Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Transcriptome-wide characterization of novel WRKY family genes Engaged in Crocin biosynthesis in Crocus Sativus

Guangchun Gao
Jiaxing University
Jun Li
lijun@jxvtc.edu.cn
Jiaxing Vocational and Technical College https://orcid.org/0000-0002-3326-4577
Ping Wang
Zhejiang University of Technology
Jiming Wu
Jiaxing University
Jinbiao Xu
Jiaxing University
Lisa Mao
Jiaxing University

Research Article

Keywords: Crocus sativus, RNA-seq, WRKY transcription factors, crocin biosynthesis

Posted Date: May 10th, 2024

DOI: https://doi.org/10.21203/rs.3.rs-4335844/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Abstract

Crocins, exhibiting remarkable pharmacological value were generated in significant quantities in stigma of *Crocus sativus*. Although the biosynthesis pathway of crocins has been elucidated to a great degree, there is still little information about the regulating mechanism of TFs on the biosynthesis of crocins in stigma development process. WRKY TFs were reported to play a role in modulating carotenoids/apocarotenoids metabolism. CsWRKY genes were identified from RNA sequencing database of stigma at different developmental stages. Phylogenetic analysis was employed to work out their evolutionary relation. Coexpression profile of CsWRKY genes and crocin biosynthesis-related genes was performed on Majorbio Cloud Platform. Quantitative real-time PCR was used to investigate the expression level of CsWRKY TFs in various tissue and developmental stages. A total of 34 CsWRKY TFs were identified from stigma of *C. Sativus*. Phylogenetic analysis of these CsWRKY TFs together with their orthologs from Arabidopsis clustered them into group Ⅰ, Ⅱ and Ⅲ. Coexpression network between CsWRKY TFs and crocin biosynthesis-related genes (CsBCH, CsCCD2L, CsALDH and CsUGT) revealed that CsWRKY1, -2, -8, -10, -15 and − 28 genes showed strong correlation with above structure genes. CsWRKY2, -15 and − 28 have identical motifs, belonging to group Ⅱd. The expression levels of candidate CsWRKY genes are highest in stigma comparing with other tissues. Furtherly, the expression patterns of candidate CsWRKY genes were in parallel to the accumulation of crocins. Our study established an extensive overview of the WRKY family in *C. sativus* and screened six candidate CsWRKY (1, -2, -8, -10, -15 and − 28) involved in the crocin biosynthesis in saffron.

Introduction

Crocus sativus L. (Iridaceae), commonly known as saffron, is a sterile triploid plant propagated vegetatively through corms [1]. The dried red stigma of *C. sativus* is considered the world's most expensive spice due to its diversity medicinal values for human health [2–4]. The apocarotenoids crocins, picrocin and safranal which derive from crocin biosynthesis pathway are considered to be major valuable metabolites in stigma of *C. sativus* [5]. The apocarotenoids were accumulated in stigma of the flower at specific developmental stages. Over the past decade, the biosynthesis pathway of crocins, the characteristic apocarotenoids in stigma of *C. sativus*, were almost elucidated [6]. For example, the apocarotenoids are originated from the cleavage of zeaxanthin at 7,8:7',8' position, in a reaction catalyzed by a specific carotenoid cleavage dioxygenase enzyme(CsCCD2L) [7]. As result of this cleavage reaction, crocetin dialdehyde is created and further catalyzed by CsALDH to produce crocetin which is the crucial precursor of crocins [8]. Finally, the glycosylation of crocetin under the catalysis of various UDP-glucosyltransferase (CsUGT) develops diverse crocins [9]. However, so far, the mechanism that regulates the accumulation of crocins in tissue and developmental stage-specific manner in *C. sativus* still remains tricky with only a few reports [10–12].

In our previous work, we obtained diverse transcription factor families including WRKY, bHLH, MYB, AP2/ERF, and so on, from stigma of *C. sativus* by transcriptome analysis [13]. WRKY TFs are among the largest families of transcriptional regulator in plants, which play crucial roles in regulation of secondary

metabolites biosynthesis, defense signaling, plant growth and development and responses to biotic and abiotic factors [14–16]. All known WRKY proteins are characterized by one or two WRKY domain. This domain is approximately 60 conserved amino acid residues containing the heptapeptide WRKYGQK in the N-terminal region and C2H2 or C2HC zinc-finger motif at the C-terminal region. The WRKY proteins display regulatory function by binding to cis-regulatory elements located in the 5' upstream region of the target gene in a sequence-specific manner [17]. In recent years, there are many reports where WRKY TFs illustrated regulation function in biosynthesis of varied secondary metabolites such as flavonoid [18], alkaloid [14, 19], etc. Recently, increasing evidence in other species has demonstrated that WRKY TFs are involved in the regulation of carotenoids/apocarotenoids metabolism as well by modulating the expression of functional genes in their metabolic pathways [20]. For instance, *SlWRKY35* can regulate carotenoid biosynthesis by directly activating the expression of SlDXS1gene in tomato. Meanwhile coexpression of *SlWRKY35* and *SlLCYE* can promote lutein production in transgenic tomato [21]. In Osmanthus fragrans, *OfWRKY3* was characterized as a positive regulator of *OfCCD4* gene [22]. Another WRKY TF, *CrWRKY42* in citrus was found to accelerate the accumulation of carotenoids by directly binding to the promoters of *CrBCH1*, *CrPDS* and *CrLCYB2* genes and activating their expression [23]. These work advocate possibility of involvement of WRKY genes in regulating carotenoid/apocarotenoid biosynthesis in *C. sativus*.

Present research aimed to demonstrate an extensive overview of the WRKY family transcription factors in stigma of *C. sativus* and provide candidate WRKY regulatory genes involved in crocin biosynthesis. We characterized 34 CsWRKY genes according to RNA-seq data from *C. sativus* [13]. Subsequently conserved domains and motifs were predicted using online software. Further Phylogenetic and MSA analysis were performed to elucidate the functional and evolutionary aspects. To find out the regulatory function of CsWRKY with crocin biosynthesis-related genes, we analyzed coexpression correlation of these genes on Majorbio Cloud Platform (https://cloud.majorbio.com/page/tools/). In addition, the expression of candidate CsWRKYs genes in various tissues as well as different developmental stages were determined by qRT-PCR. Our research provides an insight into the understanding of CsWRKYs in stigma of *C. sativus* and obtains several candidate regulators of Crocin biosynthesis.

Materials and Methods

Identification and sequence annotation of WRKY genes

Saffron was planted at Jiaxing city, Zhejiang Province, China. Stigmas at different developmental stages (-3da, -2da, da, +1da, +2da) were collected and sequenced as our previous study [13]. In brief, RNA sequencing was performed by Shanghai Majorbio Co., Ltd. (Shanghai, China) using the Illumina Hiseq™ 2000 system. Transcriptome-wide identification of WRKY genes were screened from *C. sativus* stigmas RNA sequencing data. All the sequences were revalidated using Uniprot protein database (https://www.uniprot.org/) and Pfam (https://pfam.xfam.org/), re-examined for the presence of WRKY domains using conserved domain database (https://www.ncbi.nlm.nih.gov/cdd/) and through HMMScan (https://www.ebi.ac.uk/Tools/hmmer/search/hmmscan). The molecular weight (MW), Theoretical

isoelectric point (pI), instability index, aliphatic index, Grand average of hydropathicity (GRAVY) of CsWRKY proteins were predicted via the ProtParam (http://web.expasy.org/protparam/). Additionally, subcellular localization was also predicted by an advanced protein subcellular localization prediction tool WoLFPSORT (https://wolfpsort.hgc.jp/).

Investigations of conserved domain and motif of CsWRKYs

The conserved domains of 34 CsWRKYs were detected using NCBI CD database. MEME suite was employed to analyze the conserved motifs of 34 CsWRKYs (http://meme-suite.org/tools/meme).

Multiple sequence alignment, phylogenetic analysis and classification of CsWRKYs

The multiple sequence alignment (MSA) of 106 WRKY proteins was performed using 34 WRKY proteins of *Crocus Sativus* (CsWRKY), 72 from A. thaliana (AtWRKYs). The protein sequences of Arabidopsis were downloaded from TAIR (http://www.Arabidopsis.org/). The conserved regions of 34 amino acids for the WRKY proteins were searched using HMMScan and aligned using CLUSTALW for the construction of the phylogenetic tree. For the CsWRKY based phylogenetic tree, complete protein sequences were used. The tree was constructed using MEGA 7.0 with neighbor-joining method using JTT substitution model and pair-wise deletion method with 1000 bootstrap value. The 34 amino acid conserved region of MSA of *Crocus Sativus* was visualized using DNAMAN. The MSA included conserved region of WRKY members representing each group and subgroup from A. thaliana as reference.

Co-expression analysis of CsWRKYs and crocin biosynthesis-related genes in stigma developmental stages

Eight crocin biosynthesis-related genes (CsCCD2L, CsBCH, CsUGT74AD1, CsUGT709G1, CsALDH3I1, CsALDH3898, CsALDH20158 and CsALDH54788) were obtained from papers reported previously [2, 7, 24]. The whole CDS of these crocin biosynthesis-related genes CsCCD2L (GenBank:KP887110.1), CsBCH (GenBank: AJ416711.2), CsUGT74AD1 (GenBank:MF596166.1), CsUGT709G1 (GenBank:KX385186.1), CsALDH3I1 (GenBank:MF596165.1), CsALDH3898 (GenBank:KU577905.2), CsALDH20158 (GenBank:KU577906.2) and CsALDH54788 (GenBank:KU577904.2) were obtained from NCBI (https://www.ncbi.nlm.nih.gov/). The RNA sequencing unigenes, which were annotated as WRKY genes in NR database, were blasted with the download whole CDS respectively. The 34 unigenes matching with the whole CDS were used to further analysis. For the coexpression network, we used the FPKM of the 34 CsWRKY and 8 crocin biosynthesis-related genes from the stigma developmental stages with a spearman correlation coefficient. The correlation matrix heatmap was conducted on the online tool of Majorbio Cloud Platform (https://cloud.majorbio.com/page/tools/). The co-expression network analysis was then done on Majorbio Cloud Platform with a spearman correlation coefficient cutoff of ≥ 0.7. The interaction network was furtherly visualized using Cytoscape 3.9.1.

RNA extraction and quantitative real-time reverse transcription PCR (qRT-PCR).

Total RNA of *C. sativus* stigmas from different developmental stages (-3da, -2da, da, +1da, +2da) and different tissues (petal, leaf, corn, stamen and stigma) were extracted using TRIzol reagent (TIANGEN, Shanghai, China), and the quality and quantity of the total RNA were determined using a NanoDrop 2000C spectrophotometer (Thermo Scientific, USA). Total RNA was used to generate complementary DNA using M-MLV Reverse Transcriptase (TaKaRa Bio, Beijing, Chian) following the manufacturer's instructions. qRT-PCRs were performed using the SYBR Premix Ex Taq™ Kit (Takara, Japan) and then conducted on an ABI 7300 Real-Time PCR System (Applied Biosystems, Foster City, CA, USA), each tissue sample were carried out in triplicate. Gene-specific RT primers were designed using Primer premier 6 according to the full-length base sequence from RNA-seq data. The amplicons size ranged between 150 to 250 bp. Actin gene was selected as an internal reference gene. All the PCRs were performed under following conditions: 95°C for 30s, 95°C for 5s, 55°C for 30s, and 72°C for 30s for a total of 40 cycles. The relative mRNA expression levels of CsWRKY genes were calculated using the $2^{-\Delta\Delta Ct}$ method. The expression level of key CsWRKY was showed in heatmap done by GraphPad Prism 8. To compare the expression profiles of multiple experiment designs with each other, Tukey's multiple comparisons test was performed using GraphPad Prism 8 (*, $P < 0.05$; **, $P < 0.01$; ***, $P < 0.001$).

Results

Transcriptome-wide analysis and characterization of Crocus Sativus WRKY TFs

We have performed the transcriptomics analysis of stigma of *Crocus Sativus* and mined the data for the WRKY transcription factor. Among the 63 sequences that were annotated and revalidated to WRKY genes, 34 CsWRKYs had complete CDS, and 30 gene sequences were partial (Table 1). The CsWRKYs sequences were submitted to NCBI, and their accession numbers are given in (Table 1, Supplementary File S1). The sizes of the CsWRKY proteins varied dramatically. The CDS length of the 34 CsWRKY genes ranged from 519 to 2088 bp, and the amino acid length ranged from 172 to 695 aa (Table 1). Out of these, CsWRKY23 was the longest, while the shortest was CsWRKY29, and their molecular weight (MW) varied between 19099.46 Da (CsWRKY29) to 75517.75 Da (CsWRKY23) (Table 2). The isoelectric point (pI) of 17 CsWRKYs were acidic, and the remaining 17 were basic proteins. In the present study all of the CsWRKYs were found to be unstable, having maximum instability index of 69.46 (CsWRKY17) with the minimum of 41.53 (CsWRKY12) (Table 2), based on the instability index which is believed as unstable if it is higher than 40.0 [25]. Additionally, the WoLFPSORT prediction showed that 33 CsWRKY proteins were localized in nucleus, suggesting that they play regulatory role predominantly in cell nucleus, while CsWRKY 11 had cytoplasmic subcellular localization (Table 2). 32 CsWRKYs had a distinctive hepta-peptide DNA binding sequence (WRKYGQK), the identifying character of the WRKY family, with exception of three CsWRKYs, namely CsWRKY6 and 34 substituted with WRKYGEK domain, and CsWRKY11 substituted with WRKYGKK domain (Table 1).

Table 1
Sequence features of WRKY genes in stigma of *Crocus Sativus*.

CsWRKYs	Gene ID	CDS (bp)	ORF (aa)	Group	Conserved motif	Domain pattern	Zinc finger
CsWRKY1	OP375600	1521	506	1	2(WRKYGQK)	C-X_4-C-X_{22}-HNH	C_2H_2
CsWRKY2	OP375601	1080	359	2d	WRKYGQK	C-X_5-C-X_{23}-HNH	C_2H_2
CsWRKY3	OP375602	888	295	2d	WRKYGQK	C-X_5-C-X_{23}-HRH	C_2H_2
CsWRKY4	OP375603	900	299	2a	WRKYGQK	C-X_5-C-X_{23}-HNH	C_2H_2
CsWRKY5	OP375604	1062	353	1	WKKYGQK/ WRKYGQK	C-X_4-C-X_{22}-HNH	C_2H_2
CsWRKY6	OP375605	522	173	3	WRKYGEK	C-X_7-C-X_{24}-HSC	C_2HC
CsWRKY7	OP375606	1896	631	1	2(WRKYGQK)	C-X_4-C-X_{22}-HDH/ C-X_4-C-X_{23}-HNH	C_2H_2
CsWRKY8	OP375607	864	287	2c	WRKYGQK	C-X_4-C-X_{23}-HTH	C_2H_2
CsWRKY9	OP375608	627	208	2a	WRKYGQK	C-X_4-C-X_{23}-HNH	C_2H_2
CsWRKY10	OP375609	1542	513	1	2(WRKYGQK)	C-X_4-C-X_{22}-HSH/ C-X_4-C-X_{23}-HNH	C_2H_2
CsWRKY11	OP375610	549	182	2c	WRKYGKK	C-X_4-C-X_{23}-HNH	C_2H_2
CsWRKY12	OP375611	1530	509	2b	WRKYGQK	C-X_5-C-X_{23}-HNH	C_2H_2
CsWRKY13	OP375612	1428	475	1	2(WRKYGQK)	C-X_4-C-X_{22}-HNH/ C-X_4-C-X_{23}-HNH	C_2H_2
CsWRKY14	OP375613	579	192	2e	WRKYGQK	C-X_5-C-X_{23}-HNH	C_2H_2

CsWRKYs	Gene ID	CDS (bp)	ORF (aa)	Group	Conserved motif	Domain pattern	Zinc finger
CsWRKY15	OP375614	897	298	2d	WRKYGQK	C-X_5-C-X_{23}-HRH	C_2H_2
CsWRKY16	OP375615	822	273	3	WRKYGQK	C-X_7-C-X_{22}-HTC	C_2HC
CsWRKY17	OP375616	567	188	2c	WRKYGQK	C-X_4-C-X_{23}-HTH	C_2H_2
CsWRKY18	OP375617	546	181	2c	WRKYGQK	C-X_4-C-X_{23}-HTH	C_2H_2
CsWRKY19	OP375618	1131	376	1	2(WRKYGQK)	C-X_4-C-X_{22}-HNH/ C-X_4-C-X_{22}-HSH	C_2H_2
CsWRKY20	OP375619	1080	359	2d	WRKYGQK	C-X_5-C-X_{23}-HNH	C_2H_2
CsWRKY21	OP375620	2085	694	1	2(WRKYGQK)	C-X_4-C-X_{22}-HNH/ C-X_4-C-X_{23}-HNH	C_2H_2
CsWRKY22	OP375621	1416	471	1	2(WRKYGQK)	C-X_4-C-X_{22}-HNH/ C-X_4-C-X_{23}-HNH	C_2H_2
CsWRKY23	OP375622	2088	695	1	2(WRKYGQK)	C-X_4-C-X_{22}-HNH/ C-X_4-C-X_{23}-HNH	C_2H_2
CsWRKY24	OP375623	2031	676	1	2(WRKYGQK)	C-X_4-C-X_{22}-HSH/ C-X_4-C-X_{23}-HNH	C_2H_2
CsWRKY25	OP375624	1521	506	1	2(WRKYGQK)	C-X_4-C-X_{22}-HTH/ C-X_4-C-X_{23}-HNH	C_2H_2

CsWRKYs	Gene ID	CDS (bp)	ORF (aa)	Group	Conserved motif	Domain pattern	Zinc finger
CsWRKY26	OP375625	1266	421	1	2(WRKYGQK)	C-X_4-C-X_{22}-HNH/ C-X_4-C-X_{23}-HNH	C_2H_2
CsWRKY27	OP375626	1197	398	1	2(WRKYGQK)	C-X_4-C-X_{22}-HNH/ C-X_4-C-X_{23}-HNH	C_2H_2
CsWRKY28	OP375627	888	295	2d	WRKYGQK	C-X_5-C-X_{23}-HRH	C_2H_2
CsWRKY29	OP375628	519	172	2c	WRNYGQK	C-X_4-C-X_{23}-HNH	C_2H_2
CsWRKY30	OP375629	864	287	2c	WRKYGQK	C-X_4-C-X_{23}-HTH	C_2H_2
CsWRKY31	OP375630	1212	403	1	2(WRKYGQK)	C-X_4-C-X_{22}-HNH/ C-X_4-C-X_{23}-HNH	C_2H_2
CsWRKY32	OP375631	624	207	2c	WRKYGQK	C-X_4-C-X_{22}-HSH	C_2H_2
CsWRKY33	OP375632	948	315	NG	WRKYGQK	Zinc cluster	
CsWRKY34	OP375633	858	285	3	WRKYGEK	C-X_7-C-X_{24}-HSC	C_2HC

Table 2
Physical parameters of CsWRKY genes.

CsWRKYs	pI	Mw (Da)	Instability Index	Aliphatic index	GRAVY	Subcellular calization
CsWRKY1	5.90	54273.61	65.74	59.76	-0.708	Nucleus
CsWRKY2	9.85	40592.10	56.84	58.97	-0.894	Nucleus
CsWRKY3	9.68	32093.69	49.20	66.24	-0.553	Nucleus
CsWRKY4	6.90	32734.59	48.84	62.68	-0.694	Nucleus
CsWRKY5	5.59	39484.71	43.96	60.76	-0.848	Nucleus
CsWRKY6	9.07	20108.68	63.15	53.47	-1.060	Nucleus
CsWRKY7	5.90	68849.82	51.94	58.83	-0.806	Nucleus
CsWRKY8	6.61	31932.89	56.92	62.47	-0.691	Nucleus
CsWRKY9	8.86	23940.25	47.56	68.41	-0.790	Nucleus
CsWRKY10	8.15	55355.56	58.95	61.21	-0.671	Nucleus
CsWRKY11	4.92	20367.40	46.16	64.67	-0.635	Cytoplasm
CsWRKY12	6.44	54628.61	41.53	59.51	-0.649	Nucleus
CsWRKY13	8.32	53101.46	59.29	55.39	-0.963	Nucleus
CsWRKY14	5.27	21567.99	43.91	61.41	-0.816	Nucleus
CsWRKY15	9.91	32466.01	42.94	69.70	-0.543	Nucleus
CsWRKY16	5.60	30098.55	44.00	67.18	-0.594	Nucleus
CsWRKY17	9.71	21838.68	69.46	50.27	-1.056	Nucleus
CsWRKY18	7.82	20969.23	48.61	58.62	-1.038	Nucleus
CsWRKY19	8.86	41366.42	40.80	67.31	-0.740	Nucleus
CsWRKY20	9.85	40592.08	54.90	59.25	-0.885	Nucleus
CsWRKY21	5.73	75290.51	55.70	53.43	-0.770	Nucleus
CsWRKY22	6.06	51381.83	51.56	52.59	-0.828	Nucleus
CsWRKY23	5.89	75517.75	57.61	53.19	-0.774	Nucleus
CsWRKY24	5.77	73013.86	56.00	53.25	-0.727	Nucleus
CsWRKY25	5.71	54313.54	66.10	59.57	-0.705	Nucleus
CsWRKY26	7.90	45870.62	71.17	58.36	-0.776	Nucleus

CsWRKYs	pI	Mw (Da)	Instability Index	Aliphatic index	GRAVY	Subcellular calization
CsWRKY27	8.81	43495.04	48.60	54.40	-0.860	Nucleus
CsWRKY28	9.68	32093.69	49.20	66.24	-0.553	Nucleus
CsWRKY29	6.37	19099.46	46.52	72.50	-0.544	Nucleus
CsWRKY30	6.33	31879.84	55.19	62.47	-0.667	Nucleus
CsWRKY31	7.55	44108.07	52.13	55.71	-0.938	Nucleus
CsWRKY32	5.01	22722.64	53.18	45.70	-0.983	Nucleus
CsWRKY33	10.20	35043.06	53.24	71.62	-0.671	Nucleus
CsWRKY34	9.21	31870.09	54.04	66.07	-0.710	Nucleus

Conserved domain and motif prediction of CsWRKY genes

Conserved domain analysis showed that 10 CsWRKYs possessed additional domains besides the WRKY domain (Fig. 1). For example, CsWRKY1, CsWRKY25 and CsWRKY26 possessed PHA03247 superfamily domain (170 amino acid residues) at the N-terminal, CsWRKY9 had lipocalin FABP super family domain (130 amino acid residues) at the N-terminal while 6 CsWRKYs (-2, -3, -15, -20, -28, -33) had Plant Zn cluster domain(36 amino acid residues) at the C-terminal (Fig. 1). The conserved motifs of CsWRKY proteins were predicted based on MEME online software. In all, fifteen conserved motifs, named motif 1 to motif 15, were elucidated (Fig. 2). The numbers of motifs in CsWRKYs varied from 1 to 7, and the length of motifs ranged from 21 to 107 amino acids. CsWRKY21, -22, -23 and − 24, belonging to group IC, have the largest number of motifs, while CsWRKY6 and − 34 had only one motif (Fig. 2a). Motif 1 formed the main structure of the WRKY domain (WRKYGQK), which is distributed in all of the CSWRKYs proteins. Motif 1 and motif 2 were present in group □C, □C, □a+□b and □+□d. Motif 3 and Motif 5 were only found in group □N. Motif 4 and Motif 9 were only observed in CsWRKY2 and − 20 belonging group □+□d, while Motif 15 was merely detected in CsWRKY8 and − 30, belonging group □C (Fig. 2a and 2b). Based on these results, the same group owned similar motif compositions and arrangement which demonstrated a highly functional conservation.

Phylogenetic analysis and Multiple sequence alignment of the identified WRKY proteins

To investigate the evolutionary relationships of CsWRKY genes, phylogenetic analysis was carried out using 34 CsWRKY proteins and 72 WRKYs identified from Arabidopsis thaliana (Fig. 3). The results showed that all WRKYs could be categorized into three groups (group □, □ and □). Among these CsWRKYs, 15 were clustered in group □, and 13 CsWRKYs were assigned to group □ that was further sub-divided into five sub-groups, IIa + IIb (3), IIc (5), IId + IIe (5). Finally, 4 CsWRKYs were classified into group-III. The multiple sequence alignment of 60 amino acids conserved region of all the 34 CsWRKY proteins were clustered in 8 different groups and sub-groups with very high homology (> 70%) as shown in

(Fig. 4). Group IN displayed conserved motif 1 (DGYNWRKYGQK) and zinc finger pattern of C-X4-C-X23-HXH showing conservancy with 13 CsWRKY proteins. While Group IC had 13 CsWRKYs displayed conserved motif 2 (GPN[N/H/Y]PRSYY) and motif 3 (VRKHVERA), zinc finger pattern of C-X4-C-X23-HXH. The third group IIa had 2 CsWRKY proteins (CsWRKY4 and − 9) displaying motif 4 (KDG[F/Y]QWRKYGQK[V/I]TRDNP), motif 5 (VKKKVQRS) and a zinc finger motif pattern of C-X5-C-X23-HNH. Group IIb had 1 CsWRKY (CsWRKY12) with three conserved motifs, namely motif 6 (DGCQWRKYGQK), motif 7 (PPAYYRC), motif 8 (CPVRKQVQRC) and a zinc finger motif pattern of C-X5-C-X23-HNH, while Group IIc comprised of 5 CsWRKYs (CsWRKY8, -11, -17, -18, -30) with conserved motif 9 (C[N/G/S/T]VKK[Q/R]V[Q/E]R) and a zinc finger motif pattern of C-X4-C-X23-HXH. 5 CsWRKYs (CsWRKY2, -3, -15, -20 and − 28), containing conserved motif 10 (YSWRKYGQKPIKGSP[H/Y]PRGYYKCS), motif 11 (RGCPARKHVER) and a zinc finger motif pattern of C-X5-C-X23-HXH, were categorized to Group IId, While CsWRKY14, clustered in group IIe, showed conserved motif 12 (WRKYGQKPIK[S/G]SPYPR), motif 13 (SKGC[F/L/S]A[R/K]KQV[E/D]R) and a zinc finger motif pattern of C-X5-C-X23-HXH. 3 CsWRKY proteins (CsWRKY6, -16 and − 34) were eventually clustered to group III with a zinc finger motif C-X7-C-X24-HXC and no conserved motif.

Coexpression profile of CsWRKY and crocin biosynthesis-related genes and prediction of key CsWRKYs

WRKY transcription factors have been identified as regulators of secondary metabolites biosynthesis (Yuan et al. 2022). To gain a better understanding of the transcriptional regulation landscape of CsWRKY in crocin biosynthesis, we conducted the coexpression analysis of 34 CsWRKY with eight crocin biosynthesis-related genes (CsCCD2L, CsBCH, CsUGT74AD1, CsUGT709G1, CsALDH3I1, CsALDH3898, CsALDH20158 and CsALDH54788) on the online tool of Majorbio Cloud Platform (https://cloud.majorbio.com/page/tools/). The results were showed in the correlation matrix heatmap (Fig. 5a, Supplementary File S2). CsBCH showed significantly strong positive correlation with CsWRKY2, -28 and − 30($P < 0.001$) and negative correlation with CsWRKY1, -3, -8, -10, -15, -16 and − 19 ($P < 0.001$). Similar strong correlations were also found between CsUGT74AD1 and CsWRKY1, -2, -3, -8, -10, -15, -16, -28, -30, except negative correlation ($P < 0.01$) with CsWRKY5 and positive correlation correlation ($P < 0.001$) with CsWRKY20. CsALDH54788 exhibited remarkable positive correlation ($P < 0.001$) with CsWRKY7, negative correlation ($P < 0.01$) with CsWRKY11. There were no statistically significant correlations found in other crocin biosynthesis-related genes and CsWRKYs.

To seek the key correlation between structural genes and CsWRKYs, the spearman correlation coefficient ≥ 0.7 and the q-value ≤ 0.05 were chosen on the online analysis of Majorbio Cloud Platform (https://report.majorbio.com/) and the coexpression profile was visualized by Cytoscape 3.9.1(Fig. 5b). In total, forty-seven correlations (23 nodes and 47 edges) were screened including eighteen correlations between CsWRKY and crocin biosynthesis-related genes, twenty-eight correlations in CsWRKYs and one correlation in crocin biosynthesis-related genes. CsBCH exhibited nine correlations with CsWRKY1, -2, -3, -5, -10, -15, -16, -28 and − 30. CsCCD2L showed three correlations with CsWRKY8, -15 and − 32, and CsUGT74AD1 also took on correlations with Cs WRKY8 and − 15. Inspiringly, CsWRKY15 was found to

significantly coexpress with CsBCH, CsCCD2L and CsUGT74AD1 with *r*=-0.771, 0.714 and − 0.704 respectively ($P < 0.05$). CsWRKY8 showed remarkably correlations with CsCCD2L ($r = 0.778$) and CsUGT74AD1 (*r*=-0.778) ($P < 0.05$). Four strongest correlations were to be screened between CsBCH and CsWRKY1 (*r*=-0.869), -2 ($r = 0.907$), -10 (*r*=-0.9) and − 28 ($r = 0.869$) at extremely significant level ($p < 0.001$). CsALDH20158, CsALDH3I1 and CsALDH54788 exhibited two independent correlations network with CsWRKY4($r = -0.725$), -13($r = -0.822$), -17 ($r = -0.704$) and − 34 ($r = -0.714$) base on coexpression analysis (Supplementary File S3). Taken together, all these results indicated that CsWRKYs, particularly CsWRKY15, -8, -1, -2, -10 and − 28, can modulate the biosynthesis of crocins by regulating the expression level of CsBCH, CsCCD2L, CsUGT74AD1 and CsALDHs.

Expression profile of key CsWRKY genes in different tissues and stigma developmental stages

In order to understand the roles of CsWRKY in the regulation of the contents of crocins which had strong correlations with crocin biosynthesis-related genes, we studied the expression levels of fifteen key CsWRKYs ($r \geq 0.7$; q-value ≤ 0.05) in different tissues (petal, leaf, corn, stamen and stigma) by quantitative real time PCR. Six CsWRKYs (CsWRKY1, -2, -8, -10, -15 and − 28) showed relatively high expression in stigma comparing with other four tissues. In addition, the expression pattern of other CsWRKYs presented relatively low expression in the stigma. Previous evidence showed that crocins accumulate in stigma in developmental stage specific manner. We furtherly evaluated the expression pattern of high-expressed CsWRKYs (CsWRKY1, -2, -8, -10, -15 and − 28) in stigma in different developmental stages. Six CsWRKYs exhibited increasing expression from stages-3da to da, then subsequently decreased from stage da to + 2da (Fig. 6b). These results indicated that the expression profile of these six key CsWRKYs are consistent with the accumulation of crocins in Crocus stigma, therefore, indicating possible involvement of these CsWRKYs TF in regulating crocins.

Discussion

In plants, secondary metabolites play predominant role in multiple biological functions in the whole life. The biosynthesis of these metabolites is accurately tuned at the spatiotemporal level by means of the transcriptional regulation of functional genes involved in these pathways. This coordinated regulation generally rely on the interplay of DNA-related mechanisms and the action of corresponding transcription factors [26]. The expression of genes involved in metabolic pathways has evolved to become highly correlated temporally and spatially through the process of natural selection. In *C. sativus*, the apocarotenoids such as crocins, picrocrocin and safrananl were predictably accumulated in temporal and spatial manner [10]. The biosynthesis pathway of apocarotenoids has been elucidated to a considerable extent [9, 27–28]. However, there is dearth of information about the regulating mechanism of transcription factors on expression patterns of function genes till now [10–12, 29].

In recent years, WRKY TFs were demonstrated to act in apocarotenoids metabolism in many plants [21, 30–31]. Here, we reported 34 CsWRKY TFs based on transcriptome data developed in our lab [13] in

stigma of *C. sativus* (Table 1). The number of CsWRKYs obtained in this study were smaller than those reported in *C. sativus* (40) treated with MeJA [29], which confirmed that the expression of CsWRKYs varied in various physiological status. Meanwhile these 34CsWRKYs are not the actual number of WRKYs in this crop because our transcriptome data were created from not the whole plant but stigma tissue. According to the sequence similarity and the conserved domains, 34 CsWRKY genes were clustered in group Ⅰ, Ⅱ and Ⅲ (Fig. 3). In accordance with the classification of WRKYs in other plant species, especially Arabidopsis, Glycyrrhiza glabra and Lonicera macranthoides [32, 33], Group Ⅱ was further divided into five subgroups. An evolution analysis in group Ⅱ showed that subgroup IIa and IIb were closely related, and the same went between IId and IIe (Fig. 3), revealing that these two pairs were evolved from a common ancestor separately. Moreover, group IId and IIe illustrated a higher divergence than group IIa and IIb, IIc in the phylogenetic analysis and were closer to group Ⅲ (Fig. 3), and this result is in consistent with several species previously reported, such as Cynanchum thesioides [34] Caragana korshinskii [35] and Taxus chinensis [14]. Most of known WRKY family members are characterized by a 60 highly conserved amino acid sequence with the WRKYGQK motif in the N-terminal. Earlier studies have also reported several variants including WRKYGKK, WKKYRQK [36], WRKYGEK, WRKYEDK, and WKKYCEDK [37]. In this study, 3 out of 34 CsWRKY proteins had WRKYGE/KK domain instead of the characterized WRKYGQK domain, which also occurred in Banana [38] and Arabidopsis [39]. Mutations of the conserved WRKYGQK domain may change flexibility for binding to the W-box element of downstream structure genes, thus inhibiting or disordering DNA binding activities [17](Goyal et al. 2023). In a word, the conserved motif distribution patterns of CsWRKY genes differ among different subgroups, while in the same subgroup the CsWRKY genes have similar conserved motif distribution patterns (Table 1, Fig. 1–4). The extremely conserved CsWRKY TFs may be functionally conserved with other related species, exhibiting similar activities in various physiological processes.

The coexpression profile is a widely used approach for the identification of new metabolic regulators in recent years [21, 40]. Based on correlation analysis of the expression pattern of CsCCD2L, CsBCH, CsUGT74AD1, CsUGT709G1, CsALDH3I1, CsALDH3898, CsALDH20158 and CsALDH54788, structure genes involved in the crocins biosynthesis pathway and 34CsWRKY genes, we found strong correlations between some structure genes and CsWRKY genes. For instance, CsWRKY2, -15 and − 28, which were all clustered in group IId, exhibited strong correlations with CsBCH gene that is the first enzyme of apocarotenoids biosynthesis. These three WRKY genes are orthologous to Arabidopsis WRKY11 (Figs. 3 and 4). The ortholog of AtWRKY11 was found to modulate the biosynthesis of second metabolites including anthocyanins [41] and flavonoids [42]. In addition, CsWRKY8, -15 and − 32 showed significant correlations with CsCCD2L which is the key enzyme to produce crocins. CsWRKY8 showed close evolutionary relationship with AtWRKY23 based on phylogenetic analysis. WRKY23 was found to play a role in accumulation of flavonols in A. thaliana [43]. So, we suppose CsWRKY8 may regulate the production of corcins by modulating the expression level of CsCCD2L gene along with the development of stigma. In saffron, CsALDHs can oxidize crocetin dialdehyde into crocetin which is a direct precursor of crocins. Coexpression analysis presented that CsWRKY4, -13, -17 and − 34 correlated with CsALDH, suggesting these WRKY TFs may influence the expression of CsALDH genes. CsWRKY17 was clustered

in group IIc as CsWRKY8, and possessed identical motifs (motif1, 8 and 13) (Fig. 2), which indicated that CsWRKY and CsWRKY8 had similar function in stigma development. Finally, there were also remarkable correlations which were found between CsWRKY8, -15 and CsUGT74AD1, the last enzyme in the biosynthesis pathway of crocins. Interestingly, CsWRKY8 and − 15 demonstrated simultaneous correlations with two or three structure genes respectively, consistent with the result from previous study [44, 45], suggesting these CsWRKY genes may be the key regulator of biosynthesis of crocins. Subsequently, the expression profiles of all CsWRKY genes which were obtained from coexpressin analysis in different tissue provided information that CsWRKY1, -2, -8, -10, -15 and − 28 were highly expressed in stigma in contrast with other tissues. Furthermore, the expression pattern of these six CsWRKY genes, as expected, were paralleled to the content of crocins during stigma development. On the basis of these results, we speculate that these six WRKY TFs (CsWRKY1, -2, -8, -10, -15 and − 28) might play roles in accumulation of crocins in stigma at specific developmental stages through regulating the expression level of structure genes in the biosynthesis pathway. However, the molecular mechanism of these CsWRKY in modulating crocin biosynthesis-related genes requires further elucidation via a series of in vivo and in vitro experiment.

Conclusions

In conclusion, 34 CsWRKY TFs from stigma of C.Sativus were characterized by transcriptome-wide analysis. These CsWRKY TFs were divided into group □, □ and □ supported by phylogenetic tree analysis. Coexpression network between CsWRKY TFs and crocin biosynthesis-related genes revealed that CsWRRKY1, -2, -8, -10, -15 and − 28 genes were potentially involved in the regulation of crocin biosynthesis, which were validated by their expression profile in various tissue and developmental stages. This work provides six candidate CsWRKY TFs and make a very convenient shortcut way for future further research into regulating mechanism of crocins biosynthesis.

Declarations

Author contributions conceptualization: GCG and JL; Writing, review and editing: GCG, JMW and JL; Resources: PW; Data analysis: GCG, JL; Experiment: LSM, GCG and JBX.

Funding This study was funded by Jiaxing Science and Technology Planning Project (No. 2020AY10023 and 2023AY11048), the National Natural Science Foundation of China (No. 81703667) and Special Project of Jiaxing Science and Technology Commissioner (2021K117).

Data Availability The datasets generated and/or analyzed during the current study are available from the corresponding author upon reasonable request.

Conflicts of interest The authors have no relevant financial or non-financial interests to disclose.

Ethical Approval There are no human and animal subjects in this article and informed consent is not applicable.

Consent to participate and Consent to publish All authors read and approved the final version of the manuscript. All authors consent to participate and consent to publish this manuscript.

References

1. Husaini AM, Haq SA ul, Jiménez AJL (2022) Understanding saffron biology using omics- and bioinformatics tools: stepping towards a better Crocus phenome. Mol Biol Rep 49:5325–5340. https://doi.org/10.1007/s11033-021-07053-x
2. Diretto G, Ahrazem O, Rubio-Moraga Á, et al (2019) UGT709G1: a novel uridine diphosphate glycosyltransferase involved in the biosynthesis of picrocrocin, the precursor of safranal in saffron (*Crocus sativus*). New Phytol 224:725–740. https://doi.org/10.1111/nph.16079
3. El Midaoui A, Ghzaiel I, Vervandier-Fasseur D, et al (2022) Saffron (Crocus sativus L.): A Source of Nutrients for Health and for the Treatment of Neuropsychiatric and Age-Related Diseases. Nutrients 14:597. https://doi.org/10.3390/nu14030597
4. Moradi S, Kafi M, Aliniaeifard S, et al (2022) Monochromatic blue light enhances crocin and picrocrocin content by upregulating the expression of underlying biosynthetic pathway genes in saffron (Crocus sativus L.). Front Hortic 1:960423. https://doi.org/10.3389/fhort.2022.960423
5. Ahrazem O, Diretto G, Rambla JL, et al (2022) Engineering high levels of saffron apocarotenoids in tomato. Hortic Res 9:uhac074. https://doi.org/10.1093/hr/uhac074
6. Demurtas OC, Frusciante S, Ferrante P, et al (2018) Candidate Enzymes for Saffron Crocin Biosynthesis Are Localized in Multiple Cellular Compartments. Plant Physiol 177:990–1006. https://doi.org/10.1104/pp.17.01815
7. Frusciante S, Diretto G, Bruno M, et al (2014) Novel carotenoid cleavage dioxygenase catalyzes the first dedicated step in saffron crocin biosynthesis. Proc Natl Acad Sci 111:12246–12251. https://doi.org/10.1073/pnas.1404629111
8. Gómez-Gómez L, Pacios LF, Diaz-Perales A, et al (2018) Expression and Interaction Analysis among Saffron ALDHs and Crocetin Dialdehyde. Int J Mol Sci. https://doi.org/ 10.3390/ijms19051409
9. López-jimenez AJ, Frusciante S, Niza E, et al (2021) A New Glycosyltransferase Enzyme from Family 91, UGT91P3, Is Responsible for the Final Glucosylation Step of Crocins in Saffron (Crocus sativus L.). Int J Mol Sci 22:8815. https://doi.org/10.3390/ijms22168815
10. Ashraf N (2015) Identification, cloning and characterization of an ultrapetala transcription factor CsULT1 from Crocus: a novel regulator of apocarotenoid biosynthesis. BMC Plant Biol. https://doi.org/ 10.1186/s12870-015-0423-7
11. Malik AH, Ashraf N (2017) Transcriptome wide identification, phylogenetic analysis, and expression profiling of zinc-finger transcription factors from Crocus sativus L. Mol Genet Genomics 292:619–633. https://doi.org/10.1007/s00438-017-1295-3
12. Bhat ZY, Mohiuddin T, Kumar A, et al (2021) Crocus transcription factors CstMYB1 and CstMYB1R2 modulate apocarotenoid metabolism by regulating carotenogenic genes. Plant Mol Biol 107:49–62.

https://doi.org/10.1007/s11103-021-01180-6

13. Gao G, Wu J, Li B, et al (2021) Transcriptomic analysis of saffron at different flowering stages using RNA sequencing uncovers cytochrome P450 genes involved in crocin biosynthesis. Mol Biol Rep 48:3451–3461. https://doi.org/10.1007/s11033-021-06374-1
14. Zhang M, Chen Y, Nie L, et al (2018) Transcriptome-wide identification and screening of WRKY factors involved in the regulation of taxol biosynthesis in Taxus chinensis. Sci Rep 8:5197. https://doi.org/10.1038/s41598-018-23558-1
15. Wani SH, Anand S, Singh B, et al (2021) WRKY transcription factors and plant defense responses: latest discoveries and future prospects. Plant Cell Rep 40:1071–1085. https://doi.org/10.1007/s00299-021-02691-8
16. Wang H, Chen W, Xu Z, et al (2023) Functions of WRKYs in plant growth and development. Trends Plant Sci 28:630–645. https://doi.org/10.1016/j.tplants.2022.12.012
17. Goyal P, Devi R, Verma B, et al (2023) WRKY transcription factors: evolution, regulation, and functional diversity in plants. Protoplasma 260:331–348. https://doi.org/10.1007/s00709-022-01794-7
18. Zhang J, Zhao H, Chen L, et al (2023) Multifaceted roles of WRKY transcription factors in abiotic stress and flavonoid biosynthesis. Front Plant Sci 14:1303667. https://doi.org/10.3389/fpls.2023.1303667
19. Huang X, Jia A, Huang T, et al (2023) Genomic profiling of WRKY transcription factors and functional analysis of CcWRKY7, CcWRKY29, and CcWRKY32 related to protoberberine alkaloids biosynthesis in Coptis chinensis Franch. Front Genet 14:1151645. https://doi.org/10.3389/fgene.2023.1151645
20. Liang M-H, Li X-Y (2023) Involvement of Transcription Factors and Regulatory Proteins in the Regulation of Carotenoid Accumulation in Plants and Algae. J Agric Food Chem https://doi.org/10.1021/acs.jafc.3c05662
21. Yuan Y, Ren S, Liu X, et al (2022) SlWRKY35 positively regulates carotenoid biosynthesis by activating the MEP pathway in tomato fruit. New Phytol 234:164–178. https://doi.org/10.1111/nph.17977
22. Han Y (2016) Characterization of OfWRKY3, a transcription factor that positively regulates the carotenoid cleavage dioxygenase gene OfCCD4 in Osmanthus fragrans. Plant Mol Biol. https://doi.org/ 10.1007/s11103-016-0483-6
23. Chen H, Ji H, Huang W, et al (2024) Transcription factor CrWRKY42 coregulates chlorophyll degradation and carotenoid biosynthesis in citrus. Plant Physiol 38394457 https:///10.1093/plphys/kiae048
24. Gómez-Gómez L, Parra-Vega V, Rivas-Sendra A, et al (2017) Unraveling Massive Crocins Transport and Accumulation through Proteome and Microscopy Tools during the Development of Saffron Stigma. Int J Mol Sci 18:76. https://doi.org/10.3390/ijms18010076
25. Yang X, Li H, Yang Y, et al (2018) Identification and expression analyses of WRKY genes reveal their involvement in growth and abiotic stress response in watermelon (Citrullus lanatus). PLOS ONE

13:e0191308. https://doi.org/10.1371/journal.pone.0191308

26. Colinas M, Goossens A (2018) Combinatorial Transcriptional Control of Plant Specialized Metabolism. Trends Plant Sci 23:324–336. https://doi.org/10.1016/j.tplants.2017.12.006
27. Jain M, Srivastava PL, Verma M, et al (2016) De novo transcriptome assembly and comprehensive expression profiling in Crocus sativus to gain insights into apocarotenoid biosynthesis. Sci Rep 6:22456. https://doi.org/10.1038/srep22456
28. Pu X, Xu Z, Gao R, Song J (2021) Research progress on biosynthesis and synthetic biology of crocin. Chin Sci Bull 66:219–232. https://doi.org/10.1360/TB-2020-0875
29. Luo D, Wang T, Ye M, et al (2023) Identification and characterization of Crocus sativus WRKY and its interacting MPK involved in crocins biosynthesis based on full-length transcriptome analysis. Ind Crops Prod 197:116559. https://doi.org/10.1016/j.indcrop.2023.116559
30. Liang M-H, Jiang J-G (2017) Analysis of carotenogenic genes promoters and WRKY transcription factors in response to salt stress in Dunaliella bardawil. Sci Rep 7:37025. https://doi.org/10.1038/srep37025
31. Duan X, Jiang C, Zhao Y, et al (2022) Transcriptome and metabolomics analysis revealed that CmWRKY49 regulating CmPSY1 promotes β-carotene accumulation in orange fleshed oriental melon. Hortic Plant J 8:650–666. https://doi.org/10.1016/j.hpj.2022.07.005
32. Goyal P, Manzoor MM, Vishwakarma RA, et al (2020) A Comprehensive Transcriptome-Wide Identification and Screening of WRKY Gene Family Engaged in Abiotic Stress in Glycyrrhiza glabra. Sci Rep 10:373. https://doi.org/10.1038/s41598-019-57232-x
33. Cao Z, Wu P, Gao H, et al (2022) Transcriptome-wide characterization of the WRKY family genes in Lonicera macranthoides and the role of LmWRKY16 in plant senescence. Genes Genomics 44:219–235. https://doi.org/10.1007/s13258-021-01118-8
34. Chang X, Yang Z, Zhang X, et al (2022) Transcriptome-wide identification of WRKY transcription factors and their expression profiles under different stress in *Cynanchum thesioides*. PeerJ 10:e14436. https://doi.org/10.7717/peerj.14436
35. Liu J, Li G, Wang R, et al (2023) Genome-Wide Analysis of WRKY Transcription Factors Involved in Abiotic Stress and ABA Response in Caragana korshinskii. Int J Mol Sci 24:9519. https://doi.org/10.3390/ijms24119519
36. Baranwal VK, Negi N, Khurana P (2016) Genome-wide Identification and Structural, Functional and Evolutionary Analysis of WRKY Components of Mulberry. Sci Rep 6:30794. https://doi.org/10.1038/srep30794
37. Wu J, Chen J, Wang L, Wang S (2017) Genome-Wide Investigation of WRKY Transcription Factors Involved in Terminal Drought Stress Response in Common Bean. Front Plant Sci 8:. https://doi.org/10.3389/fpls.2017.00380
38. Goel R, Pandey A, Trivedi PK, Asif MH (2016) Genome-Wide Analysis of the Musa WRKY Gene Family: Evolution and Differential Expression during Development and Stress. Front Plant Sci 7:. https://doi.org/10.3389/fpls.2016.00299

39. Hussain RMF, Sheikh AH, Haider I, et al (2018) Arabidopsis WRKY50 and TGA Transcription Factors Synergistically Activate Expression of PR1. Front Plant Sci 9:930. https://doi.org/10.3389/fpls.2018.00930
40. Li Y, Chen Y, Zhou L, et al (2020) MicroTom Metabolic Network: Rewiring Tomato Metabolic Regulatory Network throughout the Growth Cycle. Mol Plant 13:1203–1218. https://doi.org/10.1016/j.molp.2020.06.005
41. Wang Y, Chen L, Yang Q, et al (2022) New insight into the pigment composition and molecular mechanism of flower coloration in tulip (Tulipa gesneriana L.) cultivars with various petal colors. Plant Sci 317:111193. https://doi.org/10.1016/j.plantsci.2022.111193
42. Wang Z, Luo Z, Liu Y, et al (2021) Molecular cloning and functional characterization of NtWRKY11b in promoting the biosynthesis of flavonols in Nicotiana tabacum. Plant Sci 304:110799. https://doi.org/10.1016/j.plantsci.2020.110799
43. Grunewald W, De Smet I, Lewis DR, et al (2012) Transcription factor WRKY23 assists auxin distribution patterns during *Arabidopsis* root development through local control on flavonol biosynthesis. Proc Natl Acad Sci 109:1554–1559. https://doi.org/10.1073/pnas.1121134109
44. Zhou Y, Ma Y, Zeng J, et al (2016) Convergence and divergence of bitterness biosynthesis and regulation in Cucurbitaceae. Nat Plants 2:16183. https://doi.org/10.1038/nplants.2016.183
45. Dong YM, Zhang MY, Ling ZY, et al (2020) Advances in transcription factors regulating plant terpenoids biosynthesis. Chin Bull Bot 55, 340-350. https://doi.org/ 10.11983/CBB19186

Figures

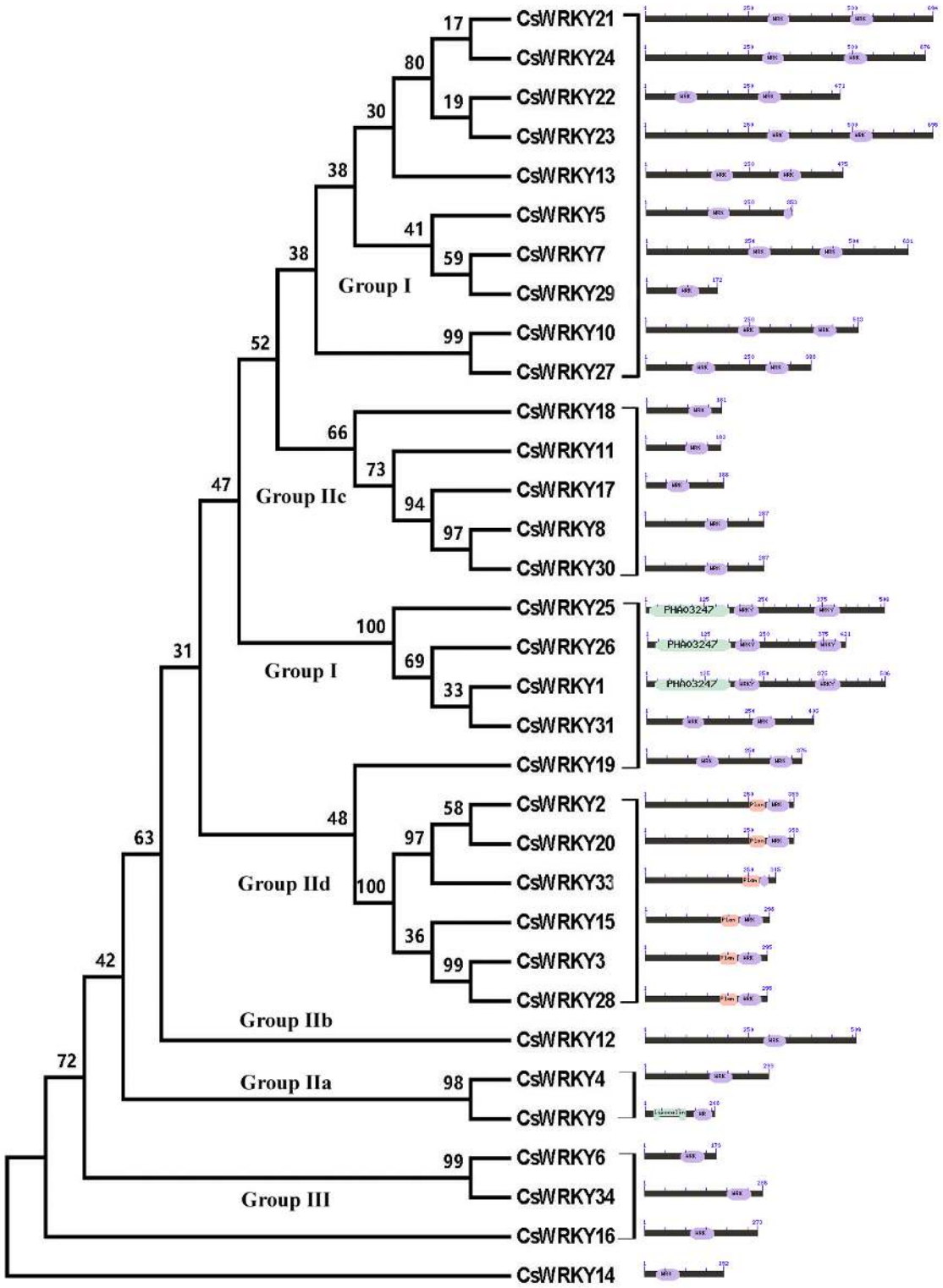

Figure 1

Classification of full-length CsWRKYs amino acid sequences with different conserved domains.

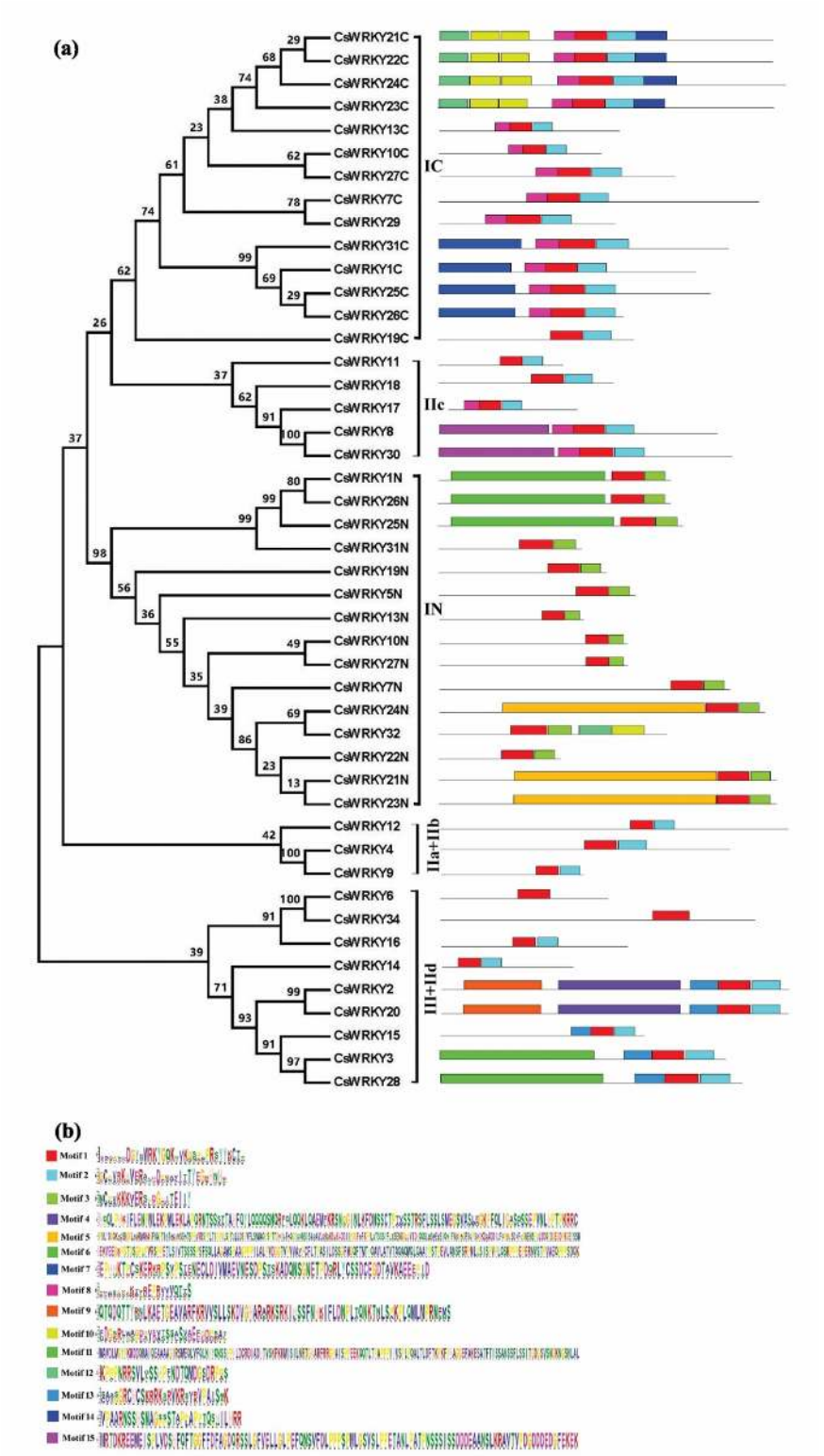

Figure 2

visualization of classification of 34 CsWRKY proteins. (a) conserved motifs distribution of the CsWRKYs. (b) The sequences of the fifteen motifs.

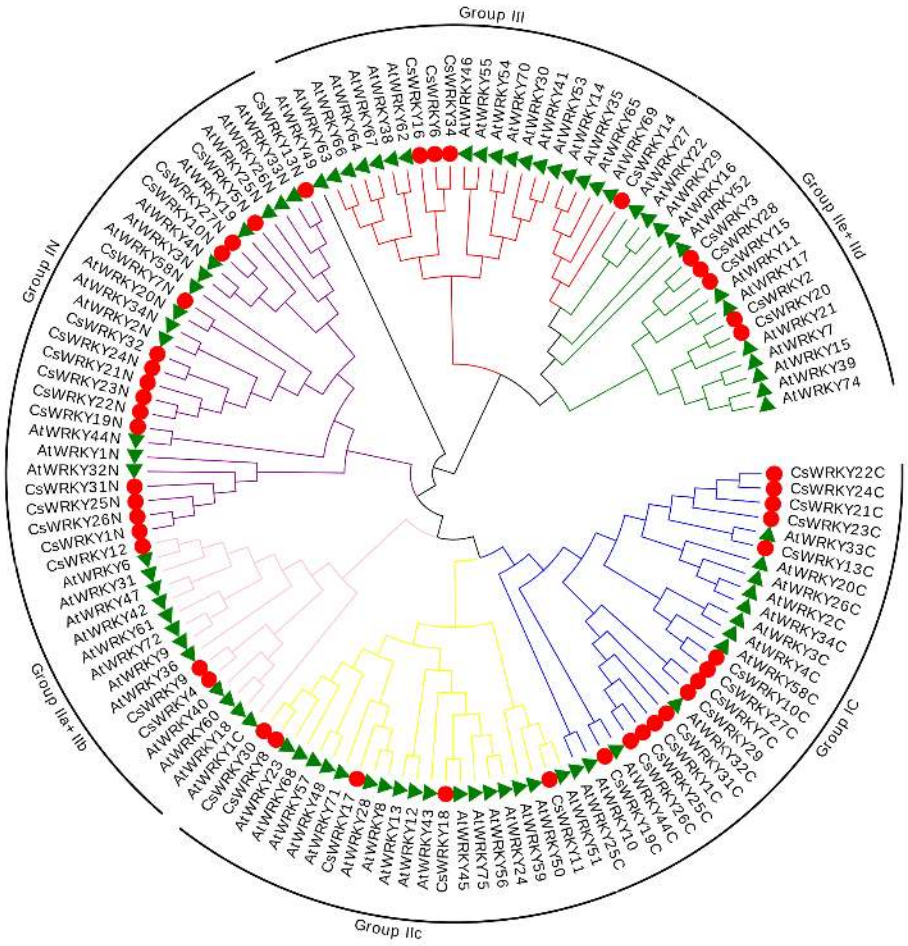

Figure 3

Neighbour-joining JTT model of phylogenetic tree.

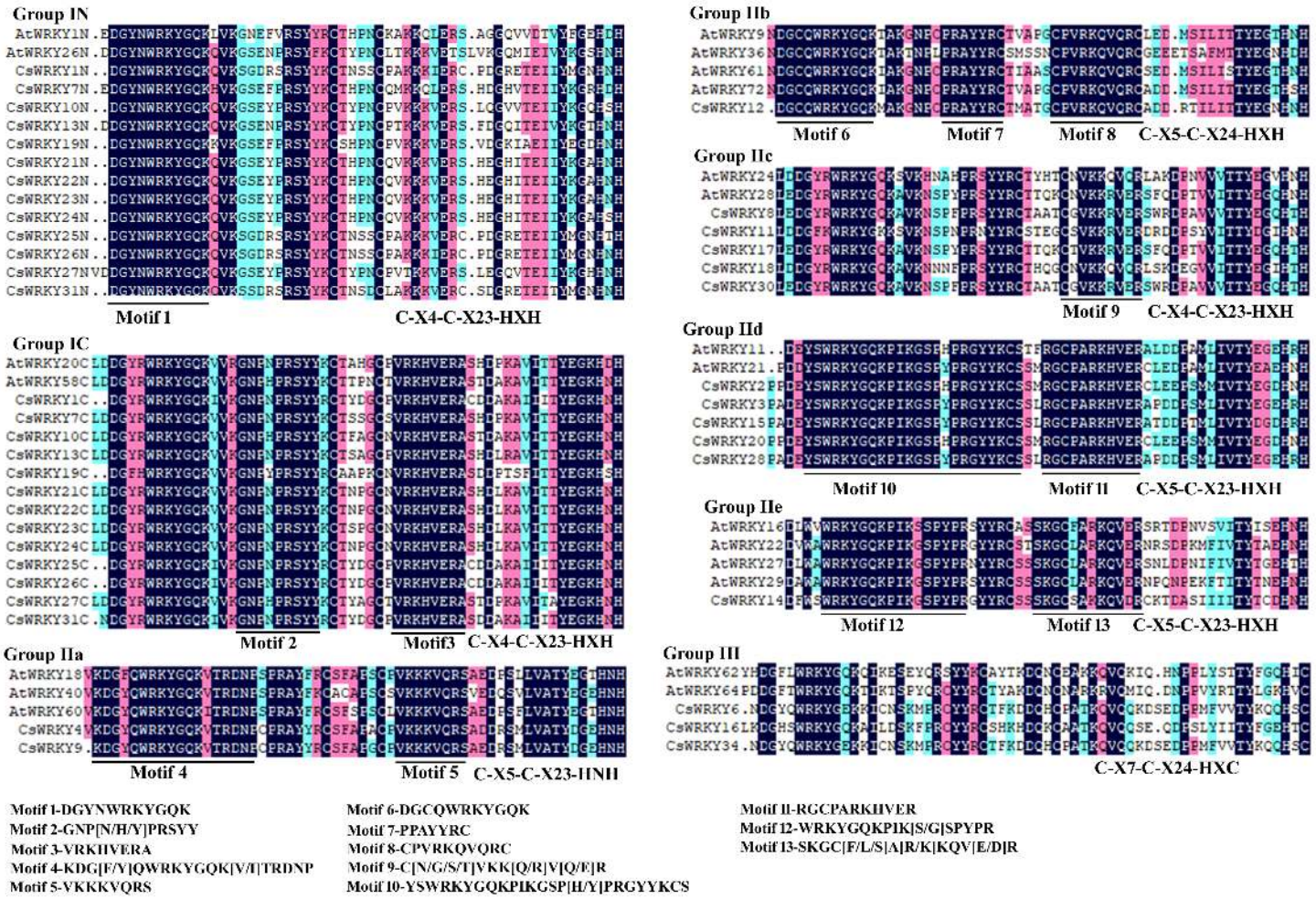

Figure 4

Multiple sequence alignment of conserved CsWRKY domain.

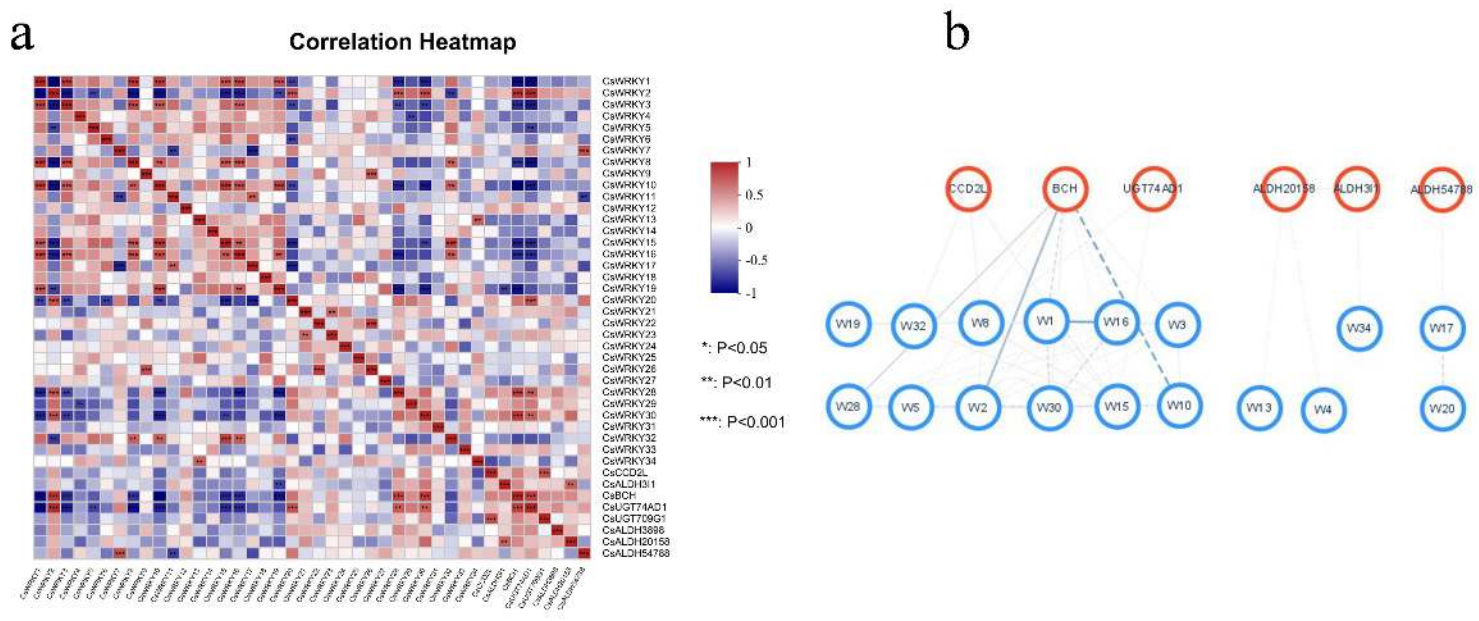

Figure 5

Coexpression analysis of CsWRKY and crocin biosynthesis-related genes. (a) correlation matrix heatmap. Red color indicates positive correlation, blue color indicates negative correlation. The correlation coefficients were calculated using spearman method. (b) coexpression network. The blue circle nodes represent CsWRKYs, and W1-34 means CsWRKY1-34 respectively. The red circle nodes represent crocin biosynthesis-related genes. Solid and dashed lines (edges) represent positive and negative coexpression correlations, respectively. The width of edges means the strength of the correlation.

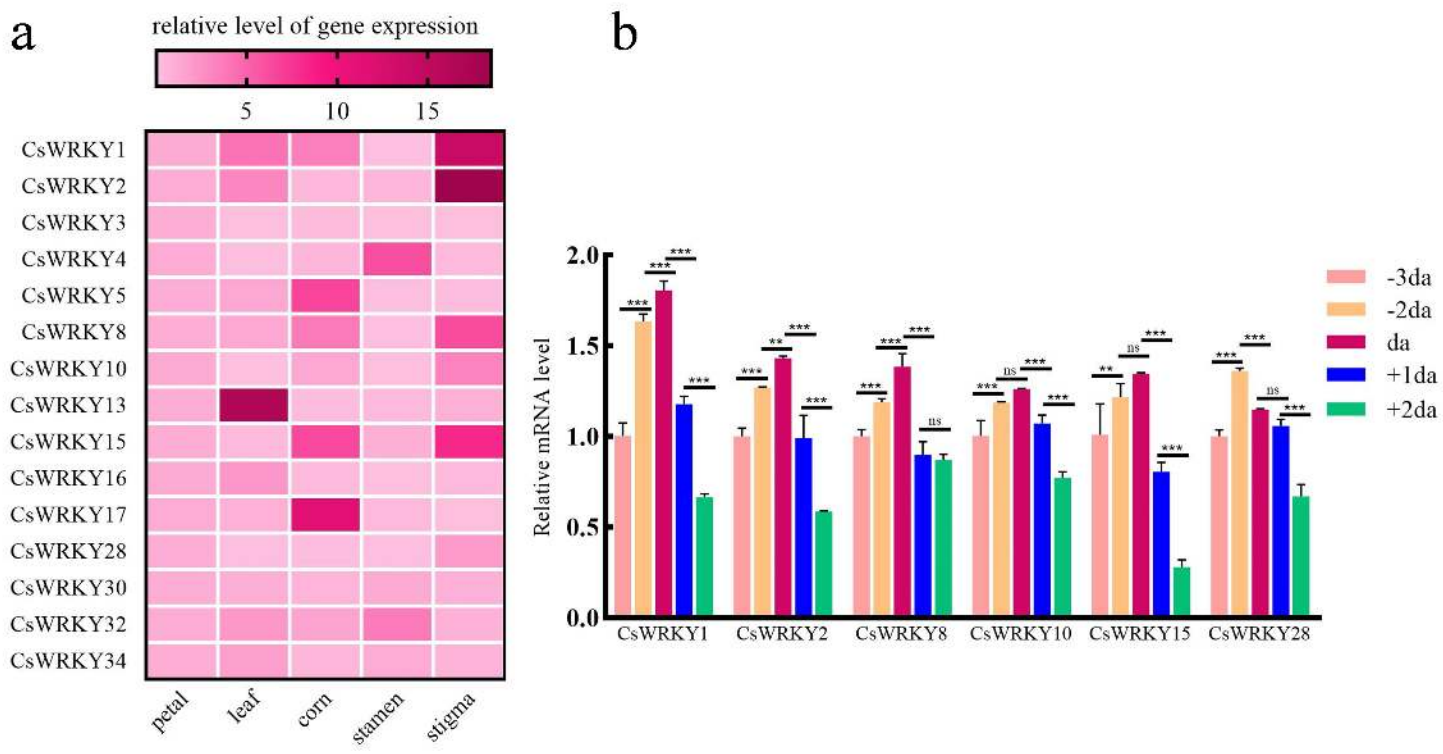

Figure 6

expression profile of key CsWRKYs in different tissues and developmental stages by RT-QPCR. (a) expression of fifteen CsWRKYs in different tissues. (b) expression of five CsWRKYs in different developmental stages.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupplementaryFileS1.docx
- SupplementaryFileS2.xls

- SupplementaryFileS3.xls