



## Research paper

# Rapid evolution of *T2/S-RNase* genes in *Fragaria* linked to multiple transitions from self-incompatibility to self-compatibility

 Wu Chen <sup>a, b</sup>, Hong Wan <sup>c</sup>, Fang Liu <sup>a</sup>, Haiyuan Du <sup>a, b</sup>, Chengjun Zhang <sup>a, \*\*\*</sup>,  
 Weishu Fan <sup>a, \*</sup>, Andan Zhu <sup>a, \*\*</sup>
<sup>a</sup> Germplasm Bank of Wild Species, Kunming Institute of Botany, Chinese Academy of Sciences, Kunming, Yunnan 650201, China

<sup>b</sup> University of Chinese Academy of Sciences, Beijing 100049, China

<sup>c</sup> Horticultural Research Institute, Yunnan Academy of Agricultural Sciences, Kunming, Yunnan 650205, China


## ARTICLE INFO

## Article history:

Received 2 April 2022

Accepted 23 April 2022

Available online 5 May 2022

## Keywords:

*Fragaria*

RNase T2

Self-(in)compatibility

*S-RNase*

## ABSTRACT

The *T2/RNase* gene family is widespread in eukaryotes, and particular members of this family play critical roles in the gametophytic self-incompatibility (GSI) system in plants. Wild diploid strawberry (*Fragaria*) species have diversified their sexual systems via self-incompatible and self-compatible traits, yet how these traits evolved in *Fragaria* remains elusive. By integrating the published and *de novo* assembled genomes and the newly generated RNA-seq data, members of the *RNase T2* gene family were systematically identified in six *Fragaria* species, including three self-incompatible species (*Fragaria nipponica*, *Fragaria nubicola*, and *Fragaria viridis*) and three self-compatible species (*Fragaria nilgerrensis*, *Fragaria vesca*, and *Fragaria iinumae*). In total, 115 *RNase T2* genes were identified in the six *Fragaria* genomes and can be classified into three classes (I–III) according to phylogenetic analysis. The identified *RNase T2* genes could be divided into 22 homologous gene sets according to amino acid sequence similarity and phylogenetic and syntenic relationships. We found that extensive gene loss and pseudogenization coupled with small-scale duplications mainly accounted for variations in the *RNase T2* gene numbers in *Fragaria*. Multiple copies of homologous genes were mainly generated from tandem and segmental duplication events. Furthermore, we newly identified five *S-RNase* genes in three self-incompatible *Fragaria* genomes, including two in *F. nipponica*, two in *F. viridis*, and one in *F. nubicola*, which fit for typical features of a pistil determinant, including highly pistil-specific expression, highly polymorphic proteins and alkaline isoelectric point (pI), while no *S-RNase* genes were found in all three self-compatible *Fragaria* species. Surprisingly, these *T2/S-RNase* genes contain at least one large intron (>10 kb). This study revealed that the rapid evolution of *T2/S-RNase* genes within the *Fragaria* genus could be associated with its sexual mode, and repeated evolution of the self-compatible traits in *Fragaria* was convergent via losses of *S-RNase*.

Copyright © 2022 Kunming Institute of Botany, Chinese Academy of Sciences. Publishing services by Elsevier B.V. on behalf of KeAi Communications Co. Ltd. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

\* Corresponding author. 132# Lanhei Road, Heilongtan, Kunming, Yunnan 650201, China.

\*\* Corresponding author. 132# Lanhei Road, Heilongtan, Kunming, Yunnan 650201, China.

\*\*\* Corresponding author. 132# Lanhei Road, Heilongtan, Kunming, Yunnan 650201, China.

E-mail addresses: [zhangchengjun@mail.kib.ac.cn](mailto:zhangchengjun@mail.kib.ac.cn) (C. Zhang), [fanweishu@mail.kib.ac.cn](mailto:fanweishu@mail.kib.ac.cn) (W. Fan), [zhuandan@mail.kib.ac.cn](mailto:zhuandan@mail.kib.ac.cn) (A. Zhu).

Peer review under responsibility of Editorial Office of Plant Diversity.

## 1. Introduction

The *RNase T2* gene family is a group of the ubiquitous enzyme RNases distributed broadly in animals, plants, protozoans, viruses, and some bacteria that catalyzes the cleavage of phosphodiester bonds in RNA to 3' mononucleotides via 2', 3' cyclic nucleotides (Taylor et al., 1993; Irie, 1999; Deshpande and Shankar, 2002). *RNase T2* genes participate in vital functions such as nutrition intake, phosphate remobilization, self-incompatibility (SI), senescence, and defense against pathogens (Irie, 1999; Deshpande and Shankar, 2002; Luhtala and Parker, 2010).

Members of the *RNase T2* family in plants are classified into *S-RNases* and *S-like RNases* based on whether they are involved in gametophytic self-incompatibility (GSI) responses (Bariola and Green, 1997; MacIntosh, 2011). Phylogenetically, they can be categorized into three distinct classes: the *T2/S-like RNases* belong to classes I and II, while *T2/S-RNases* are in class III (Igić and Kohn, 2001; Steinbachs and Holsinger, 2002; Roalson and McCubbin, 2003; MacIntosh et al., 2010; MacIntosh, 2011; Ramanauskas and Igić, 2017). *S-like RNases* show variable functions in different species. For example, they have been suggested to participate in the defense response in *Nicotiana tabacum*, while in tomato and *Arabidopsis*, they play a role in phosphate remobilization and nucleic acid scavenging (Jost et al., 1991; Köck et al., 1998; Chen et al., 2000; Kurata et al., 2002). *S-RNase* genes are the pistil determination components involved in *S-RNase*-based GSI, which could prevent self-fertilization (McClure et al., 1989; Sassa et al., 1996; Xue et al., 1996). The typical *S-RNase* genes are specifically and highly expressed in the pistil and encode alkaline, polymorphic extracellular ribonuclease. They have been characterized in various families, such as Solanaceae (McClure et al., 1989), Scrophulariaceae (Xue et al., 1996), Rosaceae (Sassa et al., 1993; Ishimizu et al., 1998), Rutaceae (Liang et al., 2020), and Cactaceae (Ramanauskas and Igić, 2021), etc. In Rosaceae, two different systems associated with GSI have been investigated: one is the self-recognition system in *Prunus* (e.g. almond, apricot, and cherry) and the other is the non-self-recognition system (e.g. apple and pear) (Sonneveld et al., 2005; Kubo et al., 2010; Greco et al., 2012; Ramanauskas and Igić, 2021). The *S-RNase* gene structure of the non-self-recognition system has one intron, while that of the self-recognition system has two introns (Sonneveld et al., 2003; Ortega et al., 2005; Mota et al., 2007; Dreesen et al., 2010). Although much attention has been given to the function of the *RNase T2* family in Rosaceae, the genomic basis and evolutionary mechanisms of the *T2/S-RNase* gene family have been less studied. A recent study among seven Rosaceae species showed that the quantity of the *RNase T2* gene in the self-incompatible species *Prunus avium* is twice than that in the self-compatible species *Prunus persica* (Zhu et al., 2020). This disparity implies that the number of the *RNase T2* genes in self-incompatible species is probably higher than that in self-compatible species in a single genus, but more evidence is needed to confirm this hypothesis on a short phylogenomic scale.

*Fragaria*, commonly known as strawberries, is a genus that belongs to the Rosaceae family. The *Fragaria* genus comprises approximately 25 wild species including 12 diploid species ( $2n = 14$ ) (Folta and Davis, 2006; Hummer and Hancock, 2009; Lei et al., 2016). All diploid *Fragaria* species are hermaphroditic and either self-incompatible or self-compatible species (Evans and Jones, 1967; Njuguna et al., 2013; Liston et al., 2014). The *Fragaria* genus engages in the *S-RNase*-based GSI system, and it has been speculated to be controlled by two independent loci (Boskovic et al., 2010; Du et al., 2019). However, in recent studies, two allelic *S-RNases* (*Sa-RNase* and *Sb-RNase*) have been identified, indicating that only one *S* locus exists in *Fragaria viridis* Duch. (Du et al., 2021). To address the evolutionary history of *S-RNases*, other *S-RNase* genes in *Fragaria* need to be identified. The released whole genome sequences of *Fragaria vesca* Lindl. (Edger et al., 2018; Li et al., 2019), *Fragaria iinumae* Makino. (Feng et al., 2021), *Fragaria nilgerrensis* Schlect. (Zhang et al., 2020), and *Fragaria nubicola* Lindl. (Feng et al., 2021) provide an opportunity to conduct the genome-wide identification of *RNase T2* genes and to characterize the *S-RNase* genes involved in the self-incompatible response in a phylogenomic framework.

In this study, six *Fragaria* species including three self-incompatible species (*Fragaria nipponica* Lindl., *F. nubicola*, *F. viridis*) (Feng et al., 2021) and three self-compatible species (*F. iinumae*, *F. nilgerrensis*, *F. vesca*) (Edger et al., 2018; Li et al., 2019; Zhang et al.,

2020; Feng et al., 2021) were selected to examine the evolution of *RNase T2* genes at the whole genome level and to identify *S-RNase* genes involved in SI. By analyzing the phylogenetic relationship, physicochemical features, conserved motifs, duplication modes, and the expression of *RNase T2* genes, we identified the *S-RNase* genes that are likely associated with SI and revealed the mechanisms underlying the rapid evolution of these genes.

## 2. Methods

### 2.1. Genomic data collection

The genome sequence and gene annotation of *Fragaria nilgerrensis* (version: WPAB01000000) (Zhang et al., 2020) was downloaded from the Genome Warehouse in the National Genomics Data Center (NGDC), Beijing Institute of Genomics (BIG), Chinese Academy of Sciences. Genome sequences of *F. vesca* (version: v4.0.a2) (Edger et al., 2018; Li et al., 2019), *F. nubicola* (version: v1.0) (Feng et al., 2021), *F. iinumae* (version: v1.0) (Feng et al., 2021), *Potentilla micrantha* (version: v1.0) (Buti et al., 2018), and *Rosa chinensis* (version: v1.0) (Saint-Oyant et al., 2018) were obtained from the Genome Database for Rosaceae (GDR, <http://www.Rosaceae.org/>). Genome sequences of *F. viridis* (unpublished) and *F. nipponica* (unpublished) were *de novo* assembled.

### 2.2. Identification and re-annotation of *RNase T2* genes

A seed alignment file of the ribonuclease T2 domain (PF00445) was downloaded from the Pfam v35.0 database (<http://pfam.janelia.org/>) (Bateman et al., 2004) and was used to identify annotated *RNase T2* proteins in the six *Fragaria* species and the two outgroups (*R. chinensis* and *P. micrantha*) through HMMER v3.3.1 (e-value <  $e^{-10}$ ) (Finn et al., 2011). Then, the *Fragaria*-specific *RNase T2* HMM file was constructed using hmmbuild from the HMMER v3.3.1 based on the identified *RNase T2* protein sequences, and a second-round identification of *RNase T2* genes in the genomes was performed using HMMER v3.3.1 (e-value <  $e^{-10}$ ) and confirmed by HMMSCAN and the Conserved Domains Database (CDD v3.19) (<https://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi>) (Lu et al., 2020).

The identified *RNase T2* genes were re-annotated using our unpublished RNA sequencing (RNA-seq) data of four *Fragaria* species (*F. nilgerrensis*, *F. vesca*, *F. viridis* and *F. nipponica*). For those *RNase T2* genes without available RNA-seq data, a homology-based method was applied for re-annotation. The pseudogene identification was defined as more than one-third truncation relative to the intact cognate CDS, while the other identified genes were defined as functional genes. The deleterious mutation pattern resulting in pseudogenes was verified by next-generation resequencing data.

### 2.3. Phylogenetic analysis

A phylogeny was inferred from 5570 orthologs, which were identified by Orthofinder v2.5.2 (Emms and Kelly, 2015) with default parameters. The sequences were aligned by MAFFT v7.487 (Katoh, 2002) and Gblocks v0.91b (Castresana, 2000) was used to trim the poorly aligned regions with default parameters. The phylogenetic tree was constructed with the maximum likelihood (ML) method by PhyML v3.1 (Guindon et al., 2010) with 1000 bootstrap values.

The phylogeny of the *RNase T2* gene family was generated as follows: complete coding nucleotide sequences of *RNase T2* gene family members in the six *Fragaria* species were aligned by ClustalW (Thompson et al., 2003) with default parameters. The

phylogenetic tree was constructed with the ML method implemented in MEGA 7.0.26 (Kumar et al., 2016) with 1000 bootstrap values.

#### 2.4. Sequence features and analysis of protein properties

The intron-exon structures were plotted using the Gene Structure Display Server (GSDS 2.0, <http://www.gsds.cbi.pku.edu.cn/>) based on pairwise alignment of the genomic sequences and the CDS sequences (Hu et al., 2015). The conserved motifs from the full-length amino acid sequences of RNase T2 in the six *Fragaria* species were identified by the Multiple Expectation-Maximization for Motif Elicitation (MEME v5.0.5) tool (Bailey et al., 2009) with the following parameters: maximum number of different motifs, 15; minimum motif width, 6; and maximum motif width, 50. The results were plotted with TBtools v1.098696 (Chen et al., 2020). The molecular weight (MW) and isoelectric point (pI) were calculated using the ProtParam tool (<https://web.expasy.org/protparam/>) implemented in ExPASy (Gasteiger et al., 2005).

#### 2.5. Gene location and synteny analysis

All RNase T2 genes of the six *Fragaria* species were mapped to their corresponding chromosomes, and their physical locations were plotted using MapGene2Chrom v2.0 ([http://mg2c.iask.in/mg2c\\_v2.0/](http://mg2c.iask.in/mg2c_v2.0/)) (Jiangtao et al., 2015). The Multiple Collinearity Scan toolkit (MCS-X) (Wang et al., 2012) was adopted to analyze syntenic relationships and the gene duplication events of RNase T2 genes among the six *Fragaria* genomes with the default parameters.

#### 2.6. RNA isolation, sequencing and gene expression analysis

Plants of *Fragaria viridis* and *F. nipponica* were grown in a greenhouse at Kunming Institute of Botany. Total RNA was separately isolated from styles, anthers, and mixed tissues (calyx and petals) at the balloon stage (Hollender et al., 2012) from the two selected *Fragaria* species by using an RNAprep Pure Plant Kit (polysaccharide- and polyphenolic-rich) (Tiangen, Beijing). Three biological replicates were generated for each tissue. The isolated RNA was used for high-throughput RNA-seq library construction, and sequenced (paired-end 150 bp) on an Illumina HiSeq 2500 platform. Clean data were *de novo* assembled using Trinity v2.9.1 (Grabherr et al., 2011) with default parameters, and were mapped to the genome using HISAT2 v2.0.5 (Kim et al., 2015). Gene expression values (TPM: transcripts per million) of the RNase T2 genes were calculated using StringTie v2.1.4 (Pertea et al., 2015).

#### 2.7. Reverse transcription PCR analysis

To verify the tissue-specific expression of the identified *S*-RNase genes, multiple tissues from *F. nipponica* and *F. viridis* were collected, including roots, stems, leaves, calyxes, petals, anthers and pistils. Reverse transcription PCR (RT-PCR) analysis was performed using the PrimeScript One Step RT-PCR Kit (Takara, Dalian, China). The *Actin* gene of the two species was selected as the control. RT-PCR primers are collected in Table S6.

### 3. Results

#### 3.1. Phylogenetic relationship of six diploid species suggests multiple transitions from self-incompatibility to self-compatibility in *Fragaria*

To investigate the origin of the self-compatible and self-incompatible species within *Fragaria*, three self-compatible species

(*F. nilgerrensis*, *F. vesca* and *F. iinumae*) and three self-incompatible species (*F. nipponica*, *F. nubicola* and *F. viridis*) were selected for phylogenetic analysis, with *R. chinensis* and *P. micrantha* used as outgroups. A maximum likelihood (ML) phylogenetic tree was constructed using the concatenated dataset of 5570 single-copy genes (Fig. 1A). This resulted in a non-monophyletic group of self-incompatible species, showing that the self-incompatible species *F. viridis* is a sister lineage to the self-compatible species *F. vesca* with maximal bootstrap support (100%), and the self-compatible species *F. nilgerrensis* is sister to the two self-incompatible species (*F. nipponica* and *F. nubicola*). Thus, the self-compatible traits likely evolve independently in *Fragaria*.

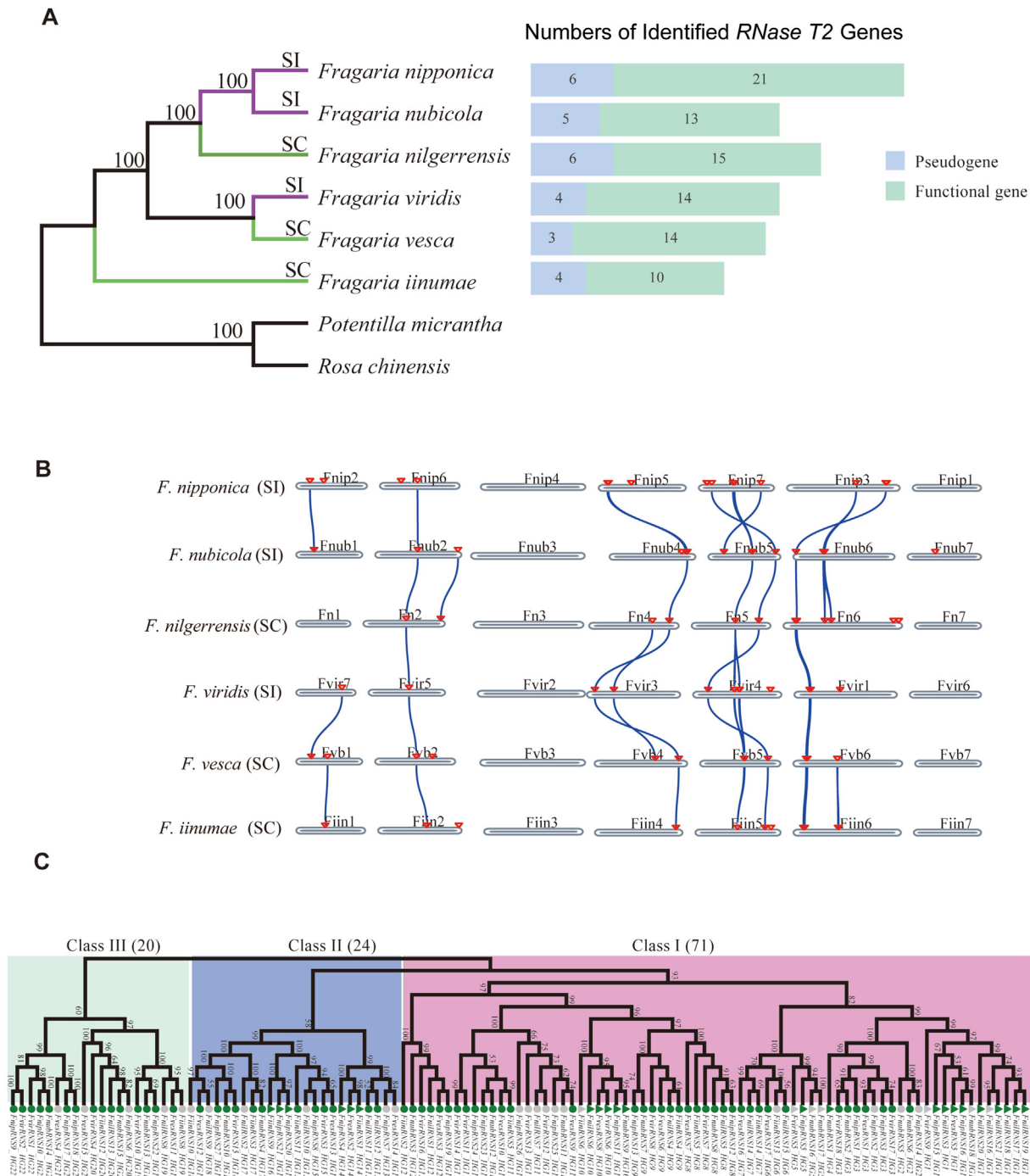
#### 3.2. Identification and classification of RNase T2 genes in *Fragaria* genomes

A total of 115 RNase T2 genes were identified in the six *Fragaria* genomes including 28 homologous pseudogenes (Tables S1 and S2). Pseudogenized RNase T2 genes were mainly caused by frame-shift mutations (67.86%), the gain of stop codons (14.28%), and exon loss (17.86%) (Table S2). We identified 27 RNase T2 genes (21 functional genes and six pseudogenes) in the self-incompatible species *F. nipponica*, ranking as the highest gene number. In contrast, only 14 RNase T2 genes were identified in the self-compatible species *F. iinumae* (10 functional genes and four pseudogenes). The total number of identified RNase T2 genes in *F. nubicola*, *F. nilgerrensis*, *F. viridis* and *F. vesca* was similar, with 18, 21, and 17 RNase T2 genes, respectively (Fig. 1A). The results revealed that the number of RNase T2 genes varied among *Fragaria* species.

A total of 22 homologous gene sets (HG1–22) were identified according to protein sequence similarity and phylogenetic relationships. These homologs were further verified by collinearity relationship (Fig. 1B). A phylogenetic tree inferred from all the 115 identified RNase T2 genes showed that they were separated into three distinct classes (I–III), among which class I comprised the largest gene set (71 genes) distributed in 12 HGs (HG1–12) including 54 functional genes (presenting 76.06%), and 17 pseudogenes (presenting 23.94%); classes II and III contained 24 and 20 RNase T2 genes, respectively. Since T2/*S*-RNase genes are known to belong to Class III (Ramanauskas and Igić, 2017; Du et al., 2021; Ramanauskas and Igić, 2021), we identified 15 functional RNase T2 genes in this class as the candidate *S*-RNase genes.

#### 3.3. Motif composition, functional conserved domains and gene structure of RNase T2 genes in *Fragaria* genomes

The functional motifs of RNase T2 proteins were investigated, and a total of 15 conserved motifs were identified (Fig. S1). Class I possesses 14 motifs including six unique motifs (motif 3, 8, 9, 11, 14, 15) while class II and class III showed highly similar motif compositions with nine and eight motifs, respectively. Motif 10 existed in class II and class III but not in class I. Motif 2 and motif 1 corresponded to the two functional conserved domains (CAS I and II) of the RNase T2 protein, respectively. Multiple sequence alignments of the full-length amino acid sequences of 115 RNase T2 proteins indicated that most of these genes (71.3%) contained two acting histidine residues (Fig. S1B), and 33 (28.7%) members including 24 pseudogenes, lost at least one functionally conserved domain, probably suggesting the loss of gene function. Specifically, 11 RNase T2 proteins including eight pseudogenes lost both CAS I and CAS II (six in class I, four in class II, one in class III). Another two subsets of genes have lost either the CAS I or CAS II domain. Five RNase T2 proteins including four pseudogenes (four in class I, one in class III), lost the CAS I domain, and 17 RNase T2 proteins, including 12



**Fig. 1. Phylogenetic relationship of six *Fragaria* species and the number of *RNase T2* genes identified in each species. (A)** An ML phylogenetic tree (on the left) of the selected *Fragaria* species inferred from 5570 single copy genes. The bootstrap support values were indicated on each node. The SC and SI species are highlighted with green and purple, respectively. Numbers of the identified *RNase T2* genes (on the right) in the six *Fragaria* genomes are shown. **(B)** Synteny analysis of the 115 *RNase T2* genes among the six *Fragaria* genomes. The species were ordered according to the inferred evolutionary relationship shown in Fig. A. The red triangles indicate the location of *RNase T2* genes, and the blue lines link orthologous *RNase T2* gene sets. The chromosome number is shown at the top of each *Fragaria* chromosome and ordered according to the collinearity to *F. vesca*. **(C)** An ML phylogenetic tree of the identified 115 *RNase T2* genes. The bootstrap support values higher than 50 were labeled on each node. Class I, II, and III are denoted by pink, blue and green, respectively. The *RNase T2* gene number in each class is shown in parentheses. The gene names were named according to their species, the order on chromosomes and the order of homologous gene sets. The prefixes “Fiin”, “Fnil”, “Fnip”, “Fnub”, “Fves” and “Fvir” corresponded to *F. iinumae*, *F. nilgerrensis*, *F. nipponica*, *F. nubicola*, *F. vesca* and *F. viridis*, respectively. Genes with RNA data supported are signed with solid circles, while genes without RNA data supported are signed with solid triangles. Functional genes are colored by green, and pseudogenes are colored gray.

pseudogenes, lost the CAS II domain (eleven in class I, four in class II, two in class III) (Fig. S1B). The loss of the CAS domain in *RNase T2* proteins may cause partial or complete loss of their biological function.

Meanwhile, the intron-exon structure of *RNase T2* genes showed substantial variations (Fig. S2). We found that the intact genes clustered into class I (HG1–12) harboring two to eight exons, among which HG1–4 and HG6–7 possess two exons, HG8 and HG10

possess three exons, HG5, HG9 and HG11 possess four exons and HG12 possesses eight exons. For those genes that were clustered into class II (HG13–18) and class III (HG19–22), two to three exons were commonly identified. Specifically, HG13–20 possesses two exons, and HG21–HG22 possesses three exons (Fig. S2A and S2B). Exon loss may also affect the gene structures. For example, only one exon was found in *FvirRNS9-HG1*, a homologous pseudogene that belongs to HG1, among which intact genes had two exons (Fig. S2A).

### 3.4. Sequence features of RNase T2 genes in the *Fragaria* genomes

Genic characteristics, including the length of the protein sequence, molecular weight (MW) and isoelectric point (pI), of RNase T2 genes were analyzed (Table S1 and Fig. S3). The length of the 115 RNase T2 protein sequences ranged from 47 to 278 amino acids (Table S1). Previous studies showed that the majority of MWs of a typical RNase T2 enzyme are in the range of 20–40 kDa (Deshpande and Shankar, 2002). In this study, among the 115 identified RNase T2 proteins, 33 proteins were out of this range and displayed lower MWs including all 28 pseudogenes (Table S1). Moreover, the sequence lengths of the 33 low MW RNase T2 proteins were shorter than those of the rest of the RNase T2 proteins, which is consistent with their lower MWs. Outrageous values occurred in both self-incompatible and self-compatible species (Table S1), indicating irrelevant compatibility. The pI values ranged from 4.23 to 9.86 (Table S1 and Fig. S3). Class I and II proteins generally have a slightly higher proportion of alkaline pI values than acidic pI values; 38 (53.52%) and 13 (54.17%) peptides exhibit alkaline pI values, while 33 (46.48%) and 11 (45.83%) peptides exhibit acidic pI values, respectively. However, in class III peptides, there were significantly more alkaline pI values; 16 (80%) peptides had alkaline pI values, while 4 (20%) peptides had acidic pI values. We found that only two homologous pseudogenes (*FvirRNS9-HG1* and *FnipRNS6-HG20*) showed independent pI shifts from alkaline to acidic values, revealing that deleterious mutations have little effect on pI values. Our results also suggested that pI values were not associated with the loss of the active histidine residue (pseudogene).

### 3.5. Evolutionary mechanisms of copy number variations of RNase T2 genes in *Fragaria* genomes

We systematically examined the physical positions of the identified RNase T2 genes and found that they were unevenly distributed on the seven *Fragaria* chromosomes (Figs. 1B and S4). Taking *F. nipponica* as an example, for its total number of 27 RNase T2 genes, 22 genes (81.48%) were located on chr3, chr5, and chr7, and no RNase T2 genes were on chr1 and chr4 (Fig. S4). This uneven distribution of RNase T2 genes is consistent in all five other examined *Fragaria* genomes. Specifically, the colinear chromosomes to chr3, chr5 and chr7 of *F. nipponica* in other *Fragaria* genomes also possessed more than 70% RNase T2 gene members of their species (72.22% in *F. nubicola*, 90.48% in *F. nilgerrensis*, 83.33% in *F. viridis*, 76.47% in *F. vesca*, 78.57% in *F. iinumae*). The chromosomes colinear to chr4 and chr1 in *F. nipponica*, which are chr2, chr3 and chr6 or chr7 in other genomes, contained no RNase T2 genes in all the six genomes, except for one RNase T2 gene exists on chr7 in *F. nubicola* genome (Figs. 1B and S4).

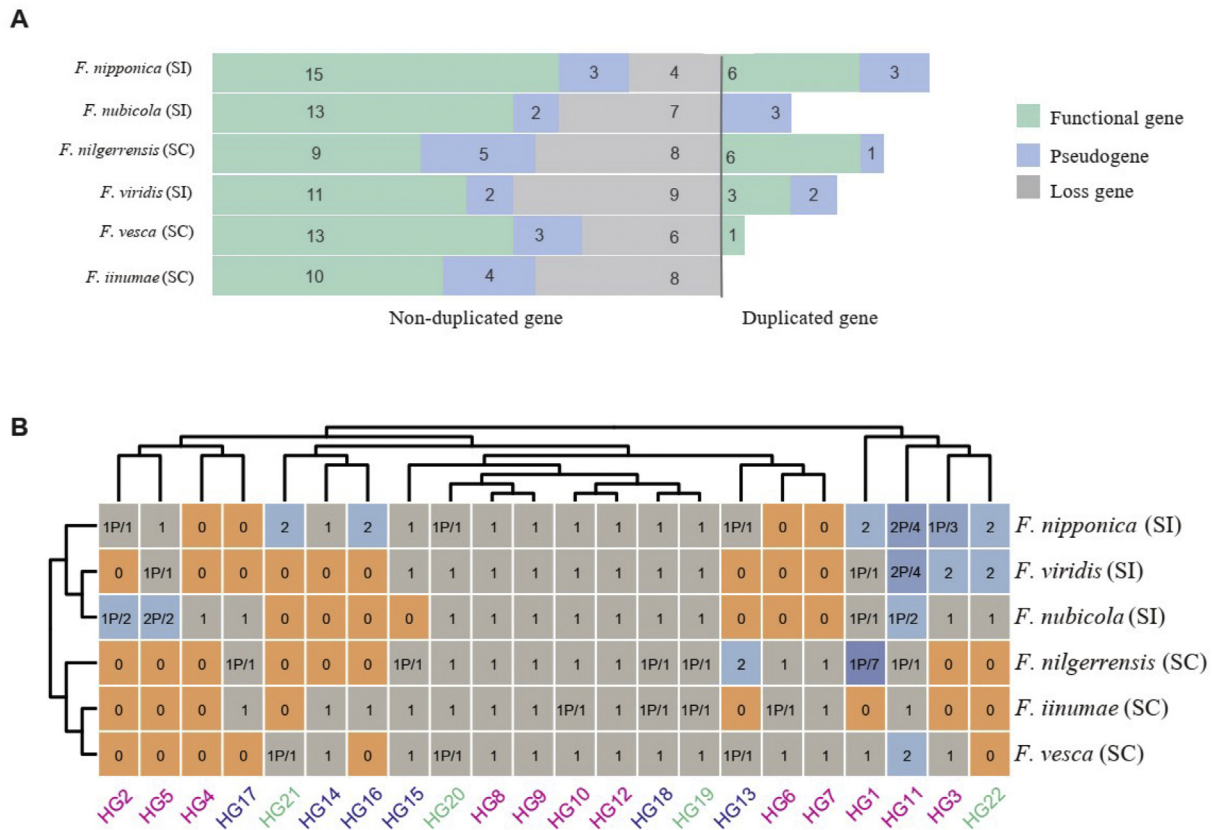
Interestingly, the total number of RNase T2 genes in *Fragaria nipponica* is approximately two-fold of that in *F. iinumae*. We thus traced the gain-and-loss mechanism for the gene number variation. All species have experienced extensive gene loss, and/or gene number increased via duplication events along with evolution (Fig. 2A). Pseudogenes could also be derived from both non-

uplicated genes and duplicated genes (Fig. 2A). *F. nipponica* showed maximum duplicated genes and minimum loss genes, making its genome contains the highest numbers of RNase T2 genes. In contrast, no duplication event occurred, and a relative high number of gene losses led *F. iinumae* to retain the fewest RNase T2 genes.

We tested whether the biased gain and loss of gene members belonging to any homologous gene set or class (class I to III) occurred (Fig. 2B). The results showed that HG2, 4, 5, 6, 7 (belonging to class I), HG13, 14, 16, 17 (class II) and HG21, 22 (class III) were lost in at least three *Fragaria* genomes, which indicated that these HGs tend to be lost in *Fragaria*, and the gene loss could occur in all three classes. For duplication of these gene family members, HG1, HG3, and HG11 (class I) were detected as duplicated genes in at least two *Fragaria* genomes, among which HG11 was duplicated in four of six *Fragaria* genomes, but 41.67% of duplicated genes of HG11 were pseudogenes. For HG22, although two copies were identified in the three SI *Fragaria* genomes, they were actually alleles rather than duplicated genes (see next Result section). Together, the interplay of gene loss, pseudogenization and duplication mainly accounted for the evolution of the RNase T2 gene family size in the *Fragaria* genomes. To further classify the RNase T2 gene copy disparity among *Fragaria* species, duplication modes of the multiple-copy genes were investigated. Two types of duplication events were observed, including tandem duplication (TD) and segmental duplication (SD). A chromosomal region within 200 kb containing two or more genes was defined as a tandem duplication event (Holub, 2001). In our study, multiple copies of HG1, HG2 and HG3 were duplicated by TD, and their distances on the chromosome ranged from 7.58 kb (between *FvirRNS17-HG3* and *FvirRNS18-HG3*) to 170.80 kb (between *FnubRNS5-HG2* and *FnubRNS6-HG2*) (Table S3). The two copies of HG5 (class I) in the *F. nubicola* genome and the two copies of HG16 (class II) in the *F. nipponica* genome were duplicated by SD. Additionally, both TD (7 genes, presenting 58.33%) and SD (5 genes, presenting 41.67%) events contributed to copy number variation of HG11 (class I) (Table S3). These results imply that TD and SD events play the major driving force on multiple copies of RNase T2 gene members in *Fragaria* genomes.

### 3.6. Identification of the candidate T2/S-RNase genes in the *Fragaria* genomes

Tissue-specific expression is one of the key features of S-RNase genes in SI. To identify the candidate S-RNase genes according to their expression, RNA-seq data of four species (*F. nilgerrensis*, *F. vesca*, *F. viridis* and *F. nipponica*) from three tissues were obtained, including anther, pistil and mixed tissue (calyx and petal). The results showed that there are four genes belonging to HG22 (*FvirRNS1-HG22*, *FvirRNS2-HG22*, *FnipRNS9-HG22* and *FnipRNS10-HG22*) were highly pistil-specific expressed (Fig. 3A and Table S4) and were regarded as candidate S-RNase genes. To further confirm the tissue-specific expression of these four candidate S-RNase genes, RT-PCR was applied in multiple tissues including root, stem, leaf, calyx, petal, anther and pistil. The results showed that the four candidate genes were all specifically expressed in the pistil, which is consistent with the expression features of style determinants (Fig. 3B), providing strong evidence that they are S-RNase genes. An additional candidate gene, *FnubRNS14-HG22*, was defined based on its homology to the *FnipRNS9-HG22* sequence due to unavailable RNA-seq data for *F. nubicola*. The amino acid sequence similarity of the candidate S-RNase genes ranged from 36.07% to 87.01% (Table S5). Furthermore, their predicted molecular masses are ranged from 26.19 kDa to 29.04 kDa and alkaline isoelectric points (pIs) ranged from 8.49 to 9.29 (Table 1). These features, including phylogenetic relationships (belonging to class III),



**Fig. 2. Copy number variations of *RNase T2* homologs.** (A) Numbers of identified non-duplicated and duplicated *RNase T2* genes in the six *Fragaria* genomes are shown. (B) Copy numbers of homologous *RNase T2* gene sets. Numbers in the rectangles indicated gene copy. HG1–HG12 belongs to class I colored by magenta, HG13–HG18 belongs to class II colored by dark blue and HG19–HG22 belongs to class III colored by green. The phylogeny of species and homologous gene sets (HGs) was inferred from cluster analysis.

pistil-specific expression, highly polymorphic proteins, a range of molecular masses and alkaline isoelectric points (pIs), are all highly consistent with the known typical *S-RNase* genes (Igic and Kohn, 2001; Ramanaukas and Igic, 2017). For three self-compatible species, no candidate *S-RNase* genes were found.

Since a recent study identified two allelic *S-RNases* (*Sa-RNase* and *Sb-RNase*) in *F. viridis* (Du et al., 2021), we aimed to verify whether these two *S-RNase* genes are the homologous genes to what we identified in this study. *Sa-RNase* and *Sb-RNase* sequences were from an assembled genome of another *F. viridis* accession, and they were both located on chr3 at approximately 468 kb in distance. According to the chromosome collinearity, *Sa-RNase* and *Sb-RNase* are orthologous genes with the five candidate *S-RNase* genes in this study (Fig. 3C). The amino acid sequence similarity between two allelic *S-RNases* (*Sa-RNase* and *Sb-RNase*) and these five *S-RNase* genes ranged from 34.72% to 92.17% (Table S5). All these results provide solid evidence to indicate that these five candidate genes are *S-RNases* in three self-incompatible *Fragaria* species.

### 3.7. Primary structural features of the candidate T2/S-RNase genes in the *Fragaria* genomes

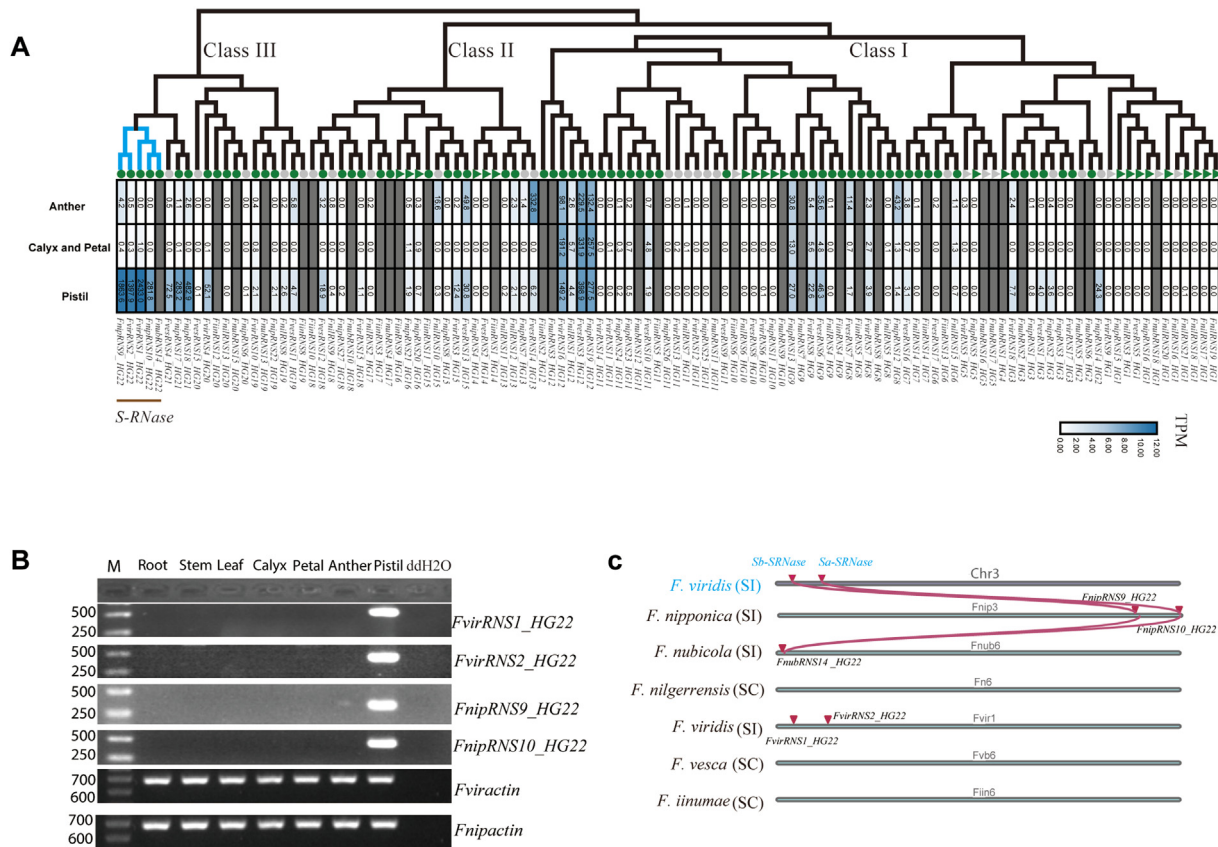
The *S-RNase* genes in *Fragaria* present exceptional structural variation. Although they all harbored three exons and two introns, the length of the first intron can range from 222 bp to 20.3 kb, and the second intron length ranged from 158 bp to 15.7 kb, which means that both introns showed approximately 100 times variation in length (Fig. 4A). Amino acid sequences of *S-RNase* genes in *Fragaria* contain four conserved regions (C1–C3, RC4, and C5) and one hypervariable region (RHV). The first intron is located between

signal peptide and C1 while the second intron is within the hypervariable region (Fig. 4B). Structural diversification illustrated that the self-incompatible species evolved fast to accommodate its compatibility.

## 4. Discussion

One diversifying feature of diploid *Fragaria* species is that they have evolved both self-compatible and self-incompatible traits (Evans and Jones, 1967; Njuguna et al., 2013; Liston et al., 2014), which have been suggested to be genetically controlled by the *S-RNase* genes in the *RNase T2* gene family (Boskovic et al., 2010; Du et al., 2019, 2021). The phylogenetic analysis in our results also supported that there were multiple transitions from SI to self-compatibility (SC) in *Fragaria* (Fig. 1). In this study, by using the whole genome sequencing together with the RNA sequencing methods, genome-wide analysis of the *RNase T2* gene family in both self-compatible and self-incompatible diploid *Fragaria* species was performed, and the evolutionary mechanisms of the rapid evolution of T2/S-RNase genes were examined.

Previous analyses of the *RNase T2* gene family in other species showed that they were clustered into three classes (Igic and Kohn, 2001; Steinbachs and Holsinger, 2002; Roalson and McCubbin, 2003; Boskovic et al., 2010; MacIntosh et al., 2010; Ramanaukas and Igic, 2017; Du et al., 2019, 2021), but a recent study documented a class IV *RNase T2* gene in seven Rosaceae species (Zhu et al., 2020), among which common introns were absent. In this study, the 115 identified *RNase T2* genes were also grouped into three classes (I–III). Notably, the phylogenetic relationship is largely affected by the correctness of gene annotation. For example,



**Fig. 3. Transcript expression, RT-PCR experiments and synteny relationship of the S-RNase genes.** (A) Expression profiles of the RNase T2 genes in three tissues. The color scale represents relative gene expression levels, and the TPM values are shown inside the rectangle. Gray rectangles represent transcriptome data unavailable for those RNase T2 genes. The five S-RNase genes (*FnipRNS9\_HG22*, *FnipRNS10\_HG22*, *FvirRNS1\_HG22*, *FvirRNS2\_HG22* and *FnubRNS14\_HG22*) are highlighted with cyan on their branch. (B) RT-PCR experiments verify the gene expression of four S-RNases in multiple tissues including root, stem, leaf, calyx, petal, anther and pistil in *Fragaria nipponica* and *F. viridis*. M: Marker (bp). (C) Synteny relationship between two published allelic S-RNases (*Sa-RNase* and *Sb-RNase*) and the five newly identified S-RNase genes. Gene locations are highlighted with purple triangles.

**Table 1**  
Genic characteristics of five S-RNase genes identified in the three self-incompatible *Fragaria* genomes.

Name	CDS (bp)	Amino acid	MW (kDa)	pI
<i>FvirRNS1_HG22</i>	762	253	29.04	9.10
<i>FvirRNS2_HG22</i>	753	250	28.50	8.90
<i>FnipRNS9_HG22</i>	744	247	28.73	9.29
<i>FnipRNS10_HG22</i>	696	231	26.78	8.64
<i>FnubRNS14_HG22</i>	696	231	26.49	8.49

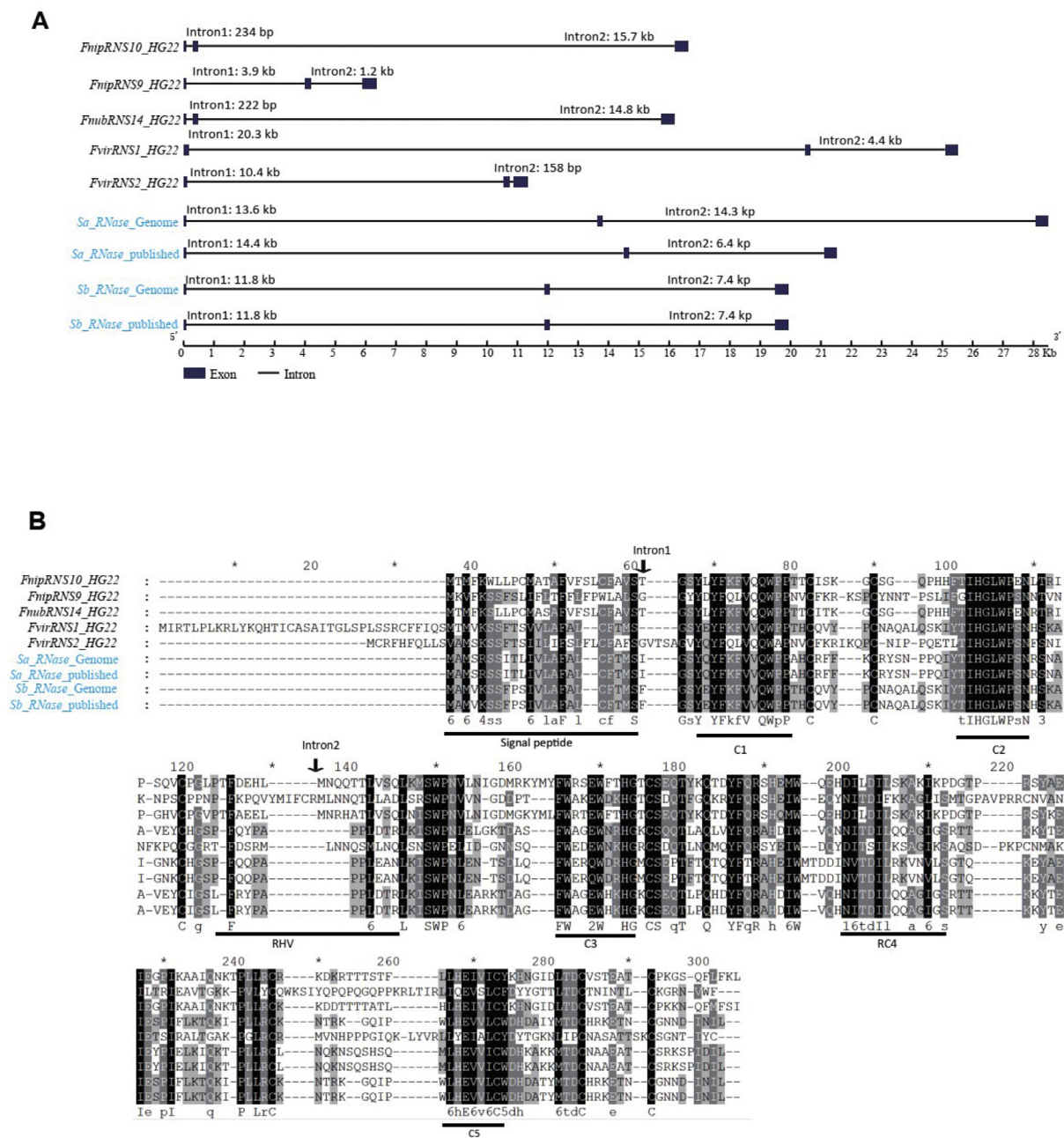
*FvesRNS9\_HG11* (*FvH4\_5g24550*) and *FvesRNS10\_HG11* (*FvH4\_5g24800*) in *F. vesca* involved in class IV were nested into class I in our study, indicating that misannotation might lead to classification confusion. Our phylogeny and classification were as robust as in *citrus* (Honscho et al., 2021) and cacti (Ramanauskas and Igić, 2021).

The identification of RNase T2 gene family members has been performed in diverse plants, and the gene numbers of this family vary among the examined plants. For example, only five RNase T2 genes were found in *Arabidopsis* (Igić and Kohn, 2001), while there were 19 members in *Rubus occidentalis* (Zhu et al., 2020). The results need to be closely investigated because the identification was based on the published annotation file, possibly ignoring the unannotated genes in the genome. For example, there were one to five members that were not found in the annotation file in each examined species in this study (Table S1). Consistent with previous

research (Zhu et al., 2020), the quantity in *Fragaria nipponica* (SI, 27 RNase T2 genes) is approximately twice than that in *F. iinumae* (SC, 14 RNase T2 genes), but this pattern cannot be applied to all self-incompatible and self-compatible species. This result suggests that total number variations may not be directly related to the transition from SI to SC, but fast-evolving mechanisms existed along with the diversification of *Fragaria* species.

Gene loss, pseudogenization, and duplication are prevalent in the *Fragaria* RNase T2 gene family, while duplications, especially TD and SD, are the main evolutionary forces. However, in some other Rosaceae species, RNase T2 duplications are dominated by dispersed duplication (DSD) or whole-genome duplication (WGD) (Qiao et al., 2018), but another case was also reported in pear with TD or PD as the main duplication events (Zhu et al., 2020). Here, we defined only the genes with multiple copies and speculated on the origins of different copies, while other studies evaluated the origin of RNase T2 genes from the entire gene family (Qiao et al., 2018; Zhu et al., 2020). Moreover, the diploid *Fragaria* species did not experience a recent WGD (Wu et al., 2013; Jiang et al., 2020; Qiao et al., 2021). Therefore, the TD could be a reason for copy variation. Fast evolved gene families can be clearly observed in large gene families but are not very common in small gene family (Meyers et al., 2003; Feng et al., 2017). Therefore, our results provide an example to illustrate the fast evolution in a relatively small gene family.

The *Fragaria* genus has been proven to engage in the S-RNase-based GSI system (Boskovic et al., 2010; Du et al., 2019, 2021), in



**Fig. 4. Characterization of primary structural features of *S*-RNase genes.** (A) The structure of the *S*-RNase genes. The genes were drawn to the scale as the scale bar is shown at the bottom, and the lengths of introns are labeled. The published references were highlighted with cyan. *Sa/Sb\_RNase\_Genome*: gene structures were drawn according to our unpublished *Fragaria viridis* genome; *Sa/Sb\_RNase\_published*: gene structures were drawn according to the gene sequences published by Du et al., (2021). (B) The five conserved regions (C1–C3, RC4, and C5) and one hypervariable region (RHV) of *S*-RNase genes.

which the pistil determination genes are used to prevent self-fertilization (McClure et al., 1989; Sassa et al., 1996; Xue et al., 1996). The distinguishing features of *S*-RNase genes include specifically high expression in pistils, possessing alkaline pI, and they were classified into class III based on the phylogenetic relationship of the entire RNase T2 gene family. The *S*-RNase genes in potato (Ye et al., 2018), citrus (Liang et al., 2020) and cacti (Ramanauskas and Igić, 2021) were identified by generating style RNA-seq data based on their highly style-specific expression features. In our study, RNA-seq data of styles in four *Fragaria* species (two self-incompatible and two self-compatible species) were obtained, and ultimately five candidate *S*-RNase genes (*FvirRNS1\_HG22*, *FvirRNS2\_HG22*, *FnipRNS9\_HG22*, *FnipRNS10\_HG22* and *FnubRNS14\_HG22*) were

found in three self-incompatible species (Fig. 3). These five candidate genes belonging to class III are highly pistil-specific (Fig. 3A and B), and their amino acid similarity (36.07%–87.01%) (Table S4), molecular masses (from 26.19 kDa to 29.04 kDa) and alkaline pI (8.49–9.29, Table S1) all fulfill the typical characteristics of the *S*-RNase gene. Moreover, we confirmed that the two reported allelic *S*-RNases (*Sa*-RNase and *Sb*-RNase) in *F. viridis* (Du et al., 2021) are the orthologous genes with the five candidate *S*-RNase genes in this study (Fig. 3C), and high amino acid sequence similarity (92.17%) was observed between *Sa*-RNase and *FvirRNS1\_HG22* (Table S4). All these results provide solid evidence to indicate that these five candidate genes are the *S*-RNases in three self-incompatible *Fragaria* species, and they may be completely lost after the transition



from SI to SC. Surprisingly, the *S-RNase* genes in *Fragaria* harbor at least one intron larger than 10 kb (10.4–20.3 kb), except for *FnipRNS9\_HG22*. The intron length also revealed the rapid evolution during the transition from SI to SC. The reason for the exceptionally long intron formation and whether it plays a role in the SI response needs to be further proven.

### Authors contributions

A.Z., W.F. and C.Z. designed and conceived this study. W.C., H.W. and A.Z. collected and maintained the plant materials. W.C., H.W., F.L. and H.D. performed the experiments and data analyses. W.C. and W.F. wrote the manuscript with input from A.Z. All authors approved this manuscript.

### Declaration of competing interest

The authors declare that they have no conflict of interest.

### Acknowledgments

The authors greatly appreciate Dr. Han Guo for enjoyable discussions and providing constructive suggestions to improve this study. We also thank Dr. Jiwei Ruan for his assistance in plant materials collection. This research was financially supported by the National Key Research and Development Program of China (2018YFD1000107) and the open research project of the “Cross-Cooperative Team” of the Germplasm Bank of Wild Species to A.Z.

### Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.pld.2022.04.003>.

### References

- Bailey, T.L., Boden, M., Buske, F.A., et al., 2009. MEME SUITE: tools for motif discovery and searching. *Nucleic Acids Res.* 37, 202–208.
- Bariola, P.A., Green, P.J., 1997. Plant ribonucleases. In: *Ribonucleases: Structure and Functions*. Elsevier, pp. 163–190.
- Bateman, A., Coin, L., Durbin, R., et al., 2004. The Pfam protein families database. *Nucleic Acids Res.* 32, 138–141.
- Boskovic, R.L., Sargent, D.J., Tobutt, K.R., 2010. Genetic evidence that two independent *S*-loci control RNase-based self-incompatibility in diploid strawberry. *J. Exp. Bot.* 61, 755–763.
- Buti, M., Moretto, M., Barghini, E., et al., 2018. The genome sequence and transcriptome of *Potentilla micrantha* and their comparison to *Fragaria vesca* (the woodland strawberry). *GigaScience* 7, 1–14.
- Castresana, J., 2000. Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. *Mol. Biol. Evol.* 17, 540–552.
- Chen, C., Chen, H., Zhang, Y., et al., 2020. TBtools: an integrative toolkit developed for interactive analyses of big biological data. *Mol. Plant* 13, 1194–1202.
- Chen, D.L., Delatorre, C.A., Bakker, A., et al., 2000. Conditional identification of phosphate-starvation-response mutants in *Arabidopsis thaliana*. *Planta* 211, 13–22.
- Deshpande, R.A., Shankar, V., 2002. Ribonucleases from T2 family. *Crit. Rev. Microbiol.* 28, 79–122.
- Dreesen, R.S.G., Vanholme, B.T.M., Luyten, K., et al., 2010. Analysis of *Malus S-RNase* gene diversity based on a comparative study of old and modern apple cultivars and European wild apple. *Mol. Breed.* 26, 693–709.
- Du, J., Ge, C., Li, T., et al., 2021. Molecular characteristics of *S-RNase* alleles as the determinant of self-incompatibility in the style of *Fragaria viridis*. *Hortic. Res.* 8, 185.
- Du, J., Lv, Y., Xiong, J., et al., 2019. Identifying genome-wide sequence variations and candidate genes implicated in self-incompatibility by resequencing *Fragaria viridis*. *Int. J. Mol. Sci.* 20, 1039.
- Edger, P.P., VanBuren, R., Colle, M., et al., 2018. Single-molecule sequencing and optical mapping yields an improved genome of woodland strawberry (*Fragaria vesca*) with chromosome-scale contiguity. *GigaScience* 7, 1–7.
- Emms, D.M., Kelly, S., 2015. OrthoFinder: solving fundamental biases in whole genome comparisons dramatically improves orthogroup inference accuracy. *Genome Biol.* 16, 1–14.
- Evans, W., Jones, J.K., 1967. Incompatibility in *Fragaria*. *Can. J. Genet. Cytol.* 9, 831–836.
- Feng, C., Wang, J., Harris, A.J., et al., 2021. Tracing the diploid ancestry of the cultivated octoploid strawberry. *Mol. Biol. Evol.* 38, 478–485.
- Feng, G., Burleigh, J.G., Braun, E.L., et al., 2017. Evolution of the 3R-MYB gene family in plants. *Genome Biol. Evol.* 9, 1013–1029.
- Finn, R.D., Clements, J., Eddy, S.R., 2011. HMMER web server: interactive sequence similarity searching. *Nucleic Acids Res.* 39, 29–37.
- Folta, K.M., Davis, T.M., 2006. Strawberry genes and genomics. *CRC Crit. Rev. Plant Sci.* 25, 399–415.
- Gasteiger, E., Hoogland, C., Gattiker, A., et al., 2005. Protein Identification and Analysis Tools on the ExPASy Server. In: *The Proteomics Protocols Handbook*, pp. 571–607.
- Grabherr, M.G., Haas, B.J., Yassour, M., et al., 2011. Trinity: reconstructing a full-length transcriptome without a genome from RNA-Seq data. *Nat. Biotechnol.* 29, 644–652.
- Greco, M., Chiappetta, A., Bruno, L., et al., 2012. In *Posidonia oceanica* cadmium induces changes in DNA methylation and chromatin patterning. *J. Exp. Bot.* 63, 695–709.
- Guindon, S., Dufayard, J.F., Lefort, V., et al., 2010. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Syst. Biol.* 59, 307–321.
- Hollender, C.A., Geretz, A.C., Slovin, J.P., et al., 2012. Flower and early fruit development in a diploid strawberry, *Fragaria vesca*. *Planta* 235, 1123–1139.
- Holub, E.B., 2001. The arms race is ancient history in *Arabidopsis*, the wildflower. *Nat. Rev. Genet.* 2, 516–527.
- Honsho, C., Ushijima, K., Anraku, M., et al., 2021. Association of T2/S-RNase with self-incompatibility of Japanese citrus accessions examined by transcriptomic, phylogenetic, and genetic approaches. *Front. Plant Sci.* 12, 121.
- Hu, B., Jin, J., Guo, A.Y., et al., 2015. GSDS 2.0: an upgraded gene feature visualization server. *Bioinformatics* 31, 1296–1297.
- Hummer, K.E., Hancock, J., 2009. Strawberry Genomics: Botanical History, Cultivation, Traditional Breeding, and New Technologies. In: *Genetics and Genomics of Rosaceae*. Springer, New York.
- Igic, B., Kohn, J.R., 2001. Evolutionary relationships among self-incompatibility RNases. *Proc. Natl. Acad. Sci. U.S.A.* 98, 13167–13171.
- Irie, M., 1999. Structure-function relationships of acid ribonucleases: lysosomal, vacuolar, and periplasmic enzymes. *Pharmacol. Ther.* 81, 77–89.
- Ishimizu, T., Shinkawa, T., Sakiyama, F., et al., 1998. Primary structural features of rosaceous S-RNases associated with gametophytic self-incompatibility. *Plant Mol. Biol.* 37, 931–941.
- Jiang, S., An, H., Xu, F., et al., 2020. Chromosome-level genome assembly and annotation of the loquat (*Eriobotrya japonica*) genome. *Gigascience* 9, gaa015.
- Jiangtao, C., Yingzhen, K., Qian, W., et al., 2015. MapGene2Chrom, a tool to draw gene physical map based on Perl and SVG languages. *Hereditas* 37, 91–97.
- Jost, W., Bak, H., Glund, K., et al., 1991. Amino acid sequence of an extracellular, phosphate-starvation-induced ribonuclease from cultured tomato (*Lycopersicon esculentum*) cells. *Eur. J. Inorg. Chem.* 198, 1–6.
- Katoh, K., 2002. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Res.* 30, 3059–3066.
- Kim, D., Langmead, B., Salzberg, S.L., 2015. HISAT: a fast spliced aligner with low memory requirements. *Nat. Methods* 12, 357–360.
- Köck, M., Theierl, K., Stenzel, I., et al., 1998. Extracellular administration of phosphate-sequestering metabolites induces ribonucleases in cultured tomato cells. *Planta* 204, 404–407.
- Kubo, K.-i., Entani, T., Takara, A., et al., 2010. Collaborative non-self recognition system in S-RNase-based self-incompatibility. *Science* 330, 796–799.
- Kumar, S., Stecher, G., Tamura, K., 2016. MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. *Mol. Biol. Evol.* 33, 1870–1874.
- Kurata, N., Kariu, T., Kawano, S., et al., 2002. Molecular cloning of cDNAs encoding ribonuclease-related proteins in *Nicotiana glauca* leaves, as induced in response to wounding or to TMV-infection. *Biosci. Biotechnol. Biochem.* 66, 391–397.
- Lei, J., Xue, L., Guo, R., et al., 2016. The *Fragaria* species native to China and their geographical distribution. *Acta Hortic.* 1156, 37–46.
- Li, Y., Pi, M., Gao, Q., et al., 2019. Updated annotation of the wild strawberry *Fragaria vesca* V4 genome. *Hortic. Res.* 6, 61.
- Liang, M., Cao, Z., Zhu, A., et al., 2020. Evolution of self-compatibility by a mutant Sm-RNase in citrus. *Nat. Plants* 6, 131–142.
- Liston, A., Cronn, R., Ashman, T.L., 2014. *Fragaria*: a genus with deep historical roots and ripe for evolutionary and ecological insights. *Am. J. Bot.* 101, 1686–1699.
- Lu, S., Wang, J., Chitsaz, F., et al., 2020. CDD/SPARCLE: the conserved domain database in 2020. *Nucleic Acids Res.* 48, 265–268.
- Luhtala, N., Parker, R., 2010. T2 Family ribonucleases: ancient enzymes with diverse roles. *Trends Biochem. Sci.* 35, 253–259.
- MacIntosh, G.C., 2011. RNase T2 family: enzymatic properties, functional diversity, and evolution of ancient ribonucleases. In: *Ribonucleases*. Springer, pp. 89–114.
- MacIntosh, G.C., Hillwig, M.S., Meyer, A., et al., 2010. RNase T2 genes from rice and the evolution of secretory ribonucleases in plants. *Mol. Genet. Genomics* 283, 381–396.
- McClure, B.A., Haring, V., Ebert, P.R., et al., 1989. Style self-incompatibility gene products of *Nicotiana glauca* are ribonucleases. *Nature* 342, 955–957.
- Meyers, B.C., Kozik, A., Griego, A., et al., 2003. Genome-wide analysis of NBS-LRR-encoding genes in *Arabidopsis*. *Plant Cell* 15, 809–834.
- Mota, M., Tavares, L., Oliveira, C.M., 2007. Identification of S-alleles in pear (*Pyrus communis* L.) cv. ‘Rocha’ and other European cultivars. *Sci. Hortic.* 113, 13–19.

- Njuguna, W., Liston, A., Cronn, R., et al., 2013. Insights into phylogeny, sex function and age of *Fragaria* based on whole chloroplast genome sequencing. *Mol. Phylogenet. Evol.* 66, 17–29.
- Ortega, E., Sutherland, B.G., Dicenta, F., et al., 2005. Determination of incompatibility genotypes in almond using first and second intron consensus primers: detection of new *S* alleles and correction of reported *S* genotypes. *Plant Breed.* 124, 188–196.
- Pertea, M., Pertea, G.M., Antonescu, C.M., et al., 2015. StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. *Nat. Biotechnol.* 33, 290–295.
- Qiao, Q., Edger, P.P., Xue, L., et al., 2021. Evolutionary history and pan-genome dynamics of strawberry (*Fragaria* spp.). *Proc. Natl. Acad. Sci. U.S.A.* 118, e2105431118.
- Qiao, X., Yin, H., Li, L., et al., 2018. 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.* 9, 161.
- Ramanauskas, K., Igić, B., 2017. The evolutionary history of plant T2/S-type ribonucleases. *PeerJ* 5, e3790.
- Ramanauskas, K., Igić, B., 2021. RNase-based self-incompatibility in cacti. *New Phytol.* 231, 2039–2049.
- Roalson, E., McCubbin, A.G., 2003. S-RNases and sexual incompatibility: structure, functions, and evolutionary perspectives. *Mol. Phylogenet. Evol.* 29, 490–506.
- Saint-Oyant, L., Ruttink, T., Hamama, L., et al., 2018. A high-quality genome sequence of *Rosa chinensis* to elucidate ornamental traits. *Nat. Plants* 4, 473–484.
- Sassa, H., Hirano, H., Ikehashi, H., 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., Nishio, T., Kowayama, Y., et al., 1996. Self-incompatibility (*S*) alleles of the Rosaceae encode members of a distinct class of the T2/S ribonuclease superfamily. *Mol. Gen. Genet.* 250, 547–557.
- Sonneveld, T., Tobutt, K.R., Robbins, T.P., 2003. Allele-specific PCR detection of sweet cherry self-incompatibility (*S*) alleles *S1* to *S16* using consensus and allele-specific primers. *Theor. Appl. Genet.* 107, 1059–1070.
- Sonneveld, T., Tobutt, K.R., Vaughan, S.P., et al., 2005. 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* 17, 37–51.
- Steinbachs, J.E., Holsinger, K.E., 2002. S-RNase-mediated gametophytic self-incompatibility is ancestral in eudicots. *Mol. Biol. Evol.* 19, 825–829.
- Taylor, C.B., Bariola, P.A., Delcardayre, S.B., et al., 1993. RNS2: a senescence-associated RNase of *Arabidopsis* that diverged from the S-RNases before speciation. *Proc. Natl. Acad. Sci. U.S.A.* 90, 5118–5122.
- Thompson, J.D., Gibson, T.J., Higgins, D.G., 2003. Multiple sequence alignment using ClustalW and ClustalX. *Curr. Protoc. Bioinformatics* 1, 2–3.
- Wang, Y., Tang, H., DeBarry, J.D., et al., 2012. MCScanX: a toolkit for detection and evolutionary analysis of gene synteny and collinearity. *Nucleic Acids Res.* 40, e49.
- Wu, J., Wang, Z., Shi, Z., et al., 2013. The genome of the pear (*Pyrus bretschneideri* Rehd.). *Genome Res.* 23, 396–408.
- Xue, Y., Carpenter, R., Dickinson, H.G., et al., 1996. Origin of allelic diversity in antirrhinum *S* locus RNases. *Plant Cell* 8, 805–814.
- Ye, M., Peng, Z., Tang, D., et al., 2018. Generation of self-compatible diploid potato by knockout of *S-RNase*. *Nat. Plants* 4, 651–654.
- Zhang, J., Lei, Y., Wang, B., et al., 2020. The high-quality genome of diploid strawberry (*Fragaria nilgerrensis*) provides new insights into anthocyanin accumulation. *Plant Biotechnol. J.* 18, 1908–1924.
- Zhu, X., Li, Q., Tang, C., et al., 2020. Comprehensive genomic analysis of the RNase T2 gene family in Rosaceae and expression analysis in *Pyrus bretschneideri*. *Plant Syst. Evol.* 306, 1–17.