Journal of Agricultural Science (2017), **155**, 777–790. © Cambridge University Press 2016
doi:10.1017/S0021859616000873

CROPS AND SOILS RESEARCH PAPER

Foxtail millet *WRKY* genes and drought stress

L. ZHANG[1]†, H. SHU[2]†, A. Y. ZHANG[3], B. L. LIU[1], G. F. XING[1], J. A. XUE[1], L. X. YUAN[1], C. Y. GAO[1] AND R. Z. LI[1]*

[1] *Institute of Molecular Agriculture & Bioenergy, Shanxi Agricultural University, Taigu 030801, China*
[2] *Program of Molecular Medicine, University of Massachusetts Medical School, Massachusetts 01605, USA*
[3] *Institute of Millet, Shanxi Academy of Agricultural Sciences, Changzhi 046011, China*

(*Received 29 March 2015; revised 22 June 2016; accepted 28 September 2016; first published online 2 November 2016*)

SUMMARY

Foxtail millet (*Setaria italica* (L.) P. Beauv.) is a naturally stress-tolerant plant, a major reserve crop and a model for panicoid grasses. The recent completion of the *S. italica* genome facilitates identification and characterization of WRKY transcription factor family proteins that are important regulators of major plant processes, including growth, development and stress response. The present study identified 103 WRKY transcription factor-encoding genes in the *S. italica* genome. The genes were named *SiWRKY*1–*SiWRKY*103 according to their order on the chromosomes. A comprehensive expression analysis of *SiWRKY* genes among four different tissues was performed using publicly available RNA sequencing data. Eighty-four *SiWRKY* genes were more highly expressed in root tissue than in other tissues and nine genes were only expressed in roots. Additionally, real-time quantitative polymerase chain reaction was performed to comprehensively analyse the expression of all *SiWRKY* genes in response to dehydration. Results indicated that most *SiWRKY* genes (over 0.8) were up-regulated by drought stress. In conclusion, genome-wide identification and expression profiling of *SiWRKY* genes provided a set of candidates for cloning and functional analyses in plants' response to drought stress.

INTRODUCTION

Foxtail millet (*Setaria italica* (L.) P. Beauv.) is a diploid, self-pollinated, C4 monocot grass in the sub-family Poaceae. It is an important strategic reserve crop (i.e. the government reserves seeds in order to ensure that some crops are available to people in case of natural disasters) owing to its high drought resistance, photosynthetic ability and nutritional quality (Muthamilarasan & Prasad 2015). Recently, the Beijing Genomics Institute (BGI) in China (Zhang *et al*. 2012) and the Joint Genome Institute (JGI) of the Department of Energy in the USA (Bennetzen *et al*. 2012) have independently sequenced the *S. italica* genome. Its size is approximately 515 Mb, larger than the rice genome (430 Mb) but smaller than pearl millet (2352 Mb) and maize (2500 Mb) genomes. The availability of *S. italica* sequence data has made it a useful model for studying related panicoid plants, including bioenergy crops such as switchgrass (*Panicum virgatum* L.) and pearl millet (*P. glaucum* L.) (Bennetzen *et al*. 2012).

The WRKY transcription factors consist of about 60 highly conserved amino acids, characterized by the WRKYGQK motif in the N-terminus and the CysCys HisCys (C_2HC)- or CysCysHisHis (C_2H_2)-type zinc finger motif in the C-terminus (Eulgem *et al*. 2000; Rushton *et al*. 2010). The WRKY proteins regulate gene expression by binding to the W-BOX motif (TTGACC/T) in gene promoters (Rushton *et al*. 1996; Eulgem *et al*. 2000). WRKY family members are divided into three groups based on the number of WRKY domains and the pattern of zinc finger motifs. The first group typically includes two WRKY domains and a C_2H_2-type zinc finger motif. Most WRKY proteins contain only one WRKY domain and belong to the second or third groups. The zinc finger motifs from the second group are identical to the first group, whereas the third group possesses the C_2HC

* To whom all correspondence should be addressed. Email: rli2001@126.com
† The first two authors contributed equally to this paper.

zinc finger motif. Group II WRKY proteins are further divided into IIa + IIb, IIc and IId + IIe, whereas group III WRKYs are classified as IIIa and IIIb (Zhang & Wang 2005; Rushton *et al*. 2010).

WRKY transcription factors play key roles in plant defence against various biotic stresses, including bacterial, fungal and viral pathogens (Yang *et al*. 1999; Marchive *et al*. 2007; Mukhtar *et al*. 2008; Tao *et al*. 2009; Gallou *et al*. 2012; Wang *et al*. 2013). Recent evidence also supports their roles in response to abiotic stressors, such as drought, salinity and cold (Jiang & Deyholos 2009; Zou *et al*. 2010; Jiang *et al*. 2012; Okay *et al*. 2014; Wang *et al*. 2014; Yan *et al*. 2014). Moreover, they are involved in the regulation of major plant growth and development processes, such as seed development, trichome initiation and senescence (Ülker & Somssich 2004; Rushton *et al*. 2010). Therefore, there has been immense interest in studying WRKY transcription factors as potential candidates for agricultural enhancement. Several WRKY families have been identified in a variety of plants, including *Arabidopsis thaliana* (L.) Heynh. (Eulgem *et al*. 2000), rice (*Oryza sativa* L.; Xie *et al*. 2005), maize (*Zea mays* L.; Wei *et al*. 2012) and bread wheat (*Triticum aestivum* L.; Okay *et al*. 2014). Recently, a global analysis of the WRKY family was conducted in *S. italica*, demonstrating the involvement of four genes in abiotic stress (Muthamilarasan *et al*. 2015).

The present study describes a more extensive analysis of 103 *WRKY* genes (*SiWRKYs*) from *S. italica* under drought stress. First, a phylogenetic tree was constructed and the genes were classified into three main groups. In addition, the conserved motifs of individual SiWRKYs were analysed and genomic distribution of the genes was estimated. Moreover, tissue-specific gene expression patterns were profiled with publicly available RNA sequencing data. A comprehensive analysis of *SiWRKY* expression in response to dehydration was also performed, to understand how drought affects gene expression in a stress-tolerant plant. The present study aims to provide an overview of *SiWRKY* genes and generate data that will help identify potential candidates for improving selection of stress tolerance in crops.

MATERIALS AND METHODS

Database queries

Two different approaches were applied to identify putative WRKY proteins from *S. italica*. First, the *S. italica* genome sequence database in Phytozome version 9.1 (www.phytozome.net) was searched using 'WRKY' and '*Setaria italica*' as keywords. In addition, 472 WRKY transcription-factor protein sequences of four species (*A. thaliana* [90], rice [109], sorghum: *Sorghum bicolor* (L.) Moench [110] and maize [163]) were downloaded from the Plant Transcription Factor Database version 3.0 (http://planttfdb.cbi.pku.edu.cn/family.php?fam=WRKY). These sequences were used to identify homologous peptides from *S. italica* through a Basic Local Alignment Search Tool for Proteins (BLASTP) search in the Phytozome database, using default parameters (Goodstein *et al*. 2012). All hits with Expect (*E*) values <1·0 were retrieved and redundant sequences were removed manually. Each non-redundant sequence was checked in the Simple Modular Architecture Research Tool (SMART) database (http://smart.emblheidelberg.de) for presence of the conserved WRKY domain.

The hits with *E* value $<1 \times 10^{-5}$ and over 80% identity were considered orthologous. The comparative orthologous relationships of *WRKY* genes among foxtail millet, sorghum, maize and rice were illustrated using the circular visualization software Circos (Krzywinski *et al*. 2009). To estimate synonymous (*Ks*) and non-synonymous (*Ka*) substitution rates, the corresponding amino acid and cDNA sequences of orthologous WRKY proteins were analysed using PAL2NAL (http://www.bork.embl.de/pal2nal/) (Suyama *et al*. 2006).

Setaria italica and *Arabidopsis* WRKY protein sequences alignment and phylogenetic tree construction

The full protein sequences of WRKY transcription factors from *Arabidopsis* and *S. italica* were used in a multiple sequence alignment. The alignment was then used to construct a neighbour-joining phylogenetic tree in the Molecular Evolutionary Genetics Analysis (MEGA) 5·0 program. Bootstrap values were calculated with 1000 iterations. Based on the multiple sequence alignment and the previously reported classification of *Arabidopsis* AtWRKY proteins (Eulgem *et al*. 2000), the SiWRKY proteins were assigned to three different groups.

Analysis of the conserved amino acid sequence domains for SiWRKY transcription factors

Protein motifs were identified using the Multiple EM (Expectation Maximization) for Motif Elicitation

https://doi.org/10.1017/S0021859616000873 Published online by Cambridge University Press

(MEME) (http://meme.nbcr.net/meme/). The analysis was performed with the following settings: number of repetitions, any; maximum number of motifs, 15; and optimum width of the motif, 10–60.

Chromosomal localization of the *SiWRKYs* and estimation of their genomic duplication

For each *SiWRKY* gene, the chromosome number plus the gene start and end positions were identified in Phytozome version 9.1 (Table S1, available from https://www.cambridge.org/core/journals/journal-of-agricultural-science). The *SiWRKY* genes were then plotted onto the respective chromosomes according to the ascending order of their physical position (in base pairs, bp). The resultant physical map was created using MapChart version 2.2 (Voorrips 2002).

Tandem duplication was defined as gene clusters located within 30 kb of each other on a chromosome (Shiu & Bleecker 2003; Puranik *et al*. 2013). Segmental duplication was analysed via a BLASTP search against the complete peptide sequences of *S. italica*, and the first five matches with *E* value $<1 \times 10^{-5}$ were chosen as potential anchors. Collinear blocks were assessed in Multiple Collinearity Scan (MCScan) version 0.8 and alignments with *E* value $<1 \times 10^{-5}$ were considered significant matches (Tang *et al*. 2008; Puranik *et al*. 2013).

In silico expression profiling

RNA sequencing data from four *S. italica* tissues types, namely spica, stem, leaf and root, were retrieved from the European Nucleotide Archive (http://www.ebi.ac.uk/ena SRX128226 [spica], SRX128225 [stem], SRX128224 [leaf], SRX128223 [root)]) (Cochrane *et al*. 2013). All paired-end Illumina reads from the four different tissue samples were mapped onto the foxtail millet gene sequences using Bowtie2 (Langmead & Salzberg 2012), and the number of mapped reads was normalized with the reads per kilobase per million (RPKM) method (Pepke *et al*. 2009). Based on the RPKM value of each *SiWRKY* gene in the respective tissues, a heat map showing tissue-specific expression was generated using The Institute for Genomic Research (TIGR) Multi Experiment Viewer (MeV4·9·0) (Saeed *et al*. 2006). Approximately 2 kb of genomic DNA sequence, upstream of each *SiWRKY* gene, was retrieved from Phytozome, and the cis-regulatory elements were identified using the plant cis-acting regulatory DNA elements (PLACE) database (http://www.dna.affrc.go.jp/PLACE/signalup.html).

Plant material and stress treatments

Seeds of foxtail millet 'Jingu21' were obtained from the Shanxi Agriculture University, Shanxi, China, and grown in pots (diameter 20 cm and height 50 cm) containing vermiculite, in a greenhouse under a 14-h photoperiod. The temperature was set to 28 °C during the day and 20 °C at night, with relative humidity set to 70%. For stress treatments, the plants were carefully pulled out and the roots of seedlings were immersed in solutions containing 0·2 g/ml polyethylene glycol 6000 for 1, 6, 12 and 24 h. After the treatments, roots were collected, frozen immediately in liquid nitrogen, and stored at −80 °C until RNA isolation. Three independent replicates were collected for each time point to ensure precision and reproducibility.

RNA isolation and quantitative real-time polymerase chain reaction analysis

Total RNA was isolated from collected samples using a Plant Total RNA Isolation Kit (TIANGEN, Beijing, China) and treated with RNase-free DNase I (RQ1, Promega, Madison, USA). SuperScript III Reverse Transcriptase (Invitrogen) with Oligo(dT)18 (Promega) was used to synthesize cDNA, following manufacturer protocol. The quantitative real-time polymerase chain reaction (qRT-PCR) primers (Table S2, available from https://www.cambridge.org/core/journals/journal-of-agricultural-science) for *SiWRKY* genes were designed with the GenScript Real-time PCR Primer Design tool (https://www.genscript.com/ssl-bin/app/primer). The StepOne™ Real-Time PCR System (Applied Biosystems, Delaware, USA) was used for PCR reactions, and the thermocycling procedure was as follows: 40 cycles (95 °C for 15 s and 60 °C for 1 min). The size of PCR products was checked with agarose gel electrophoresis. A constitutive *18sRNA* gene was used as an endogenous control. Three biological replicates were performed.

RESULTS

Identification of *SiWRKY* genes in *Setaria italica*

The keyword search in Phytozome yielded 109 *SiWRKY* genes. Verification of WRKY domain presence using SMART resulted in the exclusion of four

https://doi.org/10.1017/S0021859616000873 Published online by Cambridge University Press

genes lacking the conserved WRKY domain (Si032748m, Si003876m, Si013326m and Si027982m) and two genes with incomplete C-terminal zinc finger motifs (Si002957m and Si014936m). The BLASTP search against the *S. italica* genome database using 472 WRKY protein sequences from *A. thaliana, O. sativa, S. bicolor* and *Z. mays* identified 106 SiWRKYs (E value <1·0). Transcripts Si036581m, Si029764m and Si030012m were redundant with Si036565m, Si029245m and Si029342m, respectively. The final number of *SiWRKYs* genes obtained from the two methods was identical.

The final 103 transcripts were named *SiWRKY*1 to *SiWRKY*103, following their order on the chromosomes (Fig. 1; Table S1, available from https://www.cambridge.org/core/journals/journal-of-agricultural-science). The predicted sizes of 75 out of 103 SiWRKY proteins were between 200 amino acids (aa) to 400 aa. Si028710m was predicted to be the longest protein sequence (1291 aa), whereas Si038955m was predicted to be the shortest protein (94 aa).

Genomic distribution of *SiWRKY* genes reveals gene clustering and tandem duplications.

All 103 *SiWRKYs* genes were physically mapped onto nine chromosomes of *S. italica* (Fig. 1). The map revealed a non-random distribution of *SiWRKYs* genes in the genome. Two to five *SiWRKY* gene clusters were identified (black boxes, Fig. 1). The largest cluster of five *SiWRKY* genes was on chromosomes 5 and 7. In total, 30 *SiWRKY* genes were tandem duplicates. Twenty-six of these were in group III, suggesting that tandem duplication contributed to the expansion of the SiWRKY family.

Phylogenetic classification of SiWRKYs and identification of motif conservation

To examine the phylogenetic relationships of SiWRKYs, a phylogenetic tree was constructed based on the whole protein sequences of 103 SiWRKYs and seven AtWRKYs (*Arabidopsis* WRKY proteins). Seven groups (I, IIa, IIb, IIc, IId, IIe and III) were categorized in *Arabidopsis* and seven AtWRKYs were selected: group I (AtWRKY33), group IIa (AtWRKY40), group IIb (AtWRKY47), group IIc (AtWRKYY50), group IId (AtWRKY69), group IIe (AtWRKY27) and group III (AtWRKY53). The 103 SiWRKY proteins were categorized into seven groups according to the *Arabidopsis* categories. The largest group in *S. italica* was group III, which was further divided into two subgroups (IIIa and IIIb; Fig. 2). Group I contained two WRKY domains and SiWRKY10, with one WRKY domain, was also categorized into group I. The group IIc proteins were close to group I in the phylogenetic tree, but contained only one WRKY domain. Among groups II, IIa and IIb shared a close evolutionary distance, whereas subgroups IId and IIe were closer (Fig. 2).

From the protein sequences of 103 SiWRKYs, 15 conserved motifs were identified (Table S3, available from https://www.cambridge.org/core/journals/journal-of-agricultural-science). The motif distribution corresponding to the SiWRKY gene family revealed that SiWRKY proteins in the same sub-family had similar motifs (Fig. 3). All SiWRKY proteins contained motifs 1, 2 and 3. Motif 4 was observed in the more closely related groups I and IIc (Fig. 3). Motifs 5 and 6 were specifically found in SiWRKY73, 81, 83, 84, 85, 86 and 88 of group III. Motif 7 was the N-terminal WRKY domain of group I proteins and was therefore primarily found in that group.

Syntenic relationships of *SiWRKY* genes among foxtail millet, rice, sorghum and maize

To obtain the syntenic relationships of *SiWRKYs*, comparative maps were constructed with *SiWRKY* genes, as well as orthologous rice, sorghum and maize genes (Fig. 4). The ratios of *Ka v. Ks* substitution (*Ka/Ks*) were estimated for the orthologous genes of *SiWRKY* with those of sorghum (25), maize (19) and rice (11). This analysis showed that the SiWRKY family experienced strong purifying selection, as *Ka/Ks* ratios of the duplicated genes equalled one. Among the orthologous gene pairs of *SiWRKYs* and *WRKYs* of other grass species, the average *Ka/Ks* value was highest between foxtail millet and rice (0·24) and lowest between foxtail millet and maize (0·13; Table S4, available from https://www.cambridge.org/core/journals/journal-of-agricultural-science).

In silico tissue-specific expression profiling of *SiWRKY* genes

Gene expression with high tissue specificity is generally considered a strong indication for a specific role in the tissue involved. The strong drought resistance of *S. italica* means that *SiWRKY* gene specificity in roots is of special interest, because roots are the key organ perceiving drought signals. The results of *SiWRKY* gene specificity analyses in the root, leaf, stem and spica revealed that the expression of 84

https://doi.org/10.1017/S0021859616000873 Published online by Cambridge University Press

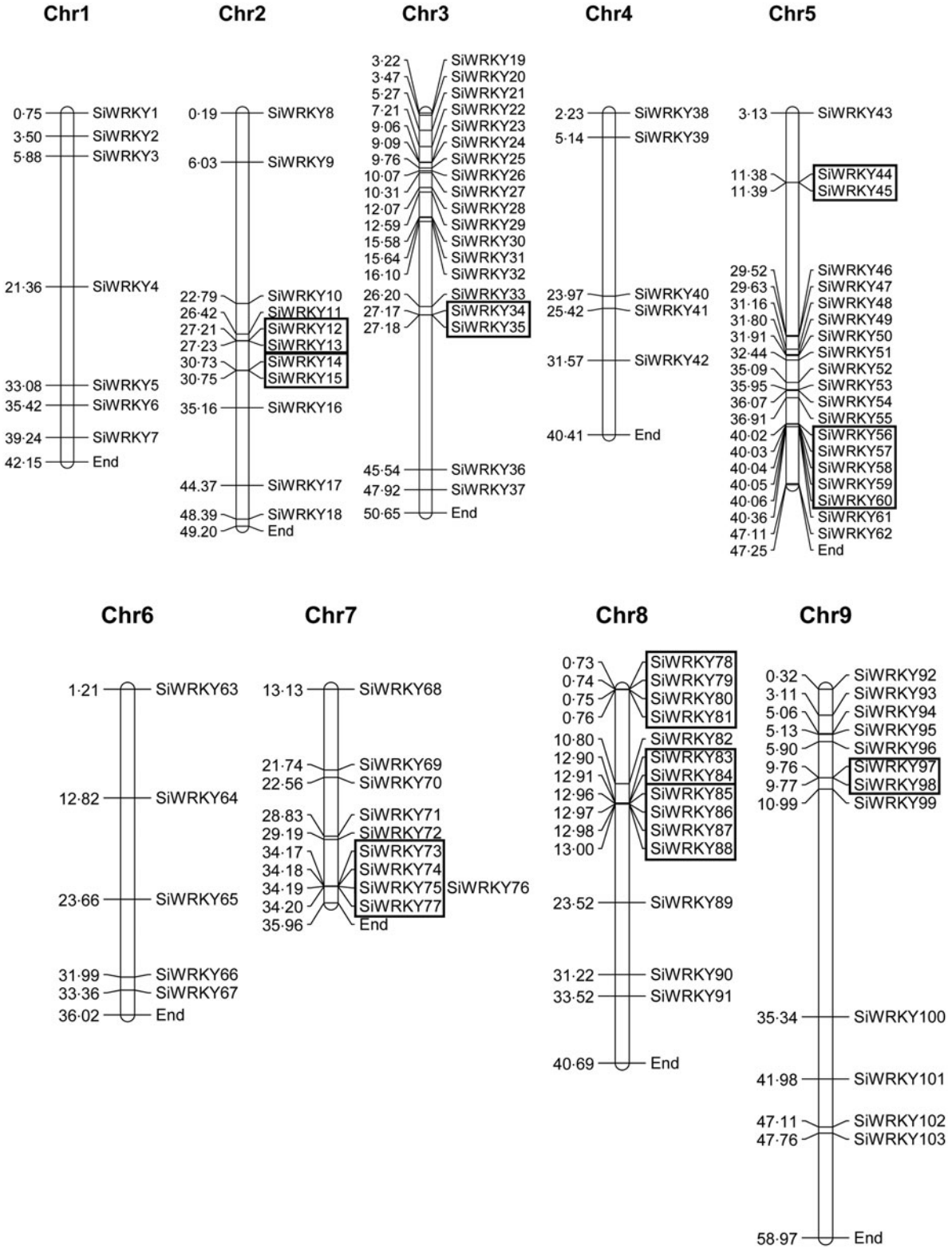

Fig. 1. Distribution of 103 *SiWRKY* genes on nine foxtail millet chromosomes. Graphical representation of physical locations for each *SiWRKY* gene on foxtail millet chromosomes (numbered Chr1–9). Tandem-duplicated genes on a particular chromosome are indicated with black boxes. Chromosomal distances are given in Mb.

https://doi.org/10.1017/S0021859616000873 Published online by Cambridge University Press

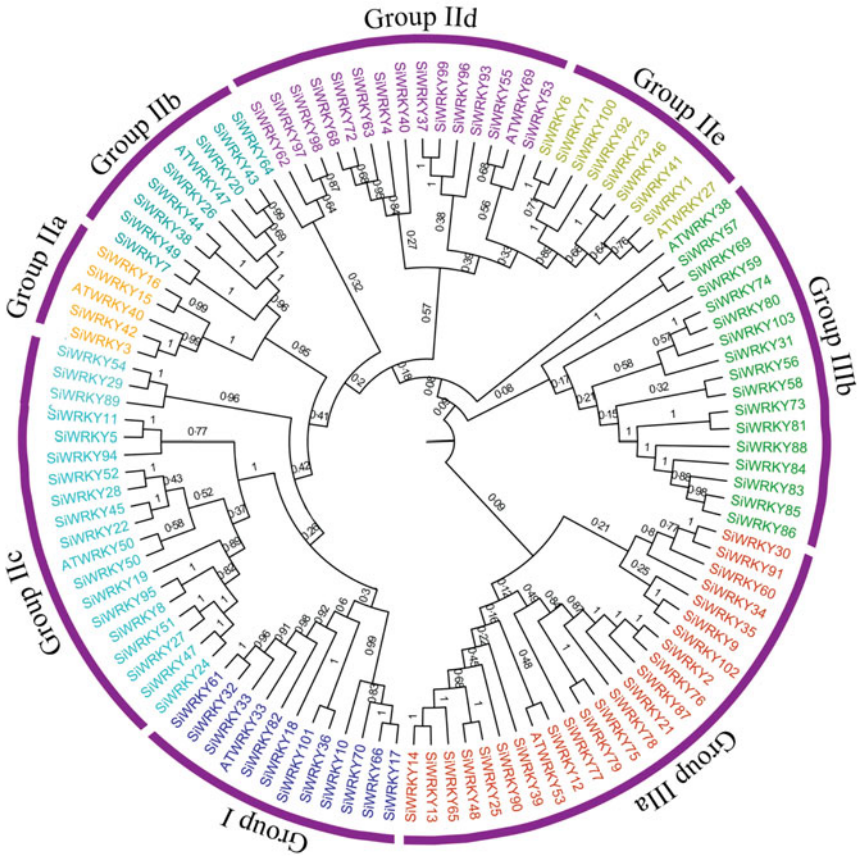

Fig. 2. Phylogenetic tree of WRKY proteins from foxtail millet (SiWRKYs) and *Arabidopsis* (AtWRKYs). The sequences were aligned with Clustal W in MEGA5 and the phylogenetic tree was constructed using the neighbour-joining method. Proteins were classified into three distinct clusters and each group was assigned a different colour. Group (I, II and III) and subgroup (IIa, IIb, IIc, IId and IIe) names are indicated around the outside of the circle. Colour online.

SiWRKY genes was higher in the root than in the other tissues examined. Among those, *SiWRKY*19, 56, 83, 84, 85, 86, 87, 88 and 91 were specifically expressed in the root (Fig. 5; Table S5, available from https://www.cambridge.org/core/journals/journal-of-agricultural-science). However, the heat map showing tissue-specific expression revealed that over half of the genes were expressed in all four tissues, suggesting constitutive expression (Fig. 5; Table S5, available from https://www.cambridge.org/core/journals/journal-of-agricultural-science).

Analysis of *SiWRKY* gene expression in response to drought stress

To further investigate *SiWRKYs* in response to drought stress, the roots of seedlings treated with 1-h drought stress were used for qRT-PCR expression analysis of all 103 *SiWRKYs*. The results showed that most genes were up-regulated after drought stress; however, several genes were down-regulated (Fig. 6). To identify putative drought-response genes, eight genes (*SiWRKY*9, 19, 44, 59, 85, 87, 96 and 102) with expression two-fold higher than the control were further examined with qRT-PCR at 0, 1, 6, 12 and 24 h of dehydration. These *SiWRKY* genes exhibited tissue-specific responses to drought stress (Fig. 7) with increased expression in roots but little change in leaves. Most gene expression (*SiWRKY*19, 59, 85, 87 and 96) in roots peaked at 6 h, but became down-regulated afterwards. In contrast, the remaining genes exhibited a more variable expression pattern with double peaks: one at 1 h and the other at 12 h (Fig. 7). Interestingly, the expression of root-specific genes (*SiWRKY*85 and *SiWRKY*87) was low under normal environmental conditions but became highly up-regulated once drought was detected.

https://doi.org/10.1017/S0021859616000873 Published online by Cambridge University Press

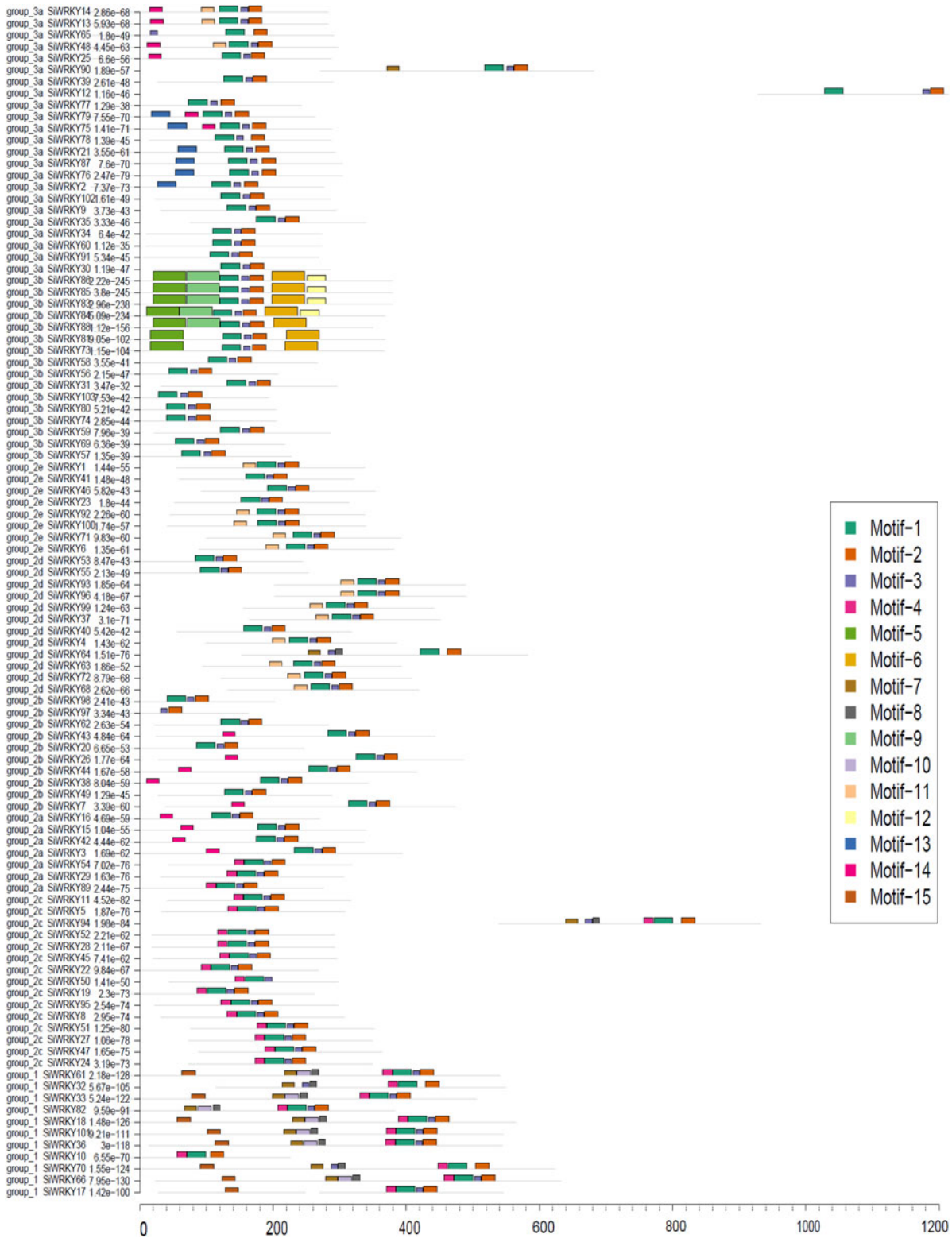

Fig. 3. Schematic representation of amino acid motifs in SiWRKY proteins from different groups. Motif analysis was performed using Meme 4·0 software as described in the Materials and Methods. The selected WRKY proteins are listed on the left. The black solid line represents the corresponding WRKY protein and its length. The differently colour boxes represent separate motifs and their position in each WRKY sequence. Colour online.

https://doi.org/10.1017/S0021859616000873 Published online by Cambridge University Press

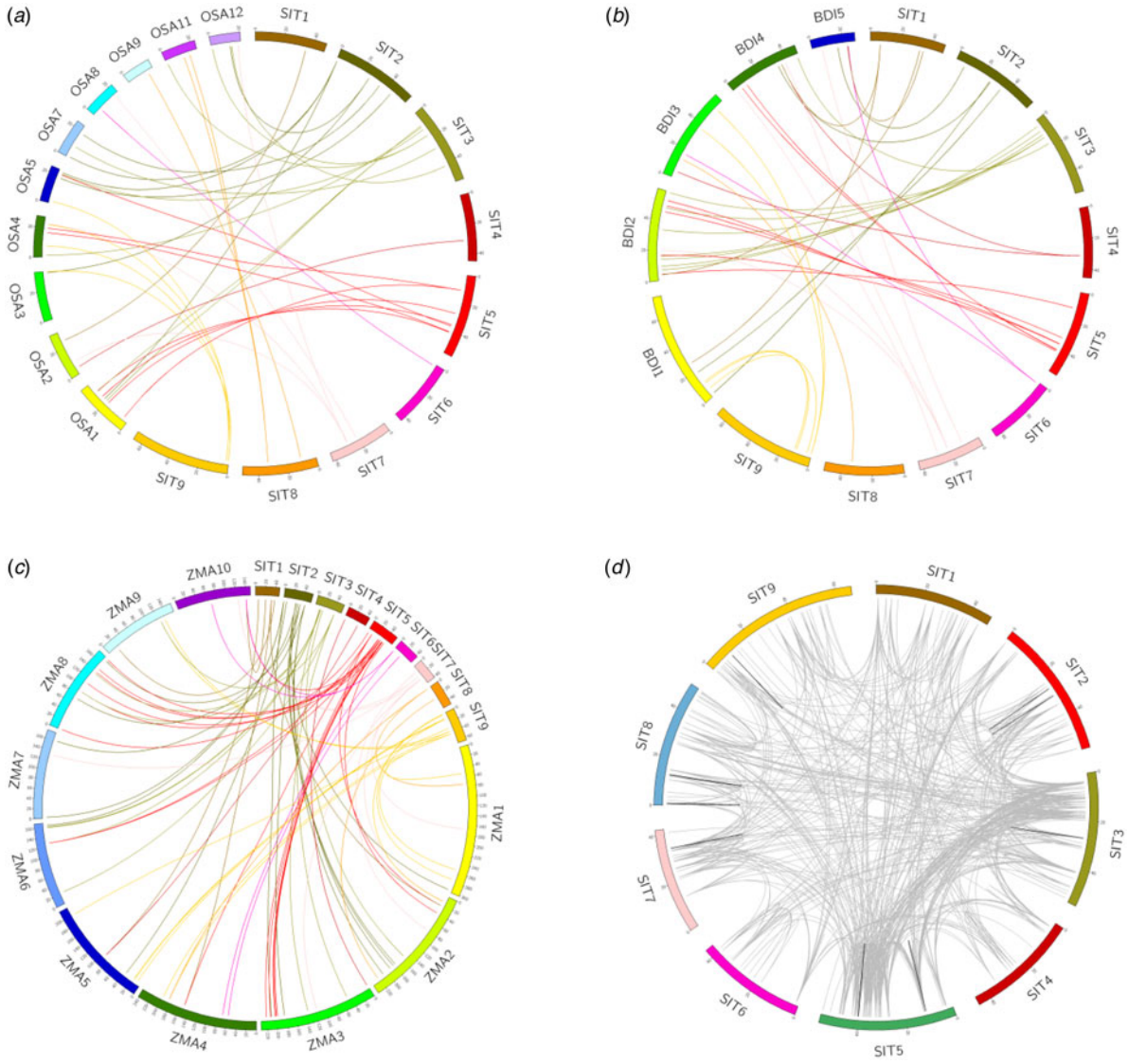

Fig. 4. Comparative physical mapping showing the degree of orthologous relationships of *SiWRKY* genes located on nine chromosomes of foxtail millet with (a) sorghum, (b) maize, (c) rice, (d) *S. italica*. Colour online.

DISCUSSION

Characterization and phylogenetic relationships of WRKY proteins in *Setaria italica*

The present study facilitated the characterization and phylogenetic relationships of *S. italic* WRKY transcription factors family proteins, which are important regulators of major plant processes such as growth, development and stress responses. Thus, these proteins have been subjected to intensive investigations in various crop plants. In the present study, 103 *SiWRKY* genes were identified, after excluding two genes with incomplete C-terminal zinc finger motifs (Si002957m and Si014936m). In addition, a phylogenetic tree was constructed via a multiple sequence alignment using whole-protein sequences of SiWRKYs and seven *Arabidopsis* WRKYs (AtWRKYs). Phylogenetic analysis revealed that SiWRKY and AtWRKY proteins can be classified into three major groups (I, II and III). Group III contained only 14 AtWRKYs out of 90 in the AtWRKY family (Eulgem *et al*. 2000), but 39 out of 103 SiWRKY proteins were in group III, indicating an expansion of *WRKY* genes in this group during *S. italica* evolution. The current results are in agreement with a recent global analysis of WRKY transcription factors in *S. italica*, which provided an overview of SiWRKY transcription factors structure and potential function in

https://doi.org/10.1017/S0021859616000873 Published online by Cambridge University Press

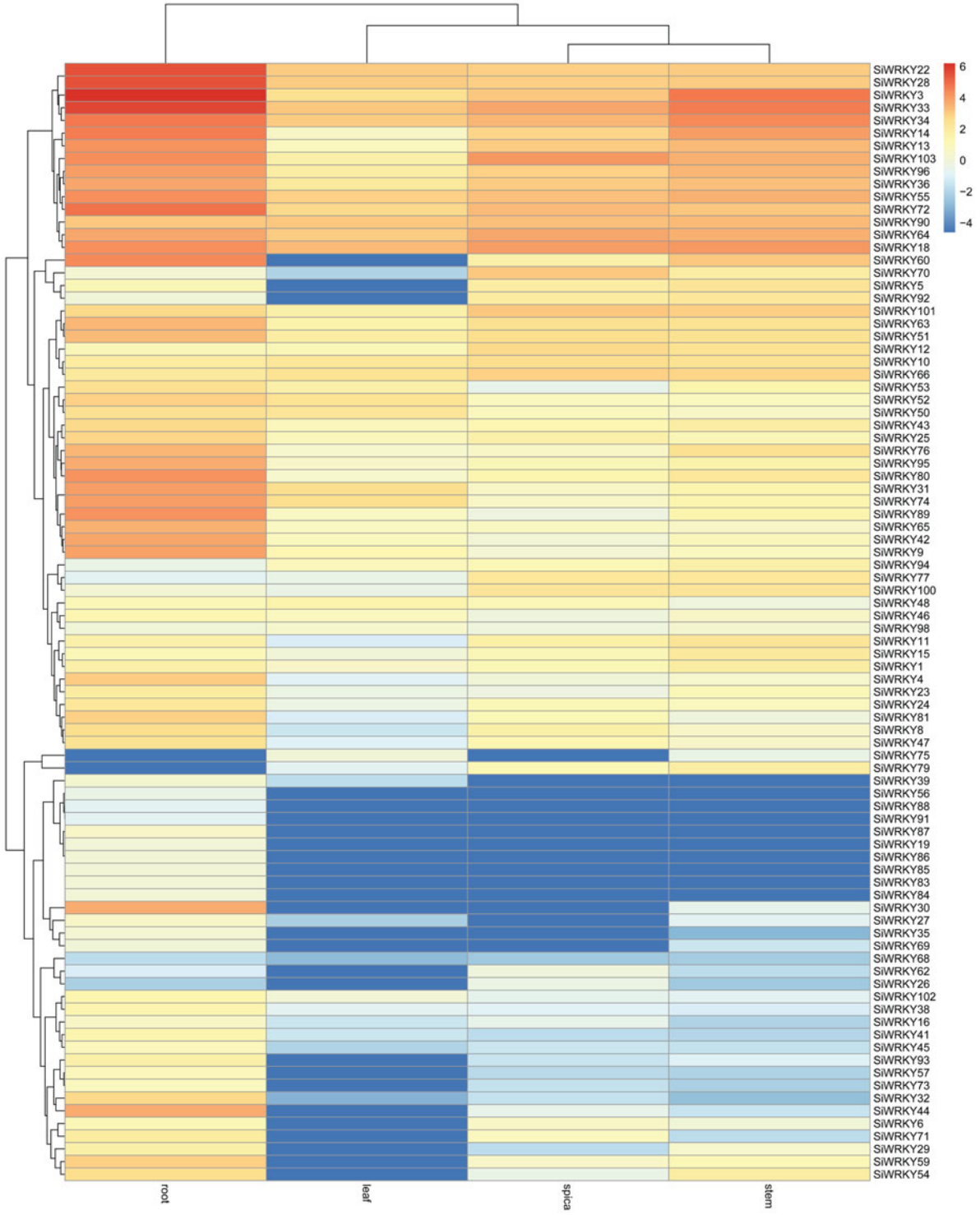

Fig. 5. Heat map showing the expression patterns of *SiWRKY* genes in four tissues: leaf, root, stem, and spica. The colour scales for fold-change values are on the right. Eighty-four out of 103 *SiWRKYs* were highly expressed in root tissue. Note that expression values mapped onto a colour gradient from low (blue) to high (orange). Colour online.

abiotic stress signalling (Muthamilarasan *et al*. 2015). Muthamilarasan *et al*. (2015) identified 105 *SiWRKY* genes and also noted the expansion of *SiWRKY* genes in one group. The mechanism behind SiWRKY family expansion may at least partially involve tandem duplication of paralogous group III genes, as most of the 30 tandemly duplicated genes identified in the current work were from group III.

https://doi.org/10.1017/S0021859616000873 Published online by Cambridge University Press

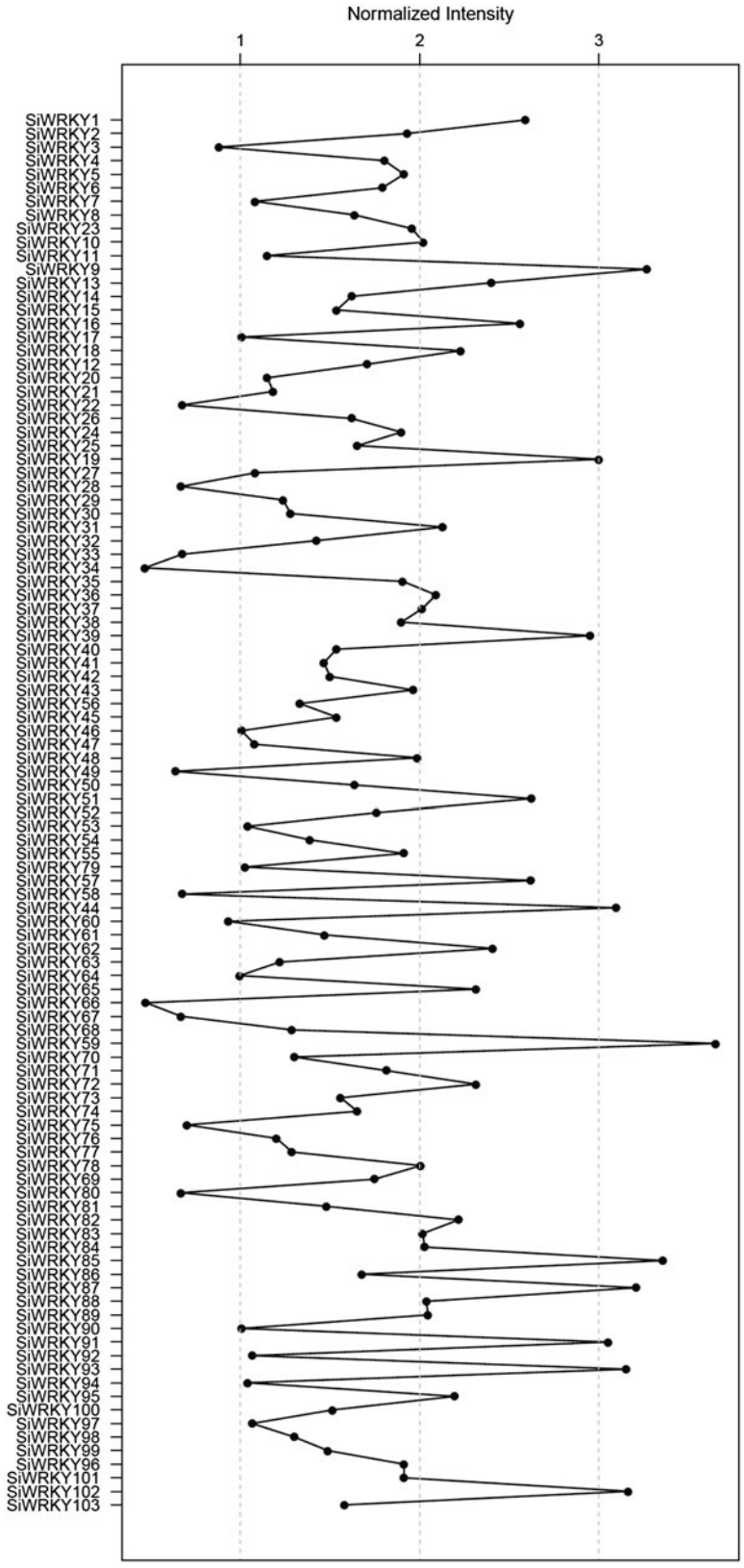

Fig. 6. Pairwise comparisons of *SiWRKY* gene expression profiles in response to drought. The signal intensities of qRT-PCR results were normalized to the mean expression values and plotted in log-scale for all *SiWRKY* genes.

https://doi.org/10.1017/S0021859616000873 Published online by Cambridge University Press

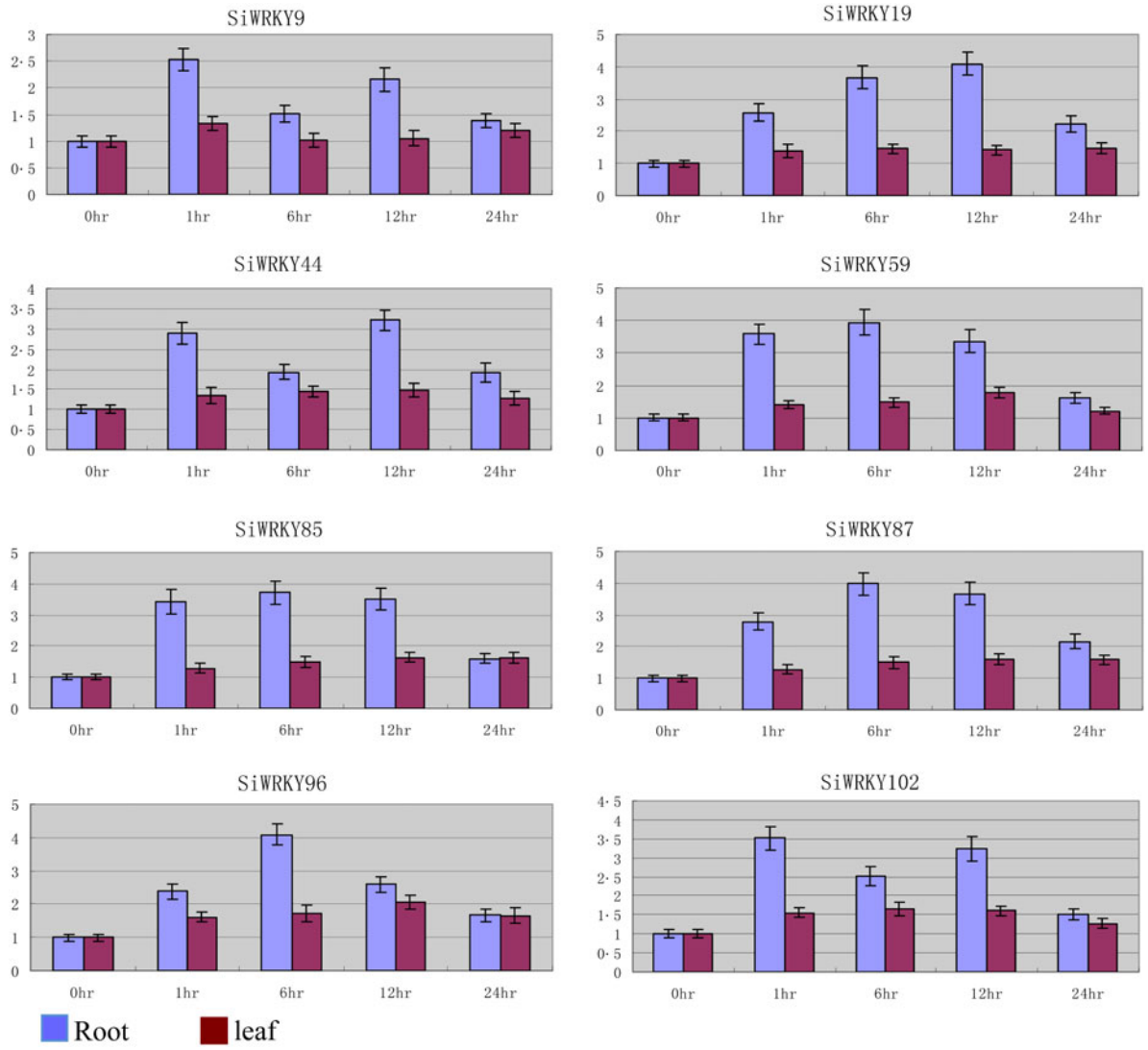

Fig. 7. The relative expression ratio of eight *SiWRKY* genes analysed with qRT-PCR at 0, 1, 6, 12 and 24 h of dehydration. For all graphs, relative amounts of selected RNA were evaluated for gene expression using the $2^{(-\Delta\Delta C_t)}$ method. Error bars indicate standard errors of the mean among replicates. Significant differences between treated samples and control sample (0 h under stress) were examined with *t*-tests. If $P < 0{\cdot}01$, *SiWRKY* genes were considered differentially expressed. *18sRNA* was used as an internal control to normalize the data. Colour online.

Tandem duplication of *WRKY* genes also occurs in maize and *Brachypodium* (Tripathi *et al*. 2012; Wei *et al*. 2012), suggesting that it may be a common mechanism in grasses for the expansion of this gene family.

The present study found that several *SiWRKY* genes were syntenic with *WRKY* genes in rice, sorghum and maize. A previous investigation aiming to infer grass genome evolution uncovered key chromosome shuffling events between these three crops, after using whole-genome sequences to examine collinear relationships between *S. italica, Brachypodium*, rice, sorghum and maize (Zhang *et al*. 2012). Similarly, comparative mapping of *WRKY* orthologues between *S. italica*, maize, switchgrass and sorghum demonstrated syntenic relationships between them (Muthamilarasan *et al*. 2015). Together, these results suggest that chromosomal rearrangements predominantly shaped *WRKY* gene distribution and organization in these genomes, providing insight on WRKY family evolution among grasses. In addition, the present study should be helpful in selecting candidate *WRKY* genes for the genetic improvement of related grass family members.

https://doi.org/10.1017/S0021859616000873 Published online by Cambridge University Press

Expression profiles of *SiWRKY* genes under drought stress

The function of WRKY transcription factors in drought stress tolerance is fairly well established across various species (Golldack *et al*. 2011; Tripathi *et al*. 2014), but none of the examined crops is naturally drought-tolerant plants. Thus, drought-resistant *S. italica* is a useful model plant for understanding *WRKY* genes involvement in the genetic mechanisms of improved dehydration tolerance. In the present study, most *SiWRKY* genes were up-regulated in the roots after 1 h of drought stress. These results are in accord with recent findings in 21-day-old *S. italica* seedlings, which exhibited up-regulation of four *SiWRKY* genes after a 7-h drought treatment (Muthamilarasan *et al*. 2015). Together, the data strongly suggest that *SiWRKY* genes are involved in the high drought resistance of *S. italica*, with most *SiWRKY*s responding quickly to low water availability. Studies in other plants indicate a similar role of *WRKY* genes. For example, *OsWRKY*45 overexpression increased drought and salt tolerance in rice (Qiu & Yu 2009). Moreover, transgenic *Arabidopsis* overexpressing soybean *WRKYs* exhibited increased drought tolerance (Zhou *et al*. 2008; Luo *et al*. 2013), and *AtWRKY*57 has been implicated in natural *Arabidopsis* drought resistance (Jiang & Deyholos 2009).

The results of *in silico* expression profiling provide further indication that *WRKY* genes affect *S. italica* response to drought stress. Specifically, 85 out of 103 *SiWRKY* genes were highly expressed in roots and nine genes were specifically expressed in root tissues. As roots adopt structural and functional modifications during stressful periods (Ghosh & Xu 2014), root-specific expression of most *SiWRKYs* hints at the genes' potential role in abiotic stress detection and response. Previous studies in other plant species have also linked *WRKYs* function to the roots. In *Arabidopsis*, root-specific *AtWRKY*46 plays an important role during lateral root development (Ding *et al*. 2015). Additionally, physic nut (*Jatropha curcas* L.) *JcWRKYs* from group IIe were highly expressed in roots and up-regulated under drought stress (Xiong *et al*. 2013). Finally, modulation of *OsWRKY*8 gene expression in rice altered root architecture and led to improved drought response (Song *et al*. 2009). Notably, the number of *SiWRKYs* expressed in *S. italica* roots is higher than in other plants, indicating that the crop's natural stress tolerance may be attributable to these transcription factors.

CONCLUSION

The present study characterized *SiWRKY* genes in foxtail millet (*S. italica*), enhancing existing knowledge of this major reserve crop and providing potential tools for breeding varieties with improved stress tolerance in this species and in major related crops. The expression profile analysis of *SiWRKY* genes in four different tissues, under both normal growth conditions and drought stress, demonstrated that many *SiWRKY* genes were highly expressed in the roots. These root-specific *SiWRKY* genes should be useful in future investigations of mechanisms behind root development and drought tolerance. Finally, estimation of the syntenic relationship in *S. italica* and related crops clarified WRKY family evolution among grasses. These results should provide a foundation for future phylogenetic studies using a greater number of crop species, which may potentially yield important insights on how artificial selection affected WRKY evolution.

This study was funded by the Natural Science Foundation of China (NSFC) under projects (Grant numbers 31401430 and 31501323), the Natural Science Foundation of Shanxi Province (Grant no. 2013021024-1) and a Research Project Supported by the Shanxi Scholarship Council of China (Grant no. 2016-070). GF XING was the recipient of a China Postdoctoral Science Foundation grant (No. 2012M520600). In addition, we would like to thank Editage (http://www.editage.cn) for English language editing.

SUPPLEMENTARY MATERIAL

The supplementary material for this article can be found at https://doi.org/10.1017/S0021859616000873.

REFERENCES

Bennetzen, J. L., Schmutz, J., Wang, H., Percifield, R., Hawkins, J., Pontaroli, A. C., Estep, M., Feng, L., Vaughn, J. N., Grimwood, J., Jenkins, J., Barry, K., Lindquist, E., Hellsten, U., Deshpande, S., Wang, X., Wu, X., Mitros, T., Triplett, J., Yang, X., Ye, C. Y., Mauro-Herrera, M., Wang, L., Li, P., Sharma, M., Sharma, R., Ronald, P. C., Panaud, O., Kellogg, E. A., Brutnell, T. P., Doust, A. N., Tuskan, G. A., Rokhsar, D. & Devos, K. M. (2012). Reference genome sequence of the model plant *Setaria*. *Nature Biotechnology* **30**, 555–561.

Cochrane, G., Alako, B., Amid, C., Bower, L., Cerdeño-Tárraga, A., Cleland, I., Gibson, R., Goodgame, N., Jang, M., Kay, S., Leinonen, R., Lin, X., Lopez, R., McWilliam, H., Oisel, A., Pakseresht, N., Pallreddy, S., Park, Y., Plaister, S., Radhakrishnan, R., Rivière, S., Rossello, M., Senf, A., Silvester, N., Smirnov, D., Ten Hoopen, P., Toribio, A., Vaughan, D. & Zalunin, V. (2013). Facing growth in the European Nucleotide Archive. *Nucleic Acids Research* **41**, D30–D35.

Ding, Z. J., Yan, J. Y., Li, C. X., Li, G. X., Wu, Y. R. & Zheng, S. J. (2015). Transcription factor WRKY46 modulates the development of *Arabidopsis* lateral roots in osmotic/salt stress conditions via regulation of ABA signaling and auxin homeostasis. *Plant Journal* **84**, 56–69.

Eulgem, T., Rushton, P. J., Robatzek, S. & Somssich, I. E. (2000). The WRKY superfamily of plant transcription factors. *Trends in Plant Science* **5**, 199–206.

Gallou, A., Declerck, S. & Cranenbrouck, S. (2012). Transcriptional regulation of defence genes and involvement of the WRKY transcription factor in arbuscular mycorrhizal potato root colonization. *Functional and Integrative Genomics* **12**, 183–198.

Ghosh, D. & Xu, J. (2014). Abiotic stress responses in plant roots: a proteomics perspective. *Frontiers in Plant Science* **5**, 6.

Golldack, D., Lüking, I. & Yang, O. (2011). Plant tolerance to drought and salinity: stress regulating transcription factors and their functional significance in the cellular transcriptional network. *Plant Cell Reports* **30**, 1383–1391.

Goodstein, D. M., Shu, S. Q., Howson, R., Neupane, R., Hayes, R. D., Fazo, J., Mitros, T., Dirks, W., Hellsten, U., Putnam, N. & Rokhsar, D. S. (2012). Phytozome: a comparative platform for green plant genomics. *Nucleic Acids Research* **40**, D1178–D1186.

Jiang, Y. & Deyholos, M. K. (2009). Functional characterization of *Arabidopsis* NaCl inducible WRKY25 and WRKY33 transcription factors in abiotic stresses. *Plant Molecular Biology* **69**, 91–105.

Jiang, Y., Liang, G. & Yu, D. (2012). Activated expression of WRKY57 confers drought tolerance in *Arabidopsis*. *Molecular Plant* **5**, 1375–1388.

Krzywinski, M. I., Schein, J. E., Birol, I., Connors, J., Gascoyne, R., Horsman, D., Jones, S. J. & Marra, M. A. (2009). Circos: an information aesthetic for comparative genomics. *Genome Research* **19**, 1639–1645.

Langmead, B. & Salzberg, S. L. (2012). Fast gapped-read alignment with Bowtie 2. *Nature Methods* **9**, 357–359.

Luo, X., Bai, X., Sun, X., Zhu, D., Liu, B., Ji, W., Cai, H., Cao, L., Wu, J., Hu, M., Liu, X., Tang, L. & Zhu, Y. (2013). Expression of wild soybean WRKY20 in *Arabidopsis* enhances drought tolerance and regulates ABA signalling. *Journal of Experimental Botany* **64**, 2155–2169.

Marchive, C., Mzid, R., Deluc, L., Barrieu, F., Pirrello, J., Gauthier, A., Corio-Costet, M. F., Regad, F., Cailleteau, B., Hamdi, S. & Lauvergeat, V. (2007). Isolation and characterization of a *Vitis vinifera* transcription factor, VvWRKY1, and its effect on responses to fungal pathogens in transgenic tobacco plants. *Journal of Experimental Botany* **58**, 1999–2010.

Muthamilarasan, M. & Prasad, M. (2015). Advances in *Setaria* genomics for genetic improvement of cereals and bioenergy grasses. *Theoretical and Applied Genetics* **128**, 1–14.

Muthamilarasan, M., Bonthala, V. S., Khandelwal, R., Jaishankar, J., Shweta, S., Nawaz, K. & Prasad, M. (2015). Global analysis of WRKY transcription factor superfamily in *Setaria* identifies potential candidates involved in abiotic stress signaling. *Frontiers in Plant Science* **6**, 910.

Mukhtar, M. S., Deslandes, L., Auriac, M. C., Marco, Y. & Somssich, I. E. (2008). The *Arabidopsis* transcription factor WRKY27 influences wilt disease symptom development caused by *Ralstonia solanacearum*. *Plant Journal* **56**, 935–947.

Okay, S., Derelli, E. & Unver, T. (2014). Transcriptome-wide identification of bread wheat WRKY transcription factors in response to drought stress. *Molecular Genetics and Genomics* **289**, 765–781.

Pepke, S., Wold, B. & Mortazavi, A. (2009). Computation for ChIP-seq and RNA-seq studies. *Nature Methods* **6**, S22–S32.

Puranik, S., Sahu, P. P., Mandal, S. N., Suresh, B., V., Parida, S. K. & Prasad, M. (2013). Comprehensive genome-wide survey, genomic constitution and expression profiling of the NAC transcription factor family in foxtail millet (*Setaria italica* L.). *PLoS ONE* **8**, e64594. http://dx.doi.org/10.1371/journal.pone.0064594

Qiu, Y. P. & Yu, D. Q. (2009). Over-expression of the stress-induced *OsWRKY*45 enhances disease resistance and drought tolerance in *Arabidopsis*. *Environmental and Experimental Botany* **65**, 35–47.

Rushton, P. J., Torres, J. T., Parniske, M., Wernert, P., Hahlbrock, K. & Somssich, I. E. (1996). Interaction of elicitor-induced DNA-binding proteins with elicitor response elements in the promoters of parsley PR1 genes. *EMBO Journal* **15**, 5690–5700.

Rushton, P. J., Somssich, I. E., Ringler, P. & Shen, Q. J. (2010). WRKY transcription factors. *Trends in Plant Science* **15**, 247–258.

Saeed, A. I., Bhagabati, N. K., Braisted, J. C., Liang, W., Sharov, V., Howe, E. A., Li, J., Thiagarajan, M., White, J. A. & Quackenbush, J. (2006). TM4 microarray software suite. *Methods in Enzymology* **411**, 134–193.

Shiu, S. H. & Bleecker, A. B. (2003). Expansion of the receptor-like kinase/pelle gene family and receptor-like proteins in *Arabidopsis*. *Plant Physiology* **132**, 530–543.

Song, Y., Jing, S. J. & Yu, D. Q. (2009). Overexpression of the stress-induced *OsWRKY*08 improves osmotic stress tolerance in *Arabidopsis*. *Chinese Science Bulletin* **54**, 4671–4678.

Suyama, M., Torrents, D. & Bork, P. (2006). PAL2NAL: robust conversion of protein sequence alignments into the corresponding codon alignments. *Nucleic Acids Research* **34**, (Suppl. 2), W609–W612.

Tang, H., Bowers, J. E., Wang, X., Ming, R., Alam, M. & Paterson, A. H. (2008). Synteny and collinearity in plant genomes. *Science* **320**, 486–488.

Tao, Z., Liu, H. B., Qiu, D. Y., Zhou, Y., Li, X. H., Xu, C. G. & Wang, S. P. (2009). A pair of allelic WRKY genes play

opposite roles in rice-bacteria interactions. *Plant Physiology* **151**, 936–948.

Tripathi, P., Rabara, R. C., Langum, T. J., Boken, A. K., Rushton, D. L., Boomsma, D. D., Rinerson, C. I., Rabara, J., Reese, R. N., Chen, X., Rohila, J. S. & Rushton, P. J. (2012). The WRKY transcription factor family in *Brachypodium distachyon*. *BMC Genomics* **13**, 270.

Tripathi, P., Rabara, R. C. & Rushton, P. J. (2014). A systems biology perspective on the role of WRKY transcription factors in drought responses in plants. *Planta* **239**, 255–266.

Ülker, B. & Somssich, I. E. (2004). WRKY transcription factors: from DNA binding towards biological function. *Current Opinion in Plant Biology* **7**, 491–498.

Voorrips, R. E. (2002). MapChart: software for the graphical presentation of linkage maps and QTLs. *Journal of Heredity* **93**, 77–78.

Wang, Y. N., Dang, F. F., Liu, Z. Q., Wang, X., Eulgem, T., Lai, Y., Yu, L., She, J. J., Shi, Y. L., Lin, J. H., Chen, C. C., Guan, D. Y., Qiu, A. & He, S. L. (2013). *CaWRKY*58, encoding a group I WRKY transcription factor of *Capsicum annuum*, negatively regulates resistance to *Ralstonia solanacearum* infection. *Molecular Plant Pathology* **14**, 131–144.

Wang, L. N., Zhu, W., Fang, L. C., Sun, X. M., Su, L. Y., Liang, Z. C., Wang, N., Londo, J. P., Li, S. H. & Xin, H. P. (2014). Genome-wide identification of *WRKY* family genes and their response to cold stress in *Vitis vinifera*. *BMC Plant Biology* **14**, 103.

Wei, K. F., Chen, J., Chen, Y. F., Wu, L. J. & Xie, D. X. (2012). Molecular phylogenetic and expression analysis of the complete WRKY transcription factor family in maize. *DNA Research* **19**, 153–164.

Xie, Z., Zhang, Z. L., Zou, X., Huang, J., Ruas, P., Thompson, D. & Shen, Q. J. (2005). Annotations and functional analyses of the rice WRKY gene superfamily reveal positive and negative regulators of abscisic acid signaling in aleurone cells. *Plant Physiology* **137**, 176–189.

Xiong, W., Xu, X., Zhang, L., Wu, P., Chen, Y., Li, M., Jiang, H. & Wu, G. (2013). Genome-wide analysis of the *WRKY* gene family in physic nut (*Jatropha curcas* L.). *Gene* **524**, 124–132.

Yan, H. R., Jia, H. H., Chen, X. B., Hao, L. L., An, H. L. & Guo, X. Q. (2014). The cotton WRKY transcription factor GhWRKY17 functions in drought and salt stress in transgenic *Nicotiana benthamiana* through ABA signaling and the modulation of reactive oxygen species production. *Plant Cell Physiology* **55**, 2060–2076.

Yang, P. Z., Chen, C. H., Wang, Z. P., Fan, B. F. & Chen, Z. X. (1999). A pathogen- and salicylic acid-induced WRKY DNA-binding activity recognizes the elicitor response element of the tobacco class I chitinase gene promoter. *Plant Journal* **18**, 141–149.

Zhang, Y. & Wang, L. (2005). The WRKY transcription factor superfamily: its origin in eukaryotes and expansion in plants. *BMC Evolutionary Biology* **5**, 1.

Zhang, G., Liu, X., Quan, Z., Cheng, S., Xu, X., Pan, S. K., Xie, M., Zeng, P., Yue, Z., Wang, W. L., Tao, Y., Bian, C., Han, C. L., Xia, Q. J., Peng, X. H., Cao, R., Yang, X. H., Zhan, D. L., Hu, J. C., Zhang, Y. X., Li, H., Li, H., Li, N., Wang, J. Y., Wang, C. C., Wang, R., Guo, T., Cai, Y. J., Liu, C. Z., Xiang, H. T., Shi, Q. X., Huang, P., Chen, Q. C., Li, Y. G., Wang, J., Zhao, J. H. & Wang, J. (2012). Genome sequence of foxtail millet (*Setaria italica*) provides insights into grass evolution and biofuel potential. *Nature Biotechnology* **30**, 549–554.

Zhou, Q. Y., Tian, A. G., Zou, H. F., Xie, Z. M., Lei, G., Huang, J., Wang, C. M., Wang, H. W., Zhang, J. S. & Chen, S. Y. (2008). Soybean WRKY-type transcription factor genes, *GmWRKY*13, *GmWRKY*21, and *GmWRKY*54, confer differential tolerance to abiotic stresses in transgenic *Arabidopsis* plants. *Plant Biotechnology Journal* **6**, 486–503.

Zou, C. S., Jiang, W. B. & Yu, D. Q. (2010). Male gametophyte-specific WRKY34 transcription factor mediates cold sensitivity of mature pollen in *Arabidopsis*. *Journal of Experimental Botany* **61**, 3901–3914.