



OPEN

## Genome-wide identification and comprehensive analysis of WRKY transcription factor family in safflower during drought stress

Xianming Song<sup>1,2,3</sup>, Xianfei Hou<sup>1,3</sup>, Youling Zeng<sup>2✉</sup>, Donghai Jia<sup>1✉</sup>, Qiang Li<sup>1✉</sup>, Yuanguo Gu<sup>1</sup> & Haocui Miao<sup>1</sup>

The WRKY family is an important family of transcription factors in plant development and stress response. Currently, there are few reports on the WRKY gene family in safflower (*Carthamus tinctorius* L.). In this study, a total of 82 *CtWRKY* genes were identified from the safflower genome and could be classified into 3 major groups and 5 subgroups based on their structural and phylogenetic characteristics. The results of gene structure, conserved domain and motif analyses indicated that *CtWRKYs* within the same subfamily maintained a consistent exon/intron organization and composition. Chromosomal localization and gene duplication analysis results showed that *CtWRKYs* were randomly localized on 12 chromosomes and that fragment duplication and purification selection may have played an important role in the evolution of the WRKY gene family in safflower. Promoter cis-acting element analysis revealed that the *CtWRKYs* contain many abiotic stress response elements and hormone response elements. Transcriptome data and qRT-PCR analyses revealed that the expression of *CtWRKYs* showed tissue specificity and a strong response to drought stress. Notably, the expression level of the *CtWRKY55* gene rapidly increased more than eightfold under drought treatment and rehydration, indicating that it may be a key gene in response to drought stress. These results provide useful insights for investigating the regulatory function of the *CtWRKY* gene in safflower growth and development, as well as identifying key genes for future molecular breeding programmes.

Transcription factors (TFs) play an important role in plant growth and development and biotic and abiotic stress responses by regulating the transcription of downstream target genes through interaction with cis-acting elements. Since the first extraction of the WRKR gene *SPF1* from *Pomoea batatas* in 1994, researchers have identified the corresponding WRKY genes from *Acer truncatum*, *Taraxacum kok-saghyz* Rodin, *Chrysanthemum lavandulifolium*, *Linum usitatissimum* L. and *Zea mays*<sup>1–5,8</sup>. Hu Wenjing et al. found that the conserved domain of WRKY family members consists of 60–70 amino acids, including a highly conserved seven-phthalide sequence (WRKYGQK) and a zinc finger structure (C<sub>2</sub>H<sub>2</sub> or C<sub>2</sub>HC)<sup>5</sup>. Researchers discovered that members of the WRKY gene family could be divided into Groups I, II and III based on the number of conserved structural domains and the type of zinc finger structures, and group II could be further divided into five subgroups, IIa, IIb, IIc, IId and IIe, based on differences in amino acid sequences<sup>6,8</sup>. Among them, members of Group I have two conserved WRKY structural domains, and members of groups II and III have one conserved WRKY structural domain; members of Groups I and II have C–X<sub>4–5</sub>–C–X<sub>23–24</sub>–H–X<sub>1</sub>–H (C<sub>2</sub>H<sub>2</sub>) zinc finger structures, and members of group III have C–X<sub>5–7</sub>–C–X<sub>23–38</sub>–H–X<sub>1</sub>–C (C<sub>2</sub>HC) zinc finger structures<sup>6</sup>. WRKY transcription factors can specifically bind to cis-acting elements in the promoter regions of target genes to regulate the expression of downstream genes and enhance plant resistance to abiotic stresses such as low temperature, drought and salt, as well as to pathogenic bacteria<sup>4,6,7,8</sup>. For example, suppression of *GhWRKY21* has been shown to improve drought tolerance in cotton, although *GhWRKY21* plays a negative role in the drought response in cotton<sup>9</sup>. *MaWRKY80* overexpression results in improved phenotypic morphology, enhanced survival, reduced water loss

<sup>1</sup>Present address: Economic Crop Research Institute, Xinjiang Academy of Agricultural Sciences, Urumqi 830091, China. <sup>2</sup>Present address: Xinjiang Key Laboratory of Biological Resources and Genetic Engineering, College of Life Science & Technology, Xinjiang University, Urumqi 830046, China. <sup>3</sup>These authors contributed equally: Xianming Song and Xianfei Hou. ✉email: zeng\_ylxju@126.com; 545507831@qq.com; lq19820302@126.com

rate, and lower malondialdehyde (MDA) levels compared to the wild-type (WT) under drought stress. Notably, the leaves of transgenic *MaWRKY80 Arabidopsis* exhibited lower reactive oxygen species (ROS) levels than WT leaves under drought stress<sup>10</sup>. *AtWRKY11* and *AtWRKY70* can coordinate their resistance to *Bacillus* through jasmonic acid (JA) and salicylic acid (SA) signalling pathways, and their overexpression can enhance drought tolerance and promote seed germination and root growth under drought stress in *Arabidopsis*<sup>11</sup>. The sorghum WRKY transcription factor *SbWRKY30* is mainly expressed in leaves and roots and is induced by drought stress. Heterologous expression of *SbWRKY30* in *Arabidopsis* and rice confers drought tolerance by altering root architecture<sup>12</sup>. Through the ABA signalling pathway, *ZmWRKY79* can improve drought tolerance in maize<sup>13</sup>. The *ClWRKY20* expression level rose in response to treatment with salinity, drought, and phytohormones (ABA, ET, and SA). Transgenic *Arabidopsis* exhibited heightened sensitivity to ABA during seed germination and under low temperatures and salinity upon overexpression of *ClWRKY20*<sup>14</sup>. Additionally, WRKY genes, which respond to drought stress, have also been identified in other crops, providing evidence that WRKY transcription factors are crucial in responding and adapting to drought stress<sup>15–18</sup>.

Safflower (*Carthamus tinctorius* L.) is a special economic crop utilized in medicinal materials, oil, dyes and feed. It has the characteristics of strong environmental adaptability, resistance to high temperature and drought, and a short growth cycle<sup>19,20</sup>. The Xinjiang Uygur Autonomous Region is the main safflower-producing area in China. This perennially utilized area accounts for more than 80% of the area and yield of safflower in China<sup>21</sup>. Xinjiang belongs to the “temperate continental climate” zone, with a large temperature range, sufficient sunshine (the annual sunshine time reaches 2 500 to 3 500 h), low precipitation and a dry climate. Safflower has strong stress resistance, but its quality and yield are still affected by drought, high temperature and scarce rainfall<sup>22–24</sup>. In 2021, Wu Zhihua et al. sequenced the safflower genome, which provided us with a helpful foundation to further analyse the gene function of safflower<sup>23</sup>.

Safflower represents an important economic crop. Studies have shown that multiple transcription factor families have been found in safflower, such as MYB, bHLH and bZIP transcription factor families<sup>25–27</sup>. Nevertheless, the molecular mechanism governing drought tolerance regulation in safflower remains unknown. Furthermore, there has been no account of the genome-wide identification and functional analysis of the WRKY gene family in safflower. In this study 82 *CtWRKY* genes were identified from the safflower genome and their physicochemical properties, phylogeny, chromosomal location, gene structure, conserved motifs, cis-acting elements, collinearity and protein interaction networks were comprehensively analysed by bioinformatics. Transcriptome data and qRT-PCR were used to analyse *CtWRKY* gene expression patterns in different tissues and under drought stress and rewetting treatments. This study provides valuable insights for investigating the regulatory function of the *CtWRKY* gene in safflower growth and development, as well as identifying key genes for future molecular breeding programmes.

## Results

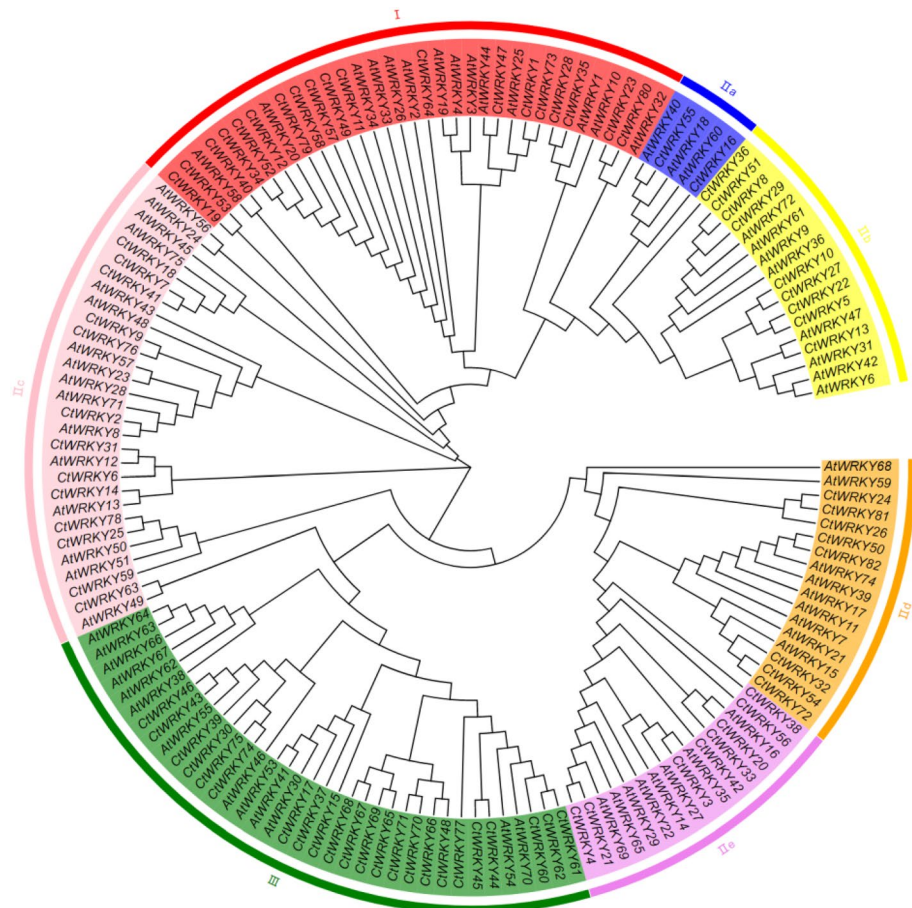
### Identification and physicochemical property analysis of WRKY gene family members in safflower

Using the protein sequence of the *Arabidopsis* WRKY gene family and the WRKY domain hidden Markov model (HMM) file (PF03106), 83 *CtWRKY* gene family members were preliminarily identified from the safflower genome by BLAST alignment, and then incomplete sequences and candidate sequences without corresponding domains were eliminated. A total of 82 *CtWRKY* gene family members were identified from the safflower genome, which were named according to their positions on the 12 safflower chromosomes, namely, *CtWRKY1–CtWRKY82* (Table S1).

As shown in Table S1, the coding sequences (CDS) length of *CtWRKY* genes ranged from 504 (*CtWRKY78*) to 4 011 bp (*CtWRKY30*). The length of *CtWRKY* proteins varied from 167 (*CtWRKY78*) to 1 366 bp (*CtWRKY30*) and the average protein sequence length was 379 residues. The molecular weight of the proteins was 19 130.99 Da (*CtWRKY78*)–150 954.83 Da (*CtWRKY30*), with an average of 42 203.72 Da. The isoelectric point was 4.92 (*CtWRKY38*)–10.15 (*CtWRKY24*). The total average hydrophilicity index revealed that *CtWRKY* proteins were hydrophilic proteins. Subcellular localization prediction showed that most *CtWRKY* genes were in the nucleus, and a few *CtWRKY* genes were in the peroxisomes (*CtWRKY6*, –8 and –48), chloroplasts (*CtWRKY33*), cytoplasm (*CtWRKY52*) and mitochondria (*CtWRKY75*).

### Phylogenetic analysis of the WRKY gene in safflower

Using the *Arabidopsis* WRKY gene family as a reference, the phylogenetic tree of safflower and *Arabidopsis* WRKY protein sequences was constructed by the neighbour-joining method (Fig. 1). The results revealed that *CtWRKY* proteins were divided into three groups (I–III) according to the grouping of *AtWRKY* gene family members, and group II was divided into five subgroups: IIa, IIb, IIc, IId, and IIe. Multiple sequence alignment revealed that the heptapeptide sequence (WRKYGQK) and zinc finger domains (C<sub>2</sub>H<sub>2</sub> and C<sub>2</sub>HC) were conserved in each group. In addition, three heptapeptide variants were found: WRKYGKK (*CtWRKY25*, –59 and –78) in group IIc, WKKYGEK (*CtWRKY24*, –26 and –81) in group IId and WRKYGHK (*CtWRKY60*) in group III (Fig. S1). Group I included 17 *CtWRKY* members, all of which had 2 WRKY conserved domains and a C<sub>2</sub>H<sub>2</sub> zinc finger structure. Group II included 42 *CtWRKY* members, all of which had a WRKY conserved domain and a C<sub>2</sub>H<sub>2</sub>-type zinc finger structure. They were further divided into 5 subgroups according to their genetic relationship and structure: IIa (2 proteins), IIb (9 proteins), IIc (15 proteins), IId (8 proteins) and IIe (8 proteins) subgroups. Group III included 23 *CtWRKY* members, all of which had a WRKY conserved domain and a C<sub>2</sub>HC zinc finger structure.



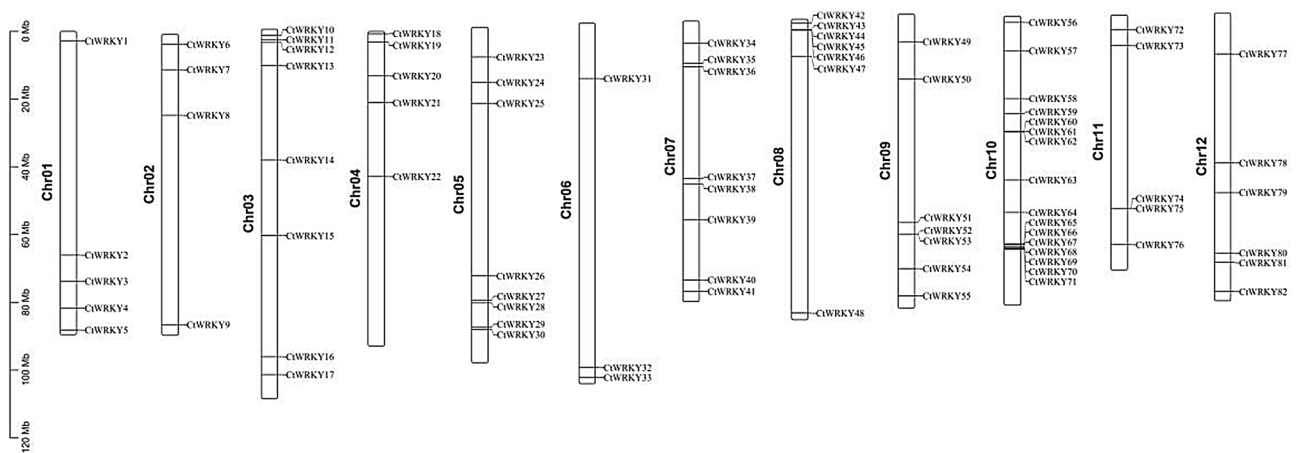
**Figure 1.** Phylogenetic analysis of safflower and *Arabidopsis* WRKY proteins. An unrooted neighbour-joining (NJ) phylogenetic tree was constructed with WRKY domains of WRKY proteins from safflower and *Arabidopsis* using MEGA7.0 with a bootstrap of 1000. Seven major clades are indicated: I, IIa, IIb, IIc, IId, IIe, and III with different coloured areas.

### Analysis of the chromosomal localization, gene structure and conserved motif of the WRKY gene in safflower

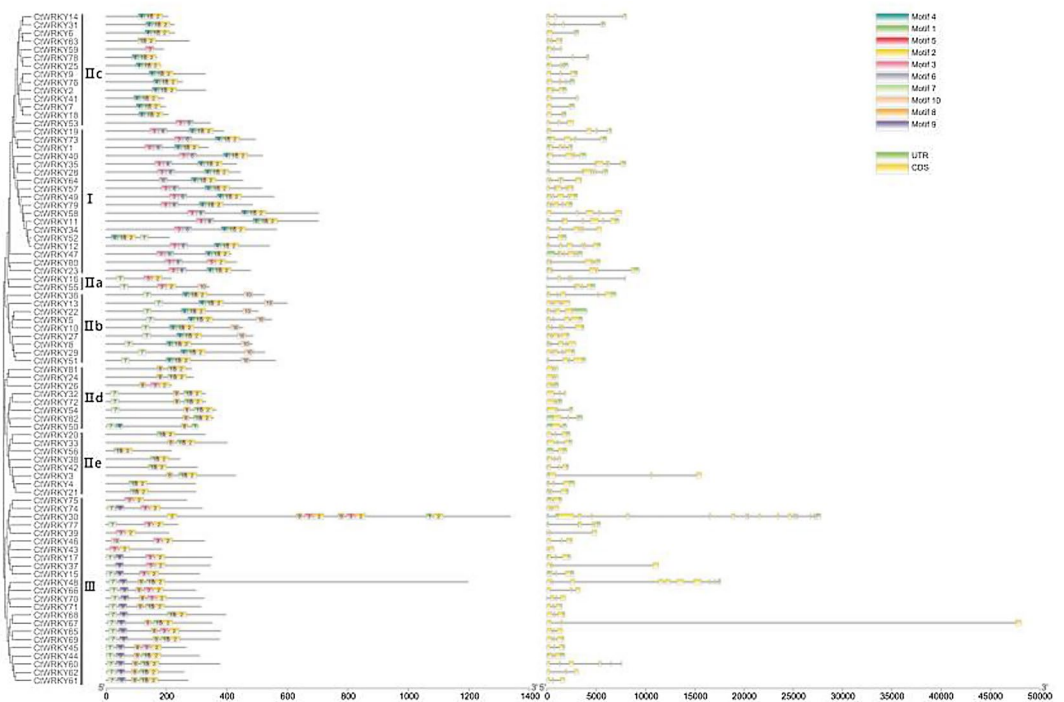
The distribution of the *CtWRKY* gene on safflower chromosomes was analysed with TBtools software (Fig. 2). The results showed that 82 *CtWRKY* genes were unevenly distributed on the 12 safflower chromosomes. The highest number of *CtWRKY* genes was found on chromosome 10 (16 genes), and the lowest number of *CtWRKY* genes was found on chromosome 6 (3 genes).

The diversity of gene structure promotes the evolution of the gene family. To further understand the evolutionary process of the *CtWRKY* gene family, the structure and conserved motifs of the *CtWRKY* gene were analysed in combination with the *CtWRKY* protein phylogenetic tree (Fig. 3). Ten conserved motifs were identified in *CtWRKY* proteins using the MEME tool, and 82 *CtWRKY* proteins were found to have specific conserved domains (Fig. S2 and Table S2). As shown in Fig. 3, most *CtWRKY* family members in the same group or subgroup shared similar conserved motifs, but there were some differences between groups. Furthermore, members of the same group or subgroup had specific conserved motifs. For example, motif 6 was present only in Group I members, motif 10 was present only in group IIa and IIb members, motif 9 was present only in group III members, and the group III member *CtWRKY30* protein contained two motif 3 s, which are rare in other species. These results suggest that these conserved motifs may play different roles in gene evolution and function. Additionally, almost all members of the *CtWRKY* family contain a WRKYGQK heptapeptide sequence (motif 1 or motif 3), indicating that WRKYGQK is the core conserved motif of the *CtWRKY* gene family.

According to the results of gene annotation, the exon–intron coding sequence structure of the *CtWRKY* gene was analysed (Fig. 3). The results revealed that *CtWRKY* family members had 2–17 exons, and a few members had no UTR. Among them, most *CtWRKY* genes contained 4 exons (2 introns), accounting for 48.78% (40/82) of all *CtWRKY* genes. The *CtWRKY48* gene contained 10 exons, and the *CtWRKY30* gene contained 17 exons. Consistent with conserved motif analysis, *CtWRKY* members in the same group or subgroup had similar gene structures and a high degree of conservation, indicating functional similarity between different members.



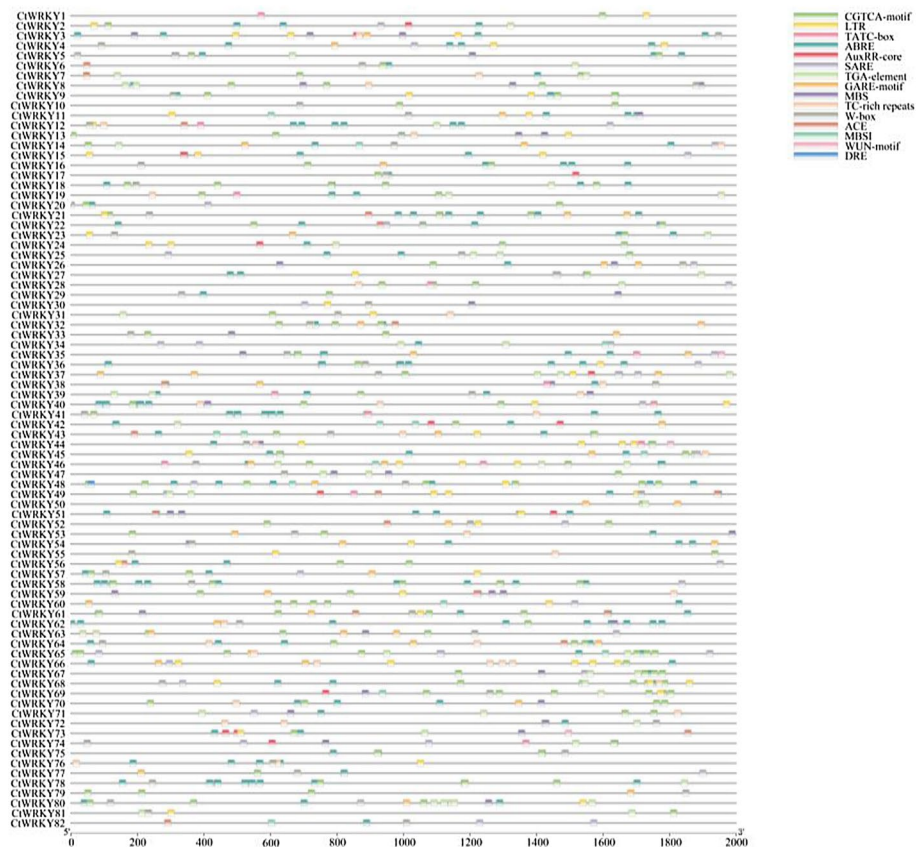
**Figure 2.** Chromosomal localization of the WRKY gene in safflower. A total of 82 *CtWRKY* genes were mapped to the ten safflower chromosomes with an uneven distribution. The gene name on the right side of each chromosome corresponds to the approximate location of each *CtWRKY* gene. The scale on the left is in megabases.



**Figure 3.** Phylogenetic tree, gene structure and conserved motif of the *WRKY* gene in safflower. The evolutionary tree was constructed from the *CtWRKY* protein sequences using MEGA 7.0 software with the neighbour-joining method. The structure of all 82 *CtWRKY* genes was obtained by TBtools software. The conserved motifs of *CtWRKY* proteins were analysed using the MEME tool.

### Analysis of cis-acting elements and collinearity of the *WRKY* gene in safflower

The 2 000 bp promoter sequence of the *CtWRKY* gene was extracted, and its cis-elements were analysed using the PlantCARE database to explore the potential function of the *CtWRKY* gene in abiotic stress response (Fig. 4). The results revealed that a total of 14 cis-acting elements associated with plant hormones and stress responses were found in the promoter region of the *CtWRKY* gene, including 1 light response element (ACE); 6 abiotic stress response elements (DRE, LTR, MBS, W-box, WUN motif and TC-rich repeats); and 7 hormone response elements (SARE, ABRE, TATC box, TGA element, AuxRR core, CGTCA motif and GARE motif). As shown in Figure S3, the promoter sequence of the *CtWRKY* gene contains the most cis-acting elements: the CGTCA motif (methyl jasmonate response element) and ABRE (abscisic acid response element). Approximately 68 genes contained these two elements. The promoter regions of 60 *CtWRKY* genes contained W-box elements, indicating



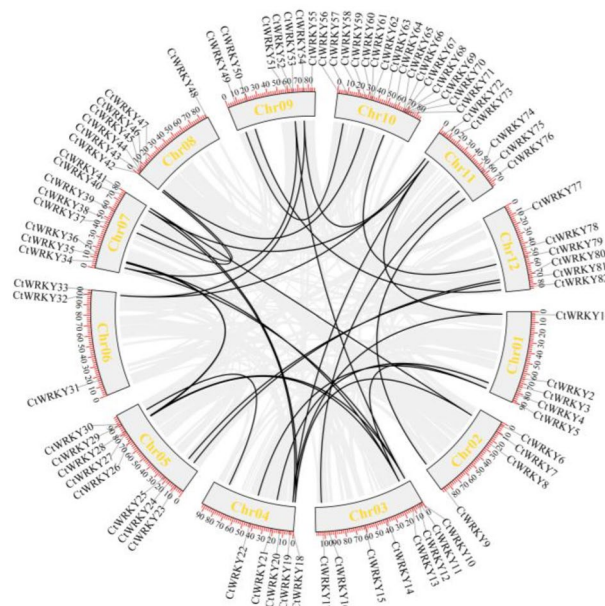
**Figure 4.** Distribution of cis-acting elements in the promoter region of the WRKY gene in safflower. The putative DRE, LTR, MBS, WUN motif, W-box, TATC box, GARE motif, AuxRR core, TGA element, CGTCA motif, TGACG motif, ABRE, SARE and MBSI core sequences are represented in different colours. DRE: cis-acting elements involved in dehydration, low temperature and salt stress responses; ABRE: cis-acting element involved in abscisic acid responsiveness; LTR: cis-acting element involved in low-temperature responsiveness; MBS: MYB binding site involved in drought inducibility; TGACG-motif and CGTCA-motif: cis-acting regulatory element involved in MeJA responsiveness; TC-rich repeats: cis-acting element involved in defence and stress responsiveness; TATC-box and GARE-motif: gibberellin-responsive element; W box: elicitation; wounding and pathogen responsiveness; AuxRR-core and TGA-element: auxin-responsive element; SARE: cis-acting element involved in salicylic acid responsiveness; MBSI: MYB binding sites involved in the regulation of flavonoid biosynthesis genes.

that these genes may be regulated by other WRKY transcription factors or by themselves. Furthermore, 39 *CtWRKY* genes contained the low temperature response element LTR, 31 *CtWRKY* genes contained the drought stress response element MBS, 35 *CtWRKY* genes contained the gibberellin response element GARE motif, and 28 *CtWRKY* genes contained defence and stress response element TC-rich repeats.

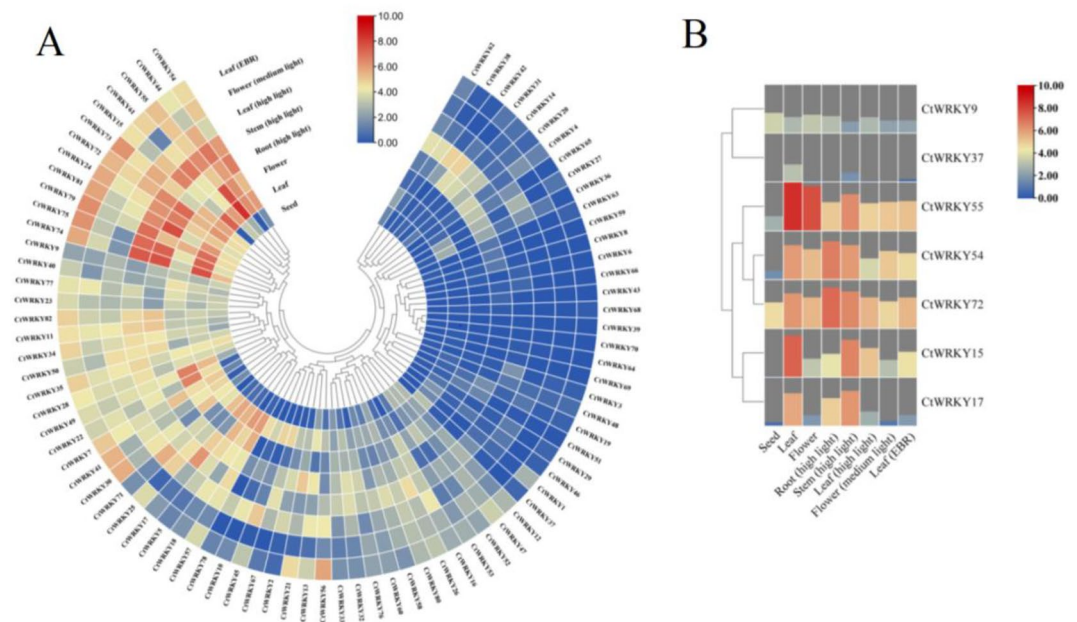
Fragment replication and tandem replication events are the main modes of gene family formation and expansion. The gene replication events of the safflower genome were analysed by the MCScanX tool (Fig. 5). The results revealed that 5 pairs of tandem duplicated gene pairs (*CtWRKY43/44*, *CtWRKY45/46*, *CtWRKY52/53*, *CtWRKY67/68* and *CtWRKY74/75*) were found in 82 *CtWRKY* genes, which were located on Chr8, Chr9, Chr10 and Chr11, respectively. Moreover, 34 pairs of fragment replication gene pairs were found. This indicates that fragment replication events may be the main driving force for the formation and expansion of the *CtWRKY* gene family. The Ka/Ks values of four *CtWRKY* gene pairs, which were NaN, were excluded. Furthermore, the Ka/Ks values of the remaining 34 gene pairs were below 0.5, and one gene pair exhibited a Ka/Ks value lower than 1, implying that these genes evolved via the process of purification selection. Tandemly duplicated genes had Ka/Ks values ranging from 0.27 to 0.62, whereas segmentally duplicated genes had Ka/Ks values ranging from 0.11 to 0.42. The mean Ka/Ks value for tandemly repeated genes (0.37) exceeded that of segmentally duplicated genes (0.22) (Table S3).

### Analysis of the expression pattern of the WRKY gene in safflower under different tissues and stress treatments

To explore the potential function of the *CtWRKY* gene in plant growth and development, TPM values were calculated by TBtools software according to *CtWRKY* transcriptome data downloaded from the NCBI database (Table S4), and an expression heatmap of the *CtWRKY* gene in different tissues was generated (Fig. 6). The



**Figure 5.** Collinear analysis of WRKY genes in safflower. The grey line represents all collinear regions in the entire safflower genome, and the black line represents the *CtWRKY* gene pair with replication events.



**Figure 6.** Expression of safflower WRKY genes in different tissues. **(A)** Expression heatmap of 82 *CtWRKY* genes in different tissues; **(B)** Expression heatmap of 7 *CtWRKY* genes with high homology to *Arabidopsis* drought-resistant WRKY genes in different tissues. The colours from blue to red represent different expression abundances. The light intensity was 40,000 lx (high light) and 20,000 lx (medium light), and the illumination time was 10 h/d. EBR treatment conditions: 0.1  $\mu\text{mol/L}$  EBR solution was sprayed evenly on the front and back of the leaves.

results showed that 82 *CtWRKY* genes were expressed in different tissues under the categories of seeds, leaves, flowers, EBR and light stress. However, the expression levels of a few genes were very low. Notably, 12 genes (*CtWRKY15*, -24, -44, -54, -55, -61, -72, -73, -74, -75, -79, and -81) showed high expression levels in 8 tissues, indicating that these genes play fundamental roles in growth and development and stress response. In addition, some *CtWRKY* genes were preferentially expressed in different tissues. For example, 4 *CtWRKY* genes (*CtWRKY11*, -34, -35 and -82) were highly expressed in different tissues under EBR and light stress; 4

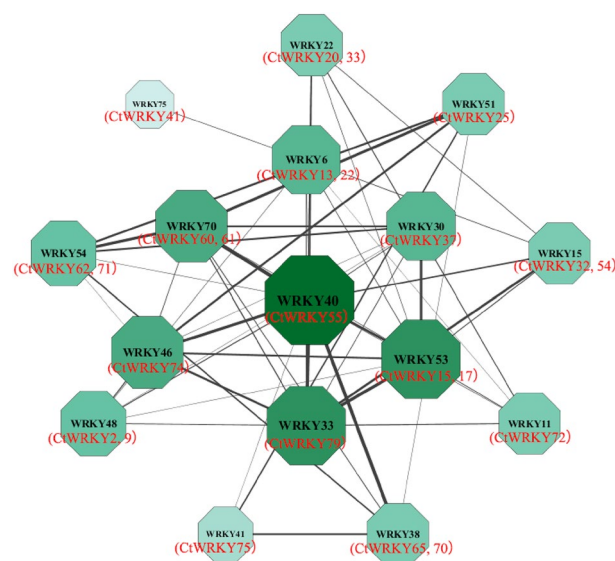
*CtWRKY* genes (*CtWRKY17*, –22, –25 and –49) were preferentially expressed in leaves, flowers and roots; and only a few genes were expressed in seeds. Analysis of the expression levels of 7 *CtWRKY* genes (*CtWRKY9*, –15, –17, –37, –54, –55, and –72) with high homology to *Arabidopsis* drought tolerance genes in different tissues revealed that all 7 genes had high expression levels in leaves (Fig. 6). Additionally, 4 *CtWRKY* genes (*CtWRKY15*, –54, –55, and –72) showed high expression levels in flowers as well as in different tissues under EBG and light stress, suggesting that these *CtWRKY* genes may play an important role in abiotic stress.

### Analysis of the protein interaction network of the WRKY gene in safflower

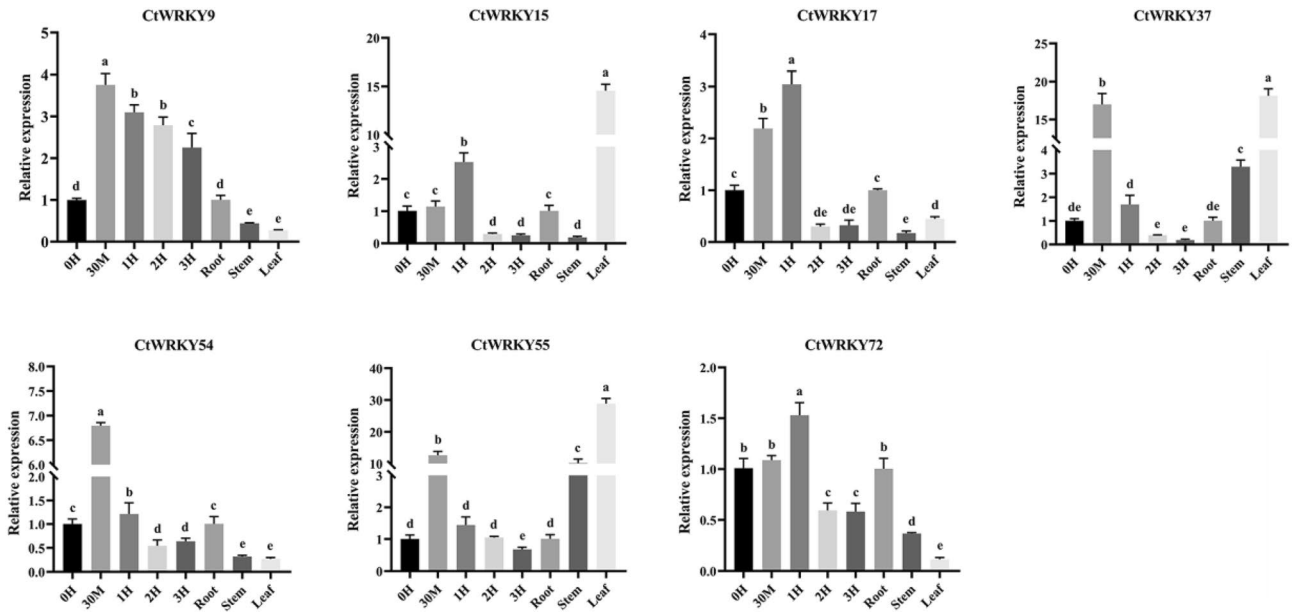
To predict the potential regulatory role of *CtWRKY* proteins, a *CtWRKY* protein interaction network was constructed using the STRING database based on *Arabidopsis* WRKY protein interaction data (Fig. 7). The findings showed that the *CtWRKY* proteins with high sequence similarity to *AtWRKY30* (*CtWRKY37*), *AtWRKY33* (*CtWRKY79*), *AtWRKY40* (*CtWRKY55*) and *AtWRKY53* (*CtWRKY15/17*) were the hubs of the protein interaction network, and the remaining *CtWRKY* proteins had strong or weak interactions with *CtWRKY15*, *CtWRKY17*, *CtWRKY37*, *CtWRKY55* and *CtWRKY79*.

### Analysis of the expression pattern of the WRKY gene in safflower under different tissues and drought stress

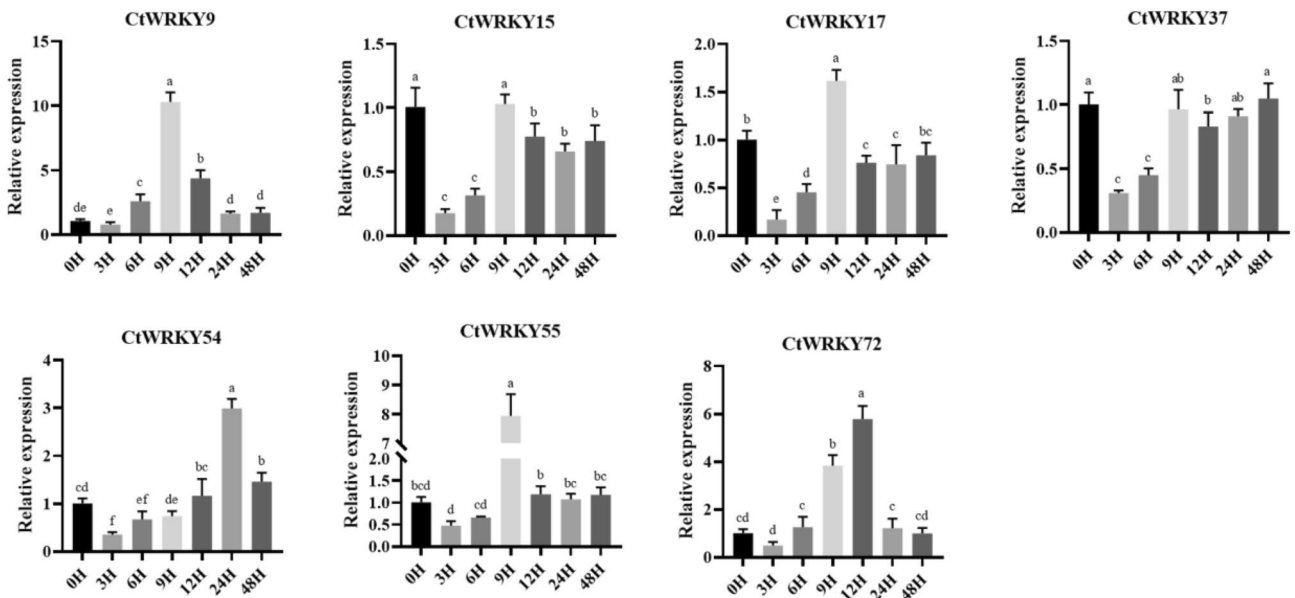
Under adverse conditions, many stress-inducible genes can help plants resist the influence of adverse factors. To analyse the stress response of the *CtWRKY* gene under drought stress, seven *CtWRKY* genes with high homology to *Arabidopsis* drought tolerance genes were selected (group II: *CtWRKY9*, –54, –55, and –72; group III: *CtWRKY15*, –17, and –37), and their expression patterns in different tissues and under drought stress and drought-rewatering treatment were measured using qRT-PCR (Figs. 8 and 9). As shown in Figure S4, seven *CtWRKY* genes had a heptapeptide core sequence WRKYGQK and a zinc finger structure C<sub>2</sub>H<sub>2</sub> or C<sub>2</sub>HC, which are very different from other drought tolerance genes except for the conserved domain. As shown in Fig. 8, seven *CtWRKY* genes were differentially expressed in roots, stems, and leaves. Three *CtWRKY* genes (*CtWRKY15*, –37, and –55) were highly expressed in leaves, and four *CtWRKY* genes (*CtWRKY9*, –17, –54, and –72) were highly expressed in roots, indicating that they play important roles in different tissues. The qRT-PCR analysis showed that the relative expression levels of the seven *CtWRKY* genes were significantly different within 3 h of drought treatment and 48 h of rewatering, and the overall expression pattern showed a trend of first increasing and then decreasing, indicating that these genes may play an important role in the drought stress response. As shown in Figs. 8 and 9, the expression levels of seven *CtWRKY* genes increased to varying degrees after drought stress and drought rehydration treatment, but the expression levels of these genes began to decrease as the treatment time increased. For example, after 30 min of drought treatment, three *CtWRKY* genes (*CtWRKY37*, –54 and –55) showed high expression levels in leaf tissues (increased by >6.5-fold), three *CtWRKY* genes (*CtWRKY9*, –15 and –17) showed high expression levels in leaf tissues (increased by >2.5-fold), and the expression level of *CtWRKY72* increased 1.5-fold. After 9 h of rehydration, three *CtWRKY* genes (*CtWRKY9*, –55 and –72) showed high expression levels (increased >5.6-fold), and the expression level of *CtWRKY17* increased 1.6-fold. The expression level of *CtWRKY54* increased 2.9-fold after 24 h of rehydration. In addition, the relative expression levels of the *CtWRKY37* and *CtWRKY55* genes rapidly increased more than 12-fold after drought treatment for 30 min, and the relative expression levels of the *CtWRKY9* and *CtWRKY55* genes rapidly increased more than



**Figure 7.** The protein interaction network of the WRKY protein from safflower. Cytoscape was used to visualize the network. The green coloured ball (node) represents the *CtWRKY* gene. The thickness of the connecting line represents the correlation between the two proteins.



**Figure 8.** The relative expression of the WRKY gene in safflower under drought stress. qRT-PCR data were normalized using the safflower EF-1a and Actin genes. X-axes represent different treatments (0 h, normal conditions; 30 min, 1 h, 2 h, and 3 h indicate hours of drought treatment), and y-axes are scales of relative expression levels. Error bars are based on three biological replicates.



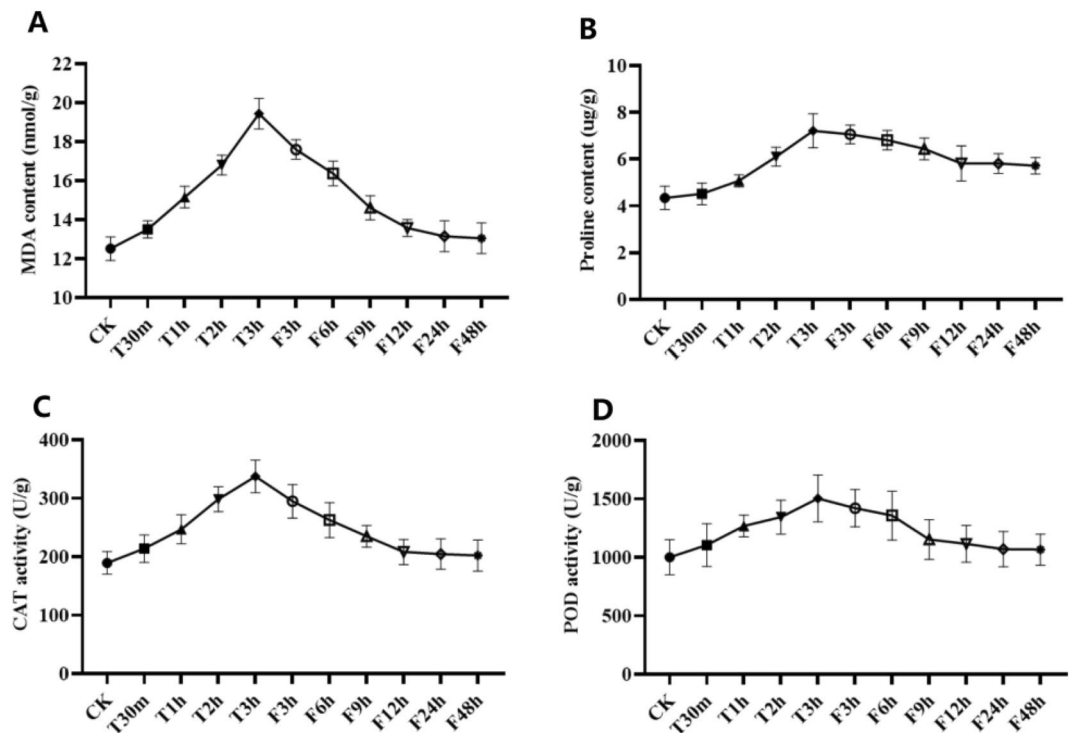
**Figure 9.** Relative expression level of the WRKY gene in safflower under drought-rewatering treatment. qRT-PCR data were normalized using the safflower EF-1a and Actin genes. X-axes represent various treatments (0 h, normal conditions; 3 h, 6 h, 9 h, 12 h, 24 h and 48 h indicate hours of drought-rewatering treatment), and y-axes are scales of relative expression levels. Error bars are based on three biological replicates.

eightfold after rehydration treatment for 9 h. This indicates that these genes may be involved more strongly or more rapidly in the response of plants to drought stress.

### Analysis of proline content, malondialdehyde content and antioxidant enzyme activity in safflower under drought stress

Osmotic stress occurs when plants have reduced water uptake ability under drought. To examine the stress response of safflower under drought, this study examined the proline content, malondialdehyde content, and antioxidant enzyme activities of safflower during drought stress and rehydration treatments (Fig. 10). As depicted in Fig. 10, the contents of Pro and MDA, as well as the activities of CAT and POD, exhibited an upwards trend during the 3-h period of drought stress, peaking at 3 h. However, as the rehydration time increased, there was a





**Figure 10.** Effects of drought stress and rewatering on proline, malondialdehyde and antioxidant enzyme activities in safflower. (A) MDA: malondialdehyde content; (B) Pro: proline content; (C) CAT: catalase activity; (D) POD: superoxide dismutase activity. T30m, T1h, T2h and T3h represent drought stress time. F3h, F6h, F9h, F12h, F24h and F48h represent the rewatering time. Error bars are based on three biological replicates.

subsequent downwards trend in these variables, resulting in an overall increasing and then decreasing pattern. Both MDA and Pro contents reached their highest point after 3 h of drought stress, with respective increases of 1.55- and 1.66-fold compared to the control (Fig. 10A,B). After rehydration, both contents began to decrease and stabilized after 12 h but did not return to the control level. During drought stress, CAT and POD activities continued to increase and reached their peak at 3 h, showing increases of 1.78- and 1.51-fold, respectively, when compared with the control (Fig. 10C,D). After rehydration, their activities began to decline at 12 h and then stabilized, although they remained slightly higher than the control level. These findings suggest that safflower maintained a balance between the production and elimination of ROS during drought stress and rehydration.

## Discussion

The WRKY transcription factor family is a plant-specific supergene family that plays a key role in the growth and development of plants as well as biotic and abiotic stresses. Currently, 64, 54, 138, 102, and 125 members of the WRKY gene family have been identified in strawberry, shantung maple, mandarin orange, flax and maize<sup>1,4,5,8,28</sup>. In this study, a total of 82 *CtWRKY* genes were identified from the safflower genome data, which were unevenly distributed on 12 chromosomes (Fig. 2)<sup>28</sup>. Furthermore, *CtWRKY43*, *-44*, *-45*, and *-46* were grouped together in the identical chromosomal segment, creating a gene cluster, which indicates that these genes might share common biological roles since they are regulated and expressed closely together<sup>5</sup>. The 82 *CtWRKY* genes can be classified into three major groups and five subgroups, which is consistent with the results of previous studies (Fig. 1 and Fig. S1)<sup>3,8</sup>. Group II contains 42 WRKY proteins, which can be further divided into five subgroups according to the variation in protein sequence, suggesting that group II may be the main driving force in the expansion of the WRKY gene family<sup>4-6,8</sup>. Among them, group IIa has only two WRKY proteins, which is probably due to its shorter evolutionary history compared to the other subgroups<sup>5,29</sup>. Furthermore, the WRKY domain at the C-terminus of Group I members was found to contain the same conserved motif as the WRKY domain of group II and group III members, supporting the hypothesis that mutation or deletion of the C-terminal domain of Group I may have been the cause of the evolution of the WRKY domain of Group II and Group III members<sup>5,30,30,31</sup>.

Multiple sequence alignment revealed that most *CtWRKY* proteins have a highly conserved heptapeptide core sequence (WRKYGQK) and zinc finger structure (C<sub>2</sub>H<sub>2</sub> or C<sub>2</sub>HC)<sup>28</sup>. Moreover, three types of heptapeptide sequence variations (WKKYGEK, WRKYGKK and WRKYGHK) were found, indicating that the variation in the domain may be one of the reasons for the expansion of *WRKY* genes in safflower. This phenomenon has been observed in the *WRKY* members of strawberry, *Isatis indigotica*, rice, and apple<sup>28,32-34</sup>. Reportedly, the WRKY domain of *GmWRKY6* and *GmWRKY21* is a heptapeptide variant sequence WRKYGKK, which has lost the ability to bind to the W-box, affecting the expression pattern of stress-responsive genes of WRKY transcription factor

targets<sup>35</sup>. It is speculated that these variations may change the binding specificity of DNA targets and influence the expression status of stress-responsive genes targeted by CtWRKY transcription factors.

Analysis of gene structure and conserved motifs can provide essential information for understanding the evolution of gene families<sup>36</sup>. As shown in Fig. 3, the conserved motifs and exon/intron distribution types of CtWRKY genes in the same group or subgroup have good similarity, indicating that the members of the same group are highly conserved and functionally similar during the evolutionary process, which is very common in plants<sup>5,37,38</sup>. In addition, 10 conserved motifs (length: 6–30 amino acid residues) were identified in 82 CtWRKY genes (Fig. S2 and Table S2). Gene replication events (fragment replication or tandem replication) are thought to be the key force in amplifying gene families in angiosperm evolution<sup>38</sup>. As shown in Fig. 5 and Table S3, 5 tandem duplicated gene pairs and 34 fragment repeated gene pairs were found in CtWRKY TFs, involving 68 CtWRKY genes, accounting for 82.93% of the members of the safflower WRKY family and indicating that fragment duplication events play a crucial role in the expansion of the CtWRKY gene family. Normally, a ratio of 1 indicates neutral selection; a Ka/Ks ratio > 1 means adaptive evolution with positive selection, while a ratio < 1 means negative or purifying selection<sup>40,41</sup>. As a result, all the Ka/Ks ratios for the 34 gene pairs were < 1 (Table S3), indicating that the CtWRKY gene family is highly conserved during evolution, consistent with previous research<sup>5</sup>.

Based on the interaction analysis of WRKY proteins in other species, the potential regulation of CtWRKY proteins, gene function and its relationship with the overall biological function can be predicted. In this study, the STRING database was used to construct a CtWRKY protein interaction network using *Arabidopsis thaliana* as a template. As illustrated in Fig. 7, four CtWRKY proteins (CtWRKY15, CtWRKY17, CtWRKY55 and CtWRKY79) exhibited high sequence similarity with AtWRKY33, AtWRKY40 and AtWRKY55 and were identified as central nodes within the interaction network. These proteins displayed varying degrees of interaction with others. Reportedly, AtWRKY40 can form dimers with AtWRKY18 or AtWRKY60 to modify pathogen resistance, regulate the osmotic and antioxidant capacity of plants, and improve the drought tolerance of plants<sup>42–44</sup>. AtWRKY33 can enhance *Arabidopsis* resistance to fungal pathogens and improve *Arabidopsis* resistance to low temperature and high salt stress<sup>45–47</sup>. AtWRKY53 can regulate stomatal movement by reducing hydrogen peroxide content or promoting starch degradation, thereby regulating drought tolerance in *Arabidopsis*<sup>48,49</sup>. It is widely accepted that homologous proteins with similar domains and sequences across different species may possess identical or similar functions<sup>5</sup>. Hence, it is conjectured that CtWRKY15, CtWRKY17, CtWRKY55, and CtWRKY79 might participate in a transcriptional regulatory mechanism, along with others, in the response of safflower to drought stress. In summary, various strong or weak interactions exist among CtWRKY proteins, indicating potential coexpression in response to biotic and abiotic stresses, with consequent regulation of safflower stress resistance.

In this study, we analysed transcriptome data for safflower retrieved from the NCBI database. The results showed that 82 CtWRKY genes had tissue-specific expression patterns under normal and stress treatments (Fig. 6), which was consistent with the results in other plants, such as *Acer truncatum*, *Linum usitatissimum* L., *Platycodon grandiflorum* and *Zea mays*<sup>1,5,8,50–53</sup>. However, the expression levels of some of these genes were low, suggesting that these genes may be pseudogenes or may have specific spatial and temporal expression patterns that were not investigated in this database<sup>5,8</sup>. As shown in Fig. 6B, 7 CtWRKY genes (CtWRKY9, –15, –17, –37, –54, –55, and –72) were highly expressed in leaves. Moreover, 4 CtWRKY genes (CtWRKY15, –54, –55, and –72) had high expression levels in different tissues and were induced by EBG and light stress, indicating that these genes may play an important role in specific tissues and abiotic stresses<sup>8,28</sup>. For instance, ZmWRKY79-mediated drought tolerance positively depends on an increase of ABA levels by triggering ABA biosynthetic genes including AAO3 genes<sup>13</sup>. AtWRKY57 stimulates the expression of RD29A and NCED3 by attaching to its W-box element, thereby boosting drought tolerance in *Arabidopsis*<sup>54</sup>. TaWRKY51 is a positive regulator of root architecture and grain yield-contributing traits<sup>55</sup>. These findings indicate that WRKY transcription factors display diverse expression profiles in different organs or tissues, regulating a range of biological and physiological metabolic processes in safflower<sup>1,3,8</sup>. The cis-acting element is an important “switch” used to regulate gene expression in plants. Promoter analysis has identified numerous abiotic stress response cis-elements in the promoter regions of WRKY genes across multiple species, indicating that most plant WRKY genes may play a role in controlling transcriptional responses to abiotic stress<sup>25,26</sup>. In this study, a variety of cis-acting elements, such as light-responsive elements, abiotic stress-responsive elements, and hormone-responsive elements, were identified in the promoter region of the CtWRKY gene, which is important for the study of the potential function of the CtWRKY gene and is in agreement with the results of previous studies (Fig. 4 and Fig. S3). Previous research has demonstrated that WRKY transcription factors participate in the response to abiotic stress via stress-related transcriptional regulatory elements. For example, various environmental factors, including drought, salinity, and high temperature, alongside defence-related plant hormones such as SA and ABA, can trigger the gene transcription of ZmWRKY65, thereby increasing its tolerance to abiotic stress<sup>56</sup>. The expression of the McWRKY57 gene was stimulated by mannitol, ABA, and MeJA. Overexpression of McWRKY57-like in *Arabidopsis* plants considerably improved plant drought tolerance<sup>57</sup>. Additionally, transgenic *Arabidopsis* lines overexpressing of EjWRKY17 exhibited increased cotyledon greening and root elongation under abscisic acid (ABA) treatment<sup>58</sup>. In summary, the research findings offer new insights for safflower in uncovering the signalling pathways in response to abiotic stress.

WRKY transcription factors are crucial in plant responses to drought, high temperatures, and salt stress<sup>6,7,50,51</sup>. It has been reported that SbWRKY30 is expressed predominantly in sorghum leaves and roots. Overexpression of this gene in *Arabidopsis* has been shown to enhance drought tolerance, likely through interference with root structure and induced expression of the AtRD19 gene<sup>12</sup>. The GhWRKY91-transformed plants displayed enhanced drought tolerance and delayed senescence of leaves under drought conditions. This was associated with increased expression of stress-responsive genes and decreased expression of senescence-associated genes (SAGs)<sup>59</sup>. To study the stress response of the CtWRKY gene under drought stress, seven CtWRKY genes with high homology to *Arabidopsis* drought tolerance genes were selected (group II: CtWRKY9, –54, –55, and –72;

group III: *CtWRKY15*, *-17*, and *-37*). qRT-PCR was used to analyse their expression patterns in different tissues and under drought stress and drought rewatering treatments. As shown in Figs. 8 and 9, the relative expression levels of these *CtWRKY* genes were differentially upregulated from 0 to 3 h under drought stress, showing a trend of increasing and then decreasing, suggesting that these genes may be crucial at the early stage of drought stress treatment and have a positive or negative regulatory effect on drought stress; however, their regulation varies with the degree of stress. Under drought-rehydration treatment, the expression levels of *CtWRKY* genes were significantly upregulated after 9 h. However, following rehydration, the downregulation of *CtWRKY* genes was relatively minor. These results suggest that these genes may play a role in the plant late-stage recovery from drought stress. This difference in the levels of expression before and after drought stress indicates that WRKY transcription factors respond to drought stress in a time-sensitive manner, consistent with previous studies<sup>60,61</sup>. The expression level of *ZmWRKY40* can peak rapidly at 10 h after drought treatment, and its overexpression can improve drought tolerance in transgenic *Arabidopsis* through the regulation of stress-related genes (*DREB2B* and *RD29A*)<sup>62</sup>. In addition, the expression levels of *CtWRKY9* and *CtWRKY55* increased rapidly under drought and rehydration treatments compared to the control, reaching a peak of more than eightfold, implying a significant role in the response to drought stress. Such observations suggest that these genes should be explored further to understand their potential as key regulators under drought conditions. However, the underlying mechanism of this phenomenon remains to be further elucidated. Combined with the results of functional studies of the WRKY family in many plants, it can be speculated that more *CtWRKY* genes have regulatory functions related to abiotic stress that are waiting to be discovered.

Drought stress exposes plants to osmotic stress, which leads to an increase in ROS production, resulting in peroxidation of plant cell membranes. Plants can actively accumulate osmotic solutes or antioxidant enzymes to reduce the damage caused by drought<sup>63,64</sup>. Wei et al. reported higher antioxidant enzyme activity in drought-tolerant safflower varieties under drought stress<sup>65</sup>. In this study, the Pro content, MDA content, CAT activity and POD activity of safflower under drought stress and rehydration treatment were determined. As shown in Fig. 10, the contents of Pro, MDA, CAT and POD in safflower under drought stress and rehydration treatments showed a trend of first increasing and then decreasing, which was similar to the findings in oak<sup>66</sup>. However, after rewatering, the MDA level did not decrease to the level of the control sample, indicating that drought stress can cause irreversible damage to plants. Furthermore, the levels of CAT and POD were still slightly higher than those in the control after 48 h, indicating that drought stress may cause damage to the antioxidant oxidase system. Çulha Erdal et al. found that drought stress had a negative effect on proline content, MDA content and antioxidant enzyme activity in safflower, which was difficult to fully recover even after rewatering<sup>67</sup>. Notably, the expression level of the *CtWRKY* gene was high after drought stress. In addition, the expression level of genes was also significantly upregulated after rewatering, showing a trend of first increasing and then decreasing. In conclusion, it is speculated that the *CtWRKY* gene may regulate the ROS level of plants and enhance the drought resistance of plants.

## Conclusion

Safflower is an important economic crop with medicinal and food utility. In this study, 82 *CtWRKY* transcription factors were identified from the safflower genome, and bioinformatics and gene expression pattern analysis were performed to reveal the important functions of *CtWRKY* genes in plant growth and development and abiotic stress response under drought treatment. The qRT-PCR analysis showed that seven *CtWRKY* genes were responsive to drought stress treatment, indicating that they had positive or negative regulatory effects on drought stress. In addition, *CtWRKY9* and *CtWRKY55* were found to be more involved or more rapidly involved in the drought stress response process, providing a basis for the future mining of candidate genes related to drought tolerance in safflower. The results provide a theoretical basis for further exploring the function of the *CtWRKY* gene in safflower growth and development and response to abiotic stress.

## Materials and methods

### Genome-wide identification and physicochemical analysis of the safflower WRKY gene family

Our study complies with relevant institutional, national, and international guidelines and legislation. The whole-genome sequence, CDS, protein sequence and annotation files for safflower were provided by the Institute of Economic Crops, Xinjiang Academy of Agricultural Sciences. Identification of the safflower WRKY gene family was performed using two methods: 1. The protein sequence of the AtWRKY gene family was downloaded from the NCBI database (<https://www.ncbi.nlm.nih.gov/>). The protein sequence of the AtWRKY gene family was used as a template, and the protein sequence of the safflower WRKY gene family was preliminarily screened by the Basic Local Alignment Search Tool (BLAST) of TBtools software<sup>1</sup>. 2. The hidden Markov model (HMM) file of the WRKY domain (PF03106), downloaded from the Pfam protein family database (<http://pfam.sanger.ac.uk/>), was used for the identification of WRKY genes from the safflower genomic database by TBtools<sup>5,8,68</sup>. Subsequently, the online tools Pfam (<http://pfam.xfam.org/search#tabview=Table> 1), SMART (<http://smart.embl.de/>) and NCBI CDD (<https://www.ncbi.nlm.nih.gov/cdd/>) were used to verify whether the *CtWRKY* protein had a conserved WRKY domain, and the candidate protein sequences without conserved domains were removed<sup>8</sup>. The ExPASy website (<https://www.ExPASy.org/>) was used to predict the physical and chemical properties of *CtWRKY* proteins, including molecular weight, isoelectric point and hydrophilicity, and the parameters were set as default values<sup>69</sup>. The online tool WoLF PSORT (<https://wolfpsort.hgc.jp/>) was used to predict the subcellular localization of *CtWRKY* family members<sup>70</sup>.

### Phylogenetic analysis and chromosome mapping of the safflower WRKY gene family

The whole genome, protein sequence, annotation file, and protein sequence of the WRKY gene family were downloaded from The *Arabidopsis* Information Resource (TAIR) database (<https://www.arabidopsis.org/>)<sup>71</sup>.

DNAMAN 9.0 software was used for multiple sequence alignment of AtWRKY gene family proteins in safflower and *Arabidopsis*. The phylogenetic tree was constructed by the neighbour-joining (NJ) method using MEGA 7.0 software, with 1000 bootstrap replicates<sup>72</sup>. The chromosome location map of the CtWRKY gene family was drawn by TBtools software according to the gene annotation file of the safflower genome.

### Gene structure, promoter and collinearity analysis of the safflower WRKY gene family

The intron and exon characteristics of the CtWRKY gene were analysed by TBtools software according to the annotation information of the CtWRKY gene<sup>73</sup>. The conserved motifs of the CtWRKY protein were identified using the MEME online tool (<http://meme.sdsc.edu/meme/>)<sup>74</sup>. The following parameters were used: any number of repetitions, maximum of 10 motifs, and an optimum motif width of 6 to 150 amino acid residues. The gene structure and conserved motif results were visualized using TBtools software.

The sequence of the promoter region 2000 bp upstream of the CDS region of the CtWRKY gene was extracted using TBtools software, and the online tool PlantCARE (<http://bioinformatics.psb.ugent.be/webtools/plantcare/html/>) predicted the cis-acting elements of the CtWRKY gene and analysed the stress and hormone response elements<sup>1,5,8,75</sup>. Collinearity data of CtWRKY genes were obtained using the MCScanX tool, and the fragment replication and tandem replication genes were identified according to the safflower genome data<sup>76</sup>. Determination of gene duplication events was performed according to the following criterion: If two genes were located on the same chromosomal region and were adjacent to each other or separated by one gene, they were considered tandemly duplicated genes<sup>39,77</sup>. Based on the CtWRKY protein and CDSs, the Ka, Ks and Ka/Ks values of the CtWRKY gene were calculated using the Simple Ka/Ks Calculator of TBtools software.

### Transcriptome analysis of the safflower WRKY gene family

The RNA-seq data for safflower under the categories of seed, leaf, flower, epibrassinolide (EBR) and light stress were downloaded from the NCBI-SRA database (Biological project: PRJNA76135; PRJNA628030; PRJNA493984)<sup>91</sup>. The light intensity was 40,000 lx (high light) and 20,000 lx (medium light), and the illumination time was 10 h/d. EBR treatment conditions: 0.1 µmol/L EBR solution was sprayed evenly on the front and back of the leaves. The expression patterns of CtWRKY gene in different tissues such as roots, stems, leaves and flowers were calculated by TBtools software. The expression values were calculated by log<sub>2</sub> (TPM: transcripts per million) and were displayed as a heatmap by TBtools software<sup>68</sup>.

### Protein interaction network analysis of the safflower WRKY gene family

The CtWRKY protein–protein interaction (PPI) network was constructed using the Search Tool for the Retrieval of Interacting Genes/Proteins (STRING) database (<https://string-db.org/>) with the *Arabidopsis* WRKY protein as a template<sup>5,78</sup>. According to the protein–protein interaction score ( $\geq 0.400$ ), the number of interactions was 5. The thicker the line between the targets in the interaction network, the stronger the interaction. The results were imported into Cytoscape software to visualize the protein–protein interaction network<sup>79</sup>.

### Plant materials, drought stress treatment and tissue collection

The safflower seeds (21TY012) used in this experiment were provided by the Institute of Economic Crops, Xinjiang Academy of Agricultural Sciences. Safflower seeds were planted in pots and cultivated under ambient temperature of 25 °C and 16 h/8 h (light/dark) light conditions for 14 days. Safflower seedlings with uniform growth were selected, and their roots were soaked in 1/2 Hoagland nutrient solution for 9 days. The roots were then soaked in 11% PEG-6000 solution to simulate drought stress. With 0 h treatment as a control, seedling samples were collected at 0 h, 30 min, 1 h, 2 h and 3 h after drought treatment and 3 h, 6 h, 9 h, 12 h, 24 h and 48 h after rehydration. Roots, stems, and leaves were immediately frozen in liquid nitrogen and stored at –80 °C until RNA extraction.

### Analysis of safflower WRKY gene expression under drought stress by qRT-PCR

The RNAprep Pure plant kit (TIANGEN, China) was used to extract total RNA from different treated seedling samples according to the manufacturer's instructions, and the concentration and quality of RNA were determined by spectrophotometry. Then, the RNAs were reverse transcribed into first-strand cDNA by a FastKing RT Kit (with gDNase) (TIANGEN, China) and stored at –20 °C.

Seven CtWRKY genes (*CtWRKY9*, –15, –17, –37, –54, –55 and –72) with high homology to *Arabidopsis* drought tolerance genes *AtWRKY11*, –17, –30, –40, –46, –53 and –57 were selected for qRT-PCR analysis<sup>11,42,43,48,54,80,81</sup>. The reference genes were *EF-1a* (Accession No.: KJ634806) and *Actin* (Accession No.: KJ634809). Gene-specific primers were designed using Primer 5.0 software, and the primer sequences are shown in Table S5. DNAMAN 9.0 software was used to perform multiple sequence alignment of 9 candidate CtWRKY proteins with drought-resistant WRKY proteins from *Arabidopsis thaliana*, *Chrysanthemum morifolium* Ramat, *Helianthus annuus* L. and *Triticum aestivum* L.<sup>59,62,82–90</sup>

Quantitative real-time PCR was performed using a CFX96 Touch™ Real-Time PCR detection system (Bio-Rad, USA). The qRT-PCR amplification program and reaction system configuration were performed according to the manufacturer's instructions for SuperReal PreMix Plus (SYBR Green) (TIANGEN, China). The quantitative data were calculated using the  $2^{-\Delta\Delta Ct}$  method<sup>92</sup>. The amplification conditions were as follows: initial denaturation at 95 °C for 15 min and 40 cycles of denaturation at 95 °C for 10 s, annealing at 50–60 °C for 20 s and extension at 72 °C for 20 s. Finally, a melting curve was used to check the amplification specificity.

## Determination of antioxidant enzyme activity and proline and malondialdehyde (MDA) contents

The CAT activity was determined by ammonium molybdate colorimetry (Boxbio, AKAO003 Assay Kit, China). The POD activity was determined by the guaiacol method (Boxbio, AKAO005C Assay Kit, China). The MDA content was determined by thiobarbituric acid colorimetry (Boxbio, AKFA013C Assay Kit, China). The Pro content was measured using the ninhydrin colorimetric method (Boxbio, AKAM003C Assay Kit, China).

## Data availability

RNA-seq fastq files were deposited in NCBI Sequence Read Archive (SRA) under accession number PRJNA76135, PRJNA628030 and PRJNA493984. Genome, proteome, and annotation files were available in *Carthamus tinctorius* database (<https://safflower.scuec.edu.cn/>).

Received: 19 May 2023; Accepted: 6 October 2023

Published online: 07 October 2023

## References

- Li, Y. *et al.* Genome-wide identification and analysis of the WRKY gene family and cold stress response in *Acer truncatum*. *Genes* **12**, 1867 (2021).
- Ishiguro, S. & Nakamura, K. Characterization of a cDNA encoding a novel DNA-binding protein, SPF1, that recognizes SP8 sequences in the 5' upstream regions of genes coding for sporamin and beta-amylase from sweet potato. *Mol. Gen. Genet.* **244**, 563–571 (1994).
- Cheng, Y. F. Identification of the WRKY gene family and characterization of stress-responsive genes in *Taraxacum kok-saghyz* Rodin. *Int. J. Mol. Sci.* **23**, 10270 (2022).
- Muhammad, A. K. *et al.* Characterization of WRKY gene family in whole-genome and exploration of flowering improvement genes in *Chrysanthemum lavandulifolium*. *Front. Plant Sci.* **13**, 861193 (2022).
- Hu, W. J. *et al.* Genome-wide identification and analysis of WRKY gene family in maize provide insights into regulatory network in response to abiotic stresses. *BMC Plant Biol.* **21**, 427 (2021).
- Khoso, M. A. *et al.* WRKY transcription factors (TFs): Molecular switches to regulate drought, temperature, and salinity stresses in plants. *Front. Plant Sci.* **13**, 1039329 (2022).
- Wani, S. H. *et al.* WRKY transcription factors and plant defense responses: Latest discoveries and future prospects. *Plant Cell Rep.* **40**, 1071–1085 (2021).
- Yuan, H. M. *et al.* Genome-wide identification and expression analysis of the WRKY transcription factor family in flax (*Linum usitatissimum* L.). *BMC Genom.* **22**, 375 (2021).
- Wang, J. Y. *et al.* GhWRKY21 regulates ABA-mediated drought tolerance by fine-tuning the expression of GhHAB in cotton. *Plant Cell Rep.* **40**, 2135–2150 (2021).
- Liu, G. Y. *et al.* MaWRKY80 positively regulates plant drought stress resistance through modulation of abscisic acid and redox metabolism. *Plant Physiol. Biochem.* **156**, 155–166 (2020).
- Ali, M. A. *et al.* Transcription factors WRKY11 and WRKY17 are involved in abiotic stress responses in *Arabidopsis*. *J. Plant Physiol.* **226**, 12–21 (2018).
- Yang, Z. *et al.* SbWRKY30 enhances the drought tolerance of plants and regulates a drought stress-responsive gene, SbRD19, in sorghum. *J. Plant Physiol.* **246**, 153142 (2020).
- Gulzar, F. *et al.* Maize WRKY transcription factor ZmWRKY79 positively regulates drought tolerance through elevating ABA biosynthesis. *Int. J. Mol. Sci.* **22**, 10080 (2021).
- Zhu, L. *et al.* Overexpression of watermelon CIWRKY20 in transgenic *Arabidopsis* improves salt and low-temperature tolerance. *Sci. Hortic.* **295**, 110848 (2022).
- Long, L. X. *et al.* Progress in the understanding of WRKY transcription factors in woody plants. *Int. J. Biol. Macromol.* **242**, 124379 (2023).
- Zhang, Y. L. *et al.* WRKY genes improve drought tolerance in *Arachis duranensis*. *Front. Plant Sci.* **13**, 910408 (2022).
- An, X. *et al.* Bioinformatics analysis of WRKY family genes in flax (*Linum usitatissimum*). *Life (Basel)* **13**, 1258 (2023).
- Huang, J. J. *et al.* The WRKY transcription factor OsWRKY54 is involved in salt tolerance in rice. *Int. J. Mol. Sci.* **23**, 11999 (2022).
- Zhang, X. L. *et al.* Pharmacological actions, molecular mechanisms, pharmacokinetic progressions, and clinical applications of hydroxysafflor yellow A in antidiabetic research. *J. Immunol. Res.* **2021**, 4560012 (2021).
- Lu, P. *et al.* Safflower extract inhibits adp-induced human platelet aggregation. *Plants (Basel)* **10**, 1192 (2021).
- Lin, H. *et al.* Species and distribution of safflower germplasm resources in China. *Biot. Resour.* **40**, 314–320 (2018).
- Wang, P. Q. *et al.* Evaluation of drought resistance of medicinal safflower in seedling stage. *Modern Chin. Med.* **23**, 660–665 (2021).
- Wu, Z. H. *et al.* The chromosome-scale reference genome of safflower (*Carthamus tinctorius*) provides insights into linoleic acid and flavonoid biosynthesis. *Plant Biotechnol. J.* **19**, 1725–1742 (2021).
- Çulha, E. Ş., Eyidoğan, F. & Ekmekçi, Y. Comparative physiological and proteomic analysis of cultivated and wild safflower response to drought stress and re-watering. *Physiol. Mol. Biol. Plants.* **27**, 281–295 (2021).
- Hong, Y. Q. *et al.* Genome-wide analysis and transcriptional reprogrammings of MYB superfamily revealed positive insights into abiotic stress responses and anthocyanin accumulation in *Carthamus tinctorius* L. *Mol. Genet. Genom.* **297**, 125–145 (2022).
- Hong, Y. Q. *et al.* Genome-wide identification, expression analysis, and subcellular localization of *Carthamus tinctorius* bHLH transcription factors. *Int. J. Mol. Sci.* **20**, 3044 (2019).
- Li, H. Y. *et al.* Genome-wide identification and expression analysis of bZIP gene family in *Carthamus tinctorius* L. *Sci. Rep.* **10**, 15521 (2020).
- Garrido-Gala, J. *et al.* A comprehensive study of the WRKY transcription factor family in strawberry. *Plants (Base)* **11**, 1585 (2022).
- Zhang, Y. X., Liu, Z. & Zhang, T. Q. Genome-wide identification and expression analysis of WRKY gene family in water spinach (*Ipomoea aquatica*). *Mol Plant Breed.* **21**, 55–66 (2023).
- Ding, X. Q., Hu, M. J. & Xiang, S. Bioinformatics analysis of WRKY gene family in *Kandelia cande*. *Jiangsu Agric. Sci.* **50**, 50–60 (2022).
- Xiang, Y. T., Wang, J. & Zeng, J. L. Genome-wide identification and bioinformatics analysis of WRKY transcription factor family in sunflower (*Helianthus annuus*). *Mol Plant Breed.* **18**, 4572–4586 (2020).
- Qu, R. J. *et al.* Identification and expression analysis of the WRKY gene family in *Isatis indigotica*. *Gene* **783**, 145561 (2021).
- Abdullah-Zawawi, M. R. *et al.* Comparative genome-wide analysis of WRKY, MADS-box and MYB transcription factor families in *Arabidopsis* and rice. *Sci. Rep.* **11**, 19678 (2021).
- Meng, D. *et al.* Genome-wide identification and characterization of WRKY transcriptional factor family in apple and analysis of their responses to waterlogging and drought stress. *Plant Physiol. Biochem.* **103**, 71–83 (2016).

35. Zhou, Q. Y. *et al.* Soybean WRKY-type transcription factor genes, *GmWRKY13*, *GmWRKY21*, and *GmWRKY54*, confer differential tolerance to abiotic stresses in transgenic *Arabidopsis* plants. *Plant Biotechnol. J.* **6**, 486–503 (2008).
36. Hu, B. *et al.* GSDS 2.0: an upgraded gene feature visualization server. *Bioinformatics* **31**, 1296–1297 (2015).
37. Waqas, M. *et al.* Genome-wide identification and expression analyses of WRKY transcription factor family members from chickpea (*Cicer arietinum* L.) reveal their role in abiotic stress-responses. *Genes Gen.* **41**, 467–81 (2019).
38. Ling, J. *et al.* Genome-wide analysis of WRKY gene family in *Cucumis sativus*. *BMC Genom.* **12**, 471 (2011).
39. De, G. A., Lanave, C. & Saccone, C. Genome duplication and gene-family evolution: The case of three OXPPOS gene families. *Gene* **421**, 1–6 (2008).
40. Ke, Y. J. *et al.* Genome-Wide Identification of the MYB gene family in *cymbidiumensifolium* and its expression analysis in different flower colors. *Int. J. Mol. Sci.* **22**, 13245 (2021).
41. Nekrutenko, A. *et al.* The K(A)/K(S) ratio test for assessing the protein-coding potential of genomic regions: An empirical and simulation study. *Genome Res.* **12**, 198–202 (2002).
42. Xu, X. P. *et al.* Physical and functional interactions between pathogen-induced *Arabidopsis WRKY18*, *WRKY40*, and *WRKY60* transcription factors. *Plant Cell.* **18**, 1310–1326 (2006).
43. Che, Y. M. *et al.* *AtWRKY40* functions in drought stress response in *Arabidopsis thaliana*. *Plant Physiol. J.* **54**, 456–464 (2018).
44. Liu, Z. Q. *et al.* A conserved double-W box in the promoter of *CaWRKY40* mediates autoregulation during response to pathogen attack and heat stress in pepper. *Mol. Plant Pathol.* **22**, 3–18 (2021).
45. Fu, Q. T. & Yu, D. Expression profiles of *AtWRKY25*, *AtWRKY26* and *AtWRKY33* under abiotic stresses. *Hereditas (Beijing)* **32**, 848–856 (2010).
46. Zheng, Z. Y. *et al.* *Arabidopsis WRKY33* transcription factor is required for resistance to necrotrophic fungal pathogens. *Plant J.* **48**, 592–605 (2006).
47. Zhou, J. *et al.* Characterization of the promoter and extended C-terminal domain of *Arabidopsis WRKY33* and functional analysis of tomato *WRKY33* homologues in plant stress responses. *J. Exp. Bot.* **66**, 4567–4583 (2015).
48. Sun, Y. D. & Yu, D. Q. Activated expression of *AtWRKY53* negatively regulates drought tolerance by mediating stomatal movement. *Plant Cell Rep.* **34**, 1295–1306 (2015).
49. Zhao, M. M. *et al.* A WRKY transcription factor, *TaWRKY42-B*, facilitates initiation of leaf senescence by promoting jasmonic acid biosynthesis. *BMC Plant Biol.* **20**, 444 (2020).
50. Cheng, Z. Y. *et al.* WRKY transcription factor response to high-temperature stress. *Plants (Basel)* **10**, 2211 (2021).
51. Song, H. *et al.* Review: WRKY transcription factors: Understanding the functional divergence. *Plant sci.* **334**, 111770 (2023).
52. Goyal, P. *et al.* WRKY transcription factors: evolution, regulation, and functional diversity in plants. *Protoplasma* **260**, 331–348 (2023).
53. Li, J. *et al.* Transcriptome-wide identification of WRKY transcription factors and their expression profiles in response to methyl jasmonate in *Platycodon grandiflorus*. *Plant Signal Behav.* **17**, 2089473 (2022).
54. Jiang, Y. J., Liang, G. & Yu, D. Q. Activated expression of *WRKY57* confers drought tolerance in *Arabidopsis*. *Mol. Plant.* **5**, 1375–1388 (2012).
55. Li, Y. Y. *et al.* Transcription factor *TaWRKY51* is a positive regulator in root architecture and grain yield contributing traits. *Front. Plant Sci.* **12**, 734614 (2021).
56. Huo, T. *et al.* Overexpression of *ZmWRKY65* transcription factor from maize confers stress resistances in transgenic *Arabidopsis*. *Sci. Rep.* **11**, 4024 (2021).
57. Bai, Y. *et al.* Overexpression of a WRKY transcription factor *McWRKY57*-like from *Mentha canadensis* L. enhances drought tolerance in transgenic *Arabidopsis*. *BMC Plant Biol.* **23**, 216 (2023).
58. Wang, D. *et al.* A WRKY transcription factor, *EjWRKY17*, from *eriobotrya japonica* enhances drought tolerance in transgenic *Arabidopsis*. *Int. J. Mol. Sci.* **22**, 5593 (2021).
59. Gu, L. J. *et al.* The cotton *GhWRKY91* transcription factor mediates leaf senescence and responses to drought stress in transgenic *Arabidopsis thaliana*. *Front. Plant Sci.* **10**, 1352 (2019).
60. Bao, F. *et al.* Genome-wide analysis of members of the WRKY gene family and their cold stress response in *Prunus mume*. *Genes* **10**, 911 (2019).
61. He, Y. J. *et al.* Genome-wide identification and expression analysis of WRKY transcription factors under multiple stresses in *Brassica napus*. *PLoS ONE* **11**, e0157558 (2016).
62. Wang, C. T. *et al.* The maize WRKY transcription factor *ZmWRKY40* confers drought resistance in transgenic *Arabidopsis*. *Int. J. Mol. Sci.* **19**, 2580 (2018).
63. Gao, S. S. *et al.* Effects of drought stress on growth, physiology and secondary metabolites of two adonis species in Northeast China. *Sci. Hortic.* **259**, 108795 (2020).
64. Zhang, X. *et al.* Effects of drought stress and water recovery on physiological responses and gene expression in maize seedlings. *BMC Plant Biol.* **18**, 1–16 (2018).
65. Wei, B. *et al.* Integrating transcriptomics and metabolomics to studies key metabolism, pathways and candidate genes associated with drought-tolerance in *Carthamus tinctorius* L. under drought stress. *Ind. Crop Prod.* **151**, 112465 (2020).
66. Xiong, S. F. *et al.* Effects of drought stress and rehydration on physiological and biochemical properties of Four Oak Species in China. *Plants (Basel)* **11**, 679 (2022).
67. Çulha Erdal, Ş., Eyidoğan, F. & Ekmekçi, Y. Comparative physiological and proteomic analysis of cultivated and wild safflower response to drought stress and re-watering. *Physiol. Mol. Biol. Plants.* **27**, 281–295 (2021).
68. Chen, C. J. *et al.* TBtools: An integrative toolkit developed for interactive analyses of big biological data. *Mol. Plant.* **13**, 1194–1202 (2020).
69. Gasteiger, E. *et al.* ExpASY: The proteomics server for in-depth protein knowledge and analysis. *Nucleic Acids Res.* **31**, 3784–3788 (2003).
70. Horton, P. *et al.* WoLF PSORT: Protein localization predictor. *Nucleic Acids Res.* **35**, W585–W587 (2007).
71. Swarbreck, D. *et al.* The *Arabidopsis* Information Resource (TAIR): Gene structure and function annotation. *Nucleic Acids Res.* **36**, D1009–D1014 (2008).
72. Kumar, S. *et al.* MEGA X: Molecular evolutionary genetics analysis across computing platforms. *Mol. Biol. Evol.* **35**, 1547–1549 (2018).
73. Gao, S. Q. *et al.* Genome-wide identification of the histone acetyltransferase gene family in *Triticum aestivum*. *BMC Genom.* **22**, 49 (2021).
74. Viana, V. E. *et al.* When rice gets the chills: Comparative transcriptome profiling at germination shows WRKY transcription factor responses. *Plant Biol.* **23**, 100–112 (2021).
75. Lescot, M. *et al.* 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**, 325–327 (2002).
76. Wang, Y. P. *et al.* MCSanX: a toolkit for detection and evolutionary analysis of gene synteny and collinearity. *Nucleic Acids Res.* **40**, e49 (2012).
77. Wang, Y. Y. *et al.* Genome-wide Identification and expression analysis of NF-YA gene family in *Gossypium arboreum*. *Mol. Plant Breed.* **19**, 4564–4573 (2021).

78. Szklarczyk, D. *et al.* The STRING database in 2021: Customizable protein-protein networks, and functional characterization of user-uploaded gene/measurement sets. *Nucleic Acids Res.* **49**, D605–D612 (2021).
79. Doncheva, N. T. *et al.* Cytoscape StringApp: Network analysis and visualization of proteomics data. *J. Proteome Res.* **18**, 623–632 (2019).
80. El-Esawi, M. A. *et al.* Overexpression of AtWRKY30 transcription factor enhances heat and drought stress tolerance in wheat (*Triticum aestivum* L.). *Genes.* **10**, 136 (2019).
81. Chen, J. N. *et al.* *Arabidopsis* WRKY46, WRKY54, and WRKY70 transcription factors are involved in brassinosteroid-regulated plant growth and drought responses. *Plant Cell.* **29**, 1425–1439 (2017).
82. He, G. H. *et al.* Drought-responsive WRKY transcription factor genes *TaWRKY1* and *TaWRKY33* from wheat confer drought and/or heat resistance in *Arabidopsis*. *BMC Plant Biol.* **16**, 116 (2016).
83. Ma, J. H. *et al.* Overexpression of *TaWRKY146* increases drought tolerance through inducing stomatal closure in *Arabidopsis thaliana*. *Front. Plant Sci.* **8**, 2036 (2017).
84. Wei, W. *et al.* *GmWRKY54* improves drought tolerance through activating genes in abscisic acid and Ca<sup>2+</sup> signaling pathways in transgenic soybean. *Plant J.* **100**, 384–398 (2019).
85. Shi, W. Y. *et al.* The WRKY transcription factor *GmWRKY12* confers drought and salt tolerance in soybean. *Int. J. Mol. Sci.* **19**, 4087 (2018).
86. Fan, Q. Q. *et al.* *CmWRKY1* enhances the dehydration tolerance of chrysanthemum through the regulation of ABA-Associated genes. *PLoS One* **11**, e0150572 (2016).
87. Jaffar, M. A. *et al.* Involvement of *CmWRKY10* in drought tolerance of chrysanthemum through the ABA-signaling pathway. *Int. J. Mol. Sci.* **17**, 693 (2016).
88. Liu, H. Y. *et al.* Ectopic expression of a grapevine transcription factor *VvWRKY11* contributes to osmotic stress tolerance in *Arabidopsis*. *Mol. Biol. Rep.* **38**, 417–427 (2011).
89. Raineri, J., Ribichich, K. F. & Chan, R. L. The sunflower transcription factor *HaWRKY76* confers drought and flood tolerance to *Arabidopsis thaliana* plants without yield penalty. *Plant Cell Rep.* **34**, 2065–2080 (2015).
90. Chu, X. Q. *et al.* The cotton WRKY Gene *GhWRKY41* positively regulates salt and drought stress tolerance in transgenic nicotiana benthamiana. *PLoS One* **10**, e0143022 (2015).
91. Jiang, C. *pCtFSG2*, a flower-specific promoter with suitable promoter activity in safflower. *J. Plant Biochem. Biotechnol.* **32**, 478–486 (2023).
92. Ma, H. Y. *et al.* qPCR and qRT-PCR analysis: Regulatory points to consider when conducting biodistribution and vector shedding studies. *Mol. Ther. Methods Clin. Dev.* **20**, 152–168 (2020).

## Acknowledgements

This work was financially supported by the youth fund project “Identification of WRKY gene family in safflower and its drought response model” (2021D01B61).

## Author contributions

X.S. submitted the manuscript. X.S., X.H., Y.Z. and D.J. designed the experiments and analysed the data. X.S., X.H., Y.Z., Q.L. and D.J. wrote the manuscript text. X.S., Q.L. and H.M. conducted the experiments, X.S. and Y.G. analysed the results. All authors reviewed the manuscript.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1038/s41598-023-44340-y>.

**Correspondence** and requests for materials should be addressed to Y.Z., D.J. or Q.L.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher’s note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023