of transitions. Once the ancestral state is determined, we can explore transitions between generalist and specialist pollination systems in related species, and we can assess if generalist species share more genes with each other than with specialized species. If so, we could identify general gene distribution patterns in generalist and specialist pollination systems and identify if there are specific genes associated with both system types. We expected that the genes that could differ between generalist and specialist systems are those related to floral traits subject to pollinator mediated-selection such us genes coding color, scent, floral size, and nectaries.

A promising group to make such comparisons is the Cleomaceae, which is a small family (270 spp approximately, Patchell *et al.* 2014) with a cosmopolitan distribution that includes both generalist and specialist pollination systems (Cane 2008a; Fleming *et al.* 2009; Higuera-Díaz *et al.* 2015; Machado *et al.* 2006; Martins and Johnson 2013; Raju and Rani 2016). Although most flowers are monosymmetric and share a similar ground plan, they display great variation in morphological traits commonly associated with advertisement for pollinator attraction including petal colour, length of stamens and gynophore (stalk structure that supports the gynoecium), nectar guides, and floral fragrances (Cane 2008a; Higuera-Díaz *et al.* 2015; Machado *et al.* 2006; Patchell *et al.* 2014; Raju and Rani 2016). In addition, the flowers also vary in traits related with rewards such as nectar and pollen production (Higuera-Díaz *et al.* 2015; Raju and Rani 2016). Despite its diversity in floral traits, only a few studies have examined the pollination biology of Cleomaceae, with research into five generalist species (Cane 2008a; Higuera-Díaz *et al.* 2015; Martins and Johnson 2013; Raju and Rani 2016), and two specialist species (Fleming *et al.* 2009; Machado *et al.* 2006).

As an important first step in elucidating genetic variation in two Cleomaceae species exhibiting a generalist pollination system (*Cleomella serrulata* (Pursh) E.H.

Roalson and J.C. Hall and *P. dodecandra* DC) *vs.* one specialist Cleomaceae species (*Melidiscus giganteus* (L.) Raf., we generated and analyzed their floral transcriptomes. These species are found in distinct Cleomaceae clades, they belong to different genera (Figure 4.1, Patchell *et al.* 2014) and have different pollination systems (Cane 2008a; Fleming *et al.* 2009; Higuera-Díaz *et al.* 2015). The three species differ in their floral traits and they attract distinctive suites of pollinators (Cane 2008a; Fleming *et al.* 2009; Higuera-Díaz *et al.* 2015). The flowers of these species also vary in size and color, nectar gland presence and shape, and stamen number, (Figure 4.1), which likely reflect differences in their pollinator communities (Table 4.1).

The main goal of this study is to compare the genetic repertoires potentially underlying the main floral traits of two species exhibiting a generalist pollination system (*C. serrulata* and *P. dodecandra*) and one specialist (*M. giganteus*) species, focusing on traits that attract different suites of pollinators such as floral color, symmetry, and nectar gland presence/absence. We used a comparative transcriptomics approach to answer the following questions: Are the floral transcriptome profiles of species exhibiting generalized traits more similar to each other than that of a species exhibiting specialized floral traits? Which putative genes encode the main floral traits responsible for pollinator attraction? Are there differences in presence/absence of transcription factors in species that have generalize and specialized floral traits? We expected to find differences in genes coding for floral color between species exhibiting generalized and specialized floral traits, because all of them differ in petal color. In addition, we expected to find differences between species with specialized and generalized traits in genes coding for floral nectary development and nectar production, because these traits are likely to be under strong pollinator-mediated selection in specialist pollination systems. We also expected that *Melidiscus giganteus* has a higher number of nectar production related genes due to bat pollination.

4.2. Materials and Methods

4.2.1. Study system

The pollination biology of *C. serrulata* and *P. dodecandra* was previously described and further details about floral rewards, attractants and frequency of pollinator visits on these species can be found in Cane (2008a) and Higuera-Díaz *et al.* (2015).

Although there is only one published reference about pollination in *M. giganteus* by Fleming *et al.* (2009), the species is visited by bats and accordingly it has modified floral traits related with bat pollinator syndrome. In addition, under greenhouse conditions the flowers of this species bloom at dawn and they only last one night, though we did not observe nectar secretion (Higuera, personal observation).

Despite different pollinators (Cane 2008a; Fleming *et al.* 2009; Higuera-Díaz *et al.* 2015) *M. giganteus*, *C. serrulata* and *P. dodecandra* share many floral features such as mono-symmetric flowers with four sepals and four petals oriented upwards. However, there are also notable differences between these species. Plants of *C. serrulata* are 20 cm to 1.5 m high, with pink, oblanceolate petals, and long stamens with greenish linear anthers (Table 4.1, Figure 4.1). The gynophore, the stalk supporting the gymnoecium on *C. serrulata* is pinkish and as long as the stamens (Moss and Packer 1983), and the flowers exhibit a green, elongated and laminar nectar gland. This species has low drought tollerance. *Polanisia. dodecandra* plants are 30 to 90 cm high with white, clawed petals, and stamens that are unequal in length, with purplish-green globular anthers (Table 4.1, Figure 4.1). The gynophore in *P. dodecandra* is short (0 – 2 mm), and the flowers have an orange, obliquely truncated and concave nectar gland (Moss and Packer 1983). This species has excellent drought tolerance. *Melidiscus giganteus* plants are 1.8 to 3 m high, with green, linear and elongated flowers, and long stamens with conspicuous purplish-yellow linear anthers. The gynophore in *M. giganteus* is reddish and as long as the stamens (Iltis and Cochrane 2014; Soares Neto 2017). *Melidiscus giganteus* flowers from the seeds that we growth lack of an observable nectar gland (Table 4.1, Figure 4.1). However, caution is needed because there are reports of bats feeding on nectar in this species (Fleming *et al.* 2009). There are not published studies about drought tolerance in *M. giganteus*. *Cleomella serrulata, M. giganteus*, and *P. dodecandra* are diploid species (2n), and their chromosome number is n=17, n=16-17, n=10 respectively (Ruiz-Zapata *et al.* 1996), and the based number of chromosomes in plant species ranged from 3 to 25 (Wanscher 1934).

4.2.2. Plant material and growth conditions

Plants of *M. giganteus* and *P. dodecandra* were grown under long day conditions at 22–24 °C with 16 h light cycle in growth chambers at the University of Alberta.

Vouchers were deposited in the Vascular Plant Herbarium at the University of Alberta (ALTA), *M. giganteus* (Mankowski and Bolton 23 Jun 2008; 814 from Hortus Botanicus), and *P. dodecandra* (Hall & Bolton 20 Feb 2008; 68456 from B&T World Seeds). *Cleomella serrulata* plants were obtained from an Edmonton home garden, because we could not grow it successfully in the greenhouse after several trials.

4.2.3. RNA extraction

We pooled apical meristematic tissue and floral tissue of different developmental stages including small, medium and large buds, and open flowers from two plants for each species (n=3). We used meristematic tissue because the genes related with floral development are expressed before the actual flower develops. We decided to pool the samples because we need a high amount of RNA and we did not obtain the required amount with flowers and floral tissue from just one plant. However, because we pooled meristematic and flower tissue, we got genes expressed in leaves too, and we could not differentiate the genes that are expressed only in the flowers from the obtained gene pool. In addition, we did not sequence biological replicates because when we designed experiment by 2011, RNA-seq projects were scarce and transcriptomics was the suitable approach. All the collected tissue was flash frozen in liquid Nitrogen and kept at -80 °C to avoid RNA degradation. Total RNA was extracted with RNeasy plant MiniKit (QIAGEN, Hilden, Germany) following the manufacturer's protocol, then treated with DNAse I (New England Biolabs, Ipswich, USA) for 30 min at 37 °C to remove residual DNA from the total RNA. We measured RNA quality and quantity by spectrophotometry on a NanoDrop ND-1000 (Thermo Scientific, Wilmington, USA) and a 2100 Bioanalyzer (Agilent Technologies, Böblingen, Germany).

4.2.4. Illumina sequencing, *de novo* assembly and assembly quality

RNA sequencing was conducted by Plate-forme d'Analyses Génomique de l' Université Laval. mRNA was purified from 3 µg of total RNA, then fragmented and converted to double-stranded cDNA using the Illumina TruSeq RNASeq library preparation kit following Illumina's guidelines. In addition, each sample was tagged with an indexed adapter that contained a barcode sequence used to distinguish the pooled samples from each other after sequencing, as libraries from the three species were combined in a single sequencing lane to reduce sequencing costs. To improve

transcriptome gene discovery, the cDNA was normalized using the Trimmer cDNA Normalization Kit (Evrogen, Moskow, Russia) according to the manufacturer's instructions. As a result of the normalization step, the abundance of highly represented transcripts decreased prior to sequencing, which increased the chance of finding rare and low abundant transcripts during the sequencing process (Shcheglov *et al.* 2007). After sequencing, the adapter sequences were then removed by Plate-forme d'Analyses Génomique de l' Université Laval following Illumina guidelines. The complete set of generated sequences will be deposited in Sequence Read Archive (SRA) at the National Center for Biotechnology Information (NCBI).

Prior to assembly, we assessed raw read quality using FASTQC v0.10.1 software (http:/bioinformatics.bbsrc.ac.uk/projects/fastqc). Results from FASTQC indicated that the trimming steps were sufficient for assembly in all the species (Appendix 4.1 – 4.3 Supplementary flies 1-3). Because all three species lack a reference genome and to select for the most accurate and complete subsequent analysis, we performed a *de novo* assembly using CLC genome workbench (v9.5- https://www.qiagenbioinformatics.com/) and Trinity (vr20131110- Haas *et al.* 2013) with the default parameters for both programs (Honaas *et al.* 2016). We obtained the N50 statistics or unigene length at which the assembled bp reaches 50% of total assembly length (Honaas *et al.* 2016). The resulting *de novo assembly* from Trinity was used to generate two transcriptomes for each species: a complete transcriptome and a reference transcriptome. The complete transcriptome consists of all contigs including paralogs and isoforms from Trinity assembly. Because Trinity assembly includes all paralogs and isoforms, we created a reference transcriptome for each species using in-house bash scripts. Thus, the scripts selected the longest contig from the complete transcriptome to represent each gene identified by Trinity; these will be referred to as genes from here on. Therefore, the reference transcriptomes are smaller, should not contain all the isoforms and may not include all gene copies. The reference transcriptomes were used for subsequent analysis.

After completing the assembly, we compared the performance and completeness of CLC and Trinity *de novo* assemblers. We conducted both analyses on the CLC assembly and Trinity complete and reference assemblies for comparative performance purposes. We used TransRate v1.0.3 (Smith-Unna *et al.* 2016) to determine assembly quality and

accuracy. TransRate provided read-based metrics to generate contig scores, which are integrated together to evaluate whole assemblies (Smith-Unna *et al.* 2016). Furthermore, we used Benchmarking Universal Single-Copy Orthologs v2.0 (BUSCO- Simão *et al.* 2015) to evaluate transcriptome completeness. We ran BUSCO with the '-m tran' setting against the 'Plants' dataset (Embryophyta_odb9 on their website), which comprised 30 species and 1440 BUSCOs.

4.2.5. Gene Ontology annotations

Once we obtained good quality assemblies, we proceeded to identify the genes and their functions. The resulting contigs from the Trinity reference transcriptome were compared against an *Arabidopsis thaliana* database (TAIR10_pep_20110103 representative gene model) using BLASTX (E-value $<10^{-6}$) in Blast2GO v4.07. The subsequent blast results were mapped and annotated in Blast2GO, in which the functional information for each contig BLAST hit was recovered from the Gene Ontology (GO) database. GO terms were selected from a pool of candidates by applying the default annotation rule in Blast2GO (Conesa *et al.* 2005). We did a Venn-diagram for main three GO categories (Biological Process, Cellular Component and Molecular Function) with the top 10 related terms by category using Jvenn (Bardou *et al.* 2014). In addition, we did Venn-diagrams for all pollination related terms and for selected terms related (e.g., floral color, nectaries development and nectar production) to identify similarities and differences in gene repertoire between species with generalized and specialized traits.

4.2.6. Orthologous gene analysis

To compare presence/absence of the orthologous groups between species, we used OrthoMCL v2.0.9 (Li *et al.* 2003) to identify orthologous proteins encoded by the reference transcriptomes of *C. serrulata, M. giganteus*, and *P. dodecandra*. Predicted protein sequences were obtained from Trinity contigs from reference transcriptome as described above using TransDecoder v3.0.1 (Haas *et al.* 2013). An all-by-all BLAST search was performed on all protein sequences according to OrthoMCL specifications and using an e-value threshold of 10^{-6}. Potential ortholog pairs were grouped together with the Markov cluster algorithm (-MCL c12-135-, Enright *et al.* 2002) according to OrthoMCL specifications and an inflation value (I) of 1.5. All other steps were performed with default parameters (Li *et al.* 2003). We did not identify paralog genes with this analysis, because to

identify with certainty the genes produced by genome duplication events or paralog genes (gene copies), it is necessary conduct a gene phylogeny for the genes of interest, and to have a reference genome, which is lacking for the studied species.

4.2.7. Transcription factor analysis

Angiosperm transcription factors (TFs) are involved in major biological processes such as development, response to environmental stress, flowering, and fruit ripening (Lehti-Shiu *et al.* 2017). TFs are also drivers of evolutionary innovations in plants including the evolution of novel floral structures (He and Saedler 2005; Panchy *et al.* 2016; Zhang *et al.* 2015a). Since TFs are potential candidate genes that encode for floral traits important for pollinator attraction and they play a main role in floral trait variation (Lehti-Shiu *et al.* 2017; Ortiz-Barrientos 2013; Sedeek *et al.* 2013; Yuan *et al.* 2013*a*), we identified TFs in the transcriptomes of the three Cleomaceae species. To detect TFs for the complete Trinity assembly of *M. giganteus*, *C. serrulata* and *P. dodecandra*, we obtained protein sequences from the reference assemblies using TransDecoder v3.0.1 with default parameters (Haas *et al.* 2013). Then, TFs families were assigned to the predicted proteins using the Transcription Factor Prediction Tool v4.0 on the Plant Transcription Factor Database (PlantTFDB) website (Jin *et al.* 2017), which employs Hidden Markov models (HMMs) to classify protein sequences.

4.2.8. Anthocyanin pathway analysis

Differences in flower color have been frequently associated with pollinator attraction and pollination syndromes (Davies 2009; Fenster *et al.* 2004; Sedeek *et al.* 2013; Willmer 2011). Anthocyanin-derived pigments are mainly responsible for pink, red, blue and purple flower colors and may be key players in determining pollinator preferences. We therefore compared the anthocyanin pathway in *C. serrulata*, *M. giganteus*, and *P. dodecandra*, which have pink, green, and white flowers, respectively. Anthocyanins were previously identified in *C. serrulata* and *T. hassleriana* (Jordheim *et al.* 2009; Nozzolillo *et al.* 2010) and we wanted to assess differences or similarities in the anthocyanin pathway between species. We examined the Kyoto Encyclopedia of Genes and Genomes (KEGG- Kanehisa and Goto 2000) to generate a list of candidate genes involved in the anthocyanin pathway from all GO terms associated with anthocyanins. We reviewed literature of floral anthocyanins to provide a preliminary list of genes that

may have encoded for the main enzymes involved in the pathway (Davies 2009; Guo *et al.* 2014; Ho and Smith 2016; Wessinger and Rausher 2012). Next, we performed presence/absence comparisons of the genes reported in each species. Finally, to compare if the main anthocyanin enzymes were present in the species exhibiting generalized and specialized traits, we obtained the anthocyanin pathway from KEGG (Appendix 4.4) to generate a simplified diagram depicting the main enzymes present in anthocyanin pathway of *C. serrulata*, *M. giganteus*, and *P. dodecandra* (Figure 4.8).

4.3. Results

4.3.1. Sequencing, *de novo* assembly and assembly quality

Three normalized cDNA libraries were constructed from meristematic and floral tissue of *C. serrulata*, *M. giganteus*, and *P. dodecandra*. After sequencing on the HiSeq 2000 (Illumina) platform, approximately 96 million paired-end reads were obtained for *C. serrulata*, whereas for *M. giganteus* and *P. dodecandra* 130 and 85 million, were obtained, respectively (Table 4. 2). The average length of the raw short reads for all the species was 100 bp (Table 4.2). Results from FASTQC specified that the sequences had enough quality for assembly (Appendix 4.1 – 4.3). CLC assemblies produced between 59220 and 79103 contigs with an average length of 669 bp, while Trinity assemblies produced contigs with an average length of 140014 and 240907 bp (Table 4.3). Since the contigs produced by Trinity contained paralogs and isoforms, which is determined by the Bruijn algorithm that the program used to reconstruct all the original transcripts (Haas *et al.* 2013), we made a reference assembly that included only the longest contig for each Trinity unigene. Thus, the reference transcriptomes range from 47090 to 76228 contigs (Table 4.3). The average contig length and the N50 values where higher in Trinity assemblies than in CLC assemblies (Table 4.3).

The quality assessment and completeness approaches indicated that the assemblies of *C. serrulata*, *M. giganteus*, and *P. dodecandra* were of good quality and relatively complete. TransRate scores ranged from 0.24 to 0.28 for CLC, while for Trinity complete and reference transcriptomes, TransRate scores ranged from 0.009 to 0.021, and 0.25 to 0.34 respectively (Table 4.4). TransRate scores above 0.22 thresholds indicate good quality of the obtained assemblies (Smith-Unna *et al.* 2016). The BUSCO results showed that the

assemblies were highly complete, with complete matches ranging from 60.7 to 61.7 % of plant orthologs for CLC, from 90.9 to 94.2 for Trinity complete transcriptome, and 74 to 76.9 for Trinity reference transcriptome (Table 4.4). Since Trinity provided more complete and accurate assemblies than CLC, we used Trinity complete and reference transcriptomes for GO annotations analyses (Table 4.4).

4.3.2. Gene Ontology annotations

Gene ontology (GO) annotations were recovered from most sequences of *C. serrulata, M. giganteus,* and *P. dodecandra*. Overall, we obtained 119403 (64.07%), 189269 (78.54%), and 100864 (72.03%) sequence annotations for the complete Trinity transcriptome of *C. serrulata, M. giganteus,* and *P. dodecandra*, respectively (Appendix 4.5). In addition, we obtained 37277 (48.9%), 9060 (52.35%), 21727 (46.13%) sequence annotations for the reference transcriptome of *C. serrulata, M. giganteus,* and *P. dodecandra*, respectively. The reference transcriptomes of the three species were classified for three main GO functional categories: biological process, molecular function, and cellular component with their related subcategories (Figure 4.2).

The three-species showed similar patterns of gene annotations. For all three species, the largest number of transcripts for the reference transcriptome was associated with biological process that includes pollination, followed by cellular component, and molecular function (Figure 4.2), with the most common GO term for each of these three categories being metabolic process, chloroplast, and protein binding respectively (Figure 4.2). In general, *Cleomella serrulata* had the most transcripts associated with GO terms, while *P. dodecandra* had the fewest genes associated with GO terms (Figure 4.2). Furthermore, we selected 14 GO terms specifically related to pollination and floral development, all of which belonged to the biological process category (Figure 4.3). We identified 1247 transcripts associated with these selected GO terms (Appendix 4.6). Moreover, *C. serrulata* had the highest number of transcripts associated with GO pollination and floral development terms, followed by *M. giganteus*, then *P. dodecandra*. Within pollination and floral development, the most abundant terms were anthocyanin accumulation in tissues in response to UV light, determination of bilateral symmetry, and polarity specification of adaxial/abaxial or dorsal/ventral axis (Figure 4.3). The pollination and floral development GO terms were generally similar across species

(Appendix 4.6), although there were differences in the presence/absence of underlying transcripts associated with pollination and floral development GO terms (Appendix 4.6). The three species shared 231 transcripts (24.29%), but *C. serrulata* had the highest number of unique transcripts (20.08%), while *P. dodecandra* and *M. giganteus* shared more transcripts (10.72%; Figure 4.4, Appendix 4.6). For the anthocyanin accumulation in tissues in response to UV light Venn-diagram, we found that *C. serrulata* has the higher number of unique transcripts, but there were few differences in the number of transcripts shared by species with generalized and specialized traits, as we expected genes related with anthocyanin accumulation in tissues was different between species (Figure 4.5a, Table 4.5). Regarding genes involved in the anthocyanin pathway, we found all the major genes in the three species despite color differences between them and the results did not meet our expectation (Table 4.5, Figure 4.8). This result may be due to the inclusion of meristematic tissue in the samples, since anthocyanins are involved in other metabolic processes and not only in color determination. We found shared core genes for nectary development (*CRC, BOP*) between generalist and specialist species. Interestingly, we found genes only shared by the generalist species (*AFO* and *YABBY2*), and we also found unique genes in the specialist species (*INO* and *YABBY5*), in this case our expectation was affirmative and the generalist species shared specific nectary development genes (Figure 4.5b, Table 4.5). Regarding nectar production, we did not find shared transcripts between generalist species (Figure 4.5c), in fact, generalist and specialist species shared more transcripts between them, but as we expected the bat pollinated species (*M. giganteus*) had more nectar production related genes and our expectation was met (Figure 4.5c, Table 4.5).

4.3.3. Orthologous gene analysis

When comparing the number of shared and unique protein families across our three species, we found 56616 proteins that were clustered into 35896 orthologous groups (Figure 4.6). Of these groups, 8041 (22.4%) orthologous groups were common to all species, while 1876 (5.22%) orthologous groups were shared by *C. serrulata* and *P. dodecandra*, 1364 (3.79%) orthologous groups were shared by *C. serrulata* and *M. giganteus*, and 1398 (3.89%) by *M. giganteus* and *P. dodecandra* (Figure 4.6). *Cleomella serrulata* had the highest number (12983; 36.16%) of unique protein families, almost twice

as many as *M. giganteus* (6495; 18.09%) and three times more than *P. dodecandra* (3739; 10.41%; Figure 4.6). However, *C. serrulata* and *P. dodecandra* shared a higher number of orthologous groups when compared with *M. giganteus* (Figure 4.6). Interestingly, patterns of shared orthologous groups were not associated with phylogenetic relatedness. Although, *M. giganteus* and *P. dodecandra* share a more recent common ancestor (Patchell *et al.* 2014), they did not share more orthologous groups (Figure 4.6).

4.3.4. Transcription factor analysis

The distribution of the TF families was similar across the three species (Figure 4.7). Overall, we identified 58 putative TF families represented in 2924 unique transcripts (1212 unique transcripts for *C. serrulata*, 864 for *M. giganteus*, and 848 for *P. dodecandra*). The number of transcripts assigned to TF families was highest in *C. serrulata*, followed by *M. giganteus* and *P. dodecandra*. For all three species, the most abundant TF families were *bHLH (BASIC HELIX-LOOP-HELIX), ERF (ETHYLENE RESPONSE FACTOR)*, *C2H2* (*CYS2-HIS2*), *NAC* (*NAM, ATAF* and *CUC*), and *MYB* (*MYELOBLASTOSIS* - Figure 4.7). Other abundantly represented TF families included *WRKY* (*WRKYGQK AMINO ACID MOTIF*), *MYB*-RELATED (*MYELOBLASTOSIS-RELATED*), *bZIP* (*BASIC-REGION LEUCINE ZIPPER*), *GRAS* (*GAI, RGA* and *SCR*), and *C3H* (*CYS3HIS*) families. On the other hand, the less abundant TF families were *LSD* (*LESION SIMULATING DISEASE*), *STAT* (*SIGNAL TRANSDUCER* and *ACTIVATOR* of *TRANSCRIPTION*), *NZZ/SPL* (*NOZZLE/SPOROCYTELESS*), *RAV* (*RELATED*-TO-*ABI3/VP1*), and *SAP* (*STERILE APETALA* - Figure 4.7). Despite these similarities of TF family distribution among the species, we found some differences in terms of present/absent families: *NZZ/SPL*, *RAV*, and *SAP* families were absent in *M. giganteus*, while, *HB-PHD* (*HOMEOBOX- PATHOGENESIS RELATED*), *HRT-Like* (*HAIRY RELATED TRANSCRIPTION FACTOR-Like*), and *SAP* families were absent in *P. dodecandra*. All the TF families were present in *C. serrulata* (Figure 4.7).

4.3.5. Anthocyanin pathway

Since the anthocyanin biosynthetic pathway is well understood in flowering plants (Wessinger and Rausher 2012), we were able to identify putative homologues of all major enzymes of the core anthocyanin pathway (Figure 4.8). We identified 90 unique transcripts involved in anthocyanin biosynthesis representing seven candidate

enzyme classes (Appendix 4.7). In general, the components of the pathway were similar across the species (Figure 4.8). However, whitish-green flowers of *M. giganteus* had a higher number of unique transcripts encoding for *CHI* (*CHALCONE ISOMERASE*), *CHS* (*CHALCONE SYNTHASE*), and F3'H (*FLAVONOID 3'-HYDROXYLASE* - Appendix 4.7), although this species has green petals, the highest number of transcripts could be related to others major anthocyanin functions as UV protection, metal chelating, and antioxidants (Gould *et al.* 2009). We also found some differences in transcripts encoding enzymes: the transcript *GSTF5* (*GLUTATIONE TRANSFERASE 5*) was present only in *M. giganteus,* while*,* the transcript GSTF6 (*GLUTATIONE TRANSFERASE 6*) was absent in *M. giganteus* (Figure 4.8, Appendix 4.7). We did not find transcripts involved in the production of pelargonidin and delphinidin, which are common components in the anthocyanin pathway in species that have red and blue flowers. These were therefore removed from the depicted KEGG metabolic pathway (Appendix 4.4). Regarding anthocyanin pathway regulation, we found similarities and differences in the presence/absence of main anthocyanin positive regulator genes belonging to *R2R3MY*B and *bHLH* transcription families. Here, *bHLH* and *R2R3MYB* represented 8.82% and 5.69%, respectively, of all transcription factors identified (Figure 4.7) and the transcripts belonging to *R2R3MY*B family *PAP1/2* (*PRODUCTION OF ANTHOCYANIN PIGMENT 1/2*), and *MYB113* (*MYELOBLASTOSIS 113*) were present among the species. Anthocyanin biosynthesis is also negatively regulated by *CPC* (*CAPRICE*) and *LBD* (*LATERAL ORGAN BOUNDARY DOMAIN*) transcription factors (Guo *et al.* 2014); *CPC* was absent only in *P. dodecandra,* while the transcript *LBD37* was present in all the species. In addition, *LBD38* was present only in *M. giganteus*, while *LBD39* was absent only in *M. giganteus.*

4.4. Discussion

We examined the gene repertoires of three Cleomaceae species with different floral traits and different pollination systems to determine if there were major differences in the transcriptomic profiles of species exhibiting generalized and specialized floral traits. We focused our analysis on the gene repertoires of flowers and meristematic tissue between species, examining candidate genes for floral traits associated with pollinator attraction

and, more specifically, genes associated with flower color, the anthocyanin pathway and nectar production. In general, the gene repertoire of all three species were similar, despite having different pollinators and belonging to relatively distant clades. Both species exhibiting a generalist pollination system (*Cleomella serrulata* and *Polanisia dodecandra*) shared more orthologous groups with each other than the specialist pollination system species (*M. giganteus*), but importantly we found presence/absence similarities and differences in the genes associated with pollination related terms (Figure 4.4, Appendix 4.6). The obtained transcriptome data resulted in identification of candidate genes for further investigation as they may be important for difference in pollinator attraction, specifically genes involved in nectar production. We found the main components of the anthocyanin pathway in all species. The color differences between species may be the result of differential gene expression and regulatory processes rather than gene repertoire. Contrary to our expectations (Table 4.5), we found the same anthocyanin pathway related genes despite color differences between species, we attributed these results to the inclusion of meristematic tissue in the analysis and other major anthocyanin roles such us UV light protection, and antioxidant activity. Based on gene presence/absence data and after reviewing core literature, we proposed candidate genes coding for flower color and nectar traits that may be subject to selection pressures exerted by pollinators. Importantly, we found that nectary development genes were shared by species exhibiting generalized floral traits (*C. serrulata* and *P. dodecandra*) and species exhibiting specialized floral traits (*M. giganteus*) have unique genes involved in nectary development. Overall, there were more similarities than differences between species belonging to generalist and specialist pollination systems in gene repertoires for the reference transcriptome, and we found differences in the genes coding for flower color, nectary development, nectar production, and other pollination related traits.

4.4.1. Gene repertoires were similar across species

We found a similar pattern in terms of gene ontology annotations, and gene repertoires across species despite phylogenetic distance (Figures 4.1, 4.2, and 4.5) and differences in pollinator communities. Interestingly, *C. serrulata* has more unique protein groups than the other species, with double and triple groups of unique proteins when compared to *M. giganteus* and *P. dodecandra*, respectively (Figure 4.6). In the

Cleomaceae phylogeny *C. serrulata* belongs to a notoriously long branch (Figure 4.1, Patchell *et al.* 2014), which means that in this clade there are more nucleotide substitutions per site and the variation in substitution rates is high. Cleomaceae experienced an ancient whole genome triplication event, which occurred approximately 24-13 Mya (Barker *et al.* 2009; Cheng *et al.* 2013; van den Bergh *et al.* 2016a; van den Bergh *et al.* 2014). This triplication event is shared amongst common ancestors of *Tarenaya* and *Gynandropsis* (van den Bergh *et al.* 2014), including the lineage for *Melidiscus giganteus* (Patchell et al. 2014). *C. serrulata* and *P. dodecandra* are not identified as any of these lineages (Feodorova *et al.* 2010), but further investigations are needed to determine number and placement of genome duplications in the Cleomaceae. However, we predict *C. serrulata* has undergone the whole genome triplication event due to the high number of unique assembly contigs and isoforms from the Trinity complete assembly, but a gene phylogeny is necessary to establish gene copies number. Morphological complexity in plants has been correlated with retention and expansion of transcriptional regulation genes after gene duplication events and morphological novelties can originate from these retained gene copies (Specht *et al.* 2012). It would be interesting to investigate the function of genes copies in the development of novel structures as androgynophores and nectar glands, because these traits are different among Cleomaceae species. Further studies, including a reference genome and gene phylogenies are needed to elucidate whether *C. serrulata* and *P. dodecandra* experienced the whole genome triplication event and the impact of these events on generating variation in gene repertoires across species.

4.4.2. Candidate genes for pollinator attraction

Floral symmetry, floral color, and nectar gland development are all traits that are predicted to be under selection by pollinators, but the magnitude and direction of selection may vary depending on whether plants have generalized vs. specialized pollination systems. Cleomaceae flowers are monosymmetric and there are two different pathways to achieve monosymmetry in this family (Patchell *et al.* 2011). *Cleomella serrulata* and *P. dodecandra* exhibit early monosymmetry, while *M. giganteus* exhibits early dissymmetry (Patchell *et al.* 2011). As in other angiosperms, genes belong to the *TEOSINTE BRANCHED 1/CYCLOIDEA/PCF* (*TCP*) family such as *CYCLOIDEA* (*CYC)*

are involved in floral symmetry determination (Busch and Zachgo 2009). In particular, genes belonging to *CYCLOIDEA 2* (*CYC2*) subgroup play a key role in the determination of floral monosymmetry and also contribute in the evolution of flower form (Busch and Zachgo 2009; Hileman and Cubas 2009; Kramer 2007; Preston and Hileman 2009; Zhang *et al.* 2010). In our analysis, we found a sequence with a high similarity to the gene *TCP1*, which is responsible for monosymmetry determination in *Iberis amara* L. (Brassicaceae) and it was present in the three species (Appendix 4.6). Although the three species have a different developmental pathway, they shared the same symmetry type and the lack of difference may be evidence that pollinators are exerting stabilizing selection to favor symmetric flowers (Neal *et al.* 1998). Further studies are necessary to identify *TCP* gene copy number and function in *C. serrulata*, *M. giganteus,* and *P. dodecandra.* Likewise, it would be interesting to determine if the role of the *TCP1* in the determination of monosymmetry of Cleomaceae species is similar to the pattern observed in *I. amara*, in which *TCP1* is strongly expressed in adaxial petals before anthesis or if the pattern is different (Busch and Zachgo 2007, 2009).

Nectar is a key reward that mediates plant-pollinator interactions and several genes are important in the development of nectar glands and the regulation of nectar production (Bender *et al.* 2012; Kram *et al.* 2009; Lin *et al.* 2014; Roy *et al.* 2017; Ruhlmann *et al.* 2010). Both *C. serrulata* and *P. dodecandra* have an evident nectar gland and they produce copious amounts of viscous nectar (Higuera-Díaz *et al.* 2015). These species shared the same number and identity of nectar gland development genes (Appendix 4.6), but *C. serrulata* had more genes associated with nectar production when compared to *P. dodecandra* (Appendix 4.6). Interestingly, flowers of *M. giganteus* lacked an apparent nectary and *M. giganteus* plants grown for this study produced no nectar, though it has nectar production genes (e.g. *CWINV2-6*), and is visited by nectarivorous bats (Fleming *et al.* 2009). Moreover, as we expected, *M. giganteus* had the highest number of genes associated with nectar production (Figure 4.3, Appendix 4.6). This pattern is consistent with differences in pollinators as bat-pollinated flowers produce higher volume and low sugar concentration nectar (Fleming *et al.* 2009). Our observations suggest that the lack of nectar was likely due to greenhouse conditions and plant growth. Floral observations and dissections in wild plants will allow us to identify if the flowers have or not a nectar gland.

Overall, there were some presence/absence differences in the genes involved in nectary development and nectar production among species. The genes *CRC* and *BOP2* were present in all species, but *AFO* and *YAB2* were absent in *M. giganteus*, while *INO* and *YAB5* were present only in *M. giganteus* (Appendix 4.6). More importantly, species exhibiting generalized floral traits shared specific nectary development transcripts that were absent in the specialist species (Figure 4.5b, Table 4.5) Future studies should focus more on the candidate genes *AFO*, *BOP2*, *CRC*, *CWINV*,4 and *INO* because they could play a role in nectar gland development and nectar production in *M. giganteus*, *C. serrulata,* and *P. dodecandra*. Further quantitative gene expression studies are necessary to elucidate if there are gene expression differences in these genes among the species. *Melidiscus giganteus* had the highest number of transcripts involved in nectar secretion and in nectar gland development even though we did not observe nectar in this species (Appendix 4.6). It is possible that while *CRC* and *BOP2*, which are major players in the formation and development of nectar glands in angiosperms, are present in *M. giganteus*, they have a posttranscriptional regulation process or the transcripts may have low abundances. In a recent review, the genes *BOP1*/2, *AUXIN RESPONSE FACTOR 6* (*ARF6*), *ARF8*, and *CRC* are important for nectary development and mutants fail to develop the nectary and lack of it (Roy *et al.* 2017). In our study, we found the genes mentioned above associated with nectary development GO term with the exception of *ARF6* and *ARF8*. Since *ARF6* and *ARF8* are essential for nectary development in *A. thaliana* (Nagpal et al. 2005), we advise further exploration of these genes to test their function in nectary development in *C. serrulata*, *M. giganteus,* and *P. dodecandra*. We also recommend *AFO*, *BOP2*, *CRC*, *CWINV4*, *INO*, *YAB2*, and *YAB5* as candidate genes to perform functional studies to determine the role of these genes in nectary development and nectar production in *C. serrulata*, *M. giganteus,* and *P. dodecandra*.

4.4.3. Flower color and the anthocyanin pathway

Selection on flower color is often mediated by pollinator preference (Giurfa *et al.* 1995; Hoballah *et al.* 2007; Schemske and Bradshaw 1999; Schiestl and Johnson 2013; Wessinger and Rausher 2012) and anthocyanins play an important role in the determination of flower color in many plant species (Davies 2009; Grotewold 2006; Quattrocchio *et al.* 1999; Wessinger and Rausher 2012, 2014; Whittall *et al.* 2006). In our

three species, visual inspection initially suggested that anthocyanins, or the lack of anthocyanins, may be an important floral trait under pollinator-mediated selection. Our results suggest that the anthocyanin core pathway is well conserved across all the species, despite their belonging to different clades, and the main differences related to floral color may be due to differences in gene expression and gene regulatory functions.

In our study, all three plants have different colored flowers, which may be attributed, at least in part, to pollinator-mediated selection on the anthocyanin pathway. Numerous studies have attempted to elucidate the genetic basis of transitions between pigmented to non-pigmented flowers within and among species in anthocyanins, betalains, and carotenoids derived colors (Christinet *et al.* 2004; Davies 2009; Ho and Smith 2016; Wessinger 2015; Wessinger and Rausher 2012). In general, transitions from pigmented anthocyanin derived colors to white flowers are more common than the reverse (Davies 2009; Wessinger and Rausher 2012). These transitions may involve loss of function in pigment pathway genes or regulatory genes. Loss of function has also been reported in genes encoding for betalains and carotenoids derived pigments (Christinet *et al.* 2004; Wessinger 2015), suggesting that just detecting the presence of the genes from these pathways may not provide sufficient information to understand the role of anthocyanins in expressed floral colour.

Anthocyanin pigments have been previously identified for *C. serrulata* and *T. hassleriana* (Jordheim et al. 2009; Nozzolillo et al. 2010; van den Bergh 2017). Thus, *C. serrulata* have five anthocyanin derived pigments including cyaniding-3-2-glucosyl-6-coumaroyl-glucoside-5-glucoside and four acylated derivatives compounds: caffeoyl, sinapoyl, feruloyl, and *p*-coumaroyl. *Tarenaya hassleriana* has the same five anthocyanin pigments as *C. serrulata* plus five pelargonidin derived pigments (Jordheim *et al.* 2009; Nozzolillo *et al.* 2010). In *T. hassleriana*, purplish flowers have anyhocyanin derived compounds, while, pinkish flowers have pelargonidin derived pigments (Jordheim *et al.* 2009; Nozzolillo *et al.* 2010). Pelargonidin derived pigments are synthetized by the enzymes dihydroflavonol-4-reductase and anthocyanidin synthase (ANS) and they were present in the three-species (Appendix 4.7). However, we did not find any GO term related to pelargonidin production; instead, we found GO terms related to all the anthocyanidin-derived pigments (Appendix 4.6 and 4.7).

In our study, we identified 90 unique transcripts involved in anthocyanin biosynthesis (Appendix 4.7). We were surprised that *P. dodecandra*, which has white flowers, did not exhibit the absence of specific anthocyanin transcripts. This could be explained by the presence of anthocyanins in vegetative tissues, which suggests that the lack of color may be the result of gene down-regulation rather than gene loss in this species. In addition, anthocyanins play major roles in vegetative tissues as UV protection, and stress response (Hatier and Gould 2009). In *M. giganteus*, petals are mostly green, which could be attributed to the presence of chlorophyll in flowers (Ohmiya *et al.* 2014), and because the flowers also exhibit some pinkish color, we were not surprised to find anthocyanin-related transcripts (Appendix 4.7). We also found transcripts involved in chlorophyll accumulation in petals as *STAY-GREEN RICE-like* (SGR-like), *MAGNESIUM-CHELATASE SUBUNIT CHLI* 1 (CHLI1) and 7-HYDROXYMETHYL CHLOROPHYLL A REDUCTASE (HCAR); these transcripts have been reported for chlorophyll accumulation in carnation petals (Ohmiya *et al.* 2014). Although, anthocyanin pathway transcripts were present in *M. giganteus*, it is probable that the expression level of these transcripts is low, because the predominant color of the flowers is green. In other plant species, including *T. hassleriana*, transcripts from the following candidate genes are highly abundant: *CHALCONE SYNTHASE*, *CHALCONE ISOMERASE*, *DIHYDROFLAVONOL 4-REDUCTASE*, *ANTHOCYANIDIN 3-O-GLUCOSYLTRANSFERASE* and *ANTHOCYANIDIN SYNTHASE* (Bhide *et al.* 2014; Guo *et al.* 2014; Ho and Smith 2016; Wessinger and Rausher 2012). All these genes were present in the transcriptome of *M. giganteus*, *C. serrulata* and *P. dodecandra*. We did find that the F3'5'H transcript involved in the synthesis of delphinidin-derived pigments, which gives blue color to the flowers, was not expressed in our sample. The complete loss of this transcript has been reported in Rosaceae, most Asteraceae, *Antirrhinum majus*, *Arabidopsis thaliana*, *Ipomoea* spp, *Matthiola*, *Ophrys exaltata*, *O. sphegodes*, *O. garganica*, and *Tulipa* (Sedeek *et al.* 2013; Wessinger and Rausher 2012). Further investigation into gene expression levels based on RNA-seq experiments and qPCR are required to improve our knowledge of flower color and pollinator attraction in *M. giganteus*, *C. serrulata* and *P. dodecandra*, due to study limitation time, we did not perform these analyses. The next step should be to perform qPCR experiments to

measure gene expression levels in the candidate genes. Probably, more than differences in gene repertoires, the color differences observed among species may be the result of spatial and temporal gene expression in the core genes of the anthocyanin pathway. For example, gene repertoires and differential gene expression have been studied for Cleomaceae species with different mechanism of carbon fixation as C3 and C4 species. The main difference between them is differential gene expression and posttranscriptional regulation of many genes rather than gene repertoire, because C3 and C4 species shared the same relevant genes (Fankhauser and Aubry 2017; Külahoglu *et al.* 2014; Williams *et al.* 2016).

In addition to anthocyanins, transcription factors also play a key role in the determination of flower color. Transcription factors related to pollinator specificity and pollinator shifts among closely related species with differences in coloration have been identified in *Antirrhinum, Clarkia, Ipomoea, Mimulus, Ophrys, Petunia*, and *Phlox* species (Des Marais and Rausher 2010; Sedeek *et al.* 2013; Yuan *et al.* 2013*a*). In these systems, transcription factors belonging to the MYB family seem to play a prominent role in the diversification of pollinator-associated floral traits (Sedeek *et al.* 2013; Yuan *et al.* 2013*a;* Yuan *et al.* 2013*b*). MYB transcription factors are involved in the determination and regulation of flower color. For example, CPC can act as a negative regulator of anthocyanin production, which is what we anticipated to see in the white-flowered *P. dodecandra.* Surprisingly, this transcript was absent in this species, but present in the other focal species. Frequently, white phenotypes have a R2R3MYB gene inactivation, however, anthocyanins are still detected in vegetative traits. Although *P. dodecandra* have white flowers, we did not find a R2R3MYB gene inactivation in this species (Wessinger and Rausher 2012). We should also explore the MYB transcription factors in *C. serrulata* and determine their function in flower color regulation, as this species has many more MYB transcription factors compared to the other species (Figure 4.7). Further study is needed to determine the regulation function of MYB transcription factors in *M. giganteus*, *C. serrulata* and *P. dodecandra*.

4.4.4. Implications for understanding differences between generalist and specialist pollination systems

The evolution of specialist pollination systems has been widely studied, particularly in terms of the selective pressure of specialist pollinators on floral traits and pollinator shifts between closely related plant species (Clare *et al.* 2013; Fernandez-Mazuecos and Glover 2017; Gübitz *et al.* 2009; Hoballah *et al.* 2007). In these specialist systems, pollinators exert directional and stabilizing selection on floral traits, which frequently include corolla size, flower shape and color, nectar guides, stigma length, and flowering phenology (Armbruster 2012; Conner *et al.* 1996; Johnston 1991; José M. Gómez *et al.* 2006; Medel *et al.* 2007; Nattero *et al.* 2010; Sletvold *et al.* 2016). Researchers looking for genes underlying floral traits important for pollinator attraction have mainly focused on differences in genes coding for flower color, and the production of flower fragrances. These genes often differ between plants exhibiting different pollination syndromes promoting specialization, and they have been identified in a wide range of plants (Hoballah *et al.* 2007; Kessler *et al.* 2008; Kramer 2009*a;* Schiestl and Schlüter 2009; Sedeek *et al.* 2013; Sletvold *et al.* 2016; Spitzer-Rimon *et al.* 2012; Verdonk *et al.* 2005; Xu *et al.* 2012; Yuan *et al.* 2013*a*). On the other hand, the genetic basis of generalist pollination systems has been mostly studied to explain transitions between specialist and generalist pollination systems, and to study evolution of flower development (Alexandre *et al.* 2015). Currently, there are few studies that compare genomics and/or transcriptomics data obtained through NGS between plants exhibiting specialist or generalist pollination systems (Sedeek *et al.* 2013; Xu *et al.* 2012). Taken together, we found candidate genes coding for flower color, nectar glands, and nectar production; while we might predict that these traits are under directional and stabilizing selection in the specialist *M. giganteus*, the repertoire of genes involved in these floral traits was similar across specialist and generalist species. This suggests that selection is not leading to variable gene number in these different pollination systems.

Interestingly, we found some differences regarding presence/absence of the above genes, but they could not be attributed for sure to the different pollination system, because our sample sized was small (n=3) and specialist and generalist species were not well represented in our sampling. Moreover, these differences may be attributed to the

genetic pool of the species because they belong to different Cleomaceae clades and not to different pollination systems. Similar to this study, other researchers also found similarities in the repertoire of genes involved in flower color, nectary development, and floral scent between generalist and specialist species (Bender *et al.* 2012; Clegg and Durbin 2000; Kram *et al.* 2009; Młodzińska 2009; Ruhlmann *et al.* 2010; Winkel-Shirley 2001). It is important to clarify that the genetic bases of floral traits and the differences between generalist and specialist plant species are not always caused by the presence/absence of genes or gene repertoires. Gene differential, temporal and spatial expression, and regulatory networks also could play a role in the differences exhibited by generalist and specialist species.

Finally, to better identify candidate genes coding for pollinator mediated-selection traits in different pollination systems, it is necessary to perform quantitative gene expression studies as RNA-seq from only floral tissue (to get only the genes expressed in floral tissues, which are important for pollinator attraction). In addition, we need to include biological replicates to detect individual variation in the traits under study, and it is necessary to increase the number of species exhibiting both kind of pollination systems to find general patterns about gene expression in generalist and specialist systems.

4.5. Conclusions and future directions

We identified some interesting patterns using comparative transcriptomics in *M. giganteus*, *C. serrulata* and *P. dodecandra*. In general, the gene repertoires were similar among the species in terms of presence/absence. Overall, the transcriptome of species with generalized floral traits was more similar to each other than the species with specialized traits. However, we found differences in candidate genes encoding for flower color and nectary traits between species exhibiting generalized and specialized traits which could be subjected to selective pressure by pollinators. Although, the genes encoding for major floral traits are similar in specialist and generalist species, and the differences between floral traits is likely due to other genetic factors.

The next step in this work would be to complete quantitative gene expression in candidate genes and to determine if there are different expression levels in the candidate genes associated with key floral traits, which could suggest selection by pollinators. These

candidate genes could be the focus of functional manipulative experiments to test assumptions of pollinator-mediated selection on floral traits, which was the original aim. To better understand the evolution of pollination systems, these gene expression studies should incorporate both generalist and specialist pollination systems and be carried out in a wide taxonomic range of plants. The information derived from these studies will allow us to expand our knowledge about the evolution of pollination systems and to find the link between genes underlying important floral traits and pollinator-mediated selection of floral traits.

4.6. References

Alexandre H, Vrignaud J, Mangin B, Joly S. 2015. Genetic architecture of pollination syndrome transition between hummingbird-specialist and generalist species in the genus *Rhytidophyllum* (Gesneriaceae). *PeerJ* **3**: e1028.

Armbruster WS. 2012. Evolution and ecological implications of "specialized" pollinator rewards. In: Patiny S, ed. *Evolution of plant-pollinator relationships*. Cambridge; New York: Cambridge University Press, 44-67.

Bardou P, Mariette J, Escudié F, Djemiel C, Klopp C. 2014. jvenn: an interactive Venn diagram viewer. *BMC Bioinformatics* **15**: 293.

Barker MS, Vogel H, Schranz ME. 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.

Bender R, Klinkenberg P, Jiang Z, Bauer B, Karypis G, Nguyen N, Perera MADN, Nikolau BJ, Carter CJ. 2012. Functional genomics of nectar production in the Brassicaceae. *Flora - Morphology, Distribution, Functional Ecology of Plants* **207**: 491-496.

Bey M, Stuber K, Fellenberg K, Schwarz-Sommer Z, Sommer H, Saedler H, Zachgo S. 2004. Characterization of *Antirrhinum* petal development and identification of target genes of the class B MADS box gene DEFICIENS. *Plant Cell* **16**: 3197-3215.

Bhide A, Schliesky S, Reich M, Weber A, 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.

Busch A, Zachgo S. 2007. Control of corolla monosymmetry in the Brassicaceae *Iberis amara*. *Proceedings of the National Academy of Sciences of the United States of America* **104**: 16714-16719.

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

Cane JH. 2008. Breeding biologies, seed production and species-rich bee guilds of *Cleome lutea* and *Cleome serrulata* (Cleomaceae). *Plant Species Biology* **23**: 152-158.

Chanderbali AS, Berger BA, Howarth DG, Soltis PS, Soltis DE. 2016. Evolving ideas on the origin and evolution of flowers: New perspectives in the genomic era. *Genetics* **202**: 1255-1265.

Cheng S, van den Bergh E, Zeng P, Zhong X, Xu J, Liu X, Hofberger J, de Bruijn S, Bhide AS, Kuelahoglu C, Bian C, Chen J, Fan G, Kaufmann K, Hall JC, Becker A, Brautigam A, Weber AP, Shi C, Zheng Z, Li W, Lv M, Tao Y, Wang J, Zou H, Quan Z, Hibberd JM, Zhang G, Zhu XG, Xu X, Schranz ME. 2013. The *Tarenaya hassleriana* genome provides insight into reproductive trait and genome evolution of crucifers. *Plant Cell* **25**: 2813-2830.

Christinet L, Burdet FX, Zaiko M, Hinz U, Zrÿd J-P. 2004. Characterization and functional identification of a novel plant 4,5-extradiol dioxygenase involved in betalain pigment biosynthesis in *Portulaca grandiflora*. *Plant Physiology* **134**: 265-274.

Clare EL, Schiestl FP, Leitch AR, Chittka L. 2013. The promise of genomics in the study of plant-pollinator interactions. *Genome Biology* **14**: 207.

Clegg MT, Durbin ML. 2000. Flower color variation: A model for the experimental study of evolution. *Proceedings of the National Academy of Sciences* **97**: 7016-7023.

Conesa A, Götz S, García-Gómez JM, Terol J, Talón M, Robles M. 2005. Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics* **21**: 3674-3676.

Conner JK, Rush S, Jennetten P. 1996. Measurements of natural selection on floral traits in wild radish (*Raphanus raphanistrum*). I. Selection through lifetime female fitness. *Evolution* **50**: 1127-1136.

Davies KM. 2009. Modifying Anthocyanin Production in Flowers. In: Winefield C, Davies K, Gould K, eds. *Anthocyanins: Biosynthesis, Functions, and Applications*. New York, NY: Springer New York, 49-80.

Des Marais DL, Rausher MD. 2010. Parallel evolution at multiple levels in the origin of hummingbird pollinated flowers in *Ipomoea*. *Evolution* **64**: 2044-2054.

Enright AJ, Van Dongen S, Ouzounis CA. 2002. An efficient algorithm for large-scale detection of protein families. *Nucleic Acids Research* **30**: 1575-1584.

Fankhauser N, Aubry S. 2017. Post-transcriptional regulation of photosynthetic genes is a key driver of C4 leaf ontogeny. *Journal of Experimental Botany* **68**: 137-146.

Fenster CB, Armbruster WS, Wilson P, Dudash MR, Thomson JD. 2004. Pollination syndromes and floral specialization. *Annual Review of Ecology Evolution and Systematics* **35**: 375-403.

Feodorova TA, Voznesenskaya EV, Edwards GE, Roalson EH. 2010. Biogeographic patterns of diversification and the origins of C4 in *Cleome* (Cleomaceae). *Systematic Botany* **35**: 811-826.

Fernandez-Mazuecos M, Glover BJ. 2017. The evo-devo of plant speciation. *Nature Ecology & Evolution* **1**: 110.

Fleming TH, Geiselman C, Kress WJ. 2009. The evolution of bat pollination: A phylogenetic perspective. *Annals of Botany* **104**: 1017-1043.

Giurfa M, Núñez J, Chittka L, Menzel R. 1995. Colour preferences of flower-naive Honeybees. *Journal of Comparative Physiology A* **177**: 247-259.

Gould K, Davies KM, Winefield C. 2009. Anthocyanins biosynthesis, functions, and applications. New York: Springer.

Grotewold E. 2006. The genetics and biochemistry of floral pigments. *Annual Review of Plant Biology* **57**: 761-780.

Gübitz T, Hoballah ME, Dell'Olivo A, Kuhlemeier C. 2009. *Petunia* as a model system for the genetics and evolution of pollination syndromes. In: Gerats T, Strommer J, eds. *Petunia*. New York: Springer 29-49.

Guo N, Cheng F, Wu J, Liu B, Zheng S, Liang J, Wang X. 2014. Anthocyanin biosynthetic genes in *Brassica rapa*. *BMC Genomics* **15**: 426.

Haas BJ, Papanicolaou A, Yassour M, Grabherr M, Blood PD, Bowden J, Couger MB, Eccles D, Li B, Lieber M, MacManes MD, Ott M, Orvis J, Pochet N, Strozzi F, Weeks N, Westerman R, William T, Dewey CN, Henschel R, Leduc RD, Friedman N, Regev A. 2013. *De novo* transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. *Nature Protocols* **8**: 1494-1512.

Hatier J-HB, Gould KS. 2009. Anthocyanin Function in Vegetative Organs. In: Winefield C, Davies K, Gould K, eds. *Anthocyanins: Biosynthesis, Functions, and Applications*. New York, NY: Springer New York, 1-19.

He C, Saedler H. 2005. Heterotopic expression of MPF2 is the key to the evolution of the Chinese lantern of *Physalis*, a morphological novelty in Solanaceae. *Proceedings of the National Academy of Sciences of the United States of America* **102**: 5779-5784.

Herrera CM. 1996. Floral traits and plant adaptation to insect pollinators: A devil's advocate approach. In: Lloyd DG, Barrett SCH, eds. *Floral biology*. Boston, MA: Springer, 65-87.

Herrera CM. 2005. Plant generalization on pollinators: Species property or local phenomenon? *American Journal of Botany* **92**: 13-20.

Herrera CM, Pellmyr O. 2002. *Plant animal interactions: An evolutionary approach.* Oxford: Blackwell Science.

Higuera-Díaz M, Manson JS, Hall JC. 2015. Pollination biology of *Cleomella serrulata* and *Polanisia dodecandra* in a protected natural prairie in southern Alberta, Canada. *Botany* **93**: 745-757.

Hileman L, Cubas P. 2009. An expanded evolutionary role for flower symmetry genes. *Journal of Biology* **8**: 90.

Ho WW, Smith SD. 2016. Molecular evolution of anthocyanin pigmentation genes following losses of flower color. *BMC Evolutionary Biology* **16**: 98.

Hoballah ME, Gubitz T, Stuurman J, Broger L, Barone M, Mandel T, Dell'Olivo A, Arnold M, Kuhlemeier C. 2007. Single gene-mediated shift in pollinator attraction in *Petunia*. *Plant Cell* **19**: 779-790.

Honaas LA, Wafula EK, Wickett NJ, Der JP, Zhang Y, Edger PP, Altman NS, Pires JC, Leebens-Mack JH, dePamphilis CW. 2016. Selecting superior *De Novo* transcriptome assemblies: Lessons learned by leveraging the best plant genome. *PLoS ONE* **11**: e0146062.

Iltis HH, Cochrane TS. 2014. Cleomaceae. In: Davidse G, Sousa Sánchez S, Knapp S, Cabrera FC, eds. *Flora Mesoamericana*. St. Louis, Missouri: Missouri Botanical Garden Press, 256 -267.

Jin J, Tian F, Yang D-C, Meng Y-Q, Kong L, Luo J, Gao G. 2017. PlantTFDB 4.0: Toward a central hub for transcription factors and regulatory interactions in plants. *Nucleic Acids Research* **45**: D1040-D1045.

Johnson SD, Steiner KE. 2000. Generalization versus specialization in plant pollination systems. *Trends in Ecology & Evolution* **15**: 140-143.

Johnston MO. 1991. Natural selection on floral traits in two species of *Lobelia* with different pollinators. *Evolution* **45**: 1468-1479.

Jordheim M, Andersen ØM, Nozzolillo C, Amiguet VT. 2009. Acylated anthocyanins in inflorescence of spider flower (*Cleome hassleriana*). *Phytochemistry* **70**: 740-745.

José M. Gómez, Francisco Perfectti, Juan Pedro M. Camacho. 2006. Natural selection on *Erysimum mediohispanicum* flower shape: Insights into the evolution of zygomorphy. *The American Naturalist* **168**: 531-545.

Kanehisa M, Goto S. 2000. KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Research* **28**: 27-30.

Kessler D, Baldwin IT. 2011. Back to the past for pollination biology. *Current Opinion in Plant Biology* **14**: 429-434.

Kessler D, Bhattacharya S, Diezel C, Rothe E, Gase K, Schottner M, Baldwin IT. 2012. Unpredictability of nectar nicotine promotes outcrossing by hummingbirds in *Nicotiana attenuata*. *Plant Journal* **71**: 529-538.

Kessler D, Gase K, Baldwin IT. 2008. Field experiments with transformed plants reveal the sense of floral scents. *Science* **321**: 1200-1202.

Kram BW, Xu WW, Carter CJ. 2009. Uncovering the *Arabidopsis thaliana* nectary transcriptome: Investigation of differential gene expression in floral nectariferous tissues. *BMC Plant Biology* **9**: 92-92.

Kramer EM. 2007. Understanding the genetic basis of floral diversity. *BioScience* **57**: 479-487.

Kramer EM. 2009*a*. *Aquilegia*: A new model for plant development, ecology, and evolution. *Annual Review of Plant Biology* **60**: 261-277.

Külahoglu C, Denton AK, Sommer M, Mass J, Schliesky S, Wrobel TJ, Berckmans B, Gongora-Castillo E, Buell CR, Simon R, De Veylder L, Brautigam A, Weber AP. 2014. Comparative transcriptome atlases reveal altered gene expression modules between two Cleomaceae C3 and C4 plant species. *Plant Cell* **26**: 3243-3260.

Lehti-Shiu MD, Panchy N, Wang P, Uygun S, Shiu S-H. 2017. Diversity, expansion, and evolutionary novelty of plant DNA-binding transcription factor families. *Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms* **1860**: 3-20.

Li L, Stoeckert CJ, Roos DS. 2003. OrthoMCL: Identification of ortholog groups for Eukaryotic genomes. *Genome Research* **13**: 2178-2189.

Lin IW, Sosso D, Chen LQ, Gase K, Kim SG, Kessler D, Klinkenberg PM, Gorder MK, Hou BH, Qu XQ, Carter CJ, Baldwin IT, Frommer WB. 2014. Nectar secretion requires sucrose phosphate synthases and the sugar transporter SWEET9. *Nature* **508**: 546-549.

Ma H. 2001. Regulatory Genes in Plant Development: MADS. *eLS*John Wiley & Sons, Ltd.

Machado IC, Lopes AV, Leite AV, de Brito Neves C. 2006. *Cleome spinosa* (Capparaceae): Polygamodioecy and pollination by bats in urban and Caatinga areas, northeastern Brazil. *Botanische Jahrbücher* **127**: 69-82.

Martins DJ, Johnson SD. 2013. Interactions between hawkmoths and flowering plants in East Africa: Polyphagy and evolutionary specialization in an ecological context. *Biological Journal of the Linnean Society* **110**: 199-213.

Mayer C, Adler LS, Armbruster WS, Dafni A, Eardley C, Huang SQ, Kevan PG, Ollerton J, Packer L, Ssymank A, Stout JC, Potts SG. 2011. Pollination ecology in

the 21st Century: Key questions for future research. *Journal of Pollination Ecology* **3**: 8-23.

Medel R, Valiente A, Botto-Mahan C, Carvallo G, Pérez F, Pohl N, Navarro L. 2007. The influence of insects and hummingbirds on the geographical variation of the flower phenotype in *Mimulus luteus*. *Ecography* **30**: 812-818.

Mitchell RJ, Irwin RE, Flanagan RJ, Karron JD. 2009. Ecology and evolution of plant-pollinator interactions. *Annals of Botany* **103**: 1355-1363.

Młodzińska E. 2009. Survey of plant pigments: molecular and environmental determinants of plant colors. *Acta Biologica Cracoviensia. Series Botanica* **51**: 7-16.

Moss EH, Packer JG. 1983. *Flora of Alberta: A manual of flowering plants, conifers, ferns, and fern allies found growing without cultivation in the Province of Alberta, Canada.* Toronto; Buffalo: University of Toronto Press.

Nagpal P, Ellis CM, Weber H, Ploense SE, Barkawi LS, Guilfoyle TJ, Hagen G, Alonso JM, Cohen JD, Farmer EE, Ecker JR, Reed JW. 2005. Auxin response factors ARF6 and ARF8 promote jasmonic acid production and flower maturation. *Development* **132**: 4107-4118.

Nattero J, Cocucci AA, Medel R. 2010. Pollinator-mediated selection in a specialized pollination system: Matches and mismatches across populations. *Journal of Evolutionary Biology* **23**: 1957-1968.

Neal PR, Dafni A, Giurfa M. 1998. Floral symmetry and its role in plant-pollinator systems: Terminology, distribution, and hypotheses. *Annual Review of Ecology and Systematics* **29**: 345-373.

Nozzolillo C, Amiguet VT, Bily AC, Harris CS, Saleem A, Andersen OM, 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, Hirashima M, Yagi M, Tanase K, Yamamizo C. 2014. Identification of genes associated with chlorophyll accumulation in flower petals. *PLoS ONE* **9**: e113738.

Ollerton J, Alarcón R, Waser NM, Price MV, Watts S, Cranmer L, Hingston A, Peter CI, Rotenberry J. 2009. A global test of the pollination syndrome hypothesis. *Annals of Botany* **103**: 1471-1480.

Ortiz-Barrientos D. 2013. The color genes of speciation in plants. *Genetics* **194**: 39-42.

Owen CR, Bradshaw HD. 2011. Induced mutations affecting pollinator choice in *Mimulus lewisii* (Phrymaceae). *Arthropod-Plant Interactions* **5**: 235-244.

Panchy N, Lehti-Shiu M, Shiu S-H. 2016. Evolution of gene duplication in plants. *Plant Physiology* **171**: 2294-2316.

Patchell MJ, Bolton MC, Mankowski P, Hall JC. 2011. Comparative floral development in Cleomaceae reveals two distinct pathways leading to monosymmetry. *International Journal of Plant Sciences* **172**: 352-365.

Patchell MJ, Roalson EH, Hall JC. 2014. Resolved phylogeny of Cleomaceae based on all three genomes. *Taxon* **63**: 315-328.

Preston JC, Hileman LC. 2009. Developmental genetics of floral symmetry evolution. *Trends in Plant Science* **14**: 147-154.

Proctor MCF, Yeo P, Lack A. 1996. *The natural history of pollination.* Portland: Timber Press.

Quattrocchio F, Wing J, van der Woude K, Souer E, de Vetten N, Mol J, Koes R. 1999. Molecular analysis of the anthocyanin2 gene of *Petunia* and its role in the evolution of flower color. *Plant Cell* **11**: 1433-1444.

Raju AJS, Rani DS. 2016. Reproductive ecology of *Cleome gynandra* and *Cleome viscosa* (Capparaceae). *Phytologia Balcanica* **22**: 15-28.

Roy R, Schmitt AJ, Thomas JB, Carter CJ. 2017. Review: Nectar biology: From molecules to ecosystems. *Plant Science* **262**: 148-164.

Ruhlmann JM, Kram BW, Carter CJ. 2010. CELL WALL INVERTASE 4 is required for nectar production in *Arabidopsis*. *Journal of Experimental Botany* **61**: 395-404.

Ruiz-Zapata T, Huérfano AA, Enrech N. 1996. Contribución al estudio citotaxonómico del género *Cleome* L. (Capparidaceae). *Phyton* **59**: 85-94.

Sapir Y. 2009. Pollination genetics: Using molecular genetic underlying floral traits to study pollination ecology in an evolutionary context. *Israel Journal of Plant Sciences* **57**: 141-149.

Schemske DW, Bradshaw HD. 1999. Pollinator preference and the evolution of floral traits in Monkey flowers (*Mimulus*). *Proceedings of the National Academy of Sciences of the United States of America* **96**: 11910-11915.

Schiestl FP, Johnson SD. 2013. Pollinator-mediated evolution of floral signals. *Trends in Ecology & Evolution* **28**: 307-315.

Schiestl FP, Schlüter PM. 2009. Floral isolation, specialized pollination, and pollinator behavior in orchids. *Annual Review of Entomology* **54**: 425-446.

Sedeek KE, Qi W, Schauer MA, Gupta AK, Poveda L, Xu S, Liu ZJ, Grossniklaus U, Schiestl FP, Schluter PM. 2013. Transcriptome and proteome data reveal candidate genes for pollinator attraction in sexually deceptive orchids. *PLoS ONE* **8**: e64621.

Shcheglov AS, Zhulidov PA, Bogdanova EA, Shagin DA. 2007. Normalization of cDNA libraries. In: Buzdin AA, Lukyanov SA, eds. *Nucleic Acids Hybridization Modern Applications*. Dordrecht: Springer Netherlands, 97-124.

Simão FA, Waterhouse RM, Ioannidis P, Kriventseva EV, Zdobnov EM. 2015. BUSCO: Assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics* **31**: 3210-3212.

Sletvold N, Trunschke J, Smit M, Verbeek J, Ågren J. 2016. Strong pollinator-mediated selection for increased flower brightness and contrast in a deceptive orchid. *Evolution* **70**: 716-724.

Smith-Unna R, Boursnell C, Patro R, Hibberd JM, Kelly S. 2016. TransRate: Reference-free quality assessment of *de novo* transcriptome assemblies. *Genome Research* **26**: 1134-1144.

Soares Neto RL. 2017. Flora das cangas da Serra dos Carajás, Pará, Brasil: Cleomaceae. *Rodriguésia* **68**: 931-933.

Specht C, Yockteng R, Almeida A, Kirchoff B, Kress WJ. 2012. Homoplasy, pollination, and emerging complexity during the evolution of floral development in the tropical Gingers (Zingiberales). *The Botanical Review* **78**: 440-462.

Spitzer-Rimon B, Farhi M, Albo B, Cna'ani A, Ben Zvi MM, Masci T, Edelbaum O, Yu Y, Shklarman E, Ovadis M, Vainstein A. 2012. The R2R3-MYB-like regulatory factor EOBI, acting downstream of EOBII, regulates scent production by

activating ODO1 and structural scent-related genes in *Petunia*. *Plant Cell* **24**: 5089-5105.

van den Bergh E. 2017. *Comparative genomics and trait evolution in Cleomaceae, a model for ancient polyploidy,* PhD, Wageningen University

van den Bergh E, Hofberger JA, Schranz ME. 2016. Flower power and the mustard bomb: Comparative analysis of gene and genome duplications in glucosinolate biosynthetic pathway evolution in Cleomaceae and Brassicaceae. *American Journal of Botany* **103**: 1212-1222.

van den Bergh E, Külahoglu C, Bräutigam A, Hibberd JM, Weber AP, Zhu X-G, Schranz 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 Niet T, Peakall R, Johnson SD. 2014. Pollinator-driven ecological speciation in plants: New evidence and future perspectives. *Annals of Botany* **113**: 199-212.

Verdonk JC, Haring MA, van Tunen AJ, Schuurink RC. 2005. ODORANT1 regulates fragrance biosynthesis in *Petunia* flowers. *Plant Cell* **17**: 1612-1624.

Wanscher JH. 1934. The Basic Chromosome Number of the Higher Plants. *New Phytologist* **33**: 101-126.

Waser NM, Chittka L, Price MV, Williams NM, Ollerton J. 1996. Generalization in pollination systems, and why it matters. *Ecology* **77**: 1043-1060.

Wessinger CA. 2015. A genetic route to yellow flowers. *New Phytologist* **206**: 1193-1195.

Wessinger CA, Rausher MD. 2012. Lessons from flower colour evolution on targets of selection. *Journal of Experimental Botany* **63**: 5741-5749.

Wessinger CA, Rausher MD. 2014. Predictability and irreversibility of genetic changes associated with flower color evolution in *Penstemon barbatus*. *Evolution* **68**: 1058-1070.

Whittall JB, Voelckel C, Kliebenstein DJ, Hodges SA. 2006. Convergence, constraint and the role of gene expression during adaptive radiation: Floral anthocyanins in *Aquilegia*. *Molecular Ecology* **15**: 4645-4657.

Williams BP, Burgess SJ, Reyna-Llorens I, Knerova J, Aubry S, Stanley S, Hibberd JM. 2016. An untranslated cis element regulates the accumulation of multiple C4 enzymes in *Gynandropsis gynandra* mesophyll cells. *The Plant Cell* **28**: 454-465.

Willmer P. 2011. *Pollination and floral ecology.* Princeton: Princeton University Press.

Winkel-Shirley B. 2001. Flavonoid biosynthesis. A colorful model for genetics, biochemistry, cell biology, and biotechnology. *Plant Physiology* **126**: 485-493.

Xu SQ, Schluter PM, Grossniklaus U, Schiestl FP. 2012. The genetic basis of pollinator adaptation in a sexually deceptive orchid. *Plos Genetics* **8**.

Yuan YW, Byers KJRP, Bradshaw HD. 2013*a*. The genetic control of flower-pollinator specificity. *Current Opinion in Plant Biology* **16**: 422-428.

Yuan YW, Saqawa JM, Young RC, Christensen BJ, Bradshaw HDJ. 2013*b*. Genetic dissection of a major anthocyanin QTL contributing to pollinator-mediated reproductive isolation between sister species of *Mimulus*. *Genetics* **194**: 255-263.

Zhang S, Zhang J-S, Zhao J, He C. 2015. Distinct subfunctionalization and neofunctionalization of the B-class MADS-box genes in *Physalis floridana*. *Planta* **241**: 387-402.

Zhang W, Kramer EM, Davis CC. 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

Table 4.1. Floral differences among *Cleomella serrulata*, *Melidiscus giganteus,* and *Polanisia dodecandra. Cleomella serrulata* and *Polanisia dodecandra* are pollinated by different pollinator guilds, while *Melidiscus giganteus* is pollinated by bats.

	Cleomella serrulata	***Melidiscus giganteus***	***Polanisia dodecandra***
Aestivation	Closed	Closed	Open
Inflorescence type	Bracteated raceme	Bracteated raceme	Bracteated raceme
Flower length	3 cm	7.5	2.5 cm
Petals			
curvature	Upward	Upward	Upward
orientation	Upward	Upward	Upward
differentiation	Adaxial petals shorter	Similar size	Adaxial petals longer
shape	Oblanceolate	Oblong	Clawed
color	Purple, pink, white or mixed	Whitish-Green	White
Nectar gland			
colour	Green	N/A*	Orange
position and display	Adaxial and concealed	N/A*	Adaxial and exposed
Shape	Elongated and laminar	N/A*	Cup shape

Nectar			
appearance	Transparent, copious and lighter	N/A	Yellowish, copious and thick
sugar concentration**	40.8 ± 7.68%	N/A	43.32 ± 4.74%
UV reflection	Absent	Absent	Fluorescent nectar
volume**	0.28 – 0.85 ul	N/A	0.25 – 0.63 ul
Stamens			
stamen curvature	Upward	Upward	Upward
length	Equal	Equal	Unequal
number	6	8	8 – 12
Anthers			
colour	Green, yellow	Green, yellow	Green, purplish, yellow
shape	Oblong	Linear	Oval
pollen color	Yellow	Yellow	Green, purplish and yellow
pollen display	Conspicuous, yellow	Conspicuous, yellow	Conspicuous, green, yellow

Gynophore	As long as stamens	As long as stamens	Short (0 – 2mm)

* We did not observe a nectar gland in *Melidiscus giganteus*. Flowers of plants growing in the greenhouse did not have nectar at all.

** Adapted from Higuera-Díaz *et al.* 2015. Nectar sugar concentration (mass/mass as percentage), concentration values are population means ± standard deviation, n= 20. Minimum and maximum nectar volume range, n= 20.

N/A= No data available.

Table 4.2. Summary of Illumina sequencing from DNS normalized libraries.

	Cleomella serrulata	***Melidiscus giganteus***	***Polanisia dodecandra***
Total number of raw reads	95 616 414	130 268 202	84 853 538
Total nucleotides	9 561 641 400	13 026 820 200	8 485 353 800
Average length	100	100	100

Table 4.3. *Cleomella serrulata, Melidiscus giganteus,* and *Polanisia dodecandra* assemblies. Assembly comparison between CLC Workbench and Trinity software.

	Cleomella serrulata		***Melidiscus giganteus***		***Polanisia dodecandra***	
	CLC	**Trinity**	**CLC**	**Trinity**	**CLC**	**Trinity**
No of assembled contigs/isotigs	79103	186361	68951	240976	59220	140014
No of unique assembled contigs	79103	76228	68951	55503	59220	47090
Minimum contig length (bp)	156	201	200	201	146	201
Maximum contig length (bp)	16513	16513	16538	15678	15239	15234
Average contig length (bp)	625	1671.76	669	2202.47	696	1986.25
N50 (bp)	888	2788	1024	2855	1075	3021

Table 4.4. BUSCO notations for estimated assembly completeness and TransRate quality assessments in *Cleomella serrulata*, *Melidiscus giganteus* and *Polanisia dodecandra* CLC assembly and Trinity complete and reference assemblies.

	Cleomella serrulata			***Melidiscus giganteus***			***Polanisia dodecandra***		
		Trinity			**Trinity**			**Trinity**	
	CLC	**Complete**	**Reference**	**CLC**	**Complete**	**Reference**	**CLC**	**Complete**	**Reference**
Contig metrics									
No of transcripts	79103	186361	76228	68951	240976	55503	59220	140014	47090
BUSCO notations*, n= 1440*									
% Complete	60.7	92.3	77.2	60.7	90.9	74	61.7	94.2	76.9
Single copy	58.5	45.8	75.1	57.2	41	73.1	60.1	43.5	75.5
Duplicated	2.2	46.5	2.1	3.5	49.9	0.9	1.6	50.7	1.4
% Fragmented	19.1	2.9	11.1	21.4	4	12.6	20.9	2.9	12.5
% Missing	20.2	4.8	11.7	17.9	5.1	13.4	17.4	2.9	10.6
TransRate score									
TransRate assembly score	0.25	0.021	0.33	0.24	0.009	0.25	0.28	0.018	0.34
% Good contigs	73.65	35.57	84.67	72.57	16.5	84.71	76.47	26.99	85.81

Table 4.5. Expectations and agreement/disagreement results regarding gene repertoires distribution for selected floral traits in the generalist species *Cleomella serrulata* and *Polanisia dodecandra* and the specialist species *Melidiscus giganteus*.

Expectations	Agree/Disagree
- Genes related with anthocyanin pathway will differ between species, because the species exhibit different petal color	Disagree
- Genes related with nectary development will be more similar to each other in species exhibiting generalized floral traits than in species with specialized floral traits	Agree
- Genes related with nectar production will be more similar to each other in species exhibiting generalized floral traits than in species with specialized floral traits	Disagree
- Genes related with nectar production will be higher in the species exhibiting a bat pollination syndrome	Agree

Figure 4.1. Current phylogenetic hypothesis of Cleomaceae depicting the main clades to which belong, *Cleomella serrulata*, *Melidiscus giganteus*, and *Polanisia dodecandra*. Adapted from Patchell *et al*. 2014.

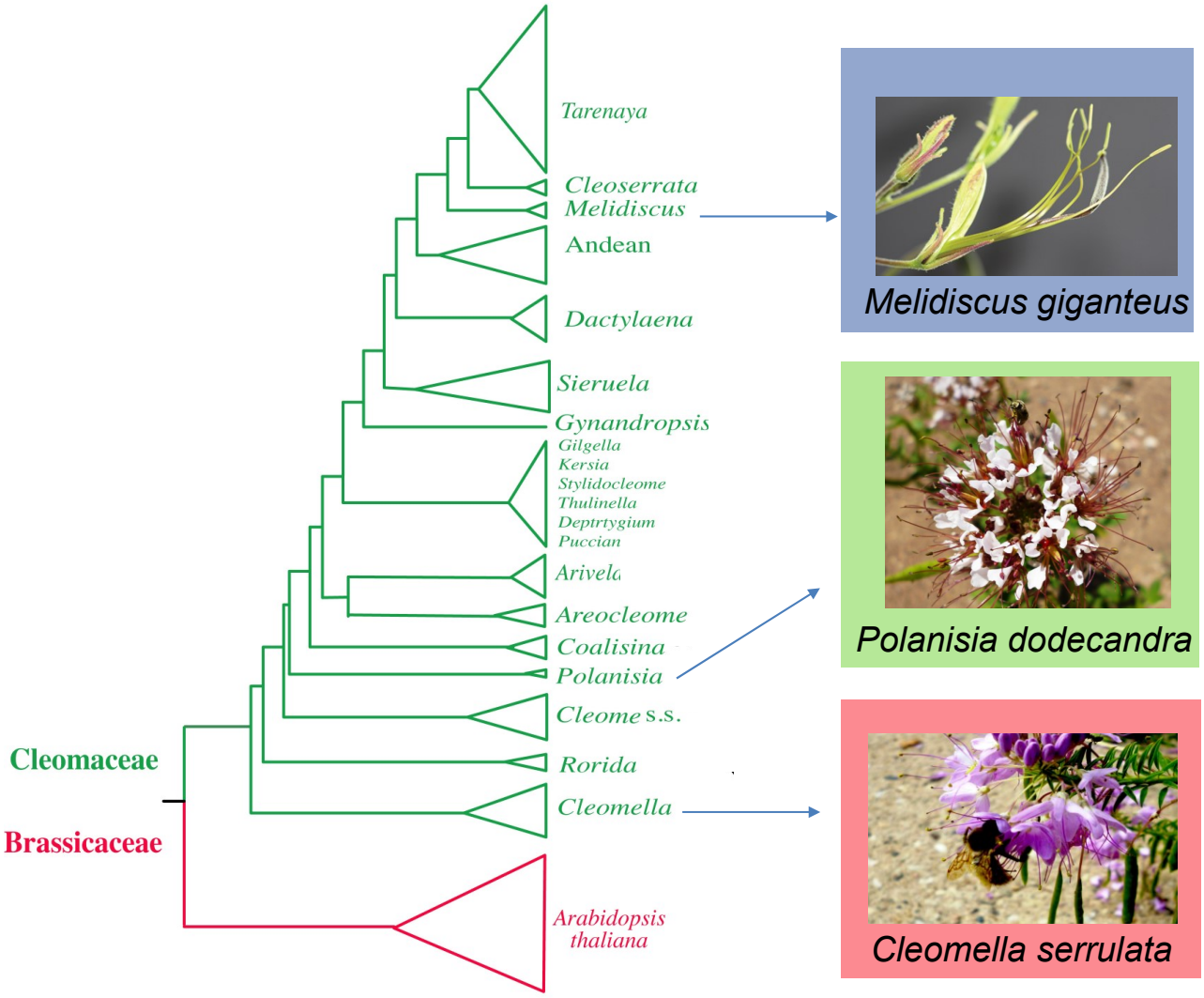


Figure 4.2. Top 10 gene ontology annotations for the reference transcriptome of *Cleomella serrulata*, *Melidiscus giganteus*, and *Polanisia dodecandra*. BP: Biological process, MF: Molecular function, CC: Cellular component.

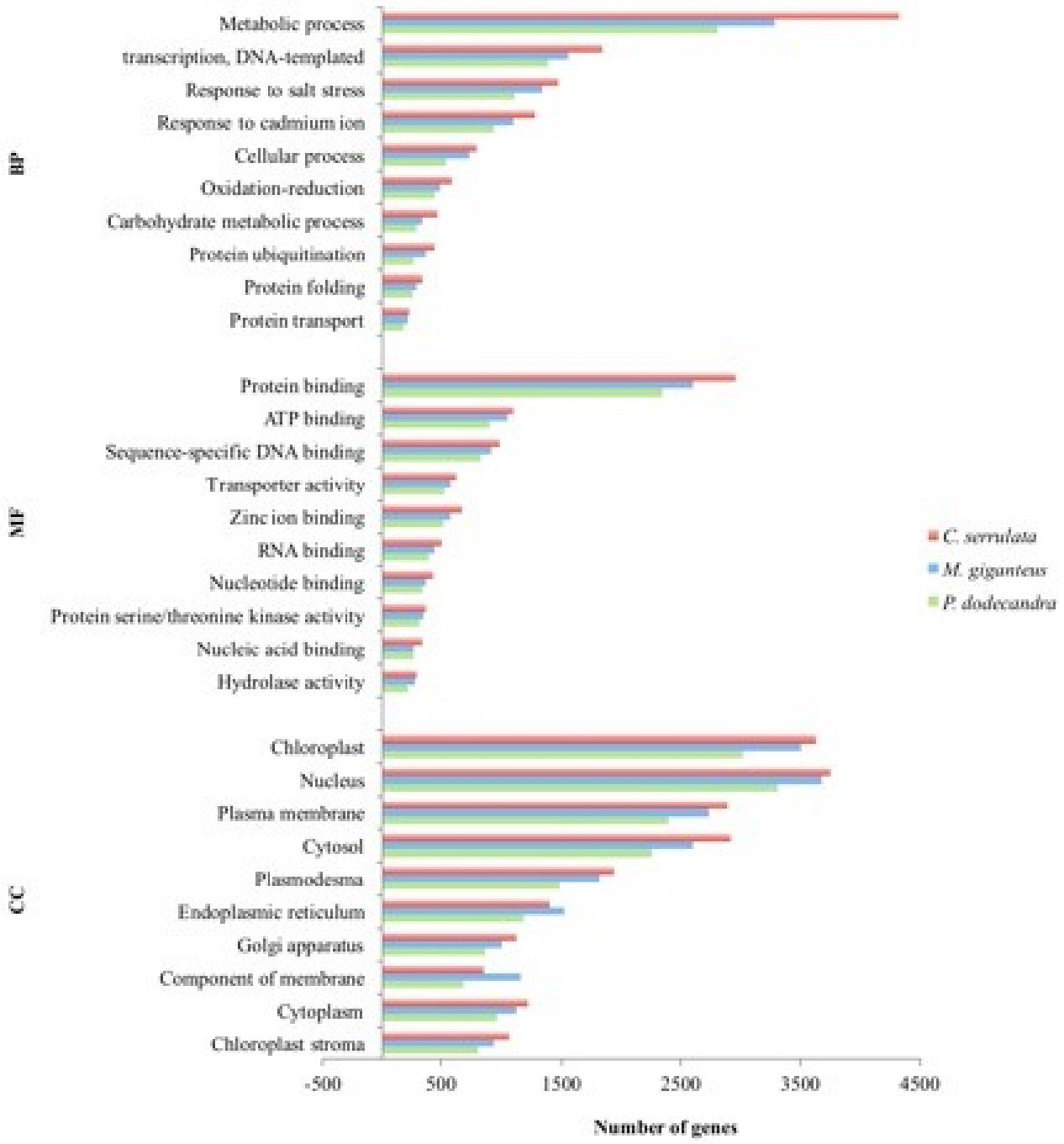


Figure 4.3. Floral and pollination related gene ontology terms for the reference transcriptome of *Cleomella*. *serrulata*, *Melidiscus giganteus*, and *Polanisia*. *dodecandra*. BP: Biological process.

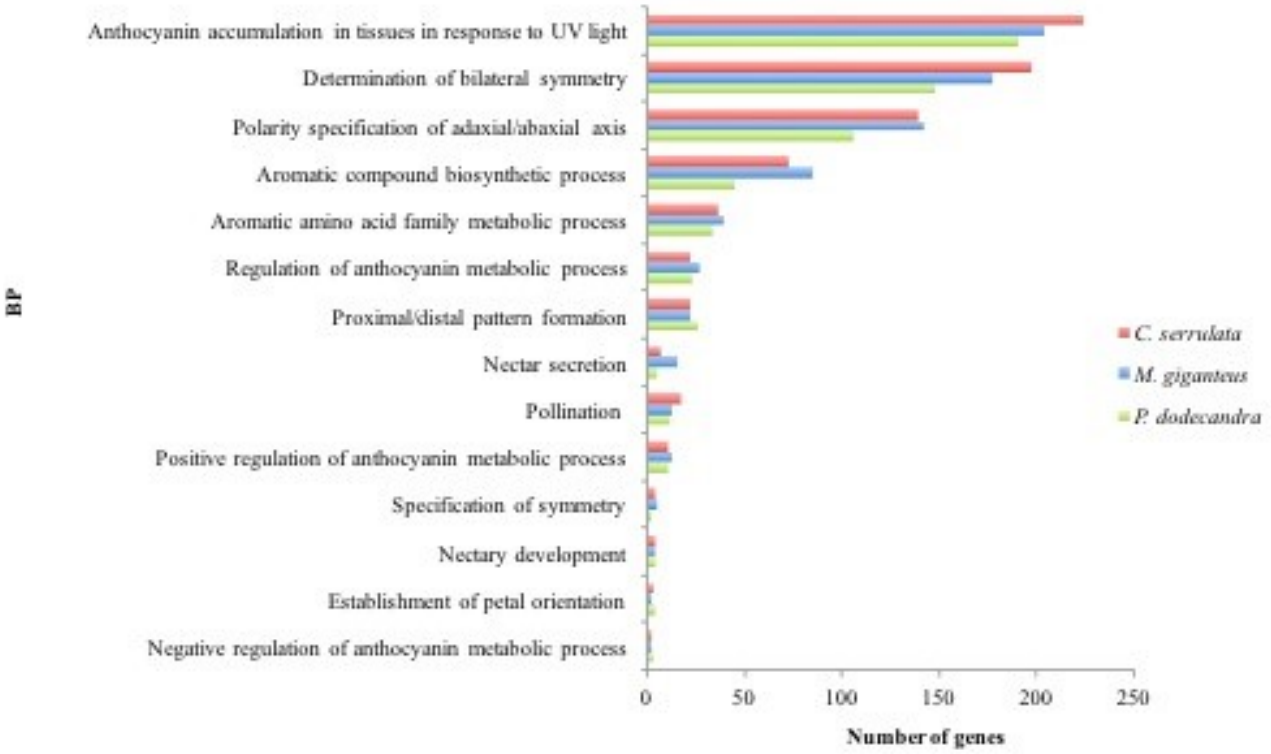


Figure 4.4. Comparison of the number of shared and specific transcripts related to pollination gene ontology terms in *Cleomella serrulata*, *Melidiscus giganteus*, and *Polanisia dodecandra*.

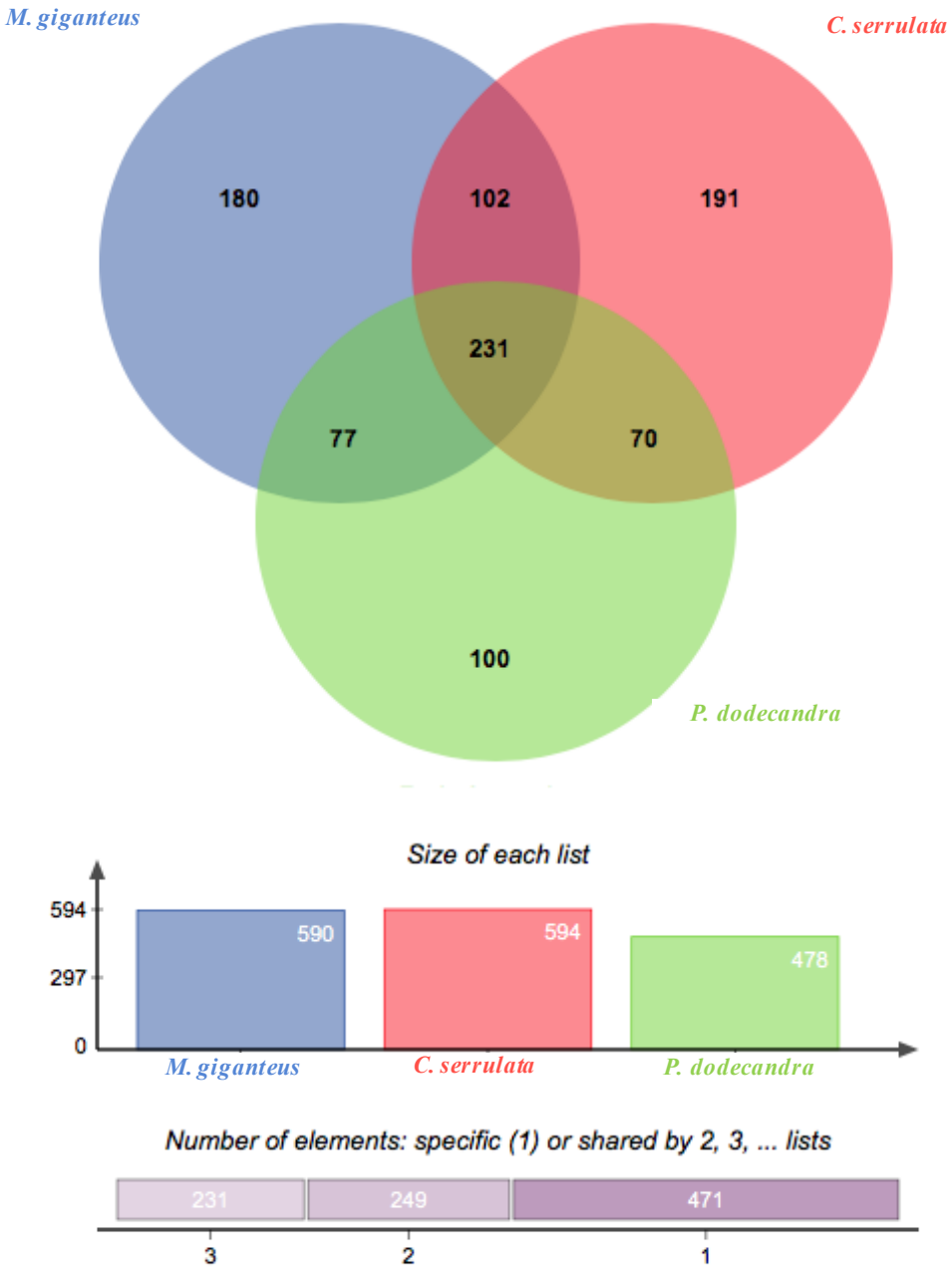


Figure 4.5. Comparison of the number of shared and specific transcripts related to specific pollination gene ontology terms in *Cleomella serrulata*, *Melidiscus giganteus*, and *Polanisia dodecandra*. (A) Anthocyanin accumulation in tissues in response to UV light GO term. (B) Nectary development GO term. (C) Nectar production GO term.

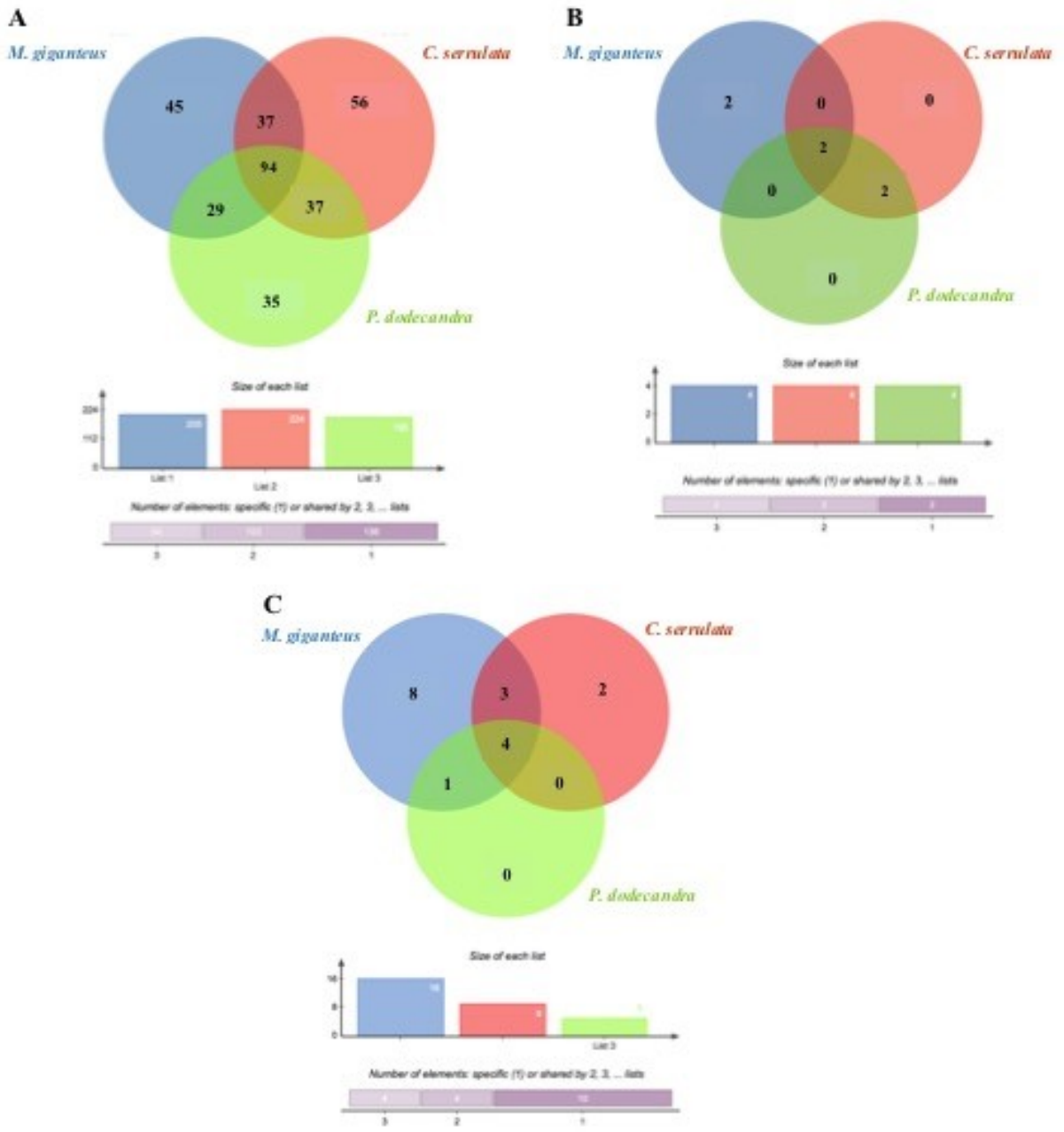


Figure 4.6. Comparison of the number of shared and species-specific orthologous protein groups from OrthoMCL analysis in *Cleomella serrulata*, *Melidiscus giganteus*, and *Polanisia dodecandra*.

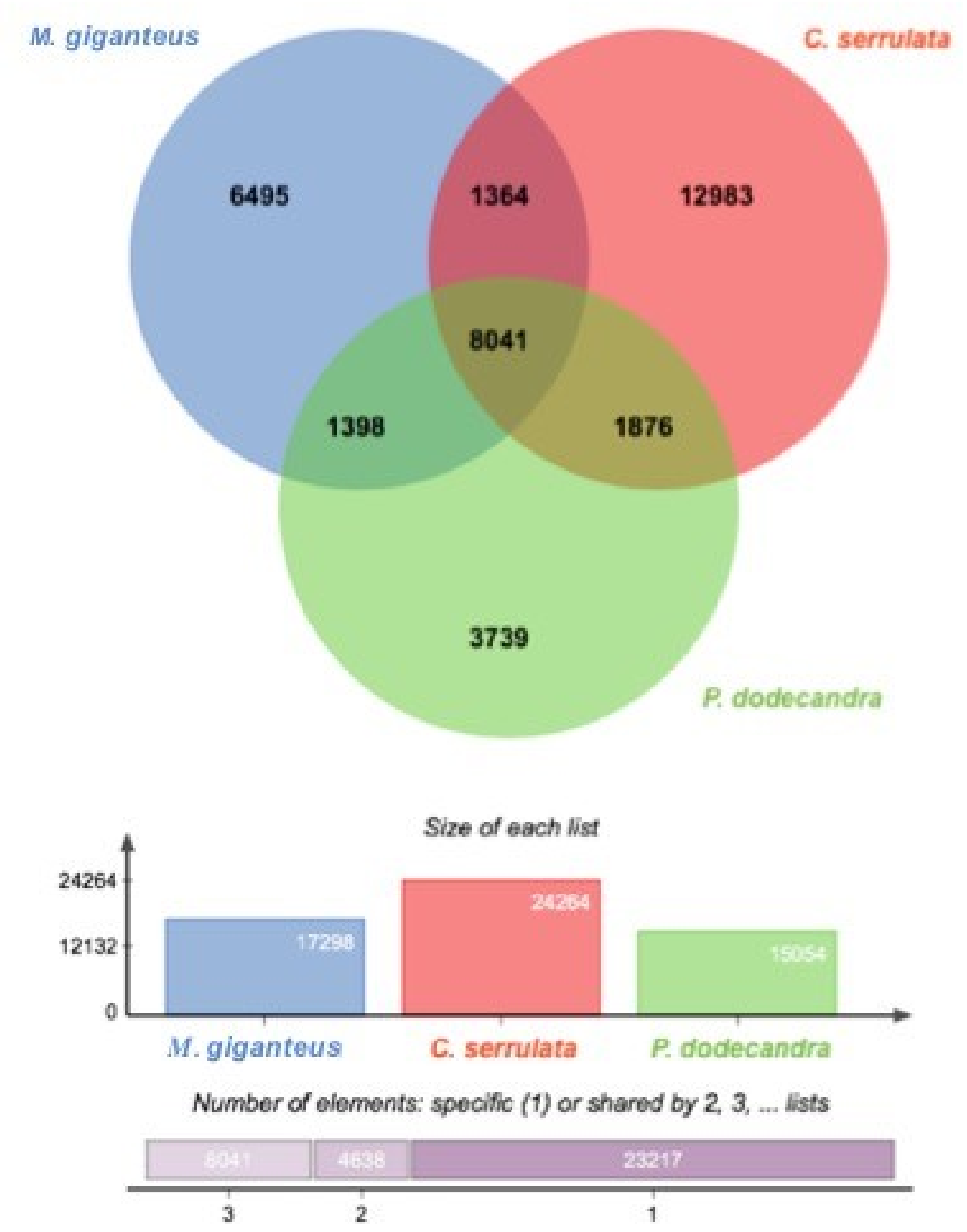


Figure 4.7. Bar plot showing the number of transcripts in different transcription factor families for Trinity reference transcriptome of *Cleomella serrulata, Melidiscus giganteus,* and *Polanisia dodecandra. NZZ/SPL, RAV, and SAP* families were absented in *Melidiscus giganteus,* while*, HB-PHD, HRT-Like,* and *SAP* families were absented *in Polanisia dodecandra.*

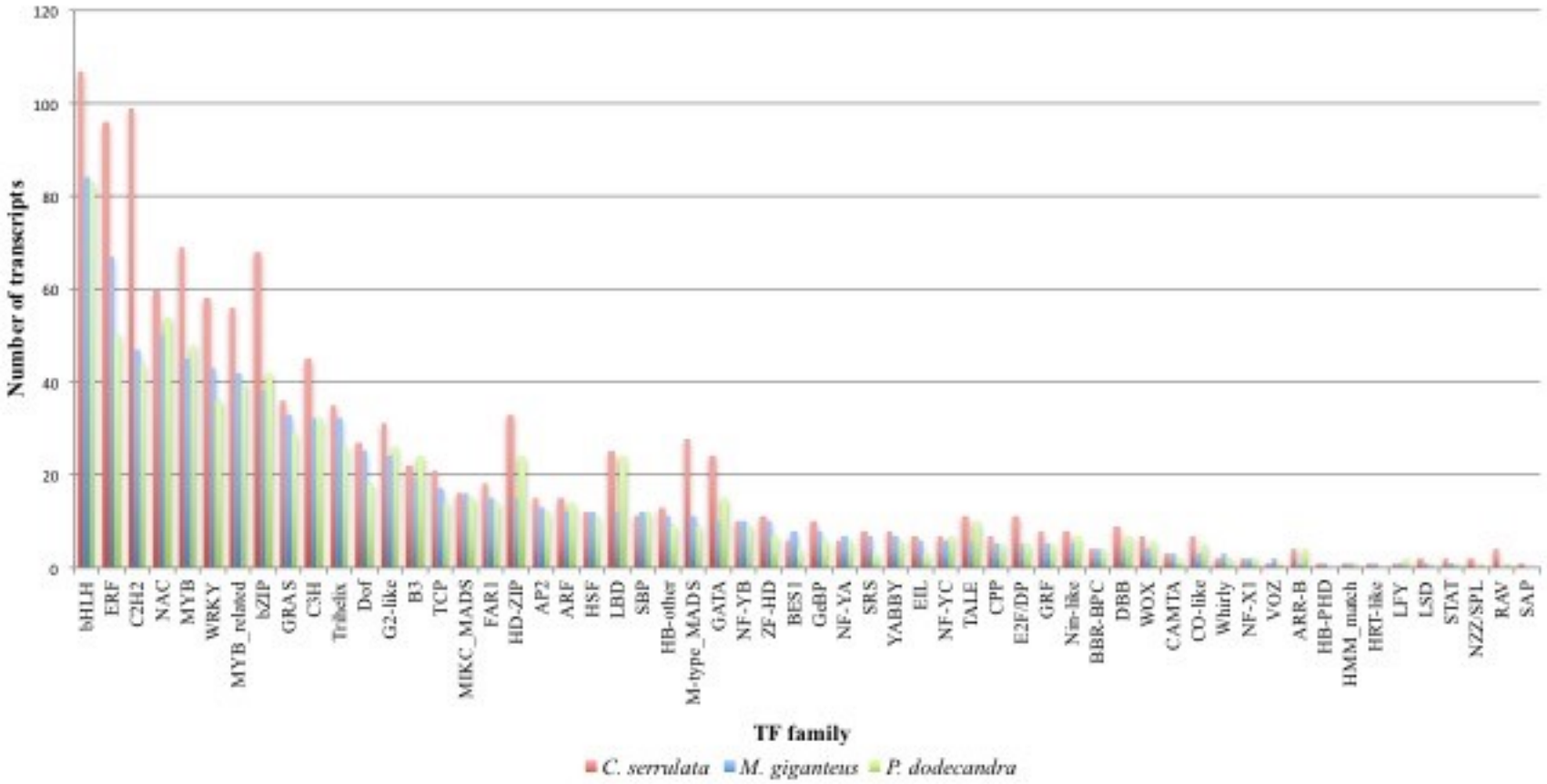
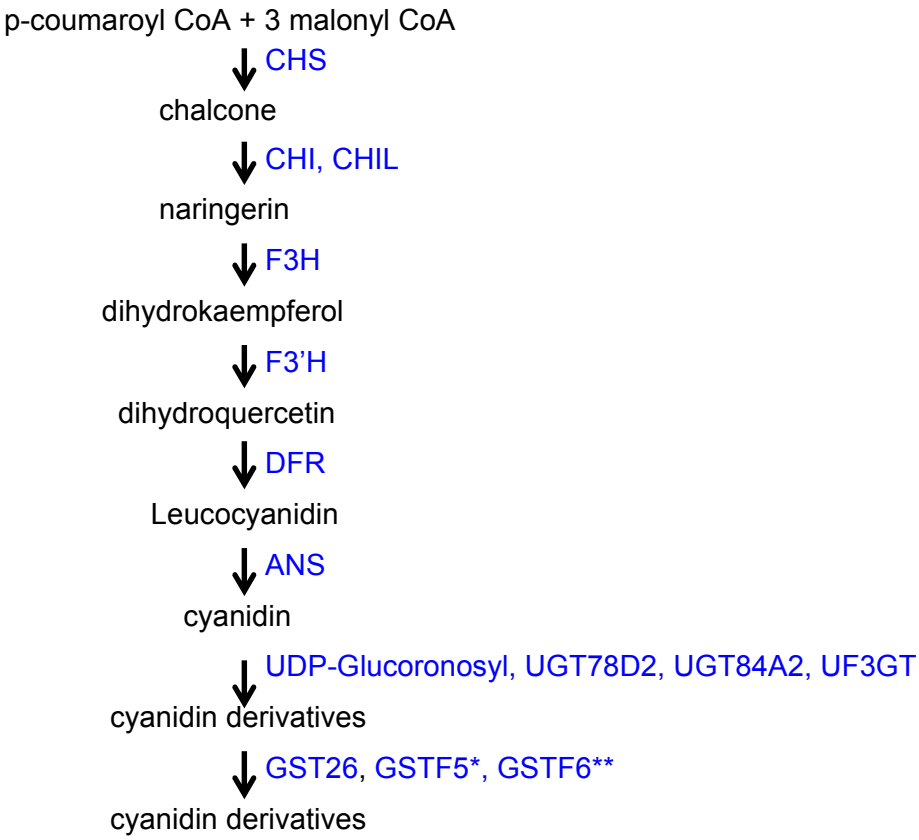


Figure 4.8. Anthocyanin pathway in *Cleomella serrulata, Melidiscus giganteus,* and *P. dodecandra.* Enzymes are indicated in black, main genes involved in the pathway are in blue. *CHS, CHALCONE SYNTHASE; CHI, CHALCONE ISOMERASE; CHIL, CHALCONE-FLAVONONE ISOMERASE: F3H, FLAVONONE 3-HYDROXYLASE; F3'H, FLAVONOID 3'-HYDROXYLASE; DFR, DIHYDROKAEMPFEROL 4-REDUCTASE; ANS, ANTHOCYANIDIN SYNTHASE; UDP-GLUCORONOSYL, ANTHOCYANIDIN 5-O-GLUCOSYLTRANSFERASE; UGT78D2, ANTHOCYANIDIN 3-O-GLUCOSYLTRANSFERASE; UGT84A2, SINAPATE1-GLUCOSYLTRANSFERASE; UF3GT, ANTHOCYANIN 3-O-GLUCOSIDE; GST26, GLUTATIONE TRANSFERASE; GSTF5, GLUTATIONE TRANSFERASE; GSTF6, GLUTATIONE TRANSFERASE.* * Present only in *M. giganteus*. ** Absent in *M. giganteus*. Adapted from Wessinger and Rausher (2012); Sedeek et al. (2013).



Chapter 5: Virus-Induced Gene Silencing in *Cleome violacea* and *Polanisia dodecandra* (Brassicales: Cleomaceae)

5.1. Introduction

Cleomaceae is a cosmopolitan plant family comprising approximately 270 species (Patchell *et al.* 2014). Because Cleomaceae species exhibit different floral traits that attract different pollinator guilds, floral diversity and evolution of pollination systems are promising areas for further research in this family. Species of Cleomaceae display a range of floral diversity with regards to organ number and morphology despite the common pattern that most flowers are open with perianth and reproductive organs curved upwards (Endress 1992; Iltis *et al.* 2011a; Patchell *et al.* 2011). Further, the genomic resources in the family are increasing rapidly, making it an excellent candidate for studying the genetic basis of floral traits. Cleomaceae therefore represents a good candidate system to investigate the evolution of pollination systems, including the genetic basis of floral traits that may influence pollinator interactions. However, some species are difficult to growth in growth chambers.

Cleomaceae is a promising group to investigate key ecological and evolutionary phenomena, and frequently functional genetic tools are needed to address questions regarding to these phenomena. Thus, the family has been the focus of a broad range of research including evolution of C_4 photosynthesis (Brautigam *et al.* 2011a; Marshall *et al.* 2007; Voznesenskaya *et al.* 2007; Williams *et al.* 2016), comparative genomics/transcriptomics including Brassicaceae comparisons (Barker *et al.* 2009; Bhide *et al.* 2014; Braütigam *et al.* 2011b; Cheng *et al.* 2013; Külahoglu *et al.* 2014; Schranz and Mitchell-Olds 2006b), floral morphology and development (Endress 1992; Nozzolillo *et al.* 2010; Patchell *et al.* 2011), glucosinolates or mustard oils evolution (Edger *et al.* 2015; van den Bergh *et al.* 2016b), and pollination biology (Cane 2008a; Fleming *et al.* 2009; Higuera-Díaz *et al.* 2015; Machado *et al.* 2006; Martins and Johnson 2013; Raju and Rani 2016). Using forward and reverse genetic tools we can generate flowers with altered phenotypes, which can help us to establish which genes are responsible for key floral traits and their effects in pollinator attraction.

Genetic methods have been successfully used in many species of plants to obtain individuals with modified phenotypes (Deng *et al.* 2012; Newell *et al.* 2010; Owen and Bradshaw 2011; Weigel and Glazebrook 2002). Forward genetics tools create alternative phenotypes, but without knowing which genes contributed to the alternative phenotype. Methods include chemical mutagenesis using a mutagen agent, usually EMS (ethyl methanesulfonate), and TILLING (targeting induced local lesions in genomes). On the other hand, reverse genetics tools are useful to test the role(s) of candidate genes underlying traits of interest. Reverse genetics are comprised of techniques such as gene silencing and insertional mutagenesis. Within Cleomaceae, stable transformation techniques have been developed in *Gynandropsis gynandra* (L.) Briq. (Newell *et al.* 2010) and *Tarenaya spinosa* (Jacq) Raf. (Tsai *et al.* 2012) but are quite time-consuming and intensive. A viable alternative approach is virus-induced gene silencing, or VIGS.

VIGS is a well-established post-transcription gene silencing (PTGS) reverse genetic tool used in model and non-model species (Becker 2013; Becker and Lange 2010; Burch-Smith *et al.* 2004; Burch-Smith *et al.* 2006; Di Stilio *et al.* 2010; Gould and Kramer 2007). VIGS uses RNA interference to exploit plant defense which prevents endogenous mRNA from being translated (Baulcombe 1999; Burch-Smith *et al.* 2004; Ruiz *et al.* 1998). This technique requires the use of a viral construct containing a fragment of the target gene to elicit the PTGS response (Baulcombe 1999; Ruiz *et al.* 1998). Infection of the viral construct is mediated by *Agrobacterium* transformation followed by infiltration into target plants (Dinesh-Kumar *et al.* 2003). Thus, VIGS provides the opportunity to knock down one or more genes and determine specific gene functions through the resulting phenotypes, which exhibit modified traits according with the function of the knock down gene(s) (Becker 2013; Becker and Lange 2010; Kirigia *et al.* 2014; Senthil-Kumar and Mysore 2014; Wang *et al.* 2015).

Although VIGS has the potential to examine gene function in a range of species, one challenge is that protocols must be tested and tailored for each new species (Becker 2013; Senthil-Kumar *et al.* 2007). Further, not all species are amenable to the VIGS technique. Here we report on testing and optimizing the VIGS methodology in *Cleome violacea* L. and *Polanisia dodecandra* (L.) DC (Figure 5.1), which are two strong candidates for using VIGS to ask questions about floral traits and pollinator preference

because they belong to different clades, have different petal shape, color and size, exhibit variation in nectar gland color and position, as well as stamen number (Figure 5.1). Though a full genome of *C. violacea* has been generated (J. C. Pires and P. Edger, unpublished), little is known about its pollinators and we could modify the floral traits to perform pollinator preference tests. In contrast, there are fewer genomic resources for *P. dodecandra* (but see chapter 4), but more is known about its pollination biology (Higuera-Díaz *et al.* 2015). Finally, both species exhibits traits that facilitate ease of study including small plant height, high percentage of successful seed germination, reproduction by selfing, and relatively short generation time.

To test feasibilities of VIGS in *C. violacea* and *P. dodecandra*, we developed constructs with endogenous genes of PHYTOENE DESATURASE (PDS) for both species. Usually, PDS is initially evaluated to determine the efficacy and optimize the VIGS technique in the species of interest, because silencing this gene results in an easy to score photo-bleached phenotype (Gould and Kramer 2007; Liu *et al.* 2002; Wege *et al.* 2007). The VIGS protocol for *C. violacea* was previously established by Mankowski (2012). Here, we further explore optimization of the VIGS technique for *C. violacea* by trying alternative types of infiltration methods that may increase penetrance. We also tested whether or not VIGS would be a viable approach to perform functional gene analysis in *P. dodecandra*, with a specific goal of examining the role of floral traits in pollinator attraction. For *P. dodecandra*, we additionally explored whether chemical mutagenesis is a viable technique to generate altered floral phenotypes.

5.2. Materials and Methods

5.2.1. Plant growth conditions

We grew *C. violacea* (813 from Hortus Botanicus) and *P. dodecandra* (68456 from B&T World Seeds) from Hall's lab seed stock. Voucher specimens were deposited in the Vascular Plant Herbarium at the University of Alberta (ALTA), *C. violacea* (Hall & Bolton 20 Feb 2008;) and *P. dodecandra* (Hall & Bolton 20 Feb 2008). Seeds of both species were sown in groups of 10 on a 2:1 mixture of sterilized SunGro sunshine mix and Terra-lite perlite, and grown under specific growth chamber conditions (24 ºC, 16

hour days). Seeds of *C. violacea* and *P. dodecandra* began to germinate approximately two weeks after sowing.

5.2.2. Cloning and construct design

Construct design and viral inoculation were carried out in accordance to previous VIGS protocols (Gould and Kramer 2007; Kramer *et al.* 2007a). We obtained Tobacco Rattle Virus vectors TRV1 and TRV2 sequences from TAIR (Dinesh-Kumar original donor). We made two endogenous constructs using *C. violacea* and *P. dodecandra* mRNA: TRV2-*Cv*PDS and TRV2-*Pd*PDS. RNA was extracted from leaves using a Concert Plant RNA reagent kit (Invitrogen, CA, USA), treated with DNAse1 (Fermentas, MD, USA), and enriched with an RNeasy Mini kit. (Qiagen, MD, USA). mRNA was isolated with a Dynabeads mRNA purification kit (Invitrogen, CA, USA), and cDNA was synthesized using SuperScript III reverse transcriptase (Invitrogen, CA, USA) and polyT primer (Kramer *et al.* 1998). Degenerate primers were designed using *Arabidopsis thaliana* L. PDS3 (GenBank: AT4G14210: F 5' TGG AAG GAR CAC TCM ATG ATW TTY GCH ATG 3' and R 5' ACR ACA TGR TAC TTS AVD ATT TTW GCY TT 3'). A 953bp and an 849 bp fragment were amplified from *C. violacea* and *P. dodecandra*, respectively. We cloned the fragments into a TOPO-TA plasmid vector (Invitrogen, Carlsbad, CA, USA) using manufacturer instructions, and sequenced using BigDye terminator sequencing (Applied Biosystems, Foster City, CA) with vector specific primers (Applied Biosystems, CA, USA) M13F (5' GTA AAA CGA CGG CCA G-3') and M13R (5'-CAG GAA ACA GCT ATG AC-3'). Using these sequences, new primers were designed adding XbaI and BamHI restriction sites to the 5' and 3' fragment ends for *C. violacea*, *Cv*-PDS-F-XbaI (5' -GGT CTA GAT AGT AGA TTT GAT TTC CCA GAT-3') and *Cv*-PDS-R-BamHI (5'-AAG GAT CCT AGA ATT TAG TCG TAC TTC CCC-3'). Similarly, we designed specific primers for *P. dodecandra* with BamHI and XhoI restriction enzymes, *Pd*-PDS-F-BamHI (5'-CGG GAT CCT AGT AGA TTT GAT TTC CCA GAT-3') and *Pd*-PDS-R-XhoI (5'-AAC TCG AGT AGA ATT TAG TCG TAC TTC-3'). For both species, the primers amplified a 429 bp fragment region of PDS gene. These smaller fragments of *Cv*PDS and *Pd*PDS were cloned into TOPO-TA plasmid. Using both XbaI and BamHI for *C. violacea*, and BamHI and XhoI for *P. dodecandra* restriction enzymes, the fragments were excised from the TOPO-TA vector

and ligated into separately digested TRV2 vectors using PGem t4 DNA ligase (Promega, WI, USA). We confirmed construct identity using BigDye Terminator sequencing with primers that spanned the TRV2 cloning site, 156F (5'-TTA CTC AAG GAA GCA CGA TGA GC-3') and156R (5'-GAA CCG TAG TTT AAT GTC TTC GGG-3'). Sequences were obtained with an ABI-3730 DNA Analyzer (Applied Biosystems, CA, USA) after being cleaned with a Performa DTR V# 96-well Short Plate Kit (Edge BioSystems, MD, USA). Nucleotide sequences of *Cv*PDS and *Pd*PDS fragments will be deposited in GenBank.

5.2.3. Preparation of *Agrobacterium tumefaciens*

VIGS technique usually involves a binary vector system, in this case TRV vectors, which consist of a vector containing a viral genome, pTRV1, and a vector containing a fragment of the target gene, pTRV2 (Kirigia *et al.* 2014). Thus, we transformed the following vectors into *Agrobacterium tumefaciens* Smith & Townsend: pTRV2-*CvPDS*, pTRV2-*PdPDS*, pTRV2, and pTRV1. For each of the four constructs, we mixed 10 µl of each construct with 250 µl of competent *A. tumefaciens*. Transformants were plated on LB media containing the following antibiotics: 50mg/ml kanamycin, 50mg/ml gentamycin, and 25mg/rifampicin. Transformants were then screened using 156F and 156R primers, glycerol stocks containing the transformed vectors were made for future use and were preserved at -80 ^{o}C.

5.2.4. Infiltration

Initially vacuum infiltration was the chosen method in *C. violacea*, and *P. dodecandra* because of its simplicity and efficiency in contrast to other methods (Becker and Lange 2010; Wang *et al.* 2006). *Cleome violacea* seedlings with 4-6 leaves and >7 true leaves (medium and large size seedling) were chosen for vacuum infiltration, after initial experiments performed by Mankowski (2012) demonstrated this stage has highest penetrance (Supplementary material 1). We performed a pilot experiment to decide the best seedling size for vacuum infiltration in *P. dodecandra*, after the experiment medium and large size seedlings were chosen for infiltration (Supplementary material 2).

To prepare for infiltration *A. tumefaciens* containing pTRV2-CvPDS, pTRV2-*PdPDS*, pTRV2 (empty vector), and pTRV1 were plated onto LB agar containing the standard antibiotic mixture and incubated for 72 hours at 26 ^{o}C. Serial inoculation, 72

hours at 150rpm, was used to obtain a final volume of 500ml for each colony. After serial inoculation, each culture was then transferred in equal volumes to 50ml falcon tubes and centrifuged at 3200g for 20 minutes at 4 ^{o}C. The supernatant was decanted from each tube, and cells were resuspended in 20ml of infiltration buffer (10 mM MES(2-(4-Morpholino)-Ethane Sulfonic Acid), 10 mM MgCl2, 0.1 M acetosyringone (3'5'-Dimethozy-4'-hydrozyamcetophenone)). Colonies from the same culture were recombined into new sterile beakers, infiltration buffer was added to OD600 = 2.0 ±.1, and beakers were left at room temperature for 4 hours to acclimatize.

Beakers containing *A. tumefaciens* for each TRV2 construct were combined in a 1:1 ratio of 250ml with TRV1. Silwet L-77 surfactant was added to each mixture at 100 µl/ L. Seedlings were uprooted and washed in distilled water to remove excess dirt and dried on paper towel. We performed two endogenous PDS treatments (pTRV2-*CvPDS* + pTRV1, pTRV2-*PdPDS* + pTRV1), empty pTRV2 + pTRV1 vector, and non-treated plants as a control for both species for a total of four treatments. Groups of medium and large seedlings were submerged into pTRV2 + pTRV1 mixtures and placed under vacuum for 2 minutes. Mixtures were reused for subsequent infiltrations, and 100 seedlings were infiltrated for each construct mixture.

We obtained high penetrance in *C. violacea* in which more than the 50% of the leaves showed the symptoms, but for *P. dodecandra* we observed few leaves showing the symptoms and these were patchy using the vacuum infiltration method. We therefore explored other infiltration methods (Agrodrench and syringe) to potentially increase VIGS success rate in *P. dodecandra* as well as produce even higher penetrance in *C. violacea*. The Agrodrench method consisted of drenching the root crown of each plant with 5 ml of *A. tumefaciens* solution (Kirigia *et al.* 2014; Ryu *et al.* 2004). For syringe infiltration, we used a 0.5 ml needle syringe and we injected the *Agrobacterium* solution along the plant stem (Wang *et al.* 2010b). However, because we did not get high penetrance with either of these methods, we tried a booster inoculation proposed by Senthil-Kumar and Mysore (2014). We did the booster inoculation to *P. dodecandra* plants exhibiting the altered symptoms 30 days after vacuum infiltration with the following protocol amendments: we chose the shorter plants due to vacuum size constraints and removed all dried leaves before the inoculation; Step A (ii) instead of

syringe infiltration we did a vacuum infiltration one plant at a time (see Senthil-Kumar and Mysore 2014 for detailed protocol). We also did a booster inoculation for *C. violacea* plants to test whether it would lead to even higher penetrance. For this species we used Agrodrench, syringe and vacuum infiltration in the booster inoculation. For both species, we did not perform the second booster inoculation because the plants looked sick and weak after the first booster inoculation and having healthy plants is a requirement to perform a second booster inoculation (Senthil-Kumar and Mysore 2014).

Phenotypes for *Cv*PDS and *Pd*PDS treated plants became apparent an average of 20 days after infiltration. Plants treated with pTRV2-*Cv*PDS + pTRV1 and pTRV2-*Pd*PDS had their leaves scored 30 days after the different infiltration methods. Leaves were scored by categories using visual inspection: Pale bleaching (<50%), variegated bleaching (>50% <80%), and strong bleaching (>80%). Leaf tissue from inoculated plants showing phenotypes, as well as control (non-treated plants) and empty vector plants, were excised, flash frozen in liquid nitrogen, and stored at -80 ^{o}C.

Unfortunately, we did not obtain sufficient high penetrance for *P. dodecandra* as just 27 infiltrated plants showed the symptoms and only one plant showed high penetrance. As such, we also investigated chemical mutagenesis using EMS to possibly generate plants with mutant flower phenotypes (Appendix 5.1). We followed protocol designed by Leyser (2000) with some amendments (see Appendix 5.1 for further details). In summary, we performed the chemical mutagenesis on 10000 seeds using different EMS concentrations and different seed immersion times in the EMS solution (Appendix 5.1). Due to space constraints, we grew between 950 to 1500 seeds of each treatment (Appendix 5.1). However, after mutagenesis we observed only one altered phenotype in the first generation, which consisted of only one plant with flowers without petals, stamens shorter than usual, style abnormally long, and the ovules appear to be external (Appendix 5.1). We were unable to produce a second mutant generation, because the seeds grown from the F1 mutant generation did not germinate and the seeds were unviable. Due to time and space constraints this experiment was not pursued further (Appendix 5.1).

Finally, we also developed the pTRV2-construct for *Cleomella serrulata* (Pursh) E.H. Roalson and J.C. Hall, although this species is challenging to grow under

greenhouse conditions. We tried a 1st VIGS round using vacuum infiltration but all the plants died shortly after the treatment (data not shown), suggesting this species is not amenable for the VIGS technique.

5.3. Results and Discussion

To avoid confounding the silencing symptoms with the viral infection, we did a control with an empty TRV2. Of the two species tested here for VIGS technique, *C. violacea* showed higher phenotypic response to *PDS* gene down regulation than *P. dodecandra*. This pattern was consistent across n=430 *C. violacea* (Table 5.1, Appendix 5.2) and n=911 *P. dodecandra* plants (Table 5.2 and 5.3, Appendix 5.3) plants tested across nine trials for each species. In both species, medium to large plants had the highest percentages exhibiting altered phenotypes with TRV2-*CvPDS* and TRV2-*PdPDS* constructs (Appendix 5.2 – 5.3) and the viral induction symptoms were displayed as early as 20 days after inoculation. Four weeks after inoculation, *C. violacea* and *P. dodecandra* plants showed different altered phenotypes. In both species, we found that most plants did not show infection symptoms, but we observed pale, variegated and strongly altered phenotypes (Figure 2 and 3). In plants that exhibited symptoms, the strong (28.4%) and variegated (7.03 %) phenotypes were the most frequent in *C. violacea* (Table 5.1), while the variegated (9.2 %) and pale (7.36 %) phenotypes were the most frequent in *P. dodecandra*. In fact, only one plant of *P. dodecandra* exhibited a strong altered phenotype (Table 5.2).

Of the three infiltration methods that we explored in *P. dodecandra*, syringe infiltration was the best method to obtain altered phenotypes, representing 52% of successfully modified plants (Table 5.3). However, for *P. dodecandra* the mortality rate (20 %) was higher under syringe infiltration when compared with other infiltration methods and the plants looked unhealthy four weeks after the infiltration procedure (Table 5.3). Although the Agrodrench method was used successfully to knock-down genes in tomato and *Striga hermonthica* Delile (Kirigia *et al.* 2014; Ryu *et al.* 2004), this approach was not very successful in *P. dodecandra* and the plants exhibited low rates of altered phenotypes (1.2 %). For these reasons, we chose vacuum infiltration over the Agrodrench and syringe infiltration methods for additional experiments. We also found a high mortality (50 %) in *P. dodecandra* plants after the booster inoculation. In addition,

the surviving plants did not exhibit the altered phenotypes (Appendix 5.4) and they recovered from the initial VIGS symptoms.

VIGS is a well-established method for down-regulation of genes in many angiosperm families and has proven to be valuable in studying plant lineages such as Ranunculales, asterids, Caryophyllales, and rosids (Deng *et al.* 2012; Gould and Kramer 2007; Hileman *et al.* 2005; Pabón-Mora *et al.* 2012; Ratcliff *et al.* 2001). Further, VIGS has helped to answer questions in regard to organ identity, sex determination, flower induction, and compound leaf development (Galimba and Di Stilio 2015; Harkess and Leebens-Mack 2017; Hsieh *et al.* 2013; LaRue *et al.* 2013; Wang *et al.* 2015). VIGS approaches are also potentially valuable to addressing multiple phenomena in some Cleomaceae species. *Cleome violacea* in particular is garnering interest due to its phylogenetic position and morphological traits. The VIGS protocol was already established for *C. violacea* by Mankowski (2012) and this species is highly amenable to VIGS technique. Although, we pursued different infiltration methods to improve penetrance through a booster inoculation, these alternative infiltration methods did not provide substantially more plants with altered phenotypes than the original vacuum infiltration method, which had the highest percentage (10.3 %) of altered phenotypes (Table 5.4). Based on our results, we recommend following the protocol developed by Mankowski (2012). In addition, we recommend using vacuum infiltration over other infiltration methods as Agrodrench or syringe in *C. violacea*, because it is a faster method to use and produced best results.

Despite our efforts to try multiple infiltration methods (Table 5.3), and a vacuum booster inoculation (Appendix 5.4) in *P. dodecandra,* this species does not appear to be amenable for VIGS, at least using the TRV vectors. We suggest that future studies explore different transformation vectors such as apple latent spherical virus (ALSV) in *P. dodecandra*. This vector had a high transformation efficacy in a wide range of plants including *A. thaliana*, tobacco, cucurbit and legume species, and for some species this vector performed better than TRV vectors (Igarashi *et al.* 2009). Another technique that is worth pursuing in *P. dodecandra* is CRISPR/cas9, which is a stable genome editing technique in which a target DNA sequence is transformed by adding, removing or replacing nucleotides (Bortesi and Fischer 2015; Kumar and Jain 2015). Finally, we

advise caution when examining the symptoms caused by the viral infection because they could be confused by altered phenotypes caused by the gene knockdown, and the experiments always should include a control treatment with an empty vector (Liu *et al.* 2002; Unver and Budak 2009). In *C. violacea* and *P. dodecandra* some plants showed abnormal growth pattern and contorted leaves, stems and pistils. Further, some plants showed a yellowish abnormal coloration in leaves and stems, but these are mild stress symptoms and they are likely caused by the viral infection.

Finally, the efficiency of VIGS technique is frequently validated via quantitative-PCR (qPCR) to compare relative transcript levels between the down-regulated target gene and untreated controls (Becker and Lange 2010; Burch-Smith *et al.* 2004; Kirigia *et al.* 2014). *Polanisia dodecandra* was excluded from the qPCR analysis because the penetrance was weak for the plants exhibiting the altered phenotypes, in which plants showed mostly pale phenotypes, and the number of plants with symptoms was low, suggesting VIGS was much less successful in this species. The qPCR validation was performed previously for *C. violacea* by (Mankowski 2012) and the target gene PDS was effectively down-regulated, with the relative expression values of the target gene being lower than untreated and empty vector plants (Appendix 5.5). These results indicated that the VIGS technique was successful in *C. violacea.* Thus, we encourage further VIGS studies with different gene targets encoding for floral traits in this species to answer pollination related questions.

5.4. Conclusion

Virus-inducing gene silencing has been shown to be a successful technique in only one of three Cleomaceae species that have been experimentally tested. This pattern does not necessarily mean that we should avoid this family for further VIGS experiments, rather that the protocol must be tailored for each species of interest. In particular, we recommend further attempts for VIGS optimization in *P. dodecandra* because the pollination biology of this species is known and the floral transcriptome is available, making it a good system to explore pollination related questions. However, different vectors such as ALSV should be tried to test if this species is amenable or not to VIGS. In addition, the VIGS protocol is working and available for *C. violacea*, making it a good

focal plant for a range of questions. Specifically, this system may be ideal for exploring questions about pollination by modifying traits important for pollinator attraction such as floral color and symmetry, nectar gland development, UV guides, and fruit development.

On a more technical note, our study found that using vacuum infiltration instead of syringes and Agrodrench infiltration was the most successful VIGS approach in Cleomaceae species. More generally, we recommend the use of VIGS technique in species where stable transformation protocols are lacking because VIGS is a fast and useful tool to study gene function.

5.5. References

Barker MS, Vogel H, Schranz ME. 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.

Baulcombe DC. 1999. Fast forward genetics based on virus-induced gene silencing. *Current Opinion in Plant Biology* **2**: 109-113.

Becker A. 2013. *Virus-Induced gene silencing.* New York: Humana Press.

Becker A, Lange M. 2010. VIGS - genomics goes functional. *Trends in Plant Science* **15**: 1-4.

Bhide A, Schliesky S, Reich M, Weber A, 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.

Bortesi L, Fischer R. 2015. The CRISPR/Cas9 system for plant genome editing and beyond. *Biotechnology Advances* **33**: 41-52.

Brautigam A, Kajala K, Wullenweber J, Sommer M, Gagneul D, Weber KL, Carr KM, Gowik U, Mass J, Lercher MJ, Westhoff P, Hibberd JM, Weber AP. 2011a. An mRNA blueprint for C4 photosynthesis derived from comparative transcriptomics of closely related C3 and C4 species. *Plant Physiology* **155**: 142-156.

Braütigam A, Mullick T, Schliesky S, Weber AP. 2011b. 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. *Journal of Experimental Botany* **62**: 3093-3102.

Burch-Smith TM, Anderson JC, Martin GB, Dinesh-Kumar SP. 2004. Applications and advantages of virus-induced gene silencing for gene function studies in plants. *Plant Journal* **39**: 734-746.

Burch-Smith TM, Schiff M, Liu Y, Dinesh-Kumar SP. 2006. Efficient virus-induced gene silencing in *Arabidopsis*. *Plant Physiology* **142**: 21-27.

Cane JH. 2008a. Breeding biologies, seed production and species-rich bee guilds of *Cleome lutea* and *Cleome serrulata* (Cleomaceae). *Plant Species Biology* **23**: 152-158.

Cheng S, van den Bergh E, Zeng P, Zhong X, Xu J, Liu X, Hofberger J, de Bruijn S, Bhide AS, Kuelahoglu C, Bian C, Chen J, Fan G, Kaufmann K, Hall JC, Becker A, Brautigam A, Weber AP, Shi C, Zheng Z, Li W, Lv M, Tao Y, Wang J, Zou H, Quan Z, Hibberd JM, Zhang G, Zhu XG, Xu X, Schranz ME. 2013. The *Tarenaya hassleriana* genome provides insight into reproductive trait and genome evolution of crucifers. *Plant Cell* **25**: 2813-2830.

Deng X, Elomaa P, Nguyen CX, Hytönen T, Valkonen JPT, Teeri TH. 2012. Virus-induced gene silencing for Asteraceae—a reverse genetics approach for functional genomics in *Gerbera hybrida*. *Plant Biotechnology Journal* **10**: 970-978.

Di Stilio VS, Kumar RA, Oddone AM, Tolkin TR, Salles P, McCarty K. 2010. Virus-induced gene silencing as a tool for comparative functional studies in *Thalictrum*. *PLoS ONE* **5**: e12064.

Dinesh-Kumar SP, Anandalakshmi R, Marathe R, Schiff M, Liu Y. 2003. Virus-Induced Gene Silencing. In: Grotewold E, ed. *Plant Functional Genomics*. New York: Humana Press, 452.

Edger PP, Heidel-Fischer HM, Bekaert M, Rota J, Glöckner G, Platts AE, Heckel DG, Der JP, Wafula EK, Tang M, Hofberger JA, Smithson A, Hall JC, Blanchette M, Bureau TE, Wright SI, dePamphilis CW, Eric Schranz M, Barker MS, Conant GC, Wahlberg N, Vogel H, Pires JC, Wheat CW. 2015. The butterfly plant arms-race escalated by gene and genome duplications. *Proceedings of the National Academy of Sciences* **112**: 8362-8366.

Endress PK. 1992. Evolution and floral diversity: The phylogenetic surroundings of *Arabidopsis* and *Antirrhinum*. *International Journal of Plant Sciences* **153**: S106-S122.

Fleming TH, Geiselman C, Kress WJ. 2009. The evolution of bat pollination: A phylogenetic perspective. *Annals of Botany* **104**: 1017-1043.

Galimba KD, Di Stilio VS. 2015. Sub-functionalization to ovule development following duplication of a floral organ identity gene. *Developmental Biology* **405**: 158-172.

Gould B, Kramer EM. 2007. Virus-induced gene silencing as a tool for functional analyses in the emerging model plant *Aquilegia* (columbine, Ranunculaceae). *Plant Methods* **3**: 6.

Harkess A, Leebens-Mack J. 2017. A century of sex determination in flowering plants. *Journal of Heredity* **108**: 69-77.

Higuera-Díaz M, Manson JS, Hall JC. 2015. Pollination biology of *Cleomella serrulata* and *Polanisia dodecandra* in a protected natural prairie in southern Alberta, Canada. *Botany* **93**: 745-757.

Hileman LC, Drea S, de Martino G, Litt A, Irish VF. 2005. Virus-induced gene silencing is an effective tool for assaying gene function in the basal eudicot species *Papaver somniferum* (Opium Poppy). *The Plant Journal* **44**: 334-341.

Hsieh M-H, Pan Z-J, Lai P-H, Lu H-C, Yeh H-H, Hsu C-C, Wu W-L, Chung M-C, Wang S-S, Chen W-H, Chen H-H. 2013. Virus-induced gene silencing unravels multiple transcription factors involved in floral growth and development in *Phalaenopsis* orchids. *Journal of Experimental Botany* **64**: 3869-3884.

Igarashi A, Yamagata K, Sugai T, Takahashi Y, Sugawara E, Tamura A, Yaegashi H, Yamagishi N, Takahashi T, Isogai M, Takahashi H, Yoshikawa N. 2009. Apple latent spherical virus vectors for reliable and effective virus-induced gene silencing among a broad range of plants including tobacco, tomato, *Arabidopsis thaliana*, cucurbits, and legumes. *Virology* **386**: 407-416.

Iltis HH, Hall JC, Cochrane TS, Sytsma KJ. 2011a. Studies in the Cleomaceae I: On the separate recognition of Capparaceae, Cleomaceae, and Brassicaceae. *Annals of the Missouri Botanical Garden* **98**: 28-36.

Kirigia D, Runo S, Alakonya A. 2014. A virus-induced gene silencing (VIGS) system for functional genomics in the parasitic plant *Striga hermonthica*. *Plant Methods* **10**: 16.

Kramer EM, Dorit RL, Irish VF. 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 EM, Holappa L, Gould B, Jaramillo MA, Setnikov D, Santiago PM. 2007a. Elaboration of B gene function to include the identity of novel floral organs in the lower eudicot *Aquilegia*. *The Plant Cell* **19**: 750-766.

Külahoglu C, Denton AK, Sommer M, Mass J, Schliesky S, Wrobel TJ, Berckmans B, Gongora-Castillo E, Buell CR, Simon R, De Veylder L, Brautigam A, Weber AP. 2014. Comparative transcriptome atlases reveal altered gene expression modules between two Cleomaceae C3 and C4 plant species. *Plant Cell* **26**: 3243-3260.

Kumar V, Jain M. 2015. The CRISPR–Cas system for plant genome editing: Advances and opportunities. *Journal of Experimental Botany* **66**: 47-57.

LaRue NC, Sullivan AM, Di Stilio VS. 2013. Functional recapitulation of transitions in sexual systems by homeosis during the evolution of dioecy in *Thalictrum*. *Frontiers in Plant Science* **4**: 487.

Leyser O. 2000. Mutagenesis. *Methods in Molecular Biology* **141**: 133-144.

Liu Y, Schiff M, Dinesh‐Kumar SP. 2002. Virus‐induced gene silencing in tomato. *The Plant Journal* **31**: 777-786.

Machado IC, Lopes AV, Leite AV, de Brito Neves C. 2006. *Cleome spinosa* (Capparaceae): Polygamodioecy and pollination by bats in urban and Caatinga areas, northeastern Brazil. *Botanische Jahrbücher* **127**: 69-82.

Mankowski P. 2012. *A functional approach to profiling candidate genes in non model Brassicales,* Master of Science, University of Alberta, Edmonton, AB.

Marshall DM, Muhaidat R, Brown NJ, Liu Z, Stanley S, Griffiths H, Sage RF, Hibberd JM. 2007. *Cleome*, a genus closely related to *Arabidopsis*, contains species spanning a developmental progression from C3 to C4 photosynthesis. *The Plant Journal* **51**: 886-896.

Martins DJ, Johnson SD. 2013. Interactions between hawkmoths and flowering plants in East Africa: Polyphagy and evolutionary specialization in an ecological context. *Biological Journal of the Linnean Society* **110**: 199-213.

Newell CA, Brown NJ, Liu Z, Pflug A, Gowik U, Westhoff P, Hibberd JM. 2010. *Agrobacterium tumefaciens*-mediated transformation of *Cleome gynandra* L., a C(4) dicotyledon that is closely related to *Arabidopsis thaliana*. *Journal of Experimental Botany* **61**: 1311-1319.

Nozzolillo C, Amiguet VT, Bily AC, Harris CS, Saleem A, Andersen OM, 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.

Owen CR, Bradshaw HD. 2011. Induced mutations affecting pollinator choice in *Mimulus lewisii* (Phrymaceae). *Arthropod-Plant Interactions* **5**: 235-244.

Pabón-Mora N, Ambrose BA, Litt A. 2012. Poppy APETALA1/FRUITFULL orthologs control flowering time, branching, perianth Identity, and fruit development. *Plant Physiology* **158**: 1685-1704.

Patchell MJ, Bolton MC, Mankowski P, Hall JC. 2011. Comparative floral development in Cleomaceae reveals two distinct pathways leading to monosymmetry. *International Journal of Plant Sciences* **172**: 352-365.

Patchell MJ, Roalson EH, Hall JC. 2014. Resolved phylogeny of Cleomaceae based on all three genomes. *Taxon* **63**: 315-328.

Raju AJS, Rani DS. 2016. Reproductive ecology of *Cleome gynandra* and *Cleome viscosa* (Capparaceae). *Phytologia Balcanica* **22**: 15-28.

Ratcliff F, Martin-Hernandez AM, Baulcombe DC. 2001. Tobacco rattle virus as a vector for analysis of gene function by silencing. *Plant Journal* **25**: 237-245.

Ruiz MT, Voinnet O, Baulcombe DC. 1998. Initiation and maintenance of virus-induced gene silencing. *Plant Cell* **10**: 937-946.

Ryu C-M, Anand A, Kang L, Mysore KS. 2004. Agrodrench: A novel and effective agroinoculation method for virus-induced gene silencing in roots and diverse solanaceous species. *The Plant Journal* **40**: 322-331.

Schranz ME, Mitchell-Olds T. 2006b. Independent ancient polyploidy events in the sister families Brassicaceae and Cleomaceae. *Plant Cell* **18**: 1152-1165.

Senthil-Kumar M, Hema R, Anand A, Kang L, Udayakumar M, Mysore KS. 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, Mysore KS. 2014. Tobacco rattle virus–based virus-induced gene silencing in *Nicotiana benthamiana*. *Nature Protocols* **9**: 1549-1562.

Tsai YT, Chen PY, To KY. 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 Reports* **31**: 1189-1198.

Unver T, Budak H. 2009. Virus-induced gene silencing, a post transcriptional gene silencing method. *International Journal of Plant Genomics* **2009**.

van den Bergh E, Hofberger JA, Schranz ME. 2016b. Flower power and the mustard bomb: Comparative analysis of gene and genome duplications in glucosinolate biosynthetic pathway evolution in Cleomaceae and Brassicaceae. *American Journal of Botany* **103**: 1212-1222.

Voznesenskaya EV, Koteyeva NK, Chuong SDX, Ivanova AN, Barroca J, Craven LA, Edwards GE. 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, Zheng Z. 2006. Optimisation of tobacco rattle virus-induced gene silencing in *Arabidopsis*. *Functional Plant Biology* **33**: 347-355.

Wang T, Li R, Wen L, Fu D, Zhu B, Luo Y, Zhu H. 2015. Functional analysis and RNA sequencing indicate the regulatory role of Argonaute1 in tomato compound leaf development. *PLoS ONE* **10**: e0140756.

Wang X, Zhu H, Shao Y, Chen A, Ma Y, Luo Y, Zhu B. 2010b. Analysis of gene functions by a syringe infiltration method of VIGS. *Russian Journal of Plant Physiology* **57**: 590-597.

Wege S, Scholz A, Gleissberg S, 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. *Annals of Botany* **100**: 641-649.

Weigel D, Glazebrook J. 2002. *Arabidopsis: A laboratory manual.* New York: Cold Spring Harbor Laboratoty Press.

Williams BP, Burgess SJ, Reyna-Llorens I, Knerova J, Aubry S, Stanley S, Hibberd JM. 2016. An untranslated cis element regulates the accumulation of multiple C4 enzymes in *Gynandropsis gynandra* mesophyll cells. *The Plant Cell* **28**: 454-465.

Table 5.1. Altered phenotypes exhibited by *Cleome violacea* in one trial of vacuum infiltrated plants.

Treatment	N	% Mortality (N)	% Phenotype
Untreated	57	14.03 (8)	N/A
Empty-TRV2	102	0	N/A
TRV2-*CvPDS:*	100	14 (14)	N/A
None	44		44
Pale	6		6
Variegated	7		7
Strong	29		29

Table 5.2. Altered phenotypes exhibited by *Polanisia dodecandra* in one trial of vacuum infiltrated plants.

Treatment	N	% Mortality	% Phenotype
Untreated	172	0	N/A
Empty-TRV2	102	0	N/A
TRV2-*PdPDS:*	163	0	N/A
None	135		82.82
Pale	12		7.36
Variegated	15		9.2
Strong	1		0.61

Table 5.3. Comparison of mortality and observed altered phenotype of *Polanisia dodecandra* across different infiltration methods in four VIGS trials.

Treatment	Infiltration method	N	% Mortality (N)	% Phenotype
Untreated	Agrodrench	74	0	N/A
	Syringe	50	30 (15)	N/A
	Vacuum	172	0	N/A
Empty-TRV2	Agrodrench	79	0	N/A
	Syringe	50	26 (13)	N/A
	Vacuum	102	0	N/A
TRV2-*PdPDS*	Agrodrench	83	1.2 (1)	13.25
	Syringe	50	20 (10)	52
	Vacuum	163	0	16.56

Table 5.4. Comparison of mortality and observed altered phenotype of *Cleome violacea* booster inoculation across different infiltration methods in three VIGS trials.

Treatment	Infiltration method	N treated	% Mortality (N)	% Phenotype
Untreated	Agrodrench	150	4.66 (7)	N/A
	Syringe	150	4.66 (7)	N/A
	Vacuum	150	4.66 (7)	N/A
Empty-TRV2	Agrodrench	150	6 (9)	N/A
	Syringe	150	5.33 (8)	N/A
	Vacuum	150	5.33 (8)	N/A
TRV2-*CvPDS*	Agrodrench	165	5.45 (9)	5.45
	Syringe	165	7.87 (13)	8.48
	Vacuum	165	9.69 (16)	10.30

Figure 5.1. *Cleome violacea* and *Polanisia dodecandra* inflorescences.

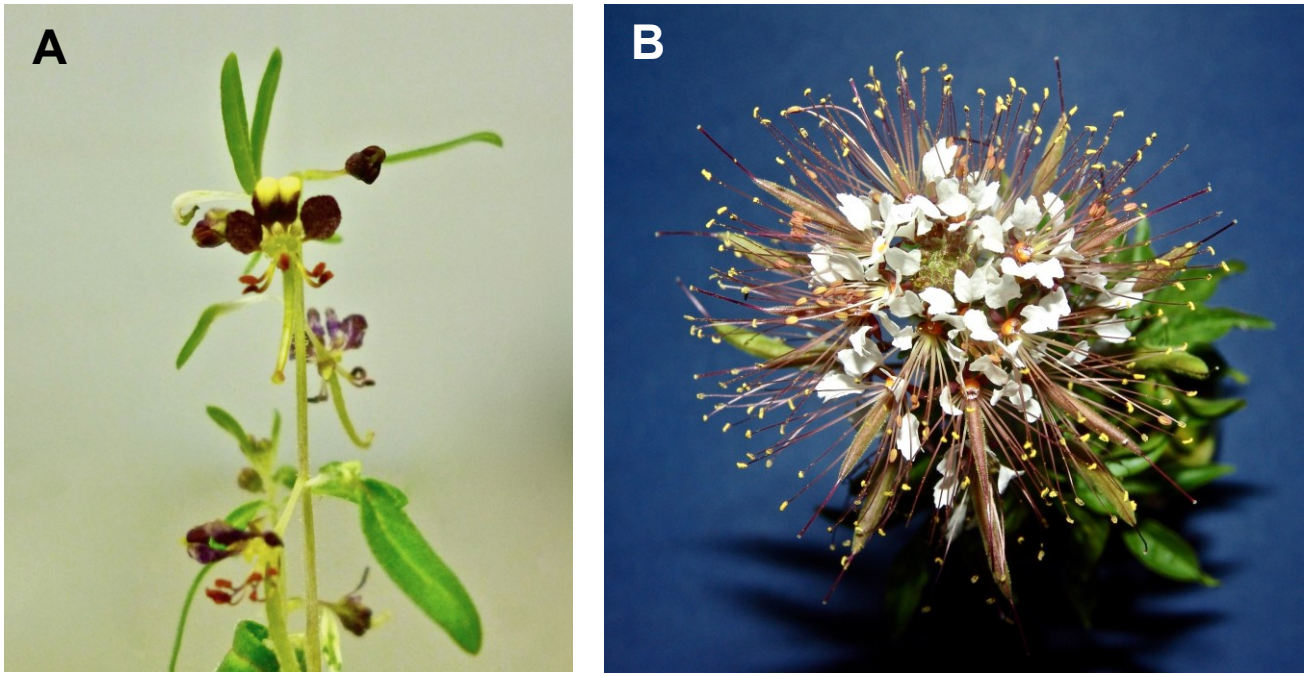


Figure 5.2. Phenotypes of *PDS* gene knocked-down in *Cleome violacea*. A) Untreated; B) Pale; C) Variegated; D) Strong.

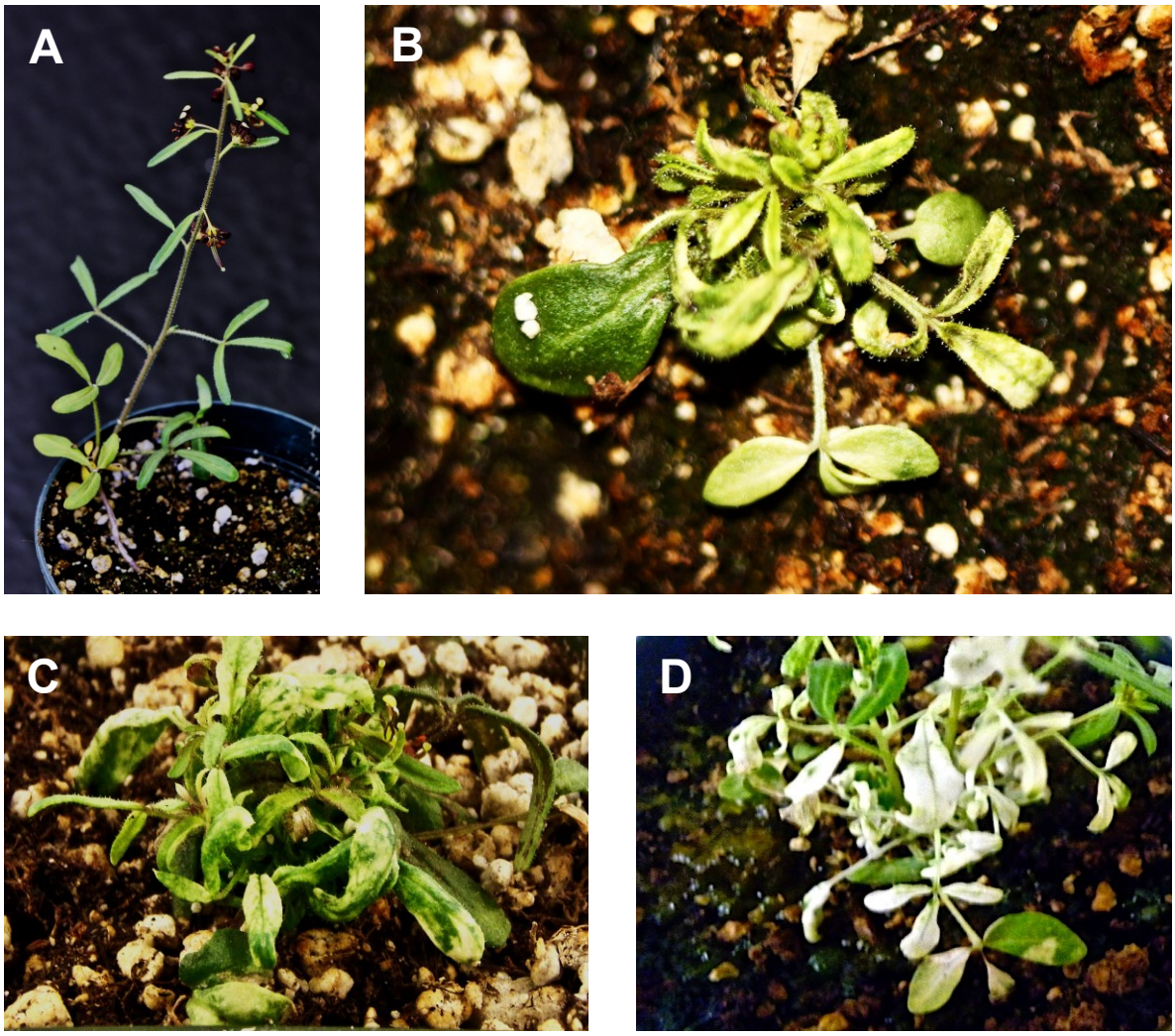
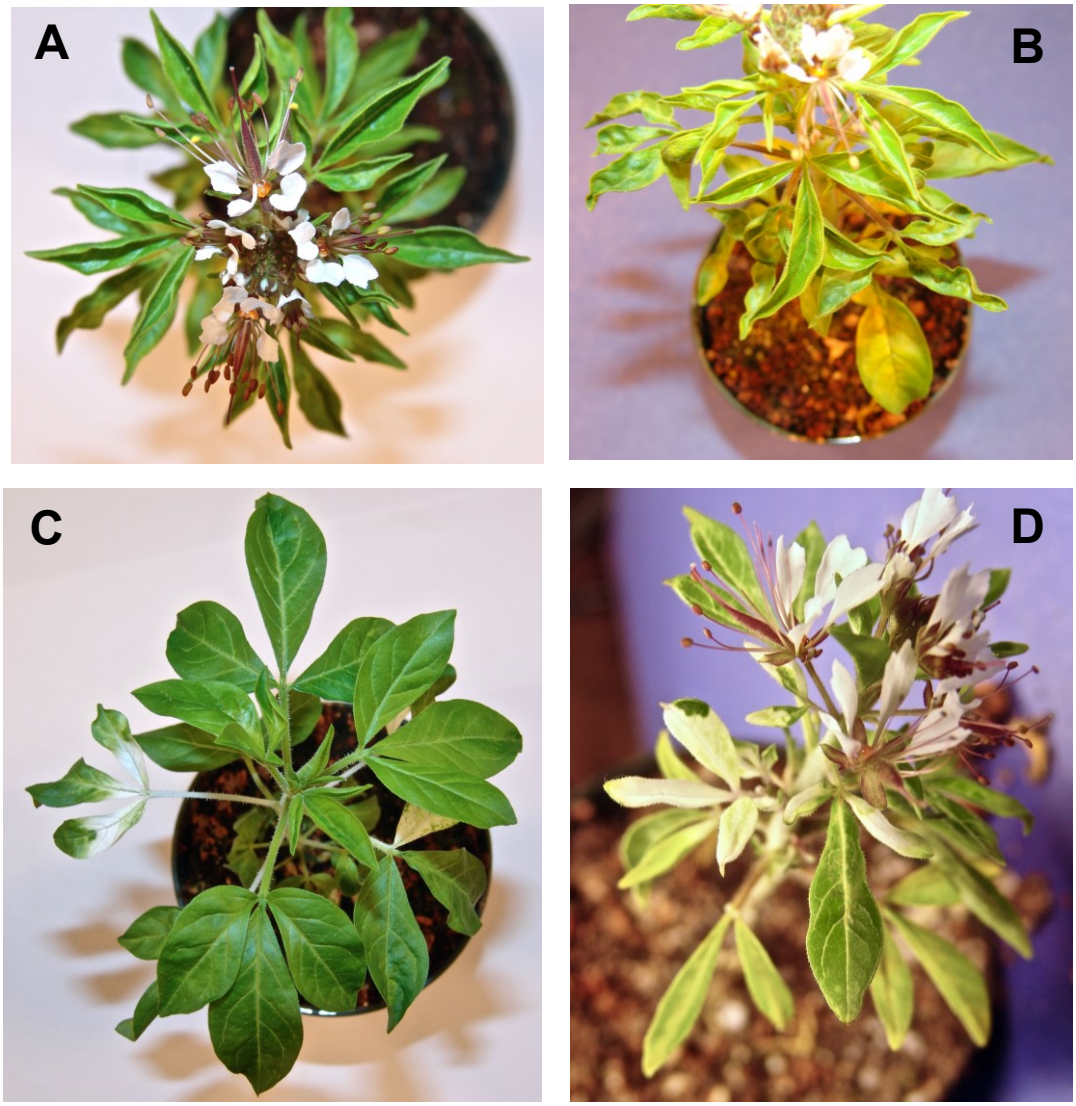


Figure 5.3. Phenotypes of *PDS* gene knocked-down in *Polanisia dodecandra*. A) Untreated; B) Pale; C) Variegated; D) Strong.



Chapter 6: Conclusion and future directions

Plant-pollinator interactions have fascinated scientists for centuries (Herrera and Pellmyr 2002b; Patiny 2012; Proctor *et al.* 1996; Willmer 2011). Recent advances in next generation sequencing (NGS) provides a new opportunity for pollination researchers to explore the evolution of floral phenotypes and pollinator-mediated selection by directly investigating the genetic basis of floral traits (Clare *et al.* 2013; Gübitz *et al.* 2009; Sedeek *et al.* 2013; Yuan *et al.* 2013*a*). The objectives of this thesis were to investigate the pollination biology of Cleomaceae species and identify putative genes encoding major floral traits significant for pollinator attraction in species exhibiting different pollination systems. These kinds of studies are complementary to ecological pollination studies and will help us to better understand the evolution of floral traits in plants exhibiting generalist and specialist pollination systems, as we could identify genes coding floral traits that are subjected to pollinator-mediated selection. I first reviewed the opportunities that NGS presents for pollination studies and propose Cleomaceae as a suitable taxon to explore eco-evo-devo studies in integrative pollination studies (Chapter 2). I used my own research as an example of this; first I obtained empirical pollination data in generalist species, then, I compared the transcriptome of species exhibiting different pollination systems. Further, I performed functional studies of candidate genes that encoded floral traits that may be subject to pollinator-mediated selection (Chapters 3 – 5). My research elucidated the genetic differences in terms of gene repertoires between two Cleomaceae species with generalized floral traits and one species exhibiting specialized floral traits.

I concluded that *Cleomella serrulata* (Pursh) E.H. Roalson and J.C. Hall and *Polanisia dodecandra* (L.) DC., exhibit a generalist pollination system, but they are pollinated by different pollinator guilds, despite that species live in close proximity. These species also differ in the floral cues and rewards offered to pollinators (Chapter 3). When comparing the floral transcriptomes of three focal species, I found that the floral transcriptome of the two-generalist species (*C. serrulata* and *P. dodecandra*) were more similar to each other than to the specialist *Melidiscus giganteus* (L.) Raf. (Chapter 4). However, I found that the floral transcriptomes had similar gene repertoires across all three species despite exhibiting different pollination systems. After the transcriptome analysis, I selected putative candidate genes that may encode main floral traits and may be responsible

for the morphological differences found among species (Chapter 4). Two different techniques (virus-inducing gene silencing and chemical mutagenesis) were performed in an attempt to obtain plants with modified floral phenotypes. These experiments were performed in *P. dodecandra*, which has a short life cycle and rapidly grows in a growth chamber environment. This species also exhibits interesting floral traits as white petals and conspicuous nectar gland with UV glowing nectar, which could be modified to perform pollinator preference tests. Unfortunately, this species was not amenable for the techniques that I attempted (Chapter 5 and Appendix). My results suggested that although Cleomaceae family has great potential for further integrative pollination studies in both generalist and specialist species, the species that we researched are not good models for moving forward in integrative pollination studies, because they were not suitable for genetic modifications, they did not growth well in greenhouse environments.

6.1. Cleomaceae as a model for integrative pollination studies

Pollination biology studies in Cleomaceae species are scarce and the pollination biology of most species is unknown, but the studied species exhibit a generalist pollination system or a specialist one (Chapters 1, 3 and 4). However, it is possible that the species exhibit a gradient between generalist and specialist pollination systems, but we need more pollination studies in the family to support this hypothesis, because the pollination biology has been explored in only 10 species (Cane 2008a; Fleming *et al.* 2009; Higuera-Díaz *et al.* 2015; Machado *et al.* 2006; Martins and Johnson 2013; Raju and Rani 2016). Due to the fact that the family has both generalist and specialist species, it is a suitable group to study if there are gradations in pollination systems. Additional pollination studies are needed to elucidated patterns and shifts between pollination systems and to determine which floral features may be of relative importance to specific pollinator guilds. In this sense, we can explore features such as floral color and display, nectaries, UV reflectance, and scents, and determine if they are correlated with specific pollinator guilds. We need to explore further pollination systems in Cleomaceae to find closely related species exhibiting different pollination systems. This would allow us to perform comparisons of pollinators, and floral traits subjected to pollinator-mediated selection under different pollination regimes.

To perform comparative studies of genes encoding for floral traits in specialist and generalist pollination systems within Cleomaceae, we should start by focusing on species that have available genomic or transcriptomic resources such as *C. violacea*, *C. serrulata*, *M. giganteus*, *P. dodecandra*, *Tarenaya hassleriana* (Chodat) Iltis, and *T. spinosa* (Bhide *et al.* 2014; Brautigam *et al.* 2011a; Braütigam *et al.* 2011b; Cheng *et al.* 2013; Feodorova *et al.* 2010; Külahoglu *et al.* 2014; Mohammadin *et al.* 2015; van den Bergh *et al.* 2016a; van den Bergh *et al.* 2014; Williams *et al.* 2016). These species exhibit differences in key floral traits and likely have different pollinator communities, making them suitable for studies comparing pollination systems. Consequently, I recommend performing integrative pollination studies first in these species to study the transitions between the generalist *C. serrulata*, *P. dodecandra* and the specialists *T. spinosa*. Additional studies on species in different clades could help us to understand the evolution and transitions of pollination systems in the family and determine which pollination system occurred first in the family. Most genetic pollination studies focus on floral traits such as color and nectaries, therefore studying the volatile organic compounds of flowers would be informative, due to their importance for the attraction of different pollinator guilds and may therefore influence the evolution of specialist pollination systems in the family. Cleomaceae is also a good model to explore the pollination role of additional floral traits that were not studied in this thesis (e.g., floral size and display, nectar guides, scents, and novel morphological structures such as androgynophores).

6.2. The value of 'omics' for studying pollination

The drastic increase in new sequencing technologies allows us to have a huge amount of genomics and transcriptomics data in a relatively short time. 'Omics' provide a great opportunity to study floral evolution as we can perform broad scale comparisons across species to find whether there are specific genes associated with each type of pollination system or pollination syndrome. Despite the value of 'omics' data derived from NGS only a few studies have embraced this approach to answer specific pollinator related questions (Sedeek *et al.* 2013; Zhang *et al.* 2015b). However, RNA-seq is a powerful tool that could be used in pollination research. We can use RNA-seq to perform comparisons not only in terms of genes presence/absence, but also to obtain differential expression

values in candidate genes of interest. In Cleomaceae this technique has already been used to compare the leaf and floral transcriptome of *T. hassleriana* (Bhide *et al*. 2014) and this study can be used as an starting point to identify genes expressed in floral tissue only, which may be relevant for the pollination process. In addition, my own research can also be useful for identifying genes important for pollinator attraction in generalist and specialist pollination systems. Furthermore, because the generation and availability of 'omics' data in Cleomaceae species, these could be used to formulate hypotheses about the evolution of pollination syndromes such as the bat pollination floral related phenotype (van den Bergh *et al.* 2016a).

In Chapter 4, I identified putative candidate genes that may be involved in morphological differences among *M. giganteus*, *C. serrulata,* and *P. dodecandra* and also genes potentially underlying major floral traits for pollinator attraction. These genes were associated with flower color, symmetry, nectar gland development, and nectar production. For example*, Melidiscus giganteus* had the highest number of genes related to nectar production and genes involved in the formation and development of nectar glands that could be related with the bat pollinator syndrome, but the specimens that I studied lack both, presumably because the plants were grown for months in greenhouse. Similarly, when I analyzed genes related to flower color, I identified the main genes associated to the anthocyanin pathway in species with green (*M. giganteus*) and white (*P. dodecandra*) petals, but I did not observe absence anthocyanin related pathway genes in white flowered species. It is possible that the differences in the floral traits of these species are defined by up or down regulation of the major genes or by other genetic factors and future studies should focus on characterizing gene expression levels.

Further investigations should be performed to determine the role of anthocyanin related genes in pollinator attraction. In addition, we could explore which genes are involved in the formation of other distinctive floral traits such as petal spots, exhibited by *C. violacea*, and their role in pollinator attraction. Once the gene or genes that underlie the floral trait of interest are identified, such as floral symmetry or flower color, we could test the function of this gene and its role in pollinator attraction by modifying plants to express alternative phenotypes. These plants with altered phenotypes could be exposed to pollinators to test the adaptive value of the trait. To obtain plants with modified phenotypes

in Cleomaceae, we could use techniques such as CRISPR rather than mutagenesis and virus-inducing gene silencing (with the exception of *C. violacea* in which VIGS techniques is optimized), because these techniques are time consuming and they were not effective at least for the species of Cleomaceae that I studied. CRISPR has not been used to modify floral traits yet, but the technique has been proposed to change floral traits in horticultural plants and apparently it has a great potential to alter floral phenotypes (Noman *et al.* 2017). Furthermore, Cleomaceae is a good model for evo-devo and pollination research, because the species could have variations in genome duplication events, and flowers and fruit types are diverse. Genome duplication events are important for evolutionary studies, because gene copy retention can be the origin of gene neofunctionalization or subfunctionalization and they can affect floral traits important for pollinator attraction. Moreover, some species exhibit morphological novelties such as androgynophores (Hall 2008) and there are opportunities to study the evolution of this trait and its role in pollinator attraction. Finally, there are species with short and fast life cycles allowing to grow the plants in greenhouses or growth chambers that can be used to harvested tissue for 'omics' studies in short time or to perform pollinator preferences test under controlled environments.

Finally, more than transcriptomics studies, we encourage to perform quantitative gene expression studies as RNA-seq, because as we showed in or results, most of the genes coding for floral traits important for pollinator attraction showed low variation in terms of presence/absence, and may be the differences between specialist and generalist pollination are based on differential gene expression rather than gene profiles. We also recommend increasing the number of studied species to find broader patterns about the evolution of pollination systems in Cleomaceae.

6.3. Future directions

Because of its diversity of pollination systems, the Cleomaceae is a promising clade to study the evolution of pollination systems outside of the common crop model plants. In addition, the small number of species in this family, polyploidy events, and its close relationship with Brassicaceae and hence with *Arabidopsis thaliana*, make Cleomaceae as an ideal group to perform ecology, genomic and transcriptomic studies of pollination systems evolution across a family clade. However, it is necessary to perform studies in

more Cleomaceae species that exhibit generalist or specialist pollination systems and determine the relative proportions of generalists and specialists in the family. With this knowledge, we could elucidate the evolutionary patterns between pollination types (e.g., are generalists or specialists the ancestral state? Who was the first pollination vector? Are there transitions of pollination systems in species that live in tropical and temperate habitats?). Furthermore, we could identify which floral features may be under directional selection (e.g., how do genes and alleles that modulate floral traits and their expression change across this family? Are the floral color and nectar glands subject to pollinator-mediated selection in specialized pollination systems?).

In general, floral color, floral size and display, and corolla tube length are the most studied traits in terms of pollinator-mediated selection (Fleming *et al.* 2009; Proctor *et al.* 1996; Willmer 2011). However, volatile organic compounds and their derivatives are important for the attraction of different pollinator guilds as moths, bats, bees, and others, and frequently these volatiles are the first signal barrier for pollination selection (Fleming *et al.* 2009; Proctor *et al.* 1996; Willmer 2011). Many species of Cleomaceae produce strong smells that can be associated with herbivore deterrence and pollinator attraction. While, other compounds such as glucosinolates have been associated with bat pollination syndrome in Cleomaceae species (van den Bergh *et al.* 2016a). Furthermore, I found differences in genes associated with volatile organic compounds in an initial examination of the studied species. These compounds could be responsible for the attraction of different pollinator guilds in Cleomaceae and we could determine which volatile compounds are responsible for attracting moths and bats pollinators. Moreover, we could explore whether these compounds are widely produced in the family or if they are restricted to a few species, however, specific studies on volatiles compounds are necessary to test these assumptions.

I also suggest performing more studies about nocturnal pollination in this family using a phylogenetic approach to determine if it is an ancestral trait in Cleomaceae. If nocturnal pollination is actually an ancestral character, it may mean the specialist pollination systems arose first in the family. I encourage the study of pollination biology in *C. violacea* because this species is amenable to performing genetic modifications through VIGS, it has an available genome and it grows easily and fast in green-house environments.

Cleome violacea also exhibits an interesting color change under UV radiation that should be explored further. Finally, it is necessary to perform genomics studies to identify if the genome triplication event happened before or after the divergence of *Polanisia* clade (van den Bergh 2017). With sequenced genomes, we could establish the copy number of genes of interest, because gene copy retention could affect floral traits important for pollinator attraction. The information derived from these studies will allow us to improve our understanding of floral evolution in Cleomaceae, while elucidating evolutionary patterns, shifts in pollination systems and define floral traits that can be under selection exerted by pollinators.

Cleomaceae could be a good model to study the evolution of specialized and generalized floral traits across an entire cosmopolitan family clade, although not all species are amenable to genetic transformations. The discoveries made using this model clade, which is relatively less complex due the small number of species, can be applied and tested to more complex taxa, generating a useful bridge to increase the general understanding of pollination systems and plant evolutionary processes.

6.4. References

Bhide A, Schliesky S, Reich M, Weber A, 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.

Brautigam A, Kajala K, Wullenweber J, Sommer M, Gagneul D, Weber KL, Carr KM, Gowik U, Mass J, Lercher MJ, Westhoff P, Hibberd JM, Weber AP. 2011a. An mRNA blueprint for C4 photosynthesis derived from comparative transcriptomics of closely related C3 and C4 species. *Plant Physiology* **155**: 142-156.

Braütigam A, Mullick T, Schliesky S, Weber AP. 2011b. 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. *Journal of Experimental Botany* **62**: 3093-3102.

Cane JH. 2008a. Breeding biologies, seed production and species-rich bee guilds of *Cleome lutea* and *Cleome serrulata* (Cleomaceae). *Plant Species Biology* **23**: 152-158.

Cheng S, van den Bergh E, Zeng P, Zhong X, Xu J, Liu X, Hofberger J, de Bruijn S, Bhide AS, Kuelahoglu C, Bian C, Chen J, Fan G, Kaufmann K, Hall JC, Becker A, Brautigam A, Weber AP, Shi C, Zheng Z, Li W, Lv M, Tao Y, Wang J, Zou H, Quan Z, Hibberd JM, Zhang G, Zhu XG, Xu X, Schranz ME. 2013. The *Tarenaya hassleriana* genome provides insight into reproductive trait and genome evolution of crucifers. *Plant Cell* **25**: 2813-2830.

Clare EL, Schiestl FP, Leitch AR, Chittka L. 2013. The promise of genomics in the study of plant-pollinator interactions. *Genome Biology* **14**: 207.

Feodorova TA, Voznesenskaya EV, Edwards GE, Roalson EH. 2010. Biogeographic patterns of diversification and the origins of C4 in *Cleome* (Cleomaceae). *Systematic Botany* **35**: 811-826.

Fleming TH, Geiselman C, Kress WJ. 2009. The evolution of bat pollination: A phylogenetic perspective. *Annals of Botany* **104**: 1017-1043.

Gübitz T, Hoballah ME, Dell'Olivo A, Kuhlemeier C. 2009. *Petunia* as a model system for the genetics and evolution of pollination syndromes. In: Gerats T, Strommer J, eds. *Petunia*. New York: Springer 29-49.

Hall JC. 2008. Systematics of Capparaceae and Cleomaceae: An evaluation of the generic delimitations of *Capparis* and *Cleome* using plastid DNA sequence data. *Botany-Botanique* **86**: 682-696.

Herrera CM, Pellmyr O. 2002b. *Plant-animal interactions : An evolutionary approach.* Oxford: Blackwell Science.

Higuera-Díaz M, Manson JS, Hall JC. 2015. Pollination biology of *Cleomella serrulata* and *Polanisia dodecandra* in a protected natural prairie in southern Alberta, Canada. *Botany* **93**: 745-757.

Külahoglu C, Denton AK, Sommer M, Mass J, Schliesky S, Wrobel TJ, Berckmans B, Gongora-Castillo E, Buell CR, Simon R, De Veylder L, Brautigam A, Weber AP. 2014. Comparative transcriptome atlases reveal altered gene expression modules between two Cleomaceae C3 and C4 plant species. *Plant Cell* **26**: 3243-3260.

Machado IC, Lopes AV, Leite AV, de Brito Neves C. 2006. *Cleome spinosa* (Capparaceae): Polygamodioecy and pollination by bats in urban and Caatinga areas, northeastern Brazil. *Botanische Jahrbücher* **127**: 69-82.

Martins DJ, Johnson SD. 2013. Interactions between hawkmoths and flowering plants in East Africa: Polyphagy and evolutionary specialization in an ecological context. *Biological Journal of the Linnean Society* **110**: 199-213.

Mohammadin S, Edger PP, Pires JC, Schranz ME. 2015. Positionally-conserved but sequence-diverged: Identification of long non-coding RNAs in the Brassicaceae and Cleomaceae. *BMC Plant Biology* **15**: 217.

Noman A, Aqeel M, Deng J, Khalid N, Sanaullah T, Shuilin H. 2017. Biotechnological advancements for improving floral attributes in ornamental plants. *Frontiers in Plant Science* **8**: 530.

Patiny S. 2012. *Evolution of plant-pollinator relationships.* Cambridge: Cambridge University Press.

Proctor MCF, Yeo P, Lack A. 1996. *The natural history of pollination.* Portland: Timber Press.

Raju AJS, Rani DS. 2016. Reproductive ecology of *Cleome gynandra* and *Cleome viscosa* (Capparaceae). *Phytologia Balcanica* **22**: 15-28.

Sedeek KE, Qi W, Schauer MA, Gupta AK, Poveda L, Xu S, Liu ZJ, Grossniklaus U, Schiestl FP, Schluter PM. 2013. Transcriptome and proteome data reveal candidate genes for pollinator attraction in sexually deceptive orchids. *PLoS ONE* **8**: e64621.

van den Bergh E. 2017. *Comparative genomics and trait evolution in Cleomaceae, a model for ancient polyploidy,* PhD, Wageningen University

van den Bergh E, Hofberger JA, Schranz ME. 2016a. Flower power and the mustard bomb: Comparative analysis of gene and genome duplications in glucosinolate biosynthetic pathway evolution in Cleomaceae and Brassicaceae. *American Journal of Botany* **103**: 1212-1222.

van den Bergh E, Külahoglu C, Bräutigam A, Hibberd JM, Weber AP, Zhu X-G, Schranz 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.

Williams BP, Burgess SJ, Reyna-Llorens I, Knerova J, Aubry S, Stanley S, Hibberd JM. 2016. An untranslated cis element regulates the accumulation of multiple C4 enzymes in *Gynandropsis gynandra* mesophyll cells. *The Plant Cell* **28**: 454-465.

Willmer P. 2011. *Pollination and floral ecology.* Princeton: Princeton University Press.

Yuan YW, Byers KJRP, Bradshaw HD. 2013*a*. The genetic control of flower-pollinator specificity. *Current Opinion in Plant Biology* **16**: 422-428.

Zhang W, Wei X, Meng H-L, Ma C-H, Jiang N-H, Zhang G-H, Yang S-C. 2015b. Transcriptomic comparison of the self-pollinated and cross-pollinated flowers of *Erigeron breviscapus* to analyze candidate self-incompatibility-associated genes. *BMC Plant Biology* **15**: 248.

Bibliography

Alexandre H, Vrignaud J, Mangin B, Joly S. 2015. Genetic architecture of pollination syndrome transition between hummingbird-specialist and generalist species in the genus *Rhytidophyllum* (Gesneriaceae). *PeerJ* **3**: e1028.

Almeida J, Rocheta M, Galego L. 1997. Genetic control of flower shape in *Antirrhinum majus*. *Development* **124**: 1387-1392.

Altieri MA. 1999. The ecological role of biodiversity in agroecosystems. *Agriculture, Ecosystems & Environment* **74**: 19-31.

Anderson JT, Mitchell-Olds T. 2011. Ecological genetics and genomics of plant defenses: Evidence and approaches. *Functional Ecology* **25**: 312-324.

Anderson JT, Willis JH, Mitchell-Olds T. 2011. Evolutionary genetics of plant adaptation. *Trends in Genetics* **27**: 258-266.

Armbruster WS. 2012. Evolution and ecological implications of "specialized" pollinator rewards. In: Patiny S, ed. *Evolution of plant-pollinator relationships*. Cambridge; New York: Cambridge University Press, 44-67.

Baker HG, Baker I, Hodges SA. 1998. Sugar composition of nectars and fruits consumed by birds and bats in the tropics and subtropics. *Biotropica* **30**: 559-586.

Barker MS, Vogel H, Schranz ME. 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.

Baucom RS, Chang SM, Kniskern JM, Rausher MD, Stinchcombe JR. 2011. Morning Glory as a powerful model in ecological genomics: Tracing adaptation through both natural and artificial selection. *Heredity (Edinb)* **107**: 377-385.

Baulcombe DC. 1999. Fast forward genetics based on virus-induced gene silencing. *Current Opinion in Plant Biology* **2**: 109-113.

Bayo-Canha A, Delgado-Benarroch L, Weiss J, Egea-Cortines M. 2007. Artificial decrease of leaf area affects inflorescence quality but not floral size in *Antirrhinum majus*. *Scientia Horticulturae* **113**: 383-386.

Becker A. 2013. *Virus-Induced gene silencing.* New York: Humana Press.

Becker A, Lange M. 2010. VIGS - genomics goes functional. *Trends in Plant Science* **15**: 1-4.

Bender R, Klinkenberg P, Jiang Z, Bauer B, Karypis G, Nguyen N, Perera MADN, Nikolau BJ, Carter CJ. 2012. Functional genomics of nectar production in the Brassicaceae. *Flora - Morphology, Distribution, Functional Ecology of Plants* **207**: 491-496.

Bey M, Stuber K, Fellenberg K, Schwarz-Sommer Z, Sommer H, Saedler H, Zachgo S. 2004. Characterization of *Antirrhinum* petal development and identification of target genes of the class B MADS box gene DEFICIENS. *Plant Cell* **16**: 3197-3215.

Bhide A, Schliesky S, Reich M, Weber A, 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.

Bolstad GH, Armbruster WS, Pélabon C, Pérez-Barrales R, Hansen TF. 2010. Direct selection at the blossom level on floral reward by pollinators in a natural population of *Dalechampia schottii*: Full-disclosure honesty? *New Phytologist* **188**: 370-384.

Bortesi L, Fischer R. 2015. The CRISPR/Cas9 system for plant genome editing and beyond. *Biotechnology Advances* **33**: 41-52.

Bradshaw HD, Schemske DW. 2003. Allele substitution at a flower colour locus produces a pollinator shift in Monkeyflowers. *Nature* **426**: 176-178.

Brautigam A, Gowik U. 2010. What can next generation sequencing do for you? Next generation sequencing as a valuable tool in plant research. *Plant Biology* **12**: 831-841.

Brautigam A, Kajala K, Wullenweber J, Sommer M, Gagneul D, Weber KL, Carr KM, Gowik U, Mass J, Lercher MJ, Westhoff P, Hibberd JM, Weber AP. 2011a. An mRNA blueprint for C4 photosynthesis derived from comparative transcriptomics of closely related C3 and C4 species. *Plant Physiology* **155**: 142-156.

Braütigam A, Mullick T, Schliesky S, Weber AP. 2011b. 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. *Journal of Experimental Botany* **62**: 3093-3102.

Burch-Smith TM, Anderson JC, Martin GB, Dinesh-Kumar SP. 2004. Applications and advantages of virus-induced gene silencing for gene function studies in plants. *Plant Journal* **39**: 734-746.

Burch-Smith TM, Schiff M, Liu Y, Dinesh-Kumar SP. 2006. Efficient virus-induced gene silencing in *Arabidopsis*. *Plant Physiology* **142**: 21-27.

Busch A, Zachgo S. 2007. Control of corolla monosymmetry in the Brassicaceae *Iberis amara*. *Proceedings of the National Academy of Sciences of the United States of America* **104**: 16714-16719.

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

Cameron SA, Lozier JD, Strange JP, Koch JB, Cordes N, Solter LF, Griswold TL. 2011. Patterns of widespread decline in North American Bumble bees. *Proceedings of the National Academy of Sciences* **108**: 662-667.

Campbell DR, Waser NM, Price MV. 1996. Mechanisms of hummingbird-mediated selection for flower width in *Ipomopsis aggregata*. *Ecology* **77**: 1463-1472.

Cane JH. 2008a. Breeding biologies, seed production and species-rich bee guilds of *Cleome lutea* and *Cleome serrulata* (Cleomaceae). *Plant Species Biology* **23**: 152-158.

Cane JH. 2008b. Pollinating bees crutial to farming wildflower seed for U.S. habitat restoration. In: James RR, Pitts-Singer TL, eds. *Bee pollination in agricultural ecosystems*. New York: Oxford University Press, 48 - 64.

Celik I, Gurbuz N, Uncu AT, Frary A, Doganlar S. 2017. Genome-wide SNP discovery and QTL mapping for fruit quality traits in inbred backcross lines (IBLs) of *Solanum pimpinellifolium* using genotyping by sequencing. *BMC Genomics* **18**: 1.

Chanderbali AS, Berger BA, Howarth DG, Soltis PS, Soltis DE. 2016. Evolving ideas on the origin and evolution of flowers: New perspectives in the genomic era. *Genetics* **202**: 1255-1265.

Cheng S, van den Bergh E, Zeng P, Zhong X, Xu J, Liu X, Hofberger J, de Bruijn S, Bhide AS, Kuelahoglu C, Bian C, Chen J, Fan G, Kaufmann K, Hall JC, Becker A, Brautigam A, Weber AP, Shi C, Zheng Z, Li W, Lv M, Tao Y, Wang J, Zou H, Quan Z, Hibberd JM, Zhang G, Zhu XG, Xu X, Schranz ME. 2013. The *Tarenaya hassleriana* genome provides insight into reproductive trait and genome evolution of crucifers. *Plant Cell* **25**: 2813-2830.

Chittka L, Kevan PG. 2005. Flower colour as advertisement. In: Dafni A, Kevan PG, Husband BC, eds. *Practical pollination biology*. Ontario, Canada: Enviroquest Ltd., 157 - 196.

Chittka L, Raine NE. 2006. Recognition of flowers by pollinators. *Current Opinion in Plant Biology* **9**: 428-435.

Christinet L, Burdet FX, Zaiko M, Hinz U, Zrÿd J-P. 2004. Characterization and functional identification of a novel plant 4,5-extradiol dioxygenase involved in betalain pigment biosynthesis in *Portulaca grandiflora*. *Plant Physiology* **134**: 265-274.

Clare EL, Schiestl FP, Leitch AR, Chittka L. 2013. The promise of genomics in the study of plant-pollinator interactions. *Genome Biology* **14**: 207.

Clegg MT, Durbin ML. 2000. Flower color variation: A model for the experimental study of evolution. *Proceedings of the National Academy of Sciences* **97**: 7016-7023.

Collier R, Fuchs B, Walter N, Kevin Lutke W, Taylor CG. 2005. *Ex vitro* composite plants: An inexpensive, rapid method for root biology. *The Plant Journal* **43**: 449-457.

Conesa A, Götz S, García-Gómez JM, Terol J, Talón M, Robles M. 2005. Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics* **21**: 3674-3676.

Conner JK, Rush S, Jennetten P. 1996. Measurements of natural selection on floral traits in wild radish (*Raphanus raphanistrum*). I. Selection through lifetime female fitness. *Evolution* **50**: 1127-1136.

Cronk Q, Ojeda I. 2008. Bird-pollinated flowers in an evolutionary and molecular context. *Journal of Experimental Botany* **59**: 715-727.

Cronk QCB, Bateman RM, Hawkins JA. 2002. *Developmental genetics and plant evolution.* New York: Taylor & Francis.

Cruden RW, Lyon DL. 1985. Patterns of biomass allocation to male and female functions in plants with different mating systems. *Oecologia* **66**: 299-306.

Dafni A. 1992. *Pollination ecology: A practical approach.* Oxford: Oxford University Press.

Dafni A, Eisikowitch D, Ivri Y. 1987. Nectar flow and pollinators' efficiency in two co-occurring species of *Capparis* (Capparaceae) in Israel. *Plant Systematics and Evolution* **157**: 181-186.

Davies KM. 2009. Modifying Anthocyanin Production in Flowers. In: Winefield C, Davies K, Gould K, eds. *Anthocyanins: Biosynthesis, Functions, and Applications*. New York, NY: Springer New York, 49-80.

Deng X, Elomaa P, Nguyen CX, Hytönen T, Valkonen JPT, Teeri TH. 2012. Virus-induced gene silencing for Asteraceae—a reverse genetics approach for functional genomics in *Gerbera hybrida*. *Plant Biotechnology Journal* **10**: 970-978.

Des Marais DL, Rausher MD. 2010. Parallel evolution at multiple levels in the origin of hummingbird pollinated flowers in *Ipomoea*. *Evolution* **64**: 2044-2054.

Devaux C, Lepers C, Porcher E. 2014. Constraints imposed by pollinator behaviour on the ecology and evolution of plant mating systems. *Journal of Evolutionary Biology* **27**: 1413-1430.

Di Stilio VS, Kumar RA, Oddone AM, Tolkin TR, Salles P, McCarty K. 2010. Virus-induced gene silencing as a tool for comparative functional studies in *Thalictrum*. *PLoS ONE* **5**: e12064.

Dinesh-Kumar SP, Anandalakshmi R, Marathe R, Schiff M, Liu Y. 2003. Virus-Induced Gene Silencing. In: Grotewold E, ed. *Plant Functional Genomics*. New York: Humana Press, 452.

Dornhaus A, Chittka L. 1999. Insect behaviour: Evolutionary origins of bee dances. *Nature* **401**: 38-38.

Dudareva N, Cseke L, Blanc VM, Pichersky E. 1996. Evolution of floral scent in *Clarkia*: Novel patterns of S-linalool synthase gene expression in the *C. breweri* flower. *The Plant Cell Online* **8**: 1137-1148.

Dudareva N, Raguso RA, Wang JH, Ross EJ, Pichersky E. 1998. Floral scent production in *Clarkia breweri* - III. Enzymatic synthesis and emission of benzenoid esters. *Plant Physiology* **116**: 599-604.

Duelli P, Obrist MK, Schmatz DR. 1999. Biodiversity evaluation in agricultural landscapes: Above-ground insects. *Agriculture, Ecosystems & Environment* **74**: 33-64.

Dyer AG, Whitney HM, Arnold SEJ, Glover BJ, Chittka L. 2007. Mutations perturbing petal cell shape and anthocyanin synthesis influence bumblebee perception of *Antirrhinum majus* flower colour. *Arthropod-Plant Interactions* **1**: 45-55.

Edger PP, Heidel-Fischer HM, Bekaert M, Rota J, Glöckner G, Platts AE, Heckel DG, Der JP, Wafula EK, Tang M, Hofberger JA, Smithson A, Hall JC, Blanchette M, Bureau TE, Wright SI, dePamphilis CW, Eric Schranz M, Barker MS, Conant GC, Wahlberg N, Vogel H, Pires JC, Wheat CW. 2015. The butterfly plant arms-race escalated by gene and genome duplications. *Proceedings of the National Academy of Sciences* **112**: 8362-8366.

Endress PK. 1992. Evolution and floral diversity: The phylogenetic surroundings of *Arabidopsis* and *Antirrhinum*. *International Journal of Plant Sciences* **153**: S106-S122.

Enright AJ, Van Dongen S, Ouzounis CA. 2002. An efficient algorithm for large-scale detection of protein families. *Nucleic Acids Research* **30**: 1575-1584.

Epperson BK, Clegg MT. 1992. Unstable white flower color genes and their derivatives in Morning Glory. *Journal of Heredity* **83**: 405-409.

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

Erhardt A, Rusterholz H-P. 1998. Do Peacock butterflies (*Inachis io* L.) detect and prefer nectar amino acids and other nitrogenous compounds? *Oecologia* **117**: 536-542.

Fankhauser N, Aubry S. 2017. Post-transcriptional regulation of photosynthetic genes is a key driver of C4 leaf ontogeny. *Journal of Experimental Botany* **68**: 137-146.

Fenster CB, Armbruster WS, Wilson P, Dudash MR, Thomson JD. 2004. Pollination syndromes and floral specialization. *Annual Review of Ecology Evolution and Systematics* **35**: 375-403.

Fenster CB, Martén-Rodriguez S, Schemske DW. 2009. Pollination syndromes and the evolution of floral diversity in *Iochroma* (Solanaceae). *Evolution* **63**: 2758-2762.

Feodorova TA, Voznesenskaya EV, Edwards GE, Roalson EH. 2010. Biogeographic patterns of diversification and the origins of C4 in *Cleome* (Cleomaceae). *Systematic Botany* **35**: 811-826.

Fernández VA, Galetto L, Astegiano J. 2009. Influence of flower functionality and pollination system on the pollen size-pistil length relationship. *Organisms Diversity & Evolution* **9**: 75-82.

Fernandez-Mazuecos M, Glover BJ. 2017. The evo-devo of plant speciation. *Nature Ecology & Evolution* **1**: 110.

Fleming TH, Geiselman C, Kress WJ. 2009. The evolution of bat pollination: A phylogenetic perspective. *Annals of Botany* **104**: 1017-1043.

Free JB. 1993. *Insect pollination of crops.* London: Academic Press.

Fuhlendorf SD, Zhang H, Tunnell TR, Engle DM, Cross AF. 2002. Effects of grazing on restoration of southern mixed Prairie soils. *Restoration Ecology* **10**: 401-407.

Galen C. 1999. Why do flowers vary? The functional ecology of variation in flower size and form within natural plant populations. *BioScience* **49**: 631-640.

Galimba KD, Di Stilio VS. 2015. Sub-functionalization to ovule development following duplication of a floral organ identity gene. *Developmental Biology* **405**: 158-172.

Gallego Romero I, Pai AA, Tung J, Gilad Y. 2014. RNA-seq: impact of RNA degradation on transcript quantification. *BMC Biology* **12**: 42.

Galloway LF, Cirigliano T, Gremski K. 2002. The contribution of display size and dichogamy to potential geitonogamy in *Campanula americana*. *International Journal of Plant Sciences* **163**: 133-139.

Garber M, Grabherr MG, Guttman M, Trapnell C. 2011. Computational methods for transcriptome annotation and quantification using RNA-seq. *Nature Methods* **8**: 469-477.

Garibaldi LA, Carvalheiro LG, Leonhardt SD, Aizen MA, Blaauw BR, Isaacs R, Kuhlmann M, Kleijn D, Klein AM, Kremen C, Morandin L, Scheper J, Winfree R. 2014. From research to action: Enhancing crop yield through wild pollinators. *Frontiers in Ecology and the Environment* **12**: 439-447.

Gaskett AC. 2011. Orchid pollination by sexual deception: Pollinator perspectives. *Biological Reviews* **86**: 33-75.

Giurfa M, Núñez J, Chittka L, Menzel R. 1995. Colour preferences of flower-naive Honeybees. *Journal of Comparative Physiology A* **177**: 247-259.

Goldblatt P, Manning JC. 2000. The long-proboscid fly pollination system in Southern Africa. *Annals of the Missouri Botanical Garden* **87**: 146-170.

Gould B, Kramer EM. 2007. Virus-induced gene silencing as a tool for functional analyses in the emerging model plant *Aquilegia* (columbine, Ranunculaceae). *Plant Methods* **3**: 6.

Grotewold E. 2006. The genetics and biochemistry of floral pigments. *Annual Review of Plant Biology* **57**: 761-780.

Gübitz T, Hoballah ME, Dell'Olivo A, Kuhlemeier C. 2009. *Petunia* as a model system for the genetics and evolution of pollination syndromes. In: Gerats T, Strommer J, eds. *Petunia*. New York: Springer 29-49.

Gumbert A. 2000. Color choices by Bumble bees (*Bombus terrestris*): Innate preferences and generalization after learning. *Behavioral Ecology and Sociobiology* **48**: 36-43.

Guo N, Cheng F, Wu J, Liu B, Zheng S, Liang J, Wang X. 2014. Anthocyanin biosynthetic genes in *Brassica rapa*. *BMC Genomics* **15**: 426.

Haas BJ, Papanicolaou A, Yassour M, Grabherr M, Blood PD, Bowden J, Couger MB, Eccles D, Li B, Lieber M, MacManes MD, Ott M, Orvis J, Pochet N, Strozzi F, Weeks N, Westerman R, William T, Dewey CN, Henschel R, Leduc RD, Friedman N, Regev A. 2013. *De novo* transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. *Nature Protocols* **8**: 1494-1512.

Hall JC. 2008. Systematics of Capparaceae and Cleomaceae: An evaluation of the generic delimitations of *Capparis* and *Cleome* using plastid DNA sequence data. *Botany-Botanique* **86**: 682-696.

Hansen DM, Van der Niet T, Johnson SD. 2012. Floral signposts: Testing the significance of visual 'nectar guides' for pollinator behaviour and plant fitness. *Proceedings Biological Sciences* **279**: 634-639.

Harder LD, Johnson SD. 2005. Adaptive plasticity of floral display size in animal-pollinated plants. *Proceedings of the Royal Society B: Biological Sciences* **272**: 2651-2657.

Harkess A, Leebens-Mack J. 2017. A century of sex determination in flowering plants. *Journal of Heredity* **108**: 69-77.

Hawkins RD, Hon GC, Ren B. 2010. Next-Generation genomics: An integrative approach. *Nature Reviews. Genetics* **11**: 476-486.

He C, Saedler H. 2005. Heterotopic expression of MPF2 is the key to the evolution of the Chinese lantern of *Physalis*, a morphological novelty in Solanaceae. *Proceedings of the National Academy of Sciences of the United States of America* **102**: 5779-5784.

Hegland SJ, Totland Ø. 2005. Relationships between species' floral traits and pollinator visitation in a temperate grassland. *Oecologia* **145**: 586-594.

Henson J, Tischler G, Ning Z. 2012. Next-generation sequencing and large genome assemblies. *Pharmacogenomics* **13**: 901-915.

Herrera CM. 1996. Floral traits and plant adaptation to insect pollinators: A devil's advocate approach. In: Lloyd DG, Barrett SCH, eds. *Floral biology*. Boston, MA: Springer, 65-87.

Herrera CM. 2005. Plant generalization on pollinators: Species property or local phenomenon? *American Journal of Botany* **92**: 13-20.

Herrera CM, Pellmyr O. 2002a. *Plant animal interactions: An evolutionary approach.* Oxford: Blackwell Science.

Herrera CM, Pellmyr O. 2002b. *Plant-animal interactions : An evolutionary approach.* Oxford: Blackwell Science.

Higuera-Díaz M, Manson JS, Hall JC. 2015. Pollination biology of *Cleomella serrulata* and *Polanisia dodecandra* in a protected natural prairie in southern Alberta, Canada. *Botany* **93**: 745-757.

Hileman L, Cubas P. 2009. An expanded evolutionary role for flower symmetry genes. *Journal of Biology* **8**: 90.

Hileman LC, Drea S, de Martino G, Litt A, Irish VF. 2005. Virus-induced gene silencing is an effective tool for assaying gene function in the basal eudicot species *Papaver somniferum* (Opium Poppy). *The Plant Journal* **44**: 334-341.

Ho WW, Smith SD. 2016. Molecular evolution of anthocyanin pigmentation genes following losses of flower color. *BMC Evolutionary Biology* **16**: 98.

Hoballah ME, Gubitz T, Stuurman J, Broger L, Barone M, Mandel T, Dell'Olivo A, Arnold M, Kuhlemeier C. 2007. Single gene-mediated shift in pollinator attraction in *Petunia*. *Plant Cell* **19**: 779-790.

Hobbs GA. 1968. Ecology of species of *Bombus* (Hymenoptera: Apidae) in southern Alberta: VII. Subgenus bombus. *The Canadian Entomologist* **100**: 156-164.

Hodges SA, Kramer EM. 2007. Columbines. *Current Biology* **17**: R992-R994.

Honaas LA, Wafula EK, Wickett NJ, Der JP, Zhang Y, Edger PP, Altman NS, Pires JC, Leebens-Mack JH, dePamphilis CW. 2016. Selecting superior *De Novo* transcriptome assemblies: Lessons learned by leveraging the best plant genome. *PLoS ONE* **11**: e0146062.

Horridge GA. 1996. The honeybee (*Apis mellifera*) detects bilateral symmetry and discriminates its axis. *Journal of Insect Physiology* **42**: 755-764.

Hsieh M-H, Pan Z-J, Lai P-H, Lu H-C, Yeh H-H, Hsu C-C, Wu W-L, Chung M-C, Wang S-S, Chen W-H, Chen H-H. 2013. Virus-induced gene silencing unravels multiple transcription factors involved in floral growth and development in *Phalaenopsis* orchids. *Journal of Experimental Botany* **64**: 3869-3884.

Ichihashi Y, Tsukaya H. 2015. Behavior of leaf meristems and their modification. *Frontiers in Plant Science* **6**: 1060.

Igarashi A, Yamagata K, Sugai T, Takahashi Y, Sugawara E, Tamura A, Yaegashi H, Yamagishi N, Takahashi T, Isogai M, Takahashi H, Yoshikawa N. 2009. Apple latent spherical virus vectors for reliable and effective virus-induced gene silencing among a broad range of plants including tobacco, tomato, *Arabidopsis thaliana*, cucurbits, and legumes. *Virology* **386**: 407-416.

Iltis HH. 1957. Studies in the Capparidaceae. III. Evolution and phylogeny of the western North American Cleomoideae. *Annals of the Missouri Botanical Garden* **44**: 77-119.

Iltis HH. 1958. Studies in the Capparidaceae—IV. *Polanisia* Raf. *Brittonia* **10**: 33-58.

Iltis HH, Cochrane TS. 2014. Cleomaceae. In: Davidse G, Sousa Sánchez S, Knapp S, Cabrera FC, eds. *Flora Mesoamericana*. St. Louis, Missouri: Missouri Botanical Garden Press, 256 -267.

Iltis HH, Hall JC, Cochrane TS, Sytsma KJ. 2011a. Studies in the Cleomaceae I: On the separate recognition of Capparaceae, Cleomaceae, and Brassicaceae. *Annals of the Missouri Botanical Garden* **98**: 28-36.

Iltis HH, Hall JC, Cochrane TS, Sytsma KJ. 2011b. Studies in the Cleomaceae I. On the separate recognition of Capparaceae, Cleomaceae, and Brassicaceae. *Annals of the Missouri Botanical Garden* **98**: 28-36.

Inda LA, Torrecilla P, Catalán P, Ruiz-Zapata T. 2008. Phylogeny of *Cleome* L. and its close relatives *Podandrogyne* Ducke and *Polanisia* Raf. (Cleomoideae, Cleomaceae) based on analysis of nuclear ITS sequences and morphology. *Plant Systematics and Evolution* **274**: 111-126.

Jersáková J, Johnson SD, Kindlmann P. 2006. Mechanisms and evolution of deceptive pollination in orchids. *Biological Reviews* **81**: 219-235.

Jin J, Tian F, Yang D-C, Meng Y-Q, Kong L, Luo J, Gao G. 2017. PlantTFDB 4.0: Toward a central hub for transcription factors and regulatory interactions in plants. *Nucleic Acids Research* **45**: D1040-D1045.

Johnson SD, Steiner KE. 2000. Generalization versus specialization in plant pollination systems. *Trends in Ecology & Evolution* **15**: 140-143.

Johnston MO. 1991. Natural selection on floral traits in two species of *Lobelia* with different pollinators. *Evolution* **45**: 1468-1479.

Jordheim M, Andersen ØM, Nozzolillo C, Amiguet VT. 2009. Acylated anthocyanins in inflorescence of spider flower (*Cleome hassleriana*). *Phytochemistry* **70**: 740-745.

José M. Gómez, Francisco Perfectti, Juan Pedro M. Camacho. 2006. Natural selection on *Erysimum mediohispanicum* flower shape: Insights into the evolution of zygomorphy. *The American Naturalist* **168**: 531-545.

Kanehisa M, Goto S. 2000. KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Research* **28**: 27-30.

Karron JD, Ivey CT, Mitchell RJ, Whitehead MR, Peakall R, Case AL. 2012. New perspectives on the evolution of plant mating systems. *Annals of Botany* **109**: 493-503.

Kay KM, Sargent RD. 2009. The role of animal pollination in plant speciation: Integrating ecology, geography, and genetics. *Annual Review of Ecology Evolution and Systematics* **40**: 637-656.

Kearns CA, Inouye DW. 1993. *Techniques for pollination biologists.* Colorado: University Press of Colorado.

Kessler D, Baldwin IT. 2011. Back to the past for pollination biology. *Current Opinion in Plant Biology* **14**: 429-434.

Kessler D, Bhattacharya S, Diezel C, Rothe E, Gase K, Schottner M, Baldwin IT. 2012. Unpredictability of nectar nicotine promotes outcrossing by hummingbirds in *Nicotiana attenuata*. *Plant Journal* **71**: 529-538.

Kessler D, Diezel C, Baldwin IT. 2010. Changing pollinators as a means of escaping herbivores. *Current Biology* **20**: 237-242.

Kessler D, Gase K, Baldwin IT. 2008. Field experiments with transformed plants reveal the sense of floral scents. *Science* **321**: 1200-1202.

Kevan PG. 1999. Pollinators as bioindicators of the state of the environment: Species, activity and diversity. *Agriculture, Ecosystems & Environment* **74**: 373-393.

Kirigia D, Runo S, Alakonya A. 2014. A virus-induced gene silencing (VIGS) system for functional genomics in the parasitic plant *Striga hermonthica*. *Plant Methods* **10**: 16.

Klahre U, Gurba A, Hermann K, Saxenhofer M, Bossolini E, Guerin PM, Kuhlemeier C. 2011. Pollinator choice in *Petunia* depends on two major genetic loci for floral scent production. *Current Biology* **21**: 730-739.

Klein A-M, Vaissière BE, Cane JH, Steffan-Dewenter I, Cunningham SA, Kremen C, Tscharntke T. 2007. Importance of pollinators in changing landscapes for world crops. *Proceedings of the Royal Society B: Biological Sciences* **274**: 303-313.

Klinkhamer PG, de Jong TJ. 1990. Effects of plant size, plant density and sex differential nectar reward on pollinator visitation in the protandrous *Echium vulgare* (Boraginaceae). *Oikos*: 399-405.

Kooke R, Wijnker E, Keurentjes JJ. 2012. Backcross populations and near isogenic Lines. In: Rifkin SA, ed. *Quantitative Trait Loci (QTL): Methods and protocols*. New York: Springer, 3-16.

Kram BW, Xu WW, Carter CJ. 2009. Uncovering the *Arabidopsis thaliana* nectary transcriptome: Investigation of differential gene expression in floral nectariferous tissues. *BMC Plant Biology* **9**: 92-92.

Kramer EM. 2007. Understanding the genetic basis of floral diversity. *BioScience* **57**: 479-487.

Kramer EM. 2009*a*. *Aquilegia*: A new model for plant development, ecology, and evolution. *Annual Review of Plant Biology* **60**: 261-277.

Kramer EM. 2009*b*. New model systems for the study of developmental evolution in plants. In: Jeffery WR, ed. *Current Topics in Developmental Biology*. Burlington: Academic Press, 67-105.

Kramer EM, Dorit RL, Irish VF. 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 EM, Holappa L, Gould B, Jaramillo MA, Setnikov D, Santiago PM. 2007a. Elaboration of B gene function to include the identity of novel floral organs in the lower eudicot *Aquilegia*. *Plant Cell* **19**: 750-766.

Kramer EM, Holappa L, Gould B, Jaramillo MA, Setnikov D, Santiago PM. 2007b. Elaboration of B gene function to include the identity of novel floral organs in the lower eudicot *Aquilegia*. *The Plant Cell* **19**: 750-766.

Külahoglu C, Denton AK, Sommer M, Mass J, Schliesky S, Wrobel TJ, Berckmans B, Gongora-Castillo E, Buell CR, Simon R, De Veylder L, Brautigam A, Weber AP. 2014. Comparative transcriptome atlases reveal altered gene expression modules between two Cleomaceae C3 and C4 plant species. *Plant Cell* **26**: 3243-3260.

Kumar V, Jain M. 2015. The CRISPR–Cas system for plant genome editing: Advances and opportunities. *Journal of Experimental Botany* **66**: 47-57.

Lange M, Yellina AL, Orashakova S, Becker A. 2013. Virus-induced gene silencing (VIGS) in plants: An overview of target species and the virus-derived vector systems. In: Becker A, ed. *Virus-induced gene silencing: Methods and protocols*. Totowa, NJ: Humana Press, 1-14.

LaRue NC, Sullivan AM, Di Stilio VS. 2013. Functional recapitulation of transitions in sexual systems by homeosis during the evolution of dioecy in *Thalictrum*. *Frontiers in Plant Science* **4**: 487.

Lehti-Shiu MD, Panchy N, Wang P, Uygun S, Shiu S-H. 2017. Diversity, expansion, and evolutionary novelty of plant DNA-binding transcription factor families. *Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms* **1860**: 3-20.

Lemmon E, Lemmon A. 2013. High-throughput genomic data in systematics and phylogenetics. *The Annual Review of Ecology, Evolution, and Systematics* **44**: 1-23.

Lewis AC. 1986. Memory constraints and flower choice in *Pieris rapae*. *Science* **232**: 863-865.

Leyser O. 2000. Mutagenesis. *Methods in Molecular Biology* **141**: 133-144.

Li L, Stoeckert CJ, Roos DS. 2003. OrthoMCL: Identification of ortholog groups for Eukaryotic genomes. *Genome Research* **13**: 2178-2189.

Lin IW, Sosso D, Chen LQ, Gase K, Kim SG, Kessler D, Klinkenberg PM, Gorder MK, Hou BH, Qu XQ, Carter CJ, Baldwin IT, Frommer WB. 2014. Nectar secretion requires sucrose phosphate synthases and the sugar transporter SWEET9. *Nature* **508**: 546-549.

Liu Y, Schiff M, Dinesh-Kumar SP. 2002. Virus-induced gene silencing in tomato. *The Plant Journal* **31**: 777-786.

Lloyd DG, Bawa KS. 1984. Modification of the gender of seed plants in varying conditions. *Evolutionary Biology* **17**: 255-338.

Lulin H, Xiao Y, Pei S, Wen T, Shangqin H. 2012. The first Illumina-based de novo transcriptome sequencing and analysis of Safflower flowers. *PLoS One* **7**: e38653.

Luo L, Zhu Y, Xiong M. 2012. Quantitative Trait Locus (QTL) analysis for Next-Generation Sequencing with the functional linear models. *Journal of medical genetics* **49**: 513-524.

Ma H. 2001. Regulatory Genes in Plant Development: MADS. *eLS*John Wiley & Sons, Ltd.

Machado IC, Lopes AV, Leite AV, de Brito Neves C. 2006. *Cleome spinosa* (Capparaceae): Polygamodioecy and pollination by bats in urban and Caatinga areas, northeastern Brazil. *Botanische Jahrbücher* **127**: 69-82.

Mankowski P. 2012. *A functional approach to profiling candidate genes in non model Brassicales,* Master of Science, University of Alberta, Edmonton, AB.

Maple J, Moller SG. 2007. Mutagenesis in *Arabidopsis*. *Methods in Molecular Biology* **362**: 197-206.

Marshall DM, Muhaidat R, Brown NJ, Liu Z, Stanley S, Griffiths H, Sage RF, Hibberd JM. 2007. *Cleome*, a genus closely related to *Arabidopsis*, contains species spanning a developmental progression from C3 to C4 photosynthesis. *The Plant Journal* **51**: 886-896.

Martin LB, Fei Z, Giovannoni JJ, Rose JK. 2013. Catalyzing plant science research with RNA-seq. *Frontiers in Plant Science* **4**: 1-10.

Martin NH, Sapir Y, Arnold ML. 2008. The genetic architecture of reproductive isolation in Louisiana irises: Pollination syndromes and pollinator preferences. *Evolution* **62**: 740-752.

Martins DJ, Johnson SD. 2013. Interactions between hawkmoths and flowering plants in East Africa: Polyphagy and evolutionary specialization in an ecological context. *Biological Journal of the Linnean Society* **110**: 199-213.

Mayer C, Adler LS, Armbruster WS, Dafni A, Eardley C, Huang SQ, Kevan PG, Ollerton J, Packer L, Ssymank A, Stout JC, Potts SG. 2011. Pollination ecology in the 21st Century: Key questions for future research. *Journal of Pollination Ecology* **3**: 8-23.

McAleece N, Gage JDG, Lambshead PJD, Paterson GLJ. 1997. BioDiversity Professional statistics analysis software.

Medel R, Botto-Mahan C, Kalin-Arroyo M. 2003. Pollinator-mediated selection on the nectar guide phenotype in the Andean Monkey flower, *Mimulus luteus*. *Ecology* **84**: 1721-1732.

Medel R, Valiente A, Botto-Mahan C, Carvallo G, Pérez F, Pohl N, Navarro L. 2007. The influence of insects and hummingbirds on the geographical variation of the flower phenotype in *Mimulus luteus*. *Ecography* **30**: 812-818.

Menzel R, Muller U. 1996. Learning and memory in Honeybees: From behavior to neural substrates. *Annual Review of Neuroscience* **19**: 379-404.

Metzker ML. 2010. Applications of Next-Generation Sequencing technologies - the next generation. *Nature Reviews Genetics* **11**: 31-46.

Miller JS, Diggle PK. 2007. Correlated evolution of fruit size and sexual expression in andromonoecious *Solanum* sections *Acanthophora* and *Lasiocarpa* (Solanaceae). *American journal of botany* **94**: 1706-1715.

Mitchell RJ, Irwin RE, Flanagan RJ, Karron JD. 2009. Ecology and evolution of plant-pollinator interactions. *Annals of Botany* **103**: 1355-1363.

Młodzińska E. 2009. Survey of plant pigments: molecular and environmental determinants of plant colors. *Acta Biologica Cracoviensia. Series Botanica* **51**: 7-16.

Mlot C. 1990. Restoring the prairie. *BioScience* **40**: 804-809.

Mohammadin S, Edger PP, Pires JC, Schranz ME. 2015. Positionally-conserved but sequence-diverged: Identification of long non-coding RNAs in the Brassicaceae and Cleomaceae. *BMC Plant Biology* **15**: 217.

Moldenke AR. 1976. California pollination ecology and vegetation types. *Phytologia* **34**.

Møller AP. 1995. Bumblebee preference for symmetrical flowers. *Proceedings of the National Academy of Sciences* **92**: 2288-2292.

Møller AP. 2000. Developmental stability and pollination. *Oecologia* **123**: 149-157.

Mondragón-Palomino M. 2013. Perspectives on MADS-box expression during orchid flower evolution and development. *Frontiers in Plant Science* **4**: 377.

Moss EH, Packer JG. 1983. *Flora of Alberta: A manual of flowering plants, conifers, ferns, and fern allies found growing without cultivation in the Province of Alberta, Canada.* Toronto; Buffalo: University of Toronto Press.

Nagpal P, Ellis CM, Weber H, Ploense SE, Barkawi LS, Guilfoyle TJ, Hagen G, Alonso JM, Cohen JD, Farmer EE, Ecker JR, Reed JW. 2005. Auxin response factors ARF6 and ARF8 promote jasmonic acid production and flower maturation. *Development* **132**: 4107-4118.

Nattero J, Cocucci AA, Medel R. 2010. Pollinator-mediated selection in a specialized pollination system: Matches and mismatches across populations. *Journal of Evolutionary Biology* **23**: 1957-1968.

Neal PR, Dafni A, Giurfa M. 1998. Floral symmetry and its role in plant-pollinator systems: Terminology, distribution, and hypotheses. *Annual Review of Ecology and Systematics* **29**: 345-373.

Newell CA, Brown NJ, Liu Z, Pflug A, Gowik U, Westhoff P, Hibberd JM. 2010. *Agrobacterium tumefaciens*-mediated transformation of *Cleome gynandra* L., a C(4) dicotyledon that is closely related to *Arabidopsis thaliana*. *Journal of Experimental Botany* **61**: 1311-1319.

Noman A, Aqeel M, Deng J, Khalid N, Sanaullah T, Shuilin H. 2017. Biotechnological advancements for improving floral attributes in ornamental plants. *Frontiers in Plant Science* **8**: 530.

Nozzolillo C, Amiguet VT, Bily AC, Harris CS, Saleem A, Andersen OM, 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.

Ó'Maoiléidigh DS, Graciet E, Wellmer F. 2014. Gene networks controlling *Arabidopsis thaliana* flower development. *New Phytologist* **201**: 16-30.

Ohmiya A, Hirashima M, Yagi M, Tanase K, Yamamizo C. 2014. Identification of genes associated with chlorophyll accumulation in flower petals. *PLoS ONE* **9**: e113738.

Ollerton J, Alarcón R, Waser NM, Price MV, Watts S, Cranmer L, Hingston A, Peter CI, Rotenberry J. 2009. A global test of the pollination syndrome hypothesis. *Annals of Botany* **103**: 1471-1480.

Ollerton J, Rech AR, Waser NM, Price MV. 2015. Using the literature to test pollination syndromes — Some methodological cautions. *Journal of Pollination Ecology* **16**: 119-125.

Ortiz-Barrientos D. 2013. The color genes of speciation in plants. *Genetics* **194**: 39-42.

Ouborg NJ, Vriezen WH. 2007. An ecologist's guide to ecogenomics. *Journal of Ecology* **95**: 8-16.

Owen CR, Bradshaw HD. 2011. Induced mutations affecting pollinator choice in *Mimulus lewisii* (Phrymaceae). *Arthropod-Plant Interactions* **5**: 235-244.

Ozsolak F, Milos PM. 2011. RNA sequencing: Advances, challenges and opportunities. *Nature Reviews Genetics* **12**: 87-98.

Pabón-Mora N, Ambrose BA, Litt A. 2012. Poppy APETALA1/FRUITFULL orthologs control flowering time, branching, perianth Identity, and fruit development. *Plant Physiology* **158**: 1685-1704.

Panchy N, Lehti-Shiu M, Shiu S-H. 2016. Evolution of gene duplication in plants. *Plant Physiology* **171**: 2294-2316.

Pashley CH, Ellis JR, McCauley DE, Burke JM. 2006. EST databases as a source for molecular markers: Lessons from *Helianthus*. *Journal of Heredity* **97**: 381-388.

Patchell MJ, Bolton MC, Mankowski P, Hall JC. 2011. Comparative floral development in Cleomaceae reveals two distinct pathways leading to monosymmetry. *International Journal of Plant Sciences* **172**: 352-365.

Patchell MJ, Roalson EH, Hall JC. 2014. Resolved phylogeny of Cleomaceae based on all three genomes. *Taxon* **63**: 315-328.

Patiny S. 2012. *Evolution of plant-pollinator relationships.* Cambridge: Cambridge University Press.

Pflieger S, Lefebvre V, Causse M. 2001. The candidate gene approach in plant genetics: A review. *Molecular Breeding* **7**: 275-291.

Policha T, Davis A, Barnadas M, Dentinger BT, Raguso RA, Roy BA. 2016. Disentangling visual and olfactory signals in mushroom-mimicking *Dracula* orchids using realistic three-dimensional printed flowers. *New Phytologist* **210**: 1058-1071.

Pootakham W, Jomchai N, Ruang-areerate P, Shearman JR, Sonthirod C, Sangsrakru D, Tragoonrung S, Tangphatsornruang S. 2015. Genome-wide SNP discovery and identification of QTL associated with agronomic traits in oil palm using genotyping-by-sequencing (GBS). *Genomics* **105**: 288-295.

Preston JC, Hileman LC. 2009. Developmental genetics of floral symmetry evolution. *Trends in Plant Science* **14**: 147-154.

Proctor MCF, Yeo P, Lack A. 1996. *The natural history of pollination.* Portland: Timber Press.

Quattrocchio F, Wing J, van der Woude K, Souer E, de Vetten N, Mol J, Koes R. 1999. Molecular analysis of the anthocyanin2 gene of *Petunia* and its role in the evolution of flower color. *Plant Cell* **11**: 1433-1444.

R Development Core Team. 2013. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing.

Raguso RA. 2008. Start making scents: The challenge of integrating chemistry into pollination ecology. *Entomologia Experimentalis et Applicata* **128**: 196-207.

Raju AJS, Rani DS. 2016. Reproductive ecology of *Cleome gynandra* and *Cleome viscosa* (Capparaceae). *Phytologia Balcanica* **22**: 15-28.

Ratcliff F, Martin-Hernandez AM, Baulcombe DC. 2001. Tobacco rattle virus as a vector for analysis of gene function by silencing. *Plant Journal* **25**: 237-245.

Rausher MD. 2008. Evolutionary transitions in floral color. *International Journal of Plant Sciences* **169**: 7-21.

Rausher MD, Fry JD. 1993. Effects of a locus affecting floral pigmentation in *Ipomoea purpurea* on female fitness components. *Genetics* **134**: 1237-1247.

Rifkin SA. 2012. *Quantitative trait loce (QTL): Methods and protocols.* New York: Springer.

Riser JP, Cardinal-McTeague WM, Hall JC, Hahn WJ, Sytsma KJ, Roalson EH. 2013. Phylogenetic relationships among the North American cleomoids (Cleomaceae): A test of Iltis's reduction series. *American Journal of Botany* **100**: 2102-2111.

Roy R, Schmitt AJ, Thomas JB, Carter CJ. 2017. Review: Nectar biology: From molecules to ecosystems. *Plant Science* **262**: 148-164.

Ruhlmann JM, Kram BW, Carter CJ. 2010. CELL WALL INVERTASE 4 is required for nectar production in *Arabidopsis*. *Journal of Experimental Botany* **61**: 395-404.

Ruiz MT, Voinnet O, Baulcombe DC. 1998. Initiation and maintenance of virus-induced gene silencing. *Plant Cell* **10**: 937-946.

Ruiz-Zapata T, Huérfano AA, Enrech N. 1996. Contribución al estudio citotaxonómico del género *Cleome* L. (Capparidaceae). *Phyton* **59**: 85-94.

Ryu C-M, Anand A, Kang L, Mysore KS. 2004. Agrodrench: A novel and effective agroinoculation method for virus-induced gene silencing in roots and diverse solanaceous species. *The Plant Journal* **40**: 322-331.

Samson F, Knopf F. 1994. Prairie conservation in North America. *BioScience* **44**: 418-421.

Sandring S, Agren J. 2009. Pollinator-mediated selection on floral display and flowering time in the perennial herb *Arabidopsis lyrata*. *Evolution* **63**: 1292-1300.

Sapir Y. 2009a. Effects of floral traits and plant genetic composition on pollinator behavior. *Arthropod-Plant Interactions* **3**: 115-129.

Sapir Y. 2009b. Pollination genetics: Using molecular genetic underlying floral traits to study pollination ecology in an evolutionary context. *Israel Journal of Plant Sciences* **57**: 141-149.

Sapir Y, Armbruster WS. 2010. Pollinator-mediated selection and floral evolution: From pollination ecology to macroevolution. *New Phytologist* **188**: 303-306.

Schemske DW, Bradshaw HD. 1999. Pollinator preference and the evolution of floral traits in Monkey flowers (*Mimulus*). *Proceedings of the National Academy of Sciences of the United States of America* **96**: 11910-11915.

Schiestl FP. 2005. On the success of a swindle: Pollination by deception in orchids. *Naturwissenschaften* **92**: 255-264.

Schiestl FP, Johnson SD. 2013. Pollinator-mediated evolution of floral signals. *Trends in Ecology & Evolution* **28**: 307-315.

Schiestl FP, Schlüter PM. 2009. Floral isolation, specialized pollination, and pollinator behavior in orchids. *Annual Review of Entomology* **54**: 425-446.

Schmid-Hempel P, Speiser B. 1988. Effects of Inflorescence Size on Pollination in *Epilobium angustifolium*. *Oikos* **53**: 98-104.

Schranz ME, Mitchell-Olds T. 2006a. Independent ancient polyploidy events in the sister families Brassicaceae and Cleomaceae. *Plant Cell* **18**: 1152-1165.

Schranz ME, Mitchell-Olds T. 2006b. Independent ancient polyploidy events in the sister families Brassicaceae and Cleomaceae. *The Plant Cell Online* **18**: 1152-1165.

Scopece G, Cozzolino S, Johnson Steven D, Schiestl Florian P. 2010. Pollination efficiency and the evolution of specialized deceptive pollination systems. *The American Naturalist* **175**: 98-105.

Sedeek KE, Qi W, Schauer MA, Gupta AK, Poveda L, Xu S, Liu ZJ, Grossniklaus U, Schiestl FP, Schluter PM. 2013. Transcriptome and proteome data reveal candidate genes for pollinator attraction in sexually deceptive orchids. *PLoS ONE* **8**: e64621.

Senthil-Kumar M, Hema R, Anand A, Kang L, Udayakumar M, Mysore KS. 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, Mysore KS. 2011. New dimensions for VIGS in plant functional genomics. *Trends in Plant Science* **16**: 656-665.

Senthil-Kumar M, Mysore KS. 2014. Tobacco rattle virus–based virus-induced gene silencing in *Nicotiana benthamiana*. *Nature Protocols* **9**: 1549-1562.

Sharma B, Kramer EM. 2013. Virus-induced gene silencing in the rapid cycling columbine *Aquilegia coerulea* "Origami". In: Becker A, ed. *Virus-Induced Gene Silencing: Methods and Protocols*. Totowa, NJ: Humana Press, 71-81.

Shcheglov AS, Zhulidov PA, Bogdanova EA, Shagin DA. 2007. Normalization of cDNA libraries. In: Buzdin AA, Lukyanov SA, eds. *Nucleic Acids Hybridization Modern Applications*. Dordrecht: Springer Netherlands, 97-124.

Sheehan H, Hermann K, Kuhlemeier C. 2012. Color and scent: How single genes influence pollinator attraction. *Cold Spring Harbor Symposia on Quantitative Biology* **77**: 117-133.

Sheffield CS, Frier SD, Dumesh S. 2014. The bees (Hymenoptera: Apoidea, Apiformes) of the Prairies Ecozone, with comparisons to other Grasslands of Canada. In: Giberson DJ, Carcamo HA, eds. *Arthropods of Canadian Grasslands*. Canada: Biological Survey of Canada, 427-467.

Simão FA, Waterhouse RM, Ioannidis P, Kriventseva EV, Zdobnov EM. 2015. BUSCO: Assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics* **31**: 3210-3212.

Sletvold N, Trunschke J, Smit M, Verbeek J, Ågren J. 2016. Strong pollinator-mediated selection for increased flower brightness and contrast in a deceptive orchid. *Evolution* **70**: 716-724.

Smith FS, Maywald PD, Lloyd-Reilley J, Maher SD, Pawelek KA, Scott Jr AW, Garza J. 2010. Zapata germplasm rio grande Clammyweed. *Native Plants Journal* **11**: 269-273.

Smith SD. 2010. Using phylogenetics to detect pollinator-mediated floral evolution. *New Phytologist* **188**: 354-363.

Smith SD. 2016. Pleiotropy and the evolution of floral integration. *New Phytologist* **209**: 80-85.

Smith SD, Ane C, Baum DA. 2008. The role of pollinator shifts in the floral diversification of *Iochroma* (Solanaceae). *Evolution* **62**: 793-806.

Smith-Unna R, Boursnell C, Patro R, Hibberd JM, Kelly S. 2016. TransRate: Reference-free quality assessment of *de novo* transcriptome assemblies. *Genome Research* **26**: 1134-1144.

Soares Neto RL. 2017. Flora das cangas da Serra dos Carajás, Pará, Brasil: Cleomaceae. *Rodriguésia* **68**: 931-933.

Sobral M, Veiga T, Domínguez P, Guitián JA, Guitián P, Guitián JM. 2015. Selective pressures explain differences in flower color among *Gentiana lutea* populations. *PLoS ONE* **10**: e0132522.

Specht C, Yockteng R, Almeida A, Kirchoff B, Kress WJ. 2012. Homoplasy, pollination, and emerging complexity during the evolution of floral development in the tropical Gingers (Zingiberales). *The Botanical Review* **78**: 440-462.

Spitzer-Rimon B, Farhi M, Albo B, Cna'ani A, Ben Zvi MM, Masci T, Edelbaum O, Yu Y, Shklarman E, Ovadis M, Vainstein A. 2012. The R2R3-MYB-like regulatory factor EOBI, acting downstream of EOBII, regulates scent production by activating ODO1 and structural scent-related genes in *Petunia*. *Plant Cell* **24**: 5089-5105.

Stellari GM, Jaramillo MA, Kramer EM. 2004. Evolution of the APETALA3 and PISTILLATA lineages of MADS-box-containing genes in the basal angiosperms. *Molecular Biology and Evolution* **21**: 506-519.

Stratmann JW, Hind SR. 2011. Gene silencing goes viral and uncovers the private life of plants. *Entomologia Experimentalis Et Applicata* **140**: 91-102.

Strickler SR, Bombarely A, Mueller LA. 2012. Designing a transcriptome next-generation sequencing project for a nonmodel plant species. *American Journal of Botany* **99**: 257-266.

Swarbreck D, Wilks C, Lamesch P, Berardini TZ, Garcia-Hernandez M, Foerster H, Li D, Meyer T, Muller R, Ploetz L, Radenbaugh A, Singh S, Swing V, Tissier C, Zhang P, Huala E. 2008. The *Arabidopsis* Information Resource (TAIR): Gene structure and function annotation. *Nucleic Acids Research* **36**: D1009-D1014.

Tsai YT, Chen PY, To KY. 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 Reports* **31**: 1189-1198.

Tucker GC, Vanderpool SS. 2010. Cleomaceae. In: Committee FNAE, ed. *Flora of North America North of Mexico*. New York: Oxford University Press, 199-223.

Ungerer MC, Johnson LC, Herman MA. 2007. Ecological genomics: Understanding gene and genome function in the natural environment. *Heredity* **100**: 178-183.

Unver T, Budak H. 2009. Virus-induced gene silencing, a post transcriptional gene silencing method. *International Journal of Plant Genomics* **2009**.

van den Bergh E. 2017. *Comparative genomics and trait evolution in Cleomaceae, a model for ancient polyploidy,* PhD, Wageningen University

van den Bergh E, Hofberger JA, Schranz ME. 2016a. Flower power and the mustard bomb: Comparative analysis of gene and genome duplications in glucosinolate biosynthetic pathway evolution in Cleomaceae and Brassicaceae. *American Journal of Botany* **103**: 1212-1222.

van den Bergh E, Hofberger JA, Schranz ME. 2016b. Flower power and the mustard bomb: Comparative analysis of gene and genome duplications in glucosinolate biosynthetic pathway evolution in Cleomaceae and Brassicaceae. *American Journal of Botany* **103**: 1212-1222.

van den Bergh E, Külahoglu C, Bräutigam A, Hibberd JM, Weber AP, Zhu X-G, Schranz 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 Niet T, Peakall R, Johnson SD. 2014. Pollinator-driven ecological speciation in plants: New evidence and future perspectives. *Annals of Botany* **113**: 199-212.

van Dijk EL, Auger H, Jaszczyszyn Y, Thermes C. 2014. Ten years of next-generation sequencing technology. *Trends in Genetics* **30**: 418-426.

Verdonk JC, Haring MA, van Tunen AJ, Schuurink RC. 2005. ODORANT1 regulates fragrance biosynthesis in *Petunia* flowers. *Plant Cell* **17**: 1612-1624.

Voznesenskaya EV, Koteyeva NK, Chuong SDX, Ivanova AN, Barroca J, Craven LA, Edwards GE. 2007. Physiological, anatomical and biochemical characterisation of photosynthetic types in genus *Cleome* (Cleomaceae). *Functional Plant Biology* **34**: 247-267.

Wallis C. 2001. *Proposed protocols for inventories of rare plants of the Grassland Natural Region.* Edmonton: Alberta Sustainable Resource Development, Fish and Wildlife Service.

Wang C, Cai X, Wang X, Zheng Z. 2006. Optimisation of tobacco rattle virus-induced gene silencing in *Arabidopsis*. *Functional Plant Biology* **33**: 347-355.

Wang L, Li PH, Brutnell TP. 2010a. Exploring plant transcriptomes using ultra high-throughput sequencing. *Briefings in Functional Genomics* **9**: 118-128.

Wang T, Li R, Wen L, Fu D, Zhu B, Luo Y, Zhu H. 2015. Functional analysis and RNA sequencing indicate the regulatory role of Argonaute1 in tomato compound leaf development. *PLoS ONE* **10**: e0140756.

Wang X, Cairns MJ. 2013. Gene set enrichment analysis of RNA-Seq data: Integrating differential expression and splicing. *BMC Bioinformatics* **14**: 1-10.

Wang X, Zhu H, Shao Y, Chen A, Ma Y, Luo Y, Zhu B. 2010b. Analysis of gene functions by a syringe infiltration method of VIGS. *Russian Journal of Plant Physiology* **57**: 590-597.

Wang Z, Gerstein M, Snyder M. 2009. RNA-Seq: a revolutionary tool for transcriptomics. *Nature Reviews Genetics* **10**: 57-63.

Ward JA, Ponnala L, Weber CA. 2012. Strategies for transcriptome analysis in nonmodel plants. *American Journal of Botany* **99**: 267-276.

Waser NM, Chittka L, Price MV, Williams NM, Ollerton J. 1996. Generalization in pollination systems, and why it matters. *Ecology* **77**: 1043-1060.

Wege S, Scholz A, Gleissberg S, 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. *Annals of Botany* **100**: 641-649.

Weigel D, Glazebrook J. 2002. *Arabidopsis: A laboratory manual.* New York: Cold Spring Harbor Laboratoty Press.

Wellmer F, Riechmann JL. 2010. Gene networks controlling the initiation of flower development. *Trends in Genetics* **26**: 519-527.

Wessinger CA. 2015. A genetic route to yellow flowers. *New Phytologist* **206**: 1193-1195.

Wessinger CA, Hileman LC, Rausher MD. 2014. Identification of major quantitative trait loci underlying floral pollination syndrome divergence in *Penstemon*. *Philosophical Transactions of the Royal Society B: Biological Sciences* **369**.

Wessinger CA, Rausher MD. 2012. Lessons from flower colour evolution on targets of selection. *Journal of Experimental Botany* **63**: 5741-5749.

Wessinger CA, Rausher MD. 2014. Predictability and irreversibility of genetic changes associated with flower color evolution in *Penstemon barbatus*. *Evolution* **68**: 1058-1070.

Whibley AC, Langlade NB, Andalo C, Hanna AI, Bangham A, Thébaud C, Coen E. 2006. Evolutionary paths underlying flower color variation in *Antirrhinum*. *Science* **313**: 963-966.

Whittall JB, Hodges SA. 2007. Pollinator shifts drive increasingly long nectar spurs in Columbine flowers. *Nature* **447**: 706-709.

Whittall JB, Voelckel C, Kliebenstein DJ, Hodges SA. 2006. Convergence, constraint and the role of gene expression during adaptive radiation: Floral anthocyanins in *Aquilegia*. *Molecular Ecology* **15**: 4645-4657.

Wiens D. 1984. Ovule survivorship, brood size, life history, breeding systems,and reproductive success in plants. *Oecologia* **64**: 47-53.

Williams BP, Burgess SJ, Reyna-Llorens I, Knerova J, Aubry S, Stanley S, Hibberd JM. 2016. An untranslated cis element regulates the accumulation of multiple C4 enzymes in *Gynandropsis gynandra* mesophyll cells. *The Plant Cell* **28**: 454-465.

Willmer P. 2011. *Pollination and floral ecology.* Princeton: Princeton University Press.

Wilson P, Castellanos MC, Hogue JN, Thomson JD, Armbruster WS. 2004. A multivariate search for pollination syndromes among Penstemons. *Oikos* **104**: 345-361.

Winkel-Shirley B. 2001. Flavonoid biosynthesis. A colorful model for genetics, biochemistry, cell biology, and biotechnology. *Plant Physiology* **126**: 485-493.

Xu SQ, Schluter PM, Grossniklaus U, Schiestl FP. 2012. The genetic basis of pollinator adaptation in a sexually deceptive orchid. *Plos Genetics* **8**.

Yant L, Collani S, Puzey J, Levy C, Kramer EM. 2015. Molecular basis for three-dimensional elaboration of the *Aquilegia* petal spur. *Proceedings of the Royal Society of London B: Biological Sciences* **282**.

Yuan YW, Byers KJRP, Bradshaw HD. 2013*a*. The genetic control of flower-pollinator specificity. *Current Opinion in Plant Biology* **16**: 422-428.

Yuan YW, Saqawa JM, Young RC, Christensen BJ, Bradshaw HDJ. 2013*b*. Genetic dissection of a major anthocyanin QTL contributing to pollinator-mediated reproductive isolation between sister species of *Mimulus*. *Genetics* **194**: 255-263.

Zhang S, Zhang J-S, Zhao J, He C. 2015a. Distinct subfunctionalization and neofunctionalization of the B-class MADS-box genes in *Physalis floridana*. *Planta* **241**: 387-402.

Zhang W, Kramer EM, Davis CC. 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.

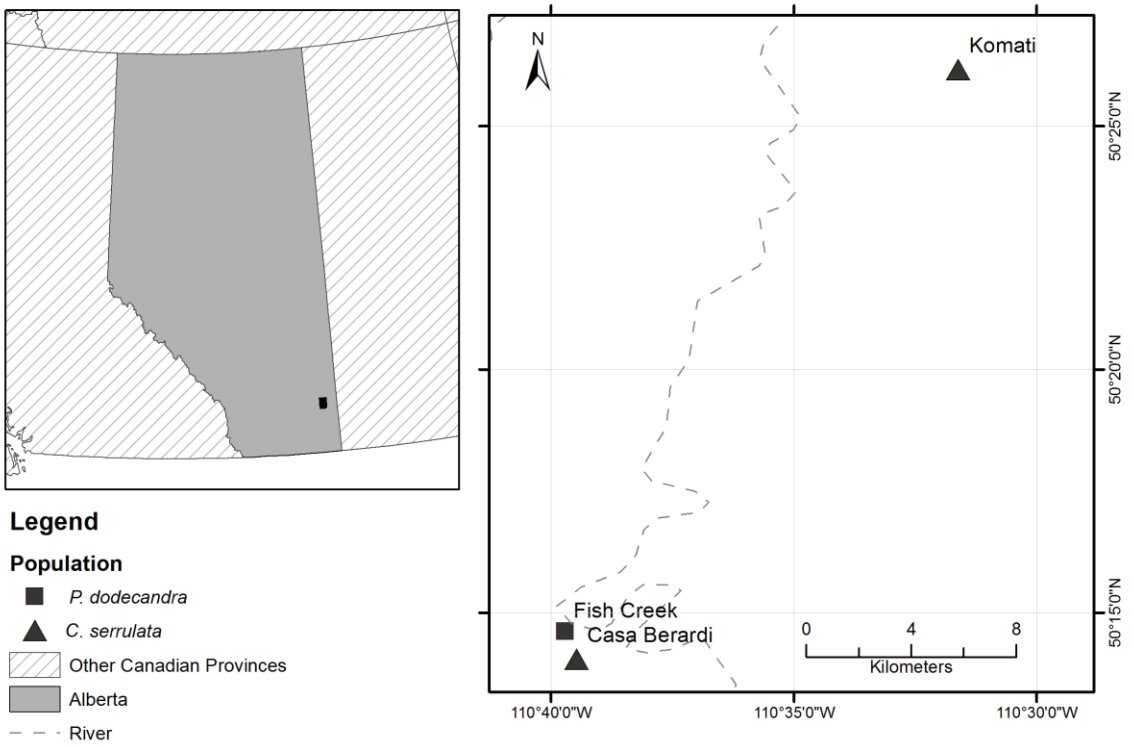
Zhang W, Wei X, Meng H-L, Ma C-H, Jiang N-H, Zhang G-H, Yang S-C. 2015b. Transcriptomic comparison of the self-pollinated and cross-pollinated flowers of *Erigeron breviscapus* to analyze candidate self-incompatibility-associated genes. *BMC Plant Biology* **15**: 248.

Zhao Z-G, Wang Y-K. 2015. Selection by pollinators on floral traits in generalized *Trollius ranunculoides* (Ranunculaceae) along Altitudinal Gradients. *PLoS ONE* **10**: e0118299.

Zou Z, Ishida M, Li F, Kakizaki T, Suzuki S, Kitashiba H, Nishio T. 2013. QTL analysis using SNP markers developed by Next-Generation Sequencing for identification of candidate genes controlling 4-Methylthio-3-Butenyl Glucosinolate contents in roots of radish, *Raphanus sativus* L. *PLoS ONE* **8**: e53541.

Appendices

Appendix 3.1. Map representing the studied natural populations in Suffield NWA. *Cleomella serrulata* on Casa Berardi and Komati areas. *Polanisia dodecandra* on Fish Creek area.



Appendix 3.2. Student t-test across plant populations and plant species between sampling periods. CB=Casa Berardi, FC=Fish creek, KM=Komati. (N=100 per site, except for nectar variables N=25 per site). Inflorescences per plant N/A=data did not fit into a normal distribution. N/A=missing data. Significant P-values in **bold**.

	Cleomella **CB 2013 – CB2014**		***Cleomella*** **CB 2013 – KM 2013**		***Cleomella*** **KM 2013** ***Polanisia*** **FC 2013**		***Cleomella*** **CB 2013** ***Polanisia*** **FC 2013**		***Cleomella*** **CB 2014-** ***Polanisia*** **FC 2014**		***Polanisia*** **FC 2013-FC 2014**	
	t-value	***P*-value**	**t-value**	***P*-value**	**t-value**	***P*-value**	**t-value**	***P*-value**	**t-value**	***P*-value**	**t-value**	***P*-value**
Plant height	1.965	**0.050**	1.524	0.129	4.421	**1.67E-05**	6.253	**2.63E-09**	18.9	**2.20E-16**	15.622	**2.20E-16**
Inflorescences per plant	3.243	**0.001**	0.065	0.948	N/A	N/A	N/A	N/A	9.176	**4.56E-15**	N/A	N/A
Flowers per inflorescence	9.784	**2.20E-16**	0.110	0.912	4.885	**2.71E-06**	8.729	**1.46E-15**	6.41	**2.28E-09**	3.437	**0.0007**
Fruit set	4.465	**1.37E-05**	4.304	**2.68E-05**	1.564	0.120	6.530	**1.62E-09**	3.631	**0.0004**	4.165	**4.65E-05**
Nectar sugar concentration	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	4.601	**3.654e-05**	N/A	N/A
Nectar volume	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	0.172	0.864	N/A	N/A
Number of visits	1.979	**0.050**	0.851	0.396	2.289	**0.024**	2.793	**6.08E-03**	2.207	**0.029**	1.186	0.240

Appendix 3.3. Mann-Whitney U test across plant populations and plant species between sampling periods. CB=Casa Berardi, FC=Fish creek, KM=Komati. (N=100 per site). Significant P-values in **bold.**

	***Cleomella* KM 2013** ***Polanisia* FC 2013**		***Cleomella* CB 2013** ***Polanisia* FC 2013**		***Polanisia* FC 2013-FC 2014**	
	U-value	***P*-value**	**U-value**	***P*-value**	**U-value**	***P*-value**
Inflorescences per plant	6708	**2.54E-05**	8028	**1.04E-13**	2899.5	**1.46E-07**

Appendix 3.4. Insect and spider visitors found on *Cleomella serrulata* Casa Berardi and Komati populations, and *Polanisia dodecandra* Fish Creek populations. F=females. M=males.

Order/Family/Species	Casa Berardi *C. serrulata*	Komati *C. serrulata*	Fish Creek *P. dodecandra*	Specimens collected
Araneae				
Thomisidae	3		5	1
Coleoptera				
Coleoptera sp. 1	19			1
Coleoptera sp. 2	4			1
Coleoptera sp. 4	1			1
Chrysomelidae				
Chrysomelidae sp. 2	1			1
Disonycha sp.	1		1	1
Coccinellidae				
Coccinellidae sp. 1	1	1		1
Coccinellidae sp. 2	1	3		1
Meloidae				
Coleoptera sp. 1			4	1
Coleoptera sp. 3	1			1
Lytta nuttalli (Say, 1824)		2		2
Nitidulidae				
Nitidulidae sp.	243	19	28	4
Diptera				
Diptera sp. 1	21	10	9	2
Diptera sp. 10	1			1
Diptera sp. 12	1			1
Diptera sp. 2	1			1
Diptera sp. 3	8	8		1
Diptera sp. 4		2		1

Diptera sp. 5		4		1
Diptera sp. 6		1		1
Diptera sp. 7		1		1
Diptera sp. 8	2	3		1
Diptera sp. 9	1			1
Asilidae				
Asilidae sp. 1			2	1
Bombyliidae				
Bombyliidae sp. 1	4		3	2
Bombyliidae sp. 2	2		1	1
Bombyliidae sp. 3	1			1
Poecilanthrax sp.	5			4
Villa nigropecta Cresson, 1916		1	2	1
Calliphoridae				
Calliphoridae sp. 1	3		1	1
Calliphoridae sp. 2	1			1
Conopidae				
Physocephala texana (Williston, 1882)		2		1
Culicidae				
Culicidae sp. 1	23	2	6	10
Sarcophagidae				
Sarcophagidae sp. 1	10	1	5	2
Sarcophagidae sp. 2	7		1	1
Sarcophagidae sp. 3	6	1		2
Syrphidae				
Syrphidae sp. 2	3	5	1	1
Syrphidae sp. 3	11	11	6	1
Syrphidae sp. 4	6	3	12	3
Syrphidae sp. 5	3	2	2	1
Syrphidae sp. 6		6	1	2
Syrphidae sp. 7	1	8	1	2
Syrphidae sp. 8	5	4		1

Toxomerus marginatus (Meigen, 1822)	5	17	8	4
Tephritidae				
Tephritidae sp. 1	2	1		1
Ulidiidae				
Oedopa capito Loew, 1868	1	2		1
Ephemeroptera				
Ephemeroptera sp. 1			1	1
Hemiptera				
Hemiptera sp. 1		1		3
Hemiptera sp. 2			1	1
Aphididae				
Aphididae sp. 1	29		12	1
Cercopidae				
Cercopidae sp. 1	1			1
Cicadellidae				
Cicadellidae sp. 1	1			1
Miridae				
Miridae sp. 1	14			1
Miridae sp. 2	7	1	2	1
Pentatomidae				
Banasa sp.	9	2		3
Reduviidae				
Phymata americana (Melin, 1930)	5	69		4
Tingidae				
Tingidae sp. 1	3			1
Hymenoptera (unidentified families)				
Hymenoptera sp. 1			1	1F
Hymenoptera sp. 10	1			1F
Hymenoptera sp. 11	1			1F
Hymenoptera sp. 16	1			1F
Hymenoptera sp. 17	1			1F

Hymenoptera sp. 18	3			1F
Hymenoptera sp. 19	1			1F
Hymenoptera sp. 7		1		1F
Hymenoptera sp. 8	1			1F
Andrenidae				
Andrena prunorum (Cockerell, 1896)		2	2	3F
Andrena sp.	1			1F
Perdita sp. 1	71	3	6	1F
Perdita sp. 2	4			1F
Apidae				
Anthophora occidentalis Cresson, 1869	7	1	3	1F
Apis mellifera L, 1758	1		1	1F
Bombus borealis Kirby, 1837	23	16	3	4F, 2M
Bombus fervidus (F, 1798)	5	1		1M
Bombus griseocollis (Degeer, 1773)	9	1		1F
Bombus huntii Greene, 1860	3	5		3F
Bombus nevadensis Cresson, 1874	4			1F
Bombus ternarius (Say, 1837)	133	7	4	3F
Melissodes sp.	1			1F
Braconidae				
Braconidae sp. 2	8	1		2
Braconidae sp. 3		1	1	1
Braconidae sp. 4	2			2
Chrysididae				
Chrysididae sp. 2	1	2		2
Chrysididae sp. 4	1		1	1
Chrysis sp.	2	1	1	1
Hedychrum cupicicollis	1	1		1
Colletidae				
Colletes sp. 1	4			1M

Colletes sp. 2	2	6	2	2F
Crabronidae				
Bembix americana F, 1793	1			1
Cerceris sp.	40	1		1
Ectemnius rufifemur (Packard, 1866)		4		1
Eucerceris superba (Cresson, 1865)	1	1		1
Microbembex monodonta (Say, 1824)	6	4	5	1F, 3M
Philanthus gloriosus Cresson, 1865		1		1
Nysson plagiatus (Cresson, 1882)	3	2		2
Tachysphex sp. 1	1	3	2	1
Tachysphex sp. 2	11	1		1
Formicidae				
Formicidae sp. 1	52	6	3	5
Formicidae sp. 2	194	11		1
Formicidae sp. 3	176	15		8
Formicidae sp. 4	29	3	2	2
Formicidae sp. 5	150	6	3	1
Halictidae				
Agapostemon texanus Cresson, 1872	1	1		1
Halictus confusus Smith, 1853	2		1	1
Lasioglossum sp. 1	2	1		1
Lassioglossum sp. 2			1	1
Lassioglosum sp. 3	1			1
Lassioglossum sp. 4			1	1
Lassioglossum sp. 5			1	1
Ichneumonidae				
Ichneumonidae sp. 1	2		2	1
Ichneumonidae sp. 2			1	1
Megachilidae				

Dianthidium sp.			1	1F
Megachile dentitarsus Sladen, 1919	4	5	1	2
Megachile texana Cresson, 1878	11	3	2	1
Mutilidae				
Dasymutilla bioculata (Cresson, 1865)	1	1	1	1F, 2M
Perilampidae				
Perilampus sp. 1	1		1	1
Perilampus sp. 2	3			1
Pompilidae				
Anoplius sp.	1	4		
Pompilidae sp. 1	2	2		2
Pompilidae sp. 2	10	3	1	2
Sphecidae				
Ammophila azteca Cameron, 1888	2	9		1
Ammophila sp.	1	1		1
Podalonia valida (Cresson, 1865)	4	6		2
Sphecidae sp. 6		1		1
Sphex ichneumoneus (L, 1758)	9	54		5
Tiphiidae				
Tiphia sp.	9	4	1	1
Vespidae				
Euodynerus auranus (Cameron, 1906)		5		2
Pterocheilus quinquefasciatus Say, 1824	1	1		1
Vespidae sp. 3		2	1	1
Vespidae sp. 9		1		1
Lepidoptera (unidentified families)				
Moth sp. 4	1			1
Moth sp. 5	2			2
Arctiidae				
Estigmene acrea (Drury, 1773)		13		1

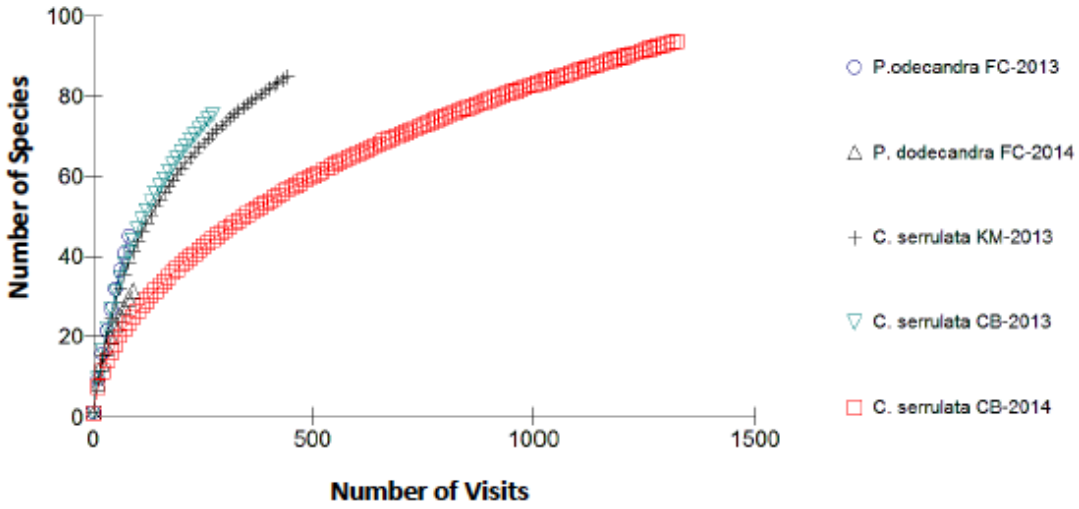
Crambidae				
Loxostege sp.	1		1	1
Erebidae				
Caenurgina sp.			2	1
Hesperiidae				
Erynnis sp.	1			1
Hesperia comma L, 1758	7			1
Pyrgus communis (Grote, 1872)	7	3		1
Lycaenidae				
Lycaeides melissa (Edwards, 1873)	15			1F, 3M
Phycioides tharos (Drury, [1773])	1			4
Noctuidae				
Cryptocala acadiensis (Bethune, 1870)	3			1
Schinia meadi Grote, 1873	2			1
Nymphalidae				
Cercyonis pegala F, 1775	5	4		2
Phycioides tharos (Drury, [1773])	3	2		4
Speyeria callippe (Boisduval, 1852)	3			1F
Vanessa cardui L, 1758	1			1
Pieridae				
Colias philodice Godart, [1819]	8	9	1	1
Pontia occidentalis (Reakirt, 1866)	3	1	1	1F, 3M
Orthoptera				
Acrididae				
Acrididae sp.	25	4		1
Tettigoniidae				
Phaneropterinae sp. 1		1	1	1M
Tettigoniidae sp. 1	1	1	9	2M

Thysanoptera		
Thripidae		
Thripidae sp. 1	1	1

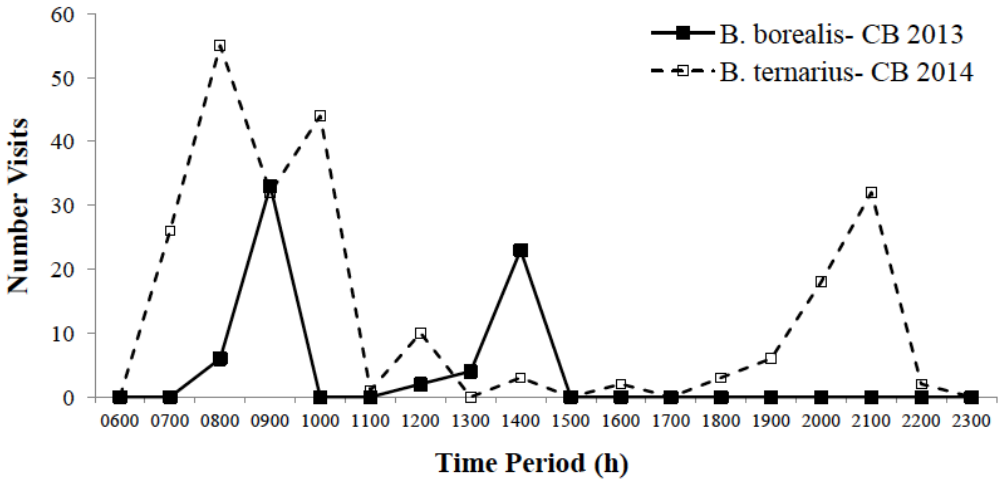
Appendix 3.5. Percentage frequency of the top five insect species visiting *Cleomella serrulata* and *Polanisia dodecandra* excluding permanent residents. CB= Casa Berardi, KM=Komati, and FC=Fish Creek area. Data collected in 2013 and 2014. n=100.

Species	**CB 2013 *C. serrulata***	**CB 2014 *C. serrulata***	**KM 2013 *C. serrulata***	**FC 2013 *P. dodecandra***	**FC 2014 *P. dodecandra***
Ammophila azteca			3.03		
Andrena prunorum				3.33	
Anthophora occidentalis					6.25
Bombus borealis	23.94		8.35		
Bombus ternarius		34.31	4.30		6.25
Cerceris sp.		6.15			
Colletes sp. 2				3.33	
Coleoptera sp. 1		2.93			
Culicidae sp. 1		4.10			7.50
Megachile dentitarsus	4.92				
Megachile texana	5.63				
Microbembex monodonta				8.33	
Perdita sp. 1		14.22			11.25
Sphex ichneumoneus	3.52		22.27		
Syrphidae sp. 3	3.16				
Syrphidae sp. 4				3.33	18.75
Toxomerus marginatus			4.55	6.67	

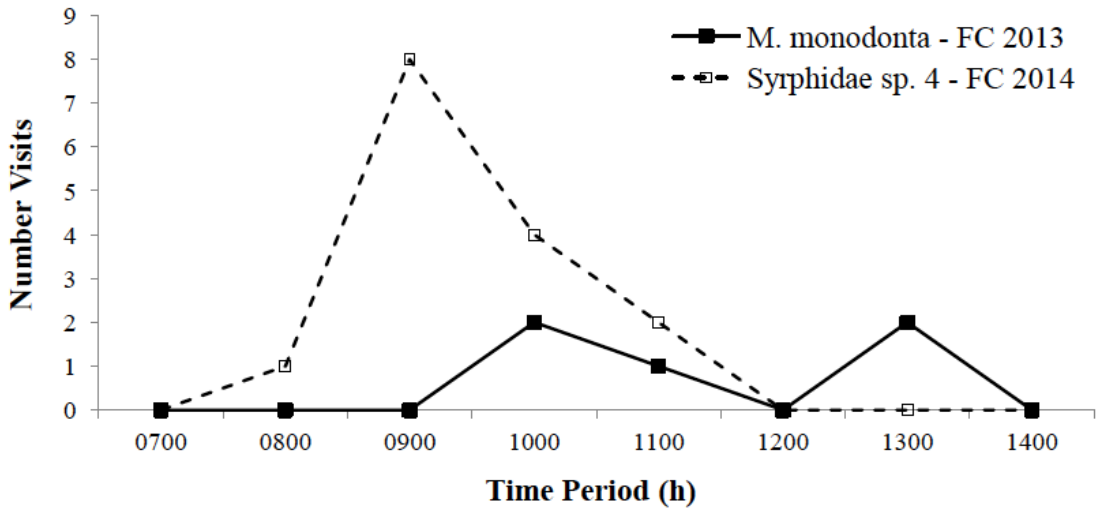
Appendix 3.6. Rarefactions plot across plant populations and plant species between sampling periods. CB=Casa Berardi, FC=Fish creek, KM=Komati.



Appendix 3.7. Number of visits of the most frequent species in *Cleomella serrulata.* CB= Casa Berardi. Data collected in 2014 and 2014.



Appendix 3.8. Number of visits of the most frequent species in *Polanisia dodecandra.* FC= Fish Creek. Data collected in 2014 and 2014.



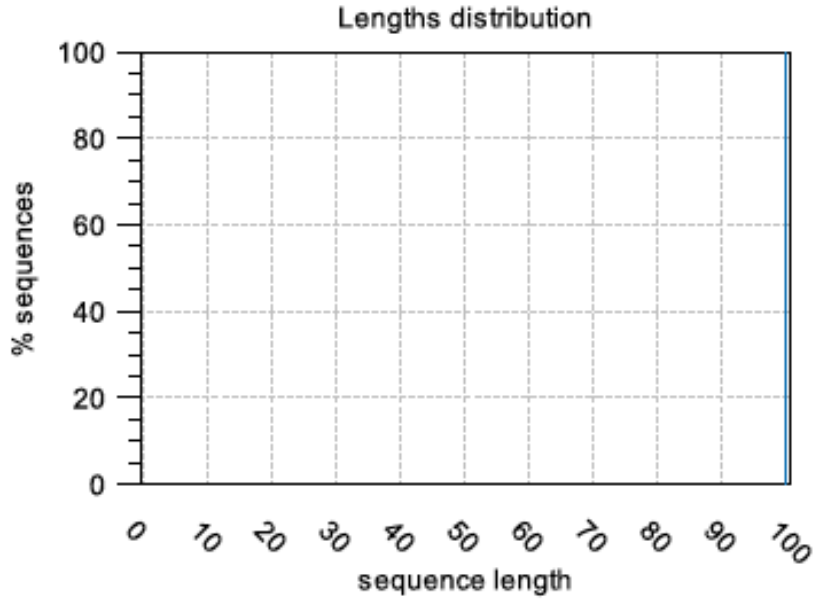
Appendix 4.1. *Cleomella serrulata* FastQC report.

1. Summary

Creation date:	Mon Mar 07 16:36:34 MST 2016
Generated by:	Deyholos
Software:	CLC Genomics Workbench 7.5
Based upon:	1 data set
CS_R1 (paired):	95,616,414 sequences in pairs
Total nucleotides in data set	9,561,641,400 nucleotides

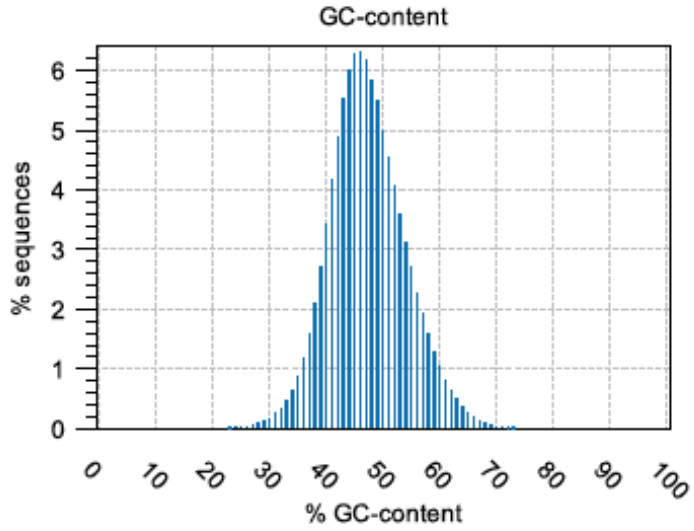
2. Per-sequence analysis

2.1 Lengths distribution



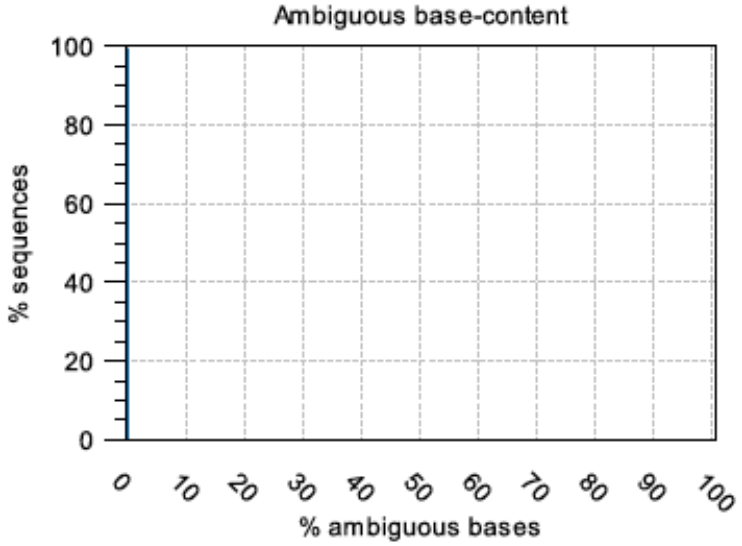
Distribution of sequence lengths. In cases of untrimmed Illumina or SOLiD reads it will ju st contain a single peak.
x: sequence length in base-pairs
y: number of sequences featuring a particular length normalized to the total number of seq uences

2.2 GC-content



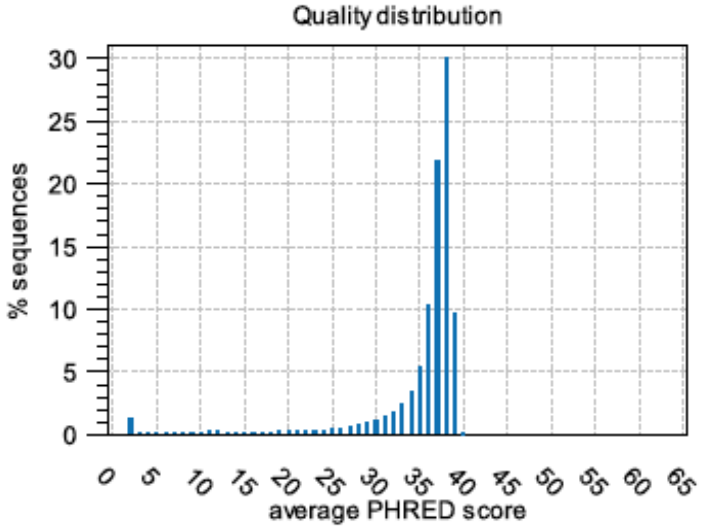
Distribution of GC-contents. The GC-content of a sequence is calculated as the number of G C-bases compared to all bases (including ambiguous bases).
x: relative GC-content of a sequence in percent
y: number of sequences featuring particular GC-percentages normalized to the total number of sequences

2.3 Ambiguous base-content



Distribution of N-contents. The N-content of a sequence is calculated as the number of amb iguous bases compared to all bases.
x: relative N-content of a sequence in percent
y: number of sequences featuring particular N-percentages normalized to the total number o f sequences

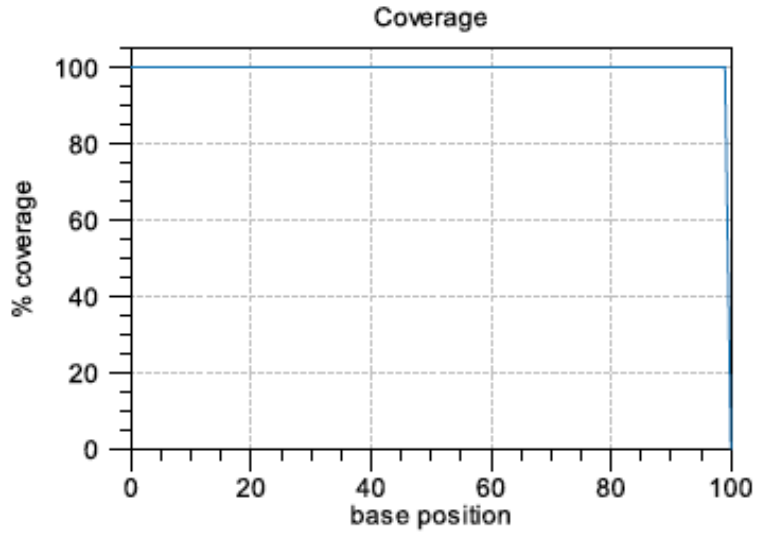
2.4 Quality distribution



Distribution of average sequence qualitie scores. The quality of a sequence is calculated as the arithmetic mean of its base qualities.
x: PHRED-score
y: number of sequences observed at that qual. score normalized to the total number of sequ ences

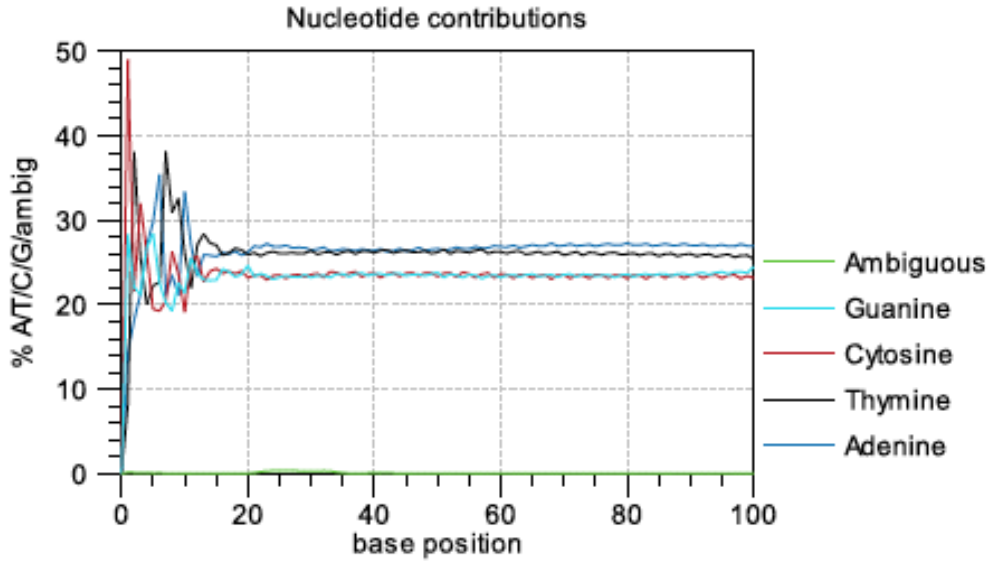
3. Per-base analysis

3.1 Coverage



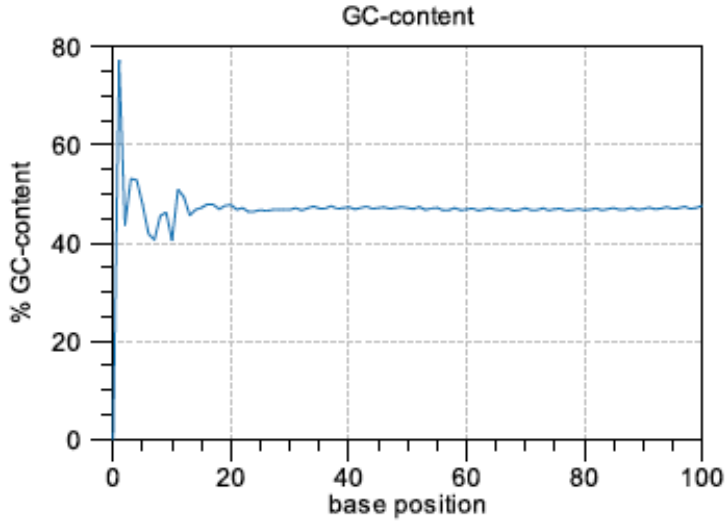
The number of sequences that support (cover) the individual base positions. In cases of un trimmed Illumina or SOLiD reads it will just contain a rectangle.
x: base position
y: number of sequences covering individual base positions normalized to the total number o f sequences

3.2 Nucleotide contributions



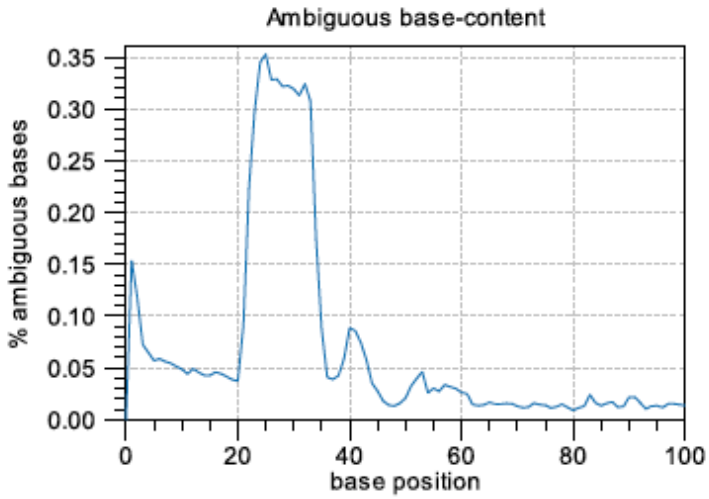
Coverages for the four DNA nucleotides and ambiguous bases.
x: base position
y: number of nucleotides observed per type normalized to the total number of nucleotides o bserved at that position

3.3 GC-content



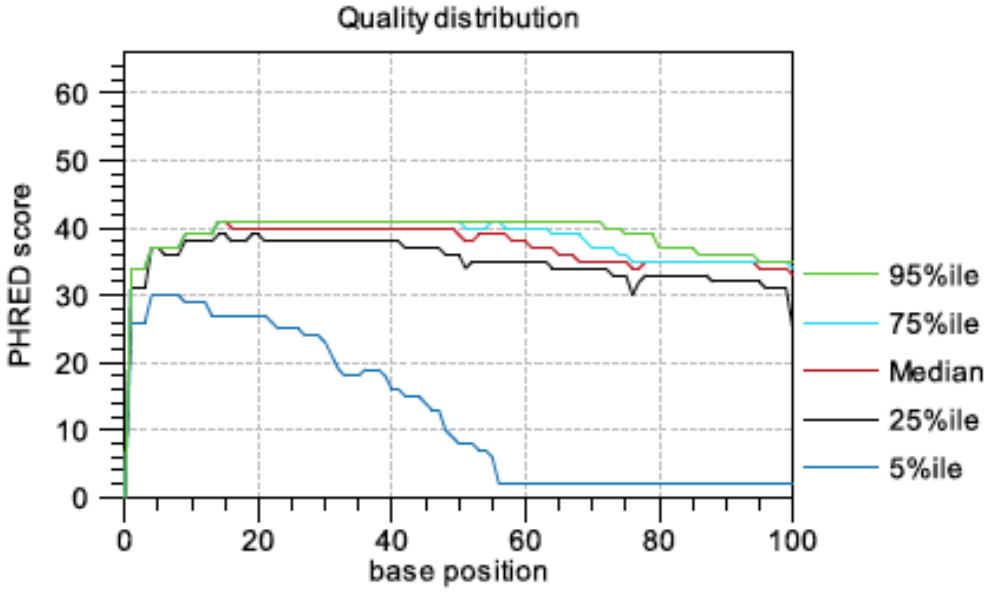
Combined coverage of G- and C-bases.
x: base position
y: number of G- and C-bases observed at current position normalized to the total number of bases observed at that position

3.4 Ambiguous base-content



Combined coverage of ambiguous bases.
x: base position
y: number of ambiguous bases observed at current position normalized to the total number o f bases observed at that position

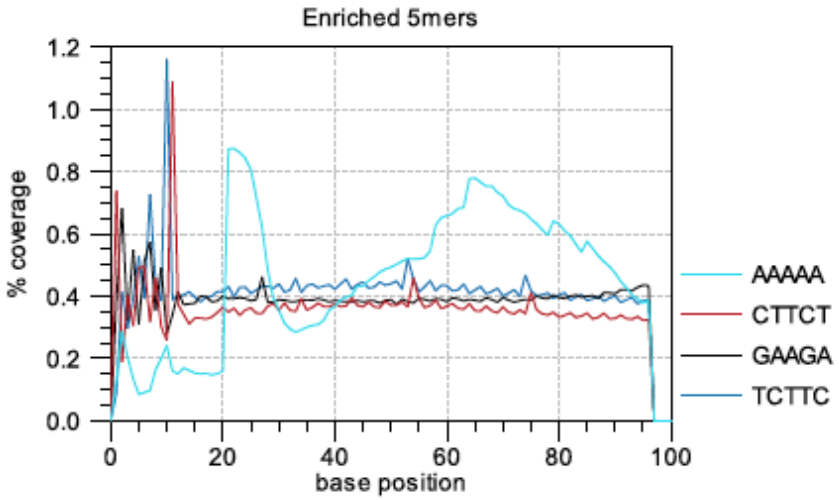
3.5 Quality distribution



Base-quality distribution along the base positions.
x: base position
y: median & percentiles of quality scores observed at that base position

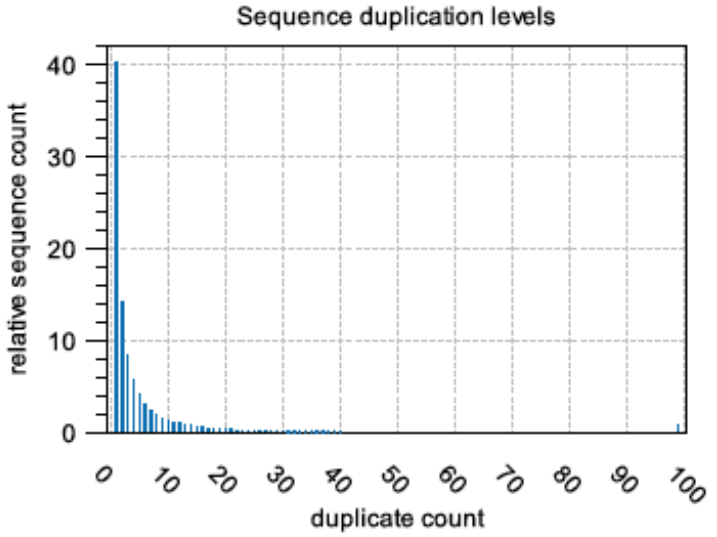
4. Over-representation analyses

4.1 Enriched 5mers



The five most-overrepresented 5mers. The over-representation of a 5mer is calculated as th e ratio of the observed and expected 5mer frequency. The expected frequency is calculated as product of the empirical nucleotide probabilities that make up the 5mer. (5mers that contain ambiguous bases are ignored)
x: base position
y: number of times a 5mer has been observed normalized to all 5mers observed at that posit ion

4.2 Sequence duplication levels



Duplication level distribution. Duplication levels are simply the count of how often a par ticular sequence has been found.
x: duplicate count
y: number of sequences that have been found that many times normalized to the number of un ique sequences

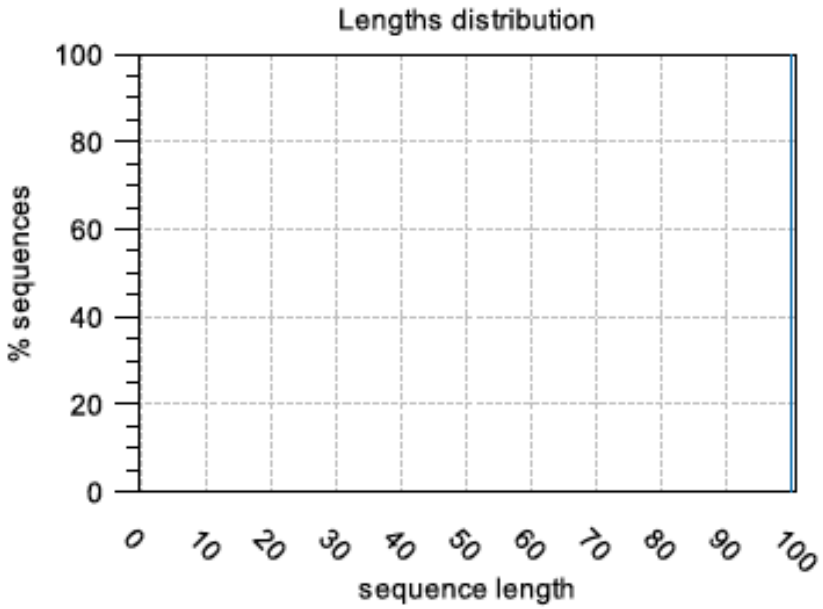
Appendix 4.2. *Melidiscus giganteus* FastQC report**.**

1. Summary

Creation date:	Fri Mar 04 15:52:43 MST 2016
Generated by:	Deyholos
Software:	CLC Genomics Workbench 7.5
Based upon:	1 data set
cvr1 (paired):	130,268,202 sequences in pairs
Total nucleotides in data set	13,026,820,200 nucleotides

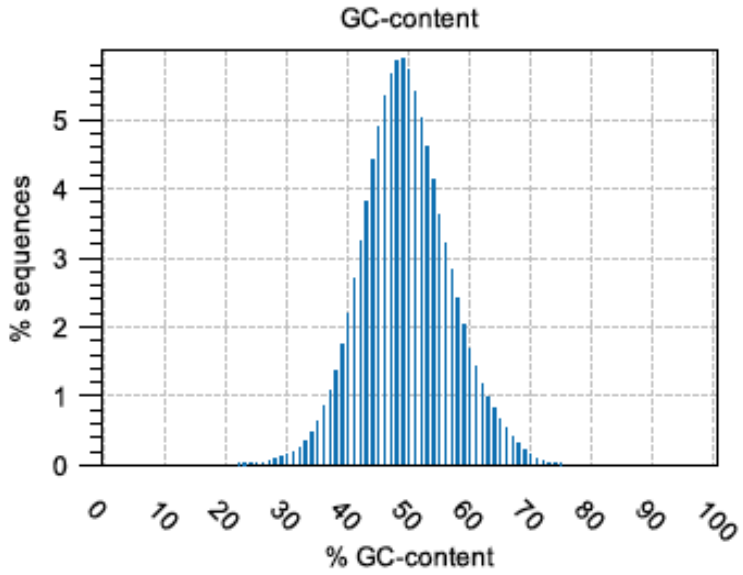
2. Per-sequence analysis

2.1 Lengths distribution



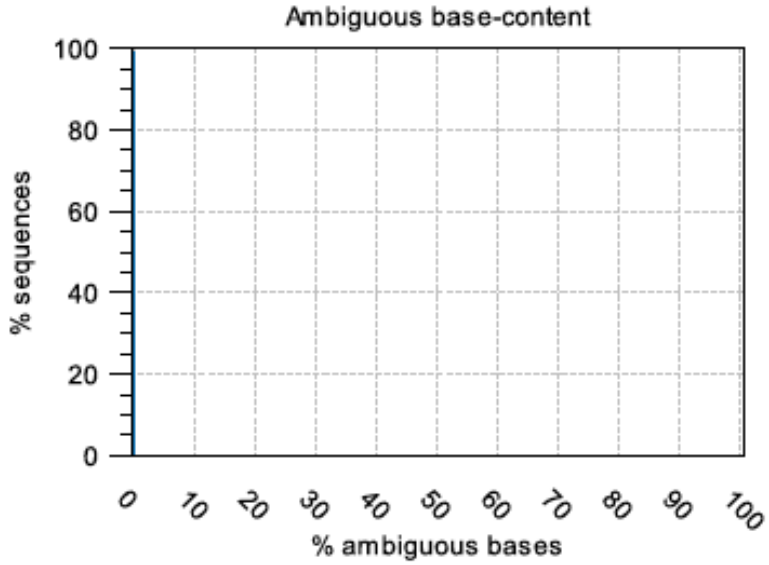
Distribution of sequence lengths. In cases of untrimmed Illumina or SOLiD reads it will ju st contain a single peak.
x: sequence length in base-pairs
y: number of sequences featuring a particular length normalized to the total number of seq uences

2.2 GC-content



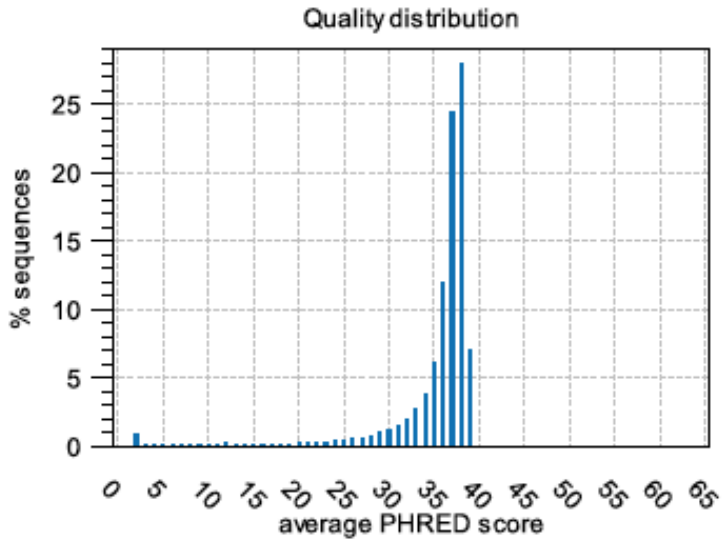
Distribution of GC-contents. The GC-content of a sequence is calculated as the number of G C-bases compared to all bases (including ambiguous bases).
x: relative GC-content of a sequence in percent
y: number of sequences featuring particular GC-percentages normalized to the total number of sequences

2.3 Ambiguous base-content



Distribution of N-contents. The N-content of a sequence is calculated as the number of amb iguous bases compared to all bases.
x: relative N-content of a sequence in percent
y: number of sequences featuring particular N-percentages normalized to the total number o f sequences

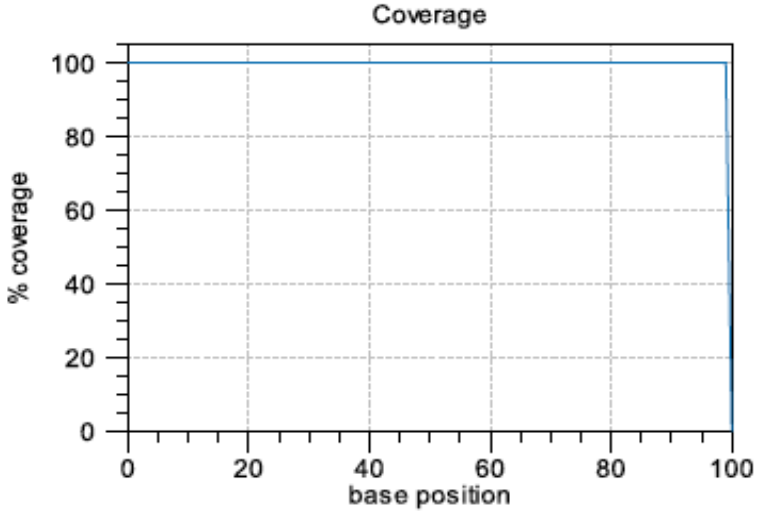
2.4 Quality distribution



Distribution of average sequence qualitie scores. The quality of a sequence is calculated as the arithmetic mean of its base qualities.
x: PHRED-score
y: number of sequences observed at that qual. score normalized to the total number of sequ ences

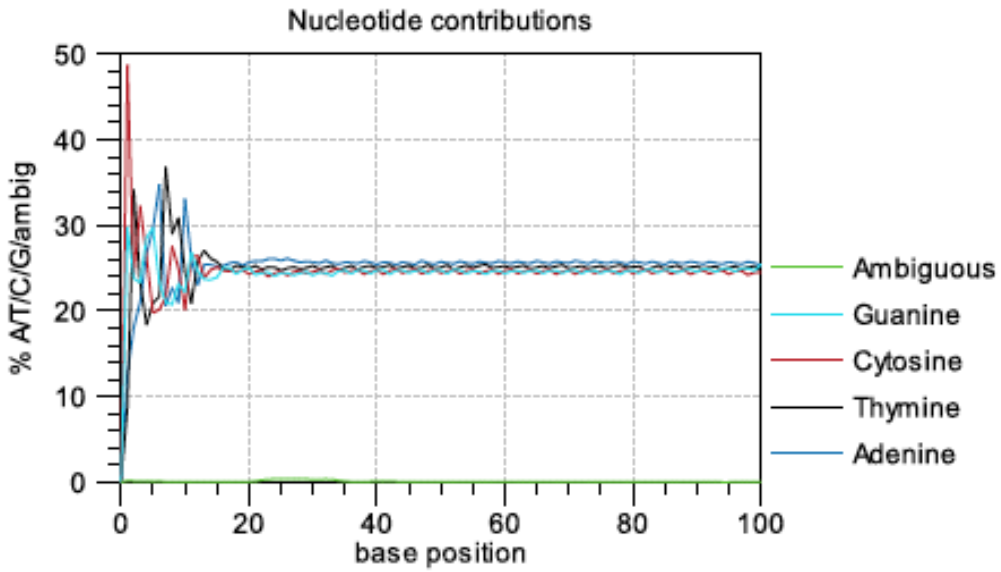
3. Per-base analysis

3.1 Coverage



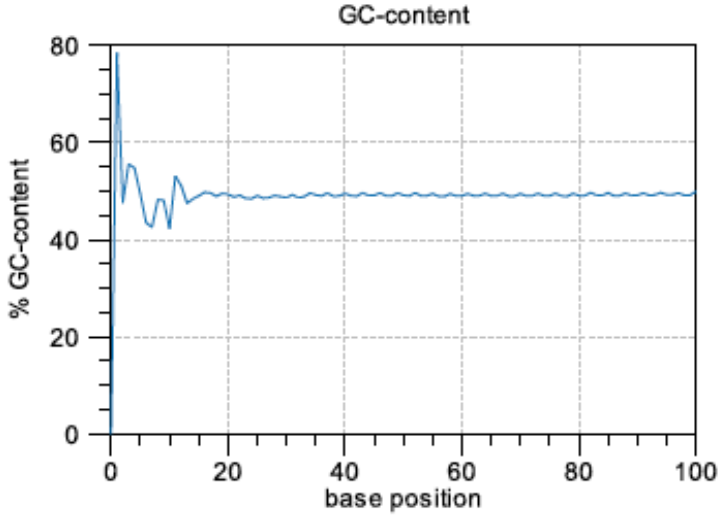
The number of sequences that support (cover) the individual base positions. In cases of un trimmed Illumina or SOLiD reads it will just contain a rectangle.
x: base position
y: number of sequences covering individual base positions normalized to the total number o f sequences

3.2 Nucleotide contributions



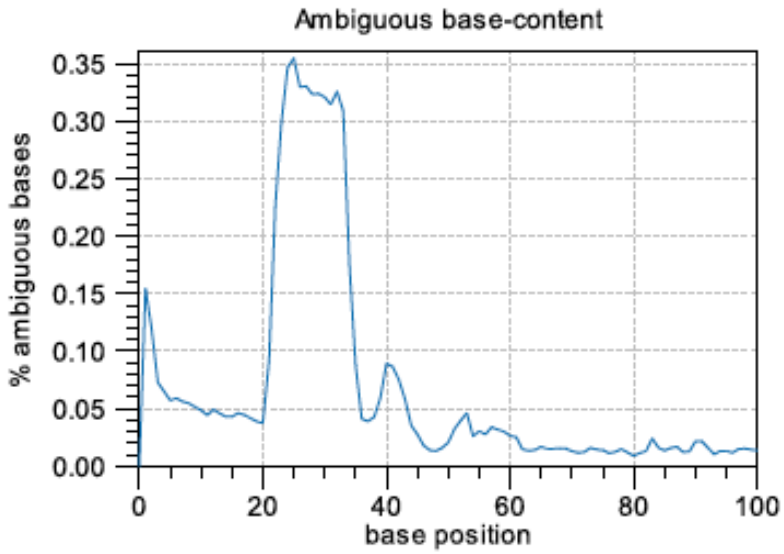
Coverages for the four DNA nucleotides and ambiguous bases.
x: base position
y: number of nucleotides observed per type normalized to the total number of nucleotides o bserved at that position

3.3 GC-content



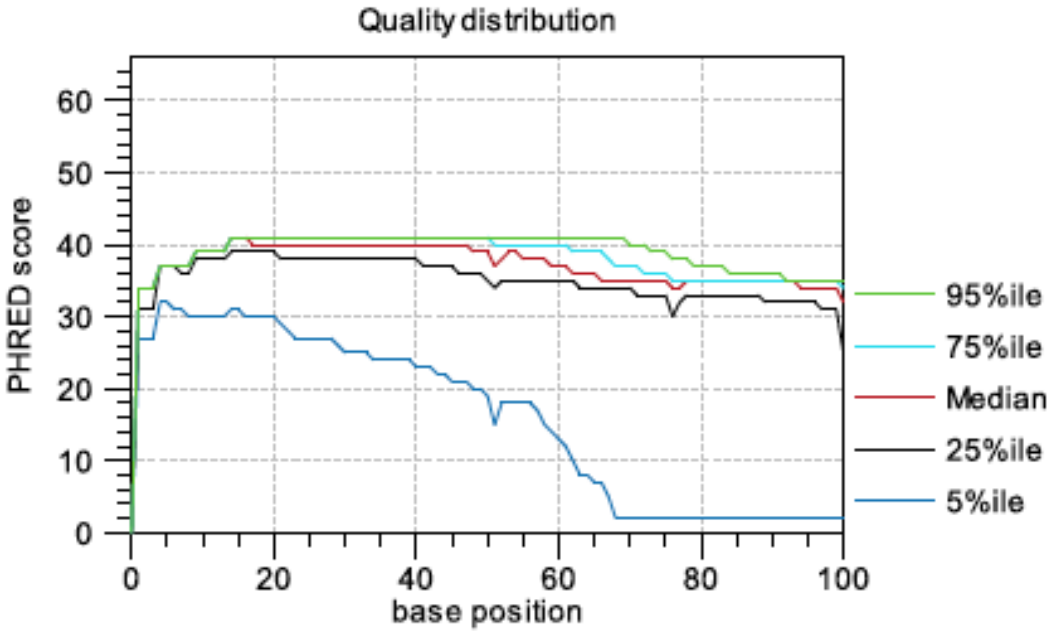
Combined coverage of G- and C-bases.
x: base position
y: number of G- and C-bases observed at current position normalized to the total number of bases observed at that position

3.4 Ambiguous base-content



Combined coverage of ambiguous bases.
x: base position
y: number of ambiguous bases observed at current position normalized to the total number o f bases observed at that position

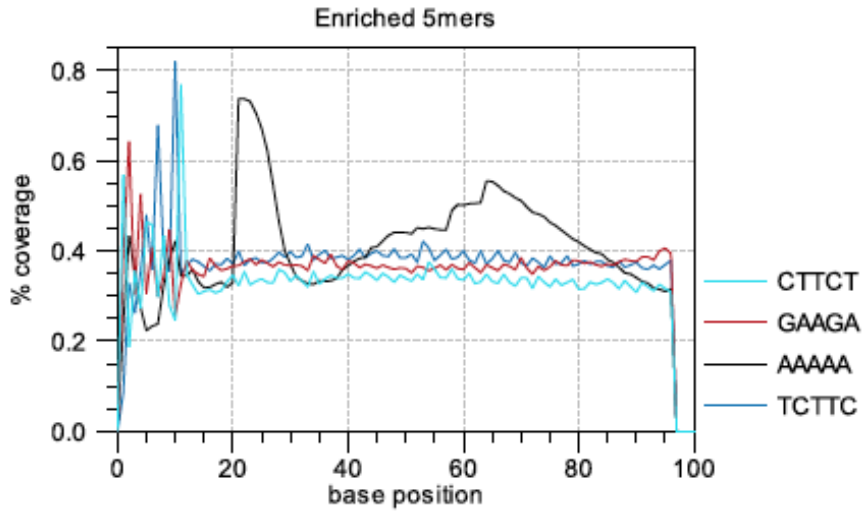
3.5 Quality distribution



Base-quality distribution along the base positions.
x: base position
y: median & percentiles of quality scores observed at that base position

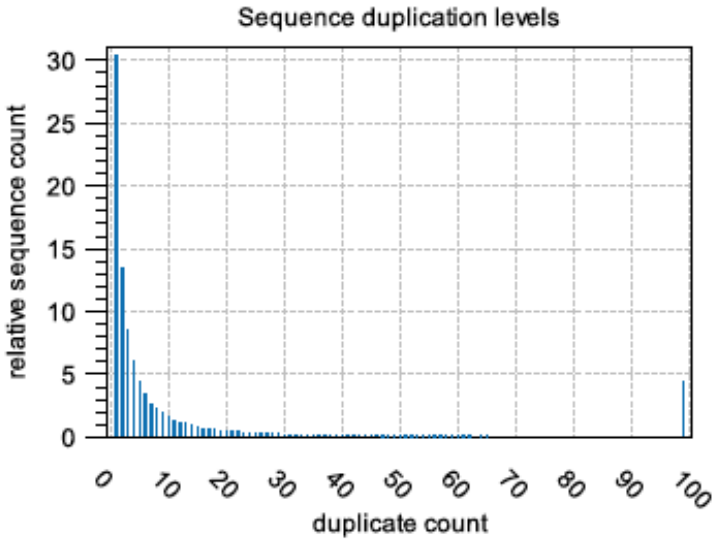
4. Over-representation analyses

4.1 Enriched 5mers



The five most-overrepresented 5mers. The over-representation of a 5mer is calculated as th e ratio of the observed and expected 5mer frequency. The expected frequency is calculated as product of the empirical nucleotide probabilities that make up the 5mer. (5mers that contain ambiguous bases are ignored)
x: base position
y: number of times a 5mer has been observed normalized to all 5mers observed at that posit ion

4.2 Sequence duplication levels



Duplication level distribution. Duplication levels are simply the count of how often a par ticular sequence has been found.
x: duplicate count
y: number of sequences that have been found that many times normalized to the number of un ique sequences

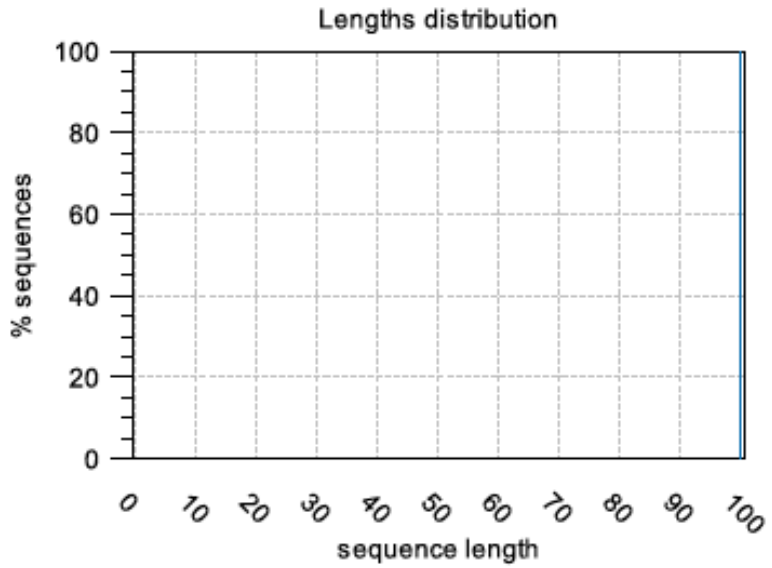
Appendix 4.3. *Polanisia dodecandra* FastQC report.

1. Summary

Creation date:	Thu Mar 10 15:39:51 MST 2016
Generated by:	Deyholos
Software:	CLC Genomics Workbench 7.5
Based upon:	1 data set
PD_R1 (paired):	84,853,538 sequences in pairs
Total nucleotides in data set	8,485,353,800 nucleotides

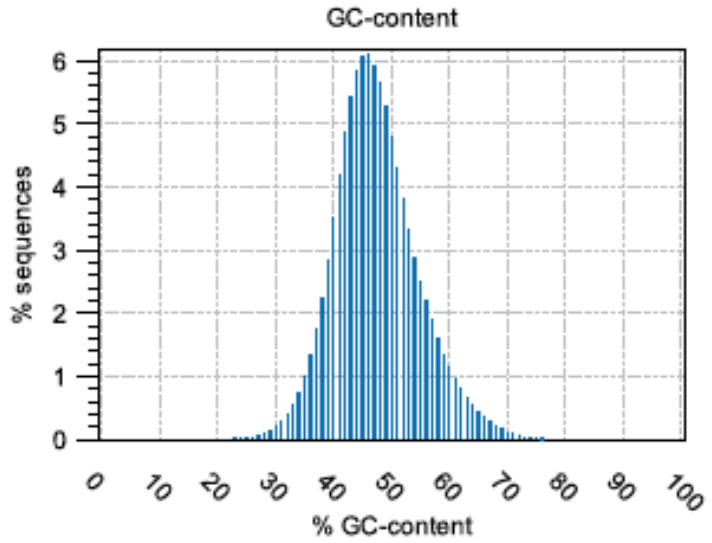
2. Per-sequence analysis

2.1 Lengths distribution



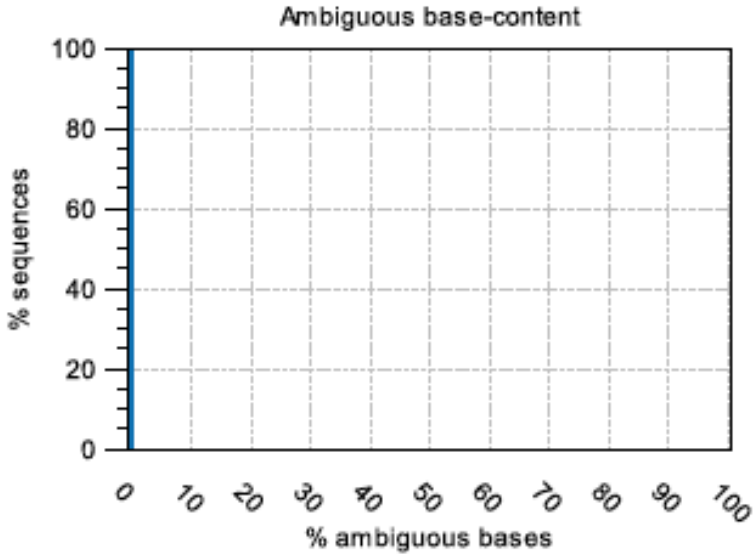
Distribution of sequence lengths. In cases of untrimmed Illumina or SOLiD reads it will ju st contain a single peak.
x: sequence length in base-pairs
y: number of sequences featuring a particular length normalized to the total number of seq uences

2.2 GC-content



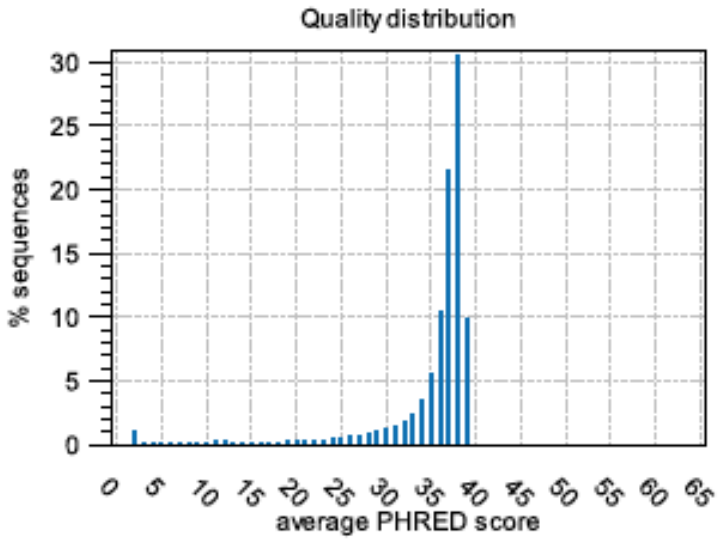
Distribution of GC-contents. The GC-content of a sequence is calculated as the number of G C-bases compared to all bases (including ambiguous bases).
x: relative GC-content of a sequence in percent
y: number of sequences featuring particular GC-percentages normalized to the total number of sequences

2.3 Ambiguous base-content



Distribution of N-contents. The N-content of a sequence is calculated as the number of amb iguous bases compared to all bases.
x: relative N-content of a sequence in percent
y: number of sequences featuring particular N-percentages normalized to the total number o f sequences

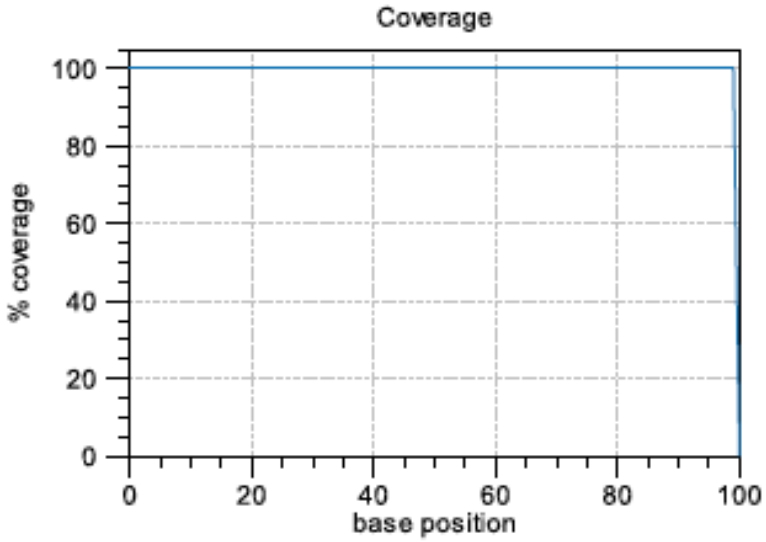
2.4 Quality distribution



Distribution of average sequence qualitie scores. The quality of a sequence is calculated as the arithmetic mean of its base qualities.
x: PHRED-score
y: number of sequences observed at that qual. score normalized to the total number of sequ ences

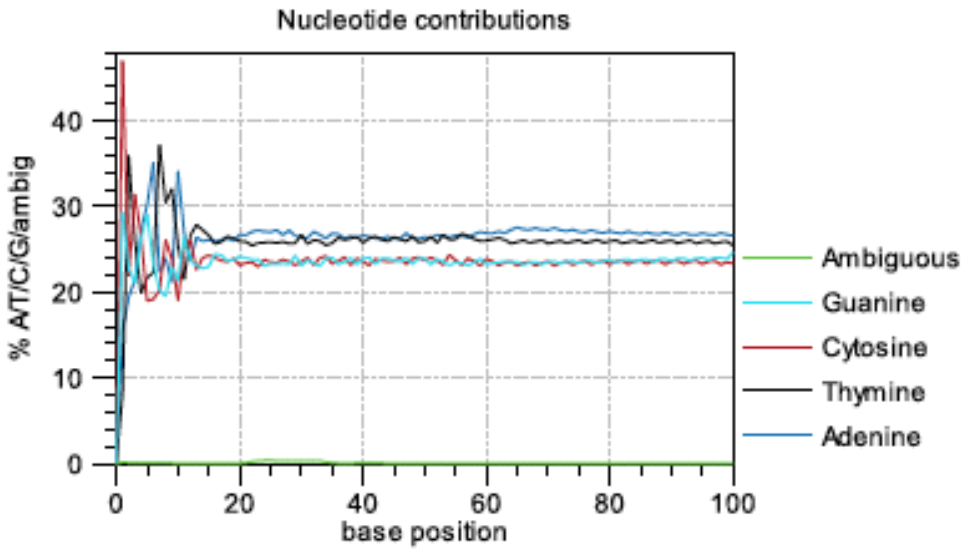
3. Per-base analysis

3.1 Coverage



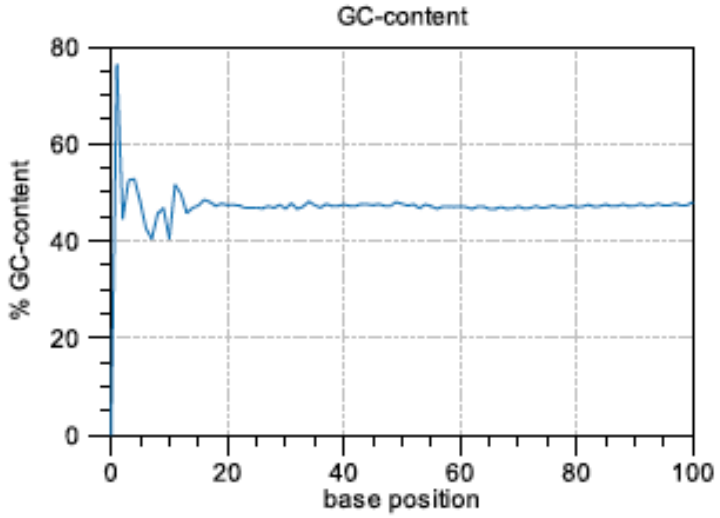
The number of sequences that support (cover) the individual base positions. In cases of un trimmed Illumina or SOLiD reads it will just contain a rectangle.
x: base position
y: number of sequences covering individual base positions normalized to the total number o f sequences

3.2 Nucleotide contributions



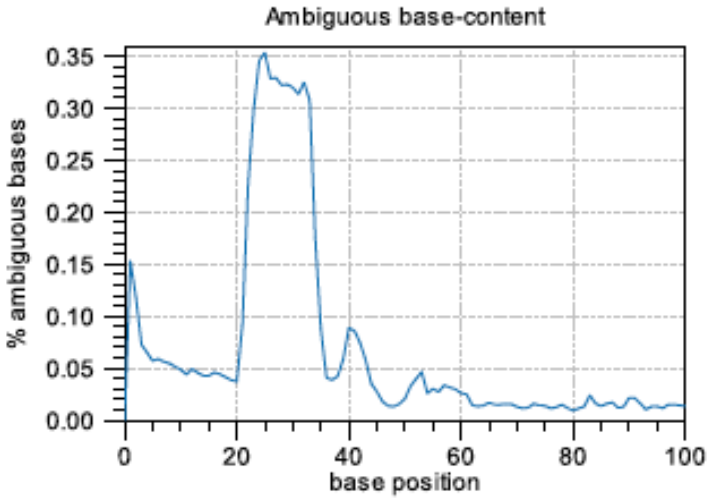
Coverages for the four DNA nucleotides and ambiguous bases.
x: base position
y: number of nucleotides observed per type normalized to the total number of nucleotides o bserved at that position

3.3 GC-content



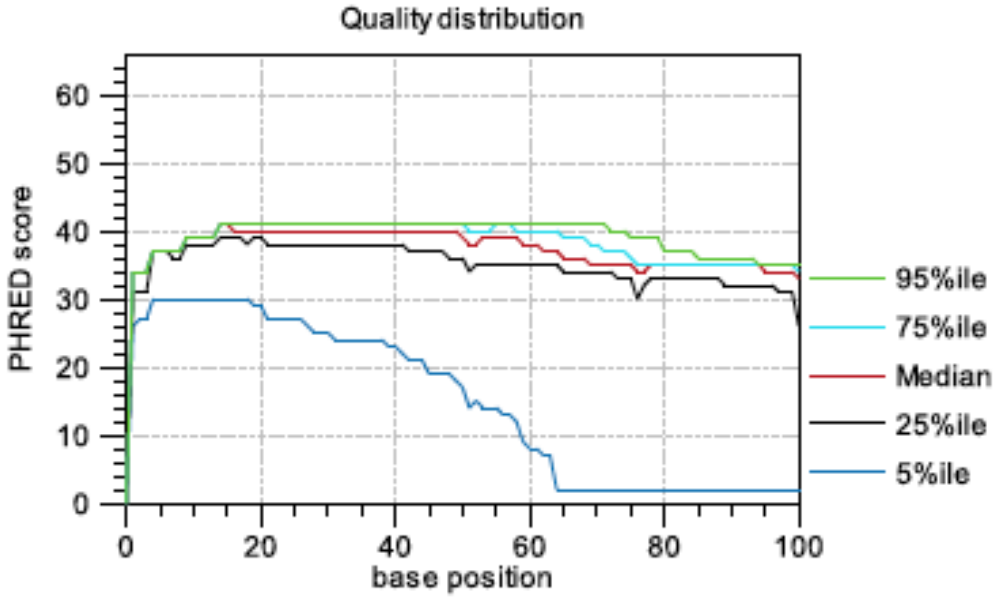
Combined coverage of G- and C-bases.
x: base position
y: number of G- and C-bases observed at current position normalized to the total number of bases observed at that position

3.4 Ambiguous base-content



Combined coverage of ambiguous bases.
x: base position
y: number of ambiguous bases observed at current position normalized to the total number o f bases observed at that position

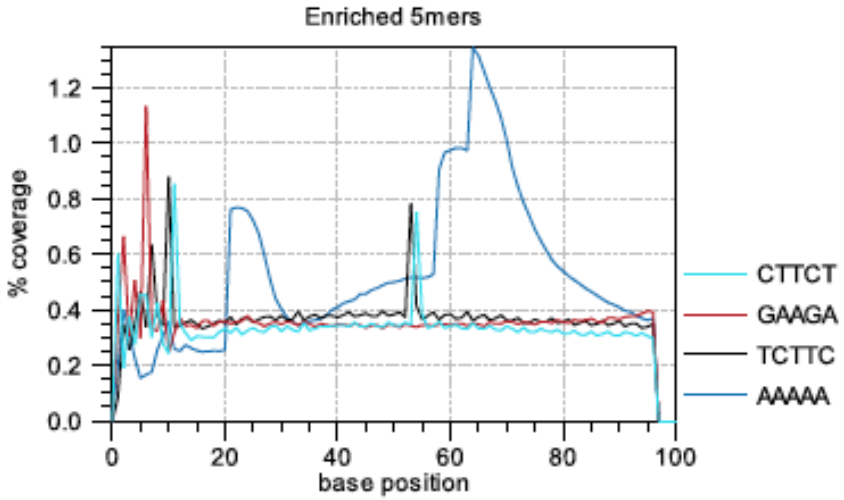
3.5 Quality distribution



Base-quality distribution along the base positions.
x: base position
y: median & percentiles of quality scores observed at that base position

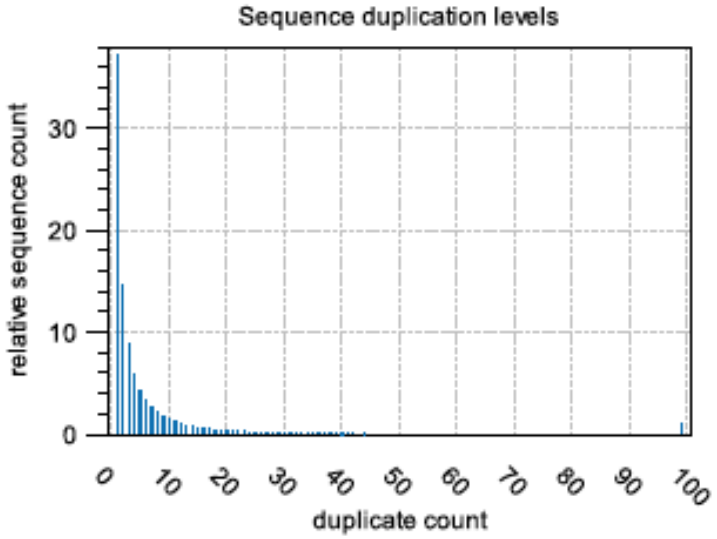
4. Over-representation analyses

4.1 Enriched 5mers



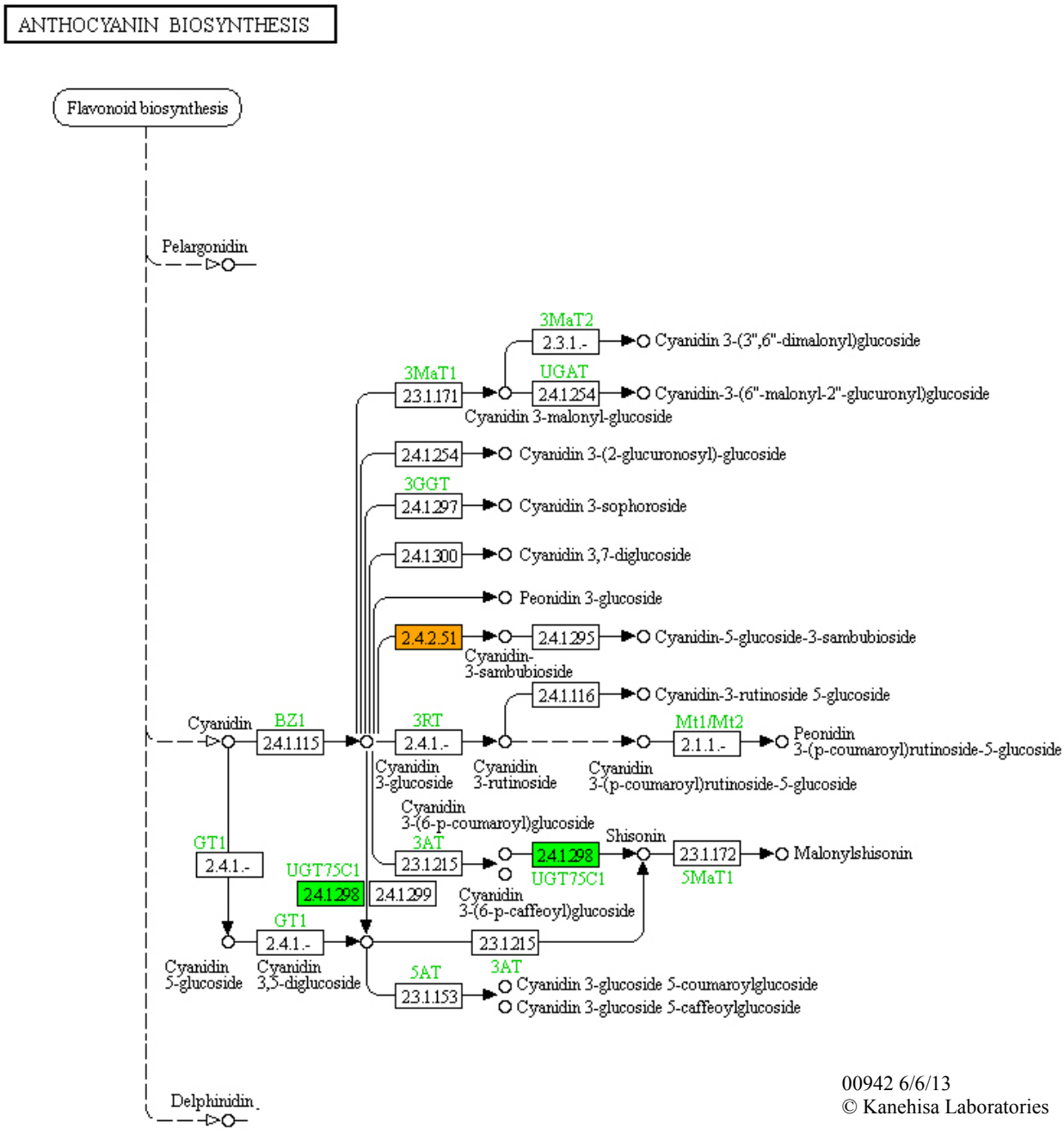
The five most-overrepresented 5mers. The over-representation of a 5mer is calculated as th e ratio of the observed and expected 5mer frequency. The expected frequency is calculated as product of the empirical nucleotide probabilities that make up the 5mer. (5mers that contain ambiguous bases are ignored)
x: base position
y: number of times a 5mer has been observed normalized to all 5mers observed at that posit ion

4.2 Sequence duplication levels

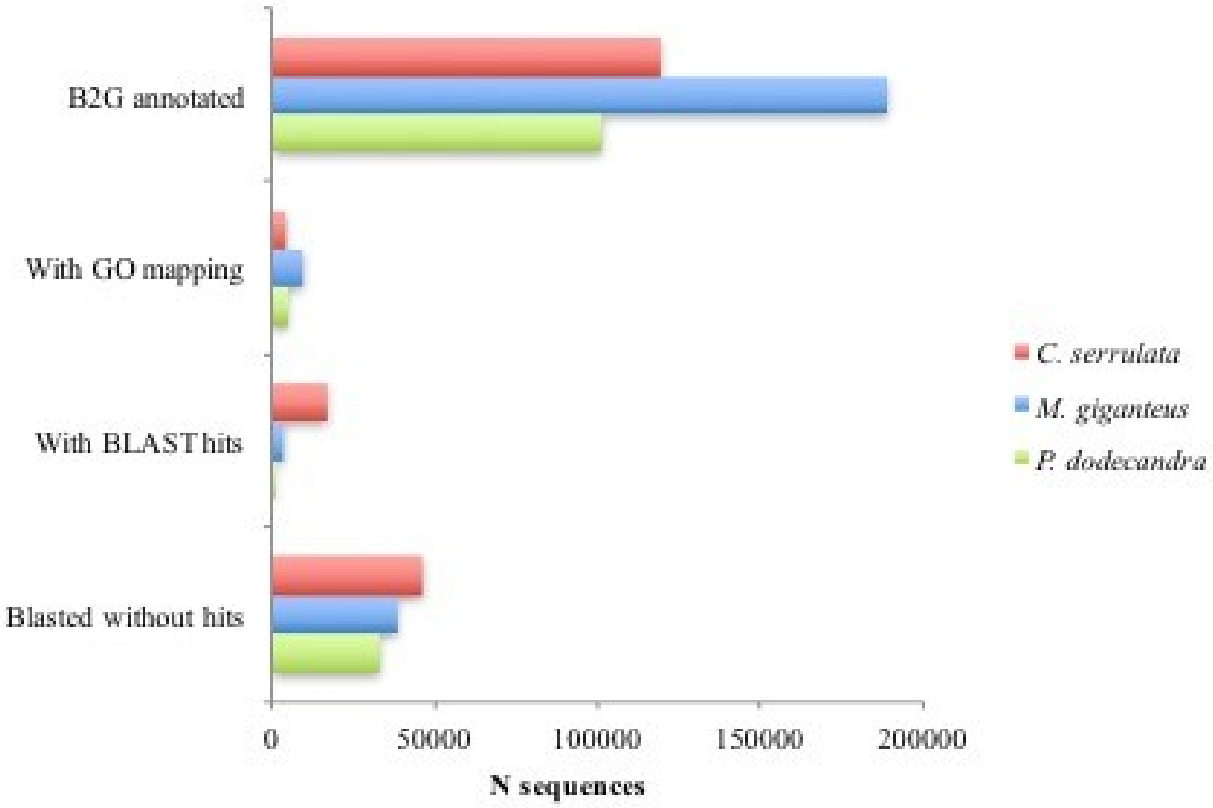


Duplication level distribution. Duplication levels are simply the count of how often a par ticular sequence has been found.
x: duplicate count
y: number of sequences that have been found that many times normalized to the number of un ique sequences

Appendix 4.4. Modified KEGG anthocyanin pathway depicting cyanidin derivatives present in *Cleomella serrulata, Melidiscus giganteus* and *Polanisia dodecandra*. In orange anthocyanidin 3-O-glucoside 2'''-O-xylosyltransferase. In green anthocyanidin 3-O-glucoside 5-O-glucosyltransferase.



Appendix 4.5. Blast2GO report statistics for the complete Trinity transcriptome of *C. serrulata, Melidiscus giganteus* and *Polanisia dodecandra*.



Appendix. 4.6. Pollination and floral development GO terms and their related genes in *Cleomella serrulata, Melidiscus giganteus*, and *Polanisia dodecandra.* Species code: *Cleomella serrulata* (CS), *Melidiscus giganteus* (MG), and *P. dodecandra* (PD).

GO term	Genes	CS	MG	PD
Nectary development	AFO (ABNORMAL		X	X
GO:0010254	BOP2 (BLADE ON PETIOLE2)		X	X
	CRC (CRABS CLAW)		X	X
	INO (INNER NO OUTER)	X		
	YAB2 (YABBY2)	X	X	X
	YAB5 (YABBY5)	X		
Specification of symmetry	AS2 (ASYMMETRIC LEAVES 2)			
GO:0009799	LOB (LATERAL ORGAN BOUNDARIES)	X	X	
	LBD2 (LOB DOMAIN-CONTAINING PROTEIN 2)			X
	LBD20 (LOB DOMAIN-CONTAINING PROTEIN 20)		X	
	LBD24 (LOB DOMAIN-CONTAINING PROTEIN 24)		X	
	LBD27 (LOB DOMAIN-CONTAINING PROTEIN 27)	X	X	X
	LBD37 (LOB DOMAIN-CONTAINING PROTEIN 37)	X		
Nectar secretion	ATBETAFRUCT4	X	X	X
GO:0071836	ATBFRUCT1	X	X	X
	AtcwINV2	X	X	X
	AtcwINV4	X	X	
	AtcwINV5		X	
	AtcwINV6	X	X	
	ATSPS1F (sucrose phosphate synthase 1F)		X	
	ATSPS2F (SUCROSE PHOSPHATE SYNTHASE 2F)		X	
	ATSPS4F	X	X	

	AtVEX1 (VEGETATIVE CELL EXPRESSED1)		X	
	BFRUCT3 (beta-fructosidase)		X	
	geranyl diphosphate synthase, putative	X		
	MTN3		X	
	nodulin MtN3 family protein	X	X	X
	SAG29 (SENESCENCE-ASSOCIATED PROTEIN 29)		X	X
	SPS1 (solanesyl diphosphate synthase 1)		X	
	SPS2 (Solanesyl diphosphate synthase 2)	X		
	SUS3 (sucrose synthase 3)		X	
Establishment of petal orientation GO:0048498	AT-GTL1 (GT2-LIKE 1)	X		X
	GT2		X	X
	trihelix DNA-binding protein, putative	X		X
	PTL (PETAL LOSS)	X	X	X
Regulation of anthocyanin metabolic process GO:0031537	AHK2 (ARABIDOPSIS HISTIDINE KINASE 2)	X	X	
	AHK3 (ARABIDOPSIS HISTIDINE KINASE)	X	X	X
	AHP1 (HISTIDINE-CONTAINING PHOSPHOTRANSMITTER 1)	X	X	X
	AHP2 (HISTIDINE-CONTAINING PHOSPHOTRANSMITTER 2)		X	X
	AHP3 (HISTIDINE-CONTAINING PHOSPHOTRANSMITTER 3)	X		X
	AHP4 (HPT PHOSPHOTRANSMITTER 4)	X	X	X
	AHP5 (HISTIDINE-CONTAINING PHOSPHOTRANSFER FACTOR 5)			X
	AHP6 (ARABIDOPSIS HISTIDINE PHOSPHOTRANSFER PROTEIN 6)			X

APRR2		X	X
APRR3 (ARABIDOPSIS PSEUDO-RESPONSE REGULATOR 3)			X
APRR5 (ARABIDOPSIS PSEUDO-RESPONSE REGULATOR 5)	X	X	X
APRR8 (PSEUDO-RESPONSE REGULATOR 8)	X		
APRR9 (ARABIDOPSIS PSEUDO-RESPONSE REGULATOR 9)	X	X	X
ARR1 (ARABIDOPSIS RESPONSE REGULATOR 1)	X	X	X
ARR2 (ARABIDOPSIS RESPONSE REGULATOR 2)	X		
ARR4 (RESPONSE REGULATOR 4)		X	
ARR5 (ARABIDOPSIS RESPONSE REGULATOR 5)		X	
ARR7 (RESPONSE REGULATOR 7)		X	X
ARR9 (RESPONSE REGULATOR 9)	X	X	
ARR10 (ARABIDOPSIS RESPONSE REGULATOR 10)	X	X	
ARR11 (RESPONSE REGULATOR 11)		X	X
ARR12 (ARABIDOPSIS RESPONSE REGULATOR 12)	X	X	X
ARR14 (ARABIDOPSIS RESPONSE REGULATOR 14)		X	
ARR16 (ARABIDOPSIS RESPONSE REGULATOR 16)	X		
ARR17 (ARABIDOPSIS RESPONSE REGULATOR 17		X	
ARR18 (ARABIDOPSIS RESPONSE REGULATOR 18)			X
ARR20 (ARABIDOPSIS RESPONSE REGULATOR 20)	X		
ARR21 (ARABIDOPSIS RESPONSE REGULATOR 21)	X		
armadillo/beta-catenin repeat family protein		X	
ATPDIL1-1 (PDI-LIKE 1-1)			X
ATRR3 (RESPONSE REGULATOR 3)		X	X
DNA repair protein, putative			X
ECT10	X		
GLK2 (GOLDEN2-LIKE 2	X	X	X

	GPRI1 (GBF'S PRO-RICH REGION-INTERACTING FACTOR 1)	X	X	
	myb family transcription factor	X	X	
	PCL1 (PHYTOCLOCK 1)		X	X
	PRR7 (PSEUDO-RESPONSE REGULATOR 7)		X	
	protein kinase family protein		X	
	TOC1 (TIMING OF CAB EXPRESSION 1)		X	
	WOL (WOODEN LEG)	X		X
Proximal/distal pattern formation GO:0009954	acetyl-CoA C-acyltransferase, putative / 3-ketoacyl-CoA thiolase, putative		X	
	ankyrin repeat family protein	X	X	X
	AS2 (ASYMMETRIC LEAVES 2)	X	X	
	ASL1	X	X	X
	ASL5		X	X
	ASL9 (ASYMMETRIC LEAVES 2 LIKE 9)		X	
	ATPOB1	X	X	X
	BOP2 (BLADE ON PETIOLE2)		X	X
	BT3 (BTB AND TAZ DOMAIN PROTEIN 3)	X	X	X
	BT4 (BTB AND TAZ DOMAIN PROTEIN 4)		X	X
	BT5 (BTB AND TAZ DOMAIN PROTEIN 5)	X	X	X
	BTB/POZ domain-containing protein	X	X	X
	glycosyl hydrolase family 5 protein / cellulase family protein		X	
	LBD2 (LOB DOMAIN-CONTAINING PROTEIN 2)	X		X
	LBD4 (LOB DOMAIN-CONTAINING PROTEIN 4)	X	X	X
	LBD10 (LOB DOMAIN-CONTAINING PROTEIN 10)	X		X
	LBD11 (LOB DOMAIN-CONTAINING PROTEIN 11)	X		X
	LBD13 (LOB DOMAIN-CONTAINING PROTEIN 13)	X		X
	LBD15 (LOB DOMAIN-CONTAINING PROTEIN 15)	X	X	X
	LBD16 (LATERAL ORGAN BOUNDARIES-DOMAIN 16)	X		

	LBD19 (LOB DOMAIN-CONTAINING PROTEIN 19)			
	LBD20 (LOB DOMAIN-CONTAINING PROTEIN 20)		X	X
	LBD21 (LOB DOMAIN-CONTAINING PROTEIN 21)		X	X
	LBD22 (LOB DOMAIN-CONTAINING PROTEIN 22)	X		X
	LBD24 (LOB DOMAIN-CONTAINING PROTEIN 24)		X	X
	LBD25 (LOB DOMAIN-CONTAINING PROTEIN 25)	X	X	X
	LBD27 (LOB DOMAIN-CONTAINING PROTEIN 27)	X	X	X
	LBD31 (LOB DOMAIN-CONTAINING PROTEIN 31)			X
	LBD33 (LOB DOMAIN-CONTAINING PROTEIN 33)	X		
	LOB (LATERAL ORGAN BOUNDARIES)	X	X	
	phototropic-responsive NPH3 family protein		X	X
	RPT2 (ROOT PHOTOTROPISM 2)		X	X
	speckle-type POZ protein-related	X	X	
	TraB protein-related			X
Aromatic compound biosynthetic process GO:0019438	1-aminocyclopropane-1-carboxylate oxidase, putative / ACC oxidase, putative	X		
	29 kDa ribonucleoprotein, chloroplast, putative		X	X
	4-coumarate-CoA ligase		X	
	ACA8 (AUTOINHIBITED CA2+ -ATPASE, ISOFORM 8)		X	
	ACS10 (ACC SYNTHASE 10)			X
	AGT2 (ALANINE: GLYOXYLATE) AMINOTRANSFERASE 2)		X	
	alanine--glyoxylate aminotransferase, putative		X	
	alcohol dehydrogenase, putative			X
	aldo/keto reductase family protein	X		
	amine oxidase/ copper ion binding / quinone binding	X	X	
	aminotransferase class I and II family protein	X		
	AMP-dependent synthetase and ligase family protein			X
	APUM3 (Arabidopsis Pumilio 3)		X	

armadillo/beta-catenin repeat family protein		X	
AT3BETAHSD/D1 (3BETA-HYDROXYSTEROID-DEHYDROGENASE/DECARBOXYLASE ISOFORM 1)		X	
ATAO1 (ARABIDOPSIS THALIANA AMINE OXIDASE 1)		X	X
AtFAAH (Arabidopsis thaliana fatty acid amide hydrolase)	X	X	X
ATFP3		X	
ATGSTF11 (GLUTATHIONE S-TRANSFERASE F11)		X	
ATGSTF8 (ARABIDOPSIS THALIANA GLUTATHIONE S-TRANSFERASE PHI 8)		X	X
ATGSTT1 (GLUTATHIONE S-TRANSFERASE THETA 1)	X		
ATL2	X		
ATLIP1 (Arabidopsis thaliana lipase 1)	X		
ATMPK4 (ARABIDOPSIS THALIANA MAP KINASE 4)	X		
ATNFXL1 (ARABIDOPSIS THALIANA NF-X-LIKE 1)	X		
ATMYB21 (ARABIDOPSIS THALIANA MYB DOMAIN PROTEIN 21)		X	
ATP-dependent Clp protease proteolytic subunit		X	X
ATP synthase gamma chain, mitochondrial (ATPC)	X		
ATPER1	X		
AtPPa6 (Arabidopsis thaliana pyrophosphorylase 6)	X		
ATPRX Q	X		
ATRFNR1 (ROOT FNR 1)			
BCDH BETA1 (BRANCHED-CHAIN ALPHA-KETO ACID DECARBOXYLASE E1 BETA SUBUNIT)	X		
BEN1		X	
BGLU31 (BETA GLUCOSIDASE 31)		X	
BGLU33 (BETA GLUCOSIDASE 33)	X		X
BGLU40 (BETA GLUCOSIDASE 40)	X		
binding / catalytic/ coenzyme binding		X	

binding / ubiquitin-protein ligase	X		
BIP1		X	
BIP2		X	
BIP3		X	
CAC2	X	X	
calmodulin-binding family protein		X	
cinnamoyl-CoA reductase-related		X	
cinnamyl-alcohol dehydrogenase family / CAD family		X	
cinnamyl-alcohol dehydrogenase, putative (CAD)		X	
CLPP3			X
copine-related			X
copper amine oxidase, putative	X	X	X
CP33	X		
CPK1 (CALCIUM DEPENDENT PROTEIN KINASE 1)	X		
CYP710A1 (cytochrome P450, family 710, subfamily A, polypeptide 1)			X
CYP706A3	X		
CYP711A1		X	
CYP94B2			X
CYP94B3	X		X
CYSD1 (CYSTEINE SYNTHASE D1)		X	
delta-OAT	X		
DFL2 (DWARF IN LIGHT 2)	X		
disease resistance protein (NBS-LRR class), putative			X
DJ-1 family protein			X
ECA1 (ER-TYPE CA2+-ATPASE 1)	X		
ECA3 (ENDOPLASMIC RETICULUM-TYPE CALCIUM-TRANSPORTING ATPASE 3)	X		
ELI3-2 (ELICITOR-ACTIVATED GENE 3-2)	X		

EDA39 (embryo sac development arrest 39)		X	
elongation factor 1-alpha / EF-1-alpha		X	
elongation factor family protein	X		
emb2444 (embryo defective 2444)			X
emb2726 (embryo defective 2726)		X	
EMB3009 (embryo defective 3009)		X	
eukaryotic translation initiation factor 2 family protein / eIF-2 family protein	X		
ferredoxin hydrogenase	X		
FRS7 (FAR1-related sequence 7)		X	
FRS8 (FAR1-related sequence 8)	X		X
fructose-1,6-bisphosphatase family protein	X	X	
fructose-1,6-bisphosphatase, putative	X	X	X
GA20OX1	X		
gibberellin 20-oxidase-related	X		
glycine-rich RNA-binding protein, putative	X		
heat shock protein 70, putative / HSP70, putative	X		
heat shock cognate 70 kDa protein 2 (HSC70-2) (HSP70-2)			X
HEME1	X		
heavy-metal-associated domain-containing protein		X	
HSP91	X		
IAA2 (INDOLE-3-ACETIC ACID INDUCIBLE 2)		X	
IAA13		X	
immunophilin, putative / FKBP-type peptidyl-prolyl cis-trans isomerase, putative	X		
invertase/pectin methylesterase inhibitor family protein		X	
KAB1 (POTASSIUM CHANNEL BETA SUBUNIT)	X	X	X
L-ascorbate oxidase			X
L-galactose dehydrogenase (L-GalDH)		X	

LACS6 (long-chain acyl-CoA synthetase 6)	X	X	X
leucine-rich repeat protein kinase, putative	X		
light repressible receptor protein kinase		X	
MAB1 (MACCI-BOU)	X	X	X
MAP3KA		X	
MCCA	X		
MYB3 (MYB DOMAIN PROTEIN 3)		X	
NAD(P)H dehydrogenase complex assembly		X	
NCLPP7 (NUCLEAR-ENCODED CLP PROTEASE P7)	X	X	
NLP1 (NITRILASE-LIKE PROTEIN 1)			X
OASC (O-ACETYLSERINE (THIOL) LYASE ISOFORM C)	X		
OBP2		X	
oxidoreductase, 2OG-Fe(II) oxygenase family protein		X	
oxidoreductase, zinc-binding dehydrogenase family protein	X	X	
PAB4 (POLY(A) BINDING PROTEIN 4)	X		
PCB2 (PALE-GREEN AND CHLOROPHYLL B REDUCED 2)		X	
PDR10 (PLEIOTROPIC DRUG RESISTANCE 10)		X	X
pentatricopeptide (PPR) repeat-containing protein		X	
peptide chain release factor, putative	X	X	
peptidyl-prolyl cis-trans isomerase cyclophilin-type family protein		X	
peptidyl-prolyl cis-trans isomerase, chloroplast		X	
peptidyl-prolyl cis-trans isomerase, putative		X	X
peroxidase 27 (PER27) (P27) (PRXR7)		X	
phosphoribulokinase/uridine kinase family protein	X		
PIL5 (PHYTOCHROME INTERACTING FACTOR 3-LIKE 5)		X	
PME1 (PECTIN METHYLESTERASE INHIBITOR 1)		X	
POP2 (POLLEN-PISTIL INCOMPATIBILITY 2)	X	X	X

protein kinase family protein	X		
PSBO2 (PHOTOSYSTEM II SUBUNIT O-2)		X	
PTEN1		X	
PUB9 (PLANT U-BOX 9)	X		
pyridoxal-5'-phosphate-dependent enzyme, beta family protein	X		
RAN1 (RESPONSIVE-TO-ANTAGONIST 1)	X		
ribose 5-phosphate isomerase-related	X		
ribose-phosphate pyrophosphokinase, putative			X
ribosomal protein L1 family protein		X	
RNA binding / nucleic acid binding / nucleotide binding	X		
RNA recognition motif (RRM)-containing protein	X	X	X
RNA-binding protein 45 (RBP45), putative		X	X
scarecrow-like transcription factor 11 (SCL11)		X	
SHM6 (serine hydroxymethyltransferase 6)	X		
SHM7 (serine hydroxymethyltransferase 7)	X		
short-chain dehydrogenase		X	
sks4 (SKU5 Similar 4)		X	
sks5 (SKU5 Similar 5)		X	X
SKS6 (SKU5-SIMILAR 6)		X	X
sks15 (SKU5 Similar 15)			X
sks17 (SKU5 Similar 17)			X
sorbitol dehydrogenase, putative / L-iditol 2-dehydrogenase, putative	X		
SOT16 (SULFOTRANSFERASE 16)		X	
splicing factor-related		X	
ST2A (SULFOTRANSFERASE 2A)			X
sulfotransferase family protein	X	X	X
tatD-related deoxyribonuclease family protein	X		

	terpene cyclase/mutase-related		X	
	TIC55 (TRANSLOCON AT THE INNER ENVELOPE MEMBRANE OF CHLOROPLASTS 55)		X	X
	TOC33 (TRANSLOCON AT THE OUTER ENVELOPE MEMBRANE OF CHLOROPLASTS 33)			X
	TPA_exp: actin-related protein 2	X		
	TPA_exp: actin-related protein 3	X		
	TPA_exp: PDR3 ABC transporter			X
	TPA_exp: PDR4 ABC transporter		X	
	transferase family protein		X	
	UDP-glucoronosyl/UDP-glucosyl transferase family protein		X	
	vestitone reductase-related		X	
	WD-40 repeat family protein	X		X
	WRKY18		X	
	WRKY30		X	
	WRKY38		X	
	WRKY54		X	
	XLG1 (EXTRA-LARGE G-PROTEIN 1)			X
	zinc finger (B-box type) family protein			X
	zinc finger (C3HC4-type RING finger) family protein		X	X
Aromatic amino acid family metabolic process GO:0009072	ABC transporter family protein	X	X	X
	ACS10 (ACC SYNTHASE 10)			X
	acyl-(acyl-carrier-protein) desaturase, putative / stearoyl-ACP desaturase, putative	X	X	X
	AGD2 (ABERRANT GROWTH AND DEATH 2); L,L-diaminopimelate aminotransferase/ transaminase			X
	aldo/keto reductase family protein	X		
	aminotransferase class I and II family protein	X		

antiporter/ drug transporter		X	
ATGSTF11 (GLUTATHIONE S-TRANSFERASE F11)		X	
ATGSTF12 (ARABIDOPSIS THALIANA GLUTATHIONE S-TRANSFERASE PHI 12)		X	
ATGSTF8 (ARABIDOPSIS THALIANA GLUTATHIONE S-TRANSFERASE PHI 8)		X	X
ATGSTZ1 (ARABIDOPSIS THALIANA GLUTATHIONE S-TRANSFERASE ZETA 1)		X	
ATLIP1 (Arabidopsis thaliana lipase 1)	X		
ATNAP9	X	X	
ATPC1	X		
ATPC2	X	X	X
ATP synthase gamma chain, mitochondrial (ATPC)	X		
Cul3-RING ubiquitin ligase complex	X		
BLH5 (BELL1-like homeodomain 5)		X	
CPZ			X
Cul4-RING E3 ubiquitin ligase complex			X
CYP77B1		X	
CYSD1 (CYSTEINE SYNTHASE D1)	X		
CYP82F1	X		
cytochrome c oxidase assembly protein CtaG / Cox11 family	X		
DJ-1 family protein			X
DXR (1-DEOXY-D-XYLULOSE 5-PHOSPHATE REDUCTOISOMERASE)	X	X	X
EDS5 (ENHANCED DISEASE SUSCEPTIBILITY 5)	X	X	X
EMB3003 (embryo defective 3003)	X	X	
FAB1 (FATTY ACID BIOSYNTHESIS 1)	X	X	X
FAD6 (FATTY ACID DESATURASE 6)		X	
fructose-bisphosphate aldolase, putative	X	X	

GLB1 (GLNB1 HOMOLOG)			X
HDS (4-HYDROXY-3-METHYLBUT-2-ENYL DIPHOSPHATE SYNTHASE)			X
HFR1 (LONG HYPOCOTYL IN FAR-RED)		X	
HPT1 (HOMOGENTISATE PHYTYLTRANSFERASE 1)	X	X	X
hydroxyproline-rich glycoprotein family protein		X	X
ISPD	X	X	X
KAS I (3-KETOACYL-ACYL CARRIER PROTEIN SYNTHASE I)	X	X	X
KCS18	X		
KCS19 (3-KETOACYL-COA SYNTHASE 19)		X	
KCS21 (3-KETOACYL-COA SYNTHASE 21)	X		
KCS5 (3-KETOACYL-COA SYNTHASE 5)	X		
KCS6 (3-KETOACYL-COA SYNTHASE 6)		X	
leucine-rich repeat family protein			X
LIP2 (LIPOYLTRANSFERASE 2)			X
lipoxygenase, putative	X		
LOX1	X		
LOX2 (LIPOXYGENASE 2)	X	X	X
LOX3		X	
LOX5		X	
LTA2		X	X
LTA3	X		
malate dehydrogenase (NAD), mitochondrial		X	X
NADH dehydrogenase subunit 4L	X		
NPQ1 (NON-PHOTOCHEMICAL QUENCHING 1)			X
OASC (O-ACETYLSERINE (THIOL) LYASE ISOFORM C)	X		
oxidoreductase, zinc-binding dehydrogenase family protein	X	X	
PDS1 (PHYTOENE DESATURATION 1)		X	X

	pentatricopeptide (PPR) repeat-containing protein			X
	peptidyl-prolyl cis-trans isomerase, putative / cyclophilin, putative / rotamase, putative		X	X
	PMDH2 (peroxisomal NAD-malate dehydrogenase 2)		X	
	potassium channel tetramerisation domain-containing protein	X		
	pyridoxal-5'-phosphate-dependent enzyme, beta family protein	X		
	ROC3			X
	SCPL3	X		
	scpl14 (serine carboxypeptidase-like 14)			X
	scpl6 (serine carboxypeptidase-like 6)	X		X
	SCPL19		X	X
	serine carboxypeptidase S10 family protein		X	X
	SNG1 (SINAPOYLGLUCOSE 1)		X	
	SSI2	X	X	
	TIL1 (TILTED 1)		X	X
Polarity specification of adaxial/abaxial axis GO:0009944	(1-4)-beta-mannan endohydrolase		X	X
	ABIL1 (Abi-1-like 1)			X
	ADOF1		X	
	AFH1 (FORMIN HOMOLOGY 1)		X	X
	AFO (ABNORMAL FLORAL ORGANS)	X	X	X
	AGD1 (ARF-GAP domain 1)	X	X	X
	AGD6	X		
	agenet domain-containing protein		X	
	AGO1 (ARGONAUTE 1)	X	X	
	AGO5 (ARGONAUTE 5)	X	X	X
	AIL5 (AINTEGUMENTA-LIKE 5)	X		X
	AIL6 (AINTEGUMENTA-LIKE 6)	X	X	X
	AIL7 (AINTEGUMENTA-like 7)	X	X	

ANT (AINTEGUMENTA)	X	X	X
AP2 domain-containing transcription factor		X	
AP2 (APETALA 2)			X
ARA4	X	X	
ARA-5 (ARABIDOPSIS RAS 5)	X		X
ARF6 (AUXIN RESPONSE FACTOR 6)	X	X	
ARF8 (AUXIN RESPONSE FACTOR 8)	X	X	
ARF17 (AUXIN RESPONSE FACTOR 17)	X		
ARF19 (AUXIN RESPONSE FACTOR 19)	X	X	X
AS1 (ASYMMETRIC LEAVES 1)		X	
AS2 (ASYMMETRIC LEAVES 2)	X	X	X
ASL1	X	X	
ASL5	X		X
ASL9 (ASYMMETRIC LEAVES 2 LIKE 9)	X		
aspartic-type endopeptidase/ peptidase	X	X	
aspartyl-tRNA synthetase		X	
aspartyl-tRNA synthetase, putative / aspartate--tRNA ligase, putative	X		
ATCEL2	X	X	X
ATCEL3 (ARABIDOPSIS THALIANA CELLULASE 3)		X	X
ATER	X	X	X
ATFP8		X	
AtGH9A4 (Arabidopsis thaliana Glycosyl Hydrolase 9A4)		X	
ATGH9B1 (ARABIDOPSIS THALIANA GLYCOSYL HYDROLASE 9B1)	X	X	
AtGH9B8 (Arabidopsis thaliana glycosyl hydrolase 9B8)		X	
AtGH9B13 (Arabidopsis thaliana glycosyl hydrolase 9B13)	X		X
AtGH9B15 (Arabidopsis thaliana glycosyl hydrolase 9B15)	X		X
AtGH9B17 (Arabidopsis thaliana glycosyl hydrolase 9B17)	X		

AtGH9B18 (Arabidopsis thaliana glycosyl hydrolase 9B18)	X		
AtGH9C3 (Arabidopsis thaliana glycosyl hydrolase 9C3)			X
AT-GTL1 (GT2-LIKE 1)	X		X
ATHB-15	X	X	
ATHB-8 (HOMEOBOX GENE 8)	X	X	
ATL5 (ATL5)			X
ATL8		X	
ATMYB21 (ARABIDOPSIS THALIANA MYB DOMAIN PROTEIN 21)	X	X	X
ATP binding / ATP-dependent DNA helicase/ DNA binding / hydrolase		X	
ATRAB1A	X		
ATRAB1C	X	X	
ATRAB8	X		X
ATRAB18 (ARABIDOPSIS RAB GTPASE HOMOLOG B18)	X		
ATRABA2B (ARABIDOPSIS RAB GTPASE HOMOLOG A2B)		X	
ATRABA3 (ARABIDOPSIS RAB GTPASE HOMOLOG A3)		X	
AtRABA5a (Arabidopsis Rab GTPase homolog A5a)	X	X	X
AtRABA5b (Arabidopsis Rab GTPase homolog A5b)	X	X	X
AtRABA5d (Arabidopsis Rab GTPase homolog A5d)		X	
ATRABH1D (ARABIDOPSIS RAB GTPASE HOMOLOG H1D)	X		
AtRLP44 (Receptor Like Protein 44)		X	
AtRLP45 (Receptor Like Protein 45)			X
ATRPS5B (RIBOSOMAL PROTEIN 5B)			X
ATS (ABERRANT TESTA SHAPE)	X	X	
BAM3 (BARELY ANY MERISTEM 3)		X	
BBM (BABY BOOM)	X		

BIM1	X		
BIM2 (BES1-interacting Myc-like protein 2)	X		
BIM3 (BES1-interacting Myc-like protein 3)			X
BLH1 (BEL1-LIKE HOMEODOMAIN 1)		X	
BLH5 (BELL1-like homeodomain 5)			X
BLH8 (BEL1-LIKE HOMEODOMAIN 8)		X	
BLH11 (BEL1-LIKE HOMEODOMAIN 11)	X		
BOP2 (BLADE ON PETIOLE2)	X	X	X
BRI1 (BRASSINOSTEROID INSENSITIVE 1)	X		
BST1 (BRISTLED 1)	X		X
C2 domain-containing protein	X	X	X
calcium-binding EF hand family protein	X	X	X
calcium-binding EF hand family protein, putative	X		X
CLAVATA1 receptor kinase (CLV1)	X		
COB (COBRA)	X	X	
COBL1 (COBRA-LIKE PROTEIN 1 PRECURSOR)	X	X	X
COBL2 (COBRA-LIKE PROTEIN 2 PRECURSOR)		X	X
COBL5 (COBRA-LIKE PROTEIN 5 PRECURSOR)	X	X	
CRC (CRABS CLAW)	X	X	X
CRN (CORYNE)			X
CVP2 (COTYLEDON VASCULAR PATTERN 2)	X	X	
DAG1 (dof affecting germination 1)	X		
DEAD/DEAH box helicase		X	
Dof-type zinc finger domain-containing protein	X	X	X
DRT100 (DNA-DAMAGE REPAIR			X
DUO1 (DUO POLLEN 1)			X
endomembrane protein 70		X	
endonuclease/exonuclease/phosphatase family protein	X	X	

ER (ERECTA)	X	X	
ERL1 (ERECTA-LIKE 1)		X	
ERL2 (ERECTA-LIKE 2)	X		
FAMT (farnesoic acid carboxyl-O-methyltransferase)		X	
FEI1 (FEI 1)	X		
formin homology 2 domain-containing protein / FH2 domain-containing protein	X	X	
GAMT2 (GIBBERELLIC ACID METHYLTRANSFERASE 2)		X	
glycosyl hydrolase family 5 protein / cellulase family protein		X	
glycosyl hydrolase family protein 5 / cellulase family protein / (1-4)-beta-mannan endohydrolase		X	
glycine-rich protein			X
GT2		X	X
HD2B (HISTONE DEACETYLASE 2B)	X		
HD2C (HISTONE DEACETYLASE 2C)	X	X	X
HDA3 (HISTONE DEACETYLASE 3)	X	X	X
heat shock protein-related	X	X	X
homogentisate farnesyltransferase/ homogentisate geranylgeranyltransferase/ homogentisate solanesyltransferase	X	X	
HPT1 (HOMOGENTISATE PHYTYLTRANSFERASE 1)	X	X	X
HST (HASTY)	X	X	X
IAMT1 (IAA CARBOXYLMETHYLTRANSFERASE 1)	X	X	X
INO (INNER NO OUTER)	X	X	X
inositol polyphosphate 5-phosphatase		X	
IRX6	X	X	X
KAN (KANADI)	X	X	
KAN2 (KANADI 2)	X	X	X
KAN3 (KANADI 3)	X	X	
kelch repeat-containing protein	X	X	

KNAT1 (KNOTTED-LIKE FROM ARABIDOPSIS THALIANA)	X	X	
KNAT6		X	
KOR2			X
LBD1 (LOB DOMAIN-CONTAINING PROTEIN 1)		X	
LBD2 (LOB DOMAIN-CONTAINING PROTEIN 2)	X		X
LBD4 (LOB DOMAIN-CONTAINING PROTEIN 4)	X	X	X
LBD10 (LOB DOMAIN-CONTAINING PROTEIN 10)	X		X
LBD11 (LOB DOMAIN-CONTAINING PROTEIN 11)	X		X
LBD13 (LOB DOMAIN-CONTAINING PROTEIN 13)	X		X
LBD15 (LOB DOMAIN-CONTAINING PROTEIN 15)	X	X	X
LBD16 (LATERAL ORGAN BOUNDARIES-DOMAIN 16)	X		X
LBD18 (LOB DOMAIN-CONTAINING PROTEIN 18)	X	X	
LBD20 (LOB DOMAIN-CONTAINING PROTEIN 20)		X	
LBD21 (LOB DOMAIN-CONTAINING PROTEIN 21)		X	X
LBD22 (LOB DOMAIN-CONTAINING PROTEIN 22)	X		X
LBD24 (LOB DOMAIN-CONTAINING PROTEIN 24)		X	X
LBD25 (LOB DOMAIN-CONTAINING PROTEIN 25)	X	X	X
LBD27 (LOB DOMAIN-CONTAINING PROTEIN 27)	X	X	X
leucine-rich repeat family protein	X	X	
leucine-rich repeat family protein / protein kinase family protein	X	X	
leucine-rich repeat transmembrane protein kinase	X	X	X
LOB (LATERAL ORGAN BOUNDARIES)	X	X	
LRP1 (LATERAL ROOT PRIMORDIUM 1)			X
LSH1 (LIGHT-DEPENDENT SHORT HYPOCOTYLS 1)	X	X	
LSH3 (LIGHT SENSITIVE HYPOCOTYLS 3)		X	X
LSH4 (LIGHT SENSITIVE HYPOCOTYLS 4)	X	X	
LSH6 (LIGHT SENSITIVE HYPOCOTYLS 6)	X	X	X

LSH7 (LIGHT SENSITIVE HYPOCOTYLS 7)	X		X
LSH9 (LIGHT SENSITIVE HYPOCOTYLS 9)	X	X	
LSH10 (LIGHT SENSITIVE HYPOCOTYLS 10)	X	X	
MP (MONOPTEROS)	X	X	X
MRLK (MERISTEMATIC RECEPTOR-LIKE KINASE)	X		
myb family transcription factor	X	X	X
MYB2 (MYB DOMAIN PROTEIN 2)	X	X	X
MYB62 (myb domain protein 62)	X		X
MYB108 (myb domain protein 108)	X		
MYB305 (myb domain protein 305)	X		X
NIK1 (NSP-INTERACTING KINASE 1)	X	X	
NIK3 (NSP-INTERACTING KINASE 3)	X	X	X
NPH4 (NON-PHOTOTROPHIC HYPOCOTYL)	X	X	X
OBP1 (OBF BINDING PROTEIN 1)		X	
OBP2	X		
OBP4	X	X	
peroxidase 17 (PER17) (P17)	X	X	X
peroxidase 22 (PER22) (P22) (PRXEA) / basic peroxidase E		X	
permease, putative			X
phagocytosis and cell motility protein ELMO1-related	X	X	X
PHB (PHABULOSA)	X	X	X
phototropic-responsive NPH3 family protein	X	X	X
phototropic-responsive protein, putative	X		
PHV (PHAVOLUTA)	X		
PIN1 (PIN-FORMED 1)	X	X	X
PIN3 (PIN-FORMED 3)	X		X
PIN4 (PIN-FORMED 4)	X	X	X
PIN5 (PIN-FORMED 5)	X		

PIN7 (PIN-FORMED 7)			X
PIN8 (PIN-FORMED 8)	X		
protein binding / zinc ion binding	X		
protein phosphatase 2C, putative / PP2C, putative		X	
PTL (PETAL LOSS)	X	X	X
RABA4D (RAB GTPASE HOMOLOG A4D)		X	
RABA5E (RAB GTPASE HOMOLOG A5E)	X	X	X
RAP2.7 (RELATED TO AP2.7)		X	
Ras-related GTP-binding family protein	X		
REV (REVOLUTA)	X	X	X
RHA1 (RAB HOMOLOG 1)	X		
RHA3A			X
ribonuclease III family protein	X	X	
RLK1 (RECEPTOR-LIKE PROTEIN KINASE 1)			X
RNA-dependent RNA polymerase family protein	X	X	X
RPL (REPLUMLESS)			X
RPT2 (ROOT PHOTOTROPISM 2)		X	
S-adenosyl-L-methionine:carboxyl methyltransferase family protein	X	X	
SFC (SCARFACE)	X	X	X
SHI (SHORT INTERNODES)		X	
SRS4 (SHI-RELATED SEQUENCE 4)		X	
SRS5 (SHI-RELATED SEQUENCE 5)	X		
SRS6 (SHI-RELATED SEQUENCE 6)		X	
STY1 (STYLISH 1)		X	
STY2 (STYLISH 2)		X	
TET2 (TETRASPANIN2)	X		
TET8 (TETRASPANIN8)	X		

	TET13 (TETRASPANIN13)			X
	TMKL1 (transmembrane kinase-like 1)	X	X	X
	TRN2 (TORNADO 2)	X	X	X
	ubiquitin carboxyl-terminal hydrolase family protein			X
	WRI1 (WRINKLED 1)		X	X
	xanthine/uracil permease family protein	X	X	
	YAB2 (YABBY2)		X	X
	YAB3 (YABBY3)	X		
	YAB5 (YABBY5)	X	X	
	zinc finger (C3HC4-type RING finger) family protein	X	X	X
	ZLL (ZWILLE)	X	X	X
	ZPR2 (LITTLE ZIPPER 2)		X	
Pollination	acetyl-CoA C-acyltransferase, putative / 3-ketoacyl-CoA thiola			X
GO:0009856	ANNAT2 (Annexin Arabidopsis 2)		X	X
	ANNAT3 (ANNEXIN ARABIDOPSIS 3)			X
	annexin, putative	X		
	ARPN (PLANTACYANIN)	X		X
	calmodulin, putative	X		
	CDC48B	X	X	
	EDA10 (embryo sac development arrest 10)		X	
	emb2411 (embryo defective 2411)		X	
	fimbrin-like protein, putative	X		
	fringe-related protein		X	
	FtsH protease, putative		X	
	FTSH12 (FTSH PROTEASE 12)		X	
	guanine nucleotide exchange family protein			X
	LBA1 (LOW-LEVEL BETA-AMYLASE 1)		X	X
	leucine-rich repeat transmembrane protein kinase, putative	X	X	

	MLO4 (MILDEW RESISTANCE LOCUS O 4)			X
	MLO13 (MILDEW RESISTANCE LOCUS O 13)		X	
	MLO15 (MILDEW RESISTANCE LOCUS O 15)	X		
	MORN (Membrane Occupation and Recognition Nexus)	X		
	NSF (N-ethylmaleimide sensitive factor)	X		
	pectinesterase family protein			X
	PEX6 (PEROXIN 6)	X		
	phosphatidylinositol-4-phosphate 5-kinase family protein		X	
	PKT2 (PEROXISOMAL 3-KETO-ACYL-COA THIOLASE 2)	X		X
	PKT3 (PEROXISOMAL 3-KETOACYL-COA THIOLASE 3)	X		
	plastocyanin-like domain-containing protein / mavicyanin, putative		X	X
	polcalcin, putative / calcium-binding pollen allergen, putative		X	X
	protein kinase family protein	X		
	protein kinase, putative	X		
	RPT5A (REGULATORY PARTICLE TRIPLE-A ATPASE 5A)	X		
	self-incompatibility protein-related			X
	SKD1 (SUPPRESSOR OF K+ TRANSPORT GROWTH DEFECT1)	X		
	TPA_exp: callose synthase	X		
Determination of bilateral symmetry GO:0009855	(1-4)-beta-mannan endohydrolase, putative			X
	ABIL1 (Abi-1-like 1)			X
	ADOF1		X	
	AFH1 (FORMIN HOMOLOGY 1)	X	X	X
	AGC1.5 (AGC KINASE 1.5)	X	X	
	AGC1.7 (AGC KINASE 1.7)	X	X	X
	AGD1 (ARF-GAP domain 1)	X	X	X
	AGD2 (ABERRANT GROWTH AND DEATH 2	X	X	

AGD4 (ARF-GAP domain 4)	X	X	X
AGD6	X		
AGD7		X	
agenet domain-containing protein	X	X	
AGO1 (ARGONAUTE 1)	X	X	X
AGO4 (ARGONAUTE 4)	X	X	X
AGO5 (ARGONAUTE 5)	X	X	X
AGO6 (ARGONAUTE 6)	X		
AIL5 (AINTEGUMENTA-LIKE 5)	X		X
AIL6 (AINTEGUMENTA-LIKE 6)	X	X	X
AIL7 (AINTEGUMENTA-like 7)	X	X	
ALD1 (AGD2-LIKE DEFENSE RESPONSE PROTEIN1)	X	X	
ANAC038 (ARABIDOPSIS NAC DOMAIN CONTAINING PROTEIN 38)	X		
ANAC058	X	X	
anac074 (Arabidopsis NAC domain containing protein 74)			X
ANAC087	X	X	X
ANAC100 (ARABIDOPSIS NAC DOMAIN CONTAINING PROTEIN 100)	X		
ANT (AINTEGUMENTA)	X	X	X
AP2 (APETALA 2)	X		X
AP2 domain-containing transcription factor, putative		X	
ARA4	X	X	
ARA-5 (ARABIDOPSIS RAS 5)	X		X
ARF1 (AUXIN RESPONSE FACTOR 1)	X	X	
ARF4 (AUXIN RESPONSE FACTOR 4)			X
ARF6 (AUXIN RESPONSE FACTOR 6)		X	X
ARF8 (AUXIN RESPONSE FACTOR 8)		X	X
ARF9 (AUXIN RESPONSE FACTOR 9)		X	

ARK3 (ARMADILLO REPEAT KINESIN 3)		X	X
ARF18 (AUXIN RESPONSE FACTOR 18)		X	
ARF17 (AUXIN RESPONSE FACTOR 17)	X		
ARF19 (AUXIN RESPONSE FACTOR 19)	X	X	X
ARF1A1C	X		
ARF3 (ADP-RIBOSYLATION FACTOR 3)	X		
ARF6 (AUXIN RESPONSE FACTOR 6)	X		
ARF8 (AUXIN RESPONSE FACTOR 8)	X		
ARK3 (ARMADILLO REPEAT KINESIN 3)	X		X
armadillo/beta-catenin repeat family protein / kinesin motor family protein	X	X	X
AS2 (ASYMMETRIC LEAVES 2)	X		
ASL1	X	X	X
ASL5	X		X
ASL9 (ASYMMETRIC LEAVES 2 LIKE 9)	X		
aspartic-type endopeptidase/ peptidase	X	X	
aspartyl-tRNA synthetase, putative / aspartate--tRNA ligase, putative	X	X	X
ATARFA1B (ADP-ribosylation factor A1B)	X		
ATARFA1E (ADP-ribosylation factor A1E)	X		
ATARFA1F (ARABIDOPSIS THALIANA ADP-RIBOSYLATION FACTOR A1F)	X	X	
ATARFD1B (ADP-ribosylation factor D1B)	X		
ATCEL2	X	X	X
ATCEL3 (ARABIDOPSIS THALIANA CELLULASE 3)		X	X
ATFP8		X	
AtGH9A4 (Arabidopsis thaliana Glycosyl Hydrolase 9A4)		X	
ATGH9B1 (ARABIDOPSIS THALIANA GLYCOSYL HYDROLASE 9B1)	X	X	

AtGH9B8 (Arabidopsis thaliana glycosyl hydrolase 9B8)		X	
AtGH9B13 (Arabidopsis thaliana glycosyl hydrolase 9B13)	X		X
AtGH9B15 (Arabidopsis thaliana glycosyl hydrolase 9B15)	X		X
AtGH9B17 (Arabidopsis thaliana glycosyl hydrolase 9B17)	X		
AtGH9B18 (Arabidopsis thaliana glycosyl hydrolase 9B18)	X		
AtGH9C2 (Arabidopsis thaliana glycosyl hydrolase 9C2)		X	
AtGH9C3 (Arabidopsis thaliana glycosyl hydrolase 9C3)			X
AT-GTL1 (GT2-LIKE 1)	X		X
ATHMG (ARABIDOPSIS THALIANA HIGH MOBILITY GROUP)	X		
ATK1 (ARABIDOPSIS THALIANA KINESIN 1)	X		
ATK5 (ARABIDOPSIS THALIANA KINESIN 5)	X		
ATL5 (ATL5)			
ATL8		X	
ATMYB21 (ARABIDOPSIS THALIANA MYB DOMAIN PROTEIN 21)	X	X	X
ATNAC3 (ARABIDOPSIS NAC DOMAIN CONTAINING PROTEIN 3)			X
ATP binding / ATP-dependent DNA helicase/ DNA binding / hydrolase		X	X
ATP binding / cAMP-dependent protein kinase regulator/ catalytic/ protein kinase/ protein serine/threonine phosphatase	X		
ATP binding / DNA binding / DNA-directed DNA polymerase/ nucleoside-triphosphatase/ nucleotide binding	X		
ATP-binding region, ATPase-like domain-containing protein-related	X	X	
ATPK7	X	X	X
ATRAB8	X		X
ATRAB18 (ARABIDOPSIS RAB GTPASE HOMOLOG B18)	X		

ATRAB1A	X		
ATRAB1C	X	X	
ATRABA2B (ARABIDOPSIS RAB GTPASE HOMOLOG A2B)		X	
ATRABA3 (ARABIDOPSIS RAB GTPASE HOMOLOG A3)		X	
AtRABA5a (Arabidopsis Rab GTPase homolog A5a)	X	X	X
AtRABA5b (Arabidopsis Rab GTPase homolog A5b)	X	X	X
AtRABA5d (Arabidopsis Rab GTPase homolog A5d)		X	
ATRABH1D (ARABIDOPSIS RAB GTPASE HOMOLOG H1D)	X		
ATTOC120	X	X	
BAM1 (BARELY ANY MERISTEM 1)	X	X	X
BAM2 (BARELY ANY MERISTEM 2)	X		
BAM3 (BARELY ANY MERISTEM 3)	X	X	X
basic helix-loop-helix (bHLH) protein-related	X		
BBM (BABY BOOM)	X		
beta-amylase activity		X	
BIM1	X		
BIM2 (BES1-interacting Myc-like protein 2)	X		
BIM3 (BES1-interacting Myc-like protein 3)			X
BLH1 (BEL1-LIKE HOMEODOMAIN 1)		X	
BLH5 (BELL1-like homeodomain 5)			X
BLH8 (BEL1-LIKE HOMEODOMAIN 8)		X	
BLH11 (BEL1-LIKE HOMEODOMAIN 11)	X		
BRI1 (BRASSINOSTEROID INSENSITIVE 1)	X		
BST1 (BRISTLED 1)	X		
C2 domain-containing protein	X	X	X
calcium-binding EF hand family protein	X	X	X
calcium-binding EF hand family protein, putative	X		X

CHR24 (chromatin remodeling 24)	X	X	X
CLAVATA1 receptor kinase (CLV1)	X		
CLV2 (clavata 2)	X	X	X
COB (COBRA)	X	X	
COBL1 (COBRA-LIKE PROTEIN 1 PRECURSOR)	X	X	X
COBL2 (COBRA-LIKE PROTEIN 2 PRECURSOR)		X	X
COBL5 (COBRA-LIKE PROTEIN 5 PRECURSOR)	X	X	
COL3 (CONSTANS-LIKE 3)	X		
CPSF73-I (CLEAVAGE AND POLYADENYLATION SPECIFICITY FACTOR 73-I)	X	X	X
CRN (CORYNE)			X
CRR6 (chlororespiratory reduction 6)	X		
CUC2 (CUP-SHAPED COTYLEDON 2)		X	X
CUC3 (CUP SHAPED COTYLEDON3)	X		
Cul4-RING E3 ubiquitin ligase complex	X		
CVP2 (COTYLEDON VASCULAR PATTERN 2)	X	X	
D6PK (D6 PROTEIN KINASE)	X	X	X
D6PKL1 (D6 PROTEIN KINASE LIKE 1)	X	X	
D6PKL2 (D6 PROTEIN KINASE LIKE 2)	X	X	X
DAG1 (dof affecting germination 1)	X		
DCL2 (DICER-LIKE 2)	X	X	X
DEAD/DEAH box helicase, putative		X	
Dof-type zinc finger domain-containing protein	X	X	X
DOT2 (DEFECTIVELY ORGANIZED TRIBUTARIES 2)	X	X	X
DUO1 (DUO POLLEN 1)			X
EDA10 (embryo sac development arrest 10)			X
endomembrane protein 70, putative		X	
endonuclease/exonuclease/phosphatase family protein	X	X	X

EOL1 (ETO1-LIKE 1)	X	X	X
EOL2 (ETO1-LIKE 2)	X	X	
ETO1 (ETHYLENE OVERPRODUCER 1)	X	X	X
ETT (ETTIN)	X	X	X
FEI1 (FEI 1)	X		X
Fh5 (FORMIN HOMOLOGY5)	X		
formin homology 2 domain-containing protein / FH2 domain-containing protein	X	X	X
glycosyl hydrolase family 5 protein / cellulase family protein		X	
glycosyl hydrolase family protein 5 / cellulase family protein / (1-4)-beta-mannan endohydrolase, putative		X	
glycine-rich protein			X
GT2		X	X
HDG9 (HOMEODOMAIN GLABROUS 9)	X		
heat shock protein-related	X	X	X
helicase-related		X	X
hexose transporter, putative	X	X	X
inositol polyphosphate 5-phosphatase, putative		X	
IRX6	X	X	X
kelch repeat-containing protein	X	X	
kinase interacting family protein	X	X	X
kinase interacting protein-related			X
kinesin heavy chain, putative	X	X	
KIPK (KCBP-interacting protein kinase)	X	X	
KNAT1 (KNOTTED-LIKE FROM ARABIDOPSIS THALIANA)	X	X	
KNAT3 (KNOTTED1-LIKE HOMEOBOX GENE 3)	X		
KNAT6		X	X
KOR2			X

LBD1 (LOB DOMAIN-CONTAINING PROTEIN 1)		X	
LBD4 (LOB DOMAIN-CONTAINING PROTEIN 4)	X	X	X
LBD10 (LOB DOMAIN-CONTAINING PROTEIN 10)	X		X
LBD11 (LOB DOMAIN-CONTAINING PROTEIN 11)	X		X
LBD13 (LOB DOMAIN-CONTAINING PROTEIN 13)	X		X
LBD15 (LOB DOMAIN-CONTAINING PROTEIN 15)		X	X
LBD16 (LATERAL ORGAN BOUNDARIES-DOMAIN 16)	X		X
LBD18 (LOB DOMAIN-CONTAINING PROTEIN 18)		X	
LBD21 (LOB DOMAIN-CONTAINING PROTEIN 21)		X	X
LBD22 (LOB DOMAIN-CONTAINING PROTEIN 22)	X		X
LBD24 (LOB DOMAIN-CONTAINING PROTEIN 24)			X
LBD25 (LOB DOMAIN-CONTAINING PROTEIN 25)	X	X	X
LBD27 (LOB DOMAIN-CONTAINING PROTEIN 27)			X
leucine-rich repeat family protein	X	X	X
leucine-rich repeat family protein / protein kinase family protein	X	X	X
leucine-rich repeat transmembrane protein kinase, putative	X	X	X
LHW (LONESOME HIGHWAY)	X	X	X
LOB (LATERAL ORGAN BOUNDARIES)	X		
LRP1 (LATERAL ROOT PRIMORDIUM 1)	X		X
LSH1 (LIGHT-DEPENDENT SHORT HYPOCOTYLS 1)	X	X	
LSH3 (LIGHT SENSITIVE HYPOCOTYLS 3)		X	X
LSH4 (LIGHT SENSITIVE HYPOCOTYLS 4)	X	X	
LSH6 (LIGHT SENSITIVE HYPOCOTYLS 6)	X	X	X
LSH7 (LIGHT SENSITIVE HYPOCOTYLS 7)	X		X
LSH9 (LIGHT SENSITIVE HYPOCOTYLS 9)	X	X	
LSH10 (LIGHT SENSITIVE HYPOCOTYLS 10)	X	X	
LUG (LEUNIG)	X	X	X
MP (MONOPTEROS)	X	X	X

MRH2 (MORPHOGENESIS OF ROOT HAIR 2)		X	X
MYB2 (MYB DOMAIN PROTEIN 2)	X	X	X
MYB62 (myb domain protein 62)	X		X
MYB108 (myb domain protein 108)	X		
MYB305 (myb domain protein 305)	X		X
myosin heavy chain-related		X	
NAD(P)H dehydrogenase complex assembly		X	
NIK1 (NSP-INTERACTING KINASE 1)	X	X	
NIK3 (NSP-INTERACTING KINASE 3)	X	X	X
NPH4 (NON-PHOTOTROPHIC HYPOCOTYL)	X	X	X
OBP1 (OBF BINDING PROTEIN 1)		X	
OBP2	X		
OBP4	X	X	
PARP2 (POLY(ADP-RIBOSE) POLYMERASE 2)	X		X
PAZ domain-containing protein / piwi domain-containing protein		X	
peptidyl-prolyl cis-trans isomerase cyclophilin-type family protein		X	
permease, putative	X	X	X
peroxidase 17 (PER17) (P17)	X	X	X
peroxidase 22 (PER22) (P22) (PRXEA) / basic peroxidase E		X	
phagocytosis and cell motility protein ELMO1-related	X	X	X
PHB (PHABULOSA)			X
PHOT1 (PHOTOTROPIN 1)			X
PHOT2 (PHOTOTROPIN 2)	X	X	
phototropic-responsive NPH3 family protein	X	X	X
phototropic-responsive protein, putative	X		
PID (PINOID)		X	
PID2 (PINOID2)	X		

PIN1 (PIN-FORMED 1)	X	X	X
PIN3 (PIN-FORMED 3)	X		X
PIN4 (PIN-FORMED 4)	X	X	X
PIN5 (PIN-FORMED 5)	X		X
PIN6 (PIN-FORMED 6)	X	X	X
PIN7 (PIN-FORMED 7)			X
PIN8 (PIN-FORMED 8)	X	X	
PK1 (PROTEIN-SERINE KINASE 1)	X		
protein binding / zinc ion binding		X	
protein kinase family protein	X	X	X
protein kinase family protein / C-type lectin domain-containing protein	X		X
protein phosphatase 2C, putative / PP2C, putative	X		
PTL (PETAL LOSS)	X	X	X
RABA4D (RAB GTPASE HOMOLOG A4D)		X	
RABA5E (RAB GTPASE HOMOLOG A5E)	X	X	X
RAP2.7 (RELATED TO AP2.7)		X	
Ras-related GTP-binding family protein	X		
RDR2 (RNA-DEPENDENT RNA POLYMERASE 2)	X	X	X
remorin family protein		X	
RHA1 (RAB HOMOLOG 1)	X		
RHA3A			X
ribonuclease III family protein	X	X	
RLK1 (RECEPTOR-LIKE PROTEIN KINASE 1)			X
RNA-dependent RNA polymerase family protein	X	X	X
RPL (REPLUMLESS)			X
RPT2 (ROOT PHOTOTROPISM 2)		X	
SFC (SCARFACE)	X	X	X

SGB1 (SUPPRESSOR OF G PROTEIN BETA1)		X	
SHI (SHORT INTERNODES)		X	
SNF2 domain-containing protein		X	X
SPK1 (SPIKE1)	X	X	
SRF7 (STRUBBELIG-RECEPTOR FAMILY 7)	X		
SRS4 (SHI-RELATED SEQUENCE 4)		X	
SRS5 (SHI-RELATED SEQUENCE 5)	X		
SRS6 (SHI-RELATED SEQUENCE 6)		X	
STM (SHOOT MERISTEMLESS)	X	X	
STY1 (STYLISH 1)		X	
STY2 (STYLISH 2)		X	
SUS2 (ABNORMAL SUSPENSOR 2)	X	X	X
SUS2 (SUCROSE SYNTHASE 2)	X		
SUS3 (sucrose synthase 3)	X	X	
SUS4	X	X	X
SUS5	X		X
SUS6 (SUCROSE SYNTHASE 6)	X	X	X
SUVH4 (SU(VAR)3-9 HOMOLOG 4)	X	X	X
SYD (SPLAYED)			X
TCP1 (TCP1)	X	X	X
TET2 (TETRASPANIN2)	X		
TET8 (TETRASPANIN8)	X		
TET13 (TETRASPANIN13)			X
tetratricopeptide repeat (TPR)-containing protein	X		X
TMKL1 (transmembrane kinase-like 1)	X	X	X
TOC132 (MULTIMERIC TRANSLOCON COMPLEX IN THE OUTER ENVELOPE MEMBRANE 132)	X	X	X
TOC159 (TRANSLOCON AT THE OUTER ENVELOPE	X	X	X

	MEMBRANE OF CHLOROPLASTS 159)			
	TPA_exp: actin-related protein 4		X	
	TPR3 (TOPLESS-RELATED 3)		X	
	transducin family protein / WD-40 repeat family protein	X	X	
	TRN1 (TORNADO 1)	X	X	X
	TRN1 (TRANSPORTIN 1)	X	X	X
	TRN2 (TORNADO 2)	X	X	X
	TTN7 (TITAN7)		X	
	TTN8 (TITAN8)	X	X	X
	ubiquitin carboxyl-terminal hydrolase family protein			X
	VND7 (VASCULAR RELATED NAC-DOMAIN PROTEIN 7)		X	
	WAG1 (WAG 1)			X
	WD-40 repeat family protein	X		
	WRI1 (WRINKLED 1)		X	X
	xanthine/uracil permease family protein	X	X	X
	zinc finger (C3HC4-type RING finger) family protein		X	X
	ZLL (ZWILLE)	X	X	X
	ZPR2 (LITTLE ZIPPER 2)		X	
Anthocyanin accumulation in tissues in response to UV light GO:0043481	26S protease regulatory complex subunit 4, putative	X		
	26S proteasome regulatory complex subunit p42D, putative	X		
	ABC transporter family protein	X	X	X
	ABCB1 (ATP BINDING CASSETTE SUBFAMILY B1)	X	X	X
	ABCB19	X	X	X
	ACC2 (ACETYL-COA CARBOXYLASE 2)	X		
	acid phosphatase class B family protein	X	X	X
	ACL5 (ACAULIS 5)	X	X	X

ACLA-1	X	X	X
ACLA-2	X		X
ACLA-3	X	X	X
ACLB-1	X		
ACT2 (ACTIN 2)	X	X	X
ACT3 (actin 3)	X		
ACT4 (ACTIN 4)			X
ACT7 (ACTIN 7)	X		X
ACT11 (actin-11)	X		X
ACT12 (ACTIN-12)		X	
ACYB-2	X	X	X
ADT1 (arogenate dehydratase 1)	X		X
ADT2 (arogenate dehydratase 2)	X	X	X
ADT4 (arogenate dehydratase 4)	X	X	X
ADT5 (arogenate dehydratase 5)			X
ADT6 (arogenate dehydratase 6)	X	X	X
AFB2 (AUXIN SIGNALING F-BOX 2)	X	X	
AFB3 (AUXIN SIGNALING F-BOX 3)		X	
AFH1 (FORMIN HOMOLOGY 1)	X	X	X
AG (AGAMOUS)	X		X
AGC1.5 (AGC KINASE 1.5)		X	
AGL6 (AGAMOUS-LIKE 6)	X	X	X
AGL8 (agamous-like 8)	X		X
AGL12 (AGAMOUS-LIKE 12)	X		
AGL14 (agamous-like 14)		X	X
AGL16 (AGAMOUS-LIKE 16)		X	
AGL18	X	X	X
AGL19 (AGAMOUS-LIKE 19)	X		

AGL20 (AGAMOUS-LIKE 20)	X	X	X
AGL21			X
AGL29 (AGAMOUS-LIKE 29)		X	
AGL42 (AGAMOUS LIKE 42)	X	X	X
AGL44 (AGAMOUS-LIKE 44)	X		
AGL48 (AGAMOUS-LIKE 48)			X
AGL57		X	
AGL65 (AGAMOUS-LIKE 65)		X	
AGL71 (AGAMOUS-LIKE 71)	X		
AGL97 (AGAMOUS-LIKE 97)		X	
AGL104 (AGAMOUS-LIKE 104)	X		
AHK5 (ARABIDOPSIS HISTIDINE KINASE 5)	X	X	X
AIR1	X		
amino acid permease, putative		X	X
ANL2 (ANTHOCYANINLESS 2)	X	X	
AP1 (APETALA1)	X		X
apocytochrome B	X	X	X
ARAC1	X	X	
ARAC2 (ARABIDOPSIS RAC-LIKE 2)	X		X
ARAC3 (ARABIDOPSIS RAC-LIKE 3)	X	X	
ARAC5 (RAC-LIKE GTP BINDING PROTEIN 5)	X		X
ARAC9	X	X	X
ARAC10		X	
AT59	X		X
ATATH12			X
ATCAP1 (ARABIDOPSIS THALIANA CYCLASE ASSOCIATED PROTEIN 1)	X	X	X
ATCBL3 (ARABIDOPSIS THALIANA CALCINEURIN B-LIKE 3)	X		

ATCSLD4	X		
ATGCN1	X		
ATGH9A1 (ARABIDOPSIS THALIANA GLYCOSYL HYDROLASE 9A1)	X	X	X
ATGH9A3 (ARABIDOPSIS THALIANA GLYCOSYL HYDROLASE 9A3)		X	
AtGH9A4 (Arabidopsis thaliana Glycosyl Hydrolase 9A4)		X	
AtGH9B15 (Arabidopsis thaliana glycosyl hydrolase 9B15)		X	
AtGH9B7 (Arabidopsis thaliana glycosyl hydrolase 9B7)	X	X	X
AtGH9B8 (Arabidopsis thaliana glycosyl hydrolase 9B8)		X	
AtGH9C2 (Arabidopsis thaliana glycosyl hydrolase 9C2)		X	
ATMRP13			X
ATNAP8	X	X	X
ATNAP9	X		
ATOPT2	X	X	X
ATOPT3 (OLIGOPEPTIDE TRANSPORTER)	X	X	X
ATOPT9 (ARABIDOPSIS THALIANA OLIGOPEPTIDE TRANSPORTER 9)		X	
ATSIK		X	
AUX1 (AUXIN RESISTANT 1)	X	X	X
AXR3 (AUXIN RESISTANT 3)		X	
BAM1 (BARELY ANY MERISTEM 1)	X	X	X
BAM1 (BETA-AMYLASE 1)		X	X
BAM2 (BARELY ANY MERISTEM 2)	X		
BAM2 (BETA-AMYLASE 2)		X	X
BAM3 (BARELY ANY MERISTEM 3)	X	X	X
BAM4 (BETA-AMYLASE 4)		X	
BAM6 (BETA-AMYLASE 6)			X
BAM7 (BETA-AMYLASE 7)	X	X	

BCCP2 (BIOTIN CARBOXYL CARRIER PROTEIN 2)		X	X
beta-hydroxyacyl-ACP dehydratase, putative	X	X	X
BETA-TIP (BETA-TONOPLAST INTRINSIC PROTEIN)	X	X	
BMY2 (BETA-AMYLASE 2)	X	X	
BRI1 (BRASSINOSTEROID INSENSITIVE 1)		X	X
BRL2 (BRI1-LIKE 2)			X
BURP domain-containing protein / polygalacturonase, putative		X	X
CAC1 (CHLOROPLASTIC ACETYLCOENZYME A CARBOXYLASE 1)		X	X
CAC2	X	X	X
calmodulin-binding protein-related		X	
CBL1 (CALCINEURIN B-LIKE PROTEIN 1)	X	X	
CBL5 (CALCINEURIN B-LIKE PROTEIN 5)			X
CBL8 (CALCINEURIN B-LIKE PROTEIN 8)	X		X
CBL10 (CALCINEURIN B-LIKE 10)	X	X	X
cell division cycle protein 48, putative			X
CESA1 (CELLULOSE SYNTHASE 1)	X	X	X
CESA2 (CELLULOSE SYNTHASE A2)	X		X
CESA4 (CELLULOSE SYNTHASE A4)	X		X
CESA6 (CELLULOSE SYNTHASE 6)	X	X	X
CESA9 (CELLULOSE SYNTHASE A9)	X		
CESA10 (CELLULOSE SYNTHASE 10)			X
CEV1 (CONSTITUTIVE EXPRESSION OF VSP 1)	X	X	X
CLASP (CLIP-ASSOCIATED PROTEIN)	X	X	
CLS (CARDIOLIPIN SYNTHASE)	X	X	
CLV1 (CLAVATA 1)	X	X	X
COB (COBRA)	X	X	
COBL1 (COBRA-LIKE PROTEIN 1 PRECURSOR)	X	X	

COBL2 (COBRA-LIKE PROTEIN 2 PRECURSOR)		X	X
COBL5 (COBRA-LIKE PROTEIN 5 PRECURSOR)	X	X	
coenzyme Q biosynthesis Coq4 family protein			X
CPD (CONSTITUTIVE PHOTOMORPHOGENIC DWARF)	X	X	
CSLD2 (CELLULOSE-SYNTHASE LIKE D2)		X	
CSN5B (COP9-SIGNALOSOME 5B)			X
CT-BMY (CHLOROPLAST BETA-AMYLASE)	X	X	X
CYP90D1		X	
D6PK (D6 PROTEIN KINASE)		X	X
dehydration-responsive family protein			X
dehydration-responsive protein-related		X	X
DELTA-TIP	X		
DER1 (DERLIN-1)	X	X	X
DNA binding / protein binding / transcription regulator	X		
DRT112	X	X	X
DWF1 (DWARF 1)	X	X	X
DWF4 (DWARF 4)			X
EDM2		X	X
EIR1 (ETHYLENE INSENSITIVE ROOT 1)		X	
ELP1 (EDM2-LIKE PROTEIN1)			X
EMB25 (EMBRYO DEFECTIVE 25)			X
emb2731 (embryo defective 2731)		X	
esterase/lipase/thioesterase family protein		X	
EXGT-A1 (ENDOXYLOGLUCAN TRANSFERASE)	X	X	X
EXGT-A4 (ENDOXYLOGLUCAN TRANSFERASE A4)		X	
FEI1 (FEI 1)	X		X
Fh5 (FORMIN HOMOLOGY5)	X		
FLC (FLOWERING LOCUS C)	X		

formin homology 2 domain-containing protein / FH2 domain-containing protein	X	X	X
galactosyl transferase GMA12			X
galactosyl transferase GMA12/MNN10 family protein		X	
GAMMA-TIP (GAMMA TONOPLAST INTRINSIC PROTEIN)	X	X	X
GASA1 (GAST1 PROTEIN HOMOLOG 1)	X	X	X
GDSL-motif lipase			X
GDSL-motif lipase/hydrolase family protein	X	X	X
gibberellin-regulated family protein			X
gibberellin-responsive protein, putative	X		
GL2 (GLABRA 2)	X		X
GPAT1 (GLYCEROL-3-PHOSPHATE ACYLTRANSFERASE 1)	X	X	
GPAT4 (GLYCEROL-3-PHOSPHATE ACYLTRANSFERASE 4)	X		X
GPAT6 (GLYCEROL-3-PHOSPHATE ACYLTRANSFERASE 6)	X	X	X
GPAT8 (glycerol-3-phosphate acyltransferase 8)		X	
GRH1 (GRR1-LIKE PROTEIN 1)	X	X	
GSTL2	X		
GTB1			X
HAE (HAESA)	X		X
HB-7 (HOMEOBOX-7)	X		X
HDG1 (HOMEODOMAIN GLABROUS 1)	X	X	
HDG2 (HOMEODOMAIN GLABROUS 2)		X	X
HDG7 (HOMEODOMAIN GLABROUS 7)	X	X	
HDG8 (HOMEODOMAIN GLABROUS 8)	X		
HDG9 (HOMEODOMAIN GLABROUS 9)	X		

HDG10 (HOMEODOMAIN GLABROUS 10)	X		
HDG11 (HOMEODOMAIN GLABROUS 11)	X	X	X
HEXO3 (BETA-HEXOSAMINIDASE 3)		X	
HSI2 (HIGH-LEVEL EXPRESSION OF SUGAR-INDUCIBLE GENE 2)	X	X	
HSL1 (HAESA-Like 1)	X		X
HSL1 (HSI2-LIKE 1)	X		X
IAA7 (INDOLE-3-ACETIC ACID 7)	X		X
IAA8			X
IAA16	X	X	X
IRX1 (IRREGULAR XYLEM 1)			X
IRX3 (IRREGULAR XYLEM 3)		X	X
IRX6	X	X	X
JP630		X	
KCS2 (3-KETOACYL-COA SYNTHASE 2)	X		
KCS4 (3-KETOACYL-COA SYNTHASE 4)		X	
KCS6 (3-KETOACYL-COA SYNTHASE 6)	X		X
KCS11 (3-KETOACYL-COA SYNTHASE 11)	X		
KOR2	X	X	X
KUP3 (K+ UPTAKE TRANSPORTER 3)	X		
LAX3 (LIKE AUX1 3)	X	X	X
LCR68 (LOW-MOLECULAR-WEIGHT CYSTEINE-RICH 68)	X	X	
LCR69 (LOW-MOLECULAR-WEIGHT CYSTEINE-RICH 69)	X	X	X
legume lectin family protein		X	X
leucine-rich repeat family protein / protein kinase family protein	X	X	X
LIM domain-containing protein	X	X	X
LNG1 (LONGIFOLIA1)		X	X
LNG2 (LONGIFOLIA2)	X		

lyase/ pectate lyase	X	X	
MADS-box protein (AGL60)		X	
MADS-box protein (AGL72)		X	X
MADS-box protein (AGL100)			X
MAF1 (MADS AFFECTING FLOWERING 1)		X	
MAF3 (MADS AFFECTING FLOWERING 3)			X
MAF4 (MADS AFFECTING FLOWERING 4)	X		
MAP1A (METHIONINE AMINOPEPTIDASE 1A)	X	X	X
MAP1C (METHIONINE AMINOPEPTIDASE 1B)	X		
MCCA	X		
NCRK	X		
NIK1 (NSP-INTERACTING KINASE 1)	X	X	X
NIK2 (NSP-INTERACTING KINASE 2)			X
NIK3 (NSP-INTERACTING KINASE 3)		X	X
nodulin MtN21 family protein	X	X	X
NPH3 (NON-PHOTOTROPIC HYPOCOTYL 3)		X	X
OPT1 (OLIGOPEPTIDE TRANSPORTER 1)	X	X	
OPT4 (OLIGOPEPTIDE TRANSPORTER 4)	X		X
OPT6 (OLIGOPEPTIDE TRANSPORTER 1)	X		X
OPT5 (OLIGOPEPTIDE TRANSPORTER 5)	X	X	
OPT7 (OLIGOPEPTIDE TRANSPORTER 7)	X	X	X
PAP2 (PHYTOCHROME-ASSOCIATED PROTEIN 2)			X
PAS2 (PASTICCINO 2)	X	X	X
PDF2 (PROTODERMAL FACTOR 2)	X	X	X
pEARLI 1			X
pectate lyase family protein	X	X	X
pentatricopeptide (PPR) repeat-containing protein			X
PEP (PEPPER)	X	X	

PGP1 (PHOSPHATIDYLGLYCEROLPHOSPHATE SYNTHASE 1)	X	X	
PGP2 (P-GLYCOPROTEIN 2)	X	X	X
PGP6 (P-GLYCOPROTEIN 6)		X	X
PGP9 (P-GLYCOPROTEIN 9)	X	X	X
PGP10 (P-GLYCOPROTEIN 10)		X	
PGP11 (P-GLYCOPROTEIN 11)	X	X	
PGP13 (P-GLYCOPROTEIN 13)	X	X	X
PGP20 (P-GLYCOPROTEIN 20)	X	X	
PGP21 (P-GLYCOPROTEIN 21)	X		X
PGPS2 (phosphatidylglycerolphosphate synthase 2)	X		
phototropic-responsive NPH3 family protein		X	
PI (PISTILLATA)	X		X
PID (PINOID)	X	X	X
PIN1 (PIN-FORMED 1)	X	X	X
PIN3 (PIN-FORMED 3)	X		X
PIN4 (PIN-FORMED 4)	X	X	X
PIN5 (PIN-FORMED 5)	X	X	X
PIN6 (PIN-FORMED 6)	X		
PIN7 (PIN-FORMED 7)			X
PIN8 (PIN-FORMED 8)	X	X	
PMR6 (powdery mildew resistant 6)	X		X
POM1 (POM-POM1)	X	X	X
PP2AA2 (PROTEIN PHOSPHATASE 2A SUBUNIT A2)	X	X	X
PP2AA3 (PROTEIN PHOSPHATASE 2A SUBUNIT A3)	X		X
proline-rich family protein	X	X	
protease inhibitor/seed storage/lipid transfer protein (LTP) family protein	X		
protein binding / structural molecule	X		

protein kinase family protein	X	X	
protein kinase, putative		X	
RCN1 (ROOTS CURL IN NPA)	X	X	X
ROP1 (RHO-RELATED PROTEIN FROM PLANTS 1)	X		X
ROP2 (RHO-RELATED PROTEIN FROM PLANTS 2)	X		
ROP9 (RHO-RELATED PROTEIN FROM PLANTS 9)	X		
ROP10 (RHO-RELATED PROTEIN FROM PLANTS 10)	X		
ROP10 (RHO-RELATED PROTEIN FROM PLANTS 10)	X		
ROPGEF1		X	X
ROPGEF2		X	
ROPGEF4 (RHO GUANYL-NUCLEOTIDE EXCHANGE FACTOR 4)		X	X
ROPGEF5 (ROP GUANINE NUCLEOTIDE EXCHANGE FACTOR 5)	X	X	X
ROPGEF6	X		
ROPGEF7	X	X	X
ROPGEF8	X		
ROT3 (ROTUNDIFOLIA 3)	X	X	
RPT2 (ROOT PHOTOTROPISM 2)		X	
RPT2a (regulatory particle AAA-ATPase 2a)	X		
RPT3 (REGULATORY PARTICLE TRIPLE-A ATPASE 3)	X		X
RPT5A (REGULATORY PARTICLE TRIPLE-A ATPASE 5A)	X		
RPT6A (REGULATORY PARTICLE TRIPLE-A ATPASE 6A)	X		
RTE1 (REVERSION-TO-ETHYLENE SENSITIVITY1)		X	
SEP1 (STRESS ENHANCED PROTEIN 1)		X	
SEP2 (STRESS ENHANCED PROTEIN 2)	X	X	X
SEP2 (SEPALLATA 2)			X
SEP3 (SEPALLATA3)		X	
SEP4 (SEPALLATA 4)	X	X	X

serine/threonine protein phosphatase 2A (PP2A) regulatory subunit B', putative		X	
SKS1 (SKU5 SIMILAR 1)	X	X	X
SKS2 (SKU5 SIMILAR 2)	X		
sks3 (SKU5 Similar 3)	X	X	X
sks4 (SKU5 Similar 4)	X	X	X
sks5 (SKU5 Similar 5)	X	X	X
SKU5			X
SKS6 (SKU5-SIMILAR 6)	X	X	X
sks7 (SKU5 Similar 7)		X	
sks9 (SKU5 Similar 9)	X		
sks11 (SKU5 Similar 11)		X	
sks12 (SKU5 Similar 12)	X	X	X
sks13 (SKU5 Similar 13)	X	X	
sks14 (SKU5 Similar 14)		X	X
sks15 (SKU5 Similar 15)		X	X
sks17 (SKU5 Similar 17)	X		X
SKU5	X	X	
SNL1 (SIN3-LIKE 1)	X		
SNL2 (SIN3-LIKE 2)	X	X	X
SNL3 (SIN3-LIKE 3)	X		
SNL4 (SIN3-LIKE 4)	X	X	
SNL5 (SIN3-LIKE 5)	X		X
SNL6 (SIN3-LIKE 6)	X		
SOS3 (SALT OVERLY SENSITIVE 3)	X	X	X
SP1L1 (SPIRAL1-LIKE1)		X	
SP1L4 (SPIRAL1-LIKE4)	X		
SP1L5 (SPIRAL1-LIKE5)		X	X

STK (SEEDSTICK)			X
TCH4 (Touch 4)	X		
thioesterase family protein	X		
TIP1	X	X	X
TIP2 (TONOPLAST INTRINSIC PROTEIN 2)		X	
TIP4		X	
TIR1 (TRANSPORT INHIBITOR RESPONSE 1)		X	X
TPA_exp: actin-related protein 2			X
TT16 (TRANSPARENT TESTA16)		X	
TUA1 (ALPHA-1 TUBULIN)	X	X	
TUA3	X	X	X
TUA4	X		
TUA6	X		X
TUB1	X	X	X
TUB2	X	X	X
TUB4	X		
TUB5	X		
TUB6 (BETA-6 TUBULIN)	X	X	X
TUB7	X	X	X
TUB8	X	X	X
TUB9	X	X	
ubiquitin family protein			X
UPL7	X		
WAG1 (WAG 1)			X
WLIM1	X		X
XT1 (XYLOSYLTRANSFERASE 1)	X		
XT2 (UDP-XYLOSYLTRANSFERASE 2)	X	X	X
XTH21 (XYLOGLUCAN		X	

	ENDOTRANSGLUCOSYLASE/HYDROLASE 21)			
	XTR6 (XYLOGLUCAN ENDOTRANSGLYCOSYLASE 6)		X	X
	XXT5 (XYLOGLUCAN XYLOSYLTRANSFERASE 5)		X	X
	xyloglucan:xyloglucosyl transferase, putative	X	X	X
Negative regulation of anthocyanin metabolic process GO:0031538	ABO1 (ABA-OVERLY SENSITIVE 1)	X	X	X
	DNAJ heat shock N-terminal domain-containing protein			X
	Paxneb protein-related	X	X	X
Positive regulation of anthocyanin metabolic process GO:0031539	basic helix-loop-helix (bHLH) family protein	X	X	
	BIM3 (BES1-interacting Myc-like protein 3)		X	
	BZIP17		X	X
	chromosome scaffold protein-related		X	
	gibberellic acid mediated signaling pathway		X	
	HFR1 (LONG HYPOCOTYL IN FAR-RED)	X		
	HY5 (ELONGATED HYPOCOTYL 5)	X	X	X
	HYH (HY5-HOMOLOG)			X
	PAP3 (PURPLE ACID PHOSPHATASE 3)	X		
	PAP7 (PURPLE ACID PHOSPHATASE 7)	X	X	X
	PAP17	X	X	
	PIF3 (PHYTOCHROME INTERACTING FACTOR 3)		X	X
	PIF7 (PHYTOCHROME-INTERACTING FACTOR7)			X
	PIL2 (PHYTOCHROME INTERACTING FACTOR 3-LIKE 2)	X	X	
	PIL5 (PHYTOCHROME INTERACTING FACTOR 3-LIKE 5)	X	X	X
	PIL6 (PHYTOCHROME INTERACTING FACTOR 3-LIKE 6)			X
	purple acid phosphatase family protein	X	X	X
	SPT (SPATULA)	X	X	

UNE10 (unfertilized embryo sac 10)	X	X
UNE12 (unfertilized embryo sac 12)		X

Appendix 4.7. List of selected candidate genes encoding enzymes putatively involved in core anthocyanin biosynthesis, sorted alphabetically. TAIR ID corresponds to the *Arabidopsis thaliana* gene identifier and enzyme code is the enzyme commission number assigned for a candidate protein. Species code: *Cleomella serrulata* (CS), *Melidiscus giganteus* (MG), and *P. dodecandra* (PD).

Candidate gene/encoded enzyme	TAIR ID	Enzyme code	Unigene name	Length (bp)	Species
ANS/Anthocyanidin synthase	AT4G22880	1.14.11.19	comp22553_c0_seq2	1358	MG
			comp46010_c0_seq1	1298	CS
			comp26676_c0_seq1	1467	PD
CHI/Chalcone isomerase	AT3G55120	5.5.1.6	comp11006_c0_seq1	973	MG
			comp51966_c0_seq1	246	MG
			comp53156_c0_seq1	238	MG
			comp29236_c0_seq1	1286	CS
			comp26683_c0_seq1	1149	PD
CHIL/Chalcone-flavanone isomerase	AT5G05270	5.5.16	comp32379_c0_seq1	933	MG
			comp21708_c0_seq1	793	CS
			comp15176_c0_seq2	1011	PD
CHS/Chalcone synthase	AT5G13930	2.3.1.7.4	comp1710_c0_seq1	232	MG
			comp1710_c1_seq1	510	MG
			comp16254_c0_seq1	516	MG
			comp16837_c1_seq1	1561	MG

			comp26818_c0_seq2	1710	MG
			comp29685_c6_seq1	265	MG
			comp48382_c0_seq1	280	MG
			comp49539_c0_seq1	256	MG
			comp54679_c0_seq1	274	MG
			comp55198_c0_seq1	215	MG
			comp21766_c0_seq1	1859	CS
			comp19312_c0_seq1	1305	PD
			comp22421_c1_seq3	2298	PD
			comp23611_c0_seq2	800	PD
			comp23611_c1_seq1	402	PD
			comp23611_c2_seq1	213	PD
DFR/Dihydrokaempferol 4-reductase	AT5G42800	1.1.1.219	comp18420_c0_seq1	274	MG
			comp32899_c0_seq1	1358	MG
			comp49946_c0_seq1	259	MG
			comp34607_c0_seq1	1501	CS
			comp63206_c0_seq1	294	CS
			comp72193_c0_seq1	204	CS
			comp11743_c0_seq2	1618	PD
F3H/Flavanone 3-hydroxylase	AT3G51240	1.14.11.9	comp31032_c0_seq1	1652	MG

			comp25260_c1_seq1	218	CS
			comp39685_c0_seq1	1593	CS
			comp48072_c0_seq1	409	CS
			comp25445_c0_seq1	1562	PD
F3'H/Flavonoid 3'-hydroxylase	AT5G07990	1.14.13.21	comp10092_c0_seq1	505	MG
			comp10314_c0_seq1	398	MG
			comp27205_c2_seq1	1856	MG
			comp53606_c0_seq1	317	MG
			comp48378_c0_seq1	431	CS
			comp58182_c0_seq1	304	CS
			comp24938_c0_seq5	1784	PD
			comp42068_c0_seq1	252	PD
GST26/Glutatione transferase	AT5G17220	2.5.1.18	comp45787_c0_seq	315	MG
			comp27627_c0_seq1	527	CS
			comp23636_c0_seq2	1152	PD
GSTF5/Glutathione transferase	AT1G02940	2.5.1.18	comp9715_c1_seq1	254	MG
GSTF6/Glutathione transferase	AT1G02930	2.5.1.18	comp41619_c1_seq1	852	CS
			comp26912_c0_seq1	750	PD
UDP-Glucoronosyl/Anthocyanidin 5-O-glucosyltransferase	AT4G14090		comp22380_c0_seq1	1776	MG

			comp46668_c0_seq1	1844	CS
			comp27770_c0_seq1	1841	PD
UF3GT/Anthocyanin 3-O-glucoside	AT5G54060	2.6.1.5	comp17955_c0_seq2	1415	MG
			comp36093_c0_seq1	3173	CS
			comp10729_c0_seq1	1645	PD
UGT84A2/Sinapate 1-glucosyltransferase	AT3G21560	2.4.1.120	comp14618_c0_seq1	879	MG
			comp85431_c0_seq1	226	CS
UGT78D2/Anthocyanidin 3-O-glucosyltransferase	AT5G17050	2.4.1.115	comp31910_c0_seq1	1649	MG
			comp32980_c0_seq1	1494	MG
			comp22461_c0_seq1	228	CS
			comp42173_c0_seq2	2110	CS
			comp5856_c0_seq1	1551	PD

Appendix 5.1. Mutagenesis experiment.

I performed a mutagenesis experiment on approximately 10000 seeds of *Polanisia dodecandra* inbreeding line. The procedure was performed in the lab fume hood using a medium two-hand zipper-lock AtmosBag (Sigma). I followed the protocol proposed by Leyser (2000) with the following amendments:

Step 1. I tried different ethyl methanesulfonate (EMS) concentrations: 60mM, 80mM, and 100mM. I used approximately 9900 seeds in total, 1110 seeds per treatment. For each EMS concentration, I left the seeds immersed in the EMS solution during 2, 4 and 6 hours. I performed nine EMS treatments in total.

Step 5. After washing the seeds, they were air dried for two days in petri dishes covered with 3 mm Whatman paper. During the two days, I kept the seeds in the lab fume hood.

Table. Characterization of *Polanisia dodecandra* M1 generation by EMS treatment. Control represents non-mutagenized plants.

Treatment	M1 seeds planted	M1 surviving plants	%Mortality
Control	144	119	17.36
60mM 2hours	1403	810	42.27
4hours	1508	840	44.30
6hours	1207	754	37.53
80mM 2hours	968	617	36.26
4hours	1053	632	39.98
6hours	928	510	45.04
100mM 2hours	1152	622	46.01
4hours	996	614	38.35
6hours	1374	724	47.31

General considerations

Unfortunately, the mutagenesis experiment failed in *P. dodecandra* plants. In theory, when the mutagenesis experiment is successful the germination percentage is close to 75% and albino phenotypes are exhibited by 0.5% of the plants Leyser (2000). In our study, the germination percentage was 57.82%, so the mortality rate was high and I

did not observe any albino phenotypes. Further, I observed just one mutated plant (Figure 1) but this plant was unfertile and the fruits were not produced. I planted the seeds from M1 generation hoping to observe mutations in the M2 generation, but all the seeds were inviable. Perhaps the poor results on the mutagenesis screen were due to EMS concentrations. To perform further mutagenesis experiments in *P. dodecandra*, I recommend starting with a lower EMS concentration between 20 mM and 30 mM and use at least 10000 seeds by treatment when possible.

Figure. *Polanisia dodecandra* mutant phenoptype. The plant does not have petals, the stamens are shorter than usual, the style is abnormally long, and the ovules appear to be external.

Appendix 5.2. Collection records of eight VIGS trials for *Cleome violacea*. Percentage altered phenotype including pale, variegated and strong phenotypes. Number of individual plants treated grouped by treatment and size. 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 greater than 7 true leaves. Modified from Mankowski, 2012.

Treatment	**Plants treated**			**% Phenotype**		
	S	M	L	S	M	L
Untreated	42	44	57	N/A	N/A	N/A
Empty-TRV2	84	129	21	N/A	N/A	N/A
TRV2-*CvPDS*	41	175	114	0	14.28	24.56

Appendix 5.3. Collection records of 4 VIGS trials for *Polanisia dodecandra*. Percentage altered phenotype including pale, variegated and strong phenotypes. Number of individual plants treated grouped by treatment and size. 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 greater than 7 true leaves.

Treatment	**Plants treated**			**% Phenotype**		
	S	M	L	S	M	L
Untreated	116	141	179	N/A	N/A	N/A
Empty-TRV2	107	113	165	N/A	N/A	N/A
TRV2-*PdPDS*	141	170	141	7.80	18.82	9.21

Appendix 5.4. Booster inoculation of large plants of *P. dodecandra* 30 days after vacuum infiltration.

Treatment	Infiltration method	N treated	% Mortality (N)
Control	Vacuum	7	0
Mock-control	Vacuum	2	0
TRV2-*PdPDS*	Vacuum	18	50 (9)

Appendix 5.5. PDS expression in *Cleome violacea*. Each bar represents the relative quantitation (RQ) of PDS expression in experimental groups averaged across three to five biological replicates. ACTIN and RAN4 were used as reference genes, and Control was used as the reference tissue (RQ=1). Expression is upregulated if RQ > 1 and downregulated if RQ < 1. Error bars represent the maximum and minimum RQ values calculated across all biological replicates within a 95% confidence level. Significance between experimental groups (α = 0.01) calculated using multiple two tailed t-tests (a,b).