PCP
PLANT & CELL PHYSIOLOGY

Use of Domain-Swapping to Identify Candidate Amino Acids Involved in Differential Interactions between Two Allelic Variants of Type-1 S-Locus F-Box Protein and S_3-RNase in *Petunia inflata*

Lihua Wu[1], Justin S. Williams[2,3], Ning Wang[1,4], Wasi A. Khatri[2], Daniele San Román[2,5] and Teh-hui Kao[1,2,*]

[1]Intercollege Graduate Degree Program in Plant Biology, The Pennsylvania State University, University Park, PA 16802, USA
[2]Department of Biochemistry and Molecular Biology, The Pennsylvania State University, University Park, PA 16802, USA
[3]Present address: St. Jude Children's Research Hospital, Department of Computational Biology, MS 1135, Room I6104, 262 Danny Thomas Place, Memphis, TN 38105, USA
[4]Present address: BioVision Inc., 155 S. Milpitas Blvd., Milpitas, CA 95035, USA
[5]Present address: Washington and Lee University School of Law, 1 Denny Circle, Lexington, VA 24450, USA
*Corresponding author: E-mail, txk3@psu.edu; Fax, +1-814-863-7024
(Received October 2, 2017; Accepted November 11, 2017)

Regular Paper

Petunia inflata **possesses a self-incompatibility (SI) mechanism, which involves *S-RNase* and multiple *S-locus F-box* (*SLF*) genes at the polymorphic *S*-locus. For a given *S*-haplotype, each SLF is thought to interact with some of its non-self S-RNases, but not with its self S-RNase. In this work, we studied an allelic pair of SLF1, S_2-SLF1 and S_3-SLF1, which differ in 44 amino acids and show differential interactions with S_3-RNase. We first used an in vivo transgenic assay to determine whether four chimeric proteins of S_2-SLF1 and S_3-SLF1, each with one of the three functional domains swapped, interact with S_3-RNase. The results narrowed the candidate amino acids for specific interaction of S_2-SLF1 with S_3-RNase to the 16 in domain FD3. We then examined seven additional chimeric proteins by dividing FD3 into two subdomains and four mini-domains (A, B, C and D). The results further narrowed the candidate amino acids to four in mini-domain A and four in mini-domain D. Molecular modeling of interactions between S_3-RNase and S_2-SLF1 revealed that three of these eight are at the interaction surface, and all three are conserved in S_1-SLF1 and S_{6a}-SLF1, both of which interact with S_3-RNase based on the in vivo transgenic assay. Three of the chimeric proteins were used for the in vivo transgenic assay to determine whether FD3 alone contains the amino acids required for S_2-SLF1 to interact with S_7-RNase and S_{13}-RNase. The results revealed the diversity and complexity of interactions between SLF proteins and S-RNases.**

Keywords: Chimeric S-locus F-box proteins • *Petunia inflata* • Pollen specificity determinant • Self-incompatibility • Self/non-self recognition • S-RNases.

Abbreviations: Co-IP, co-immunoprecipitation; FD, functional domain; GFP, green fluorescent protein; SCF complex, Skp1–Cullin1–F-box protein complex; SI, self-incompatibility; SLF, S-locus F-box.

Introduction

Flowering plants producing bisexual flowers have adopted various strategies to counteract the tendency to self-fertilize, as consequent inbreeding invariably leads to impaired fitness in the progeny and reduced genetic diversity in the species. Self-incompatibility (SI) is one such strategy that allows the pistil to reject self-pollen from the same plant or genetically related plants, and to accept non-self pollen from genetically unrelated plants (De Nettancourt 2001). In Solanaceae, self/non-self recognition between pollen and pistil is regulated by the highly polymorphic *S*-locus. When the *S*-haplotype of pollen matches either haplotype of the pistil, the pollen is recognized as self-pollen and their tube growth in the pistil is inhibited. Pollen carrying an *S*-haplotype different from both *S*-haplotypes of the pistil is recognized as non-self pollen, and their tubes are allowed to grow down the pistil to effect fertilization. In *Petunia*, the *S*-locus houses the *S-RNase* gene, which encodes the pistil determinant, and multiple *S-locus F-box* (*SLF*) genes, which collectively specify the pollen determinant (Lee et al. 1994, Sijacic et al. 2004, Kubo et al. 2010, Williams et al. 2014a, Williams et al. 2014b, Kubo et al. 2015). For example, the same 17 polymorphic *SLF* genes, *SLF1* (or Type-1 *SLF*) to *SLF17* (or Type-17 *SLF*), have been identified in both the S_2-haplotype and S_3-haplotype of *P. inflata* (Sijacic et al. 2004, Williams et al. 2014a).

S-RNases are produced in the transmitting tissue and secreted into the transmitting tract of the pistil, where they are taken up by pollen tubes in a non-*S*-haplotype-specific manner (Luu et al. 2000, Goldraij et al. 2006). As the RNase activity of S-RNases (allelic variants of S-RNase) is essential for their function in SI (Huang et al. 1994), S-RNases are thought to exert their function by degrading pollen tube RNAs. A model, named collaborative non-self recognition, was proposed to explain why an allelic variant of S-RNase is cytotoxic to its self-pollen tubes, but not to its non-self pollen tubes (Kubo et al. 2010). According to this model, each SLF protein is a component of

Plant Cell Physiol. 0(0): 1–14 doi:10.1093/pcp/pcx176, Advance Access publication on 15 November 2017, available online at www.pcp.oxfordjournals.org

an SCF complex, a multi-subunit E3 ubiquitin ligase, which also contains Skp1, Cullin1 and Rbx1 (Bai et al. 1996, Stone and Callis 2007). The SCF complex, together with E1 (ubiquitin-activating enzyme) and E2 (ubiquitin-conjugating enzyme), catalyzes the transfer of polyubiquitin chains to the S-RNase(s) with which the particular SLF protein interacts. This model predicts that SLF proteins produced by pollen of a given S-haplotype collectively interact with all their non-self S-RNases, but do not interact with their self S-RNase. As a result, non-self S-RNases are ubiquitinated and degraded by the 26S proteasome, and only self S-RNase is able to exert its cytotoxicity.

Results supporting the collaborative non-self recognition model have been obtained. First, each of the 17 SLF proteins of the S_2-haplotype of *Petunia inflata*, as well as each of their allelic variants in the S_3-haplotype, was found to be the F-box protein component of an SCF (Skp1–Cullin1-F–box) complex that contains a pollen-specific Cullin1 (named PiCUL1-P), a pollen-specific Skp1 protein (named PiSSK1) and a conventional Rbx1 (named PiRBX1) (Li et al. 2014, Li et al. 2016). Secondly, in *Petunia*, interaction relationships between a number of SLF proteins and S-RNases have been established genetically by an in vivo transgenic assay, and all the results obtained so far are consistent with the prediction by the model. This assay examines whether expression of an SLF protein of a particular *S*-haplotype in pollen of the same or different *S*-haplotypes causes breakdown of SI in transgenic pollen. For example, expression of S_2-SLF1 of *P. inflata* in S_3 pollen caused breakdown of SI in S_3 transgenic pollen (Sijacic et al. 2004), suggesting that S_2-SLF1 interacts with and detoxifies S_3-RNase in the S_3 transgenic pollen tube to allow the S_3 transgenic pollen to be compatible with pistils carrying the S_3-haplotype. In a few cases where the interaction between a pair of SLF and S-RNase, established from this transgenic assay, has subsequently been examined by co-immunoprecipitation (Co-IP), all the genetic interactions have been verified (Kubo et al. 2010, Sun and Kao 2013). Based on this assay, none of the SLF proteins of a particular *S*-haplotype caused breakdown of SI in pollen of the same *S*-haplotype (i.e. self *S*-haplotype), suggesting that they do not interact with their respective self S-RNases. For example, expression of S_2-SLF1 did not cause breakdown of SI in S_2 pollen (Sijacic et al. 2004). Moreover, each SLF of a particular *S*-haplotype only caused breakdown of SI in pollen of some of the non-self *S*-haplotypes examined, suggesting that an SLF interacts with a subset of its non-self S-RNases (Sijacic et al. 2004, Kubo et al. 2010, Williams et al. 2014b, Kubo et al. 2015). For example, expression of S_2-SLF1 did not cause breakdown of SI in S_5 or S_{11} pollen, suggesting that S_2-SLF1 does not interact with S_5-RNase or S_{11}-RNase (Williams et al. 2014b).

To address the biochemical basis for differential interactions between SLF proteins and S-RNases, we have chosen to study an allelic pair of SLF1, S_2-SLF1 and S_3-SLF1, as expression of S_2-SLF1, but not S_3-SLF1, caused breakdown of SI in S_3, S_7 and S_{13} pollen, suggesting that S_2-SLF1, but not S_3-SLF1, interacts with S_3-, S_7- and S_{13}-RNases (Hua et al. 2007, Kubo et al. 2010, Williams et al. 2014b). In the case of S_3-RNase, the interaction with S_2-SLF1 has been verified by Co-IP (Sun and Kao 2013). Interestingly, expression of S_2-SLF1 or S_3-SLF1 in S_{12} pollen caused breakdown of SI, suggesting that both interact with S_{12}-RNase (Sun and Kao 2013). The deduced amino acid sequences of S_2-SLF1 (389 amino acids) and S_3-SLF1 (388 amino acids) are 88.7% identical, differing in 44 of their aligned amino acid positions.

In this work, we used the domain-swapping approach to narrow down the candidate amino acids of S_2-SLF1 that are required for the interaction with S_3-RNase. We previously divided S_2-SLF1 into three functional domains (FDs), FD1, FD2 and FD3, based on the results of in vitro binding between S-RNases and various truncated forms of S_2-SLF1 (Hua et al. 2007). Here, we made chimeric genes of S_2-*SLF1* and S_3-*SLF1* by first swapping the coding sequences for one of these three domains to determine which domain(s) contain(s) amino acids required for the interaction with S_3-RNase. If a chimeric protein behaves as S_2-SLF1, its expression in S_3 pollen of S_2S_3 transgenic plants should allow S_3 transgenic pollen to detoxify S_3-RNase and render the transgenic plants self-compatible. This result would suggest that the amino acids that are different between S_2-SLF1 and S_3-SLF1 in the swapped out domain of S_2-SLF1 are not required for the interaction with S_3-RNase. If expression of a chimeric protein in S_3 pollen does not alter the SI behavior of the S_2S_3 transgenic plants, this result would suggest that the domain of S_2-SLF1 swapped out contains amino acids required for the interaction with S_3-RNase. We found that the third domain (FD3), but not the first two domains (FD1 and FD2), of S_2-SLF1, is required for the interaction with S_3-RNase, suggesting that one or more of the 16 amino acids different between S_2-SLF1 and S_3-SLF1 in FD3 is (are) involved. We then made additional chimeric genes for fine dissection of FD3, first dividing this domain into two subdomains (each containing eight of the 16 amino acids different between S_2-SLF1 and S_3-SLF1), and then further dividing this domain into two mini-domains (each containing four of the eight amino acids different between S_2-SLF1 and S_3-SLF1). Analysis of the SI behavior of the S_2S_3 transgenic plants expressing each of these chimeric proteins revealed that the first mini-domain and the fourth mini-domain of FD3 contain the amino acids required for interactions with S_3-RNase. We also used some of the chimeric proteins to examine whether FD3 alone contains the amino acids required for S_2-SLF1 to cause breakdown of SI in S_7 and S_{13} pollen. The results reveal the diversity and complexity of the interactions between SLF proteins and S-RNases. Finally, we used modeling and molecular docking to predict the interaction surface between S_2-SLF1 and S_3-RNase. The results showed that the first mini-domain and the last mini-domain are at the interface of these two proteins, and further identified the amino acids of S_2-SLF1 most likely to be involved in the specific interaction with S_3-RNase.

Results

A single domain, FD3, spanning the C-terminal one-third of S_2-SLF1 contains amino acids required for the interaction with S_3-RNase in vivo

To determine which of the 44 amino acids different between S_2-SLF1 and S_3-SLF1 are required for the specific interaction

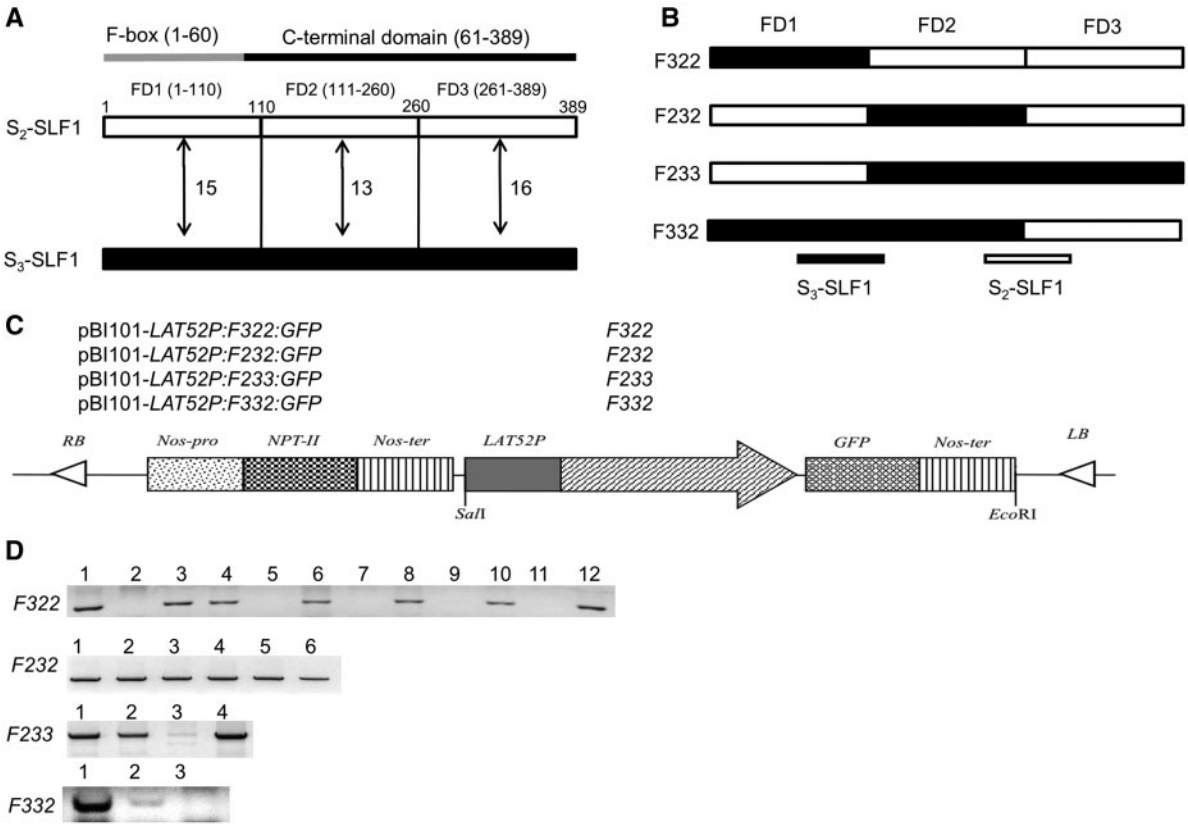

Fig. 1 Four chimeric proteins of S_2-SLF1 and S_3-SLF1, their transgene constructs and the presence of the transgene in each transgenic line. (A) Schematics of S_2-SLF1 (389 amino acids) and S_3-SLF1 (388 amino acids), with the three functional domains, FD1, FD2 and FD3, demarcated. The first and last amino acid residues of the region covered by each domain of S_2-SLF1 are indicated. The number of amino acids different between S_2-SLF1 and S_3-SLF1 in each domain is indicated by a double-headed vertical arrow. (B) Schematics of four chimeric proteins, F322, F232, F233 and F332, of S_2-SLF1 and S_3-SLF1, each with one of the three domains swapped. Each chimeric protein is denoted by 'F' (for SLF) followed by three digits (e.g. 322), each representing one domain (FD1, FD2 or FD3). Each digit indicates whether S_2-SLF1 (2) or S_3-SLF1 (3) contributes the domain. (C) Schematics of transgene constructs of chimeric genes *F322*, *F232*, *F233* and *F332*, in the Ti plasmid pBI101. Abbreviations: *pLAT52*, the promoter of *LAT52* of tomato; *GFP*, the coding sequence for the green fluorescent protein; *NOS-ter*, transcription termination signal of the nopaline synthase gene. (D) PCR analysis of genomic DNA isolated from T_0 plants of each transgenic line to assess the presence of the respective transgene. The primers used are specific to the *GFP* sequence included in each transgene; the DNA band detected is of the size expected for the amplified DNA fragment. Plant number is indicated above each gel lane.

between S_2-SLF1 and S_3-RNase, we first constructed four chimeric genes by swapping the coding sequences for one of the three FDs, FD1, FD2 and FD3, between S_2-SLF1 and S_3-SLF1. FD1 (amino acids 1–110) contains the N-terminal F-box domain, and FD2 (amino acids 111–260) and FD3 (amino acids 261–389) constitute the remainder of SLF1 (**Fig. 1A**). These four chimeric proteins were designated F322, F232, F233 and F332 (**Fig. 1B**), with 'F' denoting SLF, and each digit indicating whether the particular domain is contributed by S_2-SLF1 (2) or S_3-SLF1 (3). For example, F322 denotes a chimeric protein that contains FD1 of S_3-SLF1 and FD2 and FD3 of S_2-SLF1. The four transgene constructs (**Fig. 1C**) were separately introduced into *P. inflata* plants of the S_2S_3 genotype via *Agrobacterium*-mediated transformation. PCRs, using a primer pair designed based on the *GFP* (green fluorescent protein) sequence (Supplementary Table S1), were performed on genomic DNA isolated from at least three of the transgenic plants obtained for each construct. A DNA fragment of the expected size (0.5 kb) was observed for multiple plants in each transgenic line (the results of representative plants shown in **Fig. 1D**). For each transgenic line, at least one transgenic plant was observed to express GFP in in vitro germinated pollen tubes. The dark field and bright field images of a representative plant from all except the *F322* transgenic line are shown in Supplementary Fig. S1.

We first used pollen from each T_0 transgenic plant (S_2S_3) to pollinate a wild-type S_2S_3 plant (self-cross) to see whether expression of a particular chimeric gene caused breakdown of SI in S_3 pollen, but not in S_2 pollen, produced by the T_0 plant. A previously generated *S_2-SLF1:GFP*/S_2S_3 transgenic plant (Hua et al. 2007) was used as a control for self-compatible pollination. For each transgenic line, the SI behavior of all the T_0 plants that expressed the transgene was examined, and at least three pollinations were performed for each plant. As the results from different T_0 plants were identical, only the results of one representative plant are shown in **Table 1**. Pollination of the wild-type S_2S_3 plant by pollen from *F233:GFP*/S_2S_3 did not set any

Table 1 Effect of expressing each of 11 chimeric genes of S_2-*SLF1* and S_3-*SLF1* on the self-incompatibility behavior of S_2S_3 transgenic plants

Transgene	Set fruit when pollinate S_2S_3 pistil (Yes/No)	SI behavior of T_0 plant	T_1 progeny genotype (S_2S_3 : S_3S_3)	Expected ratio	χ^2 (1 : 1)	*P*-value (1 : 1)	χ^2 (1 : 2 : 1) (S_2S_2 : S_2S_3 : S_3S_3)	*P*-value (1 : 2 : 1)
F322:GFP	Yes	SC	24 : 23	1 : 1	0.02	0.88	22.532	<0.001
F232:GFP	Yes	SC	17 : 14	1 : 1	0.29	0.60	12.935	0.002
F233:GFP	No	SI						
F332:GFP	Yes	SC	8 : 10	1 : 1	0.22	0.64	11.333	0.003
F23(23):GFP	No	SI						
F23(32):GFP	No	SI						
F33(3222):GFP	No	SI						
F33(2322):GFP	Yes	SC	10 : 22	1 : 1	4.50	0.03	34.75	<0.001
F33(2232):GFP	Yes	SC	15 : 17	1 : 1	0.13	0.72	18.188	<0.001
F33(2223):GFP	No	SI						
F33(2332):GFP	Yes	SC	14 : 17	1 : 1	0.29	0.60	18.935	<0.001

Pollen from each transgenic plant was used to pollinate wild-type S_2S_3 plants, and for each self-compatible (SC) cross, the presence/absence of the transgene and the S-genotypes of the progeny were determined. All the plants in each T_1 progeny inherited the transgene.
A χ^2 analysis was used to test the null hypothesis of a 1 : 1 ratio of S_2S_3 : S_3S_3, in comparison with a 1 : 2 : 1 ratio (S_2S_2 : S_2S_3 : S_3S_3).

fruit, suggesting that expression of F233 did not cause breakdown of SI in either S_2 or S_3 pollen. In contrast, pollination of the wild-type S_2S_3 plant by pollen from *F232:GFP*/S_2S_3, *F322:GFP*/S_2S_3 and *F332:GFP*/S_2S_3 led to all setting large fruits with seed numbers comparable with those obtained from pollination of the wild-type S_2S_3 plant by pollen from S_2-*SLF1:GFP*/S_2S_3, suggesting that expression of these three chimeric genes also causes breakdown of SI in S_3 pollen. To confirm this interpretation, T_1 plants were raised from each self-crossed progeny, and their S-genotypes and presence/absence of the transgene were determined by PCR, using S_2-*RNase*-, S_3-*RNase*- and *GFP*-specific primers (Supplementary Table S1). For *F232:GFP*/S_2S_3, all 47 T_1 plants inherited the transgene; the S_3-*RNase*-specific fragment was detected in all of them, and the S_2-*RNase*-specific fragment was detected in 24 of them, indicating that 24 T_1 plants were S_2S_3 and the other 23 were S_3S_3 (results for 11 representative plants shown in **Fig. 2A**). The *P*-value of a χ^2 analysis of a 1:1 ratio null hypothesis of S_2S_3:S_3S_3 was 0.88, whereas the *P*-value for a 1:2:1 null hypothesis of S_2S_2:S_2S_3:S_3S_3 was <0.001, thus supporting the absence of the S_2S_2 genotype in the progeny (**Table 1**). Based on a similar analysis, all 31 T_1 plants of *F322:GFP*/S_2S_3 inherited the transgene: 17 were S_2S_3 and 14 were S_3S_3 (results for 12 representative plants shown in **Fig. 2B**); all 18 T_1 plants of F332:GFP/S_2S_3 inherited the transgene: eight were S_2S_3 and 10 were S_3S_3 (results for 10 representative plants shown in **Fig. 2C**). The χ^2 analysis also supported the absence of S_2S_2 genotype in both progeny (**Table 1**).

Absence of S_2S_2 plants in the progeny obtained from self-crosses of *F232:GFP*/S_2S_3, *F322:GFP*/S_2S_3 and *F332:GFP* suggests that both wild-type S_2 pollen and S_2 pollen carrying *F232:GFP*, *F322:GFP* or *F332:GFP* are rejected by the wild-type S_2S_3 pistil. Moreover, the finding that all progeny plants carried the respective transgene suggests that S_3 pollen carrying each transgene, but not wild-type S_3 pollen, is compatible with the wild-type S_2S_3 pistil. Thus, expression of F232:GFP, F322:GFP and F332:GFP caused breakdown of SI in S_3 pollen, suggesting that these three chimeric proteins interact with S_3-RNase. That is, replacing FD1 of S_2-SLF1 with FD1 of S_3-SLF1 in F322, replacing FD2 of S_2-SLF1 with FD2 of S_3-SLF1 in F232 and replacing both FD1 and FD2 of S_2-SLF1 with FD1 and FD2 of S_3-SLF1 in F332 did not affect the ability of S_2-SLF1 to interact with S_3-RNase. Thus, none of the 15 amino acids different between S_2-SLF1 and S_3-SLF1 in FD1, and none of the 13 amino acids different between S_2-SLF1 and S_3-SLF1 in FD2, are required for the interaction of S_2-SLF1 with S_3-RNase. We conclude that FD3 alone is responsible for allele specificity of S_2-SLF1 with respect to its interaction with S_3-RNase.

Fine dissection of FD3 reveals two regions that contain amino acids required for the interaction of S_2-SLF1 with S_3-RNase in vivo

FD3 spans approximately the C-terminal one-third of S_2-SLF1 (**Fig. 1A**), and contains 16 of the 44 amino acids that are different between S_2-SLF1 and S_3-SLF1. To narrow further the amino acids that are required for the interaction with S_3-RNase, we first divided FD3 into two subdomains, amino acids 261–325 and 326–389, with each subdomain containing eight of the 16 different amino acids (**Fig. 3A**), to determine whether all the amino acids required are located in one of the subdomains. We made two chimeric genes, designated *F23(23):GFP* and *F23(32):GFP*, with each digit in parentheses denoting the allele of *SLF1* (2 for S_2-*SLF1* and 3 for S_3-*SLF1*) that contributes the sequence for a particular subdomain (**Fig. 3B**). Both transgene constructs were separately introduced into S_2S_3 plants. Analysis of each line of transgenic plants by PCR to identify those that carried the transgene was carried out as described for the four chimeric genes shown in **Fig. 1B**; the results of representative plants are shown in **Fig. 3C**. The SI behavior of all the T_0 plants that expressed the transgene was analyzed similarly to as described above. None of the pollinations resulted in fruit sets (**Table 1**), suggesting that neither chimeric protein interacts with and

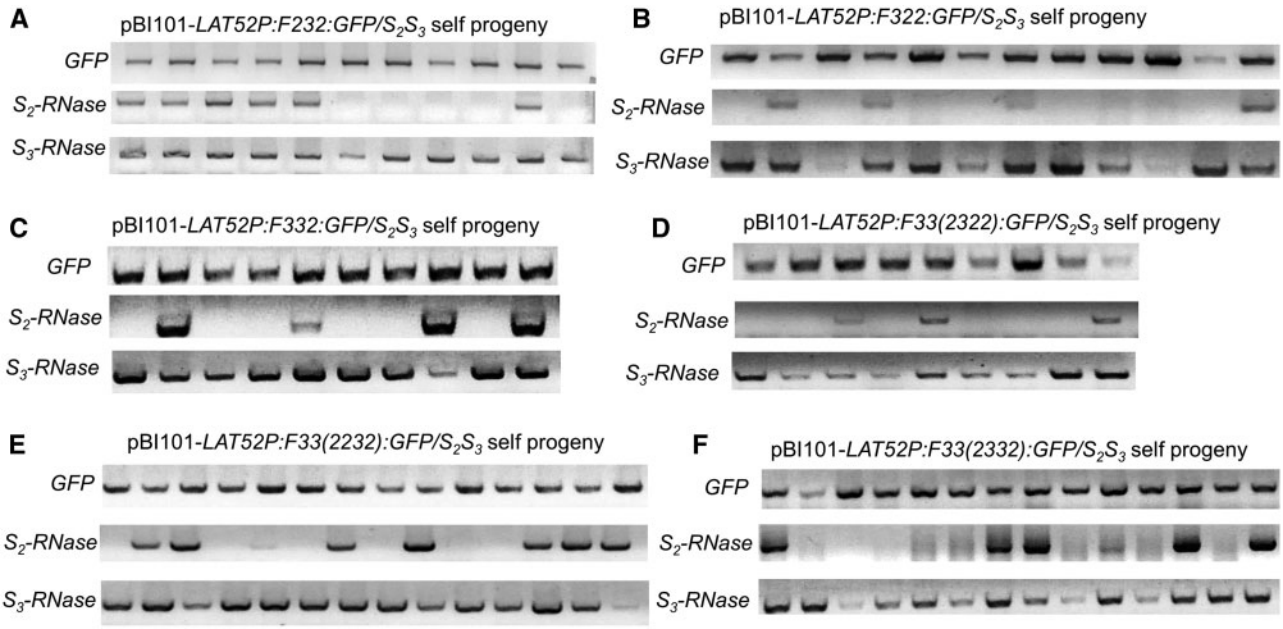

Fig. 2 PCR analysis of progeny plants obtained from self-crosses of T_0 plants from six transgenic lines for the presence of the transgene and for their *S*-genotypes. Each progeny was obtained from pollination of a wild-type S_2S_3 plant by pollen of a T_0 plant from the transgenic line denoted above the gel image in each panel. In (A), plant #1 of the T_0 plants was used; in (B), plant #1 was used; in (C), plant #1 was used; in (D), plant #7 was used; in (E), plant #4 was used; in (F), plant #1 was used. These T_0 plants are shown in **Fig. 1D** or **Fig. 3C**. Genomic DNA isolated from T_1 progeny plants was amplified by *GFP* primers (top panel), *S_2-RNase* (middle panel) and *S_3-RNase* (bottom panel).

detoxifies S_3-RNase to allow the S_3 transgenic pollen to be compatible with the S_2S_3 pistil. Thus, one or more amino acid residues in both subdomains are required for S_2-SLF1 to interact with S_3-RNase.

We next divided FD3 into four mini-domains, A (amino acids 261–275), B (amino acids 276–310), C (amino acids 311–360) and D (amino acids 361–389), with each mini-domain containing four of the 16 different amino acids between S_2-SLF1 and S_3-SLF1. We made five chimeric constructs, all having FD1 and FD2 of S_3-SLF1, and FD3 of S_2-SLF1 except for the replacement of one or two of these four mini-domains with the corresponding mini-domain(s) of S_3-SLF1. These chimeric genes were designated *F33(3222)*, *F33(2322)*, *F33(2232)*, *F33(2223)* and *F33(2332)*, and they are schematically shown in **Fig. 3A**. Each digit in parentheses denotes the allele of *SLF1* (2 for *S_2-SLF1* and 3 for *S_3-SLF1*) that contributes the sequence for a particular mini-domain. The transgene constructs for these five chimeric genes (**Fig. 3B**) were made similarly to as described above, and separately introduced into *P. inflata* plants of the S_2S_3 genotype. Analysis of each line of transgenic plants by PCR to identify those that carried the transgene, and subsequent analysis of GFP fluorescence to determine whether the transgenes were expressed, were carried out as described above. The results showed that most of the plants in each transgenic line carried the transgene (**Fig. 3C**), and the transgenes were expressed in pollen of most of these plants; the dark field and bright field images of a representative plant from each transgenic line are shown in **Supplementary Fig. S1**. For each transgenic line, the SI behavior of all the T_0 plants that expressed the transgene was examined, and at least three pollinations were performed for each plant. As the results from different T_0 plants were identical, only the results of one representative plant are shown in **Table 1**. *F33(2322):GFP/S_2S_3*, *F33(2232):GFP/S_2S_3* and *F33(2332):GFP/S_2S_3*, but not *F33(3222):GFP/S_2S_3* or *F33(2223):GFP/S_2S_3*, set large fruits with seed number comparable with those obtained from pollination of the wild-type S_2S_3 plant with pollen from *S_2-SLF1:GFP/S_2S_3*.

To confirm that the breakdown of SI in three of the five transgenic lines was due to expression of F33(2322):GFP, F33(2232):GFP and F33(2332):GFP in S_3 pollen, T_1 plants were raised from each self-crossed progeny to determine their *S*-genotypes and inheritance of the transgene by PCR, using the *S_2-RNase-*, *S_3-RNase-* and *GFP*-specific primers (**Supplementary Table S1**). The results are shown in **Table 1**. For *F33(2322):GFP/S_2S_3* plant #7 (**Fig. 3C**), all 32 T_1 plants inherited the transgene; the *S_3-RNase*-specific fragment was detected in all of them and the *S_2-RNase*-specific fragment was detected in 10 of them, indicating that 10 T_1 plants were S_2S_3 and the other 22 plants were S_3S_3 (representative genotyping results shown in **Fig. 2D** and **Table 1**). Although the *P*-value for the null hypothesis of a 1 : 1 ratio of S_2S_3:S_3S_3 (0.03) was less than the minimum measure of significance (0.05), it is important to note the absence of S_2S_2 in the progeny and the presence of the transgene in all progeny (see below). Using similar PCR analyses to those described above, for *F33(2232):GFP/S_2S_3* plant #4 (**Fig. 3C**), all 32 T_1 plants inherited the transgene: 15 were S_2S_3 and 17 were S_3S_3 (representative genotyping results shown in **Fig. 2E** and **Table 1**); for *F33(2332):GFP* plant #1 (**Fig. 3C**), all 31 T_1 plants inherited the transgene: 14 were S_2S_3 and 17 were S_3S_3 (representative genotyping results shown in **Fig. 2F**; **Table 1**). The null hypothesis for a 1 : 1 ratio of S_2S_3:S_3S_3 in the progeny analysis of $S_2S_3\times$*F33(2232):GFP* and $S_2S_3\times$*F33(2332):GFP* was supported, while the null hypothesis for a 1 : 2 : 1 ratio of S_2S_2: S_2S_3:S_3S_3 was rejected (**Table 1**).

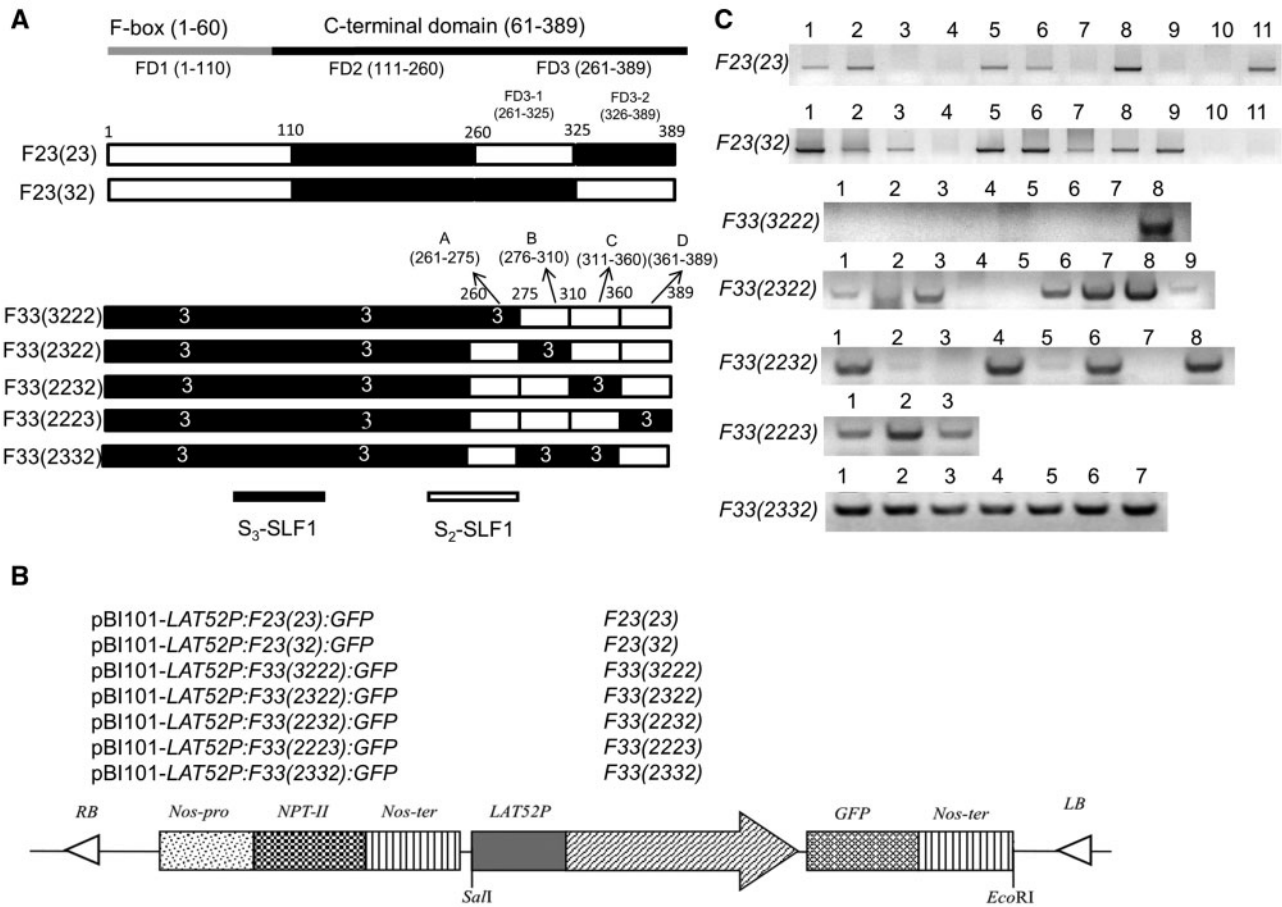

Fig. 3 Seven chimeric proteins of S_2-SLF1 and S_3-SLF1, their transgene constructs and the presence of the transgene in each transgenic line. (A) Schematics of chimeric proteins F23(23), F23(32), F33(3222), F33(2322), F33(2232), F33(2223) and F33(2332) of S_2-SLF1 and S_3-SLF1. The 'F' in the name of each chimeric protein denotes 'SLF', and the first two digits (e.g. 23) represent FD1 and FD2, with each digit indicating whether S_2-SLF1 (2) or S_3-SLF1 (3) contributes the domain. The subsequent digits in parentheses indicate the chimeric nature of FD3. The chimeric protein with the first half (subdomain) of FD3 contributed by S_2-SLF1 and second half (subdomain) contributed by S_3-SLF1 is designated '23', and the chimeric protein with the first half of FD3 contributed by S_3-SLF1 and second half contributed by S_2-SLF1 is designated '32'. For the chimeric proteins that contain one or two of the four mini-domains swapped between S_2-SLF1 and S_3-SLF1, the four digits indicate whether S_2-SLF1 (2) or S_3-SLF1 (3) contributes each of the four mini-domains. The first and last amino acid residues in each subdomain of FD3 and in each mini-domain of FD3 are indicated. (B) Schematics of transgene constructs of chimeric genes *F23(23)*, *F23(32)*, *F33(3222)*, *F33(2322)*, *F33(2232)*, *F33(2223)* and *F33(2332)* in the Ti plasmid pBI101. Abbreviations are the same as defined in **Fig. 1C**. (C) PCR analysis of genomic DNA isolated from T_0 plants of each transgenic line to assess the presence of the respective transgene. The primers used are specific to the *GFP* sequence included in each transgene; the DNA band detected is of the size expected for the amplified DNA fragment. Plant number is indicated above each gel lane.

Absence of plants of the S_2S_2 genotype in the progeny obtained from self-crosses involving *F33(2322):GFP*/S_2S_3, *F33(2232):GFP*/S_2S_3 and *F33(2332):GFP*/S_2S_3 suggests that both wild-type S_2 pollen and S_2 pollen carrying *F33(2322):GFP*, *F33(2232):GFP* or *F33(2332):GFP* are rejected by S_2S_3 pistils. Moreover, the finding that all the progeny plants carried the respective transgenes suggests that S_3 pollen carrying one of these transgenes is compatible with wild-type S_2S_3 pistils. Thus, expression of F33(2322):GFP, F33(2232):GFP and F33(2332):GFP caused breakdown of SI in S_3 pollen, but not in S_2 pollen, suggesting that these three chimeric proteins interact with S_3-RNase. That is, replacing mini-domain B, mini-domain C or both mini-domains B and C of S_2-SLF1 with the corresponding mini-domain(s) of S_3-SLF1 did not affect the ability of S_2-SLF1 to interact with S_3-RNase. Thus, the four amino acids that are different between S_2-SLF1 and S_3-SLF1 in each of these two mini-domains are not required for the interaction with S_3-RNase. In contrast, replacing mini-domain A or mini-domain D of S_2-SLF1 with the corresponding mini-domain of S_3-SLF1 resulted in the inability of *F33(3222):GFP* and *F33(2223):GFP* to cause breakdown of S_3 pollen (**Table 1**), suggesting that both mini-domains are required for the breakdown of SI in S_3 pollen. Thus, at least one of the four amino acids different between S_2-SLF1 and S_3-SLF1 in each of these two mini-domains are required for S_2-SLF1 to interact with S_3-RNase.

FD2 and FD1 + FD2 of S_2-SLF1 contain amino acids required for the breakdown of SI in S_7 and S_{13} pollen, respectively

We previously found that (i) expression of S_2-SLF1 caused breakdown of SI in S_7, S_{12} and S_{13} pollen, suggesting that S_2-SLF1 interacts with S_7-, S_{12}- and S_{13}-RNases, whereas expression of

S_3-SLF1 only caused breakdown of SI in S_{12} pollen, suggesting that S_3-SLF1 interacts with S_{12}-RNase, but not with S_7-RNase or S_{13}-RNase; (ii) expression of neither S_2-SLF1 nor S_3-SLF1 caused breakdown of SI in S_5, S_{6a}, S_{11} or S_{16} pollen, suggesting that S_2-SLF1 and S_3-SLF1 do not interact with S_5-, S_{6a}-, S_{11}- or S_{16}-RNase (Hua et al. 2007, Kubo et al. 2010, Sun and Kao 2013, Williams et al. 2014b).

To determine whether FD3 alone also contains the amino acids of S_2-SLF1 required for the breakdown of SI in S_7 and S_{13} pollen, we first used pollen from T_0 plants *F232:GFP*/S_2S_3 and *F322:GFP*/S_2S_3 to pollinate wild-type plants of S_7S_{13} and S_7S_{16} genotypes, and used pollen from the T_0 plant *F332:GFP*/S_2S_3 to pollinate wild-type plants of S_7S_7 and $S_{13}S_{13}$ genotypes (**Table 2**). As controls, we also used wild-type plants of S_5S_{11}, $S_{6a}S_{6a}$ and $S_{12}S_{12}$ genotypes as females in pollination involving *F232:GFP*/S_2S_3 and *F322:GFP*/S_2S_3, and wild-type plants of S_5S_5, $S_{6a}S_{6a}$, $S_{11}S_{11}$, $S_{12}S_{12}$ and $S_{16}S_{16}$ genotypes as females in pollination involving *F332:GFP*/S_2S_3 (**Table 2**). As expected, all these crosses were compatible, setting large size fruits. For plants in each progeny, analysis of the presence/absence of the transgene and determination of the *S*-genotypes were carried out as described for the self-crosses between these T_0 plants and wild-type S_2S_3 plants. Representative genotyping results are shown in **Supplementary Fig. S2A–C**. We chose the T_1 transgenic plants that contained the S_2-haplotype and one of the following *S*-haplotypes, S_5, S_{6a}, S_7, S_{11}, S_{12}, S_{13} or S_{16}, to examine whether expression of *F232:GFP*, *F322:GFP* or *F332:GFP* caused breakdown of SI in S_5, S_{6a}, S_7, S_{11}, S_{12}, S_{13} and S_{16} pollen. Pollen from at least two T_1 transgenic plants of each *S*-genotype was used in self-crosses with wild-type plants of their respective *S*-genotypes, and, as the results were identical, **Table 2** only lists the result of one of the plants for all these self-crosses.

For example, when pollen from *F232:GFP*/S_2S_5 T_1 plants was self-crossed with a wild-type plant of the S_2S_5 genotype, no fruits were set, and thus these T_1 plants were self-incompatible. As expression of F232:GFP did not cause breakdown of SI in S_2 pollen (**Table 1**), this finding suggests that F232:GFP, like S_2- SLF1, does not cause breakdown of SI in S_5 pollen and thus does not interact with S_5-RNase. *F322:GFP*/S_2S_5 and *F332:GFP*/S_2S_5 were also found to be self-incompatible, suggesting that F322 and F332 behave like S_2-SLF1 and S_3-SLF1 in terms of their lack of interactions with S_5-RNase. Our findings that *F232:GFP*/S_2S_{6a}, *F322:GFP*/S_2S_{6a},

Table 2 Self-incompatibility phenotype of T_1 plants carrying chimeric genes *F232*, *F322*, *F332* or *F33(2332)* in various *S*-genotype backgrounds

Transgene	Cross involving T_0 plant	T_1 plant background	SI phenotype of T_1 plant
F232:GFP	$S_5S_{11}\times$*F232:GFP* /S_2S_3	S_2S_5	SI
F232:GFP	$S_5S_{11}\times$*F232:GFP* /S_2S_3	S_2S_{11}	SI
F232:GFP	$S_{6a}S_{6a}\times$*F232:GFP* /S_2S_3	S_2S_{6a}	SI
F232:GFP	$S_{12}S_{12}\times$*F232:GFP* /S_2S_3	S_2S_{12}	SC
F232:GFP	$S_7S_{13}\times$*F232:GFP* /S_2S_3	S_2S_7	SI
F232:GFP	$S_7S_{13}\times$*F232:GFP* /S_2S_3	S_2S_{13}	SI
F232:GFP	$S_7S_{16}\times$*F232:GFP* /S_2S_3	S_2S_{16}	SI
F322:GFP	$S_7S_{13}\times$*F322:GFP*/S_2S_3	S_2S_7	SC
F322:GFP	$S_7S_{13}\times$*F322:GFP*/S_2S_3	S_2S_{13}	SI
F322:GFP	$S_{12}S_{12}\times$*F322:GFP*/S_2S_3	S_2S_{12}	SC
F322:GFP	$S_5S_{11}\times$*F322:GFP*/S_2S_3	S_2S_5	SI
F322:GFP	$S_5S_{11}\times$*F322:GFP*/S_2S_3	S_2S_{11}	SI
F322:GFP	$S_7S_{16}\times$*F322:GFP*/S_2S_3	S_2S_{16}	SI
F322:GFP	$S_{6a}S_{6a}\times$*F322:GFP*/S_2S_3	S_2S_{6a}	SI
F332:GFP	$S_5S_5\times$*F332:GFP* /S_2S_3	S_2S_5	SI
F332:GFP	$S_{6a}S_{6a}\times$*F332:GFP* /S_2S_3	S_2S_{6a}	SI
F332:GFP	$S_7S_7\times$*F332:GFP* /S_2S_3	S_2S_7	SI
F332:GFP	$S_{11}S_{11}\times$*F332:GFP* /S_2S_3	S_2S_{11}	SI
F332:GFP	$S_{12}S_{12}\times$*F332:GFP* /S_2S_3	S_2S_{12}	SC
F332:GFP	$S_{13}S_{13}\times$*F332:GFP* /S_2S_3	S_2S_{13}	SI
F332:GFP	$S_{16}S_{16}\times$*F332:GFP* /S_2S_3	S_2S_{16}	SI
F33(2332):GFP	$S_5S_5\times$*F33(2332):GFP* /S_2S_3	S_2S_5	SI
F33(2332):GFP	$S_{11}S_{11}\times$*F33(2332):GFP* /S_2S_3	S_2S_{11}	SI
F33(2332):GFP	$S_7S_{13}\times$*F33(2332):GFP* /S_2S_3	S_2S_7	SI
F33(2332):GFP	$S_7S_{13}\times$*F33(2332):GFP* /S_2S_3	S_2S_{13}	SI
F33(2332):GFP	$S_{6a}S_{12}\times$*F33(2332):GFP* /S_2S_3	S_2S_{6a}	SI
F33(2332):GFP	$S_{6a}S_{12}\times$*F33(2332):GFP* /S_2S_3	S_2S_{12}	SC
F33(2332):GFP	$S_{16}S_{16}\times$*F33(2332):GFP* /S_2S_3	S_2S_{16}	SI

F332:GFP/S$_2$S$_{6a}$, *F232:GFP/S$_2$S$_{11}$*, *F322:GFP/S$_2$S$_{11}$*, *F332:GFP/S$_2$S$_{11}$*, *F232:GFP/S$_2$S$_{16}$*, *F322:GFP/S$_2$S$_{16}$* and *F332:GFP/S$_2$S$_{16}$* were all self-incompatible suggest that F232, F322 and F332 also behave like S$_2$-SLF1 and S$_3$-SLF1 in terms of their lack of interactions with S$_{6a}$-, S$_{11}$- and S$_{16}$-RNase.

When pollen from T$_1$ plants *F232:GFP/S$_2$S$_{12}$*, *F322:GFP/S$_2$S$_{12}$* and *F332:GFP/S$_2$S$_{12}$* was used to pollinate a wild-type plant of the *S$_2$S$_{12}$* genotype, large fruits were observed. As expression of these three chimeric *SLF* genes did not cause breakdown of SI in *S$_2$* pollen (**Table 1**), these results suggest that F232, F322 and F332 all interact with and detoxify S$_{12}$-RNase to allow *S$_{12}$* transgenic pollen to be compatible with *S$_2$S$_{12}$* pistils. The finding that these chimeric proteins, with one or two of the domains of S$_2$-SLF1 swapped with the corresponding domain(s) of S$_3$-SLF1, still interact with S$_{12}$-RNase is consistent with the previous finding that both S$_2$-SLF1 and S$_3$-SLF1 interact with S$_{12}$-RNase (Sun and Kao 2013).

Pollination of a wild-type *S$_2$S$_7$* plant by pollen from T$_1$ plants *F322:GFP/S$_2$S$_7$* set large size fruits, but pollination of this wild-type *S$_2$S$_7$* plant by pollen from T$_1$ plants *F232:GFP/S$_2$S$_7$* and *F332:GFP/S$_2$S$_7$* did not set fruits. These results suggest that, unlike interaction with S$_3$-RNase, FD2 of S$_2$-SLF1 contains the amino acids that are required for the breakdown of SI in *S$_7$* pollen. When pollen from T$_1$ plants *F232:GFP/S$_2$S$_{13}$*, *F322:GFP/S$_2$S$_{13}$* and *F332:GFP/S$_2$S$_{13}$* was used to pollinate a wild-type plant *S$_2$S$_{13}$*, no fruit set was observed, suggesting that both FD1 and FD2 of S$_2$-SLF1 contain the amino acids required for the breakdown of SI in *S$_{13}$* pollen.

As chimeric protein F33(2332) was found to interact with S$_3$-RNase (**Table 1**), we used T$_0$ plants *F33(2332):GFP/S$_2$S$_3$* to examine further the findings described above that FD2 and FD1 + FD2 contain the amino acids required for S$_2$-SLF1 to cause breakdown of SI in *S$_7$* pollen and *S$_{13}$* pollen, respectively. Pollen from these T$_0$ plants was used to pollinate wild-type plants of *S$_5$S$_5$*, *S$_{6a}$S$_{12}$*, *S$_7$S$_{13}$*, *S$_{11}$S$_{11}$* and *S$_{16}$S$_{16}$* genotypes. T$_1$ plants containing the transgene *F33(2332):GFP* and of the *S$_2$S$_5$*, *S$_2$S$_{6a}$*, *S$_2$S$_7$*, *S$_2$S$_{11}$*, *S$_2$S$_{12}$*, *S$_2$S$_{13}$* or *S$_2$S$_{16}$* genotype were identified by PCR as described before. Representative results are shown in Supplementary Fig. S2D. We then used pollen from these T$_1$ transgenic plants to pollinate wild-type plants of their respective *S*-genotypes. No fruit set was observed from these self-crosses, except for the self-cross with a wild-type *S$_2$S$_{12}$* plant (**Table 2**). These results are consistent with the findings that amino acids contained in FD2 and FD1 + FD2 are required for S$_2$-SLF1 to cause breakdown of SI in *S$_7$* pollen and *S$_{13}$* pollen, respectively.

Computational modeling of interaction between S$_2$-SLF1 and S$_3$-RNase reveals the involvement of amino acids in mini-domain A and mini-domain D at the interaction surface

We further used computational modeling to examine the possible role of the 16 amino acids that are different between S$_2$-SLF1 and S$_3$-SLF1 in FD3 in the specific interaction of S$_2$-SLF1 with S$_3$-RNase. We first used the I-TASSER server (Zhang 2008, Yang et al. 2015) to predict the tertiary structure of S$_2$-SLF1. After iterative simulations, the top five models were generated. Their C-scores and QMEAN norm scores were examined (**Supplementary Table S2**), and the final refined best S$_2$-SLF1 structure model was obtained, which contained several α-helices in its N-terminal region (the first 60 amino acid residues representing the F-box domain) and many β-sheets in its C-terminal region (representing FD2 and FD3) (**Fig. 4A**). We then used VADAR (http://vadar.wishartlab.com/; Willard et al. 2003) and ProSa-web server (https://prosa.services.came.sbg.ac.at/prosa.php; Wiederstein and Sippl 2007) to validate the quality and reliability of the modeled structure. We found that the distribution of the ψ/φ angles of 90% of the amino acid residues was in the core or allowed regions of the Ramachandran plot, and the overall model quality, *Z*-score (−4.2, 389 amino acids), was within the range of X-ray or nuclear magnetic resonance (NMR)-solved structures of proteins with a similar size (**Supplementary Fig. S3**). We further carried out a structural comparison using the Dali server (http://ekhidna.biocenter.helsinki.fi/dali_server/; Holm and Rosenström 2010), and found that S$_2$-SLF1 is most similar to human F-box/WD repeat-containing protein-7 (Fbxw7, 2ovp-B) and yeast Cdc4/Skp1 (3mks-B), suggesting that its structure is very similar to the WD40-repeat β-propeller domain (**Supplementary Fig. S4**).

Fig. 5 shows an alignment of the amino acid sequences in FD3 of S$_2$-SLF1 and S$_3$-SLF1, along with those of S$_{6a}$-SLF1 and S$_1$-SLF1, both of which interact with S$_3$-RNase based on the in vivo transgenic assay (Sijacic et al. 2004, Kubo et al. 2010, Shu Li and Teh-hui Kao, unpublished results). The four mini-domains are demarcated, and the four amino acids different between S$_2$-SLF1 and S$_3$-SLF1 in each mini-domain are indicated with asterisks above the alignment. Interestingly, of the 16 amino acids that are different between S$_2$-SLF1 and S$_3$-SLF1, 10 are conserved among all three SLFs (S$_2$-SLF1, S$_{6a}$-SLF1 and S$_1$-SLF1) that interact with S$_3$-RNase: three each in mini-domain A, mini-domain B and mini-domain D, and one in mini-domain C. To examine further the role of these conserved amino acids in the interaction with S$_3$-RNase, we first predicted the tertiary structure of S$_3$-RNase using the same method as used for S$_2$-SLF1 (**Supplementary Figs. S5, S6**), and then used the web-based molecular docking server, ClusPro 2.0 (https://cluspro.org; Comeau et al. 2004, Kozakov et al. 2017), to predict the interaction between S$_2$-SLF1 and S$_3$-RNase, as no crystal structure of this protein complex is yet available. The tertiary structure of S$_3$-RNase (the ligand) was docked onto that of S$_2$-SLF1 (the receptor) without the first 95 amino acids containing the F-box domain. Using the scoring scheme for the balanced mode, which takes electrostatic attractions, surface hydrophobicity and van der Waals interactions into consideration, a docked structure with the largest cluster size of 62 is obtained, with the center lowest energy as − 882.7 (**Fig. 4B**). This docked structure was visualized by PyMOL (The PyMOL Molecular Graphics System, Version 1.8.6.0 Schrödinger, LLC.), revealing that FD3 of S$_2$-SLF1 and the N-terminal part of S$_3$-RNase were at the interaction surface (**Fig. 4B**), and that both mini-domain A and mini-domain D were in contact with the surface of S$_3$-RNase (**Fig. 4C**). These findings are consistent with the results from the domain-swapping experiments

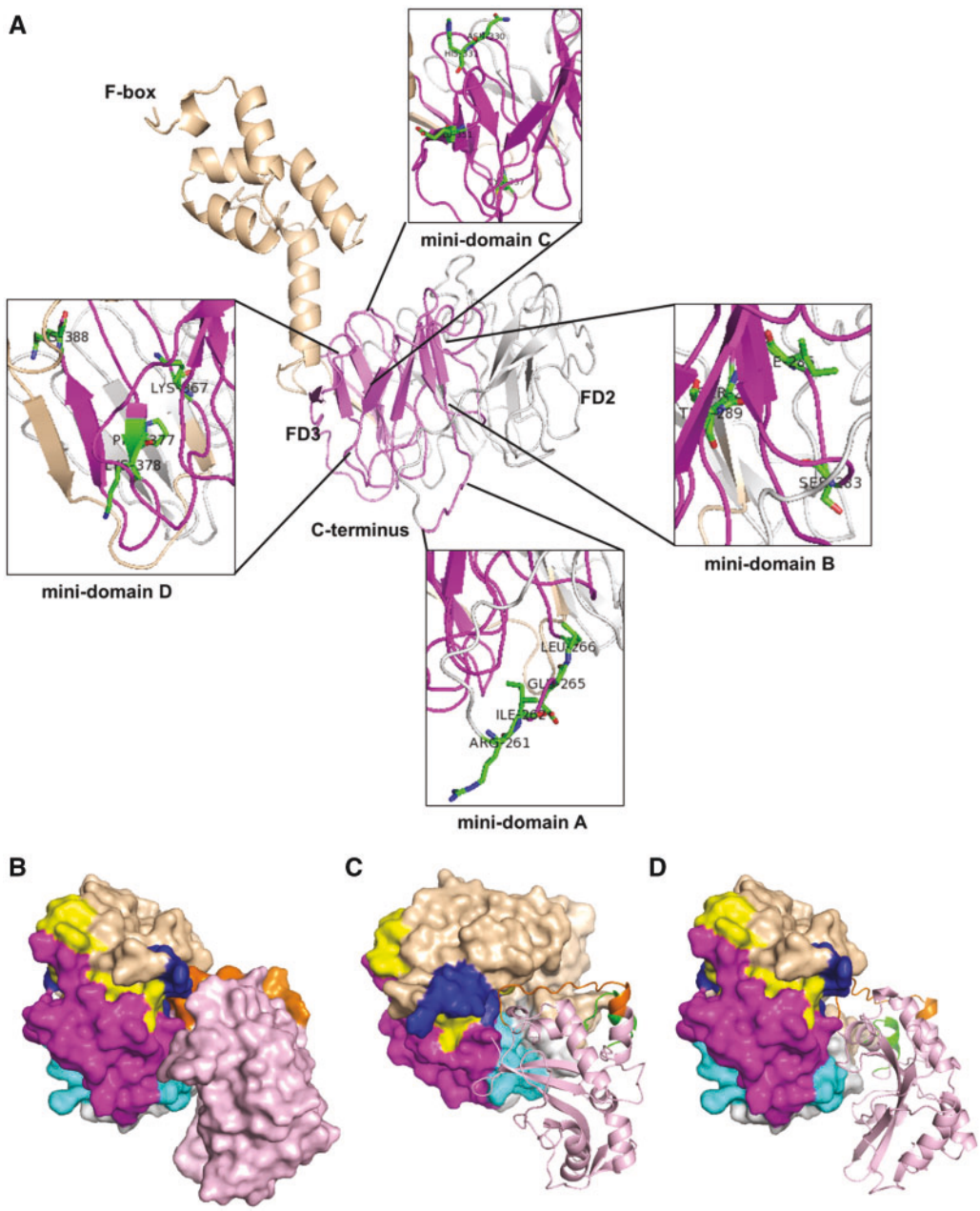

Fig. 4 Computational modeling of S_2-SLF1 and molecular docking of S_3-RNase onto S_2-SLF1, as visualized in PyMOL. (A) Predicted tertiary structure of S_2-SLF1 shown as a ribbon diagram. FD1, FD2 and FD3 are shown in beige, gray and purple colors, respectively. The four mini-domains of FD3 are enlarged in the boxes. In each mini-domain, the four amino acids that are different between S_2-SLF1 and S_3-SLF1 are labeled and shown in stick mode. (B) Predicted docked structure of S_2-SLF1 (receptor) and S_3-RNase (ligand) through the ClusPro molecular docking web server using a balanced-mode scoring scheme. The structure is shown in surface mode. FD2 and FD3 are shown in beige and gray, respectively. The four amino acids different between S_2-SLF1 and S_3-SLF1 in mini-domain A are shown in blue; the four amino acids different between S_2-SLF1 and S_3-SLF1 in mini-domain D are shown in cyan. S_3-RNase is shown in light pink, with hypervariable regions A and B (HVA and HVB) labeled in orange and green, respectively. (C) S_2-SLF1 in the docked structure shown in surface mode, with the four amino acids different between S_2-SLF1 and S_3-SLF1 in mini-domain A highlighted in blue. S_3-RNase is shown as a ribbon diagram. (D) S_2-SLF1 in the docked structure shown in surface mode, with the four amino acids different between S_2-SLF1 and S_3-SLF1 in mini-domain D highlighted in cyan. S_3-RNase is shown as a ribbon diagram.

showing that mini-domain A and mini-domain D are required for interactions with S_3-RNase. Of the eight amino acids different between S_2-SLF1 and S_3-SLF1 in these two mini-domains, Glu265 and Leu266 are predicted to be in contact with S_3-RNase (**Fig. 4C**), and Lys378 is predicted to be very close to the surface of S_3-RNase (**Fig. 4D**). Interestingly, these three amino acids are among the 10 amino acids that are conserved among S_2-SLF1, S_1-SLF1 and S_{6a}-SLF1 that interact with S_3-RNase, supporting their possible involvement in the specific interaction between S_2-SLF1 and S_3-RNase.

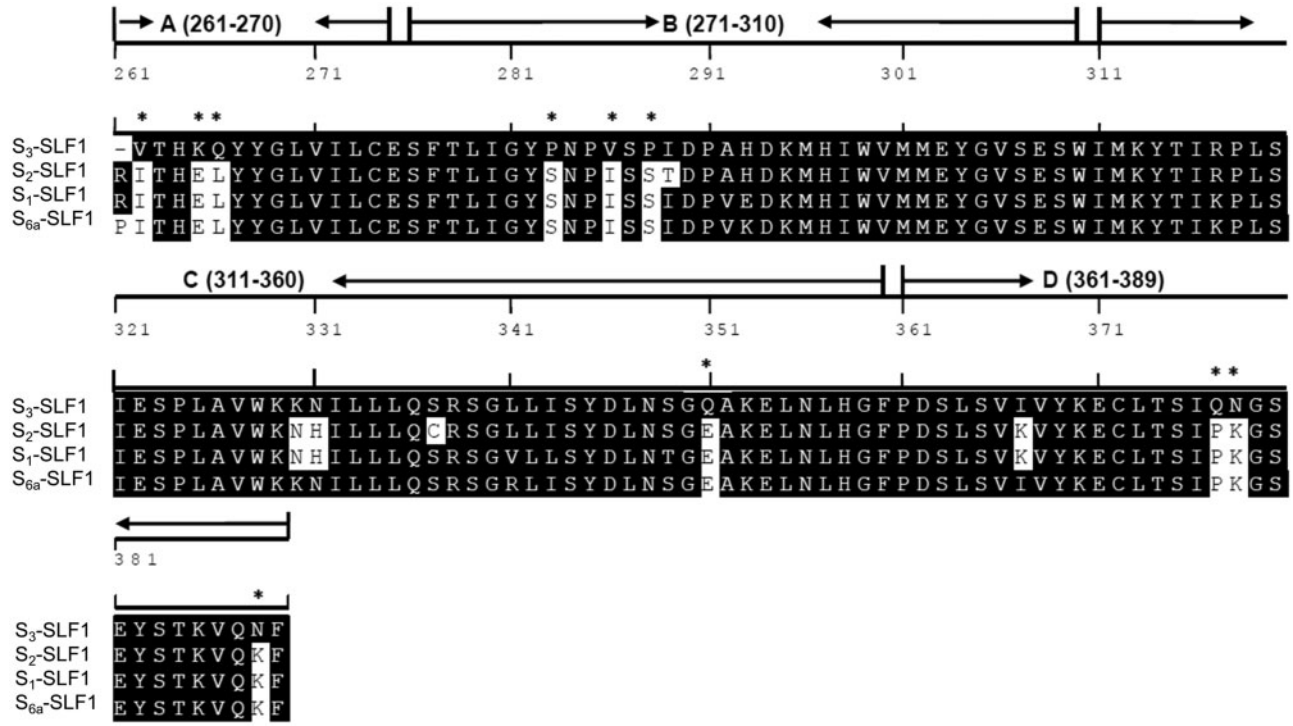

Fig. 5 Alignment of deduced amino acid sequences in FD3 of three allelic variants of SLF1 that interacts with S_3-RNase and one allelic variant that does not, based on genetic or biochemical evidence. The four mini-domains are demarcated and labeled, A, B, C and D. The first and last amino acid residues of each mini-domain are indicated. The 16 amino acids that are different between S_2-SLF1 and S_3-SLF1 are marked in a white background. Each mini-domain contains four of these 16 amino acids. Ten of these 16 amino acids are conserved in two other allelic variants of SLF1, S_1-SLF1 and S_{6a}-SLF1, both of which interact with S_3-RNase, based on in vivo transgenic assay. These 10 amino acid residues are indicated with asterisks.

Discussion

The complex interaction relationships between multiple SLF proteins produced by pollen of a given *S*-haplotype (e.g. 17 produced by both the S_2-haplotype and S_3-haplotype of *P. inflata*; Williams et al. 2014b) and S-RNases produced by pistils of a large number of *S*-haplotypes (e.g. 32 have been reported for *Petunia*; Sims and Robbins 2009) provide a good opportunity for studying biochemical and structural bases of protein–protein interactions. Results of the in vivo transgenic assay suggest that differential interactions exist between different SLF proteins of a given *S*-haplotype and S-RNases, and between allelic variants of each SLF and S-RNases (Sijacic et al. 2004, Kubo et al. 2010, Sun and Kao 2013, Williams et al. 2014b). S_2-SLF1 and S_3-SLF1 (S_2 and S_3 allelic variants of SLF1) is an interesting pair of SLF proteins to study, as, on the one hand, they show differential interactions with several S-RNases, but on the other hand, they also share the same interaction relationships with a number of other S-RNases. Specifically, our previous results suggest that (i) S_2-SLF1, but not S_3-SLF1, interacts with S_1-, S_3-, S_7- and S_{13}-RNases; (ii) both S_2-SLF1 and S_3-SLF1 interact with S_{12}-RNase; and (iii) neither S_2-SLF1 nor S_3-SLF1 interacts with S_2-, S_5-, S_{6a}-, S_{11}- or S_{16}-RNase (Williams et al. 2014b). Moreover, we subsequently used Co-IP to show that S_2-SLF1 interacted with S_3-RNase (Sun and Kao 2013), thus confirming the genetic interaction suggested from the in vivo transgenic assay. The differential interactions of S_2-SLF1 and S_3-SLF1 with the S-RNases examined perhaps are not surprising given that S_2-SLF1 and S_3-SLF1 are 88.7% identical in their amino acid sequences. The question is then, among the S-RNases we have examined, which of the 44 amino acids that are different between S_2-SLF1 and S_3-SLF1 are required for S_2-SLF1 to interact with the four additional S-RNases? In this work, we have adopted the domain-swapping approach to begin to address this question, with the goal of narrowing the candidate amino acids involved in the specific interaction of S_2-SLF1 with S_3-RNase.

We previously divided S_2-SLF1 into three functional domains, FD1, FD2 and FD3, with each domain containing approximately equal numbers of amino acids that are different between S_2-SLF1 and S_3-SLF1 (Hua et al. 2007). It is thus convenient to use the demarcation of these three domains (**Fig. 1A**) to construct chimeric *SLF* genes, by swapping coding sequences for one of these three domains, to see which one(s) contain(s) the amino acids required for the interaction between S_2-SLF1 and S_3-RNase. We have constructed a total of 11 chimeric genes of S_2-*SLF1* and S_3-*SLF1*, and generated S_2S_3 transgenic lines for all of them. The results from the first series, involving *F322*, *F232* and *F233*, suggest that (i) F322 and F232, but not F233, interact with and detoxify S_3-RNase in S_3 pollen to allow S_3 transgenic pollen to be compatible with the S_2S_3 pistil; and (ii) FD3 of S_2-SLF1 is required for its interaction

with S_3-RNase (**Table 1**). However, the involvement of FD1, or FD2, in conjunction with FD3, could not be ruled out. That is, the interaction of F322 with S_3-RNase might also require FD2 of S_2-SLF1, and the interaction of F232 with S_3-RNase might also require FD1 of S_2-SLF1. To examine these possibilities, we tested *F332* and found that it could still break down SI in S_3 pollen (**Table 1**), suggesting that F332, containing FD1 and FD2 of S_3-SLF1, can still interact with S_3-RNase. Thus, studying this series of four chimeric genes, *F322*, *F232*, *F233* and *F332*, has allowed us to narrow down the candidate amino acids required for the interaction between S_2-SLF1 and S_3-RNase from 44 down to the 16 in FD3.

In the second series, we wished to narrow down further the candidate amino acids required for the interaction of S_2-SLF1 with S_3-RNase. As F232 could interact with S_3-RNase, we used it as the backbone and divided FD3 into two subdomains, each containing eight of the 16 different amino acids. We found that the resulting chimeric genes, *F23(23)* and *F23(32)*, could not break down SI in S_3 pollen (**Table 1**). Thus, both subdomains of FD3 contain the amino acids required for the interaction between S_2-SLF1 and S_3-RNase. In the third series, we used *F332* as the backbone and divided FD3 into four mini-domains, A, B, C and D (each containing four of the 16 different amino acids), to make four chimeric genes, *F33(3222)*, *F33(2322)*, *F33(2232)* and *F33(2223)*. *F33(2322)* and *F33(2232)*, but not *F33(3222)* or *F33(2223)*, caused breakdown of SI in S_3 pollen (**Table 1**), suggesting that one or more of the four amino acids different between S_2-SLF1 and S_3-SLF1 in mini-domain A and in mini-domain D are required for S_2-SLF1 to interact with S_3-RNase, and that none of the four different amino acids in either mini-domain B or mini-domain C is required. To confirm this interpretation, we made an additional chimeric gene, *F33(2332)*, and found that it caused breakdown of SI in S_3 pollen (**Table 1**). This series of studies has allowed us to conclude that both mini-domain A and mini-domain D contain amino acids that are required for S_2-SLF1 to interact with S_3-RNase.

In the final series, we asked the question of which domain(s) of S_2-SLF1 is (are) required for the breakdown of SI in S_7 pollen and S_{13} pollen. We examined three chimeric genes, *F322*, *F232* and *F332*, all of which caused breakdown of SI in S_3 pollen. We included S_{12} pollen as a positive control, because we would expect that expression of any chimeric protein of S_2-SLF1 and S_3-SLF1 should cause breakdown of SI in S_{12} pollen. We included S_5, S_{6a}, S_{11} and S_{16} pollen as negative controls, because we would expect that expression of chimeric proteins of S_2-SLF1 and S_3-SLF1 should not cause breakdown of SI in pollen of these *S*-haplotypes. All the results obtained were as expected, suggesting that these chimeric proteins of S_2-SLF1 and S_3-SLF1 behave as S_2-SLF1 and S_3-SLF1 in their ability, or inability, to cause breakdown of SI in pollen of the *S*-haplotypes tested. That is, the overall structure of the chimeric proteins remains unaltered. F322, but not F232 or F332, caused breakdown of SI in S_7 pollen, suggesting that FD2, but not FD1, of S_2-SLF1 is required for the interaction with S_7-RNase. However, whether the interaction also requires FD3, in conjunction with FD2, would have to be addressed by examining an additional chimeric protein, F323, to see whether it could cause breakdown of SI in S_7 pollen. If it did, then this would rule out the involvement of the 16 amino acids that are different between S_2-SLF1 and S_3-SLF1 in FD3. None of the chimeric proteins, F322, F232 and F332, caused breakdown of SI in S_{13} pollen, suggesting that both FD1 and FD2 domains are required for interaction with S_{13}-RNase. Similarly, the role of FD3, if any, would have to be addressed by examining another additional chimeric protein, F223, to see whether it can cause breakdown of SI in S_{13} pollen. We further examined *F33(2332)* in this series, as we showed in the previous series that F33(2332) interacted with S_3-RNase (**Table 1**). Consistent with the results of *F332*, *F33(2332)* could only cause breakdown of SI in S_{12} pollen.

The pair of SLF proteins, S_2-SLF1 and S_3-SLF1, chosen for this study is interesting in that even though they are allelic variants of the same type of SLF protein, they differ in >10% of their amino acid sequences. Many other allelic pairs of the same type of SLF protein share higher degrees of amino acid sequence identity, e.g. S_2-SLF5 and S_3-SLF5 differ in only eight of the 388 amino acids (but only S_2-SLF5 interacts with S_{12}-RNase based on the in vivo transgenic assay), and S_2-SLF6 and S_3-SLF6 differ in only four of the 390 amino acids (but only S_3-SLF6 interacts with S_2-RNase based on the in vivo transgenic assay) (Williams et al. 2014b). The approach of site-directed mutagenesis could be used to identify the amino acid(s) that differentiate(s) such an allelic pair in their interactions with S-RNases. Differential interactions with S-RNases also exist between different types of SLF proteins of a given *S*-haplotype, and, in this case, the amino acid sequences of paralogous SLF proteins differ over a wide range, e.g. 45–88% for the 17 SLF proteins of the S_2-haplotype (Williams et al. 2014b). The series of domain-swapping experiments performed in this work demonstrates the feasibility of using this approach as a first step towards addressing the biochemical basis of differential interactions with S-RNases for a pair of SLF proteins that differ by a large number of amino acids.

In this work, we have narrowed down the candidate amino acids required for the specific interaction of S_2-SLF1 with S_3-RNase from 44 amino acids to eight, with four in mini-domain A and four in mini-domain D. We further examined the role of these eight candidate amino acids by using I-TASSER to predict possible tertiary structures of S_2-SLF1 and S_3-RNase, and using the molecular docking program ClusPro 2.0 to predict the interaction between S_2-SLF1 and S_3-RNase. The docking result indicated that FD3 of S_2-SLF1 and the N-terminal part of S_3-RNase are at the interaction interface. The docked structure also revealed that both mini-domains are in contact with S_3-RNase. More specifically, Glu265 and Leu266 in mini-domain A are predicted to be in direct contact with S_3-RNase, and Lys378 in mini-domain D is predicted to be very close to S_3-RNase. One could examine additional allelic variants of SLF1 for their interaction relationships with S_3-RNase, and assess whether any of the eight amino acids different between S_2-SLF1 and S_3-SLF1 in mini-domain A and mini-domain D is conserved among all the allelic variants that interact with S_3-RNase, but divergent among all of those that do not. To date, 20 alleles of *SLF1* of *P. inflata* have been identified and sequenced (including S_2-*SLF1* and S_3-*SLF1* studied in this work), and their deduced amino acid sequences are 85.8–99.7%

identical (Williams et al. 2014a). The in vivo transgenic assay used in this study could be used to determine whether expression of any of the as yet untested *SLF1* alleles could cause specific breakdown of SI in S_3 pollen. As an example, we present in **Fig. 5** an amino acid alignment of four of these SLF1 allelic variants for which we have obtained the interaction relationships with S_3-RNase based on the in vivo transgenic assay. S_1-SLF1, S_{6a}-SLF1 and S_2-SLF1, but not S_3-SLF1, interact with S_3-RNase (Sijacic et al. 2004, Kubo et al. 2010, Shu Li and Tehhui Kao, unpublished results). For the 16 amino acids in FD3 that are different between S_2-SLF1 and S_3-SLF1, 10 are conserved among S_1-SLF1, S_2-SLF1 and S_{6a}-SLF1, three are located in mini-domain A, including Glu265 and Leu266, and three are located in mini-domain D, including Lys378. Interestingly, three of the conserved amino acids are located in mini-domain B. As we have shown that the amino acid differences between S_2-SLF1 and S_3-SLF1 in mini-domain B are not required for the interaction of S_2-SLF1 with S_3-RNase, the conservation of these residues between S_1-SLF1, S_2-SLF1 and S_{6a}-SLF1 may implicate their role in the interaction of these three allelic variants with the same S-RNase(s) other than S_3-RNase. Consistent with this possibility are our findings that FD2 of S_2-SLF1 is required for the breakdown of SI in S_7 pollen, and both FD1 and FD2 of S_2-SLF1 are required for the breakdown of SI in S_{13} pollen. This highlights the diversity and complexity of interactions between SLF proteins and S-RNases.

Previously, a similar domain-swapping approach was used by Li et al. (2017) to study an allelic pair of SLF1 proteins of *P. hybrida*, (Ph)S_3-SLF1 and (Ph)S_{3L}-SLF1. These two SLF proteins (with 89.7% sequence identity) differ in 40 amino acids, and the results of the in vivo transgenic assay suggest that (Ph)S_{3L}-SLF1, but not (Ph)S_3-SLF1, interacts with (Ph)S_3-RNase. It should be noted that, even though the allele number, 3, was also used to denote (Ph)S_3-RNase and (Ph)S_3-SLF1, (Ph)S_3-RNase is not the same protein as S_3-RNase of *P. inflata*, and neither (Ph)S_3-SLF1 nor (Ph)S_{3L}-SLF1 is the same protein as S_3-SLF1 (or S_2-SLF1) of *P. inflata*. (Ph)S_3-RNase and S_3-RNase share 78.7% amino acid identity. (Ph)S_3-SLF1 and (Ph)S_{3L}-SLF1 are 90.5% and 86.6% identical to S_3-SLF1, respectively, and (Ph)S_3-SLF1 and (Ph)S_{3L}-SLF1 are 91.3% and 94.6% identical to S_2-SLF1, respectively. Li et al. (2017) also divided (Ph)S_{3L}-SLF1 and (Ph)S_3-SLF1 into the three domains (FD1, FD2 and FD3) defined in Hua et al. (2007), and used the in vivo transgenic assay to examine the ability of six chimeric proteins to interact with S_3-RNase. The results suggest that chimeric proteins 3-L-L, 3-3-L and L-3-L, but not L-3-3 or L-L-3, interact with S_3-RNase. These results seemed to suggest that FD3 alone contains the amino acids required for (Ph)S_{3L}-SLF1 to interact with S_3-RNase. However, inexplicably, Li et al. (2017) found that chimeric protein 3-L-3, not containing FD3 of (Ph)S_{3L}-SLF1, also caused breakdown of SI in S_3 pollen (and thus would interact with S_3-RNase). This internal inconsistency is puzzling. They further identified amino acid residue 293 that might determine the interaction specificity of (Ph)S_{3L}-SLF1. This residue is in mini-domain B based on our definition (**Fig. 5**), and it is glutamate in PhS_{3L}-SLF1 and histidine in (Ph)S_3-SLF1. Interestingly, when His293 in (Ph)S_3-SLF1 was changed to Glu293, the resulting SLF acquired the ability of (Ph)$_{S3L}$-SLF1 to cause the breakdown of SI in S_3 pollen. That a single amino acid might be involved in interaction specificity is consistent with our having narrowed down the amino acids required for the specific interaction of S_2-SLF1 with S_3-RNase from the 44 that are different between S_2-SLF1 and S_3-SLF1 down to three.

Materials and Methods

Plant materials

S_2S_2, S_3S_3 and S_2S_3 genotypes of *P. inflata* were described by Ai et al. (1990), and S_5S_{11} and S_7S_{13} genotypes were described by Wang et al. (2001). The S_{12}-haplotype was identified from a population of plants germinated from seeds obtained from the Ornamental Plant Germplasm Center of the United States Department of Agriculture. The S_{12}-carrying plants were bud-selfed to produce plants of the $S_{12}S_{12}$ genotype. The S_6S_6 genotype was described by Quattrocchio et al. (1999), and S_6 was designated S_{6a} to distinguish it from our previously reported and genetically distinct S_6-haplotype (Wang et al. 2001). Plants of $S_{6a}S_{6a}$, $S_{12}S_{12}$, S_5S_{11} and S_7S_{13} genotypes were crossed with plants of the S_2S_3 genotype to obtain S_2S_5, S_2S_{6a}, S_2S_7 and S_2S_{12} genotypes (Sun and Kao 2013). All *S*-genotypes are verified by PCR, as described in Sun and Kao (2013).

Generation of Ti plasmid constructs and plant transformation

A total of 11 Ti plasmid constructs were made in pBI101, and each contained the *Late Anther Tomato 52* promoter (*LAT52P*) (Twell et al. 1990) driving the expression of a particular chimeric *SLF* gene fused at its last codon with the coding sequence of GFP. These 11 constructs (schematically represented in **Figs. 1C and 3B**) were pBI101-*LAT52P:F33(3222):GFP*, pBI101-*LAT52P:F33(2322):GFP*, pBI101-*LAT52P:F33(2232):GFP*, pBI101-*LAT52P:F33(2223):GFP*, pBI101-*LAT52P:F33(2332):GFP* and pBI101-*LAT52P:F332:GFP*. They were made using the In-Fusion HD Cloning Kit (Clontech) as described below, and all the primers used in PCRs are listed in **Supplementary Table S1**. Each individual chimeric *SLF* gene was amplified by first performing a series of PCRs to generate DNA fragments for *FD3* from S_2-*SLF1*, *FD1* and *FD2* from S_3-*SLF1*, and *FD3-A*, *FD3-B*, *FD3-C* and *FD3-D* from both S_2-*SLF1* and S_3-*SLF1*. These 11 fragments contained overlapping regions, 14–16 bp in length, designed to allow hybridization between DNA fragments in the correct order. Appropriate fragments were used as a mixed template in overlapping PCRs to generate each of the following chimeric DNA fragments of approximately 1.2 kb in length: *F33(3222)*, *F33(2322)*, *F33(2232)*, *F33(2223)* and *F33(2332)*. Using *F33(3222)* as an example, the fragment containing FD1, FD2 and FD3-A from S_3-SLF1, and the fragment containing FD3-B, FD3-C and FD3-D from S_2-SLF1, were first generated by PCRs, and then pooled for use as template for overlapping PCR, using the LAT52 F332-ATGFW (In-Fusion) and F332-S2SLF1TAGREV (In-Fusion) primer pairs, to produce the chimeric DNA fragment *F33(3222)*. The *LAT52* promoter region was amplified using primers INFUSION PBI LAT52FOR/INFUSION LAT52REV, resulting in a 0.6 kb fragment. The *GFP* coding sequence and the *Nos* transcription terminator region were amplified using primers INFUSION GFP FOR/INFUSION PBI NOS REV, resulting in a 1.1 kb fragment. A four-way ligation was performed to join the 0.6 kb *LAT52* promoter fragment, the 1.2 kb fragment containing a specific chimeric *SLF* gene, the 1.1 kb fragment containing *GFP* and *Nos*, and *Sal*I-*Eco*RI-digested pBI101, in one reaction according to the In-Fusion HD cloning protocol and as described by Sun and Kao (2013). All the 11 Ti plasmid constructs were electroporated into competent cells of *A. tumefaciens* (LBA4404), and subsequently used in *Agrobacterium*-mediated transformation of leaf tissues of *P. inflata* as described by Lee et al. (1994).

Amino acid sequence alignment and analysis

Sequence alignments were performed using MEGA 6 (Tamura et al. 2013) and ClustalW (Thompson et al. 1994). Alignments of deduced amino acid sequences were produced using default settings.

Visualization of GFP fluorescence

As described by Meng et al. (2009), mature pollen was germinated in pollen germination medium (Lee et al. 1996) for a maximum period of 2 h. Pollen tubes were visualized using a Nikon Eclipse 90i epifluorescence microscope.

Genomic DNA isolation and genotyping by PCR

Isolation of genomic DNA was performed using Plant DNAzol® Reagent (Invitrogen) following the manufacturer's protocol as described in Meng et al. (2011). S-haplotype-specific primers for *S-RNase* or *SLF1* were used to genotype all of the plants used in these experiments. The primer pairs used are shown in **Supplementary Table S1**.

Protein structure prediction and protein–protein docking analysis

S_2-SLF1 and S_3-RNase protein structures were modeled using the I-TASSER server (http://zhanglab.ccmb.med.umich.edu/I-TASSER/) (Zhang 2008, Yang et al. 2015). Backbone and overall model quality were evaluated by the VADAR version 1.8 program (http://vadar.wishartlab.com/) (Willard et al. 2003) and the ProSa-web program (https://prosa.services.came.sbg.ac.at/prosa.php). The refined protein model of S_2-SLF1 and S_3-RNase using SWISS-MODEL (https://swissmodel.expasy.org/qmean/) were used for protein–protein docking analysis by ClusPro (https://cluspro.bu.edu/login.php) (Comeau et al. 2004). The docked results were visualized and analyzed using PyMOL. All the structural images were produced by the PyMOL molecular visualization package (The PyMOL Molecular Graphics System, Version 1.8 Schrödinger, LLC.).

Supplementary data

Supplementary data are available at PCP online.

Funding

This work was supported by the U.S. National Science Foundation [IOS-1146182 and IOS-1645557 to T.-h.K.].

Acknowledgments

We thank C.A. Natale for help with some of the experiments, L.S. Burghard for greenhouse management, and H. Hao for general laboratory help.

Disclosures

The authors have no conflict of interests to declare.

References

Ai, Y., Singh, A., Coleman, C.E., Ioerger, T.R., Kheyr-Pour, A. and Kao, T.-h. (1990) Self-incompatibility in *Petunia inflata*: isolation and characterization of cDNAs encoding three S-allele-associated proteins. *Sex. Plant Reprod.* 3: 130–138.

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

Comeau, S.R., Gatchell, D.W., Vajda, S. and Camacho, C.J. (2004) ClusPro: a fully automated algorithm for protein–protein docking. *Nucleic Acids Res.* 32: 96–99.

De Nettancourt, D. (2001) The genetics of self-incompatibility. *In* Incompatibility and Incongruity in Wild and Cultivated Plants. pp. 26–27. Spinger-Verlag, Berlin.

Goldraij, A., Kondo, K., Lee, C.B., Hancock, C.N., Sivaguru, M., Vazquez-Santana, S., et al. (2006) Compartmentalization of S-RNase and HT-B degradation in self-incompatible *Nicotiana*. *Nature* 439: 805–810.

Holm L and Rosenström P. (2010) Dali server: conservation mapping in 3D. *Nucleic Acids Res.* 38: W545–W549.

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.

Kozakov, D., Hall, D.R., Xia, B., Porter, K.A., Padhorny, D., Yueh, C., et al. (2017) The ClusPro web server for protein–protein docking. *Nature Protoc.* 12: 255–278.

Kubo, K.I., Entani, T., Takara, A., Wang, N., Fields, A.M., Hua, Z., et al. (2010) Collaborative non-self recognition system in S-RNase-based self-incompatibility. *Science* 330: 796–799.

Kubo, K.I., Paape, T., Hatakeyama, M., Entani, T., Takara, A., Kajihara, K., et al. (2015) Gene duplication and genetic exchange drive the evolution of S-RNase-based self-incompatibility in *Petunia*. *Nat. Plants* 1: 14005.

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

Lee, H.-S., Karunanandaa, B., McCubbin, A., Gilroy, S. and Kao, T.-h. (1996) PRK1, a receptor-like kinase of *Petunia inflata*, is essential for postmeiotic development of pollen. *Plant J.* 9: 613–624.

Li, J., Zhang, Y., Song, Y., Zhang, H., Fan, J., Li, Q., et al. (2017) Electrostatic potentials of the S-locus F-box proteins contribute to the pollen S specificity in self-incompatibility in *Petunia hybrida*. *Plant J.* 89: 45–57.

Li, S., Sun, P., Williams, J.S. and Kao, T.-h. (2014) Identification of the self-incompatibility locus F-box protein-containing complex in *Petunia inflata*. *Plant Reprod.* 27: 31–45.

Li, S., Williams, J.S., Sun, P. and Kao, T.-h. (2016) All 17 types of S-locus F-box proteins of S_2- and S_3-haplotypes of *Petunia inflata* are assembled into similar SCF complexes with specific function in self-incompatibility. *Plant J.* 87: 606–616.

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

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* 2011: plr016.

Meng, X., Hua, Z., Wang, N., Fields, A.M., Dowd, P.E. and Kao, T.-h. (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.

Quattrocchio F., Wing J., van der Woude K., Souer E., de Vetten N., Mol J., et al. (1999) Molecular analysis of the anthocyanin2 gene of petunia and its role in the evolution of flower color. *Plant Cell* 11: 1433–1444.

Sijacic, P., Wang, X., Skirpan, A.L., Wang, Y., Dowd, P.E., McCubbin, A.G., et al. (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. Edited by Gerats, T. and Strommer, J. pp. 85–106. Springer, New York.

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

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., Stecher, G., Peterson, D., Filipski, A. and Kumar, S. (2013) MEGA6: molecular evolutionary genetics analysis version 6.0. *Mol. Biol. Evol*. 30: 2725–2729.

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.

Twell, D., Yamaguchi, J. and McCormick, S. (1990) Pollen-specific gene expression in transgenic plants: coordinate regulation of two different tomato gene promoters during microsporogenesis. *Development* 109: 705–713.

Wang, X., Hughes, A.L., Tsukamoto, T., Ando, T., and Kao, T.-h. (2001) Evidence that intragenic recombination contributes to allelic diversity of the S-RNase gene at the self-incompatibility (*S*) locus in *Petunia inflata*. *Plant Physiol*. 125: 1012–1022.

Wiederstein, M. and Sippl, M.J. (2007) ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins. *Nucleic Acids Res*. 35: W407–W410.

Willard L., Ranjan A., Zhang H.Y., Monzavi H., Boyko R.F., Sykes B.D., et al. (2003) VADAR: a web server for quantitative evaluation of protein structure quality. *Nucleic Acids Res*. 31: 3316–3319.

Williams, J.S., Der, J.P. and Kao, T.-h. (2014a) 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.

Williams, J.S., Natale, C.A., Wang, N., Li, S., Brubaker, T.R., Sun, P., et al. (2014b) Four previously identified *Petunia inflata S-locus F-box* genes are involved in pollen specificity in self-incompatibility. *Mol. Plant* 7: 567–569.

Yang, J., Yan, R., Roy, A., Xu, D., Poisson, J. and Zhang, Y. (2015) The I-TASSER suite: protein structure and function prediction. *Nature Methods* 12: 7–8.

Zhang, Y. (2008) I-TASSER server for protein 3D structure prediction. *BMC Bioinformatics* 9: 40.