Screening and functional analysis of potential *S* genes in *Chrysanthemum morifolium*

Fan Wang[1,2], Sujuan Xu[1,2], Ze Wu[1,2], Xinghua Zhong[1,2], Weimin Fang[1], Fadi Chen[1], and Nianjun Teng[1,2*]

[1] *College of Horticulture, Nanjing Agricultural University, Key Laboratory of Landscape Design, Ministry of Agriculture and Rural Affairs, Nanjing 210095, China*
[2] *Nanjing Agricultural University-Nanjing Oriole Island Modern Agricultural Development Co., Ltd. Jiangsu Graduate Workstation, Nanjing 210043, China*
These authors contributed equally: Fan Wang, Sujuan Xu
* Corresponding author, E-mail: njteng@njau.edu.cn

Abstract

S genes are the key genes that cause plant self-incompatibility, to find out the key *S* genes and understand molecular mechanism of self-incompatibility in chrysanthemum, the stigmas and anthers at different developmental stages of 'Q10-22-2'—a self-incompatible chrysanthemum cultivar, were used for RNA sequencing. After bioinformatics analysis, 13 candidate pistil *S* genes and five candidate pollen *S* genes were excavated. A potential pistil *S* gene was cloned and named as *CmSRK1*. Meanwhile, a potential pollen *S* gene was cloned and named as *CmPCP1*. qRT-PCR revealed that *CmSRK1* was specifically expressed in mature stigmas, and *CmPCP1* was specifically expressed in anthers 3 d before maturation. Subcellular localization showed that both CmSRK1 and CmPCP1 were located in the nucleus and the cell membrane. Transcriptional activation activity analysis indicated that both of the two proteins had no transcriptional activation activity. Yeast two hybrid assay showed that there was no interaction between CmSRK1 and CmPCP1. *CmSRK1* was constructed on the expression vector containing stigma-specific promoter, and *CmPCP1* was constructed on the expression vector containing pollen-specific promoter, they were then transformed into *Arabidopsis thaliana*. Artificial hybridization was performed with transgenic lines containing *CmSRK1* as the female parents, and transgenic lines containing *CmPCP1* as the male parents. The hybridization results showed that seed sets of two transgenic lines were 19.62% and 11.64%, respectively, while cross-pollinated seed sets of Col-0 was 84.43%. Therefore, it was speculated that *CmSRK1* and *CmPCP1* might be pistil and pollen *S* genes of chrysanthemum, respectively, and SI of chrysanthemum belonged to SSI.

INTRODUCTION

Self-incompatibility (SI) is a complicated system which enables plants to avoid inbreeding by self-pollination and promote hybridization[1]. SI is also an ideal model for studying signal recognition and transduction, intercellular interaction, and gene spatiotemporal expression. Investigation of SI has important theoretical and practical significances in plant reproductive biology, cross breeding and utilization of heterosis[2]. Therefore, SI research has long been an important area of plant reproductive and developmental biology, and significant progress has been made in recent years.

The majority of SI systems are regulated by the *S*-locus[3,4]. *S*-locus includes at least two closely related polymorphic genes, one determines pollen specificity, and the other determines pistil specificity. Futhermore, many other genes associated or not associated with *S*-locus also play key roles in SI[5,6]. SI systems are mainly divided into gametophytic self-incompatibility (GSI) and sporophytic self-incompatibility (SSI). In GSI, there are two main systems. One is S-RNase-based GSI in Solanaceae[7,8], Rosaceae[9,10] and Plantagina-ceae[11,12], the pistil *S* gene is *S-RNase*, the pollen *S* gene is *SFB/SLF*. The other is Ca[2+]-dependent GSI in Papaveraceae[13,14], the pistil *S* gene is *PrsS*, and the pollen *S* gene is *PrpS*. However, as far as is known, it is unclear how many different

SSI systems exist in plants. Among them, SSI in Brassicaceae is the most studied, where the pistil *S* gene is *SRK*, and the pollen *S* gene is *SCR/SP11*[15,16]. The interaction between SRK and SCR contributes to the failure of pollen grain germination on the stigmas, which leads to SI in plants.

In *Brassica*, SRK is a plant receptor kinase with approximately 857 amino acids and a high degree of polymorphism. It mainly consists of three domains, containing an extracellular S domain, a transmembrane domain, and an intracellular domain with Ser/Thr kinase activity. Extracellular S domain is the binding site of SCR and contains 12 Cys residues and three hypervariable regions[17]. The presence of hypervariable regions is responsible for the polymorphism of SRK. If SRK domains are further divided, they can be categorised into the following domains based on the N-terminal to C-terminal[18]: LLD1 (Lectin-like 1) domain, DR (Delectable Region) domain, LLD2 (Lectin-like 2) domain, EGF-like domain, PAN-APPLE domain, TM domain, JM domain, Ser/Thr kinase domain and C-terminal. LLD1 domain affects the activation of SRK[19,20]. DR domain is a linker sequence with variable lengths[18]. LLD2 and EGF-like domains can bind to SCR[21], they contain three hypervariable regions which determine haplotype specificity of SRK[22,23]. PAN-APPLE domain determines homodimeri-zation and heterodimerization of SRK[18]. The functions of TM

and JM domains remain unclear, but SRK lacking JM domain will lose kinase activity[24]. Kinase domain transmits extracellular signals through catalyzing phosphorylation of Ser/Thr residues[25,26]. *SRK* is specifically expressed in stigma papilla cells, and its protein is mainly located in the cell membrane of stigma papilla cells. The distinct expression of *SRK* occurs at the early stage of flower bud development, and the expression level increases gradually as the stigma grows, the expression peak is reached on the day of flowering[27]. SCR/SP11 is a small alkaline hydrophilic protein in the defensin-like PCP (pollen coat protein) family with about 74–77 amino acids. It is highly polymorphic and has a higher polymorphism than that of SRK[28,29]. *SCR/SP11* is specifically expressed only in the anther tapetum and pollen[30–32].

Chrysanthemum is a representative species in Asteraceae, and many cultivars are self-incompatible, which provides abundant materials for the study of SI. However, there are few reports on SI in Asteraceae. The present study was intended to search and identify candidate *S* genes in chrysanthemum and pursue a better understanding of the SI mechanism in chrysanthemum. As the expression of *S* genes is tissue-specific and developmentally regulated, the stigmas and anthers at different developmental stages of 'Q10-22-2'—a self-incompatible chrysanthemum cultivar, were sampled for RNA sequencing. Plants in the Asteraceae are considered to be sporophytic SI. Thus, genes encoding proteins homologous to SRK and belonging to PCP family were selected and cloned from 'Q10-22-2', respectively, and the potential functions were verified through transformation into *A. thaliana*.

RESULTS

RNA sequencing and read assembly

To search for differentially expressed genes between mature stigmas (MS) and immature stigmas (IS) of 'Q10-22-2', two cDNA libraries constructed from them were subjected to RNA-Seq. After filtering out any low-quality reads, each sample had about 44 Mb clean reads. Then, high-quality reads were clustered, 98,595 unigenes were finally obtained, the total length of them was 78,399,181 bp, the average length was 795 bp, and the percentage of GC was 39.96% (Supplemental Table 1).

To investigate differential gene expression between mature anthes (MA) and immature anthes (IA) of 'Q10-22-2', two cDNA libraries constructed from them were subjected to RNA-seq. After filtering out any low-quality reads, each sample had about 44 Mb clean reads. High-quality reads were clustered, 100,512 unigenes were finally obtained, the total length of them was 77,001,891 bp, the average length was 766 bp, and the percentage of GC was 40.08% (Supplemental Table 2).

Unigene function annotation

Function annotation was performed on all assembled unigenes using seven function databases to predict their functions in stigma. As a result, 55,084 unigenes were annotated in the seven databases (Supplemental Table 3). Firstly, 17,851 unigenes were classified and annotated by COG database. In 25 COG sorts, the most was 'general function prediction only' (5924, 33.19%), next was

'transcription' (3235, 18.12%), and the least was 'nuclear structure' (6, 0.034%) (Fig. 1a). Then, GO annotation was used to classify unigene function. Consequently, 21,300 unigenes were divided into 54 function groups, which belonged to three main categories: biological process, cellular component, and molecular function. In the biological process category, the main function groups were 'metabolic process' and 'cellular process'. In the cellular component category, 'cell' and 'cell part' were the most abundant groups. In the molecular function category, 'catalytic activity' and 'binding' were remarkable (Fig. 1b). At last, KEGG annotation was used to authenticate the biological pathways activated in 'Q10-22-2's stigmas. Altogether, 37,674 unigenes were clustered into 135 pathways. Among them, the three main pathways were 'metabolic pathways [ko01100]' (7933, 21.06%), 'biosynthesis of secondary metabolites [ko01110]' (4507, 11.96%) and 'plant-pathogen interaction [ko04626]' (1505, 3.99%).

To predict possible functions of unigenes in 'Q10-22-2's anthers, function annotation was performed on all assembled unigenes using the seven function databases mentioned above. As a result, 56,879 unigenes were annotated in the databases (Supplemental Table 4). Firstly, 18,889 unigenes were classfied and annotated by COG database. In 25 COG sorts, the most was 'general function prediction only' (6251, 33.09%), next was 'transcription' (3399, 17.99%), and the least was 'nuclear structure' (6, 0.032%) (Fig. 1c). Then, GO annotation was used to classify unigene functions. Consequently, 22,095 unigenes were divided into 55 function groups, which belonged to the three main categories mentioned above. In the biological process category, the main function groups were 'metabolic process' and 'cellular process'. In the cellular component category, 'cell' and 'cell part' were the most abundant groups. And in the molecular function category, 'catalytic activity' and 'binding' were remarkable (Fig. 1d). At last, KEGG annotation was used to authenticate the biological pathways activated in 'Q10-22-2's anthers. Altogether, 38,994 unigenes were clustered into 135 pathways. Among them, the three main pathways were 'metabolic pathways [ko01100]' (8234, 21.12%), 'biosynthesis of secondary metabolites [ko01110]' (4635, 11.89%) and 'plant-pathogen interaction [ko04626]' (1614, 4.14%).

Candidate S genes in stigmas and anthers of 'Q10-22-2'

Based on FPKM values, unigene expression levels in MS and IS were studied. Compared with IS, expression levels of 835 unigenes were up-regulated, and expression levels of 614 unigenes were down-regulated in MS. To confirm possible functions of these DEGs, GO and KEGG analyses were conducted. Consequently, 1,099 DEGs were annotated to GO functional groups, and 931 DEGs were annotated to 117 KEGG pathways. Based on FPKM values, unigene expression levels in MA and IA were studied. Compared with IA, expression levels of 4,105 unigenes were up-regulated and expression levels of 5,430 unigenes were down-regulated in MA. To confirm possible functions of these DEGs, GO and KEGG analyses were conducted. Consequently, 7,834 DEGs were annotated to GO functional groups, and 6,211 DEGs were annotated to 134 KEGG pathways. After in-depth analysis and screening of the above unigenes and DEGs, 13 candidate pistil *S* genes (Table 1) and five candidate pollen *S*
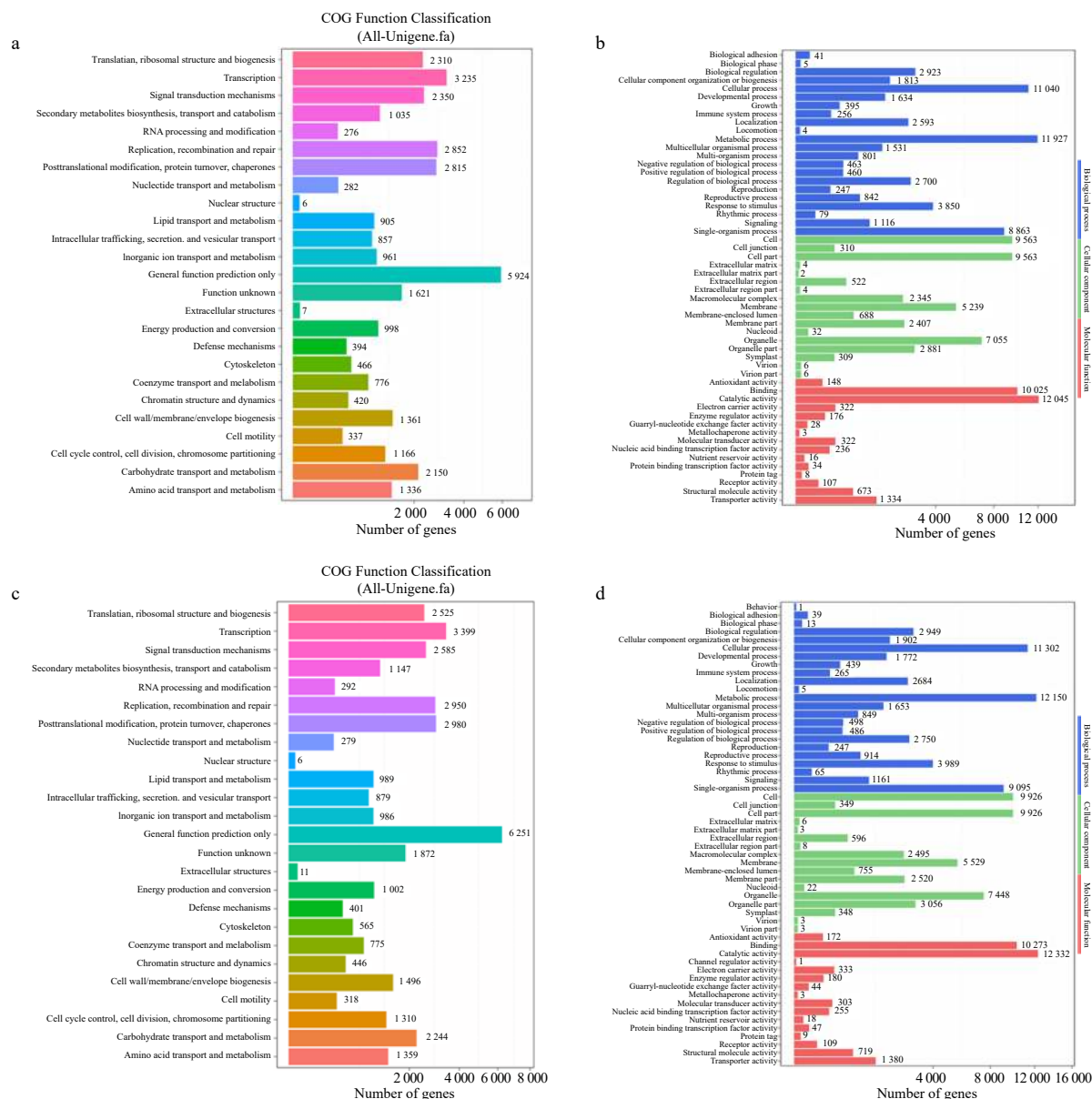
Fig. 1 (a) COG function annotation of stigma transcriptome. (b) GO function annotation of stigma transcriptome. (c) COG function annotation of anther transcriptome. (d) GO function annotation of anther transcriptome.

Table 1. Candidate pistil *S* genes in 'Q10-22-2's stigmas.

Unigene ID	MS-FPKM	IS-FPKM	Annotation
CL7408.Contig2	72.1	35.61	*S*-receptor-like serine/threonine-protein kinase RLK1
CL5423.Contig1	6.76	4.03	*S*-receptor-like serine/threonine-protein kinase At1g11410
Unigene50549	5.71	0	*S*-receptor-like serine/threonine-protein kinase SD1-1
CL6070.Contig3	5.43	4.42	*S*-receptor-like serine/threonine-protein kinase At4g27290
CL8545.Contig1	5.38	1.58	*S*-receptor-like serine/threonine-protein kinase SD1-1
CL9433.Contig3	4.57	1.73	*S*-receptor-like serine/threonine-protein kinase At2g19130
Unigene17909	3.42	1.6	*S*-receptor-like serine/threonine-protein kinase At1g34300
Unigene10651	3.37	1.36	*S*-receptor-like serine/threonine-protein kinase SD3-1
CL1678.Contig13	3.21	1.1	*S*-receptor-like serine/threonine-protein kinase B120
CL1678.Contig3	3.18	1.49	*S*-receptor-like serine/threonine-protein kinase At1g67520
Unigene54285	2.97	0	*S*-receptor-like serine/threonine-protein kinase At4g27290
CL6328.Contig1	2.9	1.25	*S*-receptor-like serine/threonine-protein kinase RLK1
Unigene5472	2.46	0.82	*S*-receptor-like serine/threonine-protein kinase At5g24080

genes (Table 2) were selected. All the candidate *S* genes were in the same family as the *S* genes in Brassicaceae plants, which belonged to SSI.

qRT-PCR validation of transcriptome data

To detect the reliability of RNA-Seq data, ten unigenes were selected from the stigma library for qRT-PCR. The results showed that expression trends of all unigenes were consistent with the sequencing results (Fig. 2a). Most selected unigenes were associated with pistil *S* genes of other plants,

such as the genes encoding *S*-receptor serine /threonine protein kinase, ribonuclease T2 family protein and epidermal-specific secretory glycoprotein.

Similarly, ten unigenes were selected from the anther library for qRT-PCR. The results showed that expression trends of all unigenes were also consistent with the sequencing results (Fig. 2b). Selected unigenes were either associated with pollen germination and pollen development, or associated with pollen *S* genes in other plant families, such as genes encoding PCP.

Table 2. Candidate pollen *S* genes in 'Q10-22-2's anthers.

Unigene ID	MA-FPKM	IA-FPKM	Annotation
CL12121.Contig1	0.33	262.6	*S* locus-related glycoprotein 1 binding pollen coat protein
CL12224.Contig1	29.61	589.34	pollen coat-like protein
CL12224.Contig2	12.22	156.65	pollen coat-like protein
CL12493.Contig1	2.06	621.99	*S* locus-related glycoprotein 1 binding pollen coat protein
CL12896.Contig1	113.15	816.33	pollen coat-like protein
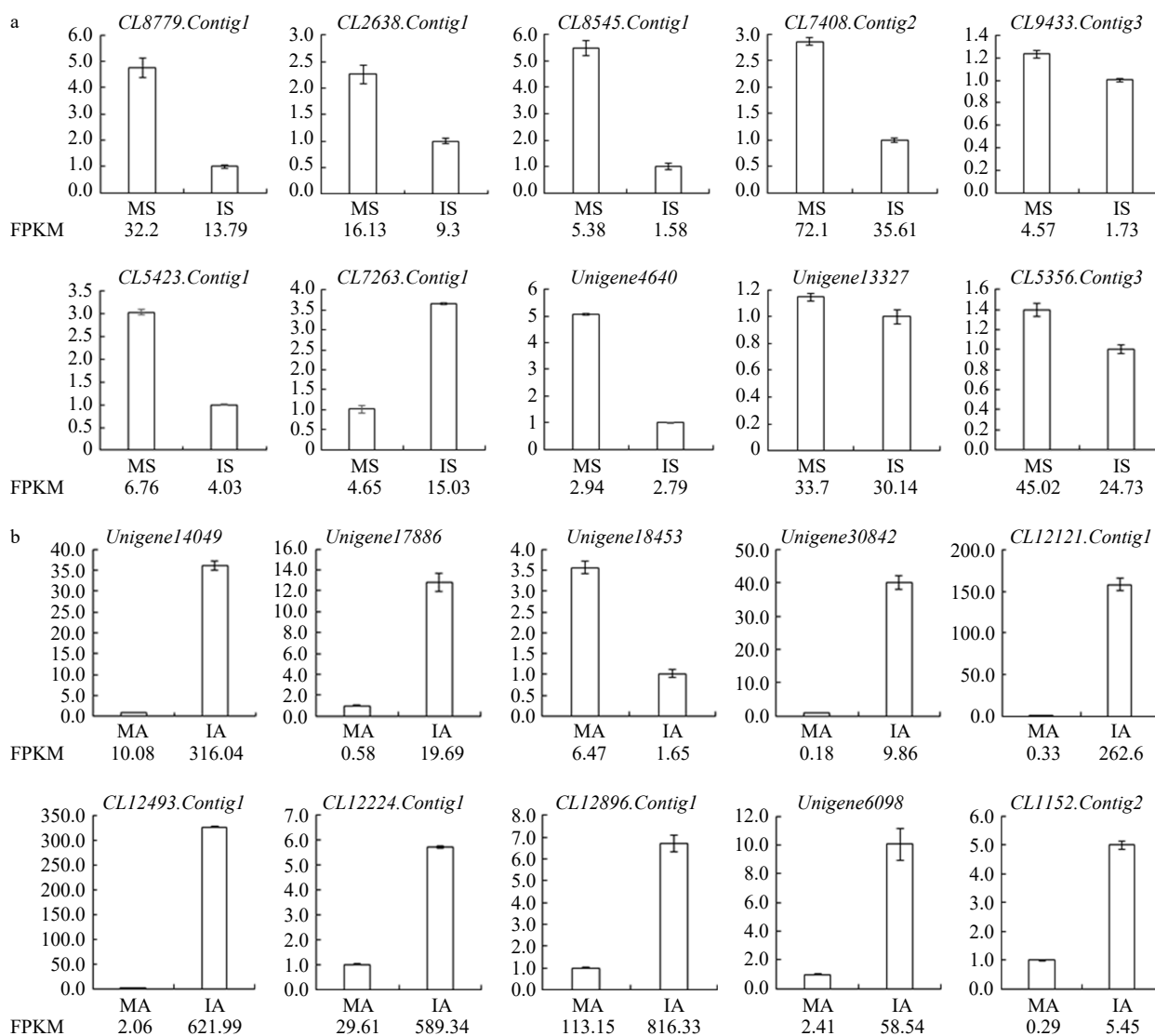
Fig. 2 (a) qRT-PCR verification of RNA-Seq results of stigma transcriptome. (b) qRT-PCR verification of RNA-Seq results of anther transcriptome.

## Sequence analysis of CmSRK1 and CmPCP1

A potential pistil *S* gene, *CmSRK1*, was cloned according to the sequence of CL5423.Contig1, which was annotated as SRK in 'Q10-22-2' stigma library. The ORF length of *CmSRK1* was 2,568 bp, and the gene was predicted to encode 855 amino acid residues, 96.24 kDa peptide with a pI of 5.36. When the amino acid sequence was aligned on NCBI (National Center of Biotechnology Information), it was found that CmSRK1 contained classic domains of SRK, such as Lectin, *S*-locus glycoprotein, PAN-APPLE and protein kinase domain (Fig. 3a), this indicated that CmSRK1 was a homologous protein of SRK. Meanwhile, a potential pollen *S* gene, *CmPCP1*, was cloned according to the sequence of CL12224.Contig1, which was annotated as PCP in 'Q10-22-2' anther library. The ORF length of *CmPCP1* was 201 bp, and the gene was predicted to encode 66 amino acid residues, 6.82 kDa peptide with a pI of 4.88. The amino acid sequence was aligned on NCBI and it was found that CmPCP1 did not contain any conserved domains.

Tissue expression characteristics of CmSRK1 and CmPCP1

The expression of *CmSRK1* was tissue-specific, it was specifically expressed in stigmas, especially in mature stigmas. However, its expression levels in anthers, flower buds and leaves were very low (Fig. 3b). Similarly, the expression of *CmPCP1* was also tissue-specific, it was specifically expressed in anthers, especially in anthers 3 d before maturation. But its expression in stigmas, flower buds and leaves were very low (Fig. 3b).

Subcellular localization and transactivation activity of CmSRK1 and CmPCP1

In transiently transformed onion epidermal cells, GFP signals of pMDC43 empty vector distributed throughout the cells, but GFP signals induced by the *35S::GFP-CmSRK1* and *35S::GFP-CmPCP1* were located in the nucleus and the cell membrane (Fig. 3c), indicating that both CmSRK1 and CmPCP1 proteins localized to the nucleus and the cell membrane. With the transactivation activity assay, neither
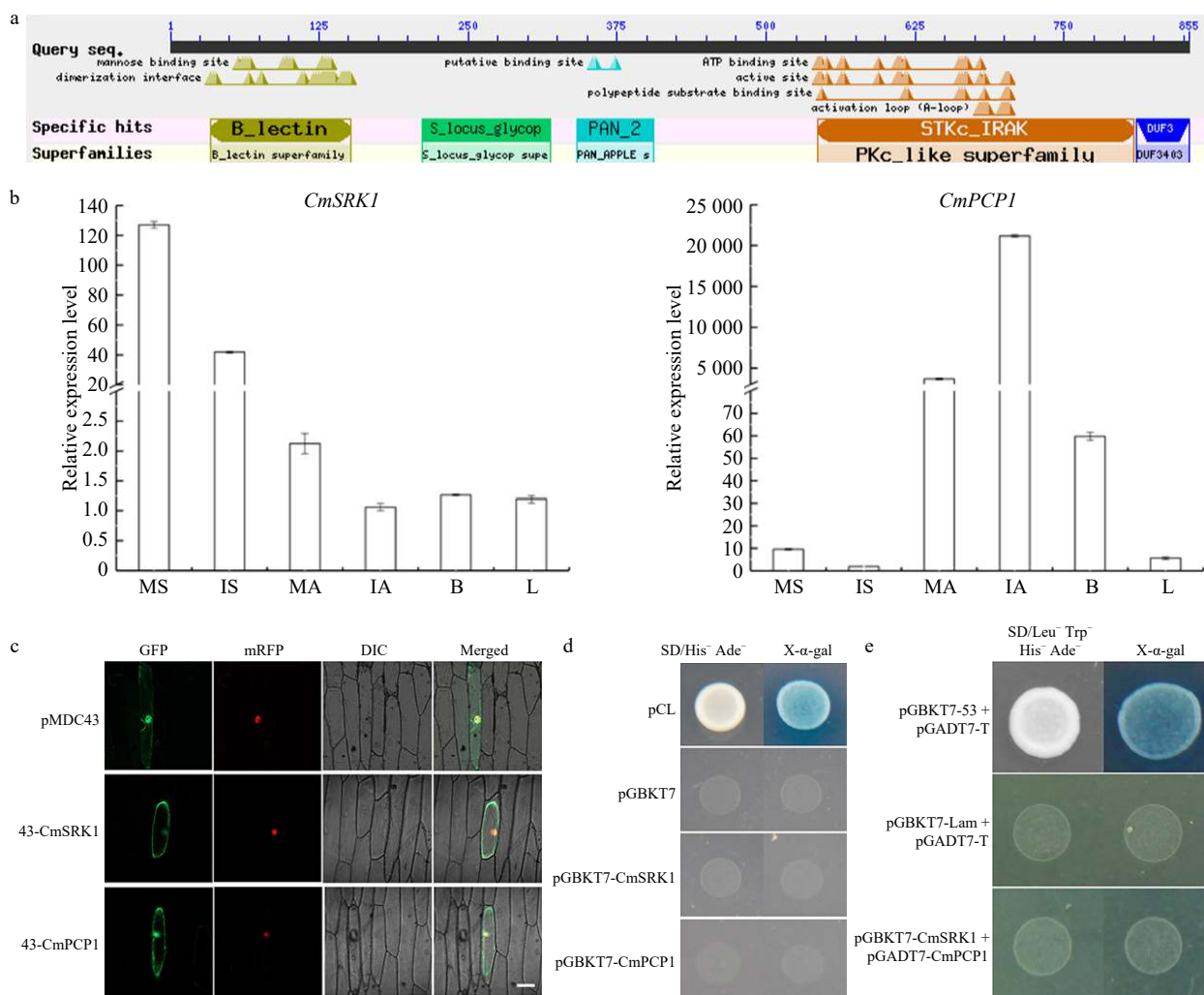
Fig. 3 (a) Blast result of CmSRK1. (b) Expression profiles of *CmSRK1* and *CmPCP1* in different tissues of 'Q10-22-2'. MS: mature stigmas; IS: immature stigmas; MA: mature anthers; IA: immature anthers; B: flower buds; L: leaves. (c) Subcellular localization of CmSRK1 and CmPCP1. Bar=100 μm. (d) Transcriptional activation activity of CmSRK1 and CmPCP1. (e) Yeast two hybrid between CmSRK1 and CmPCP1.

pGBKT7-*CmSRK1* nor pGBKT7-*CmPCP1* raised well on SD/His⁻ Ade⁻ medium with X-α-gal and lacking X-α-gal. Negative control pGBKT7 was also unable to grow on the two media, while the positive control pCL1 was able to grow on the two media (Fig. 3d). The above results indicate that both CmSRK1 and CmPCP1 had no transactivation activity.

Yeast interaction verification between CmSRK1 and CmPCP1

To verify the interaction between CmSRK1 and CmPCP1, yeast cells co-expressing pGADT7-*CmPCP1* and pGBKT7-*CmSRK1* were unable to grow on the SD-Leu/-Ade/-His/-Trp/ screening medium with X-α-gal and lacking X-α-gal. Negative control pGBKT7-*lam* + pGADT7-*T* was also unable to grow on the two media, while the positive control pGBKT7-*53* + pGADT7-*T* was able to grow on the two media (Fig. 3e). These results showed that there was no interaction between CmSRK1 and CmPCP1.

Positive seedling identification of transgenic Arabidopsis

To verify whether the stigma-specific promoter *SLR1* could initiate the expression of *CmSRK1* in Arabidopsis stigmas, stigmas of Col-0 and T₁ generation were observed under a confocal laser scanning microscope. The results showed that it was difficult to determine whether there were GFP signals in the stigmas of the transgenic Arabidopsis due to strong autofluorescence of the stigmas (Fig. 4a).

Similarly, to verify whether the pollen-specific promoter *LAT52* could initiate the expression of *CmPCP1* in Arabidopsis pollens, pollen grains of Col-0 and T₁ generation were also observed under a confocal laser scanning microscope. It was found that there was no GFP signal in pollen grains of Col-0, while GFP signals were observed in pollen grains of transgenic Arabidopsis (Fig. 4b), indicating that *LAT52* could initiate the expression of *CmPCP1* in Arabidopsis pollens.

The T₃ generation homozygous transgenic lines were

sampled, the inflorescence RNA was extracted, and semi-quantitative analysis was carried out. It was found that all transgenic lines had targeted bands (Fig. 4c–d), indicating that the promoters initiated the expression of genes. According to the results of the expression assay, two independent transgenic lines of each gene were selected for the following artificial hybridization.

Hybridization fertility characteristics of transgenic Arabidopsis

Before artificial hybridization of *A. thaliana*, floral organs of Col-0 and transgenic lines were observed. It was found that both Col-0 and transgenic Arabidopsis had tetradynamous stamens, the pistils and stamens grew normally (Fig. 5), which could ensure normal occurrence of pollination.

As shown in Table 3, the self-pollinated seed set of Col-0 was 93.58 ± 2.17%, while the cross-pollinated seed set of Col-0 was slightly lower, which was 84.43 ± 5.01%. The self-pollinated seed set of the female parent *CmSRK1* #1 and the male parent *CmPCP1* #1 were 84.89 ± 3.85% and 94.87 ± 2.25%, respectively, but their cross-pollinated seed set was significantly reduced to 19.62 ± 6.70%, and some fruit pods even had no seed (Fig. 6h). Likewise, the self-pollinated seed set of the female parent *CmSRK1* #2 and the male parent *CmPCP1* #2 were 78.26 ± 7.12% and 50.51 ± 6.60%, respectively, but their cross-pollinated seed set was significantly reduced to 11.64 ± 3.68%, and some fruit pods also had no seed (Fig. 6l). These results indicated that when artificial hybridization was conducted with transgenic lines containing *CmSRK1* as the female parents, and transgenic lines containing *CmPCP1* as the male parents, seed sets were significantly reduced, and some even had no seeds. Thus, it was possible that there was an interaction between CmSRK1 and CmPCP1, resulting in low cross-pollinated seed sets. *CmSRK1* and *CmPCP1* were likely to be pistil and pollen *S* genes, respectively, which determined the SI of chrysanthemum.
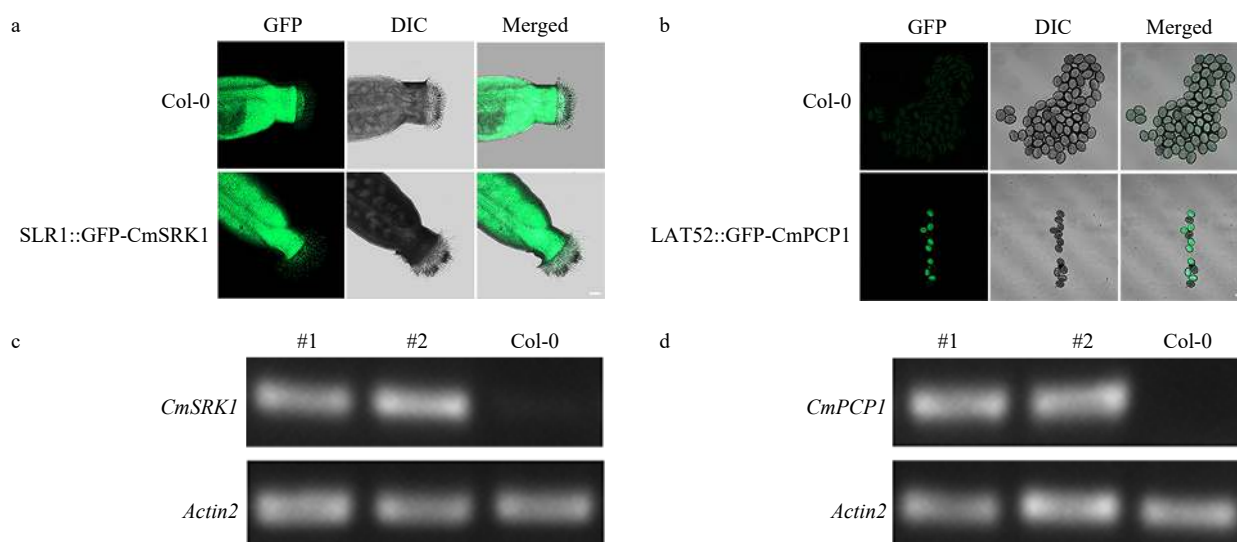
Fig. 4 a,b GFP signals in stigmas and pollens of Arabidopsis. (a) GFP signal in stigmas of Arabidopsis; (b) GFP signals in pollens of Arabidopsis. Bar = 20 μm. (c) RT-PCR identification of *CmSRK1* in Arabidopsis; (d) RT-PCR identification of *CmPCP1* in Arabidopsis. #1 and #2: two transgenic lines.
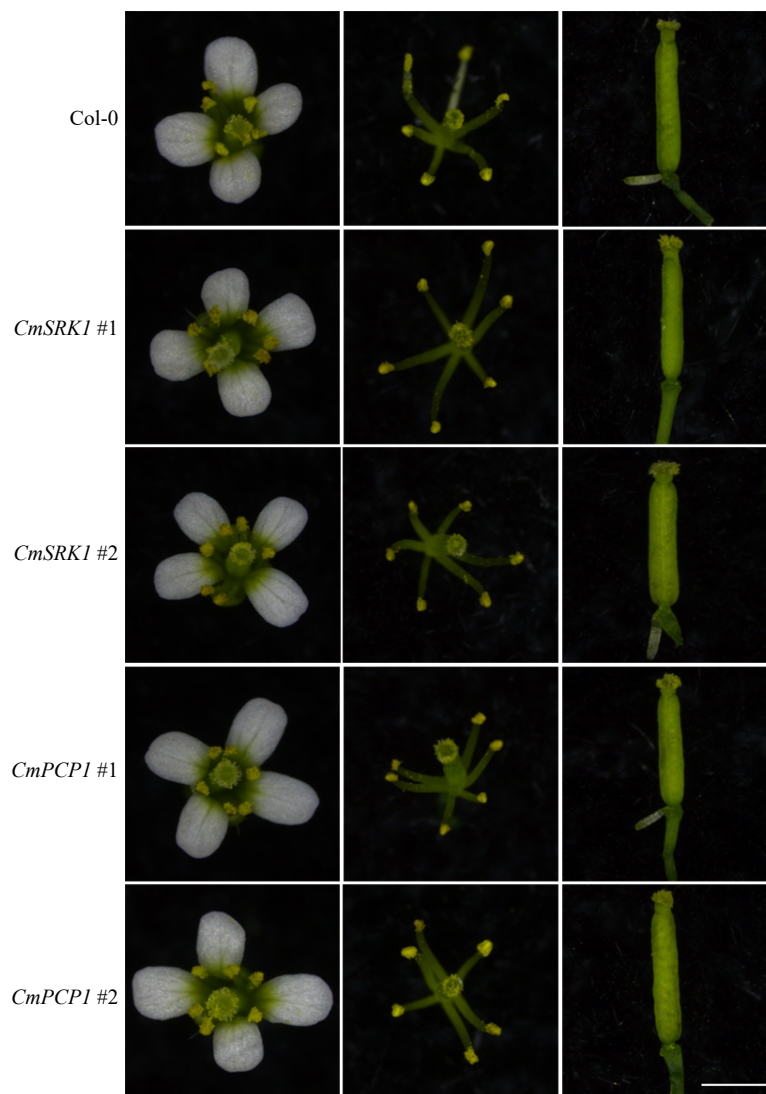
Fig. 5 Floral organ observation of Arabidopsis after transformation of *CmSRK1* and *CmPCP1*. The first column was the whole flowers, the second was the pistils and stamens, and the third column was the pistils. The arabidopsis used for observation was 30 d after planting. Bar = 1 mm.

Table 3. Self-pollinated and cross-pollinated seed sets of *Arabidopsis*.

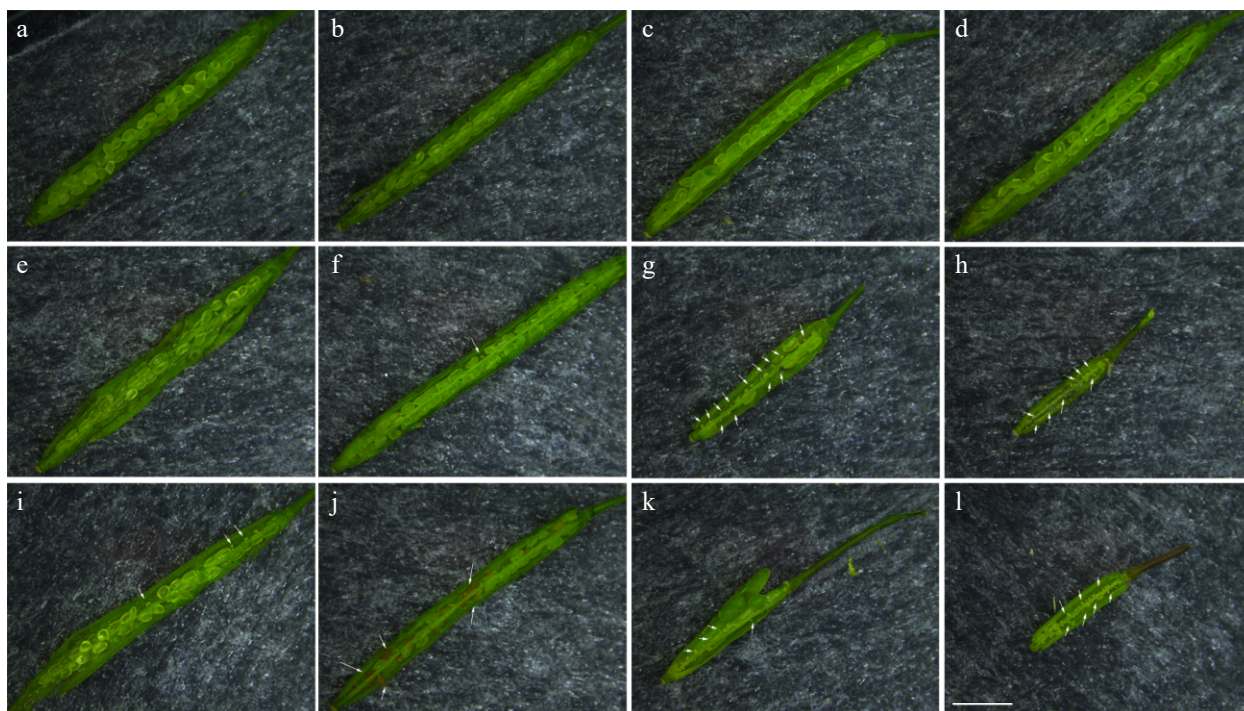| Self-pollination/cross-pollination	Seed set (%)
Col-0	93.58 ± 2.17[a]
Col-0 × Col-0	84.43 ± 5.01[a]
CmSRK1 #1	84.89 ± 3.85[a]
CmPCP1 #1	94.87 ± 2.25[a]
CmSRK1 #1 × *CmPCP1* #1	19.62 ± 6.70[c]
CmSRK1 #2	78.26 ± 7.12[a]
CmPCP1 #2	50.51 ± 6.60[b]
CmSRK1 #2 × *CmPCP1* #2	11.64 ± 3.68[c]

Values given were mean ± standard error. Values with different superscript indicated significant differences at $p \leq 0.05$ according to Tukey's test.

DISCUSSION

To date, SSI has been found mainly in Brassicaceae, Convolvulaceae and Asteraceae[3], of which Brassicaceae has

been the most extensively researched. In Brassicaceae, the female *S* determinant is SRK, and the male *S* determinant is SCR/SP11—a kind of pollen coating protein (PCP)[33], such information provides a crucial clue for the analysis of chrysanthemum SI. During the analysis of stigma transcriptome data, we found that the expression levels of some unigenes encoding SRK were up-regulated in MS, such as CL7408.Contig2 (MS 72.1, IS 35.61), Unigene17909 (MS 3.42, IS 1.6), and CL1678.Contig13 (MS 3.21, IS 1.1). The analysis of anther transcriptome data found that the expression levels of some unigenes encoding PCP were down-regulated in MA, such as CL12121.Contig1 (MA 0.33, IA 262.6), CL12224.Contig1 (MA 29.61, IA 589.34), and CL12493.Contig1 (MA 2.06, IA 621.99). Expression levels of these unigenes in MA were lower than those in IA, a possible reason might be that pollen grains had been fully mature in MA, the content of PCP had reached the highest, and the *PCP* did not need to be expressed. Therefore, we proposed that SSI system of

Fig. 6 Anatomical observation of self-pollinated and cross-pollinated fruit pods of Arabidopsis. (a, b) Col-0⊗; (c, d) Col-0 × Col-0; (e) *CmSRK1* #1⊗; (f) *CmPCP1* #1⊗; (g, h) *CmSRK1* #1 × *CmPCP1* #1; (i) *CmSRK1* #2⊗; (j) *CmPCP1* #2⊗; (k, l) *CmSRK1* #2 × *CmPCP1* #2. The white arrows identify the defected position in transgenic Arabidopsis. Bar = 2 mm.

Brassicaceae also played a role in chrysanthemum.

In the stigma transcriptome, some unigenes involved in pollen recognition were discovered, such as CL8785.Contig1 and Unigene19117. Some unigenes encoding definite proteins that had been reported in other plants' stigma transcriptome were also found[34,35]. For example, CL8779. Contig1 and CL9665.Contig3 which encode the stigma-specific peroxidase, Unigene41728 and CL1151.Contig2 encode the pistil-specific extensin-like protein. In reference to previous research[36,37], the analysis of anther transcriptome identified some unigenes involved in pollen germination, pollen exine formation, and pollen tube reception and growth. For instance, Unigene5436 and Unigene10201 participated in pollen germination, Unigene15286 and CL12901.Contig1 were implicated in pollen exine formation, CL4167.Contig1 and Unigene22620 took part in pollen tube reception, Unigene1360 and CL10302.Contig1 were involved in pollen tube growth. Taken together, this may suggest that these unigenes might play a vital part in chrysanthemum SI.

CmSRK1 was the gene annotated as '*S*-receptor-like serine/threonine-protein kinase' in the stigma transcriptome. Because CmSRK1 had classic domains of SRK, CmSRK1 and SRK in Brassicaceae were considered to be homologous proteins. Tissue quantification revealed that *CmSRK1* was specifically expressed in MS, which was consistent with the expression characteristic of stigma *S* genes. *CmPCP1* was the gene annotated as 'pollen coat-like protein' in the anther transcriptome. CmPCP1 did not have any conserved domain, and SCR in Brassicaceae also did not have any conserved domain with high sequence polymorphism. Tissue quantification revealed that *CmPCP1* was specifically expressed in IA, which was consistent with the expression

characteristic of pollen *S* genes.

The interaction between SRK and SCR in the same haplotype causes SI in Brassicaceae. When yeast two hybrid assay was performed between SCR_{B3} and $eSRK_{B3}$ (extracellular domain of SRK_{B3}) carrying complete signal peptides from *Brassica oleracea* L. B3, no blue positive clone was found, this might be because the signal peptides carried by the two proteins affected the secretion of mature peptides[38]. A previous study proved that SCR and eSRK could recognize and interact with each other[31]. However, when the tobacco expression system and pull-down method were used for interaction detection, it was found that only the full-length SRK_8 and $mSRK_8$ (extracellular domain and transmembrane domain of SRK_8) could interact with SCR_8, while $eSRK_8$ could not interact with SCR_8[39]. These different results might be related to *S* haplotype specificity of the SCR-SRK complex, and might also be related to the spatial structures of SCR and SRK in different detection systems. In our study, we found CmSRK1 and CmPCP1 did not interact with each other in the pGBKT7-pGADT7 yeast system, and this result might also be affected by the signal peptides or the protein spatial structures.

Self-pollinated seed sets in Col-0 were 93.58 ± 2.17%, and the cross-pollinated seed set was slightly lower, which was 84.43 ± 5.01%, but their difference was small, indicating that artificial hybridization was successful. Self-pollinated seed sets of *CmSRK1* #1 and *CmPCP1* #1 were 84.89 ± 3.85% and 94.87 ± 2.25%, respectively, and their cross-pollinated seed set was 19.62 ± 6.70%, indicating that the stigmas of the female parent and the pollens of the male parent developed normally, and the low cross-pollinated seed set was probably due to protein interaction *in vivo*. Likewise, self-pollinated

seed sets of *CmSRK1* #2 and *CmPCP1* #2 were 78.26 ± 7.12% and 50.51 ± 6.60%, respectively, and their cross-pollinated seed set was 11.64 ± 3.68%, this result confirmed the above hypothesis. It was likely that CmSRK1 and CmPCP1 were pistil and pollen S proteins with the same haplotype, and there was an interaction between them, resulting in low cross-pollinated seed sets of transgenic lines. Furthermore, self-pollinated seed sets of *CmSRK1* #2 and *CmPCP1* #2 were lower than those of *CmSRK1* #1 and *CmPCP1* #1, respectively, and cross-pollinated seed sets of *CmSRK1* #2 and *CmPCP1* #2 were lower than that of *CmSRK1* #1 and *CmPCP1* #1, indicating that to some extent, developmental state of parents had an effect on cross-pollinated seed sets. Notably, although cross-pollinated seed sets of transgenic lines were very low, there were still seeds in the fruit pods. It was speculated that not only *CmSRK1* and *CmPCP1* played a role in chrysanthemum SI, but there were other genes controlling chrysanthemum SI.

MATERIALS AND METHODS

Plant materials and growing conditions

Spray cut chrysanthemum 'Q10-22-2' is a self-incompatible cultivar with great ornamental traits[40]. Uniform rooted cuttings of 'Q10-22-2' were planted into a 1:1 (v/v) mixture of soilrite and vermiculite and grown in a greenhouse with an 8 h light period under controlled conditions (day/night temperature 25℃/18℃) with a relative humidity of 70%.

The expression of *S* genes is tissue-specific and controlled by development. As a consequence, mature stigmas (MS) and stigmas 3 d before maturation (IS) were collected from 'Q10-22-2' as a pair of comparable samples. Meanwhile, mature anthers (MA) instead of pollen and anthers 3 d before maturation (IA) of 'Q10-22-2' were collected as another pair of comparable samples. The four samples were frozen in liquid nitrogen and stored at −80 ℃.

RNA extraction, sequencing (RNA-Seq) and de novo assembly

Total RNA was extracted as described previously[41]. RNA sequencing was performed on the BGISEQ-500 platform (BGI, Shenzhen, China) to yield 150 bp paired-end reads. Each cDNA library was established with a mixture of RNA from three biological replicates. Before downstream analysis, the raw reads were filtered, and clean reads were obtained[36]. Then, the clean reads were assembled with the Trinity program[42], after which Tgicl was used to cluster them according to their redundancy, finally, the unigenes were obtained[43].

Unigene function annotation and screening of DEGs

Unigene sequences were aligned using blastn to NT, and aligned by blastx to protein databases such as NR, Swiss-Prot, COG and KEGG with a cutoff e-value < 10^{-5}[44]. Moreover, we used Blast2 GO[45] and WEGO[46] software to obtain GO annotation and GO function classification of unigenes. The unigenes were annotated against the KEGG database.

To determine the expression levels of unigenes in different samples, we used the method of fragments per kb per million fragments[47], which could eliminate the influence of different gene lengths and sequencing levels. Next, referring to the method of Audic and Claverie[48], DEGs (differentially expressed genes) between two samples were identified. DEGs need to meet two criteria, namely FDR (false discovery rates) < 0.001 and an absolute value of $\log_2$ ratio ≥ 1. Finally, GO and KEGG enrichment analysis were performed on DEGs.

qRT-PCR validation

To identify the quality of RNA-Seq data, ten unigenes were chosen from each library for qRT-PCR validation. Gene-specific primers (Supplemental Table 5 and Supplemental Table 6) were designed using Primer Premier 5.0 software and the *Elongation Factor 1α* (*CmEF1α*) gene (KF305681) was used as a normalization control[49].

Beverage Plant Research

Based on the sequence of CL5423.Contig1 in stigma transcriptome, specific primers (CmSRK1-ORF-F and -R, Supplemental Table 7) were designed to amplify the ORF (open reading frame) of *CmSRK1*. Based on the sequence of CL12224.Contig1 in anther transcriptome, specific primers (CmPCP1-ORF-F and -R, Supplemental Table 7) were designed to amplify the ORF of *CmPCP1*. High-fidelity PCR was conducted with stigma and anther cDNA as templates, respectively, they were then sub-cloned into pMD19-T (TaKaRa, Tokyo, Japan) for sequencing.

qRT-PCR of CmSRK1 and CmPCP1

To determine the tissue expression characteristics of *CmSRK1* and *CmPCP1*, RNA was isolated from flower buds, leaves, stigmas and anthers at different developmental stages using RNAiso reagent. Quantitative primers (CmSRK1-Q-F, -R and CmPCP1-Q-F, -R, Supplemental Table 7) were designed on the basis of *CmSRK1* and *CmPCP1* sequences, and qRT-PCR was processed. The reference gene was *CmEF1α*.

Subcellular localization of CmSRK1 and CmPCP1

According to the sequences of *CmSRK1* and *CmPCP1*, primers containing restriction sites (CmSRK1-*Bam*H I-F, -*Not* I-R and CmPCP1-*Bam*H I-F, -*Not* I-R, Supplemental Table 7) were designed. High-fidelity PCR was conducted to amplify the ORF of *CmSRK1* and *CmPCP1* and the fragments were inserted into pMDC43 vector carrying *35S::GFP*, respectively[50]. A transient assay was then performed to determine the subcellular localization of CmSRK1 and CmPCP1, by transforming the construct into onion (*Allium cepa*) epidermal cells as previously described[51].

Transactivation activity assay of CmSRK1 and CmPCP1

LR Clonase™ II enzyme mix was used to insert *CmSRK1* and *CmPCP1* ORF fragments into yeast expression vector pGBKT7. The two constructs were transformed into yeast strain *Y2H*. In parallel, the yeast cells transformed with pCL1 and pGBKT7 separately served as positive and negative controls. Transactivation activity assay was performed by growing yeast cells on two media SD/His⁻Ade⁻ added X-α-gal and lacking X-α-gal. All operations were performed as per the manufacturer's protocol (Clontech, Mountain View, CA, USA).

Yeast two hybrid assay between CmSRK1 and CmPCP1

Using LR Clonase™ II enzyme mix, *CmPCP1* ORF fragment was inserted into pGADT7 vector. pGADT7-*CmPCP1* construct and pGBKT7-*CmSRK1* construct were co-transformed into

yeast strain *Y2H*. Then a yeast two-hybrid (Y2H) assay was performed as previously described[52].

Construction of expression vectors containing tissue-specific promoters

The DNA of Arabidopsis and tomato was extracted using the CTAB method. To clone the stigma-specific promoter *SLR1*[53], high-fidelity PCR was performed using Arabidopsis DNA as the template. At the same time, to clone the pollen-specific promoter *LAT52*[54], high-fidelity PCR was performed using tomato DNA as the template. The primer sequences required for promoters cloning (SLR1-F, -R and LAT52-F, -R) and containing restriction sites (SLR1-Pme I-F, -Kpn I-R and LAT52-Hind III-F, -Kpn I-R) were listed in Supplemental Table 8. To construct expression vectors containing both tissue-specific promoters and target genes, *2×35S* promoter in pMDC43-*CmSRK1* construct was replaced by *SLR1* promoter using restriction enzymes *Pme* I and *Kpn* I, and *2×35S* promoter in pMDC43-*CmPCP1* construct was replaced by *LAT52* promoter using restriction enzymes *Hind* III and *Kpn* I.

Arabidopsis transformation

The *SLR1::CmSRK1* and *LAT52::CmPCP1* constructs were separately transformed into *A. thaliana* Col-0 by the floral dip method mediated by *Agrobacterium tumefaciens EHA105*. To detect whether the promoters could initate the expression of genes in specific tissues, the stigmas and pollen grains of the T_1 generation were observed under a confocal laser scanning microscope for GFP signals, and the inflorescence RNA of the T_3 generation was extracted for RT-PCR. The primers for RT-PCR were CmSRK1-Q-F, -R and CmPCP1-Q-F, -R (Supplemental Table 7), and the reference gene was *Actin2*.

Hybridization of transgenic lines

When the T_3 generation was obtained, aritifical hybridization was performed with transgenic lines containing *CmSRK1* as the female parents, and transgenic lines containing *CmPCP1* as the male parents. In parallel, to exclude the effects of human manipulation, artificial hybridization between wild type Col-0 was performed as the control. After 7 d, the seed sets of hybridization were calculated. In addition, self seed sets of the female and male parents were also calculated to exclude the effects of their own factors. For each combination, seed sets of 15−20 fruit pods were counted.

Conflict of interest

The authors declare that they have no conflict of interest.

Supplementary Information accompanies this paper at (http://www.maxapress.com/article/doi/10.48130/OPR-2021-0006)

REFERENCES

1. Takayama S, Isogai A. 2005. Self-incompatibility in plants. *Annual Review of Plant Biology* 56:467−89
2. Iwano M, Takayama S. 2012. Self/non-self discrimination in angiosperm self-incompatibility. *Current Opinion in Plant Biology* 15:78−83
3. Yamamoto M, Tantikanjana T, Nishio T, Nasrallah ME, Nasrallah JB. 2014. Site-specific *N*-glycosylation of the *S*-locus receptor kinase and its role in the self-incompatibility response of the Brassicaceae. *The Plant Cell* 26:4749−62
4. Sun P, Li S, Lu D, Williams JS, Kao TH. 2015. Pollen *S*-locus F-box proteins of *Petunia* involved in S-RNase-based self-incompatibility are themselves subject to ubiquitin-mediated degradation. *The Plant Journal* 83:213−23
5. Hee-Jeong J, Uddin AN, Jong-In P, Kumar TS, Hye-Ran K, et al. 2014. Analysis of *S*-locus and expression of *S*-alleles of self-compatible rapid-cycling *Brassica oleracea* 'TO1000DH3'. *Molecular Biology Reports* 41:6441−48
6. Kitashiba H, Nasrallah JB. 2014. Self-incompatibility in Brassicaceae crops: lessons for interspecific incompatibility. *Breeding Science* 64:23−37
7. Goldraij A, Kondo K, Lee CB, Hancock CN, Sivaguru M, et al. 2006. Compartmentalization of S-RNase and HT-B degradation in self-incompatible *Nicotiana*. *Nature* 439:805−10
8. Williams JS, Wu L, Li S, Sun P, Kao TH. 2015. Insight into S-RNase-based self-incompatibility in *Petunia*: recent findings and future directions. *Frontiers in Plant Science* 6:41
9. Wang C, Xu G, Jiang XT, Chen G, Wu J, et al. 2009. S-RNase triggers mitochondrial alteration and DNA degradation in the incompatible pollen tube of *Pyrus pyrifolia in vitro*. *The Plant Journal* 57:220−29
10. Wu J, Qu H, Shang Z, Tao S, Xu G, et al. 2011. Reciprocal regulation of Ca^{2+}-activated outward K^+ channels of *Pyrus pyrifolia* pollen by heme and carbon monoxide. *New Phytologist* 189:1060−68
11. Qiao H, Wang F, Zhao L, Zhou J, Lai Z, et al. 2004. The F-box protein AhSLF-S_2 controls the pollen function of S-RNase-based self-incompatibility. *The Plant Cell* 16:2307−22
12. Yang Q, Zhang D, Li Q, Cheng Z, Xue Y. 2007. Heterochromatic and genetic features are consistent with recombination suppression of the self-incompatibility locus in *Antirrhinum*. *The Plant Journal* 51:140−51
13. Thomas SG, Franklin-Tong VE. 2004. Self-incompatibility triggers programmed cell death in *Papaver* pollen. *Nature* 429:305−9
14. Wheeler MJ, de Graaf BHJ, Hadjiosif N, Perry RM, Poulter NS, et al. 2009. Identification of the pollen self-incompatibility determinant in *Papaver rhoeas*. *Nature* 459:992−95
15. Tarutani Y, Shiba H, Iwano M, Kakizaki T, Suzuki G, et al. 2010. *Trans*-acting small RNA determines dominance relationships in *Brassica* self-incompatibility. *Nature* 466:983−86
16. Tsuchimatsu T, Suwabe K, Shimizu-Inatsugi R, Isokawa S, Pavlidis P, et al. 2010. Evolution of self-compatibility in *Arabidopsis* by a mutation in the male specificity gene. *Nature* 464:1342−46
17. Kusaba M, Nishio T, Satta Y, Hinata K, Ockendon D. 1997. Striking sequence similarity in inter- and intra-specific comparisons of class I *SLG* alleles from *Brassica oleracea* and *Brassica campestris*: implications for the evolution and recognition mechanism. *PNAS* 94:7673−78
18. Naithani S, Chookajorn T, Ripoll DR, Nasrallah JB. 2007. Structural modules for receptor dimerization in the *S*-locus receptor kinase extracellular domain. *PNAS* 104:12211−16

19. Cabrillac D, Cock JM, Dumas C, Gaude T. 2001. The *S*-locus receptor kinase is inhibited by thioredoxins and activated by pollen coat proteins. *Nature* 410:220−3

20. Ivanov R, Gaude T. 2009. Endocytosis and endosomal regulation of the *S*-receptor kinase during the self-incompatibility response in *Brassica oleracea*. *Plant Cell* 21:2107−17

21. Kemp BP, Doughty J. 2007. *S* cysteine-rich (SCR) binding domain analysis of the *Brassica* self-incompatibility *S*-locus receptor kinase. *New Phytologist* 175:619−29

22. Miege C, Ruffio-Châble V, Schierup MH, Cabrillac D, Dumas C, et al. 2001. Intrahaplotype polymorphism at the *Brassica S* locus. *Genetics* 159:811−22

23. Boggs NA, Dwyer KG, Nasrallah ME, Nasrallah JB. 2009. *In vivo* detection of residues required for ligand-selective activation of the *S*-locus receptor in *Arabidopsis*. *Current Biology* 19:786−91

24. Germain H, Houde J, Gray-Mitsumune M, Sawasaki T, Endo Y, et al. 2007. Characterization of ScORK28, a transmembrane functional protein receptor kinase predominantly expressed in ovaries from the wild potato species *Solanum chacoense*. *FEBS Letters* 581:5137−42

25. Goring DR, Rothstein SJ. 1992. The *S*-locus receptor kinase gene in a self-incompatible *Brassica napus* line encodes a functional serine/threonine kinase. *The Plant Cell* 4:1273−81

26. Giranton JL, Dumas C, Cock JM, Gaude T. 2000. The integral membrane *S*-locus receptor kinase of *Brassica* has serine/threonine kinase activity in a membranous environment and spontaneously forms oligomers *in planta*. *Proceedings of the National Academy of Sciences of the United States of America* 97:3759−64

27. Stein JC, Dixit R, Nasrallah ME, Nasrallah JB. 1996. SRK, the stigma-specific S locus receptor kinase of *Brassica*, is targeted to the plasma membrane in transgenic tobacco. *The Plant Cell* 8:429−45

28. Schopfer CR, Nasrallah ME, Nasrallah JB. 1999. The male determinant of self-incompatibility in *Brassica*. *Science* 286:1697−700

29. Suzuki G, Kai N, Hirose T, Fukui K, Nishio T, et al. 1999. Genomic organization of the *S* locus: identification and characterization of genes in *SLG/SRK* region of *S⁹* haplotype of *Brassica campestris* (syn. *rapa*). *Genetics* 153:391−400

30. Stephenson AG, Doughty J, Dixon S, Elleman C, Hiscock S, et al. 1997. The male determinant of self-incompatibility in *Brassica oleracea* is located in the pollen coating. *The Plant Journal* 12:1351−59

31. Kachroo A, Schopfer CR, Nasrallah ME, Nasrallah JB. 2001. Allele-specific receptor-ligand interactions in *Brassica* self-incompatibility. *Science* 293:1824−26

32. Shiba H, Takayama S, Iwano M, Shimosato H, Funato M, et al. 2001. A pollen coat protein, SP11/SCR, determines the pollen *S*-specificity in the self-incompatibility of *Brassica* species. *Plant Physiology* 125:2095−103

33. Tabah DA, Mcinnis SM, Hiscock SJ. 2004. Members of the *S*-receptor kinase multigene family in *Senecio squalidus* L. (Asteraceae), a species with sporophytic self-incompatibility. *Sexual Plant Reproduction* 17:131−40

34. Quiapim AC, Brito MS, Bernardes LAS, DaSilva I, Malavazi I, et al. 2009. Analysis of the *Nicotiana tabacum* stigma/style transcriptome reveals gene expression differences between wet and dry stigma species. *Plant Physiology* 149:1211−30

35. Allen AM, Lexer C, Hiscock SJ. 2010. Comparative analysis of pistil transcriptomes reveals conserved and novel genes expressed in dry, wet, and semidry stigmas. *Plant Physiology* 154:1347−60

36. Iaria D, Chiappetta A, Muzzalupo I. 2016. *De novo* transcriptome sequencing of *Olea europaea* L. to identify genes involved in the development of the pollen tube. *The Scientific World Journal* 2016:4305252

37. Li Z, Zhang P, Lv J, Cheng Y, Cui J, et al. 2016. Global dynamic transcriptome programming of rapeseed (*Brassica napus* L.) anther at different development stages. *PLoS One* 11:e0154039

38. Yang H, et al. 2011. Study on the interactions between the truncated fragments of SCR and eSRK from *Brassica oleracea* L. by a yeast two-hybrid system. *Scientia Agricultura Sinica* 44:1953−62

39. Shimosato H, Yokota N, Shiba H, Iwano M, Entani T, et al. 2007. Characterization of the SP11/SCR high-affinity binding site involved in self/nonself recognition in *Brassica* self-incompatibility. *The Plant Cell* 19:107−17

40. Wang F, Zhang F, Chen F, Fang W, Teng N. 2014. Identification of chrysanthemum (*Chrysanthemum morifolium*) self-incompatibility. *The Scientific World Journal* 2014:625658

41. Wang F, Zhong X, Huang L, Fang W, Chen F, et al. 2018. Cellular and molecular characteristics of pollen abortion in chrysanthemum cv. Kingfisher. *Plant Molecular Biology* 98:233−47

42. Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, et al. 2011. Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nature Biotechnology* 29:644−52

43. Pertea G, Huang X, Liang F, Antonescu V, Sultana R, et al. 2003. TIGR Gene Indices clustering tools (TGICL): a software system for fast clustering of large EST datasets. *Bioinformatics* 19:651−52

44. Hou X, Guo Q, Wei W, Guo L, Guo D, et al. 2018. Screening of genes related to early and late flowering in tree peony based on bulked segregant RNA sequencing and verification by quantitative real-time PCR. *Molecules* 23:689

45. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, et al. 2000. Gene Ontology: tool for the unification of biology. *Nature Genetics* 25:25−9

46. Ye J, Fang L, Zheng H, Zhang Y, Chen J, et al. 2006. WEGO: a web tool for plotting GO annotations. *Nucleic Acids Research* 34:W293−W297

47. Mortazavi A, Williams BA, McCue K, Schaeffer L, Wold B. 2008. Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nature Methods* 5:621−8

48. Audic S, Claverie JM. 1997. The significance of digital gene expression profiles. *Genome Research* 7:986−95

49. Wang F, Zhong X, Wang H, Song A, Chen F, et al. 2018. Investigation of differences in fertility among progenies from self-pollinated chrysanthemum. *International Journal of Molecular Sciences* 19:832

50. Earley KW, Haag JR, Pontes O, Opper K, Juehne T, et al. 2006. Gateway-compatible vectors for plant functional genomics and proteomics. *The Plant Journal* 45:616−29

51. Guan Y, Ding L, Jiang J, Shentu Y, Zhao W, et al. 2021. Overexpression of the *CmJAZ1-like* gene delays flowering in *Chrysanthemum morifolium*. *Horticulture Research* 8:87

52. Wang J, Guan Y, Ding L, Li P, Zhao W, et al. 2019. The *CmTCP20* gene regulates petal elongation growth in *Chrysanthemum morifolium*. *Plant Science* 280:248−57

53. Haffani YZ, Gaude T, Cock JM, Goring DR. 2004. Antisense suppression of thioredoxin *h* mRNA in *Brassica napus cv*. Westar pistils causes a low level constitutive pollen rejection response. *Plant Molecular Biology* 55:619−30

54. Gerola PD, Mol CA, Newbigin E, Lush WM. 2000. Regulation of *LAT52* promoter activity during pollen tube growth through the pistil of *Nicotiana alata*. *Sexual Plant Reproduction* 12:347−52