Identification of yellowhorn (*Xanthoceras sorbifolium*) WRKY transcription factor family and analysis of abiotic stress response model

Zhi Liu¹,² · Saiyinduleng³ · Qiaoying Chang¹,² · Chuwen Cheng⁴ · Zhimin Zheng¹ · Song Yu¹

Received: 24 December 2019 / Accepted: 18 February 2020 / Published online: 7 May 2020
© The Author(s) 2020

Abstract WRKY transcription factors are widely distributed in higher plants and play important roles in many biological processes, including stress resistance. The recently published genome sequence of yellowhorn, an oil tree with robust resistance to cold, drought, heat, salt and alkali, provides an excellent opportunity to identify and characterize the entire yellowhorn WRKY protein family and a basis for the study of abiotic stress resistance of WRKY gene family in forest species. In the present comprehensive analysis of WRKY transcription factors in yellowhorn, 65 WRKY genes were identified and defined based on their location on the chromosome. According to their structure and phylogenetic relationships, XsWRKY genes clustered into WRKY groups I–III. Segmental duplication events played a significant role in the expansion of WRKY gene family. Furthermore, transcriptomic data and real-time quantitative PCR analysis showed that expression of XsWRKY genes responding to salt and drought stresses and a hormone treatment. We also determined structures of the encoded proteins, cis-elements of the promoter region, and expression patterns. These results provide a foundation for the study of the biological function of WRKY transcription factors in yellowhorn.

Keywords Yellowhorn (*Xanthoceras sorbifolium*) · WRKY transcription factor · Stress · ABA · Gene expression

Introduction

Transcription factors are proteins that bind specifically to gene promoter regions and thus regulate transcription and gene expression in response to internal and external stimuli. Unique to plants, the WRKY family is ubiquitous one of the largest transcription factor families in higher plants. WRKY transcription factors are major regulators of many plant biological processes, including abiotic and biotic stress responses, growth and development, carbohydrate synthesis, senescence and secondary metabolite synthesis (Eulgem et al. 2000; Robatzek and Somssich 2002; Sun et al. 2003; Wang et al. 2010a, b). WRKY transcription factors are...
defined by the conserved DNA binding domain (DBD) and the WRKY domain, which comprises approximately 60 residues, consisting of an N-terminal DNA binding domain and a zinc finger motif at the C-terminal (Eulgem et al. 2000). The DBD sequence is highly conserved, but the residues of the WRKY domain are sometimes replaced by WRRY, WKRY, WSKY, WVKY, or WKKY (Xie et al. 2005). The zinc finger motif of the WRKY family forms two modes during evolution: C2H2 (C-X4-5-C-X22-23-H-X1-H) and C2HC (C-X7-C-X23-H-X1-C) (Eulgem et al. 2000). According to the amino acid sequences, WRKY transcription factors can be divided into three groups (Eulgem et al. 2000; Zhang and Wang 2005; Rushton et al. 2010, 2012). The first group of WRKY proteins have two DBD domains and one C2-H2 zinc finger motif (C-X4-5-C-X22-23-H-X1-H), the second and third groups have only one DBD domain and a zinc finger motif. However, the zinc finger motifs in the second group are C-X4-5-C-X22-23-H-X1-H (Kumar et al. 2016; Gu et al. 2018a, b) and in the third group, C2-HC (C-X7-C-X23-H-X1-C) (Kumar et al. 2016; Gu et al. 2018a, b). In addition to the WRKY domain and zinc finger motif, some WRKY transcription factors also contain other structures, such as nuclear localization signal, leucine zipper, glutamic acid enrichment region, proline-rich region and kinase domain (Tian et al. 2006; Agarwal et al. 2011).

Numerous studies have shown that WRKY transcription factors are involved in plant responses to various biotic and abiotic stresses (Rushton et al. 2012; Wang et al. 2014, 2016). In Arabidopsis thaliana, AtWRKY22 and AtWRKY29 are induced as part of a defense response to an elicitor from a bacterial plant pathogen, perhaps through activation of a mitogen-activated protein kinase (MAPK) cascade or calcium-dependent protein kinase (CDPK) (Göhre et al. 2012), while overexpression of AtWRKY38 and AtWRKY62 in double mutants of A. thaliana reduces disease resistance (Kim et al. 2008). AtWRKY48 gain-of-function transgenic overexpression mutants had impaired resistance to the bacterial pathogen Pseudomonas syringae and loss-of-function T-DNA insertion mutants had enhanced resistance associated with increased of a salicylic acid-regulated defense-related protein (PR1) (Xing et al. 2008). Similarly, WRKY genes such as BdWRKY8, BdWRKY34, BdWRKY50, BdWRKY69, and BdWRKY70 in barley were upregulated after inoculation with Fusarium graminearum, causal fungus of Fusarium head blight. It can cause diseases such as ear rot, stem rot, stem base rot and root rot, and can also infect other plants (Wen et al. 2014). Studies have also shown that various plants overexpressing WRKY genes had strong resistance to salt stress (e.g., TaWRKY10 in wheat (Wang et al. 2013) ZmWRKY23 in Zea mays, (Jiang and Yu 2009); OsWRKY45 and OsWRKY72 in Oryza sativa (Yu et al. 2010; Uji et al. 2019). WRKY transcription factors are also essential in signaling pathways mediated by salicylic acid (SA) and by abscisic acid (ABA) (Qiu and Yu 2009) and in plant growth and development. AtWRKY6, AtWRKY22 (Zhou et al. 2011; Zhang et al. 2018), AtWRKY57 (Jiang et al. 2014), AtWRKY53 (Miao et al. 2013), AtWRKY54 and AtWRKY70 (Besseau et al. 2012) are involved in the regulation of senescence. Also, WRKY transcription factors regulate the synthesis of the secondary metabolites, including terpenes, lignin, flavanol, kaempferol, and quercetin (Yokotani et al. 2013).

In addition to the study of the WRKY gene family in model species, studies on WRKY transcription factors in tree species have indicated that Populus euphratica the formation of new WRKY family members and adaptive evolution of other genes may have played an important role in the evolution of high salt tolerance in this species (Ma et al. 2015). EgWRKY18 and EgWRKY64 in Elaeis guineensis are conducive to improving their cold resistance, while the induced expression of EgWRKY07 and EgWRKY52 is beneficial to resistance to salt stress (Xiao et al. 2017). SoWRKY1 is involved in SA and MeJA-mediated abiotic or biological response pathways, and its overexpression improves the salt tolerance of transgenic A. thaliana (Yan et al. 2019).

So far, the studies of WRKY transcription factors have mainly focused on model plants or horticultural plants that are sensitive to stress. So studies of WRKY transcription factors in the stress-tolerant plants will help us to understand the molecular mechanism of the WRKY transcription factors in the processes of stress resistance. Yellowhorn (Xanthoceras sorbifolium) has robust resistance to low temperature, drought, heat, salt and alkali stresses, and it also has high potential for use an edible oil and biofuel because of the high oil content in its seeds. Making use of the recently sequenced yellowhorn genome (Bi et al. 2019), we here identified 56 WRKY genes in the yellowhorn genome and explored their composition, gene structure and gene replication events. Using transcriptomic data for yellowhorn, we also identified some WRKY genes that respond to various stresses such as NaCl, cold and ABA since WRKY genes have been reported to play a vital role in regulating plant tolerance to NaCl and cold stress and in regulating ABA signaling (Wang et al. 2014; Yan et al. 2014; Zhang et al. 2016).

Materials and methods

Genome-wide identification of WRKY genes in yellowhorn

The genomic and protein sequences of yellowhorn and annotated gene models were obtained from the Yellowhorn Genome Project of the Institute of Beijing Academy of Forestry Sciences (Bi et al. 2019). WRKY family of proteins from A. thaliana, O. sativa, Malus domestica, Z.
mays, and Glycine max was downloaded from the plant TFDB (https://planttfdb.cbi.pku.edu.cn) (Jin et al. 2017), and BLASTP was used to search for homologous WRKY genes. The targeted genes (E-value less than 1e-5) were retained for the following analysis. The WRKY domain (PF03106) from Pfam (https://pfam.xfam.org/) (El-Gebali et al. 2019) and WRKY gene were subsequently identified from among the genes of the yellowhorn genome using HMMER 3.1 (Mistry et al. 2013). The target gene was screened according to e-value < 1.0, and the integrity of the WRKY domain (e-value < 0.1) was further confirmed using the online software SMART (https://smart.embl-heidelberg.de/) (Letunic and Bork 2018) and National Center of Biotechnology Information CD-search (https://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi) (Marchler-Bauer et al. 2017). Finally, the repetitive sequence and the incomplete sequence were manually removed, and the confirmed gene was designated as the yellowhorn WRKY gene family. The isoelectric point of the XsWRKY protein was identified using the ExPasy online website (https://web.expasy.org/protparam/) (Gasteiger et al. 2003). MEGA6 (Tamura et al. 2013) software was used for multiple sequence alignment analysis of the WRKY domain. The subcellular localization of XsWRKY in yellowhorn was predicted using the subcellular localization online prediction tool MBC (https://cello.life.nctu.edu.tw/) (Yu et al. 2006).

Phylogenetic analyses of the WRKY proteins

Yellowhorn phylogenetic tree was constructed from the WRKY domain and amino acid sequence of A. thaliana using the maximum likelihood (ML) MEGA6 (Tamura et al. 2013), and the bootstrap test was set as 1,000 replications. Visualization and beautification of the phylogenetic tree were used iTOL (https://itol.embl.de/) (Letunic and Bork 2019).

Gene structure of WRKY genes

Information on the structure of WRKY genes was extracted from the general feature format file of the yellowhorn genome. The TBtools platform (https://github.com/CJ-Chen/TBtools) (Chen et al. 2018) was used to display the exon/intron structures, conservative motifs, and WRKY domain of each predicted XsWRKY gene, and meme (https://meme.nbcr.net/meme/intro.html) (Bailey et al. 2009) was used to identify other conserved motifs (Motif number was set to 20; the rest were the full-length protein using default settings,) in sequences of WRKY family members. The WRKY domain was annotated by the Pfam database (https://pfam.xfam.org/) (El-Gebali et al. 2019).

Promoter analysis

For this study, the 2-kb upstream sequence of the coding region of the WRKY genes was scanned using PlantCARE (https://bioinformatics.psb.ugent.be/webtools/plantcare/html/) (Lescot et al. 2002) to identify transcription factor binding sites. Low temperature, defense responses, hormone responses, drought responses, heat stress, and salt response motif were selected for analysis of promoter specificity for abiotic stress response. TBtools (Chen et al. 2018) was used for visual display.

Chromosome localization and synteny analysis

BLASTP was used for all-against-all comparison (e-value: 1e−5) for information on gene pairs. Next, MCScan (https://chibba.agtec.uga.edu/duplication/mcscan) (Wang et al. 2012) was adopted to analyze syntenic blocks and gene duplication events (using default settings) using gene pairs and gff files. Tandem duplications of WRKY genes were used to estimate nonsynonymous (Ka) and synonymous (Ks) and calculate their ratio in KaKs_Calculator 2.0 (Wang et al. 2010a, b). Circos (Krzywinski et al. 2009) was used to draw a circular map of the yellowhorn genome according to gene position and genome collinearity information in the gff file.

Plant materials and treatments

Yellowhorn seedlings (4-weeks-old) were grown in vermiculite–humus (1:3) soil in a growth chamber at 22–24 °C, 60% relative humidity, and 12 h/12 h photoperiod for 28 days as a non-stress treatment. Seedlings for stress treatments were grown similarly, with the following differences. For the salt treatment, plants were watered with 150 mM NaCl. For ABA, 100 μM was sprayed through the foliage. Cold-stressed plants were grown at 4 °C in a growth chamber. For each treatment, three plants were sampled (biological repeats) with three corresponding untreated controls, and leaves were collected at 3, 6, 9, 12, 24, 48 and 72 h. All leaf samples were immediately frozen in liquid nitrogen and stored at − 80 °C until analyzed.

Transcriptomics analysis

The expression pattern of stress-related WRKY genes in yellowhorn after treatment with 150 mM NaCl, 4 °C, or 100 mM ABA for 24 h and 48 h was analyzed using RNA-seq analysis of transcriptome data in the laboratory (unpublished). The program rnacocktail (Sahraeian et al. 2017) by Salmon was used to calculate the expression level TPM
(per million transcripts), and DESeq2 (Love et al. 2014) was used to analyze differential expression of genes between the control with treatment groups and calculate the fold-change (FDR \leq 0.05, fold change \geq 2). A heatmap was constructed using the TBtools according to the expression matrix of XsWRKY.

### Quantitative RT-PCR analysis

The total RNA of yellowhorn was extracted using the Ominiplant RNA kit (DNase 1) (CWBIO, Beijing, China). RNA quality and concentration was determined using agarose gel electrophoresis and a DHS NanoPro 2010 spectrophotometer (Ding Haoyuan, Beijing, China). Template cDNA was synthesized using 500 ng of high-quality total RNA and the TaKaRa PrimeScript™ RT Reagent Kit with gDNA Eraser (Perfect Real Time) (TaKaRa, Dalian, China) according to the manufacturer’s instructions. Quantitative real-time PCR (qRT-PCR) primers were designed according to the CDS sequence of XsWRKY by Primer 5 software (Table S1), and primer mass was tested using PCR amplification, agarose gel electrophoresis, and melting curve analysis. qRT-PCR was performed using a Bio-Rad/CFX Connect real-time PCR detection system and UltraSYBR mixture (CWBIO, Beijing, China). Relative expression level of the XsWRKY gene was analyzed by the $2^{-\Delta\Delta Ct}$ method and normalized to the average transcript amount of the XsACTIN gene (EVM0013888.1). The qRT-PCR experiments were done three times.

### Results

#### Identification of WRKY proteins in yellowhorn

In this systematic analysis to identify WRKY proteins in yellowhorn, 67 potential WRKY proteins were identified using BLAST (Mount 2007) and HMMER (Mistry et al. 2013) for protein domain analysis. Two of these 67 (EVM0005841.1, EVM0022387.1) did not contain the complete amino acid sequence of the WRKY domain and were removed from the list. The remaining 65 WRKY genes in the yellowhorn genome (Table 1) were denoted as XsWRKY through XsWRKY 59 (Redundant protein numbers such as XsWRKY38, XsWRKY38.1, XsWRKY38.2, and XsWRKY38.3, are four genes with the same protein sequence and were sorted according to the order of genes on the chromosome.) based on their order on the chromosome, but four protein lines not mapped on the chromosome were classified by their number of contigs.

The length of the amino acid sequence in XsWRKY proteins ranged from 122 (XsWRKY24) to 759 (XsWRKY46), with an average length of 372 residues. The molecular weight (MW) of XsWRKY proteins ranged from 14.81 kDa (XsWRKY24) to 83.18 kDa (XsWRKY46), and the isoelectric point (pI) of the XsWRKY proteins ranged from 4.98 (XsWRKY42) to 9.78 (XsWRKY39), with 35 XsWRKYs having a pI < 7 and the rest having pI > 7 (Table 1). Transcription factors are usually localized in nucleus and bind to the target’s promoter regions. Using the prediction tool, we found that all XsWRKYs were localized in the nucleus, except for XsWRKY14, which is expected to be localized in the plasma membrane (Table 1).

#### Phylogenetic analysis of the XsWRKY family

In the phylogenetic analysis of XsWRKY transcription factors using multiple sequence alignment of the 60-aa WRKY domain of XsWRKY protein, the WRKY domain from each WRKY transcription family and subfamily from A. thaliana, Populus trichocarpa, Z. mays was randomly selected as a representative of each family and aligned with that of XsWRKY. In yellowhorn, we found that 61 XsWRKY proteins have the full conserved sequence WRKYGKK, while the other XsWRKY proteins (XsWRKY15, XsWRKY16, XsWRKY21, XsWRKY36) are WRKYGGK (Fig. 1). Subsequently, in a preliminary classification based on the structure of the XsWRKY protein, 12 XsWRKY proteins with two WRKY domains and a C2H2-type zinc finger motif (C-X4-C-X22-23-HXH) were classified into group I. Among the remaining 53 XsWRKY proteins, 45 XsWRKY proteins having C2H2-type zinc finger motif were classified as group II, and eight XsWRKY proteins containing the C2HC-type zinc finger were classified into group III (Fig. 1).

To further classify XsWRKY family members, three WRKY domains of each family and subfamily from A. thaliana, P. trichocarpa, and Z. mays were randomly selected to construct a phylogenetic tree together with the XsWRKY domain (Fig. 2). The phylogenetic tree shows that the WRKY domain of the second group can be divided into five subgroups (IIa–IIe). The C2H2-type zinc finger motif of IIa, IIb, IIc, IIe is C-X5-C-X23-HXH. However, the C2H2-type zinc finger motif of 16 XsWRKY in IIc is C-X4-C-X23-HXH, and the remaining two are C-X4-C-X16-HXH (Fig. 2).

#### Gene structure of yellowhorn WRKY gene family

To further understand the gene structure and motif composition of WRKY proteins, we further analyzed the gene structure of XsWRKY by using MEME. Twenty motifs were predicted in XsWRKY, and the details for the 20 conserved motifs are shown in the supplementary file (Fig. 3, Table S2). The length of motifs is approximate 10 amino acids to 50 amino acids, and each
| WRKY ID   | Gene ID       | PI   | WM   | Subcellular localization | Domain          | No. of aa | ORF length (bp) | Chromosome no | Group |
|----------|---------------|------|------|----------------------------|------------------|-----------|-----------------|---------------|-------|
| XsWRKY1  | EVM0000134.1  | 5.71 | 54,686.71 | Nuclear                   | WRKY 1 C2H2     | 499       | 1500            | LG1           | IIc   |
| XsWRKY2  | EVM0005353.1  | 5.32 | 18,451.12 | Nuclear                   | WRKY 1 C2H2     | 160       | 483             | LG1           | IIc   |
| XsWRKY3  | EVM0001821.1  | 6.61 | 35,239.82 | Nuclear                   | WRKY 1 C2H2     | 322       | 969             | LG1           | IIe   |
| XsWRKY4  | EVM0020675.1  | 5.71 | 39,259.25 | Nuclear                   | WRKY 1 C2H2     | 350       | 1053            | LG1           | IIc   |
| XsWRKY4.1| EVM0020675.1  | 5.71 | 39,259.25 | Nuclear                   | WRKY 1 C2H2     | 350       | 1053            | ctg1541       | IIc   |
| XsWRKY5  | EVM0007069.1  | 6.19 | 33,986.29 | Nuclear                   | WRKY 1 C2H2     | 306       | 921             | LG1           | IIb   |
| XsWRKY6  | EVM003466.1   | 8.45 | 38,425.92 | Nuclear                   | WRKY 1 C2CH     | 344       | 1035            | LG1           | III   |
| XsWRKY7  | EVM0022523.1  | 8.76 | 37,776.86 | Nuclear                   | WRKY 1 C2H2     | 338       | 1017            | LG1           | IIc   |
| XsWRKY8  | EVM0013898.1  | 5.77 | 40,402.79 | Nuclear                   | WRKY 1 C2H2     | 373       | 1122            | LG1           | IIe   |
| XsWRKY9  | EVM0021161.2  | 9.11 | 29,439.52 | Nuclear                   | WRKY 1 C2H2     | 260       | 783             | LG1           | IIId  |
| XsWRKY10 | EVM0019044.1  | 9.25 | 19,963.32 | Nuclear                   | WRKY 1 C2H2     | 172       | 519             | LG1           | IIc   |
| XsWRKY10.1| EVM0019044.1 | 9.25 | 19,963.32 | Nuclear                   | WRKY 1 C2H2     | 172       | 519             | LG1           | IIc   |
| XsWRKY11 | EVM0015874.1  | 9.49 | 17,911.28 | Nuclear                   | WRKY 1 C2H2     | 156       | 471             | LG1           | IIb   |
| XsWRKY12 | EVM0018786.1  | 6.31 | 41,371.66 | Nuclear                   | WRKY 2 C2H2     | 371       | 1116            | LG1           | I     |
| XsWRKY12.1| EVM0018786.1 | 6.31 | 41,371.66 | Nuclear                   | WRKY 2 C2H2     | 371       | 1116            | ctg2419       | I     |
| XsWRKY13 | EVM0009795.1  | 8.43 | 55,897.2  | Nuclear                   | WRKY 1 C2H2     | 511       | 1536            | LG1           | IIc   |
| XsWRKY14 | EVM0006339.1  | 9.38 | 37,533.58 | Plasma membrane            | WRKY 1 C2CH     | 333       | 1002            | LG2           | III   |
| XsWRKY15 | EVM0016791.1  | 5.36 | 39,564.73 | Nuclear                   | WRKY 1 C2H2     | 354       | 1065            | LG2           | IIc   |
| XsWRKY16 | EVM0016046.1  | 8.14 | 37,425.33 | Nuclear                   | WRKY 1 C2H2     | 339       | 1020            | LG2           | IIc   |
| XsWRKY17 | EVM0018382.1  | 9.13 | 33,114.21 | Nuclear                   | WRKY 1 C2H2     | 300       | 903             | LG2           | IIb   |
| XsWRKY18 | EVM0015392.1  | 8.83 | 14,886.81 | Nuclear                   | WRKY 1 C2H2     | 122       | 369             | LG3           | IIc   |
| XsWRKY19 | EVM0014416.1  | 6.65 | 59,836.26 | Nuclear                   | WRKY 1 C2H2     | 539       | 1620            | LG3           | Ia    |
| XsWRKY20 | EVM0009035.1  | 5.71 | 55,257.41 | Nuclear                   | WRKY 1 C2H2     | 504       | 1515            | LG3           | Ia    |
| XsWRKY21 | EVM0022609.1  | 6.56 | 37,890.63 | Nuclear                   | WRKY 1 C2H2     | 347       | 1044            | LG3           | IIc   |
| XsWRKY22 | EVM0013326.2  | 9.59 | 34,986.39 | Nuclear                   | WRKY 1 C2H2     | 319       | 960             | LG4           | IIId  |
| XsWRKY23 | EVM0012322.1  | 9.11 | 21,811.69 | Nuclear                   | WRKY 1 C2H2     | 191       | 576             | LG5           | IIe   |
| XsWRKY24 | EVM0005716.1  | 8.84 | 14,810.71 | Nuclear                   | WRKY 1 C2H2     | 122       | 369             | LG5           | IIc   |
Table 1 (continued)

| WRKY ID   | Gene ID   | PI    | WM    | Subcellular localization | Domain          | No. of aa | ORF length (bp) | Chromosome no | Group |
|-----------|-----------|-------|-------|--------------------------|-----------------|-----------|-----------------|---------------|-------|
| XsWRKY25  | EVM0013194.1 | 7.07  | 70,712.6 | Nuclear                  | WRKY 1 C2H2     | 649       | 1950            | LG5           | Ile   |
| XsWRKY26  | EVM0001816.1 | 6.99  | 68,769.07 | Nuclear                  | WRKY 2 C2H2     | 625       | 1878            | LG5           | I     |
| XsWRKY27  | EVM0004495.1 | 9.37  | 36,426.59 | Nuclear                  | WRKY 1 C2H2     | 324       | 975             | LG5           | IIb   |
| XsWRKY28  | EVM0020928.1 | 6.94  | 22,986.19 | Nuclear                  | WRKY 2 C2H2     | 201       | 606             | LG6           | I     |
| XsWRKY29  | EVM0009790.1 | 6.72  | 42,494.57 | Nuclear                  | WRKY 1 C2H2     | 383       | 1152            | LG6           | IId   |
| XsWRKY30  | EVM0004255.1 | 5.12  | 30,994.11 | Nuclear                  | WRKY 1 C2H2     | 269       | 810             | LG8           | IIc   |
| XsWRKY31  | EVM0020089.1 | 5.88  | 67,790.72 | Nuclear                  | WRKY 2 C2H2     | 622       | 1869            | LG8           | I     |
| XsWRKY32  | EVM0021666.1 | 5.99  | 59,926.62 | Nuclear                  | WRKY 1 C2H2     | 554       | 1665            | LG8           | IIb   |
| XsWRKY33  | EVM0003950.1 | 6.16  | 32,151.41 | Nuclear                  | WRKY 1 C2H2     | 293       | 882             | LG8           | IIc   |
| XsWRKY34  | EVM0009513.1 | 7.69  | 58,116.99 | Nuclear                  | WRKY 2 C2H2     | 534       | 1605            | LG9           | I     |
| XsWRKY35  | EVM0021601.1 | 5.72  | 54,197.23 | Nuclear                  | WRKY 1 C2H2     | 496       | 1491            | LG9           | IIa   |
| XsWRKY36  | EVM0015892.1 | 5.19  | 27,764.88 | Nuclear                  | WRKY 1 C2H2     | 247       | 744             | LG10          | IIc   |
| XsWRKY37  | EVM0007134.1 | 7.24  | 65,102.88 | Nuclear                  | WRKY 1 C2H2     | 591       | 1776            | LG10          | IIc   |
| XsWRKY38  | EVM0009349.1 | 9.44  | 23,414.96 | Nuclear                  | WRKY 1 C2H2     | 208       | 627             | LG10          | IId   |
| XsWRKY38.1| EVM0009349.1 | 9.44  | 23,414.96 | Nuclear                  | WRKY 1 C2H2     | 208       | 627             | ctg2862       | IId   |
| XsWRKY38.2| EVM0009349.1 | 9.44  | 23,414.96 | Nuclear                  | WRKY 1 C2H2     | 208       | 627             | ctg3040       | IId   |
| XsWRKY38.3| EVM0009349.1 | 9.44  | 23,414.96 | Nuclear                  | WRKY 1 C2H2     | 208       | 627             | ctg3040       | IId   |
| XsWRKY39  | EVM0017381.1 | 9.78  | 39,603.94 | Nuclear                  | WRKY 1 C2H2     | 355       | 1068            | LG11          | IIb   |
| XsWRKY40  | EVM0000723.1 | 9.65  | 38,220.38 | Nuclear                  | WRKY 1 C2H2     | 349       | 1050            | LG11          | IIe   |
| XsWRKY41  | EVM0002462.1 | 8.55  | 27,915.54 | Nuclear                  | WRKY 1 C2CH      | 241       | 726             | LG11          | III   |
| XsWRKY42  | EVM0010266.1 | 4.98  | 35,380.09 | Nuclear                  | WRKY 1 C2H2     | 318       | 957             | LG11          | IId   |
| XsWRKY43  | EVM0006293.1 | 5.36  | 33,837.8 | Nuclear                  | WRKY 1 C2H2     | 300       | 903             | LG12          | IIe   |
| XsWRKY44  | EVM0007056.1 | 6.52  | 61,996.35 | Nuclear                  | WRKY 1 C2CH     | 569       | 1710            | LG12          | III   |
| XsWRKY45  | EVM002590.1  | 9.74  | 39,138.46 | Nuclear                  | WRKY 1 C2CH     | 348       | 1047            | LG13          | III   |
| XsWRKY46  | EVM0013857.1 | 5.38  | 83,179.43 | Nuclear                  | WRKY 2 C2H2     | 759       | 2280            | LG13          | I     |
| XsWRKY47  | EVM0009650.1 | 5.84  | 46,304.85 | Nuclear                  | WRKY 1 C2H2     | 418       | 1257            | LG13          | IIb   |
| XsWRKY48  | EVM0022427.1 | 6.5   | 43,023.12 | Nuclear                  | WRKY 2 C2H2     | 392       | 1179            | LG13          | I     |
Identification of yellowhorn (Xanthoceras sorbifolium) WRKY transcription factor family...

The identification of WRKY transcription factor family proteins contains 3–12 motifs. The same group of proteins have a similar motif composition, but all XsWRKY proteins contain motifs 1 and 2, so motifs 1 and 2 that are present in WRKY are important parts of the WRKY domain. The highly similar motif 3 and motif 14 are part of the zinc finger motif domain, and with motifs 2 and 4 form the C2H2-type zinc finger. Motif 9 belongs to the N-terminal WRKY domain of XsWRKY proteins. Motif 10 is unique to IIa and IIb. Motif 8 is specific to IId, and motif 20 is specific to III (Fig. 3). The motifs in different subfamilies may be related to the specific functions of the family members.

Regarding the gene structure of the XsWRKY family, all XsWRKY genes have 2 to 6 exons (Fig. 3). XsWRKY genes in the same group have a similar gene structure. For example, members of IIa and IIb except for XsWRKY19 contain only phase 0 introns, and all members of group I have N-terminal domains that do not contain introns (Fig. 3).

### Distribution and synonymy analysis of XsWRKY genes in genome

Regarding the distribution of XsWRKY genes in the yellowhorn genome, we found that in addition to XsWRKY4.1, 12.1, 38.2, 38.3, 56, 57, 58, 59 (on the contig), the XsWRKY gene is unevenly distributed on 14 chromosomes except chromosome 7 (Fig. 4). Chromosome 1 contains the most WRKY genes. Although most XsWRKY genes are distributed on several chromosomes, there is no positive correlation between the number of XsWRKY genes on a chromosome and its length (Fig. 4).

To understand the expansion of the XsWRKY family, we identified two tandem repeat gene pairs, XsWRKY54/55 and XsWRKY19/20 (Table S3) and 23 segmental duplication events based on MCScanX for collinearity analysis of yellowhorn.

To explore the phylogenetic mechanisms of the XsWRKY family, we studied the collinear relationship between yellowhorn four dicotyledons (A. thaliana, G. max, P. trichocarpa, M. domestica) and two monocotyledons (O. sativa and Z. mays) (Fig. 5). There were the most homologous gene pairs between WRKY and XsWRKY in G. max (51), followed by P. trichocarpa (49), A. thaliana (43), M. domestica (43), O. sativa (26) and Z. mays (22) (Table S4). The orthologous logarithms between the other six species (G. max, P. trichocarpa, A. thaliana, M. domestica, O. sativa, and Z. mays) were 151, 88, 53, 68, 43 and 36, respectively. Some XsWRKY genes have at least three related homologous gene pairs with other species (such as XsWRKY5 and XsWRKY32, XsWRKY47). Interestingly, a collinear block in O. sativa and Z. mays with yellowhorn contained more than two XsWRKY genes, and the span of any two XsWRKY genes in the block was less than 14 genes. This phenomenon does not exist in collinear blocks in G. max, P. trichocarpa, M. domestica with yellowhorn.

| WRKY ID | Gene ID | PI | WM | Subcellular localization | Domain | No. of aa | ORF length (bp) | Chromosome no | Group |
|---------|---------|----|----|--------------------------|---------|-----------|-----------------|---------------|-------|
| XsWRKY49 | EVM0014759.1 | 8.56 | 57,470.8 | Nuclear | WRKY 1 C2H2 | 526 | 1581 | LG13 | IIc |
| XsWRKY50 | EVM00022939.2 | 5.18 | 27,804.9 | Nuclear | WRKY 2 C2H2 | 246 | 741 | LG13 | I |
| XsWRKY51 | EVM0014166.1 | 7.76 | 25,870.39 | Nuclear | WRKY 2 C2H2 | 225 | 678 | LG14 | I |
| XsWRKY52 | EVM0017304.1 | 6.03 | 66,351.1 | Nuclear | WRKY 2 C2H2 | 614 | 1845 | LG14 | I |
| XsWRKY53 | EVM0011943.1 | 5.53 | 40,993.73 | Nuclear | WRKY 1 C2H2 | 357 | 1074 | LG14 | IIb |
| XsWRKY54 | EVM0002771.1 | 5.9 | 38,440.76 | Nuclear | WRKY 1 C2CH | 340 | 1023 | LG15 | III |
| XsWRKY55 | EVM0007812.1 | 5.78 | 45,788.63 | Nuclear | WRKY 1 C2CH | 404 | 1215 | LG15 | III |
| XsWRKY56 | EVM0011128.1 | 5.81 | 80,606.47 | Nuclear | WRKY 1 C2H2 | 744 | 2235 | ctg1408 | IIe |
| XsWRKY57 | EVM0005411.1 | 8.46 | 63,283.94 | Nuclear | WRKY 2 C2H2 | 575 | 1728 | ctg2606 | I |
| XsWRKY58 | EVM0005607.1 | 8.72 | 51,530.29 | Nuclear | WRKY 1 C2H2 | 467 | 1404 | ctg2924 | IId |
| XsWRKY59 | EVM0012687.1 | 5.83 | 53,611.21 | Nuclear | WRKY 1 C2CH | 492 | 1479 | ctg3055 | III |
Fig. 1 Alignment of multiple XsWRKY proteins and selected AtWRKY domain amino acid sequences. Group IN and Group IC indicate the N-terminal and C-terminal WRKY domains of group I WRKY proteins.
On the other hand, we calculated the number of XsWRKY between the two ends of XsWRKY genes in the collinear block between yellowhorn and G. max (40), P. trichocarpa (32), M. domestica (21), A. thaliana (6), O. sativa (2) and Z. mays (0). There are some unique collinear gene pairs between yellowhorn and dicotyledons (A. thaliana, G. max, P. trichocarpa, and M. domestica), but there are 10 common collinear gene pairs in the yellowhorn and these 6 species.
Cis-acting regulatory element analysis of the WRKY promoter

To identify the cis-regulatory elements in the XsWRKY promoter, we downloaded the 2-kb sequence upstream of the initiation codons of the XsWRKY genes and analyzed them with the PLACE server (Table S5). Interestingly, numerous cis-acting elements involved in the responses to plant hormone, biotic, and abiotic stress were identified in the promoter of the XsWRKY transcription factors. According to the annotations, the elements belong to eight categories: ABA, JA, GA, SA, light, low temperature, drought, anaerobic. The elements related to light response are the most enriched, e.g., Box 4, ATCT-motif, GT1-motif, GATA-motif, G-box, G-Box (Fig. 6). Other types of cis-elements such as CGTCA-motif, TGACG-motif are MeJA response elements; GARE-motif, TATC-motif, P-box are gibberellin response elements; ARE is involved in anaerobic induction; LTR linked to plant response to low temperatures related; TC-rich repeats are involved in defense and stress response; the TCA-element is involved in the salicylic acid response.

Expression profile of XsWRKY genes after NaCl or cold stress and ABA exposure

To study the role of the XsWRKY genes in response to various environmental stresses and related signal transduction, we analyzed the transcriptional responses of the XsWRKY genes to salt, cold and ABA treatments (Table S6).
Identification of yellowhorn (*Xanthoceras sorbifolium*) WRKY transcription factor family…

Treatment with 150 mM NaCl for 24 h, most of the differentially expressed gene (DEGs) were significantly upregulated (e.g., *XsWRKY1*, *XsWRKY24*, *XsWRKY4*), but *XsWRKY6* was significantly downregulated after treatment for 24 h, and significantly upregulated after 48 h. After 48 h, *XsWRKY53*, *XsWRKY24*, *XsWRKY27*, *XsWRKY4*, *XsWRKY5*, *XsWRKY1*, and other genes were significantly upregulated. At 24 h and 48 h of salt treatment, 10 and 16 genes were differentially expressed genes (Fig. 7). To validate the transcriptomic data, we analyzed some *XsWRKY* genes by qRT-PCR and found that *XsWRKY22* was induced by NaCl at 6 h. The expression of *XsWRKY22* peaked at 6 h and then slowly decreased. The expression levels of *XsWRKY6* and *XsWRKY33* showed a downward trend at the initial stage of induction but began to rise after 6 h, peaked by 24 h, then decreased by 48 h (Fig. 8). Other studies have shown that *AtWRKY46* can regulate stomatal closure and enhance plant salt stress resistance (Ding et al. 2015). In conclusion, *XsWRKY* genes are closely related to the salt stress response in yellowhorn.

Most plants native to cold areas have specific cold resistance. However, the mechanism of action of WRKYS in cold stress response is not precise. *VaWRKY12* is expressed in the nucleus and cytoplasm at average temperature, but only in the nucleus after cold treatment (Zhang et al. 2019). Overexpression of *VaWRKY12* enhances the cold tolerance of *A. thaliana* and grape callus and significantly increases expression of antioxidant-related genes. Overexpression of *CsWRKY46* increased cold tolerance in cucumber and actively regulates cold signaling pathways in an ABA-dependent manner (Zhang et al. 2016). After 24 h of cold treatment in the present study, the expression of nine *XsWRKY* genes was significantly upregulated, for example, *XsWRKY4*, *XsWRKY24*, *XsWRKY35*, and *XsWRKY46* (Fig. 7), but no downregulated genes were found. After 48 h of cold treatment, some genes were somewhat upregulated (*XsWRKY4*, *XsWRKY46*, *XsWRKY40*), while most of the other DEGs were significantly upregulated, and *XsWRKY6* and *XsWRKY54* were significantly downregulated. The
expression of XsWRKY24 as determined by qRT-PCR decreased first and then peaked at 48 h (Fig. 8). The expression of XsWRKY34 increased continuously after cold induction. The expression of XsWRKY6 decreased significantly after 48 h. The expression of XsWRKY54 decreased after cold stress. As for the WRKY genes expression trend, the expression level was the highest at 6 h and was reduced to the lowest after 48 h (Fig. 7).

ABA is closely related to plant response to abiotic stress and plays a crucial role in response to environmental challenges. A. thaliana chloroplast protein ABRR interacts with WRKY transcription factors (WRKY40, WRKY18, and WRKY60) to act as ABA signals. WRKYs can also act as negative regulators. For example, WRKY40 inhibits the expression of the ABA response gene ABI5 (Shang et al. 2010). GhWRKY17 in cotton can respond to drought and salt stress through ABA signaling and regulation of cell ROS production in plants exposed to drought and salt stress (Yan et al. 2014). The expression of XsWRKY55 was significantly upregulated, and the expression of XsWRKY8 was downregulated at 24 h after ABA treatment. The expression of XsWRKY55 in the treated group was four times higher than in the control group. The expression of XsWRKY25 was significantly downregulated after 48 h of ABA treatment. qRT-PCR showed that XsWRKY54 was induced by ABA, and its expression increased gradually in the early stage, but began to decrease after 6 h (Fig. 8). The results suggest that these XsWRKY genes may be involved in ABA signaling.

In conclusion, the expression pattern of XsWRKY genes under NaCl and cold conditions indicated that different XsWRKY genes might be involved in different stress responses. ABA treatment demonstrated that XsWRKY gene may respond to different stresses through ABA transduction pathway.
Discussion

Using the recently published yellowhorn genome (Bi et al. 2019), our laboratory found here that the genome contains 65 members of WRKY transcription factors. Based on their distribution on the chromosome, we defined them as XsWRKY1 to XsWRKY59; six sequences are redundant (XsWRKY4.1, XsWRKY10.1, XsWRKY12.1, XsWRKY38.1, XsWRKY38.2, XsWRKY38.3).

In our evaluation of the conserved domain of the XsWRKY protein, multi-sequence alignment of different subfamilies showed that the conservation of the WRKY domain differed among the subfamilies; conservation of the domain in subfamily groups IN, Ic, and IIc was weaker than in the other groups. The WRKY domain of four members (XsWRKY15, XsWRKY16, XsWRKY21, XsWRKY36) in the Ic subfamily changed from WRKYGQK to WRKYGKK. Most XsWRKY proteins had a binding preference for the homologous homeopathic element W-box. Variation in the WRKYGQK motif in the WRKY domain may affect the interaction of the WRKY genes with downstream target genes, and this mutation also exists in other species such as pineapple (Xie et al. 2018), Caragana intermedia (Wan et al. 2018), grapes (Romero et al. 2019). So these four proteins merit further study of their function and binding specificity. Changes in the WRKY domain provide the impetus for the expansion of the WRKY gene family (Xie et al. 2018). The XsWRKY proteins in group I in yellowhorn have two WRKY domains, and groups II and III have only one, indicating that XsWRKY proteins evolved different characteristics. Subfamilies II and III apparently arose from the N-terminal deletion in subfamily I, and the C-terminal domain in subfamily I members is more highly conserved than the N-terminal domain (Zhang and Wang 2005). Consistent with results of previous studies (Zhang and Wang 2005), the WRKY domains of subfamilies II and III contain introns consistent with the C-terminal WRKY domain of subfamily I. Subfamily I has the most primitive WRKY gene...
type, which transitions to phase II due to loss of differentiation of the WRKY domain. Group III may have evolved through a variation in the C-terminal zinc finger. Each group member has the same or similar phase composition, which also implies this evolutionary relationship of the WRKY transcription factor family.

In the analysis of the XsWRKY genes of the yellowhorn genome, except for 2 tandem repeats and 23 fragment repeats, the contigs and chromosomes are highly similar; XsWRKY pairs may be potential tandem repeats or segment repeats. At the same time, 25 repeat events were identified in 33 XsWRKY genes. These results indicate that the specific XsWRKY gene was generated by gene duplication and segment repeat events as the main driving force of XsWRKY evolution. At the same time, in the collinear module of yellowhorn and the other six plants analyzed, some XsWRKY gene pairs were spaced 100 or even 200 genes apart from the other species, which may be due to genetic differences between yellowhorn with other species, genomic doubling events, transposon jumping and other factors.

Compared to the genomes of dicotyledons such as M. domestica (Meng et al. 2016), A. thaliana (Dong
Identification of yellowhorn (*Xanthoceras sorbifolium*) WRKY transcription factor family…

et al. 2003), *G. max* (Bencke-Malato et al. 2014), and *P. trichocarpa* (Tuskan et al. 2006), yellowhorn has a relatively small number of *XsWRKY* genes. Segmental and tandem repeat events play an important role in the expansion of the *XsWRKY* gene family, while genome-wide doubling events play a key role in the expansion of the gene family. Yellowhorn shares γ whole-genome duplication events with other dicots (Murat et al. 2017), but based on the distribution of *WRKY* genes on the chromosomes and the collinearity analysis with other species, the lower number of *XsWRKY* genes in yellowhorn may be due to genomic structural variation. The collinear *XsWRKY*
gene pairs between the six species and yellowhorn appear simultaneously in multiple species, which may indicate that these homologous genes may play an important role in these plants.

According to the evolutionary hypothesis of the WRKY transcription factors, the most primitive WRKY proteins only contain a WRKY domain at the C-terminal, then group I was generated by doubling, while group I lost the N-terminal WRKY domain during evolution to produce group II (Zhang and Wang 2005). The group III WRKY gene is based on group II. Yellowhorn has fewer group III XsWRKY genes (8) than in most other dicotyledons, consistent with it having fewer members of the WRKY family. This may be related to the fact that yellowhorn is the only species of yellowhorn and has not been subjected to lineage-specific replication.

Yellowhorn is usually grown in impoverished land with scarce water, which confer resistance to various stresses including drought, cold, salt, and alkali. It is a green pioneer species. Based on the biological function of WRKY transcription factors in another species, it is reasonable to conclude that XsWRKY transcription factors are also involved in physiological processes and stress responses. Combined with public transcriptomic data and phylogenetic analysis, our studies have shown that XsWRKY proteins have a biological function similar to those of WRKY proteins in other species. For example, our transcriptomic data showed that the expression of 18 XsWRKY genes was significantly induced under salt stress.

Because AtWRKY33 in A. thaliana responds to salt stress, we selected its homolog in yellowhorn, XsWRKY34, for further study. Consistently, the qRT-PCR analysis indicated that expression of XsWRKY34 was strongly induced by 150 mM NaCl. The results of the qRT-PCR of XsWRKY genes showed that some XsWRKY genes were upregulated or downregulated by around 6 h. This time point may indicate a new mechanism by which XsWRKY genes participate in conferring resistance to stress in yellowhorn. These results indicate that XsWRKY proteins play a vital role in the regulation of salt stress. Yellowhorn is mainly distributed in northern China and has cold resistance. The WRKY transcription factor is a crucial component of the cold stress regulation network. In O. sativa, OsWRKY76 (Yokotani et al. 2013) is induced by low temperature. Yellowhorn XsWRKY20 was also induced by cold stress, indicating that XsWRKY20 has a potential role in the cold response of yellowhorn (Fig. 5). Increasing studies have shown that the WRKY genes are regulated by hormones such as ABA, SA, and JA and play a crucial role in the plant hormone signal transduction network. Low temperature in the present study induced XsWRKY20 and its ortholog gene OsWRK76 in rice, indicating that they have similar functions in cold tolerance and ABA signaling network. Consistent with previous studies in other species, our results indicate that differential expression of some XsWRKY genes after various abiotic stresses and hormone treatments highlights the extensive involvement of WRKY genes in environmental adaptation.

Collectively, our findings provide new insights into the potential function of XsWRKY genes in yellowhorn. A comprehensive analysis will help select candidate WRKY genes for further functional identification and genetic improvement of forest species.

Conclusion

In this analysis of 65 full-length WRKY genes identified from yellowhorn, the XsWRKY transcription factors represented seven subfamilies with high similarity in motif composition and exon–intron structure. Homology analysis and phylogenetic comparison of WRKY genes from several different plant species provided valuable clues on duplication events during the evolution of the WRKY genes. The response of XsWRKY genes is vital to the high tolerance of yellowhorn to stresses and thus to its growth and development. These valuable insights into the function of the WRKY transcription factors highlight their potential for breeding programs to increase tolerance to poor environments.

Acknowledgements We are very grateful to Prof. Zhimin Zheng and Dr. Yi Li of the University of Southern California for their detailed comments and suggestions that improved the manuscript. We also thank the students and teachers of the Zheng Lab for their support and help with this article.

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/.

References

Agarwal P, Reddy MP, Chikara J (2011) WRKY: its structure, evolutionary relationship, DNA-binding selectivity, role in stress tolerance and development of plants. Mol Biol Rep 38(6):3883–3896 Bailey TL, Boden M, Buske FA, Frith M, Grant CE, Clementi L, Ren J, Li WW, Noble WS (2009) MEME SUITE: tools for motif discovery and searching. Nucleic Acids Res 37(suppl_2):W202–W208 Bencke-Malato M, Cabreira C, Wiebke-Strohm B, Bucker-Neto L, Mancini E, Osorio MB, Homrich MS, Turchetto-Zolet AC, De Carvalho MCCG, Stolf R, Weber RLM, Westergaard G,
Identification of yellowhorn (Xanthoceras sorbifolium) WRKY transcription factor family...
Shen HS, Liu CT, Zhang Y, Meng XP, Zhou X, Chu CC, Wang XP (2012) OsWRKY30 is activated by MAP kinases to confer drought tolerance in rice. Plant Mol Biol 80(3):241–253

Sun C, Palmqvist S, Olsson H, Borén M, Ahllandsberg S, Jansson C (2003) A novel WRKY transcription factor, SUSIBA2, participates in sugar signaling in barley by binding to the sugar-responsive elements of the iso1 promoter. Plant Cell 15(9):2076–2092

Tamura K, Stecher G, Peterson D, Filipski A, Kuma S (2013) MEGA6: molecular evolutionary genetics analysis version 6.0. Mol Biol Evol 30(12):2725–2729

Tian Y, Lu XY, Peng LS, Fang J (2006) The structure and function of plant WRKY transcription factors. Yi Chuan 28(12):1607–1612

Tuskan GA, Difazio S, Jansson S, Bohlmann J, Grigoriev I, Hellsten U, Putnam N, Ralph S, Rombauts S, Salamov A, Schein J, Sterck L, Aerts A, Bhalerao RR, Bhalerao RP, Blaudez D, Boerjan W, Brun A, Brunner A, Busov V, Campbell M, Carlson J, Chalot M, Chapman J, Chen GL, Cooper D, Coutinho PM, Couturier J, Covert S, Cronk Q, Cunningham R, Davis J, Degroeve S, Déjardin A, Depamphilis C, Detter J, Dirks B, Dubchak I, Duplesis S, Ehling J, Ellis B, Gendler K, Goodstein D, Gribkov M, Grimm J, Groover A, Gunter L, Hamberger B, Heinzre H, Helariutta Y, Henrisst B, Holligan D, Holt R, Huang W, Islam-Faridi N, Jones S, Jones-Rhoades M, Jorgensen R, Joshi C, Kangasjärvi J, Karlsson J, Kelleher C, Kirkpatrick R, Kist M, Kohler A, Kalluri U, Larimer F, Lefebvre-Mack J, Leplé JC, Locascio P, Lou Y, Lucas S, Martin F, Montanini B, Napolì C, Nelson DR, Nelson C, Nieminen K, Nilsson O, Pereda V, Peter G, Philippe R, Pilate G, Poljakov A, Razumovskaya J, Richardson P, Rinaldi C, Ritland K, Rouze P, Ryaboy D, Schmutz J, Schrader J, Segerman B, Shen QJ, Siddiqui A, Sterky F, Terry A, Tsai CJ, Uberbacher E, Unneberg C, Vahala J, Wall K, Wessler S, Yang G, Yin T, Douglas C, Marga M, Sandberg G, Van de Peer Y, Yohsok D (2006) The genome of black cottonwood, Populus trichocarpa, and its relationship to rice and maple. Proc Natl Acad Sci 107(51):22338–22343

Mutation of WRKY transcription factors initiates pith second-wall formation and increases stem biomass in dicotyledonous plants. Proc Natl Acad Sci 107(51):22338–22343

Wan YQ, Mao MZ, Wan DL, Yang Q, Yang FY, Mandlaa LGJ, Wang RG (2018) Identification of the WRKY gene family and functional analysis of two genes in Caragana aragihima. BMC Plant Biol 18(1):31

Wang DP, Zhang YB, Zhang Z, Zhu J, Yu J (2010a) KaKs, Calculator 2.0: a toolkit incorporating gamma-series methods and sliding window strategies. Genom Proteom Bioinform 8(1):77–80

Wang HZ, Avci U, Nakashima J, Hahn MG, Chen F, Dixon RA (2010b) Mutation of WRKY transcription factors initiates pith secondary wall formation and increases stem biomass in dicotyledonous plants. Proc Natl Acad Sci 107(51):22338–22343

Wang YP, Tang HB, Debarry JD, Tan X, Li JP, Wang XY, Lee TH, Jin HZ, Marler B, Guo H, Kissingr JC, Paterson AH (2012) MASCANX: a toolkit for detection and evolutionary analysis of gene synteny and collinearity. Nucleic Acids Res 40(7):e49–e49

Xia X, Zhou LX, Lei XT, Cao HX, Wang Y, Dou YH, Tang WQ, Xia W (2017) Genome-wide identification of WRKY genes and their expression profiles under different abiotic stresses in Elaeis guineensis. PLOS One 12(12):e0189224

Xie Z, Zhang ZL, Zou XL, Huang J, Ruas P, Thompson D, Shen QJ (2005) Annotations and functional analyses of the rice WRKY gene superfamily reveal positive and negative regulators of abscisic acid signaling in aleurone cells. Plant Physiol 137(1):176–189

Wen F, Zhu H, Li P, Jiang M, Mao WQ, Ong C, Chu ZQ (2014) Genome-wide evolutionary characterization and expression analyses of WRKY family genes in Brachypodium distachyon. DNA Res 21(3):327–339

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.