ABSTRACT

SULLIVAN, BRIAN JAMES. Complex Transcriptional Dynamics Underlie the Development of a Vital Reproductive Meristem in *Arabidopsis thaliana*. (Under the direction of Dr. Robert Franks).

Flowering plants develop from meristems. Meristems are groups of pluripotent stem cells that give rise to the diverse multitude of organs throughout the life of the plant. Defining the molecular genetic regulation of cell divisions and cell fate decisions within meristems is a fundamental biological problem of interest to a wide variety of life scientists. Within the gynoecium, the female reproductive structure of flowering plants, meristematic cells give rise to ovules and other structures that are critical for reproductive success.

In the model angiosperm *Arabidopsis thaliana,* the gynoecium is composed of two fused carpels. Early in the development of the gynoecium two meristematic structures form at the margins of the fused carpels that give rise to the placental tissues, septum, transmitting tract, and the ovules. These structures are called the Carpel Margin Meristems (CMM). The gene regulatory events that control the development of the CMM are incompletely characterized. In the CMM, a pool of pluripotent stem cells is maintained for a short period of time until it is allowed to differentiate into mature tissues of the gynoecium. Elucidating the gene regulatory events that direct the development of the CMM and its derived structures is important to understand floral development, as well as to understand the mechanisms of meristematic development more generally.

To better understand the complex transcription network that regulates CMM development, we identified spatially- and temporally-distinct transcriptomic profiles from the developing

gynoecium. From these transcriptomes, we identified differentially expressed genes (DEGs) from the medial and lateral domains. We identified 328 genes that are differentially expressed in the medial domain and 827 genes that are differentially expressed in the lateral domain. Analysis of the expression patterns of these genes suggested that a significant gene regulatory transition occurred between 6 and 7 days after induction of flowering (DAI). To better understand CMM development, we employed an unsupervised method to create gene regulatory network representations of this developmental transition. Using this modeling approach we identified a network of 11 transcription factors that may regulate this transition. Our results suggest that the TGACG motif binding protein 9 (TGA9) and lob domain-containing protein (LBD21) function as critical regulators of this transcriptional shift.

To better understand the transcriptional interactions governing CMM development, we have optimized a CRISPR based system to generate higher order mutants in putative floral regulators. We endeavored to create higher order mutant combinations for six genes that display highly enriched expression in the medial domain of the flower compared to the rest of the floral tissue. This work was conducted in both *seuss* (*seu*) and *aintegumenta* (*ant*) mutants, as these mutant backgrounds are likely sensitized and more likely to reveal defects in CMM development. This use of a sensitized system to enhance the discovery of a mutant phenotype has been long established in developmental genetics. To increase the chances of discovering a mutant phenotype in the gynoecium, we used a CRISPR/Cas9 approach in an effort to make a line where all six of these genes are mutated. This process is expected to generate many other combinations of mutants in our six genes of interest. Further analysis of these mutant lines, as well as mutants in the genes, predicted through our modeling approach should provide further evidence into the

dynamics of this vital reproductive meristem, while providing a framework for existing knowledge of CMM development.

Complex Transcriptional Dynamics Underlie the Development of a Vital Reproductive Meristem in *Arabidopsis thaliana*

by
Brian James Sullivan

A dissertation submitted to the Graduate Faculty of
North Carolina State University
in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

Plant Biology

Raleigh, North Carolina
2020

APPROVED BY:

_____	_____
Dr. Robert Franks Committee Chair	Dr. Colleen Doherty
_____	_____
Dr. Marcela Rojas-Pierce	Dr. Steffen Heber

DEDICATION

I would like to dedicate this work to my parents who convinced me anything is possible. Thanks, Dad for being a role model, your work ethic has been inspirational. Thank you Mom for talking me through the stressful times and keeping me focused on the important things. Without you, I would never have been in a position to pursue this degree and certainly would never have been able to complete it.

BIOGRAPHY

Attended Lehigh University, graduating with a Bachelor of Science in Biology in 2013.

ACKNOWLEDGMENTS

I would like to acknowledge Miguel Flores Vergara, a friend, and mentor for showing me what science looks like and for his unwavering support.

I would like to thank all of the undergraduates who helped me with this work. Thank you Jack, Katie, Michael, and Xavious for your hard work, I'm wishing you all the best for the future.

Bob, I would be remiss if I did not thank you for your time and efforts in helping me sort out career goals, resolve problems with experiments, and especially for helping me grow as a student and scientist, Thank You!

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

One of the most fascinating and unique features of plant development is the ability of plants to continue patterning and forming new organs long after embryogenesis. This is enabled by plant meristems. Plant meristems, as defined by modern developmental biologists "are cell populations that retain some of the properties of embryonic cells, principally the capacity to divide, long after embryogenesis is over" and "produce daughter cells that differentiate and give rise to different organ structures" (Fosket 1994) (Dello Ioio et al. 2007). These cell populations are located in stem cell niches which have been traditionally limited to the shoot apical meristem (SAM), root apical meristem (RAM), and vascular meristems, however, there are other groups of cells besides these classical meristems that exhibit many of the same features (Greb and Lohmann 2016). In order to provide a historical context for the development of the meristem concept, here I present a review of seminal studies of plant meristems, as well as the study of cellular differentiation and specification from these meristematic regions.

The Historical Context of the Concept of the Plant Meristem

In 1759 the German professor of medicine and embryologist Caspar Friedrich Wolff first described the shoot apical meristem as he investigated the theory of preformationism. The theory of preformationism posits that the embryo does "contain all its parts in little from the beginning, unfolding like a Japanese paper flower in water" (Speybroeck et al. 2002). This theory arose based on the writings of Descartes who argued that there should be no intangible force driving the development of new organisms and thus it was most likely that they were encapsulated in miniature waiting to unfold into mature adults (Lawrence 2008). This theory was opposed by

proponents of the theory of epigenesis, which contradicted preformationism stating instead that "different organs form by a cascade of gradual changes in an undifferentiated mass, leading to a well-organized whole, that is, the embryo" (Speybroeck et al. 2002). One prominent early advocate for the theory of epigenesis was Aristotle who studied anatomy by looking at mature chickens but found that in order to properly understand why organisms grow to a specific form it was necessary to look at development as well. To this end, he studied the development of fertilized chicken eggs over 28 days. His observations contradicted both the theory of pangenesis, where each part of the adult contributes a specific material to the seed and also the theory of preformation, leading to the formalization of the theory of epigenesis.

In the 17th and 18th centuries, preformationism was the predominant accepted theory describing the development of plants and animals, even though it conflicted with Aristotle's concept of epigenesis. Following the path taken by Aristotle, Caspar Friedrich Wolff chose to investigate preformationism by studying development in both plant and animal systems. In plants, he chose to look at the developing shoot tip and while searching the shoot apex for miniature, pre-formed, encapsulated organisms he discovered minute undifferentiated globules that grew and differentiated into stems, leaves, and other plant tissues, supporting the theory of epigenesis. He also studied flower development and suggested that the floral organs were modified leaves. After these observations by Wolff, Karl Wilhelm von Nageli in 1858 coined the term "meristem" referring to a group of cells that are always capable of division. In fact, he chose the word meristem from the greek *merizein*, which means divide. He also believed this to be the main site of growth in plants and continued to study the pattern in which cell divisions occur in this tissue.

Thus began the investigation of the way that these meristems are maintained and contribute to plant growth.

The shoot apical meristem is the meristem that governs shoot growth. All above-ground organs (except for the cotyledon leaves) are formed from the cells of the shoot apical meristem. Three prominent theories have arisen to describe the development of the shoot apical meristem. In 1844, while working with both bryophytes and chlorophytes, Nageli proposed the "apical cell theory" to describe the action of the meristem. In the bryophytes and chlorophytes, an apical cell sits at the shoot apex, which will divide continuously (Figure 1). Nageli believed that the order and orientation of the cell divisions of the apical cell was the key to understanding the development of plants, a trend that would persist as researchers continued to study SAM development. This theory remains applicable to bryophytes and ferns which contain a large, highly-vacuolated apical cell which is often tetrahedral in shape. Through surgical experiments, it was shown that this cell produces an inhibitor that blocks the formation and development of lateral buds. When the apical cell was punctured and destroyed, existing leaf primordia would continue to develop in a normal pattern, but at most only one additional bud would form, indicating that the apical cell functions as a shoot apical meristem required for additional bud formation in these systems (Wardlaw 1951) (Wardlaw 1965). In angiosperms and gymnosperms, there is no large tetrahedral apical cell and the theory is not as applicable. A separate theory quickly developed which better described the cell layers found at the shoot tip in the seed plants.

For seed plants, the "histogen theory" was put forward by Hanstein in 1868. This theory assigned functions to each of the three layers of cells that he observed that surrounded an unorganized core in the shoot apex. Hanstein proposed that the location of the cell was responsible for the identity of the cells. The outermost cell layer (L1) was named the dermatogen which he proposed gave rise to the epidermis. The next layer, located just interior to the dermatogen, is the periblem (L2) which gave rise to the cortex cells. The most interior cell layer was the pierome (L3) from which the stele and all of the vasculature was derived. This aspect of the theory is not entirely consistent with modern data. While some features of the outer layer of the SAM show the properties Hanstein attributed to the dermatogen, the inner layers do not possess the characteristics he attributed to the periblem and the pierome (Foster 1939). Specifically, they are often indistinguishable from the surrounding tissue and do not confer cells in those layers of the SAM, cortex, or stele cell fates.

One aspect of this theory has been supported by current research; his holistic view on SAM growth, that overall SAM growth was more important than that of the fates of the individual cells, has garnered particular support in recent years. Instead of tracking the fate of individual cells, Hanstein focused on the overall trends in growth in different regions of the SAM. In addition, Hanstein noted that a single mode of action for the development of the SAM of all plants was overly reductionist. He focused on the difference in the development of vascular cryptogams and spermatophytes. The former does have an apical cell that divides to produce the rest of the shoot cells and the latter which have a more complicated layered arrangement.

One contentious area was the shoot apex of the gymnosperms which neither Nageli's apical cell model or Hanstein's histogen theory accurately described. It was not until the work of Koch that the shoot apex of conifers was accurately described (Koch 1891). The structure of this meristem consists of an outer mantle of densely cytoplasmic cells and an interior central tissue consisting of larger, more vacuolated cells. These different layers do not conform to the histogen model because the only tissue produced by the central tissue was the pith, all of the other tissues were produced by the outer mantle. Koch also disputed claims that the mantle and central tissue were the product of a single apical cell although he conceded that due to the lack of a clearly defined epidermis and the occasional enlargement of "one of the terminal surface cells" it is understandable why researchers had assumed that gymnosperms fit the apical cell model.

A Modern Conceptualization of the Angiosperm Shoot Apical Meristem

The model that currently best describes shoot apical meristem development is called the Tunica-corpus theory. First proposed in 1928 by Schmidt, this theory is similar to the mantle-core theory in that it describes two distinct zones. The tunica consists of the surface layer and may include another layer called L2 found interior to L1, this layer is formed through periclinal divisions (i.e., divisions parallel to the organ's surface) (Cunninghame et al. 1986). In the tunica anticlinal divisions, (i.e., those perpendicular to the organ surface) lead to surface growth and the maintenance of the distinct layers. Traditionally the third layer from the outside (L3) is the corpus. In the L3, tissue growth leads to an increase in volume which reflects the unorganized nature of the divisions in this area of the apex. The balance between the division of the tunica and corpus is what allows the SAM to change shape and to give rise to leaf primordia.

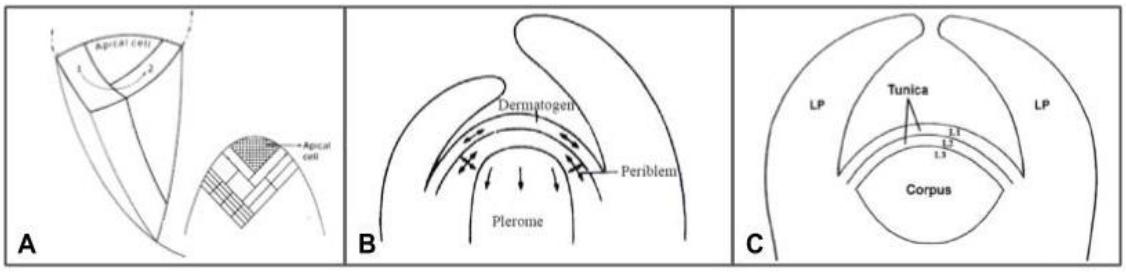


Figure 1.1 Three Models of SAM development. A. Nageli proposed a model of shoot development that centered on an apical cell that gave rise to successive cells, this model best applies to Bryophytes. B. Hanstein's histogen theory focused on three cell layers which gave rise to the major tissues of the plant; the dermatogen, periblem, and plerome. C. The Tunica Corpus Model divides the tissues of the SAM into the outer Tunica which divides primarily anticlinally while the inner Corpus divides periclinally.

The SAM has been divided into distinct regions in several different ways beyond the tunica and the corpus. One approach was to divide the SAM into distinct regions based on specific characteristics such as the rate of cell division or the fates of cells produced there. The central zone contains stem cells and the slowly dividing organizing center and provides initials that give rise to cells in the peripheral and rib zones. The peripheral zone flanks the central zone and is formed from impermanent initials located in the central zone (Medford 1992) (Ruth et al. 1985). The peripheral zone has higher mitotic activity and can be divided into two different zones. The outer peripheral zone (OPZ) contains cells that differentiate while the inner peripheral zone (IPZ) contains cells that can revert to stem cell fate (Reddy and Meyerowitz 2005, Reddy 2008). The major function of the peripheral zone is the production of lateral organs through the production of leaf primordia, the simplest set of cells capable of triggering the growth of the leaf (Bhatia 2015). Basal to the central zone, but produced in the same manner, is the rib meristem which forms cells for the middle of the stem (Medford 1992). The central zone contains cells that divide at the slowest rate and function as stem cells, although unlike in mammalian systems they

are transient and do not persist (Ruth et al 1985). In contrast, the peripheral zone cells divide more rapidly, they divide in several different directions depending on their location. Those that are located in the outermost layer, L1, divide only in the anticlinal plane which allows them to produce epidermal tissues in the mature regions of the plant, while those in the second layer, L2, divide primarily in the anticlinal plane, but also can divide in the periclinal plane when organ primordia are being initiated. The outer two layers of the peripheral zone divide at relatively similar rates as shown by Lyndon's observations in Pea plants, with L1 cells dividing more rapidly at the beginning of a new plastochron (the time interval between two successive leaf initiation events), while L2 cells divide rapidly later in the plastochron.

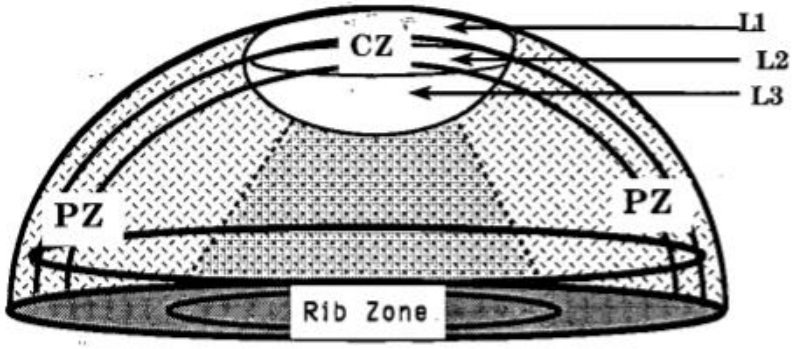


Figure 1.2. SAM.
Labeled is the central zone (CZ) consisting of three layers; L1 L2, and L3 and flanked by the peripheral zone (PZ). Adapted from J. Medford 1991 Plant Cell.

The second way that the SAM is divided mirrors the manner used by Hanstein, each of the cell layers was given a name L1, L2, and L3 and a prescribed function. This theory was described by Satina et al. in 1940 and arose after chimeric shoots were used to investigate the cell layers described by Hanstein. By treating seeds with colchicine, which inhibits cell division and causes increases in chromosome number, and then investigating ploidy in the cells of the three outermost layers of shoot tips, this group was able to show that the layers are independent. They

were also able to determine the ultimate cell fates of the different layers in a way that Hanstein was never able to prove. As stated previously, L1 produces the epidermal tissues in the mature parts of the plant. L2 is responsible for providing cells that will form the lateral organs, while L3 cells will become the internal portions of the shoot.

The maintenance of the meristem relies on specific patterns of cell division, which result in the cells that will make up the entirety of the shoot. In Priestley's observations of meristem development, he concluded that "*From the nature of the activities of the meristem cell a close connection may be anticipated between cell division and cell size*" (Priestley 1929). This suggests that as cell size increases the chances of cell division occurring also increase, this is a simple statement but due to the complex nature of the meristem it is not sufficient to describe SAM development. In this structure, there are distinct regions which contain cells of different sizes which divide at different rates. In his work, Priestley documents the cellular processes which lead up to division; an increase in the amount of protoplasm in the cell, an increase in cell size, the nucleus divides, and then in the space furthest from the two nuclei protoplasm ceases to be formed and instead forms a new cell wall. He proposes that this is due to the ratio of surface area to volume and as the size of the cell increases the amount of "food supplies" required to maintain cell growth is too great to be brought into the cell based on the amount of surface area. Despite this sized based rule, Priestley observes that there is a clearly defined pattern of cell divisions which should lead to the ability to predict the fate of each cell.

Several rules were developed by Sachs to characterize cell division in the SAM. The first says that without interference from outside debris, cells will divide into two equally sized daughter

cells. The process of karyokinesis is an internal process that senses the direction of cell growth and then forms the mitotic spindle parallel to the current direction of cell growth, this leads to the formation of the cell plate in a direction perpendicular to the direction of cell growth. This occurs because the cell plate forms in between the two new nuclei. The exact angle of the new plane of division is hard to predict in the inner corpus region of the meristem. In the tunica layers, anticlinal divisions allow for the maintenance of distinct layers, but in the corpus, the cell divisions are less organized and tend to follow Errera's law which states that "the cell plate, at the time of its formation, adopts the geometry that a soap film would take under the same conditions" (Wright and Smith 2007). Sach's second rule attempts to describe the orientation of these new cell walls stating that each succeeding plane of division tended to occur at right angles to the preceding wall. This law is inaccurate, however, one portion has been accepted; how the cells divide is the result, not the cause of the shape of the dividing structure. Priestley adds one interesting observation, due to the existing forces within the tissue the new cell tends to grow into the same shape as the parent cell.

Studying the patterns of cell division in the shoot apical meristem helps understand the factors that determine the shape of the meristem, phyllotaxy, and the characteristics of the meristem, but misses out on an equally important factor; the dynamics of development. The main role of shoot apical meristem is to maintain a population of pluripotent cells while continuing to produce cells that will go on to form the lateral organs of the shoot. This allows the activity of the SAM to be measured by tracking the formation of new cell walls. This is especially informative as the different regions of the meristem can be observed independently and more mature tissues,

located further and further away from the apex of the shoot can be observed as well to give an idea about the temporal dynamics of shoot development.

Central and Peripheral Zones of the Shoot Apical Meristem

Cells in the SAM can be divided into distinct regions that have distinct characteristics and develop in different patterns and at different rates. The central zone can be described by the Tunica-corpus theory, it contains stem cells, and the slowly dividing organizing center and gives rise to cells in the peripheral and rib zones (Jacobs and Morrow 1961). The peripheral zone flanks the central zone and is formed from impermanent initials produced by the stem cells in the central zone (Harrison et al. 2007). The major function of this zone is the production of lateral organs by producing leaf primordia. Basal to the central zone but produced in the same manner is the rib meristem which forms cells for the middle of the stem. The central zone contains cells that divide at the slowest rate and function as stem cells although unlike in mammalian systems they are transient and do not persist. In contrast, the peripheral zone cells divide more rapidly, they divide in several different directions depending on their location. Those cells that are located in the outermost layer L1, divide only in the anticlinal plane which allows them to produce epidermal tissues in the mature regions of the plant, while those in the second layer L2 divide primarily in the anticlinal plane but divide in the periclinal plane when organ primordia are being initiated. The outer two layers of the peripheral zone divide at relatively similar rates as shown by Lyndon's observations in Pea plants, with L1 cells dividing more rapidly at the beginning of a new plastochron while L2 cells divide rapidly later in the plastochron.

Carpel Margin Meristem as Compared to the Canonical Stem Cell Niches

In plants, there are three main stem cell niches, the shoot apical meristem, root apical meristem, and vascular meristems. In the shoot of the model angiosperm *Arabidopsis thaliana*, the shoot apical meristem can give rise to the three primary meristems, the protoderm, procambium, or the ground meristem and can transition into a floral meristem which gives rise to flowers at maturity. These are not the only locations of cells with meristematic characteristics in the shoot. One example of tissue with meristematic properties is the Carpel Margin Meristem (CMM)(Alvarez and Smyth 2002)(Azhakanandam et al. 2008)(Bowman et al 1999)(Liu et al 2000).

In Arabidopsis flowers, there are four different types of organs arranged in concentric circles or whorls (Figure 1.3) (Coen and Meyerowitz 1991). The outermost whorl consists of sepals that protect the developing reproductive organs. Interior to the sepals are the petals that often function to attract pollinators. The inner two whorls are the male and female reproductive structures, the stamens and carpels respectively (Kunst et al 1989). In Arabidopsis, the gynoecium, the female reproductive structure is the innermost whorl and is formed from two fused carpels (Okada et al 1989). The component carpels arise congenitally-fused during the development of the gynoecium. This fusion of the component carpels along the carpel margins results in a medial domain of the gynoecium that is somewhat confusingly composed of the marginal portions of the component carpels. An important reproductive meristem develops within the medial region of the developing gynoecium. This reproductive meristem is termed the carpel margin meristem (CMM), named because it arises from the fused margins of the two carpels. The CMM is responsible for producing many of the internal tissues in the gynoecium, specifically the placenta, ovules, septum, transmitting tract, style, and the stigma. Due to its role

in the reproduction of seed plants, understanding the processes that underlie the development of the CMM is of importance.

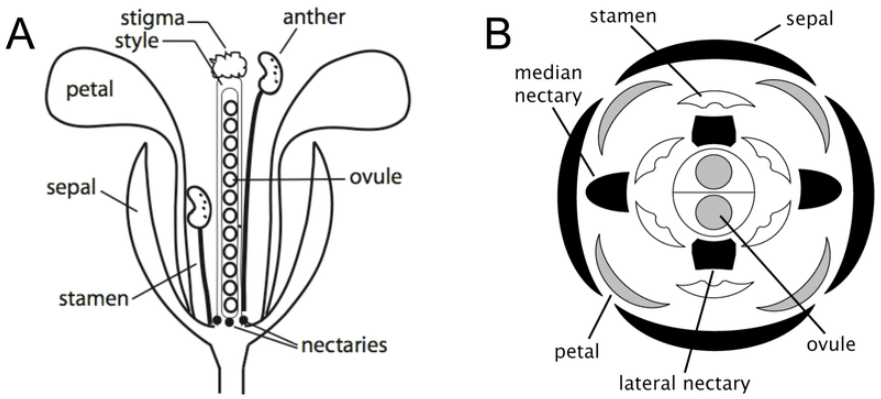


Figure 1.3 Anatomy of a *Brassicaceae* flower A. Viewed in longitudinal section flowers contain four different types of organs are arranged in whorls. The outermost whorl contains the sepals which enclose the developing flower. Interior to the sepals are the petals that may function to attract pollinators. The stamens are the third whorl and contain anthers which produce pollen, the stamens are the male reproductive structures of the flower. The fourth whorl is composed of two fused carpels, collectively called the gynoecium which produces ovules and develops into the fruit. **B.** A cross section of a Brassicaceae flower shows that there are four sepals, four petals, six stamens, and two fused carpels. The ovules are located in each of the two fused carpels and are produced by a tissue that arises at the margins of these two carpels, the CMM. (Adapted from Hampton et al. 2010)

The two fused carpels form a tube that extends as the gynoecium matures. The CMMs form early in the development of the carpels and are located on the interior of this tube and function similarly to the shoot apical meristem, producing lateral organs as the carpels develop. This is accomplished by balancing the production of new cells with differentiation of older cells. One of the areas where the CMM and the SAM differ is that the CMM is more transient and does not maintain stem cells in an undifferentiated state as long as the SAM. It has been noted that flowers are composed of organs formed from modified leaves, so it is not surprising that the flower shows similar developmental patterns and mechanisms as the shoot (Arber 1946). These

two meristems share many characteristics and even use many of the same transcription factors and hormone signaling pathways, however, the main difference in the development is that the CMM is a temporary structure while the SAM is maintained through the life of the plant (Bowman et al. 1999)(Alvarez and Smyth 2002). Several of the transcription factors shared between the CMM and SAM are *SHOOT MERISTEMLESS* (*STM*), *SPATULA* (*SPT*), and *CUP-SHAPED COTYLEDONS 1* (*CUC1*) and *CUC2*. These two different meristematic structures share additional characteristics beyond several shared genes. In order to best understand the CMM and its similarities and differences from the SAM is to take into account its development.

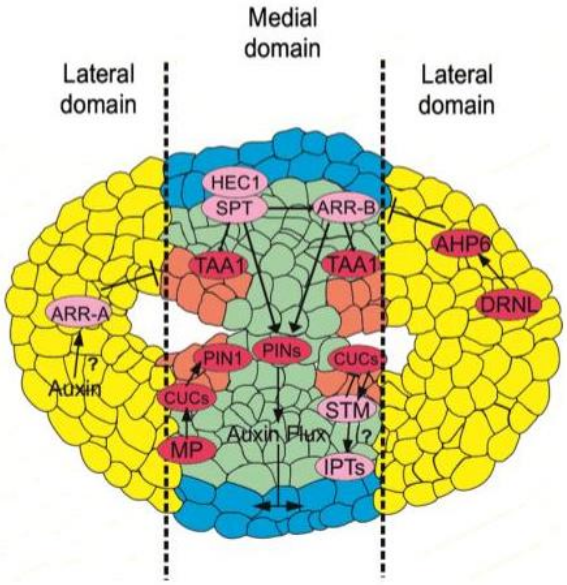


Figure 1.4 Distinct Domains Form Early in the Development of the Gynoecium. At stage eight of flower development the developing gynoecium contains three domains. Two lateral domains which will develop into the valves in mature gynoecia and a medial domain. The medial domain contains two medial ridges or CMMs which contain meristematic stem cells which will divide and differentiate to produce the medially derived structures. A complex set of interactions involving transcriptional regulation as well as hormone regulation maintains the meristematic identity of the CMM in early development and allows it to terminate to allow for proper development and maturation of the gynoecium. (Modified from Reyes-Olalde and de Folter 2019)

The development of the CMM can be broken down into two distinct phases, the first encompasses two stages of early gynoecium development when the CMM is a pluripotent zone and is described as a neogenic state, referencing that the structure was recently formed (Alvarez and Smyth 2002). This is followed by a determination and differentiation phase which lasts from stages 7 or 8 to stage 13 (Alvarez and Smyth 2002). These states are defined by the cellular activity happening in each phase. As is typical in development, the neogenic state is responsible for producing many new cells while later these cells develop and differentiate to their final cell fates. Two phytohormones play a large role in regulating the balance between the production of new cells and the differentiation of those cells. Auxin and cytokinin are critical in regulating both the SAM and the gynoecium (Reyes Olalde et al 2017)(Nemhauser et al 2000)(Sundberg and Ostergaard 2009). In the SAM, auxin is produced in the central zone of the meristem and is moved by polar auxin transporters to the incipient organ primordia where these temporary auxin maxima lead to the differentiation of cells and the formation of new organs (Stepanova et al 2008). In SAM development, cytokinins are important in ensuring that the meristem is maintained. This is done by promoting the division of stem cells in the central zone. The balance of these two hormones plays an important role in the development of the SAM. Auxin plays a very similar role in both the SAM and the CMM (Larrson et al 2014)(Muller et al 2017). In the latter, auxin is produced in the center of the medial domain and then moved outwards to the lateral domains by polar auxin transporters. Similarly to the SAM, these outer regions of the gynoecium become auxin maxima and the differentiation of these cells occurs to produce the valves, outer organs which will split off of the rest of the fruit post-fertilization. Auxin has also been implicated in apical/basal patterning of the gynoecium again through polar auxin transport. It has been shown that auxin is produced in the apical tissues of the gynoecium and moves

towards the basal portions and that this movement is critical for proper patterning and development of the gynoecium (Nemhauser et al. 2000, Benkova et al. 2003, Aloni et al. 2006).

Transcriptional Control of Carpel Margin Meristem Development

In addition to phytohormone control of the CMM, there are a series of transcription factors that are known to be involved in regulating gynoecium development. *STM* is a homeodomain-containing transcription factor expressed in the early stages of gynoecium development and its expression is required for proper development. Strong homozygous *stm* loss-of-function mutant individuals do not form a floral meristem or produce flowers due to embryonic SAM defects. However, through the use of partial loss-of-function alleles and inducible RNAi lines, it has been shown that when *STM* is not fully functional in the CMM, that less placenta and ovules form, the transmitting tract fails to form, and the carpels fail to fuse (Scofield et al., 2007, Endrizzi et al. 1996). In RNAi lines with higher levels of expression, flowers are produced but carpels fail to form at all.

A second homeodomain transcription factor is expressed in the CMM, *REPLUMLESS* (*RPL*). *RPL* is associated with meristem maintenance by repressing organ boundary genes and maintaining cell proliferation in the meristem (Khan et al 2015). It also plays a role in determining the orientation of cell division (Bencivenga et al 2016). In addition to homeodomain transcription factors, several others play important roles in gynoecium development. *CUC1* and *CUC2* encode NAC transcription factors that are necessary to form a septum and without the expression of these two genes, there are defects in ovule, septum, and replum development. These transcription factors bind the *STM* promoter and promote

transcription which increases expression of other meristem related genes (Hibara et al. 2003). *WUSCHEL* (*WUS*) and its paralog *WUSCHEL RELATED HOMEOBOX 13* (*WOX13*) play distinct roles in gynoecium, with *WUS* being expressed later in development, probably in style development. *WOX13* is expressed in the gynoecium throughout its development but the expression is restricted to the medial domain in later stages although its exact function is not understood due to a lack of a meristematic phenotype in *wox13* mutants, this has been hypothesized to be due to redundancy, specifically by *WOX6*.

Transcription factors from the basic helix-loop-helix (bHLH) (Gremski and Yanofsky. 2007, Alvarez and Smyth, Groszmann et al. 2010), auxin response factor (ARF) (Aida et al. 2002, Moller et al. 2017, Przemech et al. 1996), *AP2* (Elliott et al. 1996, Klucher et al. 1996), Basic Leucine Zipper (bZIP) (Chuang et al. 1999, Wynn et al. 2014), and MADS-box families play a role in the development of the gynoecium (Larsson et al. 2014, Villarino et al. 2016). The specific genes involved in this process can be broken down into two different groups based on their mutant phenotype. The first set of these have defects in septum formation, transmitting tract formation, and unfused carpels. There is a genetic interaction between *SPATULA (SPT)* and the *CUCs* to promote differentiation of carpel tissues. In addition, *SPATULA* is coexpressed with *HECATE1*, *HECATE2*, and *HECATE3* which in higher order mutants have similar carpel fusing phenotypes to *SPT*. *MONOPTEROS*(*MP*), an ARF transcription factor, *AINTEGUMENTA* (*ANT*) an *AP2* transcription factor, *PERIANTHIA* (*PAN*), *SHATTERPROOF1*, and *SHATTERPROOF2* are all expressed in the CMM and these transcription factors have roles in the development of the ovules and the surrounding tissues.

Additionally, there are non-transcription factor genes that play an important role in gynoecium development. The *CLAVATA3/ESR-RELATED (CLE)* family is a group of polypeptides that are important in controlling plant development. Other non-transcription factors include *LEUNIG (LUG), SEUSS (SEU),* and *SEUSS-LIKE (SLK).* These are transcriptional adapters that have been shown to repress *AG* expression. *LUG* has many roles in plant development including the initiation of the embryonic SAM and the early stages of floral meristem development where it represses *AG*. In *lug ant* double mutants there is a homeotic transformation of the outer whorl organs into carpelloid structures (Liu et al. 2000). *SEUSS* (*SEU*) has a similar function, it represses *AG* expression in the perianth. The *seu* mutants display a partial homeotic transformation of the outer two whorls of the flower. The *seu ant* double mutant differs from the *lug ant* double mutant in that the most dramatic phenotypes are associated with the inner organs. In the *seu ant* double mutant, fewer sepals and petals form than in wild type and those are narrower than wild type. When stamens form in the double mutant they contain only two locules instead of the four found in wild type. In the gynoecium, there is no change in organ number but the gynoecium fails to grow as long as it does in either single mutant or wild type (Azhakanandam et al. 2008). The functions of these genes highlight the complex nature of the transcriptional regulation of gynoecium which relies on transcription factors, transcriptional adapters, and epigenetics. Some attempts have been used to understand this complex network by comparing it to other well studied meristematic structures.

The CMM shows resemblance to the SAM even expressing many of the same genes, however, the CMM is not considered to be one of the canonical meristems in plant development. The primary difference between the CMM and the SAM is the determinate nature of the CMM. This

meristematic structure is produced in some *Brassicaceae* species long after embryogenesis and patterning and is only active for several days while in some species the SAM can remain active for thousands of years. Despite the difference in the lifespan of these two meristematic structures, they share a similar mechanism regulating their meristematic identity. This mechanism relies on the expression of WUS and as that expression is decreased the SAM transitions to a floral meristem. A similar trend is seen in the FM where WUS is repressed by AG to prevent overproliferation of carpels. In the CMM a similar pattern of development may be seen using WOX13 and WOX6 instead of WUS. Further investigation of the developmental dynamics of the CMM is needed to both understand the process that forms ovules in angiosperms and to understand how the development of the CMM compares to other more traditional meristems.

OUTLINE OF THE THESIS:

The next four chapters will: describe our approach to model the gene regulatory network (GRN) governing the development of the meristematic medial domain (the CMM) of the gynoecium (Chapter 2); outline two approaches to generate higher order mutants in *Arabidopsis thaliana* (Chapter 3); describe the progress that we have made on functionally characterizing medial domain enriched genes (Chapter 4); discuss our current model of CMM development, the key regulators, and the top approaches to identify critical genes in developmental processes (Chapter 5).

REFERENCES

Aichinger, Ernst, Noortje Kornet, Thomas Friedrich, and Thomas Laux. 2012. "Plant Stem Cell Niches." *Annual Review of Plant Biology* 63 (February): 615–36.

Aida, Mitsuhiro, Teva Vernoux, Masahiko Furutani, Jan Traas, and Masao Tasaka. 2002. "Roles of PIN-FORMED1 and MONOPTEROS in Pattern Formation of the Apical Region of the Arabidopsis Embryo." *Development* 129 (17): 3965–74.

Aloni, Roni, Erez Aloni, Markus Langhans, and Cornelia I. Ullrich. 2006. "Role of Auxin in Regulating Arabidopsis Flower Development." Planta 223 (2): 315–28.

Alvarez, John, and David R. Smyth. 2002. "CRABS CLAW and SPATULA Genes Regulate Growth and Pattern Formation during Gynoecium Development in Arabidopsis Thaliana." *International Journal of Plant Sciences* 163 (1): 17–41.

Azhakanandam S, Nole-Wilson S, Bao F, Franks RG (2008) SEUSS and AINTEGUMENTA mediate patterning and ovule initiation during gynoecium medial domain development. Plant Physiol 146:1165–1181.

Bencivenga, S., Serrano-Mislata, A., Bush, M., Fox, S., and Sablowski, R. (2016). Control of oriented tissue growth through repression of organ boundary genes promotes stem morphogenesis. *Dev. Cell* 39, 198–208. doi: 10.1016/j.devcel.2016.08.013

Benková, Eva, Marta Michniewicz, Michael Sauer, Thomas Teichmann, Daniela Seifertová, Gerd Jürgens, and Jirí Friml. 2003. "Local, Efflux-Dependent Auxin Gradients as a Common Module for Plant Organ Formation." *Cell* 115 (5): 591–602.

Bhatia, Saurabh. 2015. "Chapter 2 - Plant Tissue Culture." In *Modern Applications of Plant Biotechnology in Pharmaceutical Sciences*, edited by Saurabh Bhatia, Kiran Sharma, Randhir Dahiya, and Tanmoy Bera, 31–107. Boston: Academic Press.

Bowman JL, Baum SF, Eshed Y, Putterill J, Alvarez J. (1999) Molecular genetics of gynoecium development in Arabidopsis. Curr Top Dev Biol 45:155–205

Chuang, C. F., M. P. Running, R. W. Williams, and E. M. Meyerowitz. 1999. "The PERIANTHIA Gene Encodes a bZIP Protein Involved in the Determination of Floral Organ Number in Arabidopsis Thaliana." *Genes & Development* 13 (3): 334–44.

Coen, E.S., and Meyerowitz, E.M. (1991). The war of the whorls: genetic interactions controlling flower development. Nature 353, 31–37.

Cunninghame, Morag E., and R. F. Lyndon. 1986. "The Relationship Between the Distribution of Periclinal Cell Divisions in the Shoot Apex and Leaf Initiation." *Annals of Botany* 57 (6): 737–46.

Dello Ioio, Raffaele, Francisco Scaglia Linhares, Emanuele Scacchi, Eva Casamitjana-Martinez, Renze Heidstra, Paolo Costantino, and Sabrina Sabatini. 2007. "Cytokinins Determine Arabidopsis Root-Meristem Size by Controlling Cell Differentiation." *Current Biology: CB* 17 (8): 678–82.

Elliott, R. C., A. S. Betzner, E. Huttner, M. P. Oakes, W. Q. Tucker, D. Gerentes, P. Perez, and D. R. Smyth. 1996. "AINTEGUMENTA, an APETALA2-like Gene of Arabidopsis with Pleiotropic Roles in Ovule Development and Floral Organ Growth." The Plant Cell 8 (2): 155–68.

Endrizzi, K., B. Moussian, A. Haecker, J. Z. Levin, and T. Laux. 1996. "The SHOOT MERISTEMLESS Gene Is Required for Maintenance of Undifferentiated Cells in Arabidopsis Shoot and Floral Meristems and Acts at a Different Regulatory Level than the Meristem Genes WUSCHEL and ZWILLE." *The Plant Journal: For Cell and Molecular Biology* 10 (6): 967–79.

Fosket, Donald E., and Others. 1994. *Plant Growth and Development: A Molecular Approach*. Academic Press Inc.

Foster, Adriance S. 1939. "Structure and Growth of the Shoot Apex of Cycas Revoluta." American Journal of Botany 26 (6): 372–85.
Harrison, C. Jill, Mohi Rezvani, and Jane A. Langdale. 2007. "Growth from Two Transient Apical Initials in the Meristem of Selaginella Kraussiana." Development 134 (5): 881–89.

Greb, Thomas, and Jan U. Lohmann. 2016. "Plant Stem Cells." *Current Biology: CB* 26 (17): R816–21.

Gremski, Kristina, Gary Ditta, and Martin F. Yanofsky. 2007. "The HECATE Genes Regulate Female Reproductive Tract Development in Arabidopsis Thaliana." *Development* 134 (20): 3593–3601.

Groszmann, Michael, Yasmin Bylstra, Edwin R. Lampugnani, and David R. Smyth. 2010. "Regulation of Tissue-Specific Expression of SPATULA, a bHLH Gene Involved in Carpel Development, Seedling Germination, and Lateral Organ Growth in Arabidopsis." *Journal of Experimental Botany* 61 (5): 1495–1508.

Harrison, C. Jill, Mohi Rezvani, and Jane A. Langdale. 2007. "Growth from Two Transient Apical Initials in the Meristem of Selaginella Kraussiana." *Development* 134 (5): 881–89.

Hampton, Marshall, Wayne W. Xu, Brian W. Kram, Emily M. Chambers, Jerad S. Ehrnriter, Jonathan H. Gralewski, Teresa Joyal, and Clay J. Carter. 2010. "Identification of Differential Gene Expression in Brassica Rapa Nectaries through Expressed Sequence Tag Analysis." *PloS One* 5 (1): e8782.

Hibara, Ken-Ichiro, Shinobu Takada, and Masao Tasaka. 2003. "CUC1 Gene Activates the Expression of SAM-Related Genes to Induce Adventitious Shoot Formation: Role of CUC1 in Adventitious SAM Formation." *The Plant Journal: For Cell and Molecular Biology* 36 (5): 687–96.

Jacobs, W. P., and I. B. Morrow. 1961. "A Quantitative Study of Mitotic Figures in Relation to Development in the Apical Meristem of Vegetative Shoots of Coleus." *Developmental Biology* 3 (October): 569–87.

Khan, M., Ragni, L., Tabb, P., Salasini, B. C., Chatfield, S., Datla, R., et al. (2015). Repression of lateral organ boundary genes by PENNYWISE and POUND- FOOLISH is essential for meristem maintenance and flowering in Arabidopsis. *Plant Physiol.* 169, 2166–2186. doi: 10.1104/pp.15.00915

Klucher, K. M., H. Chow, L. Reiser, and R. L. Fischer. 1996. "The AINTEGUMENTA Gene of Arabidopsis Required for Ovule and Female Gametophyte Development Is Related to the Floral Homeotic Gene APETALA2." *The Plant Cell* 8 (2): 137–53.

Koch, L. 1891. "Ueber bau und Wachsthum der sprossspitze der phanerogamen. I. Die Gymnospermen." Jahrb. f. Wiss. Bot. 22: 481-680.

Kunst, L., J. E. Klenz, J. Martinez-Zapater, and G. W. Haughn. 1989. "AP2 Gene Determines the Identity of Perianth Organs in Flowers of Arabidopsis Thaliana." *The Plant Cell* 1 (12): 1195–1208.

Larsson, Emma, Christina J. Roberts, Andrea R. Claes, Robert G. Franks, and Eva Sundberg. 2014. “Polar Auxin Transport Is Essential for Medial versus Lateral Tissue Specification and Vascular-Mediated Valve Outgrowth in Arabidopsis Gynoecia.” *Plant Physiology* 166 (4): 1998–2012.

Lawrence, Cera R. 2008. “Preformationism in the Enlightenment.” *Arizona State University. School of Life Sciences. Center for Biology and Society. Embryo Project Encyclopedia.*, November.

Liu Z, Franks RG, Klink VP (2000) Regulation of gynoecium marginal tissue formation by LEUNIG and AINTEGUMENTA. Plant Cell 12:1879–1892

Medford, J. I. 1992. “Vegetative Apical Meristems.” *The Plant Cell* 4 (9): 1029–39.

Möller, Barbara K., Colette A. Ten Hove, Daoquan Xiang, Nerys Williams, Lorena González López, Saiko Yoshida, Margot Smit, Raju Datla, and Dolf Weijers. 2017. “Auxin Response Cell-Autonomously Controls Ground Tissue Initiation in the Early Arabidopsis Embryo.” *Proceedings of the National Academy of Sciences of the United States of America* 114 (12): E2533–39.

Müller, Christina Joy, Emma Larsson, Lukáš Spíchal, and Eva Sundberg. 2017. “Cytokinin-Auxin Crosstalk in the Gynoecial Primordium Ensures Correct Domain Patterning.” *Plant Physiology* 175 (3): 1144–57.

Nemhauser, J. L., L. J. Feldman, and P. C. Zambryski. 2000. “Auxin and ETTIN in Arabidopsis Gynoecium Morphogenesis.” *Development* 127 (18): 3877–88.

Okada, K., M. K. Komaki, and Y. Shimura. 1989. “Mutational Analysis of Pistil Structure and Development of Arabidopsis Thaliana.” *Cell Differentiation and Development: The Official Journal of the International Society of Developmental Biologists* 28 (1): 27–37.

Przemeck, G. K., J. Mattsson, C. S. Hardtke, Z. R. Sung, and T. Berleth. 1996. “Studies on the Role of the Arabidopsis Gene MONOPTEROS in Vascular Development and Plant Cell Axialization.” *Planta* 200 (2): 229–37.

Priestley, J. H. 1929. “Cell Growth and Cell Division in the Shoot of the Flowering Plant.” *The New Phytologist* 28 (1): 54–81.

Reddy, G. Venugopala. 2008. “Live-Imaging Stem-Cell Homeostasis in the Arabidopsis Shoot Apex.” *Current Opinion in Plant Biology* 11 (1): 88–93.

Reddy, G. Venugopala, and Elliot M. Meyerowitz. 2005. “Stem-Cell Homeostasis and Growth Dynamics Can Be Uncoupled in the Arabidopsis Shoot Apex.” *Science* 310 (5748): 663–67.

Reyes-Olalde, J. Irepan, and Stefan de Folter. 2019. “Control of Stem Cell Activity in the Carpel Margin Meristem (CMM) in Arabidopsis.” *Plant Reproduction* 32 (2): 123–36.

Reyes-Olalde, J. Irepan, Víctor M. Zúñiga-Mayo, Nayelli Marsch-Martínez, and Stefan de Folter. 2017. “Synergistic Relationship between Auxin and Cytokinin in the Ovary and the Participation of the Transcription Factor SPATULA.” *Plant Signaling & Behavior* 12 (10): e1376158.

Ruth, Jennifer, Edward J. Klekowski, and Otto L. Stein. 1985. “Impermanent Initials of the Shoot Apex and Diplontic Selection in a Juniper Chimera.” *American Journal of Botany* 72 (7): 1127–35.

Sachs. N. D. *Lectures on the Physiology of Plants*. Oxford: The Clarendon Press.

Scofield, Simon, Walter Dewitte, and James A. H. Murray. 2007. “The KNOX Gene SHOOT MERISTEMLESS Is Required for the Development of Reproductive Meristematic Tissues in Arabidopsis.” *The Plant Journal: For Cell and Molecular Biology* 50 (5): 767–81.

Ståldal, Veronika, and Eva Sundberg. 2009. “The Role of Auxin in Style Development and Apical-Basal Patterning of the Arabidopsis Thaliana Gynoecium.” *Plant Signaling & Behavior* 4 (2): 83–85.

Stepanova, Anna N., Joyce Robertson-Hoyt, Jeonga Yun, Larissa M. Benavente, De-Yu Xie, Karel Dolezal, Alexandra Schlereth, Gerd Jürgens, and Jose M. Alonso. 2008. “TAA1-Mediated Auxin Biosynthesis Is Essential for Hormone Crosstalk and Plant Development.” *Cell* 133 (1): 177–91.

Sundberg, Eva, and Lars Østergaard. 2009. “Distinct and Dynamic Auxin Activities during Reproductive Development.” *Cold Spring Harbor Perspectives in Biology* 1 (6): a001628.

Van Speybroeck, Linda, Dani De Waele, and Gertrudis Van de Vijver. 2002. “Theories in Early Embryology: Close Connections between Epigenesis, Preformationism, and Self-Organization.” *Annals of the New York Academy of Sciences* 981 (December): 7–49.

Villarino, Gonzalo H., Qiwen Hu, Silvia Manrique, Miguel Flores-Vergara, Bhupinder Sehra, Linda Robles, Javier Brumos, et al. 2016. "Transcriptomic Signature of the SHATTERPROOF2 Expression Domain Reveals the Meristematic Nature of Arabidopsis Gynoecial Medial Domain." *Plant Physiology* 171 (1): 42–61.

Wardlaw, C. W. 1951. "COMPARATIVE MORPHOGENESIS IN PTERIDOPHYTES AND SPERMATOPHYTES BY EXPERIMENTAL METHODS*." *The New Phytologist* 50 (1): 127–34.

Wardlaw, C. W., Cornelia A. Reinders-Gouwentak, P. Champagnat, J. Doorenbos, and A. Allsopp. 1965. "Organisation Und Ausgestaltung Des Sprosses. — Organization and Development of the Shoot." In *Differentiation and Development / Differenzierung Und Entwicklung: Part 1 / Teil 1*, edited by Allan Allsopp, Robert Bloch, John Tyler Bonner, Martin Bopp, Jean Brachet, Robert Brown, Erwin Bünning, et al., 966–1255. Berlin, Heidelberg: Springer Berlin Heidelberg.

Wright, Amanda J., and Laurie G. Smith. 2008. "Division Plane Orientation in Plant Cells." In *Cell Division Control in Plants*, edited by Desh Pal S. Verma and Zonglie Hong, 33–57. Berlin, Heidelberg: Springer Berlin Heidelberg.

Wynn, April N., Andrew A. Seaman, Ashley L. Jones, and Robert G. Franks. 2014. "Novel Functional Roles for PERIANTHIA and SEUSS during Floral Organ Identity Specification, Floral Meristem Termination, and Gynoecial Development." *Frontiers in Plant Science* 5 (April): 130.

CHAPTER 2

A Gene Regulatory Network Representation of a Vital Reproductive Meristem in the *Arabidopsis thaliana* Gynoecium

Brian Sullivan[1], Ryan Spurney[1,2], Rosangela Sozzani[1], Robert Franks[1]

[1]Plant and Microbial Biology Department, North Carolina State University, Raleigh, NC 27695

[2]Electrical and Computer Engineering Department, North Carolina State University, Raleigh, NC 27695

Abstract

Ovules are produced in the two fused carpels which make up the gynoecium of the *Arabidopsis thaliana* flower. Ovules are produced from a meristematic structure that forms at the margin of the two fused carpels. The Carpel Margin Meristem (CMM) is responsible for the production of medial produced tissues including the placenta, ovules, and septum. The gynoecium primordia are produced during stage 5 of floral development, the carpels fuse at stage 6, the CMM becomes visible at stage 7, and the two CMMs fuse at stage 9 producing the septum and at the same time, placental tissue is produced which will later initiate ovules laterally. Despite this morphological characterization, there has not been a comprehensive study of gene expression of this developmental process. We have isolated transcriptomes from the medial and lateral domains of the gynoecium over five days of development which cover approximately stage 7 to stage 12. Through analysis of these transcriptomes, we have determined that the tissue of origin is more important to transcriptome identity than the developmental stage of the tissue. Further, these transcriptomes change over time with a key shift occurring between 6 and 7 days after floral induction (6 to 7 DAI). Through modeling of this transition, we have identified a network of 11

transcription factors that may control a transcriptional shift during CMM development. Analysis of these genes shows that there are members of TGA and NAC families, which include the regulators of CMM development *PERIANTHIA* and *CUP-SHAPED COTYLEDONS 1* and *2* respectively. This study presents a model of a transcriptional shift in the CMM which may describe a shift from more to less meristematic identity and allow for the integration of many studies into the development of the CMM.

Introduction

In plants, stem cells are located in meristems that maintain themselves and divide to provide cells for the growing organs. There are two types of meristems in *Arabidopsis thaliana*; indeterminate and determinate (Melzer et al. 2008). The shoot and root apical meristems (SAM and RAM) are indeterminate in that they contain a set of completely undifferentiated and slowly dividing cells that do not possess a defined end status. In contrast, the floral meristem is determinate in that it produces specific structures and possesses a defined end status. Both types of meristems must be able to maintain themselves and make other cells. They differ in the length of time that they are maintained and the types of tissues eventually produced.

The pathways regulating meristem identity and maintenance are regulated by the complex network of interactions governed by two types of genes; those that promote meristematic identity and those that function to specify cell fate. Three *CLAVATA* genes (*CLV1*,*2*,*3*) interact with one another in a receptor complex to determine meristem size through a signal transduction pathway that promotes cell differentiation and the movement of cells from the meristem into the organ primordia (Clark et al. 1993, Clark et al. 1996, Kayes and Clark 1998).

Meanwhile, *SHOOTMERISTEMLESS* (*STM*) functions in a parallel pathway to competitively promote stem cell identity through the direct binding of the promoters of several Myb family genes as well as the *CLV* genes (Long et al. 1996, Byrne et al. 2000, Kayes et al. 1998). Also, *WUSCHEL* (*WUS*) is a transcription factor that functions to promote stem cell fate. *WUS* is regulated by the *CLV* signaling pathway and signaling through this pathway leads to the repression of *WUS* and differentiation (Laux et al. 1999).

In determinate meristems, *AGAMOUS* (*AG*) plays a critical role in ensuring determinacy and acts in opposition to *WUS*. The interactions between LFY, WUS, and *AG* highlight the complex pathways regulating meristem development; the binding of LFY and WUS to the second intron of the *AG* genic sequence enhances *AG* expression (Lohmann et al. 2001). *AG* then downregulates the expression of *WUS* which allows for stem cells to differentiate and the carpels of the flower to form. In *ag* mutant plants, floral meristems are indeterminate and continue to make concentric whorls of only sepals and petals due to *AG*'s role in regulating the floral meristem determinacy and in specifying floral organ identity. Ectopic expression of *AG* causes plants to form only a single flower, due to the premature termination of the inflorescence meristem. In contrast to these phenotypes, *wus* mutants fail to develop properly in that many meristematic structures are produced but fail to develop properly because WUS cannot promote stem cell identity, which is required to maintain a meristem.

The complex interactions that govern meristem development have long been investigated through mutant analyses and molecular genetics. These analyses have been able to identify important regulators of meristem and floral development with a focus on those genes that cause

floral homeotic transformation or prevent the proper development of the SAM. To understand these complex interactions, modeling the gene regulatory networks involved in these processes may provide a more broad view in which to frame our current understanding.

Gene Regulatory Network Representations of the CMM

Generating a model of the gene regulatory networks (GRNs) that drives meristem development over time provides a scaffold from which future analyses can be built. This type of analysis uses transcriptomic data to generate a map of regulatory interactions which can allow researchers to make predictions on which genes may represent key regulators of development. There are a variety of methods for generating GRNs that use different algorithms to build a network of predicted interactions (edges) between nodes (genes). These methods have been able to predict interactions among differentially expressed genes in the different stem cell niches produced by the *Arabidopsis thaliana* root apical meristem (de Luis Balaguer et al. 2017), to model the development of the floral meristem and floral organ differentiation (Chen et al. 2018), and to identify genes regulating senescence in *Triticum aestivum* (Borrill et al. 2019).

Due to the complex nature of meristem development, it has been very challenging to generate a complete model of the molecular genetic interactions that govern this process. Furthermore, the different types of plant meristems likely employ a mixture of common and unique regulatory mechanisms that are required to achieve their different biological functions. To investigate meristem development, it is important to take into account both spatial and temporal aspects of the developmental process. The carpel margin meristem (CMM), a meristematic structure found at the margin of the two fused carpels in *Arabidopsis,* presents an ideal structure for this study as

this tissue is a vital reproductive meristem, producing the ovules, septum, transmitting tract, and placenta (Alvarez and Smyth 2002). In addition, it maintains its meristematic identity for a relatively short period and then terminates as it differentiates into a defined set of CMM-derived structures (Nole-Wilson et al. 2010, Reyes-Olalde et al. 2014). Another advantage is that the CMM is positioned adjacent to the lateral domain of the gynoecium, which loses its meristematic potential more quickly during the development of the gynoecium than does the medially-positioned CMM (Liu et al. 2000) and thus allows for the comparison of meristematic and less meristematic tissues over time.

To investigate these developmental processes, transcriptome data from distinct spatial domains must be combined with temporal information. Generating a representation of these gene regulatory networks allows us to integrate temporal and spatial data into a single model. In addition, it presents an ideal approach to modeling a developmental process in a tissue where regulation is controlled by numerous regulators both genetic and otherwise that show redundancy (Moore et al. 2005). This is due to several factors, first, most modern analyses of the developing CMM have focused on genes that cause a large phenotype when they are mutated (Franks et al. 2002), that are highly differentially expressed (Villarino et al 2016), or that are involved in hormone-mediated signaling pathways. These genes have detectable mutant phenotypes and have been easier to identify and characterize than genes that might encode redundant functions during CMM development. A GRN prediction approach may allow for the identification of novel gynoecium developmental regulators that have been missed by these other approaches. A second novel aspect to GRN prediction is that it allows for an unsupervised approach which removes

cutoffs traditionally selected by the researcher and identifies regulators that would have been missed using a traditional cutoff (Spurney et al. 2020).

To investigate the dynamics of the gene regulatory networks regulating the development of the meristematic CMM, we identified tissue enriched, differentially-expressed genes over five developmental stages of gynoecium development. We determined the sets of genes that are differentially expressed within the medial and lateral gynoecial domains across the gynoecial developmental progression. Using the transcriptomics data, we determined that the transcriptomic profiles differed significantly and consistently between the medial and lateral domains at all time points analyzed. Additionally, we noted that the differences between the two spatial domains were more dramatic than the differences within a given spatial domain when the different developmental time points were compared. Tissue played a more important role in the identity of the transcriptomes than did the time point. In addition, we were able to identify a transcriptional shift in the medial domain which may represent a transition from a determination phase to one of differentiation. Using this transcriptomics data we created a network representation of the gene regulatory relationships within the medial domain. This GRN predicts putative critical regulatory nodes that may orchestrate the transition of this important reproductive meristem and thus enable the biologically appropriate generation of the ovule and other critical reproductive floral structures.

Results

Isolation of Temporally and Spatially Distinct Transcriptomes from the Medial and Lateral Gynoecial Domains across a Developmental Time Series

To better understand the developmental dynamics of a reproductive meristem we collected and characterized transcriptomic data sets from distinct spatial domains across a time course of gynoecial development. We chose to study the medial and lateral gynoecial domains, two adjacent tissues with different levels of meristematic identity, to allow us to identify mechanisms that regulate the meristematic character of the CMM. We combined this spatial separation with a time course to ensure that we could assay the dynamics of the gene regulatory networks driving the development of the medial and lateral domains.

To accomplish this, we first used a modified, dual-reporter system that marks the approximate medial and lateral domains of the gynoecium and allows for their physical isolation at distinct time points during the development of the gynoecium (Villarino et al. 2015). To mark the medial domain cells, we used a *pSHP2-YFP* two-component reporter system (pUAS-3xYPET; pSHP2-GAL4) as previously described (Villarino et al. 2016). To mark the lateral gynoecial domains, we used a second reporter, *pFUL:mCherry* transcriptional reporter. *FRUITFUL* (*FUL*) encodes a MADS box-containing transcription factor that is expressed in the lateral domain where it functions to mediate cell differentiation. We refer to the domain marked by the *pFUL:mCherry* reporter as the lateral domain for the rest of this paper (Balanza et al. 2019, Gu et al. 1998).

To study the temporal aspect of gene regulation within the developing Arabidopsis gynoecium, we used a floral synchronization system (Wellmer et al. 2006) to generate large quantities of flowers at similar stages of development (Villarino et al, 2016). We combined this genetic background with a fluorescence-activated cell sorting (FACS) strategy which sampled five time points at 24 hour intervals to create a sampling strategy that allowed for the isolation of protoplasts from two domains over five time points covering five days of development. Through FACS we can separate the medial meristematic domain and the lateral domain from the rest of the floral tissue. Illumina-based RNA sequencing procedures resulted in the generation of transcriptomic datasets for each of the two expression domains (medial and lateral) across the 5 developmental time points, that were captured in triplicate replicates. Analysis of the expression levels of the endogenous *SHP2* and *FUL* transcripts indicated strong expression enrichment for these RNAs within the transcriptomes of the medial and lateral domains, respectively (Supplemental Figures 2.1,2.2).

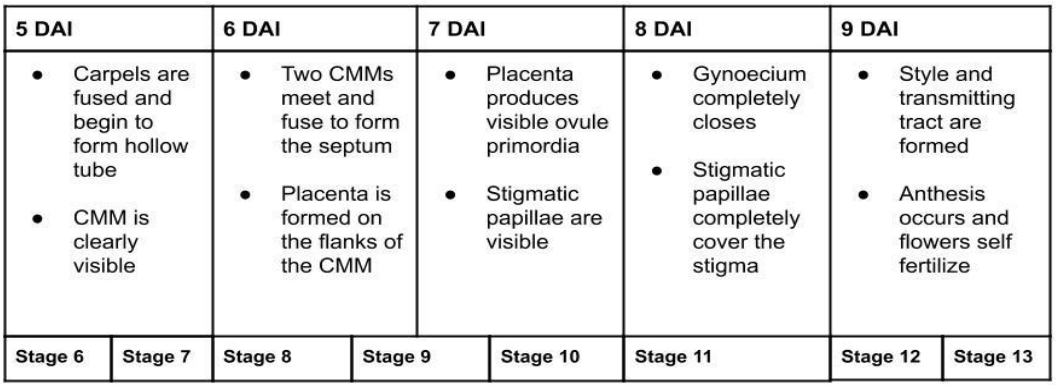


Figure 2.1 CMM Development. A system for isolating and analyzing the transcriptomes of two domains of the gynoecium as they develop over 120 hours. Staging is approximate. Cells were isolated from the medial and lateral domains of the gynoecium through FACS. *Confocal microscopy experiments were underway to confirm the staging and provide information on reporter expression but these analyses were disrupted by the COvid-19 pandemic.

The Medial and Lateral Domains of the *A. thaliana* Gynoecium Are Transcriptionally Distinct Across the Developmental Progression of the Organ

To validate our experimental technique we then performed a principal component analysis (PCA) using all expressed genes to ensure that biological replicates clustered close to one another (Supplemental Figure 2.3). The variability of gene expression values between the time points was less than the variability between the spatial domains, suggesting that the two spatial domains were indeed transcriptionally distinct throughout the developmental period that we assayed (Supplemental figure 2.4).

To build a GRN representation of the transcriptional development of the CMM we identified genes that were differentially expressed in the domain to be modeled and assayed their expression over time. Modified Shannon entropy (MSE) was used to select differentially expressed genes (DEGs). For a gene to be included in one of these selections, first it must meet the entropy selection criteria (Zambelli et al. 2018); it is assigned an optimal outlier value for each time point in each tissue and is then selected if it is only differentially expressed in only one of the two tissues for that time point. By applying this selection method, we can extract the genes that are both tissue enriched and vary significantly across the time points. This approach allowed us to analyze the two spatial domains separately and highlight changes of expression across the time series from each spatial domain separately. Using this MSE approach we identified 328 genes whose expression was enriched in the medial domain and that also varied significantly across the time course (Figure 2.2 A, Supplemental Table 2.1). The expression values of these genes provide a starting dataset for the construction of a GRN representation of the medial domain. We also identified 827 genes that were enriched in the lateral domain and that varied significantly across the time course in the lateral domain (Supplemental Figure 2.2), but analysis of this lateral domain gene set will be presented separately.

Using the expression values of the 328 genes identified as both medial domain enriched and time point variable, we performed PCAs to analyze the multidimensional variability between these samples (Figure 2.2B). In this analysis of the medial domain data, we found that there was a separation between early and late time points that was identified by the first principle component (PC1) which captured 42% of the expression variation. The transcriptomes of 5 DAI and 6 DAI time points are clustered together with separation from the 7, 8, and 9 DAI time points. This supports the idea that there are differences between the early and late transcriptomes in the developing medial domain and suggests that a gene regulatory shift occurs between the 6 and 7 DAI time points. Also, the 7 DAI time point is separated from both 6 DAI and 8 DAI time points which suggests that there is a distinct transition state that occurs between early and late medial domain development (Figure 2B).

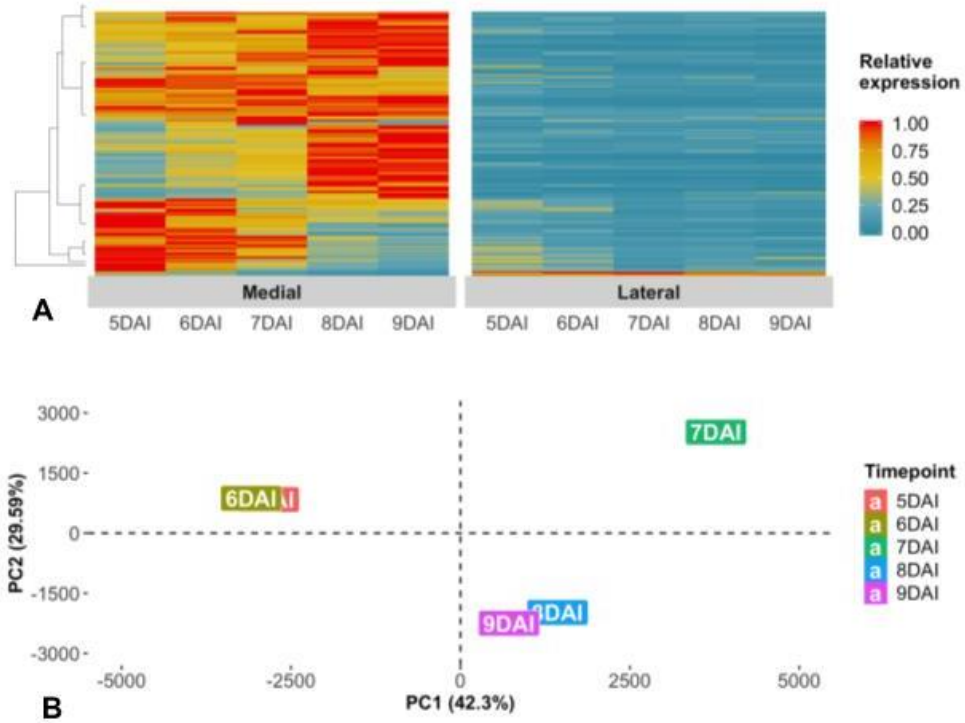


Figure 2.2 An MSE-based approach identifies a set of 328 medial domain-enriched genes. A) A heatmap representation of the expression levels of the set of 328 differentially-expressed medial domain genes. Hierarchical clustering shows four major groups of genes with distinct expression patterns, one of which appears to be expressed early in development and turned off after 7DAI, while the second group of genes is relatively highly expressed with expression peaking at 7DAI and then tapering off. **B)** A PCA representation of the expression levels of the set of 328 differentially-expressed, medial domain genes. The 7 DAI time point is well separated from both the 5/6 DAI grouping and the 8/9 DAI grouping, suggesting a significant gene regulatory shift during the 7 DAI time point.

Gene Expression Reprogramming Separates Early and Late Medial Domain Development

To further characterize regulatory shifts during medial domain development, we identified expression state changes during the developmental progression of the medial domain. We did this by enumerating instances of MSE expression-state transitions. An MSE expression state transition is defined here as a shift from an MSE state of an outlier to a non outlier or the reciprocal. For

example between DAI6 and DAI7, 277 genes exhibited an MSE expression state transition (Supplemental Table 2.2.). To determine where the greatest number of expression state transitions occurred, we counted the transitions observed between all adjacent time points pairs (Figure 3A). A total of 604 transitions occur for these 728 genes over the entire time course in the medial domain. The majority of these transitions occur from 6DAI to 7DAI (46%) and 7DAI to 8DAI (27%). This analysis again suggests that major gene expression transitions occur between these time points.

A Gene Regulatory Network Representation of the Medial Domain Development Defines Key Regulatory Nodes

To better understand the transcriptional reprogramming events that occurred between the 6 DAI and 7 DAI time points, we modeled the GRN regulating this transition. GRN inference can reveal genes of high importance in a biological system. To create a GRN representation of medial domain development, we chose to focus on the set of 277 genes that exhibited an expression-state transition between the 6 DAI and 7 DAI samples. We used RTP-STAR (Spurney et al. 2019, Clark et al. 2019), a decision tree machine learning algorithm, to infer a GRN for these 277 genes (Methods, Supplemental Table 2.3). The gene regulatory network modeling this transition is primarily composed of two transcription factors regulating many other genes, most of which are not transcription factors, and accordingly, they are not predicted to regulate other genes. These two predicted regulators of the six-seven DAI transition, *LOB DOMAIN-CONTAINING PROTEIN 21* (*LBD21*) and *TGACG* (TGA) *MOTIF-BINDING PROTEIN 9* (*TGA9*), are differentially expressed beginning at 7 DAI. These two nodes were chosen as potential key regulators in the network because of their high number of predicted regulations.

To best predict key nodes in our model we also performed motif scoring which takes into account whole network architecture to choose regulators that are predicted to have the largest impact on the developmental process. We chose motif scoring because it is a method to select key nodes more rigorously than simply selecting genes with the highest number of outgoing regulations. We found that there were no significant motifs and therefore chose those genes with the most predicted outgoing regulations (edges) to focus on for further investigation.

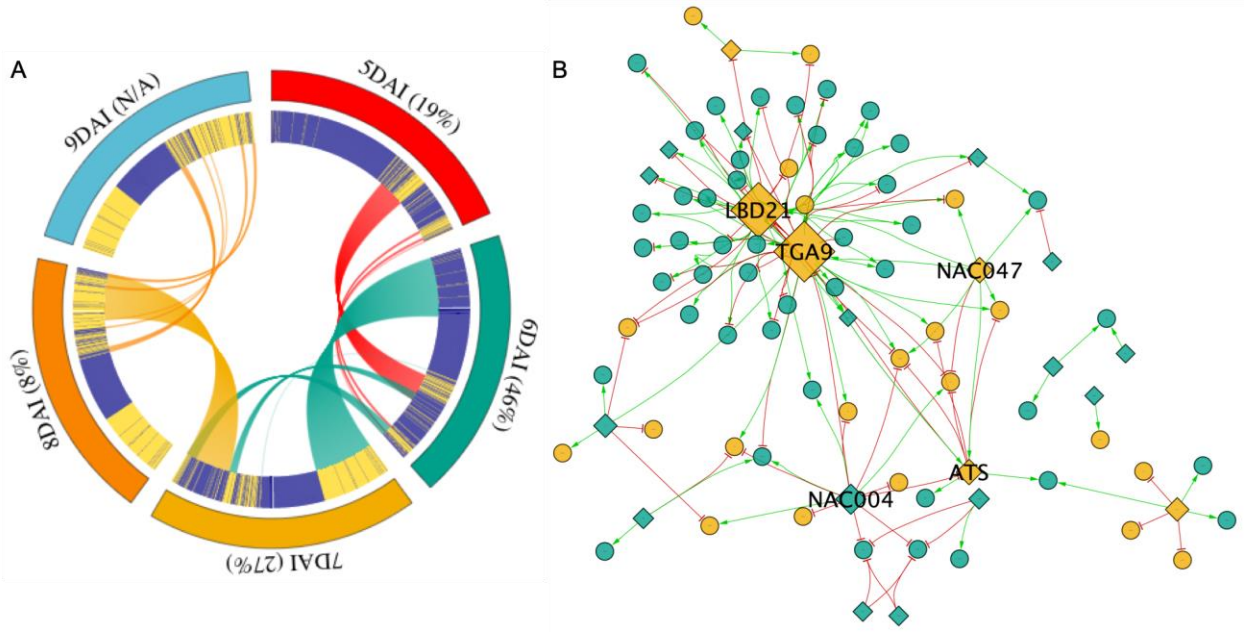


Figure 2.3 Transcriptional Shift Between 6DAI and 7DAI. A transcriptional shift between 6 and 7 DAI highlights the temporal dynamics of medial domain development. (A) Plot showing expression state transitions for genes differentially expressed in at least one medial domain time point (728 genes). Blue genes are those that are outliers which indicates that they are differentially expressed while those in yellow are not outliers (not DEGs). (B) GRN inferred using only genes that undergo an MSE expression state transition between 6DAI and 7DAI (277 genes). Green nodes represent genes that are differentially expressed at 6DAI and not at 7DAI while yellow nodes represent genes that are differentially expressed at 7DAI but not at 6DAI. The diamonds represent transcription factors while the circles represent all other genes. The size of the diamond represents the number of outgoing edges.

To limit the size of the GRN and to allow for the selection of priority nodes for functional selection, we limited the gene regulatory network to transcription factors with a calculated motif score of greater than 0.3 (Figure 4, Supplemental Table 2.4). This generated a list of 11 high priority nodes/genes for functional validation.

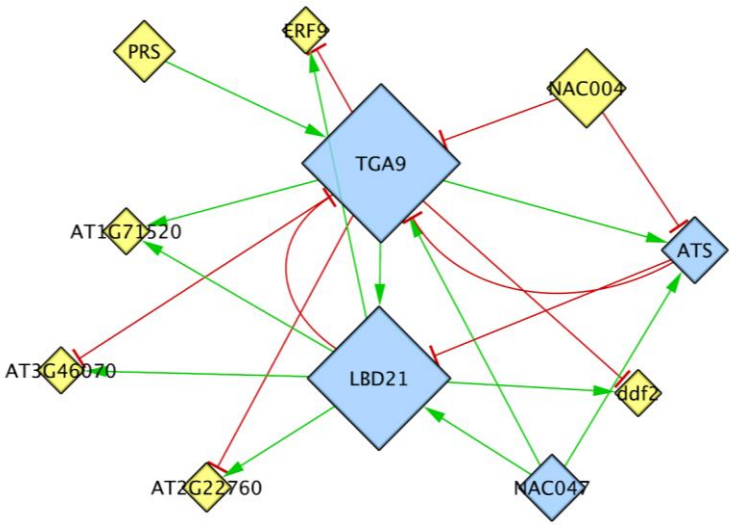


Figure 2.4 Model Gene Regulatory Network Controlling the 6 to 7 DAI Transcriptional Shift. Using motif scoring allows for the identification of 11 genes which may be be important regulators of the 6 to 7 Dai transition in gene expression. In this model generated using expression values for these 11 genes at all time points, all nodes are diamonds as they are transcription factors. The yellow nodes represent those genes that are differentially expressed 6 DAI while those in blue are expressed 7 DAI. The size of the node represents the number of outgoing regulations of that node.

Discussion

The gene regulatory networks driving gynoecium development are dynamic and reflect carefully regulated gene expression changes over time. The medial and lateral domains are fundamentally distinct, in morphology, meristematic potential, and gene expression. This is confirmed through PCA of samples taken from each domain of the gynoecium over time, which found that tissue

type contributed to a larger portion of the variation between samples than did time point (Supplemental Figure 3). Despite this finding, we did notice gene expression changes over time in both the medial and lateral domains of the gynoecium. This is highlighted by a large transition in gene expression occurring in the medial domain between 6 and 7 days after floral induction. We investigated changes in differentially expressed genes within a single domain over time and found that there were 608 expression state transitions. Between 6 DAI and 7 DAI, over 46% of transitioning genes underwent a transition with around 75% of these shifting from differentially expressed to no longer differentially expressed between 6 DAI and 7 DAI. To better understand the transcriptional dynamics specific to this developmental time point we used the DEGs exhibiting an expression shift at this time point to model a gene regulatory network. We identified two main transcription factors driving this change; TGA9 and LBD21 and found that these two genes were differentially expressed at 7 DAI and not differentially expressed 6 DAI, specifically, both TGA9 and LBD21 expression peaks at 7DAI and then tapers off. This is not the only pattern seen in the network controlling this transition, *ETHYLENE RESPONSE FACTOR 9* (*ERF9)* expression drops 63% from 6 to 7 DAI (Supplemental Table 1). This latter phenomenon is repeated in six of the 9 other genes. These newly identified nodes represent a starting point for the functional validation of this network, which may lead to a better understanding of the transcriptional regulatory logic of the meristematic transition within a vital reproductive meristem.

TGA9 as a Putative Key Regulator of Gynoecium Development

TGA9 is predicted to be a key regulator of the transcriptional shift which occurs between six and seven DAI. This transcription factor is upregulated at 7DAI in the medial domain despite

relatively high expression throughout the time course (Supplemental Table 1). TGA9 is a member of the TGA family of transcription factors which are divided into five clades (Gatz 2013). TGA8/9 are members of the fourth clade and are involved in regulating stamen development through interactions with two floral glutaredoxins; ROXY1 and ROXY2 (Murmu et al. 2010). There is one other member of the TGA family that has been shown to interact with ROXY1 and ROXY2 (Li et al. 2009), PERIANTHIA (PAN). The PAN protein is post-translationally modified at a cysteine residue at the n-terminus by ROXY1 and ROXY2. PAN is a TGA family transcription factor (TGA8) clustered in the fifth clade and is the only other TGA protein known to regulate floral development.

PAN was discovered due to its role in determining floral organ number and was found to regulate gynoecium development through the upregulation of *AG* expression (Running and Meyerowitz 1996, Das et al. 2009, Gutsche and Zachgo 2016). In *pan seu* double mutants there is a floral indeterminacy phenotype, which results in the formation of a fifth whorl (Wynn et al. 2014). This is due to the inability of the mutant allele of *PAN* to upregulate *AG* transcription which results in greater *WUS* expression and an indeterminate meristem (Liu et al. 2011). In contrast, in the sepals, PAN acts to repress *AG* expression, which suggests varied roles for TGA transcription factors in the different whorls of the flower (Wynn et al. 2014). Given this information, we suggest that TGA9 is a key regulator behind the transcriptional shift seen between six and seven DAI. Future phenotypic and molecular genetic analyses of *tga9* mutant gynoecia will be necessary to further determine the role of *TGA9* during CMM development.

Neogenic to Determination and Differentiation Phase Transition

In the paper in which they first defined the CMM (Alvarez and Smyth 2002), the authors hypothesize that the development of the CMM occurs in two stages. They propose the existence of an early "neogenic" phase where separate transcriptomic identities of the medial and lateral domains have been specified and plastic or meristematic (neogenic) potential is maintained specifically in the cells of the medial domain. During the neogenic phase, several key transcription factors regulate the expression of other genes and thus the development of the medial and lateral domains. In the medial domain, *SPATULA* (*SPT*) and SHOOT MERISTEMLESS (STM) along with two NAC family transcription factors, CUC1 and CUC2 function to maintain the meristematic potential of the medial domain through STM activity. Alvarez and Smyth hypothesized that CUC1 and CUC2 act to integrate auxin signaling into the gene regulatory network promoting the meristematic identity of the medial domain during the neogenic phase (Nahar et al. 2012). This was supported by the reduced expression of these genes in the later phase of development. In our model of the transition between six and seven DAI, we identify a different pair of NAC transcription factors, one of which (NAC004) has high expression early in development which is drastically reduced later, while the other (NAC047) peaks in expression 7 DAI which is then reduced. Future work will be necessary to determine if *NAC047* and *NAC004* are part of a network required for CMM development, similar to the network involving PAN, STM, SPT, CUC1, and CUC2.

In the Alvarez and Smyth model of medial domain development (Alvarez and Smyth 2002) a "determination and differentiation" phase follows the neogenic phase. In this second phase, cells that already possess medial domain identity undergo a period of cell division, followed by a

period of definition, during which they begin to develop into the different medially derived structures (i.e. ovules, septum, transmitting tract). Our research builds upon this model by predicting key interactions regulating a transcriptional shift which may be related to the transition of medial domain cells from a more plastic state into differentiated cell types. In contrast to the original model proposed by Alvarez and Smyth, the shift we have identified appears to occur between stages 8 and 9 of gynoecial development rather than stages 7 and 8 (Alvarez and Smyth, 2002). One reason for this difference is that we may not be modeling the transition between the neogenic phase and the determination and differentiation phase but are instead looking at a shift in tissue identity from more to less meristematic that occurs after the neogenic phase has concluded. Despite questions about the exact biological meaning of the transcriptional shift, our data show that there are two distinct phases of medial domain development between stage 7 and stage 12.

Transcriptional Shifts are Key Features of Development and May Share Characteristics

With the rise of RNA sequencing and the proliferation of studies elucidating GRNs, transcriptional shifts have been discovered during many different areas of plant development (Palumbo et al. 2014, Cruz-Ramirez et al. 2012, Liebers et al. 2017). These shifts may share characteristics despite differing biological backgrounds. Two of these characteristics associated with transcriptional shifts are that shift genes are often upregulated specifically at the time of the shift with a separate set of genes being downregulated after the shift and that switch genes are often transcription factors including NAC family transcription factors, although that may be due to the size of that particular gene family. We found these trends in our data when focusing on the 11 genes predicted to regulate the 6 to 7 DAI transition. Specifically, those genes that are

expressed early in the transition tend to decrease after the transition while those that are upregulated during the transition peak at that time point.

Methods

Experimental Design

Material for FACS was grown for 3 to 4 weeks in a growth chamber under 24 hour light. The temperature was maintained at 22 degrees. Once plants began to bolt, 20 micromolar dexamethasone in 0.015% (v/v) Silwet was sprayed onto developing inflorescence tissue via a spray bottle. 120 hrs later inflorescences were collected selecting from plants grown in at least 3 different locations within the growth chamber to limit effects due to environmental heterogeneity. Protoplasts were generated and sorted once per day for 5 consecutive days from synchronized material (according to the method of Villarino et al 2016. Protoplasts were sorted into three populations YFP+, mCherry+, and negative (non-fluorescent) protoplasts. The sorted protoplasts were collected in Trizol LS, then RNA was isolated immediately or samples were frozen in liquid nitrogen and kept at -80 degrees until RNA could be isolated. Experiments were carried out until 3 bio-replicates were generated for all time points making sure that samples within a time point came from the same bio-replicate. Of note libraries for negative samples were only generated 7 DAI (Figure Below). This allowed for differential gene expression analysis to be conducted between the medial domains, lateral domains, and other floral tissue, however, due to the use of Modified Shannon Entropy to determine DEGs we did not use the data from this time point.

Table 2.1 Sample Collection Strategy. Above are shown all samples that were sequenced. The first number represents biological replicate, the second represents time point and Y represents medial domain while 7 represents lateral domain cells.

5 DAI	6 DAI	7 DAI	8 DAI	9 DAI
9-1-Y, 9-1-7	7-2-Y, 7-2-7	9-3-Y, 9-3-7, 9-3-(-)	9-4-Y, 9-4-7	9-5-Y, 9-5-7
4-1-Y, 4-1-7	8-2-Y, 8-2-7	8-3-Y, 8-3-7, 8-3-(-)	8-4-Y, 8-4-7	7-5-Y, 7-5-7
7-1-Y, 7-1-7	9-2-Y, 9-2-7	7-3-Y, 7-3-7, 7-3-(-)	7-4-Y, 7-4-7	6-5-Y, 6-5-7
	5-2-Y, 5-2-7	6-3-Y, 6-3-7, 6-3-(-)	6-4-Y, 6-4-7	

Material

S. No. 2080 T2.33.3 (S. No. 2080) was used for this experiment. S. No. 2080 uses the AP1/Cal inducible system to synchronize flower development and increase the amount of tissue collected (Welmer et al. 2006). This system has been used in many studies and relies on the use of dexamethasone which binds to the glucocorticoid receptor (GR) to induce the transition from a vegetative meristem to a floral meristem. This allows for a rough synchronization of the flowers among all of the plants used. The *ap1/cal* mutant produces many inflorescence like meristems which allows for the collection of much more tissue than would be produced in the same number of wild type flowers. Also, this line was crossed into the background containing the dual reporter

construct with pSHP2-Gal4; pUAS-3xYpet and pFUL-mCHERRY. The pFUL-mCHERRY construct (alternatively named pVRR:Cherry) contains a 500 bp fragment from the *FUL* gene that was fused to the mCherry reporter in the vector pGreenII (Ripoll *et al.*, unpublished). This 500 bp fragment contains a gene regulatory element that conditions the expression of the mCherry reporter in the gynoecial valve domain.

Protoplast Recovery and Cell Sorting

Protoplasts were generated from S. No 2080 plants following the protocol in (Villarino et al 2016). The only adjustment was that we chose to induce flowering five days before our first time point of the sorting experiment. Also, we planted more material and prepared protoplasts from plants sprayed at the same time every 24 hours beginning five DAI.

RNA Extraction

RNA was isolated according to the protocol in (Villarino et al. 2016) using the Plant RNeasy Micro Kit (Qiagen) instead of the Plant RNeasy Mini Kit (Qiagen).

Sequencing

The sample quality was confirmed using the Agilent 2100 Bioanalyzer. We chose samples that had a RIN value of at least 7.0 and at least 10ng of RNA. Strand-specific cDNA libraries were constructed from at least 10ng of total RNA using an NEB Ultra II Directional Library Prep Kit for Illumina (New England Biolabs)(Supplemental Table 2.5).

Bioinformatics Analysis

In order to identify DEGs we chose to use an unbiased approach, which removed the need for fold change and p value cutoffs. Modified Shannon entropy (MSE) uses an entropy score to select a preliminary set of genes. In order to determine which specific set each gene is considered a DEG we employed outlier scoring. This allows for a combined approach to identify differentially expressed genes and to determine when and in which tissue they are differentially expressed.

Traditional Shannon Entropy was used to generate entropy scores for each gene based on expression in each of the samples using the method described by Shug et al. (Shug et al. 2005). Then these entropy values are used in entropy value thresholding to make a selection of differentially expressed genes. A description of the equations used to calculate entropy scores can be found in Kadota et al. 2006.

Next, to provide further resolution on the temporal and spatial specifics of each DEG we used an outlier analysis. Outlier and non-outlier scores were assigned to each of our preselected DEGs in all of our samples (Every time point in each tissue). If a sample for a particular gene receives an outlier score of 1 or -1 that means that when compared to all of the other samples the gene is over- or under-expressed; a score of 1 or -1 indicates differential expression of that gene in that sample. A score of 0 represents non-outliers and 'normal' expression and can be interpreted as a sample where the gene is not differentially expressed. To calculate outlier scores every combination of outlier and non-outlier scores are created. They are then evaluated by Equation 1 and the combination of scores that results in the lowest value of U is selected (Kadota et al. 2006).

$$U = n \log \sigma + \sqrt{2} \times s \times \frac{\log n!}{n} \text{ , (1)}$$

In the above equation s is the number of outlier candidates, n is the number of non-outlier candidates, and σ is the standard deviation of the expression values of the non-outlier candidates. Through outlier scoring we are able to identify samples where a specific gene is differentially expressed, this has allowed us to identify genes that are differentially expressed in specific tissues and at specific time points.

Determination of Gene Transition

MSE was applied to select genes with differential expression in at least one medial domain time point, resulting in a selection of 728 genes. Next, transitions between highly and lowly expressed genes in adjacent time points (5DAI to 6DAI, 6DAI to 7DAI, 7DAI to 8DAI, and 8DAI to 9DAI) were quantified and hierarchically clustered. A transition was defined as a gene changing from being differentially expressed (an outlier) to not differentially expressed (a non-outlier) in a subsequent time point, or vice versa. This corresponded to tracking significantly positive or negative changes in expression for a given gene.

Modeling Gene Regulatory Networks

MSE was applied to select genes with differential expression in at least one medial domain time point, resulting in a selection of 728 genes. Next, transitions between highly and lowly expressed genes in adjacent time points (5DAI to 6DAI, 6DAI to 7DAI, 7DAI to 8DAI, and 8DAI to 9DAI) were quantified and hierarchically clustered to pinpoint the largest transcriptional shift. 277 transitions (46%) occurred between 6DAI and 7DAI. The RTP-STAR section of TuxNet was

applied to infer a GRN using the list of 6DAI to 7DAI transitioning genes and the medial domain temporal expression dataset according to the protocol in Spurney et al. 2019.

REFERENCES

Alvarez, John, and David R. Smyth. 2002. "CRABS CLAW and SPATULA Genes Regulate Growth and Pattern Formation during Gynoecium Development in Arabidopsis Thaliana." *International Journal of Plant Sciences* 163 (1): 17–41.

Azhakanandam, Sridevi, Staci Nole-Wilson, Fang Bao, and Robert G. Franks. 2008. "SEUSS and AINTEGUMENTA Mediate Patterning and Ovule Initiation during Gynoecium Medial Domain Development." *Plant Physiology* 146 (3): 1165–81.

Balanzá, Vicente, Marisa Navarrete, Marina Trigueros, and Cristina Ferrándiz. 2006. "Patterning the Female Side of Arabidopsis: The Importance of Hormones." *Journal of Experimental Botany* 57 (13): 3457–69.

Benfey, P. N., and B. Scheres. 2000. "Root Development." Current Biology: CB 10 (22): R813–15.

Borrill, Philippa, Sophie A. Harrington, James Simmonds, and Cristobal Uauy. 2019. "Identification of Transcription Factors Regulating Senescence in Wheat through Gene Regulatory Network Modelling." *Plant Physiology* 180 (3): 1740–55.

Bowman, J. L., G. N. Drews, and E. M. Meyerowitz. 1991. "Expression of the Arabidopsis Floral Homeotic Gene AGAMOUS Is Restricted to Specific Cell Types Late in Flower Development." *The Plant Cell* 3 (8): 749–58.

Bowman, John L., Stuart F. Baum, Yuval Eshed, Joanna Putterill, and John Alvarez. 1999. "4 Molecular Genetics of Gynoecium Development in Arabidopsis." In *Current Topics in Developmental Biology*, edited by Roger A. Pedersen and Gerald P. Schatten, 45:155–205. Academic Press.

Byrne, M. E., R. Barley, M. Curtis, J. M. Arroyo, M. Dunham, A. Hudson, and R. A. Martienssen. 2000. "Asymmetric leaves1 Mediates Leaf Patterning and Stem Cell Function in Arabidopsis." *Nature* 408 (6815): 967–71.

Chen, Dijun, Wenhao Yan, Liang-Yu Fu, and Kerstin Kaufmann. 2018. "Architecture of Gene Regulatory Networks Controlling Flower Development in Arabidopsis Thaliana." *Nature Communications* 9 (1): 4534.

Clark, Natalie M., Eli Buckner, Adam P. Fisher, Emily C. Nelson, Thomas T. Nguyen, Abigail R. Simmons, Maria A. de Luis Balaguer, et al. 2019. "Stem-Cell-Ubiquitous Genes Spatiotemporally Coordinate Division through Regulation of Stem-Cell-Specific Gene Networks." Nature Communications 10 (1): 5574.

Clark, S. E. 2001. "Cell Signalling at the Shoot Meristem." *Nature Reviews. Molecular Cell Biology* 2 (4): 276–84.

Clark, S. E., M. P. Running, and E. M. Meyerowitz. 1993. "CLAVATA1, a Regulator of Meristem and Flower Development in Arabidopsis." *Development* 119 (2): 397–418.

Clark, Steven M., Mark P. Running, and Elliot M. Meyerowitz. 1995. "CLAVATA3 Is a Specific Regulator of Shoot and Floral Meristem Development Affecting the Same Processes as CLAVATA1." *Development* 121 (7): 2057–67.

Cruz-Ramírez, Alfredo, Sara Díaz-Triviño, Ikram Blilou, Verônica A. Grieneisen, Rosangela Sozzani, Christos Zamioudis, Pál Miskolczi, et al. 2012. "A Bistable Circuit Involving SCARECROW-RETINOBLASTOMA Integrates Cues to Inform Asymmetric Stem Cell Division." *Cell* 150 (5): 1002–15.

Das, Pradeep, Toshiro Ito, Frank Wellmer, Teva Vernoux, Annick Dedieu, Jan Traas, and Elliot M. Meyerowitz. 2009. "Floral Stem Cell Termination Involves the Direct Regulation of AGAMOUS by PERIANTHIA." *Development* 136 (10): 1605–11.

de Luis Balaguer, Maria Angels de, Adam P. Fisher, Natalie M. Clark, Maria Guadalupe Fernandez-Espinosa, Barbara K. Möller, Dolf Weijers, Jan U. Lohmann, Cranos Williams, Oscar Lorenzo, and Rosangela Sozzani. 2017. "Predicting Gene Regulatory Networks by Combining Spatial and Temporal Gene Expression Data in Arabidopsis Root Stem Cells." *Proceedings of the National Academy of Sciences of the United States of America* 114 (36): E7632–40.

Franks, Robert G., Chunxin Wang, Joshua Z. Levin, and Zhongchi Liu. 2002. "SEUSS, a Member of a Novel Family of Plant Regulatory Proteins, Represses Floral Homeotic Gene Expression with LEUNIG." *Development* 129 (1): 253–63.

Gatz, Christiane. 2013. "From Pioneers to Team Players: TGA Transcription Factors Provide a Molecular Link between Different Stress Pathways." *Molecular Plant-Microbe Interactions: MPMI* 26 (2): 151–59.

Gilbert, Scott F. 2016. *Developmental Biology*. Sinauer.

Graham, L. E., M. E. Cook, and J. S. Busse. 2000. "The Origin of Plants: Body Plan Changes Contributing to a Major Evolutionary Radiation." Proceedings of the National Academy of Sciences of the United States of America 97 (9): 4535–40.

Gu, Q., C. Ferrándiz, M. F. Yanofsky, and R. Martienssen. 1998. "The FRUITFUL MADS-Box Gene Mediates Cell Differentiation during Arabidopsis Fruit Development." *Development* 125 (8): 1509–17.

Gutsche, Nora, and Sabine Zachgo. 2016. "The N-Terminus of the Floral Arabidopsis TGA Transcription Factor PERIANTHIA Mediates Redox-Sensitive DNA-Binding." *PloS One* 11 (4): e0153810.

Jacobs, W. P., and I. B. Morrow. 1961. "A Quantitative Study of Mitotic Figures in Relation to Development in the Apical Meristem of Vegetative Shoots of Coleus." *Developmental Biology* 3 (October): 569–87.

Kadota, Koji, Jiazhen Ye, Yuji Nakai, Tohru Terada, and Kentaro Shimizu. 2006. "ROKU: A Novel Method for Identification of Tissue-Specific Genes." *BMC Bioinformatics* 7 (June): 294.

Kayes, J. M., and S. E. Clark. 1998. "CLAVATA2, a Regulator of Meristem and Organ Development in Arabidopsis." *Development* 125 (19): 3843–51.

Keller, G. M. 1995. "In Vitro Differentiation of Embryonic Stem Cells." Current Opinion in Cell Biology 7 (6): 862–69.

Klucher, K. M., H. Chow, L. Reiser, and R. L. Fischer. 1996. "The AINTEGUMENTA Gene of Arabidopsis Required for Ovule and Female Gametophyte Development Is Related to the Floral Homeotic Gene APETALA2." *The Plant Cell* 8 (2): 137–53.

Laux, T., K. F. Mayer, J. Berger, and G. Jürgens. 1996. "The WUSCHEL Gene Is Required for Shoot and Floral Meristem Integrity in Arabidopsis." *Development* 122 (1): 87–96.

Li, Shutian, Andrea Lauri, Mark Ziemann, Andrea Busch, Mrinal Bhave, and Sabine Zachgo. 2009. "Nuclear Activity of ROXY1, a Glutaredoxin Interacting with TGA Factors, Is Required for Petal Development in Arabidopsis Thaliana." *The Plant Cell* 21 (2): 429–41.

Liebers, Monique, Björn Grübler, Fabien Chevalier, Silva Lerbs-Mache, Livia Merendino, Robert Blanvillain, and Thomas Pfannschmidt. 2017. "Regulatory Shifts in Plastid Transcription Play a Key Role in Morphological Conversions of Plastids during Plant Development." *Frontiers in Plant Science* 8 (January): 23.

Liu, Chang, Zhonghui Thong, and Hao Yu. 2009. "Coming into Bloom: The Specification of Floral Meristems." *Development* 136 (20): 3379–91.

Liu, Z., and E. M. Meyerowitz. 1995. "LEUNIG Regulates AGAMOUS Expression in Arabidopsis Flowers." *Development* 121 (4): 975–91.

Liu, Z., R. G. Franks, and V. P. Klink. 2000. "Regulation of Gynoecium Marginal Tissue Formation by LEUNIG and AINTEGUMENTA." *The Plant Cell* 12 (10): 1879–92.

Lohmann, J. U., R. L. Hong, M. Hobe, M. A. Busch, F. Parcy, R. Simon, and D. Weigel. 2001. "A Molecular Link between Stem Cell Regulation and Floral Patterning in Arabidopsis." *Cell* 105 (6): 793–803.

Long, J. A., E. I. Moan, J. I. Medford, and M. K. Barton. 1996. "A Member of the KNOTTED Class of Homeodomain Proteins Encoded by the STM Gene of Arabidopsis." *Nature* 379 (6560): 66–69.

Lu, C. C., J. Brennan, and E. J. Robertson. 2001. "From Fertilization to Gastrulation: Axis Formation in the Mouse Embryo." *Current Opinion in Genetics & Development* 11 (4): 384–92.

Ly, S. 2011. "Embryonic Differentiation in Animals". *Embryo Project Encyclopedia*

Melzer, Siegbert, Frederic Lens, Jerôme Gennen, Steffen Vanneste, Antje Rohde, and Tom Beeckman. 2008. "Flowering-Time Genes Modulate Meristem Determinacy and Growth Form in Arabidopsis Thaliana." *Nature Genetics* 40 (12): 1489–92.

Moore, Richard C., Sarah R. Grant, and Michael D. Purugganan. 2005. "Molecular Population Genetics of Redundant Floral-Regulatory Genes in Arabidopsis Thaliana." *Molecular Biology and Evolution* 22 (1): 91–103.

Murmu, Jhadeswar, Michael J. Bush, Catherine DeLong, Shutian Li, Mingli Xu, Madiha Khan, Caroline Malcolmson, Pierre R. Fobert, Sabine Zachgo, and Shelley R. Hepworth. 2010. "Arabidopsis Basic Leucine-Zipper Transcription Factors TGA9 and TGA10 Interact with Floral Glutaredoxins ROXY1 and ROXY2 and Are Redundantly Required for Anther Development." *Plant Physiology* 154 (3): 1492–1504.

Nahar, Most Altaf-Un, Tetsuya Ishida, David R. Smyth, Masao Tasaka, and Mitsuhiro Aida. 2012. "Interactions of CUP-SHAPED COTYLEDON and SPATULA Genes Control Carpel Margin Development in Arabidopsis Thaliana." *Plant & Cell Physiology* 53 (6): 1134–43.

Nole-Wilson, Staci, Elizabeth E. Rueschhoff, Huda Bhatti, and Robert G. Franks. 2010. "Synergistic Disruptions in Seuss cyp85A2 Double Mutants Reveal a Role for Brassinolide Synthesis during Gynoecium and Ovule Development." *BMC Plant Biology* 10 (September): 198.

Ó'Maoiléidigh, Diarmuid S., Bennett Thomson, Andrea Raganelli, Samuel E. Wuest, Patrick T. Ryan, Kamila Kwaśniewska, Cristel C. Carles, Emmanuelle Graciet, and Frank Wellmer. 2015. "Gene Network Analysis of Arabidopsis Thaliana Flower Development through Dynamic Gene Perturbations." *The Plant Journal: For Cell and Molecular Biology* 83 (2): 344–58.

Palumbo, Maria Concetta, Sara Zenoni, Marianna Fasoli, Mélanie Massonnet, Lorenzo Farina, Filippo Castiglione, Mario Pezzotti, and Paola Paci. 2014. "Integrated Network Analysis Identifies Fight-Club Nodes as a Class of Hubs Encompassing Key Putative Switch Genes That Induce Major Transcriptome Reprogramming during Grapevine Development." *The Plant Cell* 26 (12): 4617–35.

Reyes-Olalde, J. Irepan, Victor M. Zuñiga-Mayo, Ricardo A. Chávez Montes, Nayelli Marsch-Martínez, and Stefan de Folter. 2013. "Inside the Gynoecium: At the Carpel Margin." *Trends in Plant Science* 18 (11): 644–55.

Rivera-Pérez, Jaime A., and Anna-Katerina Hadjantonakis. 2014. "The Dynamics of Morphogenesis in the Early Mouse Embryo." *Cold Spring Harbor Perspectives in Biology* 7 (11).

Running, M. P., and E. M. Meyerowitz. 1996. "Mutations in the PERIANTHIA Gene of Arabidopsis Specifically Alter Floral Organ Number and Initiation Pattern." *Development* 122 (4): 1261–69.

Schug, Jonathan, Winfried-Paul Schuller, Claudia Kappen, J. Michael Salbaum, Maja Bucan, and Christian J. Stoeckert Jr. 2005. "Promoter Features Related to Tissue Specificity as Measured by Shannon Entropy." *Genome Biology* 6 (4): R33.

Spurney, Ryan J., Lisa Van den Broeck, Natalie M. Clark, Adam P. Fisher, Maria A. de Luis Balaguer, and Rosangela Sozzani. 2020. "Tuxnet: A Simple Interface to Process RNA Sequencing Data and Infer Gene Regulatory Networks." *The Plant Journal: For Cell and Molecular Biology* 101 (3): 716–30.

Surani, M. Azim, Katsuhiko Hayashi, and Petra Hajkova. 2007. "Genetic and Epigenetic Regulators of Pluripotency." *Cell* 128 (4): 747–62.

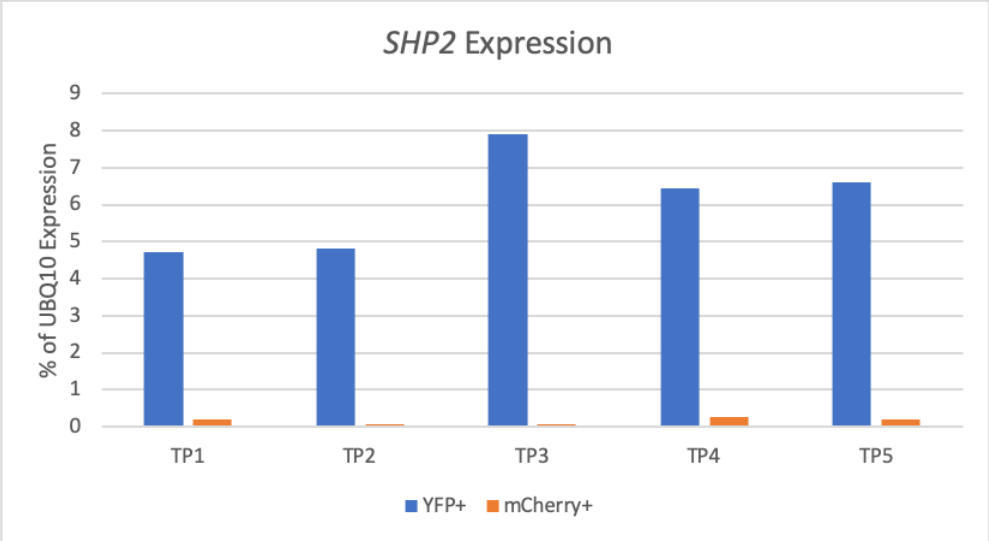
Villarino, Gonzalo H., Qiwen Hu, Silvia Manrique, Miguel Flores-Vergara, Bhupinder Sehra, Linda Robles, Javier Brumos, et al. 2016. "Transcriptomic Signature of the SHATTERPROOF2 Expression Domain Reveals the Meristematic Nature of Arabidopsis Gynoecial Medial Domain." *Plant Physiology* 171 (1): 42–61.

Wellmer, Frank, Márcio Alves-Ferreira, Annick Dubois, José Luis Riechmann, and Elliot M. Meyerowitz. 2006. "Genome-Wide Analysis of Gene Expression during Early Arabidopsis Flower Development." *PLoS Genetics* 2 (7): e117.

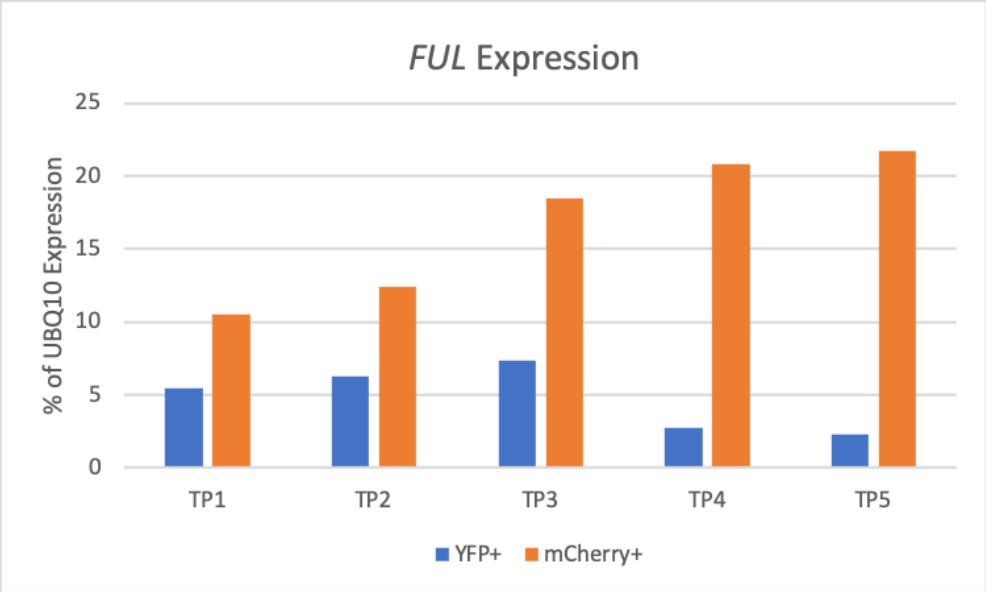
Wynn, April N., Andrew A. Seaman, Ashley L. Jones, and Robert G. Franks. 2014. "Novel Functional Roles for PERIANTHIA and SEUSS during Floral Organ Identity Specification, Floral Meristem Termination, and Gynoecial Development." *Frontiers in Plant Science* 5 (April): 130.

Zambelli, Federico, Francesca Mastropasqua, Ernesto Picardi, Anna Maria D'Erchia, Graziano Pesole, and Giulio Pavesi. 2018. "RNentropy: An Entropy-Based Tool for the Detection of Significant Variation of Gene Expression across Multiple RNA-Seq Experiments." *Nucleic Acids Research* 46 (8): e46.

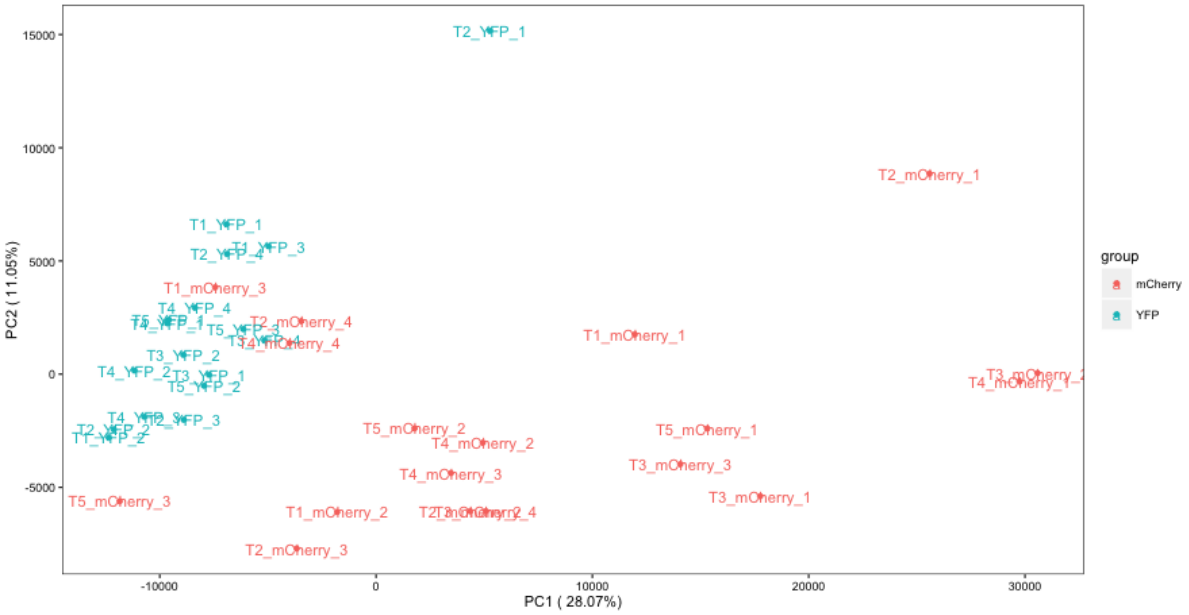
Supplemental Figures



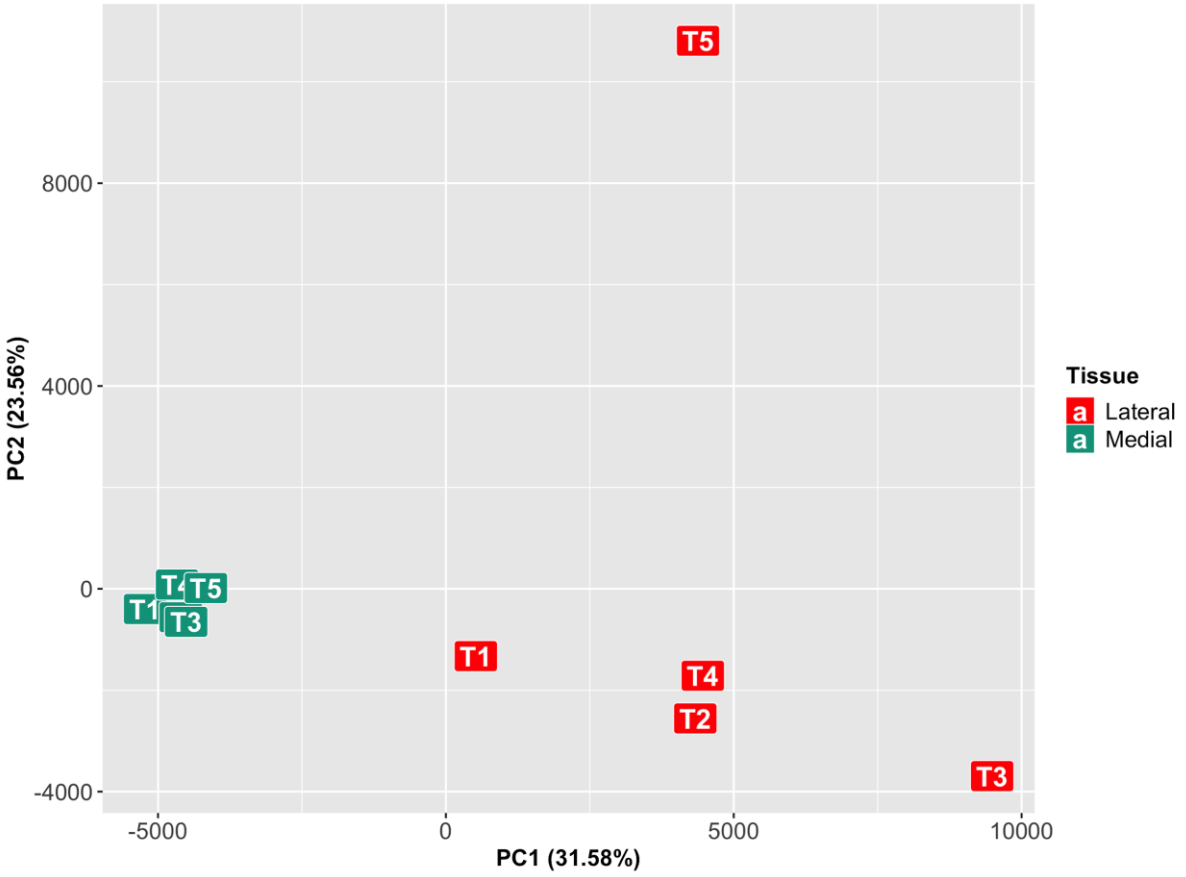
Supplemental Figure 2.1 *SHP2* Expression. Averaged FPKM values normalized as a percentage of *UBQ10* expression over all time points in each domain show that *SHP2* expression is enriched in the medial domain when compared to the lateral domain.



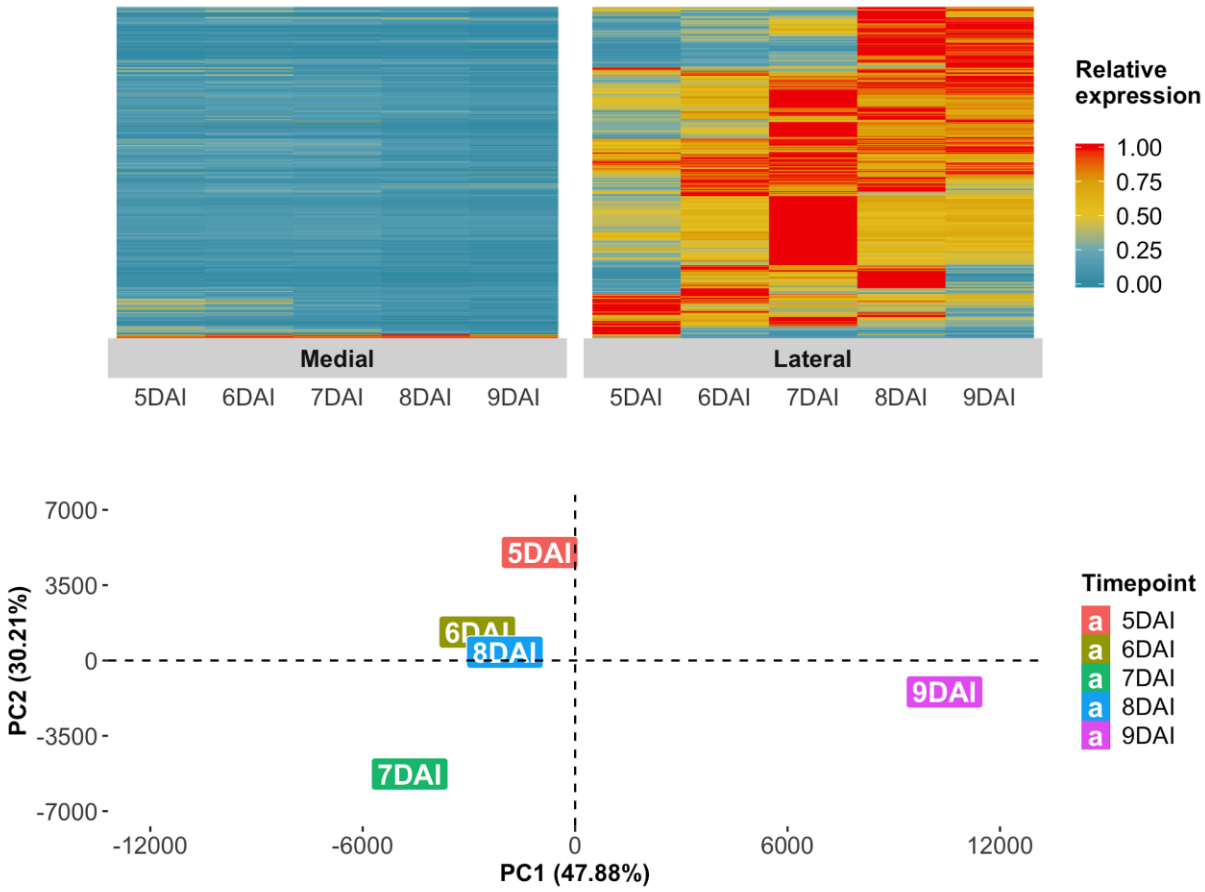
Supplemental Figure 2.2 *FUL* Expression. Averaged FPKM values normalized as a percentage of *UBQ10* expression at all time points in each tissue show that *FUL* expression is enriched in the lateral domain of the gynoecium.



Supplemental Figure 2.3 PCA of All Samples. PCA of samples from all time points in both domains with separate biological replicates with data from all expressed genes. Medial domain samples cluster tightly together while lateral domain samples show more variability.



Supplemental Figure 2.4 PCA of Averaged Bio-Replicates of All Time Points in Both Domains. Analyzing differentially expressed genes from both domains at all five time points with the bioreplicates averaged shows that the two domains cluster (except for T5 in the lateral domain) and are separated along PC1 which accounts for 31.58% of the variance among samples. This shows that cells isolated from the same tissue are more similar to one another than those isolated from different tissues at the same time.



Supplemental Figure 2.5 Differentially Expressed Genes in the Lateral Domain. **A** 827 genes were identified as differentially expressed in the lateral domain using MSE. These genes form groups based on the time of peak expression. **B.** Using the 827 genes identified as DEGs in the lateral domain we used a PCA to investigate the separation of the transcriptome of each time point, showing that TP5-8 separated over time along PC2 with TP9 separated along PC1.

Supplemental Tables

Supplemental Table 2.1 Differentially Expressed Genes in the Medial Domain. List of genes called differentially expressed through MSE in the gynoecial medial domain.

https://docs.google.com/spreadsheets/d/19MYr8SHlzy0EAdMjEpn_ODb1HbQ1Mpx8BypEWhVvxr8/edit?usp=sharing

Supplemental Table 2.2 Differentially Expressed Genes in the Lateral Domain. List of genes called differentially expressed through MSE in the lateral domain of the gynoecium.

https://docs.google.com/spreadsheets/d/1AzKNjmeLJdR61fYN2nq2pJLhBXQYu8wu6DxLFrqOaFI/edit?usp=sharing

Supplemental Table 2.3 Transitioning Genes. Genes undergoing a transition between 6 and 7 DAI in the medial domain.

https://docs.google.com/spreadsheets/d/1HaDE1zUqPW4xAfOoHrDPCeOhJjx2Z6JZP4y4NX4APtQ/edit?usp=sharing

Supplemental Table 2.4 Expression Values for Key Nodes in the 6-7DAI Network.

Gene ID	Gene Name	Medial T1	Medial T2	Medial T3	Medial T4	Medial T5
AT1G08320	TGA9	7.1358967	9.7884175	13.6694033	8.8707263	6.8786375
AT5G44210	ERF9	8.4823433	6.6012600	2.4400583	2.4648300	2.0144375
AT1G02230	NAC004	9.6987112	5.2345838	0.9528503	1.8606121	1.2534298
AT2G28610	PRS	8.8404917	6.6966588	2.0008153	0.8662384	0.9576518
AT5G42630	ATS	0.3360777	0.2107811	0.9621610	0.5099960	0.3625535
AT1G63030	ddf2	8.3008422	8.7326063	2.8942333	3.0445488	1.1944383
AT1G71520	AT1G71520	24.5513500	17.5865750	0.4863983	3.0798831	0.2566553
AT3G11090	LBD21	3.0693562	2.8625038	6.4291933	3.1126663	2.4007350
AT3G04070	NAC047	1.2446925	1.5741690	3.0110453	1.1710946	1.7513718
AT3G46070	AT3G46070	5.9025667	3.3287750	1.2501013	1.5395383	0.2824140
AT2G22760	AT2G22760	2.3488450	2.9360485	0.8126058	0.6475845	0.3726240

Supplemental Table 2.5 GRN Motifs. Characteristics of the gene regulatory network regulating the transition between 6DAI and 7DAI. When scoring networks motifs, those motifs that are expressed more in the network than in a randomly generated set of networks are used to calculate a score. When generating a score, motifs are normalized against the most prevalent motif. These motifs include Feed-forward loops (FF), where one gene regulates a target, and those two genes both regulate a third. Bifan motifs (BF) where two genes cross regulate two other genes, our networks had two other motifs that were overrepresented. Biparallel motifs occur when a gene regulates a target through two different intermediates (BP) and a four cascade motif (4C) is a straight line of regulations in four nodes. Outdegree is not a motif but denotes the number of out regulations from a single node.

Gene	FF Norm	BF Norm	BP Norm	4C Norm	Outdegree Norm	Score
TGA9	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**5.00**
LBD21	**0.78**	**0.94**	**0.90**	**0.94**	**0.87**	**4.42**
ATS	**0.29**	**0.06**	**0.58**	**0.81**	**0.18**	**1.92**
NAC047	**0.24**	**0.05**	**0.51**	**0.41**	**0.18**	**1.48**
NAC004	**0.16**	**0.06**	**0.07**	**0.9**	**0.29**	**1.01**

Supplemental Table 2.6 Sequencing Results. Number of reads and the number of mapped reads for all samples.

https://docs.google.com/spreadsheets/d/18VMTriJfdgTROpC8_UMMFlpTyR1EnqRUgR5Aq5jWDtA/edit?usp=sharing

CHAPTER 3

CRISPR/Cas9 as an Alternative to T-DNA Insertion Mutants in Reverse Genetics Screens

Introduction

Reverse Genetics: Identifying gene function

Genetic screens have long been used to identify relationships between phenotype and genotype. With the advancement of high throughput sequencing technologies reverse genetic screens have never been more relevant. Traditionally forward genetic screens were conducted after a specific phenotype was discovered. To determine the genetic basis of that phenotype of interest, a screen was conducted. These forward genetic screens begin by mutagenizing organisms with well defined genetic backgrounds. In plants, this has traditionally been the model organism *Arabidopsis thaliana* although it has also been pursued in maize (Page and Grossniklaus 2002, Candela and Hake 2008). Plants were treated with a mutagenizing agent such as ethyl methanesulfonate (EMS) which causes mutations through nucleotide substitution often through guanine alkylation or radiation. The offspring are then screened for interesting phenotypes.

Following the discovery of an interesting phenotype, a search for the responsible mutation is conducted. This has occurred in several different ways depending on the mutagenesis method and when the screen was conducted. In the past this process was conducted in several steps; genetic mapping to find the broad region responsible for the phenotype of interest and then a targeted search for candidate mutations within that region (Hazen et al. 2005). This latter stage includes genotyping in mutant and wild type backgrounds at putative causative loci and functional analyses such as complementation with wild type alleles.

With the advent of next generation sequencing, the first step of genetic mapping has become much more streamlined. Instead of using bulk segregant analysis (BSA) to map location of the mutation based on the over-representation of linked alleles in phenotypically mutant marker lines, next generation sequencing allows for mapping by sequencing approaches (Schneeeberger et al. 2009). These approaches typically skip the generation of mapping populations and rely on direct sequencing of individual mutant genomes. This approach relies on whole genome sequencing (WGS) and can be effective in identifying potential causal mutations (Schneeberger et al. 2009, Lister et al. 2009). One problem with this approach is that there are often many different genes mutated in the plants with the phenotype of interest, especially if a chemical mutagenesis approach was employed. There are other modern methods of mapping by sequencing which address this problem through the use of a crossing strategy or other specific genetic background to generate a population where the mutations are segregating or can be mapped by the presence of homozygosity (Scheeberger et al. 2009, Cuperus et al. 2010, Manavela et al. 2012). Also, traits are often controlled by multiple loci so a different approach must be taken to identify the function of a single gene of interest.

In contrast to forward genetic screens that identify a phenotype and search for the causative loci, reverse genetic screens search for the function of a gene by mutagenzing the gene and then looking for a phenotype. The rise of these techniques can be attributed in part to the development of high throughput sequencing techniques and the popularity of RNA sequencing approaches to studying different biological processes. These experiments and earlier microarray experiments use differential expression analyses between different developmental stages, organs, genetic backgrounds, or other experimental conditions to identify genes of interest which are

often considered to be key components or regulators of a given process. This is where reverse genetic screens are extremely useful as they allow researchers to discover the role or function of specific differentially-expressed genes in the development of the plant.

This is critical because even with the advancement of sequencing technology only about 12% of Arabidopsis genes have undergone *in vivo* characterization (O'Malley et al. 2015). Up to 60% of genes have an inferred function, but this is mostly based on shared sequence similarity with a functionally-characterized gene and does not present an accurate enough picture to assign a gene a role in a developmental process where spatial and temporal specificity are important.

Reverse genetic screens have benefitted from several advances in the past two decades. Specifically, the sequencing of many plant genomes and the development of libraries of *Arabidopsis thaliana* lines that contain mutations in specific genes. The sequencing of plant genomes has allowed for the annotation of whole genomes and comparisons between genomes in different plant species. This has led to the identification of domains with different structures and even predicted functions of genes, but it has also led to a challenging question: how do we figure out the function of this many genes? This is where reverse genetics has gained popularity. Gene silencing technologies, transgene-induced ectopic expression, gene targeting, T-DNA/transposons insertional mutagenesis, and target-induced local mutations are all techniques that have been used to investigate the function of specific genetic sequences. Two main stock centers house a library of T-DNA insertional mutagenesis lines which are commonly used in the analysis of genes of interest. The Arabidopsis Biological Research Center (ABRC) in the United

States currently contains 688,136 seed stocks for mutant lines in *Arabidopsis* which makes it an excellent source for material to validate gene function.

Potential Drawbacks of T-DNA Insertion Mutants

Despite a large number of different transgenic lines available from stock centers, it is not guaranteed that there will be a mutant available in your gene of interest that shows a mutant phenotype. One flaw with this approach is the lack of availability of mutants in all genes, 12% of genes have no mutant allele and 8% of genes have only a single allele (O'Malley et al. 2010). When you consider only alleles with the insertion in an exon, the percentage of genes with a mutant line in the stock center drops drastically. The percentage decreases even more if you limit potential lines to only insertions in the first 3 exons. While there are mutations in introns that correlate to a phenotype, one approach to generating lines with a high probability relies on the presence of a mutation in an exon which leads to a point-nonsense mutation and a truncated non-functional protein.

One downside to this approach, especially when investigating the function of important regulators of floral development, is redundancy. There are several families of transcription factors in Arabidopsis that control floral development and have high levels of sequence homology and are redundant in function. One of these is the REM family of transcription factors which play an important role in early flower development but have been challenging to study due to their high number and redundancy (Mantegazza et al. 2014).

To generate a line with an observable phenotype it is often necessary to investigate plants with non-functional alleles of several key genes, these so-called higher order mutants negate functional redundancy of several genes by mutating all functionally-redundant paralogues. The creation of higher order mutants is aided by *in silico* analyses that detect similar expression patterns and high levels of sequence similarity between genes. By crossing parents containing two different tDNA insertion lines, it is possible to generate a higher order mutant line where both redundant genes contain insertions that disrupt the gene product. In some cases it is necessary to make mutations in at least three genes to see a phenotype, this is the case in the REM family of genes with respect to early flower development, it should be noted that this is not always the case and a single mutant in the REM gene VERDANDI (*VDD*) affects embryo sac differentiation (Matías-Hernández et al. 2010). Here arises potential pitfalls of this technique; the availability of mutant lines, the inefficiencies in creating double mutants in genes that are located close to one another, and the amount of time that it takes to generate a higher order mutant through traditional crossing experiments.

While crossing existing mutant lines to generate a higher order mutant is a common method of overcoming genetic redundancy, it is not an efficient option when genes are located close to one another on the same chromosome. Due to the random segregation of homologous chromosomes, when crossing plants with mutations on different chromosomes it is possible to generate plants with mutant alleles at both loci. These plants will all be heterozygous for each allele and must be self crossed to create homozygous lines. When genes are on the same chromosome the only way to generate a double mutant occurs through meiosis-based crossing over. As the distance between the two loci decreases so does the chance of a crossing over event that could produce a

double mutant. This is referred to as genetic linkage and the tighter the genes are linked the lower the recombination frequency and the lower the odds are of a cross leading to the desired alleles at each locus. With 27,000 genes in the *Arabidopsis thaliana* genome and only 5 chromosomes, this creates a significant challenge when creating higher order mutant plants. As I have already discussed common sources of mutant alleles and their drawbacks, as well as challenges in creating a double mutant with genes located close to one another, I will next discuss the time requirements for a crossing experiment to generate a hextuple mutant.

Time and Effort Required to Generate Higher Order Mutants in *Arabidopsis thaliana*

The seed-to-seed time in Arabidopsis ranges with the average time to set seed between 7 and 12 weeks after plants are germinated, this is the major limiting factor on the development of higher order mutants through the crossing of existing transgenic lines. When working with transgenic seeds from a stock center the first step is to select lines that are homozygous for the mutant alleles. O'Malley et al. describe an excellent method for confirming homozygous lines which uses young plants and requires only DNA isolation and two PCR reactions. This procedure does not require additional time and after confirming that a plant is homozygous it can be crossed to another confirmed homozygous line. The offspring must then be crossed again to create a quadruple mutant and again one final time to create a hextuple mutant. The hextuple mutant will need to be homozygosed before phenotyping can occur. In the interest of speed, it is best to cross heterozygous plants to avoid an extra generation of homozygosing in the double mutant and quadruple mutant generations. However, to best investigate the interactions among these genes and their role in the plant, homozygosing, and phenotyping these intermediate mutants is worth the additional time. In the process of generating a hextuple mutant at least three different

double mutants and one quadruple mutant will be generated. It would be advantageous to generate the other 12 double mutants and quadruple mutants to best investigate the genetic interactions of the different genes. An alternative method to generate higher order mutants in Arabidopsis utilizes the CRISPR/Cas9 system.

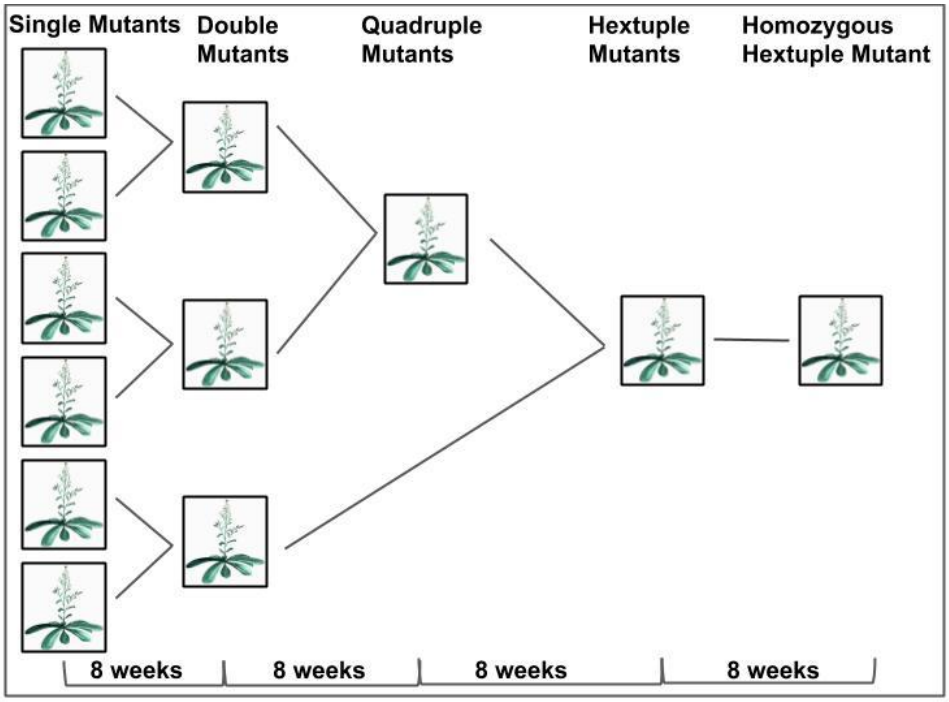


Figure 3.1 Timing of a Crossing Strategy to Generate Hextuple Mutants. Crossing Strategy to create a hextuple mutant using existing mutant lines from a stock center.

CRISPR as a Bacterial Pathogen Defense Mechanism

Clustered regularly interspaced short palindromic repeats (CRISPR) were discovered in the 1980s in the genome of *E. coli* and were named in 2002. CRISPR consists of a series of clustered repeats separated by spacer sequences and they are found in the genomes of most bacteria and archaea (Barrangou et al. 2007). In these systems, CRISPR and CRISPR-associated proteins (Cas) were shown to degrade exogenous genes delivered from a bacteriophage or a

plasmid (Ishino et al. 1987). This is possible because the CRISPR spacer sequence shows high homology to phage and plasmid DNA sequences, which is a result of the uptake of foreign DNA sequences by the host genome (Bolotin et al. 2005)(Mojica et al. 2005)(Pourcel et al. 2005). CRISPR consists of three stages; an adaptation phase, an expression phase, and an interference stage. In the adaptation phase, an exogenous sequence is digested by endonucleases and integrated into the CRISPR locus which creates a record of the attempted infection (Barrangou et al. 2007). The expression phase involves transcribing CRISPR RNAs (crRNAs) from the entire CRISPR locus. In the interference stage, mature crRNAs are loaded into effector complexes with Cas enzymes and then targeted to exogenic DNA where Cas cleaves the foreign DNA (Terns and Terns 2011). There are three major types of CRISPR/Cas systems based on gene signatures and organization of the different loci (Makarova and Coonin 2018). These systems can be broken down into many sub-types and some believe there are 5 major types. Of these type II CRISPR/Cas systems are most commonly used as a gene editing tool. They are characterized by a single protein, Cas9 which contains multiple domains responsible for the cleavage of foreign DNA and a co-processed dual-tracrRNA:crRNA molecule (Karvelis et al 2013).

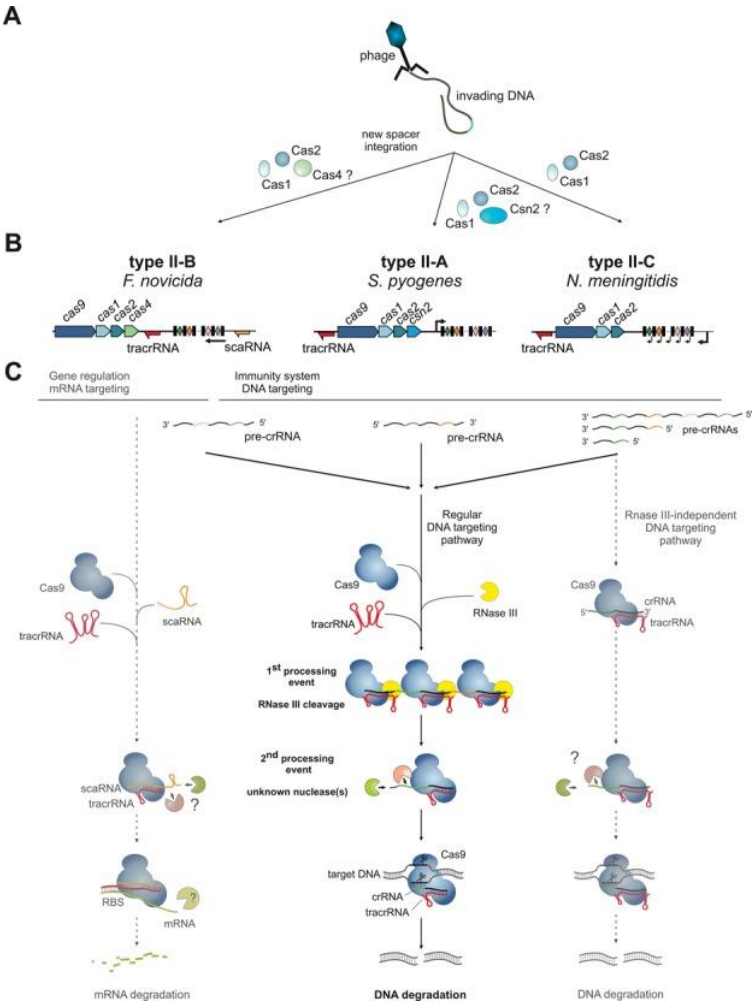


Figure 3.2 A Schematic of Type II CRISPR/Cas Systems. From Chylinski et al. 2014

In type II CRISPR/Cas systems two different RNA molecules are required (figure 3.2). The first is a crRNA formed through the transcription of the CRISPR array which includes both repeats and spacers. The product of this transcription is called a pre-crRNA. In the maturation process, the pre-crRNA is cleaved by RNase III (Deltcheva et al. 2011). This cleavage requires the presence of trans activating CRISPR RNA (tracrRNA) which is transcribed from a different locus and shares some homology with the (Chylinski et al 2014) pre-crRNA (Deltcheva et al. 2011). These two RNA molecules hybridize with a tracrRNA hybridizing to each repeat. A

cleavage then occurs in the spacer regions separating repeats which releases these hybrid tracrRNA-crRNA molecules which result in the formation of three different complexes being created from a single pre-crRNA molecule. After crRNA processing, the hybrid RNA molecules are processed again to modify the spacer portion of the RNA duplex (Deltcheva et al. 2011). The final product of crRNA maturation is a mature crRNA tracrRNA hybrid molecule that joins a complex with Cas9 where it directs the complexes targeting foreign DNA where that DNA is then degraded by Cas9.

Adaptations of CRISPR

In 2012 a group led by Emmanuelle Charpentier and Jennifer Doudna targeted and cleaved DNA *in vitro* using a modified CRISPR/Cas9 system (Jinek et al. 2012). In this system, the components had been limited to the Cas9 enzyme and a single guide RNA (gRNA). A guide RNA fuses the tracrRNA and crRNA into a single molecule. In their initial study, a gRNA was designed that was capable of targeting Cas9 to the protospacers in the green fluorescent protein (GFP) coding sequence. Found adjacent to these regions are protospacer adjacent motifs (PAMs), which are sequences of between two and six base pairs found three to four nucleotides downstream of the site that Cas9 will cleave the DNA. Cas9 was capable of cleaving the double stranded DNA in this portion of the GFP coding sequence.

This technology is capable of disrupting eukaryotic genes through two methods. Non-homologous end joining (NHEJ) takes advantage of the double-stranded breaks (DSBs) produced by Cas9 activity. A mutation is created when DNA repair mechanisms perform improperly adding additional nucleotides to create an insertion or leaving out nucleotides

creating deletions, together these are referred to as indels and they often result in frameshift mutations which can lead to premature stop codons and truncated gene products. The other main approach to create mutations using CRISPR relies on homology directed repair (HDR) and inserts a sequence into the genome by providing a donor sequence with homology to the areas surrounding the DSB which allows for a recombination event which inserts a sequence of interest into the target genome (Zhang et al. 2017). This approach has many variants in both the size of the homologous sequences in the HDR donor and the type of donor; single-stranded DNA templates (ssDNAs) or double-stranded DNA plasmid templates (Quadros et al. 2017, Richardson et al. 2016). Each of these has its benefits and experiments are ongoing to increase the editing efficiency of this technique.

Considerations When Using CRISPR to Generate Mutations

gRNA Structure and function

gRNAs are critical to CRISPR efficacy as these RNAs facilitate DNA targeting regardless of the downstream application. gRNA structure depends on several factors; the specific Cas9 enzyme used and the strategy to increase efficiency. There are many Cas9 and other enzymes in use with distinct functions and benefits of which the most common is *Streptococcus pyogenes* Cas9 (SpCas9)(Miller et al. 2020). Traditionally, gRNAs in use with SpCas9 target sequences between 19 and 20 base pairs and function in conjunction with a 3' PAM sequence of NGG (Wang et al. 2014, Peterson et al. 2016). The gRNA anneals to the antisense strand of the target sequence and guides Cas9 to cut at the -3 position relative to the PAM sequence (Liu et al. 2016). Therefore when designing gRNAs it is important to make sure that a PAM sequence is immediately downstream of the target sequence on the opposite strand (Anders et al. 2014). In

addition to making sure that an appropriate PAM sequence is found directly downstream of the target sequence, it is important to make sure that the target sequence is the appropriate length.

As stated earlier, traditionally the spacer portion of gRNAs are 19-20 nt long, however, this is not the only option. Truncated gRNAs (tru-gRNAs) are between 17 and 18 nucleotides long and increase the specificity of targeting without losses of efficiency (Liu et al. 2016). When determining the target sequence and the corresponding gRNA it is important to note the GC content of the putative target site which can affect both the efficiency and specificity of editing events. It has been shown in mammal systems that a GC content between 40-60 percent is favorable for optimal efficiency and specificity, and this has been recently confirmed in Aspen (*Populus tremula*) and Grape (*Vitis vinifera*) (Bruegmann et al. 2019, Liu et al. 2016, Ren et al. 2019). It has also been shown that it is possible to target several different sequences within the human genome through the construction of a gRNA with multiple spacer sequences (Cong et al. 2013) (Mali et al. 2013). In *Arabidopsis thaliana*, Zachary Nimchuk's group targeted 15 distinct loci and were successful in editing 13 of those (Peterson et al. 2016)

Cas Nuclease Options

There are several options available when choosing a Cas9 enzyme. The enzyme from *Streptococcus pyogenes* has traditionally been used as it was the first one to be characterized and recognizes a PAM sequence of NGG which is more plentiful than those recognized by Cas9 enzymes from other species. Also, SpCas9 recognizes a weaker PAM sequence of "NAG" (Jinek et al 2012, Jiang et al. 2013). Modified versions include a plant optimized enzyme, plant codon optimized Cas9 (pcoCas9) with a higher GC content, dual nuclear localization signals, and

FLAG-tag which increase editing efficiency (Jiang et al. 2013, Alonso and Stepanova 2015). Despite studies which have shown extreme editing specificity of SpCas9 in plants there have been attempts to increase this specificity through enzyme optimization which has produced high fidelity Cas9 SpCas9-HF1, enhanced specificity (eSpCas9(1.1), and hyper-accurate Cas9 (HypaCas9) (Kleinstiver et al. 2016, Slaymaker et al. 2016, Chen et al. 2017).

In some plant species, these enzymes are redundant as SpCas9 is already remarkably accurate, however, one approach to increase specificity that has other applications is the paired nickase approach. This approach uses SpCas9 enzymes with mutations in the RuvC or HNH catalytic sites which results in enzymes that exclusively cut the plus or minus DNA strand (Gasiunas et al. 2012). By using paired nickases in a system where each enzyme targets the same base pair on one strand, it is possible to reduce off-target effects because the enzymes are targeted to the cut site independently so for an off-target mutation to be induced both enzymes must target the same incorrect base pair. This occurs because with a nick in only a single strand the NHEJ mechanism is not employed to fix the damage (Fauser et al. 2014). In contrast, it is possible to create an insertion using the HDR pathway with nickases but as this does not occur without the presence of a template to be recombined with the host DNA then there is little chance for off-target effects through this process.

Creating Cas9 fusion proteins provides versatility and the potential to conduct assays beyond genome editing. Through the fusion of a Cas9 with other proteins, through end-to-end fusion or domain insertion, scientists have created multidomain proteins to manipulate Cas9s properties (Chen et al. 2013, Polstein and Gersbach 2015, Bolukbasi et al. 2015, Maji et al. 2017). Here are

three important ways that the CRISPR/Cas9 system has been modified to perform other functions. CRISPR/Cas9 has been used to control gene expression through several methods which rely on a protein fused to a version of Cas9 where the endonuclease activity has been removed through point mutations in that domain (Chen et al. 2013). The function of the multidomain Cas9 protein depends on which proteins are fused to disabled Cas9 (dCas9). When light-inducible heterodimerizing cryptochrome 2 (*CRY2*), and calcium and integrin-binding protein 1 (*CIB1*) are fused to dCas9 in the presence of a transcriptional activator, it is possible to induce transcriptional activation of a target gene with blue light irradiation. In this system blue light causes CIB1 and CRY2 to heterodimerize which recruits the transcriptional activator to dCas9/CIB1/CRY2 complex that was targeted to a specific locus by the gRNA (Nihongaki et al. 2015). Chemical induction of transcriptional activation has also been used in mammalian systems by fusing the human estrogen receptor 2 to either end of the Cas9 coding sequence, in the presence of the tamoxifen analog 4-HT, Cas9 was activated and in the absence of 4-HT, there was a low occurrence of editing events (K. Liu et al. 2016). In addition to chemical and light induction, there is one system that has used both, this system fuses an activator domain (VPR) to an inducible heterodimerization domain. Six different constructs were developed to respond to different stimuli ranging from hormone treatment to certain spectra of light when the stimulus was applied the VPR fusion dimerized with the dCas9 fusion protein leading to activation of expression of GFP which had been integrated into the genome (Gao et al. 2016). These technologies are a diverse toolset that allows for a variety of experiments to investigate many different processes. One of the most interesting is the ease and simplicity with which they allow investigators to interrogate transcription factor function with temporal specificity.

The Efficiency of Gene Editing

The benefits of CRISPR are lost if you are not able to efficiently create mutants. Soyars et al. 2018 argue that the most accurate way to measure CRISPR/Cas9 efficiency is to base it on the generation of stable mutant lines in the species of interest. Working with this definition it is clear that in *Arabidopsis thaliana* SpCas9 is efficient enough in most circumstances to generate many stable units in most genes. The major considerations to optimize efficiency are gRNA design which was discussed earlier and Cas9/gRNA delivery. There are different methods to deliver the enzyme and guide into your model system, in *Arabidopsis thaliana, Agrobacterium tumefaciens* mediated transformation is easy and efficient enough that in most cases alternative methods need not be considered, therefore the major consideration in optimizing Cas9 efficiency in *Arabidopsis* is the gRNA. It may be beneficial to stack several gRNAs that target the same gene to ensure the efficient targeting of that gene. A final technique that can improve editing efficiency is heat shock treatment of transformants at 37 degrees celsius which has been shown to increase efficiency up to 100 fold (Leblanc et al. 2018).

Genotyping Strategies

Once a plant has been transformed with a CRISPR/Cas9 vector and T1 seeds have been collected the next step is to genotype those progeny to determine in which siblings mutations have occurred, to this end many strategies have been developed. One of the simpler approaches isolates DNA from T1 seedlings and sequences across the CRISPR target sequence. An alignment with wild type is then performed and those siblings which show mutations are then self-crossed to isolate individuals homozygous for the desired mutation. It is important to note that in some systems it is possible to generate homozygous mutations in T1 through CRISPR

approaches, however, these occur very infrequently in Arabidopsis even when using germ-line specific promoters to drive Cas9 expression (Fauser et al. 2014, Feng et al. 2014). Other approaches to identify mutations include a modified cleaved amplicon polymorphic sequence (CAPS) which has been developed to detect small insertions and deletions (indels); this technique is called indCAPS. The indCAPS approach amplifies DNA around the Cas9 target site and then treats with an endonuclease with a restriction recognition site that includes the base targeted by Cas9. When an indel is produced by Cas9, the amplified DNA will not be digested by the restriction enzyme because the specific sequence recognized by the restriction enzyme is no longer present. This strategy allows for the rapid screening of many seedlings in a lab setting without sequencing, however, one challenge is the initial amplification of the sequences flanking the Cas9 target. This approach relies on a primer with at least one mismatch which allows for the creation of a restriction site in the amplified sequences. This primer must also be long enough that the cleaved sequence is easily detected through gel electrophoresis under standard conditions, which creates a primer of at least 20 base pairs with a mismatch and in our hands, this PCR is variable and often does not always produce amplified DNA to digest.

Phenotyping Considerations

Phenotyping mutations generated through CRISPR/Cas9 requires that special attention be taken to ensure that mutations are stable. This is due to the high occurrence of mosaicism in T1 plants, where Cas9 is generating mutations independently in each cell through the repeated cutting of the target DNA site until a misrepair occurs. It is not until T2 in Arabidopsis that it is common to generate stable homozygous mutant lines and it is in this generation that phenotyping should occur. Mosaic-induced phenotypes occur when a mutation is generated in individual somatic

cells but has not yet occurred in a germline cell leading to stably integrated mutations. These mosaic-induced phenotypes vary between transgenic lines so to ensure an accurate role of the gene of interest in the tissue of interest it is important to verify that mutations have been made stably.

Some considerations must be addressed before beginning an assay. CRISPR provides researchers with a tool to quickly generate mutants for reverse genetic analyses or to generate higher order mutants more rapidly than could be done using traditional crossing approaches.

REFERENCES

Anders, Carolin, Ole Niewoehner, Alessia Duerst, and Martin Jinek. 2014. "Structural Basis of PAM-Dependent Target DNA Recognition by the Cas9 Endonuclease." *Nature* 513 (7519): 569–73.

Barrangou, R., Fremaux, C., Deveau, H., Richards, M., Boyaval, P., Moineau, S., et al. 2007. "CRISPR provides acquired resistance against viruses in prokaryotes." Science 315: 1709–1712.

Bolotin, Alexander, Benoit Quinquis, Alexei Sorokin, and S. Dusko Ehrlich. 2005. "Clustered Regularly Interspaced Short Palindrome Repeats (CRISPRs) Have Spacers of Extrachromosomal Origin." *Microbiology* 151 (Pt 8): 2551–61.

Bolukbasi, Mehmet Fatih, Ankit Gupta, Sarah Oikemus, Alan G. Derr, Manuel Garber, Michael H. Brodsky, Lihua Julie Zhu, and Scot A. Wolfe. 2015. "DNA-Binding-Domain Fusions Enhance the Targeting Range and Precision of Cas9." *Nature Methods* 12 (12): 1150–56.

Bruegmann, Tobias, Khira Deecke, and Matthias Fladung. 2019. "Evaluating the Efficiency of gRNAs in CRISPR/Cas9 Mediated Genome Editing in Poplars." *International Journal of Molecular Sciences* 20 (15).

Candela, Héctor, and Sarah Hake. 2008. "The Art and Design of Genetic Screens: Maize." Nature Reviews. Genetics 9 (3): 192–203.

Chen, Baohui, Luke A. Gilbert, Beth A. Cimini, Joerg Schnitzbauer, Wei Zhang, Gene-Wei Li, Jason Park, et al. 2013. "Dynamic Imaging of Genomic Loci in Living Human Cells by an Optimized CRISPR/Cas System." *Cell* 155 (7): 1479–91.

Cong, Le, F. Ann Ran, David Cox, Shuailiang Lin, Robert Barretto, Naomi Habib, Patrick D. Hsu, et al. 2013. "Multiplex Genome Engineering Using CRISPR/Cas Systems." *Science* 339 (6121): 819–23.

Cuperus, Josh T., Taiowa A. Montgomery, Noah Fahlgren, Russell T. Burke, Tiffany Townsend, Christopher M. Sullivan, and James C. Carrington. 2010. "Identification of MIR390a Precursor Processing-Defective Mutants in Arabidopsis by Direct Genome Sequencing." *Proceedings of the National Academy of Sciences of the United States of America* 107 (1): 466–71.

Deltcheva, Elitza, Krzysztof Chylinski, Cynthia M. Sharma, Karine Gonzales, Yanjie Chao, Zaid A. Pirzada, Maria R. Eckert, Jörg Vogel, and Emmanuelle Charpentier. 2011. "CRISPR RNA Maturation by Trans-Encoded Small RNA and Host Factor RNase III." *Nature* 471 (7340): 602–7.

Fauser, Friedrich, Simon Schiml, and Holger Puchta. 2014. "Both CRISPR/Cas-Based Nucleases and Nickases Can Be Used Efficiently for Genome Engineering in Arabidopsis Thaliana." *The Plant Journal: For Cell and Molecular Biology* 79 (2): 348–59.

Feng, Z., Mao, Y., Xu, N., Zhang, B., Wei, P., Yang, D.-L., et al. (2014) Multigeneration analysis reveals the inheritance, specificity, and patterns of CRISPR/Cas-induced gene modifications in Arabidopsis. Proc. Natl. Acad. Sci. USA 111: 4632–4637.

Gao, Yuchen, Xin Xiong, Spencer Wong, Emeric J. Charles, Wendell A. Lim, and Lei S. Qi. 2016. "Complex Transcriptional Modulation with Orthogonal and Inducible dCas9 Regulators." *Nature Methods* 13 (12): 1043–49.

Gasiunas, Giedrius, Rodolphe Barrangou, Philippe Horvath, and Virginijus Šikšnys. 2012. "Cas9-crRNA Ribonucleoprotein Complex Mediates Specific DNA Cleavage for Adaptive Immunity in Bacteria." *Proceedings of the National Academy of Sciences of the United States of America* 109 (39): E2579–86.

Hazen, Samuel P., Justin O. Borevitz, Frank G. Harmon, Jose L. Pruneda-Paz, Thomas F. Schultz, Marcelo J. Yanovsky, Sarah J. Liljegren, Joseph R. Ecker, and Steve A. Kay. 2005. "Rapid Array Mapping of Circadian Clock and Developmental Mutations in Arabidopsis." *Plant Physiology* 138 (2): 990–97.

Ishino, Y., H. Shinagawa, K. Makino, M. Amemura, and A. Nakata. 1987. "Nucleotide Sequence of the Iap Gene, Responsible for Alkaline Phosphatase Isozyme Conversion in Escherichia Coli, and Identification of the Gene Product." Journal of Bacteriology 169 (12): 5429–33.

Jiang, Wenzhi, Huanbin Zhou, Honghao Bi, Michael Fromm, Bing Yang, and Donald P. Weeks. 2013. "Demonstration of CRISPR/Cas9/sgRNA-Mediated Targeted Gene Modification in Arabidopsis, Tobacco, Sorghum and Rice." *Nucleic Acids Research* 41 (20): e188.

Jiang, Wenyan, Inbal Maniv, Fawaz Arain, Yaying Wang, Bruce R. Levin, and Luciano A. Marraffini. 2013. "Dealing with the Evolutionary Downside of CRISPR Immunity: Bacteria and Beneficial Plasmids." *PLoS Genetics* 9 (9): e1003844.

Jinek, Martin, Krzysztof Chylinski, Ines Fonfara, Michael Hauer, Jennifer A. Doudna, and Emmanuelle Charpentier. 2012. "A Programmable Dual-RNA-Guided DNA Endonuclease in Adaptive Bacterial Immunity." *Science* 337 (6096): 816–21.

Karvelis, Tautvydas, Giedrius Gasiūnas, Algirdas Miksys, Rodolphe Barrangou, Philippe Horvath, and Virginijus Šikšnys. 2013. "crRNA and tracrRNA Guide Cas9-Mediated DNA Interference in Streptococcus Thermophilus." *RNA Biology* 10 (5): 841–51.

Kleinstiver, Benjamin P., Vikram Pattanayak, Michelle S. Prew, Shengdar Q. Tsai, Nhu T. Nguyen, Zongli Zheng, and J. Keith Joung. 2016. "High-Fidelity CRISPR-Cas9 Nucleases with No Detectable Genome-Wide off-Target Effects." *Nature* 529 (7587): 490–95.

LeBlanc, Chantal, Fei Zhang, Josefina Mendez, Yamile Lozano, Krishna Chatpar, Vivian F. Irish, and Yannick Jacob. 2018. "Increased Efficiency of Targeted Mutagenesis by CRISPR/Cas9 in Plants Using Heat Stress." *The Plant Journal: For Cell and Molecular Biology* 93 (2): 377–86.

Lister, Ryan, Brian D. Gregory, and Joseph R. Ecker. 2009. "Next Is Now: New Technologies for Sequencing of Genomes, Transcriptomes, and beyond." *Current Opinion in Plant Biology* 12 (2): 107–18.

Liu, Kaiwen Ivy, Muhammad Nadzim Bin Ramli, Cheok Wei Ariel Woo, Yuanming Wang, Tianyun Zhao, Xiujun Zhang, Guo Rong Daniel Yim, et al. 2016. "A Chemical-Inducible CRISPR-Cas9 System for Rapid Control of Genome Editing." *Nature Chemical Biology* 12 (11): 980–87.

Liu, Xiaoxi, Ayaka Homma, Jamasb Sayadi, Shu Yang, Jun Ohashi, and Toru Takumi. 2016. "Sequence Features Associated with the Cleavage Efficiency of CRISPR/Cas9 System." *Scientific Reports* 6 (January): 19675.

Maji, Basudeb, Christopher L. Moore, Bernd Zetsche, Sara E. Volz, Feng Zhang, Matthew D. Shoulders, and Amit Choudhary. 2017. "Multidimensional Chemical Control of CRISPR-Cas9." *Nature Chemical Biology* 13 (1): 9–11.

Makarova, Kira S., and Eugene V. Koonin. 2015. "Annotation and Classification of CRISPR-Cas Systems." Methods in Molecular Biology 1311: 47–75.

Mali, Prashant, Luhan Yang, Kevin M. Esvelt, John Aach, Marc Guell, James E. DiCarlo, Julie E. Norville, and George M. Church. 2013. "RNA-Guided Human Genome Engineering via Cas9." *Science* 339 (6121): 823–26.

Manavella, Pablo A., Jörg Hagmann, Felix Ott, Sascha Laubinger, Mirita Franz, Boris Macek, and Detlef Weigel. 2012. "Fast-Forward Genetics Identifies Plant CPL Phosphatases as Regulators of miRNA Processing Factor HYL1." Cell 151 (4): 859–70.

Mantegazza, Otho, Veronica Gregis, Marta Adelina Mendes, Piero Morandini, Márcio Alves-Ferreira, Camila M. Patreze, Sarah M. Nardeli, Martin M. Kater, and Lucia Colombo. 2014. "Analysis of the Arabidopsis REM Gene Family Predicts Functions during Flower Development." *Annals of Botany* 114 (7): 1507–15.

Matias-Hernandez, Luis, Raffaella Battaglia, Francesca Galbiati, Marco Rubes, Christof Eichenberger, Ueli Grossniklaus, Martin M. Kater, and Lucia Colombo. 2010. "VERDANDI Is a Direct Target of the MADS Domain Ovule Identity Complex and Affects Embryo Sac Differentiation in Arabidopsis." *The Plant Cell* 22 (6): 1702–15.

Mojica, Francisco J. M., César Díez-Villaseñor, Jesús García-Martínez, and Elena Soria. 2005. "Intervening Sequences of Regularly Spaced Prokaryotic Repeats Derive from Foreign Genetic Elements." *Journal of Molecular Evolution* 60 (2): 174–82.

O'Malley, Ronan C., Cesar C. Barragan, and Joseph R. Ecker. 2015. "A User's Guide to the Arabidopsis T-DNA Insertion Mutant Collections." *Methods in Molecular Biology* 1284: 323–42.

O'Malley, Ronan C., and Joseph R. Ecker. 2010. "Linking Genotype to Phenotype Using the Arabidopsis Unimutant Collection." *The Plant Journal: For Cell and Molecular Biology* 61 (6): 928–40.

Nihongaki, Yuta, Shun Yamamoto, Fuun Kawano, Hideyuki Suzuki, and Moritoshi Sato. 2015. "CRISPR-Cas9-Based Photoactivatable Transcription System." *Chemistry & Biology* 22 (2): 169–74.

Page, Damian R., and Ueli Grossniklaus. 2002. "The Art and Design of Genetic Screens: Arabidopsis Thaliana." *Nature Reviews. Genetics* 3 (2): 124–36.

Peterson, Brenda A., David C. Haak, Marc T. Nishimura, Paulo J. P. L. Teixeira, Sean R. James, Jeffery L. Dangl, and Zachary L. Nimchuk. 2016. "Genome-Wide Assessment of Efficiency and Specificity in CRISPR/Cas9 Mediated Multiple Site Targeting in Arabidopsis." *PloS One* 11 (9): e0162169.

Polstein, Lauren R., and Charles A. Gersbach. 2015. "A Light-Inducible CRISPR-Cas9 System for Control of Endogenous Gene Activation." *Nature Chemical Biology* 11 (3): 198–200.

Pourcel, C., G. Salvignol, and G. Vergnaud. 2005. “CRISPR Elements in Yersinia Pestis Acquire New Repeats by Preferential Uptake of Bacteriophage DNA, and Provide Additional Tools for Evolutionary Studies.” *Microbiology* 151 (Pt 3): 653–63.

Quadros, Rolen M., Hiromi Miura, Donald W. Harms, Hisako Akatsuka, Takehito Sato, Tomomi Aida, Ronald Redder, et al. 2017. “Easi-CRISPR: A Robust Method for One-Step Generation of Mice Carrying Conditional and Insertion Alleles Using Long ssDNA Donors and CRISPR Ribonucleoproteins.” *Genome Biology* 18 (1): 92.

Ren, Fengrui, Chong Ren, Zhan Zhang, Wei Duan, David Lecourieux, Shaohua Li, and Zhenchang Liang. 2019. “Efficiency Optimization of CRISPR/Cas9-Mediated Targeted Mutagenesis in Grape.” *Frontiers in Plant Science* 10 (May): 612.

Richardson, Christopher D., Graham J. Ray, Mark A. DeWitt, Gemma L. Curie, and Jacob E. Corn. 2016. “Enhancing Homology-Directed Genome Editing by Catalytically Active and Inactive CRISPR-Cas9 Using Asymmetric Donor DNA.” *Nature Biotechnology* 34 (3): 339–44.

Slaymaker, Ian M., Linyi Gao, Bernd Zetsche, David A. Scott, Winston X. Yan, and Feng Zhang. 2016. “Rationally Engineered Cas9 Nucleases with Improved Specificity.” *Science* 351 (6268): 84–88.

Terns, Michael P., and Rebecca M. Terns. 2011. “CRISPR-Based Adaptive Immune Systems.” *Current Opinion in Microbiology* 14 (3): 321–27.

Wang, Tim, Jenny J. Wei, David M. Sabatini, and Eric S. Lander. 2014. “Genetic Screens in Human Cells Using the CRISPR-Cas9 System.” *Science* 343 (6166): 80–84.

Zhang, Jian-Ping, Xiao-Lan Li, Guo-Hua Li, Wanqiu Chen, Cameron Arakaki, Gary D. Botimer, David Baylink, et al. 2017. “Efficient Precise Knockin with a Double Cut HDR Donor after CRISPR/Cas9-Mediated Double-Stranded DNA Cleavage.” *Genome Biology* 18 (1): 35.

Chapter 4

A Reverse Genetic Approach to Understanding the Molecular Mechanisms of Carpel Margin Meristem Development

Introduction

Historically forward genetic screens have been used to identify regulators of floral development (Endrizzi et al. 1996, Long et al. 1996, Chuang et al. 1999). Here we review genes regulating Carpel Margin Meristem (CMM) development and present a method to use reverse genetic screening to identify genes involved in CMM development.

In *Arabidopsis thaliana*, flowers are composed of four concentric whorls; from outermost to innermost they are the calyx (sepals), corolla (petals), androecium (stamens), and gynoecium which is composed of two fused carpels (Irish 2017). These fused carpels form a hollow tube inside which the ovules are formed and then mature into seeds. One of the early processes in gynoecium development is that a CMM forms along the margins between the two fused carpels (Bowman et al. 1999, Balanza et al. 2006, Azhakanandam et al. 2008). This meristem gives rise to many of the interior tissues of the gynoecium including the placenta, ovules, septum, and transmitting tract. Identifying and characterizing the function of the genes that maintain this meristem provides insight into the development of ovules and has potential agricultural value.

Many regulators of gynoecium development show subtle or no mutant phenotype when the gene is mutagenized. To this end it has been popular to look for second site enhancers of existing mutant lines; this technique is conducted in an existing characterized mutant background and uses various techniques (EMS mutagenesis, tDNA insertion, transposon tagging) to generate

mutants in this background (Liu and Meyerowitz 1995, Ishida et al. 2000). Progeny are screened for enhanced mutant phenotypes and then those plants are genotyped to discover the gene or genes responsible for this enhanced phenotype. The *leunig* mutant was identified in 1995 in a screen for second site enhancers of the *apetala2* (*ap2*) mutant. *AP2* encodes a transcription factor of the AP2/EREBP family that is involved in floral meristem identity, meristem maintenance, and floral organ identity. *LEUNIG* was then shown to regulate the expression of other major floral development regulators *APELATA1*(*AP1*), *APELATA3*(*AP3*), and *PISTILLATA* (*PI*) through the negative regulation of *AGAMOUS* (AG) transcription; in the *lug* mutant there is ectopic *AG* expression in the outer two whorls of the flower.

The functional role and the expression pattern of the MADS domain-containing transcription factor *AG* is complex and changes as the plant moves from maintaining a floral meristem to producing floral organs (Bowman et al. 1991). *LEUNIG* functions as a transcriptional repressor of *AG* in the outer two whorls of the flower (Conner and Liu. 2000, Liu and Meyerowitz. 1995). *SEUSS* (*SEU*), a second regulator of *AG* expression, was identified in a screen for enhancers of the transcription factor UNUSUAL FLORAL ORGANS (*UFO*) and again in a screen for enhancers of the MADS domain transcription factor CRABS CLAW (Franks et al. 2002, Levin et al. 1998, Samach et al. 1999, Eshed et al. 1999, Alvarez and Smyth. 1999). *SEU* encodes a transcriptional corepressor that, together with *LUG*, is involved in the repression of *AG* in the outer whorls(Conner and Liu. 2000, Franks et al. 2002, Liu and Meyerowitz 1995, Sridhar et al. 2004). The discovery of this corepressor highlights both the lack of strong phenotypes in many regulators of floral development and the effectiveness of screens looking for enhancers of previously characterized mutant phenotypes.

AINTEGUMENTA encodes a transcription factor that belongs to the AP2/EREBP family that was characterized in 1996 through transposon tagging, a type of forward genetic screen where plants are screened for a mutant phenotype after being transformed with a sequence which activates transposon movement (Elliot et al. 1996, Klucher et al. 1996, Krizek et al. 1999, Mizukami and Fischer. 2000). The line showing the mutant was then outcrossed to ensure that the phenotype was associated with a single locus and then the DNA flanking the transposon was amplified by polymerase chain reaction (PCR) to identify the location of the insertion. Using a polymorphism located in the flanking DNA this was determined to be at the bottom of chromosome four and to possess sequence homology to a previously identified ovule mutant which allowed it to be identified as a gene formally known as *DRAGON*, which was then renamed *ANT*.

There are eight *AINTEGUMENTA-LIKE* (*AIL*) genes in the *Arabidopsis thaliana* genome that share sequence similarity with ANT and have also been shown to be key regulators of many different developmental pathways (Horstman et al. 2014, Krizek. 2015, Krizek et al. 2016). Of these, three are involved in the shoot and floral meristem maintenance; *ANT*, *PLETHORA3* (*PLT3*), and *PLETHORA7* (*PLT7*) (Mudunkothge and Krizek. 2012, Prasad et al. 2011). In the floral meristem, these genes are involved in cellular proliferation, a function that *ANT* performs in the CMM. In the case of *ant ap2* double mutant, sepals, stamens, and petals fail to form at all (Elliot et al. 1996). These higher order mutants have been used to unravel the complex gene regulatory networks driving SAM, floral meristem (FM), and floral organ development.

Higher Order Mutants Provide Insights into the Gene Network Driving Gynoecium Development

While many of the previously discussed genes were identified through second site enhancer assays, it is through the generation of other double mutant combinations that some of the genetic interactions controlling gynoecium development have been identified. The transcriptional adapter *SEU* and the transcriptional repressor *LUG* both work to repress *AG* in the outer whorls of the flower (Liu and Meyerowitz 1995, Franks et al. 2002). To investigate the similar phenotypes of the two genes a *seu lug* double mutant was generated which showed an enhanced phenotype compared to the *lug* single mutants. The *seu lug* double mutant was characterized by reduced flower size, lower floral organ number, and homeotic transformation of the sepals into a carpelloid structure. It was later shown through *in vitro* assays that the LUG and SEU proteins physically interact directly (Sridhar et al. 2004). Also, assays conducted in yeast show that *LUG* can repress the expression of a reporter construct while *SEU* is unable to repress that reporter without *LUG*, which confirms that there is a genetic interaction between SEU and LUG which act together to repress *AG* expression.

LUG is not the only gene that interacts genetically with *SEU*, *ANT* is partially redundant with *LUG* and functions to repress *AG* (Lieu et al. 2000, Azhakanandam et al. 2008). In the *seu ant* double mutant, gynoecium development is disrupted and patterning fails early on resulting in no ovules being formed and a reduction in the CMM produced tissues. (Wynn et al. 2011, Wynn et al. 2014, Nole-Wilson et al. 2010)

Due to SEUSS' role as a transcriptional adapter and its role in varied pathways through confirmed interactions with LUG and ANT and suspected interactions with the PHABULOSA and CRABSCLAW, it presents an ideal mutant background to screen for other important floral regulators (Azhakanandam et al. 2008). One such screen was able to identify *CYP85A*, a cytochrome P450 involved in Brassinosteroid synthesis which was identified in a screen for genetic enhancers of the *seu* gynoecium phenotype (Nole-Wilson et al. 2010). This presents another pathway that SEU is involved in. In an attempt to better understand gynoecium development, we have developed an experimental approach which combines both traditional forward genetics and reverse genetics, and endeavors to identify genes that, when mutated, enhance the *seu* mutant phenotype in the hopes of identifying novel regulators of CMM development while clarifying the role of SEU in gynoecium development.

Results

Selection of Genes with Enriched Expression in the Gynoecium Allows for Targeted Screen for Enhanced CMM Phenotypes in a Sensitized Genetic Background

Discovering a gross morphological phenotype associated with a genetic mutation remains a critical aspect of reverse genetics. The phenotype lends credence to the work and provides a foundation to begin to elucidate the network underlying a developmental event. It is common though for a mutation in a single gene to not demonstrate an obvious morphological phenotype. This can be due to many reasons including redundancy at the network or gene level. One method to overcome this redundancy is to create higher order mutant lines, where several or all members of a gene family are mutated, in cases where there are many genes in a family capable and sufficient to accomplish a specific task this is an appropriate technique. An alternative

approach to discovering a phenotype associated with a particular gene is to utilize a sensitized background (Nole-Wilson et al. 2010, Ripoll et al. 2011, Wynn et. al 2011). By screening mutants in a background where a gene is already knocked out, it is possible to see a mutant phenotype in this sensitized background where in a wild type background you would not. This has been shown in *Arabidopsis thaliana* flowers in the *seu ant* double mutant which shows an enhanced gynoecium phenotype compared to the single mutants (Azhakanandam et al. 2008). It is possible to screen any mutagenized population in this way, however, we have taken a different approach by limiting the number of genes that we are focusing on and choosing to investigate not only double mutants but higher order mutants, up to septuple mutants.

We chose to conduct a screen of six mutant genes in *seu* and *ant* mutant backgrounds using a CRISPR approach. We hoped that this would allow us to generate higher order mutants, to examine mutants without tDNA insertion lines available, and to examine the many different higher order mutant combinations of our seven genes. When selecting candidate genes we relied on several criteria, they must be expressed in the *Arabidopsis* gynoecium and display enriched expression in the medial domain of the gynoecium, be highly expressed, preferably encode transcription factors, and not have a previously characterized mutant floral phenotype. When determining whether a gene was expressed in the flower, we relied on a two-pronged approach. First, the EFP browser on TAIR (Winter et al. 2007) was employed to determine the approximate expression pattern. This interactive tool visualizes gene expression data generated through the gene expression map of Arabidopsis development (Nakabayashi et al. 2005, Schmid et al. 2005, Waese et al. 2017). After this, a literature search was conducted on each gene to determine if the expression had been characterized through *in situ* hybridization or reporter constructs. In some

cases, there was no literature available on the genes of interest in that case we relied on the data from the EFP browser, as well as some transcriptomics work from our lab (Villarino et al. 2015). To determine if the genes were highly expressed in the gynoecium we used the work of Villarino *et al.* (2016) who published a comparative transcriptomics experiment that compared gene expression data in the medial domain of the gynoecium of approximately stage 8 gynecia to non-medial domain floral tissue. This work provided a list of 94 genes with enriched expression in the medial domain vs the rest of the floral tissue. We narrowed this list down to four transcription factors and then added two additional genes; At4G16230 a GDSL-motif lipase/esterase was included due to its high enrichment (24 fold) in the medial domain, *LAX1* showed 10 fold enrichment in the medial domain and has been shown to regulate phyllotaxy and other developmental processes through its role in promoting the polar transport of auxin (Akoh et al. 2004, Lai et al. 2017, Kramer 2004, Boot et al. 2016, Robert et al. 2015, Swarup and Peret 2012). Before proceeding with these genes we again reviewed the literature to ensure that no gynoecium phenotype had been described for these genes as we were looking for novel phenotypes.

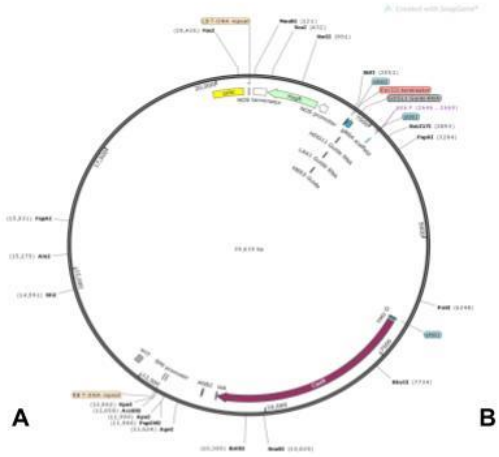
Table 4.1 Candidate Genes for CRISPR/Cas9 Mutagenesis. In order of enrichment in the medial domain vs the rest of the floral tissue the six target genes, their common gene names, and a source for their expression data if available.

Gene Identifier	Gene Name	Reference for expression	Fold Expression Difference	FDR
AT1G66950	PDR11	Wuest et. al 2010	45.22	2.02E-72
AT3G59220	PRN1	Orozco-Nunnelly et al. 2014	25.18	9.68E-38
AT4G16230	GDSL	NA	24.5	2.61E-70
AT2G21650	MEE3	Pagnusset et al. 2005	13.87	4.37E-15
AT5G01240	LAX1	NA	9.71	2.45E-56
AT1G73360	HDG11	Nakamura et al. 2006	4.2	9.41E-25

CRISPR/Cas9 Mediated Knockout of 7 Putative Floral Development Regulators

Using pMTN3164, a gateway compatible plasmid containing the coding sequence for Cas9, three different vectors were generated containing gRNAs for the different target genes (Denbow et al. 2017). The first contained the gRNA for *MEE3* alone, as this was the only gene of our targets without a tDNA insertion line where there was an insertion in an exon. The second construct contains the gRNAs for the three transcription factors; *PRN1*, *HDG11*, and *MEE3* and *PDR11*. The third vector contains gRNAs for all six targets. The gRNA sequences are shown in Figure 1. These have been optimized to work with a PAM sequence of NGG. Additionally, they target the first exon of the floral genes to increase the odds of a premature stop codon and a nonfunctional protein. To maximize the screen, these constructs were transformed into both *ant* and *seu* mutants through a floral dip.

A key to the effectiveness of any CRISPR assay is the efficiency of the gRNAs. CRISPR activity was assessed in T1 seedlings and to increase the efficiency of editing events T1 seedlings were treated with heat stress (Leblanc et al. 2017). Gene editing efficiencies vary between the different guides with MEE3 being the least efficient guide with only rare editing events occurring. This decreases the chances of generating a septuple mutant, but due to the strategy of this experiment we are still able to screen this and the other target genes for enhanced floral phenotypes. T2 seed was collected from the lines showing the highest CRISPR activity in T1. These T2 seeds were then genotyped and screens began for phenotypes in this generation.



B

Gene ID	Gene Name	gRNA Sequence
AT1G66950	PDR11	TTGGCAATGGCCGCAATGCT
AT3G59220	PRN1	AGTTCTAGCAAAGCTTGAGA
AT4G16230	GDSL	TTCTCCGAGGTGTGCCTTGC
AT2G21650	MEE3	TCAGGCTCAATGTCTTCTTA
AT5G01240	LAX1	AGTAGTCTCCGGGGAGGATG
AT1G73360	HDG11	GGCGTCGGCGGAAGTGGTAG

Figure 4.1 A Vector for the Mutation of Six Target Genes. A. pMTN3164 is a vector compatible with gateway cloning to allow for easy cloning of variable gRNAs. B. Guide RNAs for the six distinct genetic targets.

Genotyping of CRISPR Mutants is the Major Hurdle to Generating Stable Mutants

With efficient cloning through the gateway system and easy and efficient transformation in *Arabidopsis thaliana* using the floral dip method, the rate-limiting step of generating mutants through CRISPR/Cas9 is genotyping (Clough and Bent 2008). When generating mutants

through CRISPR in *Arabidopsis thaliana* genotyping must be performed twice. In the first genotyping event we confirmed the presence of Cas9 and assayed the efficiency of the gRNAs for our different targets. We first chose to genotype using an indel derived cleaved amplified polymorphic sequence (inDCAPS) method; first amplifying a fragment around the target cut site using a forward primer with a mismatch at the 3' end to introduce a restriction enzyme site that overlaps the Cas9 cut site (Hodgens et al. 2017). In cells where an indel was generated, the restriction site was disrupted and upon treatment, there would be no digestion. In practice, we found that getting amplification with these primers was challenging, as out of necessity they are long to provide a noticeable shift in fragment size after digestion and contain a mismatch. While amplification proved challenging we were able to screen 4 of the 6 targets in T1 plants using this method and selected the lines with activity in the highest number of guides to move to the next generation. Additionally, we confirmed the presence of the Cas9 gene through PCR amplification and then sequenced across the CRISPR cut site to confirm the activity of the guide and enzyme.

In the T2 generation, target plants may not be homozygous but should contain a stable mutation in all cells (*i.e.* not be chimeric). We screened seedlings for the presence of Cas9 and in this generation discarded those containing Cas9. The presence of Cas9 confounds the genotyping process and may lead to a higher rate of off-target cuts. In Cas9 negative seedlings, we chose to genotype by sequencing despite its higher cost compared to inDCAPS because of the ease of the experiment. Amplification of gene fragments from around the target cut sites was performed using primers without mismatches and these were far more efficient allowing for successful amplification far more often than the inDCAPS primers. Then PCR products were cleaned to

remove primer dimers and sequenced using Sanger sequencing. To assay whether or not there had been a mutation aligned the sequenced gene fragments against the wild type sequence. In mutants, we expected that mismatches would begin to occur three base pairs upstream of the PAM sequence (NGG) and continue if there were an indel and frameshift. We were able to detect these in T1 but were not successful in identifying any T2 lines where mutations had occurred and Cas9 was no longer present. This project was stalled due to the COVID-19 pandemic. At this point in the project, T2 samples must be genotyped to identify mutant lines and those then need to be homozygosed and phenotype. This will continue after quarantine is lifted.

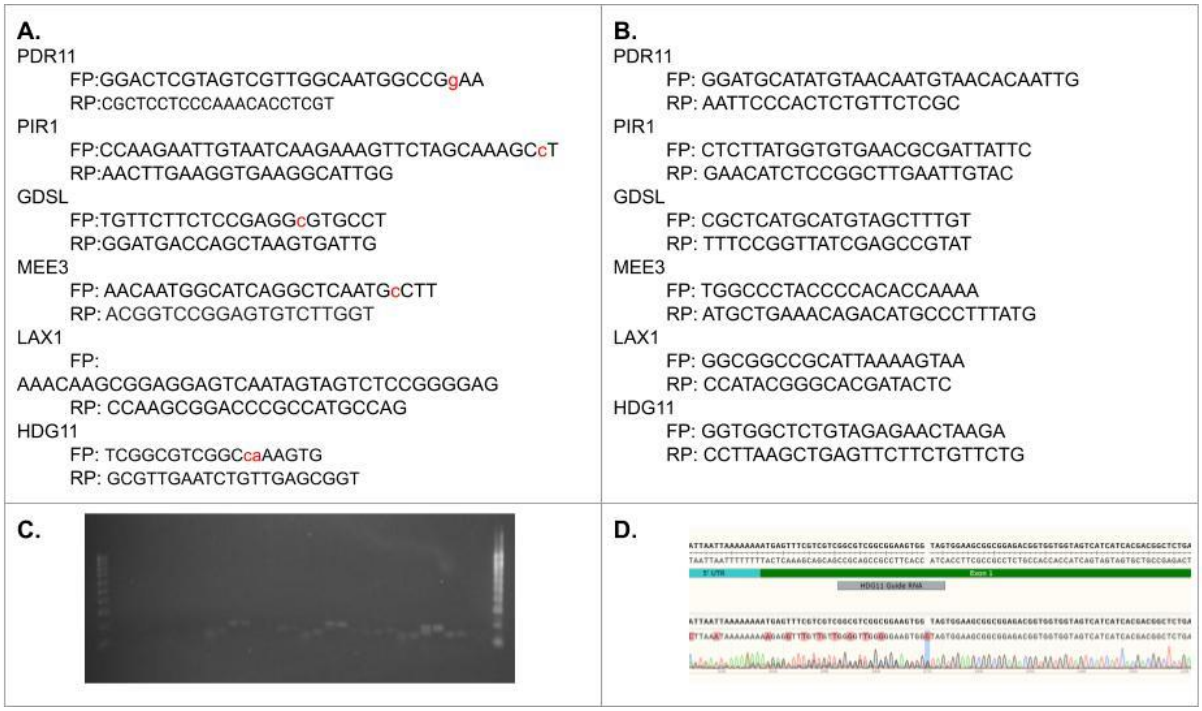


Figure 4.2 Genotyping of Cas9 Induced Mutations. A. inDCAPS primers contain a mismatch with the genomic sequence that creates a restriction enzyme recognition site and allows the amplified sequence to be cleaved unless there has been a CRISPR/Cas9 induced indel which disrupts the site. B. With a Sanger sequencing-based approach standard primers can be used for amplification and sequencing. C. Gel-electrophoresis of endonuclease treated DNA amplified with inDCAPS primers shows which lines have CRISPR activity. This can be identified by those where the enzyme does not cleave the amplified sequence. D. For those lines with potential CRISPR activity, the sequence around the cut site was amplified and then sequenced using Sanger sequencing. This follows the same procedure as used when genotyping by sequencing.

Discussion

Effectiveness of Screening a Small Number of Putative Floral Regulators

The efficacy of this assay is tied to two selections; the mutant background was chosen and the list of genes targeted for mutation. When choosing a background to screen multiple regulators it is beneficial to choose genes that may interact with many other genes, without this the full

potential of this assay is not being utilized. To learn the most about the targeted genes it would be ideal to discover distinct mutant phenotypes in each combination of mutants. This would allow for the genetic interactions between the target genes to be identified. If the gene was chosen as the mutant background only interacts with a single other gene then the likelihood of discovering how all of the genes being targeted relate to one another decreases drastically. Also, because the goal of the assay is to utilize a sensitized background to generate a genotype with a discernible phenotype the background must provide significant perturbation to the normal development of the plant. This ensures that when the target genes are mutated a new phenotype is shown with some regularity. To this end, the background must be a mutant for a gene with a significant role in the development of the tissue being studied.

Due to the nature of this assay, the selection of the genes to be targeted is of critical importance. In the Arabidopsis flower, screening for second site enhancer of mutant phenotypes is common because single mutants often have no detectable floral phenotype. Our technique modifies this screen by generating all of the different combinations of mutants possible given 6 target genes (6 factorial combinations). In examining the target genes with this level of specificity and detail the number of genes screened must be reduced. The power of this technique lies in its targeted approach. The target genes must be selected with care. In our approach, we chose first to ensure that target genes were expressed highly in the tissue of interest (the gene mutated in the background was also expressed highly in this tissue). There is a chance that genes expressed in different tissues may interact and play a role in the development of the tissue of interest but since we were limited in the number of genes that could be screened limiting genes to those that were highly expressed in the tissue of interest increased the chances of producing mutant phenotypes

that could then be studied. For this same reason, we chose to focus primarily on transcription factors that interact directly with many other genes and often form hubs in complex transcriptional regulatory networks (Palumbo et al. 2014). Finally, we chose genes that hadn't been characterized because we were looking for novel mutant phenotypes. However, if the goal of the researchers is different from ours (i.e. to link a redundant gene family to a phenotype), then different criteria for selecting genes should be employed. In the case of a redundant gene family, at least all of the redundant members of the family should be targeted. To be successful using this assay both the mutant background and the genes to be screened should be careful to ensure that they may interact and align with the goals of your experiment.

Effectiveness of CRISPR in Generating Higher Order mutants

When screening potential regulators by creating higher order mutants in a well-characterized mutant background I found two areas of concern. The first is related to gRNA efficacy, in our screen we chose to build a construct with 6 different guide RNAs all targeting different genes. The problem with this technique is that gRNA efficiencies vary and since we are generating a guide with 6 different RNAs we need each of those to be efficient. Therefore several different gRNAs must be screened individually to assess their efficacy and then assembled into a construct to generate the higher order mutants. An alternative approach would be to limit the number of genes targeted and to generate multiple gRNAs to target each gene.

The second area of concern is genotyping. To most efficiently identify stable mutations the gRNAs must be efficient. I suggest genotyping by sequencing as it is the simplest method although it is more expensive than other methods. From our experiences combining thorough

testing of gRNA efficiencies before generating a construct with multiple gRNAs and a genotyping by sequencing approach should optimize chances of success when generating higher order mutants in *Arabidopsis thaliana* using a CRISPR based system.

Materials and Methods

Cloning

gRNAs were designed to target each of the six target genes. Targets within each gene were selected using eCRISP. Cloning was accomplished using a SLIM PCR approach and the pDONOR207 plasmid (Invitrogen) which was used in combination with the binary vector pMTN3164. This vector was transformed into *Agrobacterium tumefaciens* and transformed into plants by a floral dip.

Transformation

Plants were transformed using the floral dip method described by Zhang et al. 2006. These transformations were conducted in Col-0, *ant-1* mutant, and *seu-3* mutant backgrounds.

Isolation of DNA

DNA isolations were performed in tissue collected from seedlings after they had produced four leaves according to the protocol in Edwards et al. 1991 with several modifications. Samples were frozen in liquid nitrogen before fragmentation which was done using two glass beads in the Eppendorf tube where the tissue was collected. Samples were then shaken in a Silamat S6 mixing amalgam. After the addition of Edwards Buffer, an equal amount of chloroform was added, working under a chemical hood. Then samples were mixed using a vortexer for 10

seconds. After allowing samples to sit for 10 minutes 300 microliters of the upper was taken off and isopropanol precipitation was performed according to Green and Sambrook.

inDCAPS Genotyping

inDCAPS genotyping was performed according to Hodgins et al. 2017.

Genotyping by Sequencing

When genotyping by sequencing we amplified regions around the predicted Cas9 site ensuring that amplified fragments were at least 150 base pairs long. Samples were cleaned using a capillary run using exosap for purification. Then samples were sequenced using sanger sequencing by the genome sciences lab at NCSU.

REFERENCES

Alvarez, John, and David R. Smyth. 1999. "Development in Parallel with AGAMOUS." *Development* 126: 2377–86.

Akoh, Casimir C., Guan-Chiun Lee, Yen-Chywan Liaw, Tai-Huang Huang, and Jei-Fu Shaw. 2004. "GDSL Family of Serine Esterases/lipases." *Progress in Lipid Research* 43 (6): 534–52.

Azhakanandam, Sridevi, Staci Nole-Wilson, Fang Bao, and Robert G. Franks. 2008. "SEUSS and AINTEGUMENTA Mediate Patterning and Ovule Initiation during Gynoecium Medial Domain Development." *Plant Physiology* 146 (3): 1165–81.

Balanzá, Vicente, Marisa Navarrete, Marina Trigueros, and Cristina Ferrándiz. 2006. "Patterning the Female Side of Arabidopsis: The Importance of Hormones." *Journal of Experimental Botany* 57 (13): 3457–69.

Bao, Fang, Sridevi Azhakanandam, and Robert G. Franks. 2010. "SEUSS and SEUSS-LIKE Transcriptional Adaptors Regulate Floral and Embryonic Development in Arabidopsis." *Plant Physiology* 152 (2): 821–36.

Boot, Kees J. M., Sander C. Hille, Kees R. Libbenga, Lambertus A. Peletier, Paulina C. van Spronsen, Bert van Duijn, and Remko Offringa. 2016. "Modelling the Dynamics of Polar Auxin Transport in Inflorescence Stems of Arabidopsis Thaliana." *Journal of Experimental Botany* 67 (3): 649–66.

Bowman, J. L., G. N. Drews, and E. M. Meyerowitz. 1991. "Expression of the Arabidopsis Floral Homeotic Gene AGAMOUS Is Restricted to Specific Cell Types Late in Flower Development." *The Plant Cell* 3 (8): 749–58.

Bowman, John L., Stuart F. Baum, Yuval Eshed, Joanna Putterill, and John Alvarez. 1999. "4 Molecular Genetics of Gynoecium Development in Arabidopsis." In *Current Topics in Developmental Biology*, edited by Roger A. Pedersen and Gerald P. Schatten, 45:155–205. Academic Press.

Clough, Steven J., and Andrew F. Bent. 1998. "Floral Dip: A Simplified Method forAgrobacterium-Mediated Transformation of Arabidopsis thaliana: Floral Dip Transformation of Arabidopsis." *The Plant Journal: For Cell and Molecular Biology* 16 (6): 735–43.

Conner, J., and Z. Liu. 2000. “LEUNIG, a Putative Transcriptional Corepressor That Regulates AGAMOUS Expression during Flower Development.” *Proceedings of the National Academy of Sciences of the United States of America* 97 (23): 12902–7.

Denbow, Cynthia J., Samantha Lapins, Nick Dietz, Raelynn Scherer, Zachary L. Nimchuk, and Sakiko Okumoto. 2017. “Gateway-Compatible CRISPR-Cas9 Vectors and a Rapid Detection by High-Resolution Melting Curve Analysis.” *Frontiers in Plant Science* 8: 1171.

Edwards, K., C. Johnstone, and C. Thompson. 1991. “A Simple and Rapid Method for the Preparation of Plant Genomic DNA for PCR Analysis.” *Nucleic Acids Research* 19 (6): 1349.

Elliott, R. C., A. S. Betzner, E. Huttner, M. P. Oakes, W. Q. Tucker, D. Gerentes, P. Perez, and D. R. Smyth. 1996. “AINTEGUMENTA, an APETALA2-like Gene of Arabidopsis with Pleiotropic Roles in Ovule Development and Floral Organ Growth.” *The Plant Cell* 8 (2): 155–68.

Endrizzi, K., B. Moussian, A. Haecker, J. Z. Levin, and T. Laux. 1996. “The SHOOT MERISTEMLESS Gene Is Required for Maintenance of Undifferentiated Cells in Arabidopsis Shoot and Floral Meristems and Acts at a Different Regulatory Level than the Meristem Genes WUSCHEL and ZWILLE.” *The Plant Journal: For Cell and Molecular Biology* 10 (6): 967–79.

Eshed, Y., S. F. Baum, and J. L. Bowman. 1999. “Distinct Mechanisms Promote Polarity Establishment in Carpels of Arabidopsis.” *Cell* 99 (2): 199–209.

Franks, Robert G., Chunxin Wang, Joshua Z. Levin, and Zhongchi Liu. 2002. “SEUSS, a Member of a Novel Family of Plant Regulatory Proteins, Represses Floral Homeotic Gene Expression with LEUNIG.” *Development* 129 (1): 253–63.

Green, Michael R., and Joseph Sambrook. 2017. “Precipitation of DNA with Isopropanol.” *Cold Spring Harbor Protocols* 2017 (8): db.prot093385.

Irish, Vivian. 2017. “The ABC Model of Floral Development.” *Current Biology: CB* 27 (17): R887–90.

Hodgens, Charles, Zachary L. Nimchuk, and Joseph J. Kieber. 2017. “indCAPS: A Tool for Designing Screening Primers for CRISPR/Cas9 Mutagenesis Events.” *PloS One* 12 (11): e0188406.

Horstman, Anneke, Viola Willemsen, Kim Boutilier, and Renze Heidstra. 2014. "AINTEGUMENTA-LIKE Proteins: Hubs in a Plethora of Networks." *Trends in Plant Science* 19 (3): 146–57.

Ishida, T., M. Aida, S. Takada, and M. Tasaka. 2000. "Involvement of CUP-SHAPED COTYLEDON Genes in Gynoecium and Ovule Development in Arabidopsis Thaliana." *Plant & Cell Physiology* 41 (1): 60–67.

Klucher, K. M., H. Chow, L. Reiser, and R. L. Fischer. 1996. "The AINTEGUMENTA Gene of Arabidopsis Required for Ovule and Female Gametophyte Development Is Related to the Floral Homeotic Gene APETALA2." *The Plant Cell* 8 (2): 137–53.

Kramer, Eric M. 2004. "PIN and AUX/LAX Proteins: Their Role in Auxin Accumulation." *Trends in Plant Science* 9 (12): 578–82.

Krizek, B. A. 1999. "Ectopic Expression of AINTEGUMENTA in Arabidopsis Plants Results in Increased Growth of Floral Organs." *Developmental Genetics* 25 (3): 224–36.

Lai, Chia-Ping, Li-Min Huang, Long-Fang O. Chen, Ming-Tsair Chan, and Jei-Fu Shaw. 2017. "Genome-Wide Analysis of GDSL-Type Esterases/lipases in Arabidopsis." *Plant Molecular Biology* 95 (1-2): 181–97.

LeBlanc, Chantal, Fei Zhang, Josefina Mendez, Yamile Lozano, Krishna Chatpar, Vivian F. Irish, and Yannick Jacob. 2018. "Increased Efficiency of Targeted Mutagenesis by CRISPR/Cas9 in Plants Using Heat Stress." *The Plant Journal: For Cell and Molecular Biology* 93 (2): 377–86.

Levin, J. Z., J. C. Fletcher, X. Chen, and E. M. Meyerowitz. 1998. "A Genetic Screen for Modifiers of UFO Meristem Activity Identifies Three Novel FUSED FLORAL ORGANS Genes Required for Early Flower Development in Arabidopsis." *Genetics* 149 (2): 579–95.

Liu, Z., and E. M. Meyerowitz. 1995. "LEUNIG Regulates AGAMOUS Expression in Arabidopsis Flowers." *Development* 121 (4): 975–91.

Liu, Z., R. G. Franks, and V. P. Klink. 2000. "Regulation of Gynoecium Marginal Tissue Formation by LEUNIG and AINTEGUMENTA." *The Plant Cell* 12 (10): 1879–92.

Long, J. A., E. I. Moan, J. I. Medford, and M. K. Barton. 1996. "A Member of the KNOTTED Class of Homeodomain Proteins Encoded by the STM Gene of Arabidopsis." *Nature* 379 (6560): 66–69.

Mizukami, Y., and R. L. Fischer. 2000. “Plant Organ Size Control: AINTEGUMENTA Regulates Growth and Cell Numbers during Organogenesis.” *Proceedings of the National Academy of Sciences of the United States of America* 97 (2): 942–47.

Mudunkothge, Janaki S., and Beth A. Krizek. 2012. “Three Arabidopsis AIL/PLT Genes Act in Combination to Regulate Shoot Apical Meristem Function: AIL/PLT Genes Regulate Meristem Function.” *The Plant Journal: For Cell and Molecular Biology* 71 (1): 108–21.

Nakabayashi, Kazumi, Masanori Okamoto, Tomokazu Koshiba, Yuji Kamiya, and Eiji Nambara. 2005. “Genome-Wide Profiling of Stored mRNA in Arabidopsis Thaliana Seed Germination: Epigenetic and Genetic Regulation of Transcription in Seed.” *The Plant Journal: For Cell and Molecular Biology* 41 (5): 697–709.

Nakamura, Miyuki, Hiroshi Katsumata, Mitsutomo Abe, Naoto Yabe, Yoshibumi Komeda, Kotaro T. Yamamoto, and Taku Takahashi. 2006. “Characterization of the Class IV Homeodomain-Leucine Zipper Gene Family in Arabidopsis.” *Plant Physiology* 141 (4): 1363–75.

Nole-Wilson, Staci, Elizabeth E. Rueschhoff, Huda Bhatti, and Robert G. Franks. 2010. “Synergistic Disruptions in Seuss cyp85A2 Double Mutants Reveal a Role for Brassinolide Synthesis during Gynoecium and Ovule Development.” *BMC Plant Biology* 10 (September): 198.

Orozco-Nunnelly, Danielle A., Durreshahwar Muhammad, Raquel Mezzich, Bao-Shiang Lee, Lasanthi Jayathilaka, Lon S. Kaufman, and Katherine M. Warpeha. 2014. “Pirin1 (PRN1) Is a Multifunctional Protein That Regulates Quercetin, and Impacts Specific Light and UV Responses in the Seed-to-Seedling Transition of Arabidopsis Thaliana.” *PloS One* 9 (4): e93371.

Pagnussat, Gabriela C., Hee-Ju Yu, Quy A. Ngo, Sarojam Rajani, Sevugan Mayalagu, Cameron S. Johnson, Arnaud Capron, Li-Fen Xie, De Ye, and Venkatesan Sundaresan. 2005. “Genetic and Molecular Identification of Genes Required for Female Gametophyte Development and Function in Arabidopsis.” *Development* 132 (3): 603–14.

Palumbo, Maria Concetta, Sara Zenoni, Marianna Fasoli, Mélanie Massonnet, Lorenzo Farina, Filippo Castiglione, Mario Pezzotti, and Paola Paci. 2014. “Integrated Network Analysis Identifies Fight-Club Nodes as a Class of Hubs Encompassing Key Putative Switch Genes That Induce Major Transcriptome Reprogramming during Grapevine Development.” *The Plant Cell* 26 (12): 4617–35.

Prasad, Kalika, Stephen P. Grigg, Michalis Barkoulas, Ram Kishor Yadav, Gabino F. Sanchez-Perez, Violaine Pinon, Ikram Blilou, et al. 2011. “Arabidopsis PLETHORA Transcription Factors Control Phyllotaxis.” *Current Biology: CB* 21 (13): 1123–28.

Ripoll, Juan José, Adrienne H. K. Roeder, Gary S. Ditta, and Martin F. Yanofsky. 2011. “A Novel Role for the Floral Homeotic Gene APETALA2 during Arabidopsis Fruit Development.” *Development* 138 (23): 5167–76.

Robert, Hélène S., Wim Grunewald, Michael Sauer, Bernard Cannoot, Mercedes Soriano, Ranjan Swarup, Dolf Weijers, Malcolm Bennett, Kim Boutilier, and Jiří Friml. 2015. “Plant Embryogenesis Requires AUX/LAX-Mediated Auxin Influx.” *Development* 142 (4): 702–11.

Samach, A., J. E. Klenz, S. E. Kohalmi, E. Risseeuw, G. W. Haughn, and W. L. Crosby. 1999. “The UNUSUAL FLORAL ORGANS Gene of Arabidopsis Thaliana Is an F-Box Protein Required for Normal Patterning and Growth in the Floral Meristem.” *The Plant Journal: For Cell and Molecular Biology* 20 (4): 433–45.

Schmid, Markus, Timothy S. Davison, Stefan R. Henz, Utz J. Pape, Monika Demar, Martin Vingron, Bernhard Schölkopf, Detlef Weigel, and Jan U. Lohmann. 2005. “A Gene Expression Map of Arabidopsis Thaliana Development.” *Nature Genetics* 37 (5): 501–6.

Sridhar, Vaniyambadi V., Anandkumar Surendrarao, Deyarina Gonzalez, R. Steven Conlan, and Zhongchi Liu. 2004. “Transcriptional Repression of Target Genes by LEUNIG and SEUSS, Two Interacting Regulatory Proteins for Arabidopsis Flower Development.” *Proceedings of the National Academy of Sciences of the United States of America* 101 (31): 11494–99.

Swarup, Ranjan, and Benjamin Péret. 2012. “AUX/LAX Family of Auxin Influx Carriers-an Overview.” *Frontiers in Plant Science* 3 (October): 225.

Villarino, Gonzalo H., Qiwen Hu, Silvia Manrique, Miguel Flores-Vergara, Bhupinder Sehra, Linda Robles, Javier Brumos, et al. 2016. “Transcriptomic Signature of the SHATTERPROOF2 Expression Domain Reveals the Meristematic Nature of Arabidopsis Gynoecial Medial Domain.” *Plant Physiology* 171 (1): 42–61.

Waese, Jamie, Jim Fan, Asher Pasha, Hans Yu, Geoffrey Fucile, Ruian Shi, Matthew Cumming, et al. 2017. “ePlant: Visualizing and Exploring Multiple Levels of Data for Hypothesis Generation in Plant Biology.” *The Plant Cell* 29 (8): 1806–21.

Winter, Debbie, Ben Vinegar, Hardeep Nahal, Ron Ammar, Greg V. Wilson, and Nicholas J. Provart. 2007. "An 'Electronic Fluorescent Pictograph' Browser for Exploring and Analyzing Large-Scale Biological Data Sets." *PloS One* 2 (8): e718.

Wuest, Samuel E., Kitty Vijverberg, Anja Schmidt, Manuel Weiss, Jacqueline Gheyselinck, Miriam Lohr, Frank Wellmer, Jörg Rahnenführer, Christian von Mering, and Ueli Grossniklaus. 2010. "Arabidopsis Female Gametophyte Gene Expression Map Reveals Similarities between Plant and Animal Gametes." *Current Biology: CB* 20 (6): 506–12.

Wynn, April N., Elizabeth E. Rueschhoff, and Robert G. Franks. 2011. "Transcriptomic Characterization of a Synergistic Genetic Interaction during Carpel Margin Meristem Development in Arabidopsis Thaliana." *PloS One* 6 (10): e26231.

Wynn, April N., Andrew A. Seaman, Ashley L. Jones, and Robert G. Franks. 2014. "Novel Functional Roles for PERIANTHIA and SEUSS during Floral Organ Identity Specification, Floral Meristem Termination, and Gynoecial Development." *Frontiers in Plant Science* 5 (April): 130.

Zhang, Xiuren, Rossana Henriques, Shih-Shun Lin, Qi-Wen Niu, and Nam-Hai Chua. 2006. "Agrobacterium-Mediated Transformation of Arabidopsis Thaliana Using the Floral Dip Method." *Nature Protocols* 1 (2): 641–46.

CHAPTER 5

Conclusions and Future Directions

The Carpel Margin Meristem is a Critical Reproductive Meristem

The Carpel Margin Meristem (CMM) is a meristematic structure that forms at the margin of the two fused carpels in *Arabidopsis thaliana* (Alvarez and Smyth 2002). The CMM are meristematic structures that form early in the development of the carpels and are located along the interior of the tube formed by the fusion of the two carpels. This structure functions similarly to the shoot apical meristem, producing lateral organs as the carpels develop and gives rise to the ovules, placenta, transmitting tract. Despite the similarities between the CMM and the Shoot Apical Meristem (SAM) and Floral Meristem (FM), the CMM is not considered one of the canonical plant meristems.

Traditionally there are three stem cell niches in plants; the shoot apical meristem, root apical meristem, and vascular meristem. From these structures, the primary meristems; protoderm, procambium, and ground meristem are produced. Also, the shoot apical meristem can transition into a floral meristem which is determinate and terminates after the organs of the flower are produced. Our work has investigated the development of two domains of the gynoecium, one of which contains the CMM. We have identified putative GRNs regulating CMM development and lateral domain development. Also, we have identified a transcriptional shift that occurs in the developing CMM and modeled the GRN regulating that shift. In this chapter, I will briefly compare the development of the CMM to that of a canonical meristem, the SAM, with a focus on the termination of the meristem. I will then detail potential future directions for our research in

modeling gynoecium development. Finally, I will discuss future directions for functional analysis of putative floral regulators using both CRISPR and traditional methods.

Similarities Between the Shoot Apical Meristem and Carpel Margin Meristem

Traditionally there are three stem cell niches in plants; they are found in the SAM, root apical meristem (RAM), and the vascular meristem (VM) (Aichinger et al. 2012). The term stem cell niche is one used in animal developmental biology where it refers to "A stem-cell niche is an area of a tissue that provides a specific microenvironment, in which stem cells are present in an undifferentiated and self-renewable state. Cells of the stem-cell niche interact with the stem cells to maintain them or promote their differentiation." (Stem-Cell Niche Latest Research and News, Nature). In plants two pools of pluripotent stem cells are maintained throughout the life of the plant; they are found in the shoot and root apical meristems. These two meristems are considered indeterminate in that the meristems are active throughout the life of the plant (Shishkova et al. 2007). Despite being produced by the SAM the FM is determinate and terminates after the production of the different floral organs. The process regulating FM maintenance and termination is fascinating from a developmental point of view. Similarly to the SAM, the FM must balance maintenance of the stem cells with cell division and differentiation, however, the meristem must also terminate at the correct time. When the meristem fails to terminate properly extra whorls form (Prunet et al. 2009, Wynn et al. 2014) and when the meristem terminates early certain whorls of the flower fail to form (Ikeda et al. 2009, Roth et al. 2018).

The SAM and FM share many characteristics despite the difference in determinacy. The CLAVATA-*WUSCHEL* pathway, a negative feedback loop, plays a critical role in the development of both meristems by regulating stem cell maintenance (Fletcher et al. 1999, Brand et al. 2000, Schoof et al. 2000). In the determinate FM, WUS is involved in a second mechanism regulating stem cell maintenance. This second mechanism involves the transcription factor AGAMOUS (AG). AG interacts with WUS in a CLV independent manner to terminate floral stem cells at stage six of floral development (Lohmann et al. 2001). The carpel primordia are initiated at stage five, around stage six they fuse, and at stage seven the CMM has been initiated (Bowman et al. 1999). These meristems are maintained until stage nine when they fuse to form the placental tissue from which ovule primordia arise at stage ten. The SAM and CMM differ in meristem determinacy, which is highlighted by the short period of stem cell maintenance in the CMM, however, the mechanisms that control meristem maintenance and differentiation are similar. They both require proper polar auxin transport (PAT) and the correct timing of cytokinin signaling (Reyes-Olalde and de Folter 2019).

Through modeling the development of the medial domain of the gynoecium we have discovered a transcriptional shift that occurs around stage 9 (Chapter 2). This was done through fluorescence-activated cell sorting (FACS) to isolate tissues from the medial and lateral domains of the gynoecium. This study was conducted in plants that were roughly synchronized using the *ap1/cal1* inducible system (Wellmer et al. 2006, Chapter 2). Through isolation and sequencing of the transcriptomes of two different domains of the gynoecium beginning 5 days after induction of flowering (5 DAI) to 9 DAI, we were able to investigate the transcriptional dynamics of medial and lateral domain development. This model allowed us to discover a

transcriptional shift that takes place between 6 and 7 DAI. After investigating this transition, we were able to identify 11 transcription factors that may control this shift, including members of the NAC and TGA family of transcription factors. Members of each of those transcription factor families play important roles in CMM development (Kamiuchi et al. 2014, Wynn et al. 2014). Future directions focus on identifying the function of these genes and placing them in the current model of stem cell maintenance in meristems.

Future Directions Modeling the Development of the Medial Domain of the Gynoecium

Modeling gene regulatory networks provide researchers with a framework to design novel studies or integrate new work to existing networks. Due to the nature of this approach further analyses are required. There are two reasons for this; first functional analysis of key nodes is required to confirm the validity of the network. One approach to validate these nodes is through reverse genetic analyses of critical regulators predicted by the model. This is critical as the predicted interactions are based strictly on normalized expression values and unless backed up by a previous study there is no genetic link between nodes predicted by the model. To confirm these interactions reverse genetic screens search for a mutant phenotype after perturbing the gene.

In screening for phenotypes in mutants predicted by a model GRN, there are two different types to screen for; morphological and molecular phenotypes. In screening for morphological phenotypes we are searching for perturbations in the proper development of the gynoecium, severe mutant phenotypes include failure of the carpels to fuse, indehiscence, medial produced tissues failing to form, and failure to set seed. In molecular phenotypes we are looking for a

change in the expression of target genes when a putative regulator is mutated ie; when a node which is predicted to upregulate a target is mutated we expect to see reduced target expression.

Specifically, we hope to first generate three mutants, *tga9*, *lbd21*, and the *tga9 lbd21* double mutant. These are predicted to be key regulators of the transcriptional shift that occurs between 6 and 7 DAI because of their high number of predicted regulations and high motif scores. TGA9 is a TGA family transcription factor involved in stamen development which shows a floral developmental phenotype in the *tga9 tga10* double mutant where the gynoecium fails to develop properly, this may be due to the failure of the stamens to produce viable pollen (Murmu et al. 2010). Another TGA family transcription factor, *PERIANTHIA*, also known as *TGA8* is involved in CMM development through the transcriptional adapter *SEUSS*, so it will be interesting to compare *tga9* mutants to those in *pan* and other known gynoecium mutants (Wynn et al. 2014).

Also, the transcription factors involved in this shift show certain patterns that appear to be conserved during transcriptional shifts in development (Palumbo et al. 2014). In future analyses, we hope to identify the biological context of the 6 to 7 DAI transcriptional shift that we have identified here. By framing this transition in a biological context we hope to provide more specific insight into the development of the gynoecium. Specifically, confocal microscopy of the reporter line used for FACS experiments was underway but delayed to the COVID-19 pandemic. Ultimately we hope to use our model of CMM development to build on the current understanding of genetic interactions in the medial domain of the gynoecium to build a complete

and validated network connecting the many genes known to be involved in gynoecium development.

Reverse genetic analyses provide critical links between predicted gene functions and phenotypes, which allow researchers to confirm genetic interactions. For investigating putative regulators of the 6 DAI- 7 DAI transition mutant lines are available through the Arabidopsis biological resource center. These mutant lines were generated using T-DNA insertion to disrupt gene function (Krysan et al. 1999).

Optimizing CRISPR Generation of Floral Mutants for Functional Analysis of Gene Regulatory Network

CRISPR allows for the rapid and efficient generation of mutants and provides an excellent option for generating a mutant allele for genes with no exonic mutations or for generating higher order mutants, especially when not all target genes have a mutant phenotype. We have been working on the generation of a mutant line where six putative floral regulators will have CRISPR induced mutations. To this aim, we have designed and cloned gRNAs, transformed, genotyped T1 plants, and are working on identifying T2 plants containing the mutations in the target genes. We hope to continue to optimize a CRISPR system for the functional analysis of putative floral regulators and to use this in conjunction with a modeling approach to build a complete model of the gene regulatory networks regulating CMM development.

REFERENCES

Aichinger, Ernst, Noortje Kornet, Thomas Friedrich, and Thomas Laux. 2012. "Plant Stem Cell Niches." *Annual Review of Plant Biology* 63 (February): 615–36.

Bowman, John L., Stuart F. Baum, Yuval Eshed, Joanna Putterill, and John Alvarez. 1999. "4 Molecular Genetics of Gynoecium Development in Arabidopsis." In *Current Topics in Developmental Biology*, edited by Roger A. Pedersen and Gerald P. Schatten, 45:155–205. Academic Press.

Brand, U., J. C. Fletcher, M. Hobe, E. M. Meyerowitz, and R. Simon. 2000. "Dependence of Stem Cell Fate in Arabidopsis on a Feedback Loop Regulated by CLV3 Activity." *Science* 289 (5479): 617–19.

Fletcher, J. C., U. Brand, M. P. Running, R. Simon, and E. M. Meyerowitz. 1999. "Signaling of Cell Fate Decisions by CLAVATA3 in Arabidopsis Shoot Meristems." *Science* 283 (5409): 1911–14.

Ikeda, Miho, Nobutaka Mitsuda, and Masaru Ohme-Takagi. 2009. "Arabidopsis WUSCHEL Is a Bifunctional Transcription Factor That Acts as a Repressor in Stem Cell Regulation and as an Activator in Floral Patterning." *The Plant Cell* 21 (11): 3493–3505.

Kamiuchi, Yuri, Kayo Yamamoto, Masahiko Furutani, Masao Tasaka, and Mitsuhiro Aida. 2014. "The CUC1 and CUC2 Genes Promote Carpel Margin Meristem Formation during Arabidopsis Gynoecium Development." *Frontiers in Plant Science* 5 (April): 165.

Krysan, P. J., J. C. Young, and M. R. Sussman. 1999. "T-DNA as an Insertional Mutagen in Arabidopsis." *The Plant Cell* 11 (12): 2283–90.

Lohmann, J. U., R. L. Hong, M. Hobe, M. A. Busch, F. Parcy, R. Simon, and D. Weigel. 2001. "A Molecular Link between Stem Cell Regulation and Floral Patterning in Arabidopsis." *Cell* 105 (6): 793–803.

Müller, Christina Joy, Emma Larsson, Lukáš Spíchal, and Eva Sundberg. 2017. "Cytokinin-Auxin Crosstalk in the Gynoecial Primordium Ensures Correct Domain Patterning." *Plant Physiology* 175 (3): 1144–57.

Murmu, Jhadeswar, Michael J. Bush, Catherine DeLong, Shutian Li, Mingli Xu, Madiha Khan, Caroline Malcolmson, Pierre R. Fobert, Sabine Zachgo, and Shelley R. Hepworth. 2010. "Arabidopsis Basic Leucine-Zipper Transcription Factors TGA9 and TGA10 Interact with Floral Glutaredoxins ROXY1 and ROXY2 and Are Redundantly Required for Anther Development." *Plant Physiology* 154 (3): 1492–1504.

Palumbo, Maria Concetta, Sara Zenoni, Marianna Fasoli, Mélanie Massonnet, Lorenzo Farina, Filippo Castiglione, Mario Pezzotti, and Paola Paci. 2014. "Integrated Network Analysis Identifies Fight-Club Nodes as a Class of Hubs Encompassing Key Putative Switch Genes That Induce Major Transcriptome Reprogramming during Grapevine Development." *The Plant Cell* 26 (12): 4617–35.

Prunet, Nathanaël, Patrice Morel, Ioan Negrutiu, and Christophe Trehin. 2009. "Time to Stop: Flower Meristem Termination." *Plant Physiology* 150 (4): 1764–72.

Reinhardt, Didier, Eva-Rachele Pesce, Pia Stieger, Therese Mandel, Kurt Baltensperger, Malcolm Bennett, Jan Traas, Jirí Friml, and Cris Kuhlemeier. 2003. "Regulation of Phyllotaxis by Polar Auxin Transport." *Nature* 426 (6964): 255–60.

Reyes-Olalde, J. Irepan, and Stefan de Folter. 2019. "Control of Stem Cell Activity in the Carpel Margin Meristem (CMM) in Arabidopsis." *Plant Reproduction* 32 (2): 123–36.

Roth, Ohad, John P. Alvarez, Matan Levy, John L. Bowman, Naomi Ori, and Eilon Shani. 2018. "The KNOXI Transcription Factor SHOOT MERISTEMLESS Regulates Floral Fate in Arabidopsis." *The Plant Cell* 30 (6): 1309–21.

Schoof, H., M. Lenhard, A. Haecker, K. F. Mayer, G. Jürgens, and T. Laux. 2000. "The Stem Cell Population of Arabidopsis Shoot Meristems in Maintained by a Regulatory Loop between the CLAVATA and WUSCHEL Genes." *Cell* 100 (6): 635–44.

Shishkova, S., T. L. Rost, and J. G. Dubrovsky. 2008. "Determinate Root Growth and Meristem Maintenance in Angiosperms." *Annals of Botany* 101 (3): 319–40.

"Stem-Cell Niche - Latest Research and News | Nature." 2020. Springer Nature. May 1, 2020.

Wellmer, Frank, Márcio Alves-Ferreira, Annick Dubois, José Luis Riechmann, and Elliot M. Meyerowitz. 2006. "Genome-Wide Analysis of Gene Expression during Early Arabidopsis Flower Development." *PLoS Genetics* 2 (7): e117.

Wynn, April N., Andrew A. Seaman, Ashley L. Jones, and Robert G. Franks. 2014. "Novel Functional Roles for PERIANTHIA and SEUSS during Floral Organ Identity Specification, Floral Meristem Termination, and Gynoecial Development." *Frontiers in Plant Science* 5 (April): 130.

APPENDICES

Appendix A: Forces in the Shoot Apex

The current belief is that the shape of the shoot apical meristem is defined by the orientation and rate of new cell divisions. The most important region for these defining cell divisions is between the tip of the SAM to the base of the "apical dome". These new divisions impact the shoot apex by increasing the residual stress found in that region of the plant. The residual stress is the mechanical stress generated inside a tissue. The calculation of these specific stresses is complicated; in a single cell to determine the residual stress you must take into account turgor pressure, cell wall thickness, cell size, and cell geometry and this is in a single isolated cell. To calculate the stress in an entire tissue the calculations become significantly more complex because the location and surroundings of each cell must be taken into account. Despite the complexity of this system it is possible to calculate these stresses. One aid to this research is to use a comparison between states; by looking at a system of the same size and same shape with a different number of cell walls you can predict by how much the stress would increase.

In addition to calculating the stress on the developing SAM, it is possible to determine where they are coming from and what type of forces each domain is experiencing. As stated previously stresses may be increased by differential growth between the inner and outer cell layers but they can also be generated by differential growth between adjacent regions of a single tissue. The core of the SAM, the corpus is generally faster growing than the outer tunica layers this creates compressive stress on that tissue while the increasing internal volume creates tensile stress on the outer layers. These predictions have been tested through the observation of cuts made to the surface of the SAM. If the tunica layers are under tensile stress as predicted when scored with a blade the cut will pull open while if there was a compressive force the cut would be pushed

close. The work of Snow and Snow, 1951; Hussey, 1973; Dumais and Steele, 2000 shows that there is primarily tensile stress with small areas of compressive stress.

It is important to remember that the main source of all of this stress is turgor pressure against cell walls. The amount of stress increases with the number of cells so it is important to investigate the dynamics of cell division within the SAM and the shoot as a whole. The mechanism for determining the specific location of the formation of new cell walls is unknown, however, there are rules which describe the overall tendencies of cell division in these tissues. The most popular of these is Errera's rule, the modern interpretation of which states that the plane of division corresponds to the shortest option to halve the mother cell evenly. Hofmeister's rule states that new cell walls usually form normal to the principal axis of cell elongation, this rule proves to be more challenging to investigate than it appears at first glance because the longest axis of the plant cell is not always the current principal axis of cell elongation and may reflect instead the shape of the parent cell. There are also several rules based on geometry; the plane of cell division is always perpendicular to existing cell walls, that new cell walls are formed in a way that prevents the formation of 4-way junctions, and that new cell walls are always staggered when compared to one another. Also, it has been proposed that all of the previous geometric rules are not rules but instead emergent properties as a result of chemical signals induced by physical interactions. One more proposed rule is that there is no rule, Besson and Dumais proposed that cell division in plants is inherently random. Meyerowitz et al. were able to look at cell divisions in 3d and were able to build a function that better describes this pattern of cell divisions in the shoot apical meristem than the previously mentioned theories.

References

Grandjean, Olivier, Teva Vernoux, Patrick Laufs, Katia Belcram, Yuki Mizukami, and Jan Traas. 2004. "In Vivo Analysis of Cell Division, Cell Growth, and Differentiation at the Shoot Apical Meristem in Arabidopsis." The Plant Cell 16 (1): 74–87.

Robinson, Sarah, Agata Burian, Etienne Couturier, Benoit Landrein, Marion Louveaux, Enrique D. Neumann, Alexis Peaucelle, Alain Weber, and Naomi Nakayama. 2013. "Mechanical Control of Morphogenesis at the Shoot Apex." *Journal of Experimental Botany* 64 (15): 4729–44.

Appendix B: Distinct Gene regulatory Networks Control Medial and Lateral Domain Development

Through the research described in Chapter 2, a fluorescence-activated cell sorting (FACS) time course was conducted which allowed for the temporal and spatial separation of developing gynoecial tissues; the medial and lateral domains. RNA was isolated from these tissues and sequenced. From this data, a shift in transcription was identified that occurs between 6 and 7 days after floral induction (6-7 DAI). One of the early questions we answered when investigating the development of these two adjacent gynoecium tissues was whether the two tissues were transcriptionally distinct and if they maintained this over time. We concluded that not only were they transcriptionally distinct but that the parent tissue was more significant in determining transcriptome identity than the developmental stage (Supplemental Figure 2.4).

Knowing that these two gynoecium tissues were transcriptionally distinct we then chose to generate models of the gene regulatory networks associated with the development of each of the domains. These networks were built using differentially expressed genes (DEGs) that were identified using modified shannon entropy (MSE) (Figure 2.2, Supplemental Figure 2.5). In addition to containing distinct sets of genes the models of each network also displayed distinct motifs and architecture.

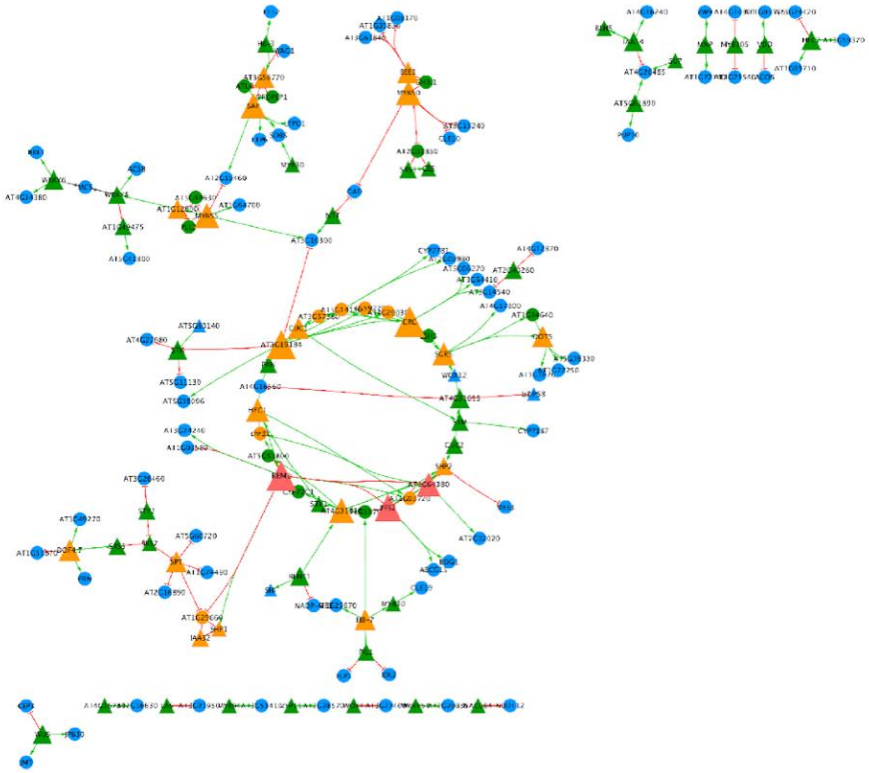


Figure 1. A Model of the Gene Regulatory Network Regulating Medial Domain Development. Transcription factors are shown as triangles and the size of the node correlates to the number of outgoing interactions it is associated with.

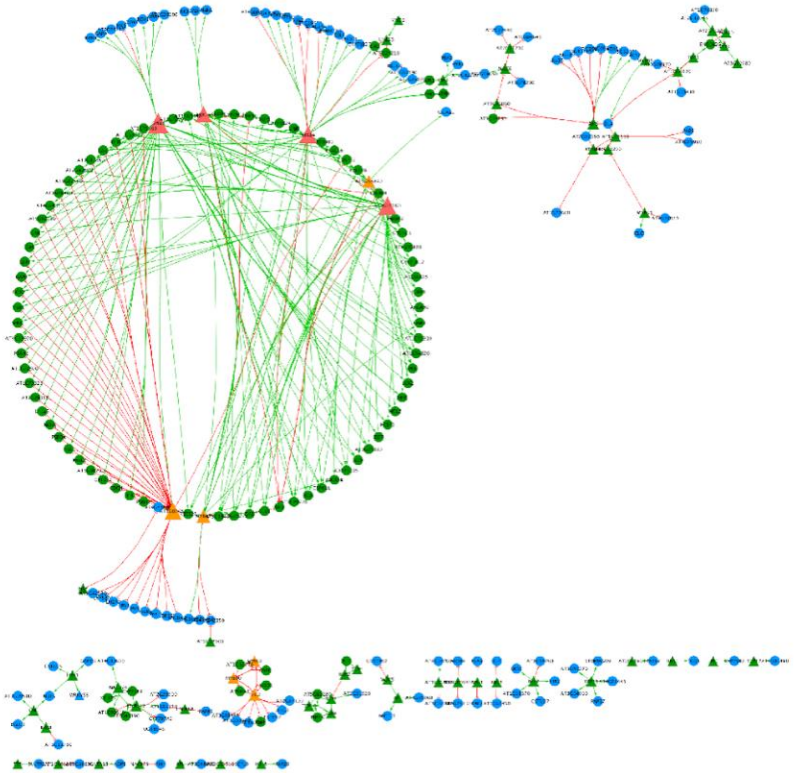


Figure 2. A Model of the Gene Regulatory Network Regulating Lateral Domain Development.