I give permission for public access to my thesis and for copying to be done at the discretion of the archives' librarian and/or the College library.

Naomi Eylath **05/07/2023**

Signature Date

Studying inflated calyx syndrome in *Physalis grisea*

by Naomi Eylath

A Paper Presented to the
Faculty of Mount Holyoke College in Partial Fulfillment of the Requirements for
the Degree of Bachelors of Arts with Honor
Department of Biological Sciences South Hadley, MA 01075

May 2023

This paper was prepared

under the direction of

Professor Amy Frary

for eight credits.

ACKNOWLEDGEMENTS

This work was funded by the National Science Foundation Plant Genome Research Program (IOS-2216612).

First and foremost, I would like to thank my advisor, Amy Frary, for her guidance throughout this project. Her expertise, support, and discerning eye were invaluable, and I learned an incredible amount under her tutelage. I began this journey as a seedling, and I now bear the fruits of my labor thanks to Amy's wonderful mentorship. I am also deeply grateful for her kindness, patience, and enthusiasm. Working with her on this thesis was both a privilege and a joy, and I could not have asked for a better advisor.

Next, I want to express my gratitude to the many people who assisted me on this project. I'd like to thank the Lippman Lab for providing the seeds and access to their sequence of the *P. grisea* genome. Without their outreach, this thesis would not have been possible. I would like to especially thank Jia He for his expertise on *P. grisea*, its genome, CRISPR-Cas9, and PCR primers. I am indebted to Jessie Blum, Thomas Clark, and the rest of the Talcott Greenhouse staff for their assistance in sowing and cultivating our *P. grisea* plants. I would like to thank Heather Hamilton for her comprehensive microscopy training, which truly came to the rescue in the face of technical difficulties. Thank you to Amy Camp for providing me with helpful sources on CRISPR-Cas9. Thank you as well to Craig Woodard and Mark McMenamin for being such wonderful members of my thesis committee, and for helping to make my thesis defense fun instead of frightening.

I would like to thank Hayley Emmons for being the most amazing collaborator in the Frary Lab. We learned together, laughed together, struggled together, and triumphed together. You were always there for me in my times of doubt, and I am so lucky to have worked alongside you. I can't wait to see what brilliant things you do in the future.

Thank you to all of my beloved friends who have been there for me throughout all my ups and downs during this process. Thank you for keeping me company, for encouraging me, and for being a comforting presence when things got tough. I am forever grateful to have you in my life.

Finally, thank you to my family. It is difficult to put into words just how grateful I am to you. Ever since I was a small child, you encouraged my passion for science and fostered my love of learning. Thank you for cultivating this spark within me, for supporting my interests and education, and for teaching me determination. Thank you for your loving nurture. Thank you so much for helping me get to where I am today. I love you.

TABLE OF CONTENTS

Page

LIST OF FIGURES

LIST OF TABLES

ABSTRACT

Physalis grisea, also known as groundcherry, is a member of the *Solanaceae* family and possesses a striking morphological novelty called inflated calyx syndrome (ICS). The inflated calyx is a modified version of sepals, which transform into a balloon-like husk that encases the fruit after the flower is fertilized. Because the genetic underpinnings of this trait remain unclear, the purpose of my project was to understand the genes controlling ICS in *P. grisea.* I focused on four homeotic genes from the MADS-box family, each one regulating an aspect of floral organ development. These genes are *MPF3, DEF, TAG1,* and *EJ2,* and previous research in other members of *Physalis* and the *Solanaceae* suggests that they are implicated in ICS development. I analyzed the floral morphology and genome sequences of four *P. grisea* CRISPR-edited mutant lines, each of which had one of the four genes of interest had been knocked out. My findings indicate that *MPF3*, *TAG1*, and *EJ2* each affect ICS differently: *mpf3* led to stunted inflated calyces; *tag1* was correlated with a lack of calyx inflation, likely due to male infertility from the transformation of stamens into petaloid organs; and *ej2* disrupted ICS development by altering sepal identity and number.

INTRODUCTION

The *Solanaceae* are a diverse and vibrant family of plants, many of which account for lucrative agricultural crops such as tomato, eggplant, potato, pepper, and tobacco. Primarily due to their economic importance, these crops have been studied extensively (Gebhardt 2016). Other members of this family have mystical, medicinal, or even dangerous associations, such as the psychoactive and poisonous mandrake, deadly nightshade, and jimson weed (Emboden 1989; Lee 2007; Devi et al. 2011). Amongst this colorful cohort lies *Physalis grisea*, commonly known as groundcherry. *P. grisea* is an annual herb native to North America (Sullivan 2004; Shenstone et al. 2020). Its upright stems and spreading branches are coated in a fine layer of small hairs, or trichomes, and it flowers from June to October (Sullivan 2004). The small, pale yellow flowers are composed of five fused petals, each with a dark purple spot that serves as a guide for pollinators (Figure 1). At the center of each flower sit five stamens surrounding a single carpel. At the base of each flower are five sepals, which expand into a balloon-like husk called an inflated calyx after the flower is fertilized (Figure 2). *P. grisea* is diploid and capable of both self- and cross-pollination, with a genome size of an estimated 1-2 Gb (He et al. 2023). Despite its round, golden-orange fruit, it is more closely related to the *Capsicum* genus— home to the chili and bell pepper— than it is to the genus *Solanum*, which includes tomato (Figure 3) (He et al. 2023).

Figure 1. *Physalis grisea* flower (Photographed by Alain Hogue).

Figure 2. The lifecycle of a *P. grisea* flower and the development of its inflated calyx (Photographed by the Lippman Lab).

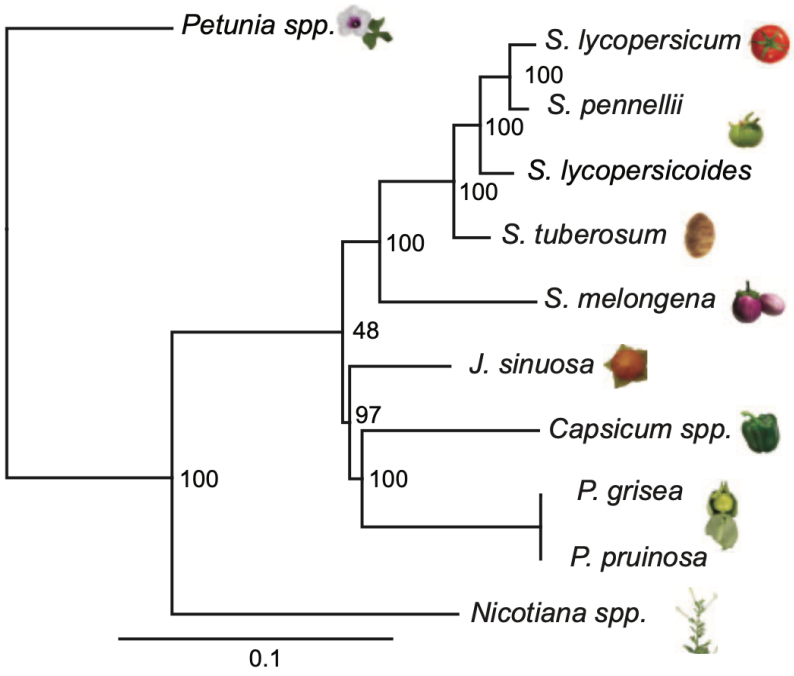

Figure 3. Molecular phylogeny of *Physalis grisea* in relation to other members of the *Solanaceae* family (He et al. 2023).

Physalis as an Orphan Crop

P. grisea, alongside multiple other members of the genus *Physalis*, is considered an orphan crop. This means that it is underutilized agriculturally on a global scale and has yet to be researched as extensively as the more popular crops amongst the *Solanaceae* (Lopez-Gomollon 2023). However, some *Physalis* are regionally important, contributing to the livelihood and food resources of local farmers and communities in low-income countries (Tadele 2019; Lopez-Gomollon 2023). The key issue with orphan crops is that, because they typically receive less scientific attention, they have not undergone intensive selection of favorable genetic traits to improve their yield and strengthen their resistance to adverse environmental conditions (Tadele 2019).

Orphan crops are a well of untapped potential both economically and nutritionally. This potential is abundant in *P. grisea*, which is easily cultivated in both fields and greenhouses, and whose fruit mature just about two months after sowing (J. He, personal communication). *Physalis* fruits also have nutritional benefits, as they are rich in essential vitamins and minerals such as vitamin C and potassium, as well as fatty acids and antioxidants (Shenstone et al. 2020). Since the *P. grisea* genome has been sequenced by the Lippman Lab at Cold Spring Harbor Laboratory (He et al. 2023), we now have the opportunity to more accurately investigate and manipulate its genes. This can facilitate targeted

selection for any number of traits to make groundcherry agriculturally viable on a global scale.

Inflated Calyx Syndrome

In addition to potentially helping introduce *P. grisea* into a wider agricultural market, the sequencing of this plant's entire genome allows us to better understand the genetic mechanisms underlying its most peculiar and striking characteristic: inflated calyx syndrome, or ICS. Post-fertilization hormonal signaling (Lu et al. 2021) triggers the expansion of the sepals into a large husk, or inflated calyx, which encases the fruit as it matures. This husk is a modified version of the sepals, floral organs that enclose developing buds and typically fan out at the base of a flower. Initially green, the inflated calyx desiccates to become papery and translucent once the fruit is fully mature (Figure 3).

The inflated calyx is not exclusive to *P. grisea*. Evidence suggests it has evolved convergently within 11 plant families, including *Malvaceae* (mallows), *Lamiaceae* (mints), and *Solanaceae* (Deanna et al. 2019). Besides the 75 species of *Physalis* (He and Saedler 2005), ICS is found in at least four other solanaceous genera: *Withania, Margaranthus, Nicandra,* and *Przewalskia* (Lu et al. 2021). A phylogenetic analysis of the Physalideae tribe performed by Deanna et al. (2019) indicates that this trait has consistently followed a directional evolutionary path. This path begins with a non-accrescent phenotype, meaning that the calyx does

not expand at all; the next stage is an accrescent calyx that presses against the fruit, either partially or fully wrapped around it; the final stage is the inflated calyx. This study of the Physalideae tribe revealed only two exceptions to the stepwise evolutionary pattern, with a pair of reversals that went from an inflated calyx back to the accrescent phenotype.

The evolutionary origins of ICS have not yet been entirely elucidated. However, various research studies approaching this question from multiple angles—genetics, phylogeny, and paleontology—have begun to unravel it. Work in *Physalis floridana* demonstrated that genes required for ICS appear to affect male fertility through the development of viable pollen (Lu et al. 2021). If these genes are indeed pleiotropic (i.e. influence these two different traits), this could indicate that the maintenance of fertility through selective pressure also brought along ICS as an extraneous trait. Then, because ICS possibly benefits fruit fitness, it could have been naturally selected over time to bring us the striking morphology we see today (Lu et al. 2021).

If its repeated evolution across taxa is any indication, ICS may confer a variety of adaptive advantages during fruit development and dispersal, such as protection from desiccation and predators (Wilf et al 2017; Deanna et al. 2018; Li et al 2019). In an attempt to ascertain the potential impact of ICS on fitness in *Physalis*, Li et al. (2019) performed several experiments in *P. floridana*. They found that the inflated calyx affects the temperature and humidity surrounding the

fruit, indicating that it may create a beneficial microclimate for fruit development. Upon measuring the chlorophyll content and photosynthetic rate of the calyx at different stages of maturity, they found the greatest indication of photosynthetic activity was between 5 and 20 days post-fertilization. This was higher than that of the mature calyx, which encapsulates the ripe fruit. These results indicate that the calyx may contribute positively to the growth of the fruit, perhaps by providing it with carbohydrates. However, the photosynthetic capabilities of the calyx were far lower than that of leaf tissue, so its contribution may be negligible and warrants further investigation before any definite conclusion can be made.

The inflated calyx may have also evolved to aid in seed dispersal, as it allows the fruit to float on water. Li et al. (2019) demonstrated this experimentally with *P. floridana*, whose intact fruits both remained buoyant and traveled along water currents. In contrast, fruits with their calyces removed immediately sank. This method of floating dispersal may date back to at least the Eocene, as fossilized progenitors of *Physalis* possessing a similar husk were discovered in Gondwanan Patagonia (Wilf et al. 2017). It is possible that ICS proved advantageous in ancient riparian environments, protecting fruits from rain and allowing them to stay afloat for days at a time (Wilf et al. 2017). Inflated calyces have also been found to facilitate dispersal in the wind, acting as a sort of sail that carries the fruit as it tumbles along the ground. Not only was this demonstrated by Li et al. with *P. floridana* (2019), it has also been observed in *Przewalskia* and *Solanum* section *Androceras* species (Knapp 2002). While the latter two examples

involve dry fruits whose seeds fall out as the calyx tumbles along the ground, they serve as further evidence that ICS confers adaptive advantages to seed dispersal across taxa.

There is a great deal more to learn about ICS, and this is not limited to investigating its possible effects on plant fitness. The genetic underpinnings of this morphological novelty remain unclear. My research goal is to understand, or at the very least further illuminate, the genetic basis of ICS in *P. grisea*. I approached this by analyzing the morphology and DNA of wild type plants in comparison with four genetically modified mutant lines— each one having been edited by the Lippman Lab using CRISPR-Cas9 to knock out a specific homeotic gene known for regulating floral organ development. These genes are *MPF3, DEF, TAG1,* and *EJ2*, and each one is a part of the MADS-box gene family. Their impact on determining floral organ identity suggests that they may be implicated in the development of ICS, because ICS itself is a floral trait.

The MADS-box and ABC Model

The MADS-box family is a group of genes that encode transcription factors which share a DNA binding domain called the MADS-box. This domain recognizes an array of similar DNA sequences, many of which are involved in determining cell identity and organ development in eukaryotes. The abbreviation "MADS" stands for representative genes in four different organisms: *MINICHROMOSOME MAINTENANCE 1* (*MCM1*) in the yeast *Saccharomyces*

cerevisiae, *AGAMOUS* (*AG*) in *Arabidopsis thaliana* (thale cress), *DEFICIENS* (*DEF*) in *Antirrhinum majus* (snapdragon), and serum response factor (*SRF*) in humans (Chen and Tye 1995; Ng and Yanofsky 2001; Irish 2017). In plants, MADS-box genes play a crucial role in the development of floral organs, as demonstrated chiefly in *Arabidopsis* and *Antirrhinum*, which have been studied extensively as model organisms (Ng and Yanofsky 2001; Irish 2017). It was through genetic research on these two plant species that the ABC model of floral organ development was first proposed (Irish 2017).

The ABC model is a useful way to visualize how different classes of MADS-box genes interact to regulate the identity of floral organs. In dicots, the group of angiosperms whose seeds have two cotyledons (embryonic leaves), most flowers comprise four concentric whorls (Ng and Yanofsky 2001). From the outside in, these whorls are as follows: Whorl 1 is the sepals, Whorl 2 is the petals, Whorl 3 is the stamens (male reproductive organs carrying pollen), and Whorl 4 is the carpels (female reproductive organs that contain ovules) (Ng and Yanofsky 2001; Irish 2017) (Figure 4). The development of these organs is governed by different homeotic genes, which are separated into classes A, B, and C.

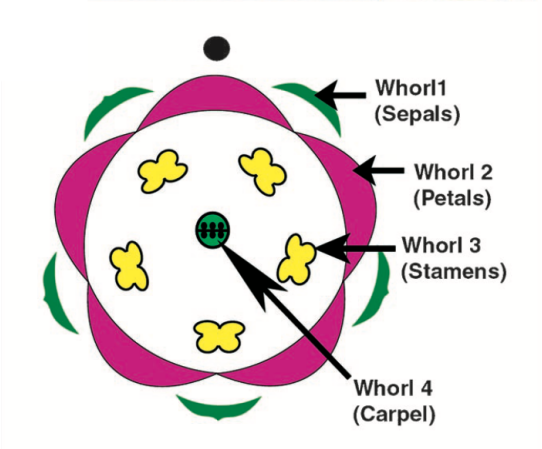

Figure 4. The four whorls of a flower (Kapoor et al. 2002).

Different combinations of ABC gene activity in each whorl inform organ identity, which is illustrated in *Arabidopsis* (Figure 5). Class A gene expression alone specifies sepals, A and B together specify petals, B and C together specify stamens, and C gene expression alone specifies carpels. When various A, B, or C genes are knocked out or silenced, organ development is altered significantly. For example, when *A* function is eliminated, the flower has no sepals or petals. Instead, it develops carpels in Whorl 1, stamens in Whorls 2 and 3, and carpels in Whorl 4. This is because A and C have a boundary setting, or antagonistic, relationship. Thus, when one does not function, the other will take over for determining floral organ development in the whorl. Another example of manipulating organ development is the elimination of *B* function, which creates a flower with only sepals and carpels. To further complicate things, class D and E genes have been discovered in the years following the establishment of the ABC model (Theißen et al. 2016). My research includes one of the Class E genes,

which work in concert with A, B, and C genes to regulate the development of the floral organs in each of the four whorls (Theißen et al. 2016; Sobral and Costa 2017).

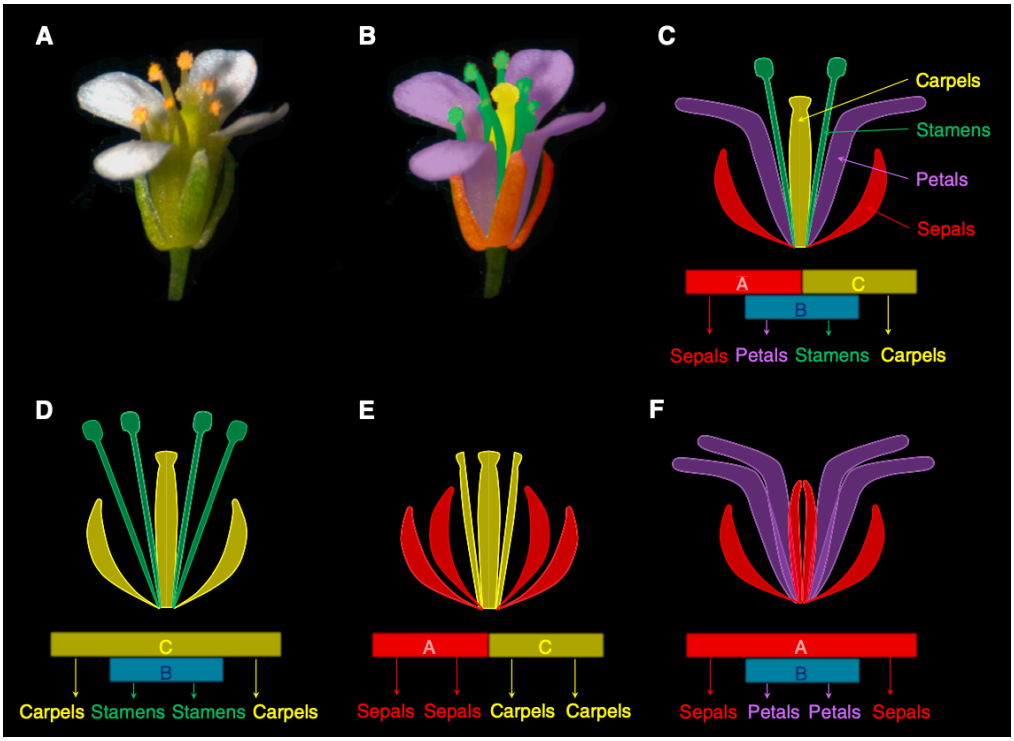

Figure 5. The ABC Model as illustrated in *Arabidopsis* (Irish 2007).

The Floral Quartet Model

How exactly do these MADS-box genes regulate floral organ development? This question is answered elegantly in the floral quartet model (FQM) (Theißen et al. 2016). According to this model, various combinations of MADS-domain transcription factors from the different classes bind together in groups of four. The dimers within these tetrameric complexes (quartets) recognize and bind to sequence motifs in the promoter regions of DNA known as CArG-boxes. When this occurs, the CArG-boxes are brought closer together, causing the DNA to form a loop during transcription. One hypothesis as to why this mechanism involves tetramers is that it increases the cooperation of DNA binding. The synchronized transcription that follows this cooperation can lead to the production of more proteins than if transcription was promoted by only dimers (Theißen et al. 2016).

Each of the four floral organs— sepals, petals, stamens, and carpel— is determined by a unique quartet of MADS-domain transcription factors from Classes A, B, C, and E (Figure 6). In Whorl 1, sepal identity is determined by a complex of two APETALA1 (AP1) Class A proteins and two SEPALLATA (SEP) Class E proteins. In Whorl 2, petal identity is determined by a complex of AP1, SEP, and the two Class B proteins APETALA3 (AP3) and PISTILLATA (PI). In Whorl 3, stamen identity is determined by SEP, AP1, PI, and the Class C protein

AGAMOUS (AG). In Whorl 4, carpel identity is determined by a pair of SEP proteins and a pair of AG proteins.

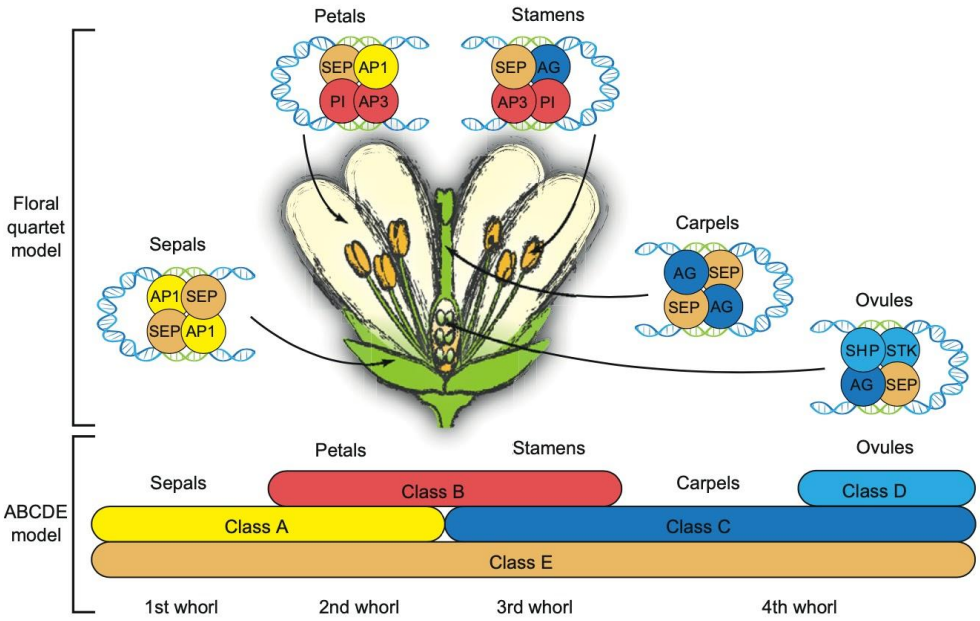

Figure 6. The floral quartet model as illustrated in *Arabidopsis* (Theißen et al. 2016).

Genes of Interest

As mentioned earlier, the four genes at the focus of my research (*MPF3, DEF, TAG1,* and *EJ2*) belong to the MADS-box family and code for transcription factors that regulate floral organ development. Each one of these genes in *P. grisea* is a homolog of a MADS-box gene found in either *Arabidopsis* or *Antirrhinum*. Homologous genes are derived from a common ancestor. They can be quite similar to one another in function but it is also possible for homologous genes to have entirely different functions (Moreira 2011). The former is the case for *P. grisea*, in which each gene of interest governs the same floral whorl as its homolog in *Arabidopsis* or *Antirrhinum*. However, though they regulate development in the same whorls, their precise effects in *P. grisea* are the subject of ongoing investigation. Using previous research on these genes and their homologs in *Physalis*, *Arabidopsis*, *Antirrhinum*, and tomato as a guide, I have predicted the phenotypes that may occur when they have been inactivated within *P. grisea*.

MPF3

MPF3 (MADS-box gene 3 as identified in *Physalis floridana*) is an ortholog of *APETALA1* in *Arabidopsis* (Zhao et al. 2021). A form of homology, orthology signifies that the gene diverged within *Physalis* and *Arabidopsis* after they speciated from their common ancestor (Moreira 2011). *MPF3* is a Class A gene that only regulates the development of sepals in *Physalis floridana* (Zhao et

al. 2013). MPF3, the protein that this gene encodes, forms dimers in the nucleus with another protein called MPF2. These dimers bind to particular CArG-boxes on the *MPF2* promoter, which regulates the gene's transcription (Zhao et al. 2013).

The close relationship between *MPF3* and *MPF2* is relevant because *MPF2* plays a role in ICS development. This was demonstrated by He and Saedler (2005) in *Solanum tuberosum* (potato). Potato does not have ICS, and its sepals are small (He et al. 2007). Its *MPF2* ortholog, *STMAS16*, is exclusively expressed in vegetative tissue. In contrast, *MPF2* is expressed in both the leaves and sepals of *Physalis peruviana* and *Physalis pubescens*— both of which have ICS (He and Saedler 2005). He and Saedler (2005) hypothesized that the lack of *STMADS16* expression in floral organs is the ancestral state, which makes sense given the stepwise evolution of ICS from a non-accresent to inflated calyx (Deanna et al. 2018). They found that the overexpression of *MPF2* and *STMADS16* in potato caused it to develop ICS. This evidence suggests that the heterotopic expression of *MPF2* is necessary for the development of ICS.

In order to determine the role of *MPF3* in ICS development, Zhao et al. (2013) downregulated this gene in *P. floridana* using two methods: virus induced gene silencing (VIGS) with tobacco rattle virus, and RNA-interference (RNAi). VIGS is a form of post-transcriptional regulation whereby host cells are inoculated with the tobacco rattle virus, which has been modified to include the

sequence of the gene of interest. This infection triggers the host plant's immune response. Under normal circumstances this would lead to the targeting and degradation of viral mRNA with double-stranded and short interfering RNA (dsRNA and siRNA) molecules and an enzyme complex called RISC (RNA-induced silencing complex). However, because the virus includes the host gene, the infected cell instead targets host mRNA to prevent the gene of interest from being translated into a protein (Lu et al. 2003, Ramegowda et al. 2014). RNAi, another form of post-transcriptional regulation, utilizes the same pathway as VIGS. It employs dsRNA, siRNA, and the RISC complex to target and destroy mRNA. However, the process of RNAi does not require viral RNA as a trigger, simply the introduction of some form of dsRNA (Agrawal et al. 2003). After downregulating *MPF3* with VIGS and RNAi, Zhao et al. (2013) observed stunted, malformed inflated calyxes in *P. floridana*. They also observed that downregulation of *MPF3* prevents the accumulation of starch in pollen, thus suggesting this gene plays a role in male fertility. Given this, I predicted that the *MPF3* knockout plants would exhibit differently-shaped inflated calyxes and have less viable pollen than the wild type.

DEF

DEF (*DEFICIENS*) in *P. grisea* is an ortholog of *APETALA3* in *Arabidopsis*, and also has a homolog in *Antirrhinum* (Ng and Yanofsky 2001). *DEF* belongs to Class B (Zhang et al. 2015), meaning it contributes to the

development of both petals and stamens. Similarly to Zhao et al. (2013), Zhang et al. (2015) used VIGS with tobacco rattle virus to downregulate *DEF* in *P. floridana*. This downregulation led to a spectrum of results. All flowers exhibited mutated petals and stamens, with some flowers resembling the wild type. In all the flowers showing an altered phenotype, the petals were transformed into sepals. In the most mild category of mutation, the stamens remained unchanged; in the second category, the stamens partially developed into carpels; and in the most extreme cases, the stamens completely developed into carpels.

However, *DEF* was not found to affect male fertility unless it was severely downregulated to the point of changing stamen identity entirely. If CRISPR had successfully knocked out *DEF* in the knockout line I cultivated, I expected to find a variety of mutations in the petals and stamens as described in *P. floridana*. In addition, because calyx inflation is likely triggered by fertilization (Lu et al. 2018), the transformation of stamens into carpels would inhibit this process and prevent the development of ICS.

TAG1

Originally identified in *Solanum lycopersicum* (tomato), *TAG1* (*TOMATO AGAMOUS 1*) is a homolog of *AGAMOUS* in *Arabidopsis* (Pnueli et al. 1994). *TAG1* is a Class C gene, meaning it is expressed in the developing stamens and carpels. In *Arabidopsis,* ectopic expression of *AGAMOUS* in Whorls 1 and 2 causes sepals to become carpels, and petals to become stamens (Mizukami et al.

1992). Pnueli et al. (1994) utilized antisense RNA in tomato to inhibit the expression of *TAG1*. This method introduces a synthesized RNA strand which binds to complementary mRNA, creating double-stranded RNA (dsRNA) that prevents protein synthesis (Mol et al. 1990). In tomato, *TAG1* antisense RNA turned stamens into petaloid organs and carpels into nested perianth flowers (Pnueli et al. 1994).

More recently, Gimenez et al. (2016) employed the alternative, yet similar, approach of RNAi for inhibiting *TAG1* expression in tomato. When *TAG1* was silenced via RNAi, the plants exhibited seedless fruits and decreased pollen viability (Gimenez et al. 2016). The mutant line I have worked with is a knockout of the *TAG1* ortholog in *P. grisea*. To my knowledge, *TAG1* has not been knocked out in *P. grisea* before. I anticipated that I would observe similar changes in the organ identity of the stamens and carpels, alongside less viable pollen and seedless fruits. Similar to the *DEF* knockout, this decrease in fertility could lead to a lack of ICS development.

EJ2

The tomato gene *EJ2* is a homolog of *SEPALLATA4* (Soyk et al. 2017), which belongs to Class E (Soza et al. 2016). As a Class E gene, it functions redundantly in all floral organ whorls alongside Classes A, B, and C. During the domestication of tomato, an *EJ2* loss-of-function mutation that leads to the development of a larger calyx and fruit was selected (Soyk et al. 2017). *EJ2* is

also a regulator of *J2*, whose mutated form results in the development of jointless pedicels which increase the ease of mechanical harvesting (Soyk et al. 2017; Reynard 1961). I predicted that the knockout of the *EJ2* ortholog in *P. grisea* could affect multiple floral organs, given its redundant function in regulating organ identity. Because the loss of function of *EJ2* in tomatoes leads to a larger calyx, I hypothesized that knocking it out may affect the calyx similarly in *P. grisea*.

CRISPR/Cas9 and Agrobacterium tumefaciens Mediated Transformation

What is particularly unique about this research project is that the knockout lines I have worked with were mutated using the CRISPR-Cas9 system. This is significant because knocking out genes via gene editing is more precise than silencing them by downregulating their expression. To "knock out" a gene is to inactivate it by altering its DNA sequence in a number of ways, such as by causing an insertion or deletion (Redei 2008). In contrast, the methods used to inhibit gene expression (VIGS, antisense RNA, RNAi) in the aforementioned studies all operate through mechanisms that occur after transcription and involve the targeting and degradation of mRNA. In other words, these techniques do not directly modify the genome, unlike CRISPR-Cas9, which edits the DNA of the target gene and leads to the synthesis of a different protein product. This leads to the question: What exactly is CRISPR-Cas9, and how was it used in *P. grisea* to create the knockout lines used in this project?

CRISPR (Clustered Regularly-Interspersed Short Palindromic Repeat) sequences were originally discovered in the *Escherichia coli* genome. These sequences consist of repeat units of 21-40 base pairs, interspersed with spacer units of 20-58 base pairs (Ishino et al. 2018). Each spacer sequence is from a fragment of foreign DNA introduced by a virus that infected the cell (or its predecessor) at some point. The bacterial cell uses Cas proteins, whose encoding sequences are adjacent to the CRISPR locus, to acquire and incorporate the fragmented DNA into its genome. This essentially stores a genetic memory of the previous infection. When the CRISPR locus is transcribed, it creates pre-CRISPR RNA (pre-crRNA). The pre-crRNA is then processed into small crRNA molecules, which form a complex with Cas proteins to target homologous DNA during any subsequent viral infections and cleave it to trigger its degradation (Mojica and Montoliu 2016; Ishino et al. 2018). Because CRISPR-Cas cleaves a recognized DNA sequence, it has been harnessed for the purpose of genetic modification in a variety of organisms. Cas9 is the most studied of the Cas proteins, and is the most popular one for use in gene editing (Ishino et al. 2018). Synthesized guide RNAs lead Cas9 to the target gene, which Cas9 then cleaves. The cell's DNA repair mechanism introduces a random base pair insertion into the genome, which may cause truncation, a premature stop codon, or some other alteration that leads to a loss of function in the target gene.

In order to introduce Cas9 into a host plant, a vector is required. This is where *Agrobacterium tumefaciens* enters the picture. *A. tumefaciens* is a soil

bacterium that infects plants with a tumor-inducing (Ti) plasmid that carries a segment of DNA which gets transferred to the host genome. This transferred DNA (T-DNA) carries genes that cause the development of tumors and the synthesis of opines, which are nitrogen-rich amino acid derivatives used by the bacterium as nutrients (Tzfira and Citovsky 2006). Van Eck et al. (2019) detail the process of *A. tumefaciens* mediated transformation in *Physalis pruinosa*: first, a binary vector is made from a modified Ti plasmid. This modified plasmid lacks the genes that cause tumors and opine production, and now contains the Cas9 sequence and guide RNAs. The vector is introduced to *A. tumefaciens*, which is then cultivated into colonies and grown in a liquid medium. Cotyledons that have been cut from seedlings are incubated in this medium, which infects them with *A. tumefaciens*. Infection leads to a high expression of Cas9 and guide RNAs in the explant tissue, increasing the chance of gene cleavage and editing. Plantlets are regenerated from this infected tissue, and then cultivated into fully grown transgenic plants called the T0 generation. Next, they are self-fertilized or, in the case of sterile phenotypes, fertilized with wild type pollen. Seeds are collected from the fruits, germinated, and cultivated into the T1 generation of plants. The Lippman Lab provided T1 seeds for this project, so I studied this generation for the effects of gene knockout in *P. grisea*.

The success of the CRISPR-Cas9 approach to knocking out gene function can be ascertained phenotypically and genotypically. My approach with *P. grisea*, therefore, was twofold. Firstly, I analyzed the phenotypes of T1 mutant lines for

any differences from the wild type. Specifically, I looked for differences in floral phenotypes, ICS, and pollen viability and germination. These features would help indicate whether a particular gene's activity had been altered. I then amplified DNA from the genic regions targeted by CRISPR-Cas9 in order to perform a genotypic analysis. Once this DNA was sequenced, I compared it to the WT sequences to see if there were insertions, deletions, splicing variants, or premature stop codons in the knockout DNA. These changes would indicate if CRISPR-mediated editing had occurred within any of the genes of interest. Following this, I would then correlate any observed changes in the sequences with any differences that I observed in the plant's floral or ICS morphology. This would help to elucidate the roles that any genes of interest may play in ICS development.

The Big Picture

My aim for completing this project was to complete a journey from *P. grisea* cultivation, to phenotypic analysis, to genotypic analysis. There were multiple points I needed to hit along the way: documenting floral and ICS morphology, extracting and amplifying DNA, and analyzing the sequenced results. I was most successful up until the final point, which was the greatest obstacle due most of the sequenced products being unreadable. There are multiple potential reasons for this, which I expand upon in the discussion. As for the other points along the journey, I documented a significant amount of morphological

data and successfully extracted and amplified DNA from all of the lines. Even though the sequencing did not come to great fruition, the data that I do have are informative and provide insight into the genetic underpinnings of ICS.

MATERIALS AND METHODS

Plant Cultivation

Seeds of the wild type *Physalis grisea* and the four *P. grisea* CRISPR-Cas9 edited lines were provided by Jia He, Lippman Lab, Cold Spring Harbor Laboratory. The *mpf3* and *ej2* lines are T1 lines and are not segregating for the CRISPR-edited alleles, whereas the *tag1* and *def* seed are F_2 populations from a cross between the wild type and the knockout mutant. F_2 populations were created for the *tag1* and *def* mutants because these mutants lack stamens or pollen, therefore wild type pollen was needed to produce seeds and maintain the lines.

Because *P. grisea* has a lifespan in the greenhouse of about four months, I performed two rounds of sowing and cultivation, one in the fall semester of 2022, and one in the spring semester of 2023. This allowed me to continue collecting data throughout the duration of the school year. For the first round, on 09/16/2022 I sowed 72 seeds per tray in pre-moistened Pro-Line C/GP Germinating Mix from Jolly Gardener. I then added ~2 mm of vermiculite on top, misted it with water until damp, and placed the trays on heating mats set to 70ºF. My watering routine was 2x per day, morning and afternoon. Once the plants grew large enough (on 10/28/2022), I transplanted them into individual 3.5’’ pots. Greenhouse staff began to fertilize them weekly starting 11/02/22 with Peters 20-20-20 fertilizer, which had 20% nitrogen, 20% phosphorous pentoxide, and 20% potassium oxide. They diluted this fertilizer to 13.5 ounces per gallon of water to achieve a target of 200 PPM of nitrogen.

I began cultivating a second batch of seeds on 02/17/2023. The main purpose of this round was to gather further morphological data on the various phenotypes, as well as document the maturation of ICS over time. Because of space constraints in the greenhouse, I sowed fewer seeds this time. I used the proportions of phenotypes that I had observed with the first batch of *P. grisea* to estimate the number of seeds needed for each line, with the intention of having multiple individuals to study and document. Following the same procedure as above, I sowed 12 seeds of the wild type, 24 seeds each of *mpf3* and *def*, and 30 seeds each of *tag1* and 30 *ej2*. I transplanted them into individual 3.5'' pots on 04/03/202. Greenhouse staff began to fertilize them weekly starting 3/23/2023 with the same fertilizer as used previously.

Documenting Phenotypic Traits

To record the variety of phenotypes present in each line, I selected plants that presented floral and ICS mutations, as well as wild type individuals to create a reference for normal morphology. In order to document flower and ICS morphology in each line, I collected samples at different stages of growth: closed bud, open bud, flower, young calyx, and older calyx. The plants did not reach the final level of maturity of a desiccated calyx by the end of each semester, so I was unable to observe this stage.

I carefully removed flowers and calyces of interest, ensuring that each plant still had at least one flower left when possible. This was done to maximize

the potential for fertilization and setting fruit. I photographed the samples using an Olympus SZ40 dissecting microscope fitted with a camera, including a ruler in each image to provide scale in millimeters. The mature calyces were too large to be photographed under the dissecting microscope, so I utilized a smartphone camera in those cases.

In addition to photographing the flowers, I created a catalog which documented the floral characteristics of every individual plant (Data in Appendix, Tables A1-A5). Each plant was assigned a name consisting of its genotype (WT, *mpf3*, *def*, *tag1*, *ej2*) and a number. In the catalog, I recorded the presence/absence of flowers, any observed floral mutations, the whorl(s) in which they occurred, the presence/absence of ICS, and any post-fertilization abnormalities. I also included notes for further detail, and whether or not tissue or pollen had been harvested for DNA/fertility analyses.

Iodine Potassium Iodide (IKI) Stain

To test for pollen starch content, I performed an IKI stain (Edwardson and Corbett 1961). I gently scraped the anthers onto glass microscope slides to collect the pollen grains from each line, and then dropped IKI onto each slide. After one minute, I added a coverslip and observed the pollen grains under a compound light microscope at 50x magnification. Darker pollen after staining indicates a higher starch content.

Germination Test

To test for pollen tube development, I incubated pollen grains from each line in a germination medium (Protocol kindly provided by Iacopo Gentile, Lippman Lab. Data in Appendix, Table A6). I observed the pollen under a compound light microscope at 10x and 50x magnification after 4 hours, 8 hours, and 12 hours. The presence of pollen tubes indicates germination.

DNA Extraction and Analysis

I harvested young leaf tissue with forceps (1-2 leaves per plant), storing it in 1.5 ml microfuge tubes on ice to prevent degradation. Young leaf tissue was used due the ease with which it can be ground down in comparison to older, tougher tissue. In order to ascertain the best method for DNA extraction, I extracted wild type DNA from the leaf tissue using a CTAB protocol from the Lippman Lab and the Qiagen DNeasy Plant Pro Kit. After this first round of extractions, I analyzed the quantity of the DNA from these two methods using the NanoDrop spectrophotometer. I then performed gel electrophoresis of the samples to estimate both the quality and quantity of the DNA. Aliquots of uncut lambda DNA were used to estimate the quantity of DNA in each sample. Moving forward, I used the CTAB protocol for DNA extraction due to its greater yield. I collected leaf tissue samples from 10 individual plants per observed phenotype in each line. When there were fewer than 10 plants exhibiting a phenotype, I collected as many samples as possible. I stored these tissue samples over winter

break in a -80ºC freezer, and extracted their DNA in small batches. After extraction, I checked the quality and quantity of the DNA using the NanoDrop spectrophotometer and gel electrophoresis. I stored the DNA samples in a -20ºC freezer for future use.

Polymerase Chain Reaction

In preparation for PCR to amplify the edited regions, I selected forward and reverse primer sequences designed for each knockout line by He *et al.* (2023) (Table 1). I utilized CSHL's online Sollab index to obtain the sequences of each wild type gene (*MPF3, DEF, TAG1,* and *EJ2*). I input these sequences into the computer platform Benchling, in which I did the following: determined the position of each primer within its respective wild type gene sequence; ensured that the primer pairs flanked the regions of each gene, determined the GC content and melting temperature of each primer, and determined the expected product size. The primers were ordered from Integrated DNA Technologies.

Table 1. Forward and reverse primers for *P. grisea* designed by He et al. (2023).

Gene Name	**Primer Name**	**Primer Sequence 5' to 3'**	**Expected Product Size (bp)**
PgMPF3	Pg.MPF3_F1	CTGTCTCTCTACCTTAAAAATTACATC	350
PgMPF3	Pg.MPF3_R1	GAGACAATAAACTTGGAAAGTCA	
PgTAG1	Pg.TAG1_F1	GGATTCTAATCCAAGTTCTTTAGC	472
PgTAG1	Pg.TAG1_R1	AGATTAGAACTCTCCGTAACCA	
PgDEF	Pg.DEF_F1	CCAAACAAACAGGCAAGTGA	455
PgDEF	Pg.DEF_R1	CCAAGTTTGAACATGAAGTGTGTA	
PgDEF	Pg.DEF_F2	GTTGAACTGTATGGACTTCCTAT	400
PgDEF	Pg.DEF_R2	ATAAGCTTCAGAGAATCGTCCA	
PgEJ2	Pg.EF2_F1	ACCCACCTTGCACTGGAAAA	306
PgEJ2	Pg.EF2_R1	TCAAAATGTCCTTTAATCTTGTAGC	

I then performed a test round of PCR on a subset of my DNA samples to ascertain the best concentration of template DNA to use. This subset consisted of 4 *mpf3* samples, 4 *tag1* samples, and 4 WT samples. These samples represented a range of NanoDrop results, from 81ng/µl to 521ng/µl, to reflect the variety of quality found across the board in the DNA I had extracted. I ran PCR reactions with three different concentrations of template DNA: undiluted, 1/5 dilution, and 1/10 dilution. The components for each reaction are listed in Table 2.

After performing PCR, I used gel electrophoresis to see if amplification was successful and to determine which amounts of template DNA yielded the best results. The undiluted DNA had the brightest bands, which meant that they had the most amplified DNA. I therefore used undiluted template DNA when

preparing my subsequent PCR reactions. Following this, I performed the same PCR procedure on all of my *def* samples, comparing the two sets of *def* primers to determine which one was more effective. The first set of primers (Pg.DEF_F1 and Pg.DEF_R1) yielded better results, so I used them moving forward.

The second round of PCR was to test if I could successfully amplify DNA from all of my samples, check if the products were of the expected size (Table 1) and see whether individuals were homozygous (one band) or heterozygous (two bands). Using the same protocol as above, I performed PCR on all of the *mpf3*, *tag1*, and *ej2* samples and ran the products through gel electrophoresis.

The third round of PCR was to create products that could be sent for sequencing. In order to have enough PCR product volume, I used 4µl of template DNA and doubled the amount of the other components needed for each reaction. I ran a portion of these products through gel electrophoresis as well to assess the aforementioned criteria.

Table 2. Reaction components used for PCR.

Component	**Volume**
2x Master Mix from New England BioLabs	12.5 µl
10 µM Forward Primer	0.5 µl
10 µM Reverse Primer	0.5 µl
ddH_2O	10.5 µl
Template DNA	1 µl

Table 3. PCR profile for amplifying *mpf3, tag1,* and *ej2*.

Step of PCR Reaction	**Temperature Profile**	**Time Profile**	**Number of Cycles**
Initial denaturation	95ºC	1 minutes	1
Denaturation	95ºC	30 seconds	30
Annealing	48ºC	30 seconds	
Extension	68ºC	1 minutes	
Final Extension	68ºC	5 minutes	1
Hold	4ºC	-	-

Table 4. PCR profile for amplifying *def*.

Step of PCR Reaction	**Temperature Profile**	**Time Profile**	**Number of Cycles**
Initial denaturation	95ºC	1 minutes	1
Denaturation	95ºC	30 seconds	30
Annealing	50ºC	30 seconds	
Extension	68ºC	1 minutes	
Final Extension	68ºC	5 minutes	1
Hold	4ºC	-	-

Gel Electrophoresis

My first PCR products were run on a 1% agarose gel prepared in a 1X TAE buffer containing ethidium bromide at 50mA (~20V) for one hour. To measure the size of the DNA molecules, I included two DNA ladders : a 100bp

ladder, and a 1Kb ladder. Because the bands on the ladders were fuzzy, I increased the concentration of agarose to 1.2% for subsequent gels, which remedied the issue. The electrical current, ladders, and run time remained the same when I repeated this procedure. I reused gels a maximum of three times before creating a new one.

Purification of PCR Products

I selected 1-3 amplified DNA samples per phenotype to be purified. My focus was getting sequences of the mutated phenotypes, so I made sure to choose 3 of each. These samples are listed below in Table 5. To purify the PCR products for sequencing, I followed the PCR clean-up protocol from NucleoSpin. Because the *def* products yielded two distinct DNA bands from gel electrophoresis, I attempted to extract the DNA from the gel for purification according to another protocol from NucleoSpin. Unfortunately, due to errors in cutting the gel, I was unable to successfully complete this procedure. As a result, I do not have sequenced *def* DNA.

Table 5. Samples of amplified DNA chosen for purification and subsequent sequencing.

Line	**Phenotype**	**Sample #**
mpf3	Short Calyx	*mpf3-3*
		mpf3-5
		mpf3-8
	Normal Calyx	*mpf3-16*
tag1	Stamens into Petals	*tag1-8*
		tag1-39
		tag1-60
	Normal Flower	*tag1-26*
		tag1-51
ej2	Too Many Sepals (TMS)	*ej2-1*
		ej2-19
		ej2-33
	Sepals into Petals (SIP)	*ej2-20*
		ej2-25
		ej2-29
	TMS+SIP	*ej2-22*
		ej2-24
		ej2-41

Analysis of DNA Sequences

Purified PCR products and their accompanying primers were sent to Eurofins Genomics for sequencing. I received two sequences per sample: one from the forward primer, and one from the reverse primer. This allowed me to study the gene sequence from both ends. Upon receiving the sequence files, I utilized SnapGene Viewer to assess each electropherogram. For any unreadable sequences, I referenced the Eurofin Genomics DNA Sequencing Results Guide to determine the possible cause. For any readable sequences, I aligned them with the corresponding WT gene in Benchling and looked for any mutations such as deletions, insertions, or substitutions. To figure out the location of a mutation in relation to the introns and exons of the gene, I referenced the sequenced *P. grisea* genome in Cold Spring Harbor Lab's Sollab server. The gene identification codes in the server are as follows (Table 6) (He et al. 2023):

Table 6. Gene ID codes in the Sollab server for each *P. grisea* gene of interest.

Gene Name	***P. grisea* Gene ID**
PgMPF3	Phygri12g018350
PgDEF	Phygri11g018450
PgTAG1	Phygri02g019350
PgEJ2	Phygri03g020760

RESULTS

Floral and ICS Morphology

In total, I cataloged the phenotypes of 257 plants across all of the lines: 32 wild type, 52 *mpf3*, 60 *def*, 68 *tag1*, and 45 *ej2*. While I had originally sowed 72 seeds per line, this number dropped throughout the cultivation process due to lack of germination or the death of seedlings and young plants. The catalog of phenotypes documented the following features: the presence/absence of flowers, any observed floral mutations and the whorl(s) in which they occurred, the presence/absence of ICS, and any post-fertilization abnormalities such as deformed or pigmented ICS. After making these observations, I chose a selection of key mutations to focus on, which are expanded upon below.

Wild Type

The wild type (WT) *P. grisea* flowers exhibited five fused pale yellow petals with dark purple pollen guides, five stamens closely huddled around a carpel (Figure 7a), and five green, hairy sepals on the underside (Figure 7b). The inflated calyx of the wild type was about 30mm long, with the ends of the modified sepals tapering together to create a narrow tip (Figure 7c).

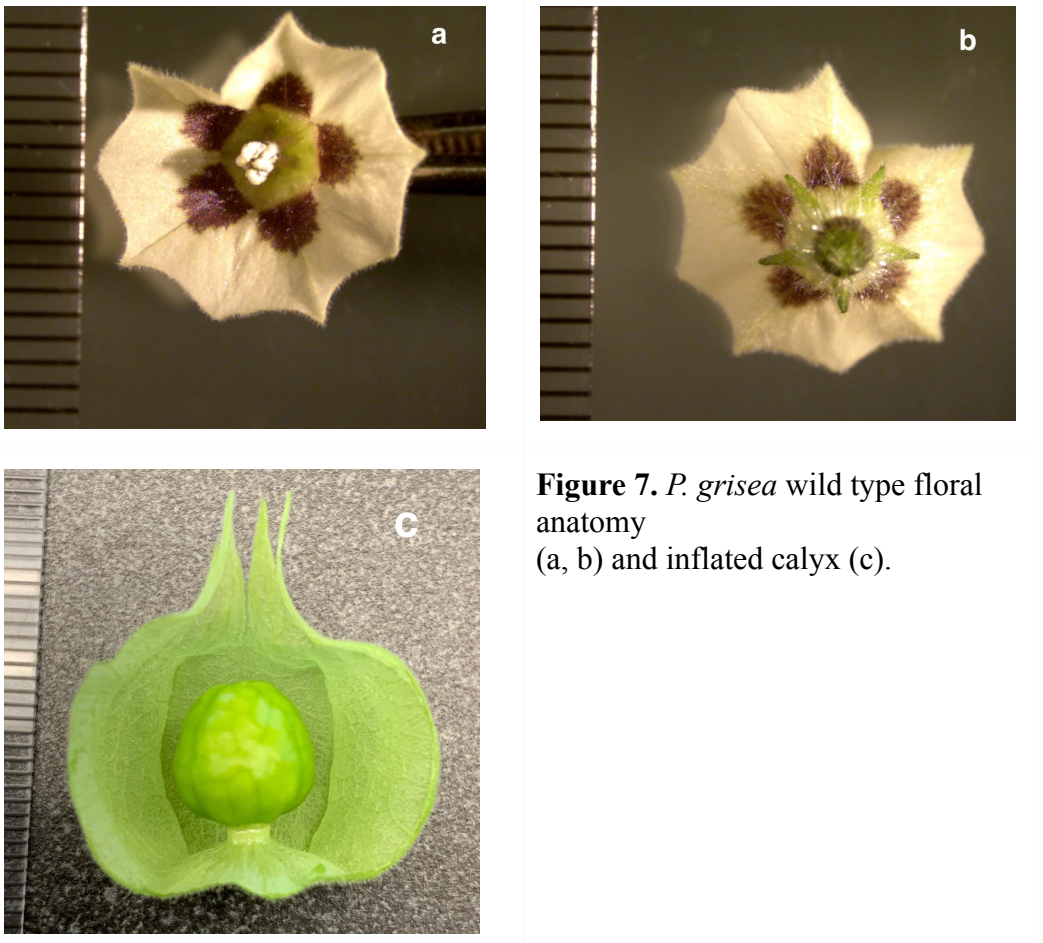

Figure 7. *P. grisea* wild type floral anatomy
(a, b) and inflated calyx (c).

mpf3

All of the plants in the *mpf3* line exhibited flowers with the WT phenotype, and about 20% of them also had normal calyces. However, approximately 80% of plants had calyces that were shorter than the WT (Table 7). Their tips did not taper together; instead, they were stunted with a squat appearance (Figure 8). The ratio of WT to mutation was about 1:4, which suggests that the short ICS phenotype is a dominant trait.

Table 7. Percentage of plants exhibiting each phenotype in the *mpf3* line (n = 52).

Phenotype	**Percentage of Plants**
Normal	20%
Short ICS	80%

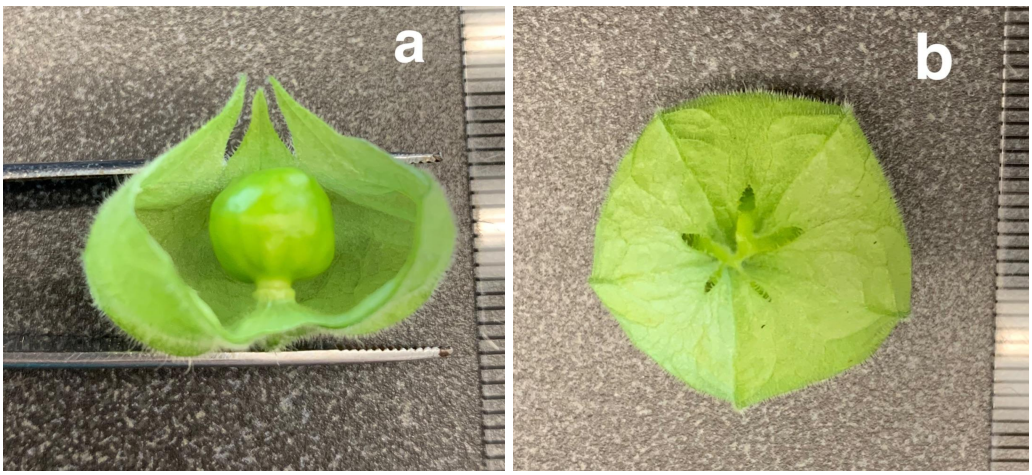

Figure 8. Short ICS exhibited by the *mpf3* line.

def

The *def* line, which was an F_2 population, exhibited both flowers and ICS with the WT phenotype. Only 5 plants were observed with a subtly different phenotype: dark brownish-purple pigmentation of the calyx tissue (unfortunately I was not able to photograph this at the time). This amounts to 8% of the plants in the *def* line presenting a pigmentation mutation (Table 8), which suggests that the dark calyx may be a recessive trait.

Table 8. Percentage of plants exhibiting each phenotype in the *def* line (n = 60)

Phenotype	**Percentage of Plants**
Normal Calyx	92%
Dark Calyx Pigmentation	8%

tag1

The *tag1* line, which was an F_2 population, exhibited the normal development of the inflated calyx, with a combination of WT and flowers with a change in floral organ identity. These mutants had petaloid organs in Whorl 3 instead of stamens. About 26% of plants in the *tag1* line exhibited this phenotype, with an approximate 3:1 ratio of WT to the mutation (Table 9). This suggests that it is a recessive trait. I observed a spectrum within this particular phenotype, with three levels of extremity in the transformation of the stamens. The mildest phenotype (Figure 9a) exhibited a combination of stamens and petaloid organs. All of these organs remained clustered closely around the carpel, which is typical

for stamens in the wild type. In the second level of mutation (Figure 9b), the tissue of the stamens was more petal-like, and no normal stamens remained. However, ridges of anther-like tissue remained on some of the petaloid organs. In addition, the petaloid organs began to fan out more from the carpel, arching away in a similar curve to the true petals. The most extreme mutated phenotype (Figure 9c, Figure 9d) resembled petals the most. While the bases of these organs were narrow and stamen-like, the tissue that emerged was virtually identical in color and texture to petals. Very distinct pigmentation resembling pollen guides was present at the base of each organ. Not all of the organs fanned out equally, but they did arch away from the carpel more significantly than in the mutants with the mildest phenotype.

Table 9. Percentage of plants exhibiting each phenotype in the *tag1* line (n = 68).

Phenotype	**Number of Plants**
Wild Type	74%
Stamens into Petals	26%

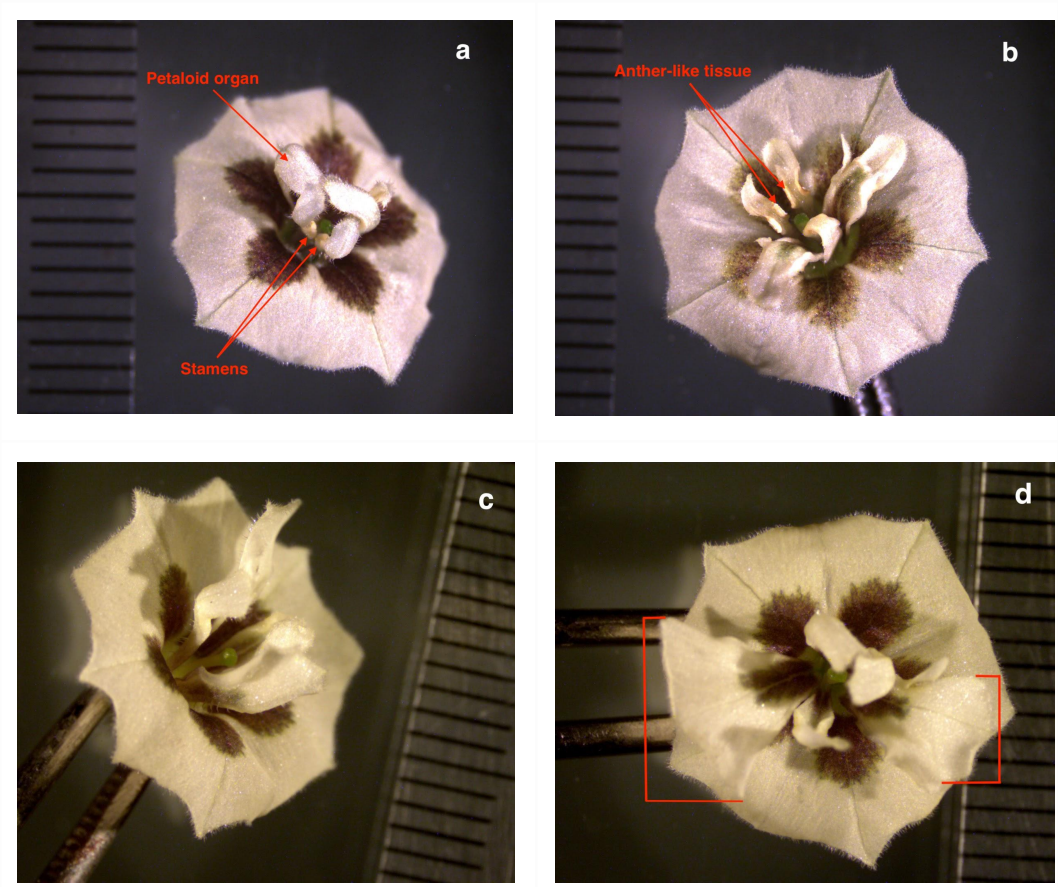

Figure 9. First (a), second (b), and third (c, d) levels of mutation of stamens into petaloid tissue in *tag1*. Brackets in (d) indicate the size of petaloid organs.

ej2

The *ej2* line showed a striking array of floral mutations, and none of the plants exhibited the WT phenotype. I identified four types of mutation, which in most cases occurred in combination with one or more of the other mutant phenotypes (Table 10). These mutations can be broken down into two main categories: A change in organ identity, and a change in organ number.

Sepals into Petals

The transformation of sepals into petals (SIP) occurred in 39 out of 46 plants, or 85% (Table 10). It mostly presented in flowers alongside changes in organ number— either too many sepals (TMS), a difference in petal number (DIP), or both. Only 3 plants (7%) exclusively exhibited the SIP phenotype (Table 10). There was a high degree of variability in how this mutation manifested. In some cases, the sepals took on a muddled, petaloid appearance. Some parts of the tissue were still green, which blended into tissue that was virtually indistinguishable from fused petals (Figure 10). In other cases, though the sepals still had petaloid tissue, they possessed a narrow and pointed shape that more closely resembled normal sepals (Figure 11). In every case of SIP, the development of ICS was disrupted, which can be seen in Figure 12. This flower had been fertilized and set a fruit, but only half of the sepals were growing into an inflated calyx. The other half were petaloid, with the true petals still attached.

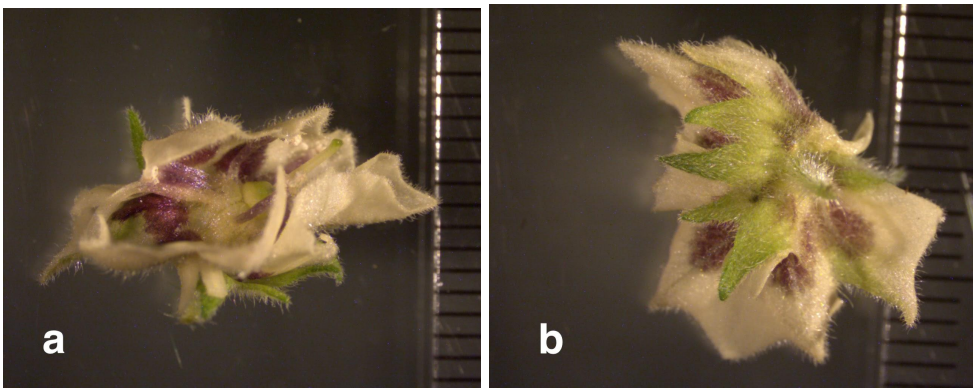

Figure 10. An example of SIP in *ej2*. The green sepals transform seamlessly into petaloid tissue.

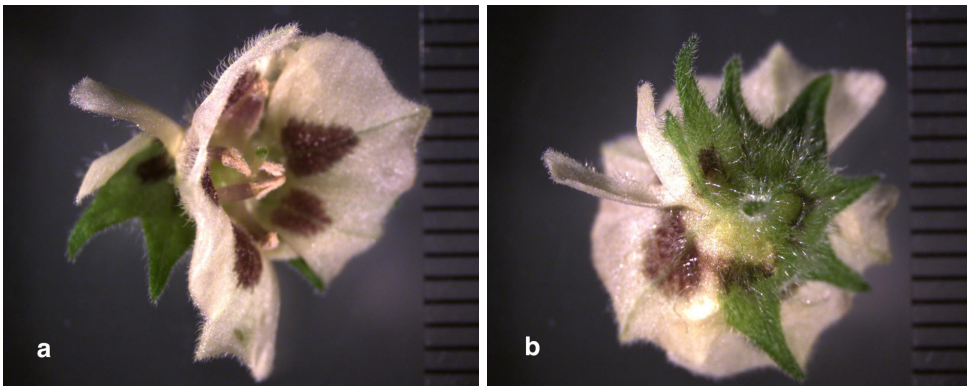

Figure 11. An example of SIP in *ej2*. The green sepals transform seamlessly into petaloid tissue.

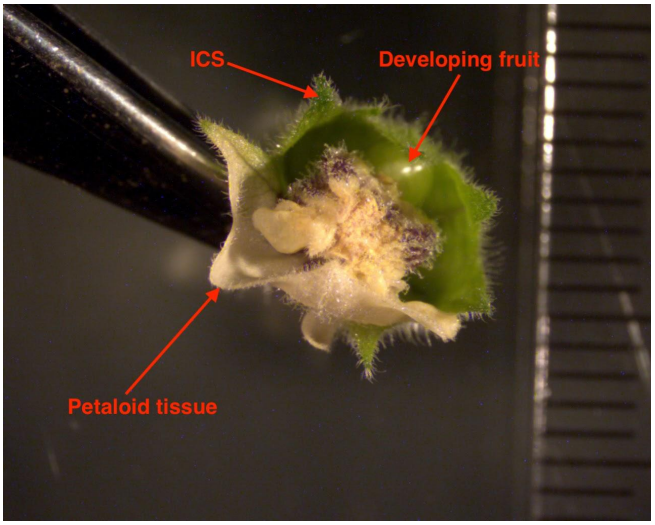

Figure 12. The disruption of ICS in the SIP phenotype.

Too Many Sepals

Flowers with more than 5 sepals (TMS) occurred in 40 out of 46 plants, or 86% (Table 10). The number of sepals in this phenotype ranged from 6 to 8, and they varied in size on each flower. They were also accompanied by an overgrowth of sepal tissue at the base of the bud or flower (Figure 13). The TMS phenotype disrupted ICS development, as demonstrated in Figure 14. In this example, the 7 sepals expanded abnormally to create a bulbous, open inflated calyx. 90% of the plants with TMS also had the SIP mutation, and TMS presented alongside DIP as well (Table 10).

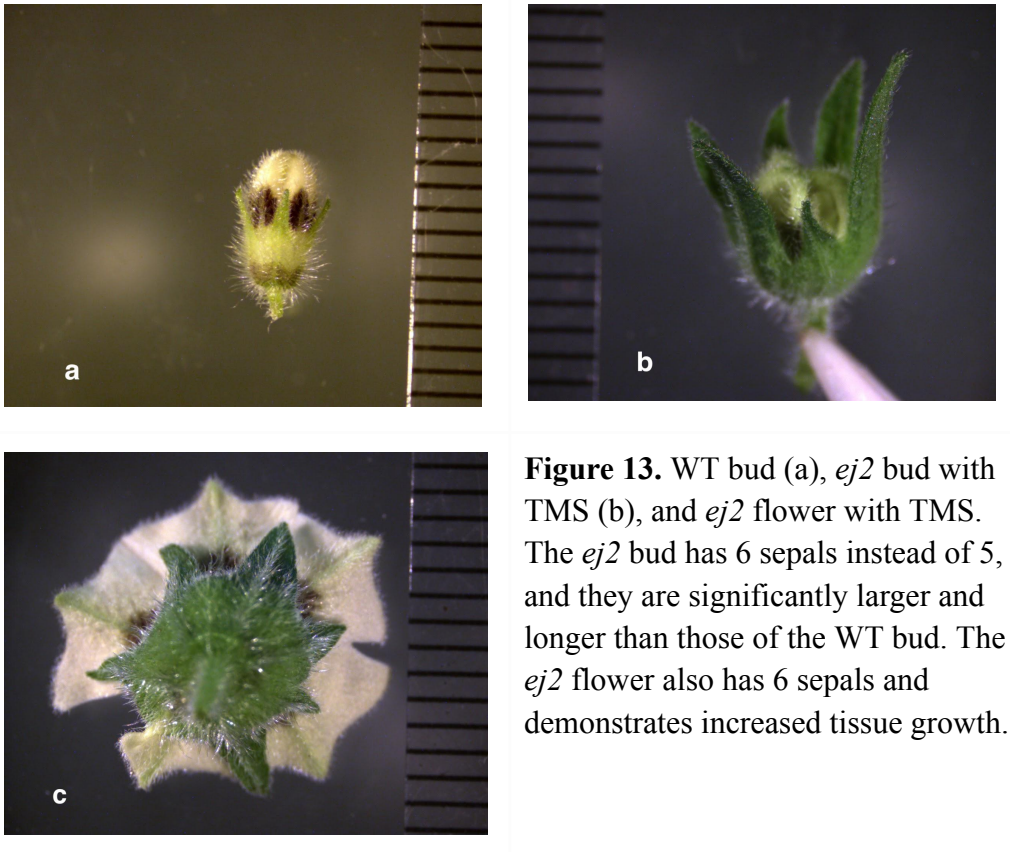

Figure 13. WT bud (a), *ej2* bud with TMS (b), and *ej2* flower with TMS. The *ej2* bud has 6 sepals instead of 5, and they are significantly larger and longer than those of the WT bud. The *ej2* flower also has 6 sepals and demonstrates increased tissue growth.

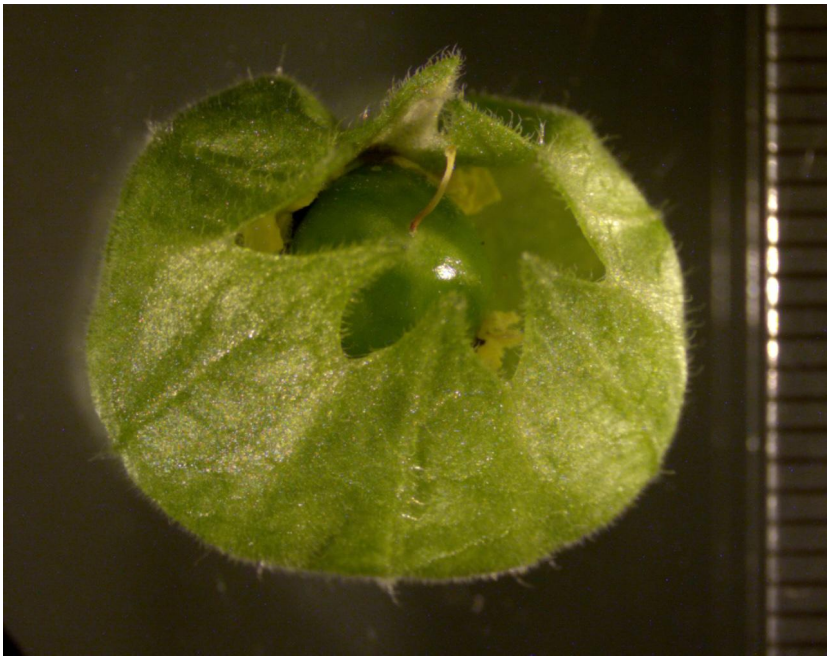

Figure 14. A malformed *ej2* inflated calyx made out of 7 sepals.

Difference in Petal Number

Flowers with differences in petal number (DIP) had either 3, 4, or 6 petals. 63% of all the *ej2* plants had this phenotype, which was observed most frequently in flowers alongside the SIP+TMS combination (Table 10). An example of DIP can be seen in Figure 15, a flower with only 4 petals. In this flower, the petals were not entirely fused together. However, when cataloging phenotypes I also observed flowers with DIP whose petals were fused. I was unable to capture a photograph of this under a dissecting microscope at the time, but I did obtain a picture with my smartphone (Figure 16).

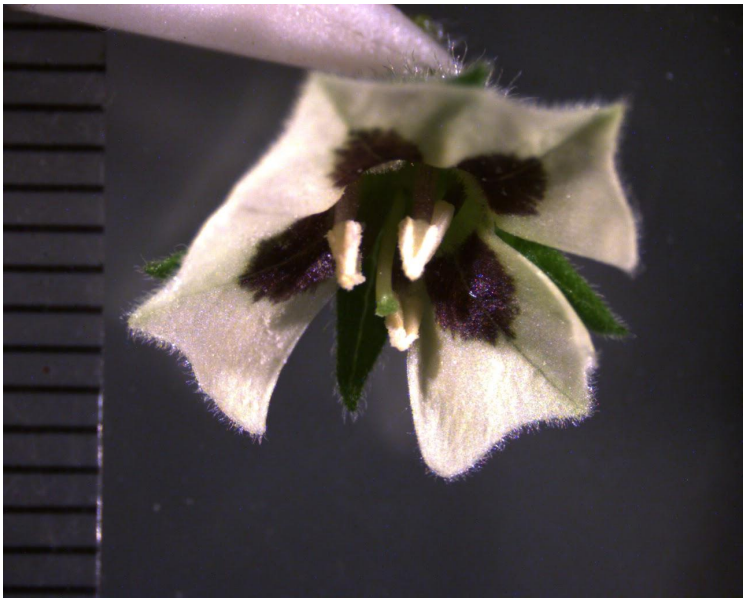

Figure 15. An *ej2* flower with only 4 petals. One petal is separated from the rest.

Figure 16. An *ej2* flower with 4 fused petals.

Table 10. Percentage of plants exhibiting each phenotype in the *ej2* line (n = 46)

Phenotype	**Percentage of Plants with Phenotype**
Wild Type	0%
Sepals into Petals (SIP)	7%
Too Many Sepals (TMS)	13%
SIP+TMS	78%
Difference in Petal Number (DIP)	63%
DIP and SIP	0%
DIP and TMS	9%
DIP and SIP+TMS	50%

Pollen Viability - IKI Stain

It is difficult to make a confident conclusion with regards to starch content in the pollen across the lines. While I was able to observe multiple pollen grains and capture images, the overall sample size was rather small. In addition, I could not get pollen from the *mpf3* line, because all of the plants had only calyces at the time.

Wild Type

Most of the WT pollen was dark in color, indicating it was well saturated with IKI (Figure 17a). This indicates high starch content, which is a sign of viability (Datta et al.). However, I observed some other pollen grains that were pale in color (Figure 17b), which suggested that they were less viable.

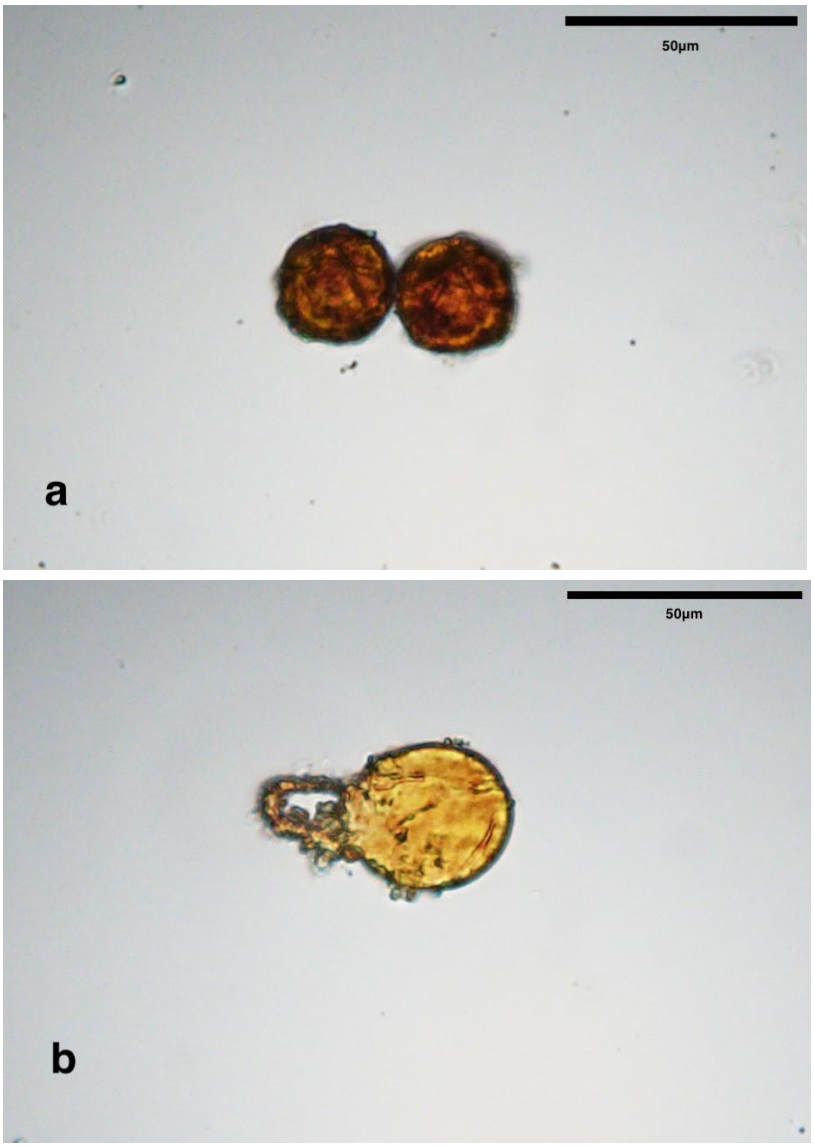

Figure 17. Pollen highly saturated with IKI (a) and minimally saturated with IKI (b).

def

The pollen from the *def* sample exhibited similar variability in starch content as the wild type; most pollen was dark, some was light (Figure 18).

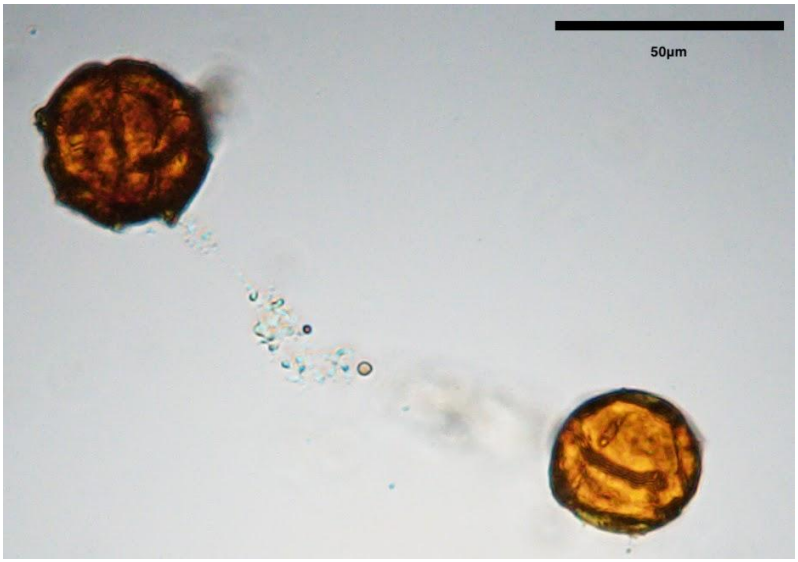

Figure 18. Two grains of *def* pollen, one more saturated with IKI than the other.

tag1

The pollen from the *tag1* sample was consistently light in color, indicating low starch content and low viability (Figure 19).

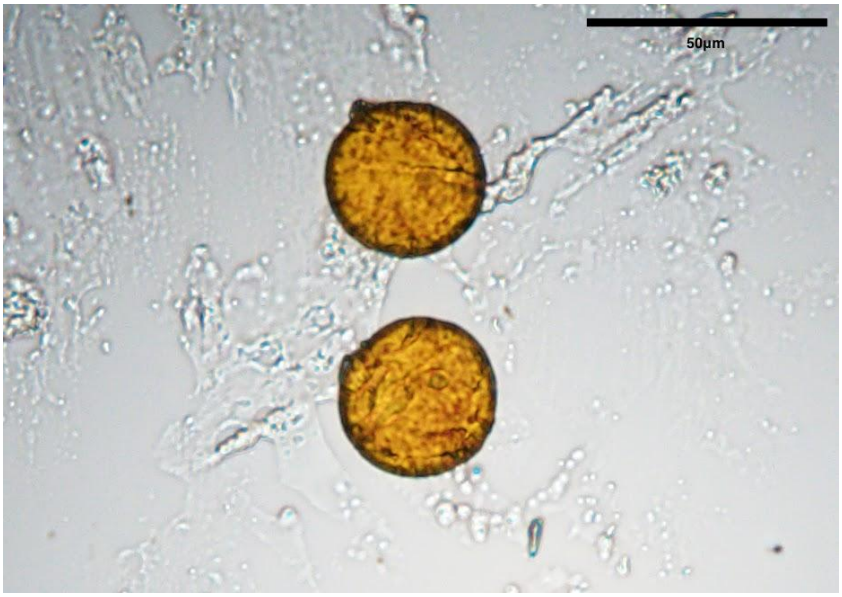

Figure 19. Two grains of *tag1* pollen with low IKI saturation.

ej2

The *ej2* pollen was similar to *tag1* in that it was consistently pale. In addition, many of the grains had an especially crumpled, deformed appearance compared to the pollen in the other lines (Figure 20).

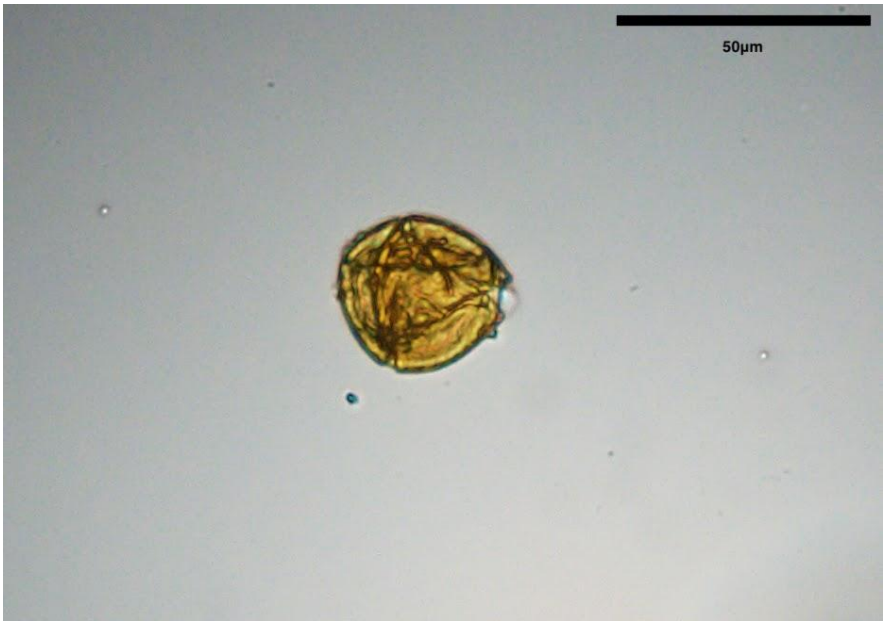

Figure 20. An *ej2* pollen grain with low IKI saturation and a deformed appearance.

Pollen Viability - Germination

The results of this test were largely inconclusive. Although I observed a handful of pollen grains with pollen tubes, most of it had not germinated across all of the lines I tested. In addition there was very little pollen in the samples overall.

Comparison of DNA Extraction Protocols

Upon analysis of my first DNA samples with the NanoDrop, the CTAB protocol yielded an average of 441 ng/µl of DNA. The Qiagen kit yielded far less, with an average of 144 ng/µl of DNA. Because the CTAB protocol yielded a greater quantity and quality of DNA than the Qiagen kit, I continued with this method for the rest of my DNA extraction.

PCR Amplification of Gene Sequences

These results are grouped into three sections based on the three rounds of PCR that I performed (see Materials and Methods). Round 1 was to determine the optimal template DNA concentration for future reactions, as well as determine the optimal primers for *def*. Round 2 was to amplify all of the DNA from each line, check product size, and identify any heterozygosity. Round 3 was to amplify DNA specifically for sequencing. The results of each gel are read from left to right, and contain the following: a 100 bp ladder, a 1kb ladder, and a series of amplified DNA bands from the PCR products. The final well at the end of a series

may show a faint band of primer-dimer from the blank sample that was loaded without DNA. This is a byproduct of the PCR reaction.

The results from Round 1 demonstrate that undiluted template DNA yielded the highest amount of amplification (Figure 21). In addition, Figure 22 shows that the first pair of *def* primers, Pg.DEF_F1 and Pg.DEF_R1, yielded products with the largest molecular size.

The results from Round 2 demonstrate successful amplification for the majority of samples from *mpf3* (Figure 23), *tag1* (Figure 24), and *ej2* (Figure 25). The sizes of the PCR products for each line were approximately the expected number of base pairs for each set of primers (Table 1). The *mpf3* products were about 300bp, the *tag1* products were about 500bp, and the *ej2* products were about 300bp.

The results from Round 3 demonstrate successful amplification once again (Figures 26, 27, 28, and 29), although the *tag1* bands were fainter than those from Round 2. The Round 3 results also clearly demonstrate heterozygosity in the *def* F2 population, which has two distinct bands of different sizes (Figure 28).

Round 1

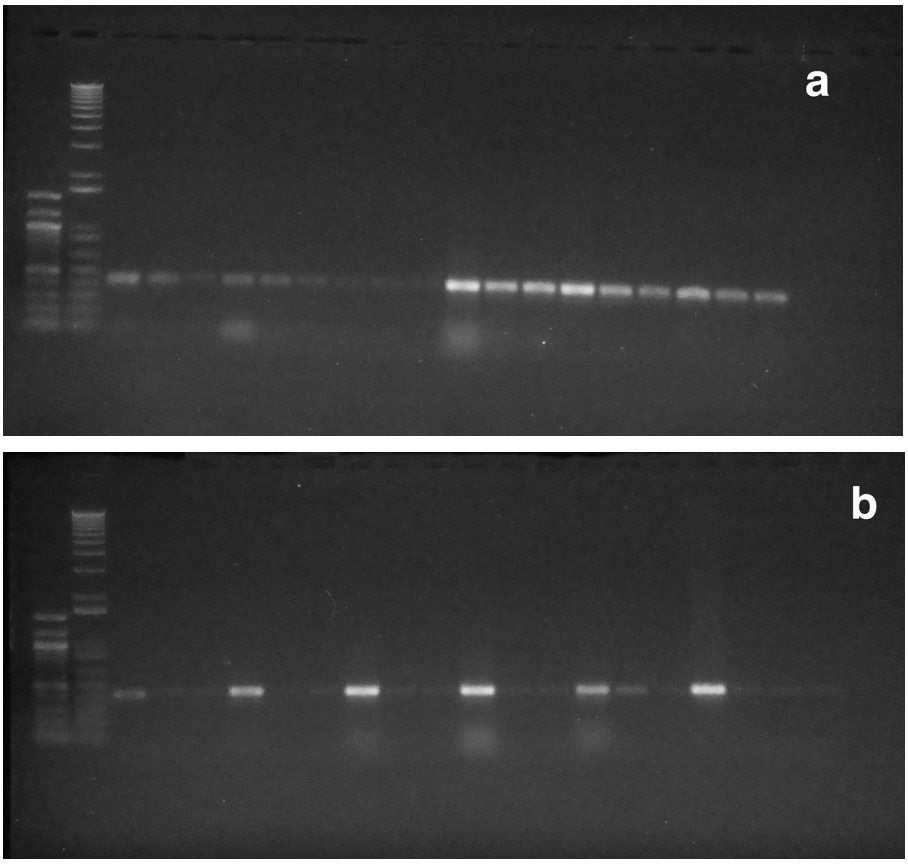

Figure 21. Gel results testing different dilutions of template DNA for PCR. Starting from the left, every set of three bands corresponds to one sample amplified with the three different dilutions: first band = undiluted, second band = 1/5 dilution, and third band = 1/10 dilution. The first bands in each set were the brightest, indicating that undiluted template DNA yielded the highest amount of amplification. This is especially clear in the bottom row (b), where the first band of every trio is significantly brighter than the second and third bands.

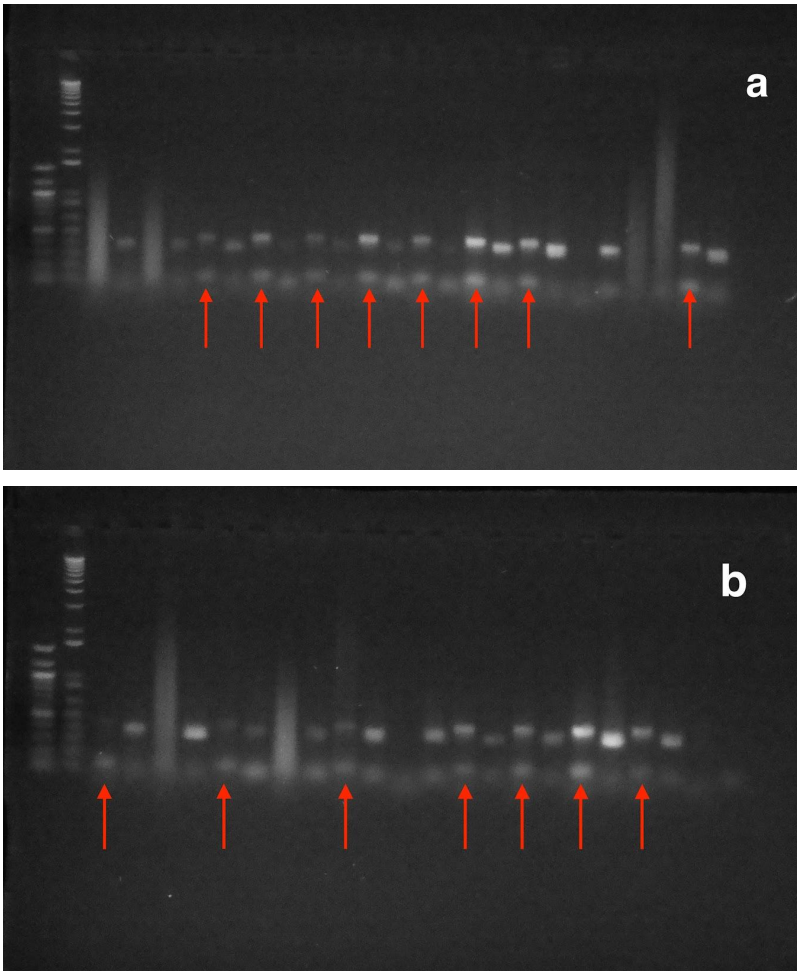

Figure 22. First row (a) and second row (b) of gel results for *def* PCR products, first round of PCR. Samples amplified with Pg.DEF_F1 and Pg.DEF_R1 yielded products of about 400bp in size, as indicated with red arrows. Samples amplified with Pg.DEF_F2 and Pg.DEF_2 yielded products that were approximately 100bp smaller.

Round 2

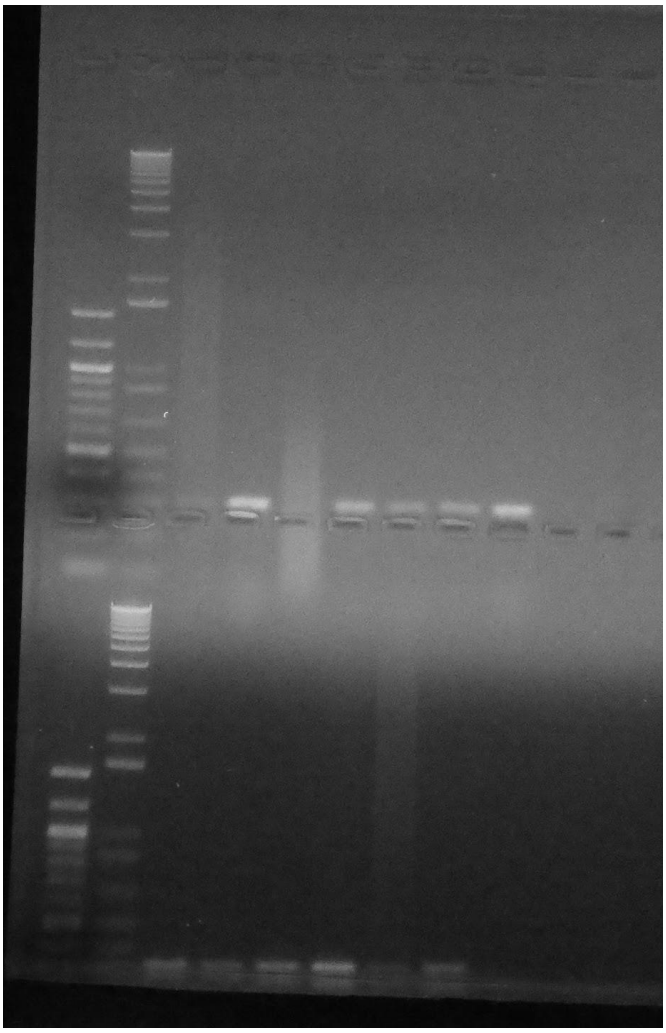

Figure 23. Gel results of *mpf3* PCR products, second round of PCR. Indicates successful amplification of products from all samples. Size of products is approximately 300bp.

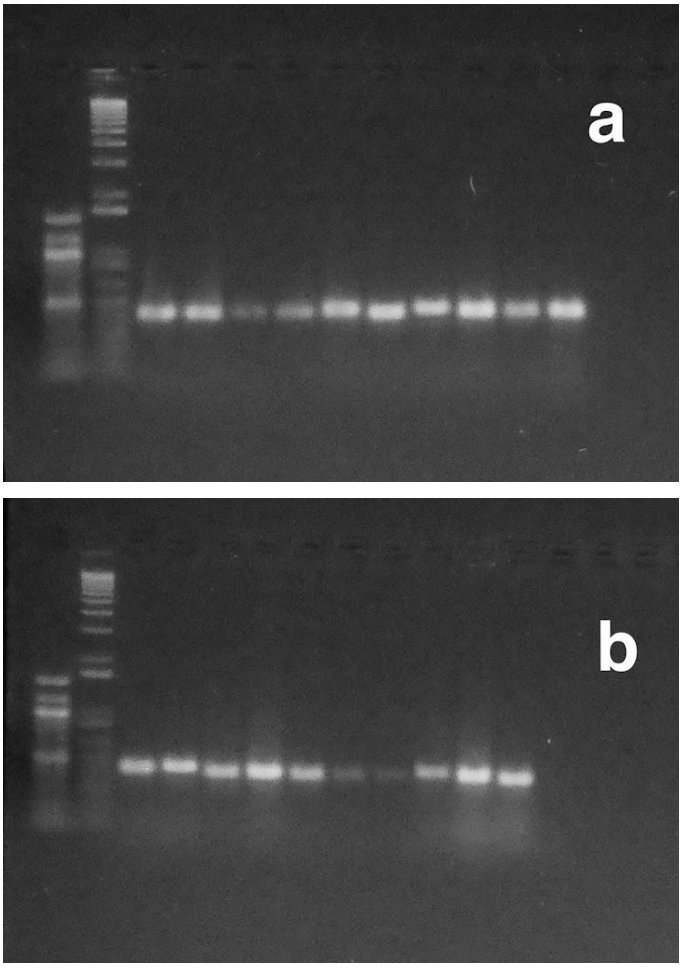

Figure 24. First row (a) and second row (b) of gel results for *tag1* PCR products PCR. Indicates successful amplification of products from all samples. Size of products is slightly under 500bp.

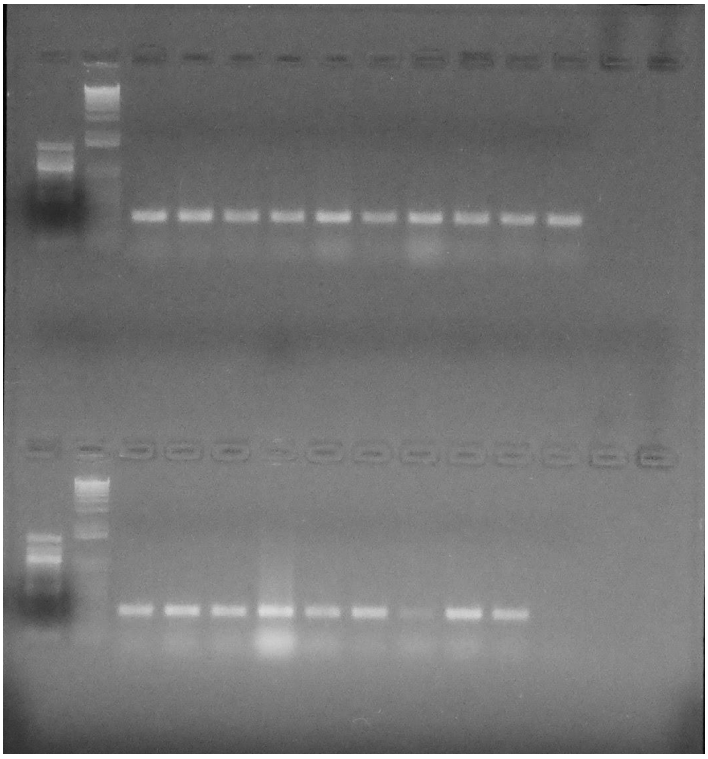

Figure 25. Gel results of *ej2* PCR products. Indicates successful amplification of products from all samples. Size is approximately 300bp.

Round 3

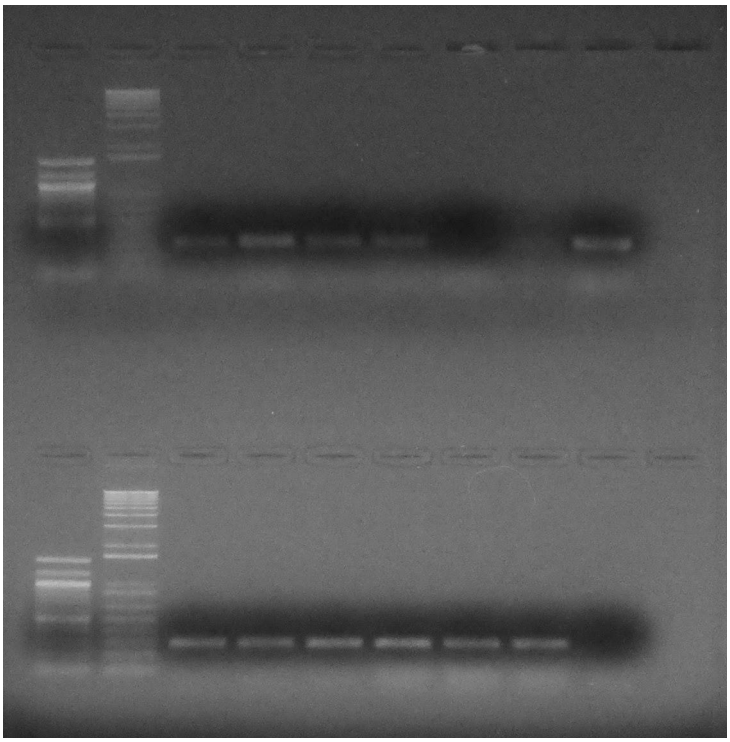

Figure 26. Gel results of *mpf3* PCR products. Indicates successful amplification from all but one individual. Size is approximately 300bp.

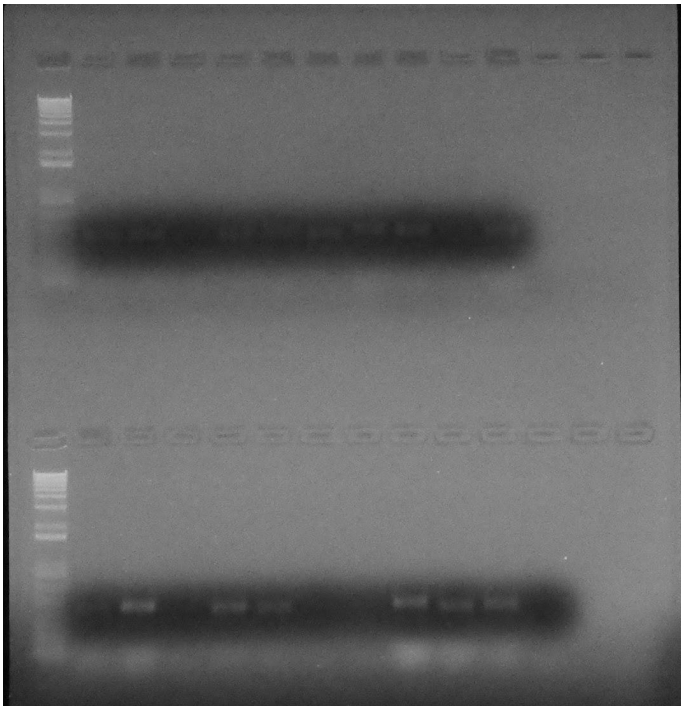

Figure 27. Gel results of *tag1* PCR products, third round of PCR. The bands here are fainter than in the previous *tag1* gel (Figure 24), indicating that less DNA was amplified. However, most of the bands do show up to an extent, which means that the PCR still worked.

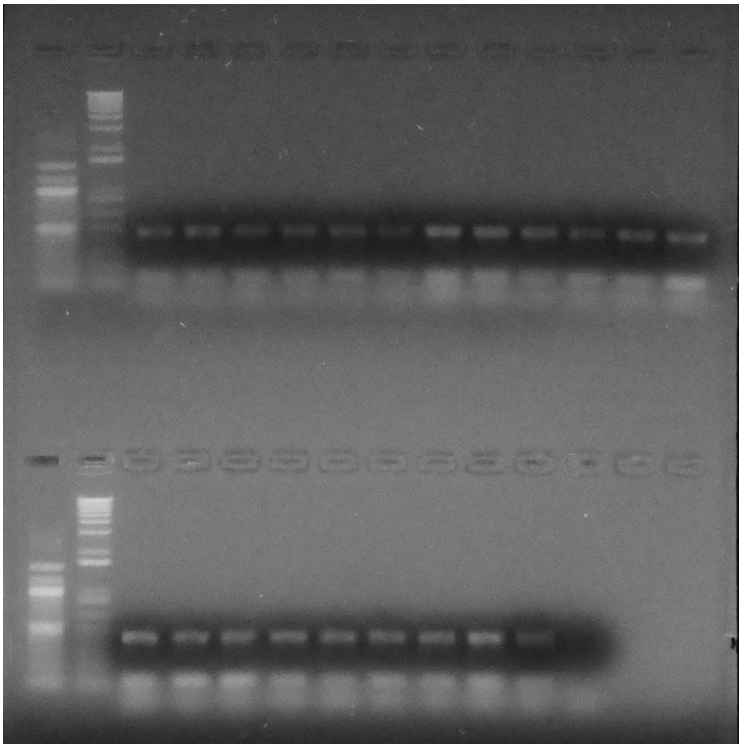

Figure 28. Gel results of *def* PCR products. Indicates successful amplification products from all samples. A clear second band indicating heterozygosity is present in each sample. Size of the first band is approximately 450bp, and the second band is somewhere between 100bp and 200bp.

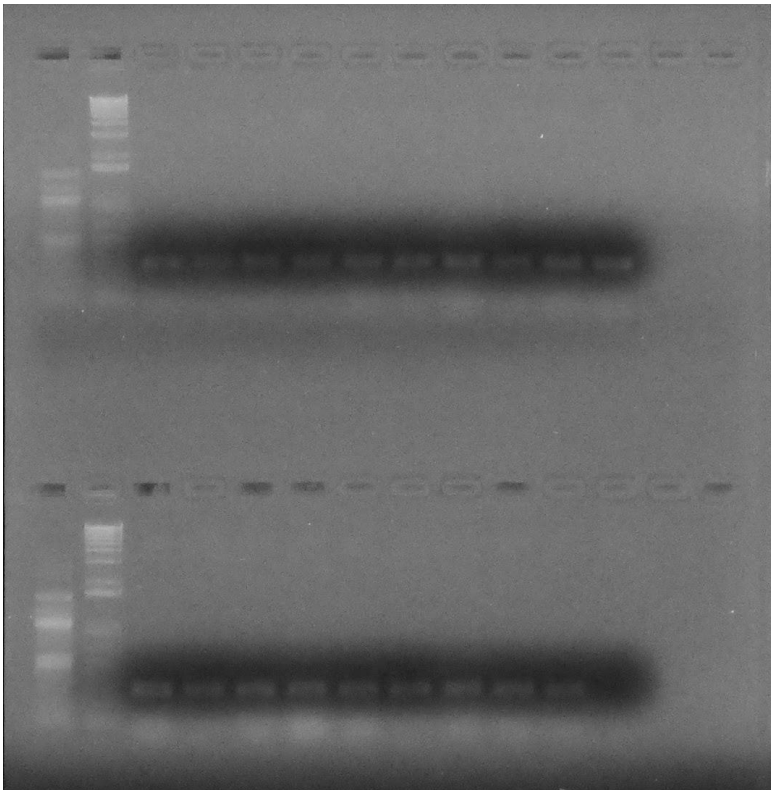

Figure 29. Gel results of *ej2* PCR products, third round of PCR. Indicates successful amplification of all products.

Sequenced PCR Products

In total, I received 38 sequences from Eurofins Genomics— two sequences per sample, one from the forward primer and one from the reverse primer. Out of those 38 sequences, 35 had completely unreadable electropherograms, which signifies that the sequencing failed. The main recurring issues in these electropherograms were unidentified bases and multiple sequence signals (Figure 30). According to the Eurofins Genomics DNA Sequencing Results Guide, multiple sequence signals can be caused by poor quality primer, which is the likely culprit in this case. Due to time constraints, I had to use my

PCR primers for sequencing. Only one of them had the correct G/C content (50%) and melting temperature (between 55ºC-60ºC) for sequencing.

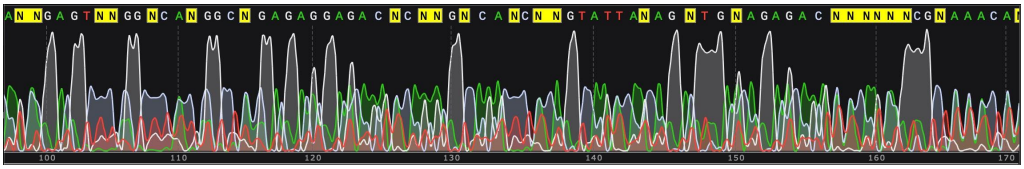

Figure 30. Electropherogram from *tag1-60* demonstrating multiple sequence signals (overlapping peaks) and unknown bases ("N").

Despite these losses, three of my sequence electropherograms were mostly readable. The first two (both from *tag1-26*, which had the WT phenotype) had a few substitutions, insertions, and deletions, but they did not appear meaningful upon alignment with the WT gene in Benchling. It's possible that they were simply artifacts of the sequencing process, especially given their proximity to regions towards the beginning and end that had more signal overlap.

Thankfully, the third sequence proved to be not only readable, but also had a significant deletion in a clean region of the electropherogram (Figure 31). This sequence was from *mpf3-3*, which had the short calyx mutation. The deletion is 41 bases long, from base 129-169 in *MPF3*. It occurs within the third exon in *MPF3* (exon:Phygri_1.1_032861.1.3) (Figure 32). This exon is 71 bases long, from base 113-183, which means the deletion removed 30 base pairs. Given that this is a large deletion that occurred within an *MPF3* exon, it appears to be meaningfully correlated with the mutated phenotype of short ICS.

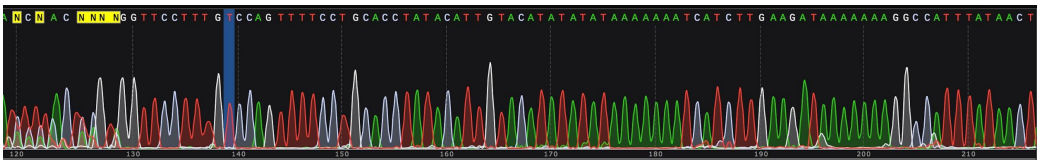

Figure 31. Electropherogram of the sequence from the reverse primer of *mpf3-3*. The base directly preceding the deletion is highlighted in blue.

Figure 32. Alignment of *MPF3* (top), mpf3-3 (middle), and exon:Phygri_1.1_032861.1.3 (bottom) sequences.

DISCUSSION

The main objective of my research was to see if there is an association between any of the ABC MADS-box genes of interest (*MPF3, DEF, TAG1,* and *EJ2*) and inflated calyx syndrome. In addition, I sought to determine if these genes have a conserved function in *Physalis*, the *Solanaceae*, or both. I found support for several of my hypotheses, and was also left with new questions that would be worth exploring in future research.

Support for the Hypotheses

mpf3

Based on previous research on *MPF3* in *P. floridana* (Zhao et al. 2013), I predicted that the *P. grisea mpf3* line would exhibit abnormally shaped ICS and have reduced pollen viability due to the prevention of starch accumulation. I was unable to ascertain pollen viability, because by the time I was ready to collect pollen, all of the flowers had set fruit. However, the documentation of floral morphology showed that 80% of the *mpf3* plants had shorter, squatter ICS in comparison with the wild type (Table 7). This suggests that the short calyx is a dominant trait. Most mutations in diploid organisms (such as *P. grisea*) are recessive, and the most common dominant mutations are typically developmental (Wilkie 1994). The short ICS phenotype may be an example of the latter

phenomenon. A possible explanation for its dominance is that CRISPR-Cas9 editing induced altered mRNA expression, which can occur in any gene with a regulatory domain (Wilkie 1994). I was able to analyze an *mpf3* sequence from one of the mutant plants, and identified a large deletion in the third exon (Figure 32). This provides one confirmation that CRISPR-Cas9 was successful in editing the *MPF3* gene sequence. Knocking out *MPF3* therefore resulted in deformed ICS, supporting the hypothesis that *MPF3* plays a role in ICS development in not just *P. grisea*, but also the genus *Physalis*.

def

The downregulation of *DEF* in *P. floridana* (Zhang et al. 2015) led to changes in petal and stamen identity, with petals becoming sepals and stamens either partially or completely becoming carpels. I anticipated that the *P. grisea def* line would display a similar floral phenotype and a potential lack of ICS due to male infertility. However, the results demonstrated the contrary. All of the floral organs developed normally, and ICS was present in every plant (Data in Appendix, Table 1A). The only notable difference in phenotype was found in 8% of the plants, which had dark pigmentation of the inflated calyx (Table 8). The pigmentation was likely anthocyanin, due to its purplish hue and the presence of anthocyanin in five wild *Physalis* genotypes (Herrera et al. 2021). Because this is such a small portion of the population, it is difficult to say whether this was caused by gene knockout or some other condition. It is also possible that this ratio

of phenotypes is due to the fact that the *def* line was an F_2 population that came from the cross of the wild type and the knockout mutant. As a result, some plants would be expected to display the wild type phenotype. Gel electrophoresis confirmed *def*'s heterozygosity (Figure 28), which demonstrates the segregation of WT and edited alleles. Human error during the purification of the *def* PCR products meant that I was unable to get them sequenced. Consequently, there is no available evidence to indicate that CRISPR editing was successful. My next course of action was to look for scientific literature indicating any relationship between *DEF* or Class B genes and anthocyanin, but my search was inconclusive. This leaves the results of the *def* line as somewhat of a mystery.

tag1

The downregulation of *TAG1* in tomato (*S. lycopersicum*) resulted in the change of stamens into petaloid organs and carpels into nested flowers. It also resulted in decreased pollen viability and seedless fruits (Pnueli et al. 1994; Gimenez et al. 2016). I hypothesized that the knockout of *tag1* would affect *P. grisea* in similar ways, with a potential lack of ICS development due to decreased male fertility. It is important to note that, although they are in the same family, *P. grisea* and tomato are in different genera. Their more distant relationship could be a reason for differences in their respective *tag1* phenotypes. On the other hand, any observed similarities between them could indicate a conserved function of *TAG1* within the *Solanaceae*. With regards to changes in floral organ identity,

26% of the *tag1* plants exhibited stamens that had developed into petaloid tissue (Table 9). Because the *tag1* line was also an F_2 population, this could be evidence of segregated alleles. There was a range of extremity in this phenotype (Figure 9), which suggests that additional genes may play a role. The IKI pollen stain revealed a lack of starch accumulation (Figure 19), though the sample size was very low so this conclusion is tentative. All of these mutated plants lacked ICS, with only one exception. This further supports the hypothesis that ICS is linked to male fertility and fertilization (Lu et al. 2021). It is worth noting that, while the single outlier may have simply been pollinated from an adjacent plant, another study with CRISPR-edited *TAG1* knockouts tells a different story. He et al. (2023) found that their *tag1* mutants in *P. grisea* exhibited the same stamens-into-petals phenotype, and were therefore incapable of self-fertilizing. However, they still grew ICS, but with an underdeveloped or absent fruit. This calls into question the ICS-fertility link, and demonstrates a more complex situation of ICS development within *P. grisea.*

ej2

For *EJ2*, I hypothesized that its knockout could alter multiple floral organs due to its redundancy as a Class E gene in regulating organ identity (Theißen et al. 2016; Sobral and Costa 2017). I also predicted that the calyces of the *ej2* plants would be larger than in the WT, based on the fact that enlarged sepals are a

consequence of *ej2* in tomato (Soyk et al. 2017). The *ej2* line generated the richest variety of phenotypes by far. The alterations that I focused on involved the transformation of sepals into petals (SIP) (Figures 10-12), an increase in sepal number (TMS) (Figures 13, 14), and differences in petal number (DIP) (Figures 15, 16). A whopping 78% of *ej2* plants had a combination of SIP and TMS, 63% had DIP, and 50% had all three mutations together (Table 10). Much smaller percentages displayed only one of the phenotypes alone. This may be a result of *EJ2*'s redundant function in floral organ development, such that knocking it out leads to combined effects in more than one whorl. The question is, what caused this array of mutations to occur?

The floral quartet model (Figure 6) may help to answer this question for the SIP phenotype. According to the floral quartet model, a tetramer of two AP1 and two SEP transcription factors regulate sepal development in Whorl 1. To make petals in Whorl 2, AP1 bonds with another transcription factor, AP3, which belongs to Class B. It is possible that, in the absence of SEP, AP3 could take its place and bind to AP1 in Whorl 1 instead. What if this phenomenon occurred in *P. grisea* when *EJ2*, a *SEP4* homolog, was knocked out? Evidence of partial AP3 redundancy was discovered in *Thalictrum thalictroides*, a petal-less plant whose sepals resemble white petals (Galimba et al. 2018). Using VIGS to downregulate *AP3* homologs, Galimba et al. (2018) observed that the sepals became smaller, curved, and developed areas of green tissue. This suggests that the heterotopic expression of *AP3* may contribute to sepal petaloidy. Because *EJ2* was knocked

out in *P. grisea*, not just downregulated, the extremity of the SIP phenotype may be further evidence to support the hypothesis of *AP3* redundancy in Whorl 1.

In the *ej2* line, not only did sepals transform into petals, they were also larger than those of the WT (Figure 13). This is in line with my prediction about calyx size, which was based on the *ej2* phenotype in tomato. An unexpected result was changes in both sepal and petal number. In seeking to understand the cause of this phenomenon, I found a research study by Biewers (2014) on the role of *SEPALLATA* genes during the development of floral organs in *Arabidopsis*. One of the experiments in this study demonstrates that *SEP4* affects floral organ number. Biewers (2014) used *A. tumefaciens* mediated transformation to create two mutant lines of *SEP4*: *sep4-1* (which produces no RNA) and *sep4-2*. Remarkably, both of these mutants exhibited an increase in both sepal and petal number, revealing a new phenotype in *Arabidopsis*. Given these results, it follows logically that *EJ2* may also also play a role in regulating floral organ number. A key difference is that in *Arabidopsis*, only an increase in the number of perianth organs was observed. In contrast, I observed both a decrease and increase in petal number in *P. grisea*.

One of the most exciting results within the context of this project is that all of the inflated calyces in the *ej2* line were deformed, their development in some cases severely disrupted by the SIP mutation (Figure 14). This indicates that *EJ2* plays an important role in the genetic underpinnings of ICS. It also suggests a

difference in developmental potential between petals and sepals, as the petaloid sepals did not inflate to form a husk. It's possible that there are unique factors regulating sepal and petal tissue that affect the ability of their cells to grow or expand in certain ways.

Review of Methods

Certain parts of my methodology were successful in this project, and others were not. Beginning with what worked well, the catalog of phenotypes was both comprehensive and incredibly helpful for documenting morphology, identifying recurring mutations, and determining the frequency of those mutations in each knockout line. The use of the dissecting scope for obtaining images proved to be invaluable, as it allowed me to record and closely study a variety of phenotypes at different stages of development. DNA extraction and amplification also yielded successful results.

The greatest obstacle in my research was the failed sequencing of most of the DNA. This prevented me from identifying evidence of CRISPR edits on the molecular level for all but one of my samples. A likely cause for this failure is the primers that I used. These were the same primers that were used for PCR, and their melting temperatures and G/C contents were not ideal for DNA sequencing. Another issue arose due to my own error when cutting the bands of *def* PCR products out of the gel for purification. I did not cut away enough agarose, which made the volume too large for purification. The pollen viability analyses were

also an area that needed improvement. I was able to observe very few pollen grains overall, and germination was consistently sparse in all of the lines. It's possible that the pollen was not mature enough, or did not take well to the -80ºC freezer. The method with which I collected the pollen by scraping the anthers may have been ineffective as well.

There are several things that I would do differently if given the time and the opportunity. For the phenotypic analyses, I would obtain a more detailed image series of floral and ICS development in each line. This series would consist of more incremental stages, and I would record specific dates along the way to see how long it takes for traits such as ICS to fully mature. In addition, I would attempt to improve the methods for obtaining pollen and testing its viability in order to receive more comprehensive data. Finally, I would observe the fruit development in each line and look for any seedless phenotypes, as this was a prediction I had made for *tag1*.

For the genotypic analyses, I would design primers suited for sequencing. The next step would be to clone the PCR products, which is a time-consuming process. This involves inserting the PCR product into a plasmid cloning vector, introducing the vector to *E. coli*, cultivating the *E. coli*, isolating the plasmid, and then using vector-specific forward and reverse primers to sequence the entire amplified regions.

Future Work

There are multiple directions that this research could go in the future based off of the conclusions drawn in this thesis. A fundamental place to start would be the proper sequencing of DNA from each line. This would help paint a better picture of what is occurring on the genetic level in the plants exhibiting mutated phenotypes. Delving deeper, I am most interested in unraveling the role that *EJ2* plays in the development of ICS, and determining what exactly caused the various *ej2* phenotypes to manifest. Further research could investigate the expression of the *EJ2*, *MPF3* (*AP1* homolog) and the *DEF* (*AP3* homolog) in the floral whorls. This could be accomplished with a northern blot, a technique which detects the presence of mRNA in tissue. He and Saedler (2005) used this method to study heterotopic expression of *MPF2* in *P. pubescens* and *P. peruviana*. It would be interesting to see if, in the *ej2* SIP mutant, heterotopic expression of *DEF* occurs alongside *MPF3* in the sepals when *EJ2* is absent. If it does, that could point to interaction between *MPF3* and *DEF*, possibly in the formation of tetramers, to transform sepals into petals. I am also curious about *EJ2'*s effects on floral organ number, as the combined results from this study and research in *Arabidopsis* (Biewers 2014) indicate that it has an instrumental role. Investigating the mechanisms underlying the determination of floral organ number, and how *EJ2* may factor into these mechanisms, could be another avenue for future research.

Conclusion

In summary, the results of this project suggest that *MPF3, TAG1,* and *EJ2* are implicated in the development of ICS in *P. grisea*. The *mpf3* mutation exhibited an alteration in ICS morphology, with shorter calyces than the wild type. The *tag1* mutation had an absence of male fertility due to a change in stamen identity, and this was strongly correlated with an absence of ICS. However, one observed exception could indicate that infertility does not prevent ICS entirely, a hypothesis supported by research done in *P. floridana* (He et al. 2023). The *ej2* line exhibited multiple mutations in floral organ identity and number. This knockout was found to alter sepal and ICS morphology in the TMS phenotype, and disrupt ICS development completely in the SIP phenotype. These findings provide exciting insight into the genetic underpinnings of ICS in *P. grisea*, and can provide a stepping stone for future research.

LITERATURE CITED

Agrawal, Neema, et al. "RNA interference: biology, mechanism, and applications." *Microbiology and molecular biology reviews* 67.4 (2003): 657-685.

Biewers, Sandra. SEPALLATA *genes and their role during floral organ formation.* The University of Leeds School of Biology, PhD dissertation. *White Rose*, https://etheses.whiterose.ac.uk/8965/1/Sepallata%20genes%20and%20their%20role%20during%20floral%20organ%20development.pdf

Chen, Yanru, and Bik-Kwoon Tye. "The yeast Mcm1 protein is regulated posttranscriptionally by the flux of glycolysis." *Molecular and cellular biology* 15.8 (1995): 4631-4639.

Deanna, Rocío, et al. "Repeated evolution of a morphological novelty: a phylogenetic analysis of the inflated fruiting calyx in the Physalideae tribe (Solanaceae)." *American Journal of Botany* 106.2 (2019): 270-279.

Devi, Maibam Rasila, et al. "Neurotoxic and medicinal properties of Datura stramonium L.–review." *Assam University Journal of Science and Technology* 7.1 (2011): 139-144.

Emboden, William. "The sacred journey in dynastic Egypt: shamanistic trance in the context of the narcotic water lily and the mandrake." *Journal of psychoactive drugs* 21.1 (1989): 61-75.

Galimba, Kelsey D., Jesús Martínez-Gómez, and Verónica S. Di Stilio. "Gene duplication and transference of function in the paleo AP3 lineage of floral organ identity genes." *Frontiers in Plant Science* 9 (2018): 334.

Gebhardt, Christiane. "The historical role of species from the Solanaceae plant family in genetic research." *Theoretical and Applied Genetics* 129 (2016): 2281-2294.

Gimenez, Estela, et al. "TOMATO AGAMOUS1 and ARLEQUIN/TOMATO AGAMOUS-LIKE1 MADS-box genes have redundant and divergent functions required for tomato reproductive development." *Plant molecular biology* 91 (2016): 513-531.

He, Chaoying, and Heinz Saedler. "Heterotopic expression of MPF2 is the key to the evolution of the Chinese lantern of Physalis, a morphological novelty in Solanaceae." *Proceedings of the National Academy of Sciences* 102.16 (2005): 5779-5784.

He, Chaoying, and Heinz Saedler. "Hormonal control of the inflated calyx syndrome, a morphological novelty, in Physalis." *The Plant Journal* 49.5 (2007): 935-946.

He, Jia, et al. "Establishing Physalis as a Solanaceae model system enables genetic reevaluation of the inflated calyx syndrome." *The Plant Cell* 35.1 (2023): 351-368.

Irish, Vivian. "The ABC model of floral development." *Current Biology* 27.17 (2017): R887-R890.

Ishino, Yoshizumi, Mart Krupovic, and Patrick Forterre. "History of CRISPR-Cas from encounter with a mysterious repeated sequence to genome editing technology." *Journal of bacteriology* 200.7 (2018): e00580-17.

Kapoor, Meenu, et al. "Role of petunia pMADS3 in determination of floral organ and meristem identity, as revealed by its loss of function." *The Plant Journal* 32.1 (2002): 115-127.

Knapp, Sandra. "Tobacco to tomatoes: a phylogenetic perspective on fruit diversity in the Solanaceae." *Journal of Experimental Botany* 53.377 (2002): 2001-2022.

Li, Jing, Chunjing Song, and Chaoying He. "Chinese lantern in Physalis is an advantageous morphological novelty and improves plant fitness." *Scientific reports* 9.1 (2019): 596.

Lopez-Gomollon, Sara. "Physalis: A new model crop to understand plant diversity." (2023): 338-339.

Lu, Jiangjie, et al. "The Physalis floridana genome provides insights into the biochemical and morphological evolution of Physalis fruits." *Horticulture Research* 8 (2021).

Lu, Rui, et al. "Virus-induced gene silencing in plants." *Methods* 30.4 (2003): 296-303.

Mizukami, Yukiko, and Hong Ma. "Ectopic expression of the floral homeotic gene AGAMOUS in transgenic Arabidopsis plants alters floral organ identity." *Cell* 71.1 (1992): 119-131.

Mojica, Francisco JM, and Lluis Montoliu. "On the origin of CRISPR-Cas technology: from prokaryotes to mammals." *Trends in microbiology* 24.10 (2016): 811-820.

Mol, J. N. M., et al. "Regulation of plant gene expression by antisense RNA." *FEBS letters* 268.2 (1990): 427-430.

Moreira, D. (2011). Homology. In: , *et al.* Encyclopedia of Astrobiology. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-11274-4_1724

Ng, Medard, and Martin F. Yanofsky. "Function and evolution of the plant MADS-box gene family." *Nature Reviews Genetics*2.3 (2001): 186-195.

Pnueli, Lilac, et al. "Isolation of the tomato AGAMOUS gene TAG1 and analysis of its homeotic role in transgenic plants." *The Plant Cell* 6.2 (1994): 163-173.

Ramegowda, Venkategowda, Kirankumar S. Mysore, and Muthappa Senthil-Kumar. "Virus-induced gene silencing is a versatile tool for unraveling the functional relevance of multiple abiotic-stress-responsive genes in crop plants." *Frontiers in plant science* 5 (2014): 323.

Redei (2008). Knock-Out (KO). In: Encyclopedia of Genetics, Genomics, Proteomics and Informatics. Springer, Dordrecht. https://doi.org/10.1007/978-1-4020-6754-9_9086

Reynard, George B. "New Source of the j 2 Gene Governing Jointless Pedicel in Tomato." *Science* 134.3496 (1961): 2102-2102.

Shenstone, Esperanza, Zach Lippman, and Joyce Van Eck. "A review of nutritional properties and health benefits of Physalis species." *Plant Foods for Human Nutrition* 75 (2020): 316-325.

Sobral, Rómulo, and M. Manuela R. Costa. "Role of floral organ identity genes in the development of unisexual flowers of Quercus suber L." *Scientific reports* 7.1 (2017): 1-15.

Soyk, Sebastian, et al. "Bypassing negative epistasis on yield in tomato imposed by a domestication gene." *Cell* 169.6 (2017): 1142-1155.

Soza, Valerie L., et al. "Partial redundancy and functional specialization of E-class SEPALLATA genes in an early-diverging eudicot." *Developmental Biology* 419.1 (2016): 143-155.

Sullivan, Janet R. "The genus Physalis (Solanaceae) in the southeastern United States." *Rhodora* (2004): 305-326.

Tadele, Zerihun. "Orphan crops: their importance and the urgency of improvement." *Planta* 250 (2019): 677-694.

Theißen, Günter, Rainer Melzer, and Florian Rümpler. "MADS-domain transcription factors and the floral quartet model of flower development: linking plant development and evolution." *Development* 143.18 (2016): 3259-3271

Tzfira, Tzvi, and Vitaly Citovsky. "Agrobacterium-mediated genetic transformation of plants: biology and biotechnology." *Current opinion in biotechnology* 17.2 (2006): 147-154.

Wilf, Peter, et al. "Eocene lantern fruits from Gondwanan Patagonia and the early origins of Solanaceae." *Science*355.6320 (2017): 71-75.

Wilkie, A. O. "The molecular basis of genetic dominance." *Journal of medical genetics* 31.2 (1994): 89-98.

Zhao, Jing, et al. "Multiple and integrated functions of floral C-class MADS-box genes in flower and fruit development of Physalis floridana." *Plant Molecular Biology* 107.1-2 (2021): 101-116.

Zhao, Jing, et al. "The euAP1 protein MPF3 represses MPF2 to specify floral calyx identity and displays crucial roles in Chinese lantern development in Physalis." *The Plant Cell* 25.6 (2013): 2002-2021.

Zhang, Shaohua, et al. "Distinct subfunctionalization and neofunctionalization of the B-class MADS-box genes in Physalis floridana." *Planta* 241 (2015): 387-402.

APPENDIX

Table A1. Catalog of wild type phenotypes.

Plant #	Whorl of Interest	Mutation	Identity Swap	Post Fertilization Abnormality	Presence of ICS	Flowers
WT-1	-	-	-	-	Y	Y
WT-2	-	-	-	-	Y	Y
WT-3	-	-	-	-	Y	Y
WT-4	-	-	-	-	Y	Y
WT-5	-	-	-	-	Y	Y
WT-6	-	-	-	-	Y	Y
WT-7	-	-	-	-	Y	Y
WT-8	-	-	-	-	Y	Y
WT-9	-	-	-	-	Y	Y
WT-10	-	-	-	-	Y	Y
WT-11	-	-	-	-	Y	Y
WT-12	-	-	-	-	Y	Y
WT-13	-	-	-	-	Y	Y
WT-14	-	-	-	-	Y	Y
WT-15	-	-	-	-	Y	Y
WT-16	-	-	-	-	Y	Y
WT-17	-	-	-	-	Y	Y
WT-18	-	-	-	-	Y	Y

WT-19	-	-	-	-	Y	Y
WT-20	-	-	-	-	Y	Y
WT-21	-	-	-	-	Y	Y
WT-22	-	-	-	-	Y	Y
WT-23	-	-	-	-	Y	Y
WT-24	-	-	-	-	Y	Y
WT-25	-	-	-	-	Y	Y
WT-26	-	-	-	-	Y	Y
WT-27	-	-	-	-	Y	Y
WT-28	-	-	-	-	Y	Y
WT-29	-	-	-	Berry extruded from calyx	Y	Y
WT-30	-	-	-	-	Y	Y
WT-31	-	-	-	-	Y	Y
WT-32	-	-	-	-	Y	Y

Table A2. Catalog of *mpf3* phenotypes.

Plant #	Whorl of Interest	Mutation	Identity Swap	Post Fertilization Abnormality	Presence of ICS	Flowers
mpf3-1	2	Missing petal, stamen fused to where petal should be		Open-tipped calyx	Y	Y
mpf3-2				Open-tipped calyx	Y	Y
mpf3-3				Short calyx	Y	Y
mpf3-4				Short calyx	Y	Y
mpf3-5				Short calyx	Y	Y
mpf3-6					Y	Y
mpf3-7				Short calyx	Y	N
mpf3-8				Short calyx	Y	N
mpf3-9				Short calyx	Y	Y
mpf3-10				Short calyx	Y	Y
mpf3-11				Short calyx	Y	N
mpf3-12				Short calyx	Y	Y
mpf3-13				Short calyx	Y	N
mpf3-14				Short calyx	Y	N
mpf3-15					Y	N
mpf3-16					Y	N
mpf3-17					Y	N
mpf3-18					Y	N
mpf3-19				Short Calyx	Y	Y
mpf3-20				Short Calyx	Y	Y
mpf3-21				Short Calyx	Y	Y
mpf3-22				Short Calyx	Y	Y

mpf3-23				Short Calyx	Y	Y
mpf3-24				Short Calyx	Y	Y
mpf3-25				Short Calyx; Petaloid ICS?	Y	Y
mpf3-26				Short calyx	Y	Y
mpf3-27				Short calyx	Y	Y
mpf3-28				Short calyx	Y	Y
mpf3-29				Short calyx	Y	Y
mpf3-30				Short calyx	Y	Y
mpf3-31				Short calyx	Y	N
mpf3-32					Y	Y
mpf3-33				Short calyx	Y	Y
mpf3-34				Short calyx	Y	N
mpf3-35				Short calyx	Y	N
mpf3-36				Short calyx	Y	N
mpf3-37				Short calyx	Y	Y
mpf3-38				Short calyx	Y	Y
mpf3-39				Both open and short calyx	Y	N
mpf3-40				Short calyx	Y	Y
mpf3-41				Both open and short calyx	Y	Y
mpf3-42				Short calyx	Y	N
mpf3-43				Short calyx	Y	Y
mpf3-44				Short calyx	Y	Y
mpf3-45				Short calyx	Y	Y
mpf3-46				Short calyx	Y	Y

mpf3-47				Both open and short calyx	Y	Y
mpf3-48				Open calyx	Y	N
mpf3-49				Short calyx	Y	N
mpf3-50				Both open and short calyx	Y	N
mpf3-51					Y	N
mpf3-52				Short calyx	Y	Y

Table A3. Catalog of *def* phenotypes.

Plant #	Whorl of Interest	Mutation	Identity Swap	Post Fertilization Abnormality	Presence of ICS	Flowers
def-1	2	Petals seem doubled, the parts with the pollination guides are darker and more defined on one of the flowers (but not both)			Y	Y
def-2	2	darker sections of the petals coming out from the pollination guides		very dark young calyx	Y	Y
def-3				very dark young calyx	Y	N
def-4					Y	N
def-5	2	darker sections of the petals coming out from the pollination guides		purple-ish pigmentation in younger calyx	Y	Y
def-6					Y	N
def-7					Y	Y
def-8					Y	Y
def-9					Y	Y
def-10					Y	Y
def-11					Y	Y
def-12					Y	Y

def-13					Y	N
def-14					Y	Y
def-15					Y	N
def-16					Y	Y
def-17				1 dark young calyx	Y	Y
def-18				1 dark young calyx	Y	Y
def-19					Y	N
def-20					Y	Y
def-21					Y	Y
def-22					Y	Y
def-23					Y	Y
def-24					Y	Y
def-25					Y	Y
def-26					Y	Y
def-27	2	4 fused petals instead of 5			Y	Y
def-28					Y	Y
def-29					Y	N
def-30					Y	Y
def-31					Y	Y
def-32					Y	Y
def-33					Y	Y
def-34					Y	Y
def-35					Y	Y
def-36					Y	Y
def-37					Y	Y
def-38					Y	Y
def-39					Y	N
def-40					Y	Y
def-41					Y	Y

def-42					Y	N
def-43					Y	Y
def-44					Y	Y
def-45					Y	Y
def-46					Y	Y
def-47					Y	N
def-48				most normal calyxes, one appears to be growing with one calyx inside another	Y	N
def-49					Y	Y
def-50					Y	Y
def-51					Y	Y
def-52					Y	Y
def-53					Y	Y
def-54					Y	Y
def-55					Y	Y
def-56					Y	Y
def-57					Y	Y
def-58					Y	Y
def-59					Y	Y
def-60					Y	Y

Table A4. Catalog of *tag1* phenotypes.

Plant #	Whorl of Interest	Mutation	Identity Swap	Post Fertilization Abnormality	Presence of ICS	Flowers
tag1-1	3		Stamens into Petals		N	Y
tag1-2	3		Stamens into Petals		N	Y
tag1-3					Y	N
tag1-4					Y	Y
tag1-5	3		Stamens into Petals		N	Y
tag1-6	3		Stamens into Petals		N	Y
tag1-7					Y	Y
tag1-8	3		Stamens into Petals	*Note: Both calyxes and mutated flowers	Y	Y
tag1-9					Y	Y
tag1-10					Y	Y
tag1-11					Y	Y
tag1-12					Y	N
tag1-13					Y	N
tag1-14					Y	Y
tag1-15					Y	Y
tag1-16					Y	Y
tag1-17					Y	Y
tag1-18					Y	Y
tag1-19					Y	N
tag1-20					Y	Y
tag1-21					Y	Y
tag1-22					Y	N
tag1-23					Y	Y

tag1-24					Y	N
tag1-25					Y	Y
tag1-26					Y	Y
tag1-27					Y	Y
tag1-28					Y	Y
tag1-29	3		Stamens into petals		N	Y
tag1-30					Y	N
tag1-31					Y	Y
tag1-32					Y	Y
tag1-33					Y	N
tag1-34					Y	Y
tag1-35					Y	N
tag1-36					Y	Y
tag1-37					Y	N
tag1-38					Y	Y
tag1-39	3		Stamens into petals		N	Y
tag1-40	3		Stamens into petals		N	Y
tag1-41					Y	Y
tag1-42	3		Stamens into petals		N	Y
tag1-43					Y	Y
tag1-44					Y	N
tag1-45					Y	N
tag1-46					Y	Y
tag1-47	3		Stamens into petals		N	Y
tag1-48	3		Stamens into petals		N	Y
tag1-49					Y	Y
tag1-50					Y	Y

tag1-51					Y	Y
tag1-52					Y	Y
tag1-53	3		Stamens into petals		N	Y
tag1-54	3		Stamens into petals		N	Y
tag1-55					Y	Y
tag1-56					Y	Y
tag1-57					Y	Y
tag1-58					Y	Y
tag1-59					Y	Y
tag1-60	3		Stamens into Petals		N	Y
tag1-61					Y	Y
tag1-62	3		Stamens into Petals		N	Y
tag1-63					Y	N
tag1-64	3		Stamens into Petals		N	Y
tag1-65	3		Stamens into Petals		N	Y
tag1-66	3		Stamens into Petals		N	Y
tag1-67					Y	Y
tag1-68					Y	Y

Table A5. Catalog of *ej2* phenotypes.

Plant #	Whorl of Interest	Mutation	Identity Swap	Post Fertilization Abnormality	Presence of ICS	Flowers
ej2-1	1	TMS		TMS leads to deformed ICS	Y	Y
ej2-2	1, 2	TMS; two flowers have 4 petals with a stunted fifth	SIP	TMS leads to deformed ICS	Y	Y
ej2-3	1, 2	TMS; one flower has too many petals, one has too few	SIP	TMS leads to deformed ICS	Y	Y
ej2-4	1	TMS	SIP		N	Y
ej2-5	1,2	TMS; one flower appears normal in petal count but is missing a pollination guide	SIP; one petal is missing and has a stamen in its place	TMS leads to deformed ICS	Y	Y
ej2-6	1,2	TMS; too many petals		TMS leads to deformed ICS	Y	Y
ej2-7	1,2	TMS	SIP; two stamen in places of	TMS leads to deformed ICS	Y	Y

			petal (unsure if fused or replacing)			
ej2-8	1,2	TMS	SIP; a stamen in place of missing petal	TMS leads to deformed ICS	Y	Y
ej2-9	1,2	TMS; one missing petal	SIP	TMS leads to deformed ICS	Y	Y
ej2-10					N	N
ej2-11	1,2	TMS; one missing petal on multiple flowers	SIP		N	Y
ej2-12	1,2	TMS; one missing petal on multiple flowers	SIP		N	Y
ej2-13	1,2,3	TMS; missing petals; stunted pollenless stamen	SIP	TMS leads to deformed ICS	Y	Y
ej2-14	1,2,3	TMS; missing petals; stunted pollenless stamen	SIP	TMS leads to deformed ICS	Y	Y

ej2-15	1,2	TMS; missing petals	SIP		N	Y
ej2-16	1,2	TMS; too many petals	SIP		N	Y
ej2-17	1	TMS	SIP		N	Y
ej2-18	1,2	TMS; missing petal	SIP	TMS leads to deformed ICS	Y	Y
ej2-19	1, 2	TMS; flower with 4 petals	SIP	TMS leads to deformed ICS	Y	Y
ej2-20	1		SIP	SIP leads to deformed ICS	Y	Y
ej2-21	1	TMS (6-7)	SIP	deformed ICS	Y	Y
ej2-22	1, 2	TMS on flower with 4 petals	SIP	SIP leads to deformed ICS; normal # sepals but open ICS	Y	Y
ej2-23	1, 2	two SIP flowers with 3 petals	SIP	SIP leads to deformed ICS	Y	Y
ej2-24	1, 2	TMS (8) on flower with 4 petals; flower with 1 stunted petal	SIP	SIP leads to deformed ICS	N	Y
ej2-25	1		SIP	SIP leads to	Y	Y

				deformed ICS		
ej2-26	1, 2	TMS (8) on flower with 4 petals	SIP	SIP leads to deformed ICS	Y	Y
ej2-27	1, 2	TMS (7) on flower with 4 petals; flower with	SIP	SIP/TMS leads to deformed ICS	Y	Y
ej2-28	1, 2	TMS (7) on flower with 3 petals	SIP		N	Y
ej2-29	1		SIP		N	Y
ej2-30	1	TMS	SIP	TMS leads to deformed ICS	Y	Y
ej2-31	1	TMS	SIP	TMS/SIP leads to deformed ICS	Y	Y
ej2-32	1, 2	TMS (6-7); flower with 3 petals has TMS	SIP	TMS leads to deformed ICS	Y	Y
ej2-33	1, 2	TMS; flower with 4 petals		TMS leads to deformed ICS	Y	Y
ej2-34	1, 2	TMS (8); flower with 3 petals an TMS (7)		TMS leads to deformed ICS	Y	Y

ej2-35	1	TMS (6)	SIP	TMS leads to deformed ICS	Y	Y
ej2-36	1, 2	TMS (6); flower with 4 petals		TMS leads to deformed ICS	Y	Y
ej2-37	1,2	TMS; one flower with 3 petals	SIP	SIP leads to deformed ICS	Y	Y
ej2-38	1	TMS	SIP	SIP leads to deformed ICS	Y	Y
ej2-39	1	TMS; one flower with 4 petals	SIP		N	Y
ej2-40	1,2	TMS; one flower with 6 petals	SIP		N	Y
ej2-41	1,2,3	TMS; two flowers with a missing petal; two flowers with all stunted pollenless stamen	SIP	SIP leads to deformed ICS	Y	Y
ej2-42	1	TMS	SIP		N	Y
ej2-43	1,2	TMS; one flower with 3 petals	SIP	SIP leads to deformed ICS	Y	Y

ej2-44	1,3	TMS; one flower with all stunted pollenless stamen	SIP	SIP leads to deformed ICS	Y	Y
ej2-45	1,2,3	TMS; one flower with 4 petals; one flower with all stunted pollenless stamen	SIP	SIP leads to deformed ICS	Y	Y

Table A6. Germination medium protocol.

Component	**[Stock]**	**For 5ml**	**For 10ml**	**For 20ml**	**For 30ml**	**For 40ml**
PEG 4000	40%	3	6	12	18	24
Boric Acid	0.1%	0.5	1	2	3	4
Sucrose	40%	0.25	0.5	1	1.5	2
HEPES buffer	0.5M, pH 6.0	0.2	0.4	0.8	1.2	1.6
$Ca(NO_3)_2$-$4H_2O$	0.1M	0.15	0.3	0.6	0.9	1.2
$MgSO_4$-$7H_2O$	2%	0.05	0.1	0.2	0.3	0.4
KNO_3	1% or 0.1M	0.05	0.1	0.2	0.3	0.4
H_2O	Double Distilled	0.8	1.6	3.2	4.8	5.6