Insights into the evolution and establishment of the *Prunus*-specific self-incompatibility recognition mechanism

2017

Takuya Morimoto

Contents

General introduction

1. Rosaceae contains important fruit and nut species with self-incompatibility

The family Rosaceae, which belongs to the order Rosales, includes 91 genera and 2,950 species (Christenhusz et al., 2016). Taxonomically, the Rosaceae can be divided into three subfamilies, Rosoideae, Dryadoideae and Amygdaloideae (Potter et al., 2007). Of these, Amygdaloideae includes a large number of economically important fruit tree species, such as apple (*Malus* × *domestica*) and pear (*Pyrus* spp.), which belong to the subtribe Malinae, and peach (*Prunus persica*), almond (*Prunus dulcis*), Japanese plum (*Prunus salicina*), European plum (*Prunus domestica*), Japanese apricot (*Prunus mume*), apricot (*Prunus armeniaca*), sweet cherry (*Prunus avium*) and sour cherry (*Prunus cerasus*), which belong to the genus *Prunus*. According to FAOSTAT (http://faostat3.fao.org/), 76,380,000 tons of apple, 23,580,000 tons of pear, 21,080,000 tons of peach, 9,520,000 tons of plum and 2,260,000 tons of cherry were produced worldwide in 2012. These rosaceous fruit trees are also economically and horticulturally important in Japan. Apple, Japanese pear (*Pyrus pyrifolia*) and peach were ranked 2nd, 3rd and 6th in the total fruit production of 2014 (http://www.maff.go.jp/) in Japan.

Pollination and fertilization are indispensable for successful fruit production in these rosaceous fruit tree species because they stimulate cell division and the enlargement of specific floral tissues that develop into the edible fruit parts. However, the self-fertilization of most rosaceous fruit tree species, including apple, pear, almond, plum, apricot and cherry, is hindered by self-incompatibility (SI). SI is defined as the inability of fertile hermaphrodite plants to produce zygotes after self-fertilization (de Nettancourt, 1997). Therefore, in cultural practices, the inter-planting of different cultivars (pollinizer) and insect pollinators is used to ensure pollination and fruit set in SI fruit tree species. Thus, the synchronous flowering of cultivated varieties and pollinizer is required for successful pollination. However, global warming and climate change have affected winter chilling (Atkinson et al., 2013; Luedeling et al., 2011) and thus the dormancy of perennial fruit tree species (Sugiura et al., 2007), which causes inconsistent flowering periods among cultivated varieties and pollinizer, resulting in pollination failure. Climate change also affects the temporal and spatial activities of the insect pollinators, in addition to the spread of pests and pathogens (Potts et al., 2010).

Although artificial pollination by hand is labor-intensive work, growers usually use this practice to ensure fruit set in SI fruit tree species under commercial production in Japan. In fact, the labor time required for artificial pollination and subsequent fruit thinning represents 22% (for Japanese plum) and 7% (for sweet cherry) of the total labor time in annual fruit production (http://www.maff.go.jp/). SI is also an obstacle to cross-breeding programs because crosses between cultivars with the same SI genotype produce no seeds.

The introduction of self-compatible (SC) cultivars is a possible solution to the SI-related issues. An SC mutant of Japanese pear, 'Osa-Nijisseiki', was found as a bud sport of 'Nijisseiki'. Because the fruit quality and growth habits of this mutant variety are similar to those of a leading cultivar, 'Nijisseiki', the production of 'Osa-Nijisseiki' has increased rapidly in Japan. In addition, 'Osa-Nijisseiki' has been used as a parent to breed new SC cultivars, such as 'Akibae', 'Natsuhime' and 'Zuishu'. SC mutants were also reported in *Prunus*. In sweet cherry, SC selections, 'JI2420' and 'JI2434', were produced at the John Innes Institute in the United Kingdom through the use of X-ray-irradiated pollen on its incompatible seed parent (Lewis and Crowe, 1954). The SC cultivar 'Stella' is derived from 'Lambert' × 'JI2430' and has been extensively used in sweet cherry breeding programs. 'Benikirari' and 'Lapins' are offspring of 'Stella'. SC mutants have also been found in Japanese plum (Watari et al., 2007), Japanese apricot (Ushijima et al., 2004), apricot (Vilanova et al., 2006) and sour cherry (Hauck et al., 2006a, b). Because Japanese apricot flowers early in the spring when pollinator insect activity is low due to low temperature, the use of SC cultivars has great advantages in ensuring fruit set even under conditions unfavorable to pollination.

2. SI systems in angiosperm

Diversity of the SI systems

SI is a major liming factor in cultural practices and breeding programs in rosaceous fruit trees. However, SI is also an important reproductive mechanism involved in maintaining genetic diversity within a species (de Nettancourt, 2001). SI has been observed not only in the Rosaceae but also in a wide range of angiosperm species. SI is found in at least 19 orders, 71 families and 250 genera, comprising approximately 60% of angiosperm species (Allen and Hiscock, 2008).

SI can be classified into two types based on the genetic control of the incompatibility phenotype in pollen: sporophytic SI (SSI) and gametophytic SI (GSI). SSI can be further divided into two types depending on its association with floral morphology: homomorphic SSI and heteromorphic SSI. Homomorphic SSI has been found in the Brassicaceae, Compositae and Convolvulaceae. Heteromorphic SSI has been reported in the Linaceae, Oxalidaceae, Polygonaceae and Primulaceae (de Nettancourt, 2001; Tao and Iezzoni, 2010). GSI is wide spread across angiosperm. It has been found in the Solanaceae, Papaveraceae, Ranunculaceae, Rubiaceae, Fabaceae, Leguminosae, Onagraceae, Plantaginaceae, Rosaceae and Poaceae (Igic and Kohn, 2001; Tao and Iezzoni, 2010).

Despite the multiple origins and the wide distribution range of diverse SI systems in angiosperm, three distant plant families, the Rosaceae, Solanaceae and Plantaginaceae, share a common SI mechanism, called S-RNase-based GSI (McClure, 2009; Tao and Iezzoni, 2010). The most recent common ancestor (MRCA) of these families is the root of approximately 75% of all dicots; therefore, S-RNase-based GSI is considered to be the ancestral state of the majority of dicots, which are thought to have arisen approximately 120 million years ago (Igic and Kohn, 2001; Steinbachs and Holsinger, 2002; Vieira et al., 2008). This implies that the loss of the S-RNase-based GSI system and the gain of other SI systems occurred frequently during the lineage speciation of dicots (Sherman-Broyles and Nasrallah, 2008). In fact, SI systems are supposed to have arisen in the basal angiosperms and are estimated to have evolved independently at least 21 times during the evolution of flowering plants (Flanklin-Tong, 2008).

S-RNase-based GSI system

The S-RNase-based GSI system is found in *Prunus* and Malinae in the Rosaceae, Petunioideae, Nicotianoideae, and Solanoideae in the Solanaceae, and Antirrhineae in the Plantaginaceae (Anderson et al., 1986; McClure et al., 1989; Sassa et al., 1992, 1993, 1996; Tao et al., 1997, 1999; Xue et al., 1996). *Coffea* (Rubiaceae) and *Fragaria* (Rosaceae) may also have S-RNase-based GSI systems (Bošković et al., 2009; Nowak et al., 2011). These families are widespread in the dicot lineage (Fig. 1).

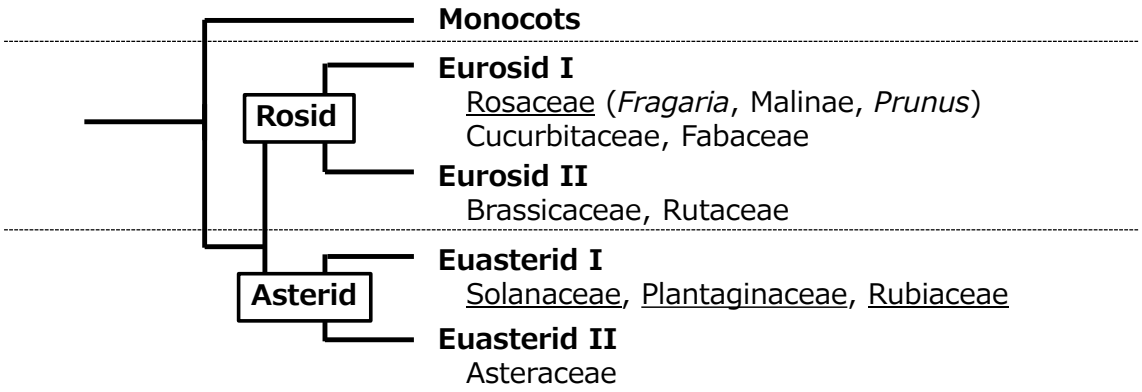

Fig. 1 Phylogenic relationship of eudicot species. Underlining represents families with S-RNase-based gametophytic self-incompatibility.

Self/nonself recognition between female (pistil) and male (pollen) organs is generally controlled by a single multi-allelic *S* locus. In the S-RNase-based GSI system, the *S* locus encodes two tightly linked genes, S-RNase and F-box genes, as pistil *S* and pollen *S* determinants, respectively. The term "*S* haplotype" is used to describe the variants of the *S* locus. The SI reaction is triggered when the genes derived from the same *S* haplotype are expressed in both the pistil and pollen, resulting in the arrest of pollen tube growth.

The identification and characterization of the specificity determinants has contributed to identify SC mutants and to develop new SC cultivars in rosaceous fruit tree species. Most of the SC mutants observed in the Rosaceae are from the mutation of the specificity determinants (Hauck et al., 2006a, b; Okada et al., 2008; Sonneveld et al., 2005; Ushijima et al., 2004). Based on the sequence polymorphisms of the specificity determinant genes, systems for the molecular *S* genotyping and marker-assisted selection of SC individuals have been established in various rosaceous fruit tree species (Yamane and Tao, 2009). However, the whole image of self-(in)compatibility reactions remains to be clarified and would provide clues to overcome the cultural and breeding difficulties caused by SI.

3. Current knowledge on S-RNase-based GSI

Identification and characterization of the pistil *S* determinant

The molecular mechanisms behind S-RNase-based GSI have been extensively

characterized in solanaceous species, such as *Nicotiana* and *Petunia*. The stylar-expressed glycoprotein, which co-segregates with its pistil *S* allele, was first identified in *Nicotiana alata* (Anderson et al., 1986, 1989). Since, the glycoprotein has been shown to be homologous to *Aspergillus* RNase-T2, and it exhibits RNase activity; therefore, it was named S-RNase (McClure et al., 1989). Transformation experiments in *Nicotiana* and *Petunia* confirmed that the inhibition of *S-RNase* expression results in the failure of transgenic plants to reject their own pollen (Lee et al., 1994; Murfett et al., 1994). In addition, gain-of-function experiment with foreign *S-RNase* conferred the ability to reject pollen of the corresponding *S* haplotype, providing direct evidence that S-RNases alone are sufficient to control the SI recognition in pistils and are presumed to function as cytotoxins against self-pollen (Lee et al., 1994; Murfett et al., 1994).

Following the identification of pistil *S* in the Solanaceae, Sassa et al., (1992, 1993, 1996) and Broothaerts et al., (1995) reported that S-RNases are also associated with the GSI of *Pyrus* and *Malus* (Malinae). Later, Xue et al. (1996) cloned the S-RNase of *Antirrhinum* (Plantaginaceae) as pistil *S*. Subsequently, Tao et al. (1997, 1999) and Ushijima et al., (1998) identified S-RNase as pistil *S* in *Prunus* (Rosaceae). Transgenic experiments in *Malus*, and loss-of-function mutants in *Pyrus* and in *Prunus*, showed that S-RNases function as pistil *S* determinants in these species (Broothaerts et al., 2004; Hanada et al., 2009; Sanzol, 2009; Sassa et al., 1997; Watari et al., 2007; Yamane et al., 2003a). Recently, an S-RNase-like gene was also reported in *Coffea* in the Rubiaceae (Nowak et al., 2011).

S-RNase sequences are highly divergent, with allelic amino acid sequence similarities ranging from 30% to 90% (Ushijima et al., 1998; McCubbin and Kao, 2000). Despite the sequence diversity, the primary structural analysis of solanaceous S-RNases revealed five conserved regions, C1 to C5, and two hypervariable regions, HVa and HVb, (Ioerger et al., 1991) (Fig. 2). The C2 and C3 regions contain conserved histidine residues required for catalytic activity. HVa and HVb are considered as primary targets for self/nonself recognition because they exhibit highly divergent sequence polymorphisms in each allele and are located on the surface of crystal structures (Ioerger et al., 1991; Ida et al., 2001). Except for the C4 region, these conserved regions were also identified in rosaceous S-RNases (Ushijima et al., 1998). The C4 region in the Rosaceae, called RC4, is located in a different position and has a different amino acid

sequence compared with that of C4 in the Solanaceae (Fig. 2). In contrast to the presence of two hypervariable regions in solanaceous S-RNases, a single hypervariable region, called RHV, was found in rosaceous S-RNases (Ishimizu et al., 1998a; Ushijima et al., 1998) (Fig. 2). Although, the hypervariable region is positively selected and considered to be responsible for self/nonself recognition, other regions are also important for the allele specificity because certain S-RNases, sharing the same RHV sequence, exhibited different specificities (Matsumoto et al., 2010; Ortega et al., 2006; Zisovich et al., 2004). Thus, the hypervariable region may be mainly responsible for self/nonself recognition but is insufficient for determining specificity at certain alleles. Regarding exon/intron structure, a single intron is present in the cording sequence of the hypervariable region of the S-RNases in the Solanaceae and Malinae (Fig. 2). In *Prunus*, an additional intron is present in the coding sequence located downstream of the signal peptide.

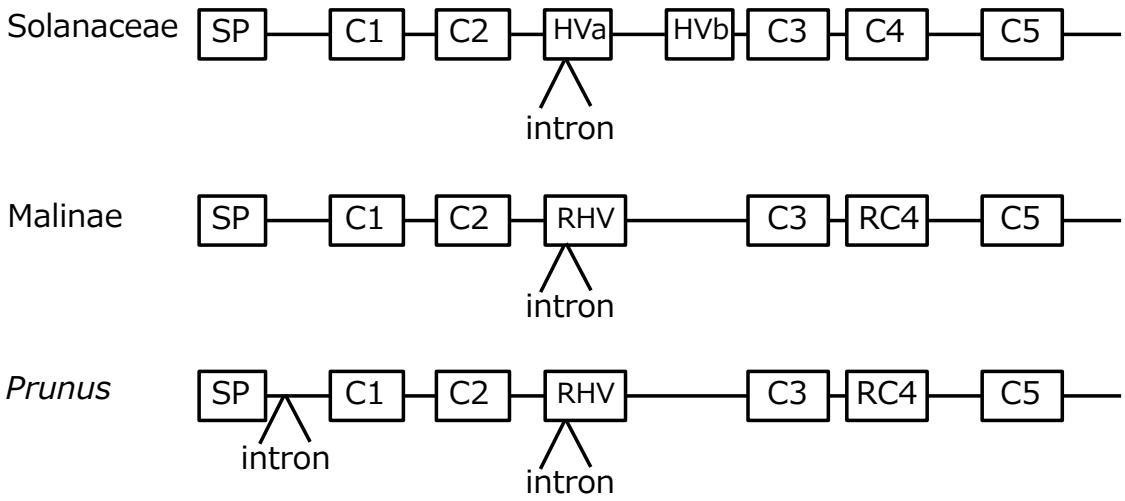

Fig. 2 Gene structures of *S-RNase*s in the Solanaceae, subtribe Malinae and genus *Prunus*. SP: signal peptide; C1–C5: conserved regions; RC4: rosaceous conserved region 4; HVa and HVb: hypervariable regions a and b; RHV: rosaceous hypervariable region.

Phylogenetic relationships of *S-RNase*s

Structural and phylogenic analyses indicated that plant T2-type RNases grouped into three major clades shared by the eudicots (Igic and Kohn, 2001). *S-RNase*s of the Rosaceae, Solanaceae and Plantaginaceae are clustered in the same clade and

share a common origin, indicating that *S-RNase* evolved once in the ancestor of the eudicots (Igic and Kohn, 2001; Steinbachs and Holsinger, 2002; Vieira et al., 2008). Despite the shared origin, *S-RNase*s show a lineage-specific clustering, and *S-RNase*s from the three families form clusters in each family (Igic and Kohn, 2001; Steinbachs and Holsinger, 2002; Vieira et al., 2008). In Rosaceae, *S-RNase*s of Malinae and *Prunus* are located in different clusters, whereas those of *Pyrus* and *Malus* (Malinae) cluster together. Sequence similarities between alleles from *Pyrus* and *Malus* are often higher than that between alleles of the same species as models of trans-specific evolution. The intron structures reflect the lineage-specific *S-RNase* clustering in the Rosaceae because a single intron is present in the *S-RNase*s of Malinae as observed in the Solanaceae and Plantaginaceae, while an additional intron is present in the *S-RNase* of *Prunus* (Fig. 2). Thus, *S-RNase*s seems to have experienced lineage specific evolution despite the single origin.

Identification and characterization of the pollen *S* determinant

After the identification of the pistil *S* determinant, the pollen *S* determinants in the S-RNase-based GSI were intensively surveyed based on their expected features, such as complete linkage to *S-RNase*, sequence divergence among *S* haplotypes and pollen expression. The first clue to the identification of the pollen *S* determinant was obtained from the Plantaginaceae. A sequence analysis of the *S* locus in *Antirrhinum hispanicum* revealed the presence of a pollen-expressed F-box gene (Lai et al., 2002). Map-based cloning and chromosome-walking approaches in the Solanaceae, Plantaginaceae and *Prunus* also identified F-box genes located on the *S* locus as pollen *S* determinant candidates (Entani et al., 2003; Ushijima et al., 2003; Wang et al., 2003; 2004; Zhou et al., 2003). The pollen F-box genes in the Solanaceae and Plantaginaceae were named *S locus F-box* (*SLF*). In *Prunus*, the pollen F-box gene was named *S locus F-box* (*SLF*) or *S haplotype-specific F-box* (*SFB*) by two different research groups (Entani et al., 2003; Ushijima et al., 2003). In this paper, the term *SFB* is used to refer to the pollen *S* F-box gene in *Prunus*, reflecting its different function compared with the *SLF* in the Solanacae and Plantaginaceae in S-RNase-based GSI. Multiple F-box genes were subsequently identified as pollen *S* determinant candidates in Malinae, and named *SFB brother*s (*SFBB*s) (Sassa et al., 2007).

Direct evidence that the *SLF* gene controls self/nonself recognition was derived from a transgenic experiment in *Petunia* in the Solanaceae (Sijacic et al., 2004). In the study, they utilized "competitive interaction", in which pollen carrying two different *S* haplotypes is compatible in pistils with cognate *S* haplotypes, resulting in the breakdown of SI (de Nettancourt, 2001; Golz et al., 1999; 2001), to demonstrate its role as the pollen *S* determinant. When SLF from the S_2 haplotype was introduced into S_1S_2 and S_2S_3 plants, it caused the breakdown of SI in S_1 and S_3 pollen, but not in S_2 pollen, as predicted by competitive interaction. Thus, it was assumed that SLF proteins specifically detoxify nonself S-RNases, but not self S-RNases.

However, *SLF* was not the sole pollen *S* determinant in the Solanaceae. Kubo et al. (2010) showed that at least three types of divergent SLF proteins function as pollen *S* determinants, each recognizing a subset of nonself S-RNases, indicating that SLFs collaboratively function as pollen *S* determinants to detoxify nonself S-RNases. A simple mathematical model confirmed that 16–20 *SLF*s would be required to recognize the majority of target S-RNases (Kubo et al., 2015), which corresponds to the observed number of *SLF*s in the *Petunia* genome (Williams et al., 2014). The nonself recognition by multiple pollen *S* determinants is consistent with observations in Malinae. Kakui et al. (2011) showed that pollen with a loss-of-function in one of the multiple *SFBB*s in the S_4 haplotype of Japanese pear is rejected by pistils, not only with the cognate S_4 haplotype, but also with the S_1 haplotype, while it is accepted by other nonself pistils. Thus, multiple SFBBs also act collaboratively to detoxify nonself S-RNases in Malinae.

In contrast to the multiple pollen F-box genes in the Solanaceae and Malinae, *SFB* has been identified as the sole pollen *S* determinant in *Prunus*. SFBs control of self/nonself recognition was determined from the molecular characterization of SC mutants in *Prunus*, in which a loss-of-function mutation in *SFB* was associated with pollen-part SC (Ushijima et al., 2004; Sonneveld et al., 2005).

Molecular mechanisms of self/nonself recognition

Since the identification of specificity determinants, the molecular basis of the interactions between S-RNase and F-box, which trigger the acceptance or rejection of the pollen tube, were of considerable interest. An important finding was a microscopic

observation showing that both self and nonself S-RNases are incorporated into pollen tubes in the Solanaceae (Luu et al., 2000). Thus, the interaction and recognition between S-RNase and pollen *S* F-box proteins should occur inside the pollen tube, and only self S-RNase can elicit its cytotoxicity.

Self/nonself recognition, based on the degradation model, has been proposed from the characterization of pollen *S* determinants encoded by F-box gene. The Skp, Cullin, F-box containing (SCF) complex participates in protein degradation through the ubiquitin 26S proteasome pathway (Zhang et al., 2009). In the SCF complex, the F-box protein determines target specificity. In the degradation model, multiple pollen *S* F-box proteins recognize all nonself S-RNases for degradation, resulting in pollen tube growth in compatible crosses (Fig. 3A). Consistent with the degradation model, *Antirrhinum* SLF (AhSLF) physically interacts with nonself S-RNases and polyubiquitinates them, but not self S-RNases (Qiao et al., 2004). AhSLF was shown to interact with SSK1 and Cul1 proteins (Huang et al., 2006), suggesting that AhSLF forms SCF complexes and participates in the degradation of nonself S-RNases. In *Petunia*, SLF preferentially interacts with nonself S-RNases rather than self S-RNases (Hua and Kao, 2006), and the selective polyubiquitination of nonself S-RNases has been observed (Kubo et al., 2010). The degradation of nonself S-RNase is consistent with the competitive interaction observed in the Solanaceae, Plantaginaceae and Malinae in which hetero-diallelic pollen bearing two different types of *S* haplotypes can detoxify all of the S-RNases. However, this model is not compatible with *Prunus*.

<u>*Prunus*-specific self/nonself recognition mechanisms</u>

The degradation of nonself S-RNases by the pollen *S* F-box protein is not compatible with *Prunus* based on the observed phenotypes of the SC mutants. In *Prunus*, the loss-of-function of the *SFB* is associated with the SC phenotype, which provides strong evidence for the role of *SFB* as the pollen *S* determinant (Ushijima et al., 2004; Sonneveld et al., 2005). Because the pollen *S* F-box functions in the detoxification of nonself S-RNases, as proposed in other species, the loss-of-function of pollen *S* F-box is expected to lead to self- and cross-incompatibility. Consistently, no SC mutants conferred by the loss-of-function of pollen *S* F-box have been reported in species with S-RNase-based GSI, except for *Prunus* (De Franceschi et al., 2012; Tao

and Iezzoni, 2010). Golz et al. (2001) conducted a large-scale screening of SC mutants in the Solanaceae using mutagenized pollen and recovered SC mutants resulting from competitive interactions but not by mutations of the pollen *S* F-box gene, suggesting that pollen *S* F-box is essential for pollen tube growth in the Solanaceae.

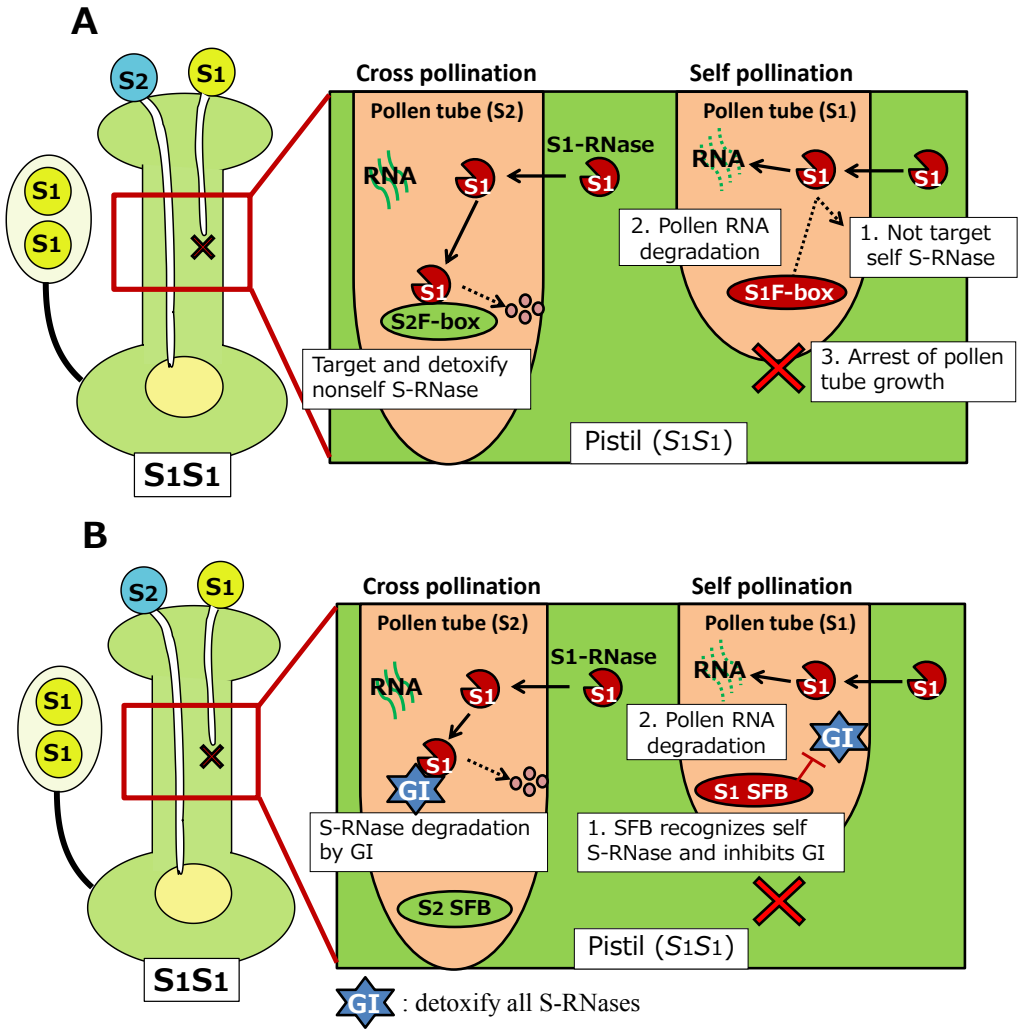

Fig. 3 Molecular mechanism for self/ nonself recognition in the Solanaceae and subtribe Malinae (A), and *Prunus* (B). (A) In cross-pollination, the pollen *S* F-box protein targets and detoxifies nonself S-RNases, resulting in pollen tube growth in the nonself pistil. In self-pollination, the pollen *S* F-box is unable to target self S-RNases, leading to pollen RNA degradation and the arrest of pollen tube growth. (B) In cross-pollination, S-RNases are targeted and degraded by a putative general inhibitor (GI), resulting in pollen tube growth in the nonself pistil. In self-pollination, pollen *S* F-box recognizes self S-RNases and inhibits GI to release the S-RNase cytotoxicity, leading to pollen RNA degradation and the arrest of pollen tube growth.

Additionally, a different outcome was also observed in the phenotype of hetero-diallelic pollen. In the Solanaceae, the polyploid-associated breakdown of GSI is caused by competitive interactions, in which pollen tubes containing two different *S* haplotypes grow normally, regardless of the *S* haplotype in the styles. In contrast, tetraploid sour cherry in *Prunus* exhibits a different phenotypic consequence because they retain SI when carrying two functional *S* haplotypes (Hauck et al., 2002; 2006a). Later, genetic and molecular studies indicated that SC mutations in sour cherry only arose by the accumulation of non-functional *S* haplotypes (Hauck et al., 2006b; Tsukamoto et al., 2006; 2008). Based on the distinct outcome from hetero-diallelic pollen in *Prunus*, the one-allele match model has been proposed to explain the self/nonself recognition in *Prunus* (Hauck et al., 2006b; Tao and Iezzoni, 2010). In this model, a match between the pollen *S* and its cognate pistil *S* causes the SI reaction, and it is plausible that competitive interactions are absent in *Prunus*. Thus, in *Prunus*, the self-interaction (self-recognition) between SFB and its cognate S-RNase would trigger an SI reaction, whereas in the Solanaceae, Plantaginaceae and Malinae, nonself-interactions (nonself-recognition) between SLF/SFBBs and S-RNases are involved in self/nonself recognition (Fig. 3A, B).

Because *Prunus* SFB acts to release the cytotoxicity of self S-RNases and is not involved in S-RNase detoxification, a putative general inhibitor (GI) that functions in detoxifying all of the S-RNases may be present in the pollen tubes of *Prunus* (Tao and Iezzoni, 2010; Matsumoto and Tao, 2016a). In this model, in cross-pollination (compatible cross), nonself S-RNases would be recognized and detoxified by GI through unknown mechanisms, resulting in pollen tube growth in the nonself pistils (Fig. 3B). In self-pollination (incompatible cross), SFB would recognize self S-RNases, protect them from detoxification by GI and allow them to exert RNase-activities, resulting in self-cytotoxicity and the arrest of pollen tube growth. Thus, *Prunus* SFB has a fundamentally different function from pollen *S* F-box in the Solanaceae, Plantaginaceae and Malinae.

4. Objectives of the study

The molecular basis of self/nonself recognition and resulting SI reactions in the S-RNase-based GSI have been extensively studied in the Solanaceae. Additionally, the

molecular mechanisms of SI have been investigated in the Rosaceae. However, the SI recognition system in *Prunus* is reported to be distinct from those of the Malinae and Solanaceae. Although the SC mutants of pollen *S* F-box genes resulted in clues to the distinct molecular function of SFB and the presence of a putative GI in *Prunus*, the evolutionally path generating *SFB* and the identity of the GI have yet to be elucidated. The characterization of the molecular mechanisms of, and evolutionally dynamics on, SI-determinant genes would be valuable for the future development of artificial SI control and for SC cultivar breeding in cultivated rosaceous fruit tree species.

This thesis consists of three chapters. Chapter 1 focuses on the evolutionary paths of the pollen *S* F-box genes in the Solanaceae, Plantaginaceae and Rosaceae, and provides evidence of the establishment of the *Prunus*-specific SI recognition system. Based on the indication in Chapter 1 that identify *Prunus* SLFLs (*S locus F-box gene with low allelic sequence polymorphism* or *S locus F-box-like*) as good GI candidates, a functional characterization of *Prunus* SLFLs was conducted in Chapter 2. Chapter 3 focuses on the evolutionary paths of the *S-RNase* genes by tracking their duplication patterns and shows *S* locus duplications in the ancestral genome of the Rosaceae, which may have caused the genus (or subtribe)-specific SI recognition mechanism in the Rosaceae. Based on these results, possible breeding strategies to produce SC cultivars in rosaceous fruit tree species are discussed.

Chapter 1. Insights into the *Prunus*-specific S-RNase-based self-incompatibility system from a genome-wide analysis of the evolutionary radiation of *S* locus-related F-box genes

1.1. Introduction

In S-RNase-based GSI, it is generally accepted that S-RNase exerts its cytotoxicity to degrade pollen tube RNA, resulting in the arrest of pollen tube growth. Pollen *S* F-box proteins are thought to target these cytotoxic S-RNases for degradation through the ubiquitin 26S proteasome pathways. In *Petunia*, in the Solanaceae, multiple *SLF* genes are present in tandem repeats that were generated by lineage-specific duplications (Kubo et al. 2010, Williams et al. 2014, Kubo et al. 2015). Multiple SLFs encoded by these genes collaboratively recognize and target nonself S-RNases for degradation, but are unable to target cognate S-RNases for degradation. It has been suggested that similar pathways involving the collaborative recognition and degradation of nonself S-RNases by multiple F-box proteins (SFBBs) are also present in the subtribe Malinae (Sassa et al. 2007, Kakui et al. 2011, Okada et al. 2011).

In contrast to the models proposed for the Solanaceae, and subtribe Malinae, the SFB protein from *Prunus* is thought to have a different function that acts through self- but not nonself-recognition. It is hypothesized that a putative GI in the pollen tube targets all of the S-RNases for degradation or inactivation. SFB, which encodes a different function from the functions in pollen *S* in the other species, recognize self S-RNases to protect them from inactivation by the GI to exert cytotoxicity (Ushijima et al. 2004, Tao and Iezzoni 2010, Matsumoto and Tao 2016a). It is suggested that the *Prunus SFB* may has a different lineage from those of *SLF/SFBB*s (Ushijima et al. 2003, Wang et al. 2004, Matsumoto et al. 2008, Aguiar et al. 2015) because *Prunus SFB* seems to be located far from the *SLF/SFBB* of the Solanaceae, Plantaginaceae and subtribe Malinae in a phylogenetic tree (Ushijima *et al*., 2003; Matsumoto *et al*., 2008). However, detailed phylogenetic analysis has yet to be conducted for pollen *S* F-box proteins of the S-RNase-based GSI.

Recent advances in genome sequencing have enabled the investigation of angiosperm genome evolution using genome-wide comparative analyses. According to the Phytozome database v9.1 (http://www.phytozome.net/), 41 sequenced and annotated

plant genomes, which have been clustered into 20 evolutionary significant nodes, are available. In this chapter, phylogenic and evolutionary analyses were conducted to retrace the duplication and functional diversification histories of the pollen *S* F-box genes. Based on the data obtained, establishment of a *Prunus*-specific self-recognition mechanism is discussed.

1.2. Materials and Methods

1.2.1. Gene collection

The putative full-length sequences of 408 genes with significant homology to *SFB*/*SLFL* genes were identified from the genomes of 31 angiosperms and two species outside the angiosperms, *Selaginella* and *Physcomitrella*, using the BLASTP algorithm in Phytozome (JGI release version 9.1, http://www.phytozome.net/) (Table 1.1). Genes having a significant homology (e^{-19} cutoff for all eudicots and e^{-10} cutoff for all others) to SFB (SFB-S_4; AB111521) or SLFL2 (SLFL2-S_4; AB280954) from *P. avium* were selected. These e-value cutoffs have been used to identify F-box genes from transcripts and amino acid sequences from a variety of plant species (Childs et al. 2007, Yang et al. 2008). Amino acid sequences with significant BLASTP hits were subjected to a gene ontology (GO) analysis by scanning the Pfam database (Finn et al., 2014) for signatures of the F-box domain (PF00646) and F-box-associated domains (PF04300, PF07734, PF07735 and PF08268), as described previously (Xu et al. 2009).

1.2.2. Construction of evolutionary topology

An amino acid sequence alignment was performed using MAFFT version 7 (Katoh and Standley, 2013). Custom Perl and Python scripts were used to adjust the amino acid alignments to correspond to the nucleotide alignments. Raw alignments were adjusted manually, using SeaView version 4 (Gouy et al. 2010), and then used for the assignment of the best available evolutionary model in jModeltest 2.1 (Darriba et al. 2012), using unequal base frequencies, a proportion of invariable sites, rate variation among sites and number of categories (-i -f -g 4 -s 11, parameters in jModeltest), based on both the Akaike and Bayesian information criteria.

Table 1.1 Gene list of *SFB/SLFL*-like F-box genes in angiosperms

Category[a]			Family	Species	Accessions[b]	Genome version[c]	Note/Reference
Embryophyto			Funariaceae	Physcomitrella patens	Pp1s67_164V6.1	v1.6	for out group
Tracheophyto			Selaginellaceae	Sellanginella moellendorfii	123864	v1.0	for out group
Angiosperm	Monocots		Poaceae	Brachypodium distachyon	Bradi5g07810.1	JGI v1.0 and MIPS/JGI v1.2	
				Oryza sativa	LOC_Os09g27570.1	v7.0	Derived from cv. Nipponbare
				Zea mays	GRMZM2G047777_T01	v2	Derived from cv. B73
				Sorghum bicolor	Sb04g005990.1	v1	
				Panicum virgatum	Pavirv00024414m	v2.6	
				Setaria italica	Si011696m	v2.1	
	Eudicots		Ranunculaceae	Aquilegia coerulea	Aquca_005_00044	v1.1	
					Aquca_033_00136		
					Aquca_020_00421		
					Aquca_004_00613		
					Aquca_004_00611		
					Aquca_028_00196		
					Aquca_037_00088		
					Aquca_037_00083		
					Aquca_049_00096		
					Aquca_002_00735		
					Aquca_002_00739		
					Aquca_002_01335		
					Aquca_002_01170		
					Aquca_002_00733		
					Aquca_055_00034		
					Aquca_055_00072		
					Aquca_022_00266		
					Aquca_016_00144		
	Core Eudicots	Euasterid I	Scrophulariaceae	Mimulus guttatas	mgv1a019337m	v1.0 and v1.1	Derived from IM62 inbred line
					mgv1a019037m		
					mgv1a018041m		
					mgv1a018794m		
					mgv1a025864m		
					mgv1a021964m		
					mgv1a022458m		
					mgv1a021233m		
					mgv1a007991m		
					mgv1a018074m		
					mgv1a020756m		
					mgv1a018184m		
					mgv1a018819m		
					mgv1a024128m		
					mgv1a024378m		
					mgv1a026138m		
			Solanaceae	Petunia hybrida	AB568390	NCBI database	PhSLF1-S5
					AB568394	NCBI database	PhSLF2-S5
					AB568399	NCBI database	PhSLF3-S5
					AB568405	NCBI database	PhSLF4-S5
					AB568411	NCBI database	PhSLF5-S5
					AB568417	NCBI database	PhSLF6-S5
				Solanum lycopersicum	Solyc10g008580.1.1	ITAG2.3 assembly	
					Solyc07g044920.1.1		
					Solyc01g067030.2.1		
					Solyc09g072930.1.1		
					Solyc04g063440.1.1		
					Solyc01g055160.1.1		
					Solyc01g057190.1.1		
					Solyc01g056280.2.1		
					Solyc01g057010.1.1		
					Solyc09g091710.1.1		
			Plantaginaceae	Antirrhinum hispanicum	AJ515535	NCBI database	AhisSLF-S1
		Rosid	Vitaceae	Vitis vinifera	GSVIVT01013506001	12X March 2010	
					GSVIVT01030585001		
			Myrtaceae	Eucalyptus grandis	Eucgr.I00158.1	v1.0	Derived from clone BRASUZ1
					Eucgr.I00164.1		
					Eucgr.C01780.1		
					Eucgr.D00965.1		
					Eucgr.D00965.2		
					Eucgr.E03976.1		
					Eucgr.H02696.1		
					Eucgr.H02715.1		
					Eucgr.H02731.1		
					Eucgr.H02734.1		
					Eucgr.H02735.1		
					Eucgr.H02929.1		
					Eucgr.I00068.1		
					Eucgr.I00070.1		
		Eurosid I	Cucurbitaceae	Cucumis sativus	Cucsa.108000.1	Roche/JGI v1	Derived from Gy14 gynoecious inbred line
					Cucsa.267690.1		
					Cucsa.307320.1		
					Cucsa.307340.1		
					Cucsa.307360.1		
					Cucsa.338770.1		
					Cucsa.338790.1		
					Cucsa.338810.1		
					Cucsa.378790.1		
					Cucsa.378910.1		
					Cucsa.378930.1		

Rosaceae	Prunus persica	ppa018725m	peach v1.0 by GDR	Derived from cv. Rovell
		ppa017733m		
		ppa021167m		
		ppa016785m		
		ppa017043m		
		ppa025214m		
		ppa011628m		
		ppa021936m		
		ppa024138m		
		ppa023942m		partial sequences in peach genome database
		ppa024149m		
		ppa021853m		
		ppa015105m		
		ppa022860m		
		ppa018335m		
		ppa005568m		
		ppa005203m		
		ppa021716m		SLFL1
		ppa027205m		
		ppa020444m		
		ppa006542m		
		ppa024694m		
		ppa025351m		
		ppa016207m		
		ppa025849m		SLFL2
		ppa019333m		
		ppa016317m		
		ppa023668m		
		ppa011628m		partial sequences in peach genome database
		ppa026586m		SLFL3
		JQ322648		PavFB from *P. avium*
		AB252410		PsSFBa from *P. salicina* [d]
	Pyrus pyrifolia	AB621597	NCBI database	PyrusSFBB1
		AB621601		PyrusSFBB2
		AB621608		PyrusSFBB4
		AB621615		PyrusSFBB6
		AB297934		PyrusSFBB8
	Malus x domestica	MDP0000257050	v1.0 by GDR	Derived from cv. Golden Delicious
		MDP0000188009		
		MDP0000130570		
		MDP0000239868		
		MDP0000124943		
		MDP0000314371		
		MDP0000214655		
		MDP0000778153		
		MDP0000532686		
		MDP0000608169		
		MDP0000393954		
		MDP0000203221		
		MDP0000642465		
		MDP0000598315		
		MDP0000890198		
		MDP0000279765		
		MDP0000311908		
		MDP0000433152		
		MDP0000291659		
		MDP0000811085		
		MDP0000157854		
		MDP0000264736		
		MDP0000136463		
		MDP0000302221		
		MDP0000775116		
		MDP0000321866		
		MDP0000252213		
		MDP0000210735		
		MDP0000214657		
		MDP0000122064		
		MDP0000309369		
		MDP0000129533		
		MDP0000266067		
		MDP0000250548		
		MDP0000312342		
		MDP0000150027		
		MDP0000314464		
		MDP0000749480		
		MDP0000203079		
		MDP0000260506		
		MDP0000454730		
		MDP0000668824		
		MDP0000273622		
		MDP0000612137		
		MDP0000237680		
		MDP0000283975		
		MDP0000500222		
		MDP0000195030		
		MDP0000294286		
		MDP0000302136		
		MDP0000250455		

		MDP0000305369		
		MDP0000306587		
		MDP0000482907		
		AB270793	DDBJ database	MdSFBB9-alpha
		AB270794	DDBJ database	MdSFBB9-beta
		AB270795	DDBJ database	MdSFBB3-alpha
		AB270796	DDBJ database	MdSFBB3-beta
	Fragaria vesca	mrna31388.1-v1.0-hybrid	v1.1	
		mrna31398.1-v1.0-hybrid		
		mrna31402.1-v1.0-hybrid		
		mrna31397.1-v1.0-hybrid		
		mrna31525.1-v1.0-hybrid		
		mrna31401.1-v1.0-hybrid		
		mrna31539.1-v1.0-hybrid		
		mrna31393.1-v1.0-hybrid		
		mrna31384.1-v1.0-hybrid		
		mrna31386.1-v1.0-hybrid		
		mrna31383.1-v1.0-hybrid		
		mrna13013.1-v1.0-hybrid		
		mrna26444.1-v1.0-hybrid		
		mrna31387.1-v1.0-hybrid		
		mrna14729.1-v1.0-hybrid		
		mrna13009.1-v1.0-hybrid		
		mrna31391.1-v1.0-hybrid		
		mrna31385.1-v1.0-hybrid		
		mrna31524.1-v1.0-hybrid		
		mrna04844.1-v1.0-hybrid		
		mrna21362.1-v1.0-hybrid		
		mrna00224.1-v1.0-hybrid		
		mrna22196.1-v1.0-hybrid		
		mrna22194.1-v1.0-hybrid		
		mrna24284.1-v1.0-hybrid		
		mrna22195.1-v1.0-hybrid		
		mrna09750.1-v1.0-hybrid		
		mrna00225.1-v1.0-hybrid		
		mrna24289.1-v1.0-hybrid		
		mrna24280.1-v1.0-hybrid		
		mrna22201.1-v1.0-hybrid		
		mrna22203.1-v1.0-hybrid		
		mrna22199.1-v1.0-hybrid		
		mrna01639.1-v1.0-hybrid		
		mrna10599.1-v1.0-hybrid		
		mrna31914.1-v1.0-hybrid		
		mrna25137.1-v1.0-hybrid		
		mrna25136.1-v1.0-hybrid		
		mrna31349.1-v1.0-hybrid		
		mrna26892.1-v1.0-hybrid		
		mrna22691.1-v1.0-hybrid		
		mrna22693.1-v1.0-hybrid		
		mrna06423.1-v1.0-hybrid		
		mrna31347.1-v1.0-hybrid		
		mrna32334.1-v1.0-hybrid		
		mrna27192.1-v1.0-hybrid		
		mrna00156.1-v1.0-hybrid		
		mrna10433.1-v1.0-hybrid		
		mrna25301.1-v1.0-hybrid		
		mrna19616.1-v1.0-hybrid		
		mrna27211.1-v1.0-hybrid		
		mrna28084.1-v1.0-hybrid		
		mrna19775.1-v1.0-hybrid		
		mrna15870.1-v1.0-hybrid		
		mrna21884.1-v1.0-hybrid		
		mrna07530.1-v1.0-hybrid		
		mrna24040.1-v1.0-hybrid		
		mrna06911.1-v1.0-hybrid		
		mrna07341.1-v1.0-hybrid		
		mrna07389.1-v1.0-hybrid		
		mrna00161.1-v1.0-hybrid		
		mrna03045.1-v1.0-hybrid		
		mrna00155.1-v1.0-hybrid		
		mrna27186.1-v1.0-hybrid		
		mrna30709.1-v1.0-hybrid		
Fabaceae	Medicago truncatula	Medtr3g014540.1	Mt3.0	
		Medtr2g021770.1		
		Medtr2g034700.1		
		Medtr6g079810.1		
		Medtr5g037680.1		
		Medtr8g063440.1		
		Medtr8g046310.1		
		Medtr3g014540.1		
		Medtr2g021770.1		
		Medtr2g034700.1		
		Medtr6g079810.1		
		Medtr5g037680.1		
		Medtr8g063440.1		
		Medtr8g046310.1		
		AC233662_36.1		
		AC233685_66.1		
		Medtr4g061340.1		
		Medtr4g133170.1		

		Glycine max	Glyma19g23910.1	v1.1	
			Glyma15g12190.4		
			Glyma09g01330.4		
			Glyma16g06890.1		
			Glyma16g06880.2		
		Phaseolus vulgaris	Phvul.003G084100.1	v1.0	
			Phvul.008G148100.1		
	Euphorbiaceae	Manihot esculenta	cassava4.1_031610m	Cassava4	
			cassava4.1_034084m		
			cassava4.1_031818m		
			cassava4.1_027730m		
			cassava4.1_024381m		
			cassava4.1_010406m		
			cassava4.1_008049m		
			cassava4.1_028851m		
			cassava4.1_009366m		
			cassava4.1_030908m		
			cassava4.1_024183m		
			cassava4.1_008606m		
		Ricinus communis	28152.m000886	TIGR/JCVI v0.1	
			29806.m000946		
			29206.m000142		
			29206.m000146		
			29904.m002893		
			30131.m007064		
	Linaceae	Linum usitatissimum	Lus10027824	BGI v1.0	
			Lus10013873		
			Lus10026587		
	Malpighiaceae	Populus trichocarpa	Potri.017G058000.1	v3	
			Potri.011G121200.1		
			Potri.017G011100.1		
			Potri.017G012200.1		
			Potri.006G013000.1		
			Potri.006G013200.1		
			Potri.006G170300.1		
			Potri.006G011900.1		
			Potri.015G013500.1		
			Potri.T135000.1		
			Potri.005G114900.1		
			Potri.001G035500.1		
			Potri.012G099800.1		
Eurosid II	Brassicaceae	Arabidopsis thaliana	AT3G06240	TAIR release 10	FBA1 domain
		Arabidopsis lyrata	Araly_928248	JGI v1.0	
			Araly_477924		
		Capsella rubella	Carubv10015426m	initial 22x	
		Thellungiella halophila	Thhalv10020779m	initial 8X mapped	
	Caricaceae	Carica papaya	evm.TU.supercontig_791.2	version 1.1	
			evm.TU.supercontig_370.6		
			evm.TU.supercontig_51.32		
			evm.TU.supercontig_75.108		
			evm.TU.contig_27218.3		
	Rutaceae	Citrus clementina	Ciclev10003888m	v1.0	
			Ciclev10003618m		
			Ciclev10003412m		
			Ciclev10003993m		
			Ciclev10006809m		
			Ciclev10013838m		
			Ciclev10011950m		
			Ciclev10011877m		
			Ciclev10011921m		
			Ciclev10012020m		
			Ciclev10027392m		
			Ciclev10027434m		
			Ciclev10027518m		
			Ciclev10026927m		
			Ciclev10027421m		
			Ciclev10027079m		
			Ciclev10027004m		
			Ciclev10027528m		
			Ciclev10027251m		
			Ciclev10027534m		
			Ciclev10027082m		
			Ciclev10033474m		
			Ciclev10033971m		
			Ciclev10033888m		
			Ciclev10023509m		
			Ciclev10028647m		
			Ciclev10030186m		
			Ciclev10017792m		
		Citrus sinensis	orange1.1g038188m	v1.0	
			orange1.1g039590m		
			orange1.1g041840m		
			orange1.1g042926m		
			orange1.1g036455m		
			orange1.1g045569m		

		orange1.1g036281m	
		orange1.1g048178m	
		orange1.1g041631m	
		orange1.1g044822m	
		orange1.1g048294m	
		orange1.1g038747m	
		orange1.1g037345m	
		orange1.1g016751m	
		orange1.1g046902m	
		orange1.1g046787m	
		orange1.1g017748m	
		orange1.1g016891m	
		orange1.1g036467m	
		orange1.1g044435m	
		orange1.1g045675m	
		orange1.1g042454m	
		orange1.1g037087m	
		orange1.1g041235m	
		orange1.1g039662m	
		orange1.1g015535m	
		orange1.1g045735m	
		orange1.1g036621m	
		orange1.1g016752m	
		orange1.1g016057m	
		orange1.1g016018m	
		orange1.1g016291m	
Malvaceae	Gossypium raimondii	Gorai.010G186900.4	v2.1
		Gorai.010G187100.2	
		Gorai.004G021600.1	
		Gorai.004G022100.1	
		Gorai.011G276100.1	
		Gorai.011G276500.1	
		Gorai.011G136300.1	
		Gorai.011G179400.1	
		Gorai.011G135900.1	
		Gorai.011G281200.1	
		Gorai.011G276700.1	
		Gorai.011G281100.1	
		Gorai.011G274400.1	
		Gorai.008G098100.1	
		Gorai.012G050800.1	
	Theobroma cacao	Thecc1EG031361t3	V1.1
		Thecc1EG026354t1	
		Thecc1EG041351t1	
		Thecc1EG001772t1	
		Thecc1EG030371t1	
		Thecc1EG030369t1	
		Thecc1EG046230t1	
		Thecc1EG045362t1	
		Thecc1EG045366t3	
		Thecc1EG045350t1	
		Thecc1EG045356t1	
		Thecc1EG045363t1	
		Thecc1EG045348t1	
		Thecc1EG045352t1	
		Thecc1EG043439t1	
		Thecc1EG045358t1	
		Thecc1EG015405t1	
		Thecc1EG018681t1	
		Thecc1EG018682t1	
		Thecc1EG020306t1	
		Thecc1EG046272t1	

a Categorization according to Angiosperm Phylogeny Group III. **b** The accessions used as transcript names in Phytozome v9.1 (http://www.phytozome.net/) are indicated, except for genes that did not originate from whole genome reference sequences. For the six *SLF* genes from *Petunia*, one *SLF* from *Antirrhinum*, five *SFBB*s from *Pyrus*, and four *SFBB*s from *Malus*, accession numbers from the NCBI/DDBJ databases are indicated. **c** For genes that did not originate from whole genome reference sequences present in Phytozome v9.1, the source databases are indicated. **d** The *SFB* gene from the peach genome is disrupted (Tao et al., 2007).

Nucleotide alignments of the *SFB/SLFL*-like genes from the *Prunus* genome and two Rosaceae-specific subclades α and ε were used to construct phylogenies. In jModeltest, the general-time reversible (GTR) + gamma + proportion of invariable sites model was selected for the *SFB/SLFL*-like genes from the *Prunus* genome; whereas transversion model 1 and a three-parameter model, with unequal base frequencies + gamma + proportion of invariable sites, were adopted for the analysis of the two Rosaceae-specific subclades α and ε, respectively. The three-parameter model is a submodel of the GTR, and the substitution pattern of the transversion model (numbers of substitutions = 4) is not implemented in MrBayes 3.1.2 (Ronquist and Huelsenbeck 2003). Thus, in this study, the GTR with a site-specific model, which is considered a codon model, was used to construct the evolutionary tree. The optimal tree and estimation of branch support were determined using a Bayesian Markov chain Monte Carlo (MCMC) approach using MrBayes. MCMC analyses used flat priors and were run for 3,000,000–10,000,000 generations and four Markov chains (using the default heating values), and sampled every 1,000 generations. The chains converged on a stable set of parameters were inferred by calculating the potential scale reduction factor using MrBayes.

For the alignment of amino acid sequences, MAFFT with the L-INS-i model and SeaView version 4 were used as described above. In all of the alignments, unnecessary long gap sequences disturbing proper alignment construction and genes showing significant homology, but with apparently different structures from other *SLFL/SFB*-like genes, were removed. The Neighbor–joining approach was applied to the resulting alignment file for a phylogenetic analysis, using MEGA v5.05 (Tamura et al., 2011) with 1,000 bootstrap replications.

1.2.3. Estimation of the time of gene divergence

The amino acid sequences of each pair of homologs were aligned using MAFFT with the L-INS-i model, and the amino acid alignments were converted to nucleotide alignments using PAL2NAL (Suyama et al., 2006). The transversion rate at four-fold-degenerate sites (4DTv) between gene pairs was calculated. Here, 4DTv are third codons of amino acid residues G, A, T, P, V, R, S and L. 4DTv values were calculated for F-box gene pairs that retained at least eight four-fold synonymous sites. For calculation of 4DTv values from *SLFL* genes to a specific gene group (i.e. *SLF* genes in *Petunia*), the 4DTv values of each pairwise gene combination between the two groups were averaged. For the estimation of 4DTv values between two species (or genera), values of the putatively orthologous *SFB/SLFL*-like F-box gene sets between

the two taxa were averaged.

1.2.4. Gene expression analysis

Mature pollen grains of *P. avium* cv. Satonishiki, and *P. mume* cvs. Nanko and Kairyo-uchida-ume were collected from anthers sampled from flower buds at the balloon stage of development in the experimental orchard of Kyoto University. Total RNA was extracted using a cold-phenol extraction method, as previously reported (Tao et al. 1999). All procedures for sample preparation were conducted using TruSeq RNA Sample Prep Kits (Illumina). Sequencing was performed using the TruSeq SBS Kit v3-HS on the Illumina HiSeq 2000 (Illumina) as paired-end reads of 100 bp.

One-fifth of a sequencing lane was dedicated to each of the three cultivars, and 29,812,682, 35,162,899 and 32,728,911 informative reads were obtained for 'Satonishiki', 'Kairyo-uchida-ume' and 'Nanko', respectively. All of the bioinformatics and statistical analyses were performed on local servers at the UC Davis Genome Center. Raw reads without adaptor sequences were subjected to trimming procedures by using custom Python scripts (http://comailab.genomecenter.ucdavis.edu/index.php/Barcoded_data_preparation_tools). The reads were trimmed at the first point where nucleotide sequences showed a Phred quality score of < 20 over mean sliding windows of 5 bp. The paired-end reads with either end showing a length < 35 bp were removed. Reads were then mapped to the 28,689 predicted mRNA transcript sequences in the *P. persica* genome database (from the GDR, Prunus_persica_v1.0) using the Burrows–Wheeler Aligner tools (Li and Durbin 2009), either using default parameters or allowing up to 8 bp of mismatches per read. The number of reads mapping to each gene were calculated from the alignment file produced by the Sequence Alignment/Map (SAM) tool (Li et al., 2009). The expression levels were normalized as reads per kilobase per million reads (RPKM).

1.2.5. Identification of selective pressure on the F-box genes

For *SFB* and the three *SLFL*s (Ushijima et al., 2003), full-length allele sequences were obtained from the NCBI or Phytozome databases or by the sequencing of PCR products, mainly in *P. avium* and *P. mume* cultivars, that showed different allele types. All of the primers were designed outside the open reading frame region in each locus, using sequences from the peach genome database (GDR, Prunus_persica_v.1.0) (Table 1.2). PCR products were sequenced after cloning into the pGEM-T vector (Promega).

Table 1.2 Primers used for RT-PCR, and amplifications of full-length alleles

Experiment	Target genes	Primer names	Primer sequences
RT-PCR in *P.avium* (cv. Satonishiki)	SLFL3	P.av-SLFL3-F	AGCCAAGCACCTTCACTTGT
		P.av-SLFL3-R	TCAGCATTGCTTTCGTTGTC
	ppa019333m	P.av-19333-F	ACTTCAGTTCCACTCGACAC
		P.av-19333-R	AAACGGGACGGAGGTGAT
	ppa021167m	P.av-21167-F	TGTTGACGATCCATTTGTGA
		P.av-21167-R	AAATGCAAAGTACCCCATTG
	ppa023668m	P.av-23668-F	CATCGATCCCCATAATGGAC
		P.av-23668-R	TTTCGGTGCGGAGGTTATAG
	ppa024694m	P.av-24694-F	CATTTTATGGTTCGCAATGG
		P.av-24694-R	ATTTTCAGGCCAAAGGAAAG
RT-PCR in *P.mume* (cv. Nanko)	SLFL3	P.m-SLFL3-F	CCTCTTCAAGCGTTCTGTCC
		P.m-SLFL3-R	CCCATAGAACGTGGGAAATG
	ppa019333m	P.m-19333-F	GAATTCTTTCGTTCGACTTTGGTG
		P.m-19333-R	GCAACGCTTTCCTTCCACAC
	ppa021167m	P.m-21167-F	TGTTGACGACCCTTTTGTGA
		P.m-21167-R	AAATGCAAAGTACCCCGTTG
	ppa023668m	P.m-23668-F	CATCGATCCCCACAATGGAC
		P.m-23668-R	TTTCGGTGCGGAGGTTATAG
	ppa024694m	P.m-24694-F	CAAGTGTACACGCTGGGAAG
		P.m-24694-R	CCTCATCCACGAATTCCTTT
RT-PCR in *P.avium* and *P.mume*	SFB	P.av.m-SFB-F	AGGACCACTCCAATGAGCAC
		P.av.m-SFB-R	AGGCATTTTTGTTGGTACGC
	SLFL1	P.av.m-SLFL1-F	CCAAATCAGGCCTTCCAGAT
		P.av.m-SLFL1-R	ATGTCCGCCTGATAAGATGC
	SLFL2	P.av.m-SLFL2-F	AGAATCGAACCCGAAATGTG
		P.av.m-SLFL2-R	GCAGCCAAGCTGTAAACCTC
	ppa023942m	P.av.m-23942-F	CCATCCCTTGGGGATCAAGA
		P.av.m-23942-R	ATTTCCTAACCGAGGGGTTC
	ppa005507m (Ubiquitin)	P.av.m-ubiquitin-F	TGATCCTTGTGGTTCCATCC
		P.av.m-ubiquitin-R	CATCCATCAGCCAAGTACGA
Amplification of full-length alleles	SFB[a]	PruSFB-1F	CTACGTAAGAAAGAAATCTTAATCGACAT
		PruSFB-2F	GTTCGGTTTCTTTGTACATGCAA
		PruSFB-1R	CCAAACTTTCTATGTAAGTAATTGCAAAC
		PruSFB-2R	AAAATAAAAGGACACAATTGTTTCC
	SLFL1	PruSLFL1-F	ACAACGTTTTGAGATTTGGGATTAC
		PruSLFL1-R	GAGAAATTATATTGCTGGAATTATGTATCAA
	SLFL2	PruSLFL2-F	TATAGAGGAAACAAGCACAGAGGACG
		PruSLFL2-R	CTATTACGATTGATTCGATCGTTGTAA
	SLFL3	PruSLFL3-F	ATTGTTTTCCACCTACATTGTCAGC
		PruSLFL3-R	GAACTTAGGCTCTTAGATCCTTTAGCC

a The full-length *SFB* genes were amplified and sequenced using four primer sets, including two forward and two reverse primers because of the high sequence diversity in *SFB* alleles from *P. avium* and *P. mume* cultivars.

Amino acid sequence alignments were constructed with MAFFT and converted to nucleotide alignments. Aligned nucleotide sequences were used to determine putative ancestral sequences with MEGA v5.05. To calculate the selective pressure in *SLF*/*SFBB* genes in *Petunia* and the subtribe Malinae, sequences from databases (NCBI, http://www.ncbi.nlm.nih.gov/; and DDBJ, http://www.ddbj.nig.ac.jp/index-e.html) were aligned. Informative single-nucleotide polymorphisms in alleles were analyzed by DnaSP 5.1 (Librado and Rozas, 2009) and used to calculate an indicator of selective pressure, the nonsynonymous to synonymous substitution (Ka/Ks) ratio. Window-averaged Ks/Ks values were calculated from the start codon (ATG) in a 150-bp window with a 30-bp step size in DnaSP 5.1, until the walking window reached the stop codon.

1.3. Results

1.3.1. Duplications and evolutionary patterns of the *Prunus SFB/SLFL*-like F-box genes

1.3.1.1. *Gene collection and phylogenetic analysis*

Prunus species carry two kinds of F-box genes around the *S* locus: a single pollen *S* determinant, *SFB*, and at least three *SLF-like* genes (*SLFL1–SLFL3*) (Ushijima et al. 2003, Entani et al. 2003, Matsumoto et al. 2008). The functions of the SLFL genes remains to be determined, although they show high sequence similarities to *SLF/SFBB* genes and pollen expression (Ushijima et al. 2004, Matsumoto et al. 2012). To obtain a broad phylogenetic context for the F-box genes across angiosperms, comprehensive searches for genes showing significant homology to the *Prunus SFB/SLFL* genes were performed using the BLAST algorithm and a GO analysis. The search was performed in three genomes from the Rosaceae: peach (*P. persica*, *Prunus*), diploid wild strawberry (*Fragaria vesca*, Rosoideae) and apple (*M.* × *domestica*, Malinae), as well as 28 other angiosperm species. Along with the F-box genes previously identified as potential pollen *S* determinants from the Solanaceae, Plantaginaceae and subtribe Malinae, 408 other genes were identified (Table 1.1).

Bayesian MCMC phylogenetic analyses of *SLF* genes from *Petunia* and *Antirrhinum*, *SFBB* genes from the subtribe Malinae and 29 *SFB/SLFL*-like F-box genes identified from the peach genome (all with one gene per locus) revealed three main clades, S, A and B, each with significant statistical support (Fig. 1.1), and strongly supporting monophyly for *SLF*, *SFBB*, the three *Prunus SLFL* genes and five other peach *SLFL* homologs (ppa023668m, ppa024694m, ppa019333m, ppa016207m and ppa016317m) (Clade S, Bayesian posterior probability > 0.95). The *Prunus SFB* gene did not cluster with genes in the *SLF/SFBB/SLFL* clade, suggesting that it was derived from another F-box gene family (Fig. 1.1). Similar results were obtained after a phylogenetic analysis that included all of the sequences obtained from the 31 angiosperm species (Fig. 1.2A shows the tree topology). A single *SFB/SLFL*-like F-box gene was present in the MRCA of angiosperms (monocots and eudicots, blue circle in Fig. 1.2A), according to the Angiosperm Phylogeny Group III (APGIII) classification (The Angiosperm Phylogeny Group. 2009). In contrast, obtained results indicated the presence of some *SFB/SLFL*-like F-box genes in the MRCA of eudicots (red circles in Fig. 1.2A), one of which corresponded to the divergence of Clade S, suggesting that *SFB* and *SLF/SFBB/SLFL* were derived from different genes already present in the eudicot's ancestor (Fig. 1.2A). The divergence pattern of the Clade S genes was consistent with the APGIII classification. Thus, the functions of SLF, SFBB and SLFL in the S-RNase-based GSI system may have been derived from a common original

function, whereas *Prunus* SFB has a distinct evolutionary path.

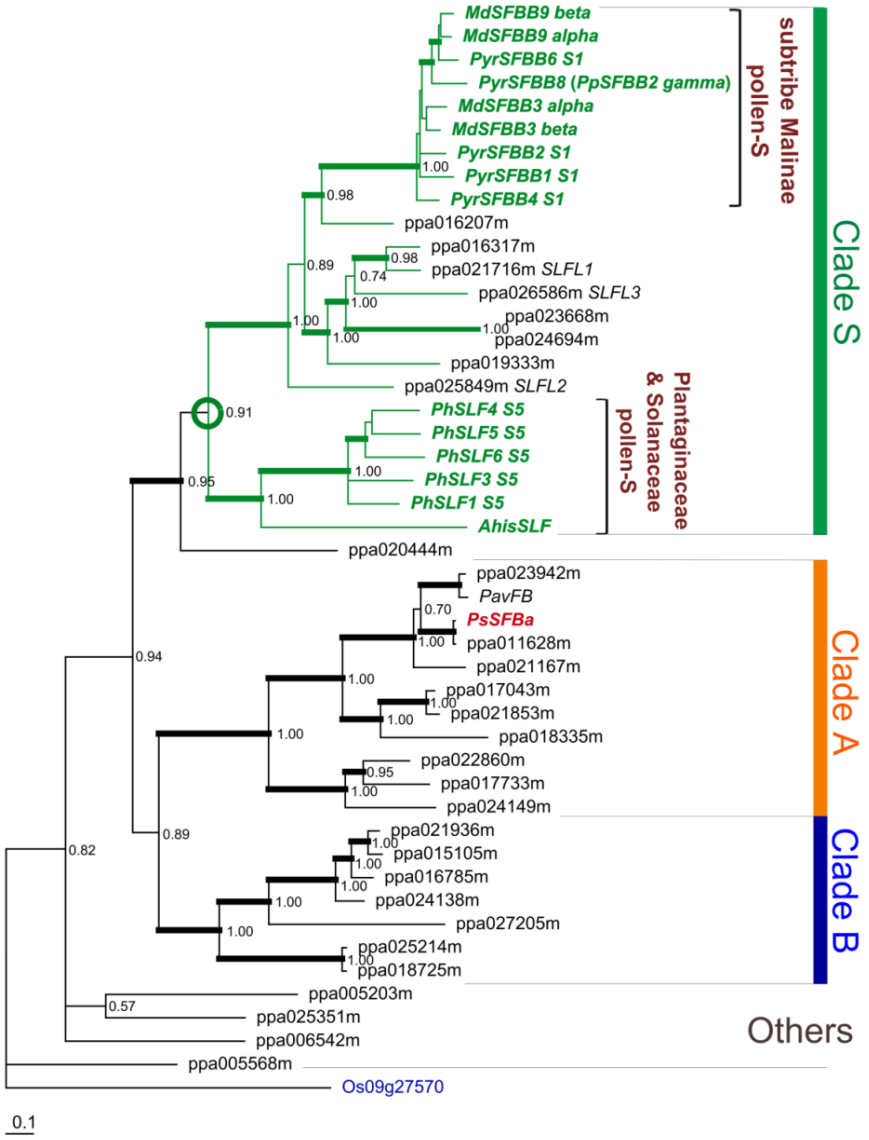

Fig. 1.1 Phylogenetic tree of *SFB/SLFL*-like F-box genes from the *Prunus persica* genome. Bayesian MCMC phylogenetic tree for *SFB/ SLFL*-like F-box genes from *P. persica* (prefixes of "ppa") along with pollen *S* F-box genes from *Antirrhinum*, *Petunia* and subtribe Malinae. The position of the most recent common ancestor of the original pollen *S* is indicated by a green circle. The pollen *S* F-box gene in *Prunus*, *SFB*, is highlighted in red. Thick branches indicate clades with strong statistical support (Bayesian posterior probability > 0.95).

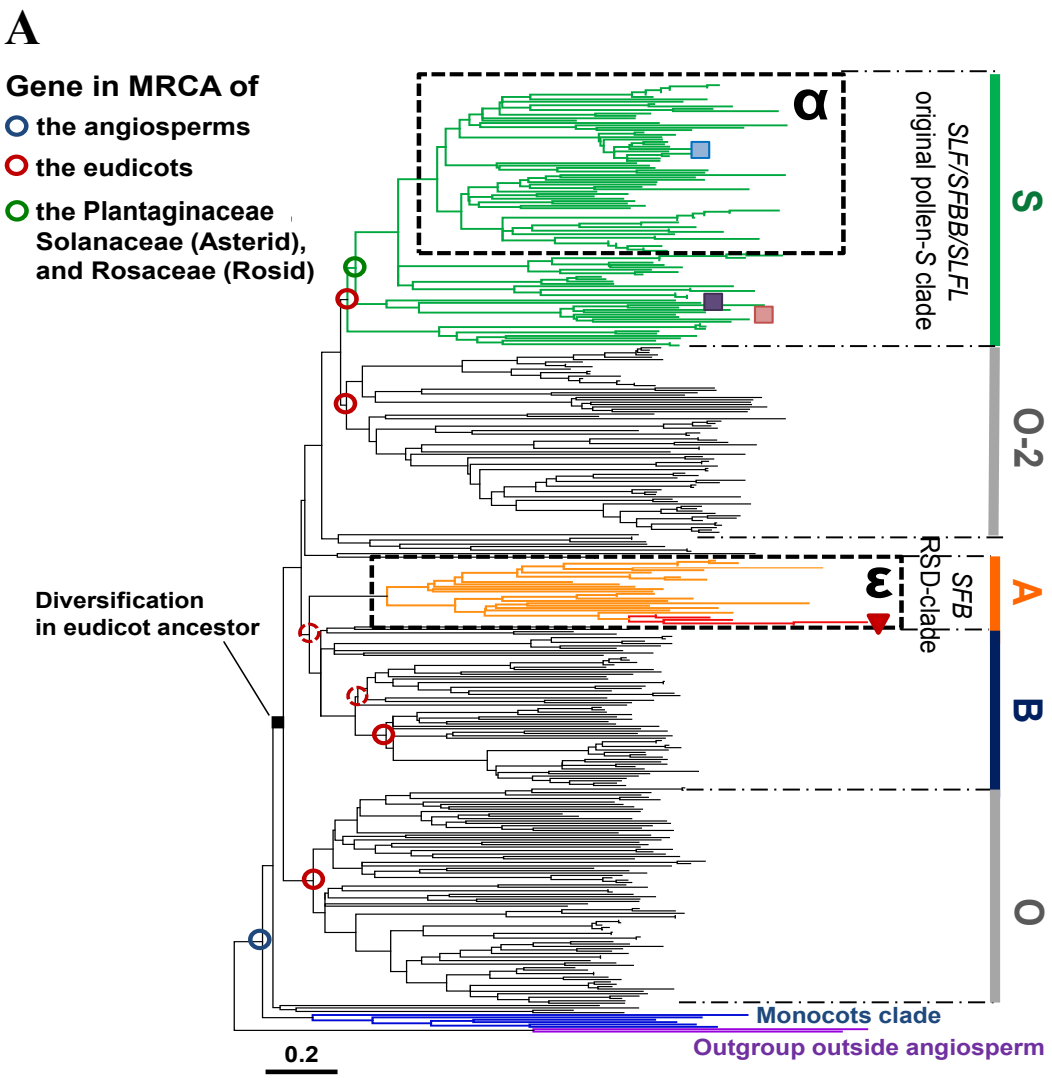

Fig. 1.2 Phylogenetic trees of *SFB/SLFL*-like F-box genes from angiosperm species. (A) Phylogenetic trees of *SFB/SLFL*-like F-box genes from 31 angiosperm genomes based on amino acid sequences and the NJ method. The pollen *S* F-box genes from *Antirrhinum*, *Petunia* and Malinae are shown in purple, pink and blue squares, respectively. The pollen *S* F-box gene from *Prunus* is indicated by a red triangle. The clade containing a Rosaceae-specific duplication is colored in orange, and within the clade, the *Prunus*-specific duplication is shown in red. The genes in the most recent common ancestor of the angiosperms, the eudicots, and the species having S-RNase-based GSI are indicated with blue, red and green circles, respectively.

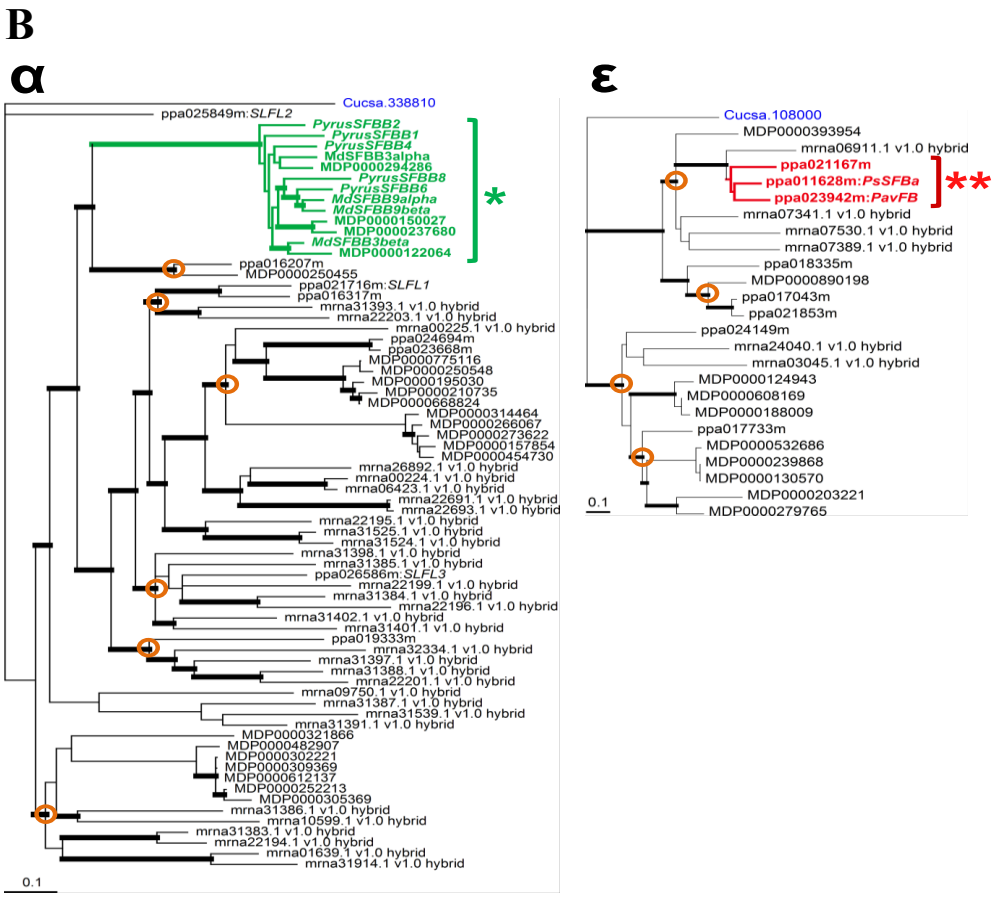

Fig. 1.2 (Continue) (B) Bayesian MCMC phylogenetic tree for Rosaceae-specific subclades α and ε, as indicated in Fig. 1.2A. The closest orthologs in the *Cucumis* genome were used as outgroups. The SFBB branches and the *Prunus*-specific duplication branches are shown in green with a single asterisk and red with double asterisks, respectively. The putative genes in the most recent common ancestor of the Rosaceae are indicated by orange circles and include at least two species from *Fragaria*, *Malus* and *Prunus*. Thick branches indicate clades with strong statistical support (Bayesian posterior probability of > 0.95).

This analysis also suggested drastic duplications of the F-box genes in the Rosaceae lineage in Clades S and A (Fig. 1.2B). Two Rosaceae-specific subclades, α and ε, corresponding to clades S and A, respectively, were defined (Fig. 1.2B). The topology of subclade α, which includes the *SLFL/SFBB* genes, was statistically well supported and suggested that the *SFBB* genes in the subtribe Malinae were established by a lineage-specific gene duplication and that they do not share orthologs in *Prunus* and *Fragaria* (Fig. 1.2B). Such lineage-specific gene duplications were also observed

for other genes in the Rosaceae lineage in subclades α and ε (Fig. 1.2B). In subclade ε, it further indicates that *SFB* was derived from *Prunus*-specific duplication (PSD) events, as no true ortholog was detected for *Prunus SFB* in *Malus* or *Fragaria* (Fig. 1.2B; red branches in Fig. 1.2A). Subclade α and ε also contained multiple paralogs in the Rosaceae. This finding suggests multiple Rosaceae-specific duplication events (Fig. 1.2B; shown in orange) prior to the diversification of *Fragaria*, *Malus* and *Prunus* in subclades α and ε.

1.3.1.2. *Genomic organization*

An analysis of the genomic organization of the F-box genes also indicated their evolutionary path (Fig. 1.3). The previous comprehensive exploration of F-box genes, in *Arabidopsis* and *Oryza* genomes, suggested that 36% and 43%, respectively, of this gene family had been generated from tandem duplications (Yang et al. 2008). This characteristic is thought to be conserved widely in plant genomes (Gagne et al. 2002, Jain et al. 2007, Yang et al. 2008, Xu et al. 2009). Thus, most F-box genes in Clade S, including the *SLFL* genes, were located in tandem at the end of chromosome (chr) 6 in an approximately 200-kb region of the peach genome that also contains the *S* locus (Fig. 1.3). The F-box genes from Clade A, including those in the PSD-containing clade (Fig. 1.2A, B, shown in red), are also organized in tandem but are located on chr 3, except for ppa018335m (chr 1), ppa024149m (chr 2) and the *SFB* gene at the *S* locus on chr 6 (Fig. 1.3).

1.3.1.3. *Estimation of the time of gene divergence*

The pairwise nucleotide diversity between the *SFB/SLFL*-like F-box genes was calculated to estimate their divergence times. The average pairwise 4DTv values between the *SLFL* genes from *Prunus* and the *SFBB* genes from the subtribe Malinae (0.271 ± 0.006) or between the *SLFL* and *SLF* genes from *Petunia* (0.366 ± 0.006) were significantly lower than those between the *SLFL* genes and the *SFB* from *Prunus* (0.417 ± 0.015) (Fig. 1.4A). In fact, the 4DTv values between the *SLFL* genes and the *Prunus SFB* genes were similar to the 4DTv values between *SLFL* genes and the single closest F-box gene from monocots (0.439 ± 0.031), which was used as the outgroup gene in the phylogenetic analysis. These results were consistent with the topology obtained (Fig. 1.2A) and suggested that the *SFB* (Clade A) and *SLFL* (Clade S) genes diverged shortly after the establishment of the eudicots.

The average 4DTv values between the *SFBB* genes in the subtribe Malinae and any *SLFL* genes in *Prunus* (0.271 ± 0.006) were higher than those among the closest

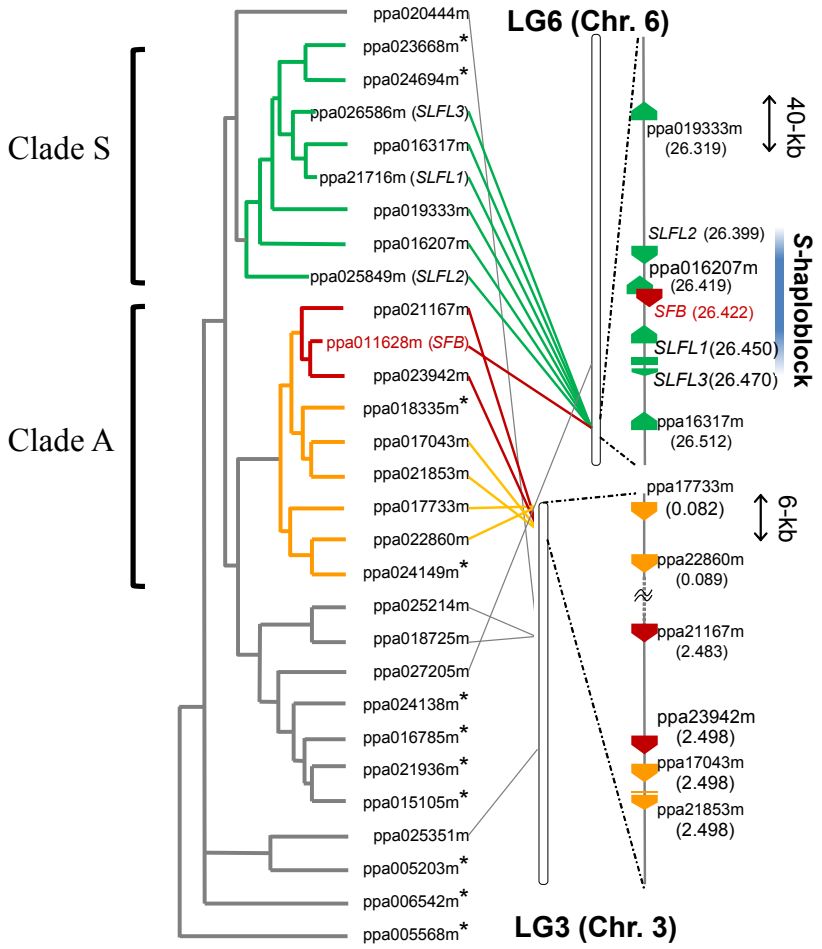

Fig. 1.3 Relationship between the genomic location and phylogeny of *SFB/SLFL*-like F-box genes from *Prunus persica*. The genomic position (Mb) of the genes located on chr 3 and 6 are indicated in parentheses and gene orientations are indicated by arrowheads. The genes not located on either chr 3 or 6 are marked by asterisks.

gene sets in *Prunus* and *Malus* (0.146 ± 0.013) (Fig. 1.4A). Thus, the *SFBB* and *SLFL* genes diverged from each other before the divergence of *Prunus* and *Malus*, and they are not true orthologs. Furthermore, most of the pairwise 4DTv values between the *SLFL* genes were significantly greater than those between the closest gene sets in *Prunus* and the other two rosaceous species (Fig. 1.4B), indicating that the tandem duplication that gave rise to the *SLFL* genes occurred before the diversification of the three rosaceous species. Also, the pairwise 4DTv values suggested that the tandemly repeated *SFBB* genes were lineage specific and were established after the divergence of *Prunus* and the subtribe Malinae (Fig. 1.4B).

The pairwise 4DTv values in clade A indicated that most of the gene duplication events in this clade occurred after the divergence between the Rosales and

Cucurbitales. The gene duplication events that generated *SFB* in the PSD clade post-dated the divergence of *Prunus* from the other rosaceous species (Fig. 1.4C; the asterisk indicates the putative timing of the duplication event that generated *SFB*).

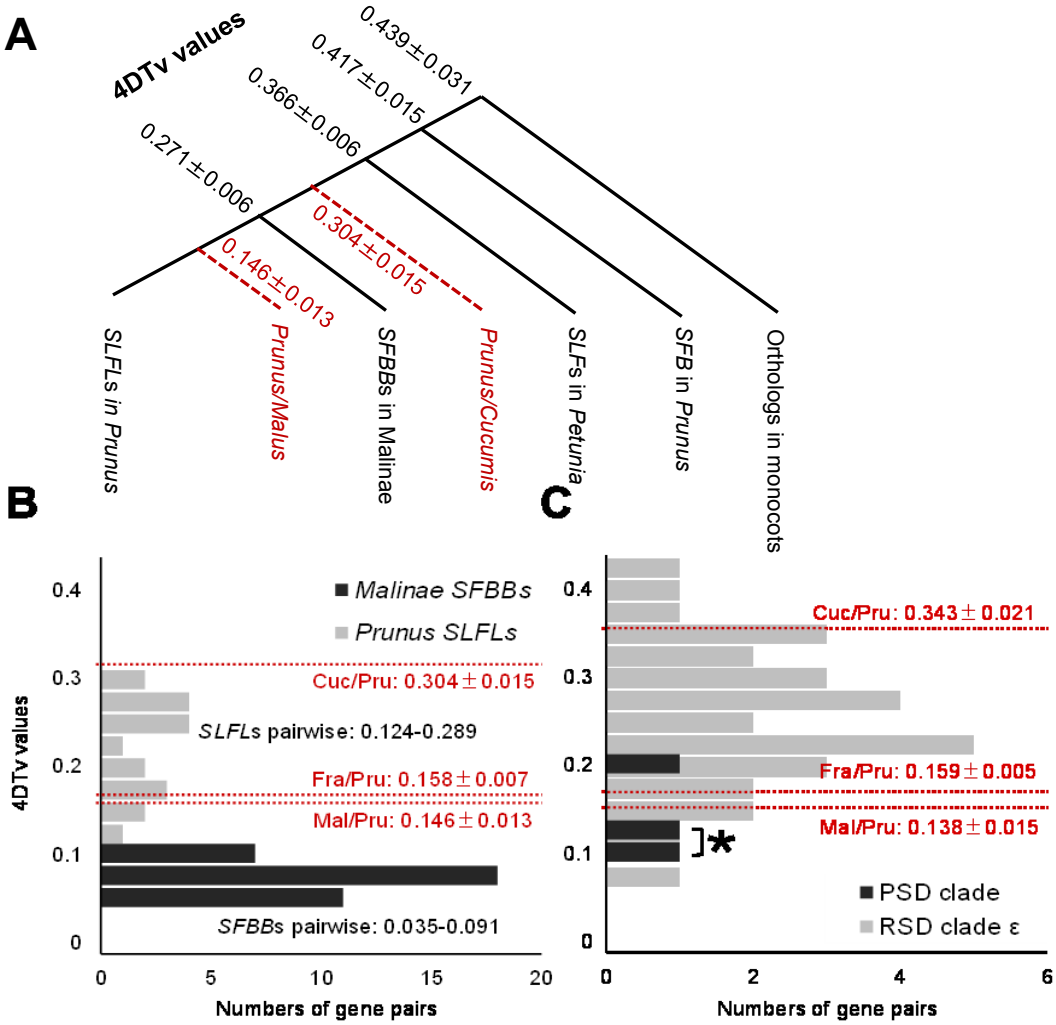

Fig. 1.4 Estimations of the gene divergence times. (A) The average values of the 4DTv distance between two gene groups are shown in black. The values for pairs of species are shown in red and were calculated using the average values of the closest gene sets between the two species across all *SFB/SLFL*-like genes. All 4DTv values are given with standard errors. (B) Pairwise 4DTv values between tandemly duplicated genes in SFBB and SLFL clusters. *Prunus SLFL* genes are represented by six genes located around the *Prunus S* locus, except for *SFB*. 4DTv values for pairs of species in this gene family are indicated by dotted lines. (C) Pairwise 4DTv values for genes within the Rosaceae-specific duplication (clade ε) and subsequent *Prunus*-specific duplication. Asterisks indicate pairwise values for *SFB* and the other two genes in the *Prunus*-specific duplication-containing clade.

Based on the phylogenic analysis and timing of gene divergence, a model for the establishment of *Prunus SFB* is illustrated in Fig. 1.5. In this model, Clade A, which includes the *Prunus SFB*, diverged from Clade S, which includes the original pollen *S*, early in the history of the eudicots. The pollen *S* function in *Prunus* should have been maintained at least until the divergence between the *SFBB* genes and their orthologs in *Prunus*. The *SFBB* genes were generated by lineage-specific duplications after the divergence from *SLFL* genes. *SFB*, in contrast, was generated by a PSD on the ancestral chr 3. A later inter-chromosomal duplication or a transposition event could have resulted in the relocation of *SFB* to the distal part of chr 6 where the *SLFL* genes and the *S-RNase* are located (Fig. 1.5).

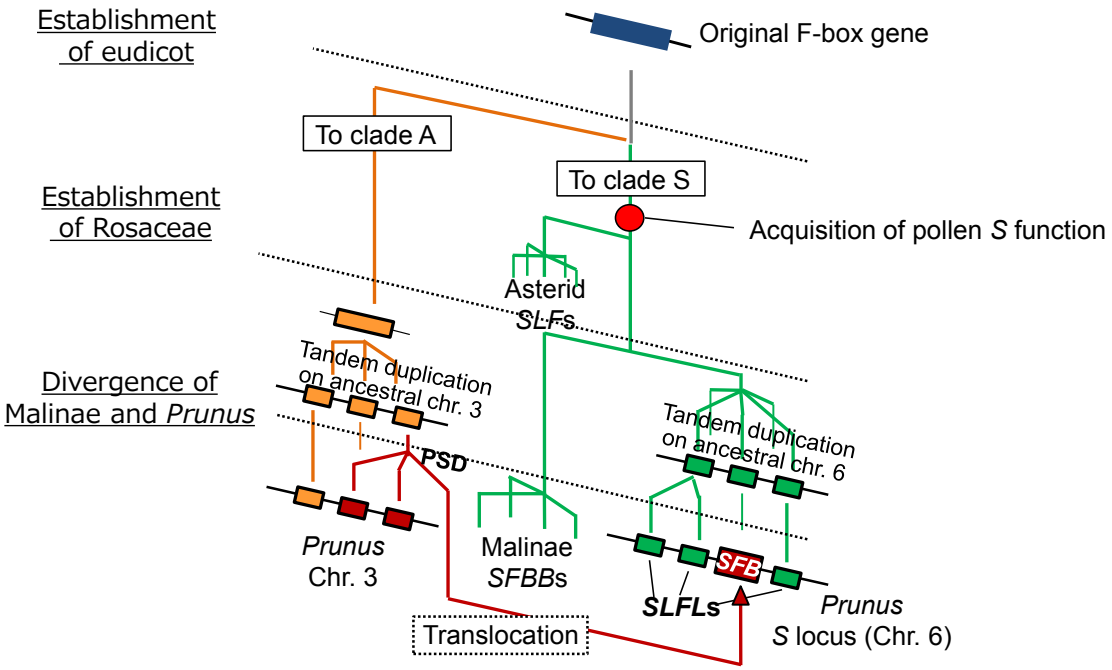

Fig. 1.5 Evolutionary model for the establishment of the *SFB/SLFL*-like genes in *Prunus*. Clade A, which includes the *Prunus SFB*, diverged from Clade S, which includes the original pollen *S*, early in the history of the eudicots. The pollen *S* function should have been maintained at least in until the divergence between the *SFBB* genes and their orthologs in *Prunus*. The *SFBB* genes were generated by lineage-specific duplications after the divergence from *SLFL* genes. *SFB*, in contrast, was generated by a *Prunus*-specific duplication on ancestral chr 3. A later inter-chromosomal duplication or a transposition event could have resulted in the relocation of *SFB* to the distal part of linkage group 6 with the *SLFL* genes and the *S-RNase*.

1.3.1.4. *Expression of SFB/SLFL-like F-box genes in Prunus*

Mature pollen grains from three cultivars in two *Prunus* species, *P. avium* and *P. mume*, were subjected to mRNA-Seq analyses. Among 28,689 mRNA transcript sequences predicted in the *P. persica* genome, 8,780 showed significant expression (RPKM > 2.0) in the pollen of one or more of the three cultivars. Genes in Clade S generally showed high expression levels in pollen grains, although there was considerable variation between genes (Fig. 1.6A). For example, two of the Clade S genes, ppa016207m and ppa016317m, showed no significant expressions in pollen. Three genes in the PSD-containing clade, possibly generated by PSD events, were expressed at similar levels. Most of the genes from clades other than the S- and PSD-containing clades showed very low expression levels in pollen. A more detailed expression analysis of the nine *SFB/SLFL*-like genes that showed significant expression levels in pollen grains was performed (Fig. 1.6B) by evaluating the expression levels in various organs of *P. avium* (cv. Satonishiki) and *P. mume* (cv. Nanko). All of the genes showed significant expression mainly in pollen grains, while the expression of some genes (*SLFL1*, ppa023668 and ppa024694) was also detected in the styles (Fig. 1.6B).

1.3.2. Selective pressure on the evolution of the *SFB/SLFL* genes

Based on the results of phylogenic and expression analyses, four genes located around the *S* locus and expressed in pollen grains, *SLFL1*, *SLFL2*, *SLFL3* and *SFB*, were tested for signs of selective pressure. Codon-based models of sequence evolution showed that Ka, Ks, and the Ka/Ks ratios in the three *Prunus SLFL* genes (Ks = 0.036–0.178, Ka = 0.010–0.061 and Ka/Ks = 0.223–0.343) were lower than those of the *Prunus S-RNase* (Ks = 0.252, Ka = 0.144 and Ka/Ks = 0.571). Similar results were found for pollen *S* from *Petunia* (Ka/Ks = 0.312 ± 0.030 for *SLF* genes and 0.485 for the *S-RNase*) and *Pyrus* (Ka/Ks = 0.335 ± 0.069 for the *SFBB* genes, and 0.878 for the *S-RNase)*. In contrast, in *Prunus*, the *SFB* gene showed a Ka/Ks ratio (0.438) closer to that of the *Prunus S-RNase* genes (Ka/Ks = 0.571), as discussed previously (Newbigin et al. 2008).

Next, a sliding window analysis was used to assess the trend of selective pressure along the coding regions. No common Ka/Ks transition pattern could be detected in the *SLFL/SFBB/SLF* genes (Fig. 1.7). Differences between *SLFL/SFBB/SLF* and the *Prunus SFB* gene were observed in the Ka/Ks ratios for the C-terminal region, especially for the 930–1,050-bp region (Ka/Ks < 0.5 in *SLFL/SFBB/SLF* genes). Importantly, the C-terminal region of the *SFB* gene, which corresponds to the hypervariable regions HVa and HVb (Ikeda et al. 2004), showed significantly higher

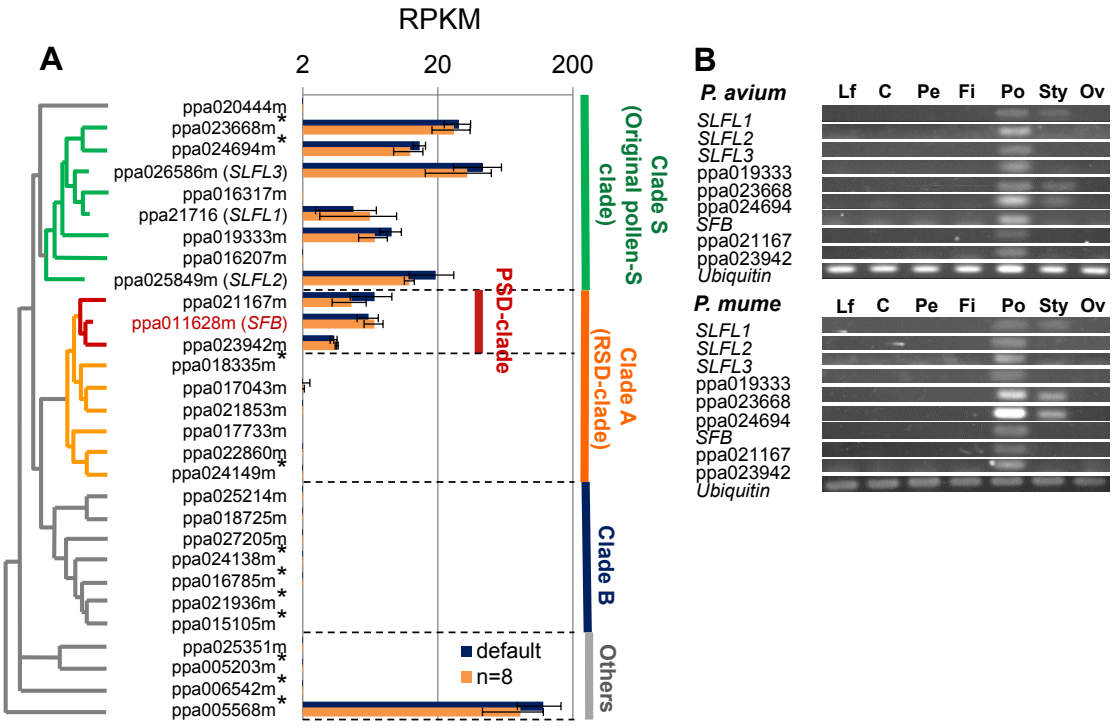

Fig. 1.6 Expression profiles of *SFB/SLFL*-like F-box genes in *Prunus*. (A) Gene expression levels of *SFB/SLFL*-like F-box genes in *Prunus* pollen. Average expression values across three *Prunus* accessions (*P. avium* 'Satonishiki' and *P. mume* 'Nanko' and 'Kairyo-uchida-ume') are expressed as RPKM with standard errors. Read mapping was performed twice, using the default parameters or allowing up to 8 bp of mismatches per read. (B) RT-PCR analysis of nine *SFB/SLFL*-like F-box genes that showed significant expression levels in pollen (Fig. 1.6A). Expression was assessed in leaves (Lf), calyxes (C), petals (Pe), filaments (Fi), pollen grains (Po), styles (Sty) and ovaries (Ov). The *ubiquitin* gene (ppa005507m) was used as the reference.

(0.9–1.5) Ka/Ks ratios than the rest of the sequence, suggesting that positive selection affected this region (Fig. 1.7).

Next, the PSD event that generated *SFB*, ppa021167m and ppa023942m, was investigated. Putative ancestral sequences were reconstructed for these genes. The Ka/Ks ratios for the full-length sequences showed no clear differences among the genes (Fig. 1.8A, Ka/Ks = 0.382–0.596). However, the sliding window analysis indicated distinct segmental Ka/Ks transition patterns in the path to the establishment of the *SFB* alleles from the ancestral gene.

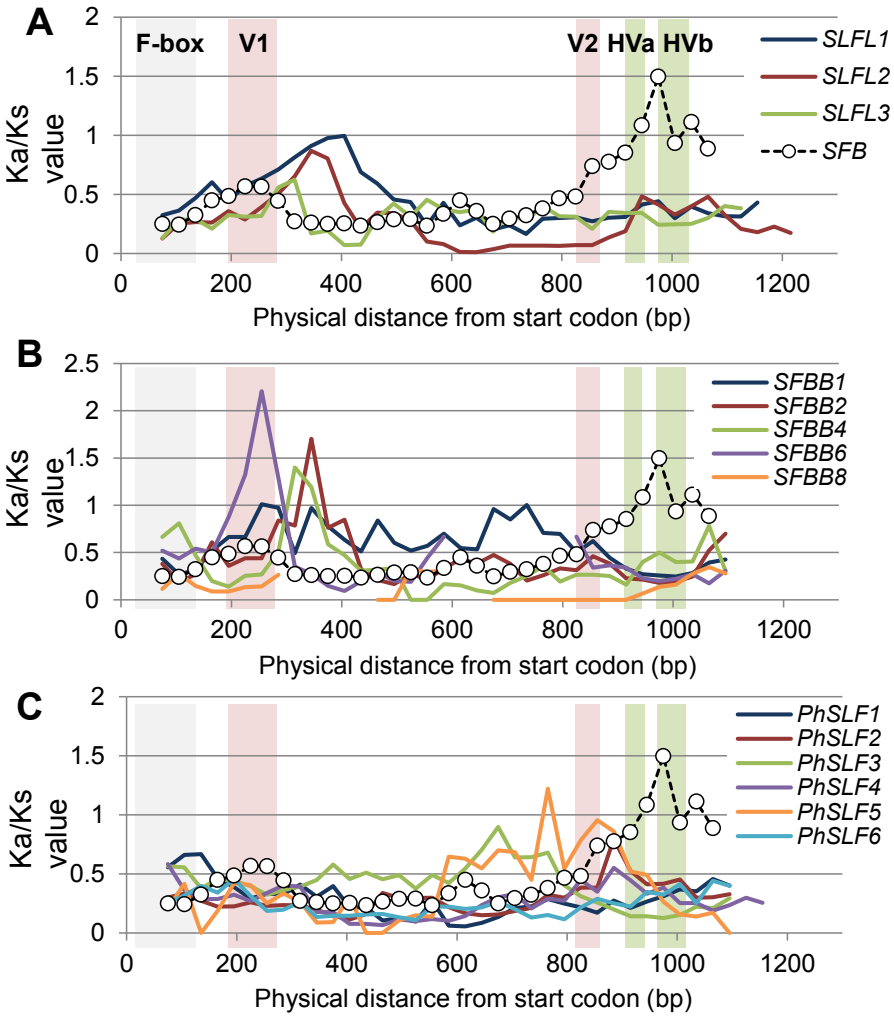

Fig. 1.7 Selective pressure on *SFB/SLFL*-like F-box genes in *Prunus*, *Pyrus* and *Petunia*. Mean pairwise Ka/Ks ratios between alleles from (A) *Prunus SFB* and three *SLFL*s located around the *Prunus S* locus, (B) *Prunus SFB* and five *SFBB*s from *Pyrus* and (C) *Prunus SFB* and six *SLF*s from *Petunia*. The F-box region, two variable regions (V1 and V2), and two hypervariable regions (HVa and HVb) of the *SFB* gene are colored in gray, red and green, respectively. *SFB* data is shown by the dotted line with white circles.

From the Rosaceae *SFB* origin (black circle in Fig. 1.8A) to the original *Prunus SFB* (red circle in Fig. 1.8A), the C-terminal region showed significantly low (< 0.3) Ka/Ks ratios, indicating a purifying selection on this region until the establishment of the *SFB* (Fig. 1.8B). In contrast, the Ka/Ks ratios in this region increased (Ka/Ks > 1.0) after the divergence of *SFB* alleles from the other members of the lineage, suggesting a release from the purifying selective pressure (Fig. 1.8B). Ka/Ks ratios between PSD and

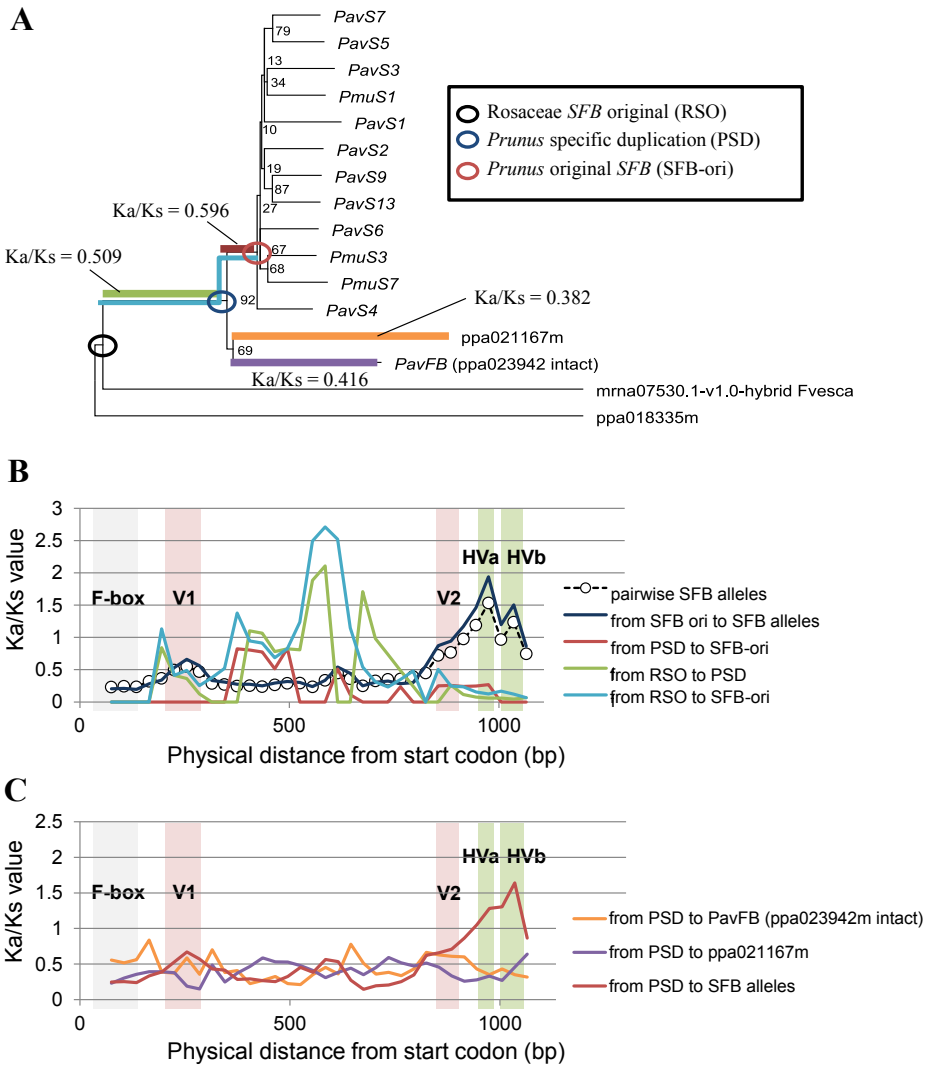

Fig. 1.8 Selective pressure during the establishment of *SFB*. (A) The phylogenetic tree was constructed based on the maximum likelihood method. Ka/Ks ratios for the *Prunus*-specific duplication (PSD)-containing clade genes are shown on each branch. Ka/Ks ratios are shown for the gene diversification path from the Rosaceae *SFB* origin (black circle) to PSD (blue circle), and to three major branches, including paths to the establishment of original *SFB* (red circle) and two other *SFB*-like F-box genes (ppa021167m and ppa023942m) from the PSD. (B) Window-average Ka/Ks ratios from the *SFB* origin to the current *SFB* alleles, from the PSD to the *SFB* origin, from Rosaceae *SFB* origin to *Prunus*-specific duplication, and from Rosaceae *SFB* origin to the *SFB* origin are shown in blue, red, green and light blue, respectively. Dotted line represents pairwise Ka/Ks ratios in current *SFB* alleles. (C) Window-average Ka/Ks ratios from PSD to ppa023942m, ppa021167m and the current *SFB* alleles, are shown in orange, purple and red, respectively.

the two *SFB* sisters (ppa021167m and ppa023942m) appeared to be low throughout the gene, indicating a purifying selection in the coding region for the C-terminal region (Ka/Ks = approximately 0.3–0.6) (Fig. 1.8C). Collectively, this results suggest that *SFB* experienced positive selection specifically on the C-terminal region, which corresponds to the hypervariable regions (HVa and HVb). This trend was not observed in the paths generating the original *SFB* or that of the other two sister genes generated by the PSD event.

1.4. Discussion

It has been suggested that the *Prunus SFB* gene is phylogenetically distant from the pollen *S* determinants of other plant species with S-RNase-based GSI systems (Ushijima et al. 2004, Wang et al. 2004, Matsumoto et al. 2008, Aguiar et al. 2015). Genome-wide analysis of the *SFB/SLFL*-like F-box genes in this Chapter supports this hypothesis and provides new clues to the evolution of *SFB* in the *Prunus* genome. The divergence between the ancestors of the *Prunus SFB* and the other pollen *S* F-box genes is supposed to have occurred early in the establishment of eudicots, and *SFB* was supposed to have been generated from PSD. These findings suggest that the *Prunus* species came to use the *SFB* gene as a pollen *S* determinant after *Prunus*-specific F-box gene duplications, which probably occurred around the time of the *Prunus* divergence from its common ancestor with *Malinae* and *Fragaria*.

This model may raises a question to be answered: what is the function of the *Prunus SLFL* genes, which belongs to Clade S, along with the pollen *S* determinants from the Solanaceae and Malinae, and is still physically close to the *S* locus? A clue to the answer might be obtained from gene expression analyses and selective pressures on *SLFL* genes. Given that *SLFL* genes are expressed mainly in the pollen (Fig. 1.6) and that they retain low Ka/Ks ratios (or purifying selection) (Fig. 1.7), they may still retain functions present in the original pollen *S* determinant. Matsumoto et al. (2012) showed that SLFLs interact with a Skp1-like homolog from *P. avium* (PavSSK1) that is proposed to be a component of the SCF complex, which is involved in the polyubiquitination of proteins for degradation. Thus, it is possible that the *Prunus* SLFLs are still involved in the degradation of S-RNase in a process similar to that of SLF/SFBB (Kubo et al. 2010, Kakui et al. 2011, Williams et al. 2014, Kubo et al. 2015). In this case, the *SLFL* genes could be good GI candidates that putatively target all of the S-RNases for degradation (Tao and Iezzoni 2010, Matsumoto and Tao 2016a).

This evolutionary scenario suggests that SFB has taken over the function of discrimination for self/nonself S-RNases from the original pollen *S* determinant. The

original ancestral *Prunus* pollen *S* determinant must have carried the function to recognize nonself-S-RNases for degradation, which is the function collaboratively carried out by the current pollen *S* proteins of the Solanaceae and Malinae. After the PSD, the *SFB* gene, which was newly duplicated in the *Prunus S* locus, took over the function of discriminating self/nonself S-RNases, acting as the pollen *S* determinant in *Prunus*. *Prunus SFB* contains two variable (V1 and V2) and two hypervariable (HVa and HVb) regions (Ikeda et al., 2004). These variable and hypervariable regions appear to be hydrophilic, which suggests that these regions may be exposed on the surface and function in self/nonself recognition (Ikeda et al., 2004; Tao and Iezzoni, 2010). Evidence that *SFB* experienced positive selective pressure in the hypervariable regions after PSD may be consistent with the acquisition of the current SFB function. Genes in the *SLF/SFBB/SLFL* clade (Clade S) in *Prunus* might have escaped from purifying selection on the function to distinguish self/nonself S-RNases. This could have resulted in the acquisition of the GI function to degrade all S-RNases (Tao and Iezzoni 2010, Matsumoto and Tao 2016a).

1.5. Summary

SI is an important plant reproduction mechanism that facilitates the maintenance of genetic diversity within a species. Three plant families, the Rosaceae, Solanaceae, and Plantaginaceae, share an S-RNase-based GSI system that involves a single S-RNase as the pistil *S* determinant and several F-box genes as pollen *S* determinants that act through nonself-recognition. Previous evidence suggested a specific self-recognition mechanism in *Prunus* (Rosaceae), raising questions about the generality of the S-RNase-based GSI system. It is hypothesized that a putative GI in the pollen tube targets all of the S-RNases for degradation in *Prunus*. In this chapter, the evolution of the pollen *S* determinants and their orthologs in other angiosperm species were investigated. The results indicate that the *Prunus* Pollen *S*, *SFB*, does not cluster with the pollen *S* of other plants and diverged early after the establishment of the eudicots. Furthermore, multiple F-box gene duplication events, specifically in the Rosaceae lineage, are indicated and the *Prunus SFB* gene originated in a recent *Prunus*-specific gene duplication event. Transcriptomic and selective pressure analyses of the *Prunus S* paralogs are consistent with the establishment of a *Prunus*-specific GSI system based on self-recognition by the newly generated *SFB*. SLFLs, which belong to original pollen *S*-containing clade, may retain functions that were present in the original pollen *S* and still might participate in SI reactions in *Prunus* as the GI.

Chapter 2. Functional characterizations of *Prunus* SLFLs in transgenic *Petunia*

2.1. Introduction

In the S-RNase-based GSI system, S-RNases and F-box proteins are specificity determinants on the pistil and pollen sides, respectively. Although these specificity determinant genes are considered to be common among different taxa, several lines of evidence indicate that the SI recognition mechanism in *Prunus* in the Rosaceae is distinct from those in the other species. The pollen *S* F-box proteins are assumed to target nonself S-RNases for degradation in the Solanaceae, Plantaginaceae, and subtribe Malinae, allowing successful pollen tube growth in nonself styles and consequent fertilization (Kakui et al., 2011; Kubo et al., 2010; Minamikawa et al., 2010; Sassa et al., 2007; Zhou et al., 2003). In contrast, SFB, the pollen *S* F-box protein in *Prunus*, is thought to function in the release of self S-RNase cytotoxicity (Matsumoto and Tao, 2016a; Tao and Iezzoni, 2010). A hypothetical GI that is involved in the detoxification of all of the S-RNases is supposed to be present in the *Prunus* pollen tube (Fig. 3B) (Tao and Iezzoni, 2010; Matsumoto and Tao, 2016a). This hypothesis is supported by the different phenotypic consequences of hetero-diallelic pollen and pollen *S* mutations in *Prunus* as compared with those in other species (Hauck et al., 2006b; Ushijima et al., 2004).

Phylogenetic, expression and selective pressure analyses in Chapter 1 confirmed previous speculation regarding the functions of SFB and the SLFLs (Aguiar et al., 2015; Matsumoto et al., 2012). In Chapter 1, it was clearly shown that *Prunus SFB* is phylogenetically distant from the pollen *S* F-box genes of other species, while *SLFL*s, which are encoded by regions neighboring the *Prunus S* locus, cluster closely with *SLF* and *SFBB* (Fig. 1.2). Furthermore, *SLFL*s are expressed mainly in pollen, as observed in pollen *S* F-box genes (Ushijima et al., 2004), and they retained low Ka/Ks ratios (or purifying selection) (Fig. 1.7). Thus, SLFLs may function to degrade S-RNases as the GI. In this chapter, the functional characterizations of *SLFL*s were performed by transforming SI *Petunia* (*Petunia* × *hybrid*, Solanaceae) with *Prunus* SLFLs to determine if they can function as GI.

2.2. Materials and Methods

2.2.1 Plant material

The SI *Petunia* cultivar 'NIL Mitchell', an SI near-isogenic line of 'Mitchell', was produced by introducing S_{3L} haplotype and *HT-B* of *P. inflata* into 'Mitchell' through backcrossing, using 'Mitchell' as the recurrent parent (Puerta et al., 2009).

2.2.2. Vector construction

The pollen-specific LAT52 promoter of tomato was used to express *SLFL*s in pollen (Twell et al., 1991). *Prunus SSK1*, which is orthologous to *Petunia SSK1*, was also used in this study. *Petunia* SSK1 is reported to interact with SLF and form an SCF complex to recognize and polyubiquitinate nonself S-RNases for degradation (Hua and Kao, 2006; Kubo et al., 2010). *Prunus* SSK1 is reported to interact with SLFLs in the same way (Matsumoto et al., 2012; Matsumoto and Tao, 2016b). The LAT52 promoter was amplified from pPK100-52AL using primers (LAT52pro-vector-F and -R) with *Hind*III and *Xba*I + *Sac*I sites at the 5′ and 3′ ends, respectively (Table 2.1). The *35S* promoter in the binary vector pGWB2 was replaced with the LAT52 promoter to yield pGWB2-LAT52. Full-length cDNAs of *SLFL*s were amplified from pollen cDNA of sweet cherry 'Satonishiki' (S_3S_6) using gene-specific primers (PavSLFL1-vector-F and -R, and PavSLFL2-vector-F and -R) with *Xba*I and *Sac*I sites at the 5′ and 3′ ends, respectively (Table 2.1). *SLFL1* and *SLFL2* cDNA sequence was inserted under the control of LAT52 to yield pGWB2-LAT52::*SLFL1* and pGWB2-LAT52::*SLFL2*, respectively. Similarly, the cDNA sequence from *P. avium SSK1* was introduced into pGWB2-LAT52. Because the coding sequence of *SLFL3* includes a *Sac*I digestion site, the InFusion system (Takara) was used to yield pGWB2-LAT52::*SLFL3* according to the manufacturer's protocol. These constructs were separately introduced into an *Agrobacterium* strain, EHA101.

2.2.3. Transformation

Petunia leaf discs (1 cm), cut from *in vitro* cultured plants were gently mixed with the *Agrobacterium* culture (diluted to optical density of 0.5 at 600 nm) for 20 min, and blotted on filter paper to remove the extra suspension. The leaf discs were then returned to a half-strength Murashige and Skoog (MS) medium supplemented with 5 μM 6-benzylaminopurine (BA) and 0.5 μM 1-naphthalene acetic acid (NAA) for 2 days of co-cultivation. After two days, the leaf discs were transferred to a disinfection medium, which was half-strength MS medium, 5 μM BA, 0.5 μM NAA and 100 mg L^{-1} meropenem, for 5 days. Leaf discs were then transferred to a selective medium, which consisted of half-strength MS medium, 5 μM BA, 0.5 μM NAA, 100 mg L^{-1} meropenem and 50 mg L^{-1} kanamycin, and were transferred to fresh selective medium every 2 weeks until shoots developed. Regenerated shoots were excised from calli and cultured on a rooting medium, which was half-strength MS medium, 100 mg L^{-1} meropenem and 50 mg L^{-1} kanamycin. Rooted shoots were planted in soil in pots and grown for further analyses.

Table 2.1 Primers used for vector construction and RT-PCR analysis.

LAT52pro-vector-F	AAGCTTGTCGACATACTCGACTCAGA
LAT52pro-vector-R	GAGCTCTCTAGAAATTTTTTTTTTGGTGTG
PavSLFL1-vector-F	TCTAGAATGTGGGAAGAGATGGCGTT
PavSLFL1-vector-R	GAGCTCCTAACTCTCAAGCTTGTTGC
PavSLFL2-vector-F	TCTAGAATGGCAACGTTGAGCAAATT
PavSLFL2-vector-R	GAGCTCCTATATGCTTTCAATCTTGT
PavSLFL3-infusion-F	CCAAAAAAAAAATTTCTAGAATGACATACTTTTGCAAAAT
PavSLFL3-infusion-R	GAACGATCGGGGAAAAGCTCCTAATAATCTTTGGTCTCGA
PavSSK1-vector-F	TCTAGAATGTCGGCCGAGGAGGAGAA
PavSSK1-vector-R	GAGCTCTCAGTCCTCATCAACTCCTT
PavSLFL1-expression-F	TGGGTGTTGGATGACTTTGA
PavSLFL1-expression-R	AGCTTGATGCCTCCCTATGA
PavSLFL2-expression-F	TTGACATGAGCGAAGAGGTG
PavSLFL2-expression-R	CGTAAGGAACCGCTTGGATA
PavSLFL3-expression-F	CCTCAGCAGGACAGAACACA
PavSLFL3-expression-R	ACCCCAAAGTGCATGAGAAG
PavSSK1-expression-F	GGCGTTCTTCCAAGATGAAG
PavSSK1-expression-R	AAGTAATCGGCGGCTAGGAT
PhTublin-expression-F	CCCTATCCCAGGATCCACTT
PhTublin-expression-R	AACATCCTTTGGCACCACAT

2.2.4. RT-PCR

Total RNA was isolated from mature anthers of transgenic lines using a Plant RNA reagent (Thermo Fisher Scientific). Wild-type 'NIL Mitchell' and 'Satonishiki' were used as controls. First-strand cDNA was synthesized from 100 ng of total RNA using the ReverTra Ace qPCR RT Master Mix with gDNA Remover (Toyobo) according to the manufacturer's protocol. Gene-specific primers for *SLFL*s and *SSK1* were designed using the online software Primer3 (http://frodo.wi.mit.edu/primer3/) (Table 2.1). *PhTublin* was used as the internal standard. RT-PCR was performed in a 20-μL reaction volume containing 1× Ex-*Taq* buffer (Takara), 0.2 mM dNTPs, 100 nM each of the forward and reverse primers, 0.5 U Ex-*Taq* (Takara), and cDNA equivalent to the amount synthesized from 10 ng of total RNA. The PCR cycle was as follows: initial denaturation at 94°C for 1 min; then 35 cycles of 94°C for 30 s, 55°C for 30 s and 72°C

for 1 min; followed by a final elongation at 72°C for 5 min. The PCR products were electrophoresed through a 1.5% agarose gel, stained with ethidium bromide and visualized under ultraviolet light.

2.2.5. Pollination test

The self-(in)compatibility phenotype was investigated by the microscopic observation of pollen tube growth after self-pollination. Ten flowers from each transgenic plant were self-pollinated in a greenhouse. Pollinated pistils were collected 24 h after pollination, and fixed in chloroform:99% ethanol:acetic acid (1:3:1) overnight. The fixed pistils were washed three times with distilled water, softened in 10% sodium hydroxide for 5 h, washed three times with distilled water, and stained with 0.1% aniline blue in 0.1 N K_3PO_4. Fluorescent microscopic observations of pollen tube growth were performed on squash samples using a fluorescence microscope BX60 (Olympus), equipped with ultraviolet epifluorescence. After investigating pollen germination on the stigma, pollen tube growth was scored as the percentage of the style length traveled by the longest pollen tube. A one-way analysis of variance was conducted on pairwise comparisons using the Tukey–Kramer method.

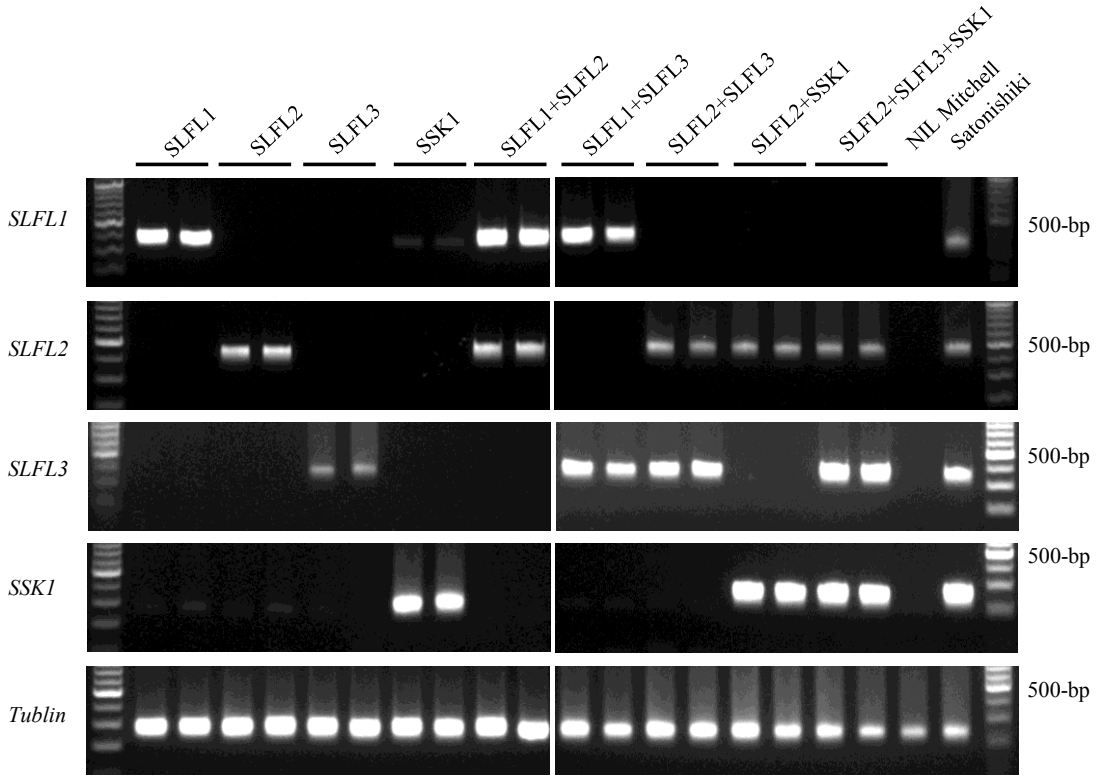

Fig. 2.1 Transgene expression in transgenic *Petunia* pollen. Wild-type 'NIL Mitchell' and 'Satonishiki' were used as the negative and positive controls, respectively.

2.3. Results and Discussion

Prunus SLFLs may function as the GI whose presence in the *Prunus* pollen tube explain the *Prunus*-specific SI reaction (Tao and Iezzoni, 2010; Matsumoto and Tao, 2016a). Because the *Prunus* genus is recalcitrant to transformation, the functional characterizations of *Prunus* SLFLs were conducted in transgenic *Petunia* (Solanaceae), which also has S-RNase-based GSI. Transgenic 'NIL Mitchell' plants with *SLFL1*, *SLFL2*, *SLFL3* or *SSK1* were successfully obtained by *Agrobacterium*-mediated transformations. Transgenic plants with two different types of *SLFL*s, or *SLFL*s together with *SSK1*, were obtained by bud pollination between transgenic plants containing the respective genes.

Each transgenic *Petunia* expressed the transgene(s) in pollen (Fig. 2.1). In the pollination experiment, seed formation was not observed after self-pollination in any of the transgenic plants, even in transgenic plants with two different types of *SLFL* genes or with *SSK1*. The effects of SLFLs and SSK1 expression on pollen tube growth in SI *Petunia* were then evaluated using a pollen tube growth test. The self pollen tube growth of wild-type 'NIL Mitchell' was arrested at ~60% from the top of the style (Fig. 2.2 and 2.3). Similarly, the self-pollen tube growth of most transgenic plants was arrested in the style and there were no significant differences in pollen tube growth rates between transgenic and wild-type *Petunia* plants. Several transgenic lines (L1+L2-#1 and -#2, L2+L3-#1 and -#4, L2+SSK1-#3 and L2+L3+SSK1-#5) showed reduced pollen tube growth rates compared with those of the wild-type plant. Transformations with two different types of *SLFL*s and *SSK1* also showed no phenotypic changes. In contrast, self-pollen tubes reached the bases of the styles by 24 h after pollination in SC *Petunia* 'Mitchell' (Fig. 2.3). Thus, *Prunus* SLFLs did not affect the self-(in)compatibility phenotype or pollen tube growth in *Petunia*, at least under experimental conditions in this Chapter.

Recent genome-wide phylogenic analyses of *S* locus-related F-box genes and *S-RNase* genes have suggested that *S* locus duplication events occurred in the ancestral genome of the Rosaceae, and plants from the genus *Prunus* and subtribe Malinae came to use different paralogs as SI determinants (Aguiar et al., 2015; see Chapter 3). Thus, the specificity for self/ nonself recognition may be changed in duplicated *S* loci, such as those in the genus *Prunus* and subtribe Malinae, because the duplicated *S* loci have accumulated mutations in their respective lineages. Changes in specificity could explain the lack of phenotypic changes associated with *Prunus* SLFLs in *Petunia*, which belongs to the Solanaceae, in the Asterid.

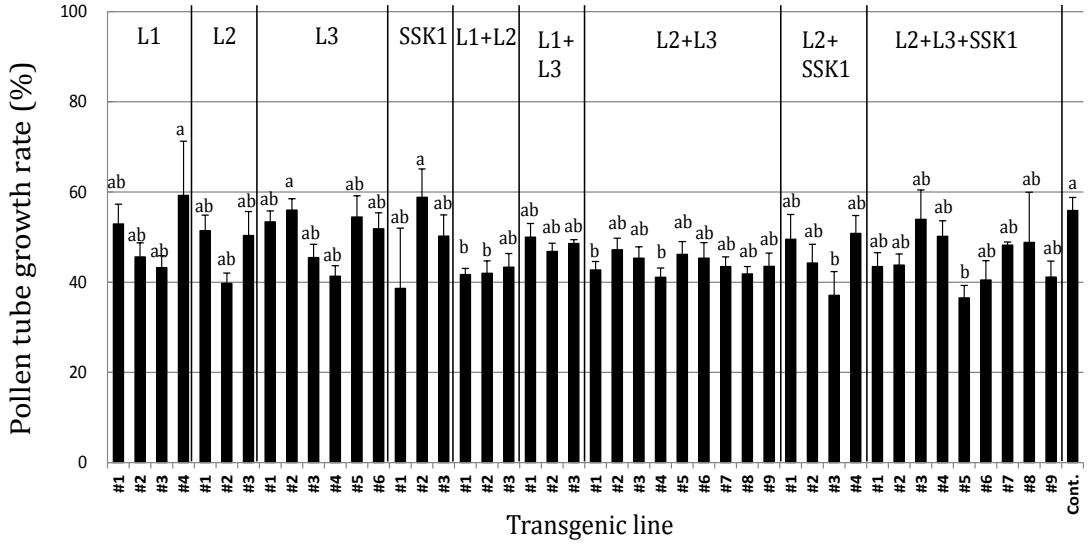

Fig. 2.2 Pollen tube growth after self-pollination in transgenic *Petunia*. The pollen tube growth rate represents the percentage of the style length traveled by the longest pollen tube. Different lowercase letters indicate significant differences ($P < 0.01$) based on the Tukey–Kramer test. Cont. (control) indicates the pollen tube growth rate for wild-type 'NIL Mitchell' after self-pollination.

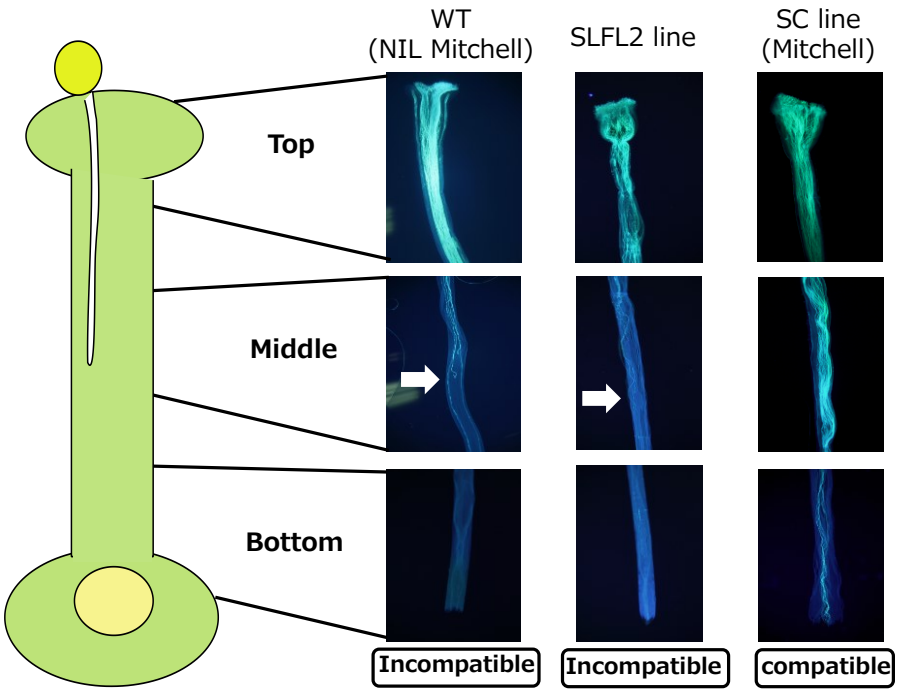

Fig. 2.3 Pollen tube growth at 24 h after self-pollination for the wild-type 'NIL Mitchell' (left), a transgenic line with SLFL2 (center) and the SC *Petunia* line 'Mitchell' (right). Arrows indicate the positions where pollen tube growth ceased in the pistil.

During the preparation of this thesis, it was reported that the SCF complex selectively uses the specific Cullin1 gene for the detoxification of nonself S-RNases in *Petunia*, and *Cullin1* may have evolved independently in the Solanaceae and Rosaceae (Kubo et al., 2016). Further transgenic experiments using taxa closely related to *Prunus*, such as *Fragaria* (Rosaceae), and Cullin1 from *Prunus* will be needed to confirm the functions of the SLFLs.

2.4. Summary

The Rosaceae, Solanaceae and Plantaginaceae have S-RNase-based gametophytic self-incompatibility that uses S-RNases and F-box proteins as the pistil *S* and pollen *S* determinants, respectively. It is generally accepted that pollen tube growth is arrested in the style by S-RNase cytotoxicity. In the Solanaceae, Plantaginaceae and subtribe Malinae, pollen *S* F-box proteins are assumed to target nonself S-RNases for degradation, which allows successful pollen tube growth in nonself pistils for fertilization. In contrast, SFB, the pollen *S* F-box protein in *Prunus* in the Rosaceae, is thought to have a distinct function in the release of self S-RNase cytotoxicity. In contrast to the intrinsic function of the pollen *S* determinant in the degradation of nonself S-RNases in plants other than *Prunus*, a GI that is involved in the detoxification of all S-RNases may be present in *Prunus* pollen tubes. Although the identity of the GI is unknown, phylogenetic, expression, and selective pressure analyses in Chapter 1 indicated that *SLFL*s, which are located in the regions flanking the *Prunus S* locus, are good GI candidates. In this chapter, functional characterizations of *SLFL*s were performed by transforming SI *Petunia* to test the functions of the SLFLs. Although the substantial expression of *Prunus SLFL*s was observed in transgenic *Petunia* pollen, no change in the SI phenotype was obtained. Furthermore, the co-expression of *Prunus SSK1*, which interacts with SLFLs to form the SCF complex, did not induce any SI phenotypic changes. Possible explanations for these results were discussed.

Chapter 3. Evolutionary analysis of genes for S-RNase-based self-incompatibility reveals *S* locus duplications in the ancestral Rosaceae

3.1. Introduction

The Solanaceae, Rosaceae, and Plantaginaceae have the S-RNase-based GSI system. Previous studies have indicated that the S-RNase-based GSI is monophyletic and evolved once in eudicots before the divergence of asterids and rosids (Igic and Kohn, 2001; Steinbachs and Holsinger, 2002; Vieira et al., 2008). Despite the indication of a single origin of the S-RNase-based GSI and the commonality of specificity determinant genes, several lines of evidence indicate that the SI recognition mechanism in *Prunus* in the Rosaceae differs from those in the other species (Hauck et al., 2006b; Tao and Iezzoni, 2010; Ushijima et al., 2004).

Several *S-RNase*-like genes, called *non-S-RNase*s, have been identified from certain plant species with S-RNase-based GSI systems. *Non-S-RNase*s, showing homology to *S-RNase*, have been reported in *P. inflata* (Lee et al., 1992), *N. alata* (Kuroda et al., 1994) and *P. avium* (Yamane et al., 2003b). The *non-S-RNase*s found in the Solanaceae are assumed to have arisen from the duplication of the *S* locus, but are not associated with SI (Golz et al., 1998; Igic and Kohn, 2001; Lee et al., 1992). Although *Prunus non-S-RNase* was reported to be a possible candidate for the ancestral form of S-RNase (Yamane et al., 2003b), its physiological functions have not been clarified. Detailed evolutionary analyses of *S-RNase*s and pollen *S* F-box genes may provide clues to the distinct recognition mechanisms of SI in different lineages and insights into the molecular basis of the S-RNase-based GSI in rosaceous species.

Gene duplication has played an important role in the evolution of plants (Flagel and Wendel, 2009). A prevailing theory predicts gene loss (pseudogenization) or functional diversification (sub- or neofunctionalization) as the main fates for one of the two copies of duplicated genes (Blanc and Wolfe, 2004; Cusack and Wolfe, 2007; Moore and Purugganan, 2005). Of the two functional diversification mechanisms, subfunctionalization, in which each duplicate gene develops a distinct expression pattern as models of cis-evolution (Carroll, 2008), seems to be the most common phenomenon and takes place soon after gene duplication (Papp et al., 2003). In contrast, neofunctionalization is the result of diversification over the long term (Rastogi and Liberles, 2005).

In this chapter, the evolutionary paths of *S-RNase* and its homologs were investigated mainly by tracking their duplication patterns in genomes. The results indicated that *S-RNase* and its homologs were duplicated in the ancestral genome of the

Rosaceae, and the *S* loci in the current rosaceous species have evolved independently from the duplicated *S* loci. The results of this study on rosaceous *S* locus evolution could explain the genus (or subtribe)-specific self/nonself recognition mechanisms present in the Rosaceae and will be valuable for the future development of artificial SI control and SC breeding in rosaceous fruit tree species.

3.2. Materials and Methods

3.2.1. Gene collection

Putative full-length sequences of 38 *S-RNase*-like genes were identified from the genomes of seven angiosperms and two outgroup species, *Selaginella* and *Physcomitrella*, using the BLASTP algorithm in Phytozome (version 9.1, http://www.phytozome.net/). The genes showing significant homology (e^{-19} cutoff for all angiosperms and e^{-10} cutoff for all others) to S_3-RNase (AB010306) from *P. avium* were selected. The amino acid sequences with significant BLAST hits were subjected to a GO analysis by scanning the Pfam database (Finn et al., 2014) for the presence of a ribonuclease T2 motif. Information concerning the physical locations and structures of genes were based on the genome sequence obtained from the Phytozome database.

For a detailed phylogenic analysis of S-RNase lineage in the Rosaceae, *S-RNase*-like sequences from several diploid *Fragaria* species, including *Fragaria iinumae* (unknown for SI/SC phenotype), *Fragaria nipponica* (SI) and *Fragaria nubicola* (SI) were retrieved. Genomic sequences that showed significant homology (e^{-19} cutoff) to S_3-RNase (AB010306) from *P. avium* or S_1-RNase (D50837) from *M.* × *domestica* were selected using tBLASTn at Strawberry GARDEN (http://strawberry-garden.kazusa.or.jp/). Putative amino acid sequences were predicted based on alignments against a true *S-RNase* sequence.

Pollen *S* F-box-like genes were searched for in regions containing *S-RNase*-like genes in the Rosaceae genome using Phytozome Gbrowse. F-box genes showing significant homology (e^{-19} cutoff) to pollen *S* F-box genes in asterids (*SLF*) and the subtribe Malinae (*SFBB*) were identified. The *Prunus SFB* was excluded from this analyses because it originated from a different ancestral F-box gene in Clade A, not in Clade S, which includes the original pollen *S* F-box gene (see Chapter 1).

3.2.2. Construction of evolutionary topology

Alignment analyses of amino acid sequences were conducted using MAFFT v7 with the L-INS-I model (Katoh and Standley, 2013). The raw alignments were subjected to manual revision using Sea View v4 (Gouy et al., 2010). Unnecessarily long gap

sequences, which disturb the proper alignment of orthologous sequences, and genes showing apparently different structures from other *S-RNase*-like genes were removed. The Neighbor–joining approach was applied to define the evolutionary topology using MEGA v5.05 (Tamura et al., 2011) with 1,000 bootstrap replications.

3.2.3. Estimation of the time of gene divergence

The amino acid sequences of each pair of homologs were aligned using MAFFT with the L-INS-I model, and converted to nucleotide alignments using PAL2NAL (Suyama et al., 2006). The 4DTv were calculated between the gene pairs using Microsoft Excel 2007 (Microsoft). Here, the four-fold degenerate sites were the codons of amino acid residues G, A, T, P, V, R, S, and L. The 4DTv values were calculated for gene pairs retaining at least 23 four-fold synonymous sites. For calculation of the 4DTv between specific gene groups, the 4DTv values from all gene combinations between the two groups were averaged. Aligned nucleotide sequences were also analyzed to calculate the Ks values, using DnaSP 5.1 (Librado and Rozas, 2009).

3.2.4. Gene expression analysis

To perform organ-specific RT-PCR analyses for *S-RNase*-like genes, total RNAs from leaves, calyxes, petals, filaments, pollen, styles and ovaries were isolated from *P. avium* (‘Satonishiki’, ‘Takasago’ and ‘Rainier’) and *P. mume* (‘Nanko’ and ‘Kairyo-uchida-ume’) using the cold-phenol extraction method as described by Tao et al. (1999). The total RNA of pistil tissue was also isolated from *P. armeniaca* (‘Yamagata No.3’), *P. salicina* (‘Sordum’ and ‘Santarosa’) and *P. persica* (‘Akatsuki’ and ‘Shimizu-hakuto’). cDNA was synthesized from 100 ng of total RNA using the ReverTra Ace qPCR RT Master Mix with gDNA Remover (Toyobo) according to the manufacturer’s protocol. Gene-specific primers for each *S-RNase*-like gene were designed using the online software Primer3 (http://frodo.wi.mit.edu/ primer3/): 5'-GTTGCCCAAGGAAAAGACAA-3' and 5'-GTCGCGTTTGTTGGAAAGAT-3' for ppa011133m (non-S-RNase1), and 5'-AGTGCTCCGACGACAAGTTT-3' and 5'-ATTGCTCGCAAAGGAGAAGA-3' for ppa024151m (non-S-RNase2). Primers 5'-ACCATAACGTTGGAGGTGGA-3' and 5'-GGAGACGAAGGACAAGGTGA-3' for *ubiquitin* (ppa005507m) were used as a control. RT-PCR was performed in a 20-μL reaction volume including 1× Ex-*Taq* buffer (Takara), 0.2 mM dNTPs, 100 nM each of the forward and reverse primers, 0.5 U Ex-*Taq* (Takara), and cDNA equivalent to the amount synthesized from 10 ng of total RNA. The PCR cycle was as follows: initial

denaturation at 94°C for 2 min; then 30 cycles of 94°C for 30 s, 55°C for 30 s and 72°C for 1 min; followed by a final elongation at 72°C for 5 min. The PCR products were electrophoresed through a 1.5% agarose gel, stained with ethidium bromide and visualized under ultraviolet light.

3.2.5. Identification of selective pressure on the *S-RNase*-like genes

Full-length allele sequences of *S-RNase*s in the Solanaceae, subtribe Malinae, and *Prunus*, and *non-S-RNase1* and *non-S-RNase2* in *Prunus*, were obtained from the NCBI and Phytozome databases. Gene sequences orthologous to ppa024151m (*non-S-RNase2*) in *P. avium* and *P. salicina* were also obtained by the direct sequencing of PCR products. Primers were designed outside of the open reading frame region in ppa024151m, using sequences from the peach genome database (GDR, Prunus_persica_v.1.0). Amino acid sequence alignments were constructed using MAFFT with the L-INS-I model, and the alignments were converted to nucleotide alignments using PAL2NAL. Informative single-nucleotide polymorphisms in the alleles were analyzed by DnaSP 5.1 and used to calculate the evolutionary speed (Ka/Ks). Window-average Ka/Ks ratios were calculated from the start codon (ATG) in 90-bp windows with a 20-bp walking step, until the walking window reached the stop codon. To calculate the amino acid variability level, normed variability indices (NVIs) for each residue in the aligned alleles were calculated as described by Kheyr-Pour et al. (1990) using Microsoft Excel 2007. Window-average NVI values were calculated from the first amino acid (M) in the 11-amino acid window with a 1-AA walking step.

3.3. Results

3.3.1. Evolutionary patterns of *S-RNase* and its homologs in angiosperm genomes

The comprehensive detection of genes that showed significant homology to *Prunus S-RNase* was performed by the BLAST algorithm and GO analyses. To obtain a wide range of phylogenetic data for *S-RNase*-like genes across angiosperms, representative genomes according to APGIII (The Angiosperm Phylogeny Group, 2009) were selected; *Oryza sativa* (monocot) as an outgroup of eudicots, *Mimulus guttatus* from asterids, *Arabidopsis thaliana* from eurosid II, and *Populus trichocarpa* and three genomes of the Rosaceae (*P. persica, F. vesca,* and *M.* × *domestica*) from eurosid I. Thirty-eight genes (one gene per locus) were identified as *S-RNase*-like genes (Table 3.1).

Table 3.1 Gene list of *S-RNase*-like genes in angiosperms

Category			Family	Species	Accessions	Note
Embrophyto			Funariaceae	*Physcomitrella patens*	Pp1s358_60V6.1	
Tracheophyto			Sellanginellaceae	*Sellanginella moellendorfii*	75785	
					270532	
Angiosperm	Monocots		Poaceae	*Oryza sativa*	LOC_Os01g67180.1	
					LOC_Os01g67190.1	
					LOC_Os07g43600.1	
					LOC_Os07g43640.1	
					LOC_Os07g43670.1	
					LOC_Os08g33710.1	
	Core Eudicots	Euasterid I	Scrophulariaceae	*Mimulus guttatus*	mgv1a011196m	
					mgv1a011299m	
					mgv1a012841m	
					mgv1a013386m	
			Solanaceae	*Petunia × hybrida*	AAG21384	S2-RNase
			Plantaginaceae	*Antirrhinum hispanicum*	CAC33020	S2-RNase
		Eurosid I	Rosaceae	*Fragaria vesca*	mrna00224.1-v1.0-hybrid	
					mrna00227.1-v1.0-hybrid	
					mrna22673.1-v1.0-hybrid	
					mrna26891.1-v1.0-hybrid	
				Malus × domestica	MDP0000135121	
					MDP0000160706	
					MDP0000210735	
					MDP0000213741	
					MDP0000236215	
					MDP0000250548	
					MDP0000251832	
					MDP0000345854	
					MDP0000400831	
					MDP0000682955	
					MDP0000826052	
					AAA79842	S3-RNase
				Prunus avium	AB010306.1	S3-RNase
				Prunus dulcis	AF227522.1	S-like
				Prunus mume	AB101438.1	S1-RNase
				Prunus persica	ppa009963m	
					ppa011014m	
					ppa011026m	
					ppa011133m	
					ppa018459m	
					ppa024151m	
				Prunus salicina	AB252411.1	Sa-RNase
				Pyrus × bretschneideri	XM_009362057.1	S-like
				Pyrus pyrifolia	BAA93052.1	S3-RNase
			Malpighiaceae	*Populus trichocarpa*	Potri.008G086800.1	
					Potri.010G168700.1	
					Potri.014G174400.1	
		Eurosid II	Brassicaceae	*Arabidpsis thaliana*	AT1G14220.1	
					AT1G26820.1	
					AT2G02990.1	
					AT2G39780.1	

A phylogenetic analysis of the *S-RNase* genes in the Solanaceae, Plantaginaceae, subtribe Malinae, *Prunus*, and the 38 *S-RNase*-like genes found in the angiosperm genomes, indicated that the representative *S-RNase*-like genes in angiosperms could be divided into three major classes [Angiosperm (AG) I–III], as

reported previously (Igic and Kohn, 2001) (Fig. 3.1). At least three *S-RNase*-like genes were defined in the MRCA of the eudicots (triangles in Fig. 3.1), which corresponded to the divergence of the three major classes. For the AGI and AGII classes, the divergence patterns of the genes corresponded well to the lineage speciation in APGIII. For the AGIII class, which includes the *S-RNase*s of the Solanaceae, Plantaginaceae, and Rosaceae, along with *non-S-RNase1* (Yamane et al., 2003b) from *Prunus*, only one gene in the MRCA of eudicots was defined, supporting the single origin of the S-RNase-based GSI system. In AGIII class, a gene in the MRCA of the Rosaceae was defined (circle in Fig. 3.1).

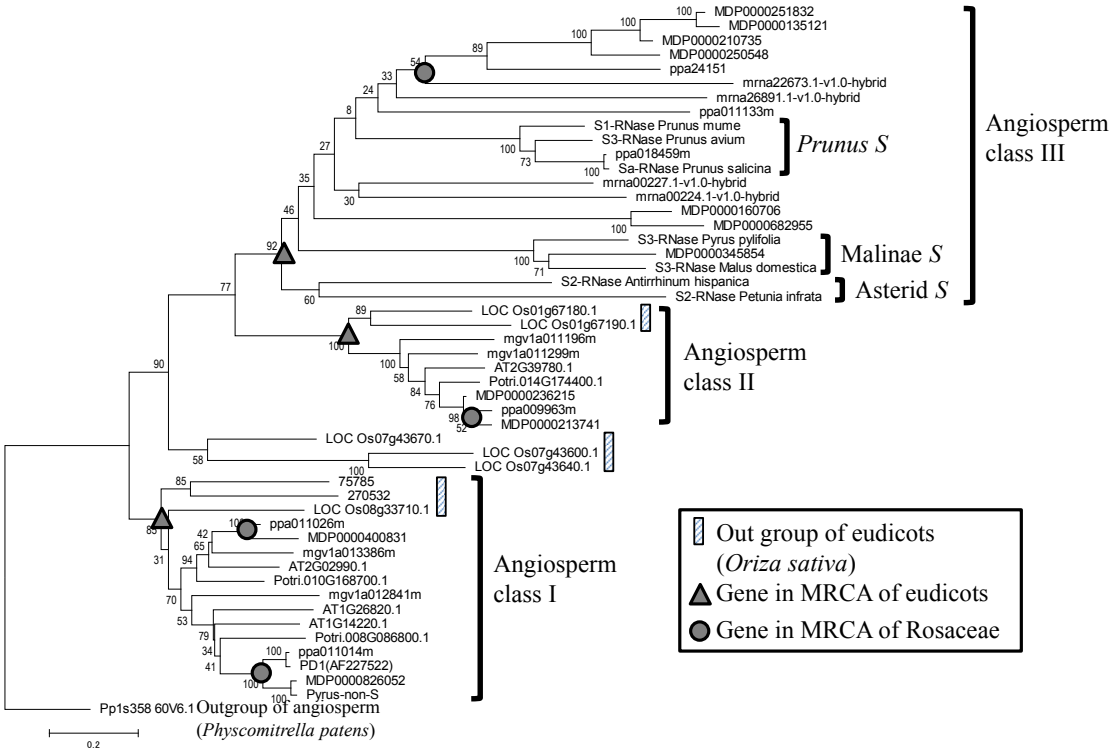

Fig. 3.1 Phylogenetic tree of *S-RNase*s and *S-RNase*-like genes in the angiosperm genome. Phylogenetic tree for *S-RNase*-like genes from seven angiosperm genomes, along with *S-RNase*s from asterids and Rosaceae, and outgroup species, *Physcomitrella patens* and *Selanginella moellendorffii*, constructed using amino acid sequences and the NJ method. The orthologous genes in eudicots are shown as hatched bars using rice (*Oryza sativa*) genes. The putative genes in the MRCA of eudicots and Rosaceae are indicated by triangles and circles, respectively. The prefixes "LOC", "mgv", "AT", "Potri", "ppa", "MDP", and "mrna" correspond to *Oryza sativa*, *Mimulus guttatus*, *Arabidopsis thaliana*, *Populus trichocarpa*, *Prunus persica*, *Malus* × *domestica* and *Fragaria vesca*, respectively.

The detailed phylogenetic tree constructed to focus on the AGIII class resulted in almost the same topology, with significant statistical support (Fig. 3.2), compared with that obtained using the all *S-RNase*-like gene sequences (Fig. 3.1). For this analysis, *S-RNase*-like sequences were retrieved from three *Fragaria* species, and several putative *S-RNase*-like sequences belonging to *Prunus S-RNase* and *non-S-RNase2* clades were detected. Again, a gene in the MRCA of the Rosaceae was defined in the *non-S-RNase2* clade, which included genes from three rosaceous species (*Fragaria, Prunus* and *Malus*) (circles in Fig. 3.2). Importantly, an additional gene in the MRCA of the Rosaceae could be defined in the *Prunus S-RNase* clade, which includes genes from *F. nubicola*. This means that there was an *S-RNase* duplication event in the ancestral genome of the Rosaceae. Moreover, the divergence of the *S-RNase*s of the subtribe Malinae and *Prunus* predated the root of the rosaceous species. However, none of the genes orthologous to the *S-RNase*s of the subtribe Malinae from the genomes of *Prunus* and *Fragaria*, or those orthologous to the S-RNase of *Prunus* from the subtribe Malinae, were identified.

	Prunus S vs. Malinae *S*	FNU99220 vs. *Prunus S*	mrna227 vs. Other genes in non-S-RNase clade2	*Prunus non-S-RNase 2* vs. Malinae *non-S-RNase 2*
Ks	0.91±0.01	0.46±0.01	0.69±0.04	0.58±0.01
4Dtv	0.39±0.03	0.20±0.01	0.43±0.02	0.27±0.02

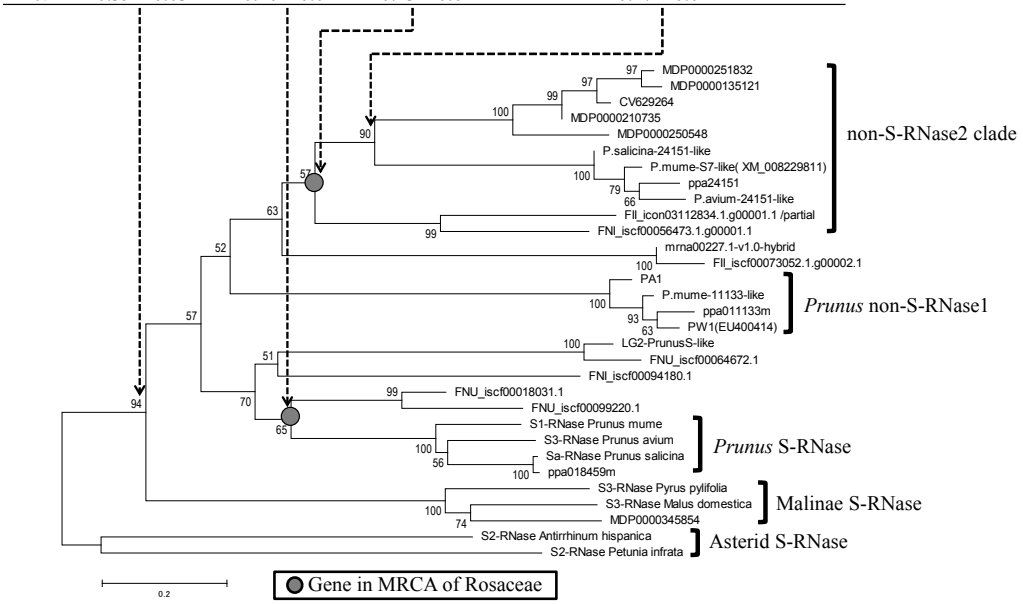

Fig. 3.2 Phylogenetic tree and genetic distance of *S-RNase*s and *S-RNase*-like genes in Angiosperm class III. Phylogenetic tree for *S-RNase*-like genes in Angiosperm class III (Fig. 3.1), along with *S-RNase*s from asterids and Rosaceae, and *non-S-RNase1* and *non-S-RNase2* genes from *Prunus*, constructed using amino acid sequences and the NJ method. The putative genes in the MRCA of the Rosaceae are indicated by circles.

3.3.2. Estimation of the time of gene divergence

The genetic distance of each gene clade in the AGIII class was calculated based on 4DTv value (Fig. 3.2). The mean 4DTv value between the *S-RNase*s from *Prunus* and the subtribe Malinae (0.39 ± 0.03) was significantly higher than the value between the *non-S-RNase2* genes from *Prunus* and the subtribe Malinae (0.27 ± 0.02), or the value between *Prunus S-RNase* and its orthologous sequence from *F. nubicola* (FNU_iscf00099220.1) (0.20 ± 0.01). In contrast, no significant difference was observed between the mean 4DTv value between the mrna00227.1-v1.0-hybrid and other genes in the *non-S-RNase2* clade (0.43 ± 0.02) and between the *S-RNase*s from *Prunus* and the subtribe Malinae (0.39 ± 0.03).

The Ks values between each gene clade showed similar results as observed in 4DTv values. The mean Ks value between the *S-RNase* genes from *Prunus* and the subtribe Malinae (0.91 ± 0.01) was significantly higher than that between the *non-S-RNase2* genes from *Prunus* and the subtribe Malinae (0.58 ± 0.01), that between the mrna00227.1-v1.0-hybrid and the other genes in the *non-S-RNase2* clade (0.69 ± 0.04), and that between *Prunus S-RNase* and its orthologous sequence from *F. nubicola* (FNU_iscf00099220.1) (0.46 ± 0.01) (Fig. 3.2). Collectively, the 4DTv and Ks values supported the topology of the phylogenetic tree (Fig. 3.2). Again, the phylogenic analysis and the time of gene divergence indicated that the current *S-RNase*s of the Rosaceae were generated by a gene duplication event that occurred before the establishment of the *Fragaria, Prunus* and Malinae, and that Malinae and *Prunus* came to use different S-RNases for their current SI systems.

*S-RNase*s in asterids and the subtribe Malinae have one intron, whereas the *Prunus S-RNase* has an additional intron located in the upstream region of the common intron present in all *S-RNase*s (Fig. 2) (Igic and Kohn, 2001). *Prunus non-S-RNase1*, however, has only one intron, at the same position as observed in the *S-RNase*s of asterids and the subtribe Malinae (Yamane et al., 2003b). The genes in the *non-S-RNase2* clade also have only one intron at the same position as in all of the *S-RNase*s, indicating that *non-S-RNase*s have maintained the original structure of *S-RNase* (data not shown). The additional intron found only in *Prunus S-RNase* may have been produced after the gene duplication that generated original *Prunus S-RNase* and *non-S-RNase*s.

3.3.3 Evolutionary patterns of pollen *S* F-box-like genes

In the regions surrounding the *S-RNase*-like genes in genomes of the Rosaceae, F-box genes that were significantly homologous to the pollen *S* F-box genes in asterids

and the subtribe Malinae were located (Fig. 3.3). The phylogenetic trees constructed from the pollen *S* F-box-like genes and their counterpart *S-RNase*s showed almost the same topologies with significant statistical support (Figs. 3.2 and 3.3), suggesting that the F-box/RNase segments of the *S* locus were duplicated in the ancestral genome of the Rosaceae. Conversely, no F-box gene was detected in the proximity of the *non-S-RNase1* gene in the *P. persica* genome.

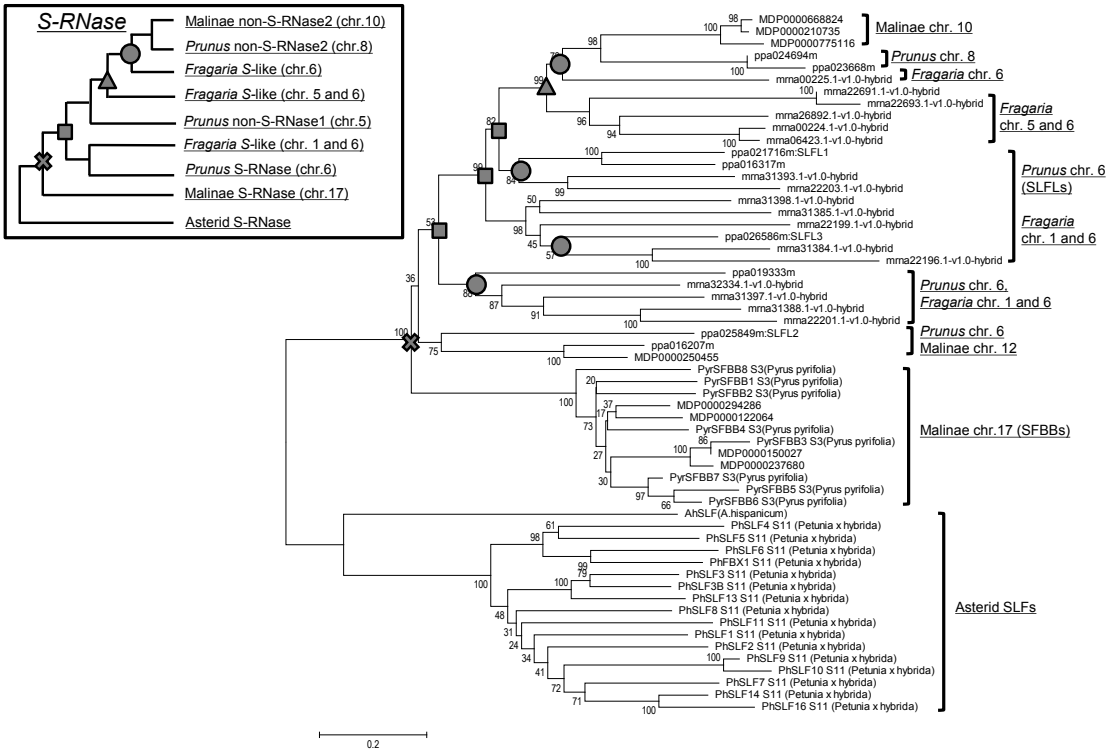

Fig. 3.3 Phylogenetic tree and gene location of pollen *S* F-box-like and pollen *S* F-box genes in *Prunus*, *Fragaria* and the subtribe Malinae in the Rosaceae, along with the asterid pollen *S* F-box gene. For pollen *S* F-box genes in the Solanaceae and subtribe Malinae, all of the *SLF*s identified in the S_{11} haplotype of *Petunia* × *hybrida* (Kubo et al., 2015) and all of the *SFBB*s from the S_3 haplotype of *Pyrus pyrifolia* (Kakui et al., 2011) were included. The phylogenetic tree was constructed using amino acid sequences and the NJ method. The pollen *S* determinant of *Prunus*, *SFB*, was not used to construct this tree because it originated from a distinct lineage, as shown in Chapter 1. The putative genes in the MRCA of the Rosaceae are indicated by circles. The symbols in the phylogenetic tree correspond to those in the schematic diagram of *S-RNase* evolution shown in the inset figure (upper left).

3.3.4. Expression of duplicated *S-RNase*-like genes in *Prunus*

To explore the possibility of changes in the cis-functions between duplicated *S-RNase*-like genes, an expression analysis was conducted with *non-S-RNase1* and *non-SRNase2* genes in *Prunus*. *Non-S-RNase1* (ppa011133m) was expressed mainly in the styles (Fig. 3.4A), but the expression patterns differed among species in *Prunus* (Fig. 3.4B). *Non-S-RNase1* expression was detected in the styles of *P. avium, P. mume*, *P. persica* and 'Santarosa' (*P. salicina*) but not in those of 'Sordum' (*P. salicina*) (Fig. 3.4A and B). The *non-S-RNase2* (ppa024151m) showed no significant expression in any tissues investigated in *P. avium* and *P. mume*, but expression was detected in the styles of *P. persica*.

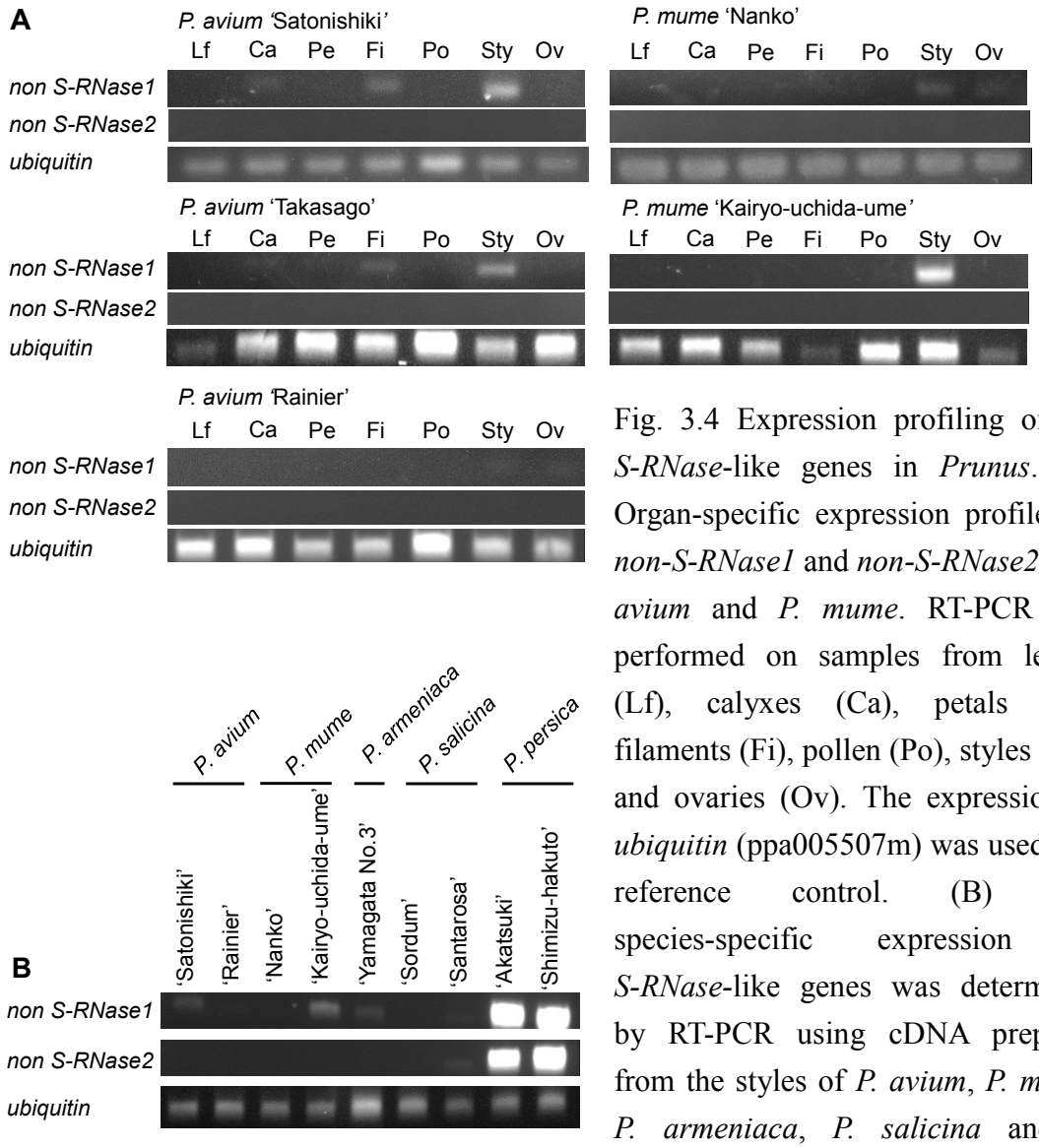

Fig. 3.4 Expression profiling of the *S-RNase*-like genes in *Prunus*. (A) Organ-specific expression profiles of *non-S-RNase1* and *non-S-RNase2* in *P. avium* and *P. mume*. RT-PCR was performed on samples from leaves (Lf), calyxes (Ca), petals (Pe), filaments (Fi), pollen (Po), styles (Sty) and ovaries (Ov). The expression of *ubiquitin* (ppa005507m) was used as a reference control. (B) The species-specific expression of *S-RNase*-like genes was determined by RT-PCR using cDNA prepared from the styles of *P. avium*, *P. mume*, *P. armeniaca*, *P. salicina* and *P. persica.*

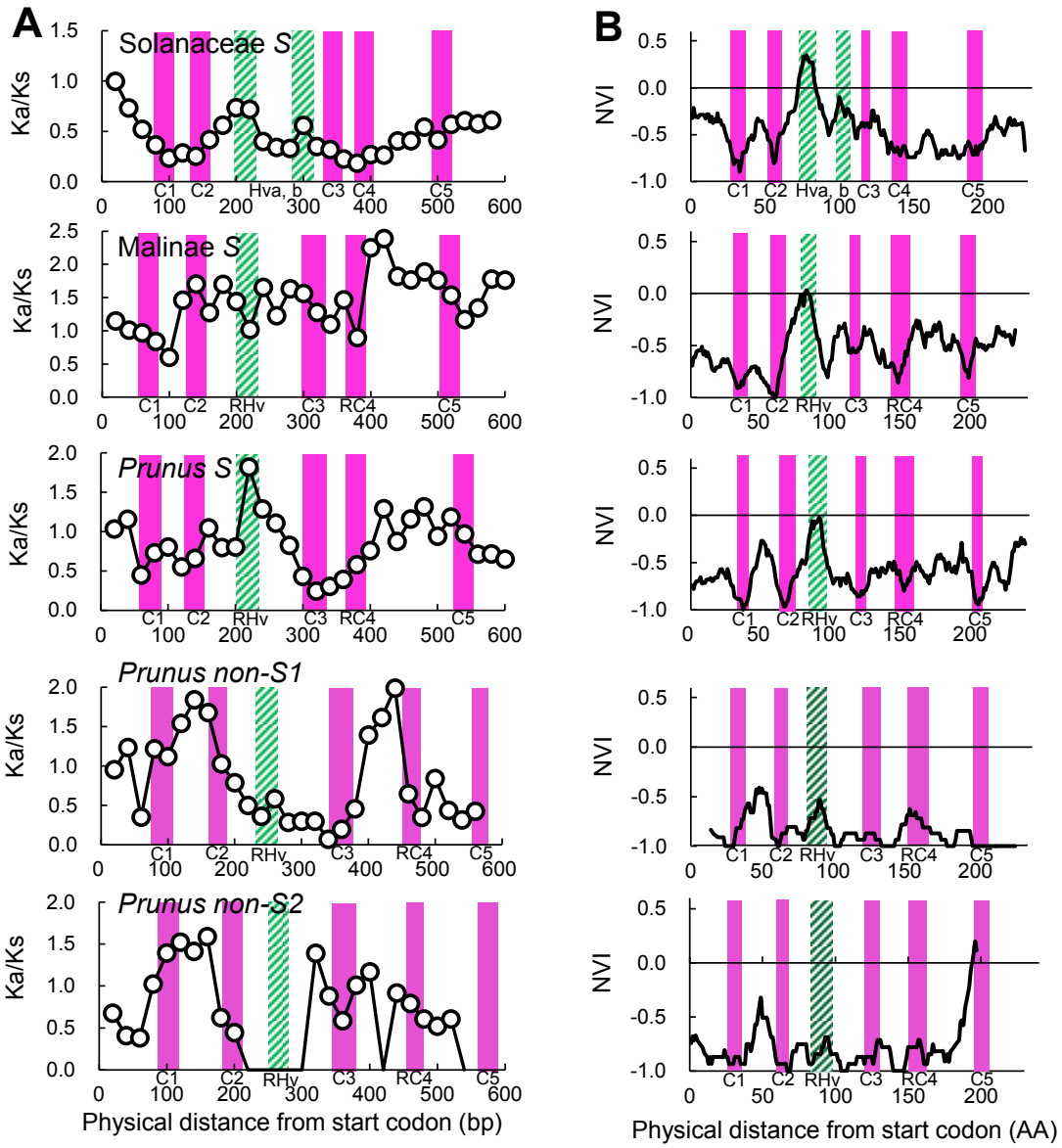

Fig. 3.5 Selective pressure on *S-RNase*s in the Solanaceae, subtribe Malinae and *Prunus* and *S-RNase*-like genes in *Prunus*. Window-average Ka/Ks ratios and NVI values for the *S-RNase* genes of the Solanaceae, subtribe Malinae and *Prunus*, and *non-S-RNase1* and *non-S-RNase2* in *Prunus*. Conserved (C1–C5) and hypervariable (HV) regions are shown by pink and green bars, respectively. For *non-S-RNase1* and *non-S-RNase2*, these regions were estimated by an alignment with the *S-RNase* sequence. (A) Window-average Ka/Ks ratios in 90-bp sliding windows with a 20-bp walking step. (B) Window-average NVI values in 11-amino acid sliding windows with a 1-AA walking step.

3.3.5. Selective pressure on duplicated *S-RNase*-like genes

Selective pressures on the *S-RNase*s in the Solanaceae, subtribe Malinae and *Prunus* and the *non-S-RNase1* and *non-S-RNase2* genes in *Prunus* were investigated in terms of the Ka/Ks and NVI. A sliding window analysis indicated that Ka/Ks ratio peaks appeared in the hypervariable region (Ishimizu et al., 1998a, b; Ushijima et al., 1998) in the *S-RNase*s of the Solanaceae and *Prunus*, while this trend was not observed in the *S-RNase* of the subtribe Malinae (Fig. 3.5A). The window-averaged plot of NVI clearly indicated that the hypervariable regions of the *S-RNase*s showed high NVI values (NVI > 0), indicating sequence diversity (or positive selection) in this region (Fig. 3.5B). In contrast, no sign of positive selection was detected in the regions of *non-S-RNase1* and *non-S-RNase2* genes that correspond to the Rosaceae hypervariable regions.

3.4. Discussion

Previous studies suggested that the S-RNase-based GSI system evolved once in eudicots (Igic and Kohn, 2001; Steinbachs and Holsinger, 2002; Vieira et al., 2008), and so do results in Chapter 3 (Figs. 3.1 and 3.2). However, segmental duplications of the *S* locus in the Rosaceae lineage were identified. Figure 3.6 shows a schematic model of the duplications of the *S-RNase* and pollen *S* F-box genes in the Rosaceae lineage. The current *S-RNase*s in asterids and the Rosaceae are thought to have originated from a single original *S-RNase*. After the divergence of the asterid and rosid ancestors, the original *S-RNase* was duplicated in the ancestral early-stage genome of the Rosaceae (duplication I in Fig. 3.6 left), producing the ancestral *S-RNase* of the subtribe Malinae (S-I). Subsequently, another ancestral *S-RNase* produced by this duplication is thought to have been duplicated again (duplication-II) to produce the ancestral *S-RNase* of *Prunus* (S-II) and the ancestral *non-S-RNase2* genes (S-III). Similar patterns of duplication and evolutionary paths have been suggested for the *S* locus F-box gene divergence. As observed in the evolutionary paths of *S-RNase*s, two duplications that generated the subtribe Malinae *SFBB* and *Prunus SLFL*s were defined by duplications-I and -II, respectively (Fig. 3.6, right). These duplication patterns indicate that the ancestral genome of the Rosaceae contained at least three *S* loci (S-I, S-II and S-III).

Recent studies suggested that duplicated genes experience subfunctionalization, mainly in their cis-functions (Liu and Adams, 2010; Roulin et al., 2013). This study demonstrated that *non-S-RNase1* and *non-S-RNase2* genes, which are *S-RNase*-like genes in *Prunus*, showed stylar-specific expression in some *Prunus* species, as do *S-RNase,* while expression was not detected in the styles of other *Prunus* species tested, indicating that pseudogenization or cis-evolution had taken place in the duplicated

genes, especially the *non-S-RNase2* gene. Regarding the maintenance of the trans-functions in duplicated *S-RNase*-like genes, no sign of positive selection on the hypervariable region was observed for *non-S-RNase1* and *non-S-RNase2* genes. Given that the hypervariable region is necessary for self/nonself recognition (Ishimizu et al., 1998a, b; Ushijima et al., 1998), the duplicated *S-RNase*-like genes are thought to have lost the self/nonself recognition function, leading to the presence a single *S* locus state in *Prunus*.

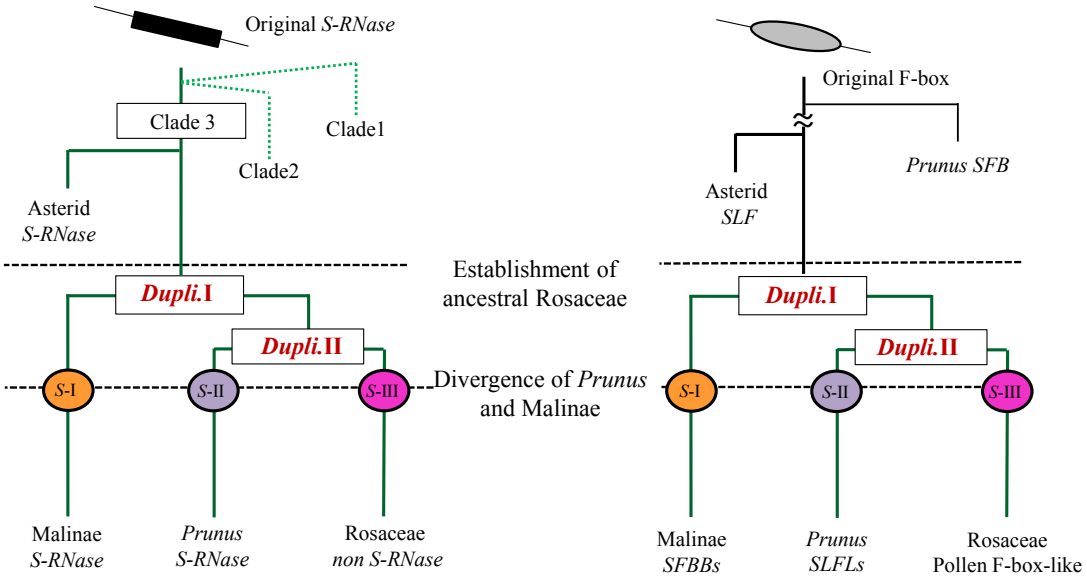

Fig. 3.6 Schematic diagram of the establishment of *S-RNase* and pollen *S* F-box genes in the Rosaceae. The current *S-RNase*s in the asterids and the Rosaceae are thought to have originated from a single original *S-RNase*. After the divergence of the asterid and rosid ancestors, the original *S-RNase* was duplicated in the ancestral early-stage genome of the Rosaceae (duplication I, left), producing the ancestral *S-RNase* of the subtribe Malinae (S-I). Subsequently, another ancestral *S-RNase* produced by this duplication is thought to have been duplicated again (duplication-II) to produce the ancestral *S-RNase* of *Prunus* (S-II) and the ancestral *non-S-RNase2* genes (S-III). Similar patterns of duplication and evolutionary paths have been suggested for the *S* locus F-box gene divergence. As observed in the evolutionary paths of *S-RNase*s, two duplications that generated the subtribe Malinae *SFBB* and *Prunus SLFL*s were defined by duplications-I and -II, respectively (right).

In conclusion, the *S* locus and its specificity determinant genes have experienced several duplications during the evolution of the Rosaceae. Furthermore, specificity for self/nonself recognition in S-RNase-based GSI systems in the Rosaceae may have evolved independently from the duplicated *S* loci, recruiting different duplicated genes, which may explain the genus (or subtribe)-specific SI recognition systems present in the Rosaceae, and the lack of phenotypic changes associated with *Prunus* SLFLs in Chapter 2. The genus (or subtribe)-specific SI recognition mechanisms, in turn, should be considered in the future development of artificial SI controls and for SC breeding.

3.5. Summary

Flowering plants have developed a genetically determined SI system to maintain genetic diversity within a species. The Solanaceae, Rosaceae and Plantaginaceae have the S-RNase-based GSI system, which uses S-RNase and F-box proteins as the pistil *S* and pollen *S* determinants, respectively. SI is associated with cultural and breeding difficulties in rosaceous fruit trees; therefore, researchers in the pomology field have long studied the mechanisms and genetics of SI to overcome these difficulties. In this Chapter, the evolutionary paths of the *S-RNase* genes were investigated by tracking their duplication patterns. A phylogenetic analysis and estimations of the time of gene divergence for *S-RNase* and its homologs in several rosaceous species showed that the divergence of *S-RNase*s in the subtribe Malinae and genus *Prunus* predated the gene in the MRCA of the Rosaceae species. Furthermore, the duplicated *S-RNase*-like genes were accompanied by duplicated pollen *S* F-box-like genes, suggesting segmental duplications of the *S* locus in the ancestral genome of the Rosaceae. An analysis of the expression patterns and selective pressure of duplicated *S-RNase*-like genes in *Prunus* suggested that these genes have lost stylar-specific expression and/or the SI recognition function, resulting in a single *S* locus state. Furthermore, the *S* loci in the current rosaceous species may have evolved independently from the duplicated *S* loci, which could explain the presence of genus (or subtribe)-specific SI recognition mechanisms in the Rosaceae. The results should be valuable for the future development of artificial SI controls and for SC breeding in rosaceous horticultural plant species.

Conclusion

Researchers in the pomology field have long studied the molecular mechanisms of rosaceous SI to overcome its associated cultural and breeding difficulties. The Rosaceae, as well as Solanaceae and Plantaginaceae, have S-RNase-based GSI, in which two tightly linked genes, S-RNase and F-box, function as pistil *S* and pollen *S* determinants, respectively. Pollen *S* F-box proteins are thought to target nonself S-RNases for degradation in the Solanaceae, Plantaginaceae and subtribe Malinae (Rosaceae). Recent studies indicated that self/nonself recognition mechanisms in *Prunus* (Rosaceae) are distinct from those in other species, raising questions about the generality of the S-RNase-based GSI system. *Prunus* pollen *S* is supposed to release the cytotoxicity of self S-RNase, which is otherwise detoxified by the GI. This study analyzed the evolution and establishment of the SI-determinant genes to further understand the molecular mechanisms underlining *Prunus*-specific SI recognition mechanisms.

In Chapter 1, the evolution of the pollen *S* determinant was investigated. A phylogenetic analysis of F-box genes with significant homology to *Prunus SFB*/*SLFL* genes showed that the *Prunus* pollen *S, SFB,* does not cluster with the pollen *S* determinants from other species (*SLF/SFBB*). This may indicate that ancestral *SFB* diverged from original pollen *S* F-box gene soon after the establishment of the eudicots. Furthermore, it has been indicated that multiple duplication events of F-box genes occurred, specifically in the Rosaceae lineage, and the *Prunus SFB* gene originate from a *Prunus*-specific gene duplication event. In contrast, *Prunus SLFL*s share the most recent common ancestor of the pollen *S* determinant with *SLF* and *SFBB*, suggesting that they might retain the functions of the original pollen *S* F-box proteins, namely S-RNase degradation. Taken together, it is concluded that a *Prunus*-specific SI recognition mechanism was established through the distinct evolution of *SFB* and that *SLFL*s are good candidates for the GI.

In Chapter 2, functional characterizations of *SLFL*s were conducted by transforming them into SI *Petunia* to test their abilities to function as the GI. Given that *Prunus* SLFLs are supposed to function in S-RNase detoxification, transgenic *Petunia* plants should become SC. Despite substantial expressions of *Prunus SLFL*s in transgenic *Petunia* pollen, no change in the SI phenotype was observed. A possible explanation for this unexpected result is that *Petunia* (Solanaceae) and *Prunus* (Rosaceae) have distinct lineages and that changes in S-RNase and/or pollen *S* F-box structures during evolution led to the non-recognition of *Petunia* S-RNase by *Prunus* SLFLs. S-RNase-like genes

may provide important information regarding the *S* locus evolution. *Non-S-RNase*s, showing homology to *S-RNase*s, are reported in *Petunia inflata* and *Nicotiana alata*. The *non S-RNase*s found in the Solanaceae are presumed to have arisen from the duplication of the *S* locus, but are not associated with current GSI system in the Solanaceae. A stylar-expressed *non-S-RNase* are also reported in *Prunus* and considered to be a possible candidate for an ancestral form of *S-RNase* in *Prunus*, although it is not involved in the current GSI system. Thus, *S-RNase*s have experienced lineage-specific evolution despite the single origin, implying that changes in SI recognition specificity may have occurred during lineage speciation.

In Chapter 3, *S* locus evolution in the rosaceous lineage was investigated to provide evidence for the divergence of SI recognition specificity. The evolution of the *S-RNase* genes was investigated by tracking their duplication patterns. A phylogenetic analysis and estimations of establishment times for *S-RNase* and its homologs in rosaceous species showed that the divergence of *S-RNase*s in Malinae and *Prunus* predated the gene in the most recent common ancestor of Rosaceae. Furthermore, the duplicated *S-RNase*-like genes were accompanied by duplicated pollen *S* F-box-like genes, indicating segmental duplications of the *S* locus in the ancestral genome of the Rosaceae. Thus, the *S* loci in the current rosaceous species may have evolved independently from the duplicated *S* loci, which could explain the presence of genus (or subtribe)-specific SI recognition mechanisms in the Rosaceae, and the lack of phenotypic changes associated with *Prunus* SLFLs in *Petunia* in Chapter 2.

Based on these results, an efficient breeding system could produce SC cultivars in rosaceous fruit tree species. Breeding individuals with loss-of-function mutations in the *SFB* is a promising approach to produce SC selections in *Prunus*, when we consider the molecular function of SFB. This could be achieved by inducing mutations in pollen grains. Pollen tubes carrying the loss-of-function mutations at *SFB* allele would have the ability to grow the full length of self-pistils because mutated SFB no longer activates or exerts self S-RNase cytotoxicity. Seedlings obtained from this cross should be SC because they carry the mutated loss-of-function *SFB* allele. Thus, the creation and selection of SC mutants can be achieved simultaneously. This strategy cannot be applied to the subtribe Malinae because loss-of-function in the *SFBB* alleles would cause self- and cross-incompatibility. Pollen *S* F-box proteins of Malinae are supposed to function in the detoxification of S-RNases. In contrast to *Prunus*, hetero-diallelic pollen with two different *S* haplotypes has the ability to detoxify all of the S-RNases in Malinae. Thus, the production of a segmental duplication of the *S* locus is the appropriate strategy to produce pollen-part SC mutants in Malinae. Taken together, we

have to take different strategy to *Prunus* and the subtribe Malinae for SC breeding.

In conclusion, the results obtained in this study support the previous indications of the presence of a *Prunus*-specific SI recognition mechanism. Furthermore, this study is the first to show dynamic *S* locus evolution, especially for the Rosaceae lineage, which may be associated with the establishment of the genus (or subtribe)-specific SI recognition mechanisms. This study should be valuable for future SC breeding programs and for developing artificial SI controls in rosaceous fruit tree species.

References

Aguiar, B., J. Vieira, A. E. Cunha, N. A. Fonseca, A. Iezzoni, S. Nocker and C. P. Vieira. 2015. Convergent evolution at the gametophytic self-incompatibility system in *Malus* and *Prunus*. PLoS One 10: e0126138.

Allen, A. M. and S. J. Hiscock. 2008. Evolution and phylogeny of self-incompatibility systems in Angiosperms. p. 73-102. In: Franklin-Tong VE, (ed.). Self-incompatibility in flowering plants. Berlin Heidelberg: Springer-Verlag.

Anderson, M. A., E. C. Cornish, S. L. Mau, E. G. Williams, R. Hoggart, A. Atkinson, I. Bönig, B. Grego, R. Simpson, P. J. Roche, J. D. Haley, J. D. Penschow, H. D. Niall, G. W. Tregear, J. P. Coghlan, R. J. Crawford and A. E. Clarke. 1986. Cloning of cDNA for a stylar glycoprotein associated with expression of self-incompatibility in *Nicotiana alata*. Nature 321: 38-44.

Anderson, M. A., G. I. McFadden, R. Bernatzky, A. Atkinson, T. Orpin, H. Dedman, G. W. Tregear, R. Fernlyey and A. E. Clarke. 1989. Sequence variability of three alleles of the self-incompatibility gene of *Nicotiana alata*. Plant Cell 1: 483-491.

Atkinson, C. J., R. M. Brennan and H. G. Jones. 2013. Declining chilling and its impact on temperate perennial crops. Env. Exp. Bot. 91: 48–62.

Blanc, G. and K. H. Wolfe. 2004. Functional divergence of duplicated genes formed by polyploidy during *Arabidopsis* evolution. Plant Cell 16: 1679–1691.

Bošković, R., D. J. Sargent and K. R. Tobutt. 2009. Genetic evidence that two independent S-loci control RNase-based self-incompatibility in diploid strawberry. J. Exp. Bot. 61: 755–763.

Brootharerts, W., G. A. Janssens, P. Proost and W. F. Broekaert. 1995. cDNA cloning and molecular analysis of two self-incompatibility alleles from apples. Plant Mol. Biol. 21: 499-511.

Brootharerts, W., J. Keulemans and I. V. Nerum. 2004. Self-fertile apple resulting from S-RNase gene silencing. Plant Cell Rep. 22: 497-501.

Carroll, S. B. 2008. Evo-Devo and an expanding evolutionary synthesis: A genetic theory of morphological evolution. Cell 134: 25-36.

Childs, K. L., J. P. Hamilton, W. Zhu, E. Ly, F. Cheung, H. Wu, P. D. Rabinowicz, C. D. Town, C. R. Buell and A. P. Chan. 2007. The TIGR plant transcript assemblies database. Nucleic Acids Res. 35: D846–D851.

Christenhusz, M. J. M. and J. W. Byng. 2016. The number of known plants species in the world and its annual increase. Phytotaxa 261: 201–217.

Cusack, B. P. and K. H. Wolfe. 2007. When gene marriages don’t work out: divorce by

subfunctionalization. Trends Genet. 23: 270–272.

Darriba, D., G. L. Taboada, R. Doallo and D. Posada. 2012. jModelTest 2: more models, new heuristics and parallel computing. Nat. Methods. 9: 772.

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 Nettancourt, D. 1997. Incompatibility in angiosperms. Sex. Plant Reprod. 10: 185-199.

de Nettancourt, D. 2001. Incompability and incongruity in wild and cultivated plants. Springer, Berlin.

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.

Finn, R. D., A. Bateman, J. Clements, P. Coggill, R. Y. Eberhardt, S. R. Eddy, A. Heger, K. Hetherington, L. Holm, J. Mistry, E. L. L. Sonnhammer, J. Tate and M. Punta. 2014. The Pfam protein families database: Nucleic Acids Res. 42: D222–D230.

Flagel, L. E. and J. F. Wendel. 2009. Gene duplication and evolutionary novelty in plants. New Phytol. 183: 557–564.

Franklin-Tong, V. E. 2008. Self-incompatibility in flowering plant. Springer, Berlin.

Gagne, J. M., B. P. Downes, S. H. Shiu, A. M. Durski and R. D. Vierstra. 2002. The F-box subunit of the SCF E3 complex is encoded by a diverse superfamily of genes in *Arabidopsis*. Proc. Natl. Acad. Sci. USA 99: 11519–11524.

Golz, J. F., A. E. Clarke, E. Newbigin and M. Anderson. 1998. A relic S-RNase is expressed in the styles of self-compatible *Nicotiana sylvestris*. Plant J. 16: 591–599.

Golz, J. F., V. Su, A. E. Clarke and E. Newbigin. 1999. A molecular description of mutations affecting the pollen components of the *Nicotiana alata S* locus. Genetics 152: 1123-1135.

Golz, J. F., H. Y. Oh, V. Su, M. Kusaba and E. Newbigin. 2001. Genetics analysis of *Nicotiana* pollen-part mutants is consistent with the presence of an S-ribonuclease inhibitor at the *S* locus. Proc. Natl. Acad. Sci. U. S. A. 98: 15372-15376.

Gouy, M., S. Guindon and O. Gascuel. 2010. SeaView version 4: a multiplatform graphical user interface for sequence alignment and phylogenetic tree building.

Mol. Biol. Evol. 27: 221–224.

Hanada, T., K. Fukakuta, H. Yamane, T. Esumi, R. Tao, T. M. Gradziel, A. M. Dandekar, M. A. Fernández, J. M. Alonso and R. Socias i Company. 2009. Cloning and characterization of a self-compatible S^f haplotype in almond [*Prunus dulcis* (Mill.) D.A. Webb. syn. *P. amygdalus* Batsch] to resolve previous confusion in its S^f*-RNase* sequence. HortScience 44: 609-613.

Hauck, N. R., H. Yamane, R. Tao and A. F. Iezzoni. 2002. Self-compatibility and incompatibility in tetraploid sour cherry (*Prunus cerasus* L.). Sex. Plant Reprod. 15, 39– 46.

Hauck, N. R., K. Ikeda, R. Tao and A. F. Iezzoni. 2006a. The mutated S_1-haplotype in sour cherry has an altered *S*-haplotype specific F-box protein gene. J. Hered. 97: 514-520.

Hauck, N. R., H. Yamane, R. Tao and A. F. Iezzoni. 2006b. Accumulation of non-functional *S*-haplotypes results in the breakdown of gametophytic self-incompatibility in tetraploid *Prunus*. Genetics 172: 1191–1198.

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

Huang, J., L. Zhao, Q. Yang and Y. Xue. 2006. AhSSK1, a novel SKP1-like protein that interacts with the S-locus F-box protein SLF. Plant J. 46: 780–793.

Ida, K., S. Norioka, M. Yamamoto, T. Kumasaka, E. Yamashita, E. Newbigin, A. E. Clarke, F. Sakiyama and M. Sato. 2001. The 1.55 Å resolution structure of *Nicotiana alata* S_{F11}-RNase associated with gametophytic self-incompatibility. J. Mol. Biol. 314: 103-112.

Igic, B. and J. R. Kohn. 2001. Evolutionary relationships among self-incompatibility *RNases*. Proc. Natl. Acad. Sci. USA 98: 13167-13171.

Ikeda, K., B. Igic, K. Ushijima, H. Yamane, N. R. Hauck, R. Nakano, H. Sassa, A. F. Iezzoni, J. R. Kohn and R. Tao. 2004. Primary structural features of the *S* haplotype-specific F-box protein, SFB, in *Prunus*. Sex. Plant Reprod. 16: 235-243.

Ioerger, T. R., J. R. Gohlke, B. Xu and T. H. Kao. 1991. Primary structural features of the self-incompatibility protein in Solananceae. Sex. Plant Reprod. 4: 81-87.

Ishimizu, T., T. Shinkawa, F. Sakiyama and S. Norioka. 1998a. Primary structural features of rosaceous S-RNases associated with gametophytic self-incompatibility. Plant Mol. Biol. 37: 931-941.

Ishimizu, T., T. Endo, Y. Yamaguchi-Kabata, K. T. Nakamura, F. Sakiyama and S.

Norioka. 1998b. Identification of regions in which positive selection may operate in S-RNase of Rosaceae: Implication for S-allele-specific recognition sites in S-RNase. FEBS Letters 440: 337-342.

Jain, M., A. Nijhawan, R. Arora, P. Agarwal, S. Ray, P. Sharma, S. Kapoor, A. K. Tyagi and J. P. Khuranal. 2007. F-Box proteins in rice. genome-wide analysis, classification, temporal and spatial gene expression during panicle and seed development, and regulation by light and abiotic stress. Plant Physiol. 143: 1467–1483.

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 pyrifolia*). Plant J. 68: 1028-1038.

Katoh, K. and D. M. Standley. 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol. Biol. Evol. 30: 772–780.

Kheyr-Pour, A., S. B. Bintrim, T. R. Ioerger and R. Remy. 1990. Sequence diversity of pistil S-proteins associated with gametophytic self-incompatibility in *Nicotiana alata*. Sex. Plant Reprod. 3: 88-97.

Kubo, K., T. Entani, A. Tanaka, 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.

Kubo, K., T. Paape, M. Hatakeyama, T. Entani, A. Takara, K. Kajihara, M. Tsukahara, R. Shimizu-Inatsugi, K. K. Shimizu and S. Takayama. 2015. Gene duplication and genetic exchange drive the evolution of S-RNase-based self-incompatibility in *Petunia*. Nature Plants 1: 14005.

Kubo, K., M. Tsukahara, S. Fujii, K. Murase, Y. Wada, T. Entani, M. Iwano and S. Takayama. 2016. Culin1-P is an essential component of non-self recognition system in self-incompatibility in *Petunia*. Plant & Cell Physiol. 57: 2403-2416.

Kuroda, S., S. Norioka, M. Mitta, I. Kato and F. Sakiyama. 1994. Primary structure of a novel stylar RNase unassociated with self-incompatibility in tobacco plant, *Nicotiana alata*. J. Protein Chem. 13: 438–439.

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: 289-296.

Lee, H. S., A. Singh and T. H. Kao. 1992. RNase X2, a pistil-specific ribonuclease from

Petunia inflata, shares sequence similarity with solanaceous *S* proteins. Plant Mol. Biol. 20: 1131−1141.

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

Lewis, D. and L. K. Crowe. 1954. The induction of self-fertility in tree fruits. J. Hort. Sci. 29: 220-225.

Li, H. and R. Durbin. 2009. Fast and accurate short read alignment with Burrows–Wheeler transform. Bioinformatics 25: 1754–1760.

Li, H., B. Handsaker, A. Wysoker, T. Fennell, J. Ruan, N. Homer, G. Marth, G. Abecasis and R. Durbin. 2009. The Sequence Alignment/Map format and SAMtools. Bioinformatics 25: 2078–2079.

Librado, P. and J. Rozas. 2009. DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. Bioinformatics 25: 1451–1452.

Liu, S.L. and K. L. Adams. 2010. Dramatic change in function and expression pattern of a gene duplicated by polyploidy created a paternal effect gene in the Brassicaceae. Mol. Biol. Evol. 27: 2817–2828.

Luedeling, E., E. H. Girvetz, M. A. Semenov and P. H. Brown. 2011. Climate change affects winter chill for temperate fruit and nut trees. PLoS ONE 6: e20155. DOI: 10.1371/journal. pone.0020155.

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

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.

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. and R. Tao. 2016a. Distinct self-recognition in the *Prunus* S-RNase-based gametophytic self-incompatibility system. Hort. J. (in press).

Matsumoto, D. and R. Tao. 2016b. Recognition of a wide-range of S-RNases by *S* locus F-box like 2, a general-inhibitor candidate in the *Prunus*-specific S-RNase-based self-incompatibility system. Plant Mol. Biol. Doi: 10.1007/s11103-016-0479-2

Matsumoto, S., K. Yamada, K. Shiratake, K. Okada and K. Abe. 2010. Structural and functional analysis of two new S-RNase alleles, *Ssi5* and *Sad5*, in apple. J. Hort. Sci. Biotech. 85: 131-136.

McClure, B. A., V. Haring, P. R. Ebert, M. A. Anderson, R. J. Simpson, F. Sakiyama and A. E. Clrke. 1989. Style self-incompatibility gene products of *Nicotiana alata* are ribonucleases. Nature 342: 955-957.

McClure, B. 2009. Darwin's foundation for investigating self-incompatibility and the progress toward a physiological model for S-RNase-based SI. J. Exp. Bot. 60: 1069–1081.

McCubbin, A. G. and T. H. Kao. 2000. Molecular recognition and response in pollen and pistil interactions. Annu. Rev. Cell Dev. Biol. 16: 333–364.

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.

Moore, R. C. and M. D. Purugganan. 2005. The evolutionary dynamics of plant duplicate genes. Curr. Opin. Plant Biol. 8: 122–128.

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

Newbigin, E., T. Paape and J. R. Kohn. 2008. RNase-based self-incompatibility: puzzled by pollen *S*. Plant Cell 20: 2286–2292.

Nowak, M. D., A. P. Davis, F. Anthony and A. D. Yoder. 2011. Expression and trans-specific polymorphism of self-incompatibility RNases in *Coffeae* (Rubiaceae). PLos One 6: e21019.

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. 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: 1887–1892.

Ortega, E., R. I. Bošković, D. J. Sargent and K. R. Tobutt. 2006. Analysis of S-RNases alleles of almond (*Prunus dulcis*): characterization of new sequences, resolution of synonyms and evidence of intragenic recombination. Mol. Gen. Genet. 276: 413-426.

Papp, B., C. Pál and L. D. Hurst. 2003. Evolution of cis-regulatory elements in duplicated genes of yeast. Trends Genet. 19: 417-422.

Potter, D., T. Eriksson, R. C. Evans, S. Oh, J. E. E. Smedmark, D. R. Morgan, M. Kerr,

K. R. Robertson, M. Arsenault, T. A. Dickinson and C. S. Campbell. 2007. Phylogenetic classification of Rosaceae. Pl. Syst. Evol. 266: 5-43.

Potts, S. G., J. C. Biesmeijer, C. Kremen, P. Neumann, O. Schweiger, and W. E. Kunin. 2010. Global pollinator declines: trends, impacts and drivers. Trends Ecol. Evol. 25: 345–353.

Puerta, A. R., K. Ushijima, T. Koba and H. Sassa. 2009. Identification and functional analysis of pistil self-incompatibility factor *HT-B* of *Petunia*. J. Exp. Bot. *60*, 1309-1318

Qiao, H., H. Wang, L. Zhao, J. Zhou, J. Huang, Y. Zhang and Y. Xue. 2004. The F-box protein Ah SLF-S_2 physically interacts with S-RNases that may be inhibited by the ubiquitin/26S proteasome pathway of protein degradation during compatible pollination in *Antirrhinum*. Plant Cell 16: 582-595.

Rastogi, S. and D. A. Liberles. 2005. Subfunctionalization of duplicated genes as a transition state to neofunctionalization. BMC Evol. Biol. 5: 28.

Ronquist, F., J. P. Huelsenbeck. 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. Bioinformatics 19: 1572-1574.

Roulin, A., P. L. Auer, M. Libault, J. Schlueter, A. Farmer, G. May, G. Stacey, R. W. Doerge and S. A. Jackson. 2013. The fate of duplicated genes in a polyploid plant genome. Plant J. 73: 143–153.

Sanzol, J. 2009. Pistil-function breakdown in a new *S*-allele of European pear, $S_{21}{}^{O}$, confer self-compatibility. Plant Cell Rep. 28: 457-467.

Sassa, H., H. Hirano and H. Ikehashi. 1992. Self-incompatibility-related RNases in styles of Japanse pear (*Pyrus serotina* Rehd.). Plant Cell Physiol. 33: 811-814.

Sassa, H., H. Hirano and H. Ikehashi. 1993. Identification and characterization of stylar glycoproteins associated with self-incompatibility genes of Japanese pear, *Pyrus serotina* Rehd. Mol. Gen. Genet. 241: 17-25.

Sassa, H., T. Nishio, Y. Kowyama, H. Hirano, T. Koba and H. Ikehashi. 1996. Self-incompatibility (*S*) alleles of the Rosacae encode members of a distinct class of the T2/S ribonuclease superfamily. Mol. Gen. Genet. 250: 547-557.

Sassa, H., H. Hirano, T. Nishio and T. Koba. 1997. Style-specific self-compatible mutation caused by deletion of the S-RNase gene in Japanese pear (*Pyrus serotina*). Plant J. 12: 223-227.

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.

Sherman-Broyles, S. and J. B. Nasrallah. 2008. Self-incompatibility and evolution of mating systems in the Brassicaceae. p. 123–147. In: Franklin-Tong V, (ed.). Self-incompatibility in flowering plants. Berlin Heidelberg: Springer-Verlag.

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.

Sonneveld, T., K. R. Tobutt, S. P. Vaughan and T. P. Robbins. 2005. Loss of pollen-*S* function in two self-compatible selections of *Prunus avium* is associated with deletion/mutation of a *S*-haplotype specific F-box gene. Plant Cell 17: 37-51.

Steinbachs, J. E. and K. E. Holsinger. 2002. *S-RNase*-mediated gametophytic self-incompatibility is ancestral in eudicots. Mol. Biol. Evol. 19: 825-829.

Sugiura, T., H. Kuroda and H. Sugiura. 2007. Influence of the current state of global warming on fruit tree growth in Japan. Hort. Res. (Japan) 6: 257–263 (In Japanese with English abstract).

Suyama, M., D. Torrents and P. Bork. 2006. PAL2NAL: robust conversion of protein sequence alignments into the corresponding codon alignments. Nucleic Acids Res. 34: W609–612.

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

Tao, R., H. Yamane, H. Sassa, H. Mori, T. M. Gradziel, A. M. Dandekar and A. Sugiura. 1997. Identification of stylar RNases associated with gametophytic self-incompatibility in almond (*Prunus dulcis*). Plant Cell Physiol. 38: 304-311.

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.

Tao, R., A. Watari, T. Hanada, T. Habu, H. Yaegaki, M. Yamaguchi and H. Yamane. 2007. Self-compatible peach (*Prunus persica*) has mutant versions of the *S* haplotypes found in self-incompatible *Prunus* species. Plant Mol. Biol. 63: 109-123.

Tao, R. and A. M. Iezzoni. 2010. The *S*-RNase-based gametophytic self-incompatibility system in *Prunus* exhibits distinct genetic and molecular features. Scientia Hort. 124: 423-433.

The angiosperm phylogeny group. 2009. An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG III.

Bot. J. Linn. Soc. 161: 105–121.

Tsukamoto, T., N. R. Hauck, R. Tao, N. Jiang and A. F. Iezzoni. 2006. Molecular characterization of three non-functional *S*-haploytpes in sour cherry (*Prunus cerasus*). Plant Mol. Biol. 62: 371-383.

Tsukamoto, T., D. Potter, R. Tao, C. P. Vieira, J. Vieira and A. F. Iezzoni. 2008. Genetic and molecular characterization of three novel *S*-haplotypes in sour cherry (*Prunus cerasus* L.). J. Exp. Bot. 59: 3169-3185.

Twell D, J. Yamaguchi, R. A. Wing, J. Ushiba and S. McCormick. 1991. Promoter analysis of genes that are coordinately expressed during pollen development reveals pollen-specific enhancer sequences and shared regulatory elements. Genes Dev. *5*, 496–507.

Ushijima, K., H. Sassa, R. Tao, H. Yamane, A. M. Dandekar, T. M. Gradziel and H. Hirano. 1998. Cloning and characterization of cDNAs encoding S-RNases from almond (*Prunus dulcis*): primery structural features and sequence diversity of the S-RNases in Rosaceae. Mol. Gen. Genet. 260: 261-268.

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, 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.

Vieira, J., N. A. Fonseca and C. P. Vieira. 2008. An *S-RNase*-based gametophytic self-incompatibility system evolved only once in eudicots. J. Mol. Evol. 67: 179-190.

Vilanova, S., M. L. Badenes, L. Burgos, J. Martinez-Calvo, G. Llacer and C. Romero. 2006. Self-compatibility of two apricot selections is associated with two pollen-part mutations of different nature. Plant Physiol. 142: 629–641.

Wang, L., L. Dong, Y. Zhang, Y. Zhang, W. Wu, X. Deng and Y. Xue. 2004. Genome-wide analysis of *S*-Locus F-box-like genes in *Arabidopsis thaliana*. Plant Mol. Biol. 56: 929-945.

Wang, Y., X. Wang, A. G. McCubbin and T. H. Kao. 2003. Genetic mapping and molecular characterization of the self-incompatibility (*S*) locus in *Petunia inflata*. Plant Mol. Biol. 53: 565-580.

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.

Watari, A., T. Hanada, H. Yamane, T. Esumi and R. Tao. 2007. A low transcriptional level of Se-RNase in the Se-haplotype confers self-compatibility in Japanese plum. J. Amer. Soc. Hort. Sci. 132: 396–406.

Williams, J. S., J. P. Der, C. W. dePamphilis and T. H. Kao. 2014. Transcriptome analysis reveals the same 17 S-locus F-box genes in two haplotypes of the self-incompatibility locus of *Petunia inflata*. Plant Cell 26: 2873–2888.

Xu, G., H. Ma, M. Nei and H. Kong. 2009. Evolution of F-box genes in plants: different modes of sequence divergence and their relationships with functional diversification. Proc. Natl. Acad. Sci. USA 106: 835–840.

Xue, Y., R. Carpenter, H. G. Dickinson and E. S. Coen. 1996. Origin of allelic diversity in *Antirrhinum S* locus RNases. Plant Cell 8: 805-814.

Yamane, H., K. Ikeda, N. R. Hauck, A. F. Iezzoni and R. Tao. 2003a. Self-incompatibility (*S*) locus region of the mutated *S^6*-haplotype of sour cherry (*Prunus cerasus*) contains a functional pollen *S* allele and a non-functional pistil *S* allele. J. Exp. Bot. 54: 2431-2437.

Yamane, H., R. Tao, H. Mori and A. Sugiura. 2003b. Identification of a non-S RNase, a possible ancestral form of S-RNases, in *Prunus*. Mol. Gen. Genomics 269: 90-100.

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.

Yang, X., U. C. Kalluri, S. Jawdy, L. E. Gunter, T. Yin, T. J. Tschaplinski, D. J. Weston, P. Ranjan and G. A. Tuskan. 2008. The F-Box gene family is expanded in herbaceous annual plants relative to woody perennial plants. Plant Physiol. 148: 1189–1200.

Zhang, Y., Z. Zhao and Y. Xue. 2009. Roles of proteolysis in plant self-incompatibility. Annu. Rev. Plant Biol. 60, 21–42.

Zhou, J., F. Wang, W. Ma, Y. Zhang, B. Han and Y. Xue. 2003. Structural and transcriptional analysis of *S*-locus F-box genes in *Antirrhinum*. Sex. Plant Reprod. 16: 165-177.

Zisovich, A. H., R. A. Stern, G. Sapir, S. Shafir and M. Goldway. 2004. The RHV region of S-RNase in European pear (*Pyrus communis*) is not required for the determination of specific pollen rejection. Sex. Plant. Reprod. 17: 151-156.

Acknowledgements

I would like to express my sincere thanks to Dr. Ryutaro Tao, Professor of Kyoto University. I wish to express my deep gratitude for his courteous supervision and invaluable discussions. I am deeply grateful to Dr. Hisayo Sawada-Yamane, Junior Associate Professor of Kyoto University, for her valuable advice. I wish to thank Dr. Takashi Akagi, Assistant Professor of Kyoto University, and Dr. Daiki Matsumoto, Assistant Professor of Yamagata University, for their technical guidance and valuable discussions. I am also grateful to Dr. Keizo Yonemori, Professor of Ryukoku University, and Dr. Kiyoshi Banno, Professor of Shinshu University, for their encouragement. Without their help, I could never have completed this study.

Special thanks must be given to all the members of the Laboratory of Pomology, Kyoto University, for their helpful assistance throughout the course of this study.

I would like to express my gratitude to Dr. Yutaka Okumoto, Professor of Kyoto University, and Dr. Ryohei Terauchi, Professor of Kyoto University, for careful reading the manuscript and for giving useful comments.

I thank to Japanese Society for the Promotion of Science for financial support for a part of this work.

Finally, I am indebted to my parents, Mr. Toshiro Morimoto and Mrs. Setsuko Morimoto, my grandfather, Shigeyoshi Morimoto, my sister, Ayako Morimoto, my brother, Keita Morimoto, and my friends for their continuous support.

Takuya Morimoto