J. Japan. Soc. Hort. Sci. 83 (2): 95–107. 2014
doi: 10.2503/jjshs1.CH-086

The Japanese Society for Horticultural Science
http://www.jshs.jp/

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

Tsuyoshi Habu[1] and Ryutaro Tao[2]*

[1]*Faculty of Agriculture, Ehime University, Matsuyama 790-8566, Japan*
[2]*Graduate 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 at 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

Received; August 5, 2013. Accepted; October 4, 2013.
First Published Online in J-STAGE on December 7, 2013.
Part of this study was presented at the 2012 Spring Meeting of the Japanese Society for Horticultural Science. This work was supported by Grants-in-Aid (nos. 20248004 and 24248007) for Scientific Research (A) from the Japan Society for the Promotion of Science (JSPS) to R.T.
* Corresponding author (E-mail: rtao@kais.kyoto-u.ac.jp).

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 ribonuclease activity for the incompatible reaction (Huang et al., 1994). It is necessary for normal pollen tube growth in compatible pollination to evade S-RNase cytotoxicity.

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 SCF^{SLF} 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^1S^7) and 'Kairyo-Uchida-Ume' (S^3S^4) 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-

pled 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 × 20 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 SP 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 $g \cdot L^{-1}$ sucrose, 125 $g \cdot L^{-1}$ PEG6000, 300 $mg \cdot L^{-1}$ casein, 10 $mg \cdot L^{-1}$ 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$\cdot min^{-1}$). 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 N_2 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/, 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: DRS012257-DRS012268, experiments: DRX012514-DRX012518, and runs: DRR013973-DRR013977). 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 three 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

Table 1. Samples and the number of reads obtained from NGS platforms.

Read set	Cultivar	Sampled organs	Pollen donor	Platform	No. of raw reads	No. of processed reads
CP_454	'Nanko'	cross-pollinated pistils	'Kairyo-Uchida-Ume'	GS FLX Titanium	231,107	212,240
SP_454	'Nanko'	self-pollinated pistils	'Nanko'	GS FLX Titanium	280,823	252,608
UP_454	'Nanko'	unpollinated pistils	—	GS FLX Titanium	205,649	186,404
NP_454	'Nanko'	pollen grains	—	GS FLX Titanium	157,615	137,856
NGP_454	'Nanko'	germinated pollen grains	—	GS FLX Titanium	141,884	121,256
KP_454	'Kairyo-Uchida-Ume'	pollen grains	—	GS FLX Titanium	133,537	114,256
KGP_454	'Kairyo-Uchida-Ume'	germinated pollen grains	—	GS FLX Titanium	163,964	140,487
CP_GAIIx	'Nanko'	cross-pollinated pistils	'Kairyo-Uchida-Ume'	GAIIx	30,987,870	27,302,728
SP_GAIIx	'Nanko'	self-pollinated pistils	'Nanko'	GAIIx	27,505,143	23,990,492
UP_GAIIx	'Nanko'	unpollinated pistils	—	GAIIx	27,634,681	25,062,983

Table 2. Results of assembly by Trinity.

Assembled reads	Total no. of processed reads	No. of unigenes	Max unigene length	Minimum unigene length	Ave. length of unigene
Seven 454 samples[z]	1,165,107	12,712	1,651	201	422
Three GAIIx samples[y]	76,356,203	39,993	7,477	201	540
All 10 samples	77,521,310	40,061	7,477	201	572

[z] CP_454 + SP_454 + UP_454 + NP_454 + NGP_454 + KP_454 + KGP_454.
[y] CP_GAIIx + SP_GAIIx + UP_GAIIx.

Table 3. Summary of the assembly and annotation of unigenes using all ten samples.

Number of unigenes	40,061
Average unigene length	572
Range of unigene length	201–7,477
Number of unigenes with at least 1 BLASTX hit against nr	29,985 (74.8%)
Number of unigenes with at least 1 BLASTX hit against TAIR10	27,898 (69.6%)

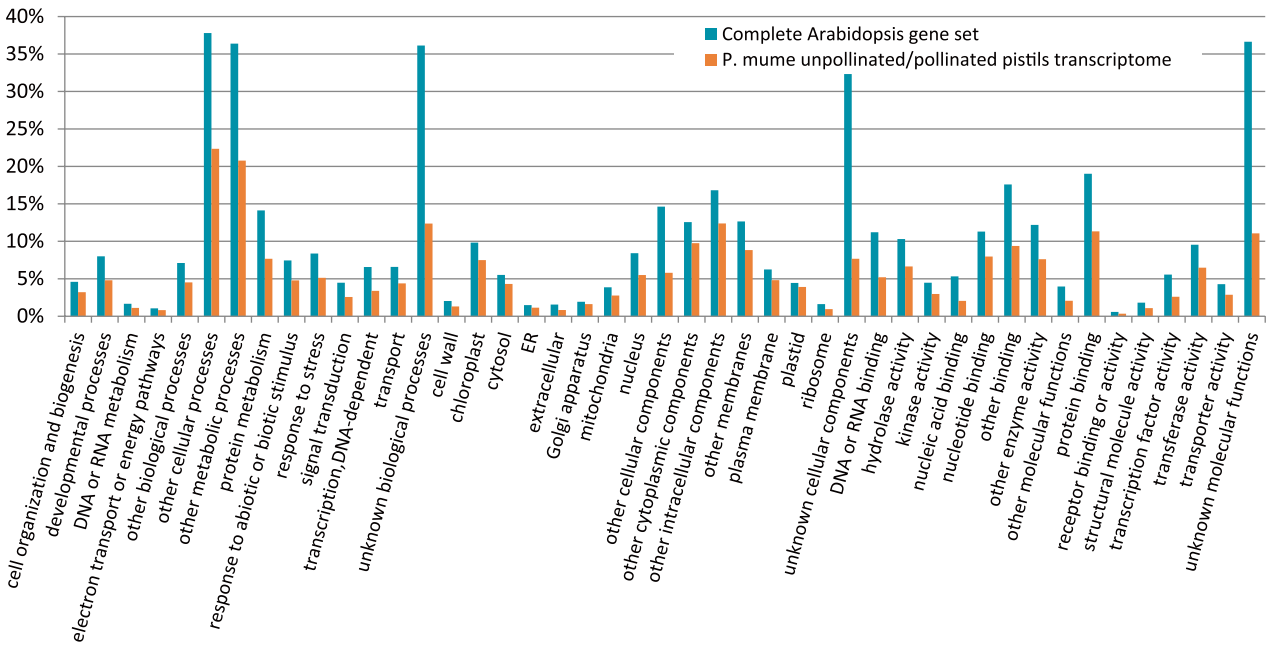

Fig. 1. Gene Ontology (GO) assignments for the unigenes. Proportions of the annotated unigenes and the complete Arabidopsis gene set that matched various gene ontology categories.

Table 4. Results of mapping to the unigenes.

Read set	Sum of mapped reads	Ave. no. of mapped reads per unigene
CP_GAIIx	22,932,076	572
SP_GAIIx	20,154,057	503
UP_GAIIx	21,536,333	538

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 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_GAIIx (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

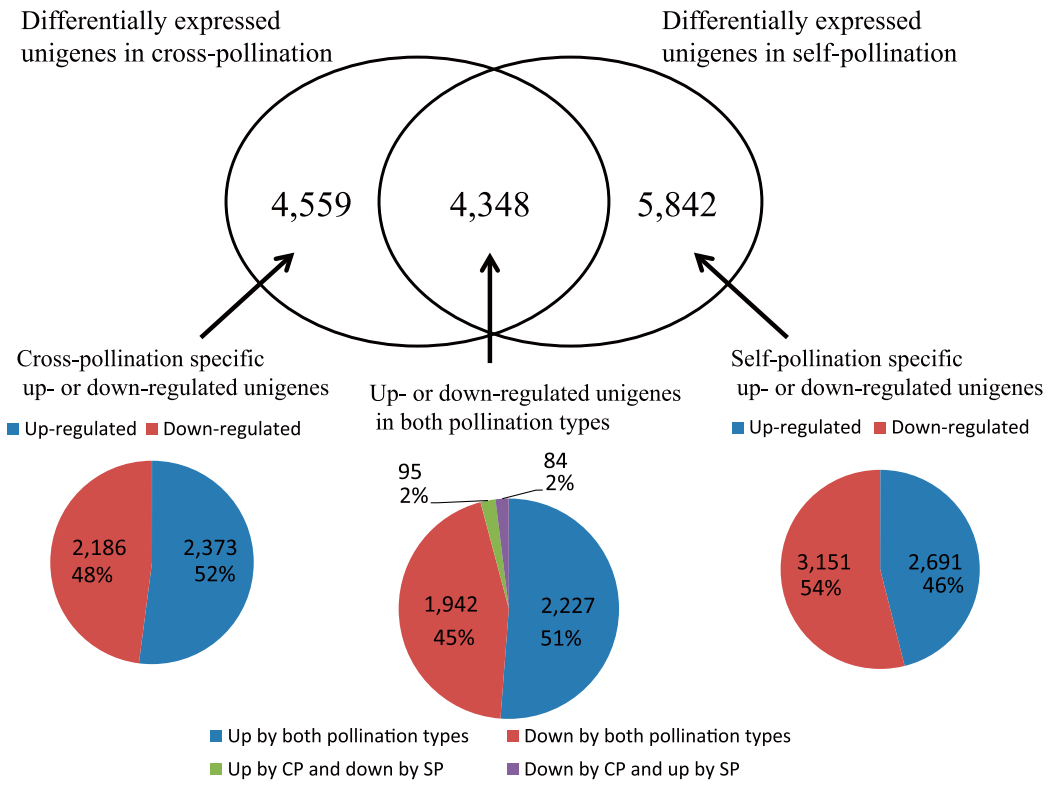

Fig. 2. Venn diagram showing the unigenes with significantly different numbers of ESTs among unpollinated and cross/self-pollinated samples. Differentially expressed unigenes were selected based on the FDR (cut-off FDR < 0.01).

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

Table 5. The results of GO enrichment analysis of unigenes differentially expressed between both pollination types.

GO ID	Term	FDR
Up-regulated Unigenes		
GO:0050896	response to stimulus	7.9e-04
GO:0010033	response to organic substance	0.0038
GO:0042221	response to chemical stimulus	0.0038
GO:0009611	response to wounding	0.0038
GO:0044281	small molecule metabolic process	0.026
GO:0009605	response to external stimulus	0.043
Down-regulated Unigenes		
GO:0006412	translation	1.7e-12
GO:0009628	response to abiotic stimulus	3.3e-07
GO:0034645	cellular macromolecule biosynthetic process	1.2e-06
GO:0010467	gene expression	1.2e-06
GO:0009059	macromolecule biosynthetic process	1.9e-06
GO:0042254	ribosome biogenesis	1.3e-04
GO:0009719	response to endogenous stimulus	1.3e-04
GO:0022613	ribonucleoprotein complex biogenesis	1.3e-04
GO:0009725	response to hormone stimulus	1.4e-04
GO:0042221	response to chemical stimulus	2.3e-04
GO:0006970	response to osmotic stress	7.9e-04
GO:0009415	response to water	8.4e-04
GO:0050896	response to stimulus	8.4e-04
GO:0009414	response to water deprivation	0.0012
GO:0009651	response to salt stress	0.0012
GO:0010033	response to organic substance	0.0015
GO:0044249	cellular biosynthetic process	0.0017
GO:0044267	cellular protein metabolic process	0.0017
GO:0006950	response to stress	0.0017
GO:0009058	biosynthetic process	0.0029
GO:0009266	response to temperature stimulus	0.0063
GO:0009737	response to abscisic acid stimulus	0.016
GO:0019538	protein metabolic process	0.026
GO:0009409	response to cold	0.035
GO:0044260	cellular macromolecule metabolic process	0.041

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

Assembled reads	Total no. of processed reads	No. of unigenes	Max unigene length	Minimum unigene length	Ave. length of unigene
UP_454 + UP_GAIIx	25,249,387	32,736	7,478	201	494
NKP[z]	56,465,821	38,797	7,698	201	808

[z] 'Nanko' and 'Kairyo-Uchida' pollen GAIIx reads.

up-regulated by either or both pollination types between *PSG3_71* and *PSG3_96*, which is defined as the *M* locus (Table 10). Among these, only comp15915_c0_seq1 was pollen-specific and up-regulated by self-pollination. On the other hand, homologous sequences to the solanaceous pistil modifier genes, such as *HT-B* and *120K*, were not included among either the pistil-specific differentially expressed unigenes or in the whole unpollinated/pollinated pistil transcriptome.

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-

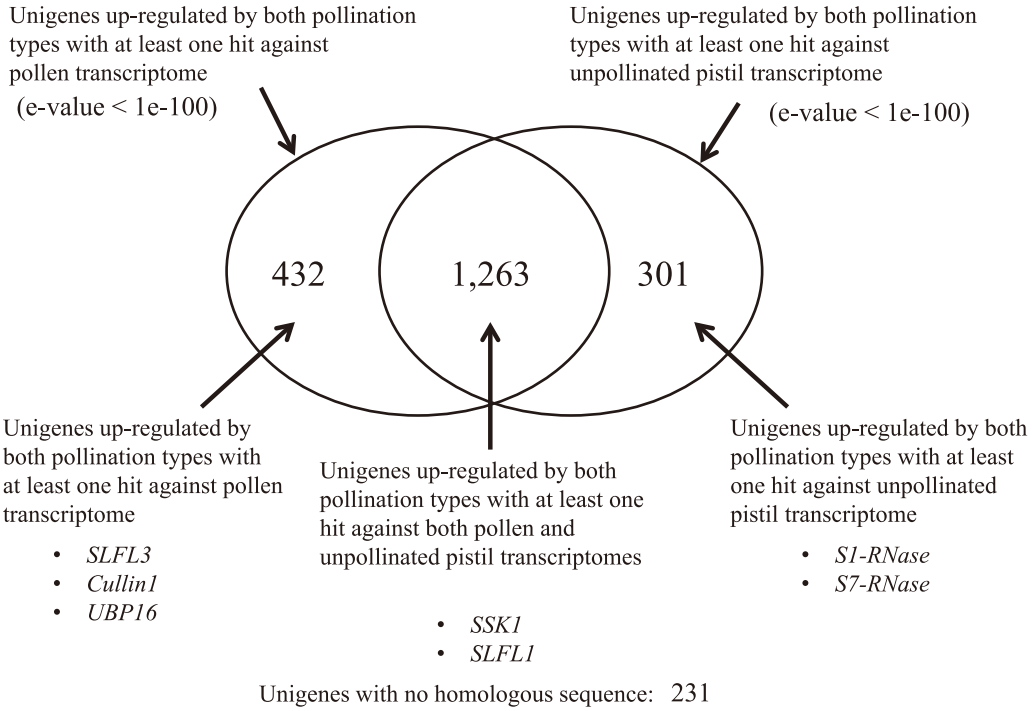

Fig. 3. Venn diagram showing the unigenes with significantly different numbers of ESTs among organs. Differentially expressed unigenes were selected based on the FDR (cut-off FDR < 0.01).

yses in non-model plants because it generates longer reads. However, the number of reads generated by the Roche/454 GS FLX is much lower than is generated by short read platforms such as the Illumina Genome Analyzer. In this study, we used both platforms to fully exploit the advantages of the two different systems. We performed preliminary testing of the *de novo* assemblies generated using either or both platforms (Table 2). Among the three assemblies, the hybrid assembly of the Roche/454 and Illumina reads gave the longest unigene length. This result is consistent with the results of Cahais et al. (2012), which showed that higher quality assemblies were obtained when the Roche/454 and Illumina data were combined, in comparison with the use of either Roche/454 or Illumina data alone. In this study, however, the assembly results of the hybrid data set, i.e. the number and average length of the unigenes, were not much different from those of the Illumina-only reads (Table 2), indicating that the Roche/454 reads contributed much less to the assembly results than the Illumina reads. This is probably because a sufficient number of reliable long unigenes could be reconstructed from the Illumina reads alone. Previously, *de novo* assembly of Illumina reads without reference sequences was more difficult than for Roche/454 reads, because of the short read length (Pop and Salzbarg, 2008; Strickler et al., 2012). Recently, however, with the improvement of read length and advances in assembly algorithms and assemblers (Grabherr et al., 2011; Martin and Wang, 2011), higher quality assemblies can be obtained from Illumina reads alone compared with Roche/454 reads (Luo et al., 2012; Xiao et al., 2013). 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.

Unigene ID	Mapped reads (reads/unigene)			UP to CP		UP to SP		NCBI nr annotation			organ
	UP	CP	SP	log2 (FC)	FDR	log2 (FC)	FDR	Acc. No.	Description	e-value	
comp22_c0_seq1	25,536	62,669	52,365	1.123	0	0.937	0	BAC56115.1	S1-RNase [Prunus mume]	1.0e-154	pistil
comp22_c0_seq2	24,700	48,925	40,861	0.814	0	0.628	0	BAC56116.1	S7-RNase [Prunus mume]	9.0e-142	pistil
comp22_c0_seq3	25,288	49,792	42,506	0.805	0	0.651	0	BAC56116.1	S7-RNase [Prunus mume]	2.0e-138	pistil
comp1704_c0_seq1	1,311	1,824	1,612	0.304	7.8e-08	0.199	1.2e-03	BAF36716.1	S locus F-box protein with the low allelic sequence polymorphism 1-Sf [Prunus mume]	0	both
comp10070_c0_seq1	19	247	74	3.528	6.2e-44	1.863	6.7e-07	BAF36718.1	S locus F-box protein with the low allelic sequence polymorphism 3-Sf [Prunus mume]	1.0e-152	pollen

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

Unigene ID	Mapped reads (reads/unigene)			UP to CP		UP to SP		NCBI nr annotation			organ
	UP	CP	SP	log2 (FC)	FDR	log2 (FC)	FDR	Acc. No.	Description	e-value	
comp38720_c0_seq1	0	0	0	0	1	0	1	BAD08320.1	S-locus F-Box protein 1 [Prunus mume]	1.0e-33	pollen
comp35131_c0_seq1	2	1	0	–1.172	1	–26.753	0.659	BAF36717.1	S locus F-box protein with the low allelic sequence polymorphism 2-Sf [Prunus mume]	3.0e-84	pollen

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

Unigene ID	Mapped reads (reads/unigene)			UP to CP		UP to SP		TAIR10 annotation			organ
	UP	CP	SP	log2 (FC)	FDR	log2 (FC)	FDR	AGI	Description	e-value	
Subunits of SCF complex											
comp2390_c0_seq1	58	2356	353	5.172	0	2.507	2.1e-46	AT5G42190.1	ASK2, SKP1B | E3 ubiquitin ligase SCF complex subunit SKP1/ASK1 family protein	6.0e-28	both
comp7228_c0_seq1	96	202	198	0.901	2.7e-06	0.946	1.1e-06	AT4G02570.4	CUL1 | cullin 1	4.0e-10	pollen
comp4940_c0_seq1	500	802	792	0.510	6.4e-09	0.565	1.2e-10	AT1G26830.1	CUL3A, ATCUL3A, ATCUL3, CUL3 | cullin 3	9.0e-166	both
comp9095_c0_seq1	196	336	289	0.606	2.8e-05	0.462	2.7e-03	AT5G46210.1	CUL4, ATCUL4 | cullin4	0	both
Ubiquitin or Ubiquitin-like protein											
comp2840_c0_seq1	533	996	1,048	0.730	2.3e-20	0.877	8.2e-30	AT5G14360.1	Ubiquitin-like superfamily protein	3.0e-40	pistil
comp18859_c0_seq1	7	82	55	3.378	3.5e-14	2.875	3.2e-08	AT5G25340.1	Ubiquitin-like superfamily protein	5.0e-17	pollen
comp438_c0_seq1	1,495	2,247	2,200	0.416	9.6e-17	0.459	6.7e-20	AT5G57860.2	Ubiquitin-like superfamily protein	8.0e-35	both
comp904_c0_seq1	1,337	1,996	2,074	0.406	2.7e-14	0.535	9.2e-25	AT4G21980.2	APG8A, ATG8A | Ubiquitin-like superfamily protein	8.0e-45	both
comp747_c0_seq1	3,006	3,800	3,476	0.166	2.3e-05	0.111	8.7e-03	AT2G17200.1	DSK2 | ubiquitin family protein	2.0e-170	both
comp2138_c0_seq1	609	899	897	0.390	2.6e-06	0.460	2.0e-08	AT5G03240.3	UBQ3 | polyubiquitin 3	6.0e-15	both
E2 ubiquitin-conjugating enzyme											
comp2792_c0_seq1	51	1,880	1,017	5.032	0	4.219	2.4e-217	AT2G18600.1	Ubiquitin-conjugating enzyme family protein	2.0e-78	both
comp155_c0_seq1	3,496	8,922	7,853	1.180	0	1.069	0	AT5G41700.5	UBC8 | ubiquitin conjugating enzyme 8	1.0e-83	both
comp9702_c0_seq1	180	281	291	0.471	3.4e-03	0.594	1.1e-04	AT2G16920.1	UBC23, PFU2 | ubiquitin-conjugating enzyme 23	1.0e-123	both
comp9059_c0_seq1	132	239	234	0.684	9.0e-05	0.727	2.8e-05	AT2G16920.1	UBC23, PFU2 | ubiquitin-conjugating enzyme 23	2.0e-44	both
E3 ubiquitin ligase											
comp3657_c0_seq1	869	1,382	1,309	0.497	2.2e-14	0.492	1.4e-13	AT1G55860.2	UPL1 | ubiquitin-protein ligase 1	4.0e-159	both
comp27124_c0_seq1	3	29	25	3.101	4.2e-05	2.960	3.5e-04	AT4G38600.1	KAK, UPL3 | HEAT repeat; HECT-domain (ubiquitin-transferase)	3.0e-18	No hit
Deuqbiquitinating enzyme											
comp16343_c0_seq1	17	48	44	1.325	3.3e-03	1.273	7.8e-03	AT4G24560.1	UBP16 | ubiquitin-specific protease 16	2.0e-13	pollen

Table 10. SP- or Cp-specific up-regulated unigenes located around the *M* locus.

Unigene ID or marker name	Homologous peach gene			Mapped reads (reads/unigene)			UP to CP		UP to SP		TAIR10 annotation			organ
	gene id	start (Mb)	end (Mb)	UP	CP	SP	log2 (FC)	FDR	log2 (FC)	FDR	AGI	Description	e-value	
PGS3_71		18.399												
comp17470_c0_seq1	ppa000048m	18.423	18.435	45	90	78	0.828	0.008	0.695	0.034	AT5G15680.1	ARM repeat superfamily protein	5.0e-87	both
comp32285_c0_seq1	ppa000048m	18.423	18.435	5	28	0	2.313	0.002	–28.075	0.135	AT5G15680.1	ARM repeat superfamily protein	2.0e-35	No hit
comp3643_c0_seq1	ppa005464m	18.436	18.439	624	844	826	0.264	0.003	0.306	4.4e-04	AT3G30390.2	Transmembrane amino acid transporter family protein	2.0e-174	both
comp3643_c0_seq2	ppa005464m	18.436	18.439	577	744	750	0.195	0.052	0.280	0.002	AT3G30390.2	Transmembrane amino acid transporter family protein	9.0e-166	both
comp20561_c0_seq1	ppa021437m	18.455	18.457	43	97	84	1.002	7.5e-04	0.867	0.007	AT3G30390.2	Transmembrane amino acid transporter family protein	1.0e-73	both
comp5700_c0_seq4	ppa007239m	18.478	18.480	25	65	25	1.206	0.002	–0.099	0.979	AT3G30380.1	alpha/beta-Hydrolases superfamily protein	7.0e-06	both
comp15915_c0_seq1	**ppa000002m**	**18.524**	**18.545**	**34**	**49**	**91**	**0.355**	**0.504**	**1.322**	**3.3e-05**	**AT3G02260.1**	**BIG, DOC1, TIR3, UMB1, ASA1, LPR1, CRM1 | auxin transport protein (BIG)**	**3.0e-15**	**pollen**
comp19579_c0_seq1	ppa000002m	18.524	18.545	25	45	73	0.676	0.168	1.447	5.1e-05	AT3G02260.1	BIG, DOC1, TIR3, UMB1, ASA1, LPR1, CRM1 | auxin transport protein (BIG)	2.0e-41	No hit
comp31231_c0_seq1	ppa000002m	18.524	18.545	3	22	14	2.702	0.002	2.124	0.038	AT3G02260.1	BIG, DOC1, TIR3, UMB1, ASA1, LPR1, CRM1 | auxin transport protein (BIG)	2.0e-28	No hit
comp5042_c0_seq1	ppa000002m	18.524	18.545	732	978	842	0.246	0.003	0.103	0.290	AT3G02260.1	BIG, DOC1, TIR3, UMB1, ASA1, LPR1, CRM1 | auxin transport protein (BIG)	0	both
comp11745_c0_seq1	ppa011450m	18.575	18.576	45	97	111	0.936	0.002	1.204	1.2e-05	No hits found	—		No hit
comp10245_c0_seq1	ppa001620m	18.582	18.589	108	136	175	0.161	0.626	0.598	0.003	AT5G38880.1	unknown protein	6.0e-75	both
comp23062_c0_seq1	ppa017665m	18.595	18.596	6	53	1	2.971	3.4e-08	–2.684	0.232	AT5G38900.1	Thioredoxin superfamily protein	4.0e-26	pistil
comp29549_c0_seq1	ppa017665m	18.595	18.596	7	69	1	3.129	3.3e-11	–2.906	0.148	AT5G38900.1	Thioredoxin superfamily protein	4.0e-21	pistil
PGS3_62 (*M*-locus)		18.612												
comp208_c0_seq1	ppa010577m	18.619	18.624	5,471	7,185	6,059	0.221	3.0e-16	0.049	0.151	AT5G15800.1	SEP1, AGL2 | K-box region and MADS-box transcription factor family protein	2.0e-95	both
comp15679_c0_seq1	ppa026503m	18.630	18.630	33	115	97	1.629	1.5e-08	1.457	2.6e-06	AT5G15802.1	unknown protein	1.0e-29	pistil
comp15212_c0_seq1	ppa007756m	18.645	18.648	37	122	32	1.549	1.3e-08	–0.308	0.633	AT3G02230.1	RGP1, ATRGP1 | reversibly glycosylated polypeptide 1	5.0e-54	pistil
comp27380_c0_seq1	ppa007756m	18.645	18.648	7	37	2	2.230	3.1e-04	–1.906	0.309	AT3G02230.1	RGP1, ATRGP1 | reversibly glycosylated polypeptide 1	8.0e-39	pistil
comp22731_c0_seq1	ppa014104m	18.664	18.665	7	23	26	1.544	0.051	1.794	0.009	AT3G29970.1	B12D protein	2.0e-27	No hit
PGS3_96		18.764												

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

Unigene ID	Mapped reads (reads/unigene)			UP to CP		UP to SP		TAIR10 annotation			organ
	UP	CP	SP	log2 (FC)	FDR	log2 (FC)	FDR	AGI	Description	e-value	
SP-specific up-regulated DUBs											
comp25740_c0_seq1	9	24	48	1.243	0.096	2.316	8.5e-06	AT3G63090.1	Ubiquitin carboxyl-terminal hydrolase family protein	5.0e-45	No hit
comp24948_c0_seq1	4	2	38	–1.172	0.839	3.149	1.4e-06	AT5G21970.1	Ubiquitin carboxyl-terminal hydrolase family protein	5.0e-08	No hit
comp13199_c0_seq1	78	57	168	–0.625	0.050	1.008	2.2e-06	AT1G04860.1	UBP2, ATUBP2 | ubiquitin-specific protease 2	6.0e-43	both
comp12545_c0_seq1	140	186	286	0.238	0.289	0.932	3.1e-09	AT2G40930.1	UBP5, ATUBP5, PDE323 | ubiquitin-specific protease 5	4.0e-91	both
comp17181_c0_seq1	**55**	**68**	**109**	**0.134**	**0.809**	**0.888**	**9.4e-04**	**AT1G32850.1**	**UBP11 | ubiquitin-specific protease 11**	**5.0e-39**	**pollen**
comp7801_c0_seq1	310	354	445	0.019	0.975	0.423	5.4e-04	AT4G01037.1	WTF1, AtWTF1 | Ubiquitin carboxyl-terminal hydrolase family protein	2.0e-35	pistil
CP-specific up-regulated DUBs											
comp20501_c0_seq1	46	127	83	1.293	6.8e-07	0.753	0.020	AT3G47890.1	Ubiquitin carboxyl-terminal hydrolase-related protein	6.0e-76	both
comp9660_c0_seq1	138	232	155	0.577	1.3e-03	0.069	0.858	AT5G22030.1	UBP8 | ubiquitin-specific protease 8	1.0e-55	both
comp1343_c0_seq1	1,323	1,796	1,388	0.269	3.1e-06	–0.030	0.755	AT5G06600.3	UBP12 | ubiquitin-specific protease 12	0	both
comp1343_c0_seq3	448	664	511	0.396	6.3e-05	0.091	0.507	AT5G06600.3	UBP12 | ubiquitin-specific protease 12	0	both
comp1343_c0_seq7	366	570	466	0.467	1.3e-05	0.250	0.040	AT5G06600.3	UBP12 | ubiquitin-specific protease 12	3.0e-84	both
comp1343_c0_seq2	360	597	471	0.558	7.7e-08	0.289	0.016	AT5G06600.3	UBP12 | ubiquitin-specific protease 12	0	both
comp1343_c0_seq5	46	100	51	0.948	1.0e-03	0.050	1	AT5G06600.3	UBP12 | ubiquitin-specific protease 12	5.0e-147	both
comp23575_c0_seq1	7	29	24	1.879	5.9e-03	1.679	0.018	AT5G65450.1	UBP17 | ubiquitin-specific protease 17	1.0e-33	No hit
comp13466_c0_seq1	48	98	53	0.858	4.3e-03	0.044	1	AT2G24640.1	UBP19 | ubiquitin-specific protease 19	6.0e-09	both
comp9414_c0_seq1	166	258	166	0.464	6.2e-03	–0.099	0.702	AT5G10790.1	UBP22 | ubiquitin-specific protease 22	2.0e-40	both
comp7266_c0_seq1	360	526	344	0.375	1.0e-03	–0.164	0.257	AT3G49600.1	UBP26, SUP32, ATUBP26 | ubiquitin-specific protease 26	2.0e-75	both
comp3345_c0_seq1	898	1172	1040	0.212	4.8e-03	0.113	0.178	AT4G17510.1	UCH3 | ubiquitin C-terminal hydrolase 3	2.0e-94	both

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 pistil 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, 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 SLFL*s are thought to be putative orthologs of the Pyrinae *SFBB*s 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 *PavSSK1*, which is a gene for the SKP1-like1 protein interacting with SLFLs and SFB to form the SCF complex with Cullin1-likes (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

Altschul, S. F., W. Gish, W. Miller, E. W. Myers and D. J. Lipman. 1990. Basic local alignment search tool. J. Mol. Biol. 215: 403–410.

Cahais, V., P. Gayral, G. Tsagkogeorga, J. Melo-Ferreira, M. Ballenghein, L. Weinert, Y. Chiari, K. Belkhir, V. Ranwez and N. Galtier. 2012. Reference-free transcriptome assembly in non-model animals from next-generation sequencing data. Molecular Ecology Resources 12: 834–845.

De Franceschi, P., L. Dondini and J. Sanzol. 2012. Molecular bases and evolutionary dynamics of self-incompatibility in the Pyrinae (Rosaceae). J. Exp. Bot. 63: 4015–4032.

De Franceschi, P., L. Pierantoni, L. Dondini, M. Grandi, J. Sanzol and S. Sansavini. 2011a. Cloning and mapping multiple *S*-locus F-box genes in European pear (*Pyrus communis* L.). Tree Genet. Genomes 7: 231–240.

De Franceschi, P., L. Pierrantoni, L. Dondini, M. Grandi, S. Sansavini and J. Sanzol. 2011b. Evaluation of candidate F-box genes for the pollen *S* of gametophytic self-incompatibility in the Pyrinae (Rosaceae) on the basis of their phylogenomic context. Tree Genet. Genomes 7: 663–683.

Du, Z., X. Zhou, Y. Ling, Z. Zhang and Z. Su. 2010. agriGO: a GO analysis toolkit for the agricultural community. Nucl. Acids Res. 38: W64–W70.

Ekblom, R. and J. Galindo. 2011. Applications of next generation sequencing in molecular ecology of non-model organisms. Heredity 107: 1–15.

Entani, T., M. Iwano, H. Shiba, F. S. Che, A. Isogai and S. Takayama. 2003. Comparative analysis of the self-incompatibility (*S*-) locus region of *Prunus mume*: identification of a pollen-expressed F-box gene with allelic diversity. Genes Cells 8: 203–213.

Fernández i Martí, A., T. Hanada, J. M. Alonso, H. Yamane, R. Tao and R. Socias I Company. 2009. A modifier locus affecting the expression of the *S*-RNase gene could be the cause of breakdown of self-incompatibility in almond. Sex. Plant Reprod. 22: 179–186.

Fu, N., Q. Wang and H.-L. Shen. 2013. *De novo* assembly, gene annotation and marker development using Illumina paired-end transcriptome sequences in celery (*Apium graveolens* L.). PLoS ONE 8: e57686. DOI: 10.1371/journal.pone.0057686. <http://www.plosone.org>.

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

Grabherr, M. G., B. J. Haas, M. Yassour, J. Z. Levin, D. A. Thompson, I. Amit, X. Adiconis, L. Fan, R. Raychowdhury, Q. Zeng, Z. Chen, E. Mauceli, N. Hacohen, A. Gnirke, N. Rhind, F. di Palma, B. W. Birren, C. Nusbaum, K. Lindblad-Toh, N. Friedman and A. Regev. 2011. Full-length transcriptome assembly from RNA-seq data without a reference genome. Nature Biotechnol. 29: 644–652.

Graciet, E. and F. Wellmer. 2010. The plant N-end rule pathway: structure and functions. Trends Plant Sci. 15: 447–453.

Habu, T., H. Yamane, K. Igarashi, K. Hamada, K. Yano and R. Tao. 2012. 454-pyrosequencing of the transcriptome in leaf and flower buds of Japanese apricot (*Prunus mume* Sieb. et Zucc.) at different dormant stages. J. Japan. Soc. Hort. Sci. 81: 239–250.

Hancock, C. N., L. Kent and B. A. McClure. 2005. The stylar 120 kDa glycoprotein is required for *S*-specific pollen rejection in *Nicotiana*. Plant J. 43: 716–723.

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

Jurka, J., V. V. Kapitonov, A. Pavlicek, P. Klonowski, O. Kohany and J. Walichiewicz. 2005. Repbase Update, a database of eukaryotic repetitive elements. Cytogenet. Genome Res. 110: 462–467.

Kakui, H., M. Kato, K. Ushijima, M. Kitaguchi, S. Kato and H. Sassa. 2011. Sequence divergence and loss-of-function phenotypes of *S locus F-box brothers* genes are consistent with non-self recognition by multiple pollen determinants in self-incompatibility of Japanese pear (*Pyrus pylifolia*). Plant J. 68: 1028–1038.

Kubo, K., T. Entani, A. Takara, N. Wang, A. M. Fields, Z. Hua, M. Toyoda, S. Kawashima, T. Ando, A. Isogai, T. H. Kao and S. Takayama. 2010. Collaborative non-self recognition system in S-RNase-based self-incompatibility. Science 330: 796–799.

Lai, Z., W. Ma, B. Han, L. Liang, Y. Zhang, G. Hong and Y. Xue. 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.

Langmead, B., C. Trapnell, M. Pop and S. L. Salzberg. 2009. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biology 10: R25.

Liu, S., W. Li, Y. Wu, C. Chen and J. Lei. 2013. *De novo* transcriptome assembly in chili pepper (*Capsicum frutescens*) to identify genes involved in the biosynthesis of capsaicinoids. PLoS ONE 8: e48156. DOI: 10.1371/journal.pone.0048156. <http://www.plosone.org>.

Luo, C., D. Tsementzi, N. Kyrpides, T. Read and K. T. Konstantinidis. 2012. Direct comparisons of Illumina vs. Roche 454 sequencing technologies on the same microbial community DNA sample. PLoS ONE 7: e30087. DOI: 10.1371/journal.pone.0030087. <http://www.plosone.org>.

Martin, J. A. and Z. Wang. 2011. Next-generation transcriptome assembly. Nature Reviews Genetics 12: 671–682.

Matsumoto, D., H. Yamane, K. Abe and R. Tao. 2012. Identification of a Skp1-like protein interacting with SFB, the pollen *S* determinant of the gametophytic self-incompatibility in *Prunus*. Plant Physiol. 159: 1252–1262.

Matsumoto, D., H. Yamane and R. Tao. 2008. Characterization of *SLFL1*, a pollen-expressed F-box gene located in the *Prunus S* locus. Sex. Plant Reprod. 21: 113–121.

McClure, B., F. Cruz-García and C. Romero. 2011. Compatibility and incompatibility in S-RNase-based systems. Ann. Botany

108: 647–658.
McClure, B. A., J. E. Gray, M. A. Anderson and A. E. Clarke. 1990. Self-incompatibility in *Nicotiana alata* involves degradation of pollen rRNA. Nature 347: 757–760.
McClure, B., B. Mou, S. Canevascini and R. Bernatzky. 1999. A small asparagine-rich protein required for S-allele-specific pollen rejection in *Nicotiana*. Proc. Natl. Acad. Sci. USA 96: 13548–13553.
Minamikawa, M., H. Kakui, S. Wang, N. Kotoda, S. Kikuchi, T. Koba and H. Sassa. 2010. Apple *S* locus region represents a large cluster of related, polymorphic and pollen-specific F-box genes. Plant Mol. Biol. 74: 143–154.
Moriya, Y., K. Okada, K. Yamamoto, H. Iwanami, H. Bessho and T. Takasaki-Yasuda. 2009. Characterization of partial self-compatibility in the European pear cultivar, 'Grand Champion'. J. Hortic. Sci. Biotechnol. 84: 77–82.
Newbigin, E., T. Paape and J. R. Kohn. 2008. RNase-based self-incompatibility: puzzled by *pollen S*. Plant Cell 20: 2286–2292.
Okada, K., N. Tonaka, Y. Moriya, N. Norioka, Y. Sawamura, T. Matsumoto, T. Nakanishi and T. Takasaki-Yasuda. 2008. Deletion of a 236 kb region around *S_4-RNase* in a stylar-part mutant *S_4^{sm}*-haplotype of Japanese pear. Plant Mol. Biol. 66: 389–400.
Okada, K., N. Tonaka, T. Taguchi, T. Ichikawa, Y. Sawamura, T. Nakanishi and T. Takasaki-Yasuda. 2011. Related polymorphic F-box protein genes between haplotypes clustering in the BAC contig sequences around the *S-RNase* of Japanese pear. J. Exp. Bot. 62: 1887–1902.
Oono, Y., F. Kobayashi, Y. Kawahara, T. Yazawa, H. Handa, T. Itoh and T. Matsumoto. 2013. Characterization of the wheat (*Triticum aestivum* L.) transcriptome by *de novo* assembly for the discovery of phosphate starvation-responsive genes: gene expression in Pi-stressed wheat. BMC Genomics 14: 77.
Pareek, C. S., R. Smoczynski and A. Tretyn. 2011. Sequencing technologies and genome sequencing. J. Appl. Genetics 52: 413–435.
Pop, M. and S. L. Salzberg. 2008. Bioinformatics challenges of new sequencing technology. Trends Genet. 24: 142–149.
Qiao, H., F. Wang, L. Zhao, J. Zhou, Z. Lai, Y. Zhang, T. P. Robbins and Y. Xue. 2004. The F-box protein AhSLF-S_2 controls the pollen function of S-RNase-based self-incompatibility. Plant Cell 16: 2307–2322.
Robinson, M. R., D. J. McCarthy and G. K. Smyth. 2010. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics 26: 139–140.
Sassa, H., H. Kakui and M. Minamikawa. 2010. Pollen-expressed F-box gene family and mechanism of S-RNase-based gametophytic self-incompatibility (GSI) in Rosaceae. Sex. Plant Reprod. 23: 39–43.
Sassa, H., H. Kakui, M. Miyamoto, Y. Suzuki, T. Hanada, K. Ushijima, M. Kusaba, H. Hirano and T. Koba. 2007. *S locus F-box brothers*: multiple and pollen-specific F-box genes with *S* haplotype-specific polymorphisms in apple and Japanese pear. Genetics 175: 1869–1881.
Sijacic, P., X. Wang, A. L. Skirpan, Y. Wang, P. E. Dowd, A. G. McCubbin, S. Huang and T.-H. Kao. 2004. Identification of the pollen determinant of S-RNase-mediated self-incompatibility. Nature 429: 302–305.
Smit, A. F. A., R. Hubley and P. Green. 1996–2010. RepeatMasker Open-3.0. <http://www.repeatmasker.org>.
Strickler, S. R., A. Bombarely and L. A. Mueller. 2012. Designing a transcriptome next-generation sequencing project for a nonmodel plant species. American Journal of Botany 99: 257–266.
Tao, R. and A. F. Iezzoni. 2010. The S-RNase-based gametophytic self-incompatibility systems in *Prunus* exhibits distinct genetic and molecular features. Sci. Hortic. 124: 423–433.
Tao, R., H. Yamane, A. Sugiura, H. Murayama, H. Sassa and H. Mori. 1999. Molecular typing of *S*-alleles through identification, characterization and cDNA cloning for *S*-RNases in sweet cherry. J. Amer. Soc. Hort. Sci. 124: 224–233.
The Gene Ontology Consortium. 2000. Gene Ontology: tool for the unification of biology. Nature Genet. 25: 25–29. <http://www.geneontology.org>.
The International Peach Genome Initiative. 2013. The high-quality draft genome of peach (*Prunus persica*) identifies unique patterns of genetic diversity, domestication and genome evolution. Nature Genet. 45: 487–494.
Ushijima, K., H. Sassa, A. M. Dandekar, T. M. Gradziel, R. Tao and H. Hirano. 2003. Structural and transcriptional analysis of the self-incompatibility locus of almond: identification of a pollen-expressed F-box gene with haplotype-specific polymorphism. Plant Cell 15: 771–781.
Ushijima, K., H. Yamane, A. Watari, E. Kakehi, K. Ikeda, N. R. Hauck, A. F. Iezzoni and R. Tao. 2004. The *S* haplotype-specific F-box protein gene, *SFB*, is defective in self-compatible haplotypes of *Prunus avium* and *P. mume*. Plant J. 39: 573–586.
Vilanova, S., M. L. Badenes, L. Burgos, J. Martínez-Calvo, G. Llácer and G. Romero. 2006. Self-incompatibility of two apricot selections is associated with two pollen-part mutations of different nature. Plant Physiol. 142: 629–641.
Wang, Y., T. Tsukamoto, K. W. Yi, X. Wang, S. Huang, A. G. McCubbin and T. H. Kao. 2004. Chromosome walking in the *Petunia inflata* self-incompatibility (*S*-) locus and gene identification in an 881-kb contig containing *S_2-RNase*. Plant Mol. Biol. 54: 727–742.
Wünsch, A. and J. I. Hormaza. 2004. Genetic and molecular analysis in Cristobalina sweet cherry, a spontaneous self-compatible mutant. Sex. Plant Reprod. 17: 203–210.
Xiao, M., Y. Zhang, X. Chen, E. J. Lee, C. J. S. Barber, R. Chakrabarty, I. Desgagné-Penix, T. M. Haslam, Y. B. Kim, E. Liu, G. MacNevin, S. Masada-Atsumi, D. W. Reed, J. M. Stout, P. Zerbe, Y. Zhang, J. Bohlmann, P. S. Covello, V. De Luca, J. E. Page, D. K. Ro, V. J. J. Martin, P. J. Facchini and C. W. Sensen. 2013. Transcriptome analysis based on next-generation sequencing on non-model plants producing specialized metabolites on biotechnological interest. J. Biotechnol. 166: 122–134.
Yamane, H. and R. Tao. 2009. Molecular basis of self-(in)compatibility and current status of *S*-genotyping in rosaceous fruit trees. J. Japan. Soc. Hort. Sci. 78: 137–157.
Yamane, H., K. Ikeda, K. Ushijima, H. Sassa and R. Tao. 2003. A pollen-expressed gene for a novel protein with an F-box motif that is very tightly linked to a gene for S-RNase in two species of cherry, *Prunus cerasus* and *P. avium*. Plant Cell Physiol. 44: 764–769.
Zhang, Q., W. Chen, L. Sun, F. Zhao, B. Huang, W. Yang, Y. Tao, J. Wang, Z. Yuan, G. Fan, Z. Xing, C. Han, H. Pan, X. Zhong, W. Shi, X. Liang, D. Du, F. Sun, Z. Xu, R. Hao, T. Lv, Y. Lv, Z. Zheng, M. Sun, L. Luo, M. Cai, Y. Gao, J. Wang, Y. Yin, X. Xu, T. Cheng and J. Wang. 2012. The genome of *Prunus mume*. Nature communications 3: 1318.
Zuriaga, E., L. Molina, M. L. Badenes and C. Romero. 2012. Physical mapping of a pollen modifier locus controlling self-incompatibility in apricot and synteny analysis within the Rosaceae. Plant Mol. Biol. 79: 229–242.