**Abstract**

Self-incompatibility (SI) substantially restricts the yield and quality of citrus. Therefore, breeding and analyzing self-compatible germplasm is of great theoretical and practical significance for citrus. Here, we focus on the mechanism of a self-compatibility mutation in ‘Guiyou No. 1’ pummelo (Citrus maxima), which is a spontaneous mutant of ‘Shatian’ pummelo (Citrus maxima, self-incompatibility). The rate of fruit set and the growth of pollen tubes in the pistil confirmed that a spontaneous mutation in the pistil is responsible for the self-compatibility of ‘Guiyou No. 1’. Segregation ratios of the S genotype in F1 progeny, expression analysis, and western blotting validated that the reduced levels of S2-RNase mRNA contribute to the loss of SI in ‘Guiyou No. 1’. Furthermore, we report a phased assembly of the ‘Guiyou No. 1’ pummelo genome and obtained two complete and well-annotated S haplotypes. Coupled with an analysis of SV variations, methylation levels, and gene expression, we identified a candidate gene (CgHB40), that may influence the regulation of the S2-RNase promoter. Our data provide evidence that a mutation that affects the pistil led to the loss of SI in ‘Guiyou No. 1’ by influencing a poorly understood mechanism that affects transcriptional regulation. This work significantly advances our understanding of the genetic basis of the SI system in citrus and provides information on the regulation of S-RNase genes.

**Introduction**

Self-incompatibility (SI) is a reproductive strategy adopted by flowering plants to prevent inbreeding and promote outcrossing; therefore, it is a reproductive strategy that promotes genetic diversity\(^1\). The SI system has been classified into two types, sporophytic SI (SSI) and gametophytic SI (GSI), based on the genetic mechanism controlling the SI phenotype of the pollen\(^2,3\). A single polymorphic locus, the S-locus, contains two tightly linked S genes—the pollen and pistil S determinants—that confer the ability of individual plants to recognize and reject their own pollen\(^1,4\). Each SI system characterized to date relies on complex genetic mechanisms. In the S-RNase-based GSI system, the pistil S determinant encodes an exocrine protein (S-RNase) acting as a cytotoxic factor that inhibits the growth of pollen tubes by degrading RNAs\(^6\). Additionally, the S-locus F-box protein (SLF) serves as a pollen S determinant. SLF is a component of a Skp1-Cullin1-F-box (SCF) complex that promotes the growth of nonself pollen tubes by ubiquitinating and degrading nonself S-RNases in a 26S proteasome-dependent manner\(^7-10\). However, SI can limit reproduction, especially during agricultural production. This is a short-term disadvantage of SI that is expected to promote frequent transitions to SC\(^5\).

Loss of SI in plants is usually caused by mutations in key genes controlling self-incompatibility, mainly pistil S determinants\(^11,12\). Studies on Petunia inflata\(^13\), Petunia axillaris\(^14\), and Solanum tuberosum\(^15\) indicated that SI could be switched to SC by directly attenuating or...
abolishing the expression of the \(S\)-RNase gene. In addition, epigenetic modifications in the 5′-flanking region of \(S\)-RNase can significantly affect its expression and lead to the loss of SI in almond\(^{16}\). In brief, the \(S\)-haplotype-specific rejection of pollen requires the maintenance of a high expression level of the \(S\)-RNase gene\(^8\). Additionally, the ribonuclease activity of \(S\)-RNases is essential for the rejection of pollen\(^8,17\). The reduction in \(S\)-RNase enzyme activity contributed to the loss of SI in Chinese pear (\(P\)yrus \(b\)retsneideri \(R\)ehd.)\(^{18,19}\) and \(P\)etunia \(i\)flata\(^{20}\). Citrus is one of the most widely cultivated and economically important fruit crops worldwide. Many citrus accessions are self-incompatible, such as Hassaku (\(C\)itrus \(h\)assaku \(h\)ort. ex \(T\)anaka)\(^{21}\), Hyuganatsu (\(C\)itrus \(t\)amurana \(h\)ort. ex \(T\)anaka)\(^{22}\), ‘Wuzishatangju’ (\(C\)itrus \(r\)eticulata \(B\)lanco)\(^{23}\), ‘Shatian’ pummelo (\(C\)itrus \(g\)randis \(O\)sbeck)\(^{24}\), Clementine mandarin (\(C\)itrus \(c\)lementina \(H\)ort. ex \(T\)an.)\(^{25}\), ‘W. Murcott Afourer’ mandarin (\(C\)itrus \(r\)eticulata)\(^{26}\), ‘Xiangshui’ lemon [\(C\). \(l\)imon (\(L\).) \(B\)urm. \(F\).]\(^{27}\), and ‘Kazgi Kalan’ lemon (\(C\)itrus \(l\)imon)\(^{28}\). Our previous studies have demonstrated that in citrus, \(S\)-RNase is a female determinant of GSI \(^{29–33}\). The full-length cDNA clones of \(S\)-RNase in citrus contain coding regions ranging from 660 to 699 bp that encode highly polymorphic proteins with predicted molecular masses between 25.1 and 26.8 kDa, and alkaline isoelectric points ranging from 7.69 to 9.44. Each amino acid sequence contains five conserved domains and two hypervariable regions (Supplementary Fig. 2). The single intron in each gene is located between the two hypervariable regions\(^{29}\). In citrus, a number of studies have been performed on the transition from SI to SC. For instance, ‘Wuzishatangju’ (\(C\). \(r\)eticulata \(B\)lanco, SI) was derived from a bud sport of ‘Shatangju’ (\(C\). \(r\)eticulata Blanco, SC)\(^{23,34}\), ‘Monreal’ (\(C\)itrus \(c\)lementina \(H\)ort. ex \(T\)an., SC) was derived from a self-compatible mutant of ‘Comune’ (\(C\)itrus \(c\)lementina \(H\)ort. ex \(T\)an., SI) that occurred naturally\(^{35}\). ‘Nishiuchi Konatsu’ (\(C\)itrus \(t\)amurana \(h\)ort. ex \(T\)anaka, SC) was derived from a bud sport variation of ‘Hyuganatsu’ (\(C\)itrus \(t\)amurana \(h\)ort. ex \(T\)anaka, SI)\(^{22}\). Nevertheless, mechanistic studies on the mutations that influence self-compatibility and self-incompatibility in citrus are rare, with the exception of \(S\)\(_m\)\(^{29}\).

In this article, we investigated the robust self-compatibility of ‘Guiyou No.1’ pummelo (\(C\)itrus \(m\)axima, hereafter referred to as \(GY\)), which is a spontaneous mutant of ‘Shatian’ pummelo (\(C\)itrus \(m\)axima, referred to as \(ST\)) (Fig. 1). Data from our pollen tube growth assays and the segregation ratios of the \(S\) genotype in the \(F\_1\) progeny indicate that the downregulated expression of \(S\)_2-RNase contributes to the loss of SI in ‘GY’. Moreover, we used a procedure that included an analysis of SV variation, methylation levels, and gene expression data to identify a candidate gene (\(C\)g\(HB\)\(_4\)) that potentially contributes to

---

**Fig. 1 Images of pummelo fruits, leaves, and floral organs. A, C ‘Guiyou No. 1’ pummelo (\(C\)itrus \(m\)axima); B, D ‘Shatian’ pummelo (\(C\)itrus \(m\)axima)**
the regulation of the promoter from S2-RNase. This work may fill a gap in our knowledge of the upstream regulation of S-RNase genes in citrus.

Materials and methods

Plant materials and sample extraction

The plant materials used in this study were from two cultivars of pummelo (Citrus maxima) that were collected from Rong County, Yulin City, Guangxi Province, China. One cultivar is the self-incompatible ‘Shatian’ pummelo (hereafter abbreviated as ST). The other cultivar is the ‘Guiyou No. 1’ pummelo (hereafter abbreviated as GY), which was derived from a bud sport of ‘ST’ (Fig. 1 and Supplementary Table 1). Fresh leaves and floral organs of ‘ST’ and ‘GY’ were collected, frozen immediately in liquid nitrogen, and then stored at −80 °C for further analysis. Fresh leaf tissues were used for genomic sequencing.

Pollination assay

To test whether ‘GY’ pummelo is an SC variety, self- and cross-pollinations were carried out with ‘ST’ and ‘GY’. Nearly opened flowers from the two cultivars were collected, and pollen was obtained by manually separating and drying the anthers at 28 °C for 12 h. One d before anthesis, floral buds were emasculated and hand-pollinated using a fine paint brush and covered with a paper pollination bag. Approximately 60–70 flowers were pollinated on 3–4 individual plants for each type of cross. Aniline blue staining was performed as described by Liang et al.36.

RT-PCR analysis

Total RNA was extracted with RNAiso Plus (TaKaRa) according to the manufacturer’s instructions. cDNA was synthesized with HiScript II Q RT SuperMix for qPCR (+gDNA wiper) (Vazyme Biotech) according to the supplier’s protocol. For RT-PCR, 100 μg of cDNA was added to each reaction system. The PCR program for all isoforms was as follows: 95 °C for 3 min, 30× (95 °C for 15 s, 57 °C for 15 s, and 72 °C for 30 s), and 72 °C for 5 min. Products were separated on 1.5% TAE-agarose gels and photographed. All samples were run in triplicate, and there were two separate runs for each plate. The pertinent primers are listed in Supplementary Table 4.

Expression and purification of recombinant S-RNase

The coding sequence from S1 and S2-RNase without the signal peptide was cloned into pGEX-6p-1 to create the GST-S-RNase fusion protein in an E. coli strain overexpressing C43 (DE3) (WeiDi Biotech) as previously described29. GST fusion proteins were overexpressed in E. coli cultures when the OD600 was 0.6–0.8 by adding 0.2 mM IPTG. Cells were harvested after 16 h at 20 °C. The recombinant S1-RNase and S2-RNase proteins were purified with GST SefnoseTM Resin (Sangon Biotech) as recommended by the manufacturer. For the RNase activity analysis, the solvent was exchanged with a solution containing 0.1 M sodium succinate and 0.01 M KCl at pH 6.0 using an Amicon centrifugation device (molecular mass cutoff, 10 kDa) (Millipore) at 4 °C. Protein concentrations of the samples were determined with an Omni-Easy™ Instant BCA Protein Assay Kit, and the samples were stored at −80 °C.

S-RNase activity analysis

RNase activity assays were performed as previously described with slight modifications37. Torula yeast RNA (10 mg ml−1, Sangon Biotech) dissolved in 0.1 M sodium succinate, and 0.01 M KCl, pH 6.0, was incubated at 55 °C for 10 min before being used as the substrate. Then, 125 μL of reaction buffer containing 0.1 M sodium succinate (pH 6.0), 0.01 M KCl, and 10 μg protein was also incubated at 55 °C for 10 min before the addition of 100 μL of RNA solution. The reaction was carried out for 30 min at 55 °C and stopped by the addition of 25 μL of ice-cold trichloroacetic acid (50%). Following a 15 min incubation on ice, the solution was centrifuged at 12,000 × g for 10 min. The supernatant was diluted 25-fold with ultrapure water. The absorbance was measured at 260 nm with an Infinite M200 Pro NanoQuant (TECAN). Specific activity was defined as the A260 units released from torula yeast RNA into an acid-soluble form from 1.0 mg of the protein in 1 min at 55 °C. The data are presented as the mean values ± SE from three independent biological replicates.

Protein extraction from styles and western blot analysis

Styles were excised from the pistils of flowers that were collected 1 d before anthesis. Total protein was extracted from styles using the Plant Protein Extraction Kit (epiTACS) according to the manufacturer’s protocol. Approximately 5–7 μg of total protein was analyzed with 10% SDS gels and then by immunoblotting using PVDF membranes. Polyclonal anti-S1-RNase antibodies (1:1,000 dilution) and anti-S2-RNase antibodies (1:1,000 dilution) were generated against the S1-RNase-GST and S2-RNase-GST fusion proteins in rabbits, respectively (FriendBio Technology, Wuhan, Hubei, China). These antibodies were used to detect S1-RNase and S2-RNase proteins, respectively, using a goat anti-rabbit IgG-horseradish peroxidase-conjugated secondary antibody (Abbkine, 1:5,000 dilution).

Genomic and transcriptomic sequencing

The sequencing of single molecules was performed on the PacBio Sequel II platform (Pacific Biosciences, USA) by Beijing Biomarker Technologies Co. Ltd. (Beijing,
China). A SMRTbell library with a size of 30 kb was constructed for genomic sequencing following the standard protocol provided by Pacific Biosciences. To survey the genome information, a 300 bp library was constructed to generate short reads (150 bp paired-end) using the BGISEQ-500 platform (BGI-Shenzhen, China).

After total RNA extraction and DNase I treatment, magnetic beads with Oligo (dT) were used to isolate mRNA. The mRNA was mixed with fragmentation buffer and fragmented into short fragments. cDNA was synthesized using mRNA fragments as templates. Short fragments were purified and resolved with EB buffer for end repair and single nucleotide A (adenine) addition. The short fragments were connected with adapters. Suitable fragments were selected and used as templates for PCR amplification. During the QC steps, the Agilent 2100 Bioanalyzer and the ABI StepOnePlus Real-Time PCR System were used for the quantification of the sample library. Subsequently, the library was sequenced using the BGISEQ-500 platform (BGI-Shenzhen, China).

**Genome assembly and S haplotype reconstruction**

A total of 220.94 Gb CLR reads were assembled using Canu (v2.1.1)\(^{39}\) with the following parameters: -- paccbo (minReadLength = 2000, minOverlapLength = 500, corOutCoverage = 150, corMinCoverage = 2 “batOptions = -dg 3 -db 3 -dr 1 -ca 500 -cp 50”, corrected error rate=0.035). Canu generated two assemblies composed of contigs and unitigs. The unitig assembly consisted of the contigs that split at any alternative path in the assembly graph. The contig assembly had longer continuity but contigs and unitigs. The unitig assembly consisted of the all-vs-all BLASTP (e-value = 0.95 identity).

The sequences of the $S_1$ and $S_2$ haplotypes were identified from the GYgv1 assembled genome using the $S$-RNase gene. The GY-$S_1$ and GY-$S_2$ loci from ‘GY’ pummelo were obtained using the conserved sequences at each end of the $S_1$ loci\(^{29}\). Gene predictions and annotations of GY-$S_1$ and GY-$S_2$ loci were made using FGENESH\(^{39}\) and non-redundant protein sequence (nr) databases\(^{40}\). Genes containing an F-box domain and an F-box-associated motif were designated SLFs. The ST-$S_1$ and ST-$S_2$ loci were obtained from publicly available databases (see Supplementary Tables 3, 5, and 6 for more details). The syntenic regions of GY-$S_1$ and GY-$S_2$ were identified using the all-vs-all BLASTP (e-value = 1e−5) method with a threshold value of 0.95 identity.

**Whole-sequence alignments and structural variation analysis**

Pairwise alignments of all $S$ loci were performed using NUCmer from MUMmer (v.4.0.0)\(^{41}\) with the --prefix option and HWB-$S_8$ loci as a reference\(^{42}\). The ‘ST’ $S_1$ and $S_2$ loci were obtained from publicly available databases (see Supplementary Table 3 for more details). Whole-sequence syntenic analysis between the ‘GY’ pummelo $S$ haplotypes was carried out with NUCmer and the --prefix option using the HWB-$S_8$ haplotype as a reference. Structural variants (SVs; >50 bp) were called using show-diff and show-snps, respectively. The SNPs were called using show-snps from MUMmer (v.4.0.0)\(^{41}\) and the -1 filter.

**Transcriptome analysis**

To gain insight into the mechanism responsible for self-compatibility in ‘GY’ pummelo, a comprehensive transcriptome analysis of the anther and style of ‘ST’ and ‘GY’ pummelo was performed to identify differentially expressed genes (DEGs) in different tissues. Anther and style samples were collected 1 d before anthesis. The reproducibility of transcriptome samples was assessed using a Spearman rank correlation. Expression levels were calculated as transcripts per kilobase per million mapped reads (TPM). The DEGs from the transcriptome data were analyzed using the R package DEseq2\(^{43}\). The significant DEGs were determined using a threshold of log2FoldChange > 1 with an adjusted $P$-value < 0.01. Furthermore, the GO (Gene Ontology) database and the KEGG (Kyoto Encyclopedia of Genes and Genomes) pathway database were utilized to identify enriched biological functions and metabolic pathways.

**Whole-genome bisulfite sequencing (WGBS) and analysis**

Genomic DNA was purified from the style tissue 1 d before anthesis using the CTAB method. For the construction of paired-end BS-seq libraries using the NEXTflex Bisulfite-Seq kit (Bioo Scientific), genomic DNA was converted to approximately 300 bp fragments using an ultrasonic homogenizer, end-repaired, and 3’ adenylated. The products were ligated with methylated adapters and subjected to size selection using the MiniElute PCR Purification Kit (QIAGEN). Bisulfite treatment was carried out using the EZ DNA Methylation-Gold kit (Zymo Research) according to the manufacturer’s instructions. The constructed libraries were validated using the Agilent Technologies 2100 bioanalyzer and paired-end sequenced with 150 bp reads on the BGISEQ-500 platform (BGI-Shenzhen, China). For data analysis, adapters and low-quality reads were removed. Clean reads were mapped to the GYgv1 genome using Bismark (v0.22.2)\(^{44}\), allowing up to three mismatches. For the identification of differentially methylated regions (DMRs), fractional DNA methylation levels were calculated using 200 bp sliding windows with 50 bp step sizes in the $S$-RNase regions of the ‘ST’ and ‘GY’ genomes. Only cytosines from both ‘ST’ and ‘GY’ that were covered by at
least four reads were considered and counted. The methylation level of individual cytosines was calculated as the ratio of mC to the total cytosines [mC/(mC+un-mC)].

Yeast one-hybrid assay
The Y1H assay was conducted as previously described.45 As bait, the promoter sequences of S2-RNase (2340 bp) were cloned into pAbAi vectors and used to transform the Y1HGol yeast strain. Then, the bait strains were transformed with the pGADT7-Cg1g003830.1 fusion plasmid. The cotransformed yeast cells were cultured on SD/-Leu agar plates with or without aureobasidin A (AbA) and incubated at 30°C for 2-3 d.

Dual-luciferase (LUC) assay
For LUC, an effector vector was constructed by cloning the full-length CgHB40 CDS into the pGreenII 62-SK vector, whereas reporter vectors were generated by fusing the pS2-RNase promoter fragment (735 bp, the second part of S2-RNase promoter) into pGreenII 0800-LUC. Effectors and reporters were cotransformed into Nicotiana benthamiana leaves and the activity of LUC and REN was measured using the Dual-Luciferase® Reporter Assay System (Promega, USA). Promoter activity was determined by the ratio of LUC/REN. LUC was imaged with the Plant Fluorescent Imaging System in Vivo after the application of luciferin on the abaxial leaf surface.

Results
A pistil-side mutation confers self-compatibility to ‘GY’ pummelo
‘GY’ is a spontaneous mutant of ‘ST’ (S1S2). The S genotype was determined to be S1S2 using specific primer pairs (Supplementary Table 1). To test the stability of the self-compatibility in ‘GY’, the fruit set rate and the average number of seeds per fruit after the self-pollination and cross-pollination of ‘GY’ and ‘ST’ were determined (Table 1). The different percentages of fruit set from the self-pollination of ‘ST’ and ‘GY’, 2.5 and 30.8%, respectively, indicated that ‘GY’ possesses self-compatability. The fruit produced by the F1 population that was derived from the ‘GY’ × ‘GY’ cross contained tenfold more seeds than the fruit produced by the F1 population that was derived from the ‘ST’ × ‘ST’ cross. The germination rate of the F1 seeds derived from the ‘GY’ × ‘GY’ cross was nearly 100%. In contrast, the seeds derived from the self-pollination of ‘ST’ barely germinated (Table 2). There was another significant difference in the outcome of the reciprocal crosses between ‘GY’ and ‘ST’. The fruit set ratio of the cross-pollinated ‘GY’ × ‘ST’ was remarkably higher than the fruit set ratio of the cross-pollinated ‘ST’ × ‘GY’, although the number of seeds per fruit from these two cross-pollination experiments was not significantly different (Table 1).

When ‘ST’ and ‘GY’ were used as sources of pollen for ‘ST’ pistils at 1 d before anthesis, virtually no pollen tube growth was observed at the middle and base of the styles at 5 d post pollination. In contrast, when ‘GY’ was used as the pollen receptor for pollen from ‘ST’ and ‘GY’, pollen tubes were observed at the bottom of the styles (Fig. 2). Therefore, ‘GY’ developed an anther-side SI phenotype but not a pistil-side SI phenotype. These observations are consistent with the fruit-set rate of ‘GY’ after the self- and cross-pollination experiments (Table 1). These data indicate that ‘GY’ is stably self-compatible and that the loss of SI was probably caused by a pistil-side mutation.

Downregulated expression of S2-RNase promotes SC in ‘GY’ pummelo
To identify the pistil-side mutation responsible for the transition from SI to SC in ‘GY’, the segregation ratios of the S genotypes in the F1 progeny from the four pollination combinations were analyzed using chi-square tests. None of the 11 seeds generated by the self-pollinated ‘ST’ germinated. In contrast, the S haplotypes of the progeny from the self-pollinated ‘GY’ segregated into two classes (S1S2:S2S2) at a ratio of 274:256 (Table 2; Fig. 3D). This ratio is consistent with an abnormality linked to the S2 locus (1:1, χ² = 0.6113, P > 0.05; Table 2). Furthermore, 309 individual F1 progeny derived from the ‘GY’ × ‘ST’ cross exhibited either the S1S2 (150 plants) or S2S2 (159 plants) genotypes in the expected 1:1 ratio (χ² = 0.2621, P = 0.6087; Table 2; Fig. 3E). Nevertheless, the S1S1 genotype was absent from the F1 progeny, although it was a possible outcome of crosses that utilized pistils from ‘GY’ (Fig. 3D, E). These data demonstrate that functional S1-RNase is present in the pistils of ‘GY’, resulting in a partial ability to reject pollen from ‘ST’ and ‘GY’. In contrast, the S2 genotype existed in all of the individual F1 progeny produced by the pistils from ‘GY’ (Fig. 3D, E). These data indicate that dysfunction of S2-RNase may disrupt SI in ‘GY’.

Therefore, we analyzed the expression of S-RNase genes in floral tissues. Not surprisingly, S1-RNase and S2-RNase

| Pollination | No. of pollinated flowers | No. of fruit | Fruit set ratio (%) | No. of total seeds | No. of seeds per fruit |
|-------------|---------------------------|--------------|---------------------|-------------------|-----------------------|
| GYG × GYS | 26                        | 8            | 30.8                | 1109              | 138.6                 |
| GYG × STS | 37                        | 5            | 13.5                | 743               | 148.6                 |
| STS × STS | 40                        | 1            | 2.5                 | 11                | 11                    |
| STS × GYS | 27                        | 1            | 3.7                 | 116               | 116                   |

The data in Table 1 were collected in 2020.
were expressed normally in the pistils of ‘ST’. In contrast, only the expression of S1-RNase was detected in the pistils of ‘GY’ (Fig. 3A, B; Supplementary Fig. 4). Subsequently, the abundance of S2-RNase proteins in the pistils of ‘GY’ was investigated. For this purpose, crude protein extracts from the style and leaf tissues of ‘GY’ and ‘ST’ were prepared. The results from the western blot analysis were consistent with the results from RT-PCR in that although the S1-RNase protein was detected in the styles of both ‘GY’ and ‘ST’, the S2-RNase protein was present only in the styles of ‘ST’ (Fig. 3C). Moreover, the analysis of RNase activity indicated that the activity of S2-RNase was higher than the activity of S1-RNase (Supplementary Fig. 3A). These data provide evidence that both the S2-RNase and S1-RNase proteins are functional S-determinants and that the attenuated expression of S2-RNase induced a transition from SI to SC in ‘GY’.

**Diversity in S loci does not attenuate the expression of S2-RNase**

To better understand the attenuated expression of S2-RNase in the style of ‘GY’, we cloned the S2-RNase gene and 2 kb of sequence upstream of the S2-RNase gene. We compared this sequence to the S2-RNase gene from ‘ST’ and found no differences in the 2 kb 5’ flanking sequence in ‘GY’ and ‘ST’ (i.e., no SNPs or other types of polymorphisms were detected) (Supplementary Fig. 2). Some structural variants (SVs) of the S-locus can be responsible for the loss of self-incompatibility. Thus, we sequenced and assembled the genome of ‘GY’. A total of 3,969,910,472 subreads and 220.94 Gb (= 552X, Table 3 and Supplementary Fig. 1) of clean data from PacBio long reads were generated from sequencing one whole flow cell. The genome was based on single-molecule real-time (SMRT) DNA sequencing and de novo assembled with Canu, which produces a partially phased diploid genome. Finally, a total of 1,429 contigs (≥100 bp) were assembled into 566.29 Mb with a contig N50 of 1.11 Mb. We used a BUSCO analysis to evaluate the integrity of the genome assembly. Among the 2,326 conserved genes, 2,252 orthologs (98.9%) of these genes were completely aligned to the assembled genome. Only 19 orthologs (0.9%) were not detected in the genome (Table 3).

The S haplotypes were identified by the presence of the S-RNase genes on the tig00000920 and tig00000926 contigs of the GYgv1 reference genome. The GY-S1 and GY-S2 loci from the ‘GY’ pummelo were obtained using the conserved sequences at each end of the S loci. In our previous work, the sequences of the ST-S loci were obtained from a BAC library (Supplementary Table 3). All of the S haplotypes were aligned to the HWB C. maxima S_haplotypes to assess structural differences and to identify S locus-specific features. The sequences of the S loci were highly polymorphic. Whole-sequence alignment analysis showed that the sequences flanking both ends—approximately 37 kb on the left flank and approximately 43 kb on the right flank—of the S1, S2, and S_haplotypes were well conserved. The S loci that determine the self-incompatibility of citrus were located in the middle of this approximately 150 kb highly polymorphic region (Fig. 4B; Supplementary Fig. 3C). We manually annotated the genes in the region of the S locus using the GYgv1 genome sequence and RNA-Seq analysis. Compared with the ST-S2 and GY-S2 loci, we found that the S1 locus from ‘GY’ was completely identical to that of ‘ST’ and that there was no SV variation between these two S loci (Fig. 4A, C; Supplementary Fig. 3B). Collinearity analysis also showed that the structure and gene content of the region flanking the GY-S locus was highly conserved and that the 7 SLF genes (SLF1-1, SLF1-3, SLF1-4, SLF1-6, SLF1-7, SLF1-8, SLF1-9).
and SLF1-9 and SLF2-2, SLF2-6, SLF2-5, SLF2-3, SLF2-7, SLF2-8, and SLF2-9) from each S locus were highly collinear (Fig. 4D and Supplementary Tables 5–7), which was consistent with the previous studies.

An analysis of the mean transcript levels from the S1 and S2 loci in the style and anther tissues of 'ST' and 'GY' pummelo indicated no significant difference between the expression of the SLF and S1-RNase genes in 'ST' relative to 'GY' in each tissue with the exception of the expression of the S2-RNase gene in the style of 'GY'. The expression of the S2-RNase gene was significantly attenuated in the 'GY' style relative to the 'ST' style (Fig. 4D). Sequence and expression analysis indicated that variation in the S loci sequences is not responsible for the attenuated expression of the S2-RNase gene in the styles of 'GY'.

Epigenetic modifications make a major contribution to gene silencing. To test whether DNA methylation might attenuate the expression of S2-RNase, we used WGBS to profile DNA methylation in the styles from both 'ST' and 'GY'. Two biological replicates from each genotype were sequenced. For each sample, at least 120 M paired-end reads (read length = 150 bp) were produced. Approximately 87% of the reads were mapped to the GYgv1 reference genome using Bismark. These data covered >99.9% of the genome. The average depth of coverage at cytosines was 16.60-fold and 16.58-fold in the

---

**Fig. 2 Fluorescence images of pollen tubes in pistils after different pollinations.** Representative fluorescence images (five of each pollination combination) of aniline blue-stained pistils at stage -1 DBA (1 d before anthesis) 5 d after pollination are shown. The accession names and their respective S- genotypes are indicated to the left and at the top. Scale bars = 1 mm. Pollen tubes (pt) are indicated with arrows. Vascular bundles (vb) are indicated with arrowheads.
styles of ‘ST’ and ‘GY’, respectively (Supplementary Tables 8 and 9).

In the styles of ‘ST’ and ‘GY’ pummelo, ~20.2 and 19.3% of the cytosines are methylated, which were defined using a binomial test as described by Zhong. The average CpG, CHG, and CHH methylation levels were 38, 35, and 26%, respectively. The largest proportion of mCs was in the CHH context. The methylation rates of ‘ST’ and ‘GY’ in the CHH context were 7.3 and 7%, respectively. The second-highest proportion of mCs was in the CHH context. The numbers of mCs in the CHG context were 142.5 and 138.1 million in ‘ST’ and ‘GY’, respectively. The corresponding methylation percentages were higher than 35%. mCs in the CpG context were methylated at the highest average levels of methylation, reaching 7.8 and 7.4% in ‘ST’ and ‘GY’, respectively (Supplementary Table 8).

We calculated the average DNA methylation rates for the S-RNase regions and found that DNA methylation gradually increased in the 3’ flanking regions of the S2-RNase gene and that the largest proportion of mCs was in the CHH context. The corresponding methylation rates were 6.85 and 9.38%, respectively. The second-highest proportion of mCs was in the CHG context, which had the highest methylation rates. The corresponding methylation rates were 72.3 and 74.2%, respectively (Fig. 5A). However, the methylation rates of the promoter and gene body were significantly lower than those of the downstream region. The average methylation rates of the
Table 3  Summary statistics for assembly of the ‘GY’ pummelo genome sequence

| Sequencing Platform   | Library size |
|-----------------------|--------------|
| PacBio Sequel II      | 30 kb        |
| Clean data (Gb)       | Coverage (X) |
| 220.94                | 552          |

| Genome assembly Assembly | Primary assembly |
|--------------------------|-------------------|
| Assembly length (bp)     | 587,752,981       |
| Number of sequences      | 1,429             |
| Average length (bp)      | 411,303           |
| Maximum length (bp)      | 18,341,039        |
| Minimum length (bp)      | 9,822             |
| N50 length (bp)          | 1,159,493         |
| GC content (%)           | 34.6%             |

| BUSCO Types               | Percentage (%)   |
|---------------------------|------------------|
| Complete BUSCOs (C)       | 98.9             |
| Complete single-copy BUSCOs (S) | 23.6 |
| Complete and duplicated BUSCOs (D) | 75.3 |
| Fragmented BUSCOs (F)     | 0.2              |
| Missing BUSCOs (M)        | 0.9              |

promoter and S2-RNase gene body in the CpG, CHG, and CHH contexts were consistent with the methylation rates that were observed in the styles of ‘ST’ and ‘GY’. Moreover, the average methylation levels in the different contexts were similar in the regions containing S1-RNase and S2-RNase in ‘ST’ and ‘GY’ (Fig. 5A and Supplementary Fig. 5C). Thus, the similar levels and patterns of methylation in the regions containing S1-RNase and S2-RNase may not strongly contribute to the attenuated expression of S2-RNase in the styles of ‘ST’ and ‘GY’.

In the HWB C. maxima genome, we identified seven genes encoding methyltransferase orthologs using BLASTP. A phylogenetic tree was constructed using the maximum-likelihood method (Supplementary Fig. 5A, B). Three of these seven methyltransferase genes (Cg2g001280, Cg6g003970, and Cg3g008940) were barely expressed [transcripts per kilobase of exon model per million mapped reads (TPM) < 1] in style tissue (Fig. 5B and Supplementary Fig. 5D). Among the other four genes, there was no significant difference in expression levels in ‘GY’ relative to those in ‘ST’. This result is consistent with the similar levels of methylation in these two cultivars (Fig. 5B). Regarding DNA demethylase genes, we identified four AtROS1 orthologs, including Cg2g0035340 (CgDME), Cg6g016280 (CgDML1), Cg8g018140 (CgDML3), and Cg5g004510 (CgDML4). Among the four pummelo DNA demethylase genes, Cg5g004510 (TPM < 3) was expressed at lower levels in styles than the other three genes (TPM > 10). The expression of Cg6g016280 in the styles of ‘GY’ was significantly lower than that in the styles of ‘ST’. The relative expression levels of the three other genes were not significantly different in the style and anther tissues of ‘ST’ and ‘GY’ (Fig. 5C). The methyltransferase and DNA demethylase genes were expressed at higher levels in the styles than in the anthers (Fig. 5B, C). The results provide evidence that variation in the levels of methylation in the styles of ‘ST’ and ‘GY’ may not be caused by changes in the expression levels of any of the genes that encode DNA methyltransferases and demethylases. In conclusion, we did not observe a significant correlation between the average levels of mC in the regions containing the S-RNase loci and the attenuated expression of S2-RNase in the styles of ‘GY’.

Differential gene expression in pistils may attenuate the expression of S2-RNase

The mechanism that attenuates the expression of S2-RNase and thus contributes to the loss of SI in ‘GY’ pummelo was explored using a transcriptome profiling approach that utilized RNA-based sequencing (RNA-Seq) of the floral tissues in ‘ST’ and ‘GY’. Approximately 40 million clean reads per sample were obtained (Supplementary Table 2). Analysis of the Spearman rank correlation demonstrated consistency among the three biological replicates. Additionally, the correlation coefficient of each sample was higher than 0.98, indicating good reproducibility (Supplementary Fig. 6). Thus, these RNA-Seq data were used for further analysis. Although principal component analysis (PCA) indicated consistent expression profiles in the anthers of ‘ST’ and ‘GY’, the transcriptomes from the styles did not cluster together (Fig. 6A), which indicates that the transcriptomes from the styles were significantly different. We assumed that the different transcriptome profiles in the ‘ST’ and ‘GY’ styles may contribute to the different expression levels of S2-RNase mRNA.

To test this hypothesis, 9791 (style vs. anther in ‘GY’, from 1 d before anthesis (-1 DB)) data were obtained. Among these 9791 differentially expressed genes (DEGs) were identified using a ≥2 fold change and a false discovery rate (FDR) ≤0.05 as the criteria for differential expression (Fig. 6B). The 585 DEGs could explain the transition from SI to SC in ‘GY’ pummelo (Supplementary Table 10). To understand the main biological functions associated with the DEGs, Gene Ontology (GO) enrichment analysis was performed against the GO database using P ≤ 0.05 as the threshold. The main biological functions associated with the 585 DEGs were categorized into 90 biological
processes (BP), 21 cellular components (CC), and 161 molecular functions (MF) (Supplementary Table 11). In the BP category, the top three terms were ‘oxidation–reduction process’ (GO:0055114, 112, 5.87%), ‘protein phosphorylation’ (GO:0006468, 101, 5.29%), and ‘regulation of transcription, DNA-dependent’ (GO: 0006355, 38, 1.99%) (Fig. 6C). To gain insight into the metabolic pathways associated with the differentially expressed genes, the genes were annotated with a Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analysis against the KEGG database. Although the main metabolic processes associated with the 585 DEGs were classified into 62 enriched pathways, only nine of these pathways were significantly enriched \( (q \leq 0.05) \). The top three significantly enriched pathways were ‘plant hormone signal transduction’ (ko04075), ‘biosynthesis of secondary metabolites’ (ko01110), and ‘diterpenoid biosynthesis’ (ko00904). These significantly enriched pathways mainly belonged to environmental information processing and metabolism in the KEGG_A class (Fig. 6D, Supplementary Table 12). GO and KEGG analyses demonstrated that the regulation of transcription biological processes and signal transduction pathways may participate in the regulation of \( S_2 \)-RNase expression.

Therefore, we tried to demonstrate a rationale for the differential expression of \( S_2 \)-RNase in ‘GY’ relative to ‘ST’...
from the perspective of transcriptional regulation. We identified 29 transcription factor genes among the 585 DEGs. Based on the expression profiles in anther and style tissues, 29 TF genes clustered into four groups (Fig. 6E). Genes in clusters 1 and 4 were mainly expressed in styles with particular tissue specificities. Importantly, the expression levels of genes in ‘ST’ and ‘GY’ were significantly different (i.e., fold change > 2). Therefore, yeast one-hybrid (Y1H) assays were performed with these ten genes to test whether they could bind to the promoter of \(S_2\)-RNase in yeast. We found that only Cg1g003830 bound to the promoter from \(S_2\)-RNase (Fig. 6F and Supplementary Fig. 7). Thus, Cg1g003830 (hereafter referred to as CgHB40) was selected for further study. A 2340 bp
segment of the $S_2$-RNase promoter was used as bait in Y1H assays. CgHB40 was expressed as a fusion to the GAL4 activation domain. The full-length $S_2$-RNase promoter and several promoter fragments ($p_{S_2}$-F1, $p_{S_2}$-F3, $p_{S_2}$-F4, and $p_{S_2}$-F5) appeared to strongly self-activate such that the strains containing the full-length $S_2$-RNase
promoter and particular promoter fragments grew well on the selective medium. Indeed, 900 ng/mL AbA could not inhibit the growth of these strains. However, although pGADT7 + pS2-F2 did not grow on the selective medium (SD/-Leu + 500 ng/mL AbA), pGAD-CgHB40 + pS2-F2 did grow on the selective medium, indicating an interaction between CgHB40 and the pS2-F2 promoter (Fig. 6F).

Transient luciferase (LUC) imaging assays were also performed to verify the binding of CgHB40 to the S2-RNase promoter in vivo. In these assays, the CaMV35S promoter was used to express CgHB40 (i.e., the effector). The plasmid (pS2-F2pro::Luc) containing the promoter of S2-RNase fused to the reporter gene luciferase, was combined with 35Spro::CgHB40 and coinfiltrated into tobacco leaves. A strong luminescence signal was detected in the coexpressed regions of 35Spro + pS2-F2pro::Luc, but much weaker and no luminescence signals were detected in 35Spro::CgHB40 + pS2-F2pro::Luc and the negative controls, respectively. Taken together, these results indicate that CgHB40 can regulate the activity of the pS2-F2 promoter (Fig. 6G). Based on these data, CgHB40 is a strong candidate for a transcription factor that regulates the transcription of S2-RNase in styles. More experiments are required to determine whether CgHB40 contributes to the loss of SI in ‘GY’.

**Discussion**

The transition from an outcrossing to a selfing mating system is one of the most prevalent evolutionary transitions in flowering plants. Hence, the number of studies on the transition from SI to SC has increased. Our understanding of the mechanisms responsible for SI has improved. In this study, we used genetic and pollination assays to demonstrate that the loss of SI in ‘GY’ was due to a pistil-side mutation. The fruit setting rate of self-pollinated ‘GY’ was remarkably higher than the fruit setting rate of self-pollinated ‘ST’ (30.8 and 2.5%, respectively). Moreover, the results from the pollination experiments were consistent with the average seed numbers from the fruit (138.6 and 11, respectively) (Table 1). Nevertheless, the self-incompatibility in the styles of ‘ST’ should have prevented the self-pollinated ‘ST’ from producing fruit with seeds (Fig. 2). However, we obtained 11 and 116 individuals from one fruit each of the ‘ST’ self-cross and ‘ST’ × ‘GY’ cross combinations, respectively. The rates of fruit set were 2.5 and 3.7%, respectively (Table 1). These unexpected results are consistent with those of previous studies and may be explained by the environment or by the accidental pollination of one or two flowers in the budding stage. Nonetheless, our work suggests that a pistil-side mutation is responsible for self-compatibility in ‘GY’ pummelo.

The mechanisms responsible for the loss of SI can be grouped into three categories that involve gene duplications and mutations in the S-RNase, SLF, or non-S determinant genes. Not surprisingly, the results of our expression analysis and the segregation ratios of the S genotype in the F1 progeny demonstrated that the attenuated expression of S2-RNase contributed to the loss of SI in ‘GY’ pummelo (Fig. 3A–C; Supplementary Fig. 4). Therefore, to better understand the mechanism that attenuates the expression of S2-RNase in the styles of ‘GY’, we prepared a phased assembly of the ‘GY’ genome (GYgv1) and obtained two complete and well-annotated S haplotypes. To date, research related to pistil dysfunction in the realm of SI has mainly focused on sequence variation that results in the reduction of RNase activity and its expression. However, a comparison of the structural variation in ‘GY’ relative to ‘ST’ did not explain the loss of SI in ‘GY’. No duplications, inversions, deletions, or translocations were observed between the same S-loci of ‘ST’ and ‘GY’. Moreover, we did not observe SNPs in 5 kb regions on the 5’ and 3’ flanks of the S-RNase genes (Fig. 4A; Supplementary Fig. 3B). Thus, we were able to completely exclude the possibility that sequence variations in the S2 locus attenuated the expression of S2-RNase in ‘GY’.

The methylation of cytosine residues was proposed to suppress gene expression. The methylation of several cytosine residues in a 4700 bp fragment in the 5’ upstream region of S2-RNase was strongly associated with the inactivation of the Sf allele in all self-compatible varieties of almond. Here, we performed WGBS for ‘ST’ and ‘GY’
styles and calculated the average DNA methylation rates of the S-RNase regions. Despite epigenetic changes downstream of the S-RNase genes, the average methylation levels in different contexts were distributed similarly in the $S_1$-RNase and $S_2$-RNase regions of ‘ST’ and ‘GY’ (Fig. 5A and Supplementary Fig. 5C). The absence of significant differences in DNA methylation rates in the regions containing the $S_1$ and $S_2$ loci of ‘GY’ and ‘ST’ provides evidence that an unknown modifier gene may affect the expression of $S_2$-RNase genes. Obviously, these observations conflict with a previous study on methylated cytosome residues. The researchers argued that the farthest site (M4) was correlated with the silencing of the $S_f$-RNase gene in almond.\(^{16}\) Currently, this conclusion seems unreliable because this study did not completely exclude the influence of different nucleotide sequences, and the analysis of methylation was limited to the CpG context and did not cover the entire $S_f$-RNase region. However, we cannot conclude that methylation does not contribute to the regulation of S-RNase genes. Indeed, changes in methylation levels were detected in the noncoding region of S-RNase genes. Nonetheless, much more work is required to test this idea.

Many years ago, similar research was performed in Petunia axillaris. The author proposed a model that included a modifier locus (MDF) that is unlinked to the S locus and that could specifically affect the expression of $S_{13}$-RNase.\(^{14}\) MDF is probably a transcription factor that specifically regulates the expression of a particular S-RNase gene. Nevertheless, due to the polymorphic nature of S loci, little is known about the regulation of the S-RNase genes. Here, by performing a transcriptome analysis, a candidate gene ($C_{gHB40}$) was identified that may be involved in the transcriptional regulation of $S_{2}$-RNase genes in styles. Additional experiments are required to test whether the candidate contributes to the loss of SI in ‘GY’. This work significantly advances our understanding of the genetic basis of the breakdown of the SI system in citrus and provides the information necessary to regulate the expression of the S-RNase genes.

Acknowledgements
This research was financially supported by the National Key Research and Development Program of China (grant no. 2018YFD01000107), the National Natural Science Foundation of China (grant nos. 31772259, 31630365, and 31521092), the Fundamental Research Funds for the Central Universities (grant no. 2662019PY044), the China Agriculture Research System of MOC and MARA and the Hubei Provincial Natural Science Foundation of China (2020CFB532).

Author details
1Key Laboratory of Horticultural Plant Biology, Ministry of Education, College of Horticulture and Forestry Sciences, Huazhong Agricultural University, Wuhan 430070, People's Republic of China. 2Guangxi Academy of Specialty Crops, Guilin 541004, People's Republic of China. 3Guangxi Subtropical Crops Research Institute, Nanning 530001, People's Republic of China. 4Horticulture Research Institute, Guangxi Academy of Agriculture Sciences, Nanning Investigation & Experiment Station of South Subtropical Fruit Trees, Ministry of Agriculture, Nanning 530007, Guangxi, People's Republic of China. 5Horticultural Institute, Hunan Academy of Agricultural Sciences, Changsha 410125, People's Republic of China.

Author contributions
J.H. and L.C. conceived the project. J.H. conducted experiments and performed transcriptomic, genomic, and W.G.B.S. sequencing, and data acquisition. C.C., C. D., B.L. and X.C. provided the plant materials. Z.W. and K.P. collected the plant samples. C.L., Z.W. and H.W. conducted the P.C.R., Y.I.H, and L.U.C. assays. Q.X. helped with western blot and S-RNase activity analysis, and revised the paper. X.D. supervised the project and revised the paper. L.C. supervised the project. J. H. and Q.X. wrote the paper with help from R.L., M.A., P.C., J.Y., and L.C. All authors contributed to the article and approved the paper before submission.

Data availability
All of the sequence data and the genome assembly have been deposited in the CNGB Sequence Archive (CNSA)\(^{59}\) of the China National GeneBank Database (CNGBdb)\(^{60}\) with accession numbers CNP0001704, CNP0001705, and CNP0001706 (https://db.cnbg.org/).

Competing interests
The authors declare no competing interests.

Supplementary information
The online version contains supplementary material available at https://doi.org/10.1038/s41438-021-00634-8.

Received: 30 March 2021 Revised: 27 May 2021 Accepted: 1 June 2021

Published online: 01 September 2021

References
1. Yamamoto, M. & Nishio, T. Commonalities and differences between Brassica and Arabidopsis self-incompatibility. Hort. Res. 1, 14054 (2014).
2. Fuji, S., Kubo, K. & Takayama, S. Non-self- and self-recognition models in plant self-incompatibility. Nat. Plants 2, 16130 (2016).
3. Tao, R. & Iezzoni, A. F. The S-RNase-based gametophytic self-incompatibility system in Prunus exhibits distinct genetic and molecular features. Sci. Hort. 124, 423–433 (2010).
4. Matsumoto, D. & Tao, R. Distinct self-recognition in the Prunus S-RNase-based gametophytic self-incompatibility system. Hort. J. 85, 289–305 (2016).
5. Goldberg, E. E. et al. Species selection maintains self-incompatibility. Science 330, 493–495 (2010).
6. Wu, L. et al. Use of domain-swapping to identify candidate amino acids involved in differential interactions between two allelic variants of type-1 S-locus F-box protein and S3-RNase in Petunia infinta. Plant Cell Physiol. 59, 234–247 (2018).
7. Guo, H., Haltitsche, R., Welsch, N., Guse, K. & Baldwin, I. T. Mate selection in self-compatible wild tobacco results from coordinated variation in homologous self-incompatibility genes. Curr. Biol. 29, 2020–2030 e6 (2019).
8. Takayama, S. & Isogai, A. Self-incompatibility in plants. Annu. Rev. Plant Biol. 56, 467–489 (2005).
9. Hua, Z. & Kao, T. H. 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 Online 18, 2531–2553 (2006).
10. Ertani, T. et al. Ubiquitin-proteasome-mediated degradation of S-RNase in a solanaceous cross-compatibility reaction. Plant J. 78, 1014–1021 (2014).
11. Li, W. et al. Molecular and genetic characterization of a self-compatible apple cultivar, ‘CAU-T’. Plant Sci. 252, 162–175 (2016).
12. Baldwin, S. J. & Schoen, D. J. Genetic variation for pseudo-self-compatibility in self-incompatible populations of Leavenworthia alabamica (Brassicaceae). New Phytol. 213, 430–439 (2017).
13. Lee, H. S., Huang, S. & Kao, T. S proteins control rejection of incompatible pollen in Petunia infinta. Nature 367, 560–563 (1994).
14. Tsukamoto, T. et al. Breakdown of self-incompatibility in a natural population of Petunia axillaris caused by a modifier locus that suppresses the expression of an S-RNase gene. Sex. Plant Reprod. 15, 255–263 (2003).
15. Ye, M. et al. Generation of self-compatible diploid potato by knockout of S-RNase. Nat. Plants 4, 651–654 (2018).
16. Fernandez I Martí, A., Gradziel, T. M. & Socas I Company, R. Methylation of the Slocus in almond is associated with S-RNase loss of function. Plant Mol. Biol. 86, 681–689 (2014).

17. Huang, S., Lee, H. S., Karunananada, B. & Kao, T. H. Ribonuclease activity of Petunia inforata S proteins is essential for rejection of self-pollen. Plant Cell 6, 1021–1028 (1994).

18. Li, Y. et al. A mutation near the active site of S-RNase causes self-incompatibility in S-RNase-based self-incompatible plants. Plant Mol. Biol. 103, 129–139 (2020).

19. Li, M. F., Li, X. F., Han Zh, H., Shu, H. R. & Li, T. Molecular analysis of two Chinese pear (Pyrus bretschneideri Rehd.) spontaneous self-compatible mutants, Yan Zhuang and Jin Zhu. Plant Mol. Biol. (Stuttg.). 11, 774–783 (2009).

20. McCubbin, A. G., Chung, Y. Y. & Kao, T. A Mutant S1 RNase of Petunia inforata lacking RNase activity has an allele-specific dominant negative effect on self-incompatibility interactions. Plant Cell 9, 85–95 (1997).

21. Yamashita, K. & Tanimoto, S. Studies on self-incompatibility of Hassaku (Citrus hassaku hort. ex Tanaka). J. Jpn Soc. Hort. Sci. 54, 178–183 (1985).

22. Honsho, C. et al. Reproductive characteristics for self-compatibility and seed-setting of the ‘Ryukyu-Shikoku’ pomelo. Acta. Hort. 1547–1551 (2009).

23. Ye, W. et al. Seedless mechanism of a new mandarin cultivar ‘Wuzihatangyu’ (Citrus reticulata Blanco). Plant Cell 177, 19–27 (2009).

24. Chai, L., Ge, X., Biswas, M. K., Xu, Q. & Deng, X. Self-stereility in the mutant ‘Zigu shatian’ pummelo (Citrus grandis Osbeck) is due to abnormal post-zygotic embryo development and not self-incompatibility. Plant Cell Tiss. Org. 104, 1–11 (2010).

25. Caruso, M. et al. Comparative transcriptome analysis of style canal cells identifies novel candidate genes implicated in the self-incompatibility response of Citrus Clementina. BMC Plant Biol. 12, 20 (2012).

26. Gambetta, G. et al. Self-incompatibility, parthenocarpy and reduction of seed presence in ‘Afonu’ mandarin. Sci. Hort. 164, 183–188 (2013).

27. Zheng, S. et al. Characterization of the ‘Yangshuli’ lemon transcriptome by de novo assembly to discover genes associated with self-incompatibility. Mol. Genet. Genomics 290, 365–375 (2015).

28. Kakade, V., Dubey, A. K., Sharma, R. M. & Malik, S. K. Gametophytic self-incompatibility causes seedlessness in ‘Kagzi Kalan’ lemon (Citrus limon). J. Hortic. Sci. Biotech. 92, 303–312 (2017).

29. Liang, M. et al. Evolution of self-compatibility by a mutant S4-RNase in citrus. Nat. Plants 6, 131–142 (2020).

30. Miao, H. X., Qin, Y. H., Teixeira da Silva, J. A., Ye, Z. X. & Hu, G. B. Cloning and molecular analysis of pollen-pistil interaction in clementine. Plant Cell Rep. 28, 1358–1367 (2011).

31. Liang, M. et al. Genome-wide identification and functional analysis of S-RNase involved in the self-incompatibility of citrus. Mol. Genet. Genomics 292, 325–341 (2016).

32. Ma, Y., Liu, Q., Hu, G. & Qin, Y. Comparative transcriptional survey between self-incompatibility and self-compatibility in Citrus reticulata Blanco. Gene 609, 52–61 (2017).

33. Lin, W. et al. Two genes (CS1 and CS2) box-involved the self-incompatibility of “Xiangshui” lemon (Citrus limon (L.) Burm. f.). Plant Mol. Biol. Rep. 37, 50–62 (2019).

34. Miao, H., Ye, Z., Hu, G. & Qin, Y. Comparative transcript profiling of gene expression between self-incompatible and self-compatible mandarins by suppression subtractive hybridization and cDNA microarray. Mol. Breed. 35, 47 (2015).

35. Distefano, G., Caruso, M., La Malfa, S., Gentile, A. & Tribulato, E. Histological and molecular analysis of pollen-pistil interaction in clementine. Plant Cell Rep. 28, 1439–1451 (2009).

36. Liang, M. et al. Genome-wide identification and functional analysis of S-RNase involved in the self-incompatibility of citrus. Mol. Genet Genomics 292, 325–341 (2017).

37. Brown, P. H. & Ho, T. H. barley aleurone layers secrete a nuclease in response to gibberellic acid purification and partial characterization of the associated ribonuclease, deoxyribonuclease, and 3’-nucleotidase activities. Plant Physiol. 82, 801–806 (1986).

38. Koren, S. et al. Carus scalable and accurate long-read assembly via adaptive k-mer weighting and repeat separation. Genome Res. 27, 722–736 (2017).

39. Solovyev, V., Kosarev, P., Seledsov, I. & Vorobyev, D. Automatic annotation of euakytic genes, pseudogenes and promoters. Genome Biol. 7, S10–11–S10 12 (2006).

40. Pruitt, K. D., Tatusova, T. A. & Maglott, D. R. NCBI Reference Sequence (RefSeq): a curated non-redundant sequence database of genomes, transcripts, and proteins. Nucleic Acids Res. 33, DS1–DS5 (2005).

41. Marcais, G. et al. MUMmer: a fast and versatile genome alignment system. PLoS Comput. Biol. 14, e1006944 (2018).

42. Wang, X. et al. Genomic analyses of primitive, wild, and cultivated citrus provide insights into asexual reproduction. Nat. Genet. 49, 765–772 (2017).

43. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data. Bioinformatics 31, 40–41 (2015).

44. Krueger, F. & Andrews, S. R. Bismark: a flexible aligner and methylation caller for Bisulft-Seq applications. Bioinformatics 27, 1571–1572 (2011).

45. Duan, Z. et al. The Brassica napus GATA transcription factor BnAS2.MLM1 is a stigma compatibility factor. J. Integr. Plant Biol. 62, 1113–1121 (2020).

46. Huang, S. X. et al. Competitive interaction between two functional S-haplotypes confer self-compatibility on tetraploid Chinese cherry (Prunus pseudocerasus Lindl. C.V. Nanjing Chushi). Plant Cell Rep. 27, 1075–1085 (2008).

47. Okada, K. et al. Deletion of a 236 kb region around SmR1-RNase of Japanese pear. Plant Mol. Biol. 66, 389–400 (2008).