

RESEARCH ARTICLE

Open Access



Long-read genome sequencing reveals the sequence characteristics of pear self-incompatibility locus

Chao Gu^{1*}, Ying Xu^{1†}, Lei Wu^{1†}, Xueping Wang¹, Kaijie Qi¹, Xin Qiao¹, Zewen Wang¹, Qionghou Li¹, Min He¹ and Shaoling Zhang^{1*} 

Abstract

The S-RNase-based self-incompatibility locus (S-locus) in *Petunia* species contains 16–20 *F-box* genes, which collaboratively function in the recognition and subsequent degradation of non-self S-RNases, while distinguishing them from self S-RNase. However, the number of S-locus *F-box* genes (*SFBBs*) physically interacted with non-self S-RNases remains uncertain in *Pyrus* species. Utilizing Pacbio long-read sequencing, we successfully assembled the genome of pear cultivar ‘Yali’ (*Pyrus bretschneideri*), and identified 19 *SFBBs* from the *Pyrus* *S*₁₇-locus spanning approximately 1.78 Mb. Additionally, we identified 17–21 *SFBBs* from other *Pyrus* and *Malus* S-loci spanning a range of 1.35 to 2.64 Mb. Based on the phylogenetic analysis, it was determined that *Pyrus* and *Malus* *SFBBs* could be classified into 22 groups, denoted as I to XXII. At amino acid level, *SFBBs* within a given group exhibited average identities ranged from 88.9% to 97.9%. Notably, all 19 *SFBBs* from the *S*₁₇-locus co-segregated with *S*₁₇-RNase, with 18 of them being specifically expressed in pollen. Consequently, these 18 pollen-specifically expressed *SFBBs* are considered potential candidates for the pollen-S determinant. Intriguingly, out of the 18 pollen-specifically expressed *SFBBs*, eight demonstrated interactions with at least one non-self S-RNase, while the remaining *SFBBs* failed to recognize any S-RNase. These findings provide compelling evidence supporting the existence of a collaborative non-self-recognition system governing self-incompatibility in pear species.

Keywords *Pyrus*, Self-incompatibility, S-Locus, *F-Box* gene, Non-self-recognition, Cross-recognition

Core

A total of 17–21 *S-locus F-box brothers (SFBBs)* were clarified in *Pyrus* and *Malus* S-loci. Of the 19 *SFBBs* in *Pyrus* *S*₁₇-locus, 18 are specifically expressed in pollen and eight

of them interact with non-self S-RNase. *SFBB.VIII* is a universe factor in recognizing non-self S-RNases.

Gene & accession numbers

PbrSFBB.Ia-S₁₇ accession: ON918613, PbrSFBB.Ib-S₁₇ accession: ON918621, PbrSFBB.II-S₁₇ accession: ON918614, PbrSFBB.III-S₁₇ accession: ON918615, PbrSFBB.IV-S₁₇ accession: ON918616, PbrSFBB.V-S₁₇ accession: ON918617, PbrSFBB.VI-S₁₇ accession: ON918618, PbrSFBB.VII-S₁₇ accession: ON918619, PbrSFBB.VIII-S₁₇ accession: ON918620, PbrSFBB.X-S₁₇ accession: ON918622, PbrSFBB.XI-S₁₇ accession: ON918623, PbrSFBB.XII-S₁₇ accession: ON918624, PbrSFBB.XIII-S₁₇ accession: ON918625, PbrSFBB.XIV-S₁₇ accession: ON918626, PbrSFBB.XV-S₁₇

[†]Chao Gu, Ying Xu and Lei Wu contributed equally to this work.

*Correspondence:

Chao Gu
guchao@njau.edu.cn
Shaoling Zhang
slzhang@njau.edu.cn

¹ Saya Institute of Nanjing Agricultural University, Nanjing Agricultural University, Nanjing 211800, China



accession: ON918627, PbrSFBB.XVI-S₁₇ accession: ON918628, PbrSFBB.XVII-S₁₇ accession: ON918629, PbrSFBB.XVIII-S₁₇ accession: ON918630.

Introduction

Self-incompatibility (SI) is a prevalent genetic mechanism that facilitates the rejection of self-pollen by the pistil as a means to prevent inbreeding. This mechanism is observed in around 40% of flowering plant species across at least 100 families (de Nattancourt, 2001; Igic et al. 2008). SI can be categorized into two main types: sporophytic SI, primarily found in Brassicaceae family (Hiscock and McInnis 2003), and gametophytic SI (GSI), extensively studied in Rosaceae, Solanaceae, and Plantaginaceae families (Fuji et al., 2016). In these three families, GSI is controlled by an *S* locus that comprises genes encoding an S-RNase expressed in pistils and multiple S-locus F-box (SLF/SFB) proteins expressed in pollen (Ushijima et al. 2003; Kao and Tsukamoto, 2004; Franklin-Tong 2008; Kubo et al., 2010; Sun et al. 2018). Upon entering the pollen tube, the S-RNase has the potential to impede its growth if the haploid pollen's S-haplotype is identical to any of the S-haplotypes present in the style (Chen et al. 2018). The SLF/SFB proteins, as constituents of Skp1-Cul1-F-box protein degradation complex, facilitate the degradation of non-self S-RNase through E3 ligase-mediated ubiquitination (Hua and Kao 2006).

During the GSI reaction, the SLF/SFB proteins exhibit two distinct mechanisms for recognizing S-RNase (Fuji et al., 2016). Self-recognition is observed specifically in *Prunus* species, where the unique pollen SLF/SFB protein can selectively identify self S-RNase, thereby impeding ubiquitination and degradation mediated by other SLFs (Li et al. 2020). Non-self-recognition is reported in *Petunia* species, where self S-RNase failed to be recognized by any of multiple SLFs, while non-self S-RNases are recognized by at least one SLF (Kubo et al., 2010; Song et al. 2023; Sun et al. 2018). It is worth noting that certain SLFs may not interact with any of the tested S-RNases (Sun et al. 2018), suggesting their lack of involvement in non-self-recognition. Furthermore, the non-self recognition mechanism is also suggested in pear (*Pyrus pyrifolia*) species, wherein pollen bearing the *S*₄-haplotype and lacking an *F-box* gene (*SFBB1-S*⁴) is rejected by the style containing *S*₁-haplotype (Kakui et al. 2011). However, the specific number of F-box proteins capable of recognizing *S*₁-RNase and the role of *SFBB1-S*⁴ in the degradation of *S*₁-RNase remain unclear.

The *S*-locus in *Pyrus* and *Malus* species is believed to encompass multiple *F-box* genes for each *S*-haplotype (Kakui et al. 2011). Through the utilization of bacterial artificial chromosome (BAC) libraries, a total of 12 and 10 *F-box* genes were identified from the *Malus* *S*₃- and

*S*₉-locus, respectively (Minamikawa et al. 2010). Similarly, 10 and six *F-box* genes were identified from the *Pyrus* *S*₂- and *S*₄-locus, respectively, and designated as *S*-locus *F-box* brother genes (*SFBBs*; Okada et al. 2011). In addition, a total of 27 *F-box* genes were identified from the two *Pyrus* *S*-locus (*S*₂₁ and *S*₃₄; Huang et al. 2023). In *Petunia* species, 17 *F-box* genes were identified from approximately 3.1 Mb within the *S*₂-locus (Wu et al. 2020), while 16–20 *SLF* genes were identified from other *S*-haplotypes (Williams et al. 2014; Kubo et al. 2015). However, the repetitive nature of Solanaceae *S*-locus sequences poses challenges in accurately assembling the *S*-locus using short-read sequences (Wu et al. 2020). A recent study successfully assembled the *Antirrhinum* genome using PacBio long-read sequencing, resulting in the identification of 32 *SLF* genes from two *S*-haplotypes spanning approximately 1.2 Mb (Zhu et al., 2023). Thus, PacBio long-read sequencing is a reliable method for determining the size of the *S*-locus region and identifying the *S*-locus *F-box* genes. In a previous investigation, the *S*₃₄-RNase detected in *Pyrus bretschneideri* 'Yali' was renamed as *S*₁₇-RNase due to identical sequences (Wang et al. 2017). In this study, the reference genome of 'Yali' were assembled by using PacBio long-read sequences to elucidate the size of the *S*₁₇-locus region and to identify the *F-box* genes within the region. Furthermore, protein–protein interaction was performed to investigate the cross-recognition between *SFBBs* and self/non-self S-RNases. The findings from this study contribute to a better understanding of the collaborative non-self recognition system in S-RNase-based GSI.

Results

Identification of *S*₁₇-locus using the genome sequencing of 'Yali'

To develop a de novo assembly of 'Yali' genome, we initially generated 63.57 Gb of Illumina paired-end reads. Based on the 17-mer spectrum, the estimated genome size was 545.39 Mb with a heterozygosity rate of 1.35%. Subsequently, 34.44 Gb of long-reads (a mean length of 9,721 bp) were generated using Illumina PacBio platform. The PacBio reads were self-corrected and assembled into genomic contigs using FALCON (Chin et al. 2016), and then polished with the Illumina paired-end reads using pilon software (Walker et al. 2014). The final assembly reached 534.45 Mb and consisted of 582 contigs, with an N50 of 1.81 Mb (Table S1). Of these contigs, 566 could be mapped to and almost cover all chromosomes in 'Cui-guan' genome (Fig. 1).

The completeness of the assembly was assessed using benchmarking universal single-copy orthologs (BUSCO; Simão et al. 2015), which revealed that 95.2% of the universal single-copy orthologs were present (Table S1),

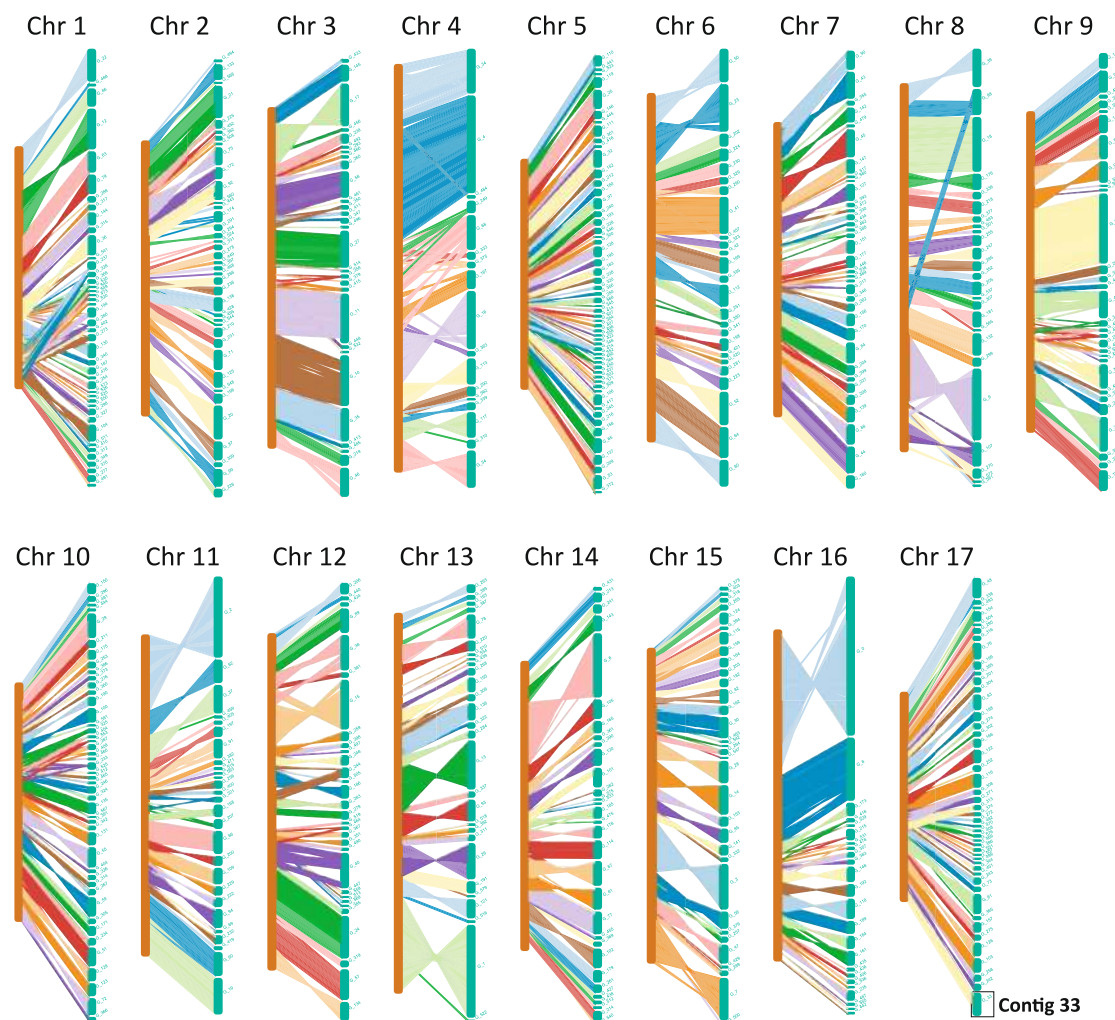


Fig. 1 Mapping of 'Yali' contigs to the 'Cuiguan' genome. The columns with brown color are the chromosomes of 'Cuiguan' genome, while the columns with green color are the contigs of 'Yali' assembly. In each chromosome, the contigs are linked with the homologous regions of 'Cuiguan' genome by the lines with different colors. A total of 566 contigs could be mapped to the 'Cuiguan' genome

indicating a reliable assembly. A total of 262.46 Mb of repetitive elements accounted for 49.11% of assembly (Tables S2), and approximately 293.95 Kb of non-coding RNAs accounted for 0.06% of the assembly (Tables S3). Moreover, the total of 2438 non-coding RNAs were identified, including 745 microRNAs, 720 transfer RNAs, 155 ribosomal RNAs, and 818 small nuclear RNAs. Furthermore, a total of 45,092 protein-coding genes were predicted from the assembly (Table S1). Notably, the *S₁₇-RNase* was annotated in the contig 33 (Fig. 1). These results suggest a reliable completeness of the assembly.

Identification of the F-box genes in *Pyrus S₁₇*-locus

To identify the *F-box* genes within the *S₁₇*-locus, a blasting search was performed in the contig 33 of the 'Yali' genomes using the conserved F-box motif (Figure S1) of

Pyrus and *Malus* *SFBBs* as query sequences. This search yielded 19 *SFBBs*, each encoding more than 150 amino acids, within the contig 33 (Fig. 2a; Table S4), which is a larger number than the 15 *SFBBs* reported in a previous study that identified *SFBBs* within the same *S*-locus (Huang et al. 2023). Notably, these *SFBBs* were found within the region from 0.94 Mb upstream and 0.83 Mb downstream of *S₁₇-RNase* (Tables S4 and S5). This finding indicates that *Pyrus* *SFBBs* are distributed within a genomic region of less than 4 Mb, probably.

Consequently, a subsequent blasting search was performed in the region from approximately 2 Mb upstream to 2 Mb downstream of the *S-RNase* in the *Pyrus* and *Malus* reference genomes that has been previously assembled using long-read sequences (Daccord et al. 2017; Linsmith et al., 2019; Zhang et al. 2019; Dong

et al. 2020; Sun et al. 2020; Gao et al. 2021). The results showed that in *Pyrus* species, 19, 17, and 20 *SFBBs* were identified from *P. communis* S_{101} -locus (PcS_{101} -locus), *P. pyrifolia* S_5 -locus ($PpyS_5$ -locus), and *P. betuleafolia* S_{67} -locus ($PbeS_{67}$ -locus), respectively (Fig. 2a; Table S4). In the case of *Malus domestica*, 21, 21, 20, and 20 *SFBBs* were identified from S_2 -, S_3 -, S_5 -, and S_9 -locus, respectively (Fig. 2a; Table S4). Furthermore, the short-read sequences from the *P. bretschneideri* cultivar ‘Dangshan-suli’ (S_7S_{17}) genome (Shi et al., 2019; Wu et al., 2013) were mapped onto contig 33, and 20 *SFBBs* were identified from the remaining S_7 -locus (Fig. 2a; Table S4). These findings indicate that *Pyrus* and *Malus* *S*-loci contain 17–21 *SFBBs*. In contrast, *Prunus salicina* S_{33} -locus, *Prunus avium* S_9 -locus, and *Prunus armeniaca* S_{24} -locus displayed the identification of 14, 12, and 12 *SLFs/SFBs*, respectively (Fig. 2a; Table S4).

Sequence analysis of protein-coding *SFBBs*

Phylogenetic classification of the 177 *SFBBs* derived from *Pyrus* and *Malus* *S*-haplotypes was performed based on their amino acid sequences. The *SFBBs* were classified into 22 distinct groups, denoted as groups I to XXII (Fig. 2b). Notably, groups I to VIII were found to correspond to the previously reported groups 1 to 8 (Kakui et al. 2011). Among these groups, group I exhibited further subdivision into sub-groups Ia, Ib, and Ic (Fig. 2b). As shown in Fig. 2a, *SFBB.XVIII* is positioned closest to the *SFBBs* in group XXI and/or XXII at one end of the *S*-locus in any *S*-haplotype, while *SFBB.XVII* is found closest to *SFBB.XVI* at another end but is absent in *Pyrus* *S*-locus. *SFBBXXI* is located between *SFBBVIII* and *SFBBXXII*, though it has been occasionally deleted. The remaining *SFBBs* appear to be randomly arranged within *Pyrus* and *Malus* *S*-loci (Fig. 2a). In contrast, the *F-box* genes within *Prunus* *S*-locus were grouped into 11 clusters, designated as *SLF1* to *SLF10* and *SFB* (Figure S2), with relatively stable positions (Fig. 2a).

The nucleotide acid sequence of all *F-box* genes present in the ‘Yali’ genome was utilized for phylogenetic analysis in conjunction with the 27 *SFBBs* previously identified in ‘Yali’ pollen (Huang et al. 2023). The results showed that all 27 *SFBBs* were grouped together with the

19 *SFBBs* identified in the *Pyrus* S_{17} -locus (Figure S3). Furthermore, the *SFBBs* previously identified from the *Pyrus* S_{34} -locus (which had been renamed as S_{17} -locus; Wang et al. 2017) were closest clustered with those identified in the ‘Yali’ genome (Figure S3). Remarkably, the bootstrap values reached 100 for *PbrSFBB.XIX-S₁₇* and *PbrSLF16-S₂₁*, *PbrSFBB.VI-S₁₇* and *PbrSLF6-S₂₁*, as well as *PbrSFBB.VIII-S₁₇* and *PbrSFBB8-S₂₁* (Figure S3). This high bootstrap value indicates that the three *SFBB* genes, namely *PbrSLF6-S₂₁*, *PbrSFBB8-S₂₁*, and *PbrSLF16-S₂₁*, are potentially located within the S_{17} -locus.

The level of sequence polymorphism observed within each group of *SFBBs* was relatively low. At nucleotide acid level, the average identities ranged from 92.1% to 98.4%, while at amino acid level, the average identities ranged from 88.9% to 97.9% (Tables 1 and S6). The sequence identities among *SFBBs* within each group were similar to those observed among *Prunus* *SLFs*, but greater than those observed among *Prunus* *SFBs* and *Pyrus* *S-RNases* (Tables S7 to S9). Notably, the inter-group identity of *SFBBs* was significantly lower than the intra-group identity of *SFBBs* (p -value = 0.0022 < 0.05). Across groups I to XXII, the average identities among *SFBBs* ranged from 12.3% to 84.0% at nucleotide acid level and from 0.06% to 71.8% at amino acid level (Tables 1 and S6). These values were closely resemble the sequence identities observed among *S-RNases* (Table 1).

Long-read sequence analysis was utilized to examine gene duplication events in the *Pyrus* and *Malus* genomes. The findings revealed the presence of dispersed, proximal, tandem, and transposed duplications within various *S*-loci, such as *Pyrus* S_{17} -locus and other *S*-locus, as well as *Malus* S_3 -locus and *Pyrus* S_5 -, S_{67} -, and S_{101} -locus (Figure S4; Table S10). For instance, in the *Pyrus* S_{17} -locus, the tandem duplicate *SFBB.X* and *SFBB.XI* were found adjacent to each other without any intervening genes. The proximal duplicate *SFBB.X* and *SFBB.Ib* were separated by a few genes, and the dispersed duplicate *SFBB.XI* and *SFBB.XIII* were not adjacent and did not share a homologous segment. In the *Pyrus* S_{67} -locus, it was observed that the transposed duplicate *SFBB.XI* and *SFBB.XIV* may have been relocated by DNA transposons. Additionally, through synteny analysis, it was determined

(See figure on next page.)

Fig. 2 Sequence analysis of the *F-box* genes in *Pyrus* and *Malus* *S*-loci. **a** Arrangement of *F-box* genes in *Pyrus*, *Malus*, and *Prunus* *S*-loci. The direction of gene transcription is indicated by arrowheads, wherein black-colored arrowheads represent the *S-RNase* gene, and other colored arrowheads represent *S*-locus *F-box* genes. The characters above the arrowheads denote the classification of *S*-locus *F-box* genes in *Pyrus*, *Malus*, and *Prunus* species. The physical distance (in kilobases, Kb) between adjacent *F-box* genes is indicated below the dotted line, while the chromosome locations of these *F-box* genes are provided in Table S4. It should be noted that the abbreviation “SLF” specifically applies to *S*-locus *F-box* genes in Solanaceae, Plantaginaceae, and *Prunus*. **b** Phylogenetic classification of *S*-locus *F-box* brother (*SFBB*) proteins in *Pyrus* and *Malus*. The I → XXII labels signify the different groups of classified *S*-locus *F-box* proteins, with Group I including the sub-groups Ia, Ib, and Ic. All groups and sub-groups are highlighted in distinct colors

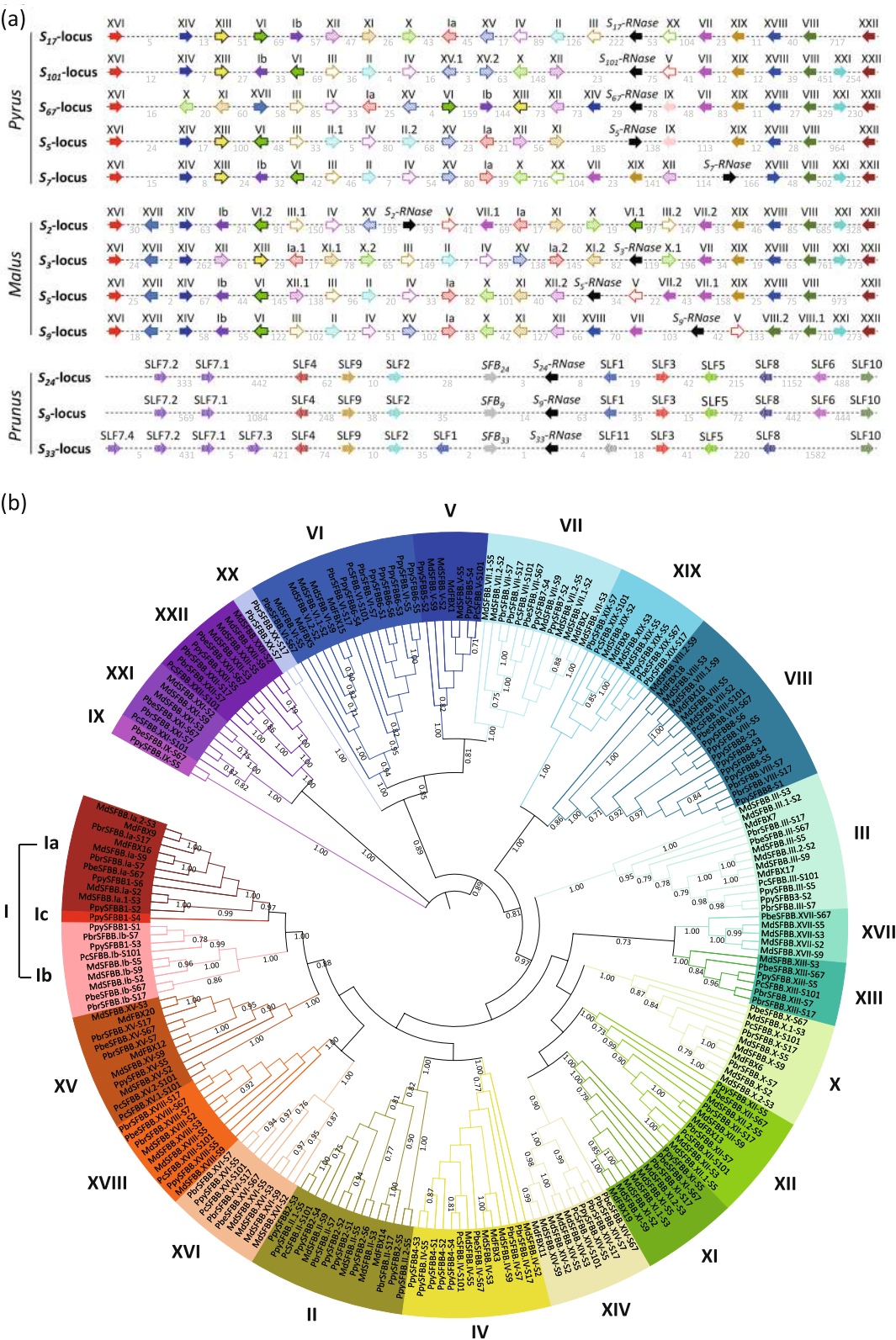


Fig. 2 (See legend on previous page.)

Table 1 Sequence similarity of *F-box* genes in *Pyrus*, *Malus*, and *Prunus* *S*-haplotypes

Group			Amino acid (%)			Nucleotide acid (%)		
Type	A	B	Min	Max	Average	Min	Max	Average
Intra-group	Ia	Ia	88.7	100.0	92.5	93.0	100.0	95.5
	Ib	Ib	93.0	100.0	95.5	96.4	99.8	97.5
	II	II	84.1	95.9	90.1	88.2	97.1	93.7
	III	III	89.5	100.0	93.5	94.3	100.0	96.4
	IV	IV	89.5	95.1	92.8	94.7	97.6	96.1
	V	V	95.1	96.6	95.7	96.6	97.2	96.8
	VI	VI	91.5	98.9	94.6	94.4	99.2	96.8
	VII	VII	78.9	100.0	90.5	86.4	100.0	93.7
	VIII	VIII	77.8	99.7	91.0	87.1	99.4	95.5
	X	X	87.7	100.0	93.3	86.4	96.9	93.3
	XI	XI	79.3	100.0	88.9	85.8	100.0	92.1
	XII	XII	86.5	100.0	91.0	90.1	100.0	93.5
	XIII	XIII	88.2	97.9	94.3	94.1	99.1	96.7
	XIV	XIV	83.0	100.0	93.3	89.7	100.0	95.5
	XV	XV	88.1	99.7	92.3	93.9	99.8	95.3
	XVI	XVI	96.4	100.0	97.6	97.5	100.0	98.4
	XVII	XVII	96.2	100.0	97.9	97.0	100.0	98.3
	XVIII	XVIII	94.6	99.2	97.5	95.7	99.3	97.8
	XIX	XIX	92.4	97.5	95.3	96.3	99.6	97.6
	XXI	XXI	95.0	100.0	96.5	94.6	99.8	97.2
	XXII	XXII	91.7	100.0	95.0	94.3	100.0	97.0
Inter-group	I→VIII	I→VIII	55.9	92.6	69.5	55.7	93.2	77.8
	I→VIII	IX	36.0	98.1	57.8	77.4	99.1	84.0
	I→VIII	X→XX	50.3	82.4	68.3	49.6	89.5	80.2
	X→XX	IX	21.1	89.9	56.0	40.9	88.6	74.4
	X→XX	X→XX	50.1	87.3	71.8	50.0	89.7	81.6
	XXI→XXII	I→VIII	4.1	13.4	8.5	18.7	33.1	21.8
	XXI→XXII	IX	5.7	37.8	20.3	9.9	33.5	21.5
	XXI→XXII	X→XX	3.7	37.8	0.06	12.5	37.5	22.3
XXI	XXII	3.4	5.4	3.9	9.8	22.8	12.3	
S-RNase			62.0	78.0	67.3	77.5	87.6	81.0

that deletions, inversions, and/or translocations of these *SFBBs* were frequently observed between any two *S*-locus (Figures S5 to S11).

Comparison analysis of the non-coding flanking sequences of *SFBBs*

In order to investigate the sequence polymorphism of the non-coding flanking sequences of *SFBBs*, an analysis was conducted on the 5 kb upstream and 5 kb downstream sequences of *SFBBs* within each *S*-locus and group (Figures S12 to S40). The results showed that the identities between either the upstream or downstream sequences of all *SFBBs* in any *S*-locus were lower (Figs. 3a and S12 to 19), compared to those between all *SFBBs* within any group (Figs. 3b and S20 to S40; Table S11). Particularly,

in group XXII, the downstream sequences displayed extremely high identities (> 90%) between any two *SFBBs* (Figs. 3b and S40d), as well as the upstream sequences between any two *Malus SFBBs* (Figure S40b), which closely resembled those found in the coding sequences of *SFBBs* (Table 1). However, although a few pairwise *SFBBs* in groups III (Figure S23b,d), VIII (Figure S28b,d), XI (Figure S30b,d), XII (Figure S31b,d), XIV (Figure S33b,d), XVI (Figure S35b,d), XVII (Figure S36b,d), XVIII (Figure S37b,d), and XIX (Figure S38b,d), as well as sub-group Ia (Figure S20b,d), exhibited extremely high identities in either the upstream or downstream sequences, the average identity in the upstream or downstream sequences between *SFBBs* in any group (excluding XXII) was lower than in the coding sequences of *SFBBs* (Figs. 3b;

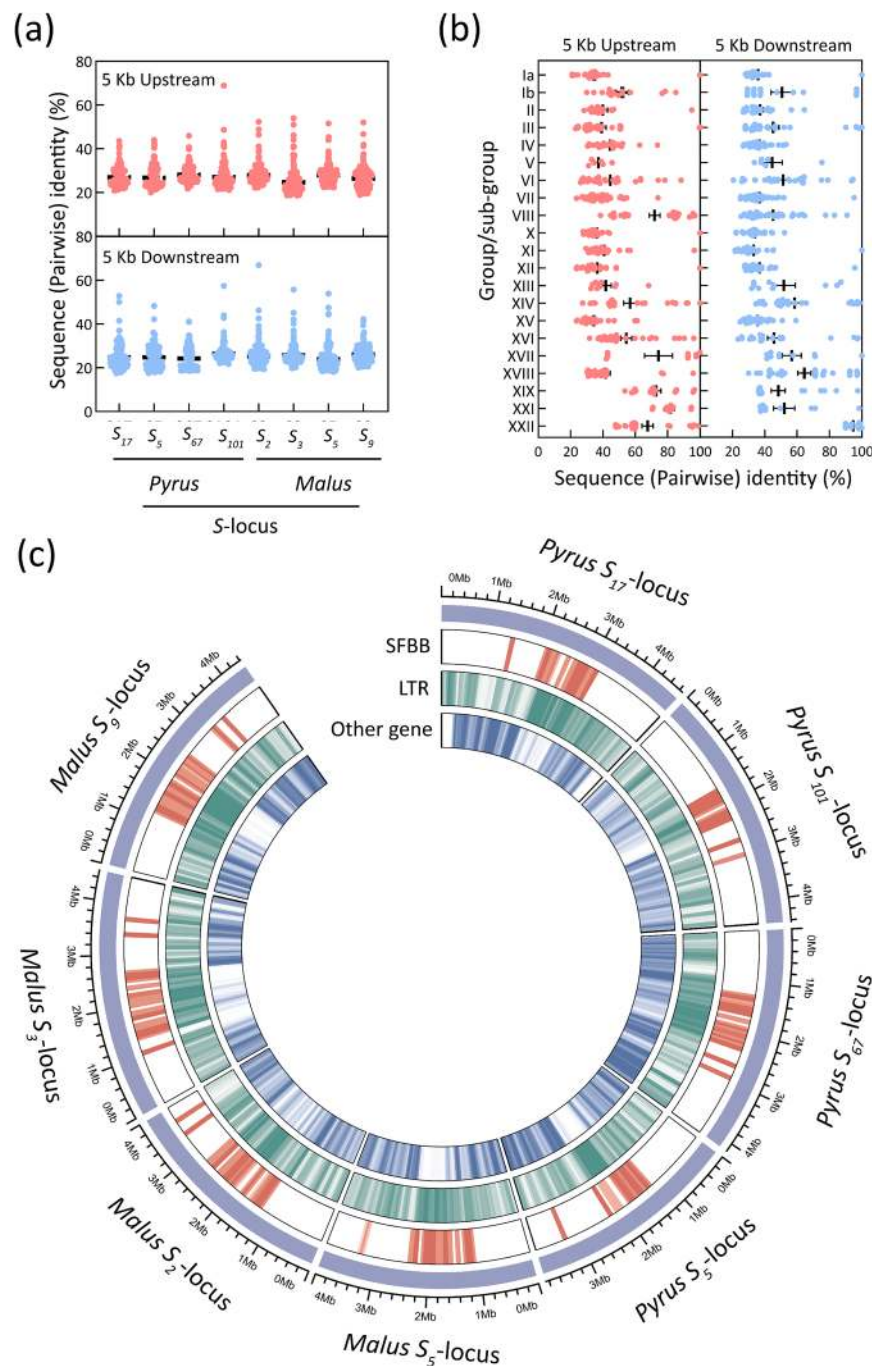


Fig. 3 Comparison analysis of the non-coding sequences in *Pyrus* and *Malus* S-loci. **a** Pairwise identity of the 5 kb upstream and downstream sequences of SFBBs within each group and sub-group. The specific data can be found in Figures S11 to S31. **b** Pairwise identity of the 5 kb upstream and downstream sequences of SFBBs in *Pyrus* and *Malus* S-loci. The corresponding data can be found in Figures S32 to S39. **c** Location of SFBBs, LTRs, and other genes within *Pyrus* and *Malus* S-loci. The outer region is the length of DNA sequences that contain S-locus. SFBBs, LTRs, and other genes are depicted in red, green, and blue colors, respectively. The phylogenetic tree was constructed using MEGA5 with the neighbor-joining method. Bootstrap values were calculated from 1000 replicates

Tables 1 and S11). These results demonstrate that genetic variation frequently occurs in the non-coding flanking sequences of *Pyrus* and *Malus* SFBBs.

The significant role of long terminal repeat (LTR) retrotransposon in genetic recombination is widely acknowledged (Kumar and Bennetzen 1999). In this study, a

survey of LTR retrotransposons within the *S*-loci of *Pyrus* and *Malus* species revealed that the total length of LTR retrotransposons accounted for approximately 13.49% of the tested regions (Table S12). A total of 568, 459, 521, and 515 LTR retrotransposons were detected in *Pyrus* *S*₁₇, *S*₅, *S*₆₇, and *S*₁₀₁-locus, respectively, and 459, 633, 665, and 650 LTR retrotransposons were detected in *Malus* *S*₂, *S*₃, *S*₅, and *S*₉-locus, respectively (Tables S12 and S13). On average, the number of LTR retrotransposons was 1.64 times (p -value < 0.001) that of the genes in the tested region (Table S12). Notably, the number of LTR retrotransposons was 4.26 times (p -value < 0.001) that of the genes in the region between *SFBB.XVI* and *SFBB.VIII*, but only 1.09 times (p -value = 0.55) that of the genes in the other region (Table S12). This result indicated that the density of LTR retrotransposons in the region between *SFBB.XVI* and *SFBB.VIII* was higher (3.89-fold; p -value < 0.005) compared to the other region (Fig. 3c). The elevated density of LTR retrotransposons between *SFBB.XVI* and *SFBB.VIII* may be associated with the genetic variations, such as gene duplication, deletion, inversion, and translocation of *SFBBs* in *Pyrus* and *Malus* species.

Inheritance and tissue specific expression of *SFBBs* in *Pyrus* *S*₁₇-locus

To investigate the inheritance pattern of *SFBBs* within the *Pyrus* *S*₁₇-locus, a total of 42 individuals were selected from the progeny resulting from a cross between ‘Yali’ (*S*₁₇*S*₂₁) and ‘Xueqing’ (*S*₃*S*₁₆). By analyzing the segregation of *S*-*RNase* alleles, four different *S*-genotypes in this progeny, which aligned with the expected 1:1:1:1 ratio ($\chi^2 = 0.286 < \chi^2_{0.05, 3} = 7.815$; Figure S41). Furthermore, 19 pairs of primer were employed to amplify the full-length sequences of the 19 *SFBBs* within the *Pyrus* *S*₁₇-locus. The resultant amplification products were subsequently cloned and sequenced. Alignment of the sequences revealed the presence of all *SFBBs* exclusively in the hybrids harboring the *S*₁₇-*RNase* allele (Fig. 4a and S42 to S60). This finding strongly indicates that the *SFBBs* within the *Pyrus* *S*₁₇-locus exhibit co-segregation with *S*₁₇-*RNase*.

To reveal the expression pattern of the 19 *SFBBs* within the *Pyrus* *S*₁₇-locus across various tissues, the transcriptome sequence reads from ‘Dangshansuli’ were utilized (Shi et al. 2017; Qiao et al. 2018; Li et al., 2019; Wu et al. 2023). These reads were previously generated and mapped to the genome of ‘Yali’. The analysis showed that 18 *SFBBs* exhibited specific expression in pollen, whereas *S*₁₇-*RNase* displayed specific expression in the style (Fig. 4b and S61). Notably, the remaining *SFBB*, *PbSFBB.XXII-S*₁₇, exhibited expression in the style, root, stem, petal, leaf, and fruit, but not in pollen (Figs. 4b and S61).

These findings suggest that the 18 *SFBBs* specifically expressed in pollen are the candidates of pollen-*S* determinant and may directly interact with non-self *S*-*RNase*.

Physical interaction between *S*-*RNases* and *SFBBs*

To examine the potential interaction between *SFBBs* and self/non-self *S*-*RNases*, yeast-two-hybrid (Y2H) and firefly luciferase complementation imaging (FLCI) assays were conducted. The Y2H assay demonstrated the absence of self-activation for any *S*-*RNase* or *SFBB* within yeast cells (Figure S62). Notably, none of the *SFBBs* expressed in the *S*₁₇-locus exhibited interaction with *S*₁₇-*RNase* (Fig. 5a; Table 2). In contrast, eight *SFBBs*, *PbrSFBB.Ia-S*₁₇, *PbrSFBB.Ib-S*₁₇, *PbrSFBB.II-S*₁₇, *PbrSFBB.VIII-S*₁₇, *PbrSFBB.X-S*₁₇, *PbrSFBB.XI-S*₁₇, *PbrSFBB.XII-S*₁₇, and *PbrSFBB.XVIII-S*₁₇, were found to interact with at least one *S*-*RNase* (Fig. 5a, b). This finding further supports the involvement of a collaborative non-self-recognition system in the regulation of *S*-*RNase*-based GSI in pear. However, it is worth noting that 10 *SFBBs* did not exhibit any interaction with the tested *S*-*RNases* (Fig. 5a). This outcome may be attributed to the limited number of non-self *S*-*RNases* utilized in this particular study.

Previous studies have showed the significance of the hypervariable (HV) region of *S*-*RNase* in mediating the recognition between *S*-*RNase* and pollen-*S* determinant (Matton et al. 1997, 1999; Wu et al. 2013). To test whether the HV region of *Pyrus* *S*-*RNase* is involved in the interaction between *SFBBs* and non-self *S*-*RNases*, an Y2H assay was conducted after removing the HV region from *S*₅- and *S*₇-*RNases*. The results showed that *PbrSFBB.VIII-S*₁₇ was unable to interact with the HV region-deleted *S*₅- and *S*₇-*RNases* within yeast cells (Fig. 5c). This outcome suggests that the HV region of *Pyrus* *S*-*RNase* plays a crucial role in facilitating the cross-recognition of non-self *S*-*RNases* by the interacting *SFBBs*.

To clarify the sub-cellular localization of the eight *SFBBs* physically interacting with non-self *S*-*RNases*, the *SFBBs* were fused with the green fluorescence protein GFP (*SFBB::GFP*). *Arabidopsis* aquaporin PIP2A fused with red fluorescence marker mCherry (*PIP2A::mCherry*) was used to show plasma membrane. Image analysis showed that in the control tobacco mesophyll cells co-transformed with *PIP2A::mCherry* and GFP empty vector, the GFP was detected in all visible tissues, while the *PIP2A::mCherry* was only detected in cell and nuclear membranes (Fig. 6). This result contributed to the presence of yellow fluorescence in cell and nuclear membranes in the merged images (Fig. 6). In the tobacco cells co-transformed with *SFBB::GFP* and *PIP2A::mCherry*, the GFP was detected in the inner side of yellow fluorescence in the merged images (Fig. 6), suggesting that all eight *SFBBs* may be localized in cytoplasm



Fig. 4 The co-segregation of pollen-specifically expressed *SFBBs* from the *Pyrus* S_{17} -locus with S_{17} -*RNase*. **a** A diagram shows the alignments of 19 *SFBBs* in the ‘Yali’ genome and in the hybrids containing S_{17} -*RNase*. **b** Integrative Genomics Viewer (IGV) snapshots show the transcript levels of *S-RNase* and *SFBB* genes from the *Pyrus* S_{17} -locus in various pear tissues. The pollen-specifically expressed *SFBBs* are highlighted in pink color. Detailed sequence alignments can be found in Figures S42 to S60

but not cell wall or cell membrane. To confirm it, protoplasts were extracted from the tobacco mesophyll cells. In the protoplast transformed with GFP empty vector, the GFP was detected in all visible tissues including cell membrane, cytoplasm, and nucleus (Fig. 6). In the protoplast transformed with *SFBB::GFP*, the GFP was only detected in cytoplasm (Fig. 6), suggesting that all eight *SFBBs* are localized in cytoplasm.

The other genes in *Pyrus* S_{17} -locus

In the genome of ‘Yali’, besides *S-RNase* and *SFBBs*, a total of 114 genes were found between *PbrSFBB.XVI-S₁₇* and *PbrSFBB.XXII-S₁₇* (Table S5). Similarly, between *SFBB.XVI* and *PbSFBB.XXII* in *Pyrus* S_{101} , *Pyrus* S_{67} , *Pyrus* S_{5-} , *Malus* S_{2-} , *Malus* S_{3-} , *Malus* S_{5-} , and *Malus* S_{9-} loci, there were 79, 153, 127, 156, 136, 138, and 115 genes, respectively (Table S5). Notably, 29 of these

genes were shared between the *Pyrus* S_{17} -locus and other *Pyrus* and *Malus* *S*-loci (Fig. 7a). Additionally, these 29 genes were consistently located between *SFBB.VIII* and *SFBB.XXII* in all *S*-locus. Transcriptome data analysis of various tissues revealed that 11 genes, *Serine/threonine-protein kinase CTR1* (evm.TU.G_33.131), *50S ribosomal protein L1* (evm.TU.G_33.130), *zuotin-like* (evm.TU.G_33.107), *Ras-related protein Rab11D* (evm.TU.G_33.106), *ATP-dependent zinc metalloprotease FTSH2* (evm.TU.G_33.103), *Receptor protein kinase TMK1* (evm.TU.G_33.90), *U-box domain-containing protein 21* (evm.TU.G_33.89), *DELLA protein GAI* (evm.TU.G_33.76), *Protease Do-like 7* (evm.TU.G_33.37.2), *Phosphatidylinositol 4-kinase beta 1* (evm.TU.G_33.36), and an *uncharacterized protein* (evm.TU.G_33.130) were expressed (RPKM > 2) in style, but not in pollen (Fig. 7b; Table S14). These genes may be associated with the style

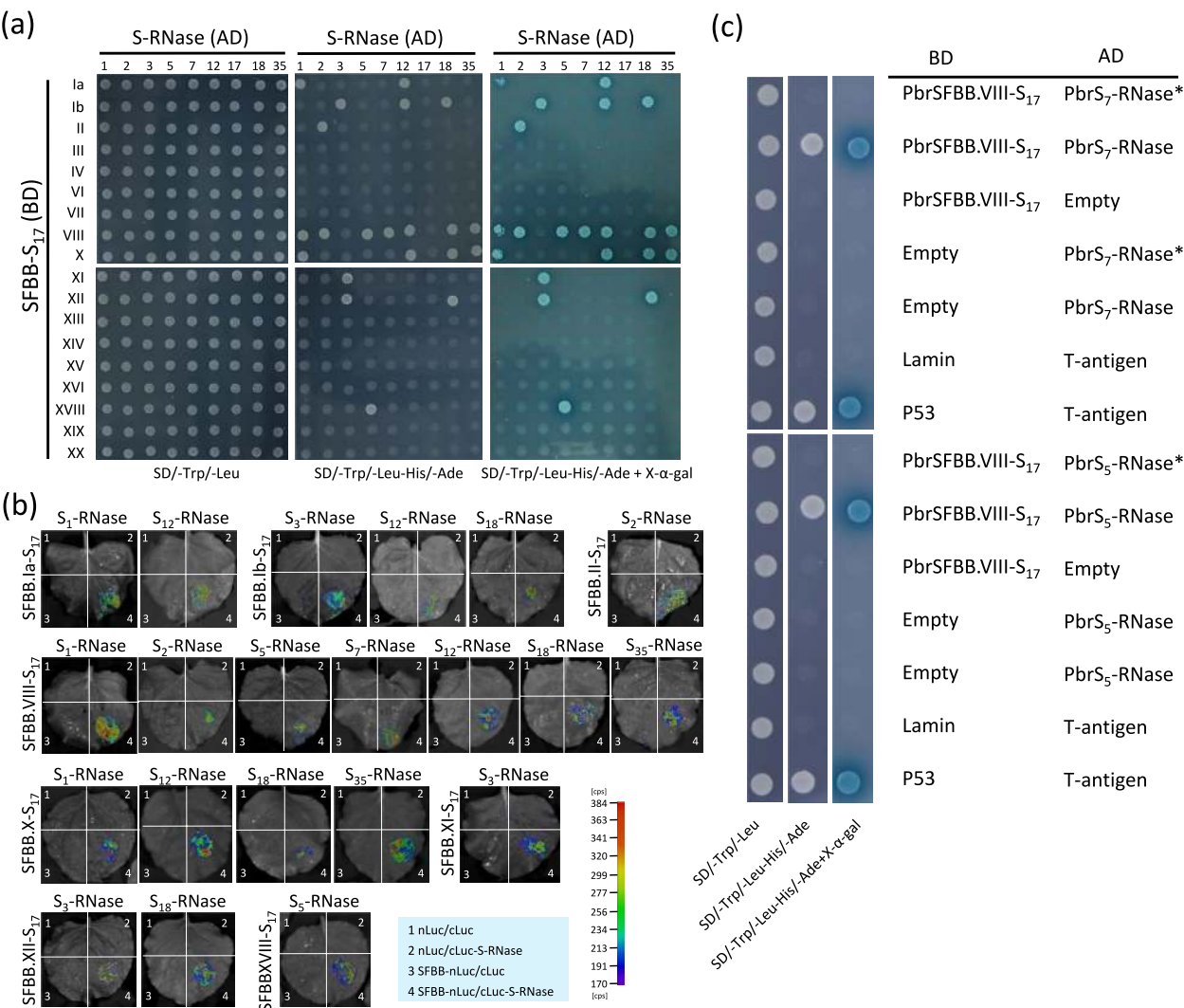


Fig. 5 Interaction of self and non-self S-RNases with the pollen-specifically expressed SFBs from the *Pyrus* S₁₇-locus. **a** Yeast-two-hybrid assay showing the interaction of 18 SFBs with 9 S-RNases. **b** Firefly luciferase complementation assay were performed between the potentially interacting proteins identified by the yeast-two-hybrid assay. The intensity of the interaction is depicted using different colors, which can be evaluated using the color column ruler. **c** Yeast-two-hybrid assay showed that S-RNases lacking the HV region (S₅-RNase* and S₇-RNase*) are unable to interact with PbSFBB.VIII-S₁₇. AD and BD represent the pGADT7 and pGBKT7 vectors, respectively. Arabic numeral indicates the serial number assigned to each S-RNase. Ia, Ib, and II → XX are the serial numbers assigned to SFBs in the *Pyrus* S₁₇-locus (SFBF-S₁₇). SD/-Trp/-Leu denotes the SD medium lacking Trp and Leu; SD/-Trp/-Leu/-His/-Ade denotes the SD medium lacking Trp, Leu, His, and Ade; SD/-Trp/-Leu/-His/-Ade + X-α-gal indicates the addition of x-a-gal to the SD medium lacking Trp, Leu, His, and Ade. P53/T-antigen represents the positive controls, while Lamin/T-antigen represents the negative control

growth and development. Moreover, five genes, *Multi-protein-bridging factor 1a* (evm.TU.G_33.125), *Shaggy-related protein kinase zeta* (evm.TU.G_33.119), *Aspartate aminotransferase* (evm.TU.G_33.115), *Nuclear poly(A) polymerase 1* (evm.TU.G_33.95), and *an uncharacterized protein* (evm.TU.G_33.81), were expressed in both pollen and style (Fig. 7b; Table S14), indicating their potential involvement in pollen tube growth and/or self-incompatibility reaction.

Discussion
Evaluation of the length of S-locus region in Maloideae species
The S-RNase-based GSI system has been extensively investigated in plants belonging to the Solanaceae, Plantaginaceae, and Rosaceae families (Hua et al. 2008). Recent studies utilizing PacBio long-read sequencing have revealed that two *Antirrhinum* S-haplotypes encompass approximately 1.2 Mb and contain 32 *SLF* genes (Zhu

Table 2 Summary of the interaction of the SFBBs expressed in *Pyrus* *S*₁₇-locus with self- and non-self S-RNases

Protein	S ₁ -RNase	S ₂ -RNase	S ₃ -RNase	S ₅ -RNase	S ₇ -RNase	S ₁₂ -RNase	S ₁₇ -RNase	S ₁₈ -RNase	S ₃₅ -RNase
SFBB.Ia-S ₁₇	+	-	-	-	-	+	-	-	-
SFBB.Ib-S ₁₇	-	-	+	-	-	+	+	-	-
SFBB.II-S ₁₇	-	-	-	-	-	-	-	-	-
SFBB.III-S ₁₇	-	-	-	-	-	-	-	-	-
SFBB.IV-S ₁₇	-	-	-	-	-	-	-	-	-
SFBB.VI-S ₁₇	-	-	-	-	-	-	-	-	-
SFBB.VII-S ₁₇	-	-	-	-	-	-	-	-	-
SFBB.VIII-S ₁₇	+	+	-	+	+	+	-	+	+
SFBB.X-S ₁₇	+	-	-	-	-	+	-	+	+
SFBB.XI-S ₁₇	-	-	+	-	-	-	-	-	-
SFBB.XII-S ₁₇	-	-	+	-	-	-	-	+	-
SFBB.XIII-S ₁₇	-	-	-	-	-	-	-	-	-
SFBBXIV-S ₁₇	-	-	-	-	-	-	-	-	-
SFBB.XV-S ₁₇	-	-	-	-	-	-	-	-	-
SFBB.XVI-S ₁₇	-	-	-	-	-	-	-	-	-
SFBB.XVIII-S ₁₇	-	-	-	+	-	-	-	-	-
SFBB.XIX-S ₁₇	-	-	-	-	-	-	-	-	-
SFBB.XX-S ₁₇	-	-	-	-	-	-	-	-	-

et al., 2023). Similarly, using a combination of BAC library and PCR-based approaches, it has been determined that the BAC contigs containing 17 *SLFs* cover roughly 3.1 Mb of the *Petunia* *S*₂-locus (Wu et al. 2020). By comparing the *Petunia* *SLFs* to the *Solanum* genomes, it has been estimated that the *S*-locus region spans approximately 17.9 Mb in tomato (*Solanum lycopersicum*) and 14.6 Mb in potato (*S. tuberosum*; Kubo et al. 2015; Wu et al. 2020). In *Pyrus* species, 10 *SFBBs* were identified from the BAC contigs covering 378 kb of the *S*₂-locus, while six *SFBBs* were identified from the BAC contigs covering 649 kb of the *S*₄-locus (Okada et al. 2011). The number of *SFBBs* in *Pyrus* *S*-locus was found to be lower compared to *Petunia* and *Antirrhinum* *S*-loci, indicating that several *SFBBs* in *Pyrus* *S*-loci remained unidentified. In this study, we used long-read sequencing to assemble the genomes of pear cultivar ‘Yali’, and we successfully identified 19 *SFBBs* from *Pyrus* *S*₁₇-locus located at the end of Chromosome 17 (Fig. 1). Furthermore, leveraging previously assembled genomes of *Pyrus* and *Malus* *S*-loci using long-read sequences (Daccord et al. 2017; Linsmith et al., 2019; Zhang et al. 2019; Dong et al. 2020; Sun et al. 2020; Gao et al. 2021), 17–21 *SFBBs* were identified from these reported *S*-loci (Fig. 2a; Table S4), which were distributed within a range of 1.35–2.64 Mb (Fig. 3c; Table S4). As a comparison, we also identified the *SLFs/SFBs* from *Prunus* *S*-locus (Figure S2), which were distributed within a range of 2.82–3.07 Mb (Table S4). These findings suggest that, on average, the *Pyrus* and *Malus* *S*-loci may have a smaller size compared to the *Prunus* *S*-locus.

The pollen-specifically expressed SFBBs are good candidates for pollen-S determinant in pear

In the GSI mechanism, the pollen-S determinant in *Prunus* species of Rosaceae family is encoded by an *SLF/SFB* (Ushijima et al. 2003, 2004; Sonneveld et al. 2005). These *Prunus* *SLFs/SFBs* exhibit considerable sequence polymorphism among different *S*-haplotypes, are specifically expressed in pollen, and are tightly linked with the corresponding *S*-RNase (Ushijima et al. 2003; Wu et al. 2013). Conversely, the pollen-S determinant in Solanaceae and some Rosaceae species consists of multiple *SLFs* (Kubo et al., 2010; Kakui et al. 2011). These *SLFs* are also specifically expressed in pollen and are tightly linked with the corresponding *S*-RNase; however, the sequence polymorphism is relatively lower compared to the *Prunus* *SLFs/SFBs* (Kakui et al. 2011; Kubo et al. 2015). In this study, a total of 19 *SFBBs* were identified from the *Pyrus* *S*₁₇-locus (Fig. 2a) and showed co-segregation with the *S*₁₇-RNase in cross-pollinated progeny (Fig. 4a). Among these *SFBBs*, *PbrSFBB.Ia-S*₁₇, *PbrSFBB.Ib-S*₁₇, *PbrSFBB.II-S*₁₇, *PbrSFBB.III-S*₁₇, *PbrSFBB.IV-S*₁₇, *PbrSFBB.VI-S*₁₇, *PbrSFBB.VII-S*₁₇, *PbrSFBB.VIII-S*₁₇, *PbrSFBB.X-S*₁₇, *PbrSFBB.XI-S*₁₇, *PbrSFBB.XII-S*₁₇, *PbrSFBB.XIII-S*₁₇, *PbrSFBB.XIV-S*₁₇, *PbrSFBB.XV-S*₁₇, *PbrSFBB.XVI-S*₁₇, *PbrSFBB.XVIII-S*₁₇, *PbrSFBB.XIX-S*₁₇, and *PbrSFBB.XX-S*₁₇ were specifically expressed in pollen (Fig. 4b). The sequence polymorphism of each of these 18 *SFBBs* was also lower than that observed among *Prunus* *SLFs/SFBs* (Tables 1 and S6 to S8). Additionally, *PbrSFBB.Ia-S*₁₇, *PbrSFBB.Ib-S*₁₇,

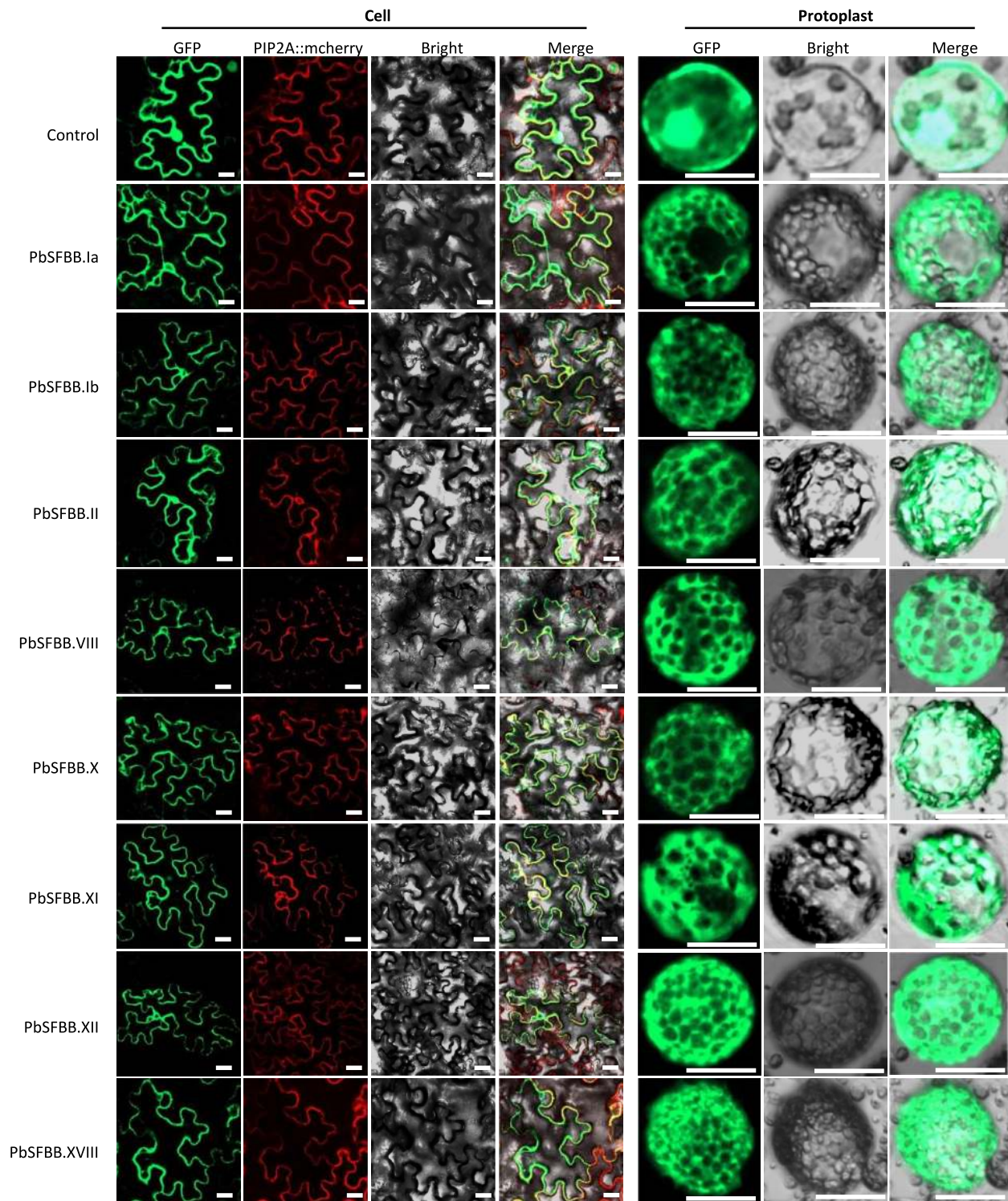


Fig. 6 Subcellular localization of eight SFBBs physically interacting with non-self S-RNases in tobacco mesophyll cell and protoplast. PbSFBB.Ia-S17, PbSFBB.Ib-S17, PbSFBB.II-S17, PbSFBB.VIII-S17, PbSFBB.X-S17, PbSFBB.XI-S17, PbSFBB.XII-S17, and PbSFBB.XVIII-S17 are the SFBBs physically interacting with non-self S-RNases. Control is the empty vector. GFP represents the GFP-fused proteins. PIP2A::mcherry is a control for plasma membrane localization. Bright and Merge show the co-localization of GFP-fused protein and PIP2A::mcherry. Bars = 20 μm

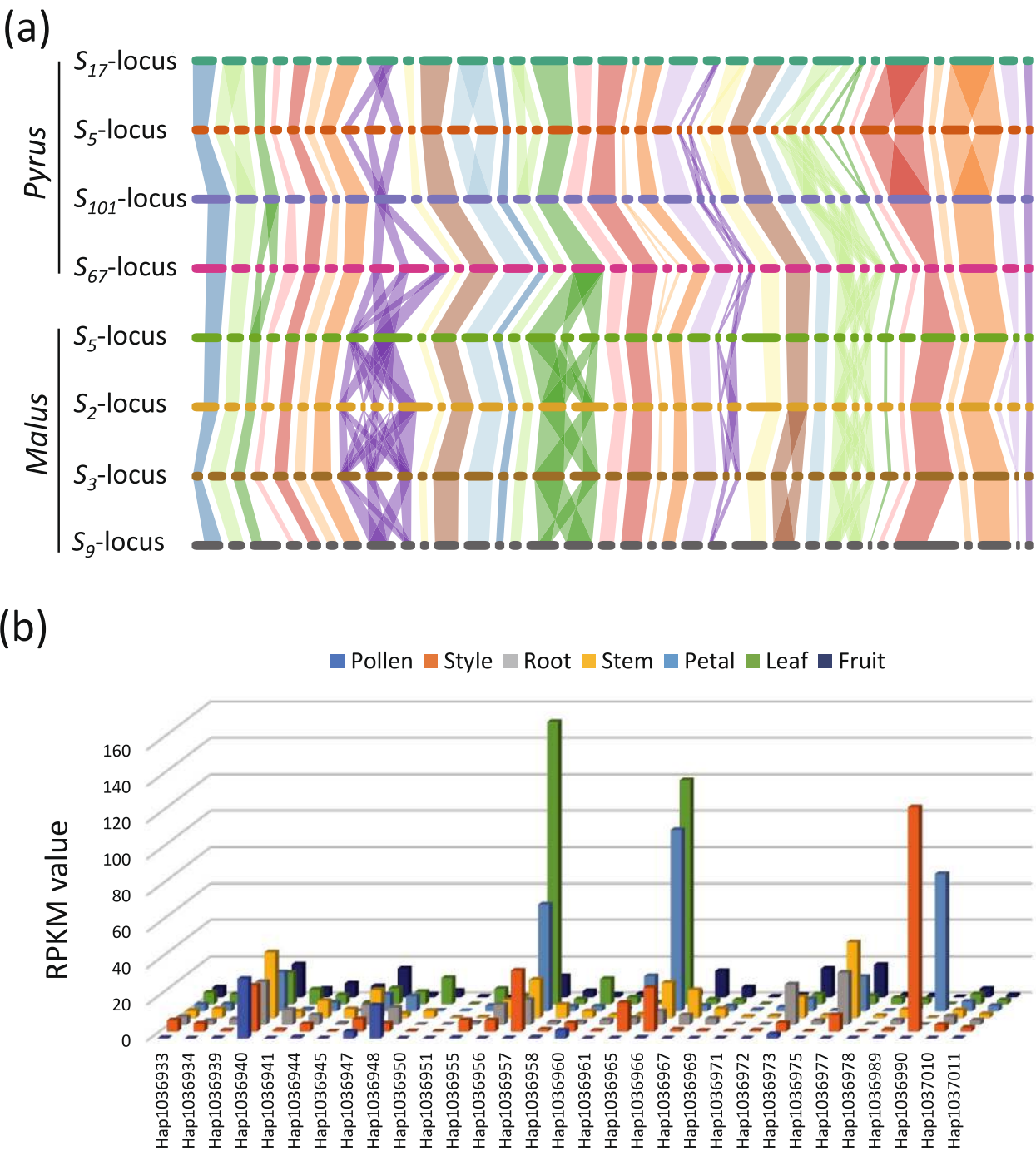


Fig. 7 Identification and expression analysis of 29 genes shared in all tested S-loci. **a** Synteny of 29 genes across the tested S-loci. **b** Transcriptome analysis showing the gene expression level in various tissues, including pollen, style, root, stem, petal, leaf, and fruit. The Y-axis represents the PRKM value, while the X-axis represents the genes shared in all tested S-loci. The order of the 29 genes in the S_{17} -locus is consistent between (a) and (b)

PbrSFBB.II- S_{17} , PbrSFBB.VIII- S_{17} , PbrSFBB.X- S_{17} , PbrSFBB.XI- S_{17} , PbrSFBB.XII- S_{17} , and PbrSFBB.XVIII- S_{17} showed direct interaction with at least one non-self S-RNase, while the absence of interaction between the remaining 10 SFBBs and non-self S-RNases may be

attributed to the limited number of non-self S-RNases used in this study (Fig. 5). Collectively, these findings indicate that these 18 SFBBs that are specifically expressed in pollen represent promising candidates for the pollen-S determinant in pear.

The cross-recognition of SFBBs with non-self S-RNases

In the non-self recognition system, it has been observed that S-RNases cannot be recognized by any SLF present in the same S-haplotype, but they can be recognized by at least one of multiple SLFs present in different S-haplotypes (Sun et al. 2018). A similar mechanism is also proposed in *Pyrus* species (Kakui et al. 2011); however, direct evidence supporting this viewpoint is lacking. In this study, both Y2H and FLCI assays demonstrated that eight out of the 19 SFBBs in the *Pyrus* *S*₁₇-locus are involved in recognizing non-self S-RNases (Fig. 5). This finding provides support for the viewpoint of the non-self recognition system in *Pyrus* species. It is important to note that PbrSFBB.VIII-*S*₁₇ is the only SFBB recognizing *S*₅- and *S*₇-RNases but does not recognize *S*₃-RNase (Fig. 5), despite the fact that there is a high amino sequence identity of 94.2% between *S*₃- and *S*₅-RNases (Table S15). If there are minor changes in the amino acid residues of *S*₅-RNase, the mutant *S*₅-RNase may not be recognized by PbrSFBB.VIII-*S*₁₇, potentially leading to cross-incompatibility. This assumption is consistent with the non-self-recognition system, as the SLFs within an S-haplotype may not recognize approximately 5% of S-RNases in a population (Kubo et al. 2015).

Among the SFBBs that interact with at least one non-self S-RNase, PbrSFBB.VIII-*S*₁₇ has the ability to recognize a majority of non-self S-RNases. On the other hand, PbrSFBB.Ia-*S*₁₇, PbrSFBB.Ib-*S*₁₇, PbrSFBB.II-*S*₁₇, PbrSFBB.X-*S*₁₇, PbrSFBB.XI-*S*₁₇, PbrSFBB.XII-*S*₁₇, and PbrSFBB.XVIII-*S*₁₇ can only recognize a limited number of non-self S-RNases (Fig. 5). It should be noted that SFBB.VIII and SFBB.XVIII are present in both *Pyrus* and *Malus* S-loci, whereas deletion events occasionally occurs in the other SFBBs (Fig. 2a). It has been determined that these gene deletions may result from ectopic recombination between two LTR retrotransposons belonging to the same family and having the same orientation on the same chromosome (Kumar and Bennetzen 1999). Consequently, the presence of enriched LTR retrotransposons in the two flanking regions of these six SFBBs may be responsible for these deletions (Table S13; Shang et al. 2017). The absence of these SFBBs may lead to cross-incompatibility.

Furthermore, we observed deletions in most genes located between SFBB.XVI and SFBB.VIII (Table S5), with the exception of SFBB.III, SFBB.IV, SFBB.XIV and SFBB.XVIII. Additionally, frequent inversions and/or translocations were detected in this genomic region (Figs. 2a and S5 to S11). It has been determined that inversion may arise from unequal recombination events between two LTR retrotransposons with the opposite orientation on the same chromosome, while translocations may result from ectopic recombination between

two LTR retrotransposons on different chromosomes (Kumar and Bennetzen 1999). Therefore, the presence of enriched LTR retrotransposons in *Pyrus* and *Malus* S-loci may be linked to the re-arrangement of SFBBs. This finding aligns with previous research indicating that transposable elements contribute to the reshaping of the *Petunia* S-locus (Wu et al. 2020).

In conclusion, based on the genomes assembled by long-read, we revealed the 17–21 SFBBs in *Pyrus* and *Malus* S-loci spanning a range of 1.35 to 2.64 Mb. These SFBBs could be classified into 22 groups, with 18 of them being specifically expressed in pollen. Notably, eight SFBBs demonstrated interactions with at least one non-self S-RNase, while the remaining SFBBs failed to recognize any S-RNase. These findings provide compelling evidence supporting the existence of a collaborative non-self-recognition system governing self-incompatibility in pear species.

Materials and methods

Plant materials

In accordance with a previous study on S-genotypes (Wang et al. 2017), a total of 10 pear cultivars, ‘Yali’ (*S*₁₇*S*₂₁), ‘Huanghua’ (*S*₁*S*₂), ‘Cuiguan’ (*S*₃*S*₅), ‘Dangshan-suli’ (*S*₇*S*₁₇), ‘Xinxing’ (*S*₄*S*₇), ‘Mantianhong’ (*S*₄*S*₁₂), ‘Jinhua’ (*S*₃*S*₁₈), ‘Hongtaiyang’ (*S*₈*S*₃₅), and ‘Nanyue’ (*S*₂*S*₉), were selected for this investigation. These cultivars were maintained in an orchard located at Hushu, Nanjing, China. During the spring season, styles and anthers from these cultivars, as well as leaves from these cultivars were collected. The collected leaves and styles were promptly frozen in liquid nitrogen and then stored –80°C until further use. The anthers, on the other hand, were enclosed in a dry container to facilitate the release of pollen grains. These pollen grains were then enclosed in sulfuric acid paper bags and stored in silica gel at –20°C until needed.

Following the cross-pollination of ‘Yali’ and ‘Xueqing’ (*S*₃*S*₁₆), a total of 270 seeds were obtained and subjected to vernalization by placing them in moist sand at 4°C. The resulting seedlings were subsequently planted in an orchard located at Jiangpu, Nanjing, China. After eliminating pseudo-hybrid plants, 113 trees reached the flowering stage. Young leaves from these trees were collected during the spring season and stored –80°C until analysis.

DNA sequencing and genome assembly

The genomic DNA of ‘Yali’ was extracted from young leaves using the DNaseq Plant Kit (Tiangen, Beijing, China) and subsequently fragmented. These fragments underwent size selection using AMPure PB

Magnetic Beads (Pacific Biosciences), were repaired for any damages, and ligated to hairpin adapters. The resulting fragments were then used to construct a long-read sequencing library with an average insert size of 20 kb. This library was sequenced using the single-molecule real-time (SMRT) DNA sequencer on the PacBio sequel platform. Moreover, a short-read sequencing library was constructed using the Illumina TruSeq library construction kit (Illumina, San Diego, CA), and a Hi-C sequencing library was constructed using the standard protocol (Belton et al. 2012). Both libraries were sequenced on the Illumina Novaseq platform (Illumina).

To improve the quality of the PacBio reads, a self-correction step was performed by aligning all sequenced reads pairwise. Subsequently, these corrected reads were assembled into contigs using an overlap-layout-consensus algorithm and FALCON (<https://github.com/PacificBiosciences/FALCON>; Chin et al. 2016). The specific parameters used for the assembly were as follows: `length_cutoff_pr=10,000`, `max_diff=95`, and `max_cov=95`. The resulting contigs were further refined through polishing steps using Quiver (Chi et al. 2013), which involved aligning the SMRT reads, as well as using Pilon software (Walker et al. 2014) and aligning Illumina paired-end reads.

In order to evaluate the quality of the assembly, several assessments were conducted. Initially, sequence consistency was examined by aligning paired-end reads to the assembly using BWA software (<http://bio-bwa.sourceforge.net/>). The assembly of the 'Yali' genome had an average sequencing depth of 101.1-fold and a coverage of 99.83%. Additionally, over 97.5% of paired-end reads could be mapped to the assemblies. Furthermore, the completeness of the assembly was assessed using the BUSCO (<http://busco.ezlab.org/>).

Genome annotation

Protein-coding genes were predicted with three different strategies. The first strategy involved *de novo* prediction, which was conducted according to the previous reported method (Shang et al. 2020) without any modifications. For the second strategy, homolog-based prediction was employed using TBLASTN (Altschul et al. 1997), aligning the protein sequences of white pear (*Pyrus x bretschneideri*), apple (*Malus domestica*), sweet cherry (*Prunus avium*), Japanese apricot (*Prunus mume*), peach (*Prunus persica*), and oriental cherry (*Prunus yedoensis*) against the assembly. In addition, transcriptome sequencing was carried out on pollen grains, styles, leaves, and/or developing fruits. The transcriptome-based prediction and function annotation of protein-coding genes were performed following the previous reported method (Shang et al. 2020) without any modification. The coding

sequencing of these protein-coding genes can be found in Files S1 and S2, and their annotations are provided in Files S3 and S4. Furthermore, non-coding RNAs and repeats were predicted using both *de novo* and homolog-based approaches, with the details mirroring those of the previous report (Shang et al. 2020). The annotations of non-coding RNAs and repeats are respectively listed in Tables S3 and S2, respectively.

Identification and sequence analysis of *Pyrus* S-locus F-box genes

The nucleotide acid sequences of *S₁₇-RNase* were utilized as a reference to determine the location of *Pyrus* *S₁₇*-locus in the 'Yali' genome. It was observed that *S₁₇-RNase* resided within the contig 33 in the 'Yali' genome. Next, the F-box motif's amino acid sequences (Figure S1) were subjected to BLAST against Chromosome 17 in the contig 33 in the 'Yali' genome. This search was performed using BioEdit software version 7.0.9.0 (<http://www.mbio.ncsu.edu/bioedit/bioedit.html>) with default parameters. Subsequently, the full-length sequences of *F-box* genes were manually predicted using Geneious Prime version 2021.1.1 (www.geneious.com). A similar approach was followed to predict the *F-box* genes in other *S*-loci of *Pyrus*, *Malus*, and *Prunus* species. The genomes used in this study were provided in Table S1.

To clone the *SFBBs* in the *Pyrus* *S₁₇*-locus, first-strand cDNA from 'Yali' pollen were employed, while the rest were obtained from the National Center for Biotechnology Information database (<https://www.ncbi.nlm.nih.gov/nucleotide/>), as listed in Table S16. Sequence alignment was conducted using ClustalW, and sequence similarity was calculated using BioEdit software. A phylogenetic tree was constructed using MEGA software version 5.04 (www.megasoftware.net) with the neighbor-joining method. Bootstrap values were calculated from 1000 replicates and recorded for values exceeding 0.6. The resulting phylogenetic tree was color-coded using the iTOL tool (<https://itol.embl.de/>). To identify conserved residues of *SFBBs*, WebLogo tool (<http://weblogo.berkeley.edu/logo.cgi>) was utilized. Furthermore, gene duplication events were predicted using the MCScanX package (Wang et al. 2012).

Yeast-two hybrid assay

Total RNAs were extracted from the styles of seven pear cultivars. The first-strand cDNA was synthesized and used as template for PCR amplification of *S-RNase* genes. A subset of nine *S-RNases* were randomly chosen for the Y2H assay. From 'Huanghua', *Pyrus* *S₁-RNase* and *S₂-RNase* were amplified. From 'Cuiguan', *S₃-RNase* and *S₅-RNase* were amplified. From 'Dangshansuli', 'Nanyue', 'Mantianhong', 'Yali', 'Jinhua', and 'Hongtaiyang', *S₇-RNase*,

S₉-RNase, *S₁₂-RNase*, *S₁₇-RNase*, *S₁₈-RNase*, and *S₃₅-RNase* were amplified, respectively. The sequence coding signal peptide were removed from these *S-RNases*, and the remaining coding sequences of these *S-RNases*, including the HV region, were amplified from the styles of pear cultivars. The amplified product was ligated into the pGADT7 vector that was digested by EcoR I and BamH I enzymes (New England Biolabs, Illinois, USA), using the 2X MultiF Seamless Assembly Mix (ABclonal, Wuhan, China) that contains a recombinase. The full-length coding sequences of *SFBBs* in the *Pyrus S₁₇*-locus were amplified from ‘Yali’ pollen. The amplified product was inserted into the pGBKT7 vector that was digested by EcoR I and Sal I enzymes (New England Biolabs), using the 2X MultiF Seamless Assembly Mix (ABclonal). The Y2H was performed in the yeast strain Y2HGold using the Matchmaker Gold Yeast Two-Hybrid System (Clontech, Palo Alto, CA). Moreover, to generate an *S-RNase* protein without the HV region, the coding sequences between the signal peptide and HV, and between HV and stop codon of the *PbS₅-RNase* were successively inserted into the same one pGADT7 vector. The Y2H assay was conducted to assess the interaction between *PbrSFBB.VIII-S₁₇* and *PbrS₅-RNase* without signal peptide and HV region. The primer sequences used for these experiments are provided in Table S17.

Firefly Luciferase complementation imaging (FLCI) assay

The interaction between eight *S-RNases* and eight *SFBBs* were assessed using the FLCI assay. The full-length coding sequences of *Pyrus S₁*, *S₂*, *S₃*, *S₅*, *S₇*, *S₁₂*, *S₁₈*, and *S₃₅-RNase* genes were individually inserted into the pCAMBIA1300-nLuc vector. Similarly, the full-length coding sequences of *PbrSFBB.Ia-S₁₇*, *PbrSFBB.Ib-S₁₇*, *PbrSFBB.II-S₁₇*, *PbrSFBB.VIII-S₁₇*, *PbrSFBB.X-S₁₇*, *PbrSFBB.XI-S₁₇*, *PbrSFBB.XII-S₁₇* and *PbrSFBB.XVIII-S₁₇* genes were individually inserted into the pCAMBIA1300-cLuc vector. These constructs were organized into 21 groups, each representing one of the 21 interaction pairs that were examined in the Y2H assay. The constructs in each group were transiently co-expressed in tobacco leaves after being introduced into *A. tumefaciens* strain GV3101. Luciferase activity was checked at 3 days after infiltration using a Luminescence & Fluorescence Imaging System (PIXIS 1024B/BUV, Teledyne Princeton Instruments, USA). The primer sequences used for this experiment are provided in Table S17.

Subcellular localization

The full-length CDS of *PbSFBB.Ia-S₁₇*, *PbSFBB.Ib-S₁₇*, *PbSFBB.II-S₁₇*, *PbSFBB.VIII-S₁₇*, *PbSFBB.X-S₁₇*, *PbSFBB.XI-S₁₇*, *PbSFBB.XII-S₁₇* and *PbSFBB.XVIII-S₁₇* genes

were individually inserted into a *GFP* vector to construct the CAMV-35S-driven fusion protein (Wu et al. 2023), *PbSFBB.Ia-S₁₇-GFP*, *PbSFBB.Ib-S₁₇-GFP*, *PbSFBB.II-S₁₇-GFP*, *PbSFBB.VIII-S₁₇-GFP*, *PbSFBB.X-S₁₇-GFP*, *PbSFBB.XI-S₁₇-GFP*, *PbSFBB.XII-S₁₇-GFP*, and *PbSFBB.XVIII-S₁₇-GFP*. Moreover, the full-length CDS of aquaporin PIP2A (At3g53420) were amplified from *Arabidopsis* and was inserted into a mCherry vector to construct the PIP2A::mCherry fusion protein as a plasma membrane control (Santiago et al., 2013). The constructs were transferred into *A. tumefaciens* strain GV3101 and were then transiently co-expressed in tobacco epidermal cells. The protoplasts of tobacco epidermal cells were isolated according to the method described in a previous study (Panda et al., 2024). The fluorescence was observed using a confocal microscope LSM780 (Zeiss, Germany). All the primer sequences are listed in Table S17.

Abbreviations

SFBB	S-locus F-box brother
BUSCO	Benchmarking universal single-copy orthologs
SLF	S-locus F-box
SFB	S-locus haplotype F-box
Y2H	Yeast two-hybrid
FLCI	Firefly luciferase complementation imaging
GFP	Green fluorescent protein
QRT-PCR	Quantitative real-time PCR
RPKM	Reads per kilobase per million
SI	Self-incompatibility

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s43897-024-00132-0>.

Supplementary Material 1: Table S1 Comparison of ‘Yali’ genome with previously published assemblies of *Pyrus* and *Malus* species. Table S2 Annotation of the repeats in ‘Yali’ genome. Table S3 Annotation of the non-coding RNAs in ‘Dananguo’ and ‘Yali’ genomes. Table S4 Identification of the *F-box* genes in *Pyrus*, *Malus* and *Prunus* S-loci. Table S5 Function annotation of the predicted genes in S-loci. Table S6 Sequence similarity (%) among *Pyrus* and *Malus SFBB* genes. Table S7 Sequence similarity (%) among *Prunus SFB* and *SLF* genes. Table S8 Sequence similarity among *Prunus SFB* and *SLF* genes. Table S9 Sequence similarity (%) among *Pyrus* and *Malus S-RNase* genes. Table S10 Prediction of gene duplication events of *Pyrus* and *Malus SFBB* genes. Table S11 Sequence similarity of the non-coding flanking sequences of *SFBBs* in *Pyrus* and *Malus* S-loci. Table S12 Analysis of number and length of LTR retrotransposon in different S-loci. Table S13 Identification of the LTR retrotransposon in different S-loci. Table S14 RPKM values of the genes commonly existed in the tested S-loci. Table S15 Sequence similarity (%) among the reported *Pyrus S-RNase* genes. Table S16 The accession numbers of *S-RNase* and S-locus *F-box* genes in *Pyrus*, *Malus*, and *Prunus*. Table S17 Primers used in this study. Figure S1 Isolation of the conserved F-box motif in the reported S-locus F-box proteins in *Pyrus* and *Malus*. The accession numbers of these F-box proteins were listed in Table S13. Figure S2 Phylogenetic classifications of S-locus *F-box* genes in *Prunus*. The S-locus F-box (SLF/SFB) proteins in *Prunus* comprised by 12 groups, SLF1→SLF11 and SFB. Each group were highlighted with different colors. Figure S3 Phylogenetic analysis of the *F-box* genes identified from this and previous studies. Cycles with black color present the *F-box* genes identified from previous study (Huang et al., 2023). The rates (%) of different types of gene duplication events (dispersed, proximal, tandem and transposed) of the S-locus *F-box* genes were calculated for three *Pyrus* (PbrS17, PcS101, and PpyS5) and three *Malus* (MdS2, MdS3,

MdS5, and MdS9) S-loci. The details of these duplication events are listed in Table S10. Figure S4 Gene duplication events of *Pyrus* and *Malus* *SFB*s in whole genome. Figures S5 and S6 Synteny analysis of *SFB*s genes in *Pyrus* and *Malus* S-loci. Arrowhead present the transcriptional direction of a gene. The arrowheads with black color present *S-RNase* gene, while the arrowheads with other colors present S-locus *F-box* genes. The characters above arrowhead are the classification of S-locus *F-box* genes in *Pyrus*, *Malus*, and *Prunus* species. The Arabic numerals below the dotted line are the physical distance (Kb) between two adjacent *F-box* genes. The Chromosome location of these *F-box* genes were listed in Table S4. Figures S7 to S11 Synteny analysis of *SFB*s genes in *Malus* S-loci and *Pyrus* *S*₁₇, *S*₁₀₁, *S*₆₇, *S*₇, and *Pyrus* *S*₅-loci. Arrowhead present the transcriptional direction of a gene. The arrowheads with black color present *S-RNase* gene, while the arrowheads with other colors present S-locus *F-box* genes. The characters above arrowhead are the classification of S-locus *F-box* genes in *Pyrus*, *Malus*, and *Prunus* species. The Arabic numerals below the dotted line are the physical distance (Kb) between two adjacent *F-box* genes. The Chromosome location of these *F-box* genes were listed in Table S4. Figures S12 to S14 Comparison analysis of the 5 kb non-coding flanking sequences of *SFB*s in *Pyrus* *S*₁₇, *S*₅, *S*₆₇, and *S*₁₀₁-loci. Figures S15 to S19 Comparison analysis of the 5 kb non-coding flanking sequences of *SFB*s in *Malus* *S*₂₇, *S*₃₇, *S*₅₇, and *S*₉₇-loci. Figures S20 to S40 Comparison analysis of the 5 kb non-coding flanking sequences of *SFB*s in groups Ia, Ib, II to VIII, X to XIX, XXI, and XXII. (a) A snapshot showing the alignment of the 5kb upstream sequences of *SFB*s. (b) Pairwise identity of the 5kb upstream sequences of *SFB*s. (c) A snapshot showing the alignment of the 5kb downstream sequences of *SFB*s. (d) Pairwise identity of the 5kb downstream sequences of *SFB*s. Figure S41 Identification of S-genotypes of the individuals in the cross-pollinated progeny of 'Yali' × 'Xueqing'. The lanes 1–42 represent the individuals. Q and Y represent the pear cultivars 'Xueqing' and 'Yali', respectively. M indicates the DNA ladder. Figures S42 to S60: The amino acid sequences of *PbrSFBB.Ia-S*₁₇, *PbrSFBB.Ib-S*₁₇, *PbrSFBB.II-S*₁₇, *PbrSFBB.III-S*₁₇, *PbrSFBB.IV-S*₁₇, *PbrSFBB.V-S*₁₇, *PbrSFBB.VII-S*₁₇, *PbrSFBB.VIII-S*₁₇, *PbrSFBB.X-S*₁₇, *PbrSFBB.XI-S*₁₇, *PbrSFBB.XII-S*₁₇, *PbrSFBB.XIII-S*₁₇, *PbrSFBB.XIV-S*₁₇, *PbrSFBB.XV-S*₁₇, *PbrSFBB.XVI-S*₁₇, *PbrSFBB.XVIII-S*₁₇, *PbrSFBB.XIX-S*₁₇, *PbrSFBB.XX-S*₁₇, and *PbrSFBB.XXII-S*₁₇ in any individual including *S*₁₇-*RNase* were identical to that in 'Yali'. YX-1, 2, 4, 5, 9, 10, 14, 19, 20, 23, 24, 28, 29, 30, 32, 34, 35, 36, 37, 38, 41, and 42 are the individuals of the cross-pollinated progeny of 'Yali' × 'Xueqing'. Figure S61 Expression analysis of *S-RNase* and *SFB*s genes in *Pyrus* *S*₁₇-locus in different tissues. Figure S62 Self-activation of *S-RNase* and *SFB*s proteins in *Pyrus* *S*₁₇-locus in yeast cells. AD and BD present the pGADT7 and pGBKT7 vectors respectively. Positive and negative controls were assigned as P53/T-antigen and Lamin/T-antigen, respectively. SD/-Trp/-Leu indicates the SD medium lacking Trp and Leu; SD/-Trp/-Leu/-His/-Ade indicates the SD medium lacking Trp, Leu, His, and Ade; SD/-Trp/-Leu/-His/-Ade + X-α-gal indicates that x-α-gal was mixed into the SD medium lacking Trp, Leu, His, and Ade. Files S1 The coding sequences of all predicted genes in 'Yali' genome. Files S2 Gene annotation of 'Yali' genome.

Acknowledgements

This work was supported by the Bioinformatics Center of Nanjing Agricultural University.

Authors' contributions

S.Z. and C.G. designed the project. C.G. performed genome sequencing and data analysis with assistance from X. W., X.Q., Q. L., and Z. W. YX. and L.W. contributed to most experiments with assistance from K. Q. and M.H.. C.G. drafted the manuscript. S. Z. revised the manuscript.

Funding

The work was supported by the National Key Research and Development Program (2022YFD1200501), the National Science Foundation of China (32172511), the Key Research and Development Program (Modern Agriculture) of Jiangsu Province (BE2022381), the Earmarked Fund for the China Agriculture Research System (CARS-28).

Data availability

Data supporting the findings of this work are available within the paper and its Supplementary Information files. The datasets are available from the corresponding author upon request. Raw genome sequencing reads and the assembly of 'Yali' genome are respectively deposited into the NCBI sequence read archive (SRA) with the BioProject number PRJNA852504. Sequence data of 18 *S*₁₇-locus *F-box* *brother* genes are deposited into the GenBank with the accession numbers ON918613–ON918630.

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interest. Prof. Shaoling Zhang is a member of the Editorial Board for *Molecular Horticulture*. He was not involved in the journal's review of, and decisions related to this manuscript.

Received: 3 April 2024 Accepted: 2 December 2024

Published online: 01 March 2025

References

- Altschul SF, Madden TL, Schaffer AA, Zhangm JH, Zhangm Z, Millermm W, Lipman DJ. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* 1997;25(17):3389–402. <https://doi.org/10.1093/nar/25.17.3389>.
- Belton JM, McCord RP, Gibcus JH, Naumova N, Zhan Y, Dekker J. Hi-C: a comprehensive technique to capture the conformation of genomes. *Methods.* 2012;58(3):268–76. <https://doi.org/10.1016/j.jmeth.2012.05.001>.
- Chen J, Wang P, de Graaf BHJ, Zhang H, Jiao H, Tang C, et al. Phosphatidic acid counteracts S-RNase signaling in pollen by stabilizing the actin cytoskeleton. *Plant Cell.* 2018;30(5):1023–39. <https://doi.org/10.1105/tpc.18.00021>.
- Chi C-S, Alexander DH, Marks P, Klammer AA, Drake J, Heiner C, et al. Non-hybrid, finished microbial genome assemblies from long-read SMRT sequencing data. *Nat Methods.* 2013;10(6):563–9. <https://doi.org/10.1038/nmeth.2474>.
- Chin C-S, Peluso P, Sedlazeck FJ, Nattestad M, Concepcion GT, Clum A, et al. Phased diploid genome assembly with single-molecule real-time sequencing. *Nat Methods.* 2016;13(12):1050–4. <https://doi.org/10.1038/nmeth.4035>.
- Daccord N, Celton J-M, Linsmith G, Becker C, Choise N, Schijlen E, et al. High-quality de novo assembly of the apple genome and methylome dynamics of early fruit development. *Nat Genet.* 2017;49(7):1099–106. <https://doi.org/10.1038/ng.3886>.
- de Nettancourt D. Incompatibility and incongruity in wild and cultivated plants. Berlin: Germany, Springer-Verlag; 2001.
- Dong X, Wang Z, Tian L, Zhang Y, Qi D, Huo H, et al. De novo assembly of a wild pear (*Pyrus betulifolia*) genome. *Plant Biotechnol J.* 2020;18(2):581–95. <https://doi.org/10.1111/pbi.13226>.
- Franklin-Tong VEE. Self-incompatibility in flowering plants: evolution, diversity, and mechanisms. Berlin: Springer-Verlag; 2008.
- Fujii S, Kubo K, Takayama S. Non-self- and self-recognition models in plant self-incompatibility. *Nature Plants.* 2016;2(9):16130. <https://doi.org/10.1038/nplants.2016.130>.
- Gao Y, Yang Q, Yan X, Wu X, Yang F, Li J, et al. High-quality genome assembly of "Cuiguan" pear (*Pyrus pyrifolia*) as a reference genome for identifying regulatory genes and epigenetic modifications responsible for bud dormancy. *Horticulture Res.* 2021;8(1):197. <https://doi.org/10.1038/s41438-021-00632-w>.

- Hiscock SJ, McInnis SM. Pollen recognition and rejection during the sporophytic self-incompatibility response: Brassica and beyond. *Trends Plant Sci.* 2003;8(12):606–13. <https://doi.org/10.1016/j.tplants.2003.10.007>.
- Hua ZH, Fields A, Kao TH. Biochemical models for S-RNase-based self-incompatibility. *Mol Plant.* 2008;1(4):575–85. <https://doi.org/10.1093/mp/ssn032>.
- Hua ZH, Kao TH. 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.* 2006;18(10):2531–53. <https://doi.org/10.1105/tpc.106.041061>.
- Huang X, Mei JW, Zhang J, Liu X, Wang CL. Identification and evolutionary characterization of *SFB* genes in 'Yali' and its spontaneous self-compatible mutant 'Jinzhui' (*Pyrus bretschneideri*). *Horticultural Plant J.* 2023;9(5):909–19. <https://doi.org/10.1016/j.hpj.2022.10.005>.
- Igic B, Lande R, Kohn JR. Loss of self-incompatibility and its evolutionary consequences. *Int J Plant Sci.* 2008;169(1):93–104. <https://doi.org/10.1086/523362>.
- Kakui H, Kato M, Ushijima K, Kitaguchi M, Kato S, Sassa H. 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.* 2011;68(6):1028–38. <https://doi.org/10.1111/j.1365-3113.2011.04752.x>.
- Kao TH, Tsukamoto T. The molecular and genetic bases of S-RNase-based self-incompatibility. *The Plant Cell.* 2004; 16 Suppl(Suppl): S72–S83. <https://doi.org/10.1105/tpc.016154>.
- Kubo K, Entani T, Takara A, Wang N, Fields AM, Hua Z, et al. Collaborative non-self recognition system in S-RNase-based self-Incompatibility. *Science.* 2010; 330(6005): 796–799. <https://www.science.org/doi/10.1126/science.1195243>.
- Kubo K, Paape T, Hatakeyama M, Entani T, Takara A, Kajihara K, et al. Gene duplication and genetic exchange drive the evolution of S-RNase-based self-incompatibility in *Petunia*. *Nature Plants.* 2015;1:14005. <https://doi.org/10.1038/nplants.2014.5>.
- Kumar A, Bennetzen JL. Plant retrotransposons. *Annu Rev Genet.* 1999;33:479–532. <https://doi.org/10.1146/annurev.genet.33.1.479>.
- Li Q, Qiao X, Yin H, Zhou Y, Dong H, Qi K, et al. Unbiased subgenome evolution following a recent whole-genome duplication in pear (*Pyrus bretschneideri* Rehder). *Horticulture Res.* 2019; 6: 34. <https://doi.org/10.1038/s41438-018-0110-6>.
- Li Y, Duan XW, Wu CB, Yu J, Liu CS, Wang J, et al. Ubiquitination of S4-RNase by S-LOCUS F-BOX LIKE2 contributes to self-compatibility of sweet cherry "Lapins". *Plant Physiol.* 2020;184(4):1702–16. <https://doi.org/10.1104/pp.20.01171>.
- Linsmith G, Rombauts S, Montanari S, Deng CH, Celton J-M, Guérif P, et al. Pseudo-chromosome-length genome assembly of a double haploid "Bartlett" pear (*Pyrus communis* L.). *Gigascience.* 2019; 8(12): giz138. <https://doi.org/10.1093/gigascience/giz138>.
- Matton DP, Maes O, Laublin G, Xike Q, Bertrand C, Morse D, Cappadocia M. Hypervariable domains of self-incompatibility RNases mediate allele-specific pollen recognition. *Plant Cell.* 1997;9(10):1757–66. <https://doi.org/10.1105/tpc.9.10.1757>.
- Matton DP, Luu DT, Xike Q, Laublin G, O'Brien M, Maes O, et al. Production of an S RNase with dual specificity suggests a novel hypothesis for the generation of new S alleles. *Plant Cell.* 1999;11(11):2087–97. <https://doi.org/10.1105/tpc.11.11.2087>.
- Minamikawa M, Kakui H, Wang S, Kotoda N, Kichu S, Koba T, Sassa H. Apple S locus region represents a large cluster of related, polymorphic and pollen-specific F-box genes. *Plant Mol Biol.* 2010;74(1–2):143–54. <https://doi.org/10.1007/s11103-010-9662-z>.
- Okada K, Tonaka N, Taguchi T, Ichikawa T, Sawamura Y, Nakanishi T, Takasaki-Yasuda T. Related polymorphic F-box protein genes between haplotypes clustering in the BAC contig sequences around the S-RNase of Japanese pear. *J Exp Bot.* 2011;62(6):1887–902. <https://doi.org/10.1093/jxb/erq381>.
- Panda D, Karmakar S, Dash M, Tripathy SK, Das P, Banerjee S, Qi Y, Samantaray S, Mohapatra PK, Baig MJ, Molla KA. Optimized protoplast isolation and transfection with a breakpoint: accelerating Cas9/sgRNA cleavage efficiency validation in monocot and dicot. *ABIOTECH.* 2024; 5:151–168. <https://doi.org/10.1007/s42994-024-00139-7>.
- Qiao X, Yin H, Li L, Wang R, Wu J, Wu J, Zhang S. Different modes of gene duplication show divergent evolutionary patterns and contribute differently to the expansion of gene families involved in important fruit traits in pear (*Pyrus bretschneideri*). *Front Plant Sci.* 2018;9:161. <https://doi.org/10.3389/fpls.2018.00161>.
- Santiago JP, Tegeder M. Connecting source with sink: the role of Arabidopsis AAP8 in phloem loading of amino acids. *Plant Physiol.* 2013;171(1):508–21. <https://doi.org/10.1104/pp.16.00244>.
- Shang J, Tian J, Cheng H, Yan Q, Li L, Jamal A, et al. The chromosome-level wintersweet (*Chimonanthus praecox*) genome provides insights into floral scent biosynthesis and flowering in winter. *Genome Biol.* 2020;21(1):200. <https://doi.org/10.1186/s13059-020-02088-y>.
- Shang Y, Yang F, Schulman AH, Zhu J, Jia Y, Wang J, et al. Gene Deletion in Barley Mediated by LTR-retrotransposon BARE. *Sci Rep.* 2017;7:43766. <https://doi.org/10.1038/srep43766>.
- Shi D, Tang C, Wang R, Gu C, Wu X, Hu S, et al. Transcriptome and phytohormone analysis reveals a comprehensive phytohormone and pathogen defence response in pear self-/cross-pollination. *Plant Cell Rep.* 2017;36(11):1785–99. <https://doi.org/10.1007/s00299-017-2194-0>.
- Shi DQ, Wu J, Tang HB, Yin H, Wang HT, Wang R, et al. Single-pollen-cell sequencing for gamete-based phased diploid genome assembly in plants. *Genome Res.* 2019; 29(11): 1889–1899. <http://www.genome.org/cgi/doi/10.1101/gr.251033.119>.
- Simão FA, Waterhouse RM, Ioannidis P, Kriventseva EV, Zdobov EM. BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics.* 2015;31:3210–2. <https://doi.org/10.1093/bioinformatics/btv351>.
- Sonneveld T, Tobutt KR, Vaughan SP, Robbins TP. Loss of pollen-S function in two self-compatible selections of *Prunus avium* is associated with deletion/mutation of an S haplotype-specific F-box gene. *Plant Cell.* 2005;17(1):37–51. <https://doi.org/10.1105/tpc.104.026963>.
- Song Z, Zhong S, Qu L. FERONIA and reactive oxygen species: regulators in the self-incompatibility response and in interspecific pollination. *Mol Hortic.* 2023;3:10. <https://doi.org/10.1186/s43897-023-00058-z>.
- Sun LH, Williams JS, Li S, Wu LH, Khatri WA, Stone PG, et al. S-locus F-box proteins are solely responsible for S-RNase-based self-Incompatibility of *Petunia* pollen. *Plant Cell.* 2018;30(12):2959–72. <https://doi.org/10.1105/tpc.18.00615>.
- Sun X, Jiao C, Schwaninger H, Chao CT, Ma Y, Duan N, et al. Phased diploid genome assemblies and pan-genomes provide insights into the genetic history of apple domestication. *Nat Genet.* 2020;52(12):1423–32. <https://doi.org/10.1038/s41588-020-00723-9>.
- Ushijima K, Sassa H, Dandekar AM, Gradziel TM, Tao R, Hirano H. 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.* 2003;15(3):771–81. <https://doi.org/10.1105/tpc.009290>.
- Ushijima K, Yamane H, Watari A, Kakehi E, Ikeda K, Hauck NR, et al. The S haplotype-specific F-box protein gene, *SFB*, is defective in self-compatible haplotypes of *Prunus avium* and *P. mume*. *The Plant Journal.* 2004; 39(4): 573–586. <https://doi.org/10.1111/j.1365-3113.2004.02154.x>.
- Walker BJ, Abeeel T, Shea T, Priest M, Abouelliel A, Sakthikumar S, et al. Pilon: an integrated tool for comprehensive microbial variant detection and genome assembly improvement. *PLoS ONE.* 2014;9(11):e112963. <https://doi.org/10.1371/journal.pone.0112963>.
- Wang GM, Gu C, Qiao X, Zhao B-Y, Ke Y-Q, Guo BB, et al. Characteristic of pollen tube that grew into self style in pear cultivar and parent assignment for cross-pollination. *Sci Hortic.* 2017;216:226–33. <https://doi.org/10.1016/j.scienta.2016.10.035>.
- Wang Y, Tang H, Debarry JD, Tan X, Li J, Wang X, et al. MCSanX: a toolkit for detection and evolutionary analysis of gene synteny and collinearity. *Nucleic Acids Res.* 2012;40(7): e49. <https://doi.org/10.1093/nar/gkr1293>.
- Williams JS, Der JP, dePamphilis CW, Kao TH. Transcriptome analysis reveals the same 17 S-Locus F-box genes in two haplotypes of the self-Incompatibility locus of *Petunia inflata*. *Plant Cell.* 2014;26(7):2873–88. <https://doi.org/10.1105/tpc.114.126920>.
- Wu J, Wang Z, Shi Z, Zhang S, Ming R, Zhu S, et al. The genome of the pear (*Pyrus bretschneideri* Rehder). *Genome Research.* 2013; 23, 396–408. <http://www.genome.org/cgi/doi/10.1101/gr.144311.112>.
- Wu LH, Williams JS, Sun LH, Kao TH. Sequence analysis of the *Petunia inflata* S-locus region containing 17 S-Locus F-Box genes and the S-RNase gene involved in self-incompatibility. *Plant J.* 2020;104(5):1348–68. <https://doi.org/10.1111/tpl.15005>.

- Wu L, Liu X, Zhang M-Y, Qi K-J, Jiang X-T, Yao J-L, et al. Self S-RNase inhibits ABF-LRX signaling to arrest pollen tube growth to achieve self-incompatibility in pear. *Plant J*. 2023;113(3):595–609. <https://doi.org/10.1111/tpj.16072>.
- Wu L, Xu Y, Qi K, Jiang X, He M, Cui Y, Bao J, Gu C, Zhang S. Self S-RNase reduces the expression of two pollen-specific COBRA genes to inhibit pollen tube growth in pear. *Mol Hortic*. 2023;3:26. <https://doi.org/10.1186/s43897-023-00074-z>.
- Zhang L, Hu J, Han X, Li J, Gao Y, Richards CM, et al. A high-quality apple genome assembly reveals the association of a retrotransposon and red fruit colour. *Nat Commun*. 2019;10(1):1494. <https://doi.org/10.1038/s41467-019-09518-x>.
- Zhu S, Zhang Y, Copsy L, Han Q, Zheng D, Coen E, Xue Y. The snapdragon genomes reveal the evolutionary dynamics of the S-Locus supergene. *Molecular Biology and Evolution*. 2023; 40(4): msad080. <https://doi.org/10.1093/molbev/msad080>.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.