Transcriptome Analysis of Self- and Cross-pollinated Pistils of Japanese Apricot (Prunus mume Sieb. et Zucc.)

Tsuyoshi Habu1 and Ryutaro Tao2*

1Faculty of Agriculture, Ehime University, Matsuyama 790-8566, Japan
2Graduate School of Agriculture, Kyoto University, Kyoto 606-8502, Japan

Solanaceae, Rosaceae, and Plantaginaceae exhibit the S-RNase-based gametophytic self-incompatibility (GSI) system. This type of GSI is controlled by a single polymorphic locus (S locus) containing the pistil S determinant gene, S-ribonuclease (S-RNase), and the pollen S determinant, the S locus F-box gene (SFB/SLF). In addition to these determinant genes, non-S factors, called modifier genes, are required for the GSI reaction. Here, we conducted large-scale transcriptome analysis of unpollinated, self-pollinated, and cross-pollinated pistils of Japanese apricot (Prunus mume Sieb. et Zucc. cv. Nanko) to capture all of the molecular events induced by the GSI reaction in Prunus, using next-generation sequencing technologies. We obtained 40,061 unigenes from 77,521,310 reads from pollinated and unpollinated pistils and pollen grains. Among these unigenes, 29,985 and 27,898 unigene sequences showed at least one hit against the NCBI nr and TAIR10 protein databases, respectively, in BLASTX searches using an E-value cutoff of 1e-6. Digital expression analysis showed that 8,907 and 10,190 unigenes were expressed at significantly different levels between unpollinated (UP) and cross-pollinated (CP) pistils and between UP and self-pollinated (SP) pistils, respectively. The expression of 4,348 unigenes in both CP and SP pollination was commonly and significantly different from that in UP, while the expression of 4,559 and 5,842 unigenes in CP and SP, respectively, was specifically and significantly different from UP. The expression of 2,227 unigenes was up-regulated both in CP and SP compared with UP. Genes supposedly involved in S-RNase-based GSI were included among the unigenes up-regulated by pollination, while no unigenes homologous to the solanaceous pistil modifiers HT-B or 120K were included among the unigenes up-regulated by pollination or in the whole unpollinated/pollinated pistil transcriptome. We discuss the distinct molecular mechanism of S-RNase-based GSI in Prunus.

Key Words: EST, gametophytic self-incompatibility, next-generation sequencing technology, RNA-Seq, S-RNase.

Introduction

Self-incompatibility (SI) is a widespread mechanism in flowering plants that prevents inbreeding and promotes out-crossing. Although there are several different SI mechanisms in flowering plants, the Solanaceae, Rosaceae, and Plantaginaceae exhibit the S-RNase-based gametophytic self-incompatibility (GSI) system. This type of GSI is controlled by a single polymorphic locus (S locus) encoding pistil S and pollen S determinants. The former is the S-ribonuclease gene (S-RNase) and the latter is a pollen-expressed F-box gene called the S haplotype-specific F-box gene (SFB) in Prunus of the Rosaceae and the S locus F-box gene (SLF) in the Solanaceae and Plantaginaceae (Entani et al., 2003; Lai et al., 2002; Qiao et al., 2004; Sijacic et al., 2004; Ushijima et al., 2003, 2004; Yamane et al., 2003; Wang et al., 2004). Although the Prunus pollen S was originally referred to by two different terms, SFB (Ushijima et al., 2003) and SLF (Entani et al., 2003), we use SFB in this article because it distinguishes the Prunus system from other systems, and recent studies have all used SFB (Newbigin et al., 2008; Sassa et al., 2010; Tao and Iezzoni, 2010). The S locus F-box brothers (SFBBs) have been identified as candidates for the pollen S determinant in the subtribe Pyrinae of the Rosaceae (De Franceschi et al., 2011a, 2011b; Kakui et al., 2011; Minamikawa et al., 2010; Okada et al., 2008, 2011; Sassa et al., 2007). Variants of the S locus defined by combinations of
S-RNase and pollen S F-box gene(s) alleles are called S haplotypes.

In the S-RNase-based GSI system, pollen with an S haplotype that matches either S haplotype of the diploid pistil is recognized as self and its tube growth is arrested in the pistil because of the cytotoxicity of S-RNase. Although how this cytotoxicity is exerted remains unclear, it has been suggested that the degradation of RNA in the pollen cytoplasm by self S-RNase results in the arrest of self-pollen tube growth, as indicated by the degradation of pollen rRNA in incompatible pollen tubes (McClure et al., 1990) and the essential role of ribonucleotide radiolysis of pollen rRNA in incompatible pollen tubes in the pollen cytoplasm by self S-RNase results in the clear, it has been suggested that the degradation of RNA although how this cytotoxicity is exerted remains unclear, it has been suggested that the degradation of RNA in the pollen cytoplasm by self S-RNase results in the arrest self pollen tube growth, as indicated by the degradation of pollen rRNA in incompatible pollen tubes (McClure et al., 1990) and the essential role of ribonucleotide radiolysis of pollen rRNA in incompatible pollen tubes in the pollen cytoplasm by self S-RNase results in the clear.

Two working models, the S-RNase degradation model and the S-RNase compartmentalization model, have been proposed for SI/SC reactions (McClure et al., 2011). The pollen determinant SFB/SLF is a member of the F-box proteins and contains an F-box domain at the N-terminus. F-box proteins typically form the SCF (SKP1/Cullin/F-box) E3 ubiquitin ligase complex, which regulates protein degradation in the ubiquitin/proteasome proteolytic pathway. In the degradation model, it is hypothesized that the SCFSLF E3 ligase specifically polyubiquitinates all non-self S-RNases. The polyubiquitinated non-self S-RNases are degraded by the 26S proteasome in the pollen tube, allowing normal pollen tube growth in compatible pollination. Recently, it was reported that at least three types of divergent SLF proteins, including the former SLF-like proteins, function as pollen determinants, each recognizing a subset of non-self S-RNases for polyubiquitination (Kubo et al., 2010). Since this finding, the S-RNase degradation model has been superseded by the collaborative non-self-recognition model, although the S-RNase degradation model and the collaborative non-self-recognition model are the same in that the pollen S determinant is involved in the degradation of cytotoxic non-self S-RNase.

The other model is the compartmentalization model, although this model does not necessarily contradict the S-RNase degradation and collaborative non-self-recognition models (Goldraij et al., 2006; McClure et al., 2011). In this model, S-RNases are taken up non-specifically into both self and non-self pollen tubes and then sequestered in the vacuole, and therefore do not exert cytotoxicity. In incompatible pollen tubes, the vacuole breaks down and S-RNases are released into the cytoplasm. HT-B, one of the modifiers of the SI reaction in Solanaceae, is presumed to be involved in this vacuole breakdown. Although genetic studies have indicated the presence of a pistil modifier in the Rosaceae (Fernández i Martí et al., 2009; Moriya et al., 2009), no genes orthologous to solanaceous modifiers, such as HT-B and 120K, have been identified in the Rosaceae.

Each model can partly explain some aspects of the GSI phenomenon but neither model fully explains the molecular mechanism of GSI, possibly because unidentified pollen and/or pistil factors play important roles in the SI/SC reaction. Furthermore, Tao and Iezzoni (2010) suggested the possible existence of a distinct molecular recognition mechanism in the GSI of Prunus based on the different SI/SC behaviors of pollen-part mutants and heterodiallelic pollen. We are still far from a full understanding of the molecular mechanism of S-RNase-based GSI. For further understanding of the molecular basis of GSI, it is essential to identify pollen- or pistil-part modifier genes.

Large-scale transcriptome analysis using next-generation sequencing (NGS) technologies has proven very useful to obtain candidates for such unidentified genes. NGS technologies are capable of generating high-throughput reads at a relatively low cost and have been used in various types of research, including genome sequencing, marker discovery, and especially transcriptome analysis (Ekblom and Galindo, 2011; Pareek et al., 2011; Strickler et al., 2012). Transcriptome analysis using NGS technologies can capture nearly all of the expressed sequences, including rare transcripts in a particular tissue at a specific developmental stage, due to the great depth of sequencing. Therefore, it is very useful, especially for transcriptome analysis in non-model organisms like Japanese apricot. In this study, we used this newly developed technology to dissect the complete molecular network underlying the GSI reaction to help identify pollen- or pistil-part modifier genes.

Although various NGS platforms, such as the Roche/454 Genome Sequencer FLX (GS FLX), the Illumina Genome Analyzer, and the ABI SOLiD, are available, the Roche/454 GS FLX is preferred for transcriptome analysis in non-model organisms (Table 2 in Ekblom and Galindo, 2011; Table 1 in Strickler et al., 2012) because of its longer read length and high accuracy. However, the number of reads generated per run by the Roche/454 GS FLX is much lower than the Illumina Genome Analyzer. The higher number of reads generated by the Illumina Genome Analyzer is preferable for more accurate gene expression measurement. In this study, we used both platforms, the Roche/454 GS FLX and the Illumina Genome Analyzer, to fully explore the advantages of each platform and conducted a large-scale transcriptome analysis of unpollinated, self-pollinated and cross-pollinated pistils of Japanese apricot to capture all of the molecular events induced by the GSI reaction in Japanese apricot, one of the fruit species in Prunus.

Materials and Methods

Plant materials and RNA extraction

Twelve-year-old Japanese apricot ‘Nanko’ (S’S’) and ‘Kairyo-Uchida-Ume’ (S’S’) trees grown at the experimental farm of Kyoto University were used in this study. During the 2010 flowering season, pistils and anthers were sampled from unopened flowers of ‘Nanko’ at the balloon stage of development. Anthers were also sam-
pleted from ‘Kairyo-Uchida-Ume’. The anthers were dehisced and the pollen grains were collected. Forty ‘Nanko’ pistils were placed on 1% (w/v) agar in each Petri dish (90 \times 20 \text{mm}) and incubated at 20°C overnight. After overnight incubation, three different pollination treatments were conducted for the pistils. Two subsets of the pistils were pollinated by ‘Nanko’ (self-pollination, SP) or ‘Kairyo-Uchida-Ume’ (cross-pollination, CP). One third of the pistils were kept unpollinated (UP). The SW and CP pistils were sampled at 3, 6, and 9 hours after pollination. The UP pistils were also sampled at the same times as the pollinated pistils. One hundred pistils, equivalent to the number of pistils incubated in two and half Petri dishes, were sampled at each time point after pollination and pooled for the SP, CP, and UP treatments. Thus, three hundred pistils were used for RNA extraction in each treatment. Two pollen samples, pollen grains and germinated pollen grains, of ‘Nanko’ and ‘Kairyo-Uchida-Ume’ were also used for RNA extraction. For the germinated pollen grain sample, pollen grains (100 mg) were germinated in 10 mL liquid pollen germination medium [50 g L^{-1} sucrose, 125 g L^{-1} PEG6000, 300 mg L^{-1} casein, 10 mg L^{-1} rifampicin, 100 mM CaCl_2, 100 mM KCl, 100 mM MgSO_4, 100 mM H_3BO_3 in 100 mM MES (pH 5.8)] at 18°C for three hours on a seesaw shaker at low speed (20 shakes min^{-1}). After three hours of incubation, the germinated pollen grains were sampled for RNA extraction. All collected pistils, pollen grains and germinated pollen grains were immediately frozen in liquid N2 and stored at −80°C until use. Total RNA was extracted by the cold phenol method as described previously (Tao et al., 1999).

Sequencing, sequence preprocessing, and assembly

The 454-pyrosequencing technique for 3’-ESTs using the GS FLX (Roche 454; Life Sciences, Branford, CT, USA) was performed with cDNA synthesized from mRNA prepared from SP pistils, CP pistils, UP pistils, pollen grains, and germinated pollen grains. Sample preparation and sequencing were conducted as per the custom service provided by TaKaRa BIO Inc. Dragon Genomics Center (Mie, Japan), as described previously (Habu et al., 2012). After sequencing, the obtained reads were processed with Seqclean software (http://compbio.dfci.harvard.edu/tgi/software/) on September 18, 2013) to trim low complexity sequences. Then, the reads were further processed using RepeatMasker (Smit et al., 1996–2010) (http://www.repeatmasker.org) with RepBase (Jurka et al., 2005) to mask repeat sequences to avoid misassembly. Finally, the masked reads were processed with a Perl script as follows: i) low quality regions were masked, ii) masked regions of both ends were trimmed, iii) reads that were shorter than 10 bases were removed, and iv) reads that contained more than 30% masked regions were removed.

Illumina Genome Analyzer IIx (GAIIX) (Illumina, San Diego, CA, USA) single-end sequencing was also performed for three samples, SP, CP, and UP pistils, as per the custom service provided by TaKaRa BIO Inc. Dragon Genomics Center. In short, sequencing libraries were constructed using the mRNA-Seq Sample Prep (Illumina) and Small RNA Sample Prep (Illumina) Kits following the Directional mRNA-Seq Library Prep Pre-Release Protocol (Illumina). Using the constructed libraries as templates, clonal clusters from molecule fragments were constructed on GAIIX flow cells by the Cluster Station (Illumina) and Single-Read Cluster Generation Kit v4 (Illumina). Sequencing was performed using the TruSeq SBS Kit v5 (Illumina). After sequencing, the obtained reads were processed with a Perl script as follows: i) low quality regions were masked, and ii) reads that contained more than 10% masked regions were removed.

The preprocessed reads obtained from 454-pyrosequencing and GAIIX sequencing were assembled by Trinity (Grabherr et al., 2011) with the default settings. In addition, the reads from unpollinated pistils obtained in this study (UP_454 and UP_GAIIx) and the reads from pollen (‘Nanko’ and ‘Kairyo-Uchida-Ume’) obtained from DRA (Acc. DRP000624) were individually preprocessed and assembled as described above to obtain unpollinated pistil and pollen transcriptomes.

Functional annotation

To annotate the unigenes, they were searched against the nonredundant protein database (nr) of the National Center for Biotechnology Information (NCBI) and the Arabidopsis protein database (TAIR10) of the Arabidopsis Information Resource (TAIR) by the BLASTX program (Altschul et al., 1990) using an E-value cutoff of 1e-6. Only best-hit results were extracted and hits against the TAIR database were used for functional gene ontology (GO) annotations (The Gene Ontology Consortium, 2000) using the TAIR Gene Ontology tool (http://www.arabidopsis.org/tools/bulk/go/index.jsp, September 13, 2013). Assigned GO terms were summarized based on “Plant GO Slim” provided by the TAIR website.

Digital Expression Analysis

Digital expression analysis was conducted using a number of reads mapped to each unigene. Three GAIIX read sets were individually mapped to the unigenes using Bowtie (Langmead et al., 2009). The obtained raw count data were used as inputs to edgeR (Robinson et al., 2010) for statistical analysis. Unigenes with less than a 1% false discovery rate (FDR) were taken as differentially expressed between samples. GO enrichment analyses of these differentially expressed unigenes were conducted by agriGO (Du et al., 2010).

Results

Sequencing, assembly and annotation

In total, 1,314,579 raw reads from seven libraries
were generated by 454-pyrosequencing and 86,127,694 raw reads from three libraries were generated by GAIIX (Table 1). All the obtained sequences from 454-pyrosequencing and GAIIX reads are available from the DDBJ Sequence Read Archive (http://trace.ddbj.nig.ac.jp/dra/index.html, September 13, 2013) under the following accession numbers (study: DRP01188, samples: DR5012257-DR5012268, experiments: DRX012514-DRX012518, and runs: DR013973-DR013977). After the preprocessing steps, we obtained 1,165,107 reads from 454-pyrosequencing and 76,356,203 reads from GAIIX (Table 1). We ran preliminary assemblies with the Trinity program using different data sets (seven 454 read sets, three GAIIX read sets, and all 10 read sets) and decided to use the unigenes generated from all 10 read sets for further analyses, because they included more and longer unigenes than those generated from the other two data sets (Table 2). This unigene data set contained 40,061 contigs (referred to as unigenes hereafter) ranging from 201 to 7,477 bases with an average length of 572 bases (Table 3). A BLASTN (1e-100) search against the transcriptome data set of dormant buds in Japanese apricot (Habu et al., 2012) revealed that 14,950 of the 40,061 unigenes were almost identical to sequences in the dormant bud transcriptome.

When the unigene sequences were queried against the NCBI non-redundant and Arabidopsis (TAIR10) protein databases using BLASTX (1e-6), 29,985 (74.8% of total unigenes) and 27,898 (69.6% of total unigenes) unigene sequences showed at least one hit against the NCBI nr and TAIR10 protein databases, respectively. The best hits from TAIR10 were used to assign gene ontology (GO) annotations. A total of 4,417 GO terms were assigned to 27,399 unique sequences, representing 68.4% of all unigene sequences and 98.2% of the BLAST-annotated sequences. GO Slim assignments for the unigene sequences and the complete Arabidopsis gene set within the three categories are summarized in Figure 1. The GO distribution indicated that genes with a wide range of functions were expressed in unpollinated and/or pollinated styles and the GO distribution trends in the P. mume pistil and pollen transcriptomes appeared to be similar to those of the complete Arabidopsis gene set.

**Digital expression analysis**

For digital expression analysis, the three GAIIX read sets were mapped to the unigenes by Bowtie (Table 4). In total, 22,932,076 CP_GAIIX reads, 20,154,057 SP_GAIIX reads, and 21,536,333 UP_GAIIX reads were mapped to the unigenes, with the average number of mapped reads per unigene being 503–572. For statistical analysis, we generated by 454-pyrosequencing and 86,127,694 raw reads from three libraries were generated by GAIIX (Table 1). All the obtained sequences from 454-pyrosequencing and GAIIX reads are available from the DDBJ Sequence Read Archive (http://trace.ddbj.nig.ac.jp/dra/index.html, September 13, 2013) under the following accession numbers (study: DRP01188, samples: DR5012257-DR5012268, experiments: DRX012514-DRX012518, and runs: DR013973-DR013977). After the preprocessing steps, we obtained 1,165,107 reads from 454-pyrosequencing and 76,356,203 reads from GAIIX (Table 1). We ran preliminary assemblies with the Trinity program using different data sets (seven 454 read sets, three GAIIX read sets, and all 10 read sets) and decided to use the unigenes generated from all 10 read sets for further analyses, because they included more and longer unigenes than those generated from the other two data sets (Table 2). This unigene data set contained 40,061 contigs (referred to as unigenes hereafter) ranging from 201 to 7,477 bases with an average length of 572 bases (Table 3). A BLASTN (1e-100) search against the transcriptome data set of dormant buds in Japanese apricot (Habu et al., 2012) revealed that 14,950 of the 40,061 unigenes were almost identical to sequences in the dormant bud transcriptome.

When the unigene sequences were queried against the NCBI non-redundant and Arabidopsis (TAIR10) protein databases using BLASTX (1e-6), 29,985 (74.8% of total unigenes) and 27,898 (69.6% of total unigenes) unigene sequences showed at least one hit against the NCBI nr and TAIR10 protein databases, respectively. The best hits from TAIR10 were used to assign gene ontology (GO) annotations. A total of 4,417 GO terms were assigned to 27,399 unique sequences, representing 68.4% of all unigene sequences and 98.2% of the BLAST-annotated sequences. GO Slim assignments for the unigene sequences and the complete Arabidopsis gene set within the three categories are summarized in Figure 1. The GO distribution indicated that genes with a wide range of functions were expressed in unpollinated and/or pollinated styles and the GO distribution trends in the P. mume pistil and pollen transcriptomes appeared to be similar to those of the complete Arabidopsis gene set.

**Digital expression analysis**

For digital expression analysis, the three GAIIX read sets were mapped to the unigenes by Bowtie (Table 4). In total, 22,932,076 CP_GAIIX reads, 20,154,057 SP_GAIIX reads, and 21,536,333 UP_GAIIX reads were mapped to the unigenes, with the average number of mapped reads per unigene being 503–572. For statistical analysis, we generated by 454-pyrosequencing and 86,127,694 raw reads from three libraries were generated by GAIIX (Table 1). All the obtained sequences from 454-pyrosequencing and GAIIX reads are available from the DDBJ Sequence Read Archive (http://trace.ddbj.nig.ac.jp/dra/index.html, September 13, 2013) under the following accession numbers (study: DRP01188, samples: DR5012257-DR5012268, experiments: DRX012514-DRX012518, and runs: DR013973-DR013977). After the preprocessing steps, we obtained 1,165,107 reads from 454-pyrosequencing and 76,356,203 reads from GAIIX (Table 1). We ran preliminary assemblies with the Trinity program using different data sets (seven 454 read sets, three GAIIX read sets, and all 10 read sets) and decided to use the unigenes generated from all 10 read sets for further analyses, because they included more and longer unigenes than those generated from the other two data sets (Table 2). This unigene data set contained 40,061 contigs (referred to as unigenes hereafter) ranging from 201 to 7,477 bases with an average length of 572 bases (Table 3). A BLASTN (1e-100) search against the transcriptome data set of dormant buds in Japanese apricot (Habu et al., 2012) revealed that 14,950 of the 40,061 unigenes were almost identical to sequences in the dormant bud transcriptome.

When the unigene sequences were queried against the NCBI non-redundant and Arabidopsis (TAIR10) protein databases using BLASTX (1e-6), 29,985 (74.8% of total unigenes) and 27,898 (69.6% of total unigenes) unigene sequences showed at least one hit against the NCBI nr and TAIR10 protein databases, respectively. The best hits from TAIR10 were used to assign gene ontology (GO) annotations. A total of 4,417 GO terms were assigned to 27,399 unique sequences, representing 68.4% of all unigene sequences and 98.2% of the BLAST-annotated sequences. GO Slim assignments for the unigene sequences and the complete Arabidopsis gene set within the three categories are summarized in Figure 1. The GO distribution indicated that genes with a wide range of functions were expressed in unpollinated and/or pollinated styles and the GO distribution trends in the P. mume pistil and pollen transcriptomes appeared to be similar to those of the complete Arabidopsis gene set.

**Digital expression analysis**

For digital expression analysis, the three GAIIX read sets were mapped to the unigenes by Bowtie (Table 4). In total, 22,932,076 CP_GAIIX reads, 20,154,057 SP_GAIIX reads, and 21,536,333 UP_GAIIX reads were mapped to the unigenes, with the average number of mapped reads per unigene being 503–572. For statistical

![](image)
down-regulated by both pollination types, the GO categories for responses to various stimuli were enriched (Table 5), while no enriched GO category in the biological process ontology was detected in the latter two subsets that were differentially regulated by self- and cross-pollinations.

Organ-specific up-regulated genes

The total mRNAs from pollinated pistils consisting of both pistil and pollen transcripts should contain transcripts from genes involved in the self-(in)compatible reaction such as self-incompatibility specificity determinant genes and modifier genes. As the expression of some of these genes could be changed by pollination, we first focused on the 2,227 unigenes whose expressions were increased by both pollination types. These unigenes were searched against the unpollinated pistil and pollen transcriptome sequences by MEGABLAST (cut-off e-value < 1e-100) to determine from which organs, pistils and/or pollen grains, the transcripts originated. Assemblies of the mixed reads of UP_454 + UP_GAIx (UP) and the mixed reads of ‘Nanko’ pollen + ‘Kairyo-Uchida-Ume’ pollen (NKP) (DRA Acc. DRP000624) yielded 32,736 and 38,797 contigs (unigenes), respectively (Table 6). Using these transcriptomes, we determined the origin of the unigenes up-regulated by both pollination types. MEGABLAST searching revealed that 1,564 and 1,695 unigenes up-regulated by both pollination types had almost identical sequences in the UP and NKP transcriptomes, respectively, indicating that the former were expressed in pistils and the latter in pollen grains. As shown in Figure 3, the expressions of 1,263 unigenes were predicted to be up-regulated commonly in

<table>
<thead>
<tr>
<th>Read set</th>
<th>Sum of mapped reads</th>
<th>Ave. no. of mapped reads per unigene</th>
</tr>
</thead>
<tbody>
<tr>
<td>CP_GAIx</td>
<td>22,932,076</td>
<td>572</td>
</tr>
<tr>
<td>SP_GAIx</td>
<td>20,154,057</td>
<td>503</td>
</tr>
<tr>
<td>UP_GAIx</td>
<td>21,536,333</td>
<td>538</td>
</tr>
</tbody>
</table>

Analyses, we used the R package edgeR. Differentially expressed unigenes were selected based on the FDR (cut-off FDR < 0.01). Comparisons of the number of mapped reads from the UP pistil with those of the CP or SP pistils showed that the expressions of 8,907 unigenes were significantly different between UP and CP pistils and those of 10,190 unigenes were significantly different between UP and SP pistils (Fig. 2). Among these, the changes in expression of 4,348 unigenes were significant in both pollination types, while the expressions of 4,559 and 5,842 unigenes were changed specifically in CP and SP, respectively. Among the CP-specific differentially expressed unigenes, the expressions of 2,373 unigenes were increased and those of 2,186 unigenes were decreased by cross-pollination. Among the SP-specific differentially expressed unigenes, the expressions of 2,691 unigenes were increased and those of 3,151 unigenes were decreased by self-pollination. All four subsets of genes showed no enriched GO categories in the biological process ontology. The expressions of 2,227 and 1,942 unigenes were increased and decreased commonly by both pollination types, respectively, while the expressions of 95 unigenes were increased by CP but decreased by SP and vice versa for 84 unigenes (Fig. 2). In the former two subsets that were commonly up- or
pollen and pistils or up-regulated at least either in pollen or pistils because they were found in both the UP- and NKP-specific unigenes, while the expressions of 432 and 301 unigenes were up-regulated specifically in pollen and pistils, respectively. The remaining 231 unigenes up-regulated by both pollination types had no homologous sequences in either of the transcriptomes.

Among the 301 pistil-specific unigenes up-regulated by both pollinations, 184 unigenes (61.1%) were annotated by TAIR10 and 199 (66.1%) were annotated by nr, including three S-RNase unigenes (Table 7). The expressions of the S-RNases were high in unpollinated pistils and were increased by both pollination types. Among 432 pollen-specific unigenes up-regulated by both pollination types, 243 unigenes (56.2%) were annotated by TAIR10 and 263 (60.9%) were annotated by nr. In contrast to the pistil determinant, S-RNase, the pollen determinant, SFB, was not included among these pollen-specific unigenes, although it was included in the combined unpollinated/pollinated pistil transcriptome. This is because the expression of SFB was not detected in all of the pistils (Table 8). Similarly, S locus F-box protein with low allelic sequence polymorphism 2, which is also called SLFL2 (Matsumoto et al., 2008), was not included among the differentially expressed unigenes but was found in the whole unpollinated/pollinated pistil transcriptome because its expression was barely detectable. Conversely, S locus F-box protein with low allelic sequence polymorphism 3 (SLFL3) (Matsumoto et al., 2008), which encodes one of the pollen-expressed F-box proteins located in the S locus but has very low allelic sequence polymorphism, was included among the pollen-specific unigenes up-regulated by both pollination types and its expression was especially increased by cross-pollination (Table 7). In addition, the pollen-specific unigenes up-regulated by both pollination types included cullin 1, a ubiquitin-like protein and ubiquitin carboxyl-terminal hydrolase16-like protein (Table 9). Among the 1,263 unigenes up-regulated in either or both organs by both pollination types, 1,079 unigenes (85.4%) were annotated by TAIR10 and 1,129 (89.4%) were annotated by nr. These unigenes included S locus F-box protein with low allelic sequence polymorphism 1 (SLFL1) (Matsumoto et al., 2008) (Table 7), the subunits of SCF complex, ASK2 (SKP1), and two types of cullins, three types of E2 ubiquitin conjugating enzymes, E3 ubiquitin-protein ligase1 (UPL1), and four types of ubiquitin or ubiquitin-like proteins (Table 9).

Searching for modifier genes in Prunus

Since the pollen-part modifier gene of Prunus (M locus) has been recently mapped to the distal part of Chr. 3 flanked by two SSR markers within an interval of 1.8 cM, corresponding to ~364 Kb in the peach genome (Zuriaga et al., 2012), we searched for differentially expressed unigenes located at this locus by MEGABLAST searching against the peach genome (The International Peach Genome Initiative, 2013). We found 19 unigenes...
Therefore, DUBs were searched against the SP- and CP-specific up-regulated unigenes; we found six and 12 deubiquitinating enzyme-like unigenes, respectively (Table 11). Among them, only UBP11-like unigene expressed specifically in pollen was found among the SP-specific up-regulated DUB-like unigenes.

Discussion

NGS technology for SI/SC study in Prunus

The application of NGS technologies to transcriptome analysis allows the characterization of nearly all of the transcripts in a particular tissue during a specific biological event. To date, the Roche/454 GS FLX system has been mainly used for transcriptome anal-

<table>
<thead>
<tr>
<th>GO ID</th>
<th>Term</th>
<th>FDR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Up-regulated Unigenes</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GO:0050896</td>
<td>response to stimulus</td>
<td>7.9e-04</td>
</tr>
<tr>
<td>GO:0010033</td>
<td>response to organic substance</td>
<td>0.0038</td>
</tr>
<tr>
<td>GO:0042221</td>
<td>response to chemical stimulus</td>
<td>0.0038</td>
</tr>
<tr>
<td>GO:0009611</td>
<td>response to wounding</td>
<td>0.0038</td>
</tr>
<tr>
<td>GO:0044281</td>
<td>small molecule metabolic process</td>
<td>0.026</td>
</tr>
<tr>
<td>GO:009605</td>
<td>response to external stimulus</td>
<td>0.043</td>
</tr>
<tr>
<td>Down-regulated Unigenes</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GO:0006412</td>
<td>translation</td>
<td>1.7e-12</td>
</tr>
<tr>
<td>GO:009628</td>
<td>response to abiotic stimulus</td>
<td>3.3e-07</td>
</tr>
<tr>
<td>GO:034645</td>
<td>cellular macromolecule biosynthetic process</td>
<td>1.2e-06</td>
</tr>
<tr>
<td>GO:0010467</td>
<td>gene expression</td>
<td>1.2e-06</td>
</tr>
<tr>
<td>GO:009059</td>
<td>macromolecule biosynthetic process</td>
<td>1.9e-06</td>
</tr>
<tr>
<td>GO:042254</td>
<td>ribosome biogenesis</td>
<td>1.3e-04</td>
</tr>
<tr>
<td>GO:009719</td>
<td>response to endogenous stimulus</td>
<td>1.3e-04</td>
</tr>
<tr>
<td>GO:0022613</td>
<td>ribonucleoprotein complex biogenesis</td>
<td>1.3e-04</td>
</tr>
<tr>
<td>GO:009725</td>
<td>response to hormone stimulus</td>
<td>1.4e-04</td>
</tr>
<tr>
<td>GO:042221</td>
<td>response to chemical stimulus</td>
<td>2.3e-04</td>
</tr>
<tr>
<td>GO:0006970</td>
<td>response to osmotic stress</td>
<td>7.9e-04</td>
</tr>
<tr>
<td>GO:009415</td>
<td>response to water</td>
<td>8.4e-04</td>
</tr>
<tr>
<td>GO:050896</td>
<td>response to stimulus</td>
<td>8.4e-04</td>
</tr>
<tr>
<td>GO:009414</td>
<td>response to water deprivation</td>
<td>0.0012</td>
</tr>
<tr>
<td>GO:009651</td>
<td>response to salt stress</td>
<td>0.0012</td>
</tr>
<tr>
<td>GO:001033</td>
<td>response to organic substance</td>
<td>0.0015</td>
</tr>
<tr>
<td>GO:044249</td>
<td>cellular biosynthetic process</td>
<td>0.0017</td>
</tr>
<tr>
<td>GO:044267</td>
<td>cellular protein metabolic process</td>
<td>0.0017</td>
</tr>
<tr>
<td>GO:0006950</td>
<td>response to stress</td>
<td>0.0017</td>
</tr>
<tr>
<td>GO:009058</td>
<td>biosynthetic process</td>
<td>0.0029</td>
</tr>
<tr>
<td>GO:009266</td>
<td>response to temperature stimulus</td>
<td>0.0063</td>
</tr>
<tr>
<td>GO:009737</td>
<td>response to abscisic acid stimulus</td>
<td>0.016</td>
</tr>
<tr>
<td>GO:019538</td>
<td>protein metabolic process</td>
<td>0.026</td>
</tr>
<tr>
<td>GO:009409</td>
<td>response to cold</td>
<td>0.035</td>
</tr>
<tr>
<td>GO:0044260</td>
<td>cellular macromolecule metabolic process</td>
<td>0.041</td>
</tr>
</tbody>
</table>

Table 6. Results of assemblies of un-pollinated pistil and pollen data sets.

<table>
<thead>
<tr>
<th>Assembled reads</th>
<th>Total no. of processed reads</th>
<th>No. of unigenes</th>
<th>Max unigene length</th>
<th>Minimum unigene length</th>
<th>Ave. length of unigene</th>
</tr>
</thead>
<tbody>
<tr>
<td>UP_454 + UP_GAIIX</td>
<td>25,249,387</td>
<td>32,736</td>
<td>7,478</td>
<td>201</td>
<td>494</td>
</tr>
<tr>
<td>NKPz</td>
<td>56,465,821</td>
<td>38,797</td>
<td>7,698</td>
<td>201</td>
<td>808</td>
</tr>
</tbody>
</table>

*Nanko* and *Kairyo-Uchida* pollen GAIIX reads.

As described above, we found that the ubiquitin carboxyl-terminal hydrolase16-like protein, a member of the deubiquitinating enzymes (DUB), was significantly up-regulated by both pollination types, suggesting that deubiquitination may be related to the SI reaction. Therefore, DUBs were searched against the SP- and CP-specific up-regulated unigenes; we found six and 12 deubiquitinating enzyme-like unigenes, respectively (Table 11). Among them, only UBP11-like unigene expressed specifically in pollen was found among the SP-specific up-regulated DUB-like unigenes.

Discussion

NGS technology for SI/SC study in Prunus

The application of NGS technologies to transcriptome analysis allows the characterization of nearly all of the transcripts in a particular tissue during a specific biological event. To date, the Roche/454 GS FLX system has been mainly used for transcriptome anal-
Therefore, Illumina platforms have recently been chosen routinely for transcriptome analysis in non-model organisms and model organisms (Fu et al., 2013; Liu et al., 2013; Oono et al., 2013). This study also confirms that Illumina reads alone can be used for transcriptome analysis of non-model organisms. In this study, we conducted a large-scale transcriptome analysis of pistils from Japanese apricot, a non-model organism, to gain insight into the molecular network of the SI/SC reaction. We obtained 40,061 unigenes from 77,521,310 reads of pollinated and unpollinated pistils and pollen grains. Among these unigenes, 29,985 had homologous genes in GenBank, while the other 10,076 appeared to be novel. We previously obtained 113,629 unigenes, including contigs and singletons, from the transcriptome of Japanese apricot dormant buds (Habu et al., 2012). Among the 40,061 unigenes obtained in this study, only 14,591 unigenes (37.2%) had at least one closest hit against the unigenes from dormant buds. The low percentage of common unigenes generated could be attributed to differences in the organs used and differences in the developmental stages of the organs. In GO analyses, none of the characteristic GO categories were found in the overall GO distribution but several GO terms involved in responses to stimuli were over-represented in the GO enrichment analysis of unigenes up-regulated by both pollination types, indicating that the stimulus of pollination triggers the up-regulation of many genes.

**Pistil modifiers**

As the pistil modifiers for S haplotype-specific pollen...
### Table 7. Differentially expressed unigenes located at the S locus.

<table>
<thead>
<tr>
<th>Unigene ID</th>
<th>Mapped reads (reads/unigene)</th>
<th>UP to CP</th>
<th>UP to SP</th>
<th>NCBI nr annotation</th>
<th>organ</th>
<th>e-value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>log2 (FC)</td>
<td>FDR</td>
<td>log2 (FC)</td>
<td>FDR</td>
<td></td>
</tr>
<tr>
<td>comp22_c0_seq1</td>
<td>25,536</td>
<td>62,669</td>
<td>52,365</td>
<td>1.123</td>
<td>0.937</td>
<td>BAC56115.1</td>
</tr>
<tr>
<td>comp22_c0_seq2</td>
<td>24,700</td>
<td>48,925</td>
<td>40,861</td>
<td>0.814</td>
<td>0.628</td>
<td>BAC56116.1</td>
</tr>
<tr>
<td>comp22_c0_seq3</td>
<td>25,288</td>
<td>49,792</td>
<td>42,506</td>
<td>0.805</td>
<td>0.651</td>
<td>BAC56116.1</td>
</tr>
<tr>
<td>comp1704_c0_seq1</td>
<td>1,311</td>
<td>1,824</td>
<td>1,612</td>
<td>0.304</td>
<td>7.8e-08</td>
<td>BAF36716.1</td>
</tr>
<tr>
<td>comp10070_c0_seq1</td>
<td>19</td>
<td>247</td>
<td>74</td>
<td>3.528</td>
<td>6.2e-44</td>
<td>BAF36718.1</td>
</tr>
</tbody>
</table>

### Table 8. Other F-box unigenes located at the S locus.

<table>
<thead>
<tr>
<th>Unigene ID</th>
<th>Mapped reads (reads/unigene)</th>
<th>UP to CP</th>
<th>UP to SP</th>
<th>NCBI nr annotation</th>
<th>organ</th>
<th>e-value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>log2 (FC)</td>
<td>FDR</td>
<td>log2 (FC)</td>
<td>FDR</td>
<td></td>
</tr>
<tr>
<td>comp38720_c0_seq1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1.172</td>
<td>1</td>
</tr>
</tbody>
</table>

### Table 9. Ubiquitination-related unigenes up-regulated by pollination.

<table>
<thead>
<tr>
<th>Unigene ID</th>
<th>Mapped reads (reads/unigene)</th>
<th>UP to CP</th>
<th>UP to SP</th>
<th>AGI</th>
<th>TAIR10 annotation</th>
<th>organ</th>
<th>e-value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>log2 (FC)</td>
<td>FDR</td>
<td>log2 (FC)</td>
<td>FDR</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Subunits of SCF complex</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>comp2390_c0_seq1</td>
<td>58</td>
<td>2356</td>
<td>353</td>
<td>5.172</td>
<td>0</td>
<td>2.507</td>
<td>2.1e-46</td>
</tr>
<tr>
<td>comp7227_c0_seq1</td>
<td>96</td>
<td>202</td>
<td>198</td>
<td>0.901</td>
<td>2.7e-06</td>
<td>0.946</td>
<td>1.1e-06</td>
</tr>
<tr>
<td>comp4940_c0_seq1</td>
<td>500</td>
<td>802</td>
<td>792</td>
<td>0.510</td>
<td>6.4e-09</td>
<td>0.565</td>
<td>1.2e-10</td>
</tr>
<tr>
<td>comp9095_c0_seq1</td>
<td>196</td>
<td>336</td>
<td>289</td>
<td>0.606</td>
<td>2.8e-05</td>
<td>0.462</td>
<td>2.7e-03</td>
</tr>
<tr>
<td>Ubiquitin or Ubiquitin-like protein</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>comp2840_c0_seq1</td>
<td>533</td>
<td>996</td>
<td>1,048</td>
<td>0.730</td>
<td>2.3e-20</td>
<td>0.877</td>
<td>8.2e-30</td>
</tr>
<tr>
<td>comp18859_c0_seq1</td>
<td>7</td>
<td>82</td>
<td>55</td>
<td>3.378</td>
<td>3.5e-14</td>
<td>2.875</td>
<td>3.2e-08</td>
</tr>
<tr>
<td>comp438_c0_seq1</td>
<td>1,495</td>
<td>2,247</td>
<td>2,200</td>
<td>0.416</td>
<td>9.6e-17</td>
<td>0.459</td>
<td>6.7e-20</td>
</tr>
<tr>
<td>comp904_c0_seq1</td>
<td>1,337</td>
<td>1,996</td>
<td>2,074</td>
<td>0.406</td>
<td>2.7e-14</td>
<td>0.535</td>
<td>9.2e-25</td>
</tr>
<tr>
<td>comp747_c0_seq1</td>
<td>3,006</td>
<td>3,800</td>
<td>3,476</td>
<td>0.166</td>
<td>2.3e-05</td>
<td>0.111</td>
<td>8.7e-03</td>
</tr>
<tr>
<td>comp2138_c0_seq1</td>
<td>609</td>
<td>899</td>
<td>897</td>
<td>0.390</td>
<td>2.6e-06</td>
<td>0.460</td>
<td>2.0e-08</td>
</tr>
<tr>
<td>E2 ubiquitin-conjugating enzyme</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>comp2792_c0_seq1</td>
<td>51</td>
<td>1,880</td>
<td>1,017</td>
<td>5.032</td>
<td>2.4e-21</td>
<td>4.219</td>
<td>2.4e-17</td>
</tr>
<tr>
<td>comp1555_c0_seq1</td>
<td>3,496</td>
<td>8,922</td>
<td>7,853</td>
<td>1.180</td>
<td>0</td>
<td>1.069</td>
<td>0</td>
</tr>
<tr>
<td>comp9702_c0_seq1</td>
<td>180</td>
<td>281</td>
<td>291</td>
<td>0.471</td>
<td>3.4e-03</td>
<td>0.594</td>
<td>1.1e-04</td>
</tr>
<tr>
<td>comp9059_c0_seq1</td>
<td>132</td>
<td>239</td>
<td>234</td>
<td>0.684</td>
<td>9.0e-05</td>
<td>0.727</td>
<td>2.8e-05</td>
</tr>
<tr>
<td>E3 ubiquitin ligase</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>comp3657_c0_seq1</td>
<td>869</td>
<td>1,382</td>
<td>1,309</td>
<td>0.497</td>
<td>2.2e-14</td>
<td>0.492</td>
<td>1.4e-13</td>
</tr>
<tr>
<td>comp27124_c0_seq1</td>
<td>2</td>
<td>29</td>
<td>25</td>
<td>3.101</td>
<td>4.2e-05</td>
<td>2.960</td>
<td>3.5e-04</td>
</tr>
<tr>
<td>Deubiquitinating enzyme</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>comp16343_c0_seq1</td>
<td>17</td>
<td>48</td>
<td>44</td>
<td>1.325</td>
<td>3.3e-03</td>
<td>1.273</td>
<td>7.8e-03</td>
</tr>
</tbody>
</table>
Table 10. SP- or CP-specific up-regulated unigenes located around the Ma locus.

<table>
<thead>
<tr>
<th>Unigene ID</th>
<th>Homologous gene</th>
<th>locus</th>
<th>Mapped reads (reads/unigene)</th>
<th>UP to CP</th>
<th>UP to SP</th>
<th>TAIR10 annotation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>PGS3_71</td>
<td>comp17470_c0_seq1 ppa000048m</td>
<td>18.423 18.435</td>
<td>45 90</td>
<td>78</td>
<td>0.828</td>
<td>0.008</td>
<td>0.695</td>
</tr>
<tr>
<td></td>
<td>comp32285_c0_seq1 ppa000048m</td>
<td>18.423 18.435</td>
<td>5</td>
<td>28</td>
<td>0</td>
<td>2.313</td>
<td>0.002</td>
</tr>
<tr>
<td></td>
<td>comp3643_c0_seq2 ppa005464m</td>
<td>18.436 18.439</td>
<td>577 744</td>
<td>750</td>
<td>0.195</td>
<td>0.052</td>
<td>0.280</td>
</tr>
<tr>
<td></td>
<td>comp20561_c0_seq1 ppa021437m</td>
<td>18.455 18.457</td>
<td>43 97</td>
<td>84</td>
<td>1.002</td>
<td>7.5e-04</td>
<td>0.867</td>
</tr>
<tr>
<td></td>
<td>comp15915_c0_seq1 ppa000002m</td>
<td>18.524 18.545</td>
<td>34 49</td>
<td>91</td>
<td>0.355</td>
<td>0.504</td>
<td>1.322</td>
</tr>
<tr>
<td></td>
<td>comp19579_c0_seq1 ppa000002m</td>
<td>18.524 18.545</td>
<td>25 45</td>
<td>73</td>
<td>0.676</td>
<td>0.168</td>
<td>1.447</td>
</tr>
<tr>
<td></td>
<td>comp5042_c0_seq1 ppa000002m</td>
<td>18.524 18.545</td>
<td>732 978</td>
<td>842</td>
<td>0.246</td>
<td>0.003</td>
<td>0.103</td>
</tr>
<tr>
<td></td>
<td>comp11745_c0_seq1 ppa011450m</td>
<td>18.575 18.576</td>
<td>45 97</td>
<td>111</td>
<td>0.936</td>
<td>0.002</td>
<td>1.204</td>
</tr>
<tr>
<td></td>
<td>comp10245_c0_seq1 ppa001620m</td>
<td>18.582 18.589</td>
<td>108 136</td>
<td>175</td>
<td>0.161</td>
<td>0.626</td>
<td>0.598</td>
</tr>
<tr>
<td></td>
<td>comp23062_c0_seq1 ppa017665m</td>
<td>18.595 18.596</td>
<td>6 53</td>
<td>1</td>
<td>2.971</td>
<td>3.4e-08</td>
<td>-2.684</td>
</tr>
<tr>
<td></td>
<td>comp29549_c0_seq1 ppa017665m</td>
<td>18.595 18.596</td>
<td>7 69</td>
<td>1</td>
<td>3.129</td>
<td>3.3e-11</td>
<td>-2.906</td>
</tr>
<tr>
<td></td>
<td>comp15679_c0_seq1 ppa026503m</td>
<td>18.630 18.630</td>
<td>33 115</td>
<td>97</td>
<td>1.629</td>
<td>1.5e-08</td>
<td>1.457</td>
</tr>
<tr>
<td></td>
<td>comp37199_c0_seq1 ppa000002m</td>
<td>18.645 18.647</td>
<td>43</td>
<td>97</td>
<td>84</td>
<td>0.381</td>
<td>1.0e-04</td>
</tr>
<tr>
<td></td>
<td>comp15212_c0_seq1 ppa007756m</td>
<td>18.645 18.648</td>
<td>37</td>
<td>122</td>
<td>32</td>
<td>1.549</td>
<td>1.3e-08</td>
</tr>
</tbody>
</table>

Table 11. Deubiquitination enzymes (DUBs) among the SP- or CP-specific up-regulated unigenes.

<table>
<thead>
<tr>
<th>Unigene ID</th>
<th>Homologous gene</th>
<th>locus</th>
<th>Mapped reads (reads/unigene)</th>
<th>UP to CP</th>
<th>UP to SP</th>
<th>TAIR10 annotation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>comp23062_c0_seq1 ppa017665m</td>
<td>18.595 18.596</td>
<td>6 53</td>
<td>1</td>
<td>2.971</td>
<td>3.4e-08</td>
<td>-2.684</td>
</tr>
<tr>
<td></td>
<td>comp29549_c0_seq1 ppa017665m</td>
<td>18.595 18.596</td>
<td>7 69</td>
<td>1</td>
<td>3.129</td>
<td>3.3e-11</td>
<td>-2.906</td>
</tr>
<tr>
<td></td>
<td>comp13199_c0_seq1</td>
<td>78</td>
<td>57</td>
<td>168</td>
<td>0.134</td>
<td>0.809</td>
<td>0.888</td>
</tr>
<tr>
<td></td>
<td>comp12545_c0_seq1</td>
<td>140</td>
<td>186</td>
<td>286</td>
<td>0.238</td>
<td>0.289</td>
<td>0.932</td>
</tr>
<tr>
<td></td>
<td>comp7801_c0_seq1</td>
<td>310</td>
<td>354</td>
<td>445</td>
<td>0.019</td>
<td>0.975</td>
<td>0.423</td>
</tr>
<tr>
<td></td>
<td>comp1343_c0_seq1</td>
<td>1,323</td>
<td>1,796</td>
<td>1,388</td>
<td>0.269</td>
<td>3.1e-06</td>
<td>-</td>
</tr>
</tbody>
</table>

T. Habu and R. Tao
rejection in the Solanaceae, the expression of HT-B and 120K in the pistil is indispensable (McClure et al., 1999; Hancock et al., 2005). In the compartmentalization model developed based on microscopic observations of pollen tube growth in Nicotiana, the pollen modifiers HT-B and 120K are thought to play a key role in vacuole breakdown in the pollen tube after SI pollination, which is supposed to release non-self S-RNase cytotoxicity (Goldraij et al., 2006; McClure et al., 2011). If the Solanaceae and Prunus share the same molecules in SI/SC recognition mechanisms, homologous sequences to HT-B and 120K should be present in the pistil transcriptome but no such sequences were found in the combined unpollinated/pollinated pistil transcriptome of Japanese apricot. It is possible that other unidentified distinct molecules are involved in the compartmentalization of S-RNases in Prunus or alternatively, the current compartmentalization model developed based on microscopic observations in the Solanaceae is not readily applicable to the Prunus GSI.

**Pollen modifiers**

Mutation in the pollen-part non-S factor, called the pollen modifier, appeared to confer SC in sweet cherry (*Prunus avium*) (Wünsch and Hormaza, 2004) and apricot (*Prunus armeniaca*) (Vilanova et al., 2006). As mutation of this pollen-part modifier gene results in self-compatibility, this modifier gene is considered to play a key role in the SI reaction and could be up-regulated upon SI pollination. Among the 19 unigenes upregulated by either or both pollination types at the M locus, only comp15915_c0_seq1 was pollen-specific and up-regulated by self-pollination (Table 10). Although this unigene was annotated as the auxin transporter BIG by TAIR10, it is thought that BIG could function as an UBR4. comp15915_c0_seq1, one of the ubiquitin-protein E3 ligase component N-recognins (Graciet and Wellmer, 2010). In fact, the putative amino acid sequence of the homologue of comp15915_c0_seq1 in peach, ppm000002m, contains the UBR4 E3 ubiquitin-protein ligase domain. Therefore, the gene from which comp15915_c0_seq1 is derived probably functions as a UBR4. Although the pollen S determinant is an F-box protein in the S-RNase-based GSI system and is thought to be involved in ubiquitination, comp15915_c0_seq1 is unlikely to interact with the pollen S F-box proteins, because the ubiquitin-protein E3 ligase component N-recognins do not interact with F-box proteins and recognize the N-terminus of substrates directly. It will be interesting to test whether comp15915_c0_seq1 is involved in the GSI reaction in *Prunus*.

**General inhibitor in Prunus SI**

A series of molecular and genetic analyses of the S-RNase-based SI mechanism in *Prunus* indicated the presence of a distinct recognition mechanism (Tao and Iezzoni, 2010). The main difference between the GSI mechanism in *Prunus* and those in Solanaceae and Plantaginaceae is considered to be the role of the pollen S determinant. It has been proposed that multiple pollen-specific F-box genes at the S locus collaboratively function to detoxify all but self S-RNase in the Solanaceae (Kubo et al., 2010). Mutations that disrupt pollen S (SLF) function in the Solanaceae and Plantaginaceae have yet to be found and therefore are thought to confer either SI or lethality. However, mutations that disrupt pollen S (SFB) function found in *Prunus* result in SC (Tao and Iezzoni, 2010; Ushijima et al., 2004; Yamane and Tao, 2009), indicating that the *Prunus* pollen S determinant may act to prohibit unknown mechanisms that inactivate the cytotoxic effects of the S-RNase. In other words, it is thought that multiple pollen S determinants (SLFs) would collaboratively detoxify the non-self S-RNases in the Solanaceae, while a single pollen S determinant (SFB) would release the cytotoxicity of the self S-RNase in *Prunus*. The role of the pollen S determinant of Pyrinae (SFBBs) is still unclear but it is suggested that the GSI mechanism of Pyrinae would be similar to that of Solanaceae (De Franceschi et al., 2012). On the other hand, three other types of pollen-expressed F-box genes, *SLFL1*, *SLFL2*, and *SLFL3* (synonymous with the S locus F-box protein with low allelic sequence polymorphism 1, 2, and 3, respectively), were also found on the S locus of *Prunus* (Entani et al., 2003; Matsumoto et al., 2008; Ushijima et al., 2003, 2004). These *Prunus* *SLFLs* are thought to be putative orthologs of the Pyrinae SFBBs based on phylogenetic analysis (Matsumoto et al., 2008). Although the distinct S-RNase detoxification mechanism in *Prunus* is still unknown, one possible explanation could be that *Prunus* *SLFLs* have lost their self-recognition function and have begun to function to detoxify all S-RNases including self S-RNase as the general inhibitor (GI) of the S-RNase. In this study, we found *SLFL1*, *SLFL2*, and *SLFL3* in the pistil transcriptome. Among these, *SLFL3* showed transcriptional pattern changes appropriate for GI. Namely, the expression of *SLFL3* was pollen-specific and highly up-regulated by pollination. It is interesting that the comp2390_c0_seq1 annotated as *ASK2* by TAIR10 is almost identical to Pav*SK1*, which is a gene for the SKP1-like protein interacting with SLFLs and SBF to form the SCF complex with Cul1-like (CUL1s) (Matsumoto et al., 2012), and was also up-regulated upon pollination, especially after compatible pollination (Table 9). We also found ubiquitin carboxyl-terminal hydrolase16-like protein, which is a member of the deubiquitinating enzymes (DUBs), among the pollen-specific unigenes up-regulated by pollination. Furthermore, an ubiquitin-specific protease, UBP11, was also found as an SP-specific up-regulated unigene expressed in pollen (Table 11). Although we have yet to clarify the SI/SC recognition mechanism of the S-RNase-based GSI in *Prunus*, it is likely that the ubiquitin/proteasome proteolytic pathway is involved in the system. Deubiquitination processes may also be
involved in the S haplotype-dependent protection and degradation of S-RNase.

In conclusion, our large scale transcriptome analyses support the presence of a distinct molecular mechanism for S-RNase-based GSI in Prunus. Further detailed studies of the genes identified in this study should lead to elucidation of the *Prunus* GSI mechanism.

Important notes: During the preparation of this paper, an article describing the whole genome sequencing of *P. mume* was published (Zhang et al., 2012). In the future, whole genome sequence data incorporated into developing bioinformatics programs will strongly promote omics studies in Japanese apricot and provide us with new insights for understanding agronomically important traits of this species.

**Literature Cited**


Lai, Z., W. Ma, B. Han, L. Liang, Y. Zhang, G. Hong and X. Yue. 2002. An F-box gene linked to the self-incompatibility (S) locus of *Antirrhinum* is expressed specifically in pollen and tapetum. Plant Mol. Biol. 50: 29–42.


