THE PENNSYLVANIA STATE UNIVERSITY
SCHREYER HONORS COLLEGE

DEPARTMENT OF BIOCHEMISTRY AND MOLECULAR BIOLOGY

DISSECTING THE INTERACTIONS BETWEEN S-RNASES AND S-LOCUS F-BOX PROTEINS OF PETUNIA DURING SELF/NON-SELF RECOGNITION IN SELF INCOMPATIBILITY

DANIELE SAN ROMÁN
SPRING 2015

A thesis
submitted in partial fulfillment
of the requirements
for a baccalaureate degree
in Biochemistry and Molecular Biology
with honors in Biochemistry and Molecular Biology

Reviewed and approved* by the following:

Teh-Hui Kao
Distinguished Professor of Biochemistry and Molecular Biology
Thesis Supervisor
Honors Adviser

Allen T. Phillips
Professor Emeritus of Biochemistry
Faculty Reader

Scott Selleck
Professor of Biochemistry and Molecular Biology
Department Head

* Signatures are on file in the Schreyer Honors College.

ABSTRACT

Genetic variability is essential in all living organisms to ensure that progeny are fit and that the species is evolving. When inbreeding occurs, genetic variance is eliminated and this can cause weaker progeny. To avoid inbreeding, animals have the ability to move around and select a mate. Plants, however, are dependent on abiotic factors to introduce new DNA with which they can be fertilized. Additionally, the bisexual structure of angiosperms, or flowering plants, makes it inevitable that self-pollen from the anther (male reproductive organ) will be introduced to the pistil (female reproductive organ). Fortunately, plants have developed a mechanism to recognize and reject self-pollen to prevent inbreeding.

The Kao Lab studies this phenomenon, known as self-incompatibility (SI), in *Petunia inflata*, a wild species of petunia. This research has discovered that the basis for SI lies within the genes of the *S*-locus region of the chromosome. The *S*-locus is polymorphic and as such, contains many variants. These variants are referred to as *S*-haplotypes and are designated S_1, S_2, and so forth. When the *S*-haplotype of the pollen matches that of the pistil, the pollen will be rejected and fertilization will not occur, thus preventing inbreeding. This mechanism is dependent on the interactions between *S*-Locus F-box (SLF) proteins produced in the pollen and S-ribonuclease (S-RNase) proteins produced in the pistil. SLF proteins produced by pollen of a given *S*-haplotype collectively interact with all their non-self S-RNases to mediate degradation of these cytotoxic enzymes, allowing the pollen tubes to grow. However, none of the SLF proteins interact with their self-S-RNase, therefore self-pollen tubes will die and fertilization will be prevented.

This study examines the biochemical basis for the differential interactions between SLF proteins and various S-RNases. Dr. Kao's lab previously divided the amino acid sequences of S_2-SLF1 (SLF1 of *S*$_2$-haplotype) and S_3-SLF1 (SLF1 of *S*$_3$-haplotype) into three domains, FD1 (Functional Domain 1), FD2 (Functional Domain 2), and FD3 (Functional Domain 3). This study will use different combinations of these domains to examine the role of each domain in interactions with S-RNases. The two combinations used are F232 and F322. This means that F232 contains FD1 and FD3 from S_2-SLF1 and FD2 from S_3-SLF1. Transgenic plants producing these chimeric SLFs were crossed with plants of varying *S*-haplotypes. The progeny were raised and tested for their SI behavior to determine whether each chimeric SLF and the S-RNases tested will interact. Based on these interactions, we can determine which domain or domains in S_2-SLF1 and S_3-SLF1 are involved in specific interactions with a particular S-RNase, and can then examine these domains further to determine which specific amino acids are responsible.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

I would like to thank Professor Teh-hui Kao for inspiring me with his incredible passion for research and exploration. I was very fortunate to have Professor Kao as my honors adviser, and even more so to have him open his lab to me. Hearing Dr. Kao talk of self-incompatibility immediately caught my interest and I was excited to work in his lab. Throughout the two years I have spent in the Kao Lab, I have learned more than I could ever have imagined, and for that I am grateful as well. Without Dr. Kao always pushing me gently forward, I would never have gained as much knowledge and experience as I have.

I would also like to thank Dr. Allen T. Phillips for not only agreeing to aid me in writing this thesis, but for giving me my first undergraduate research experience. It was with Dr. Phillips' guidance that I learned to be comfortable in a lab setting. During the year I spent working in his lab, I gained valuable skills and knowledge that I used throughout my college career.

I would also like to thank Justin Williams, the graduate student with whom I have worked most closely. Justin's patience and ability to explain complex topics helped me to understand everything that has been discovered about self-incompatibility. I am grateful for his daily guidance, as well as his constant assistance with new lab techniques.

Introduction

Background

The bisexual nature of angiosperms (flowering plants) lends itself to self-pollination, which could potentially lead to self-fertilization, resulting in inbreeding. This leads to decreased genetic diversity, which in turn could manifest itself as reduced "fitness" of the progeny. Other species, such as those in the animal kingdom, have developed other means for maintaining the fitness of their progeny; however, flowering plants do not have these abilities and must rely on pollination by wind, insects and other pollinators. The structure of angiosperms makes reproduction even more complicated. Anthers, the male organs in flowers, are located in close proximity to the female organ, the pistil. This makes it inevitable that a plant's own pollen (self-pollen) will be introduced into the pistil. Interestingly, many species of flowering plants have developed strategies to prevent self-pollen tubes from reaching the ovary so inbreeding will not occur. At the basis of several of these systems is the ability of the plant to recognize the difference between self-pollen and pollen from another, genetically distinct plant and to reject fertilization by self-pollen. This ability is referred to as self-incompatibility (SI) (de Nettancourt, 2001).

The Kao Lab uses *Petunia inflata* to study Solanaceae type SI, which is present in other wild species of the Solanaceae family such as *Capsium annuum* (peppers), *Lycopersicon esculentum* (tomatoes), and *Solanum tuberosum* (potatoes), as well as members of the Rosaceae family (*Malus domestica* [apple], *Pyrus communis* [pear], *Pyranus persica* [peach], *Prunus*

armeniaca [apricot]) and *Antirrhinum* from the Plantaginaceae family (Iwano and Takayma, 2012). The small size and ease of growing petunia plants makes the petunia the ideal model organism to study SI. The presence of Solanaceae type SI in so many species that produce food means that the discoveries made on petunias could be applied across these various food-producing species to improve agricultural practices. It has been found that there are two determinants involved in this phenomenon; the male determinant, encoded by multiple *S-Locus F-box* (SLF) genes (Kubo et al., 2010), and the female determinant, encoded for by the *S-Ribonuclease* (S-RNase) gene (Lee et al., 1994; Murfett et al., 1994). These genes are located in a genetically defined locus, referred to as the *S*-locus (Wang and Kao, 2011).

The *S*-locus is highly polymorphic with 36 documented haplotypes (Sims and Robbins, 2009). These haplotypes are designated as S_1, S_2, S_3, and so on. The self/non self-recognition that takes place within the pistil of the plant is dependent on which of these *S*-haplotypes is present in the pistil and which is present in the pollen (de Nettancourt 2001, Sijacic et al., 2004). If the same *S*-haplotype is present in both pistil and pollen, the pistil will recognize that the pollen is self-pollen. In these instances, fertilization will be prevented. When the *S*-haplotype of the pollen is different than that of the pistil, fertilization will be permitted to occur. The physical distinction of these *S*-haplotypes is dependent on the determinant proteins expressed in the pollen and the pistil, all of which are encoded by genes located in the *S*-locus. In the pistil, the female determinant protein responsible for self/non-self-recognition is called S-Ribonuclease (S-RNase) (McClure et al., 1990; Lee et al., 1994; Liu et al., 2009). S-Locus F-Box (SLF) proteins have been identified as the necessary factor for recognition present in pollen (Kubo et al., 2010). To date, there are 17 known SLF proteins (Williams et al., 2014). Fertilization is dependent on the recognition and interaction of S-RNases and all SLF proteins.

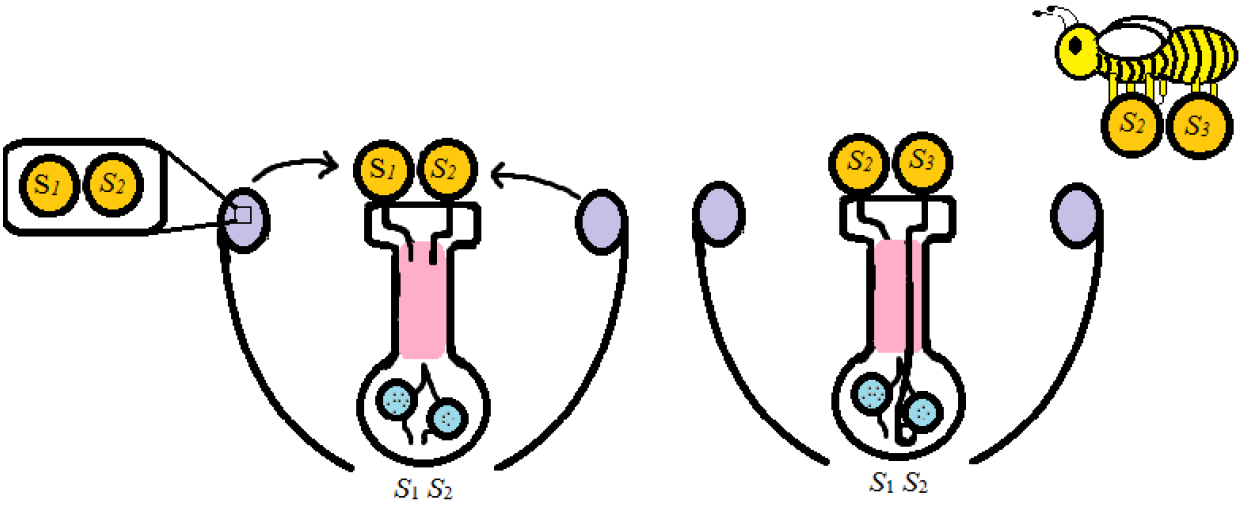


Figure 1. Self-Incompatibility in *Petunia inflata*

The S_1S_2 pistil in Figure 1 will not allow itself to be fertilized by S_1 or S_2 pollen. It will recognize that pollen of both of these *S*-haplotypes are the same as its own and prevent the growth of any self-pollen tubes. None of the SLF proteins produced by S_1 pollen will be able to recognize or interact with S_1-RNase from the pistil, and this will prevent the growth of the S_1 pollen tube. Similarly, S_2-RNase will not be recognized by any of the SLF proteins produced in S_2 pollen, and the S_2 pollen tube growth will be prevented as well. If S_2 and S_3 pollen are introduced to the same S_1S_2 pistil, the pistil will be fertilized. It will still recognize that the S_2 pollen contains the same *S*-haplotype and prevent fertilization from this pollen. However, the S_3 pollen will be recognized as non-self, and pollen tubes will be permitted to grow. This is because, among the SLF proteins produced by S_3 pollen, at least one will interact with S_1-RNase, and at least one will interact with S_2-RNase.

S-RNase acts to degrade ribonucleic acid (RNA) which is essential for cell function. The degradation of RNA in self-pollen tubes prevents these tubes from growing and fertilizing the flower in question (Huang et al., 1994). When an SLF protein recognizes a particular S-RNase, the interaction between the two causes the S-RNase to be ubiquitinated and degraded. Without

the S-RNase present, the RNA in pollen tubes will not be degraded and these tubes will continue to grow.

The SLF proteins are members of Skp1-Cullin1-F-box (SCF) E3-ubiquitin-ligase complex. This complex is involved in ubiquitin-mediated protein degradation via the 26S proteasome. Proteins that enter the complex are marked with a polyubiquitin chain which targets them for degradation by the 26S-proteasome pathway (Bai et al., 1996; Stone and Callis, 2007; Vierstra, 2009). The protein which is targeted in this mechanism is the S-RNase from the pistil. When a pollen tube grows, it takes up the S-RNases produced in the pistil (Luu et al., 2000; Goldraij et al., 2006). The SLF proteins present in the pollen tube of a given *S*-haplotype collectively mediate the degradation of all S-RNases except for their self S-RNase (Kubo et al., 2010). For example, S_2 and S_3 pollen will both germinate at the top of an S_1S_2 pistil and pollen tubes will begin to grow into the style. Each pollen tube will produce SLF proteins, but only the S_3 pollen tube will achieve successful fertilization. The S_2 pollen tube SLFs could mediate the degradation of S_1-RNase, S_3-RNase, and so on, but none of the SLFs produced by the S_2 pollen tube will be able to interact with S_2-RNase. This will prevent the S_2-RNase from entering the SCF E3-ubiquitin-ligase complex and being marked for degradation. These S-RNases will act as cytotoxins and prevent growth of the pollen tube, most likely by degrading RNA (Huang et al., 1994). The S_3 pollen tube, however, will produce at least one SLF protein capable of interacting with S_1-RNase and S_2-RNase produced the pistil. These proteins will be marked with polyubiquitin chains and degraded by the 26S proteasome. They will not be able to produce a cytotoxic effect on the growth tube and fertilization will occur (Hua and Kao, 2006; Kubo et al., 2010). Thus, S_3 pollen is compatible with an S_1S_2 pistil, and S_2 pollen is not.

In this system of recognition, if only a single SLF were involved, it would have to interact with an astonishingly wide variety of divergent S-RNases. To aid in this recognition process, pollen of each *S*-haplotype produces multiple SLF proteins. It has been shown that each SLF is responsible for recognizing only a subset of non-self S-RNases, and that all SLF proteins collectively can recognize and interact with all S-RNases besides their self S-RNase (Hua and Kao, 2006; Kubo et al., 2010). For example, in Figure 1, pollen of S_2-haplotype produces 17 SLF proteins, S_2-SLF1, S_2-SLF2, S_2-SLF3, and on to S_2-SLF17, and pollen of S_3-haplotype also produces the S_3-alleleic variants of the same 17 SLF proteins, *S_3-SLF1*, *S_3-SLF2*, and *S_3-SLF3*, to *S_3-SLF17*. S_3-SLF1 does not interact with S_2-RNase, but S_3-SLF6 does. Thus, S_3-SLF6 will mediate the degradation of S_2-RNase and allow the pollen tube to grow. Although S_3-SLF1 is ineffective against S_2-RNase, some other SLF proteins are able to mediate the degradation of this cytotoxin. In this way, the SLF proteins in the pollen are equipped to handle a wide variety of S-RNases from a wide variety of different *S*-haplotype flowers.

Project Goals

Despite the extensive body of knowledge acquired into the molecular biological mechanisms responsible for self-incompatibility (SI), there is more to discover. It has been shown that the interactions between SLF and S-RNase proteins are the basis for SI, but it is not yet understood which domains of each of these proteins are responsible for these interactions. An important area of research in SI is to discover which amino acids are responsible for the necessary SLF/non-self S-RNase interactions required for successful fertilization.

We must first learn how SLFs and S-RNases interact before we can look at the sequences of these proteins and find which amino acids are responsible for this recognition. These interactions can be examined by testing individual, isolated SLF proteins. *Petunia inflata* are transformed with a specific *SLF* gene, and the transgenic plants are crossed into a wild type *P. inflata* plant with a different *S*-haplotype. Self-pollinations of these plants can then test if the SLF in question can interact with the S-RNases produced by the wild type plant.

To create transgenic plants to test these interactions, the specific *SLF* gene must be isolated from the genomic DNA and inserted into a vector, pBI101 in this case, to produce a construct. The construct includes a promoter region, the *SLF* transgene, a Green Fluorescent Protein (GFP) marker, and a terminator region. *LAT52* is a pollen-specific promoter from tomato plants, which has been used to successfully express *SLF1* in transgenic plants. The *SLF* transgene is the gene being studied. The GFP marker is inserted into the construct to provide a means to test for the presence and expression of the transgene. The transcription terminator, *Nos-ter*, is from the gene encoding nopaline synthase and is frequently used in genetically modified plants.

Once the constructs have been made, they must be transformed into *P. inflata* plants to study their effects. This is accomplished by *Agrobacterium*-mediated transformation. This bacterium functions to induce rapid mitosis in injured plant tissue forming tumors, which are then induced to grow roots and shoots using plant growth hormones. These shoots grow into plants which can be analyzed. After confirming the presence of the transgene and viability of the plant, the transgenic plants can be tested for SI (Hua et al., 2007).

Through this method, it has been shown that SI may be broken down in transgenic plants that express an SLF protein not normally produced in their pollen. That is, when a specific *SLF*

gene is introduced into a transgenic plant, which possesses two *S*-haplotypes, one the same as the *S*-haplotype from which *SLF* is derived and the other a different *S*-haplotype, the transgenic plant may or may not remain self-incompatible (Stout and Chandler, 1941; Sijacic et al., 2004). For example, when *S_2-SLF1* is introduced to an *S_2S_3* plant, the resulting transgenic plant becomes self-compatible (SC). However, if *S_2-SLF1* is introduced to an *S_2S_2* plant, the resulting transgenic plant remains self-incompatible. In the case of *S_2S_3* transgenic plants producing S_2-SLF1, the 17 SLF proteins produced by *S_3* pollen will be able to interact with S_2-RNase and mediate its degradation, but will not be able to interact with S_3-RNase. However, as S_2-SLF1 does have the ability to interact with S_3-RNase and mediate its degradation, the *S_3* transgenic pollen that produces S_2-SLF1 will be able to interact with both S_2-RNase and S_3-RNase to effect successful fertilization, when the transgenic pollen is used to self-pollinate the pistil.

Current SI research in the Kao Lab seeks to systematically determine all the interaction relationships between the 10 SLF proteins of *S_2*-haplotype and 11 S-RNases of *Petunia inflata* available, as well as between the 10 SLF proteins of *S_3*-haplotype and these 11 S-RNases. The results obtained so far (Sijacic et al., 2004; Sun and Kao, 2013; Williams et al., 2014) are summarized in Table 1.

Table 1. Current SLF/S-RNase Interactions.

"Yes" indicates an interaction, "No" indicates no interaction.

		S-RNase										
SLF		S_2	S_3	S_5	S_{6a}	S_7	S_{11}	S_{12}	S_{13}	S_{16}	S_{22}	S_{24}
	S_2-SLF1	No	Yes	No	No	Yes	No	Yes	Yes	No	No	
	S_2-SLF2	No	No									
	S_2-SLF3	No	No	No	No	No	No	No	No	No		
	S_2-SLF4	No	No	Yes	No	No	No	No	No	No		
	S_2-SLF5	No	No	No	No	No	No	Yes	No	No		
	S_2-SLF6	No	No	No	No	No	No	No	No	No		
	S_2-SLF7											
	S_2-SLF8	No	No	No	Yes	No	No	No	No	No		
	S_2-SLF9	No	No	No	No	No	No	No	No	No	No	No
	S_2-SLF10	No	No	No	No	No	No	No	No	No	No	No
		S_2	S_3	S_5	S_{6a}	S_7	S_{11}	S_{12}	S_{13}	S_{16}	S_{22}	S_{24}
	S_3-SLF1	No	No	No	No	No	No	Yes	No	No		
	S_3-SLF2											
	S_3-SLF3											
	S_3-SLF4											
	S_3-SLF5	No	No					No				
	S_3-SLF6		No									
	S_3-SLF7											
	S_3-SLF8											
	S_3-SLF9	No	No	No	No	No	No		No	No		
	S_3-SLF10	No	No	No	No	No	No		No	No		

Comparing the consistencies and variations between different SLF interactions can be used to hypothesize which domains of SLF proteins are involved in specific interactions with S-RNases. For example, comparing the sequences of S_2-SLF1 and S_3-SLF1 reveals that, although they are approximately 88% conserved at the amino acid sequence level, they have varying interactions. Neither SLF interacts with S_2-RNase, S_5-RNase, S_{6a}-RNase, S_{11}-RNase, or S_{16}-RNase, and both SLFs interact with S_{12}-RNase. However, S_2-SLF1 interacts with S_3-RNase, S_7-RNase, and S_{13}-RNase, but S_3-SLF1 does not. It stands to reason that altering the sequence of S_2-SLF1 to be more like S_3-SLF1 could prevent these interactions from occurring. Altering the sequence of S_2-SLF1systematically and analyzing the changes in interactions caused by each part

of the sequence will demonstrate which particular segments of S_2-SLF1 are responsible for recognizing non-self S-RNases.

To study the differences between S_2-SLF1 and S_3-SLF1, a former graduate student in the Kao Lab, Ning Wang, divided the SLF1 sequence into three domains. These domains were formulated based on the hypothesized function of SLF proteins in general; F-box proteins typically consist of an F-box domain, responsible for interacting with the E3-ubiquitin ligase complex and a C-terminal domain, responsible for interacting with a protein substrate. The first domain, Functional Domain 1 or FD1, contains the F-box of the protein. This F-box domain interacts with SSK1 (a Skp1-like protein) to form the SCF-ubiquitin-ligase complex. FD1 is comprised of amino acids 1 to 130 of the 395 amino acids of SLF1. The final two domains form the C-terminal domain, which is believed to interact with S-RNase. The second domain, Functional Domain 2 or FD2, consists of amino acids 131 to 260, and the third domain, Functional Domain 3 or FD3, consists of amino acids 261 to 395. These three domains were used to create S_2-SLF1/S_3-SLF1 chimerics, which consist of various combinations of the S_2-SLF1 domains and the S_3-SLF2 domains. Each of these chimeric SLF proteins is designated by a three-digit number in which each of the three numbers depicts the *SLF1* that contributes the domain. For example, SLF(322) contains FD1 of S_3-SLF1, and FD2 and FD3 of S_2-SLF1. These chimerics are further abbreviated, with the letter "F" followed by the number, so the example above would be F322.

This study will analyze S-RNase interactions of two chimeric SLF proteins, F322 and F232. Each of these chimeric SLF proteins consists of S_2-SLF1 except for a domain from S_3-SLF1. Comparing the S-RNase interactions of these two chimeric SLF proteins to the original S_2-SLF1 and S_3-SLF1 will show the impact this swapping of domains has had. If one of these

chimeric SLFs interacts as S_3-SLF1 does, this would show that the S_3-SLF1 domain affects S-RNase recognition. In that case, it would show that the amino acids responsible for S_2-SLF1 interactions with non-self S-RNases are located within the domain that had been swapped. For example, S_2-SLF1 interacts with S_7-RNase, but S_3-SLF1 does not. If F232 does not interact with S_7-RNase, it would be acting as S_3-SLF1 would, suggesting that swapping in FD2 of S_3-SLF1 changes the way this SLF interacts with S_7-RNases. This would indicate that the amino acids that recognize S_7-RNase are located within this second domain. This method of analysis will be used to determine which domain contains amino acids of significance.

Isolating and determining the interactions of various SLFs will advance our knowledge of SI recognition as well. This study also seeks to identify the amino acid sequences of S_2-SLF5 that are involved in its specific interaction with a particular S-RNase. It is known that the amino acid sequences of S_2-SLF5 and S_3-SLF5 are 98% identical, yet S_2-SLF5 interacts with S_{12}-RNase and S_3-SLF5 does not. A number of S_2-SLF5 and S-RNase interaction relationships have been determined, but more need to be determined for S_3-SLF5 in order to continue this research. Justin Williams previously generated the *S_3-SLF5* construct to determine these interactions. This construct, shown in Figure 3, contains a *LAT52* promoter, the *S_3-SLF5* gene, a *GFP* marker, and *Nos* terminator, inserted into a pBI101 vector. To determine the interaction relationships of S_3-SLF5, the construct will be analyzed as outlined for F232 and F322. As more information is gathered about S_3-SLF5 interactions, it can be compared to S_2-SLF5. The differences between these two *SLFs* can be assessed on the genetic level as well to determine the reasons behind these differences.

Materials and Methods

Seed Germination

Forty seeds were placed in a 1.7 mL Eppendorf tube with 1 mL of .005% giberellic acid. After 30 minutes, the seeds and solution were placed onto round filter paper and placed into a petri dish. The dish was sealed with parafilm and covered completely with aluminum foil. The plate was left at 26ºC in a growth chamber with a 16 hour light cycle. After three days, the aluminum foil was removed and the plate was left under the same conditions until the seeds began to sprout roots, usually a four-day period. The seeds were transferred to soil and left under light at 26ºC as they grew into plants.

Genomic DNA Isolation

Genomic DNA was isolated using Plant DNAzol® Reagent by Invitrogen, according to the manufacturer's protocol and as described by Meng et al. (2011). Samples were then frozen at -20ºC and used for further analysis.

Genotyping by PCR and Gel Electrophoresis

Polymerase chain reaction (PCR) was used to verify the presence of the desired transgene in all samples through the green fluorescent protein (GFP) marker, as well as to determine the *S*-genotype of all plants. The primers used in these experiments to genotype all of the plants were *S*-haplotype specific for *S-RNase* or *SLF1*. PCR conditions for each primer are listed under

Polymerase Chain Reaction Conditions. The samples were then analyzed through gel electrophoresis. These experiments were conducted according to the protocol described by Meng et al. (2011)

Polymerase Chain Reaction Conditions

Table 2. PCR Conditions for 11 S-haplotypes and GFP.

S_2		
Temperature (°C)	Time	
95	5 min	
95	30 sec	Repeat 40 times
55	30 sec	
72	2 min	
72	10 min	
4	infinite	

S_3		
Temperature (°C)	Time	
95	5 min	
95	30 sec	Repeat 40 times
55	30 sec	
72	2 min	
72	10 min	
4	infinite	

S_5		
Temperature (°C)	Time	
95	5 min	
95	30 sec	Repeat 32 times
60	30 sec	
72	45 sec	
72	10 min	
4	infinite	

S_{6a}		
Temperature (°C)	Time	
95	5 min	
95	30 sec	Repeat 32 times
56	30 sec	
72	45 sec	
72	10 min	
4	infinite	

S_7		
Temperature (°C)	Time	
95	5 min	
95	30 sec	Repeat 40 times
60	30 sec	
72	2 min	
72	10 min	
4	infinite	

S_{11}		
Temperature (°C)	Time	
95	5 min	
95	30 sec	Repeat 40 times
55	30 sec	
72	2 min	
72	10 min	
4	infinite	

S_{12}		
Temperature (°C)	Time	
95	5 min	
95	30 sec	Repeat 32 times
56	30 sec	
72	45 sec	
72	10 min	
4	infinite	

S_{13}		
Temperature (°C)	Time	
95	5 min	
95	30 sec	Repeat 32 times
60	30 sec	
72	45 sec	
72	10 min	
4	infinite	

S_{16}		
Temperature (°C)	Time	
95	5 min	
95	30 sec	Repeat 40 times
55	30 sec	
72	2 min	
72	10 min	
4	infinite	

GFP		
Temperature (°C)	Time	
95	5 min	
93	30 sec	Repeat
54	30 sec	32
72	45 sec	times
72	10 min	
4	infinite	

Validate Transgene Expression

Pollen germination medium (Lee et al., 1996) was used to germinate mature pollen for a maximum of 2 hours as described by Meng et al. (2009). A JENOPTIK ProgRes C14plus camera on an Olympus S2X16 microscope was used to visualize pollen tubes at 110x magnification.

Amino Acid Sequence Alignment and Analysis

Mega 5.1 (Tamura et al., 2011) and ClustalW (Thompson et al., 1994) were used to perform all sequence alignments and the default settings of these programs were applied to produce alignments of deduced amino acid sequences.

Results

Analysis of S_2-SLF1 non-self S-RNase Recognition through Chimeric F232

In order to test which domain of S_2-SLF1 may be responsible for interacting with non-self S-RNases, and in particular the domain termed FD2, the chimeric SLF1 protein termed F232 was studied. F232 consists of FD1 from S_2-SLF1, FD2 from S_3-SLF1, and FD3 from S_2-SLF1. The gene encoding this chimeric protein was inserted into the pBI101 vector, along with a *LAT52* promoter, *GFP* tag, and *NOS* terminator region. A schematic of this construct is shown in Figure 2. Transgenic plants expressing this chimeric protein were previously crossed out into S_5, S_{6a}, S_7, S_{11}, S_{12}, S_{13}, S_{16}, S_{22}, and S_{24} genetic backgrounds by Justin Williams. The progeny of these crosses were genotyped, checked for GFP expression, and analyzed for their SI behavior. The results for S_5, S_{6a}, S_{12}, and S_{16} can be seen in Tables 3 and 6 and in Figures 4A, B, C, and 5A, B, C, D whereas the results for S_7, S_{11}, S_{13}, S_{22}, and S_{24} are still pending. Comparing these interactions with those of S_2-SLF1 and S_3-SLF1 shows that all three SLFs do interact with S_{12}-RNase and none of them interact with S_5-RNase, S_{6a}-RNase or S_{16}-RNase. This shows that swapping in FD2 of S_3-SLF1 did not change the interaction specificities of S_2-SLF1, therefore the amino acid differences between the FD2 domains of these two SLFs may not play a role in these interactions. This also suggests that the amino acids that do affect recognition are located within FD1 and/or FD3. The amino acids in question are outlined in Figure 6B, and Figure 6A displays the genetic differences between *S_2-SLF1* and *S_3-SLF1*.

Table 3. *F232* Transgenic Plants Compatibility Data.

Plants containing the constructed transgene with *SLF1 (232)* were crossed with various wild type *S*-haplotype *P. inflata* plants. The progeny from these crosses were then tested to ensure they contained the transgene, and to determine their genotype. Each plant was also analyzed to confirm that it was expressing the transgene as well as viable for successful fertilization. Finally, the plants were tested for self-incompatibility to determine if F232 interacted with their S-RNase. If the plant was determined to be self-incompatible (SI), it shows that F232 does not interact with self-pollen. In self-compatible plants, F232 is able to interact with self-pollen. Gray regions indicate that results are in progress.

Plant Name	Pistil Check	Pollen Check	PCR GFP	Pollen GFP	SI/SC	Genotype Confirmed
F232 #6			Yes	Yes	SI	S_2S_5
F232 #13			Yes	Yes	SI	S_2S_5
F232 #1A			Yes	Yes	SI	S_2S_{6a}
F232 #4			Yes	Yes	SI	S_2S_{6a}
F232 #1	Yes	Yes	Yes	Yes	SC	S_2S_{12}
F232 #7	Yes	Yes	Yes	Yes	SC	S_2S_{12}
F232 #9	Yes	Yes	Yes	Yes	SC	S_2S_{12}
F232 #21			Yes	Yes	SC	S_2S_{12}
F 232 #28			Yes	Yes	SC	S_2S_{12}
F232 #9	Yes	Yes	Yes	Yes	SI	S_2S_{16}
F232 #13	Yes	Yes	Yes	Yes	SI	S_2S_{16}

Analysis of S_2-SLF1 non-self S-RNase Interactions through Chimeric F322

S_2-SLF1 was studied to determine which domain encodes the amino acids involved in recognizing non-self S-RNases. To do so, the chimeric *F322* was generated, which is comprised of *FD1* from *S_3-SLF1*, and *FD2* and *FD3* from *S_2-SLF1*. A schematic of the transgene construct is shown in Figure 2. These domains were spliced together and inserted into a transgene containing a *LAT52* promoter, a *GFP* tag, and a *NOS* terminator. After splicing the transgene into a pBI101 vector, it was transformed into *Petunia inflata*. Justin Williams had previously crossed out transgenic plants carrying this chimeric gene into *S_5*, *S_{6a}*, *S_7*, *S_{11}*, *S_{12}*, *S_{13}*, *S_{16}*, *S_{22}*, and *S_{24}* genetic backgrounds. After being genotyped and analyzed for GFP, the progeny of these crosses were tested for their SI behavior; the results obtained are shown in Tables 4 and 6 and in Figures 4D, E, F and 5E, F. The results for *S_5*, *S_{6a}*, *S_{11}*, *S_{13}*, *S_{16}*, *S_{22}*, and *S_{24}* are still being determined, but

the results for S_7 and S_{12} show that F322 follows the interaction pattern of S_2-SLF1. All three SLFs interact with S_{12}-RNase; S_2-SLF1 and F322 interact with S_7-RNase, but S_3-SLF1 does not. This provides evidence that FD1 does not contain amino acids that are involved in the interactions with S_7-RNase and S_{12}-RNase. This also shows that the amino acids that are an integral part of these interactions may be located in FD2 and/or FD3. The amino acids in question are outlined in Figure 6B and Figure 6A displays the genetic differences between these allelic variants.

Table 4. *F322* Transgenic Plants Compatibility Data.

Various wild type *S*-haplotype plants were crossed with transgenic plants containing the *F322* construct. Progeny from this cross were grown and examined to determine if they contained the transgene as well. Those with the transgene were tested to ensure they were viable and expressing the gene. Self-pollinations were carried out to determine whether the plants were self-incompatible (SI) or self-compatible (SC). In SI plants, F322 is unable to interact with self-pollen, but the two do interact in SC plants. Gray regions indicate that results are pending.

Plant Name	Pistil Check	Pollen Check	PCR GFP	Pollen GFP	SI/SC	Genotype Confirmed
F322 #2			Yes	Yes	SC	S_2S_7
F322 #13			Yes	Yes	SC	S_2S_7
F322 #28			Yes	Yes	SC	S_2S_{12}

Analysis of S_3-SLF5 non-self S-RNase Interactions

To learn more about SLF proteins, S_3-SLF5 was tested to determine its interactions with several non-self S-RNase, and the results were compared with those of S_2-SLF5. *S_3-SLF5* was isolated from the *S_3*-haplotype *S*-locus and spliced into a construct made from the *pBI101* vector. This vector, which is shown in Figure 3, also contained the *LAT52* promotor region, as well as a *GFP* tag and a *NOS* terminator. Through *Agrobacterium*-mediated transformation, this transgene construct was introduced into *Petunia inflata*. The transgenic plants were crossed out into *S*-genotypic backgrounds *S_5*, *S_{6a}*, *S_7*, *S_{11}*, *S_{12}*, *S_{13}*, *S_{16}*, *S_{22}*, and *S_{24}*. Progeny were genotyped and

tested for GFP expression before they were assessed for their SI behavior. Tables 5 and 6 and Figures 4G, H, I, J and 5G, H, I show the results of these analyses and interactions. The results obtained so far show that S_3-SLF5 did not interact with S_5-RNase, S_7-RNase, S_{12}-RNase, or S_{13}-RNase. Analysis of the interactions with S_{6a}-RNase, S_{11}-RNase, S_{16}-RNase, S_{22}-RNase, and S_{24}-RNase are still in progress. Comparing S_3-SLF5 to S_2-SLF5 shows that neither SLF interacts with S_5-RNase, S_7-RNase, S_{12}-RNase, or S_{13}-RNase, which may be expected, as these two SLFs are 98% identical at the amino acid sequence level. This shows that the amino acid differences between these two allelic variants do not play a role in the interactions with these four S-RNases. These amino acids, as well as genetic differences, are outlined in Figure 6C, D.

Table 5. *S_3-SLF5* Transgenic Plants Compatibility Data.

The *S_3-SLF5* construct was transformed into *P. inflata* plants. These plants were crossed with a number of wild type *S*-haplotype plants to produce progeny with various genotypes. After confirming the presence of the transgene in these progeny, their genotypes were determined as well. They were tested to ensure they were expressing the transgene and were viable, and in those that were, self-pollinations were performed. In all of these crosses, S_3-SLF5 was unable to interact with self-pollen and all of the progeny were self-incompatible (SI). Gray regions and *$S_?$*-haplotypes indicate that results are still pending.

Plant Name	Pistil Check	Pollen Check	PCR GFP	Pollen GFP	SI/SC	Genotype Confirmed
S_3SLF5 #1	Yes	Yes	Yes	Yes	SI	$S_?S_5$
S_3SLF5 #6	Yes	Yes	Yes	Yes	SI	$S_?S_5$
S_3SLF5 #11	Yes	Yes	Yes	Yes	SI	$S_?S_5$
S_3SLF5 #18		Yes	Yes	Yes	SI	$S_?S_5$
S_3SLF5 #1		Yes	Yes	Yes	SI	S_3S_7
S_3SLF5 #6	Yes	Yes	Yes	Yes	SI	S_3S_7
S_3SLF5 #6			Yes	Yes	SI	S_3S_7
S_3SLF5 #8	Yes		Yes	Yes	SI	S_3S_7
S_3SLF5 #2			Yes	Yes	SI	S_3S_{13}
S_3SLF5 #11	Yes	Yes	Yes	Yes	SI	S_3S_{13}

Additional Data

pBI101-*LAT52P: SLF1 (232): GFP*
pBI101-*LAT52P: SLF1 (322): GFP*

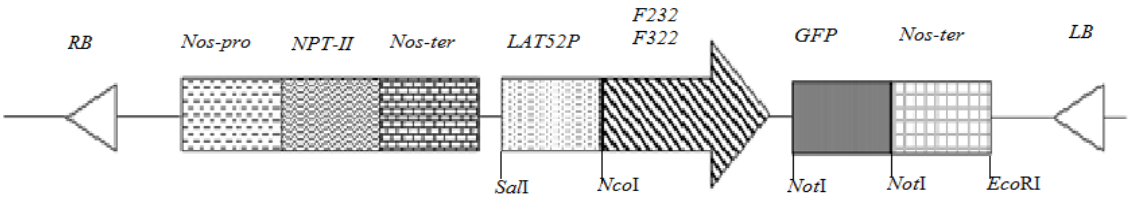


Figure 2. Structure of chimeric constructs used to genetically modify *P. inflata.*
The *F232* and *F322* constructs were generated as described by Hua et al. (2007).

pBI101-*LAT52P: S_3SLF5: GFP*

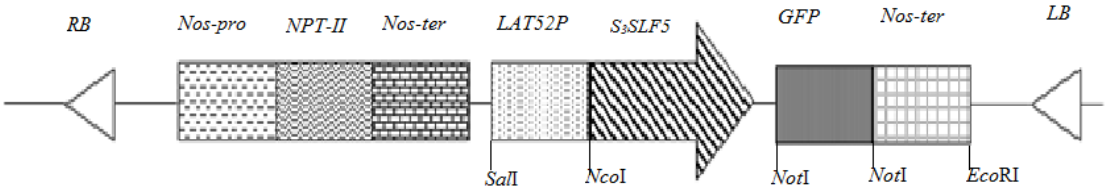
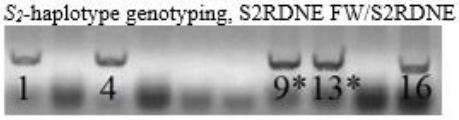


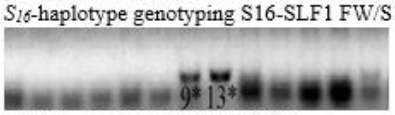
Figure 3. Structure of *S_3-SLF5* construct used to genetically modify *P. inflata.*
The In-fusion HD® Cloning kit (Clontech) was used to generate the construct, as described in Sun and Kao (2013). Right border of the DNA, *RB*; the nopaline synthase gene promoter, *Nos-pro*; the neomycin phosphotransferase II (kanamycin resistance) gene, *NPT-II*; the nopaline synthase gene transcription terminator, *Nos-ter*; pollen specific *LAT52* tomato gene promoter, *LAT52P*; Green Fluorescent Protein gene, *GFP*; left border of the T-DNA, *LB*.

A. pBI101-*LAT52P:F232:GFP/S7S16*
Transgene verification, GFP001 FW/GFP500 REV

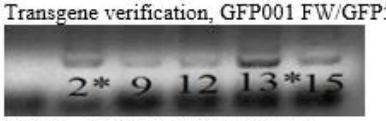
B. pBI101-*LAT52P:F232:GFP/S7S16*
S2-haplotype genotyping, S2RDNE FW/S2RDNE REV



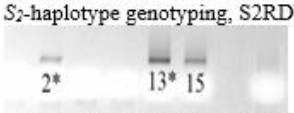
C. pBI101-*LAT52P:F232:GFP/S7S16*
S16-haplotype genotyping S16-SLF1 FW/S16-SLF1 REV



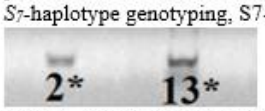
D. pBI101-*LAT52P:F322:GFP/S7S16*
Transgene verification, GFP001 FW/GFP500 REV



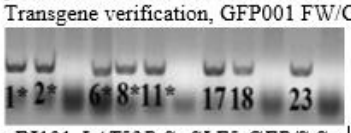
E. pBI101-*LAT52P:F322:GFP/S7S16*
S2-haplotype genotyping, S2RDNE FW/S2RDNE REV



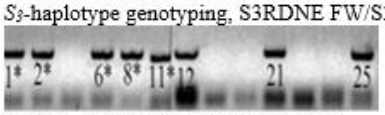
F. pBI101-*LAT52P:F322:GFP/S7S16*
S7-haplotype genotyping, S7-SLF1FW/S7-SLF1REV



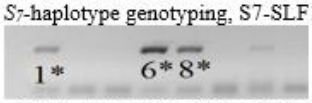
G. pBI101-*LAT52P:S3-SLF5:GFP/S7S13*
Transgene verification, GFP001 FW/GFP500 REV



H. pBI101-*LAT52P:S3-SLF5:GFP/S7S13*
S3-haplotype genotyping, S3RDNE FW/S3RDNE REV



I. pBI101-*LAT52P:S3-SLF5:GFP/S7S13*
S7-haplotype genotyping, S7-SLF1FW/S7-SLF1REV



J. pBI101-*LAT52P:S3-SLF5:GFP/S7S13*
S13-haplotype genotyping, S13-SLF1FW/S13-SLF1REV

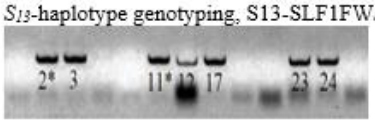


Figure 4. *GFP* PCR Results.

Samples were analyzed by PCR and gel electrophoresis to determine genotype. In this figure, the dark bands represent a positive result. The pictures are labeled by SLF X *S*-genotypic background and gene being tested for.

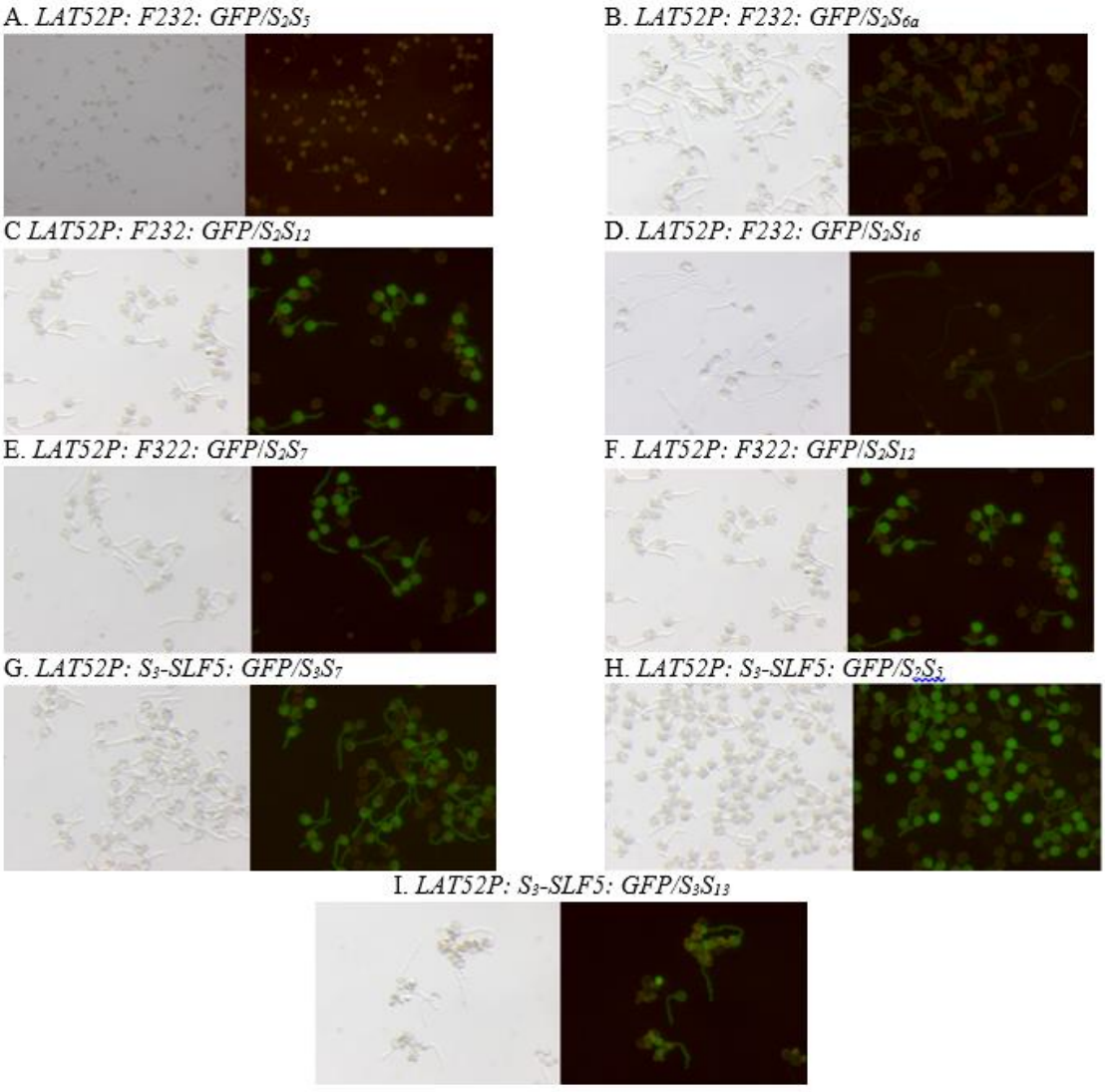
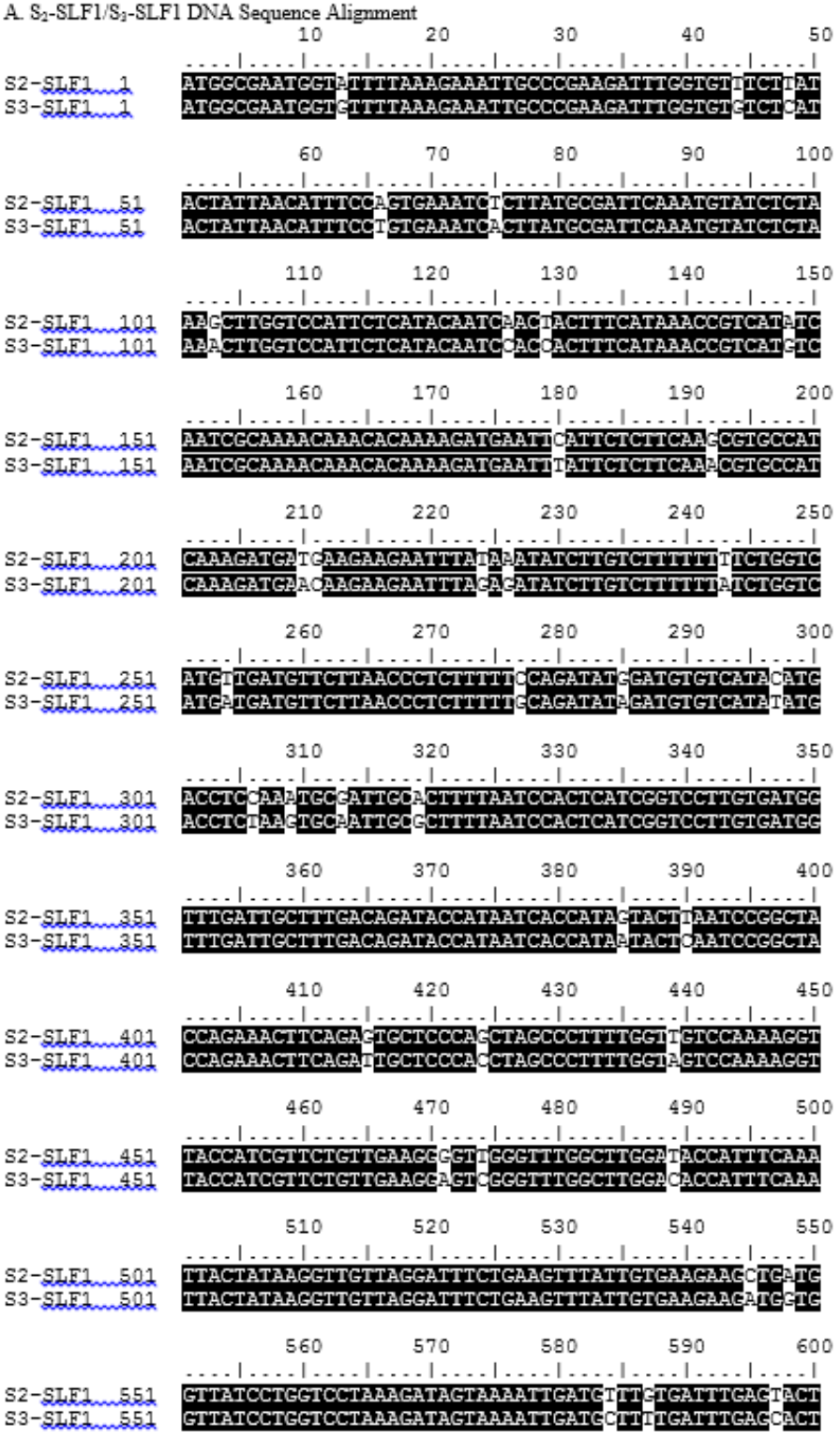
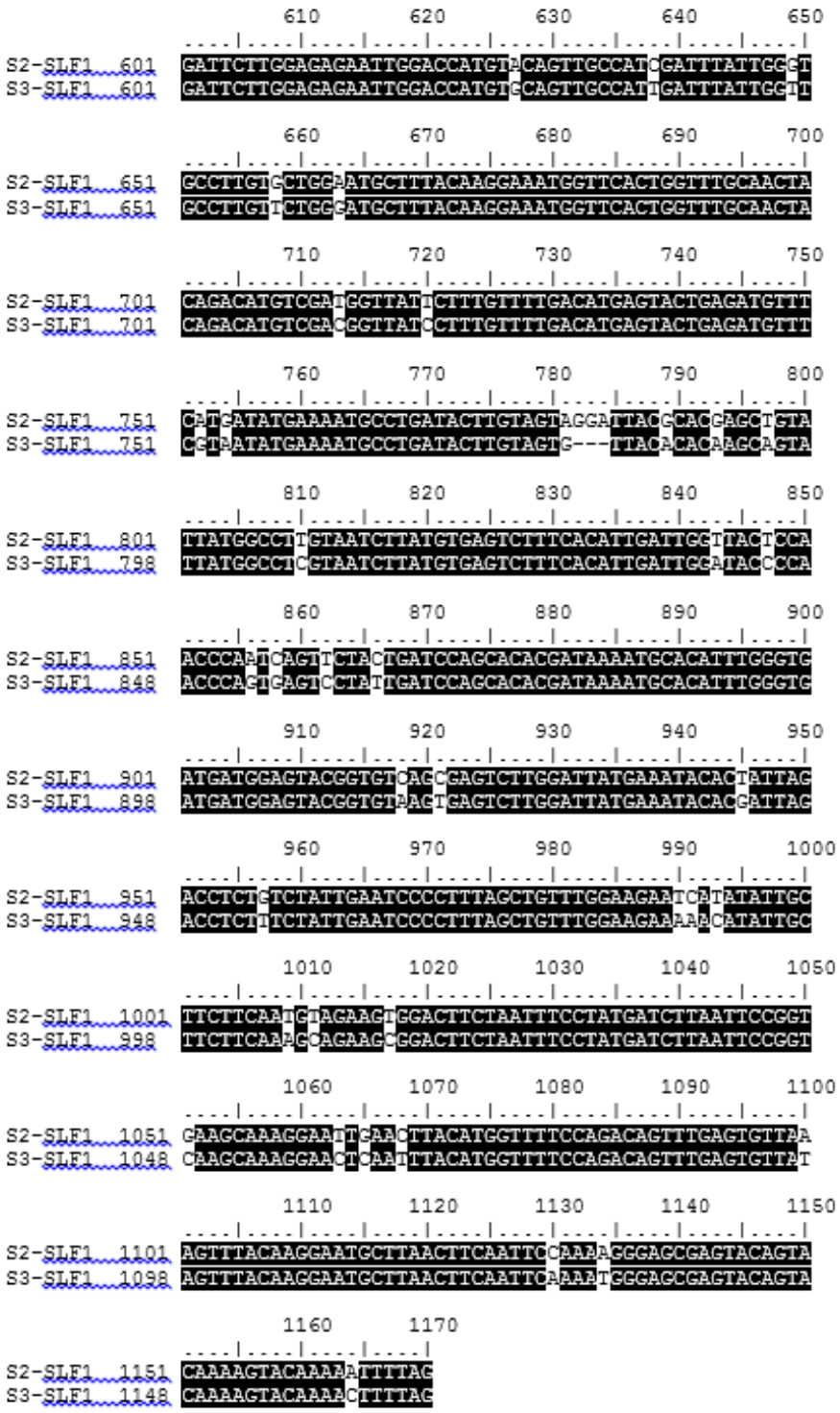


Figure 5. GFP Expression Assay

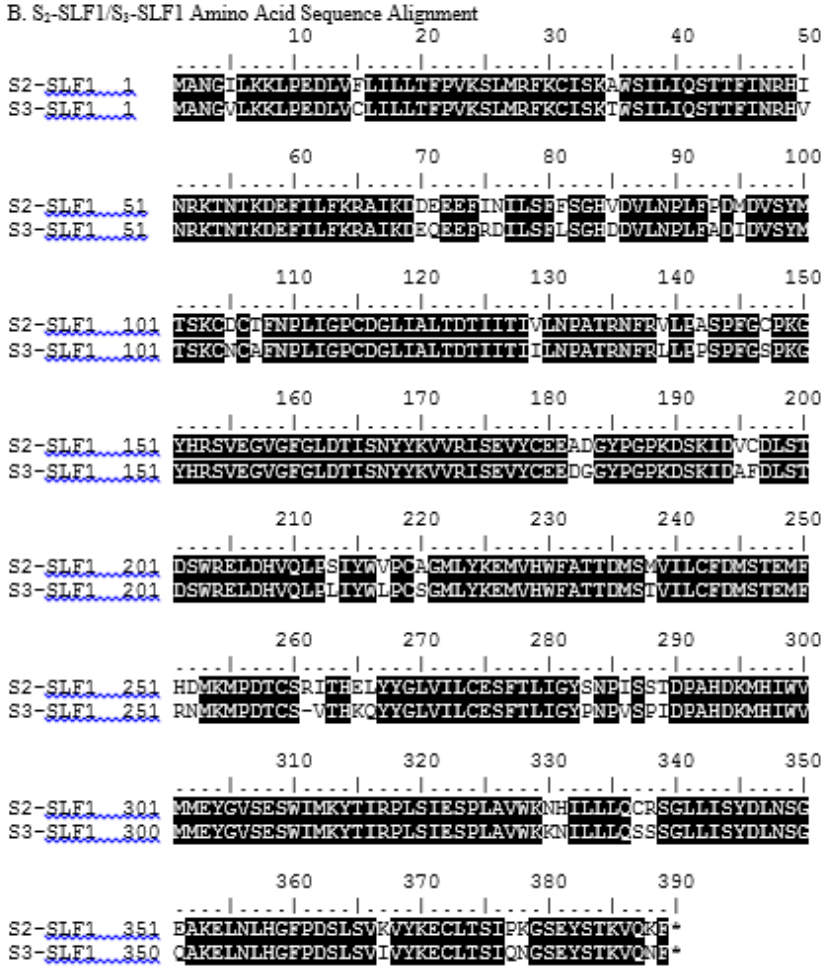
Pollen was induced to grow tubes that were examined under a fluorescence microscope. Fluorescence of these tubes confirms the expression of the transgene. This figure shows a comparison of pollen growth tubes in visible light to the same tubes under fluorescent light.

A. S_2-SLF1/S_3-SLF1 DNA Sequence Alignment

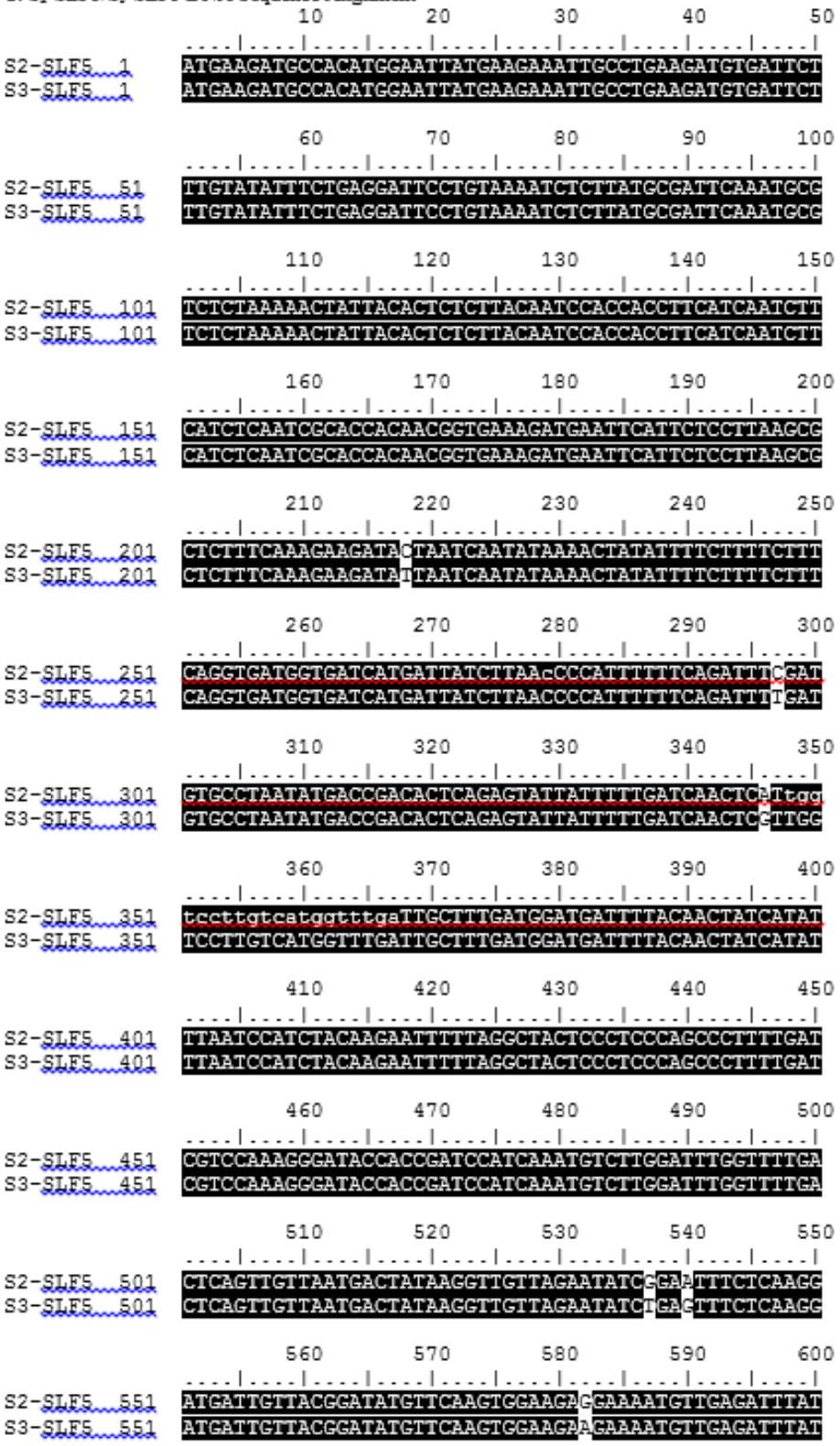


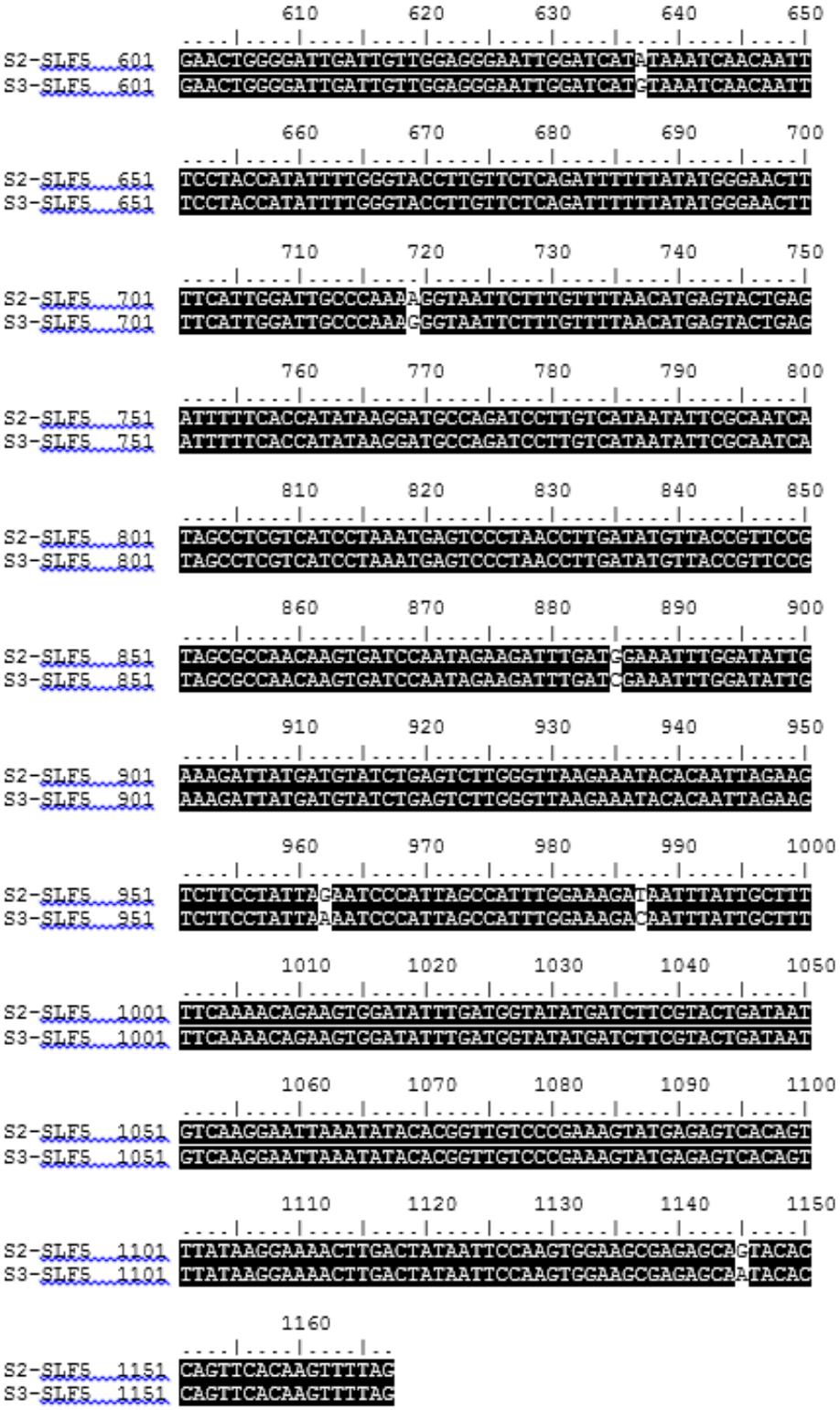

610 620 630 640 650
S2-SLF1_601 GATTCTTGGAGAGAATTGGACCATGTACAGTTGCCATCGATTTATTGGGT
S3-SLF1_601 GATTCTTGGAGAGAATTGGACCATGTGCAGTTGCCATTGATTTATTGGTT
660 670 680 690 700
S2-SLF1_651 GCCTTGTGCTGGAATGCTTTACAAGGAAATGGTTCACTGGTTTGCAACTA
S3-SLF1_651 GCCTTGTTCTGGGATGCTTTACAAGGAAATGGTTCACTGGTTTGCAACTA
710 720 730 740 750
S2-SLF1_701 CAGACATGTCGATGGTTATTCTTTGTTTTGACATGAGTACTGAGATGTTT
S3-SLF1_701 CAGACATGTCGACGGTTATCCTTTGTTTTGACATGAGTACTGAGATGTTT
760 770 780 790 800
S2-SLF1_751 CATGATATGAAAATGCCTGATACTTGTAGTAGGATTACGCAGGAGCTGTA
S3-SLF1_751 CGTAATATGAAAATGCCTGATACTTGTAGTG---TTACACACAAGCAGTA
810 820 830 840 850
S2-SLF1_801 TTATGGCCTTGTAATCTTATGTGAGTCTTTCACATTGATTGGTTACTCCA
S3-SLF1_798 TTATGGCCTCGTAATCTTATGTGAGTCTTTCACATTGATTGGATACCCCA
860 870 880 890 900
S2-SLF1_851 ACCCAATCAGTTCTACTGATCCAGCACACGATAAAATGCACATTTGGGTG
S3-SLF1_848 ACCCACTCAGTCCTATTGATCCAGCACACGATAAAATGCACATTTGGGTG
910 920 930 940 950
S2-SLF1_901 ATGATGGAGTACGGTGTCAGCGAGTCTTGGATTATGAAATACACTATTAG
S3-SLF1_898 ATGATGGAGTACGGTGTAAGTGAGTCTTGGATTATGAAATACACGATTAG
960 970 980 990 1000
S2-SLF1_951 ACCTCTCTCTATTGAATCCCCTTTAGCTGTTTGGAAGAATCATATATTGC
S3-SLF1_948 ACCTCTTTCTATTGAATCCCCTTTAGCTGTTTGGAAGAAAAACATATTGC
1010 1020 1030 1040 1050
S2-SLF1_1001 TTCTTCAATGTAGAAGTGGACTTCTAATTTCCTATGATCTTAATTCCGGT
S3-SLF1_998 TTCTTCAAAGCAGAAGCGGACTTCTAATTTCCTATGATCTTAATTCCGGT
1060 1070 1080 1090 1100
S2-SLF1_1051 GAAGCAAAGGAATTGAACTTACATGGTTTTCCAGACAGTTTGAGTGTTAA
S3-SLF1_1048 CAAGCAAAGGAACTCAATTTACATGGTTTTCCAGACAGTTTGAGTGTTAT
1110 1120 1130 1140 1150
S2-SLF1_1101 AGTTTACAAGGAATGCTTAACTTCAATTGCAAAAGGGAGCGAGTACAGTA
S3-SLF1_1098 AGTTTACAAGGAATGCTTAACTTCAATTCAAAATGGGAGCGAGTACAGTA
1160 1170
S2-SLF1_1151 CAAAAGTACAAAAATTTTAG
S3-SLF1_1148 CAAAAGTACAAAACTTTTAG

B. S_2-SLF1/S_3-SLF1 Amino Acid Sequence Alignment



C. S_2-SLF5/S_3-SLF5 DNA Sequence Alignment




610 620 630 640 650
S2-SLF5 601 GAACTGGGGATTGATTGTTGGAGGGAATTGGATCATATAAATCAACAATT
S3-SLF5 601 GAACTGGGGATTGATTGTTGGAGGGAATTGGATCATGTAAATCAACAATT
660 670 680 690 700
S2-SLF5 651 TCCTACCATATTTTGGGTACCTTGTTCTCAGATTTTTTATATGGGAACTT
S3-SLF5 651 TCCTACCATATTTTGGGTACCTTGTTCTCAGATTTTTTATATGGGAACTT
710 720 730 740 750
S2-SLF5 701 TTCATTGGATTGCCCAAAAGGTAATTCTTTGTTTTAACATGAGTACTGAG
S3-SLF5 701 TTCATTGGATTGCCCAAAGGGTAATTCTTTGTTTTAACATGAGTACTGAG
760 770 780 790 800
S2-SLF5 751 ATTTTTCACCATATAAGGATGCCAGATCCTTGTCATAATATTCGCAATCA
S3-SLF5 751 ATTTTTCACCATATAAGGATGCCAGATCCTTGTCATAATATTCGCAATCA
810 820 830 840 850
S2-SLF5 801 TAGCCTCGTCATCCTAAATGAGTCCCTAACCTTGATATGTTACCGTTCCG
S3-SLF5 801 TAGCCTCGTCATCCTAAATGAGTCCCTAACCTTGATATGTTACCGTTCCG
860 870 880 890 900
S2-SLF5 851 TAGCGCCAACAAGTGATCCAATAGAAGATTTGATGGAAATTTGGATATTG
S3-SLF5 851 TAGCGCCAACAAGTGATCCAATAGAAGATTTGATCGAAATTTGGATATTG
910 920 930 940 950
S2-SLF5 901 AAAGATTATGATGTATCTGAGTCTTGGGTTAAGAAATACACAATTAGAAG
S3-SLF5 901 AAAGATTATGATGTATCTGAGTCTTGGGTTAAGAAATACACAATTAGAAG
960 970 980 990 1000
S2-SLF5 951 TCTTCCTATTAGAATCCCATTAGCCATTTGGAAAGATAATTTATTGCTTT
S3-SLF5 951 TCTTCCTATTAAAATCCCATTAGCCATTTGGAAAGACAATTTATTGCTTT
1010 1020 1030 1040 1050
S2-SLF5 1001 TTCAAAACAGAAGTGGATATTTGATGGTATATGATCTTCGTACTGATAAT
S3-SLF5 1001 TTCAAAACAGAAGTGGATATTTGATGGTATATGATCTTCGTACTGATAAT
1060 1070 1080 1090 1100
S2-SLF5 1051 GTCAAGGAATTAAATATACACGGTTGTCCCGAAAGTATGAGAGTCACAGT
S3-SLF5 1051 GTCAAGGAATTAAATATACACGGTTGTCCCGAAAGTATGAGAGTCACAGT
1110 1120 1130 1140 1150
S2-SLF5 1101 TTATAAGGAAAACTTGACTATAATTCCAAGTGGAAGCGAGAGCAGTACAC
S3-SLF5 1101 TTATAAGGAAAACTTGACTATAATTCCAAGTGGAAGCGAGAGCAATACAC
1160
S2-SLF5 1151 CAGTTCACAAGTTTTAG
S3-SLF5 1151 CAGTTCACAAGTTTTAG

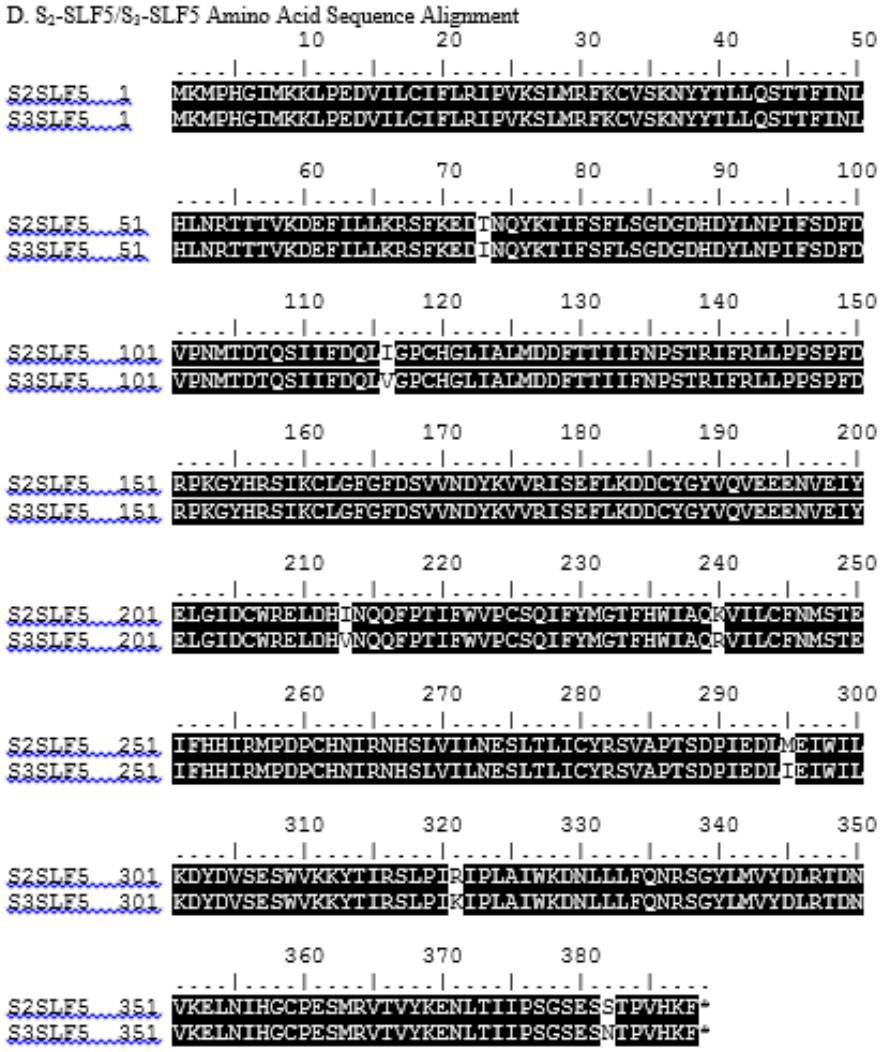


Figure 6. DNA and Amino Acid Sequence Alignments

Alignments of determined DNA and amino acid sequences for allelic variants of SLF1 (A, B) and SLF5 (C, D).

Table 6. Summary of Interactions.

The results from individual transgenic plants were compiled and compared to generate a summary of interactions. Comparing the overall results from each S-RNase reveals which SLFs and S-RNases interact, and which do not. These results can be used to determine which domains affect the recognition and interaction between SLFs and S-RNases in *Petunia inflata*. * indicates an interaction determined prior to these experiments. Gray regions indicate that the interaction has not yet been determined.

		SLF		
S-RNase		F232	F322	S_3-SLF5
	S_2	SI (-)*	SI (-)*	SI (-)*
	S_3	SC (+)*	SC (+)*	SI (-)*
	S_5	SI (-)		SI (-)
	S_{6a}	SI (-)		
	S_7		SC (+)	SI (-)
	S_{11}			
	S_{12}	SC (+)	SC (+)	SI (-)*
	S_{13}			SI (-)
	S_{16}	SI (-)		
	S_{22}			
	S_{24}			

Discussion

The first gene shown to be involved in pollen specificity in *Petunia* was *PiSLF* (*P. inflata S-Locus F-box*) (Sijacic et al., 1994). After further experiments revealed that there were additional *SLF* genes involved, *PiSLF* was renamed *SLF1*, and to date, an additional 16 *SLF* genes have been identified (Kubo et al., 2010; Williams et al., 2014). Among the SLFs examined, S_2-SLF1 has the most known interactions with non-self S-RNases. S_2-SLF1 and S_3-SLF1 have been analyzed extensively and their interactions have been determined for nearly all of the 11 S-RNases available. This makes these two SLFs strong candidates for determining which parts of these proteins are involved in S-RNase recognition. The 3 S-RNases where their interactions vary, S_3-RNase, S_7-RNase, and S_{13}-RNase, are particularly important. The goal of the research on these two SLFs is to manipulate their sequences to either cause an interaction which the original SLF did not have, or to eliminate a previously existing interaction. Obtaining either of these results would indicate that the part of the protein which had been swapped contains the amino acids responsible for S-RNase recognition. The SLF1 protein has been divided into three domains, FD1, FD2, and FD3. To determine which of these domains is essential for S-RNase interactions in S_2-SLF1, a domain from S_3-SLF1 has been swapped into the S_2-SLF1 sequence to generate SLF1 chimerics F232 and F322.

Chimeric *F232* was created by splicing *FD1* from *S_2-SLF1*, *FD2* from *S_3-SLF1*, and *FD3* from *S_2-SLF1* together. The construct was transformed into *Petunia inflata* using *Agrobacterium*. The resulting transgenic plants were crossed out with wild type plants to transfer the transgene into *S_5*, *S_{6a}*, *S_7*, *S_{11}*, *S_{12}*, *S_{13}*, *S_{16}*, *S_{22}*, and *S_{24}* genetic backgrounds. Due to time constraints, it was

not possible to complete the analysis of all of the *S*-genetic backgrounds. Full results were obtained for analysis in four of the genetic backgrounds, in which S_5-RNase, S_{6a}-RNase, S_{12}-RNase, and S_{16}-RNase are produced. Analysis of the interactions with these four S-RNases allows the chimeric to be tested in situations where both positive and negative results are expected. Neither S_2-SLF1 nor S_3-SLF1 interacts with S_5-RNase, S_{6a}-RNase, or S_{16}-RNase, yet both SLFs interact with S_{12}-RNase. Ideally, results would have been determined for an additional S-RNase that is recognized by only one of the SLFs, such as S_7-RNase or S_{13}-RNase; however, those results are still pending. It is expected that F232 would interact with S_{12}-RNase and would not interact with S_5-RNase, S_{6a}-RNase or S_{16}-RNase.

The transgenic progeny were genotyped and examined for GFP expression before being tested for their SI behavior; the results are shown in Figures 4A, B, C and 5A, B, C, D and Tables 3 and 6. These results show that the F232 chimeric protein does interact with S_{12}-RNase, but not with S_5-RNase, S_{6a}-RNase, or S_{16}-RNase. The F232 chimeric acts as expected by recognizing none of the S-RNases except for S_{12}-RNase. This tells us that despite the presence of the FD2 domain from S_3-SLF1, the chimeric is following the interaction pattern of S_2-SLF1. This implies that the amino acids responsible for S-RNase recognition are likely not present in FD2. However, S_3-SLF1 also recognizes the same S-RNases, therefore these results cannot be considered conclusive.

The F322 chimeric is comprised of FD1 from S_3-SLF1, FD2 and FD3 from S_2-SLF1, and the gene was spliced into a pBI101 vector and transformed into *Petunia inflata* using *Agrobacterium*. These transgenic plants were crossed out into *S*-genetic backgrounds S_5, S_{6a}, S_7, S_{11}, S_{12}, S_{13}, S_{16}, S_{22}, and S_{24}. Results were not obtained for all these genetic backgrounds due to time constraints; however, full results were obtained for S_7 and S_{12}. These results allow testing S-

RNase interactions in two situations where positive results are expected, but a negative result may be obtained in one instance. S_2-SLF1 interacts with both S_7-RNase and S_{12}-RNase, but S_3-SLF1 only interacts with S_{12}-RNase. Ideally, an S-RNase that does not interact with either SLF, such as S_5-RNase or S_{16}-RNase, would have yielded useful results as well, but these results are yet to be determined.

The progeny of these crosses were analyzed for GFP expression and genotyped. After which, they were tested for their SI behavior. The results of all these analyses are shown in Tables 4 and 6 and Figures 4D, E, F and 5E, F. F322 was found to interact with both S_7-RNase and S_{12}-RNase. The chimeric was expected to interact with S_{12}-RNase. The positive interaction between this SLF and S_7-RNase shows that the presence of FD1 from S_3-SLF1 does not affect the SLF proteins ability to recognize S_7-RNase. The FD2 and FD3 domains are both from S_2-SLF1, which suggests that FD1 does not contain the amino acids necessary for S_2-SLF1 and S_7-RNase interaction.

Based on the results obtained from F322 and F232, the amino acids required for S-RNase recognition are likely not located in FD1 or FD2. This implies that FD3 contains the amino acids that affect whether the SLF will recognize an S-RNase or not. Prior to this experiment, Justin Williams determined the interactions between these SLF chimerics and S_2-RNase and S_3-RNase. S_3-SLF1 interacts with neither of these S-RNases, but S_2-SLF1 interacts with S_3-RNase. F232 and F322 were determined to interact with S_3-RNase, but not S_2-RNase. The finding that swapping in FD1 or FD2 of S_3-SLF1 did not cause a change in the interaction of S_2-SLF1 with S_3-RNase provides additional evidence that FD3 contains the amino acids which affect S-RNase recognition. This is consistent with the theory that the final two domains of SLF proteins, the C-terminal domain, interact with S-RNase. Although more experiments must be completed to

confirm these results, they do correspond with current knowledge and theories regarding the SI mechanism.

The first step that must be taken to confirm these results is to determine additional interactions with chimeric F232. The results from F322 provide strong evidence that FD1 does not contain the amino acids of interest. However, the results of F232 do not provide a clear example of this chimeric interacting in a way that is specific to S_2-SLF1, and different from S_3-SLF1. S_2-SLF1 interacts with S_7-RNase and S_{13}-RNase, and S_3-SLF1 does not. If F232 were shown to interact with these S-RNases, it would provide clear evidence that this chimeric interacts as S_2-SLF1 does and that swapping in FD2 of S_3-SLF1 does not affect these interactions.

A progeny analysis must be completed to determine the genotype from all positive crosses and verify these results. The seeds must be grown into plants, and their genomic DNA must be isolated and then tested to confirm the presence of the transgene as well as their genotype. This genotyping ensures that the cross did yield the expected genotypes for testing SLF and S-RNase interactions. For example, plant F232 #1, with the genotype S_2S_{12}, was found to be self-compatible. We would expect the seeds from these crosses to yield 50% S_2S_{12} progeny and 50% $S_{12}S_{12}$ progeny, with 100% of these progeny containing the transgene. All of the plants in this experiment were stored in the same greenhouse room, and there is a possibility of accidental cross pollinations. If the progeny from these crosses consist of the expected genotype ratios, it confirms that the plant was self-compatible and eliminates the possibility of pollination by another plant.

To further explore these interactions, both the F232 and the F322 chimerics must be tested against more S-RNases. Previously, the interactions between these SLFs and S_2-RNase

and S_3-RNase were determined, and with these results, an additional six interactions have been found. There are still 12 interactions that must be tested to determine how closely each chimeric SLF follows S_2-SLF1 and to discover if either of these chimerics diverge from their expected interaction patterns. For example, neither S_2-SLF1 nor S_3-SLF1 interacts with S_{11}-RNase, and it is expected that neither F232 nor F322 would either. However, this cannot be confirmed until both of these chimerics are crossed out into an S_{11}-background and analyzed for the SI behavior of the progeny.

To confirm that FD3 contains the amino acids necessary for S-RNase recognition, more SLF1 chimerics must be made. The next two chimerics to be made are F223 and F332. F223 will contain FD1 and FD2 from S_2-SLF2 and FD3 from S_3-SLF1, and F332 will contain FD1 and FD2 from S_3-SLF1 and FD3 from S_2-SLF1. These chimerics are important because they are comprised of mostly one SLF1, but the critical FD3 domain is from a different allelic variant. If FD3 does contain the essential amino acids for these interactions, it would be expected that each chimeric would interact with the same S-RNases as the SLF1 allelic variant that contributes FD3. For example, although F332 contains two domains from S_3-SLF1, it would be expected to follow the interaction pattern of S_2-SLF1, because it contains FD3 from this allelic variant. In these experiments, F332 would be expected to yield positive results for S_3-RNase, S_7-RNase, S_{12}-RNase, and S_{13}-RNase. F223 would be expected to interact with only S_{12}-RNase and would serve as a negative control. If the interactions of F223 and F332 yield the expected results, this would strengthen the hypothesis that the amino acids involved in S-RNase recognition are located in FD3.

After further confirming that FD3 controls whether an SLF will interact with an S-RNase, it must be more deeply analyzed to determine which of the 134 amino acids in this domain are

directly responsible for these interactions. Justin Williams, a graduate student in the Kao Lab, has already begun to analyze the amino acid sequence of this domain. FD3 has been divided into four subdomains. These subdomains will be tested using a similar approach as that used to study F232 and F322. Different combinations of these subdomains from S_2-SLF1 and S_3-SLF1 will be created and then analyzed for their interactions. These subdomains will be indicated in the same manner as in the F232 and F322 chimerics, with the order of the numbers corresponding to the location of each domain and the numbers themselves indicating the *SLF* of origin. The subdomains will be written as superscripts and designated as letters A-D in the alphabet to avoid confusion. For example, chimeric $F33^{3222}$ contains FD1 and FD2 from S_3-SLF1, FD3-A is from S_3-SLF1 and FD3-B, C, and D are from S_2-SLF1. Four of these chimerics will be generated and analyzed, each containing a different subdomain from S_3-SLF1. FD1 and FD2 have been shown not to affect S-RNase interaction, and will therefore come from S_3-SLF1. The interactions of each of these four SLF chimerics will be used to determine which of these subdomains are required for S-RNase interactions.

S_3-SLF5 was isolated from the *S_3*-haplotype *S*-locus and used to generate a construct that, through *Agrobacterium* transformation, was introduced into *Petunia inflata*. These transgenic plants were then crossed out into *S_5*, *S_{6a}*, *S_7*, *S_{11}*, *S_{12}*, *S_{13}*, *S_{16}*, *S_{22}*, and *S_{24}* genetic backgrounds. After confirming that these progeny expressed the transgene, they were genotyped and analyzed for their SI behavior. The results are shown in Figures 4G, H, I, J and 5G, H, I and Table 5 and 6. These results show that S_3-SLF5 does not interact with S_5-RNase, S_7-RNase, or S_{13}-RNase. It is known that S_3-SLF5 does not interact with S_2-RNase, S_3-RNase, or S_{12}-RNase either, which can be seen in Table 1.

As allelic variants of an SLF share a high degree of amino acid similarity, there are fewer amino acid differences which could account for the differences in their interactions with S-RNases. This makes it easier to determine which domains or amino acids may be involved in SLF/S-RNase interactions. The amino acid sequences of S_2-SLF5 and S_3-SLF5 are 98% identical, yet S_2-SLF5 interacts with S_{12}-RNase and S_3-SLF5 does not. Comparing the amino acid sequences of these two SLFs to find where they differentiate will reveal the amino acids that could potentially be responsible for recognition. To determine the effect that these individual amino acids have, the *SLF* sequences must be altered and analyzed to determine if the changes cause the SLFs to interact with S-RNases they did not previously interact with, and vice versa.

Of the 388 amino acids that comprise S_2-SLF5 and S_3-SLF5, only seven of them are not conserved between these two proteins, as is displayed in Figure 6D. Analysis of the amino acid sequences of other SLF proteins eliminates the potential of six of these sites. S_2-SLF1 and S_3-SLF1 also interact with S_{12}-RNase, while S_2-SLF4 and S_2-SLF6 do not. Comparison of the amino acid sequences of these six proteins reveals that there is a methionine conserved among the three SLFs that do interact with S_{12}-RNase whereas an isoleucine amino acid is conserved among the three that do not. With this information, S_2-SLF5 and S_3-SLF5 can be altered to contain the opposing amino acid. Analysis of the nucleotide sequences of *S_2-SLF5* and *S_3-SLF5* reveals that the alteration of one base pair would result in the desired amino acid switch, shown in Figure 6C.

Site directed mutagenesis will need to be used to alter the bases on each of these two *SLF* genes. They will then be transformed into plants using *Agrobacterium*. These transgenic plants will be crossed out into wild type *$S_{12}S_{12}$* plants to transfer the transgene into the *S_{12}* genetic background. The resulting plants will be genotyped and analyzed for GFP, after which they will

be analyzed for their SI behavior. The interactions, or lack of interactions, of those plants would reveal whether the methionine is necessary and sufficient for interactions with S_{12}-RNase. If S_3-SLF5 with the isoleucine changed to methionine interacts with S_{12}-RNase, this would suggest that the methionine residue is sufficient for recognition between SLFs and S_{12}-RNase. If the mutated S_2-SLF5 no longer interacts with S_{12}-RNase, it would suggest that the methionine is essential for the interaction. The site of this particular amino acid would need to be analyzed across the known SLFs to determine if the same effects are seen with and without methionine as well.

If the expected results are not obtained, this could mean that the SLF and S-RNase interactions are controlled by the recognition of multiple amino acids. To test this theory, the *S_3-SLF5* sequence would need to be altered at all seven of the amino acid sites that are not conserved across *S_3-SLF5* and *S_2-SLF5*. Six additional mutated SLFs would need to be created, each containing the methionine, with a different amino acid alteration at one of the seven additional non-conserved sites. These mutated SLFs would be analyzed for interactions against an *S_{12}* genetic background to determine the effect each combination of mutations has on SLF and S_{12}-RNase recognition.

Interestingly, analysis of the S_2-SLF5 and S_3-SLF5 amino acid sequences has revealed that the conserved methionine in S_2-SLF5 is amino acid residue 295. If we follow the same domain demarcation of SLF1, this amino acid would fall within FD3, which contains amino acids 261-395 of SLF1. SLF1 and SLF5 are different types of SLF that have been analyzed using two different methods, yet both analyses have indicated that the amino acids required for S-RNase recognition are located within the third domain of SLF proteins.

It would seem inconceivable that a single point mutation in an *SLF* allele might hold the key to S-RNase recognition and therefore SI. This system involves a highly sophisticated mechanism, and it would be unbelievable to find that such a small change in the genetic code of these plants would have such a large effect. The Solanaceae family of plants has only one defense mechanism against inbreeding, which is dependent on the interactions between SLFs and S-RNases. Between the 17 known SLFs of S_2-haplotype and more than 30 S-RNases produced by their non-self haplotypes, there are hundreds of possible SLF and S-RNase encounters. The prevention of inbreeding hangs in the balance when these SLFs and S-RNases do or do not interact. I do not think that it is possible that this complicated process could have evolved to be dependent on the interaction of one amino acid in the SLF protein with the S-RNase protein. DNA replication systems are not perfect and can easily make single base pair mistakes. If methionine were the only amino acid involved in this recognition, one simple error in DNA replication would cause a formerly self-incompatible plant to become self-compatible. This would lead to inbreeding, and the entire purpose of this system is to prevent inbreeding. I think that the successful degradation of self-pollen tubes is too essential to be dependent on a single point mutation. I think that the interaction between SLF proteins and non-self S-RNases relies on the recognition of multiple amino acids between the two proteins. Multiple amino acid interactions would ensure that one mutation to the gene will not be sufficient to convert an SI plant to an SC plant. This would make the SI system more dependably inherited through generations of *Petunia inflata*.

References

Bai, C., Sen, P., Hofmann, K., Ma, L., Goebel, M., Harper, W., and Elledge, S. (1996). *SKP1* connects cell cycle regulators to the ubiquitin proteolysis machinery through a novel motif, the F-box. Cell **86**: 263–274.

de Nettancourt, D. (2001). Incompatibility and incongruity in wild and cultivated plants. Berlin: Springer.

Goldraij, A., Kondo, K., Lee, C.B., Hancock, C.N., Sivaguru, M., Vazquez-Santana, S., Kim, S., Phillips, T.E., Cruz-Garcia, F., and McClure B. (2006). Compartmentalization of S RNase and HT-B degradation in self-incompatible *Nicotiana*. Nature **439**: 805–810. 42

Hua, Z., and Kao, T.-H. (2006). Identification and characterization of components of a putative petunia *S*-locus F-box-containing E3 ligase complex involved in S-RNase-based self-incompatibility. Plant Cell **18**: 2531–2553.

Hua, Z., Meng, X., and Kao, T.-H. (2007). Comparison of *Petunia inflata S*-Locus F-box protein (Pi SLF) with Pi SLF like proteins reveals its unique function in S-RNase based self-incompatibility. Plant Cell **19**: 3593–3609.

Huang, S., Lee, H.S., Karunanandaa, B., and Kao, T.H. (1994). Ribonuclease activity of *Petunia inflata* S proteins is essential for rejection of self-pollen. Plant Cell **6**: 1021-1028.

Iwano, M., and Takayama, S. (2012). Self/non-self discrimination in angiosperm self-incompatibility. Curr. Opin. Plant Biol. **15**: 78–83.

Kubo, K., Entani, T., Takara, A., Wang, N., Fields, A.M., Hua, Z.H., Toyoda, M.,

Kawashima, S., Ando, T., Isogai, A., Kao, T.H., and Takayama, S. (2010). Collaborative non-self recognition system in S-RNase-based self-incompatibility. Science **330**: 796-799.

Lee, H.S., Huang, S., and Kao, T.H. (1994). S proteins control rejection of Incompatible pollen in *Petunia inflata*. Nature **367**: 560–563.

Liu, B., Morse, D., & Cappadocia, M. (2009). Compatible pollinations in *Solanum chacoense* decrease both S-RNase and S-RNase mRNA. PloS one, 4(6), e5774. doi:10.1371/journal.pone.0005774

Luu, D., Qin, X., Morse, D., and Cappadocia, M. (2000). S-RNase uptake by compatible pollen tubes in gametophytic. Nature **407**: 649–651.

McClure, B. (2004). S-RNase and SLF determine *S*-haplotype-specific pollen recognition and rejection. Plant Cell **16**: 2840–2847.

Meng, X., Hua, Z., Sun, P., and Kao, T.H. (2011). The amino terminal F-box domain of *Petunia inflata S*-locus F-box protein is involved in the S-RNase-based self-incompatibility mechanism. AoB Plants **6** doi: 10.1093/aobpla/plr016.

Meng, X., Hua, Z., Wang, N., Fields, A.M., Dowd, P.E., and Kao, T. (2009). Ectopic expression of S-RNase of *Petunia inflata* in pollen results in its sequestration and non cytotoxic function. Sex. Plant Reprod. **22**, 263–275.

Murfett, J., Atherton, T.L., Mou, B., Gasser, C.S., and McClure, B.A. (1994). S-RNase expressed in transgenic *Nicotiana* causes *S*-allele-specific pollen rejection. Nature **367**: 563–566.

Sijacic, P., Wang, X., Skirpan, A., Wang, Y., Dowd, P., Mccubbin, A.G., Huang, S., and Kao, T. (2004). Identification of the pollen determinant of S-RNase-mediated self

incompatibility. Nature **429**: 302–305.

Sims, T.L., and Robbins, T.P. (2009). Gametophytic Self-Incompatibility in *Petunia*. In *Petunia:* Evolutionary, Developmental and Physiological Genetics. T. Gerats and J. Strommer, eds. (New York: Springer), pp. 85–106.

Stone, S.L., and Callis, J. (2007). Ubiquitin ligases mediate growth and development by promoting protein death. Curr. Opin. Plant Biol. **10**: 624–632.

Stout, A.B., and Chandler, C. (1941). Change from self-Incompatibility to self-Compatibility accompanying change from diploidy to tetraploidy. Science **94**: 118.44

Sun, P., and Kao, T.H. (2013). Self-incompatibility in *Petunia inflata*: the relationship between a Self-incompatibility Locus F-Box Protein and its non-self S-RNases. Plant Cell **25**: 470–485.

Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M., and Kumar, S. (2011). MEGA5: molecular evolutionary genetics analysis using maximum likelihood , evolutionary distance, and maximum parsimony methods. Mol. Biol. Evol. **28**: 2731–2739.

Thompson, J.D., Higgins, D.G., and Gibson, T.J. (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res. **22**: 4673–4680.

Vierstra, R.D. (2009). The ubiquitin-26S proteasome system at the nexus of plant biology.Nature Rev. Mol. Cell Biol. **10**: 385–397.

Wang, N., and Kao, T.H. (2012) Self-incompatibility in *Petunia*: a self/nonself-recognition mechanism employing *S*-locus F-box proteins and S-RNase to prevent inbreeding.

WIRES Dev Biol. **1**: 267–275.

Williams, J.S., Der, J.P., dePamphilis, C.W., and Kao, T.H. (2014) Transcriptome Analysis Reveals the Same 17 S-Locus F-Box Genes in Two Haplotypes of the Self Incompatibility Locus of *Petunia inflata*. Plant Cell **26**: 2873-2888.

Academic Vita

Daniele San Román
Dos5291@psu.edu

Education

The Pennsylvania State University *Class of May 2015*
Schreyer Honors College *University Park, PA*
Bachelor of Science: Biochemistry and Molecular Biology, Molecular Cell Biology
Institute for the International Education of Students *Spring 2013*
IES Student Council *Buenos Aires, Argentina*

Professional Experience

Undergraduate Researcher, Enzymology and Molecular Genetics Laboratory
The Pennsylvania State University, University Park, PA *Spring 2012, Fall 2012*
Professor Allen T Phillips

- Prepared buffers, performed enzyme assays, cultivated bacterial cultures, performed site-directed mutagenesis

Undergraduate Researcher, Biochemical and Molecular Biology Laboratory
The Pennsylvania State University, University Park, PA *Fall 2013-Present*
Professor Teh-Hui Kao

- Conducted honors thesis research project: Dissecting the Interactions Between S-RNases and S-Locus F-Box Proteins of *Petunia inflata* during Self/Non-Self Recognition Self Incompatibility
- Perform seed germination, genomic DNA isolation, and PCR genotyping; create and analyze DNA and amino acid sequence alignments; analyze self-incompatibility of transgenic plants

Summer Experience in the Eberly College of Science (SEECoS) Mentor
Kao Lab, The Pennsylvania State University, University Park, PA *Summer 2014*

- Educated two high school students on molecular biology lab techniques and conducting experiments

Skills

Proficient in Microsoft Office Student (Word, Excel, PowerPoint)
Conversational Spanish, Argentine Dialect