HARVARD UNIVERSITY
Graduate School of Arts and Sciences

DISSERTATION ACCEPTANCE CERTIFICATE

The undersigned, appointed by the

Department of Organismic and Evolutionary Biology

have examined a dissertation entitled

"A Whole New Whorl: An Exploration of the Morphology, Genetics, and Ecological Function of the Staminodes of *Aquilegia*"

presented by

Clara Meaders

candidate for the degree of Doctor of Philosophy and hereby

certify that it is worthy of acceptance.

Signature _____

Typed name: Professor Elena Kramer

Signature _____

Typed name: Professor Charles Davis

Signature _____

Typed name: Professor Robin Hopkins

Date: April 25, 2018

A Whole New Whorl: An Exploration of the Morphology, Genetics, and Ecological Function of the Staminodes of *Aquilegia*

A dissertation presented
by
Clara Meaders
to
The Department of Organismic and Evolutionary Biology

in partial fulfillment of the requirements
for the degree of
Doctor of Philosophy
in the subject of
Biology

Harvard University
Cambridge, Massachusetts
April 2018

A Whole New Whorl: An Exploration of the Morphology, Genetics, and Ecological Function of the Staminodes of *Aquilegia*

Abstract

The organ identity genes that encode the four organs of the canonical floral body plan are well characterized. In contrast, the downstream developmental pathways involved in programming the morphological differences found in novel organ identities are poorly understood. *Aquilegia* is a genus found in the lower eudicot family Ranunculaceae. Species from this genus have a fifth organ identity: a whorl, or concentric circle, of ten staminodes (sterile stamens) located between the four-seven whorls of fertile stamens and one whorl of five carpels. Staminodes in *Aquilegia* are hypothesized to function in defense, but prior to this thesis, this had never investigated in the field. In this thesis, I have used histological approaches to compare stamens and staminodes across development in *Aquilegia* and sister genera *Semiaquilegia* and *Urophysa*. I identified key differences in *Aquilegia* staminode morphology, including a marginal "curling" phenomenon, asymmetric lignification in the adaxial epidermal layer, and differences in adaxial/abaxial cell number. I then conducted an RNA-sequencing experiment to identify the genetic pathways downstream of organ identity that are responsible for different morphologies of

staminodes and stamen filaments, and used in situ hybridization to more closely examine the expression of candidate genes identified from this analysis. Pathways enriched in staminodes included those involved in lignification, defense, wounding, and secondary cell wall development. Finally, I conducted three field studies, each using a different species of *Aquilegia*, in which I removed staminodes and counted seed sets to investigate whether removal impacted reproductive fitness. Weak decreases in seed set were observed in two species, but these results were inconclusive. These combined, diverse approaches have provided new insights into the evolutionary divergence of staminodes from stamens and facilitated the creation of more focused hypotheses for further investigation.

Contents

Listing of figures

List of Tables

This thesis is dedicated to my K-5th grade students from Science Club for Girls - their curiosity and excitement about science inspired and reinvigorated me each week during my Ph.D.

Acknowledgments

First and foremost, I would like to express my sincerest gratitude and thanks to my Ph.D. advisor, Professor Elena Kramer for all of her teaching and mentorship. Her support, motivation, enthusiasm, knowledge, and feedback were critical throughout my graduate school experience, and in particular with completing my dissertation. I have learned so much about the type of teacher and mentor I aspire to be from her, and am thankful for the example and support she gives to the people in our department and at Harvard.

Besides my advisor, I would like to thank the rest of my thesis committee: Prof. Charles Davis, Prof. Cassandra Extavour, and Prof. Robin Hopkins for their encouragement, insight, and feedback.

I want to thank my labmates (past and present) for their help with my projects. In particular, Molly Edwards and Ya Min for making the Kramer lab an empowering and fun lab complete with 2000's movie nights. A special thank you to the high school and undergraduate students I was able to mentor during my time in the Kramer lab: Gabriel Amador, Kate Freedberg, Olivia Kivel, Imtiyaz Hussain, Fabbiha Hussain, and Tom Saide. Kate contributed to initial histological studies, and my preliminary field work study.

I am also thankful for my graduate school friends, especially my co-Departmental Teaching Fellow Kira for all of our pedagogy focused lunches and discussions. I have learned so much about organismal and evolutionary biology from my peers in the department, and have been grateful to have such enthusiastic scientists and teachers for friends. In the greater Harvard and Boston community I am thankful for the Bok Center for Teaching and Learning, the Life Sciences Outreach community, Science Club for Girls, and the Journal of Emerging Investigators for providing spaces for me to learn and mentor other members of the Boston community.

Last but not least I would like to thank my husband Johnathan Meaders for his support, especially during the writing of this dissertation and my job application process, and our cat Maple for all of her snuggles during graduate school.

Chapter 1

Introduction

One of the core questions in the field of evolutionary developmental biology, or evo-devo, is how novel organs or body plans arise in multicellular organisms and how alterations in gene regulatory networks give rise to these changes. Addressing this question requires considering how we define "novel". Most formal definitions of novelty in an evolutionary context were conceptualized by animal biologists. For the purpose of this thesis, we must first determine whether these definitions are applicable to novelty in plants, which build their bodies in a fundamentally different manner from animals.

1.1 The developmental logic of plants

Plant body plans are completely modular, being comprised of repeatedly produced stems, lateral determinate organs (e.g., leaves), and axillary buds, which are

collectively termed phytomers (Figure 1.1). Modification or transformation of these subunits is the primary driver of plant morphological diversity via the expression of alternate identity programs (Sattler, 1988). For example, axillary buds may express vegetative identity in which they produce only leaves and lateral branches, or they may express reproductive identity to become inflorescence or floral meristems. Floral meristems, in turn, produce a specific set of lateral organ identity programs: sepals, petals, stamens and carpels. These organs are arranged in concentric circles, or whorls, within the flower. Flowers also represent the clearest examples of how morphological novelty in plants is closely linked with homeosis. Homeosis is broadly defined as the transformation of an organ from one segment of an organism into an organ that is typically found in another segment of the organism. The textbook example from animals is the antennapedia mutant in Drosophila melanogaster in which antennae are transformed into legs (Struhl, 1981). While this perspective focuses on the identification of homeotic mutant phenotypes, in plants, homeosis is the organizing principle of plant development, and has played an important role in plant evolution.

1.2 Defining novelty

There are many different definitions of novelty, each of which have various limitations or advantages. Ernst Mayr defined novelty as any "newly acquired structure or property that permits the performance of a new function, which in turn, will open a new adaptive zone" (Mayr, 1960, p 351). There are examples that fit this definition but it fails to include novel traits that do not open up adaptive zones, and

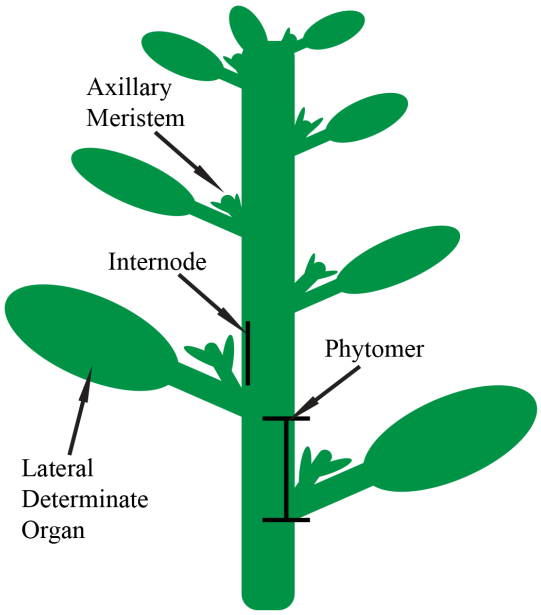


Figure 1.1: Diagram of plant vegetative phytomers. Phytomers are comprised of lateral determinate organs, axillary meristems, and internodes.

ignores traits that open up adaptive zones but are not novel (Pigliucci, 2008). This definition also does not address how the newly acquired traits arose in the first place (Moczek, 2008).

Muller defined novelty as a "qualitatively new structure with a discontinuous origin, marking a relatively abrupt deviation from the ancestral condition" (Muller, 1990, p101). This definition focused on development rather than function, and emphasized that novelties could arise as by-products of other developmental changes - for example, the enlargement of a sesamoid in the giant panda that led to evolution of a "thumb". Muller and Wagner also focused on morphological novelties, and set a narrow definition by requiring that novel structures not be homologous to structures in

ancestral species (Muller & Wagner, 1991). Muller and Newman further made a distinction between novelties and adaptations, making the claim that novelties are not derived from variation in phenotypic precursors while adaptations are improvements on already existing features derived from heritable variation and natural selection.

Unfortunately, these definitions do not lend themselves well to defining novelty in plants, where phytomer-based body plans commonly diversify via modifications of serially homologous organs. For plants, we must expand the definition of novelty to include traits that have homologous ancestral organs or gene network origins, but that have diverged in form. For instance, flowers from the Loasaceae subfamily Loasoideae have elaborate scales and staminodes (non-fertile stamens). Outer scales initiate as 3-5 stamen primordia that later fuse to form an individual staminodial scale, while inner staminodes form outgrowths that require distinct and novel developmental programs (Hufford, 2003). Staminodes of *Loasa* have bright red and yellow coloring, serving distinct roles in attracting pollinators that are not served by the white petals (Weberling, 1989). These novel organs, which are clearly derived from stamens, do not adhere to Muller's requirement for discontinuous origin. Furthermore, these staminodes would not fit Muller and Wagner's strict constraint against homologous structures.

An alternative view is presented by Wagner and Lynch (2010), who argued that there are two basic ways that novelties can arise: differentiation among serially repeated elements and *de novo* origination (Wagner & Lynch, 2010). Both require organ-specific programs of gene expression during development that result in the novelties. The networks themselves may be modified ancestral networks, or

assembled *de novo*. Differentiation requires continuous variation, while *de novo* origination represents discontinuous variation - the novelty is either present or absent (Peterson & Müller, 2016). As plants are comprised of serially repeated elements which typically differentiate via modification and individualization, Wagner and Lynch's definition seems to be the most applicable for defining novelty in plants.

Although I have argued that transformation of pre-existing components of the phytomer is the main mechanism for novelty in plants, there are examples of genuinely *de novo* origination, for instance, the corona. "Coronas" refer to sterile floral organs that typically develop in between the petals and stamens. These organs can be derived from petals, stamens, or are compound in origin. Work in *Narcissus* (daffodil), *Passiflora* (passionflower), Rafflesiaceae, and Asclepiadaceae has shown that coronas have similar expression of organ identity genes relative to stamens, despite having completely distinct morphologies (Waters et al., 2013; Hemingway et al., 2011; Nikolov et al., 2013; T. Livshultz and E. Kramer lab, pers. comm.). These disparate, independently evolved cases all appear to reflect late elaborations of stamen and/or receptacle tissue that may be controlled by genetic pathways acting downstream or in parallel to the canonical floral organ identity program described in the following section. Regardless, this type of genuine *de novo* novelty is relatively rare in plants. In the following sections I will give a broad overview of how transformations of pre-existing components of the phytomer have led to the evolution of flowers, variations in types of floral organs, and diversity within floral organs.

1.3 Flowers as novel structures

The flower itself is a novel structure in the context of the seed plants and comprises several key innovations that are thought to have facilitated the radiation of the angiosperms (Glover et al., 2015). Flowers are bisexual reproductive shoots that appear to have been evolutionarily derived from gymnosperm cones (Theissen & Melzer, 2007). Floral organs, which represent modified lateral determinate organs, are arranged in concentric circles, or whorls, with sterile organs (usually sepals and petals) surrounding the reproductive organs (male stamens and female carpels). Stamens have origins as microsporophylls, with the pollen they produce representing the haploid microgametophyte stage of the lifecycle (Wallace et al., 2011). Carpels are more complex but are likely to represent modified megasporophylls (Mathews & Kramer, 2012; Scutt et al., 2006). In contrast, gymnosperm reproductive shoots are either male or female and have elongated internodes between scales. Modifications to this plan during the evolution of flowers included the compression of the internodes and homeotic shifts in the expression of the male and female organ identity genes, which led to the evolution of the bisexual flower (Theissen & Melzer, 2007). In contrast to the fertile organs, which have clear precursors in the gymnosperms, the evolution of the sterile organs is less straightforward (see below regarding petal evolution). It is likely, however, that evolution of the attractive sterile perianth facilitated attraction of pollinators and thereby helped promote adaptive radiation in the flowering plants. It is critical to note that while the flower and several of its components are considered novel, they all require modification of ancestral

homologous organs.

The definition of the ABC model in the early 1990's, followed by its later expansion to the ABCE model, served to link floral development and genetics, and helped elucidate the origins of the floral organs and the flower itself (reviewed in Coen & Meyerowitz, 1991; Causier et al., 2010). This model depends on four classes of transcription factors - A, B, C, E - most of which function in overlapping domains to produce sepals, petals, stamens, and carpels (Figure 1.2). Altered expression of any gene class results in homeotic transformation of organ identity, and many aspects of this model are highly conserved in angiosperms. In Arabidopsis, *APETALA1* and *APETALA2* comprise the A class; *APETALA3* and *PISTILLATA* comprise the B class; *AGAMOUS* represents the C class; and *SEPALLATA 1-4* make up the E class. All of these transcription factors except for *APETALA2* are representatives of the MIKC type MADS-box family of transcription factors (Theissen, 2001).

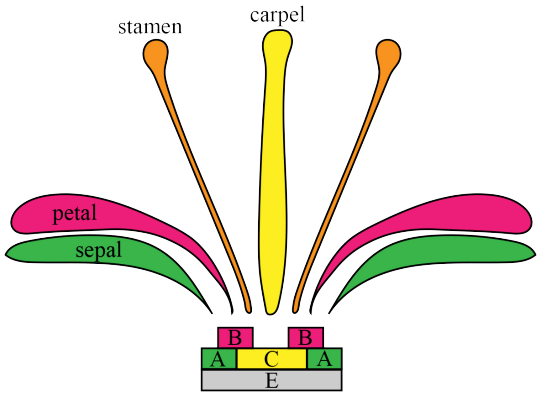


Figure 1.2: The ABC model of floral development. The four organs (sepals, petals, stamens, carpels) are arranged in whorls and patterned by overlapping domains of expression from the A, B, C, and E classes of genes.

The homeotic nature of the ABCE model fits very well with our overall

understanding of plant development. It also suggests that homeotic mutations, or shifts in the expression of these homeotic loci, allowed for evolution of bisexuality and inclusion of both male and female reproductive organs on one axis. Gymnosperm reproductive organs appear to be specified by two classes of genes - the B and C classes. Expression of both B and C results in male identity while expression of C alone results in female identity. Thus, changes in the expression domains of either class could result in bisexual cones (Theissen & Becker, 2004).

Understanding the evolution of sepals and petals has not been as clear cut. Strictly speaking, petals are defined as sterile organs positioned in the second whorl of a flower. It appears that these organs have evolved multiple times independently, either as modifications of stamens (andropetaloidy) or modifications of bract or leaf-like organs (bracteopetaloidy) (Ronse De Craene & Brockington, 2013; Ronse De Craene, 2007). Andropetals are transformed outer stamens that have lost their ability to produce pollen but maintain a single vascular trace and are found in the same parastichies (positional spirals) as stamens. These petals are often delayed in development compared to stamens. Bracteopetals on the other hand, develop before the stamens, are found in the same parastichies as bracts (surrounding leaves), and have more complex vasculature (Ronse De Craene & Brockington, 2013). It remains true that both types of petals are thought to be derived from ancestral lateral organs, but the argument could be made that they played a role in opening up new adaptive zones due to their contribution to pollinator attraction.

1.4 Novel organ identities derived from petal precursors

Not only are petals thought to be been derived many times from different precursors, once they evolved, petals have an immense range of diversity in shape, color, nectar reward, fusion, and symmetry. In some cases, this diversity in morphology reaches such extremes that botanists have considered the organs to have genuinely taken on another identity, instances best exemplified by the lodicule and the labellum. These organs do not resemble typical petals, but are still sterile organs found in the second whorl.

Within the monocot grasses, an organ dubbed the "lodicule" occupies the second whorl but plays no attractive role. Instead, in these wind-pollinated flowers, the lodicule swells late in development in order to allow the stamens to emerge and expose the carpel stigma to wind-borne pollen. Interestingly, B gene homologs control the identity of lodicules, supporting a homologous relationship between the second whorl petals of grass ancestors and the derived lodicule. In this case, it would appear that the B genes are activating a dramatically different developmental program than what is normally observed in second whorl petals (Yoshida, 2012; Whipple et al., 2007). That being said, the resultant grass flower still has only four types of floral organs, similar to most other angiosperms.

The Orchidaceae family members are also found within the monocots, but unlike grasses have floral body plans dedicated to attracting pollinators. Orchids are complex flowers with three types of sterile organs. In the first whorl, there are petaloid sepals, which are often morphologically distinct from the second whorl

sterile organs. Within the second whorl there are two very different types of sterile organs - the canonical petals as well as the labellum, which serves as both a visual attractant and a landing platform for pollinators. The labellum is thought to be derived from a fusion between a second whorl petal and a third whorl stamen. These three types of sterile organs are programmed by four different B gene paralogs that have diverged in expression, with Clades 1 and 2 expressed in the sepals and petals, and Clades 3 and 4 expressed in the petals and at higher levels in the labellum (Mondragón-Palomino, 2013; Mondragon-Palomino & Theißen, 2011). Differential expression of genetic pathways controlling floral symmetry further differentiates the labellum from petals, thereby creating two separate identity programs in the second whorl (Su et al., 2013).

1.5 Novel organ identities derived from stamen precursors

Although stamens themselves are thought to have only evolved once, the large numbers of stamens present in many flowers create a kind of raw material for evolutionary modification and the expression of novel morphologies. Staminodes are stamens that have been modified and are now sterile. This can occur during reduction of the male whorls in the transition to a female-only flower, as is seen in moneocious or dioecious plants. When such reduced staminodes do not evolve a different functional role for the flower, the vestigial organs will likely be lost. If, however, they evolve a new role, they may become highly elaborated and can perform various functions, frequently involving pollinator interactions (Walker-Larsen & Harder, 2000).

The evolution of such staminodes has occurred many times independently across the angiosperms. In magnolid-flowers, staminodes arose independently at least 5 times and function primarily to prevent self-pollination by bending inward to cover the stigma during anthesis (Walker-Larsen & Harder, 2000; Endress, 1984). Roles in visual pollinator attraction via showy structure or with provision of nectar reward are found in the petaloid staminodes of the Eupomatiaceae, Degeneriaceae, and Himantandraceae families (Figure 1.3) (Endress, 1984).

Within the monocots, functional staminodes are found within the Orchidaceae (with roles in pollinator deception) and the Zingiberales (Shi et al., 2009). In the Zingiberales, duplication and subsequent divergence of both B and C class gene homologs appears to be associated with the evolution of petaloid stamens (chimeric functional stamens) and petaloid staminodes (sterile petaloid stamens). The Zingiberales is divided into two groups: the "Ginger Group" and the "Banana Group". The Zingiberaceae, Costaceae, Cannaceae, and Marantaceae all exhibit reduced numbers of stamens in association with the evolution of petaloid staminodes (Kirchoff, 1991). In the Costaceae and Zingiberaceae families, the fusion of multiple petaloid staminodes results in a novel organ, the labellum (Rocha de Almeida et al., 2015a). One B class paralog is expressed in both petals and staminodes (Rocha de Almeida et al., 2015b), but expression of the C class homolog in the staminode and labellum was found in *Costus spicatus*, consistent with their origin from stamen tissue. Staminodes adopted a different role in the Marantaceae, where petaloid staminodes encloses the style while a trigger staminode orients the pollinator. When bees touch the trigger staminode, pollen is released in an explosive fashion

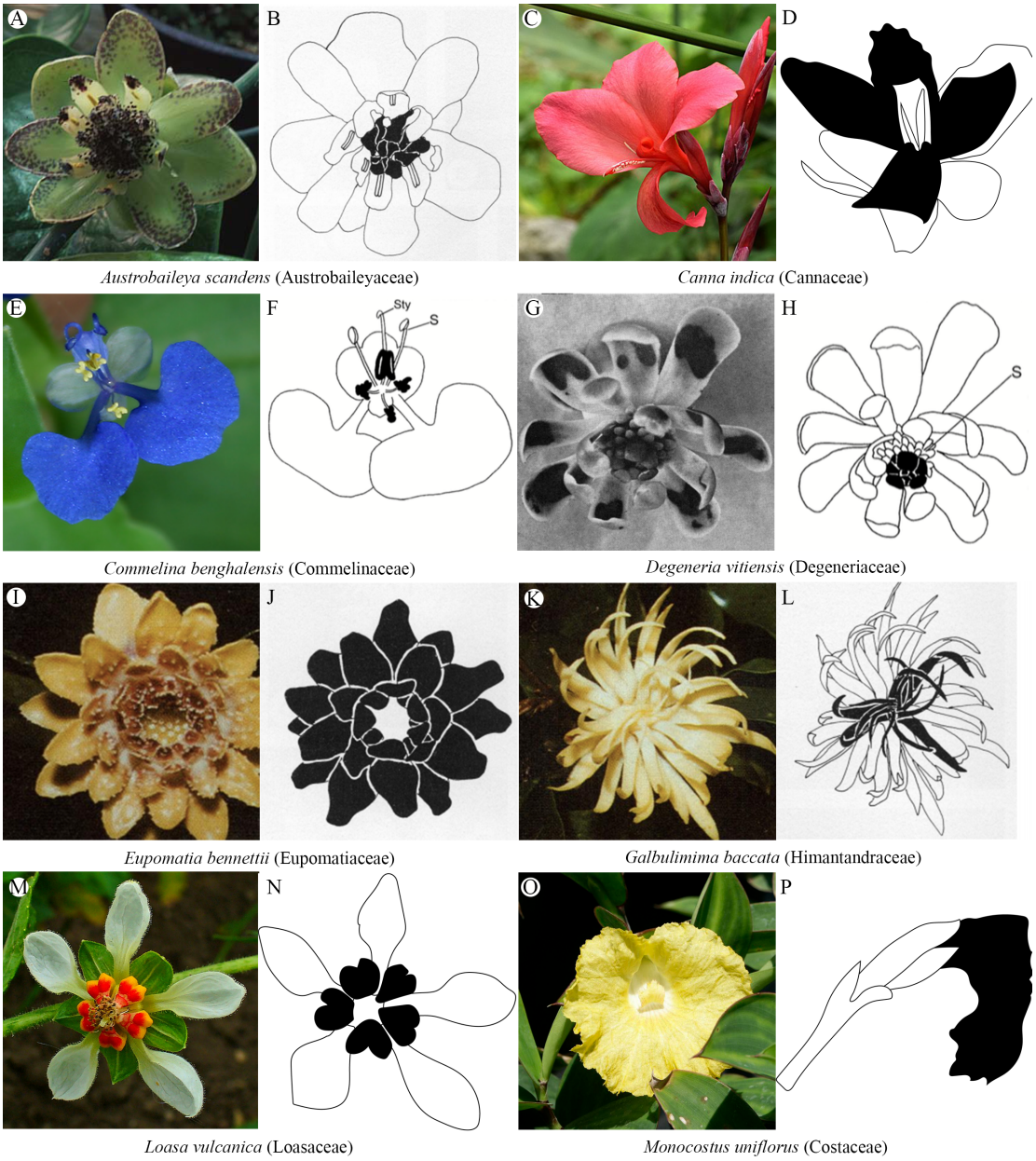


Figure 1.3: Photographs and diagrams of selected functional staminodes. (A-B, G-H) Staminodes of *Austrobaileya scandens* and *Degeneria vitiensis* hide the gynoecium. (E,F) Staminodes of *Commelina benghalensis* provide sterile or false pollen. (C-D, I-P) Staminodes of *Canna indica, Eupomatia bennettii, Galbulimima baccata, Loasa vulcanica* and *Monocostus uniflorus* participate in visual pollinator attraction. S = stamens. Sty = Style. Images compiled from: A (Elena Kramer); B, H, I, J, L (Endress, 1984); G (Thien, 1980); C, E, K, M, O Wikimedia commons; F (Walker-Larsen & Harder, 2000)

(Walker-Larsen & Harder, 2000). Members of the "Banana" group typically do not have staminodes, with the exception of the Heliconiaceae family which has one persistent small staminode in the genus *Heliconia*. The function of this staminode is unknown, but is hypothesized to play a role in guiding the pollinators towards nectar rewards (Kirchoff et al., 2009).

The core eudicots, representing 70% of all angiosperm species, also include numerous examples of staminode evolution. Many instances are associated with the evolution of floral bilateral symmetry (zygomorphy) but others represent derived, novel organ identities. In the hamamelids, staminodes are associated with both nectar production and explosive pollination (Anderson & Hill, 2002). These explosive pollination mechanisms can be quite complex such that after pollinators touch staminodes, the styles are released to contact the pollinator, followed by explosive dehiscence of the stamens to cover the pollinator with pollen. While most staminodes found in the rosids and the asterids are rudimentary and transitional in the context of unisexual flowers or zygomorphy, there are cases where staminodes have evolved roles in pollinator attraction (Linaceae), or where they mimic rewards or nectaries (Endress & Matthews, 2006). Perhaps the best understood in the context of organ identity are the independently derived petaloid staminodes of *Delosperma napiforme*. Petals were lost prior to divergence of the Caryophyllales, and second whorl petaloid organs were subsequently recruited multiple times from either bracts or stamens (Brockington et al., 2009). The inner stamen primordia of *D. napiforme* develop into functional stamens while the outer stamen primordia develop into sterile staminodes. Both stamens and staminodes express B class genes, while expression of the C class is

transient in the staminodes but persistent in stamens (Brockington et al., 2012).

1.6 Staminodes in *Aquilegia*

What is the genetic basis of the evolution of novel floral organ identity programs? I have highlighted a wide range of examples across the angiosperms but, unfortunately, most of these are either quite ancient (e.g., the origin of sepals and petals or the origin of the lodicule) and/or in taxa that are not genetically tractable. The lower eudicot genus *Aquilegia* possesses an entire whorl of sterile staminodes that represent an excellent model for exploring the evolutionary origin of a novel identity program.

Aquilegia flowers have five distinct floral organs: sepals, petals, stamens, staminodes, and carpels. The staminodes are arranged in a continuous whorl immediately adjacent to the carpels. Their primordia strongly resemble those of stamens and are arranged on the same parastichies. These staminodes appear to have evolved in the last common ancestor of *Aquilegia*, *Semiaquilegia*, and *Urophysa*. *Aquilegia* staminodes lack anthers and are laterally expanded to form a ruffled lamina, thereby diverging substantially in morphology from all the other floral organs.

The Kramer lab began to explore the identity and development of *Aquilegia* staminodes by investigating the genetic basis of their identity. In the context of the canonical ABC model, it is difficult to account for the creation of a fifth organ identity that represents an entire whorl (as opposed to the interaction of identity and zygomorphy within four whorls as in orchids). It immediately became apparent that gene duplication may have played a major role in the evolution of the staminode

identity program. Plant gene families are largely conserved, which means that novelties generally arise via a combination of gene duplication and modifications of pre-existing gene networks (Flagel & Wendel, 2009). Networks can be modified by being expressed at different times or locations, or duplicated genes may evolve novel functions.

The duplication-degeneration-complementation (DDC) model outlines three possible outcomes for duplicated genes (Force et al., 1999). Nonfunctionalization (loss of function) can occur from either accumulation of a null mutation in the coding region of a gene copy, or from loss of all of its regulatory regions. Likewise, neofunctionalization (gain of a new function) may involve the evolution of new expression domains and/or changes to coding sequences that affect biochemical function. In this case, both copies will be preserved since one copy is still necessary to confer the original function while the second copy has evolved new roles. Finally, subfunctionalization (in which each copy retains a portion of the ancestral function) can occur if each duplicate acquires degenerative mutations in different regulatory regions. It may be the case that each copy shows differential spatial or temporal expression or that their expression levels are reduced. In such a case, both copies are preserved because they are necessary to fulfill the total ancestral function (Force et al., 1999).

In *Aquilegia*, duplications of homologs of the B gene *APETALA3* and the C gene *AGAMOUS* exhibit evidence of both the DDC model and neofunctionalization. Prior to the diversification of the Ranunculids, duplication events resulted in two *APETALA3* lineages: *AP3-I/II* and *AP3-III*, which was then followed by a tandem

duplication to give rise to the separate *AP3-I* and *AP3-II* lineages (Kramer et al., 2003). These paralogous genes have undergone both sub- and neofunctionalization. It appears that members of the *AP3-III* lineage underwent an early subfunctionalization event to become petal specific, which has been confirmed by broad comparative gene expression and functional studies (Zhang et al., 2013; Sharma et al., 2011). The expression patterns of *AP3-I* and *AP3-II* homologs are quite variable across the Ranunculaceae but in the lineage leading to *Aquilegia*, the respective representatives, *AqAP3-1* and *AqAP3-2* experienced divergence in their function (Sharma & Kramer, 2013; Sharma et al., 2011. Although both paralogs are initially expressed across the stamen and staminode primordia, at the stage when carpels initiate, their expression becomes differentiated such that *AqAP3-2* remains expressed in stamens while *AqAP3-1* becomes concentrated in the staminodes. Virus-induced gene silencing of the loci separately and together demonstrated that they both contribute to stamen identity but *AqAP3-1* is the sole staminode identity gene. Further studies of the two *AGAMOUS* paralogs, *AqAG-1* and *AqAG-2*, have shown that these loci also participate in staminode identity (B. Sharma and E. Kramer, unpublished data). Thus, the fifth organ identity program of the staminodes is due to differential expression of multiple B and C gene paralogs that create a unique developmental program.

1.7 Questions addressed in this thesis

Having now determined the upstream genetic program controlling staminode identity, many questions remain regarding the basis of their developmental elaboration and ecological function. In my second chapter, I compare staminode and

stamen development and morphology within the *Aquilegia* genus and between *Aquilegia*, *Semiaquilegia*, and *Urophysa*. This study allows us to better understand the developmental programs that may be differentially expressed between the two organ types. In my third chapter, I use RNA sequencing to identify the genetic pathways downstream of organ identity that are responsible for the different morphologies of the staminode and stamen filament in *Aquilegia coerulea* 'origami', and use in situ hybridization to more closely examine the expression of two candidate genes identified from the RNA sequencing. In my fourth chapter, I conduct three ecological studies in an attempt to ascertain if the removal of staminodes affects reproductive fitness in three species of *Aquilegia*. These combined, diverse approaches have provided new insights into the evolutionary divergence of staminodes from stamens and facilitated the creation of more focused hypotheses for further investigation.

References

[1] Anderson, G. J. & Hill, J. D. (2002). Many to flower, few to fruit: the reproductive biology of hamamelis virginiana (hamamelidaceae). *American Journal of Botany*, 89(1), 67–78. doi: 10.3732/ajb.89.1.67.

[2] Brockington, S., Alexandre, R., Ramdial, J., Moore, M., Crawley, S., Dhingra, A., Hilu, K., Soltis, D., & Soltis, P. (2009). Phylogeny of the caryophyllales sensu lato: revisiting hypotheses on pollination biology and perianth differentiation in the core caryophyllales. *International Journal of Plant Sciences*, 170, 627–643.

[3] Brockington, S., Rudall, P., Frohlich, M., Oppenheimer, D., Soltis, P., & Soltis, D. (2012). 'living stones' reveal alternative petal identity programs within the core eudicots. *The Plant Journal*, 69, 193–203. doi: 10.1111/j.1365-313X.2011.04797.x.

[4] Causier, B., Schwarz-Sommer, Z., & Davies, B. (2010). Floral organ identity: 20 years of abc. *Seminars in cell and developmental biology*, 21(1), 73–79. doi: 10.1016/j.semcdb.2009.10.005.

[5] Coen, E. & Meyerowitz, E. (1991). The war of the whorls: genetic interactions controlling flower development. *Nature*, 353, 31–37. doi: 10.1038/353031a0.

[6] Endress, P. (1984). The role of inner staminodes in the floral display of some relic magnoliales. *Plant systematics and Evolution*, 146(3-4), 269–282.

[7] Endress, P. & Matthews, M. (2006). Elaborate petals and staminodes in eudicots: Diversity, function, and evolution. *Science Direct Organisms, Diversity, and Evolution*, 6, 257–293. doi: 10.1016/j.ode.2005.09.005.

[8] Flagel, L. & Wendel, J. (2009). Gene duplication and evolutionary novelties in plants. *New Phytologist*, 183(3), 557–564. doi: 10.1111/j.1469-8137.2009.02923.x.

[9] Force, A., Lynch, M., Pickett, F., Amores, A., Yan, Y., & Postlethwait, J. (1999). Preservation of duplicate genes by complementary, degenerative mutations. *Genetics*, 151(4), 1531–1545.

[10] Glover, B., Airoldi, C., Brockington, S., Fernandez-Mazuecos, M., Martinez-Perez, C., Mellers, G., Moyroud, E., & Taylor, L. (2015). How have advances in comparative floral development influenced our understanding of floral evolution? *International Journal of Plant Sciences*, 176(4), 307–323.

[11] Hemingway, C., Christensen, A., & S.T., M. (2011). B- and c-class gene expression during corona development of the blue passionflower (passiflora caerulea, passifloraceae). *American Journal of Botany*, 98(6), 923–934. doi: 10.3732/ajb.1100026.

[12] Hufford, L. (2003). Homology and developmental transformation: Models for the origins of the staminodes of loasaceae subfamily loasoideae. *International Journal of Plant Sciences*, 164(S5), S409–S439. doi: 10.1086/376873.

[13] Kirchoff, B. (1991). Homeosis in the flowers of the zingiberales. *American Journal of Botany*, 78(6), 833–837. doi: 10.2307/2445074.

[14] Kirchoff, B., Lagomarsino, L., Newman, W., Bartlett, M., & Specht, C. (2009). Early floral development of *Heliconia latispatha* (heliconiaceae), a key taxon for understanding the evolution of flower development in the zingiberales. *American Journal of Botany*, 96(3), 580–593. doi: 10.3732/ajb.0800305.

[15] Kramer, E., Di Stilio, V., & Schluter, P. (2003). Complex patterns of gene duplication in the apetala3 and pistillata lineages of the ranunculaceae. *International Journal of Plant Sciences*, 164, 1–11. doi: 10.1086/344694.

[16] Mathews, S. & Kramer, E. (2012). The evolution of reproductive structures in seed plants: a re-examination based on insights from developmental genetics. *New Phytologist*, 194(4), 910–923. doi: 10.1111/j.1469-8137.2012.04091.

[17] Mayr, E. (1960). *The emergence of evolutionary novelties*. Chicago: University of Chicago Press.

[18] Moczek, A. (2008). On the origins of novelty in development and evolution. *BioEssays*, 20, 432–447. doi: 10.1002/bies.20754.

[19] Mondragon-Palomino, M. & Theißen, G. (2011). Conserved differential expression of paralogous *DEFICIENS*- and *GLOBOSA*-like mads-box genes in the flowers of orchidaceae: refining the "orchid code". *The Plant Journal*, 66(6), 1008–1019. doi: 10.1111/j.1365-313X.2011.04560.x.

[20] Mondragón-Palomino, M. (2013). Perspectives on mads-box expression during orchid flower evolution and development. *Frontiers in Plant Science*, 4(377). doi: 10.3389/fpls.2013.00377.

[21] Muller, G. (1990). Developmental mechanisms at the origin of morphological novelty: A side-effect hypothesis. *University of Chicago Press*, (pp. 99–130).

[22] Muller, G. & Wagner, G. (1991). Novelty in evolution: Restructuring the concept. *Annual Review of Ecology and Systematics*, 22, 229–256.

[23] Nikolov, L., Endress, P., Sugumaran, M., Sasirat, S., Vessabutr, S., Kramer, E., & Davis, C. (2013). Developmental origins of the world's largest flowers, rafflesiaceae. *Proceedings of the National Academy of Sciences of the United States of America*, 110(46), 18578–18583.

[24] Peterson, T. & Müller, G. (2016). Phenotypic novelty in evodevo: The distinction between continuous and discontinuous variation and its importance in evolutionary theory. *Evolutionary Biology*, 43, 314–335. doi: 10.1007/s11692-016-9372-9.

[25] Pigliucci, M. (2008). What, if anything, is an evolutionary novelty? *Philosophy of Science*, 75(5), 887–898. doi: https://doi.org/10.1086/594532.

[26] Rocha de Almeida, A., Yockteng, R., Otoni, W., & Specht, C. (2015a). Positive selection on the k domain of the *AGAMOUS* protein in the zingiberales suggests a mechanism for the evolution of androecial morphology. *EvoDevo*, 6(7). doi: 10.1186/s13227-015-0002-x.

[27] Rocha de Almeida, A., Yockteng, R., & Specht, C. (2015b). Evolution of petaloidy in the zingiberales: An assessment of the relationship between ultrastructure and gene expression patterns. *Developmental Dynamics*, 244(9), 1121–1132. doi: 10.1002/dvdy.24280.

[28] Ronse De Craene, L. (2007). Are petals sterile stamens or bracts? the origin and evolution of petals in the core eudicots. *Annals of Botany*, 100(3), 621–630. doi: 10.1093/aob/mcm076.

[29] Ronse De Craene, L. & Brockington, S. (2013). Origin and evolution of petals in angiosperms. *Plant Ecology and Evolution*, 146(1), 5–25. doi: 10.5091/plecevo.2013.738.

[30] Sattler, R. (1988). Homeosis in plants. *American Journal of Botany*, 75(10), 1606–1617. doi: 10.2307/2444710.

[31] Scutt, C., Vinauger-Dourard, M., Fourquin, C., Finet, C., & Dumas, C. (2006). An evolutionary perspective on the regulation of carpel development. *Journal of Experimental Botany*, 57(10), 2143–2152. doi: https://doi.org/10.1093/jxb/erj188.

[32] Sharma, B., Guo, C., Kong, H., & Kramer, E. (2011). Petal-specific subfunctionalization of an apetala3 paralog in the ranunculales and its implications for petal evolution. *New Phytologist*, 191(3), 870–883. doi: 10.1111/j.1469-8137.2011.03744.

[33] Sharma, B. & Kramer, E. (2013). Sub- and neo-functionalization of apetala3 paralogs have contributed to the evolution of novel floral organ identity in aquilegia (columbine, ranunculaceae). *New Phytologist*, 197(3), 949–957. doi: 10.1111/nph.12078.

[34] Shi, L., Luo, Y., Bernhardt, P., Ran, J., Liu, Z., & Zhou, Q. (2009). Pollination by deceit in paphiopedilum barbigerum (orchidaceae): a staminode exploits the innate colour preferences of hoverflies (syrphidae). *Plant Biology*, 11(1), 17–18. doi: 10.1111/j.1438-8677.2008.00120.x.

[35] Struhl, G. (1981). A homeotic mutation transforming leg to antenna in *Drosophila*. *Nature*, 292(12).

[36] Su, C., Chen, W., Lee, A., Chen, C., Chang, Y., Chao, Y., & Shih, M. (2013). A modified abcde model of flowering in orchids based on gene expression profiling studies of the moth orchid *Phalaenopsis Aphrodite*. *PLoS ONE*, 8(11). doi: 10.1371/journal.pone.0080462.

[37] Theissen, G. (2001). Development of floral organ identity: stories from the mads house. *Current Opinion in Plant Biology*, 4, 75–85.

[38] Theissen, G. & Becker, A. (2004). Gymnosperm orthologues of class b floral homeotic genes and their impact on understanding flower origin. *Critical Reviews in Plant Sciences*, 23(2), 129–148. doi: 10.1080/07352680490433240.

[39] Theissen, G. & Melzer, R. (2007). Molecular mechanisms underlying origin and diversification of the angiosperm flower. *Annals of Botany*, 100(3), 603–619. doi: 10.1093/aob/mcm143.

[40] Thien, L. (1980). Patterns of pollination in the primitive angiosperms. *Biotropica*, 12(1), 1–13. doi: 10.2307/2387768.

[41] Wagner, G. & Lynch, V. (2010). Evolutionary novelties. *Current Biology*, 20(2), R48–R52. doi: 10.1016/j.cub.2009.11.010.

[42] Walker-Larsen, J. & Harder, L. (2000). The evolution of staminodes in angiosperms: patterns of stamen reduction, loss, and functional re-invention. *American Journal of Botany*, 87(10).

[43] Wallace, S., Fleming, A., Wellman, C., & Beerling, D. (2011). Evolutionary development of the plant and spore wall. *AoB Plants*. doi: 10.1093/aobpla/plr027.

[44] Waters, M., Tiley, A., Kramer, E., Meerow, A., Langdale, J., & Scotland, R. (2013). The corona of the daffodil narcissus bulbocodium shares stamen-like identity and is distinct from the orthodox floral whorls. *The Plant journal : for cell and molecular biology*, 74(6), 615–625. doi: 10.1111/tpj.12150.

[45] Weberling, F. (1989). *Morphology of flowers and inflorescences*. Cambridge University Press.

[46] Whipple, C., Zanis, M., Kellogg, E., & Schmidt, R. (2007). Conservation of b class gene expression in the second whorl of a basal grass and outgroups links the origin of lodicules and petals. *Proceedings of the National Academy of Sciences of the United States of America*, 104(3), 1081–1086.

[47] Yoshida, Y. (2012). Is the lodicule a petal: Molecular evidence? *Plant Science*, 184, 121–128. doi: 10.1016/j.plantsci.2011.12.016.

[48] Zhang, R., Guo, C., Zhang, W., Wang, P., Li, L., Duan, X., Du, Q., Zhao, L., Shan, H., Hodges, S., Kramer, E., Ren, Y., & Kong, H. (2013). Disruption of the petal identity gene *APETALA3-3* is highly correlated with loss of petals within the buttercup family (ranunculaceae). *Proceedings of the National Academy of Sciences*, 110(13), 5074–5079. doi: 10.1073/pnas.1219690110.

Chapter 2

Morphology and development of staminodes in *Aquilegia* and closely related genera

2.1 Introduction

2.1.1 What are staminodes?

Stamens are the male reproductive organs of the angiosperms, and are homologous to the microsporophylls of male reproductive cones in gymnosperms (D'Arcy & Keating, 1996). In the flowering plants these organs are typically comprised of a sterile basal filament, which provides support and transmits water and nutrients, and a fertile upper anther, which houses the microsporangia. The microsporangia will undergo meiosis to produce microspores and then divide by mitosis to produce the haploid vegetative and generative cells in pollen grains (Scott et al., 2004).

As reproductive organs, stamens are defined by their function. However, in

some cases, stamens become infertile and lose their reproductive capabilities. These non-functional stamens are broadly called “staminodia”, or “staminodes”. According to this botanical definition, staminodes may be vestigial, transitional, or specialized (De Craene & Smets, 2001). Some confusion arises, however, from the fact that the term “staminode” is broadly applied to many different types of sterilized stamens. For example, this broad definition can include the aborted stamens of female unisexual flowers (De Craene & Smets, 2001), or could be applied to the transitional structures that are positioned between the fully functional stamens and elaborated petals in flowers that exhibit stamen-to-petal transformational series (De Craene, 2003). I will follow the precedent set by Walker-Larsen and Harder and refer only to staminodes that exist in hermaphroditic flowers (Walker-Larsen & Harder, 2000).

Staminodes have evolved repeatedly across many different angiosperm lineages, occurring in at least one species of one third of angiosperm families, and in over half of angiosperm genera (Walker-Larsen & Harder, 2000). It is most common for staminodes to evolve during periods of evolutionary reduction of the androecium. A full whorl of stamens might lose its reproductive capacity in actinomorphic flowers, or individual stamens might lose their reproductive capacity while others retain function within the same whorl of zygomorphic flowers. Staminodes and stamens share early developmental patterns, but differentiate along distinct trajectories during later stages of development. Building floral organs requires an investment in resources, and nonfunctional staminodes that are vestigial in nature are likely to be reduced in form and eventually lost. However, there are examples of staminodes that undergo the transition from stamen to infertile staminodes but then adopt new

functional roles for the flower.

Most relevant to my work are those flowers in which persistent staminodes co-exist with fertile stamens and are relatively elaborate organs, rather than simply aborted stamens. These staminodes are commonly implicated in roles involving pollinator attraction, provision of attractants/rewards, avoidance of self-pollination, or facilitation of pollen removal and receipt (Endress & Matthews, 2006). Elaboration of staminodes is widespread among the Orchidaceae and within the Zingiberales, where they have roles ranging from mimicking pollinator breeding sites, to serving as triggers for explosive pollination (Walker-Larsen & Harder, 2000). Such staminodes represent a discrete fifth organ identity, making them difficult to resolve within the context of the traditional ABC model, which only provides a mechanism for four organ types. Moreover, they are often recently evolved and, thereby, offer a genetically tractable model for the evolution of novel organ identity programs.

2.1.2 Staminodes in *Aquilegia*

Aquilegia is a genus of 70 species found within the basal eudicot family Ranunculaceae. These species are diverse in the geography and ecology of their habitats, their morphology (including flower orientation, petal spur length, and color), are interfertile, and have undergone a recent adaptive radiation (Hodges & Arnold, 1994). The staminodes found in *Aquilegia* are an example of recently evolved organs that differ from the common pattern of evolution of staminodes involving sterilization of the flower. While one whorl of stamens has been sterilized forming the staminodes, *Aquilegia* flowers have between 40-80 stamens in 2-7 whorls. This number of

stamens is higher than that found in closely related genera without staminodes, indicating that pollen reduction is unlikely to be a factor (Table 2.1). Furthermore, the staminodes of *Aquilegia* have been described as a continuous whorl of laterally expanded organs that undergo late congenital fusion to form a cylindrical sheath that remains attached to the receptacle long after the outer floral organs abscise (Sharma et al., 2014) (Figure 2.1 A,B). Although they are sterile, the staminodes display a morphology distinct from that of stamens.

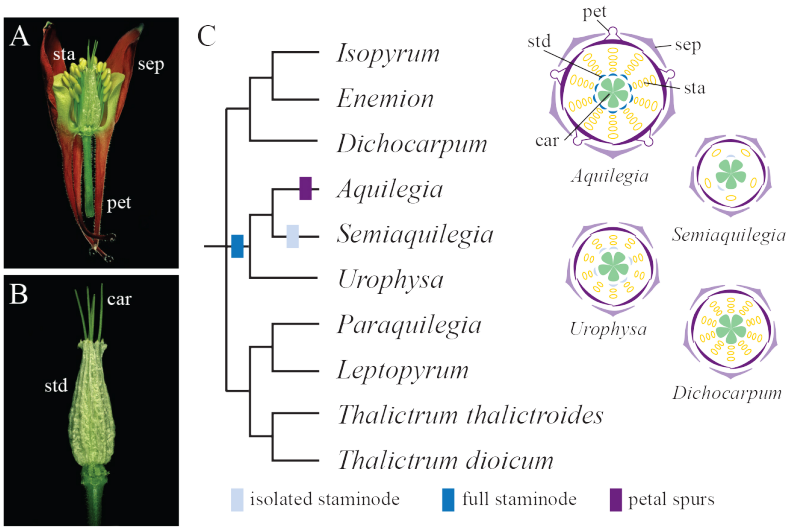


Figure 2.1: Staminodes in the Thalictroideae. (A) *Aquilegia* flower with partially removed sepals (sep), petals (pet), and stamens (sta). (B) *Aquilegia* flower with sepals, petals, stamens removed. Staminodes (std) form a sheath surrounding the carpels (car). (C) Phylogeny of the Thalictroideae subfamily of the Ranunculacae with floral diagrams of selected genera.

Along with nine other genera, the genus is a member of the monophyletic subfamily Thalictroideae, which can be further broken down into smaller clades. The monophyletic clade comprised of *Aquilegia*, *Semiaquilegia*, and *Urophysa* contains the only three genera within the Thalictroideae to have staminodes. This clade

diverged from their closest related clade of *Isopyrum*, *Enemion*, and *Dichocarpum* between 20 and 22 million years ago (mya) (Bastida et al., 2010). This pushes the date of origin of staminodes back from what was predicted in previous studies to 20 mya. Previously staminodes were thought to have become sterilized in the last common ancestor of *Aquilegia* + *Semiaquilegia* 8 mya, and elaborated in *Aquilegia* 6 mya (Sharma et al., 2014). The identification of staminodes in *Urophysa* makes it more likely that staminodes evolved in the last common ancestor of *Aquilegia* + *Semiaquilegia* + *Urophysa*, and were partially lost in *Semiaquilegia* along with the reduction in overall size in the flower.

Table 2.1: Staminodes in members of the Thalictroideae

Genus	**# of Species**	**# of stamens**	**Orthostichies**	**Staminodes (present/absent)**	**Reference**
Isopyrum	6	∼10 +	13	Absent	Rendle, 1925; Drummond & Hutchinson; 1920
Enemion	7	∼10-15 in 2-3 whorls	11	Absent	Tucker & Hodges, 2005
Dichocarpum	18	∼25 in spiral whorls	irregular	Absent	Ren et al., 2011
Aquilega	70	∼40-80 in 2-7 whorls	10	Present	Tucker & Hodges, 2005; Ren et al., 2011
Semiaquilegia	1	∼10 in 2 whorls	10	Present (variable, in inner whorl)	Tucker & Hodges, 2005
Urophysa	2	∼40 in 4 whorls	10	Present	Zhao et al., 2016
Paraquilegia	4	numerous		Absent	Fu & Robinson
Leptopyrum	1	10-15		Absent	Fu & Robinson
Thalictrum	157	∼ 34-40 in spiral whorls	irregular	Absent	Ren et al., 2011

Floral development has been described in detail for several members of the Thalictroideae (Tepfer, 1953; Endress, 1995; Tucker & Hodges, 2005; Ren et al., 2011; Zhao et al., 2016) (Table 2.1). Staminodes are irregular in number and rudimentary in form in *Semiaquilegia*, making it difficult to identify the organs early in development. Flowers from *Aquilegia* and *Urophysa* have 10 orthostichies (vertical

ranks) of stamens that initiate in whorls of five, alternating initiation with either the sepals or petals at their base (Figure 2.1). The most apical primordia in each orthostichy are smaller, similar in form to stamens, but ultimately become the staminodes, indicating that staminodes have a serially homologous relationship to stamens. All but one member of the genus have a complete whorl of staminodes. The species that has lost staminodes in conjunction with a severe reduction in floral size, *A. jonesii*, is the earliest diverging species within the North American clade. This species separated approximately 2.99 million years ago, soon after *Aquilegia* moved from Eurasia to North America (Fior et al., 2013).

In this chapter, I have conducted an in depth histological study comparing stamen and staminode morphology across selected species of *Aquilegia* as well as the sister genera *Semiaquilegia* and *Urophysa*. This analysis has yielded a number of surprising findings: 1) the two whorls of *Aquilegia* staminodes display subtle morphological distinctions in how they interact with one another along their margins, 2) these interactions appear to be directly related to late stage adhesion that develops throughout the staminodial whorl of most *Aquilegia* species, 3) the staminodes are dead at maturity with clear abaxial/adaxial differentiation in terms of cell shape and number, and 4) the adaxial epidermal cells display asymmetric patterns of lignification. These findings have implications for both the establishment of identity in the staminodial whorl and the potential ecological function of the organs.

2.2 Methods

2.2.1 Plant growth conditions and treatment

Aquilegia coerulea 'origami red and white' seeds (obtained from Swallowtail Garden Seeds, Santa Rosa, CA) were sown and grown in a growth chamber set to 18° C, 60% relative humidity, 1 μMOL of light with 16 hour days and 0 PPM of CO2. Seedlings were transplanted to individual pots after appearance of cotyledons. After 5 true leaves emerged, plants were transferred to a vernalization chamber set to 5° C, 40% relative humidity, 0 μMOL of light, and 0 PPM of CO2. After 4-5 weeks in vernalization, plants were moved to a growth chamber set to 20° C, 60% relative humidity, 2 μMOL of light, and 0 PPM of CO2. Buds and flowers were collected at varying stages of development.

Semiaquilegia buds and flowers were obtained from our collaborator Scott Hodges at the University of CA, Santa Barbara. *A. alpina* and *A. flabellata* buds and flowers were obtained from Russell's Garden Center, Wayland, MA. *Urophysa* seed was obtained from Drs. Hongzhi Kong and Rui Zhang, Chinese Academy of Science, Beijing, China, and grown under the same conditions as described above for *Aquilegia*.

2.2.2 Tissue Processing

Vacuum Infiltration

Buds and flowers were placed in 50 mL falcon tubes with 50 mL of formalin-acetic acid-alcohol (FAA): 25 mL ethanol (EtOH), 5 mL formalin, 2.5 mL acetic acid

(AcOH), and 17.5 mL water (H2O). Tissue was vacuum infiltrated for 2 x 15 min and then left on a shaker overnight at 4° C.

Dehydration

FAA solution was switched for solutions of 4° C 50% EtOH, 70% EtOH, 85% EtOH, 95% EtOH, and 100% EtOH, respectively, for 90-120 minutes each on a shaker at 4° C. Fresh 100% EtOH was added before leaving tissue on a shaker overnight at 4° C.

Infiltration and Embedding

Buds and flowers were placed in a solution of room temperature 100% EtOH on a shaker also at room temperature for 1 hour. The solution was changed to 50% EtOH and 50% Citrasolv and shaken at room temperature for 2 hours. The solution was changed to 100% Citrasolv and shaken at room temperature for another 2 hours. Plant material was then moved to scintillation vials and tissue was covered with fresh 100% Citrasolv. The vials were filled to the top with Paraplast Plus Tissue Embedding Medium chips (McCormick Scientific). The vials were placed in an oven set to 62° C. Over the next 2-4 days, 2-3 changes of melted Paraplast were made per day. Tissue was embedded in melted Paraplast using a TECTM Tissue Embedding Center (Triangle Biomedical Sciences).

2.2.3 Histology

Sectioning Tissue

Tissue was sectioned into 8 µm-thick sections using a Jung Histocut and placed on microscope slides on a slide warmer set to 42°C for 4 hours. Slides were left overnight in Drierite. Slides were then treated using the appropriate staining techniques outlined below.

Toluidine Blue

Staining protocol was adapted from both Sass's Safranin and Fast Green method, as outlined in Ruzin (1999) and Biogenic Colombia's "Toluidine Blue Staining of Paraffin Sections Used on Plant Sections-March 2013" method. Biogenic Colombia's method was modified from IHC World's "Toluidine Blue Staining Protocol for Mast Cells". All solution volumes in this protocol consisted of 300 µl. In step 1, slides were placed in a slide rack and immersed in a box of 100% Citrasolv for 10 minutes. In step 2, the slides were moved to a second box of 100% Citrasolv for another 10 minutes. In step 3, slides were moved to a box of 100% EtOH for 1 minute. In step 4, slides were moved to a second box of 100% EtOH for 1 minute. Steps 5-9 consisted of a hydration series, where slides were moved between boxes of 95%, 85%, 70%, 50%, 30% EtOH for 30 sec each. For Step 10, slides were moved to a box of 150 mM NaCl for 30 sec. In step 11, the slide rack was placed in a solution of 0.5 mg/mL toluidine blue O in H2O for 30 sec. In step 12, slides were washed several times in H2O until no more dye was visible. In step 13, slides were placed in

95% EtOH for 10 seconds. In step 14, slides were placed in 100% EtOH for 10 seconds. In step 15, slides were transferred into another box of 100% EtOH, and were incubated for 10 seconds. In step 16, slides were cleared in 100% Citrasolv for 2 minutes. In step 17, coverslips were mounted onto the slides using 75 µL of Permount. Finally, slides were scanned on the Zeiss Axio Scan.Z1 slide scanner.

2.2.4 Phloroglucinol

Phloroglucinol staining was performed according to the method outlined in Ruzin, 1999.

2.2.5 Scanning Electron Microscopy

Flowers were dissected and preserved in 100% EtOH, then dehydrated in a graded ethanol series. Samples were critical point-dried, mounted on metal stubs, and examined using a JEOL JSM-6010 LC Scanning Electron Microscope.

2.2.6 ImageJ

The cell counter plugin of imageJ was used to count individual cells from histological sections (Rueden et al., 2017).

2.2.7 Histone *in situ* hybridization

Inflorescences were collected and prepared as outlined in “infiltration and embedding”, and tissue was sectioned as outlined in “sectioning tissue”. Probe for the *HISTONEH4* locus *(AqHis4)* probe was synthesized as previously described in

Puzey et al., 2012. RNA in situ hybridization was performed according to previously described methods (Kramer, 2005). All sections were digitally photographed using a Zeiss Axio Scan.Z1 slide scanner (Harvard Imaging Center).

2.2.8 RT-qPCR

Total RNA of mature staminodes and stamens was extracted using PureLink RNA Reagent (Invitrogen) and treated with Turbo DNAse (Ambion). cDNA was synthesized from 1 μg of RNA using Superscript III reverse transcriptase (Invitrogen) and oligo (dT) primers. The resulting cDNA was diluted 1:4 as a template. Brilliant SYBR Green QPCR Master Mix, Low Rox (Stratagene) was used to carry out the qRT-PCR reactions in a Statagene Mx3005P qPCR system. Primers used are outlined in Table 2.2. Isopentenyl pyrophosphate:dimethylallyl pyrophosphate *(AqIPP2)* was used for normalization as it has previously been shown to have little quantitative transcriptional variation across tissues (Sharma et al., 2011). Primer efficiencies were evaluated before expression analysis through a six-dilution series. Expression for each tissue was assayed from three technical replicates per reaction plate. Relative gene expression levels were calculated using the 2-ΔΔCT method described in Livak and Schmittgen (2001), and formulas were adjusted with the obtained primer efficiencies for *AqIPP2*, *Aqcoe2G28520*, *Aqcoe2G018600*, *Aqcoe6G254700*, and *Aqcoe7G218800* . Paired student's t-tests were used to determine the statistical significance between stamen and staminodes values.

Table 2.2: Primers for qRT-PCR

Gene	**Primers**
Aqcoe2G28520	Forward: GGTGCCACAGTTAACTTGATT Reverse: AACATGTCTCCTCCAATATGC
Aqcoe2G018600	Forward: ACTTCTTCTATCGGTGCAGTC Reverse: CTGCTACGGTTTTTCCATAAC
Aqcoe6G254700	Forward: GTTTACTGCACATGGTGGTGG Reverse: AAAAGCCATATCAGCTCTGTC
Aqcoe7G218800	Forward: CCCCTACCTTTAAACACATTG Reverse: GAATGTTAAGTTGCTCCCTTG
AqIPP2	Forward: CAGGTGAAGACGGACTGAAGTTAT Reverse: CCAAGACTGGAAAAAAGACCACAC

2.3 Results

2.3.1 Developmental stages

Development of *Aquilegia* flowers begins with initiation of the floral meristem and concludes with the opening of the follicle fruits and subsequent release of seeds. Previous work in the Kramer lab described 16 stages of development including features unique to *Aquilegia* floral development (Min & Kramer, 2016). For instance, the sepals and petals both change in their coloration from green to pink, the petal spurs grow out from the petal, and nectaries develop in the distal tips of the spurs (Figure 2.2). I added descriptions of staminode development to the *Aquilegia* developmental stages. Scanning electron microscopy revealed that there are two distinct whorls of 5 staminodes, an outer and an inner whorl (Figure 2.3). The ten staminodes are clearly differentiated by stage 8, but by the end of stage 13, they appear to be fused into a sheath and are difficult to remove individually. After this stage, the staminodes appeared rugose or ruffled and margins of individual organs could not be differentiated by the naked eye.

Figure 2.2: Late stages of floral development in *Aquilegia*. Stages correspond to stages detailed in Table 2.3.

Table 2.3: Late stages of floral development in *Aquilegia*

Stage	Landmark Event at Beginning of Stage
11a	Carpels close, spurs not yet visible between the sepals (Figure 2.2 A,B)
11b	All organs increase in size; petal spurs become visible between the sepals (Figure 2.2 C,D)
11b'	Petal spur continues to elongate, flower bud begins to develop color (Figure 2.2 E-H) *new stage
11c	Full color, spur elongation reaches 50% of final length (Figure 2.2 I,J)
11d	Staminodes appear fused, adaxial anthers extend above carpels (Figure 2.2 K,L)
12	All organs reach full length (Figure 2.2 M,N)
13a	Sepals reflex (Figure 2.2 O, P)
13b	All stamens upright, full pollen dehiscence (Figure 2.2 Q,R)
14	Post-anthesis flower in which all the organs except the staminodes and carpels have abscised from the receptacle
15	Carpels enlarge, staminodes abscise
16	Mature fruits

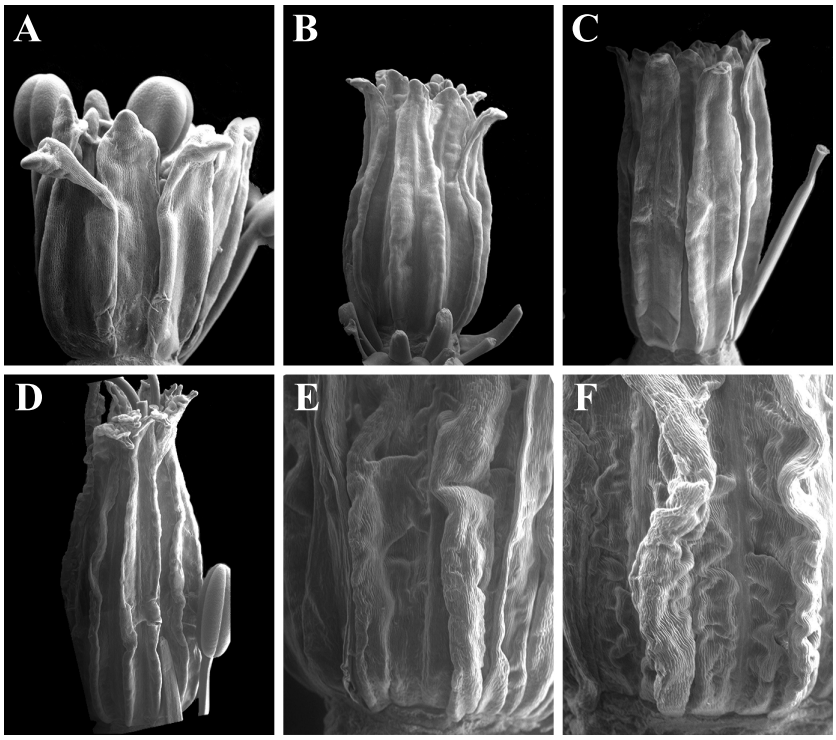


Figure 2.3: Scanning electron microscopy of various stages of *Aquilegia* development.

2.3.2 Histology

In order to closely examine the two whorls of staminodes, I used transverse sections of *Aquilegia coerulea* 'origami', sectioning from the base of the receptacle up

through the apex. I looked at flowers at three stages of development, Stage 11a, Stage 11b', and 13b (bud, pre-flower-opening, and post-flower opening), using three different stain treatments: toluidine blue, safranin and fast green (not included), and calcofluor white (not included). Immediately I noticed a new phenomenon that differentiates staminodes from stamens: the lateral margins of the outer whorl of staminodes curl around the lateral margins of the inner whorl of staminodes, which themselves curl in the opposite direction, thereby creating interlocking margins. This creates a "hand-holding" effect, which is seen at all stages of staminodes that we examined and is consistent within a flower from base to tip (Figure 2.4). The staminodes are unfused at their apices at all developmental stages.

Staminodes, like stamen filaments, have one vascular bundle. Sections of the staminodes adjacent to the receptacle show that they resemble stamens in that they are filled with mesophyll cells. However, moving apically from the base, I observed that there is little mesophyll present so the organs are entirely composed of only two epidermal cell layers.

The staminodes from stage 11a buds are curled around one another at their lateral margins, but also have clearly defined gaps between each organ. Staminodes at these stages were easy to separate individually when manually dissected. As the flowers age, it becomes more difficult to remove individual staminodes, which appear to adhere to one another. Consistent with this, the histological sections from these later stages show that the staminode margins are tightly appressed, forming a continuous sheath that lacks observable gaps between neighboring staminodes.

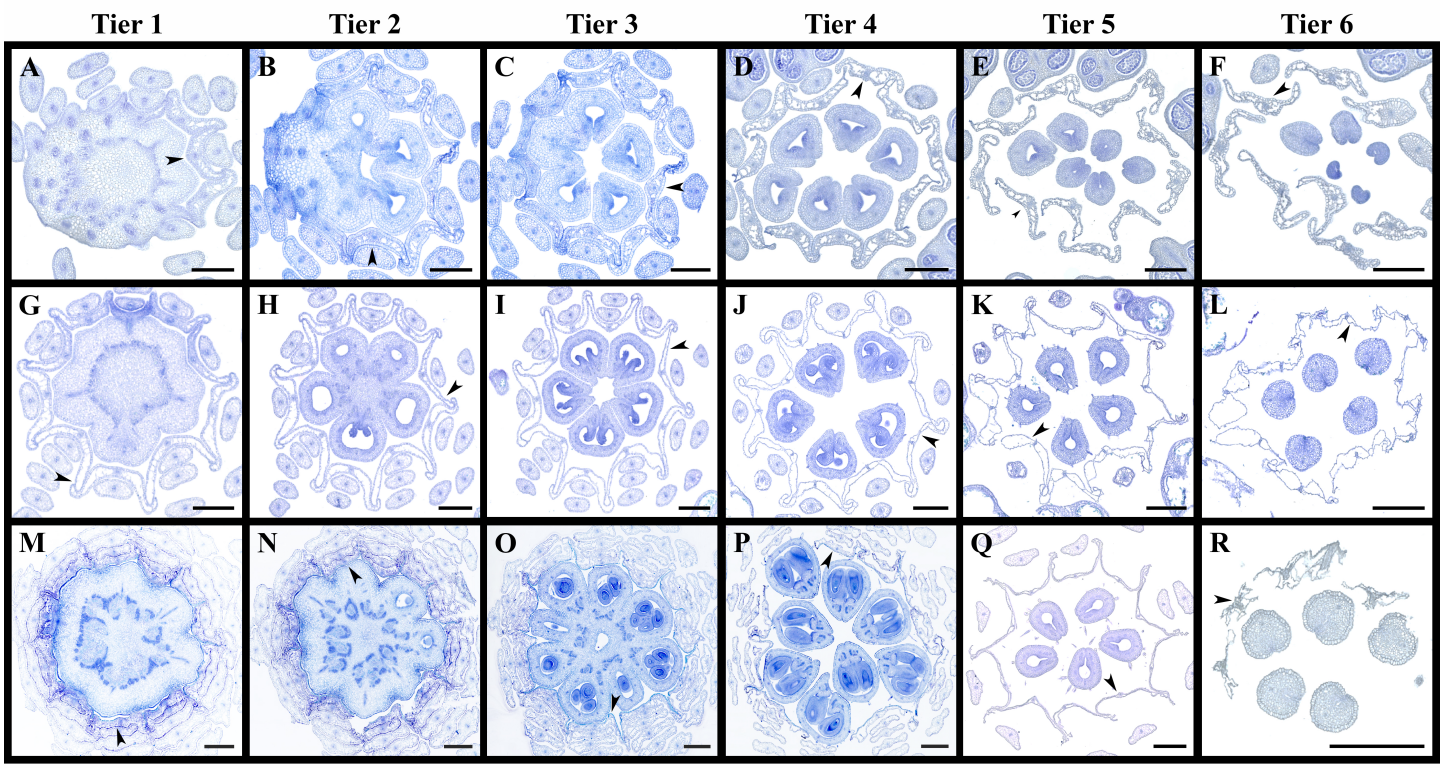


Figure 2.4: Histology of *Aquilegia* using toluidine blue. (A-F) Small flowers, stage 11a/b. Scale bars = 200 µm. (G-L) Medium flowers, stage 11b'. Scale bars = 300 µm. (M-R) Large flowers, stage 13b. Scale bars = 500 µm. Arrows indicate a staminode. Tiers correspond to locations ranging from base to apex.

2.3.3 Lignification

Toluidine blue was used for the sections shown in Figure 2.4. Toluidine blue stains bright blue in the presence of lignin, as shown in the xylem of the vascular tissue of stamens and staminodes (Figure 2.5 A,D). At high magnification, the individual staminodes from stage 13b display lignin enrichment on the adaxial walls of the adaxial epidermis. This staining was not found in the stamens, and also appeared reduced in the cells at the lateral margins of the staminodes that participate in adhesion (Figure 2.5 C).

In order to confirm *in vivo* lignin enrichment in the staminodes compared to stamens, phloroglucinol was used to stain fresh plant tissue. In Figure 2.6, there is strong pink staining in the staminodes, while the stamen filaments are not stained at all.

After lignification was demonstrated in fresh and fixed tissue, quantitative RT-PCR was used to examine expression of four genes known to have roles in the lignin biosynthesis pathway. Aqcoe7G218800 encodes a shikimate O-hydroxycinnamoyltransferase, an enzyme in the phenylpropanoid biosynthesis pathway (Hoffmann et al., 2004). Phenylpropanoids are precursors of lignin. Aqcoe2G018600 encodes a cinnamoyl-CoA reductase, which is the enzyme that catalyzes the first specific step in the synthesis of lignin monomers (Lacombe et al., 1997). Aqcoe2G285200 encodes a 3-O-methyltransferase which is also found in the lignin biosynthesis pathway (Vanholme et al., 2010). Aqcoe6G254700 encodes ferulate-5-hydroxylase, which is a member of the cytochrome P450-dependent monooxygenase family and is also found in the phenylpropanoid pathway. All of

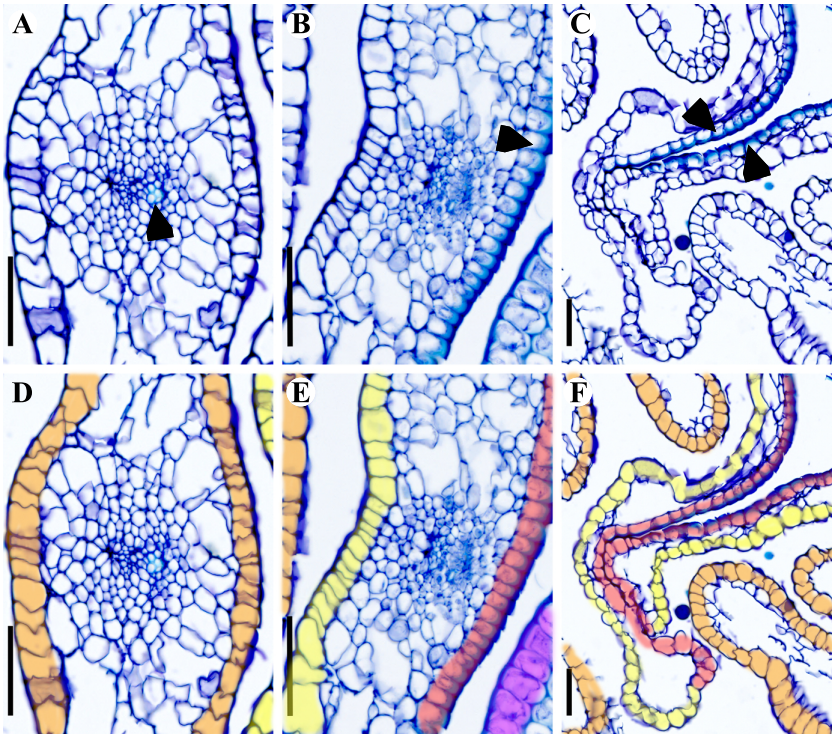


Figure 2.5: Toluidine blue staining. (A) Magnified central region of stage 13 stamen filament. Arrowhead indicates light blue stain of lignified xylem. (B) Magnified central region of stage 13 staminode. Arrowhead indicates asymmetrically lignified adaxial epidermis of staminode. (C) Magnified lateral margin region of stage 13 staminodes. Arrowheads indicate asymmetrically lignified adaxial epidermis of two adjacent staminodes. Note that asymmetric lignification disappears where the lateral margins interact. (D-F) False colored versions of images in A-C. Orange indicates stamen epidermis; yellow indicates the abaxial epidermis of staminodes; pink indicates the adaxial epidermis of staminodes; and purple indicates adjacent carpel epidermis. Scale bars: 50 µm.

these loci genes except Aqcoe6G254700 showed significantly enriched expression in staminodes compared to stamens (Figure 2.7).

2.3.4 Comparative analysis of related species and genera

To complement the initial study of the predominantly New World *A. coerulea* 'origami', I collected multiple developmental stages of flowers from both Old World (*A. alpina*, *A. flabellata*, and *A. vulgaris*) and New World (*A. canadensis*) species, as

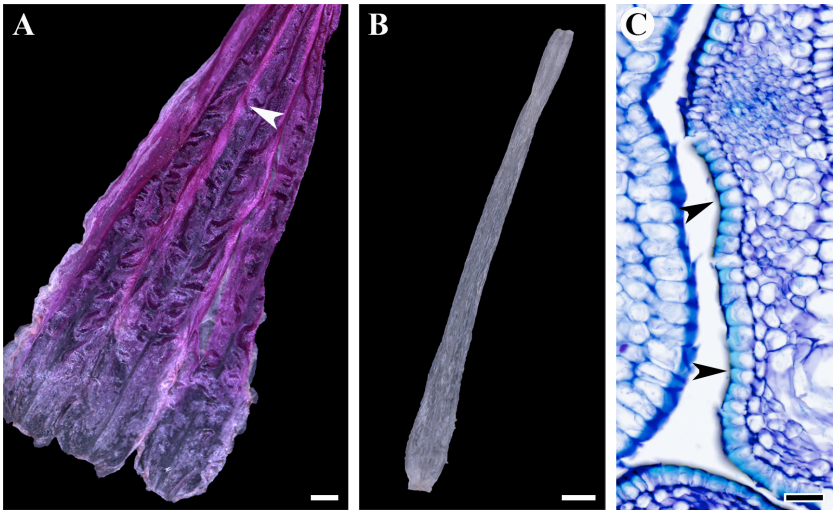


Figure 2.6: Lignin enrichment in staminodes. (A, C) vs. stamen filaments (B) as seen using phloroglucinol (A-B) and toluidine blue (C). Scale bar (A-B) 500 µm. Scale bar (C) 25 µm.

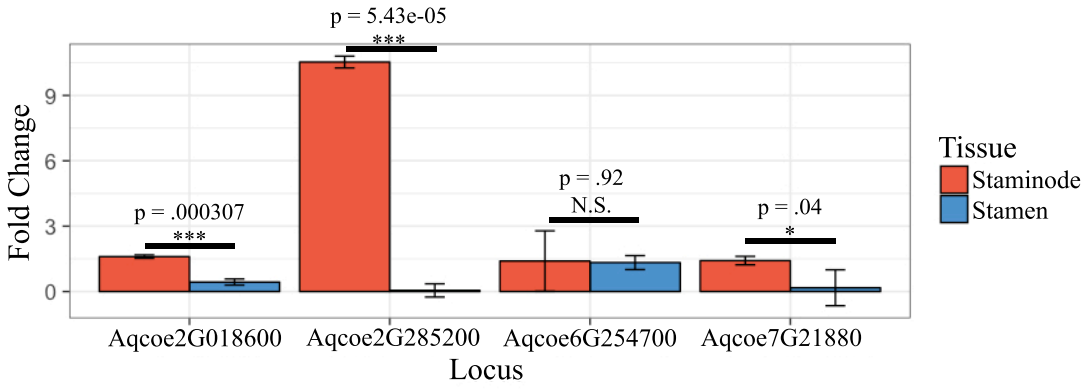


Figure 2.7: Expression of lignin synthesis-related genes in staminodes and stamens. Expression of each gene was determined by qRT in mature organs. Staminodes showed significantly higher levels of expression.

well as from the sister genera *Semiaquilegia* and *Urophysa*. At least 5 flowers were examined per species from different individual plants, with the exception of *Urophysa*, for which only 3 flowers were sectioned due to limited plant material. I found that all of the staminodes from *Aquilegia* species as well as the *Urophysa*

displayed the alternating curling morphology (Figure 2.9). The staminodes from *Semiaquilegia* were dispersed within a whorl of stamens and were generally not curled, but were identifiable due to their two cell layers and flattened morphology. However, a few individual staminodes in *A. alpina*, *A. vulgaris*, and *A. canadensis* displayed aspects of both types of curling (Figure 2.8) such that one side of the staminode curled up, while the other side curled down. In every instance, there were two adjacent staminodes displaying this "50-50" morphology, which preserved the orientations of each of the other staminodes in the sheath.

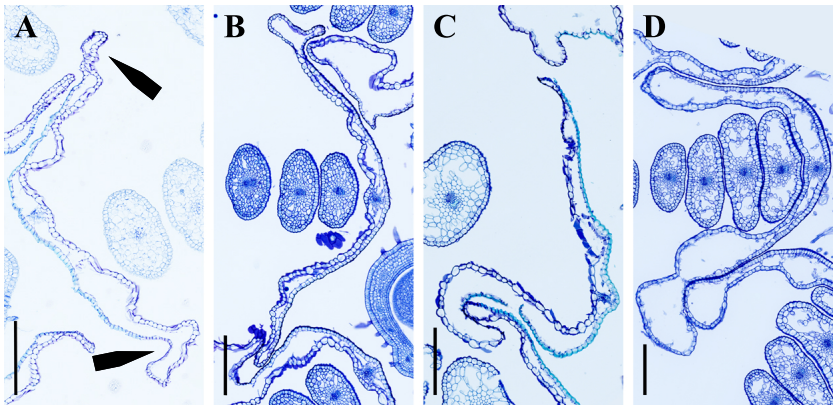


Figure 2.8: Evidence of marginal curling and adaxial lignification. (A, B) *A. alpina*. Arrowheads indicate different patterns of marginal curling. (C) *A. canadensis*. (D) *A. vulgaris*. Scale bar: 200 µm.

In every *Aquilegia* species, the staminodes displayed bright blue staining on the adaxial surface of their adaxial epidermal cells, representing lignin enrichment. This lignification was not seen in late stages of *Semiaquilegia* or *Urophysa*, indicating that lignin enrichment is unique to *Aquilegia*. I also identified gaps between adjacent late stage staminodes in *A. alpina*, *A. vulgaris*, and a few *A. canadensis*, suggesting that lateral adhesion may not be uniform across the genus (Figure 2.9).

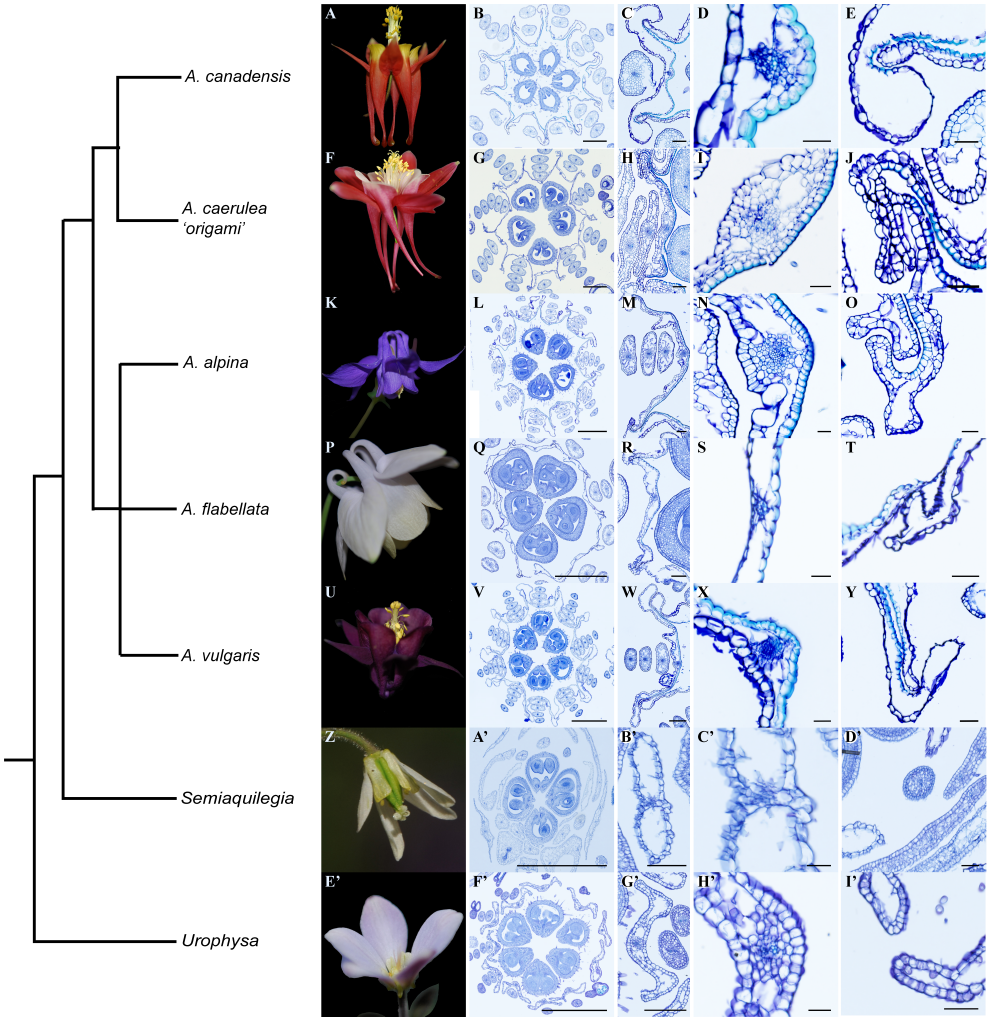


Figure 2.9: Histology of *Aquilegia, Semiaquilegia*, and *Urophysa* using toluidine blue. (A-E) *Aquilegia canadensis*. (F-J) *Aquilegia coerulea* 'origami'. (K-O) *Aquilegia alpina*. (P-T) *Aquilegia flabellata*. (U-Y) *Aquilegia vulgaris*. (Z-D') *Semiaquilegia*. (E'-I') *Urophysa*. From left to right, the first column shows photos of mature flowers. The second column contains relatively low magnification transverse sections of the androecium, staminodial whorl, and gynoecium, positioned roughly midway between the receptacal and staminode apices. The third column shows one individual staminode and adjacent stamen filaments. The fourth column focuses on the center of one staminode, showing the vascular bundle and surrounding abaxial and adaxial epidermal cells. The *Aquilegia* adaxial epidermal cells show adaxial lignification. The fifth column shows the margins of two adjacent staminodes. Scale bars: B, G, L, Q, V, A', F' = 1000 µm. C, H, M, R, W, B', G' = 100 µm. D, I, N, S, X, C', H' = 25 µm. E, J, O, T, Y, D', I' = 50 µm.

2.3.5 Abaxial/Adaxial differences

Staminodes have broad laminae while stamens are composed of narrow filaments. This difference could be due to transverse differences in cell number, size, or both. *In situ* hybridization was used to visualize patterns of cell division as indicated by *AqHISTONE4* expression on stage 6 and stage 11b flowers. By this stage of development, very few dividing cells were apparent in the stamen filaments while there was considerable more *AqHIS4* expressing cells in staminodes (Figure 2.10). Thus, it appears that the laminar expansion of staminodes is due, at least in part, to a longer period of cellular division.

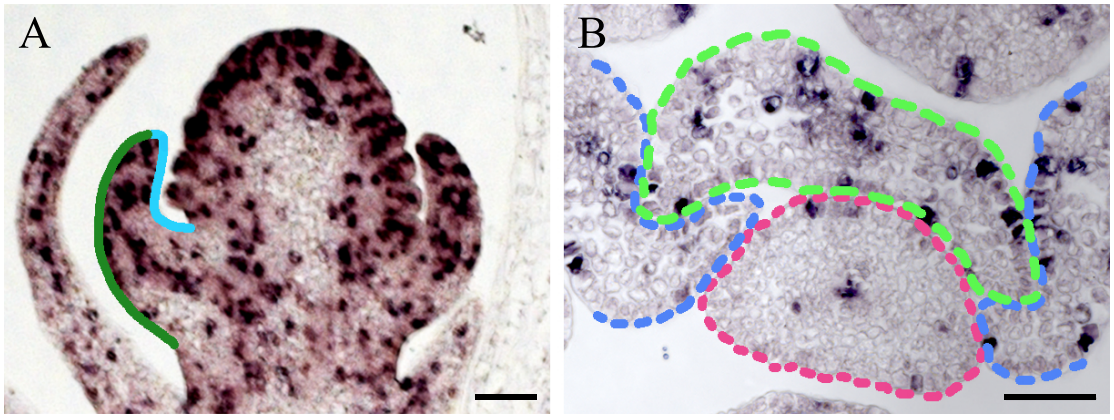


Figure 2.10: In situ hybridization of *Aquilegia coerulea* 'origami'. (A) Histone staining of a stage 6 flower. False coloring outlines the sepals. Blue indicates adaxial epidermis. Green indicates abaxial epidermis. (B) Histone staining of a stage 11b flower. Staining is seen in the stamen vasculature, but less in the epidermal layer than in the epidermal layers of the staminodes. False coloring (blue, green) outlines individual staminodes and stamen filament (pink). Scale bars: 50 µm.

The rugose surface of the staminodes, as well as the curling behavior of the lateral margins could be influenced by differences in cell shape or number between the abaxial and adaxial epidermal layers. ImageJ was used to count the number of cells present in transverse transects of the adaxial and abaxial layers in both stamen

filaments and staminodes from mature flowers. I found that there are no significant differences in cell number between the abaxial and adaxial surfaces of stamen filaments, but there are significantly more cells across the adaxial surface of staminodes (Figure 2.11). In addition, light microscopy was used to visualize the shapes of cells in the adaxial and abaxial epidermal layers (Figure 2.12). This revealed dramatic differences such that the abaxial cells are anisotropic while the adaxial cells are comparatively isotropic. These differences in cell number and cell shape, combined with the differential lignification of the adaxial epidermal cells, are likely to contribute to the rugose surface of the staminodes.

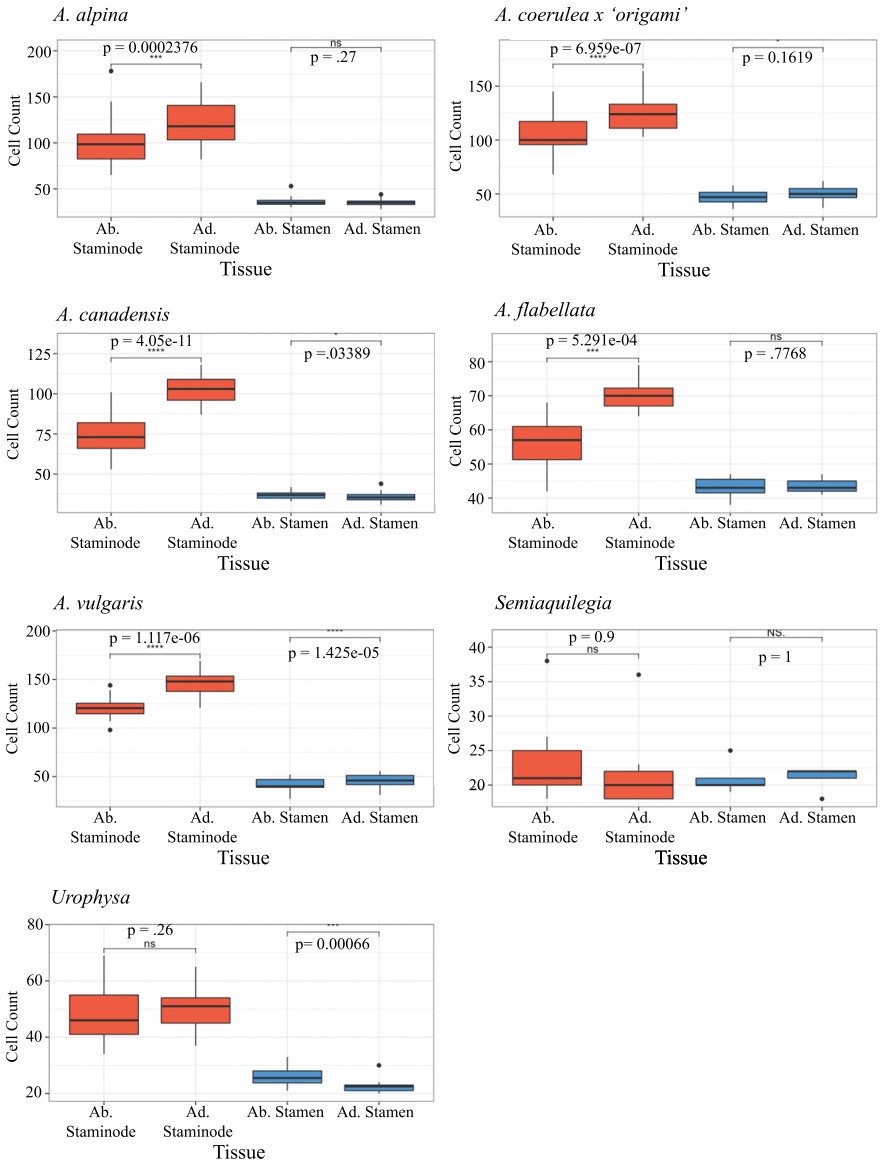


Figure 2.11: Boxplots showing the number of cells present in transverse transects of the abaxial or adaxial epidermal surfaces of staminodes (red) and adjacent stamen filaments (blue). Transects were positioned roughly midway between the receptacle and staminode apices. A total of at least 20 staminodes and their adjacent stamen filaments from 4-5 individual flowers were counted. P values were calculated from student's t-tests.

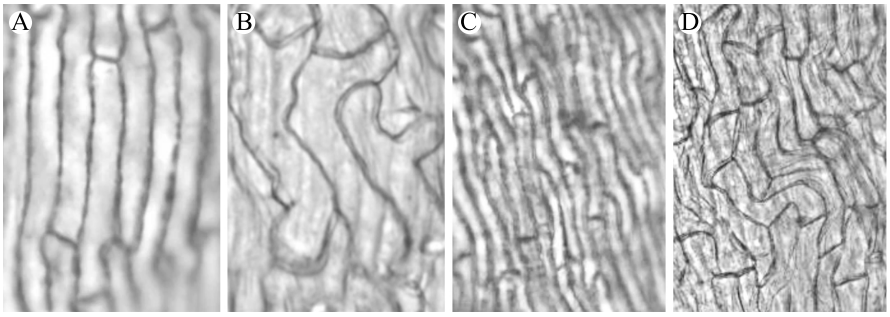


Figure 2.12: Light microscopy of abaxial/adaxial epidermal surfaces from two *Aquilegia* staminodes. (A), (C) Adaxial surface. (B), (D) Abaxial surface.

2.4 Discussion

This study is the first full characterization of the developmental and morphological differences between stamen filaments and staminodes. In addition to the differences in fertility and laminar expansion, I observed a suite of subtle and dramatic differences that may provide insight into the ecological function of *Aquilegia* staminodes.

2.4.1 Lateral interactions

Staminodes are arranged in two distinct whorls that diverge in their morphology due to relatively consistent patterns of physical interactions along their lateral margins, a phenomenon we term "curling". Outer staminodes tend to curl inward towards the carpels while the inner staminode whorl curl slightly outward towards the stamen whorl (Figure 2.13). This creates a physical interlocking. Neither whorl of staminodes displayed regularity of curling orientation in relation to orthostichies with the sepals or petals (data not shown). There is no evidence for

distinct organ identities among these two whorls, but it remains unknown how these developmental differences are controlled.

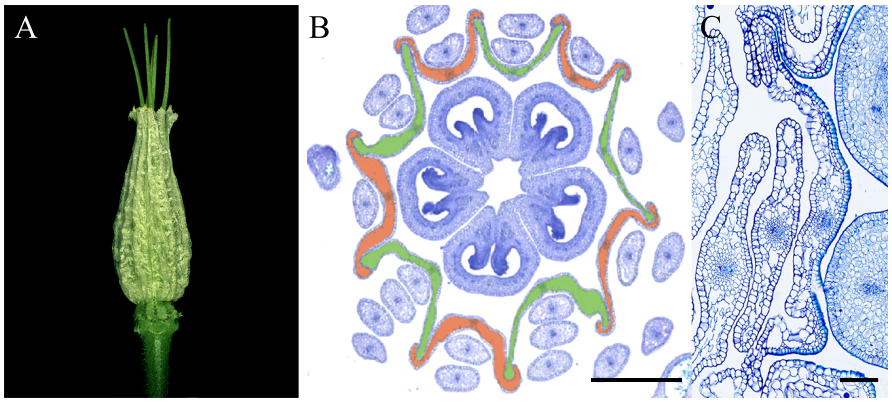


Figure 2.13: Summary. (A) *Aquilegia* flower with sepals, petals, stamens removed. Staminodes (std) form a sheath surrounding the carpels (car). (B) Transverse section from a stage 11c flower. False coloring of every other staminode. Green shows staminodes curling away from carpels. Orange shows staminodes curling in towards carpels. (C) A stage 13 staminode with asymmetric lignification. Scale bars: B: 500 µm, C: 50 µm.

It is also still an open question if the curling interactions are pre-programmed (i.e. identity) or if they are induced by their physical interactions. Removal of one staminode early in development did not appear to disrupt the later curling of adjacent staminodes, but I could only remove staminodes at stages after curling had initiated. Furthermore, I could not rule out whether the curling of adjacent staminodes when freed from their neighbors was an artifact of the fixation process (data not shown). Regardless, the two different directions of curling likely involve differential gene expression downstream of organ identity.

Many *Aquilegia* species exhibit lateral adhesion at late developmental stages. Although this was only observed for 2 out of 5 species in the current study, observations in the field suggest that the feature is quite common in other species

(pers comm S. Hodges). Further study involving broader sampling of New World and Old World species is necessary to understand the exact nature of the adhesion and its evolutionary pattern.

In my histological sections I did not see evidence of epidermal cell lysis or cell-level fusion. The adherence observed could be due to cuticle-mediated interactions. Several genes in Arabidopsis that are involved in cutin biosynthesis have been shown to be essential in preventing organ fusion - for example *LACERATA (LCR), BODYGUARD (BDG), ATP-BINDING CASSETE PROTEINs G11 and G13*, and *HOTHEAD (HTH)* (Jakobson et al., 2016; Akiba et al., 2014; Bird et al., 2007; Kurdyukov et al., Luo et al., 2007; 2006; Wellesen et al., 2001). Mutations in any of these genes impacts cuticle development and also results in adhesion between lateral organs (Ingram & Nawrath, 2017). Transmission electron microscopy of adjacent staminodes as well as stamen filament epidermal layers would allow a comparison of cuticle deposition and cell wall structure.

2.4.2 Abaxial/adaxial differentiation

Lignin is a secondary compound that provides rigidity to cell walls. In the vasculature lignin helps xylem cells withstand negative pressure from transpiration, but outside of the vasculature, lignin is used for structural support as well as herbivore or pathogen resistance (Barcelo, 1997. Lignin deposition can be induced by stress, and the compound can provide physical barriers against pathogen colonization (Miedes et al., 2014). To my knowledge, the only system in which asymmetric lignin deposition has been characterized in detail is in *Cardamine hirsuta*, where the cell

wall thickenings facilitate explosive seed dispersal (Monniaux & Hay, 2016; Hofhuis et al., 2016). Obviously *Aquilegia* staminodes do not have explosive dehiscence, but the lignification could contribute to the rigidity of the adaxial surface compared to the abaxial surface and could also be involved with pathogen resistance.

The mechanisms of polar lignification both within a cell and on selected epidermal surfaces remain an unanswered question. Much of what is known about the establishment of general cellular polarity has been through studies of apical/basal embryogenesis and auxin PIN protein polarity (Dettmer & Friml, 2011). In addition to apical basal polarity, epidermal cells establish an "outside-inside" polarity in order to properly position cells for gas exchange and cuticle for limiting water loss and protection from pathogens. Little is known about the factors involved in establishing cellular polarity beyond PIN polarity, or factors involved in polar lignin deposition.

Lignin monomer biosynthetic gene expression responds to biotic and abiotic stresses, and could be impacted by the differential patterns of stress experienced by cells on the abaxial and adaxial surfaces due to differences in cell number and shape leading to surface buckling (Barros et al., 2015). Within cells, the lignin synthesis enzymes are located within the cytoplasm or on the outer surface of the endoplasmic reticulum, and transport of lignin to the cell walls is thought to occur by passive diffusion, vesicle-mediated exocytosis, and active ATP-dependent transport (Barros et al., 2015). The mechanisms of transport have yet to be determined (Vanholme et al., 2010). Asymmetric transport or expression of enzymes involved in activating lignin monomers or in extending lignin polymers could result in asymmetric lignification within a cell.

In addition to differences in lignification between the adaxial and abaxial epidermal cells, I observed differences in cell shape and cell number between each surface. In plants, growth of lateral organs occurs in two phases: an initial period of cell proliferation, and a later period of cell expansion (Johnson & Lenhard, 2011. A longer period of cell proliferation in staminodes contributes to their broader organ size relative to the the ancestral filaments. Within the staminodes of *Aquilegia*, I hypothesize that differences in cell number and cell shape likely contribute to their their rugose lamina.

2.4.3 Evolutionary and functional implications

Evolution of staminodes always begins with sterilization of anthers and loss of functional pollen production. After this initial sterilization, the non-functional staminodes may either be lost quickly, or may evolve novel morphologies and functions (Walker-Larsen & Harder, 2000). Previously described roles of staminodes include pollinator attraction, provision of attractants/rewards, avoidance of self-pollination, or facilitation of pollen removal and receipt (Walker-Larsen & Harder, 2000).

Sterile organs evolved in the ancestor of *Aquilegia*, *Semiaquilegia*, and *Urophysa* with laminar expansion and disorganized curling. The organs appear to have been reduced in *Semiaquilegia* but elaborated in *Aquilegia* with more consistent curling patterns, adhesion, and asymmetric lignification.

The staminodes found in *Aquilegia* have distinct patterns of developmental elaboration that seem unlikely to be neutrally evolved, but we cannot rule out this

possibility. As the sepals and petals are both colorful, and the petals provide a nectar reward, pollinator attraction is an unlikely function for these staminodes. Furthermore, by stage 12 the innermost anthers extend above the staminodes, almost as far as the stigmas. This makes avoidance of self-pollination or facilitation of pollen removal equally unlikely functions for the staminodes. By stage 14, all organs except staminodes have fully absiced from the flower. The staminodes detach from the receptacle but remain surrounding the carpels due to their lateral adherence until the expanding fruits swell and tear them apart. Based on this, it has been hypothesized that the role of staminodes in *Aquilegia* is to protect the early developing fruits (Brayshaw, 1989), which would be consistent with their rugose surface, lateral adhesion, and lignification. It will be interesting to determine whether these traits hold across a more in-depth sampling of *Aquilegia* species.

References

[49] Akiba, T., Hibara, K., Kimura, F., Tsuda, K., Shibata, K., Ishibashi, M., Moriya, C., Nakagawa, K., Kurata, N., Itoh, J., & Ito, Y. (2014). Organ fusion and defective shoot development in *oni3* mutants of rice. *Plant Cell Physiology*, 55(1), 42–51. doi: 10.1093/pcp/pct154.

[50] Barcelo, A. (1997). Lignification in plant cell walls. *International Review of Cytology*, 176.

[51] Barros, J., Serk, H., Granlund, I., & Pesquet, E. (2015). The cell biology of lignification in higher plants. *Annals of Botany*, 115(7), 1053–1074. doi/: 10.1093/aob/mcv046.

[52] Bastida, J., Alcantara, J., Rey, P., Vargas, P., & Herrera, C. (2010). Extended phylogeny of aquilegia: the biogeographical and ecological patterns of two simultaneous but contrasting radiations. *Plant Systems Evolution*, 284, 171–185.

[53] Bird, D., Beisson, F., Brigham, A., Shin, J., Greer, S., Jetter, R., Kunst, L., Wu, X., Yephremov, A., & Samuels, L. (2007). Characterization of arabidopsis abcg11/wbc11, an atp binding cassette (abc) transporter that is required for cuticular lipid secretion. *The Plant Journal*, 52, 485–498.

[54] Brayshaw, C. (1989). Buttercups, waterlilies, and their relatives in british colombia. *The Plant Cell*, (pp. 1–253).

[55] D'Arcy, W. & Keating, R. (1996). *The Anther: Form, Function, and Phylogeny*. Cambridge University Press.

[56] De Craene, L. (2003). The evolutionary significance of homeosis in flowers: A morphological perspective. *International Journal of Plant Sciences*, 164(S5), S225–S235. doi: 10.1086/376878.

[57] De Craene, L. & Smets, E. (2001). Staminodes: Their morphological and evolutionary significance. *The Botanical Review*, 67(3), 351–402.

[58] Dettmer, J. & Friml, J. (2011). Cell polarity in plants: when two do the same, it is not the same... *Current Opinion in Cell Biology*, 23(6), 686–696. doi: 10.1016/j.ceb.2011.09.006.

[59] Drummond, J. & Hutchinson, J. (1920). A revision of isopyrum (ranunculaceae) and its nearer allies. *Bulletin of Miscellaneous Information (Royal Botanic Gardens, Kew)*, 5, 145–169. doi: 10.2307/4107428.

[60] Endress, P. (1995). Floral structure and evolution in ranunculanae. *Plant Systematics and Evolution*, 9, 47–61.

[61] Endress, P. & Matthews, M. (2006). Elaborate petals and staminodes in eudicots: Diversity, function, and evolution. *Science Direct Organisms, Diversity, and Evolution*, 6, 257–293. doi: 10.1016/j.ode.2005.09.005.

[62] Fior, S., Li, M., Oxelman, B., Hodges, S., Ometto, L., & Varotto, C. (2013). Spatiotemporal reconstruction of the aquilegia rapid radiation through next-generation sequencing of rapidly evolving cpdna regions. *New Phytologist*, 198(2), 579–592. doi: 10.1111/nph.12163.

[63] Fu, D. & Robinson, O. (2001). Leptopyrum. *Flora of China*, 6(276).

[64] Hodges, S. & Arnold, M. (1994). Columbines: a geographically widespread species flock. *Proceedings of the National Academy of Sciences of the United States of America*, 91(11).

[65] Hoffmann, L., Besseau, S., Geoffroy, P., Ritzenthaler, C., Meyer, D., Lapierre, C., Pollet, B., & Legrand, M. (2004). Silencing of hydroxycinnamoyl-coenzyme a shikimate/quinate hydroxycinnamoyltransferase affects phenylpropanoid biosynthesis. *The Plant Cell*, 16(6), 1446–1465. doi: 10.1105/tpc.020297.

[66] Hofhuis, H., Moulton, D., Lessinnes, T., Routier-Kierzkowska, A., Bomphrey, R., Mosca, G., Reinhardt, H., Sarchet, P., Gan, X., Tsiantis, M., Ventikos, Y., Walker, S., Goriely, A., Smith, R., & Hay, A. (2016). Morphomechanical innovation drives explosive seed dispersal. *Cell*, 166(1), 222–233. doi: 0.1016/j.cell.2016.05.002.

[67] Ingram, G. & Nawrath, C. (2017). The roles of the cuticle in plant development: organ adhesions and beyond. *Journal of Experimental Botany*, 68(19), 5307–5321. doi: 10.1093/jxb/erx313.

[68] Jakobson, L., Lindgren, L., Verdier, G., Laanemets, K., Brosché, M., Beisson, F., & Kollist, H. (2016). Bodyguard is required for the biosynthesis of cutin in arabidopsis. *New Phytologist*, 211, 614–626.

[69] Johnson, K. & Lenhard, M. (2011). Genetic control of plant organ growth. *New Phytologist*, 191, 319–333. doi: 10.1111/j.1469-8137.2011.03737.x.

[70] Kramer, E. (2005). Methods for studying the evolution of plant reproductive structures: Comparative gene expression techniques. *Methods in Enzymology*, 395, 617–636. doi: 10.1016/S0076-6879(05)95032-5.

[71] Kurdyukov, S., Faust, A., Nawrath, C., Bär, S., Voisin, D., Franke, R., Schreiber, L., Saedler, H., JP., M., & Yephremov, A. (2006). The epidermis-specific extracellular bodyguard controls cuticle development and morphogenesis in *Arabidopsis*. *The Plant Cell*, 18, 321–339.

[72] Lacombe, E., Hawkins, S., Van Doorsselaere, J., Piquemal, J., Goffner, D., Poeydomenge, O., Boudet, A., & Grima-Pettenati, J. (1997). Cinnamoyl coa reductase, the first committed enzyme of the lignin branch biosynthetic pathway: cloning, expression and phylogenetic relationships. *The Plant Journal*, 11, 429–441. doi: 10.1046/j.1365-313X.1997.11030429.x.

[73] Luo, B., Xue, X., Hu, W., Wang, L., & Chen, X. (2007). An abc transporter gene of *Arabidopsis thaliana, AtWBC11*, is involved in cuticle development and prevention of organ fusion. *Plant and Cell Physiology*, 48, 1790–1802.

[74] Miedes, E., Vanholme, R., Boerjan, W., & Molina, A. (2014). The role of the secondary cell wall in plant resistance to pathogens. *Frontiers in Plant Science*, 5(358). doi: 10.3389/fpls.2014.00358.

[75] Min, Y. & Kramer, E. (2016). The aquilegia *JAGGED* homolog promotes proliferation of adaxial cell types in both leaves and stems. *New Phytologist*, 216, 536–548. doi: 10.1111/nph.14282.

[76] Monniaux, M. & Hay, A. (2016). Cells, walls, and endless forms. *Current Opinion in Plant Biology*, 34, 114–121.

[77] Puzey, J., Gerbode, S., Hodges, S., Kramer, E., & Mahadevan, L. (2012). Evolution of spur-length diversity in *Aquilegia* petals is achieved solely through cell-shape anisotropy. *Proceedings of the Royal Society B*, 279, 1640–1645. doi: 10.1098/rspb.2011.1873.

[78] Ren, Y., Gu, T., & Chang, H. (2011). Floral development of *Dichocarpum, Thalictrum,* and *Aquilegia* (thalictroideae, ranunculaceae). *Plant Systematics and Evolution*, 292(3-4), 203–213. doi: 10.1007/s00606-010-0399-6.

[79] Rendle, A. (1925). The classification of flowering plants: Volume 2, dicotyledons. *Science*.

[80] Rueden, C., Schindelin, J., Hiner, M., DeZonia, B., Walter, A., Arena, E., & Eliceiri, K. (2017). Imagej2: Imagej for the next generation of scientific image data. *BMC Bioinformatics*, 18(529). doi: 10.1186/s12859-017-1934-z.

[81] Ruzin, S. (1999). Plant microtechnique and microscopy. *Oxford University Press*, 18(529).

[82] Scott, R., Spielman, M., & Dickinson, H. (2004). Stamen structure and function. *The Plant Cell*, 16(suppl 1), S46–S60. doi: 10.1105/tpc.017012.

[83] Sharma, B., Guo, C., Kong, H., & Kramer, E. (2011). Petal-specific subfunctionalization of an apetala3 paralog in the ranunculales and its implications for petal evolution. *New Phytologist*, 191(3), 870–883. doi: 10.1111/j.1469-8137.2011.03744.

[84] Sharma, B., Yant, L., Hodges, S., & Kramer, E. (2014). Understanding the development and evolution of novel floral form in aquilegia. *Current Opinion in Plant Biology*, 17, 22–27. doi: http://dx.doi.org/10.1016/j.pbi.2013.10.006.

[85] Tepfer, S. (1953). Floral anatomy and ontogeny in aquilegia formosa var. truncata and ranunculus repens. *University of California Press*, 25, 513–648.

[86] Tucker, S. & Hodges, S. (2005). Floral ontogeny of *Aquilegia, Semiaquilegia*, and *Enemion* (ranunculaceae). *International Journal of Plant Sciences*, 166(4), 557–574. doi: 10.1086/429848.

[87] Vanholme, R., Demedts, B., Morreel, K., Ralph, J., & Boerjan, W. (2010). Lignin biosynthesis and structure. *Plant Physiology*. doi: 10.1104/pp.110.155119.

[88] Walker-Larsen, J. & Harder, L. (2000). The evolution of staminodes in angiosperms: patterns of stamen reduction, loss, and functional re-invention. *American Journal of Botany*, 87(10).

[89] Wellesen, K., Durst, F., Pinot, F., Benveniste, I., Nettesheim, K., Wisman, E., Steiner-Lange, S., Saedler, H., & Yephremov, A. (2001). Functional analysis of the lacerata gene of arabidopsis provides evidence for different roles of fatty acid omega-hydroxylation in development. *Proceedings of the National Academy of Sciences*, 98, 9694–9699.

[90] Zhao, L., Gong, J., Zhang, X., Liu, Y., Ma, X., & Ren, Y. (2016). Floral organogenesis in *Urophysa*, a rediscovered endangered and rare species of ranunculaceae. *Botany*, 94(3), 215–224. doi: 10.1139/cjb-2015-0232.

Chapter 3

Genetic differences between staminodes and stamens

3.1 Introduction

In Chapter 2 I demonstrated that the *Aquilegia* staminodes are highly differentiated from stamen filaments, and show considerable tissue-level elaboration with their curled margins and asymmetric lignification patterns. Previous work in the Kramer lab established that duplications in the *Aquilegia* B gene *APETALA3* allowed for sub- and neo-functionalization, with the three paralogs each becoming specialized for petal, stamen, or staminode organ identity (Sharma & Kramer, 2013). While this work established the genetic code for staminode identity, it did not explore how the distinct downstream developmental programs of the stamen filament and staminode differ. In this chapter, I have characterized the divergence in the downstream genetic programs of staminode and stamen filaments.

3.1.1 Genetic control of stamen differentiation

Stamen organ identity is specified by the B (*APETALA3* and *PISTILLATA*) and C (*AGAMOUS*) homeotic genes, which function together in the third whorl (Krizek & Meyerowitz, 1996). In *Arabidopsis*, loss of function in either the *APETALA3 (AP3)* or *PISTILLATA (PI)* gene leads to homeotic conversion of the stamens into carpels, while loss of function in *AGAMOUS (AG)* leads to stamen to petal conversion (Goto & Meyerowitz, 1994; Jack et al., 1992; Bowman et al., 1991). ChIP-Seq analysis has identified more than 1,500 high confidence binding sites for both AP3 and PI in the Arabidopsis genome (Wuest et al., 2012), although other approaches identified a much smaller set of 47-200 genes as AP3 and PI regulated (Zik & Irish, 2003). This difference is likely due to the fact that the latter set was based on a microarray that covered only 25% of the genome and identified genes based on differential expression rather than DNA-binding.

After organ identity is established, stamens develop into two distinct regions: the upper fertile anther and lower sterile filament (Scott et al., 2004). The anther produces and releases pollen while the filament provides support and transports nutrients to the anther (Goldberg et al., 1993). In *Arabidopsis thaliana*, the hormones jasmonate, auxin, and gibberellin are required for stamen development (Song et al., 2013). Genes responsible for anther formation and development as well as involved in pollen development have been extensively studied in *Arabidopsis* and in crop plants such as rice and maize (Wilson & Zhang, 2009; Gomez et al., 2015). In general, anthers typically express enormous numbers of genes, such as in cotton and maize where anthers have been found to have express 20,000 loci (Wei et al., 2013;

Ma et al., 2008). There is considerable overlap of housekeeping genes between pollen and anther specific tissue, but there are also genes that are specific to each tissue type. Transcriptomics studies have identified between 3954-7235 genes that are expressed in pollen itself from *Arabidopsis* (Rutley & Twell, 2015), including enrichment of genes with roles in cell signaling, cell wall metabolism, and cytoskeleton - likely for pollen tube growth.

Stamen morphology is highly modified compared to typical lateral organs, which have one plane of symmetry - the medial/lateral plane - and are laterally expanded. Stamen filaments lack lateral expansion, however, and the anther rearranges its adaxial-abaxial polarity in order to form the two bilaterally symmetrical theca that will give rise to the microspores (Toriba et al., 2010). Both of these features can be tied to changes in how organ polarity is expressed. After leaf primordia initiate and separate from the apical meristem, they have an inherent adaxial/abaxial polarity - the side closest to the meristem is the adaxial side, while the side furthest from the meristem is the abaxial side. This polarity is reinforced by expression of distinct genetic programs in each domain. Laminar outgrowth occurs at the juxtaposition of the zones of expression of these adaxial/abaxial factors (Waites & Hudson, 1995). Establishment of polarity is required for laminar expansion, and for proper development of an adaxial (top) surface optimized for light capture, and an abaxial (bottom) surface optimized for gas exchange (Eshed et al., 2001).

In *Arabidopsis thaliana*, members of the HDZIPIII transcription factor family together with the AS1/AS2 complex promote adaxial identity while members of the KANADI and YABBY transcription factor family along with ARF3/ARF4 promote

abaxial identify (Figure 3.1 A) (Emery et al., 2003; McConnell et al., 2001; McConnell & Barton, 1998). Other interacting factors such as the small RNA mir166 or *trans*-acting small interfering RNAs (ta-siRNAs) limit the expression domains of adaxial or abaxial factors, respectively. When adaxial-abaxial factors have balanced expression, laminar outgrowth occurs at the margin (Eshed et al., 2001). Gain of function adaxial gene mutations led to adaxialization of *Arabidopsis* leaf tissue, while loss of function mutations in adaxial genes led to abaxialization (Emery et al., 2003; Kerstetter et al., 2001). In both cases, the adaxialized or abaxialized tissues lacked laminar outgrowth which resulted in radialized organs.

During stamen development adaxial/abaxial polarity is initially expressed in a similar fashion to that in leaves, with adaxial gene expression adjacent to the meristem and abaxial expression gene expression in the opposite domain (Toriba et al., 2011). However, within the anthers, adaxial/abaxial polarity quickly becomes rearranged. Formation of bilaterally symmetrical theca, which each have internal adaxial/abaxial polarity, requires changes in the zones of expression of adaxial and abaxial genes (Figure 3.1 B) (Toriba et al., 2011). Outgrowth at the junctions of these new adaxial/abaxial expression domains form pollen sacs instead of the typical laminar expansion.

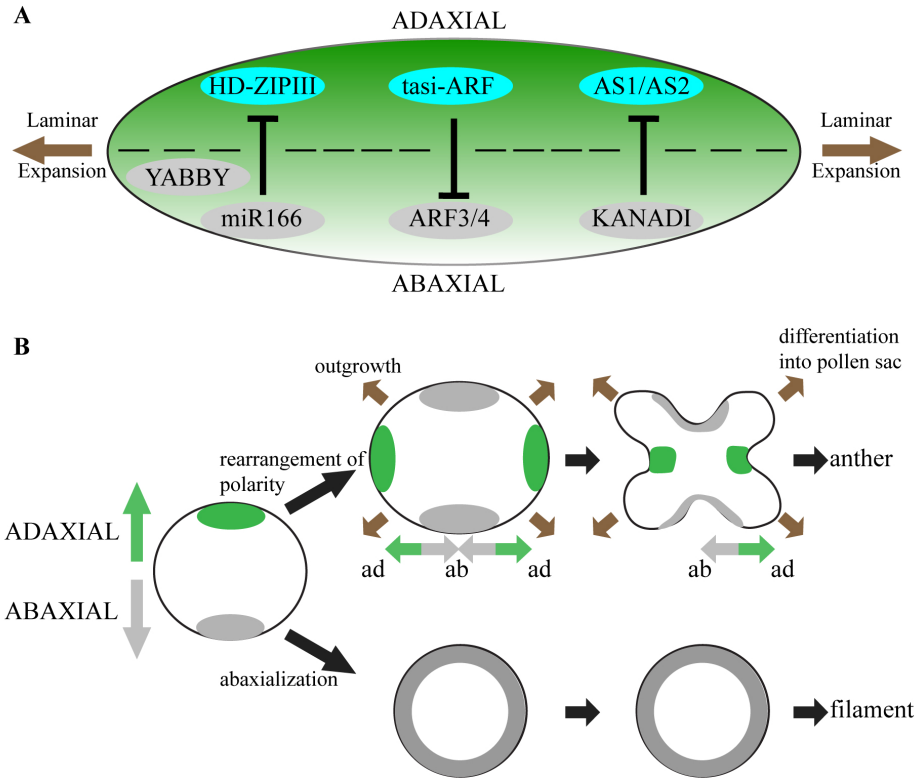


Figure 3.1: Model of adaxial-abaxial growth. (A) Model of a leaf with adaxial and abaxial identity factors. (B) Model of stamen adaxial and abaxial growth. Adapted from Husbands et al., 2009 and Toriba et al., 2011.

In the filament, organ polarity may be either maintained or lost, depending on the final stamen morphology. If a proper balance of adaxial/abaxial gene expression is maintained, laminar expansion will occur at the margins, yielding a broad stamen filament. Narrow filaments are caused by an uneven balance of either adaxial or abaxial gene expression (Eshed et al., 2001). For instance, in the Musaceae family of the Zingiberales, the ancestral state of the filaments is radially symmetric and narrow. Consistent with this, narrow filaments of *Musa acuminate* showed high expression of the abaxial YABBY2/5 genes, suggesting abaxialization. However, a representative

from the ginger family with derived petaloid, wide, staminodes, *Costus spicatus*, shows lower YABBY2/5 expression when compared to radialized filaments from *Musa acuminata* (De Almeida et al., 2014), suggesting that the abaxialization has been reversed as the expression of abaxial and adaxial factors come back into balance. Within the monocots, *Canna indica* is another representative of the Zingiberales in which the evolution of petaloid staminodes has been associated with a decrease in fertile stamens to the point of retaining only one functional stamen. RNA-seq analyses comparing the petaloid staminodes to petals found that petals had higher expression of secondary cell walls biogensis genes than the petaloid staminodes, while the staminodes had genes associated with stamen development and cell division (Tian et al., 2016). However, the *Canna* homologs of the B class PISTILLATA gene and C class *AGAMOUS* gene were identified in both the petaloid staminodes and stamens, so it remains to be determined how their differential development is controlled.

Heteranthery is another form of stamen modification that occurs when there are stamens within a flower that differ from one another in form and/or function. *Cassia bicapsularis*, within the eudicot *Fabales* order, is an example of a flower that has both different types of fertile stamens and sterile staminodes, in which the sterile staminodes function as part of a "pollen-mimicry" system to attract pollinators. RNA sequencing in *Cassia* comparing the different types of fertile stamens and staminodes revealed more differences between fertile stamens and sterile staminodes than between the various types of fertile stamens. Basic loop-helix-loop (bHLH) transcription factors and MADS box transcription factors were enriched in fertile

stamens compared to staminodes, but transcripts of two *AP3* paralogs, representatives of the core eudicot *TM6* lineage, were identified as enriched in staminodes (Luo et al., 2016), potentially suggesting a similar mechanism of gene duplicate-dependent organ identity as seen in *Aquilegia*. However, those staminodes are still radial, similar to the fertile stamens, so no difference in abaxial/adaxial gene expression was observed.

3.1.2 Staminodes in *Aquilegia*

Staminodes in *Aquilegia* are laterally expanded compared to the stamen filaments from which they are derived. They curl at their margins, forming an interlocking sheath that surrounds the carpels, and display late asymmetric adaxial lignification. Staminodes are also found in *Aquilegia*'s sister genera *Semiaquilegia* and *Urophysa*, but are rudimentary and irregular in number in *Semiaquilegia* and do not exhibit lignification in *Urophysa*.

The clade of *Aquilegia*, *Semiaquilegia*, and *Urophysa* diverged from their closest related clade of *Isopyrum, Enemion*, and *Dichocarpum* between 20 and 22 million years ago (mya) (Bastida et al., 2010). This pushes the date of origin of staminodes back from what was predicted in previous studies (6 mya, after the divergence of *Aquilegia* and *Semiaquilegia*) up to 20 mya. In Chapter 2, I demonstrated that staminodes in *Aquilegia* are similar to staminodes in *Urophysa*, but are very morphologically distinct from the stamen filaments from which they are derived.

Members of the *Aquilegia* genus are diploid, have seven chromosomes, and have a genome that is approximately 350 Mbp (Kramer, 2009). Molecular genetics

studies of *Aquilegia* have linked duplications in the B class *APETALA3* homologs to evolution of its novel fifth organ identity, the staminodes. There are four paralogs of *APETALA3* in *Aquilegia* that differ in their expression patterns: *AqAP3-1* is enriched in staminodes, *AqAP3-2* is enriched in stamens, *AqAP3-3* is petal specific, and *AqAP3-3b* is expressed at very low levels and was excluded from functional studies (Kramer et al., 2007). Viral-induced silencing (VIGS) of each of the paralogs has shown that distinct organ identity codes differentiate between stamens and staminodes, with *AqAP3-2* being particularly important in stamens while *AqAP3-1* is essential to staminodes (Sharma & Kramer, 2013). These studies established *AqAP3-1* as a key organ identity gene for staminodes, but did not explore the pathways downstream of organ identity.

A microarray study of the five floral organ types was performed on late stage pre-anthesis organs in *A. formosa* comparing gene expression profiles of each floral organ. This study identified 160 genes that were enriched specifically in the staminodes, as well as 197 genes co-up-regulated with carpels and 124 genes co-upregulated with stamens (Voelckel et al., 2010). Among the gene ontology categories of biological processes (GOBP) identified as up-regulated in staminodes were lignin biosynthesis, response to wounding, fatty acid beta-oxidation, and one carbon compound metabolic process (Voelckel et al., 2010). The microarray was designed based on preliminary EST sequencing and used 17,000 Aquilegia unigenes. While extensive, we now know that there are there are 30,023 loci containing protein-coding transcripts in the *Aquilegia* genome (JGI). RNA-seq has several advantages over DNA microarrays, including detecting genes with low expression and

higher reproducibility between biological replicates (Zhao et al., 2014).

The molecular mechanisms underlying the development of the infertile staminodes have only been explored with this one microarray. In order to flesh out our understanding of the genetic pathways that differ between staminodes and stamen filaments, I conducted an RNA-seq experiment sampling *Aquilegia coerulea* 'origami' at two developmental stages: bud stage (termed "pre" relative to lateral staminode adhesion) and pre-anthesis (termed "post" relative to this adhesion). There is low genetic variation between the 70 members of the genus (Whittall et al., 2006), making it likely that we would identify similar gene ontology categories in this study and the previous microarray analysis of late stage staminodes in *A. formosa*.

This experiment aimed to address three questions:

1. How does gene expression differ between bud stage staminodes and stamen filaments? Data from this developmental time point has not previously been collected.

2. How does gene expression differ between late stage staminodes and stamen filaments? Data comparing tissues from this developmental time point should overlap the microarray data but should be more complete using RNA-seq.

3. How does gene expression differ between early and late stage staminodes?

3.2 Methods

3.2.1 Plant growth conditions and treatment

Plant growth conditions were as outlined in section 2.2.1.

3.2.2 RNA extraction, cDNA library construction, Illumina Sequencing

Quadruplet replicates of staminodes and stamen filaments at two stages (11a, b and 13a) of floral development were removed and immediately flash frozen in liquid nitrogen. Each biological replicate contained tissue from multiple flowers, with each bioreplicate drawn from different plants, and each corresponding sample of staminodes and stamen tissue drawn from the same batch of flowers. Total RNA was isolated using the Qiagen RNeasy Plant Mini kit (Qiagen, Valencia, CA, USA), including a repeated elution step. RNA quality was analyzed using a NanoDrop 8000 spectrophotometer (Thermofisher) and an Agilent Bioanalyzer. RNA samples with a minimum 2.09 260/280, and RNA integrity number of 7.5 were used as input using the Apollo Prep X PolyA8 mRNA 200 bp beta V1 protocol for single-end library generation (Supplemental Figure). All libraries were quality confirmed for correct size distributions by TapeStation, quantified by QBIT and quantitative PCR using the ABI prism 7900 and Kappa quantification kit and pooled before running on 1 Illumina HiSeq2500 lane. One sample (Stdpre4) had technical difficulties during library generation, and is excluded from analysis.

Table 3.1: Summary of *Aquilegia* RNA sequencing

Sample	***Yield (Mbases)***	***# Reads (pre-trimming)***	***# Reads (post-QC trimming)***	***#Mapped Reads***	***% Mapped***
Std post1	775	15,508,484	15,459,777	14,718,776	95.2%
Std post 2	756	15,125,386	15,081,706	13,993,181	92.8%
Std post 3	1011	20,210,774	20,180,821	19,463,969	96.4%
Std post 4*	27*	547,347*	546,768*	530,313*	97%*
Std pre 1	807	16,147,419	16,110,314	15,570,575	96.6%
Std pre 2	640	12,802,384	12,773,571	12,237,190	95.8%
Std pre 3	847	16,939,577	16,914,131	16,397,471	96.9%
Std pre 4	985	19,703,638	19,675,567	18,997,198	96.6%
St fil post 1	916	18,318,395	18,271,047	17,151,110	93.9%
St fil post 2	600	12,008,277	11,975,120	11,363,761	94.9%
St fil post 3	952	19,044,300	19,014,631	18,347,604	96.5%
St fil post 4	940	18,801,125	18,723,453	18,000500	96.1%
St fil pre 1	497	9,946,767	9,916,960	9,429,932	95.1%
St fil pre 2	900	18,007,403	17,951,889	16,695,951	93.0%
St fil pre 3	1006	20,117,468	20,086,087	19,357,576	96.4%
St fil pre 4	1102	22,047,455	22,014,130	21,207,863	96.3%

3.2.3 RNAseq analyses

All Illumina data were assessed for basic quality with FastQC v0.11.5 (Babraham Bioinformatics). Adapters were trimmed using CutAdapt v1.8.1 (Martin, 2011). Read mapping was performed with TOPHAT v.2.0.13, using the Aquilegia V3 annotation as a reference (JGI). Numbers of mapped reads are shown in Table 3.1, and were used in subsequent analyses. Transcripts were assembled for individuals using CUFFLINKS v2.2.1 using default parameters (Trapnell et al., 2012). Uniquely mapped reads were counted in 30,023 gene models. Read counts were measured using HTSeq (Anders et al., 2015), and differential expression between tissue samples was analyzed in RStudio using EdgeR as well as DESeq2 (Chen et al., 2016; Love et al., 2014; Anders et al., 2013; Robinson et al., 2010; McCarthy et al., 2012).

3.2.4 Gene Ontology Enrichment Analysis

Out of 30,023 loci in *Aquilegia*, 20,702 matched genes described in *Arabidopsis*. Significantly upregulated genes in staminodes were identified and run through the BLASTx algorithm to identify their *A.thaliana* orthologs. Gene ontology enrichment was analyzed using Panther, which assigns significance using Fisher's Exact Test with FDR multiple test correction and a False Discovery Rate cutoff of .05 (Huaiyu et al., 2016).

3.2.5 YABBY gene tree construction

Amino acid sequences were obtained from Phytozome (JGI) and aligned using Clustal W in MacVector version 15.1.5. The obtained alignment was adjusted manually. RAxML was performed using CIPRES Science Gateway version 3.1.

3.2.6 *In situ* probe preparation and hybridization

mRNA sequences for *AqFIL* and *AqCRC* were obtained from phytozome (JGI). RNA probe synthesis and *in situ* hybridization was performed according to previously described methods (Kramer, 2005).

Table 3.2: Primers used for probe synthesis

Primer	**Sequence**
AqFIL_Forward	5'GCTGCTAATAGACCTCGGAGAA
AqFIL_Reverse	5'ACATCATGATCTCCTTCCTGCT
AqCRC_Forward	5'CAATCTTCTTCGTCATCTTCATCAAC
AqCRC_Reverse	5'CCAACTATTAGGCACCGTCAAG

3.2.7 Promoter analyses

Lists of significantly upregulated genes in staminodes were generated from the analyses described above. Gene names were matched with annotation information from the gff3 annotation file, and gene coordinates were extracted into a new file. Genes were sorted by position on the (+) or (-) strand. It is accepted that the 1000 base pair region upstream of the transcription start site (TSS) contains the promoter region (Zambelli et al., 2014). 1000 base pairs were added to the (-) strand end site, and 1000 base pairs were subtracted from the (+) strand start site in order to obtain the 5' UTR and proposed promoter regions. The coordinates were written in bed file format, and the Galaxy platform was used to extract the sequences from the fasta genome file. The sequence output was input into MEME Suite 4.12.0 using the command line (Bailey et al., 2009); Bailey & Elkan, 1994). Parameters included -dna (the input sequences were in DNA form); mod -anr (to look for any number of repeats of sequences from each gene); -minw of 4 and maxw of 16 (looking for sequences between 4-16 base pairs), and -evt of 0.01 (an E value cutoff of .01). Sequences were matched with putative TF binding sites using TOMTOM and the JASPER database (Khan et al., 2018; Mathelier et al., 2016).

3.3 Results

3.3.1 Quality Control

In order to explore the genes controlling the morphological differences between stamen filaments and staminodes, I generated transcriptomes from the two tissues at two developmental stages. Early stage "pre-fusion" flowers were defined as unopened flower buds that had not yet developed color and had staminodes that could be easily removed individually (Stage 11a, 11b flowers). Late stage "post-fusion" flowers were defined as open pre-anthesis flowers where the staminodes were removed in sheaths (Stage 12, 13a flowers). Four biological replicates were collected of each of the four sample types, for a total of 16 samples. One replicate of the late stage staminodes samples had technical errors during library prep, and was excluded from analysis.

Reads from each of the remaining replicates were mapped onto the reference genome, with percent mapped reads ranging from 92.8-97% (Table 3.1). HTSeq-count was used to count the number of aligned reads that overlapped exons in each of 30,023 gene models (Anders et al., 2015). Lowly expressed genes without at least 1 read per million reads in at least 3 samples were removed. Differentially expressed genes without a log fold change of 1 or with a p value and FDR value below .001 were removed. A final data matrix of 19,087 gene models remained after filtering out non-expressed genes across all samples.

After normalization for library size (Supplemental Table A.1), the relationships between samples were explored (Figure 3.2). All of the replicates for

each sample type formed clusters in both EdgeR and DESeq analyses. Dimension 1 separated developmental stage of the samples, while dimension 2 corresponded to tissue type. This suggests that there is strong reproducibility between samples, and that samples collected at the same developmental stage were more similar than samples of the same tissue type.

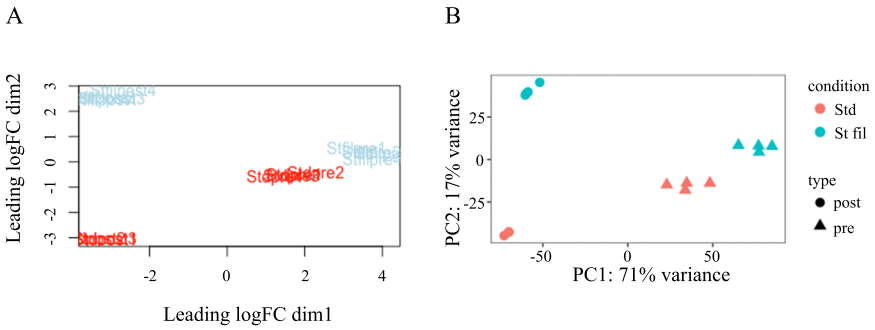


Figure 3.2: Clustering of sample replicates. (A) Multidimensional scaling plot produced using EdgeR's plotMDS showing the relationship between all pairs of samples. B) PCA plot from DESeq's plotPCA plotting rlog transformed count data. Red: Std, Staminodes; Blue Stfil, Stamen filament.

In this experiment, there were four biological replicates per sample. Due to the biological variation between replicates, there is variance in the sequencing results. This is viewed as overdispersion in sequencing data. This within-group variation is modeled by the dispersion parameter. Negative binomial dispersion was estimated using EdgeR's Cox-Reid profile-adjusted likelihood (CR) method and found to be .0519. The square root of dispersion yields the coefficient of biological variation (BCV), which was 0.22 (Figure 3.3). This indicates that expression values fluctuate by 22% between replicates.

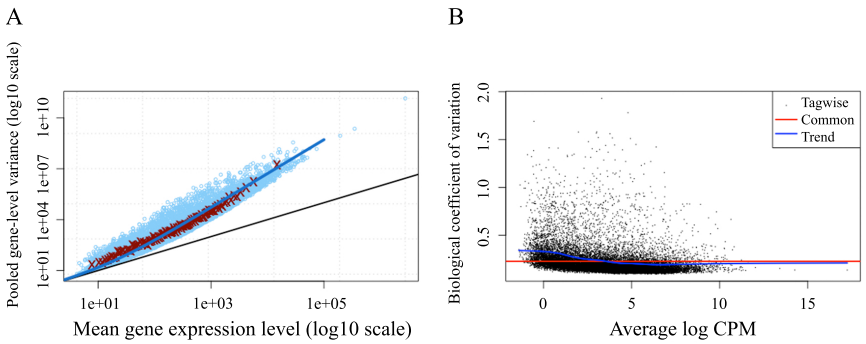


Figure 3.3: Plots of mean-variance relationship and dispersion. (A) edgeR's plotMeanVar was used to explore the mean-variance relationship; each dot represents thebestimated mean and variance for each gene, with binned variances as well as the trended common dispersion overlaid. (B) edgeR's plotBCV illustrated the relationship of biological coefficient of variation versus mean log CPM. CPM, counts per million.

3.3.2 Testing for DE genes

I began by looking at differentially expressed genes in pairwise comparisons. I identified genes that were up-regulated or down-regulated when comparing between staminodes and stamen filament tissues at both early and late stages, and then within staminodes and stamen filaments at early and late stages using both EdgeR and DESeq2 (Figure 3.4). Cutoffs of LogFC +/- 1 and p-values < 0.001 yielded similar numbers of differentially expressed genes using both pipelines. I compared the lists of genes identified in each pipeline and found that they overlapped between 91-98% (data not shown). While I relied on EdgeR for our differential gene expression analyses, I used DESeq in order to create heatmaps to visualize selected genes. Overall, gene expression differed the most over time within a tissue rather than between tissues.

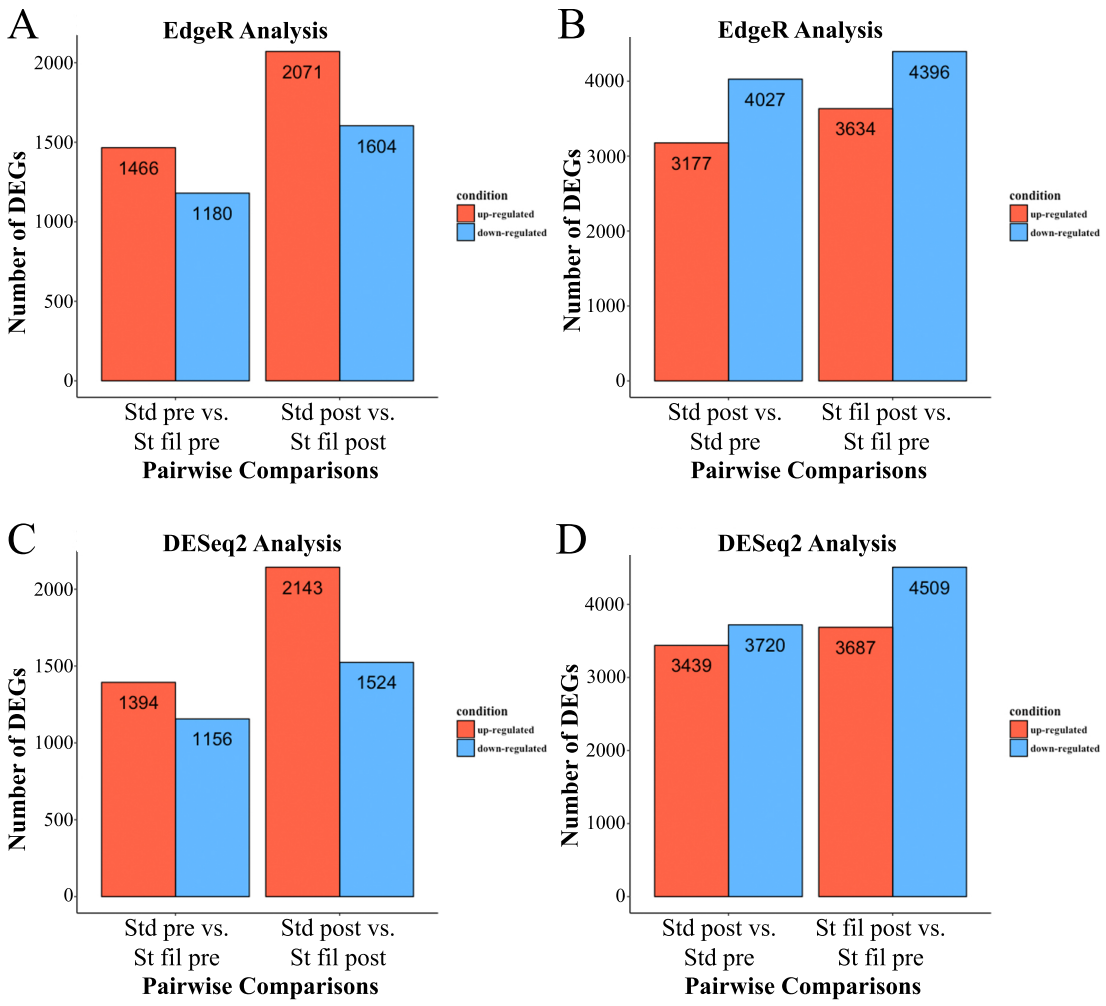


Figure 3.4: Differentially expressed genes identified from pairwise comparisons using (A) EdgeR and (B) DESeq2. Within each pairwise comparison, expression is relative to the first sample referenced (e.g. for Std pre vs St fil pre, upregulated genes are higher in Std pre). Red: upregulated genes. Blue: downregulated genes.

While the pairwise comparisons were useful for identifying differential expression between two samples, these comparisons were not useful for identifying genes specific to each sample type. I analyzed gene expression in the four sample types of *Aquilegia coerulea* 'origami' by fitting a generalized linear model to the

expression count data obtained from RNA-Seq. The four different groups of samples allowed for various contrasts to be made. In order to identify genes that were specifically up- or down- regulated to each of the four samples, I performed quasi-likelihood F tests in EdgeR by fitting four models to the data - in each model a different sample type was set as the intercept, or reference sample. Log fold change values were calculated in each of three other samples relative to the reference, and the three pairwise comparisons were combined into a single F statistic and p value. If the resulting LogFC values for a gene were negative, this indicated that expression in the reference sample was up-regulated compared to the others. If the resulting LogFC values were positive, this indicated that expression in the reference sample was down-regulated compared to the others (see below).

If Sample A - Reference = [Negative Value]

Then Reference > Sample A

If Sample B - Reference = [Positive Value]

Then Reference < Sample B

Gene lists were filtered by Benjamini and Hochberg adjusted p values < .001, and logFC < -1 and > 1. Numbers of up-regulated and down-regulated genes from each sample type compared to the others are shown in Table 3.3. Lists of differentially expressed genes specific to each sample type can be found at (https://drive.google.com/file/d/1BguMQreHOgV0-AN5DiXRIE-ZAGXNEW38/view?usp=sharing)

.

Table 3.3: Differentially expressed genes specific to each sample at LogFC 1. Four contrasts were made comparing each sample against all others. All differential expression was calculated at P values < 0.001 and log fold changes of either > 1 or < -1.

Sample	# Up-regulated genes	# and % Up-regulated that matched At genes	#Down-regulated genes	# and % Down-regulated that matched At genes
Staminodes pre	285	214 (75%)	335	243 (72.5%)
Staminodes post	1355	1078(79.6%)	811	662 (81.6%)
Stamen filaments pre	976	775 (79.4%)	1056	852(80.7%)
Stamen filaments post	811	631 (77.8%)	975	769 (78.9%)

The genes identified are uniquely up or down regulated in one sample, which filtered out genes that were up or down regulated in both staminode samples. In order to identify those genes that showed similar expression patterns across both staminode samples, I grouped the two staminode samples together and the two stamen filament samples together in our design matrix. I again performed quasi-likelihood F tests, and filtered for p adjusted genes with log fold changes above 1 or below -1. Results from this analysis are shown in Table 3.4.

Table 3.4: Differentially expressed genes identified when comparing staminodes to stamen filaments

Comparison	# Up-regulated genes	# and % Up-regulated that matched At genes	#Down-regulated genes	# and % Down-regulated that matched At genes
Staminodes compared to Stamen filaments (Log FC1)	168	137 (81.5%)	46	39 (84.5%)

3.3.3 Gene Ontology Analysis

In order to identify which gene ontology categories were enriched in staminodes, I took my differentially expressed gene lists and used BLASTx to identify their *Arabidopsis* homologs. Numbers of matches are reported in Table 3.5. The *Arabidopsis* gene IDs were used as input for gene ontology enrichment analyses. The PANTHER classification system searched database of 27502 *Arabidopsis*

reference genes and identified over-represented Gene Ontology categories for each of my samples. Reported gene ontology categories all had reported p-values < 0.05 with false discovery rate (FDR) multiple test correction.

Four gene ontology categories (secondary cell wall biogenesis, xylan metabolic process, lignin biosynthesis, oxidation-reduction) were enriched in both staminode samples, and no gene ontology categories were downregulated in both samples. Five GO categories were enriched only in staminodes pre, and 43 GO categories were enriched only in staminodes post (Table 3.5).

Notable GO that were enriched in staminodes pre include secondary metabolite biosynthetic processes. Notable GO categories that were enriched only in staminodes post included lignin biosynthetic processes, defense responses and responses to wounding, cell wall biosynthesis and organization. Notable gene ontology categories that were enriched in staminodes overall compared to stamens included cell wall and lignin biosynthesis.

Table 3.5: Enriched Gene Ontology categories. Grey colored cells indicate sample type. Yellow boxes correspond to gene ontology categories related to secondary cell wall. Pink boxes correspond to the lignin pathway. Blue boxes correspond to cell death. Green boxes correspond to defense and stress responses.

GO ID	GO Description
Staminode UP-regulated	
GO:0009834	Plant-type secondary cell wall biogenesis
GO:0045491	Xylan metabolic process
GO:0009809	Lignin biosynthetic process
GO:0055114	Oxidation-reduction process
Staminode DOWN-regulated	
N/A	N/A
Staminode pre UP-regulated	
GO:0043693	Monoterpene biosynthetic process
GO:0044550	Secondary metabolite biosynthetic process
GO:0055114	Oxidation-reduction process
GO:0016310	Phosphorylation
GO:0007275	Multicellular organism development
Staminode pre DOWN-regulated	
GO:0046323	Glucose import
GO:0035428	Hexose transmembrane transport
GO:0009813	Flavonoid biosynthetic process
GO:0045490	Pectin catabolic process
GO:0009739	Response to gibberellin
GO:0044550	Secondary metabolite biosynthetic process
GO:0055114	Oxidation-reduction process
Staminode Post UP-regulated	
GO:0009423	Chorismate biosynthetic process

Table 3.5: Enriched Gene Ontology categories (Continued)

GO:0010417	Glucuronoxylan biosynthetic process
GO:0048268	Clathrin coat assembly
GO:0009834	Plant-type secondary cell wall biogenesis
GO:2000762	Regulation of phenylpropanoid metabolic process
GO:0009718	Anthocyanin-containing compound biosynthesis process
GO:0006558	L-phenylalanine metabolic process
GO:2000652	Regulation of secondary cell wall biosynthesis
GO:0060548	Negative regulation of cell death
GO:0009809	Lignin biosynthesis process
GO:0006555	Methionine metabolic process
GO:0046323	Glucose import
GO:0035428	Hexose transmembrane transport
GO:0006730	One-carbon metabolic process
GO:0009407	Toxin catabolic process
GO:0009073	Aromatic amino acid family biosynthetic process
GO:0010089	Xylem development
GO:0030244	Cellulose biosynthetic process
GO:0010214	Seed coat development
GO:0006749	Glutathione metabolic process
GO:0010119	Regulation of stomatal movement
GO:0010224	Response to UV-B
GO:0009063	Cellular amino acid catabolic process
GO:0045490	Pectin catabolic process
GO:0010150	Leaf senescence
GO:0009611	Response to wounding
GO:0009664	Plant-type cell wall organization

Table 3.5: Enriched Gene Ontology categories (Continued)

GO:0005996	Monosaccharide metabolic process
GO:0007166	Cell surface receptor signaling pathway
GO:0009860	Pollen tube growth
GO:0080167	Response to karrikin
GO:0046777	Protein autophosphorylation
GO:0016042	Lipid catabolic process
GO:0030001	Metal ion transport
GO:0006979	Response to oxidative stress
GO:0009617	Response to bacterium
GO:0009651	Response to salt stress
GO:0055114	Oxidation-reduction process
GO:0033554	Cellular response to stress
GO:0006952	Defense response
GO:1901700	Response to oxygen-containing compound
GO:0070887	Cellular response to chemical stimulus
GO:0010033	Response to organic substance
Staminode Post DOWN-regulated	
GO:0010032	Meiotic chromosome condensation
GO:0007076	Mitotic chromosome condensation
GO:0009768	Photosynthesis, light harvesting in photosystem 1
GO:0018298	Protein-chromophore linkage
GO:0009637	Response to blue light
GO:0008202	Steroid metabolic process
GO:0006633	Fatty acid biosynthetic process
GO:0003002	Regionalization
GO:0009734	Auxin-activated signaling pathway

Table 3.5: Enriched Gene Ontology categories (Continued)

GO:0009639	Response to red or far red light
GO:0007166	Cell surface receptor signaling pathway
GO:1901615	Organic hydroxy compound metabolic process
GO:0055114	Oxidation-reduction process

3.3.4 Transcription Factor Expression

Gene ontology enrichment analyses deepened our understanding of the downstream pathways enriched in staminodes compared to stamen filaments. In order to identify key transcription factors in these pathways, I sub-sampled only the transcription factors from our lists of differentially expressed genes (Supplemental Data A.2). Of the B class genes, both *AqAP3-2* and *AqAP3-3* were expressed at higher levels in stamen filaments compared to staminodes. While *AqAP3-1* FPKM values were higher in both staminode samples, the gene was not identified as significantly enriched with our logFC cutoff (Figure 3.5). This likely reflects the relatively late developmental stage sampled in this study.

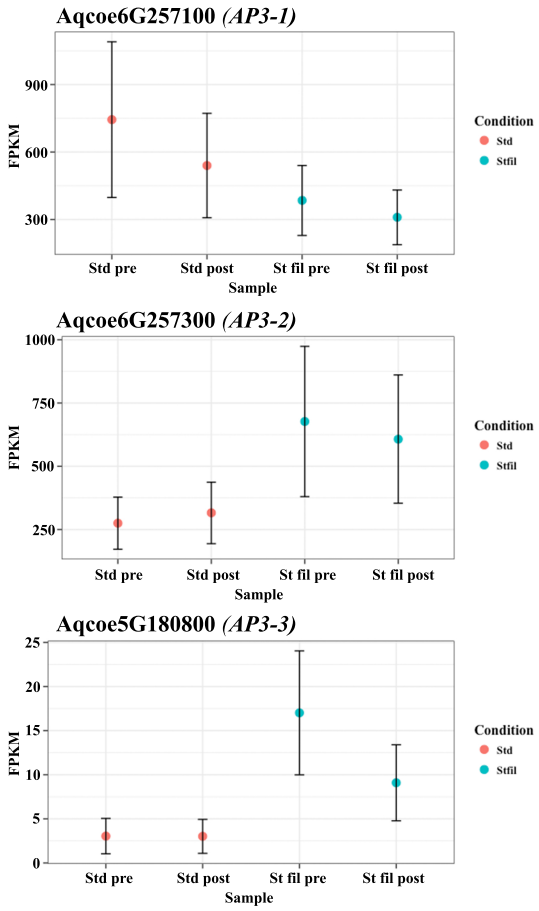


Figure 3.5: FPKM values of the three APETALA3 genes in *Aquilegia*.

There were 18 transcription factors upregulated in both staminode samples (out of 137 genes), and only 5 transcription factors downregulated (out of 39 genes) (Figure 3.6). Out of 214 genes with matches to *Arabidopsis* specifically upregulated in the "staminode pre" samples, 17 were transcription factors. Out of 243 downregulated genes, 12 were transcription factors. Finally, out of 1078 genes upregulated in the "staminode post" samples, 76 were transcription factors. Out of 662 downregulated genes, 31 were transcription factors. A large number of

transcription factors from the C2H2, MYB, NAC, and WRKY families were identified in my datasets, which is unsurprising given the large size of these transcription factor families.

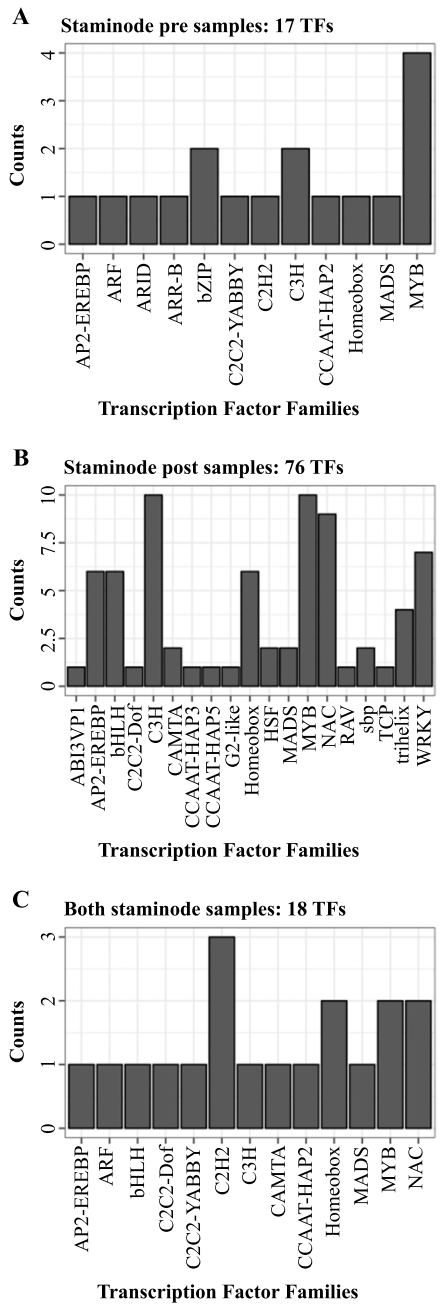


Figure 3.6: Numbers of transcription factors enriched in (A) Staminodia pre alone, (B) Staminodes post alone, or (C) both staminode samples compared to stamen filaments.

Due to the dramatic differences in lateral expansion between stamen filaments and staminodes, I was particularly interested in identifying transcription factors involved in patterning abaxial/adaxial polarity. From my lists of differentially expressed transcription factors, I identified *FILAMENTOUS FLOWER/YABBY1* as upregulated in both staminode samples, *CRABS CLAW* as enriched in early stage staminode samples, and *ARF4* and *REV* as enriched in early stage stamen filaments. In order to compare expression across the known groups of genes with roles in polarity, I visualized gene expression of a 12 selected genes in a heatmap (Figure 3.7).

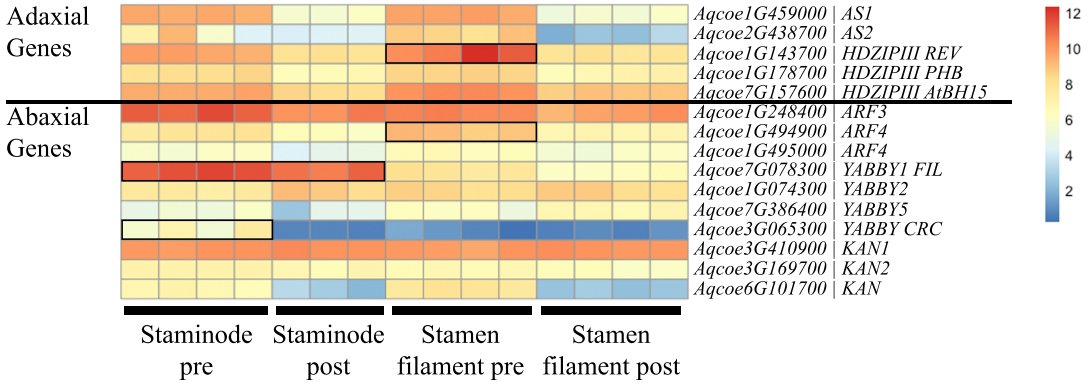


Figure 3.7: Heatmap of relative rlog-transformed values of 12 adaxial/abaxial genes across sample replicates. Sample groups are labeled "Staminode pre", "Staminode post", "Stamen filament pre", "Stamen filament post", and marked by black bars below the heatmap. Adaxial genes are represented by the 5 genes above the black line dividing the heatmap, abaxial genes are represented by the 10 genes below the black divison line. *PRS = PRESSED FLOWER, GRF = GROWTH-REGULATING FACTOR, CER = ECERIFERUM, DCR = DEFECTIVE IN CUTICULAR RIDGES, FDH= FIDDLEHEAD, GPAT = GLYCEROL-3-PHOSPHATE SN-2-ACYLTRANSFERASE, KCR = β-KETYACYL-COENZYMEA REDUCTASE*

None of the adaxial genes were significantly enriched in staminodes compared to stamen filaments, but one adaxial gene, the HDZIPIII TF family member *REVOLUTA* was enriched in early stage stamen filaments. Of the abaxial genes, two YABBY genes were enriched in staminode samples. *FILAMENTOUS FLOWER* was

strongly enriched in both samples, while *CRABS CLAW* was enriched in early stage staminode samples. These two genes were selected as candidates for the laminar expansion differences seen between staminodes and stamen filaments.

In addition to representative polarity genes, I used a second heatmap to visualize expression of loci involved in cell proliferation as well as cuticle development (Figure 3.8). I was particularly interested in identifying transcription factor candidates with roles in promoting lateral organ expansion, abaxial/adaxial differences, and lateral adhesion betweens staminodes.

PRESSED FLOWER (PRS) is a member of the WOX transcription factor family that has been shown to have roles in adaxial cell identity, petal and carpel fusion in *Petunia*, and in lateral organ development in *Arabidopsis* (Vandenbussche et al., 2009). *PRS* showed slight but significantly higher expression in early staminode samples, indicating it may be an interesting candidate for further study. I also included the *Aquilegia JAGGED* homolog in my heatmap, as previous work from the Kramer lab identified roles for this transcription factor in floral organ primordia outgrowth (Min & Kramer, 2016). There were no significant differences in expression for *AqJAG*, which was likely due to my later stages of sampling. Two growth response factors *(GRF2)* and *(GRF5)* displayed significantly higher expression levels in early stage stamen filaments, but overall expression patterns were similar between early staminodes and stamen filaments, and late staminodes and stamen filaments (Figure 3.8).

I was interested in exploring if genes involved in cuticle biosynthesis showed differential expression between staminodes and stamen filaments. *ECERIFERUM3*

(CER3), a transcription factor involved in alkane formation, was enriched in early stage stamen filaments while *DEFECTIVE IN CUTICULAR RIDGES (DCR)* and *HOTHEAD (HTH)* were both enriched in early stage staminodes.

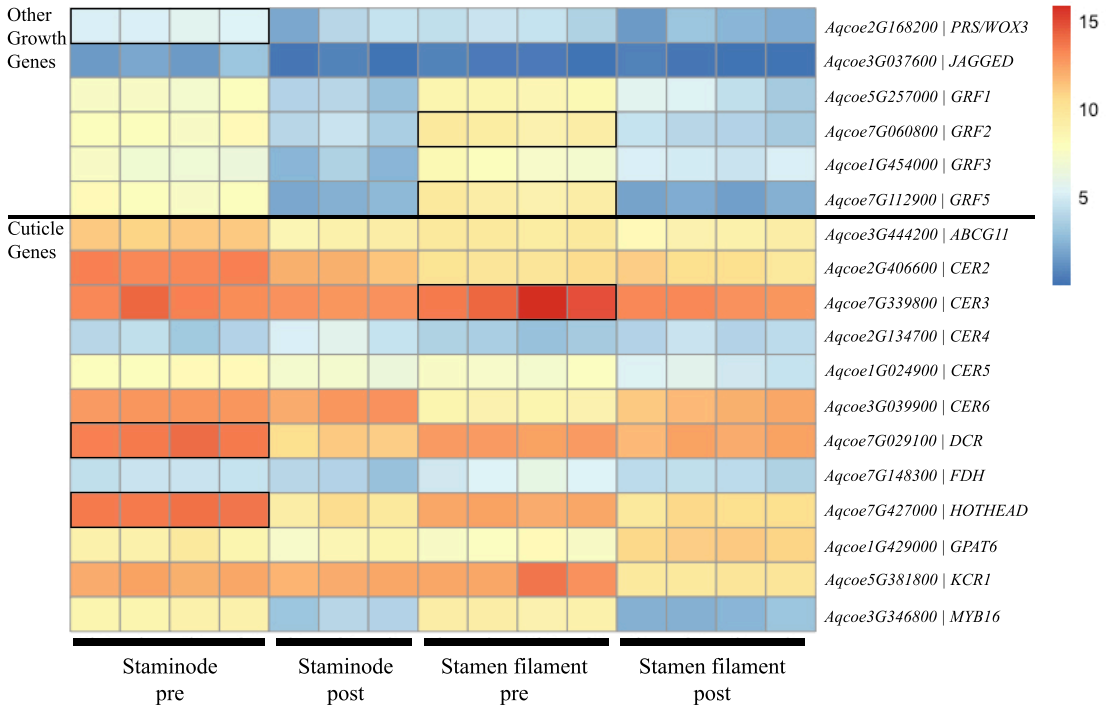


Figure 3.8: Heatmap of relative rlog-transformed values of selected growth and cuticle genes across sample replicates. Sample groups are labeled "Staminode pre", "Staminode post", "Stamen filament pre", "Stamen filament post", and marked by black bars below the heatmap. Selected growth genes are represented by the 6 genes above the black line dividing the heatmap, selected cuticle genes are represented by the 12 genes below the black divison line.

3.3.5 *YABBY in situ* hybridization

Since YABBY transcription factors have known roles in lateral organ polarity, I designed *in situ* probes to test expression of both *AqFIL* and *AqCRC* in wild-type meristems and buds. *Aquilegia* has 5 members of the YABBY transcription factor family (Figure 3.9). The *in situ* hybridization studies of *FILAMENTOUS FLOWER/YABBY1* and *CRABS CLAW (CRC)* suggest roles in floral organ polarity and carpel development, respectively. The sepal primordia emerged first and showed broad *FIL* expression (Figure 3.10 A). Petal and stamen primordia initially show abaxial *FIL* expression throughout their length (Figure 3.10 B). As the stamens differentiate into anther and filament, expression of *FIL* becomes restricted to the abaxial anther and is lost from the filament (Figure 3.10 C, D). However, expression persists in the staminodes throughout their length until Stage 10, and at late stages, it becomes localized to the abaxial surface of the basal portion of the filament (Figure 3.10 F-H). Strong abaxial expression was observed in the carpels throughout development.

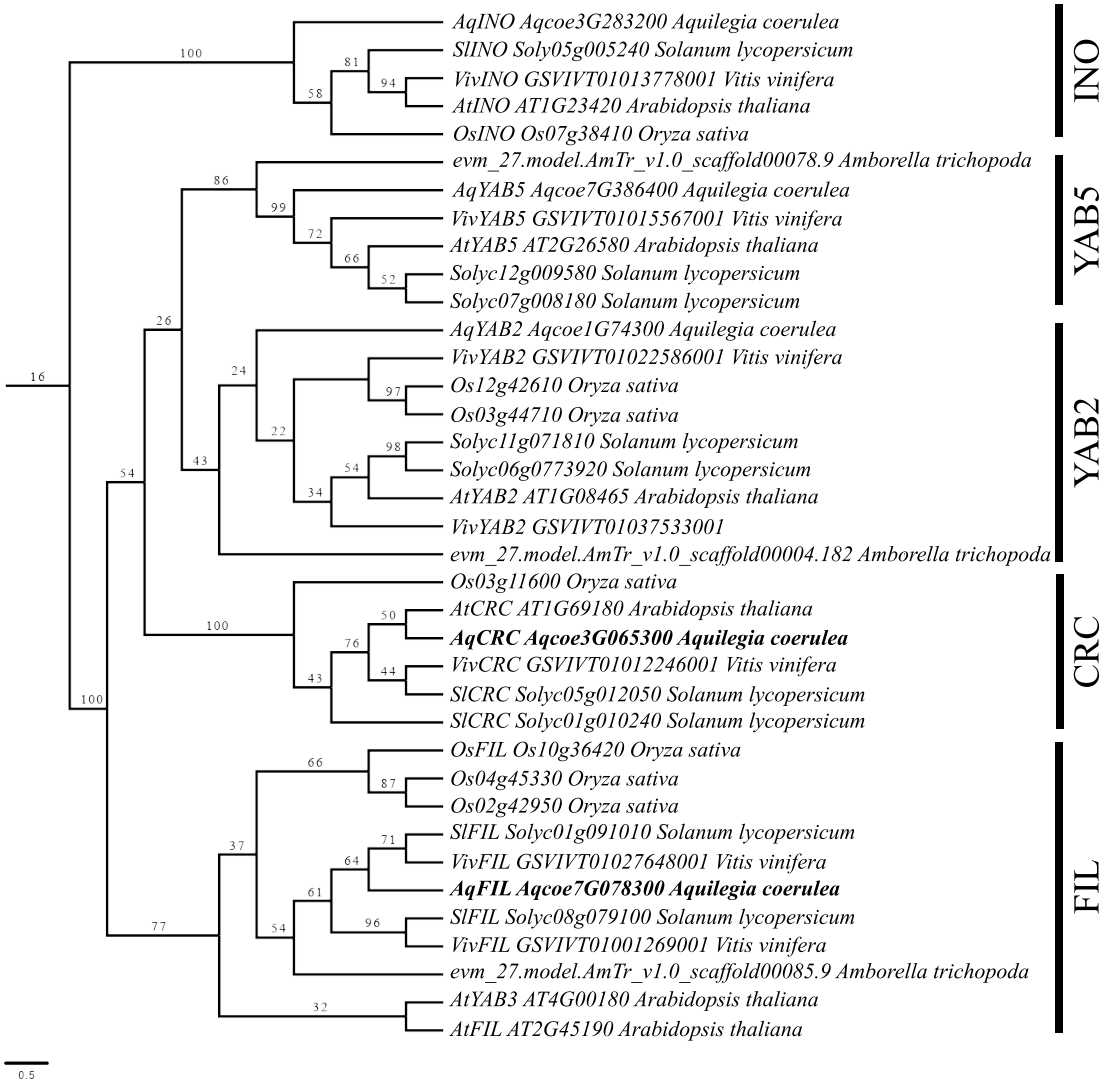


Figure 3.9: Relationships among sub-families of YAB genes. Inferred from Randomized Axelerated Maximum Likelihood.

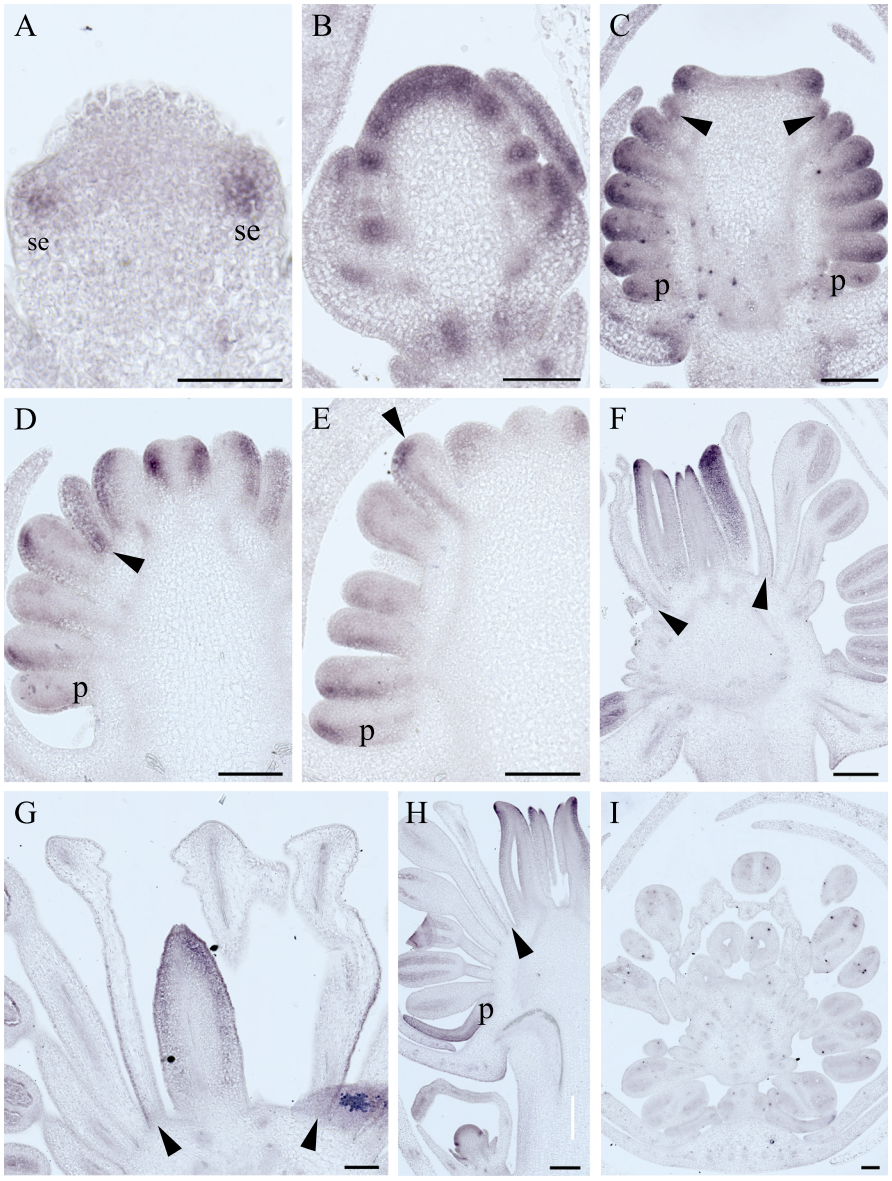


Figure 3.10: In Situ hybridization of *AqFil* in wild-type floral meristems. (A) Stage 3 floral meristems in which the sepals are just emerging. (B) Slightly later, between Stage 4 and 5, as the petal and stamen primordia are beginning to initiate. (C) Stage 7 when carpel primordia are appearing. (D) and (E) Stage 8 during which the floral organ primordia are differentiating. (F) and (G) Stage 9, staminodes that are clearly differentiated from stamens with abaxial staining. (H) Stage 10 floral bud. (I) Control probe. Scale bars: (A) 50 µm; (B), (C), (D), (G)), (I) 100 µm; (F), (H) 200 µm. Arrowheads throughout indicate staminodes.

Broad expression of *AqCRC* across the meristem becomes restricted to the carpels as the organ primordia emerged (Figure 3.11). Abaxial expression of *AqCRC* is consistent as the carpels emerged and differentiated (Figure 3.11).

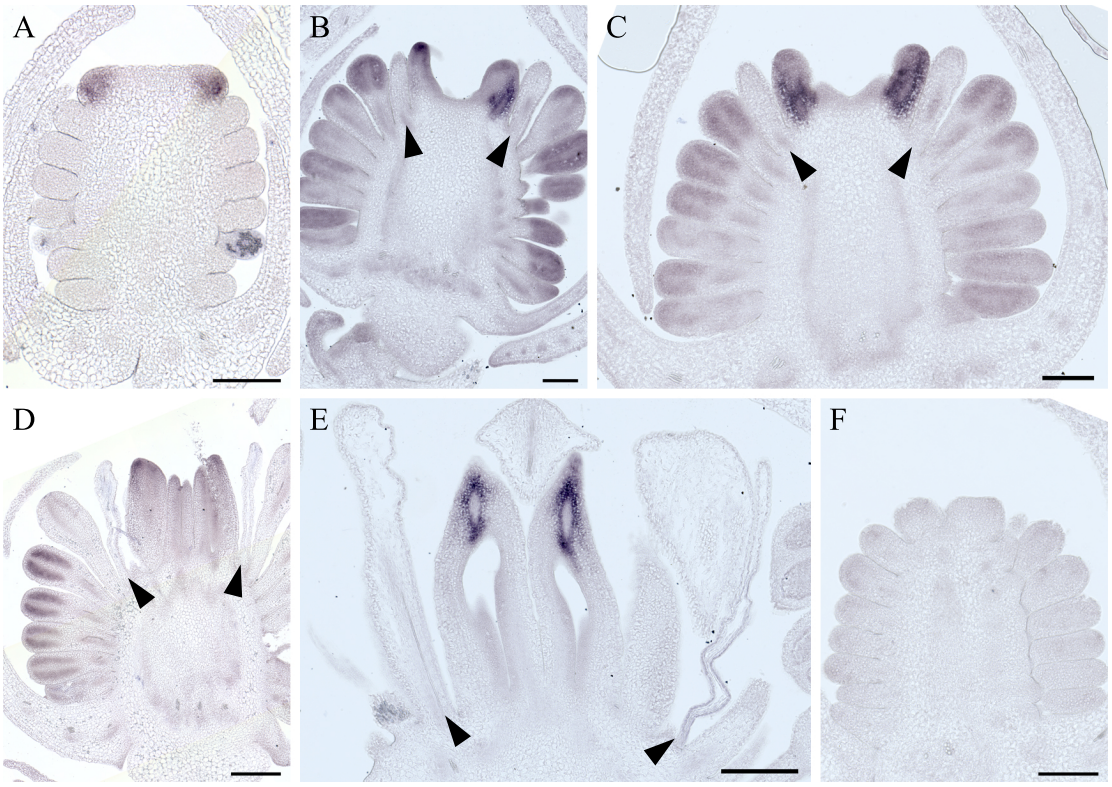


Figure 3.11: In Situ Hybridization of *AqCRC* in Wild-Type Floral Meristems. (A) Stage 6 bud, with carpel primorida appearing. (B) and (C) Stage 8 organ primordia are differentiating. (D) and (E) Staminodes are differentiated from stamen filaments during stage 9. (F) Control probe. Scale bars = 100 µm. Arrowheads throughout indicate staminodes.

3.3.6 Motif enrichment analysis

As an exploratory method of identifying other potentially important transcription factors downstream of organ identity, I took my three lists of upregulated and downregulated genes in both staminode samples and each of our individual staminodes samples at LogFC1, and, as recommended by Zambelli et al. (2014), extracted the 1000 base pairs upstream of the transcription start sites. This region is presumed to contain the promoter region of the genes, and is likely to contain short sequences, or motifs, that are bound by transcription factors. By comparing the promoters of all of the upregulated genes in staminodes I hoped to identify enriched motifs that could help me identify key transcription factors that regulated the differentially expressed genes (Zambelli et al., 2014). I identified sequences enriched in the promoters of both upregulated and downregulated genes and removed sequences that were common to both, as those sequences would be unlikely candidates for transcription factor binding sites that led to differential expression.

In order to detect enriched sequences I used the MEME platform, which uses an objective function based on the log likelihood ratio (LLR) of the occurrences of the motif. MEME calculates E values based on the likelihood of the motifs occurring compared to randomly generated input sequences. Significantly enriched sequences are reported in Supplementary Tables A.3, A.4, A.5. Transcription factors that bind to the sequences were identified by using the TOMTOM feature of MEME. TOMTOM searches the JASPAR database of known motifs in Arabidopsis and identifies transcription factors which bind to those sites (Khan et al., 2018). Repetitive

sequences such as AAAAAAAAAAAA were enriched in both upregulated and downregulated gene promoters, and were excluded from the results tables. The majority of sequences identified were not recognized as binding sites, although one AGAMOUS-LIKE 42 (AGL42) binding site was identified as enriched in both staminode samples, while a NAC binding site was identified as enriched in staminode post samples.

These analyses were limited by the number of transcription factor binding sites included in the JASPAR database, the fact that studies to identify binding sites have not been done in *Aquilegia*, identification of transcription factor binding sites is not enough to provide support that transcription factors actually bind.

3.4 Discussion

3.4.1 Gene expression differences between staminodes and stamen filaments

In this chapter I used RNA sequencing to explore the gene expression differences between stamen filaments and staminodes in *Aquilegia coerulea* 'origami' at two different developmental stages. This study expanded on a microarray study done by Voelckel et al. in 2010 that sampled late stage *Aquilegia formosa* floral organs. In that study, 160 genes were found to be enriched in staminodes and 1195 genes enriched in stamens. My more comprehensive study identified 285 genes enriched in early stage *Aquilegia* staminodes while 1355 genes were enriched at later stages. I identified similar ranges of genes enriched in stamen filaments: 976 in early stage filaments and 811 genes in late stage filaments. My gene ontology analysis

revealed an enrichment of categories related to secondary cell wall development, secondary metabolism, lignin synthesis, defense, and response to wounding as enriched in late stage staminodes, while photosynthesis and responses to light were downregulated. In my previous chapter, I identified lignin enrichment in staminodes based on histology, and used qRT-PCR to show enrichment of three genes in the lignin biosynthetic pathway. My RNA-seq dataset allowed me to confirm enrichment of the entire lignin monomer biosynthetic pathway, as well as genes involved in monomer activation and polymerization (Figure 3.12).

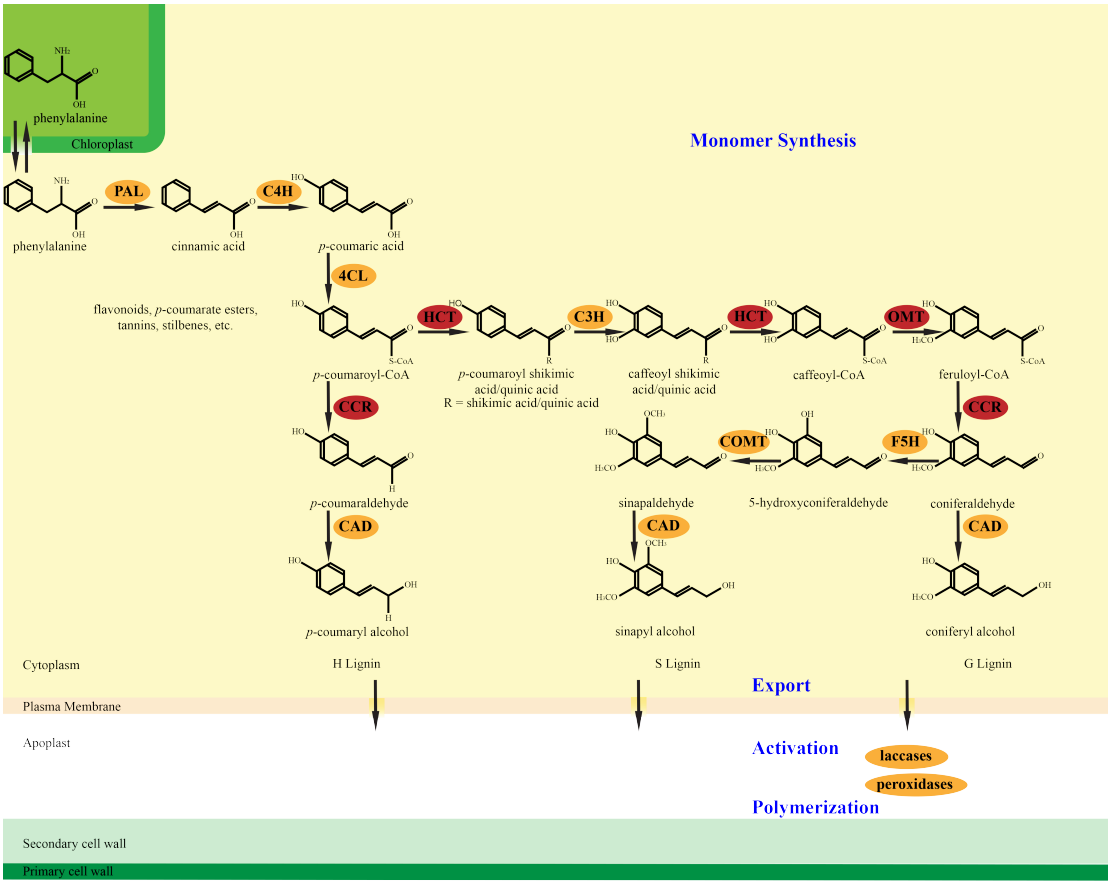


Figure 3.12: Overview of the Lignin Biosynthetic Pathway. Enzymes identified as enriched in late stage staminodes by RNA seq are encircled by orange, and enzymes further confirmed by qRT-PCR are encircled by red. Adapted from Bonawitz & Chapple, 2010

As staminodes are laterally expanded compared to stamens, and show adaxial/abaxial differences in lignification and cell number, I was interested in determining whether there are differences in expression of adaxial/abaxial identity genes. At my early time point, staminodes and stamen filaments showed similar expression patterns in most genes related to ab/ad differentiation except for the adaxial transcription factor *REVOLUTA* (higher in stamen filaments) and the abaxial transcription factors *YABBY1/FILAMENTOUS FLOWER* and *CRABS CLAW* (higher

in staminodes). At my late time point, genes that showed higher expression in staminodes included *YABBY1/FILAMENTOUS FLOWER*. This pattern suggests that *Aquilegia* stamen filaments are adaxialized while the staminodes have restored proper balance of abaxial/adaxial identity to yield laminar growth.

3.4.2 *YABBY* genes

The YABBY transcription factor family is specific to seed plants (Floyd & Bowman, 2007). There are six members of the YABBY gene family in Arabidopsis: *CRABS CLAW (CRC), FILAMENTOUS FLOWER (FIL)/YABBY1, YABBY2, YABBY3, INNER NO OUTER (INO)/YABBY4,* and *YABBY5* that encode zinc finger-like transcription factors with YAB domains (Bowman, 2000; Yamada et al., 2011). Four of the members *(FIL, YAB2, YAB3* and *YAB5)* are expressed in lateral organs with functions in abaxial cell identity and promoting laminar expansion, while *CRC* functions in regulating carpel and nectary development and INO is expressed in the outer integuments (Bowman, 2000; Bowman & Smyth, 1999).

YABBY expression is not conserved across the angiosperms - while in Arabidopsis and other dicots, members of the YABBY transcription factor family are expressed and function in abaxial identity, in rice the YABBYs do not show polarized expression, and in maize they show adaxial expression (Husbands et al., 2009).

FILAMENTOUS FLOWER has strong abaxial expression in *Arabidopsis*. Both knockout and gain of function lead to formation of filamentous organs, but with adaxialization in the case of *fil-1* and abaxialization in the case of *35S:FIL* (Sawa et al., 1999). The Kramer lab is actively working towards optimizing transgenics for

Aquilegia. The next steps for this project should involve functional work to determine if knocking down *FILAMENTOUS FLOWER* impairs lateral expansion of staminodes, or also if it has an impact on cell type polarity within staminodes.

In *Arabidopsis*, *CRABS CLAW* expression begins when the gynoecial primordium begin to differentiate, and continues until just after the ovule primordia arise. Additionally, expression is found at the base of the stamens where the nectaries will arise (Bowman and Smyth, 1999). However, in *Aquilegia formosa*, *AqCRC* expression was found only in the abaxial side of carpels and not in the nectaries (Lee et al., 2005). Our *in situs* of *Aquilegia coerulea* 'origami' also did not show expression in the nectaries. Although my RNA-sequencing data detected significant differences in expression between staminodes and stamen filaments, the differences were not as strong as those seen in *AqFIL*, and might not have been as easily detected by *in situ* hybridization.

My analyses of the promoters of genes enriched in staminodes served as an exploratory step to potentially identify binding sites of key transcription factors downstream of organ identity. While I did find enriched sequences, the majority of these sequences were not identified as known transcription factor binding sites when compared to the JASPAR database of 501 motifs (Khan et al., 2018). Furthermore, transcription factor binding sites for the YABBY gene family were not available through the JASPAR database. Putative binding sites for YABBY are unclear – an *in vitro* electrophoretic mobility shift assay (EMSA) study of the *Arabidopsis* FILAMENTOUS FLOWER (FIL/YAB1) protein indicated that FIL bound to DNA without specificity, but a protein-binding microarray showed that YAB1 and YAB5

bind to A/T rich elements, with a consensus sequence of WATNATW (Franco-Zorrilla et al., 2013; Kanaya et al., 2002). A gel shift binding assay of YAB1 binding in rice showed that the rice YAB1 bound to a GARE (TCTGTTA) site (Dai et al., 2007). A ChIP-Seq and RNA-Seq study in soybean identified three putative binding sites for YABBY in soybean: CC[CA][TC]C[TA][CT]C; GA[AG]AGAAA; and CCCCAC (Shamimuzzaman & Vodkin, 2013). In my analyses of promoters genes differentially expressed between staminodes and stamens, I did identify several CCCCAC sites, but further studies will be necessary to determine if the YABBY genes play a functional role in differences in morphology.

Previous ChIP-Sequencing analysis in *Arabidopsis* identified *CRABS CLAW* as directly regulated by APETALA3/PISTILLATA as well as AGAMOUS (Ó'Maoiléidigh et al., 2013; Wuest et al., 2012). In addition to *CRABS CLAW*, abaxial/adaxial axis specification was a process identified as a process directly regulated by AGAMOUS (Ó'Maoiléidigh et al., 2013). A microarray of *ap3-1* and *pi Arabidopsis* mutants also identified abaxial/adaxial axis specification as a process identified as a process indirectly regulated by AP3-1 and PI (Wuest et al., 2012). ChIP-sequencing of the AP3-1 and AP3-2 paralogs in *Aquilegia* would help determine how abaxial/adaxial identity is differentially expressed in stamens versus staminodes.

References

[91] Anders, S., McCarthy, D., Chen, Y., Okoniewski, M., Smyth, G., Huber, W., & Robinson, M. (2013). Count-based differential expression analysis of rna sequencing using r and bioconductor. *Nature Protocols*, 8, 1765–1786. doi: 10.1038/nprot.2013.099.

[92] Anders, S., Pyl, P., & Huber, W. (2015). Htseq-a python framework to work with high-throughput sequencing data. *Bioinformatics*, 31(2), 166–169. doi: 10.1093/bioinformatics/btu638.

[93] Bailey, T., Bodén, M., Buske, F., Frith, M., Grant, C., Clementi, L., Ren, J., Li, W., & Noble, W. (2009). Meme suite: tools for motif discovery and searching. *Nucleic Acids Research*, 37, W202–W208.

[94] Bailey, T. & Elkan, C. (1994). Fitting a mixture model by expectation maximization to discover motifs in biopolymers. *Proceedings of the Second International Conference on Intelligent Systems for Molecular Biology*, (pp. 28–36).

[95] Bastida, J., Alcantara, J., Rey, P., Vargas, P., & Herrera, C. (2010). Extended phylogeny of aquilegia: the biogeographical and ecological patterns of two simultaneous but contrasting radiations. *Plant Systems Evolution*, 284, 171–185.

[96] Bonawitz, N. & Chapple, C. (2010). The genetics of lignin biosynethsis: connecting genotype to phenotype. *Annual Review of Genetics*, 44, 337–363. doi: 10.1146/annurev-genet-102209-163508.

[97] Bowman, J. (2000). The *YABBY* gene family and abaxial cell fate. *Current Opinion in Plant Biol*, 3, 17–22.

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

[99] Bowman, J., Smyth, D., & Meyerowitz, E. (1991). Genetic interactions among floral homeotic genes of arabidopsis. *Development*, 112, 1–20.

[100] Chen, Y., Lun, A., & Smyth, G. (2016). From reads to genes to pathways: differential expression analysis of rna-seq experiments using rsubread and the edger quasi-likelihood pipeline. *F1000Research*, 5(1438). doi: 10.12688/f1000research.8987.2.

[101] Dai, M., Zhao, Y., Ma, Q., Hu, Y., Hedden, P., Zhang, Q., & Zhou, D. (2007). The rice *YABBY1* gene is involved in the feedback regulation of gibberellin metabolism. *Plant Physiology*, 144(1), 121–133. doi: 10.1104/pp.107.096586.

[102] De Almeida, A., Yockteng, R., Schnable, J., Alvarez-Buylla, E., Freeling, M., & Specht, C. (2014). Co-option of the polarity gene network shapes filament morphology in angiosperms. *Scientific Reports*, 4(6194). doi: 10.1038/srep06194.

[103] Emery, J., Floyd, S., Alvarez, J., Eshed, Y., Hawker, N., Izhaki, A., Baum, S., & Bowman, J. (2003). Radial patterning of arabidopsis shoots by class iii hd-zip and kanadi genes. *Current Biology*, 13(20), 1768–1774.

[104] Eshed, Y., Baum, S., Perea, J., & Bowman, J. (2001). Establishment of polarity in lateral organs of plants. *Current Biology*, 11(16), 1251–1260. doi: 10.1016/S0960-9822(01)00392-X.

[105] Floyd, S. & Bowman, J. (2007). The ancestral developmental tool kit of land plants. *International Journal of Plant Sciences*, 168, 1–35.

[106] Franco-Zorrilla, J., López-Vidriero, I., Carrasco, J., Godoy, M., Vera, P., & Solano, R. (2013). Dna-binding specificities of plant transcription factors and their potential to define target genes. *Proceedings of the National Academy of Sciences of the United States of America*, 111(6), 2367–2372. doi: 10.1073/pnas.1316278111.

[107] Goldberg, R., Beals, T., & Sanders, P. (1993). Anther development: Basic principles and practical applications. *The Plant Cell*, 5, 1217–1229.

[108] Gomez, J., Talle, B., & Wilson, Z. (2015). Anther and pollen development: A conserved developmental pathway. *Journal of Integrative Plant Biology*, 57, 876–891. doi: 10.1111/jipb.12425.

[109] Goto, K. & Meyerowitz, E. (1994). Function and regulation of the *Arabidopsis* floral homeotic gene *PISTILLATA*. *Genes and Development*, 8, 1548–1560. doi: 10.1101/gad.8.13.1548.

[110] Huaiyu, M., Huang, X., Muruganujan, A., Tang, H., Mills, C., Kang, D., & Thomas, P. (2016). Panther version 11: expanded annotation data from gene ontology and reactome pathways, and data analysis tool enhancements. *Nucl. Acids Res*. doi: 10.1093/nar/gkw1138.

[111] Husbands, A., Chitwood, D., Plavskin, Y., & Timmermans, M. (2009). Signals and prepatterns: new insights into organ polarity in plants. *Genes and Development*, 23, 1986–1997. doi: 10.1101/gad.1819909.

[112] Jack, T., Brockman, L., & Meyerowitz, E. (1992). The homeotic gene apetala3 of *Arabidopsis thaliana* encodes a mads box and is expressed in petals and stamens. *Cell*, 68, 683–697.

[113] Kanaya, E., Nakajima, N., & Okada, K. (2002). Non-sequence-specific dna binding by the filamentous flower protein from *Arabidopsis thaliana* is reduced by edta. *Journal of Biological Chemistry*, 277, 11957–11964. doi: 10.1074/jbc.M108889200.

[114] Kerstetter, R., Bollman, K., Taylor, R., Bomblies, K., & Poethig, R. (2001). Kanadi regulates organ polarity in arabidopsis. *Nature*, 411(6838), 706–709. doi: 10.1038/35079629.

[115] Khan, A., Fornes, O., Stigliani, A., Gheorghe, M., Castro-Mondragon, J., van der Lee, R., Bessy, A., Cheneby, J., Kulkarni, S., Tan, G., Baranasic, D., Arenillas, D., Sandelin, A., Vandepoele, K., Lenhard, B., Ballester, B., Wasserman, W., Parcy, F., & Mathelier, A. (2018). Jaspar 2018: update of the open-access database of transcription factor binding profiles and its web framework. *Nucleic Acids Research*. doi: 10.1093/nar/gkx1126.

[116] Kramer, E. (2005). Methods for studying the evolution of plant reproductive structures: Comparative gene expression techniques. *Methods in Enzymology*, 395, 617–636. doi: 10.1016/S0076-6879(05)95032-5.

[117] Kramer, E. (2009). Aquilegia: A new model for plant development, ecology, and evolution. *Annual Review of Plant Biology*, 60(1), 261–277. doi: 10.1146/annurev.arplant.043008.092051.

[118] Kramer, E., Holappa, L., Gould, B., Jaramillo, A., Setnikov, D., & Santiago, P. (2007). Elaboration of b gene function to include the identity of novel floral organs in the lower eudicot aquilegia. *The Plant Cell*, 19, 750–756. doi: 10.1105/tpc.107.050385.

[119] Krizek, B. & Meyerowitz, E. (1996). The arabidopsis homeotic genes apetala3 and pistillata are sufficient to provide the b class organ identity function. *Development*, 122(1), 11–22.

[120] Lee, J., Baum, S., Oh, S., Jiang, C., Chen, J., & Bowman, J. (2005). Recruitment of crabs claw to promote nectary development within the eudicot clade. *Development*, 132(22), 5021–5032. doi: 10.1242/dev.02067.

[121] Love, M., Huber, W., & Anders, S. (2014). Moderated estimation of fold change and dispersion for rna-seq data with deseq2. *Genome Biology*, 15, 550. doi: 10.1186/s13059-014-0550-8.

[122] Luo, Z., Hu, J., Zhao, Z., & Zhang, D. (2016). Transcriptomic analysis of hetermorphic stamens in *Cassia biscapsularis L.* *Scientific Reports*, 6. doi: 10.1038/srep31600.

[123] Ma, J., Skibbe, D., Fernandes, J., & Walbot, V. (2008). Male reproductive development: gene expression profiling of maize anther and pollen ontogeny. *Genome Biology*, 9, R181. doi: 10.1186/gb-2008-9-12-r181.

[124] Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal*, 17(1), 10–12. ISSN 2226–6089.

[125] Mathelier, A., Fornes, O., Arenilla, D., Chen, C., Denay, G., Lee, J., Shi, W., Shyr, C., Tan, G., Worsley-Hunt, R., Zhang, A., Parcy, F., Lenhard, B., Sandelin, A., & Wasserman, W. (2016). Jaspar 2016: a major expansion and update of the open-access database of transcription factor binding profiles. *Nucleic Acids Research*, 44(D1), D110–D115. doi: 10.1093/nar/gkv1176.

[126] McCarthy, D., Chen, Y., & Smyth, G. (2012). Differential expression analysis of multifactor rna-seq experiments with respect to biological variation. *Nucleic Acids Research*, 40(10), 4288–4297. doi: 10.1093/nar/gks042.

[127] McConnell, J. & Barton, M. (1998). Leaf polarity and meristem formation in arabidopsis. *Development*, 125, 2935–2942.

[128] McConnell, J., Emery, J., Eshed, Y., Bao, N., Bowman, J., & Barton, M. (2001). Role of phabulosa and phavoluta in determining radial patterning in shoots. *Nature*, 411(6838), 709–713.

[129] Min, Y. & Kramer, E. (2016). The aquilegia *JAGGED* homolog promotes proliferation of adaxial cell types in both leaves and stems. *New Phytologist*, 216, 536–548. doi: 10.1111/nph.14282.

[130] Robinson, M., McCarthy, D., & Smyth, G. (2010). edger: a bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics*, 26, 1.

[131] Rutley, N. & Twell, D. (2015). A decade of pollen transcriptomics. *Plant Reproduction*, 28(2), 73–89. doi: 10.1007/s00497-015-0261-7.

[132] Sawa, S., Watanabe, K., Goto, K., Kanaya, E., Morita, E., & Okada, H. (1999). *FILAMENTOUS FLOWER*, a meristem and organ identity gene of *Arabidopsis*, encodes a protein with a zinc finger and hmg-related domains. *Genes and Development*, 13, 1079–1088.

[133] Scott, R., Spielman, M., & Dickinson, H. (2004). Stamen structure and function. *The Plant Cell*, 16(suppl 1), S46–S60. doi: 10.1105/tpc.017012.

[134] Shamimuzzaman, M. & Vodkin, L. (2013). Genome-wide identification of binding sites for nac and yabby transcription factors and co-regulated genes during soybean seedling development by chip-seq and rna-seq. *BMC Genomics*, 14(477). doi: 10.1186/1471-2164-14-477.

[135] Sharma, B. & Kramer, E. (2013). Sub- and neo-functionalization of apetala3 paralogs have contributed to the evolution of novel floral organ identity in aquilegia (columbine, ranunculaceae). *New Phytologist*, 197(3), 949–957. doi: 10.1111/nph.12078.

[136] Song, S., Qi, T., Huang, H., & Xie, D. (2013). Regulation of stamen development by coordinated actions of jasmonate, auxin, and gibberellin in arabidopsis. *Molecular Plant*, 6(4), 1065–1073. doi: 10.1093/mp/sst054.

[137] Tian, X., Yu, Q., Liu, H., & Liao, J. (2016). Temporal-spatial transcriptiome analyses provide insight into the development of petaloid androecium in *Canna indica*. *Frontiers in Plant Science*, 7(1194). doi: 10.3389/fpls.2016.01194.

[138] Toriba, T., Ohmori, Y., & Hirano, H. (2011). Common and distinct mechanisms underlying the establishment of adaxial and abaxial polarity in stamen and leaf development. *Plant Signaling and Behavior*, 6(3), 430–433. doi: 10.4161/psb.6.3.14494.

[139] Toriba, T., Suzaki, T., Yamaguchi, T., Ohmori, Y., Tsukaya, H., & Hirano, H. (2010). Distinct regulation of adaxial-abaxial polarity in anther patterning in rice. *The Plant Cell*, 22(5), 1452–1462. doi: 10.1105/tpc.110.075291.

[140] Trapnell, C., Roberts, A., Goff, L., Pertea, G., Kim, D., Kelley, D., Pimentel, H., Salzberg, S., Rinn, J., & Pachter, L. (2012). Differential gene and transcript expression analysis of rna-seq experiments with tophat and cufflinks. *Nature Protocols*, 7(3), 562–578. doi: 10.1038/nprot.2012.016.

[141] Vandenbussche, M., Horstan, A., Zethof, J., Koes, R., Rijpkema, A., & Gerats, T. (2009). Differential recruitment of *Wox* transcription factors for lateral development and organ fusion in petunia and *Arabidopsis*. *The Plant Cell*. doi: 10.1105/tpc.109.065862.

[142] Voelckel, C., Borevitz, J., Kramer, E., & Hodges, S. (2010). Within and between whorls: comparative transcriptional profiling of aquilegia and arabidopsis. *PloS one*, 5(3). doi: 10.1371/journal.pone.0009735.

[143] Waites, R. & Hudson, A. (1995). *phantastica*: a gene required for dorsoventrality of leaves in *Antirrhinum majus*. *Development*, 121, 2143–2154.

[144] Wei, M., Song, M., Fan, F., & Yu, S. (2013). Transcriptomic analysis of differentially expressed genes during anther development in genetic male sterile and wild type cotton by digital gene-expression profiling. *BMC Genomics*, 14(97). doi: 10.1186/1471-2164-14-97.

[145] Whittall, J., Medina-Marino, A., Zimmer, E., & Hodges, S. (2006). Generating single-copy nuclear gene data for a recent adaptive radiation. *Molecular Phylogenetics and Evolution*, 39, 124–134.

[146] Wilson, Z. & Zhang, D. (2009). From arabidopsis to rice: Pathways in pollen development. *Journal of Experimental Botany*, 60, 1479–1492.

[147] Wuest, S., O'Maoileidigh, D., Rae, L., Kwasniewska, K., Raganelli, A., Hanczaryk, K., Lohan, A., Loftus, B., Graciet, E., & Wellmer, F. (2012). Floral organ specification by ap3 and pi. *Proceedings of the National Academy of Sciences*, 109(33), 13452–13457. doi: 10.1073/pnas.1207075109.

[148] Yamada, T., Yokota, S., Hirayama, Y., Imaichi, R., Kato, M., & Gasser, C. (2011). Ancestral expression patterns and evolutionary diversification of yabby genes in angiosperms. *The Plant Journal*, 67(1), 26–36. doi: 10.1111/j.1365-313X.2011.04570.x.

[149] Zambelli, F., Pesole, G., & Pavesi, G. (2014). Using weeder, pscan, and pscanchip for the discovery of enriched transcription factor binding site motifs in nucleotide sequences. *Current Protocols in Bioinformatics*. doi: 10.1002/0471250953.bi0211s47.

[150] Zhao, S., Fung-Leung, W., Bittner, A., Ngo, K., & Liu, X. (2014). Comparison of rna-seq and microarray in transcriptome profiling of activated t cells. *PloS ONE*, 9(1), e78644. doi: 10.1371/journal.pone.0078644.

[151] Zik, M. & Irish, V. (2003). Global identification of target genes regulated by *APETALA3* and *PISTILLATA* floral homeotic gene action. *Plant Cell*, 15(1), 207–222. doi: 10.1105/tpc.006353.

[152] Ó'Maoiléidigh, D., Wuest, S., Rae, L., Raganelli, A., Ryan, P., Kwasniewska, K., Das, P., Lohan, A., Loftus, B., Graciet, E., & Wellmer, F. (2013). Control of reproductive floral organ identity specification in arabidopsis by the c function regulator agamous. *The Plant Cell*, 25(7), 2482–2503. doi: 10.1105/tpc.113.113209.

Chapter 4

The Ecological Function of Staminodes in *Aquilegia*

4.1 Introduction

Ecologists ask questions about the interactions between organisms and their environments. Field-based education addressing ecological questions provides a link between the textbook and the natural world, providing a space for students to learn how to ask questions, conduct experiments, and analyze data. Conducting research provides a way for students to make connections between the different parts of the scientific process and leads to richer understandings of the discipline (Burrow, 2018). Field-based classes allow for scaffolded learning where the instructor can provide guidance and feedback and also leave opportunities for choice in student projects and approaches. Undergraduates who participate in research develop skills in critical thinking, communication, and analytic skills, and show increases in confidence

(Lewinsohn et al., 2015; Davis, 2009). In this chapter, I will provide an overview of the benefits of field-based learning as well as present my own ecological experiment as a case study of the challenges and important aspects to consider when teaching field-based learning.

4.1.1 Benefits of Field Based Learning

Making connections to the scientific process is an integral part of any scientific classroom, and is achieved not just by learning about the steps in a textbook but by practicing integrating concepts and by knowing when to apply skills (Ambrose et al., 2010). Problem-based learning with a student-centered approach is known to improve student attitudes towards biology by introducing concepts within a relevant context and by encouraging inquiry (French & Russell, 2006). Field-work projects allow for check-points at each stage of the process where instructors can assess student understanding and learning. For example, assignments can be broken into scaffolded projects with check-points after selection of topic and background research (including how to read scientific literature), formulation of scientific questions, experimental design, experimental execution, experimental analysis, and scientific writing.

At each stage, the student has some amount of freedom and ownership over their work, which are known factors to increase student motivation in a course (Patall et al., 2010). By breaking up the large experiment into multiple sections, students are not overwhelmed by the breadth of the project and can set realistic performance goals with multiple opportunities for feedback. Should experiments yield uncertain results,

division of the project into multiple sections can minimize the focus on results and reward components such as literature review, data analysis, discussion, and presentation.

There are numerous factors that can contribute to anxiety about science in the classroom, including student misconceptions about the scientific process due to the way it is taught in high school education (with an emphasis on memorization and not on experimentation), anxiety about the process of critical thinking due to lack of experience, and preconceptions about their abilities to do scientific research (Mallow, 2006). Field-based education not only provides students opportunities to learn via inquiry and critical thinking, but it also promotes an increase in confidence in ability "to do" science and "to be" scientists. Beyond field skills, students gain skills in working cooperatively, communicating their findings, and in organization and time management during the experimental process. Exposure to research can also lead students to gain confidence in conducting undergraduate research, and better prepare them for graduate school (Beck & Blumer, 2012; Laursen et al., 2006).

However, there are challenges to field education ranging from institutional hurdles such as limited resources (costs or lack of local habitats), time constraints, uncertain outcomes, and difficulty in analyzing results (Fleischner et al., 2017; Openshaw & Whittle, 1993). Depending on the type of ecological study students will be conducting, there are a range of challenges as well. Studies may fall into one of our categories: non-manipulative, non-hypothesis testing (primarily descriptive); non-manipulative, hypothesis testing; manipulative, non-hypothesis testing, and manipulative, hypothesis testing (Wilson; 2009). The last category is the most

"experimental", and requires careful thought about proper controls, randomization of treatments, which variables to measure, and how many replicates are necessary for the experiment. Instructors should guide their students to select feasible projects, and also determine which types of data analyses are appropriate given the level of the students. Additionally, in the field researchers often have to cope with unexpected challenges, and helping students think through as many potential hurdles in advance may better prepare them during their experiments.

Beyond dissemination of knowledge, instructors who engage in problem-based or field-based learning must use teaching techniques different from those in lecture based classes. Particular emphasis ought to be paid to making connections between different aspects of field work, and instructors should take care to maintain the "investigative" aspect of field-based learning (Openshaw & Whittle, 1993). This requires a time investment on the part of the instructor - both to locate information and to read said information about student centered techniques.

To ensure equitable participation and preparation, workshops or time in class should be dedicated to covering the steps of each part of the field-based project. If students engage in group-work for their research projects, the instructors should be clear about the break-down of grading, and ensure that there are clear tasks that allow for fair division of labor, as well as reward individual input (Davis, 2009). Resources should also be provided to course instructors, for example workshops on student-centered learning and on informal assessments, or with mentorship (Nordlund, 2016). Instructor motivation and reward can be incentivized by publishing results of classroom ecological studies in journals such as "Teaching Issues and

Experiments in Ecology" (TIEE), a peer-reviewed publication designed for faculty interested in linking pedagogy and ecology (D'Avanzo et al., 2006).

4.1.2 Examples of Field-based learning in the classroom

The topics and types of questions that can be addressed in a field-based unit depend on resources available for study as well as feasibility of the projects. For example, field-based learning may be as simple as a field trip to observe the habitats and phenomena covered in lecture. A group of secondary school students from Nigeria who were taken on field trips to a school farm, pond, and stream to learn ecology outperformed students who learned the same concepts but only in the classroom (Hamilton-Ekeke, 2007). The act of being in an ecosystem allows students to observe connections between differing concepts they have learned about in class.

If limited to staying within a classroom, teachers have several options of case studies or open sourced data sets that can be used when teaching about ecology. One particular curriculum, "The Truth About Science", is a 40 lesson module geared towards middle school students (Steel et al., 2004). In early lessons students learn how to formulate testable questions, and differentiate between quantitative and qualitative observations. The second section of the curriculum tackles experimental design, including treatment types (treatments and controls), replication, and randomization. Thirdly, the curriculum covers data analysis and interpretation, focusing on T-Tests and p-values. Finally, the students participate in peer review of each other's work. Teachers were provided with professional development workshops to better prepare themselves, and feedback from those teachers included astonishment

that their middle school students could complete those long-term projects (Steel et al., 2004). This type of scaffolded learning with support for both students and teachers resulted in students learning about statistical methods while conducting their own original research.

It is important for teachers to keep in mind proximity of field sites, and to take advantage of local sites. A professor from Central Connecticut State University designed a class-wide project where students were given a hypothetical scenario where a portion of a real-life park was under consideration for development. The assignment for students was to assess the park (acting as environmental consultants) and determine locations that could be sold with the fewest ecological repercussions (Tessier, 2004). Over the course of the semester, students were given choice in determining group leadership and decision-making, as well as which parameters to use in assessing the value of different park locations. Students then learned about different sampling schemes, and were given a choice in selecting the appropriate assessment methods. Finally, time was dedicated to learning about statistical methods such as Analysis of Variance and chi-square. In this case, the overarching topic was given, but students had multiple opportunities for choice in their assessment process. With this assignment, students connected ecological principles to their own lives.

4.1.3 This study as a case study for field work

In this chapter, I will use my own experience designing and executing a field study to outline the importance of proper controls, feasibility, and data analysis when leading field-based courses. At each step, I will review how a project can be adjusted

to be led at levels appropriate for high school biology, AP biology, or a higher level undergraduate course.

At any of these student levels, the first step in a field-based assignment is to identify an ecological question. In my project, I was interested in investigating whether the staminodes of *Aquilegia* serve an ecological function for the flower.

Staminodes, which are sterile stamens, commonly evolve during reduction of the androecium, and become vestigial organs and are quickly lost if they do not experience positive selection. However, there are examples of staminodes that adopt novel functions and persist in certain angiosperm lineages (Walker-Larsen & Harder, 2000). For instance, removal of staminodes from *Penstemon* flowers in the field led to a decrease in pollen receipt and removal by bee pollination, likely by increasing contact with individual pollinators (Walker-Larsen & Harder, 2001).

Aquilegia staminodes are found in all but one species, the high altitude *Aquilegia jonesii*. Given that *Aquilegia* sepals and petals are showy and provide pollinator rewards, staminodes are not thought to have roles in pollinator attraction. The staminodes are also shorter than the stamens at the time of anthesis, and therefore are incapable of serving as barriers to self-pollination. However, since the staminodes form a sheath that surrounds the carpels long after the other floral organs have abscised, and since there is adaxial lignification of the staminodes and an enrichment of defense gene ontology categories, I hypothesize that staminodes in *Aquilegia* are serving a protective role for the developing fruit.

Once I identified my ecological question, my next step was to design my study. I first conducted a preliminary study using transplanted *Aquilegia canadensis*

in raised beds. I removed staminodes both before and after flowers opened, and used seed set as a proxy for reproductive fitness. I hypothesized that if staminodes have a defensive role, removal of staminodes in the field would correlate with a decrease in fitness (seed set). Measures of either the number of seeds per carpel or total seeds per flower are commonly used measures of reproductive fitness (Younginger et al., 2017). Seed set encompasses pollination success, fertilization, and subsequent seed development. If a class does not have access to plants in natural environments, this type of artificial set-up is a cost-effective way to design an easily monitored experiment with simple data collection.

My first study provided a proof-of-concept that staminodes could easily be removed from flowers in the field. To account for side effects of handling flowers, I collected seeds from un-manipulated flowers (an overall control), and also seeds from flowers that were agitated before or after flowers opened (as controls for each time point). Recognizing that the artificial raised bed environment might not expose flowers to their natural pollinators, I identified three natural sites with wild *Aquilegia formosa* and *Aquilegia eximia* and conducted removal studies in the field. I faced a number of challenges working outdoors, such as flower loss due to herbivory, as well as experimental challenges in identifying which types of data from interacting factors to collect and in selecting models and statistical measures to fit the data. In this chapter, I will use my data to underscore the importance of including statistics in any field-based course, and show how different tests of the same results can lead to different interpretations of significance. I will begin with a simplistic analysis and progress to increasingly complex approaches.

4.2 Methods

4.2.1 Field Sites

Aquilegia canadensis, 2015

Thirty-three *Aquilegia canadensis* plants were selected from two different nurseries, and planted in raised beds in the experimental garden at Harvard University between May 13th and July 2nd, 2015 (4.1 A-C). Five plants did not survive transplantation, two plants dried out due to issues with the watering lines, two plants fell over due to windy conditions, and two plants produced limited numbers of flowers due to poor health. Flowers from the remaining twenty-two plants were included in the experiment. A soaker hose on a timer was set to provide daily water. Flowers on each of the plants were treated in one of six ways (Table 4.1). Each treatment was color-coded, and tape of the corresponding color was placed on the internode below the flower being treated. Initially, every plant bore flowers for every treatment, and a total of 487 flowers were tagged for study. After flower loss due to herbivory, 404 flowers remained in the study. For each of those flowers, the mature fruit was weighed and the number of seeds from each carpel was counted and recorded. During the study year (2015), the maximum daily temperature for the period May 13 - July 2nd was 88 F and the minimum daily temperature was 44 F.

Aquilegia canadensis, 2017

At the Concord field station (42.5081° N, 71.2949° W), 35 *Aquilegia canadensis* plants were selected from two nurseries and transplanted in a field at the

Table 4.1: Treatment conditions and number of flowers per treatment

Group	***Treatment***	***Aquilegia canadensis, 2015***	***Aquilegia canadensis, 2017***	***Aquilegia formosa, 2016***	***Aquilegia eximia, 2016 Site 1***	***Aquilegia eximia, 2016 Site 2***
1	Control. No manipulation to flowers.	94 initial	21 initial	90 initial	82 initial	11 initial
		7 lost (7.4%)	4 lost (19%)	4 lost (4.7%)	17 lost (20.7%)	1 lost (9.1%)
		87 final	17 final	86 final	65 final	10 final
2	Control. Flower poked and prodded; a few stamens removed after flower opens.	37 initial	21 initial	48 initial	58 initial	8 initial
		6 lost (16.2%)	16 lost (76.2%)	6 lost (12.5%)	23 lost (39.7%)	1 lost (12.5%)
		31 final	5 final	42 final	35 final	7 final
3	Control. Flower poked and prodded; a few stamens removed before flower opens.	46 initial	23 initial	63 initial	23 initial	7 initial
		11 lost (23.9%)	16 lost (70%)	19 lost (30.2%)	11 lost (47.8%)	2 lost (28.6%)
		35 final	7 final	44 final	12 final	5 final
4	Experimental. All staminode removed after the flower opens.	77 initial	35 initial	78 initial	86 initial	8 initial
		19 lost (24.7%)	21 lost (60%)	12 lost (15.4%)	28 lost (32.6%)	1 lost (12.5%)
		58 final	14 final	66 final	58 final	7 final
5	Experimental. All staminode removed before the flower opens.	83 initial	30 initial	82 initial	30 initial	9 initial
		26 lost (31.3%)	22 lost (73.3%)	21 lost (25.6%)	15 lost (50%)	1 lost (11.1%)
		57 final	8 final	61 final	15 final	8 final
6	Experimental. Partial staminode removal before the flower opens.	67 initial	27 initial	53 initial	18 initial	6 initial
		14 lost (20.9%)	21 lost (77.8%)	10 lost (18.9%)	9 lost (50%)	0 lost (0%)
		53 final	6 final	43 final	9 final	6 final
Total Initial Flowers		487	157	414	297	49
Total Final Flowers		404	57	336	194	43
* Lost refers to flowers that were not counted due to herbivory, broken branches, or rotted flowers						

station in rows of five plants each, with 1 foot between each plant. The array was enclosed by a 3-foot-tall wire fence. 157 flowers were tagged for study between May 25th and July 3rd, 2017, and 57 flowers survived (34%) herbivory and had seeds collected and counted between June 22nd and July 22nd, 2017.

Aquilegia formosa, 2016

At Landels Hill Big Creek Reserve in California (36.0699° N, 121.5990° W), 132 *Aquilegia formosa* plants were identified along both creek beds and trails and tagged with tape and numbers on their stems between May 12th and June 16th, 2016 (4.1 D-F). 414 total flowers were tagged for study, and 336 (81%) survived herbivory and had seeds collected and counted between May 23rd and July 18th, 2016.

Aquilegia eximia, 2016

Site 1: At Clear Creek Management Area (36.381823° N, 120.716737° W), 43 *Aquilegia eximia* plants were identified along the creek bed and tagged with tape on their stems between June 22nd and June 24th, 2016 (4.1 G-I). 297 total flowers were tagged for study, and 194 survived herbivory (65%) and had seeds collected and counted between July 15th-17th, 2016 and July 24th, 2016. Many plants were well into flowering at the beginning of this experiment, consequently fewer unopened flowers were available for removal of staminodes at early stages. Site 2: At Prefumo Canyon (35.2583082° N, -120.7010065° W), 17 *Aquilegia eximia* plants were identified and tagged with tape on their stems on June 26th, 2016. 49 total flowers were tagged for study, and 44 (88%) survived herbivory and had seeds collected and

counted on July 24th, 2016. Due to the small number of flowers collected, this site was excluded from further analysis.

4.2.2 Control and Experimental Treatments

As outlined in Table 4.1, I divided flowers into one of six treatment conditions. In order to understand the breadth of naturally occurring variation in seed set, I collected seeds from unmanipulated flowers (Group 1).

Both abiotic (non-living factors e.g. drought, temperature) and biotic (e.g. pathogens, herbivores) stressors can impact seed sets in plants. As I was physically agitating flowers by opening them and removing staminodes, I included two control conditions to account for effects of physical stress on seed set. At each of my two developmental stages of manipulation (before the flower opened and after the flower opened), I included a control condition (Groups 2 and 3) where flowers were lifted, prodded with forceps, and 1-3 stamens were removed (to account for accidental removal of stamens in some flowers during removal of staminodes).

I chose three experimental conditions. In group 4 I removed all staminodes after flowers opened. By this stage, the staminodes formed an interlocking sheath surrounding the carpels, and were easily removed as a sheath. In group 5 I removed all staminodes before the flowers opened. By applying pressure to the spurs of the petals I was able to push the petal blades open and reach between the stamens and carpels to remove individual staminodes using forceps. At these early stages staminodes were easily removed individually. In order to test whether there were a full whorl was necessary to promote full seed set, I included a condition (group 6)

Figure 4.1: Field sites and species. (A) Raised beds and *Aquilegia canadensis* plants at Harvard University. (B) Mature *Aquilegia canadensis* flower. (C) Young *Aquilegia canadensis* flower. By applying pressure to the lower petals, the top of the flower would open up and I was able to remove the staminodes with forceps. (D) Population of *Aquilegia formosa* at Landels-Hill Big Creek Reserve. (E) Mature Aquilegia formosa flower. (F) Young *Aquilegia formosa* flower. (G) *Aquilegia eximia* plant at Clear Creek Management Area. (H) Mature *Aquilegia eximia* flower. (I) Young *Aquilegia eximia* flower.

where I partially removed staminodes prior to the flowers opening.

4.2.3 Statistical analysis

Simple summary statistics

In order to visualize my dependent variable (total number of seeds) and my independent variable (treatment), I used the R package ggplot2 and generated boxplots of each of our datasets. Student's T-tests were conducted using RStudio in order to compare the means of experimental groups versus respective control groups. As RStudio is an open source interface, I recommend its usage for biology courses (RStudio Team). For an example of the R code used to produce the analyses in this chapter, see Supplemental Data B.1.

Linear Models

I fit poisson and negative binomial distributions to my data using the car and MASS packages in Rstudio. I selected the poisson distribution as appropriate for fitting my mixed models.

Advanced Mixed Models

Models with both fixed and random factors were tested using lme4. Optimal models were identified by using anova to compare between full and partial models with droppped factors. Anova was used to test for treatment effect in the full model. The lsmeans package was used to predict the total number of seeds in my final

models. Finally, the multcomp package was used for pairwise comparisons between my control and experimental groups.

4.3 Results

4.3.1 Data exploration

As a first step, I generated boxplots in order to visualize the range of variation of my dependent variable, (total number of seeds per flower) for each of my treatment groups (Figure 4.2). I used two-sample T-tests to compare means between each of my treatment groups (Supplemental Tables B.1, B.2, B.3). According to Next Generation Science Standards for high school students, high school biology students are encouraged to compare datasets for consistency, and to use digital tools to apply basic summary statistics (NGSS Lead States). The visualization and simple summary statistics outlined in this section would be executable in an introductory course.

Overall, the un-manipulated controls in each species (*Aquilegia canadensis, Aquilegia formosa, and Aquilegia eximia*) had the highest mean and median number of seeds. However, I observed differing trends in each of the datasets. In *Aquilegia canadensis*, seed set in flowers with staminodes removed late were significantly lower ($p < 0.05$) than both the un-manipulated controls and control flowers agitated after the flowers opened (Supplemental Data B.1). In *Aquilegia formosa*, seed set in flowers with staminodes removed early was significantly lower ($p < 0.05$) than corresponding controls (Supplemental Data B.3). Finally, in *Aquilegia eximia* both the experimental and manipulated control groups were significantly lower than unmanipulated controls,

but there were no significant differences between the experimental and manipulated controls (Supplemental Data B.2).

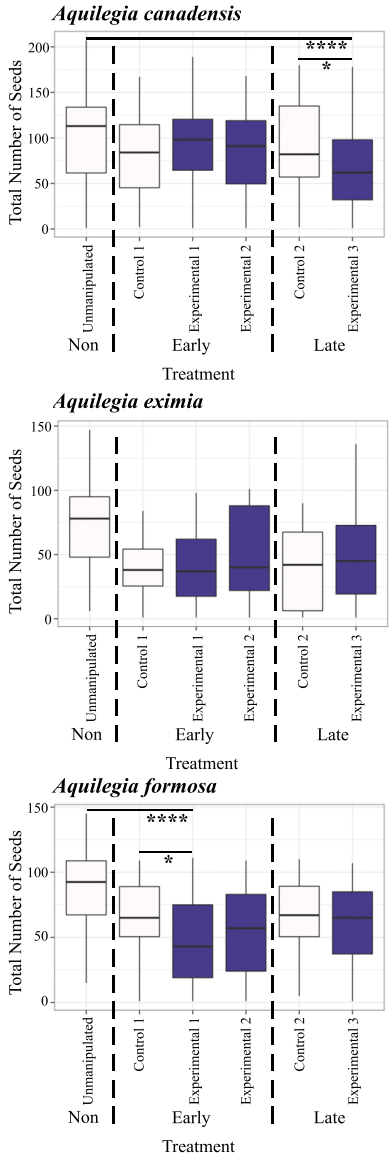


Figure 4.2: Box plots showing the total number of seeds per flower for each of my treatment conditions. 404 flowers were analyzed for *Aquilegia canadensis* in 2015. 336 flowers were analyzed for *Aquilegia formosa* in 2016. 194 flowers were analyzed for *Aquilegia eximia* in 2016. White bars, control groups. Purple bars, experimental groups. Control 1, agitated bud-stage flowers. Experimental 1, staminodes removed from bud-stage flowers. Experimental 2, partial removal of staminodes from bud-stage flowers. Control 2, agitated mature flowers. Experimental 3, staminodes removed from mature flowers.

T-tests are useful for pairwise comparisons between means of groups, and can be quickly done in excel or R. The main principal for student understanding is that they are comparing two groups and looking to either accept or reject a null hypothesis - the null hypothesis being that there is no difference between a control and experimental group. However, t-tests can only be used to compare the means of two groups at a time. Additionally, this type of statistical test cannot control for environmental or other random effects. For example, while the typical number of carpels per flower in *Aquilegia canadensis* is 5, I collected seeds from one flower with as many as 12 carpels (Figure 4.3). In order to explore if increases in carpel number were correlated with increases in the total number of seeds per flower, I generated a scatterplot of data from un-manipulated flowers (Figure 4.3). As the number of carpels increases, there is a weak increase in total number of seeds per flower. This indicates that the number of carpels is a factor that we must take into account. Additionally, the slopes of my regression lines differ widely, indicating that other factors are also contributing to variation in my dataset. As my experiment requires comparison of multiple groups with multiple explanatory variables, I turned to more complex methods of analysis.

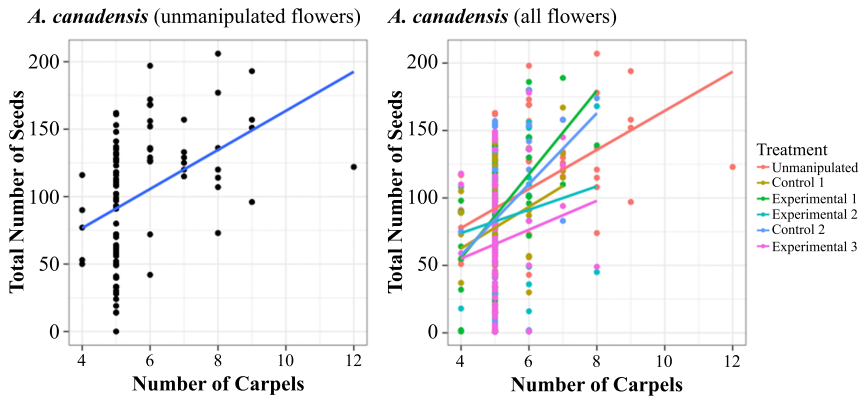


Figure 4.3: Scatterplots of the total number of seeds per flower (x-axis) flowers plotted against the number of carpels (y-axis). Left: Data from unmanipulated flowers of *Aquilegia canadensis*. Right: Data from all treatment conditions of *Aquilegia canadensis*. Lines represent fitted regression lines. Control 1, agitated bud-stage flowers. Experimental 1, staminodes removed from bud-stage flowers. Experimental 2, partial removal of staminodes from bud-stage flowers. Control 2, agitated mature flowers. Experimental 3, staminodes removed from mature flowers.

4.3.2 Probability Distribution

In the previous section, the only data necessary to include was our dependent variable, the total number of seeds, and my independent variable, the treatment. After visualizing the data I observed a large amount of variability within each treatment group. In the following sections I will use statistical models to predict expected patterns of seed output given a treatment condition. I envision linear mixed models as analyses appropriate for a high school AP biology or introductory biology undergraduate course, and generalized linear mixed models as analyses appropriate for upper level undergraduate biology courses.

In order to make informed predictions and incorporate differing explanatory variables into a model, I must include a "stochastic" portion component - a metric that will estimate variation. Variation in datasets can be due to either measurement or

process error. Measurement error is variability due to measurements and not due to the environmental system, while process errors are unmeasured but real sources of variability affecting the environmental system. Probability distributions are used to represent the random variability around the mean, and are critical for fitting models to datasets. Consequently, my first step prior to fitting models was to identity our probability distribution.

While selecting which distributions to fit to my data I narrowed down the options based on my data types. My response variable ("total number of seeds per flower") was simple count data with a non-normal distribution. Since I measured seeds, my data were discrete non-negative integers. Count data limits the available probability distributions to either poisson or negative binomial distributions. I fit each of these probability distributions to my response variable (Figure 4.4). Poisson distributions plot the number of events with the assumption that each event is independent. Concurrent with my expectations, the majority of the data points fit between the curves in the Poisson distribution, and so I used Poisson (log link) in my model generation.

4.3.3 Fitting linear mixed models to the data

Once I had identified my probability distribution I could move forward with fitting models to my datasets. Models are functions that take user-specified parameters and return predicted datasets. I have a choice between fitting simple and complex models with my data. I will begin with a simple model focusing only on our fixed effects (treatment), and in the next section will test models including random

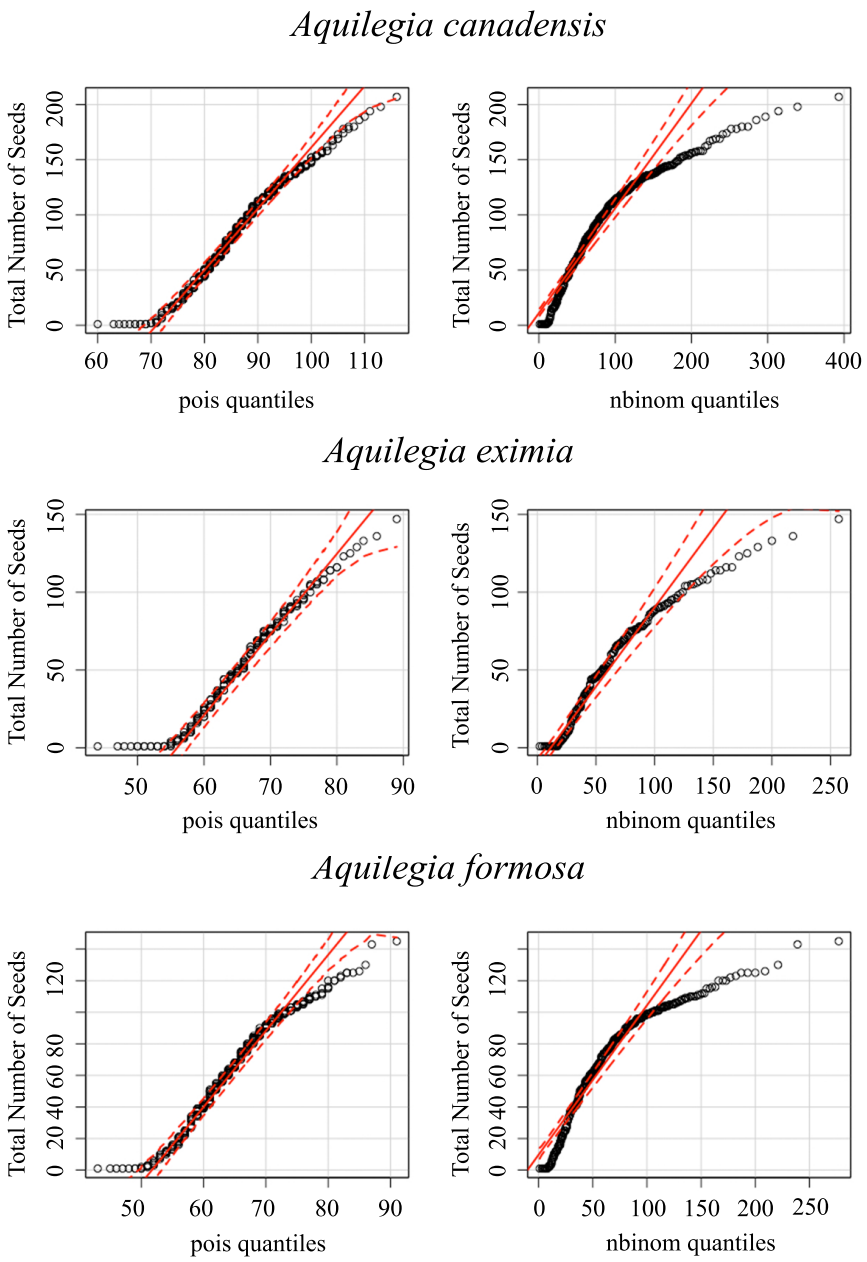


Figure 4.4: Probability distributions for each dataset Left, poisson. Right, negative binomial.

effects.

Linear regressions allow me to use treatment type to predict the total number

of seeds. Linear model output in R provides information about how well the model fits, and if I can reject the null hypothesis that treatment has no effect on seed set. F-statistic numbers that are larger than 1 indicate that there is a relationship between the dependent and independent variable. R-squared statistics measure the linear relationship between the dependent and independent variable on a scale between 0 and 1: numbers closer to 0 indicate that the independent variable does not explain the variance in the dependent variable, while numbers closer to 1 indicate that the independent variable explains a high percentage of variability.

In each of my linear models, treatment did have a significant effect on the total number of seeds (Supplemental Table B.4). For the *Aquilegia canadensis* data, $F(5, 398) = 5.279$, $p = 0.001041$ and treatment accounted for 5% of the explained variability in seed set. For the *Aquilegia eximia* data, $F(5, 191) = 6.739$, $p = 8.28e\text{-}06$ and treatment accounted for 12.8% of explained variability in seed set. Lastly, for *Aquilegia formosa*, $F(5, 336) = 14.96$, $p = 2.912e\text{-}13$ and treatment accounted for 17% of the explained variability in seed set.

While we can see that treatment does have a significant effect, the low R-squared values indicate that I am missing factors that explain large amounts of variability in our datasets. I knew that increasing the number of carpels led to a slight increase in seed set, making it likely that the number of carpels explained a portion of variation in our data. Random effects such as the plant a flower was from, or the individual flower itself could also account for variation.

4.3.4 Fitting generalized linear mixed models to the data

I chose to use generalized linear mixed models (GLMM) to fit my data because these types of models combine aspects of linear mixed models (such as their ability to incorporate random effects) and generalized linear models (which can be fit to non-normal data) (Bolker et al., 2009). My data had both fixed and random effects, and was non-normally distributed, consequently GLMM's were a good fit. GLMMs are widespread for use with ecological studies.

I followed the general framework outlined in Bolker et al., 2009 to construct my full model. I designated "treatment" as my fixed effect, while my random effects included which plant the flowers were from, inflorescence branch, and the date the treatment was added. The number of carpels was designated as an offset. My random effects were associated with individual experimental units drawn at random from the population, and accounted for variation between groups that might affect the response (mathworks). My full model showed overdispersion of my data - variation was higher than would be expected. To account for this, I included one random effect level for each flower.

I used the Laplace approximation for my model construction. For each dataset, I tested a full model including all of the random effects as well as reduced models dropping each individual random effect. The inflorescence branch location and date added contributed minimally to variation, and we dropped the effects from our models, while the plant a flower was from and the individual flower accounted for a large amount of variation [*]

[*] *Aquilegia canadensis*: flower accounted for 0.9048 of variance; plant accounted for 0.2932 of

Table 4.2: GLMM models for *Aquilegia canadensis, Aquilegia eximia* and *Aquilegia formosa*

Model	Response	Number of flowers	AIC	Fixed effects	Estimate	Std. Error	ChiSq	Pr(>Chisq)
Aquilegia canadensis	Total number of seeds	404	4585.1	Treatment	-1.1955	0.1528	9.8064	8.09E-02
Aquilegia eximia	Total number of seeds	197	4325.7	Treatment	-0.75917	0.14056	474.23	2.2e-16 ***
Aquilegia formosa	Total number of seeds	342	3539.6	Treatment	-0.50685	0.09759	43.254	3.282e-08 ***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1

I used likelihood ratio tests and Akaike information criterion (AIC) to select our final models (Supplemental Table 4). Models that excluded the number of carpels as an interaction had slightly lower AIC values (data not shown), but I chose to leave in the number of carpels as an offset.

Consistent with what I saw in my linear model, my fixed factor, treatment, had a significant effect on both *Aquilegia eximia* and *Aquilegia formosa* data sets using GLMM models (Table 4.2). I used lsmeans in order to display the predicted number of seeds under each treatment condition for each of those models (Figure 4.5, Table 4.3). Inclusion of random effects via GLMM models lessened the amount of variance within each of my treatment groups compared to my raw data.

I used multcomp to do pairwise comparisons between the associated control and experimental groups of our *Aquilegia eximia* and *Aquilegia formosa* data sets and test if, given the parameters of our models, there were differences between the controls and experimental groups. Multiple comparisons of means for my *Aquilegia eximia* dataset revealed that partial removal of staminodes early in development did

variance. *Aquilegia eximia*: flower accounted for 0.3321 of variance; plant accounted for 0.2916 of variance. *Aquilegia formosa*: flower accounted for 0.4724 of variance, plant accounted for 0.1932 of variance.

have a significant effect on seed set. For my *Aquilegia formosa* dataset, there were no significant differences between our experimental conditions and their corresponding controls.

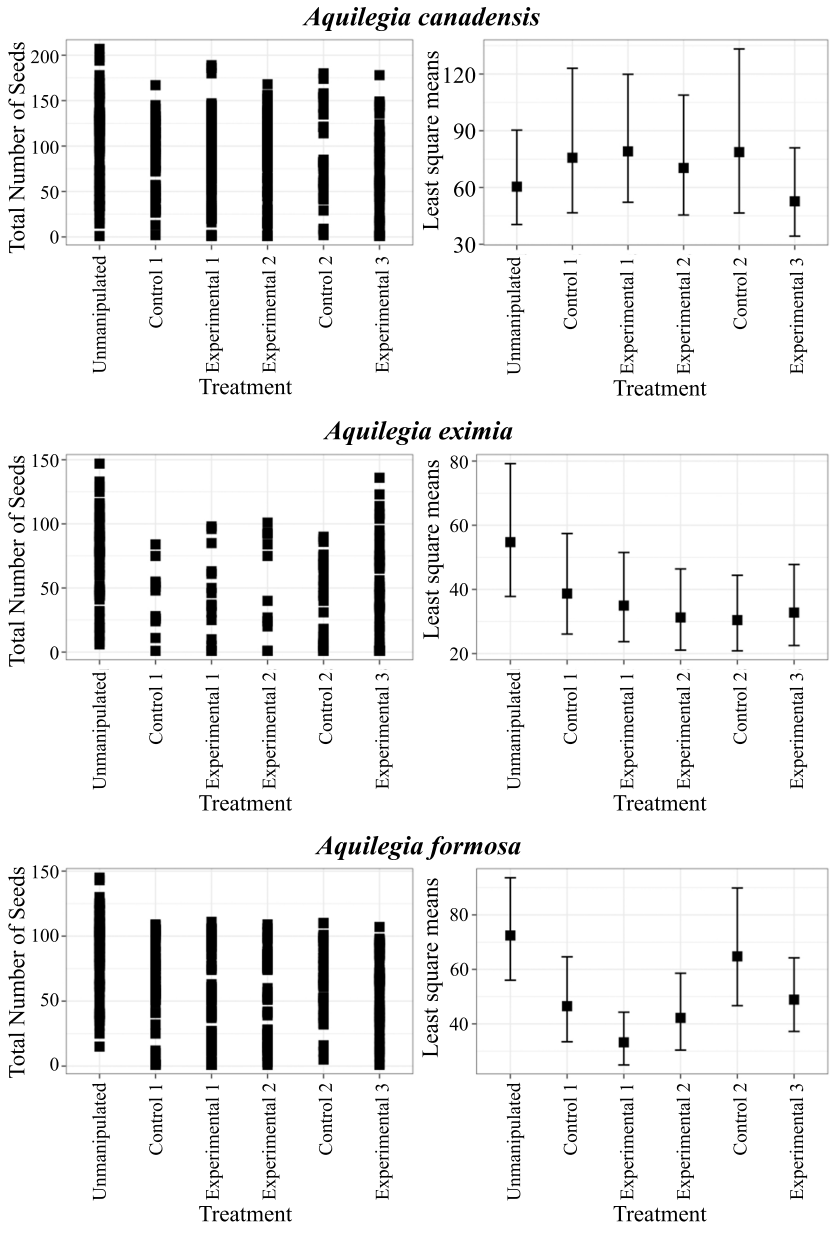


Figure 4.5: Actual and predicted seed output from each treatment condition. Left: Actual raw data values. Right: Predicted values using lsmeans. Squares indicate predicted mean values. Control 1: agitated before the flower opens. Experimental 1: full staminode removal. Experimental 2: partial staminode removal. Control 2: agitated after the flower opens. Experimental 3: full staminode removal.

Table 4.3: Predicted total number of seeds for *Aquilegia canadensis*, *Aquilegia eximia*, *Aquilegia formosa* using lsmeans. Control 1, agitated bud-stage flowers. Experimental 1, staminodes removed from bud-stage flowers. Experimental 2, partial removal of staminodes from bud-stage flowers. Control 2, agitated mature flowers. Experimental 3, staminodes removed from mature flowers.

Aquilegia canadensis			
Treatment	**Predicted number of seeds**	**Std. Error**	**95% Confidence level**
Unmanipulated	60.43557	9.233466	44.79581 -81.53365
Control 1	75.74045	13.967805	52.76509-108.71795
Experimental 1	79.09028	12.496282	58.02735-107.79932
Experimental 2	70.335	11.673937	50.80360 -97.37695
Control 2	78.72664	15.749804	53.19066-116.52392
Experimental 3	52.69771	8.605708	38.26304 -72.57561
Aquilegia eximia			
Treatment	**Predicted number of seeds**	**Std. Error**	**95% Confidence level**
Unmanipulated	54.72117	7.692	41.54401-72.07793
Control 1	38.70636	5.803104	28.85125-51.92779
Experimental 1	34.95595	5.152555	26.18502-46.66480
Experimental 2	31.2627	4.687576	23.30215-41.94275
Control 2	30.43785	4.368916	22.97398-40.32661
Experimental 3	32.79069	4.686928	24.77903-43.39271
Aquilegia formosa			
Treatment	**Predicted number of seeds**	**Std. error**	**95% Confidence level**
Unmanipulated	72.43054	7.06846	59.82096-87.69807
Control 1	46.48379	5.81842	36.37105-59.40831
Experimental 1	33.22843	3.628502	26.82624-41.15854
Experimental 2	41.18871	5.265581	33.03373-53.88091
Control 2	64.7705	8.061098	50.75044-82.66366
Experimental 3	48.89522	5.068417	39.90542-59.91022

Table 4.4: Multiple Comparisons of Means: Tukey Contrasts. Control 1, agitated bud-stage flowers. Experimental 1, staminodes removed from bud-stage flowers. Experimental 2, partial removal of staminodes from bud-stage flowers. Control 2, agitated mature flowers. Experimental 3, staminodes removed from mature flowers.

Aquilegia eximia				
Contrast	Estimate	Std. Error	Z value	Pr(<|z|)
Control 1 - Experimental 1	0.10191	0.07115	1.432	0.3702
Control 1 - Experimental 2	0.21358	0.07905	2.702	.0201 *
Control 2 - Experimental 3	-0.07446	0.04003	-1.86	0.1679
Aquilegia formosa				
Contrast	Estimate	Std. Error	Z value	Pr(<|z|)
Control 1 - Experimental 1	0.3357	0.14826	2.264	0.066
Control 1 - Experimental 2	0.09695	0.16058	0.604	0.894
Control 2 - Experimental 3	0.28117	0.14698	1.913	0.15

4.4 Discussion

4.4.1 Interpreting the data

This was the first study to address the function of *Aquilegia* staminodes in the field. The formation of a sheath surrounding the carpels led to a generally accepted hypothesis that the staminodes served a protective role during seed development (Walker-Larsen & Harder, 2000; Brayshaw, 1989). This was supported in Chapter 2 with our identification of asymmetric lignification on the adaxial epidermis of staminodes surrounding the carpels, and in Chapter 3 by our identification of gene ontology categories related to defense. In this chapter, although initial data exploration indicated weak reductions in seed sets associated with removal of staminodes at different stages for both *Aquilegia canadensis* and *Aquilegia formosa*, these differences were not significant once models that incorporated carpel number, plant number and flower number were included. Surprisingly, although initial exploration of the data from *Aquilegia eximia* indicated that there were no significant differences between the control and experimental groups, my GLMM analysis revealed a slightly significant decrease in seed set after partial removal of staminodes. However, *Aquilegia eximia* had the highest loss of flowers from herbivory and there were fewer bud-stage flowers available for manipulation and so this group had the smallest sample size (9 flowers) which may have skewed the results. Overall, in each of the three data sets, all controls and experimental treatments produced lower seed sets than unmanipulated flowers, indicating that agitating flowers may trigger a stress response that negatively impacts reproductive fitness.

Each of the species included in this study originated within the last 3 MYA after the radiation of *Aquilegia* in North America (Fior et al., 2013). It is possible that staminodes served a function in Aquilegia at their place of origin in Asia, and do not face selective pressures to maintain staminodes in America. *Aquilegia jonesii*, an early diverging species of the American clade, is the sole member of the genus to have lost staminodes. Although no other American species have lost staminodes, this may be due to the short amount of evolutionary time that has passed since the radiation in America. Another possibility is that staminodes served a function at the time of origin that is no longer needed today. If this is the case, staminodes of European and Asian species would also be expected to be vestigial and eventually lost.

There is also the possibility that seed set was too simple of a measure. Staminodes form a sheath surrounding carpels during seed development. Formation of this sheath could influence temperature of seed maturation at early stages, which could impact seed dormancy and germination. Although seed set is an accepted measure of fecundity, it does not measure fitness of offspring. During my pilot study I chose not to keep seeds after counting them - as I was conducting the study in an artificial environment, I theorized that there would be more self-pollination than outcrossing. *Aquilegia* has high inbreeding depression and we worried that taking that into account would complicate measuring fitness in offspring (Montalvo, 1994). However, looking back, it may have been valuable to collect the seeds and measure germination rates from each treatment group compared to unmanipulated flowers.

Although their thin and papery form does not deter larger herbivores,

staminodes could have a role in antimicrobial activity. Leaves from *Aquilegia vulgaris* have been shown to have antimicrobial activity against bacteria and fungi (Bylka et al., 2004), but floral tissue remains untested. Genes identified as enriched in staminodes in the RNA-seq experiment detailed in Chapter 3 have roles in defense responses to fungi, bacterium, wounding, and environmental stress, which could support a primarily biochemical function. Of course, the physical and biochemical functions are not mutually exclusive. Moreover, there is the possibility that staminodes enhance mutualist microbial communities, by providing a physical space for growth that remains long after the other floral organs abscise. Future experiments that explore microbial communities on each of the floral organs of *Aquilegia* as well as on carpels from manipulated flowers could address these possible functions.

4.4.2 What I would change about the experiment

While conducting each of these experiments I identified aspects of either experimental design or set-up that I wished I had executed differently. For example, during the preliminary study using *Aquilegia canadensis*, I used different colors of tape to label flowers with different treatment conditions. During the course of this experiment I realized that this could have an effect on pollinators, and for my subsequent two experiments I used only green tape to match the stem colors. However, I cannot discount that this could have impacted pollinator visitation.

Although I used a random number generator to assign treatment conditions, at my *Aquilegia eximia* sites flowers were late in their season and I had limited numbers of unopened flowers to treat. Furthermore, at each site I collected as many

unmanipulated flowers as possible, in order to get a sense for natural variation in seed set. This led to drastically uneven sample sizes for each condition, and did artificially reduce the number of flowers "lost" from the unmanipulated category.

At each site, I lost a number of flowers to herbivory. During my preliminary study, I lost a substantial number of flowers to rabbit herbivory. At my natural sites I had no enclosed fencing, and at my field station site I used fencing that was 3 feet tall, but I still lost the majority of my flowers. If this experiment is repeated in the future, I recommend either setting up full cages surrounding the plants that are buried well into the ground to prevent burrowing, or bagging manipulated flowers using mesh netting after pollination.

4.4.3 Tying this study in with teaching field studies

Although this was a simple study in terms of treatments and techniques, it did not provide definitive results. This study highlights the difficulties in both interpreting ecological data but also the importance of knowing which statistical measures to use. When designing experiments for students, it is important to consider their background knowledge of statistics and statistical tools, and also to consider the amount of time in the classroom that will be dedicated to both the experiment as well as analyses. If students have not had a background in statistics, than it might be appropriate to teach them about different types of analyses, but only expect them to try basic methods on their own datasets.

Whether in introductory biology, AP biology, or undergraduate biology level courses, students can conduct a simple experiment and collect data. By visualizing the

data in this experiment with boxplots, it was immediately clear that there was a large amount of natural variation in seed set. It was also apparent that the act of manipulating flowers in any way is correlated with a decrease in seed set. While the generalized linear mixed models provided a way to explain portions of the variation I observed, and revealed that what appeared initially to be significant differences were not significant, students in introductory courses can still benefit from the simple methods to learn about data collection and descriptive statistics to predict trends.

Feasibility is a primary factor when considering including field work in a course. In this project, I successfully collected and analyzed data from raised beds and from natural field sites. I lost the majority of my plants from my transplanted field site, which is an important factor to consider when preparing field sites for a course. Ensuring that a site is easily accessible and is protected from herbivores (unless herbibory is a factor of study) is essential for ultimately collecting data from a site.

In conclusion, while the function of the staminodes is still elusive, in this chapter I have described different ways that this type of research could be linked to biology courses for students at three different levels and highlighted the statistical methods needed for analysis of data at each of those levels.

References

[153] Ambrose, S., Bridges, M., DiPietro, M., Lovett, M., & Norman, M. (2010). *How Learning Works: Seven Research-Based Principles for Smart Teaching*. Jossey-Bass Publishing.

[154] Beck, C. & Blumer, L. (2012). Inquiry-based ecology laboratory courses improve student confidence and scientific reasoning skills. *Ecosphere*, 3(12), 1–11. doi: 10.1890/ES12-00280.1.

[155] Bolker, B., Brooks, M., Clark, C., Geange, S., Poulsen, J., Stevens, M., & White, J. (2009). Generalized linear mixed models: a practical guide for ecology and evolution. *Trends in Ecology and Evolution*, 24(3), ISSN 0169–5347. doi: 10.1016/j.tree.2008.10.008.

[156] Brayshaw, C. (1989). Buttercups, waterlilies, and their relatives in british colombia. *The Plant Cell*, (pp. 1–253).

[157] Burrow, A. (2018). Teaching introductory ecology with problem-based learning. *Bulletin of the Ecological Society of America*, 99(1), 137–150.

[158] Bylka, W., Szaufer-Hajdrych, M., Matlawska, I., & Goslinska, O. (2004). Antimicrobial activity of isocytisoside and estracts of *Aquilegia vulgaris* l. *Letters in Applied Microbiology*, 39(1), 93–97. doi:10.1111/j.1472-765X.2004.01553.x.

[159] D'Avanzo, C., Grant, B., Morris, D., Musante, S., Taylor, J., Riney, J., & Udovic, D. (2006). Design and evaluation of tiee, a peer-reviewed electronic teaching resource. *Frontiers in Ecology and the Environment*, 4(4), 189–195. doi: 10.1890/1540-9295(2006)004[0189:DAEOTA]2.0.CO;2.

[160] Davis, B. (2009). *Tools for Teaching*. Jossey-bass Publishing.

[161] Fior, S., Li, M., Oxelman, B., Hodges, S., Ometto, L., & Varotto, C. (2013). Spatiotemporal reconstruction of the aquilegia rapid radiation through next-generation sequencing of rapidly evolving cpdna regions. *New Phytologist*, 198(2), 579–592. doi: 10.1111/nph.12163.

[162] Fleischner, T., Espinoza, R., Gerrish, G., Greene, H., Kimmerer, R., Lacey, E., Pace, S., Parrish, J., Swain, H., Trombulak, S., Weisber, S., Winkler, D., & Zander, L. (2017). Teaching biology in the field: Importance, challenges, and solutions. *BioScience*, 67(6), 558–567.

[163] French, D. & Russell, C. (2006). Improving student attitudes toward biology. In J. Mintzes & W. Leonard (Eds.), *Handbook of College Science Teaching* (pp. 15–23). Arlington, VA: National Science Teachers Association Press.

[164] Hamilton-Ekeke, J. (2007). Relative effectiveness of expository and field trip methods of teaching on students' achievement in ecology. *International Journal of Science Education*, 15, 1869–1889. doi: 10.1080/09500690601101664.

[165] Laursen, S., Hunter, A., Seymour, E., DeAntoni, T., De Welde, K., & Thiry, H. (2006). Undergraduate research in science: Not just for scientists anymore. In J. In Mintzes & W. Leonard (Eds.), *Handbook of College Science Teaching* (pp. 55–66). Arlington, VA: National Science Teachers Association Press.

[166] Lewinsohn, T., Attayde, J., Fonseca, C., Canade, G., Jorge, L., Kollmann, J., Overbeck, G., Prado, P., Pillar, V., Popp, D., da Rocha, P., Silva, W., Spiekermann, A., & Weisser, W. (2015). Ecological literacy and beyond: Problem-based learning for future professionals. *Ambio*, 44(2), 154–162. doi: 10.1007/s13280-014-0539-2.

[167] Mallow, J. (2006). Science anxiety: Research and action. In J. In Mintzes & W. Leonard (Eds.), *Handbook of College Science Teaching* (pp. 3–13). Arlington, VA: National Science Teachers Association Press.

[168] Montalvo, A. (1994). Inbreeding depression and maternal effects in aquilegia caerulea, a partially selfing plant. *Ecology*, 75(8).

[169] NGSS Lead States (2013). *Next Generation Science Standrads: For States, By States.* The National Academies Press, Washington, DC.

[170] Nordlund, L. (2016). Teaching ecology at university - inspiration for change. *Global Ecology and Conservation*, 7, 174–182. doi: https://doi.org/10.1016/j.gecco.2016.06.008.

[171] Openshaw, P. & Whittle, S. (1993). Ecological field teaching: how can it be made more effective? *Journal of Biological Education*, 27(1), 58–66. doi: 10.1080/00219266.1993.9655305.

[172] Patall, E., Cooper, H., & Wynn, S. (2010). The effectiveness and relative importance of choice in the classroom. *Journal of Educational Psychology*, 102(4), 896–915. doi: 10.1037/a0019545.

[173] RStudio Team (2015). *RStudio: Integrated Development Environment for R*. RStudio, Inc., Boston, MA.

[174] Steel, E., Kelsey, K., & Morita, J. (2004). The truth about science: A middle school curriculum teaching about the scientific method and data analysis in an ecology context. *Environmental and Ecological Statistics*, 11, 21–29.

[175] Tessier, J. (2004). Ecological problem-based learning: An environmental consulting task. *The American Biology Teacher*, 66(7).

[176] Walker-Larsen, J. & Harder, L. (2000). The evolution of staminodes in angiosperms: patterns of stamen reduction, loss, and functional re-invention. *American Journal of Botany*, 87(10).

[177] Walker-Larsen, J. & Harder, L. (2001). Vestigial organs as opportunities for functional innovation: The example of the penstemon staminode. *Evolution*, 55(3), 477–487.

[178] Wilson, B. (2009). Nature as laboratory: Experiments in ecology and evolutionary biology. *Presented at the Second Biennial Conference of the Society for the Philosophy of Science in Practice*.

[179] Younginger, B., Sirova, D., Cruzan, M., & Ballhorn, D. (2017). Is biomass a reliable estimate of plant fitness? *Applications in Plant Sciences*, 5(2). doi: 10.3732/apps.1600094.

Chapter 5

Conclusion

In this dissertation, I used a variety of tools to study the morphological and genetic differences between staminodes and stamen filaments in *Aquilegia*.

5.1 Development and form of *Aquilegia* staminodes

The staminodes of *Aquilegia* are an example of stamen filament derivatives that have reverted back to a flattened and broad shape that is ancestral for lateral organs. After organ identity is established by protein complexes containing AqAP3-1, cell proliferation resulting in laminar expansion persists in staminodes. Laminar expansion is dependent on a balance of adaxial/abaxial gene expression, and radialized stamen filaments are either adaxialized or abaxialized. In the case of *Aquilegia*, my RNA-sequencing found higher expression of the adaxial identity gene *REVOLUTA* in stamen filaments. In staminodes, expression of adaxial genes remained at fairly high levels early in development, but the abaxial identity YABBY

transcription factor family genes *FILAMENTOUS FLOWER (FIL)* and *CRABS CLAW (CRC)* were enriched in staminodes compared to stamen filaments. I hypothesize that the stamen filaments are radialized while the staminodes have restored a overall balance of adaxial/abaxial expression. The higher abaxial expression of *AqFIL* could account for the lateral expansion of staminodes compared to stamen filaments, and potentially the polarity differences between the adaxial and abaxial epidermal surfaces of staminodes. Viral-induced gene silencing of *AqFIL* is a logical next step for confirming the role of this transcription factor in polarity and lateral organ expansion of staminodes.

My histological studies revealed two subtly distinct whorls of staminodes, with every other staminode either curling up or around one another to form an interlocking sheath with a ruffled surface. Whether this curling morphology is genetically controlled or due to physical interactions between alternating staminodes is an open question. Removal of individual staminodes at early stages did not consistently affect the curling of adjacent staminodes, but this could have been due to artifacts from the fixation process. Furthermore, the earliest stage at which I was able to remove staminodes was at stage 11a, but histology revealed that by this time the curling pattern is already established.

We currently have no evidence for differences in organ identity between the two whorls of staminodes, but qRT-PCR to examine the expression of the *APETALA3* paralogs in alternating staminodes could be used to detect subtle differences in gene expression. There is the same caveat that organ identity is established much earlier than the stages at which we can easily remove staminodes, but in my

RNA-sequencing data I did detect differences in *APETALA3* paraglog expression between later stage staminodes and stamen filaments. If there are differences in *APETALA3-1* or *APETALA3-2* expression between alternating whorls of staminodes, these patterns would likely be detectable by comparing expression data of multiple individuals.

Lateral organ shape is determined by a combination of cell number and cell shape/elongation. While I found higher numbers of adaxial epidermal cells compared to abaxial epidermal cells across five species of *Aquilegia*, I did not do an in-depth study of cell shape differences. Measurements of cell area could reveal differences in cell size, while confocal microscopy could reveal differences in cell shape. As staminodes are only comprised of two epidermal layers and one vascular strand, next steps could also involve creating models of organs with two epidermal layers and investigating whether reproducing parameters of cell shape, number, and size is sufficient to contribute to the ruffled morphology.

In addition to higher numbers of adaxial cells, I consistently found asymmetric lignification on the adaxial surfaces of the adaxial epidermal cells in mature staminodes. Consistent with my histological observations, qRT-PCR and RNA-seq confirmed that the lignin biosynthetic pathway is enriched in staminodes. I hypothesize that changes in cell number and shape, combined with lignification on the adaxial epidermal surface forces the abaxial surface to buckle to create the ruffled abaxial surface characteristic of *Aquileiga* staminodes.

My morphological and genetic work described adaxial/abaxial cell differences as well as lignification. ChIP-sequencing of APETALA3 (AP3) and AGAMOUS

(AG) in *Arabidopsis* has shown that genes involved in adaxial/abaxial polarity are directly regulated by complexes containing AP3 and/or AG (Wuest et al., 2012). ChIP-sequencing of the APETALA3-1 and APETALA3-2 paralogs would be useful to investigate if *AqFIL* is directly regulated by AP3-1, and to identify any other genes involved in the differences downstream of organ identity.

5.2 Evolution and function

In my histological studies I used five species of *Aquilegia* - two native to North America (*Aquilegia canadensis* and *Aquilegia coerulea* 'origami'), two from Europe (*Aquilegia alpina* and *Aquilegia vulgaris*), and one from Asia (*Aquilegia flabellata*). Asymmetric lignification was consistent across the five species of *Aquilegia*, but was not found in the sister genera *Semiaquilegia* or *Urophysa*, each of which have staminodes. *Semiaquilegia* has rudimentary staminodes that are irregular in number, and do not exhibit curling. While *Urophysa*, like *Aquilegia*, has ten individual staminodes that curl, the staminodes of *Urophysa* do not curl in an identifiable pattern.

In the previous section I hypothesized that lignification may contribute to the rigidity of the adaxial epidermal surface and create tension leading to buckling on the abaxial side. This rigidity may also be critical to providing structure in the epidermal cells of mature staminodes once the cells have undergone programmed cell death. Staminodes are dead at maturity, and yet they remain in a sheath surrounding the carpels. Lignification may preserve the integrity of the adhered sheath. The presence of lignin may also serve as a physical block that prevents microbial deposition on the

carpel.

While lignification and cell number differences were consistent across the five *Aquilegia* species that we sampled, a broader sampling of species is necessary to investigate the evolution of lateral adhesion across *Aquilegia*. Transmission electron microscopy of the interface between staminodes would provide a closer look at the margins and could reveal differences in cuticle structure at the adhered junction of staminodes.

Finally, there is the open question of function. In Chapter 4 I attempted several field studies using various species of North American *Aquilegia*. Although there were trends of decreasing seed sets after removal of staminodes, the data were not significant. Future studies investigating the function of staminodes could examine germination rates as another proxy for reproductive fitness, but also could involve sequencing microbial communities on staminodes as well as carpels, and carpels after staminode removal to identify if staminodes have a defensive role in microbial community structure.

In conclusion, we now have a better understanding of the morphological and genetic differences between staminodes and stamen filaments, but there are several interesting open questions left with these mysterious organs.

Appendix A

Supplemental Data Chapter 3

Table A.1: Library Normalization Factors

Sample Replicate	Library Normalization Factor
Staminodes post 1	0.807996
Staminodes post 2	0.8483057
Staminodes post 3	0.790424
Staminodes pre 1	1.1627088
Staminodes pre 2	1.1579592
Staminodes pre 3	1.1238238
Staminodes pre 4	1.1818069
Stamen filament post 1	0.835625
Stamen filament post 2	0.8758507
Stamen filament post 3	0.8440034
Stamen filament post 4	0.8494513
Stamen filament pre 1	1.2716188
Stamen filament pre 2	1.2444071
Stamen filament pre 3	1.0601729
Stamen filament 4	1.1725976

Table A.2: Transcription factors identified from DGE analysis

TF Family	Aquilegia gene ID	Arabidopsis Gene ID	Gene
Staminode Upregulated			
AP2-EREB2	Aqcoe3G269000	At5g51190	
ARF	Aqcoe5G226900	At1g19850	ARF5, IAA24, MP
bHLH	Aqcoe2G335600	At4g09820	TT8
C2C2-Dof	Aqcoe7G134900	At1g51700	ADOF1
C2C2-YABBY	Aqcoe7G078300	At2g45190	AFO, FIL, YAB1
C2H2	Aqcoe2G212600	At3g10470	
C2H2	Aqcoe2G353000	At3g13810	AtIDD11
C2H2	Aqcoe5G082300	At1g55110	AtIDD7
C3H	Aqcoe5G004800	At3g25030	
CAMTA	Aqcoe7G181400	At1g67910	
CCAAT-HAP2	Aqcoe1G076700	At5g06510	NF-YA10
Homeobox (HDG2)	Aqcoe4G229400	At1g05230	HDG2
Homeobox (HDZIPI)	Aqcoe1G251700	At5g53980	ATHB52
MADS	Aqcoe5G160100	At2g22540	AGL22, SVP
MYB	Aqcoe2G366400	At5g26660	AtMB84, AtMYB86
MYB	Aqcoe7G024100	At1g63910	AtMYB103
NAC	Aqcoe5G024200	At4g28530	ANAC074
NAC	Aqcoe5G470900	At2g46770	ANAC043, NST1
Staminode Downregulated			
C2C2-Dof	Aqcoe4G093300	At2g28510	
MADS	Aqcoe4G258900	At4g09960	AGL11, STK
MADS	Aqcoe5G180800	At3g54340	AP3-3
MADS	Aqcoe6G257300	At4g11880	AP3-2
TCP	Aqcoe3G081500	At1g69690	

Table A.2: Transcription factors identified from DGE analysis (Continued)

Staminode Pre Upregulated			
AP2-EREBP	Aqcoe4G238900	At5g51990	CBF4
ARF	Aqcoe2G315700	At5g62000	ARF1-BP
ARID	Aqcoe2G397100	At1g04880	
ARR-B	Aqcoe3G009700	At2g01760	ARR14
bZIP	Aqcoe1G242500	At3g58120	AtBZIP61
bZIP	Aqcoe3G399500	At5g06950	AHBP-1B
C2H2	Aqcoe1G124500	At5g22890	
C2H2-YABBY	Aqcoe3G065300	At1g69180	CRC
C3H	Aqcoe5G004800	At3g25030	
C3H	Aqcoe5G203400	At4g33565	
CCAAT-HAP2	Aqcoe1G076700	At5g06510	NF-YA10
Homeobox	Aqcoe2G168200	At2g28610	PRS, WOX3
MADS	Aqcoe6G153100	At2g22540	AGL22, SVP
MYB	Aqcoe3G121900	At5g14750	AtMYB66
MYB	Aqcoe4G279900	At5g53200	TRY
MYB	Aqcoe5G159600	At3g50060	MYB77
MYB	Aqcoe7G038200	At3g61250	AtMYB17
Staminode Pre Downregulated			
AP2-EREBP	Aqcoe1G245100	At3g23220	
AP2-EREBP	Aqcoe5G252800	At5g64750	ABR1
bHLH	Aqcoe3G085300	At1g74500	
bHLH	Aqcoe5G421800	At5g51780	
C3H	Aqcoe2G053700	At2g35420	
C3H	Aqcoe2G170400	At5g02750	
C3H	Aqcoe7G414000	At4g28890	

Table A.2: Transcription factors identified from DGE analysis (Continued)

MADS	Aqcoe4G258900	At4g09960	AGL11, STK
MADS	Aqcoe7G053700	At2g45650	AGL6
NLP	Aqcoe3G256800	At1g64530	
WRKY	Aqcoe1G167500	At2g34830	WRKY35
WRKY	Aqcoe1G294300	At1g62300	WRKY6
Staminode Post Upregulated			
ABI3VP1	Aqcoe1G406800	At4g32010	HSI2-L1
AP2-EREBP	Aqcoe6G063200	At1g78080	RAP2.4
AP2-EREBP	Aqcoe1G245100	At3g23220	
AP2-EREBP	Aqcoe2G371900	At3g23240	ATERF1
AP2-EREBP	Aqcoe7G364300	At5g25190	
AP2-EREBP	Aqcoe5G327100	At5g61890	
AP2-EREBP	Aqcoe5G252800	At5g64750	ABR1
bHLH	Aqcoe7G007900	At1g01260	MYC7E
bHLH	Aqcoe6G064400	At2g22750	
bHLH	Aqcoe2G335600	At4g09820	TT8
bHLH	Aqcoe5G469700	At5g46760	ATR2
bHLH	Aqcoe2G415100	At5g48560	
bHLH	Aqcoe5G421600	At5g51780	
bZIP	Aqcoe3G382800	At1g13600	AtbZIP58
bZIP	Aqcoe6G325700	At1g42990	AtbZIP60
C2C2-Dof	Aqcoe1G227500	At1g29160	COG1
C2H2	Aqcoe5G414800	At1g02040	
C2H2	Aqcoe1G429500	At2g29660	
C2H2	Aqcoe2G179500	At3g10470	
C2H2	Aqcoe1G471100	At3g46090	ZAT7

Table A.2: Transcription factors identified from DGE analysis (Continued)

C2H2	Aqcoe1G449300	At3g52800	
C2H2	Aqcoe2G179300	At3g53600	
C2H2	Aqcoe5G103800	At5g67450	AZF1
C3H	Aqcoe3G023400	At1g68200	
C3H	Aqcoe2G053700	At2g35420	
C3H	Aqcoe1G075700	At2g35910	
C3H	Aqcoe1G269800	At4g27470	
C3H	Aqcoe6G075200	At4g33565	
C3H	Aqcoe5G458800	At5g07040	
C3H	Aqcoe1G121700	At5g17600	
C3H	Aqcoe1G385200	At5g55970	
CAMTA	Aqcoe5G398700	At1g67910	
CCAAT-HAP3	Aqcoe1G379200	At1g21970	EMB212
CCAAT-HAP5	Aqcoe2G002400	At3g12480	NF-YC11
G2-like	Aqcoe3G045700	At1g13300	
Homeobox	Aqcoe1G341500	At1g62360	STM, BUM
Homeobox	Aqcoe7G052800	At1g62990	IXR11
Homeobox	Aqcoe7G002200	At3g61890	AtHB12
Homeobox	Aqcoe6G318700	At4g08150	BP
Homeobox	Aqcoe5G197900	At4g36740	AtHB40
Homeobox	Aqcoe5G202000	At4g36870	BLH2
HSF	Aqcoe7G042500	At4g11660	AT-HSFB2B
HSF	Aqcoe6G140800	At5g62020	AT-HSFB2A
MADS	Aqcoe3G373600	At1g69120	AGL7, AP1
MADS	Aqcoe1G209900	At4g18960	AG
MYB	Aqcoe7G024100	At1g63910	AtMYB103

Table A.2: Transcription factors identified from DGE analysis (Continued)

MYB	Aqcoe3G412400	At1g73410	AtMYB54
MYB	Aqcoe2G014200	At3g08500	AtMYB83
MYB	Aqcoe1G055500	At3g13540	AtMYB85
MYB	Aqcoe5G012100	At3g24310	AtMYB71
MYB	Aqcoe7G074600	At4g22680	AtMYB85
MYB	Aqcoe5G236600	At5g17800	AtMYB56
MYB	Aqcoe2G239500	At5g26660	AtMYB84
MYB	Aqcoe7G114500	At5g35550	AtMYB123
NAC	Aqcoe1G321200	At1g52890	ANAC019
NAC	Aqcoe3G345900	At1g69490	ANAC029
NAC	Aqcoe5G470900	At2g46770	ANAC043
NAC	Aqcoe1G251800	At4g28500	ANAC073, SND2
NAC	Aqcoe5G344500	At5g13180	ANAC083
NAC	Aqcoe2G001300	At5g61430	ANAC100
NAC	Aqcoe2G001600	At5g63790	ANAC102
RAV	Aqcoe3G011700	At1g50680	
SBP	Aqcoe6G327700	At1g76580	SPL16
SBP	Aqcoe2G319400	At5g18830	SPL7
TCP	Aqcoe3G048600	At1g68800	BRC2
Trihelix	Aqcoe5G171700	At1g49490	
Trihelix	Aqcoe6G099500	At2g15880	
Trihelix	Aqcoe1G418600	At2g38250	
Trihelix	Aqcoe3G276700	At5g14540	
WRKY	Aqcoe3G360300	At1g69310	AtWRKY57
WRKY	Aqcoe7G312800	At1g80840	AtWRKY40
WRKY	Aqcoe1G484100	At2g37260	AtWRKY44

Table A.2: Transcription factors identified from DGE analysis (Continued)

WRKY	Aqcoe1G392900	At2g38470	AtWRKY33
WRKY	Aqcoe1G503300	At2g40750	AtWRKY54
WRKY	Aqcoe5G143700	At4g24240	AtWRKY7
WRKY	Aqcoe3G102200	At5g01900	AtWRKY62
Staminode Post Downregulated			
ABI3VP1	Aqcoe3G217300	At3g26790	FUS3
AP2-EREBP	Aqcoe7G221500	At1g16060	
ARF	Aqcoe1G494900	At5g60450	ARF4
ARF	Aqcoe4G073400	At2g33860	ARF3
ARF	Aqcoe5G074300	At3g61830	ARF18
bHLH	Aqcoe3G395700	At1g68810	
bHLH	Aqcoe6G075100	At1g72210	
bZIP	Aqcoe4G068200	At3g58120	AtBZIP61
C2C2-Dof	Aqcoe1G464100	At3g55370	OBP3
C2C2-dof	Aqcoe5G172800	At3g52440	
C2C2-Gata	Aqcoe1G060700	At4g32890	GATA-9
C2C2-Yabby	Aqcoe3G065300	At1g69180	CRC
C2C2-YABBY	Aqcoe7G386400	At2g26580	YAB5
C2H2	Aqcoe1G046500	At3g12270	ATPRMT3
C2H2	Aqcoe1G376300	At1g75710	
C2H2	Aqcoe3G427600	At2g02080	AtIDD4
C3H	Aqcoe2G013900	At3g63530	BB
C3H	Aqcoe3G049000	At1g18470	
CPP	Aqcoe7G393500	At3g22760	SOL1
G2-like	Aqcoe6G278700	At4g37650	SGR7
GRAS	Aqcoe1G454000	At2g36400	AtGRF3

Table A.2: Transcription factors identified from DGE analysis (Continued)

GRF	Aqcoe5G457000	At2g22840	AtGRF1
GRF	Aqcoe3G272100	At4g16780	AtHB-2
Homeobox	Aqcoe5G096100	At2g22800	HAT9
Homeobox	Aqcoe2G045800	At1g46264	AT-HSFB4
HSF	Aqcoe2G279000	At5g49330	AtMYB111
MYB	Aqcoe6G244500	At2g18060	ANAC037
NAC	Aqcoe7G309900	At1g76580	SPL16
SBP	Aqcoe3G081500	At1g69690	
TCP	Aqcoe3G288500	At1g70460	
Trihelix	Aqcoe6G278700	At4g37650	SGR7

Table A.3: Enriched sequences identified in Both Staminode Samples

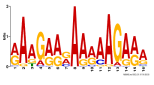
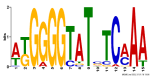
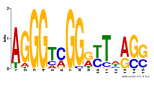
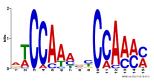
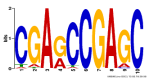
Sequence	*Width*	*#sites*	*MEME P value*	*MEME E value*	*TOMTOM matches with q and P values <.05*
	16	50	3.50E-252	2.00E-45	BPC1, BPC6, RAMOSA1, BPC5, AT3G46070
	16	48	2.10E-202	4.20E-03	AT1G69570, COG1, OBP3, AT5G02460, OBP1, Adof1, AT2G28810, AT5G66940, AT3G45610
	15	19	5.00E-94	4.90E-08	N/A
	15	20	1.30E-93	1.70E-03	N/A
	15	24	1.80E-109	4.70E-03	AGL42
	15	50	1.80E-212	1.10E-05	N/A
	10	28	7.60E-146	2.30E-23	N/A
	10	50	1.30E-212	9.70E-06	N/A
	8	40	2.70E-182	4.50E-13	RAMOSA1, BPC6, BPC1, BPC5
	8	62	2.20E-247	1.30E-03	N/A

Table A.4: Enriched sequences identified in Staminode-Pre Samples

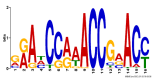
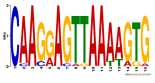
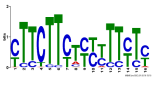
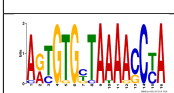
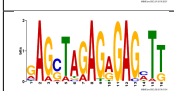
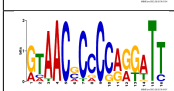
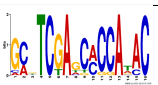
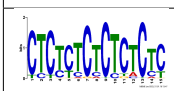
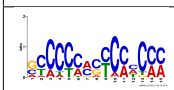
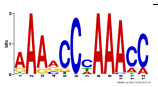
Sequence	*Width*	*Number of sites*	*P value*	*E value*	*TOMTOM matches with q and P values <.05*
	16	16	1.00E-105	4.50E-29	N/A
	16	38	4.10E-210	3.60E-41	N/A
	16	16	2.50E-99	1.10E-22	N/A
	16	16	1.30E-93	5.60E-17	N/A
	16	48	6.00E-222	3.80E-13	OBP3, AT1G69570, COG1, AT5G02460, Adof1, AT3G45610, AT5G66940, OBP1
	16	15	6.00E-81	1.00E-08	N/A
	16	26	5.40E-126	2.10E-06	N/A
	16	17	8.00E-86	8.70E-05	RAMOSA1, BPC6
	16	16	2.60E-82	1.20E-05	N/A
	16	13	2.60E-66	5.10E-03	N/A
	15	50	4.70E-302	2.20E-85	BPC1, RAMOSA1, BPC6, BPC5, AT3G46070
	15	48	3.20E-217	2.10E-08	TCP1, RAX3
	12	58	1.80E-250	2.40E-06	TRP1, TRP2, AT4G12670

Table A.5: Enriched sequences identified in Staminode-Post Samples

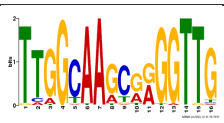
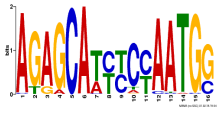
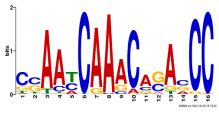
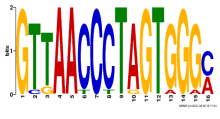
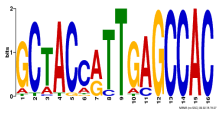
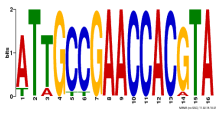
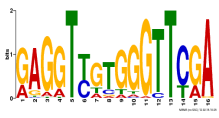
Sequence	*Width*	*Number of Sites*	*P value*	*E value*	*TOMTOM matches with q and P values <.05*
	16	36	1.8e-239	1.3e-053	N/A
	16	46	5.9e-302	3.5e-069	N/A
	16	92	1.5e-493	1.6e-055	N/A
	16	26	1.6e-179	8.1e-042	N/A
	16	15	2.5e-105	1.3e-022	N/A
	16	20	9.0e-136	9.6e-028	N/A
	16	25	1.8e-154	1.2e-021	N/A
	16	14	1.7e-098	6.3e-021	N/A
	16	37	9.7e-220	3.9e-029	N/A

Table A.5: Enriched sequences identified in Staminode-Post samples (Continued)

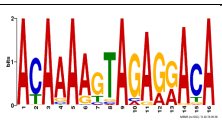
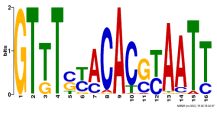
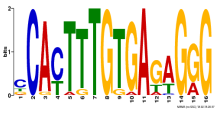
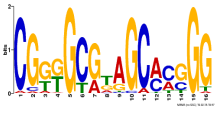
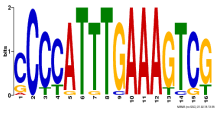
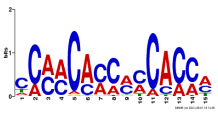
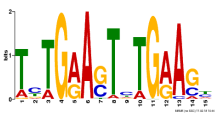
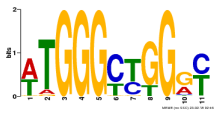
	16	38	4.5e-214	9.6e-019	N/A
	16	40	7.9e-223	4.5e-018	NAC055, NAC080
	16	18	4.0e-115	3.7e-017	N/A
	16	20	6.0e-124	6.4e-016	N/A
	16	26	1.2e-155	6.3e-018	N/A
	16	19	2.5e-115	2.5e-012	N/A
	16	16	1.0e-101	7.0e-014	N/A
	16	23	1.3e-140	1.1e-017	AT1G19210
	15	332	4.2e-1554	1.2e-173	DREB26, AT1G77640, AT1G22810, CEJ1, AT1G44830, AT5G18450, AT4G31060, RAX3, RAP21, AT4G28140, AT1G75490, DREB2
	15	109	6.7e-528	4.3e-017	N/A

Table A.5: Enriched sequences identified in Staminode-Post samples (Continued)

	15	98	6.7e-484	5.3e-020	N/A
	11	93	2.4e-453	8.6e-011	N/A
	11	199	1.2e-921	1.0e-043	N/A

Appendix B

Supplemental Data Chapter 4

B.0.1 R Code

Listing B.1: R Code example using our *Aquilegia canadensis* data

```
#load data tables into R
canadensis<-read.csv("canadensis.csv", sep=",")

#standardize the column names, set factors as characters if necessary, add 1 total
     seed count to remove 0 values
colnames(canadensis)<-c('plant', 'flower', 'treatment', 'total_number_seeds', '
    carpels')
canadensis$plant<-as.character(canadensis$plant)
canadensis$flower<-as.character(canadensis$flower)
canadensis$total_number_seeds<-canadensis$total_number_seeds+1

#boxplot of data
ggplot(canadensis,aes(x=treatment, y=total_number_seeds, fill = factor(treatment))
    )
  + geom_boxplot() + theme_bw(base_size=16) + ggtitle('Aquilegia canadensis')
  + theme(axis.text.x=element_text(angle=90, hjust=1))
  + scale_x_discrete(labels=c('Unmanipulated', 'Control 1', 'Experimental 1', '
      Experimental 2', 'Control 2', 'Experimental 3'))
  + scale_fill_manual(values = c('snow1', 'snow1', 'slateblue4', 'slateblue4', '
      snow1', 'slateblue4'), guide=FALSE)
  + theme(text=element_text(family="Times New Roman", face="bold", size=16)) +
      labs(x='Treatment', y='Total Number of Seeds')

#looking at variation based on number of carpels
CanadensisCarpels<-ggplot(canadensis, aes(x=carpels, y=total_number_seeds, colour
```

Listing B.1 (Cont.): R Code example using our *Aquilegia canadensis* data

```
    = treatment))
20  + geom_point() + stat_smooth(method=lm, se=FALSE) + theme_bw(base_size=16) +
      xlab('Number of Carpels')
21  + ylab('Total Number of Seeds') + theme(text=element_text(family='Times New
      Roman', face='bold', size=16))
22
23 CanadensisUnmanipulatedCarpels<-ggplot(Carpels, aes(x=Number_Carpels_Canadensis, y
    =Total_Seeds_Canadensis))
24  + geom_point() + stat_smooth(method=lm, se=FALSE) + theme_bw(base_size=16) +
      xlab('Number of Carpels') + ylab('Total Number of Seeds')
25
26 #fitting linear models using only fixed effects to the data
27 lm1Canadensis<-lm(total_number_seeds ~ treatment, data=canadensis)
28
29 #fitting generalized linear mixed models using both fixed and random effects
30 CanadensisGLMM1<-glmer(total_number_seeds ~ treatment + offset(carpels) + (1|plant
    ) + (1|flower), data=canadensis, family = poisson(link="log"))
31 CanadensisGLMM2<-glmer(total_number_seeds ~ treatment + offset(carpels) + (1|
    flower), data=canadensis, family = poisson(link="log"))
32 CanadensisGLMM3<-glmer(total_number_seeds ~ treatment + offset(carpels) + (1|plant
    ), data=canadensis, family = poisson(link="log"))
33
34 #Model information and selection
35 Summary(CanadensisGLMM1)
36 Anova(CanadensisGLMM1)
37 anova(CanadensisGLMM1, CanadensisGLMM2, CanadensisGLMM3)
38
39 #Lsmeans
40 lsmeansCanadensis<-lsmeans(CanadensisGLMM1, 'treatment', type = 'response')
41 #plot LSmeans to compare to raw data
42 CLDcanadensis<-cld(lsmeansCanadensis, alpha=0.05, adjust='tukey')
43 ggplot(CLDcanadensis, aes(x=treatment,y=rate,label=.group)) + geom_point(shape=15,
     size=4)
44  +geom_errorbar(aes(ymin=asymp.LCL, ymax=asymp.UCL), width=.2, size=.7)+theme_bw(
      base_size=16)
45  + xlab('Treatment') + ylab('Least square mean') + theme(axis.text.x=element_text
      (angle=90, hjust=1))
46  + theme(text=element_text(family='Times New Roman', face='bold', size=16))
47  + scale_x_discrete(labels=c('Unmanipulated', 'Control 1', 'Experimental 1', '
      Experimental 2', 'Control 2', 'Experimental 3'))
48
49 #raw data graphs
50 ggplot(canadensis, aes(x=treatment,y=total_number_seeds,label=treatment))
51  + geom_point(shape=15, size=4)+theme_bw(base_size=16) + xlab('Treatment') + ylab
      ('Total Number of Seeds')
52  + theme(axis.text.x=element_text(angle=90, hjust=1)) + theme(text=element_text(
      family='Times New Roman', face='bold', size=16))
53  + scale_x_discrete(labels=c('Unmanipulated', 'Control 1', 'Experimental 1', '
      Experimental 2', 'Control 2', 'Experimental 3'))
54
55 #multcomp                                          compCanadensis<-glht(
    CanadensisGLMM1,linfct=mcp(treatment=c('Control1 - Experimental1=0', 'Control1 -
    Experimental2=0', 'Control2 - Experimental3=0')), rank='Tukey')
```

B.0.2 Supplemental Tables

Table B.1: Two Sample T test results for *Aquilegia canadensis*. Yellow boxes indicate relevant comparisons with p values < 0.05.

	Unmanipulated (mean = 101.03)	**Control 1 (mean = 80)**	**Experimental 1 (mean = 89.2)**	**Experimental 2 (mean = 83.5)**	**Control 2 (mean = 90.73)**
Control 1 (mean = 80)	t = 2.715 df = 103.06 p = 0.0078				
Experimental 1 (mean = 89.2)	t = 1.7 df = 174.5 p = 0.09	t = -1.2 df = 100 p = 0.23			
Experimental 2 (mean = 83.5)	t = 2.36 df = 146 p = 0.019	t = -0.44 df = 103.3 p = 0.66	t = 0.77 df = 140.02 p = 0.44		
Control 2 (mean = 90.73)	t = 1.0809 df = 63.503 p = 0.2838	1.0606 df = 69.2, p - 0.29.26	t = 0.158 df = 62.7 p = 0.87	t = 0.73 df = 68.9 p = 0.46	
Experimental 3 (mean = 66.23)	t = 4.8709 d = 165.03 p = 2.586e-06	t = 1.73 df = 103.15 p = 0.0866	t = -3.23 df = 156.5 p = 0.0015	t = -2.28 df = 139.3 p = 0.02	t = 2.5294 df = 65.754 p = 0.01

Table B.2: Two Sample T test results for *Aquilegia eximia*. Yellow boxes indicate relevant comparisons with p values < 0.05

	Unmanipulated (mean = 69.9)	**Control 1 (mean = 39.3)**	**Experimental 1 (mean = 39)**	**Experimental 2 (mean = 40.4)**	**Control 2 (mean = 37.7)**
Control 1 (mean = 39.3)	t = 3.621 df = 20 p = 0.0017				
Experimental 1 (mean = 39)	t = 3.235 df = 20.5 p = 0.004	t = 0.02 df = 23.9 p = 0.98			
Experimental 2 (mean = 40.4)	t = 2.28 df = 10.2 p = 0.045	t = -0.08 df =13 p = 0.93	t = -0.0976 df = 15.47 p = 0.92		
Control 2 (mean = 37.7)	t = 4.733 df = 77.65 p = 0.75e-06	t = 0.18 df = 23 p = 0.86	t = 0.13 df = 23.4 p = 0.89	t = 0.21 df = 11 p = 0.84	
Experimental 3 (mean = 57.85)	t = 3.88 df = 120 p = 1.7e-04	t = -0.67 df = 21.4 p = 0.5095	t = -0.62 df = 21.7 p = 0.54	t = -0.35 df = 10.5 p = 0.7	t = -1.05 df = 80.6 p = 0.3

Table B.3: Two Sample T test results for *Aquilegia formosa*. Yellow boxes indicate relevant comparisons with p values < 0.05

	Unmanipulated (mean = 86.9)	**Control 1 (mean = 62.7)**	**Experimental 1 (mean = 46.6)**	**Experimental 2 (mean = 54.28)**	**Control 2 (mean = 65.74)**
Control 1 (mean = 62.7)	t = 4.22 df = 80 p = 6.3e-05				
Experimental 1 (mean = 46.6)	t = 7.47 df = 115.7 p = 1.7e-11	t = 2.47 df = 96.6 p = 0.015			
Experimental 2 (mean = 54.28)	t = 5.5 df = 74.77 p = 5.86e-07	t = 1.2 df = 84.6 p = 0.2314	t = -1.14 df = 92.2 p = 0.2573		
Control 2 (mean = 65.74)	t = 4.07 df = 87.5 p = 1e-04	t = -0.47 df = 82.8 p = 0.63	t = -3.16 df = 99 p = 0.002	t = -1.75 df = 80.17 p = 0.08	
Experimental 3 (mean = 57.85)	t = 6.15 df = 141 p = 7.4e-09	t = 0.82 df = 85.6 p = 0.42	t = -1.99 df = 117.4 p = 0.049	t = -0.58 df = 80 p = 0.57	t = 1.448 df = 91.6 p = 0.15

Table B.4: Linear Model Output

Aquilegia canadensis	***Coefficient estimates***	***Std. Error***	***t value***	***Pr(>)|t|***
intercept	102.032	4.707	21.675	<2e-16 ***
control 1	-21.054	8.212	-2.564	0.0107 *
epxerimental 1	-11.803	6.874	-1.717	0.0868 .
experimental 2	-17.495	7.297	-2.397	0.0170 *
control 2	-10.302	8.858	-1.163	0.245
experimental 3	-34.772	7.015	-4.957	1.06 e -06 ***
Residuals: Min = -101.32, 1Q = -35.648, Med = 2.1, 3Q=32.3, Max = 110.74				
Residual standard error: 33.45 on 191 degrees of freedom				
Multiple R-squared:,0.15, Adjusted R-squared: 0.1277				
F-statistic: 6.739 on 5 and 191 DF, p-value: 8.28e-06				
Aquilegia eximia	***Coefficient estimates***	***Std. Error***	***t value***	***Pr(>)|t|***
intercept	72.631	4.149	17.506	<2e-16 ***
control 1	-32.381	10.51	-3.081	0.00237 **
epxerimental 1	-29.631	9.582	-3.092	0.00228 **
experimental 2	-21.903	10.906	-2.008	0.04600 *
control 2	-34.748	7.08	-4.908	1.97e-06 ***
experimental 3	-24.581	5.989	-4.105	6.00e-05***
Residuals: Min -66.63, 1Q=-28.05, Median = 0.95, 3Q = 24.27, Max = 87.95				
Residual standard error: 33.45 on 191 degrees of freedom				
Multiple R-squared:,0.15, Adjusted R-squared: 0.1277				
F-statistic: 6.739 on 5 and 191 DF, p-value: 8.28e-06				
Aquilegia formosa	***Coefficient estimates***	***Std. Error***	***t value***	***Pr(>)|t|***
intercept	87.919	3.307	26.588	<2e-16 ***
control 1	-24.214	5.684	-4.26	2.65e-05 ***
epxerimental 1	-40.296	5.133	-7.85	5.62e-14 ***

Table B.4: Linear Model Output (Continued)

experimental 2	-32.64	5.727	-5.699	2.64 e -08 ***
control 2	-21.18	5.773	-3.669	0.000283 ***
experimental 3	-29.07	5.018	-5.793	1.59e-08 ***
Residuals: Min = -72.919, 1Q = -23.792, Med = 3.11, 3Q = 24.259, Max = 63.377				
Residual standard error: 30.66 on 336 degrees of freedom				
Multiple R-squared:,0.1821, Adjusted R-squared: 0.1699				
F-statistic: 14.96 on 5 and 336 DF, p-value: 2.912e-13				
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1				

Table B.5: Anova results of model selections

Aquilegia canadensis	**AIC**	**loglik**	**deviance**	**Chisq**	**Pr(>Chisq)**
Full model	4585.1	-2284.5	4569.1	77.49	<2.2e-16 ***
Without "plants" as a random effect	4660.6	-2323.3	4646.6		
Without "flower" as a random effect	31014.4	-15500.2	31000.4	0	1
Aquilegia eximia	**AIC**	**loglik**	**deviance**	**Chisq**	**Pr(>Chisq)**
Full model	4325.7	-2154.8	4309.7	813.54	<2.2e-16 ***
Without "plants" as a random effect	5542.9	-2764.4	5528.9		
Without "flower" as a random effect	5217.2	-2561.6	5123.2	405.67	<2.2e-16 ***
Aquilegia formosa	**AIC**	**loglik**	**deviance**	**Chisq**	**Pr(>Chisq)**
Full model	3539.6	-1761.8	3523.6	3922.5	<2.2e-16 ***
Without "plants" as a random effect	3571.5	-1778.8	3557.5		
Without "flower" as a random effect	7460.1	-3723	7446.1	0	1
Signif. codes:,0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1					