3 Biotech (2022) 12:137
https://doi.org/10.1007/s13205-022-03196-1

ORIGINAL ARTICLE

Identification and relative expression analysis of *CaFRK* gene family in pepper

Shufang Zhao[1] · Bingdiao Gou[1] · Yongfu Wang[1] · Nan Yang[1] · Panpan Duan[1] · Min Wei[1] · Gaoyuan Zhang[1] · Bingqiang Wei[1]

Received: 30 June 2021 / Accepted: 29 April 2022 / Published online: 24 May 2022

Abstract

Fructokinase is the main catalytic enzyme for fructose phosphorylation and can also act as a glucose receptor and signal molecule to regulate the metabolism of plants, which plays an important role in plant growth and development. In this study, the *CaFRK* gene family and their molecular characteristics are systematically identified and analyzed, and the specific expression of *CaFRKs* under different tissues, abiotic stresses and hormone treatments were explored. Nine *FRK* genes were authenticated in pepper genome database, which were dispersedly distributed on eight reference chromosomes and predicted to localize in the cytoplasm. Many *cis*-acting elements that respond to light, different stresses, hormones and tissue-specific expression were found in the promoters of *CaFRKs*. FRK proteins of four species including *Capsicum annuum*, *Arabidopsis thaliana*, *Solanum lycopersicum* and *Oryza sativa* were divided into four groups via phylogenetic analysis. The collinearity analysis showed that there were two collinear gene pairs between *CaFRKs* and *AtFRKs*. In addition, it was significantly found that *CaFRK9* expressed far higher in flower than other tissues, and the relative expression of *CaFRK9* was gradually enhanced with the development of flower buds in fertile accessions, 8B, R1 and F_1. Nevertheless, *CaFRK9* hardly expressed in all stages of cytoplasmic male sterile lines. Based on the quantitative real-time PCR, most of *CaFRK* genes showed significant up-regulation under low-temperature, NaCl and PEG6000 treatments. On the contrary, the expression levels of most *CaFRKs* revealed a various trend in response to hormone treatments (IAA, ABA, GA_3, SA and MeJA). This study systematically analyzed *CaFRK* gene family and studied its expression pattern, which lay the foundation of CaFRK genes cloning and functional verification response to abiotic stresses, and provides new insights into exploring the *CaFRK* genes on the pollen development in pepper.

Keywords *FRK* gene family · Gene expression · Male sterility · Pollen development · Pepper

Introduction

Pepper (*Capsicum annuum* L.) is an important commercial crops as vegetable spice crop and value-added processed products in the worldwide (Hong et al. 1998; Pino et al. 2006). Farmers have a growing demand for hybrid seeds of pepper due to their superior performance in quality, yield and stress resistance (Swamy et al. 2017). Cytoplasmic male sterility (CMS) system has an important agricultural value in the production of hybrid seeds without the need for flower emasculation, and it is one of the most effective methods on the utilization of crop heterosis (Zhang et al. 2015; Chase 2007; Hanson et al. 2004; Yang et al. 2020). Since male sterility is mainly manifested in pollen abortion, it is vitally important to have a good command of knowledge of the molecular mechanism of pollen germination and maturation becomes the basis for investigating male sterility (Chen et al. 2015). The development of pollen from archesporial cells to mature pollen involves a series of complex activities regulated by multitudinous genes (McCormick 2004). Any interruption during the period of pollen differentiation, stamens development, sporogenous cell differentiation, meiosis, microspore mitosis or anthesis may result in male sterility (Glover et al. 1998), which indicate that pollen development is a very complex and subtle process that is sensitive to mutations. As a crucial premise for success of pollination, pollen fertility is affected by a range of signal transduction,

✉ Bingqiang Wei
bqwei@gsau.edu.cn

[1] College of Horticulture, Gansu Agricultural University, Lanzhou 730070, China

biochemical changes, and complicated cellular interactions and numerous internal and external factors. Tapetum plays a role in anther development and pollen fertility, provides enzymes for callose dissolution and materials for formation and development of pollen walls (Yang et al. 2008). Some of genes influenced under CMS are those related with anther or pollen development, particularly the formation of pollen exine (Guo et al. 2017).

In most higher plants, sucrose is the major carbon resources for most of metabolic pathways. Sucrose must be cleaved into UDG-glucose and fructose by sucrose synthase (SUS), or can be cleaved into glucose and fructose by invertases for further sugar metabolism (Truernit et al. 1999; Shi et al. 2014). Before entering the metabolic pathway, the free hexoses, glucose and fructose must be phosphorylated by hexokinase (HXK, EC 2.7.1.1) or fructokinase (FRK, EC 2.7.1.4) (Ashwell 1964). Fructose can be phosphorylated to perform glycolysis and oxidized pentose pathways, which could also be used in starch synthesis. HXK and FRK could catalyze phosphorylation of fructose. Nevertheless, the affinity of FRK are much higher than hexokinase for fructose. So in most cases, FRK is the main catalytic enzyme for fructose phosphorylation (Renz and Stitt 1993). Studies have shown that the decrease of FRK activity was accompanied by a temporary halt of starch synthesis in developing tomato fruits. Further studies have shown that FRK could regulate the interconversion between sucrose and starch (Schaffer and Petreikov 1997). In addition, FRK can also act as a glucose receptor and signal molecule to regulate plant metabolism, growth and development (Rolland et al. 2006). *FRK* gene belongs to the phosphofructokinase B (PfkB), and the binding region of ATP to sugar is highly conserved (Fennington and Hughes 1996). At present, *FRK* genes have been cloned and identified in the tissues of Arabidopsis (Gonzali et al. 2001; Kaplan et al. 1997; Alexandrov et al. 2006), beet taproot (Franck et al. 1995), potato tuber (Taylor et al. 1995), tomato fruit (Kanayama et al. 1997; German et al. 2004), rice (Jiang et al. 2003), soybean (Kuo et al. 1990), maize endosperm (Zhang et al. 2003), poplar (Ralph et al. 2008a), barley leaves (Baysdorfer et al. 1989), avocado (Copeland and Tanner 1988), spinach leaves (Schnarrenberger 1990), camellia pollen (Nakamura et al. 1991), lily pollen (Nakamura et al. 1991), pea (Copeland et al. 1984), sugarcane (Hoepfner and Botha 2004), spruce (Ralph et al. 2008b) and citrus (Qin et al. 2004). Nevertheless, due to the complex pepper genetic background, there are no reports about identification and function of *FRK* genes in pepper.

Understanding the molecular structure and evolution of gene family is a critical step to search the physiological function and metabolic mechanism of its members. What's more, gene identification makes it possible to study gene expression to evaluate the potential function of gene family. In this study, nine *CaFRK* genes were identified from pepper genomics, and the gene structure characteristics, conserved motifs, chromosomal localization, composition of *cis*-acting elements and phylogenetic relationship of *CaFRKs* were comprehensively investigated. In addition, the expression patterns of *CaFRKs* were in detail explored under different tissues, abiotic stresses and hormone treatments. In our previous study on the comparative analysis between the buds of CMS accessions and fertility restorer accessions, *Capana00g002348*, designated as *CaFRK9* in this paper, was significant up-regulated in restorer accessions as compared to CMS accessions in pepper (Wei et al. 2019), and the raw data can be found in NCBI (https://www.ncbi.nlm.nih.gov/) with an accession number of SRA895207. In this study, the expression level of *CaFRK9* was also analyzed by quantitative real-time polymerase chain reaction (qRT-PCR) between male sterile accessions and fertile accessions at different developmental stage of flower buds (I, calyx closed; II, corolla is flush with the calyx; III, the ratio of corolla to calyx calyx height is about 1:1; IV, the buds are about to open.). This study suggested an overall knowledge of *CaFRK* gene family, and provided a new insight into the function of *FRK* gene on pollen development, the restoration of fertility and response to abiotic stresses.

Materials and methods

Identification and sequence characterization of *FRK* gene

The proteome and genome data of *Capsicum annuum* Zunla database (v2.0) were downloaded from Sol Genomics Network (https://solgenomics.net/ftp/genomes/Capsicum_annuum/C.annuum_zunla/assemblies/Capsicum.annuum.L_Zunla-1_Release_2.0.fasta.gz) to establish pepper local genome-wide database. The FRK genes of *Arabidopsis thaliana* were obtained from the published article (Riggs et al. 2017), and the protein sequences were downloaded from the Arabidopsis Information Resource (TAIR11) (https://www.arabidopsis.org/). The Hidden Markov Model (HMM) profiles of the *FRK* conservative domain (PfkB, serial number PF00294) was downloaded from the Pfam database (http://pfam.xfam.org/) (Sara er al. 2019). Subsequently, the FRK genes were searched from pepper local genome-wide database using HMMER3.0 (the E-value was less than 1e-10) (Finn et al. 2011). According to the results of further analysis for sequence alignment with *AtFRKs* and functional annotation of *Capsicum annuum* Zunla Genome protein sequences (v2.0) (Sol Genomics Network), the redundant protein sequences were eliminated. The candidate protein sequences were predicted by the online software SMART (http://smart.embl.de/) and Pfam database (http://pfam.

xfam.org/) to further identify whether the PfkB domain was contained.

The physicochemical properties of *FRK* genes of pepper were analyzed using the ProtParam (https://web.expasy.org/protparam/) (Gasteiger et al. 2005), including molecular weight, number of amino acids, isoelectric point. Subcellular localization of *CaFRK* genes was predicted using the Euk-mPLoc 2.0 (http://www.csbio.sjtu.edu.cn/bioinf/euk-multi-2/) (Chou and Shen 2010).

Chromosomal location, collinearity, promoter *cis*-acting element and gene structure analysis

The chromosome location information of *CaFRKs* was obtained from the pepper local genome-wide database. The mappings of physical locations of the *FRK* genes on pepper chromosomes were draw using MapChart2.3 tools (Voorrips 2002). The collinearity of *FRK* genes of pepper and *Arabidopsis* was analyzed using MCScanX from TBtools. The upstream 2000 bp sequence of the start codon of each *CaFRK* coding sequence (CDS) was extracted by TBTools, and the promoter *cis*-acting elements of *CaFRKs* were predicted by the PlantCARE server (http://bioinformatics.psb.ugent.be/webtools/plantcare/html/) (Lescot et al. 2001). The intron/exon configuration of the *FRK* genes coding sequences and genomic sequences are shown in the gene structure diagram aided by GSDS2.0 software (http://gsds.gao-lab.org/index.php) (Hu et al. 2015).

Conservated domain analysis, conservative motifs and secondary structure prediction

The conservative domains of CaFRK proteins were downloaded from Pfam v33.1-18271 PSSMs database in NCBI CDD (https://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi). The conservative motifs of CaFRK proteins were analyzed by MEME5.2.0 software (http://meme-suite.org/meme_5.2.0/) (Bailey et al. 2009), under these parameters: the Motif E-value threshold was not limited; the optimum motif width ranged from 6 to 50; and the maximum number of motifs of the conserved domain was set to 10 (Bailey et al. 2006). The prediction results of conservative domains and motifs were analyzed in visualization by TBtools (Chen et al. 2020). The online software SOPMA was used to predict the secondary structure of FRK protein in pepper, including random coil, extended strand, alpha helix and beta turn (https://npsa-prabi.ibcp.fr/cgi-bin/npsa_automat.pl?page=npsa_sopma.html).

Phylogenetic tree construction of FRK family genes

To research the phylogenetic relationship of *FRK* genes, FRK protein sequences were obtained from *Capsicum annuum* L., *Arabidopsis thaliana, Solanum lycopersicum, Oryza sativa*. Protein sequences of published *AtFRKs* were downloaded from the *Arabidopsis thaliana* Database (https://www.arabidopsis.org/) (Riggs et al. 2017). Based on the data obtained from previous searches in phytozome database (Ye and Zhou 2021), FRK protein sequences of *Solanum lycopersicum* and *Oryza sativa* were obtained from the Sol Genomics Network (https://solgenomics.net/) and Rice Whole Genome Database (http://rice.plantbiology.msu.edu/) (Feng et al. 2016), respectively. Multiple sequence alignment for FRK proteins of these four species was conducted by ClustalW (Higgins et al. 1996) in MEGA-X software with the default parameters, and the phylogenetic tree was constructed by neighbor-joining method in MEGA-X (Kumar et al. 2018; Saitou 1987). In the phylogenetic tree, execution parameters were p-distance and pairwise deletion, and the number of repeats of bootstrap was set to 1000 (Sanderson 1989), and the default option was selected for the remaining set-up. And then the result was revealed via the software Evolview version 2 (https://www.evolgenius.info//evolview/#login) (He et al. 2016).

Plant materials, cultivation and treatment conditions

The plump seeds of pepper variety "Qiangfeng 101" were wrapped in gauze, then sterilized by a 20% hypochlorous acid solution for 20 min and soaked in distilled water for 8 h. The seeds were placed in a controlled chamber (28 °C day/night, 24 h dark cycle) to germinate. After germination for 3 days, the seedlings were transferred into hydroponic boxes containing hoagland nutrient solution, and were placed in plant incubators at 28 ℃ day/23 ℃ night and a photoperiod of 16 h light/8 h dark. The six-leaf stage seedlings were used for abiotic stress treatments. For low-temperature treatment, plant samples were placed in plant incubators at 23 °C day/18 °C night with a photoperiod of 16 h light/8 h dark and a light intensity of 2000 Lx. For drought, salt and hormone treatments, plant samples were grown in nutrient solution containing 15% (w/v) polyethylene glycol (PEG6000), 300 mM NaCl, 0.2 mM auxin (IAA), 0.6 mM abscisic acid (ABA), 0.6 mM gibberellin acid (GA_3), 2 mM salicylic acid (SA) and 0.1 mM methyl jasmonate (MeJA), respectively. After treatment for 0, 1, 3, 6, 12, 24, 48, 72 h, 4–6 true leaves of pepper were collected and stored at − 80 ℃ for qRT-PCR.

Furthermore, seeds with uniform germination were selected and seeded in a nutrient bowl containing a substrate (vermiculite: nutrient soil = 3:1). The Light incubator parameters were set as 28 ℃/23 ℃ (day/night) with a photoperiod of 16 h/8 h (light/dark) and a light intensity of 20,000 Lx. The different tissues, including the roots, tender stems, fresh leaves, flower buds, young fruit placenta, big fruit placenta, ripened fruit placenta, young fruit flesh, big fruit flesh and

ripened fruit flesh (Fig. 1) were collected and directly frozen in liquid nitrogen and stored at − 80 °C for RNA extraction.

The pepper CMS line 8A and its maintainer line 8B were a pair of near-isogenic lines. 8A is absolutely sterile with no pollen on anthers at the whole flowering stage. On the contrary, 8B presents massive and active pollen grains at anthesis stage. The fertility restorer line R1 exhibits a large number pollens on the anther when the flowers are open. Furthermore, the hybrid F_1 of 8A × R1 shows abundant pollens as R1 and 8A. All accessions were cultivated in plastic tunnel, standard horticultural practices were followed for disease and pest control, which were described in a previously published article (Wei et al. 2020). Healthy flower buds at different developmental stages were sampled, immediately frozen in liquid nitrogen and stored at − 80 °C until used.

RNA extraction, qRT-PCR and statistical analysis

Total RNA was extracted using the RNA simple Total RNA Kit (TIANGEN, Beijing, China) according to the manufacturer's instruction. The quality of RNA samples was verified by agarose gel electrophoresis. The total RNA extracted from pepper flower bud and tissue was used as the templates, and were reversely transcribed into cDNA using PrimeScript™ RT reagent Kit with gDNA Eraser (TaKaRa Biotechnology, Dalian, China).

The specific primers of *CaFRKs* and internal reference gene (β-actin; Cla007792) were designed according to the gene sequences using Primer Premier 5.0, and were synthesized by Shanghai Sangon Biotechnology (Table S1). Applied Biosystems StepOnePlus Real-Time PCR System was used for detection. qRT-PCR was carried out using TB Green® Premix Ex Taq™ II (TaKaRa Biotechnology, Dalian, China) according to the manufacturer's instructions, and was operated in a 20 µL volume that included 10 µL of 2 × SYBR Green Master mix, 0.8 µL of forward primer (10 µM) and reverse primer (10 µM) respectively, 2 µL of cDNA template and 6 µL of double-distilled water. The qRT-PCR was performed with the following parameters: denaturation at 95 °C for 30 s, 40 cycles of denaturation at 95 °C for 5 s, annealing and extension at 60 °C for 30 s. Triplicate qRT-PCR experiments were performed for all samples.

The gene expression data of *CaFRKs* were calculated by $2^{-\Delta\Delta CT}$ method (Livak and Schmittgen 2000; Taylor et al. 1995). In tissue expression analysis, the mean value ΔCt of all tested tissues was used as endogenous control to calculate the relative expression amount of genes in each tissue. In abiotic stresses and hormone treatments, 0 h of treatment was used as endogenous control. The quantitative data were performed one-way analysis of variance, and Duncan's method was used for significant difference analysis by SPSS22, and $P < 0.05$ was considered as significant difference. Histograms with error lines and heatmaps were drawn using Origin 2019 and TBtools, respectively.

Fig. 1 The images of young fruit flesh, big fruit flesh and ripened fruit flesh. The length of young fruit flesh is about 6 cm; The length of big fruit flesh is about 20 cm; The flesh of large fruit that turns red but no longer becomes long and thick is called the ripened fruit flesh

Results

Identification and characterization of *FRK* gene family in pepper

Through screening and identification, a total of nine *FRK* gene family members were identified in pepper genome, which were named *CaFRK1-9* according to their location on the chromosomes. (Table 1). The conserved domain of *FRK* was identified as PfkB (phosphofructokinase B) (serial number PF00294) by SMART and Pfam database research. The results manifested that all the *FRK* gene family members in pepper had typical PfkB domains, and *CaFRK5* also contained a ArgJ domain (also known as Ornithine acetyltransferase/OAT, serial number PF01960). The position of PfkB domain was different due to the different amino acid residues of FRK family proteins in pepper. Sequences analysis of *CaFRKs* indicated that the deduced amino acid sequence lengths ranged from 187 to 676 aa, and the CDS lengths from 564 to 2031 bp. The molecular weight of *CaFRKs* ranged between 20.59 and 75.59 kDa. The theoretical isoelectric point (pI) values of most *CaFRKs* were less than 7, only that of *CaFRK6* with a value of 8.76, which indicated that all *CaFRKs* were acidic except *CaFRK6*. In addition, subcellular locations manifested that all *CaFRK* genes

Table 1 The characteristics of *FRK* gene family in pepper genome

Gene	Sequence ID	Chromosome	PfkB domain location	Number of amino acids	CDS Length(bp)	Molecular weight (kDa)	pI	Subcellular prediction
CaFRK1	*Capana02g002877*	2	73 ~ 382	394	1185	42.37	5.35	Cytoplasm, nucleus
CaFRK2	*Capana03g002428*	3	21 ~ 331	343	1032	37.00	5.09	Cytoplasm, nucleus
CaFRK3	*Capana03g003379*	3	84 ~ 409	485	1458	52.42	6.60	Cytoplasm, nucleus, chloroplast, mitochondrion
CaFRK4	*Capana06g000734*	6	9 ~ 317	328	987	34.88	5.56	Cytoplasm
CaFRK5	*Capana07g001748*	7	182 ~ 255 (ArgJ:199 ~ 266)	266	801	29.25	5.01	Cytoplasm
CaFRK6	*Capana09g001611*	9	3 ~ 70	187	564	20.59	8.76	Cytoplasm
CaFRK7	*Capana11g001996*	11	276 ~ 583	676	2031	75.59	6.24	Cytoplasm, nucleus
CaFRK8	*Capana12g002513*	12	112 ~ 447	463	1392	52.98	6.96	Cytoplasm, nucleus
CaFRK9	*Capana00g002348*	–	62 ~ 258; 266 ~ 399	412	1239	44.68	6.17	Cytoplasm, nucleus, chloroplast

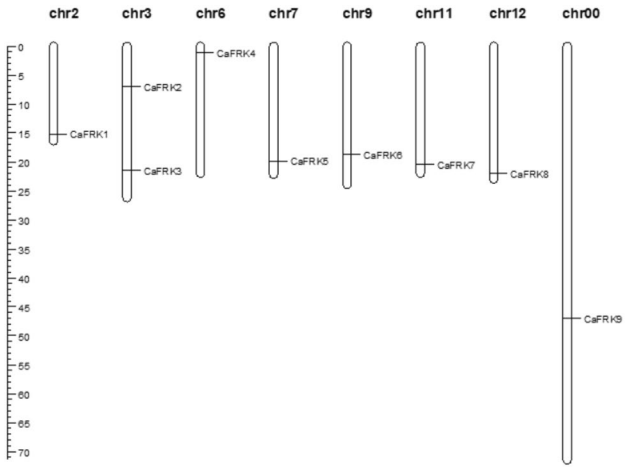


Fig. 2 Chromosomal locations of *CaFRK* genes. Chromosome numbers were showed at the top of each chromosome. The black lines on the pepper chromosomes indicated the location of the *CaFRK* genes

were predicted to localize in the cytoplasm. Among them, six *CaFRK* genes (*CaFRK1*, *CaFRK2*, *CaFRK3*, *CaFRK7*, *CaFRK8*, *CaFRK9*) were also located in the nucleus, two *CaFRK* genes (*CaFRK3*, *CaFRK9*) also in the chloroplast and *CaFRK3* also in the mitochondrion.

Chromosome localization and gene structure analysis of *CaFRKs*

The chromosomal distribution and orientation of *CaFRKs* were obtained by Mapchart2.3 tool (Fig. 2). Except *CaFRK9* was located in chromosome 0, the rest of *CaFRKs* were randomly and unequally distributed on seven out of the 12 chromosomes. There were two *CaFRKs* distributed on chromosome 3, and only one *CaFRK* gene was located on each of the remaining chromosomes. Interestingly, most of *CaFRK* genes were distributed at the end of the corresponding chromosomes, such as *CaFRK1*, *CaFRK3* and *CaFRK5-8*.

Intron/exon configurations of *CaFRKs* were constructed using the Gene Structure Display Server by aligning the cDNA sequences with the corresponding genomic DNA sequences. The full lengths of *CaFRKs* dramatically varied between 3 and 14 kb, and exon numbers ranged from two to eight, and the sizes of exons were obviously discrepant (Fig. 3). In addition, the sizes of introns have disparity. Remarkably, one intron size of *CaFRK3* was more than 9 kb. Two *CaFRKs* (*CaFRK5* and *CaFRK9*) comprised the highest exon number with a value of eight, whereas *CaFRK6* and *CaFRK8* only contained two exons. It was interesting that the numbers of exons were not consistent with the length of genes. *CaFRK8* with the shortest sequence of about 3 kb possessed two exons. However, although the sequence of *CaFRK3* was more than 14 kb and exceeded far from that of other family genes, the number of exons was only five.

Promoter *cis*-acting element analysis of *CaFRKs*

To further understand the function and regulation mode of *CaFRKs*, *cis*-elements in the promoter sequences of *CaFRKs* were studied. The 2000 bp upstream region of the start codon of *CaFRKs* was analyzed by PlantCare software. The results showed that 14 *cis*-acting elements were found in *CaFRK* promoters by screening and classifying (Fig. 4). These *cis*-acting elements were classified to four major classes including light-responsive, stress-responsive, hormone-responsive, growth and development correlation, respectively. A number of light-responsive elements in *CaFRKs* promoters were observed. Stress-responsive elements consisted of ARE, TC-rich, LTR

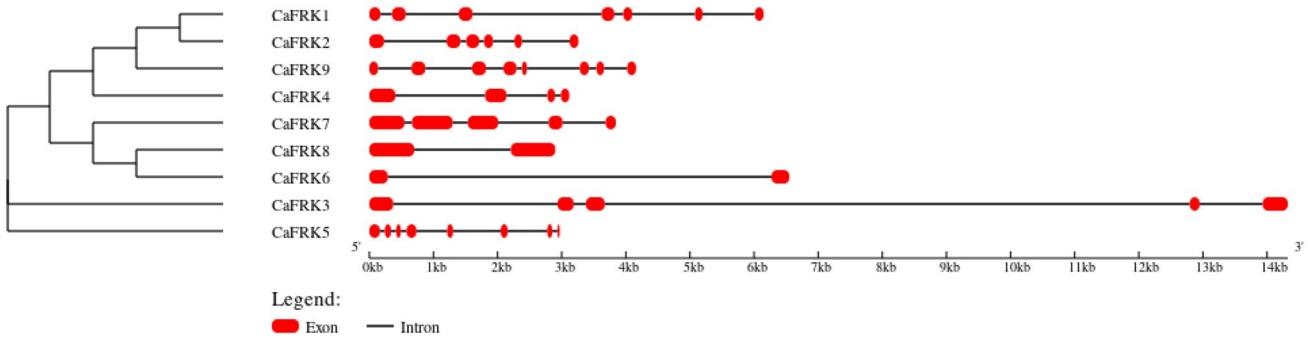


Fig. 3 Intron/exon configurations of *CaFRK* genes. Exons and introns were shown as red boxes and thin lines, respectively

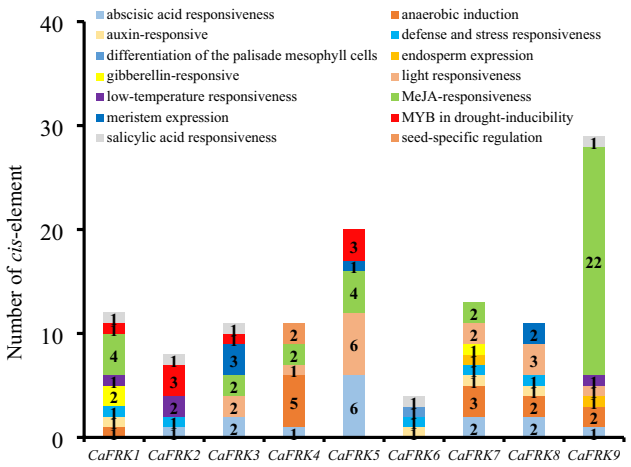


Fig. 4 *Cis*-elements in the promoters of *CaFRK* genes. The *cis*-elements with diverse functions are presented in various colors. The numbers represents the number of *cis*-elements contained each *CaFRK*

and MBS, which reflected plant responses to anaerobic induction, defense and stress responsiveness, low-temperature responsiveness and MYB-binding site involved in drought-inducibility respectively. Five types of hormone-responsive elements including MeJA (TGACG motif, CGTCA motif), abscisic acid (ABRE), salicylic acid (TCA), gibberellins (P box, GARE motif, TATC box) and auxin (TGA element) were identified, which indicated that *CaFRK* genes may be regulated by hormones. In addition, four kinds of growth and development correlative elements were discovered, including differentiation of the palisade mesophyll cells (HD-Zip 1), seed-specific regulation (RY element), meristem (CTA box) and endosperm expression (GCN4 motif). It was worth noting that MeJA-responsiveness element was found to be widely existed in *CaFRK9*, which indicated that *CaFRK9* may be regulated by MeJA. These results suggest that *CaFRKs* may play an important role in regulating the growth and development, and it respond to light, abiotic stress and hormones.

Conserved motif, domains and secondary structure prediction of CaFRKs

To further examine the structural features of CaFRKs, the conserved motifs of CaFRK protein sequences were analyzed using MEME to obtain 10 conserved motifs (Fig. 5 a). The results of conserved domains and motifs analysis were visualized by TBtools (Fig. 5 b). The length of 10 motifs ranged from 8 to 50 amino acids, and the numbers of motifs distributed on *CaFRKs* diversified from one to eight. The same conservative motifs were observed in *CaFRK1, CaFRK2, CaFRK4* and *CaFRK9*, and their domain architectures were consistent. Additionally, except for Motif 1 existing in all CaFRK proteins, Motif 8 and Motif 10 only in *CaFRK7* and *CaFRK8*, the rest of eight motifs were unequally distributed on the CaFRK proteins, including Motif 2 on eight *CaFRKs* except *CaFRK9*, Motif 3, Motif 7 and Motif 9 on six *CaFRKs*, Motif 4 on seven *CaFRKs*, Motif 5 and Motif 6 on four *CaFRKs*, respectively.

According to the prediction results of the secondary structure of CaFRK proteins, the FRK family proteins of pepper were composed of α-helix, β-corner, extended strand and random coil (Table 2). Among them, α-helix and random coil were authenticated as the main secondary structure components of CaFRK proteins, and the β-corner occupied the least proportion of all the structures. Moreover, the data showed that the proportions of the secondary structure of the six *CaFRKs* were identified as random coil > α-helix > extended strand > β-corner (*CaFRK1*, *CaFRK5-9*). *CaFRK2* was pointed out that the proportion of α-helix and random coil was equal.

Phylogenetic tree construction and collinearity analysis of FRK family genes

To investigate the phylogenetic relationships of FRK proteins between pepper and other plant species, FRK proteins from *Capsicum annuum* L., *Arabidopsis thaliana*, *Solanum lycopersicum* and *Oryza sativa* were used to

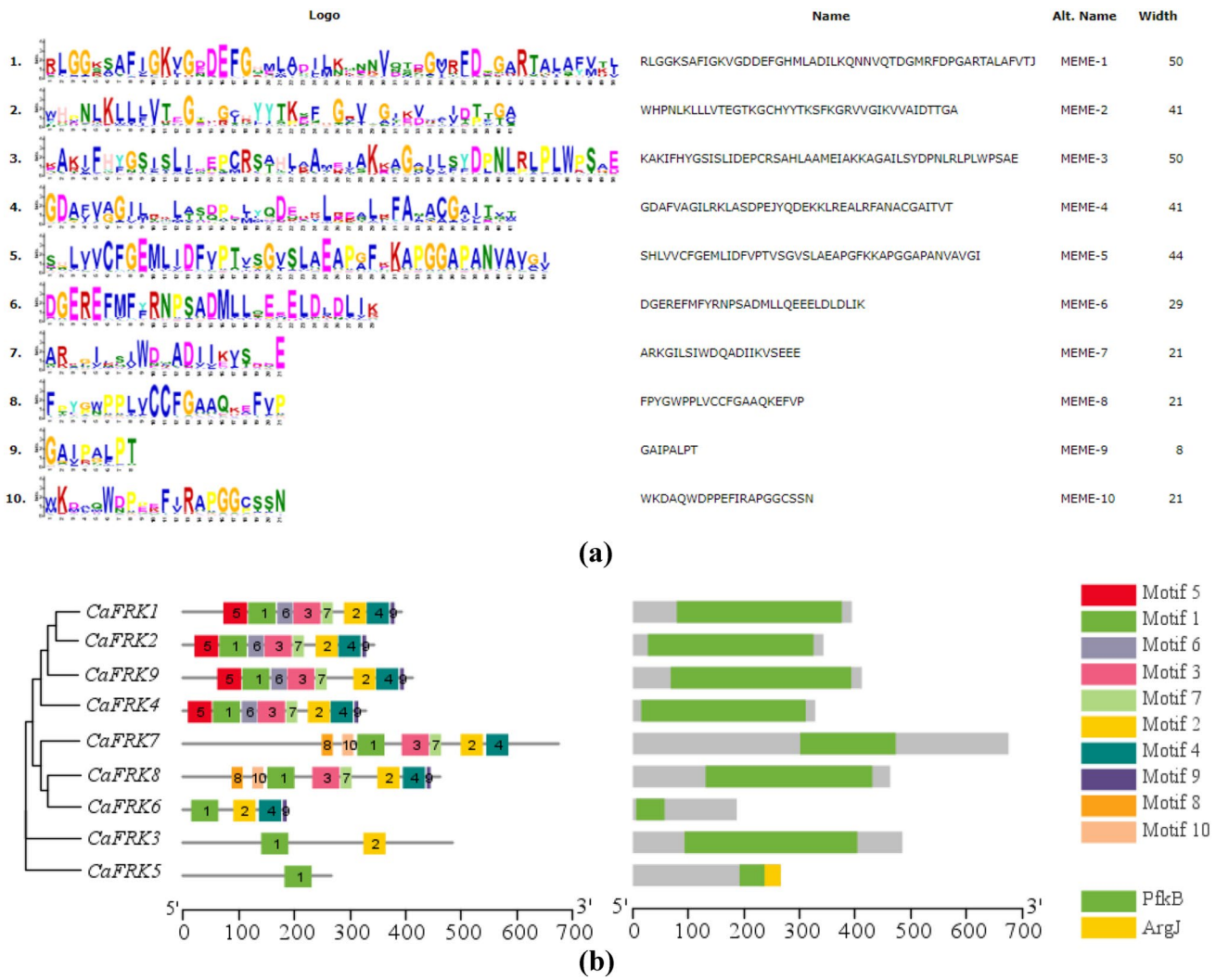


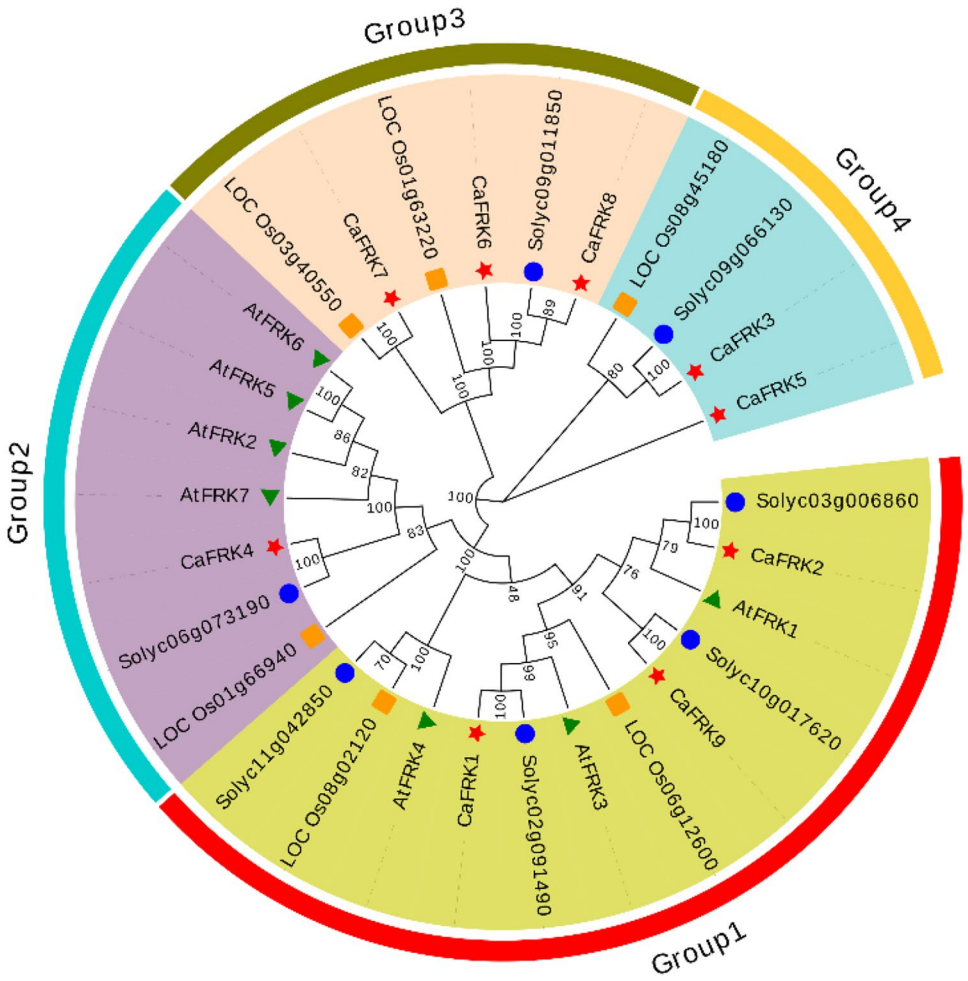
Fig. 5 Conserved domains and motifs analysis of CaFRK proteins. **a** The motif sequences and widths in *CaFRKs*, which were identified by MEME. **b** Different color boxes represent types of domains and motifs, and the number represents the name of different motifs

Table 2 The secondary structure of *FRK* gene family from *Capsicum annuum* L

Gene	α-helix (%)	β-corner (%)	Extended strand (%)	Random coil (%)
CaFRK1	33.50	5.58	21.57	39.34
CaFRK2	35.57	6.41	22.45	35.57
CaFRK3	39.38	6.60	17.32	36.70
CaFRK4	36.59	6.40	21.04	35.98
CaFRK5	29.32	5.64	11.65	53.38
CaFRK6	30.48	8.02	25.13	36.36
CaFRK7	36.98	6.66	12.87	43.49
CaFRK8	32.61	5.18	16.41	45.79
CaFRK9	35.92	6.07	21.12	36.89

build phylogenetic tree with MEGA-X software (Fig. 6). In our study, the phylogenetic tree can be separated into four main classes that referred to as Group 1–4 based on the topology and bootstrap values. The results showed that the *FRK* genes of pepper had close genetic distance with the dicotyledonous plant tomato and most bootstrap values were 100, but were distantly related to the *FRKs* of monocotyledonous plant rice according to the evolutionary relationship. Remarkably, we found that only *CaFRK4* was closely clustered with four *AtFRK* genes in Group 2. Noteworthily, the losses of *AtFRKs* were observed on the Group 3 and Group 4.

The MCScanX of TBtools software was used to analyze the collinearity of FRK genes in *Arabidopsis thaliana* and *Capsicum annuum* L. The results showed there are two collinear gene pairs between nine *CaFRKs* and seven *AtFRKs* (Fig. 7). *CaFRK1* and *CaFRK2* were collinear with *AtFRK3*



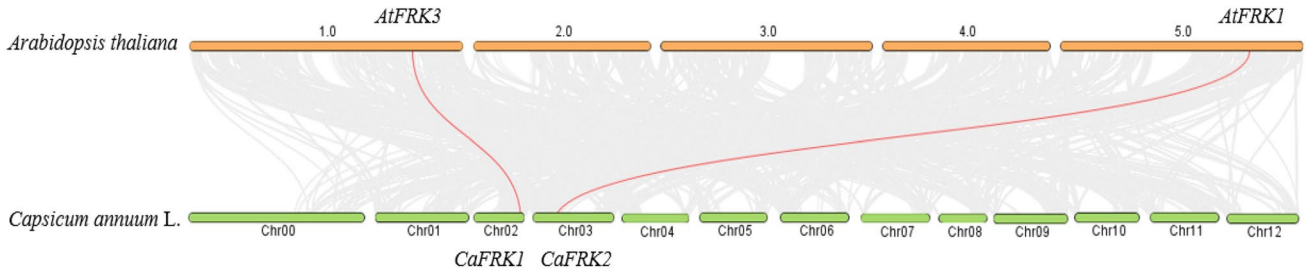


Fig. 6 Phylogenetic analysis of *FRK* family genes from *Capsicum annuum* L., *Arabidopsis thaliana*, *Solanum lycopersicum*, *Oryza sativa*. The phylogenetic tree was constructed using the approach of neighbor-joining with 1000 bootstrap replicates, and the results were displayed through Evolview Version 2

Fig. 7 Synteny analysis of *FRK* genes between *Arabidopsis thaliana* and *Capsicum annuum* L. The chromosome numbers of *Capsicum annuum* L. are Chr00 ~ Chr12, and the chromosome numbers of *Arabidopsis thaliana* are 1.0 ~ 5.0; Red curves indicate synteny relationship between *Arabidopsis thaliana* and *Capsicum annuum* L. FRK genes

and *AtFRK1* respectively (*CaFRK1*/ *AtFRK3*, *CaFRK2* /*AtFRK1*).

Expression patterns analysis of *CaFRKs* in different tissues

To characterize tissue-specific expression, the expressions of *CaFRKs* were analyzed by qRT-PCR during ten different tissues. The results show that the expressions of *CaFRKs* were identified in all tested tissues, but the expression levels of *CaFRKs* were significantly different in this ten tissues (Fig. 8). The expression levels of *CaFRK1-3* were high in the young fruit flesh, big fruit placenta and big fruit flesh, respectively, and that of *CaFRK4* was high in the roots. *CaFRK1* and *CaFRK5-8* were specifically expressed in leaves, about 1 to 5 times as much as in other tissue. It is remarkable that the expression level of *CaFRK9* in flower was more than 350-fold higher than big fruit flesh. In

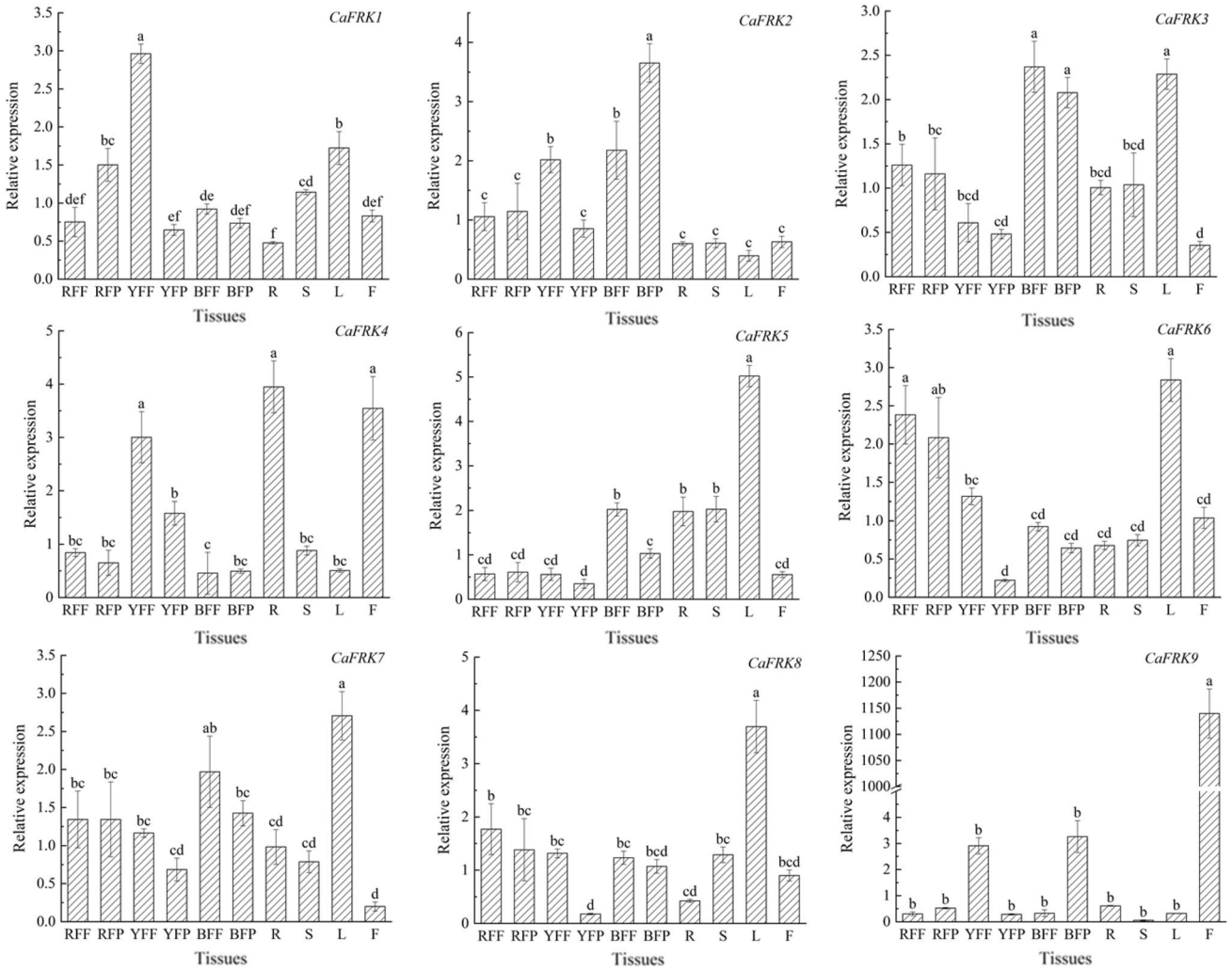


Fig. 8 The relative expression of *CaFRKs* in different tissues. *YFF* young fruit flesh, *YFP* young fruit placenta, *BFF* big fruit flesh, *BFP* big fruit placenta, *RFF* ripened fruit flesh, *RFP* ripened fruit placenta, *R* root, *S* stem, *L* leaf, *F* flower. Different lowercase letters indicate significant difference ($P < 0.05$)

addition, the expression level of *CaFRK9* was more than 320 times in flowers than *CaFRK1-8*. Therefore, *CaFRK9* may be closely related to the growth and development of flowers.

Expression analysis of *CaFRK9* in different fertility accessions

To further verify whether *CaFRK9* was related to pollen development, the relative expression of *CaFRK9* in different developmental stages of flower buds was analyzed using qRT-PCR. The results indicated that *CaFRK9* hardly expressed at stage I, II and III of all accessions except for a little expression at stage III of 8A and R1 (Fig. 9). Interestingly, the relative expression of *CaFRK9* was extremely high at stage IV in fertile accessions (8B, R1 and F_1), about 9- to 12.5-fold higher than that of the sterile accession 8A.

Expression patterns analysis of *CaFRKs* under low temperature, NaCl and PEG treatments

We conducted qRT-PCR experiments of *CaFRKs* under low-temperature, NaCl and PEG treatments to further understand the possible roles of *CaFRKs* in response to abiotic stresses. Our results indicated that all *CaFRKs* showed up-regulated expression patterns compared to untreated control under low-temperature (Fig. 10a). The expression of *CaFRK2, CaFRK3* and *CaFRK8* was up-regulated significantly and peaked at 72 h. Nevertheless, *CaFRK5* and *CaFRK7* were up-regulated slightly and peaked at 3 h. The expression level of *CaFRK4* increased most obviously, about 11–28 times.

Under NaCl treatment, the expression level of *CaFRK6* and *CaFRK8* was slightly changed, which was up-regulated in the early stage and down-regulated in the later stage (Fig. 10b). However, *CaFRK7* was down-regulated

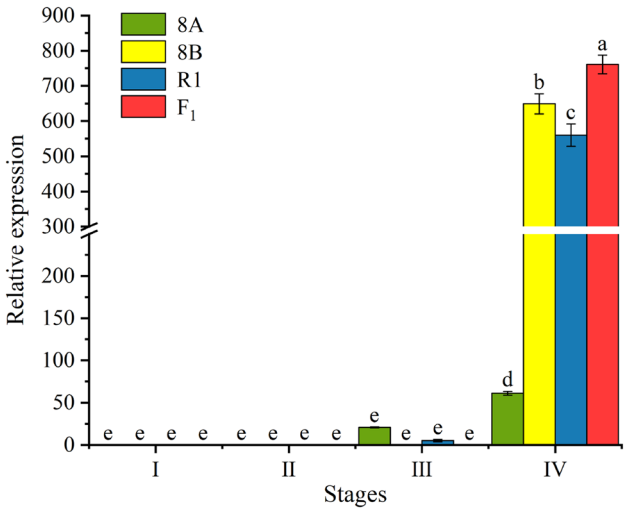


Fig. 9 The relative expression of *CaFRK9* in four different stages of flower buds in 8A, 8B, R1 and F_1. I, calyx closed; II, corolla is flush with the calyx; III, the ratio of corolla to calyx height is about 1:1; IV, the buds are about to open. Different lowercase letters indicate significant difference ($P < 0.05$)

and reached the lowest expression level at 12 h. *CaFRK2* was most significantly up-regulated compared with other *CaFRK* genes and reached its peak at 6 h, which was about 8.5 times that of the control. Under PEG treatment, the expression levels of all *CaFRK* genes were up-regulated except *CaFRK7* was significantly down-regulated fivefold change when treated for 24 h. After 48 h treatment, the expression level of *CaFRK9* was up-regulated 65 times. The rest *CaFRKs* (*CaFRK1-6* and *CaFRK8*) were up-regulated 2.5- to 13-fold change when treated for 12 or 48 h. These results suggest that *CaFRKs* may play an important role in pepper response to abiotic stress.

Expression patterns analysis of *CaFRKs* under different hormone treatments

In this experiment, we studied the expression patterns of *CaFRK* genes under different hormone treatments. The results indicated that *CaFRKs* showed different expression trends under the same hormone treatment (Fig. 11). Under IAA treatment, *CaFRK6* and *CaFRK8* were up-regulated significantly, whereas *CaFRK1* showed down-regulated obviously. The expression levels of *CaFRK2* and *CaFRK6-8* were up-regulated four- to sevenfold change when treated for 12 h. The relative expression of *CaFRK4* and *CaFRK9* displayed a trend of down-regulation and then up-regulation, and reached the peak at 48 or 72 h. Under ABA treatment, *CaFRK1* and *CaFRK5* were down-regulated significantly, whereas others genes were basically up-regulated. The peak of all *CaFRK* genes occurred almost after 12 h treatment. It

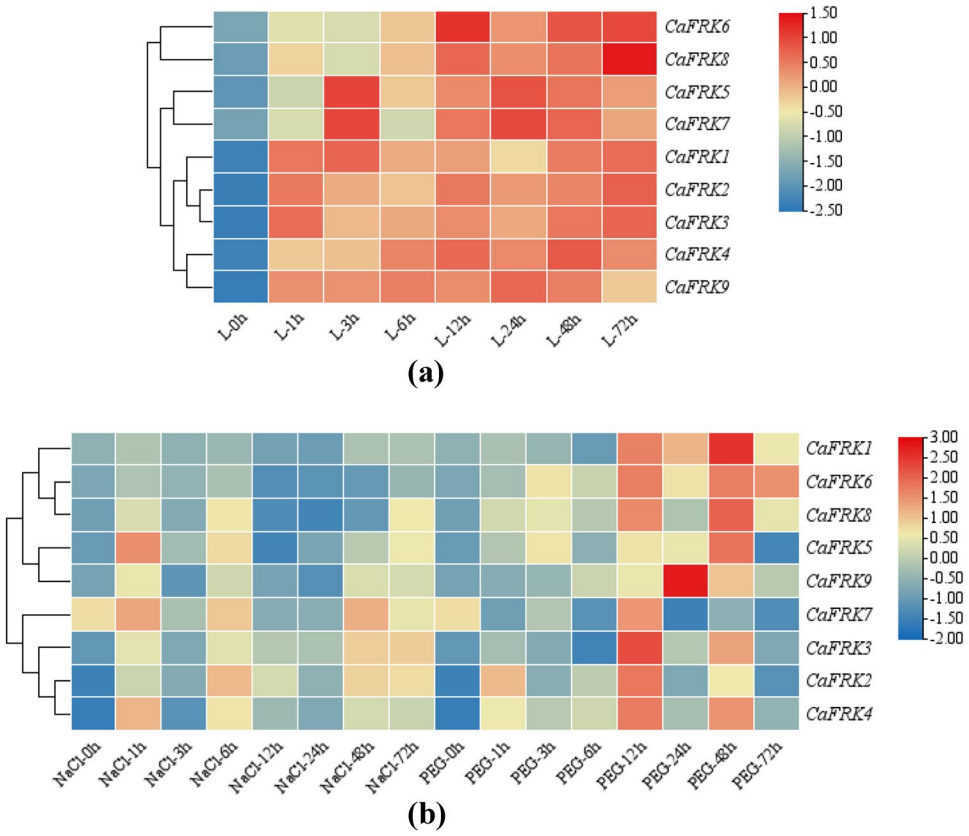


Fig. 10 Expression analyses of *CaFRKs* under low-temperature treatment, NaCl and PEG. Plant samples were treated with low temperature (23 °C day/18 °C night), 15% (w/v) PEG6000 and 300 mM NaCl. The color scale represents fold changes normalized by log2-transformed data. The red shading represents up-regulated genes and blue shading represents down-regulated genes

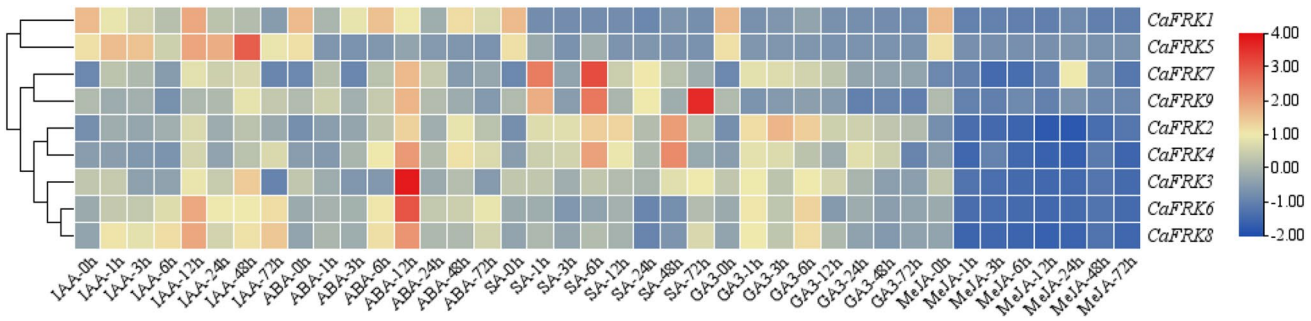


Fig. 11 Expression analyses of *CaFRKs* under different hormone treatments, containing 0.2 mM IAA, 0.6 mM ABA, 0.6 mM GA_3, 2 mM SA and 0.1 mM MeJA. The color scale represented fold changes normalized by log2-transformed data. Red and blue shading represented the up-regulated and down-regulated expression level, respectively

is noteworthy that the expression level of *CaFRK3* reached its peak at 12 h of treatment, which was 6.5 times that of the untreated level, but in other treatment periods were down-regulated relative to the control. Under SA treatment, the differences of *CaFRK* genes expression were obvious. The expression levels of *CaFRK1* and *CaFRK5* were down-regulated significantly and extremely low. Whereas, *CaFRK2*, *CaFRK4* and *CaFRK7* were up-regulated 11- to 50-fold. Under GA_3 treatment, decreased expression levels of more than twofold change were found in three *CaFRK* genes (*CaFRK1*, *CaFRK5* and *CaFRK9*). In contrast, *CaFRK2*, *CaFRK4* and *CaFRK7* displayed increased expression levels of more than threefold change when treated for 1 h. Three genes (*CaFRK3*, *CaFRK6* and *CaFRK8*) increased to peak at 6 h of treatment, and then it started to descend. The astonishing thing was that the MeJA treatment caused a large decline in the expression of almost all *CaFRKs*. The expression level of *CaFRK7* was down-regulated in the early stage of treatment, but suddenly increased by 8 times at 24 h. These results suggested that *CaFRK* genes may be involved in hormone regulation.

Discussion

To maintain their growth and development, plants have evolved various mechanisms to adapt to various environmental stresses. Recent studies indicated that FRK, as an important carbon flux-regulated kinase in plants, played an important role in plant stress. In most plants, the primary transported sugar is sucrose, which is mainly phosphorylated by FRK after being decomposed into fructose. Additionally, sugar played a momentous part in pollen development and germination, provides energy for metabolism, structure, storage and signal functions (Clement and Burrus 1996; Clément and Audran. 1995; Goetz et al. 2001; Hirsche et al. 2009). Therefore, the study of *FRK* genes will contribute to explore its influence on adversity stress and pollen development. With the development of whole genome sequencing, a good deal of potential *FRK* gene families has been identified in many species in recent years, and the function of *FRK* genes has been thoroughly studied in tomato (David-Schwartz et al. 2013; German et al. 2002, 2004; Kanayama et al. 1997) and *Arabidopsis* (Gonzali et al. 2001; Kaplan et al. 1997; Alexandrov et al. 2006). Fortunately, bioinformatics, as an independent subject emphasizing both theory and practice, has been widely used to the identification and analysis of gene family.

In previous studies, seven *MeFRKs* and *SsFRKs* were identified in cassava (*Manihot esculenta* Crantz) and sugarcane (*Saccharum* spp.) genomes respectively (Yao et al. 2017; Chen et al. 2017). However, Cao et al. have reported that 49 *FRK* genes were identified in four species of Rosaceae, including 20 *FRK* genes in Chinese white pear, and six, eight and 15 in strawberry, peach and mei, respectively (Cao et al. 2018). Differently, in this study, a total of nine *CaFRKs* were identified using bioinformatics methods in pepper. These distinct evidence manifested that the members of the *FRK* gene family were different between plant species. The molecular characteristics, gene structure, conserved motifs, chromosome distribution, secondary structure, *cis*-acting elements, subcellular localization and phylogenetic relationships of *CaFRKs* were predicted and analyzed via a series of bioinformatics analysis software. The results showed that *CaFRKs* were irregularly distributed between the chromosomes in the pepper reference genomes pepper, and tandem duplication was not observed. This result coincided with the distribution of *FRK* genes in cassava and sugarcane (Yao et al. 2017; Chen et al. 2017). The theoretical pI values indicated that all *CaFRKs* were acidic amino acid except for *CaFRK6*. This result was consistent with previous studies about FRK proteins of *Pyrus bretschneideri* (Cao et al. 2018). In this study, the molecular weight of *CaFRKs* varied dramatically between 20.59 and 75.59 kDa. Whereas, previous studies showed that the molecular weights of *FRK1* were 119 kDa in potato tuber

(Renz and Stitt 1993), 73 kDa in barley leaves (Baysdorfer et al. 1989), 84 kDa in soybean nodules (Copeland and Morell 1985), and 37 kDa in tomato (Martinez-Barajas and Randall 1996) via gel filtration chromatography, respectively. And Doehlert (1989) reported a FRK protein with a natural molecular weight of 59 kDa. All of the above evidences illustrated that the molecular weights of *FRKs* were diverse from different species to family members of the same species. In addition, subcellular location could help to forecast the potential function of gene family. In previous study, three out of four *SlFRKs* of tomato have been located in cytoplasm (Damari-Weissler et al. 2006). In this study, all *CaFRKs* were predicted to localize in the cytoplasm, which was similar to the results in tomato.

Through phylogenetic tree analysis, FRK proteins from the four species, including *Capsicum annuum* L., *Arabidopsis thaliana*, *Solanum lycopersicum* and *Oryza sativa*, were divided into four groups. Based on the clustering of the four groups, it can be concluded that pepper, tomato and *Arabidopsis* clustered together, while most of the rice was distantly related to peppers. This may be due to the fact that the tomato, pepper and *Arabidopsis* are dicotyledonous plants, while the rice is monocotyledonous. The monocotyledons and dicotyledons will cluster on their neighbor branches separately. It is speculated that *FRK* genes of dicotyledonous and monocotyledonous plants may have undergone different evolutionary processes. After years of genome evolution, gene differentiation and duplication may have occurred, and different species evolve at different rates, which makes different copies of new genes. In addition, collinearity analysis revealed that two pairs of colinear *FRK* genes (*CaFRK1*/*AtFRK3*, *CaFRK2* /*AtFRK1*) were found in *Arabidopsis* and pepper, indicating that FRK proteins from different species may have similar functions. The process of species evolution may also be accompanied by vast gain and loss events of motif or structural domain. As the main component of protein and the crucial element of functional differentiation, the functional domain was necessary for protein function and interactions (Bagowski et al. 2010). In this study, we found that CaFRK proteins of the same group had semblable conserved motifs distribution, which was coincident to the results of phylogenetic analysis. On the contrary, the differences in the number and type of motifs among different groups suggested the diversity of protein functions. Besides, the analysis of gene structure also showed that the length and intron/exon configurations of *CaFRKs* family were different, which indicated that the RNA splicing and gene function of different *CaFRK* family members were diverse.

The *cis*-elements of promoters are closely related to underlying gene function and regulatory mechanisms. The promoters region of *CaFRKs* contained various *cis*-elements that associated with hormone regulation, stress response, light response, growth and development, which suggests that *CaFRKs* may have a wide range of biological activities and may be involved in a variety of hormone and stress responses. Under PEG treatment, the expression level of *FRK1* was suppressed in three *Saccharum* species, containing *S. robustum*, *S. officinarum* and hybrid cultivar (ROC-22), but it was induced in *S. spontaneum*, whereas *FRK3* and *FRK5* were dramatically induced in *Saccharum* (Chen et al. 2017). *FRK* was up-regulated in response to drought stress in sunflower, therefore, FRK was a candidate gene for physiological and molecular studies on drought tolerance of sunflower (Fulda et al. 2011). In maize, short-term salt stress leads to up-regulation of *FRK2* which could be regarded as an early marker of salt stress (Zoerb et al. 2010). The results of this study showed that the expression levels of all *CaFRK* genes were up-regulated, except *CaFRK7* which was significantly down-regulated under PEG6000 treatment. Under NaCl treatment, the expression level of *CaFRK7* was also decreased, while *CaFRK2* was significantly increased. Plant growth and development depend on the regulation of gene expression mediated by plant hormones. With the extension of ABA treatment time, the expression level of *SsFRK1* tended to increase, while that of *SsFRK5* decreased. Under Et (Ethylene) and GA treatment, the expression of *FRK5* in most detected tissues was inhibited in *Saccharum*. The expression level of *FRKs* was not markedly affected by the IAA treatment in *Saccharum*, whereas *FRK1* in the stems of *S. officinarum* increased more than 5 times after 24 h of IAA treatment (Chen et al. 2017). In this study, we analyzed the relative expression of *CaFRKs* in pepper plants with IAA, ABA, GA_3, SA and MeJA. *CaFRK* genes expression levels were altered in response to diverse hormone treatment. Especially, the relative expressions of all *CaFRKs* showed a trend of obvious down-regulation under MeJA treatment. But the expression of minority *CaFRK* genes show no visible change under hormone treatments, such as *CaFRK3* after SA and GA_3 treatments, suggesting that these genes do not play an important role in hormone regulation. In contrast, the relative expression of some genes increased sharply under hormone treatments, for instance, *CaFRK7* was up-regulated 14.5- and 50-fold when treated with ABA and SA for 12 h and 6 h, respectively.

In our previous study, *Capana00g002348* in flower buds was expressed significantly higher in restorer accessions than cytoplasmic male sterile accession through RNA sequencing, which suggested that *Capana00g002348* presumably involved in the pollen development or the fertility restorer of CMS in pepper (Wei et al. 2019). In this study, *Capana00g002348* was named *CaFRK9* through the identification of *CaFRK* gene family. Gene expression analysis could provide the first direct evidence for studying the certain biological function of genes. The analysis

of tissue-specific expression of *CaFRKs* by qRT-PCR showed that *CaFRKs* had different levels of expression in the ten tissues. Previous studies have shown that *SlFRK1-3* are generally expressed in most tissues (German et al. 2002; Fei et al. 2004), and *MeFRK6* was expressed specially at very low levels in leaves, *MeFRK1-4* were widely expressed in flowers, leaves, stems, fruits and tubers, and the expression of *MeFRK2* was relatively low. *MeFRK3* and *MeFRK4* were highly expressed at the early stage of cassava tuber development, and were associated with high levels of FRK activity (Yao et al. 2017). Chen et al. (2015) and Cao et al. (2018) discovered that the expression levels of *SsFRK2*, *SsFRK7* and *PbFRK01* were also low in stems and leaves, indicating that these genes may not play the main role in fructose metabolism. Remarkably, some *FRK* genes were found to express specifically in flowers. For example, *SlFRK4* was specifically expressed in stamens (German et al. 2002; Fei et al. 2004) and *MeFRK5* was specially observed in flowers (Yao et al. 2017). Additionally, the analysis of different activities of *SlFRK4* promoter in anther development found that *SlFRK4* promoter was gradually activated in pollen grains throughout late anther development and pollen germination (David-Schwartz et al. 2013). In this study, *CaFRK9* was specifically expressed in flowers, compared with other tissues, so do *SlFRK4* and *MeFRK5*. The strong specifical expression of *CaFRK9* in flowers may hint toward a specific role of *CaFRK9* in pollen germination, anther development, anthesis, and perhaps pollination. To further explore the flower-specific expression pattern of *CaFRK9*, the flower buds at different development stages were used for relative expression analysis. It is worth mentioning that the relative expression of *CaFRK9* gradually increased with the development of the flower buds of fertile accession 8B, R1 and F_1. However, there is almost no expression in all stages of cytoplasmic male sterile lines except for a small amount of expression in the stage III and IV. These provided strong evidence that some *FRK* genes were involved in the pollen development and fertility.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s13205-022-03196-1.

Author contributions SZ and BW conceived and designed the experiments. SZ wrote the manuscript. BG and YW performed the qRT-PCR experiments. NY and PD and analyzed gene expression. MW and GZ shared their expertise in editing and revising the content of the manuscript. BW critically revised the manuscript. The manuscript was read and approved by all the authors.

Funding The work was supported by the National Natural Sciences Foundation of China (31760572) and the scientific research start-up funds for openly recruited doctors (GAU-KYQD-2018-17) of Science and Technology Innovation Funds of Gansu Agricultural University, China.

Declarations

Conflict of interest The authors declare that they have no conflict of interest.

References

Alexandrov NN, Troukhan ME, Brover VV et al (2006) Features of Arabidopsis genes and genome discovered using full-length cDNAs. Plant Mol Biol 60(1):69–85

Ashwell G (1964) Carbohydrate metabolism. Annu Rev Biochem 33(1):101–138

Bagowski CP, Bruins W, Te Velthuis AJW (2010) The nature of protein domain evolution: shaping the interaction network. Curr Genomics 11(5):368–376

Bailey TL, Williams N, Misleh C (2006) MEME: discovering and analyzing DNA and protein sequence motifs. Nucleic Acids Res 34:W369–W373

Bailey TL, Boden M, Buske FA et al (2009) MEME suite: tools for motif discovery and searching. Nucleic Acids Res 37:W202–W208

Baysdorfer C, Kremer DF, Sicher RC (1989) Partial purification and characterization of fructokinase activity from barley leaves. J Plant Physiol 134(2):156–161

Cao Y, Li S, Han Y et al (2018) A new insight into the evolution and functional divergence of FRK genes in *Pyrus bretschneideri*. R Soc Open Sci 5(7):171463

Chase CD (2007) Cytoplasmic male sterility: a window to the world of plant mitochondrial-nuclear interactions. Trends Genet 23(2):81–90

Chen C, Chen G, Cao B et al (2015) Transcriptional profiling analysis of genic male sterile–fertile Capsicum annuum reveal candidate genes for pollen development and maturation by RNA-Seq technology. Plant Cell Tissue Organ Cult 122(2):465–476

Chen Y, Zhang Q, Hu W et al (2017) Evolution and expression of the fructokinase gene family in *Saccharum*. BMC Genomics 18(1):197

Chen C, Chen H, Zhang Y et al (2020) TBtools: an integrative toolkit developed for interactive analyses of big biological data. Mol Plant 13(8):1194–1202

Chou KC, Shen HB (2010) A new method for predicting the subcellular localization of eukaryotic proteins with both single and multiple sites: Euk-mPLoc 2.0. PLoS ONE 5(4):e9931

Clément C, Audran JC (1995) Anther wall layers control pollen sugar nutrition in Lilium. Protoplasma 187(1):172–181

Clement C, Burrus M (1996) Floral organ growth and carbohydrate content during pollen. Am J Bot 83(4):459

Copeland L, Morell M (1985) Hexose kinases from the plant cytosolic fraction of soybean nodules. Plant Physiol 79(1):114–117

Copeland L, Tanner GJ (1988) Hexose kinases of avocado. Physiol Plant 74(3):531–536

Copeland L, Stone SR, Turner JF (1984) Kinetic studies of fructokinase I of pea seeds. Arch Biochem Biophys 233(2):748–760

Damari-Weissler H, Kandel-Kfir M, Gidoni D et al (2006) Evidence for intracellular spatial separation of hexokinases and fructokinases in tomato plants. Planta 224(6):1495–1502

David-Schwartz R, Weintraub L, Vidavski R et al (2013) The SlFRK4 promoter is active only during late stages of pollen and anther development. Plant Sci 199:61–70

Doehlert DC (1989) Separation and characterization of four hexose kinases from developing maize kernels. Plant Physiol 89(4):1042–1048

Fei Z, Tang X, Alba RM et al (2004) Comprehensive EST analysis of tomato and comparative genomics of fruit ripening. Plant J 10(1):47–59

Feng K, Jb Yu, Cheng Y et al (2016) The SOD gene family in tomato: identification, phylogenetic relationships, and expression patterns. Front Plant Sci 7:1279

Fennington GJ, Hughes TA (1996) The fructokinase from *Rhizobium* leguminosarum biovar trifolii belongs to group I fructokinase enzymes and is encoded separately from other carbohydrate metabolism enzymes. Microbiology (reading, Engl) 142(2):321–330

Finn RD, Clements J, Eddy SR (2011) HMMER web server: interactive sequence similarity searching. Nucleic Acids Res 39(Suppl 2):W29–W37

Franck C, Nsthan H, Heather AR et al (1995) Partial purification and characterization of fructokinase from developing taproots of sugar beet (*Beta vulgaris*). Plant Sci 110(2):181–186

Fulda S, Horn R, Stegmann H et al (2011) Physiology and proteomics of drought stress acclimation in sunflower (*Helianthus annuus* L). Plant Biol 13(4):632–642

Gasteiger E, Hoogland C, Gattiker A et al (2005) Protein identification and analysis tools on the ExPASy server. The proteomics protocols handbook. Humana Press, pp 571–607

German MA, Dai N, Chmelnitsky I et al (2002) LeFRK4, a novel tomato (*Lycopersicon esculentum* Mill) fructokinase specifically expressed in stamens. Plant Sci 163(3):607–613

German MA, Asher I, Petreikov M et al (2004) Cloning, expression and characterization of LeFRK3, the fourth tomato (*Lycopersicon esculentum* Mill.) gene encoding fructokinase. Plant Sci 166(2):285–291

Glover J, Grelon M, Craig S et al (1998) Cloning and characterization of MS5 from Arabidopsis: a gene critical in male meiosis. Plant J 15(3):345–356

Goetz M, Godt DE, Guivarc'h A et al (2001) Induction of male sterility in plants by metabolic engineering of the carbohydrate supply. Proc Natl Acad Sci USA 98(11):6522–6527

Gonzali S, Pistelli L, De Bellis L et al (2001) Characterization of two *Arabidopsis thaliana* fructokinases. Plant Sci 160(6):1107–1114

Guo J, Wang P, Cheng Q et al (2017) Proteomic analysis reveals strong mitochondrial involvement in cytoplasmic male sterility of pepper (*Capsicum annuum* L.). J Proteomics 168:15–27

Hanson MR, Bentolila S (2004) Interactions of mitochondrial and nuclear genes that affect male gametophyte development. Plant Cell 16(suppl 1):S154-169

He Z, Zhang H, Gao S et al (2016) Evolview v2: an online visualization and management tool for customized and annotated phylogenetic trees. Nucleic Acids Res W1:W236–W241

Higgins DG, Thompson JD, Gibson TJ (1996) Using CLUSTAL for multiple sequence alignments. Methods Enzymol 266:383–402

Hirsche J, Engelke T, Vller D et al (2009) Interspecies compatibility of the anther specific cell wall invertase promoters from Arabidopsis and tobacco for generating male sterile plants. Theor Appl Genet 118(2):235–245

Hoepfner SW, Botha FC (2004) Purification and characterization of fructokinase from the culm of sugarcane. Plant Sci 167(3):645–654

Hong ST, Chung JE, An G et al (1998) Analysis of 176 expressed sequence tags generated from cDNA clones of hot pepper by single-pass sequencing. J Plant Biol 41(2):116–124

Hu B, Jin J, Guo AY (2015) GSDS 2.0: an upgraded gene feature visualization server. Bioinformatics 31(8):1296–1297

Jiang H, Dian W, Liu F et al (2003) Isolation and characterization of two fructokinase cDNA clones from rice. Phytochemistry 62(1):47–52

Kanayama Y, Dai N, Granot D et al (1997) Divergent fructokinase genes are differentially expressed in tomato. Plant Physiol 113(4):1379–1384

Kaplan CP, Tugal HB, Baker A (1997) Isolation of a cDNA encoding an Arabidopsis galactokinase by functional expression in yeast. Plant Mol Biol 34(3):497–506

Kumar S, Stecher G, Li M et al (2018) MEGA X: molecular evolutionary genetics analysis across computing platforms. Mol Biol Evol 18(6):1547–1549

Kuo TM, Doehlert DC, Crawford CG (1990) Sugar metabolism in germinating soybean seeds: evidence for the sorbitol pathway in soybean axes. Plant Physiol 93(4):1514–1520

Lescot M, Dehais P, Thijs G (2001) PlantCARE, a database of plant cis-acting regulatory elements and a portal to tools for in silico analysis of promoter sequences. Nucleic Acids Res 30(1):325–327

Livak K, Schmittgen T (2000) Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta Ct}$ method. Methods 25(4):402–408

Martinez-Barajas E, Randall DD (1996) Purification and characterization of fructokinase from developing tomato (*Lycopersicon esculentum*) fruits. Planta 199(3):451–458

McCormick S (2004) Control of male gametophyte development. Plant Cell 16(Suppl 1):S142–S153

Nakamura N, Shimizu M, Suzuki H (1991) Characterization of hexose kinases from camellia and lily pollen grains. Physiol Plant 81(2):215–220

Pino J, Gonzalez M, Ceballos L et al (2006) Characterization of total capsaicinoids, colour and volatile compounds of Habanero chilli pepper (*Capsicum chinense* Jack) cultivars grown in Yucatan. Food Chem 104(4):1682–1686

Qin QP, Zhang SL, Chen JW et al (2004) Isolation and expression analysis of fructokinase genes from citrus. Acta Botanica Sinica 46(12):1408–1415

Ralph S, Chun HJ, Cooper D et al (2008a) Analysis of 4,664 high-quality sequence-finished poplar full-length cDNA clones and their utility for the discovery of genes responding to insect feeding. BMC Genomics 9(1):57

Ralph S, Chun H, Kolosova N et al (2008b) A conifer genomics resource of 200,000 spruce (*Picea* spp) ESTs and 6464 high-quality, sequence-finished full-length cDNAs for Sitka spruce (*Picea sitchensis*). BMC Genomics 9(1):484

Renz A, Stitt M (1993) Substrate specificity and product inhibition of different forms of fructokinases and hexokinases in developing potato tubers. Planta 190(2):166–175

Riggs JW, Cavales PC, Chapiro SM et al (2017) Identification and biochemical characterization of the fructokinase gene family in *Arabidopsis thaliana*. BMC Plant Biol 17(1):83

Rolland F, Baena-Gonzalez E, Sheen J (2006) Sugar sensing and signaling in plants: conserved and novel mechanisms. Annu Rev Plant Biol 57(1):675–709

Saitou N (1987) The neighbor-joining method. A new method for reconstructing phylogenetic tree. Mol Biol Evol 4(4):406–425

Sanderson MJ (1989) Confidence limits on phylogenies: the bootstrap revisited. Cladistics 5(2):113–129

Sara EG, Jaina M, Alex B et al (2019) The Pfam protein families database in 2019. Nucleic Acids Res 47(D1):D427–D432

Schaffer AA, Petreikov M (1997) Sucrose-to-starch metabolism in tomato fruit undergoing transient starch accumulation. Plant Physiol 113(3):739–746

Schnarrenberger C (1990) Characterization and compartmentation, in green leaves, of hexokinases with different specificities for glucose, fructose, and mannose and for nucleoside triphosphates. Planta 181(2):249–255

Shi L, Cao S, Shao J et al (2014) Relationship between sucrose metabolism and anthocyanin biosynthesis during ripening in Chinese bayberry fruit. J Agric Food Chem 62(43):10522–10528

Swamy BN, Hedau NK, Chaudhari GV et al (2017) CMS system and its stimulation in hybrid seed production of *Capsicum annuum* L. Sci Hortic 222:175–179

Taylor MA, Ross HA, Gardner A et al (1995) Characterisation of a cDNA encoding fructokinase from potato (*Solanum tuberosum* L.). J Plant Physiol 145(3):253–256

Truernit E, Stadler R, Baier K et al (1999) A male gametophyte-specific monosaccharide transporter in Arabidopsis. Plant J 17(2):191–201

Voorrips RE (2002) MapChart: software for the graphical presentation of linkage maps and QTLs. J Hered 93(1):77–78

Wei BQ, Wang LL, Bosland PW et al (2019) Comparative transcriptional analysis of Capsicum flower buds between a sterile flower pool and a restorer flower pool provides insight into the regulation of fertility restoration. BMC Genomics 20(1):837

Wei BQ, Wang LL, Bosland PW et al (2020) A joint segregation analysis of the inheritance of fertility restoration for cytoplasmic male sterility in pepper. J Am Soc Hortic Sci 145(1):3–11

Yang WQ, Lai Y, Li MN et al (2008) A novel C2-domain phospholipid-binding protein, OsPBP.1 is required for pollen fertility in rice. Mol Plant 1(5):770–785

Yang X, Ye J, Zhang L et al (2020) Blocked synthesis of sporopollenin and jasmonic acid leads to pollen wall defects and anther indehiscence in genic male sterile wheat line 4110S at high temperatures. Funct Integr Genomics 20(3):383–396

Yao Y, Geng MT, Wu XH et al (2017) Identification, expression, and functional analysis of the fructokinase gene family in Cassava. Int J Mol Sci 18(11):2398

Ye XY, Zhou WB (2021) Research advances in plant fructokinases. Chin Sci Bull 66(22):2820–2831 (**in Chinese**)

Zhang S, Nichols SE, Dong JG (2003) Cloning and characterization of two fructokinases from maize. Plant Sci 165(5):1051–1058

Zhang X, Chen B, Zhang L et al (2015) Identification of proteins associated with cytoplasmic male sterility in pepper (*Capsicum annuum* L.). S Afr J Bot 100:1–6

Zoerb C, Schmitt S, Muehling KH (2010) Proteomic changes in maize roots after short-term adjustment to saline growth conditions. Proteomics 10(24):4441–4444