

Original Article

# Transcriptomics-based identification of *WRKY* genes and characterization of a salt and hormone-responsive *PgWRKY1* gene in *Panax ginseng*

Mohammed Nuruzzaman, Hongzhe Cao, Hao Xiu, Tiao Luo, Jijia Li, Xianghui Chen, Junli Luo, and Zhiyong Luo\*

Molecular Biology Research Center, State Key Laboratory of Medical Genetics, School of Life Sciences, Central South University, Changsha 410078, China

\*Correspondence address. Tel: +86-731-84805025; Fax: +86-731-84484492; E-mail: luozhiyong@csu.edu.cn

Received 15 July 2015; Accepted 30 August 2015

## Abstract

WRKY proteins belong to a transcription factor (TF) family and play dynamic roles in many plant processes, including plant responses to abiotic and biotic stresses, as well as secondary metabolism. However, no *WRKY* gene in *Panax ginseng* C.A. Meyer has been reported to date. In this study, a number of *WRKY* unigenes from methyl jasmonate (MeJA)-treated adventitious root transcriptome of this species were identified using next-generation sequencing technology. A total of 48 promising *WRKY* unigenes encoding WRKY proteins were obtained by eliminating wrong and incomplete open reading frame (ORF). Phylogenetic analysis reveals 48 WRKY TFs, including 11 Group I, 36 Group II, and 1 Group III. Moreover, one MeJA-responsive unigene designated as *PgWRKY1* was cloned and characterized. It contains an entire ORF of 1077 bp and encodes a polypeptide of 358 amino acid residues. The *PgWRKY1* protein contains a single WRKY domain consisting of a conserved amino acid sequence motif WRKYGQK and a C<sub>2</sub>H<sub>2</sub>-type zinc-finger motif belonging to WRKY subgroup II-d. Subcellular localization of *PgWRKY1*-GFP fusion protein in onion and tobacco epidermis cells revealed that *PgWRKY1* was exclusively present in the nucleus. Quantitative real-time polymerase chain reaction analysis demonstrated that the expression of *PgWRKY1* was relatively higher in roots and lateral roots compared with leaves, stems, and seeds. Importantly, *PgWRKY1* expression was significantly induced by salicylic acid, abscisic acid, and NaCl, but downregulated by MeJA treatment. These results suggested that *PgWRKY1* might be a multiple stress-inducible gene responding to hormones and salt stresses.

**Key words:** *Panax ginseng*, transcriptome, WRKY transcription factor, abiotic stress

## Introduction

WRKY (pronounced as ‘worky’) proteins form a large family of transcription factors (TFs) with their major occurrence in the plant kingdom [1]. The first cDNA-encoding WRKY protein was cloned in 1994 from sweet potato (*Ipomoea batatas*; *SPF1*), followed by wild oat (*Avena fatua*; *ABF1*, 2), and parsley (*Petroselinum crispum*; *PcWRKY1*, 2, 3) [2–4]. To date, an increasing number of genes

encoding WRKY proteins have been experimentally identified in numerous plants, including 74 members in *Arabidopsis thaliana*, more than 109 in soybean (*Glycine max*), 109 in rice (*Oryza sativa*), 57 in cucumber (*Cucumis sativus*), and 105 members in poplar (*Populus trichocarpa*) [1,5–8]. In addition, WRKY proteins have also been isolated from lower plants (e.g. ferns and mosses), protist (*Giardia lamblia*), and slime mold (*Dictyostelium discoideum*) [1,5].

The name 'WRKY' was originated from the WRKY domain that contains a region of about 60 amino acids. The WRKY family proteins contain either one or two WRKY domains, each of which is characterized by the WRKYGQK sequence motif at their N-terminal and a Cys<sub>2</sub>His<sub>2</sub> or Cys<sub>2</sub>HisCys zinc-finger motif at their C-terminal. In general, the hallmark heptapeptide WRKYGQK is highly conserved in plant WRKY-DNA-binding domain (DBD). Both motifs of the domain are necessary for high affinity binding of a WRKY protein with the W-box in the promoter of the target gene to modulate transcription [5,9]. All the members of the WRKY family are structurally divided into three distinct groups based on both the number of WRKY domains and the pattern of their zinc-finger motif. Group I WRKYs typically contain two WRKY domains, while those with one WRKY domain belong to either Group II or III differentiated from each other according to the features of their zinc-finger motif. The Group II is further divided into five subgroups (II-a, II-b, II-c, II-d, and II-e) on the basis of additional short conserved structural motifs outside the WRKY domain. Both Group I and II have the same pattern of zinc-finger motif, namely C<sub>2</sub>H<sub>2</sub>, while the domain in Group III contains a C<sub>2</sub>HC motif [6,9].

In plants, numerous WRKY proteins appear to be associated with response to biotic and abiotic stresses of wounding, pests, pathogens, drought, heat shock, freezing, cold adaptation, as well as nutrient deficiency [10–13]. It is also evident that several members of this family might have key regulatory functions in trichome and seed coat development, embryogenesis, senescence, seed dormancy, germination, as well as metabolic pathways [14–18]. Over the past few years, WRKY TFs have been proved in the induction of natural products biosynthesis. CjWRKY1 from *Coptis japonica* regulates the production of the benzyloquinoline alkaloid, berberine [19]. A number of biosynthetic genes for the sesquiterpene lactone, such as artemisinin which is a valuable anti-malaria drug, are regulated by *Artemisia annua* WRKY1 [20]. In addition to regulating the gene expression related to triterpene ginsenoside biosynthesis [21], heterologous over-expression of American ginseng WRKY gene, *PqWRKY1* in *Arabidopsis* was also found to induce gene expression associated with salt, drought, and disease resistance, leading to the development of seedling survival to salt and drought stress. A number of WRKY genes have been identified to respond to salt and various hormones such as MeJA, SA, and abscisic acid (ABA) [21,22] associated with plant defense and secondary metabolism. Identification of hormones and salt-responsive WRKYs will thus give valuable information on plant defense and natural product regulatory networks.

Ginseng (*Panax ginseng* C.A. Meyer) is the king of herbs, which belongs to the Araliaceae family, and has long been used as one of the most precious herbal medicines in Eastern Asia and North America [23]. The main pharmaceutically active compounds of ginseng are ginsenosides, used in the treatment of cardiovascular disease, obesity, diabetes, cancer, stress, and aging-related diseases [24–26]. The biosynthesis of these compounds is regulated by various biotic and abiotic stresses [27]. Over the past several years, a number of WRKY genes have been cloned and confirmed to play a pivotal role in regulating tolerance to biotic and abiotic stresses, as well as numerous secondary metabolites biosynthesis in many other medicinal species [19,21,28]. However, due to the shortage of gene sequence information, the TF of *P. ginseng* has not been explored. More recently, the transcriptome of MeJA-treated *P. ginseng* adventitious roots has first been sequenced and analyzed using Illumina/Solexa sequencing technology [29], and results reveals a number of helpful sequence information on various TF-related unigenes for further analysis, including full-length cDNA cloning and functional studies.

In this study, the identification of 48 *PgWRKY* unigenes was first reported based on our MeJA-treated *P. ginseng* transcriptome dataset [29]. Phylogenetic analysis was also performed based on their WRKY domain amino acid sequences to classify these WRKY proteins. A possible candidate named *PgWRKY1* responding to MeJA in our transcriptome dataset was cloned. Moreover, the expression profiles of the *PgWRKY1* gene under NaCl and various hormonal treatments, including MeJA, SA, and ABA, were determined using qRT-PCR. In addition, the tissue-specific expression characteristics, subcellular localization, and 3D structure of the *PgWRKY1* protein were analyzed. To the best of our knowledge, *PgWRKY1* is the first *P. ginseng* WRKY gene to be cloned and functionally characterized.

## Materials and Methods

### Plant materials

Four-year-old *P. ginseng* C.A. Meyer (Chinese Jilin ginseng cv. Damaya) was collected from the forest of the Changbai mountain of Fusong county, Jilin Province, China. MeJA-treated (Pg-MeJA) and untreated (Pg-Con) adventitious roots were used in our previous transcriptome sequencing study through the Illumina/Solexa technology [29]. Fresh ginseng adventitious roots were used for *PgWRKY1* gene clone and to examine the expression profiles under various treatments. The roots, lateral roots, stems, seeds, and leaves were collected separately for RNA isolation and used for tissue-specific expression analysis.

### TF search from Illumina/Solexa transcriptome dataset

MeJA-treated adventitious root transcriptome generated from 4-year-old *P. ginseng* through the next-generation sequencing (NGS) technology from our previous work provided a complete view on various kinds of unigenes [29]. To identify *PgWRKY* unigenes, this Illumina/Solexa dataset was annotated in the public databases, including NCBI non-redundant protein sequence (Nr), Swiss-prot, NCBI non-redundant nucleotide sequence (Nt), and Kyoto encyclopedia of genes and genomes (KEGG) using Blastx algorithm ( $E < 1 \times 10^{-5}$ ). Subsequently, the obtained *PgWRKY* gene sequences were further analyzed via open reading frame (ORF) finder to confirm the presence of ORF. Manual annotation was performed for the incorrectly predicted genes and generated a final list of putative *PgWRKYs* (cDNAs) based on the characteristics of plant WRKYs. The molecular evolution genetics analysis (MEGA) software (version 4.0) was then used to translate the unigenes into proteins [30]. Duplicate results were removed by searching the established invariant consensus sequences WRKYGQK and known alternative WRKYGKK, WRKY-GEK, and WRKYGSK consensus sequences from the list of proteins. The finally identified *PgWRKY* TFs, including complete N-terminal and C-terminal ends, were classified into three TF groups (Groups I, II, and III) by constructing unrooted phylogenetic tree using MEGA 4.0 with the neighbor-joining (NJ) method [30,31]. Afterwards, using complete amino acid sequences of *PgWRKY* members from Group II, blast analysis against protein sequence database of different plant species in NCBI was performed to find corresponding WRKY genes similar to that of ginseng species.

### Cloning of *PgWRKY1*

To identify the unigenes from our dataset, a pair of primers was designed according to one randomly selected annotated unigene (CL495. Contig5) responding to MeJA in our transcriptome dataset as follows: 5'-CTTTGCTTTTCCTTAGTTG-3' (forward primer)

**Table 1. Primers used for cloning of 5 *PgWRKYs***

Primer name	Primer sequence (5' to 3')
PgWRKY2-F	TCTTTTGGGTGTGTGAAT
PgWRKY2-R	TTTGCGTTTGTATGGTA
PgWRKY3-F	CTTTGCTTTTCCTTAGTTG
PgWRKY3-R	ATCATTATCTCTCCTTGG
PgWRKY4-F	CTTTGCTTTTCCTTAGTTG
PgWRKY4-R	AACAAACATTTCCAACCG
PgWRKY5-F	TCTTTTGGGTGTGTGAAT
PgWRKY5-R	TGGCGTTTGTATGGTAAT
PgWRKY6-F	CTTACCAGCCCTCTCCG
PgWRKY6-R	ATACTTTTGTAAAGGAGGACTA

and 5'-GAGTTCACACTGGATCATT-3' (reverse primer). Total RNA from untreated adventitious roots was extracted by E.Z.N.A plant RNA kit (Omega Bio-Tek, Doraville, USA) according to the protocols of the manufacturer and then treated with RNase-free DNase I (Omega Bio-Tek) to eliminate potential genomic DNA contamination. The first-strand cDNA was synthesized using 1 µg of total RNA and RevertAid First Strand cDNA Synthesis kit (Thermo Fisher Scientific, Waltham, USA) in a 20 µl reaction volume. PCR was performed using 2 × *Taq* MasterMix (CW BIO, Beijing, China) in a 16 µl reaction volume containing 8 µl of master mixture, 1 µl of each primer, 5.5 µl of ddH<sub>2</sub>O, and 0.5 µl of cDNA as the template. The PCR conditions were as follows: 94°C for 1 min; 94°C for 30 s, 58°C for 30 s, 72°C for 1 min with 35 cycles; and 72°C for 7 min. PCR products were analyzed in 1% agarose gels and purified by using Wizard SV Gel and PCR Clean-Up System (Promega, Madison, USA). The generated purified fragment was then cloned into the pGEM-T Easy vector (Promega) and transferred into *Escherichia coli* DH5α according to standard procedures. The successful clone was finally sequenced by Life Technologies (Shanghai, China). The cloned cDNA was initially termed as *PgWRKY1*. Moreover, five more *PgWRKY* genes were cloned and fully sequenced. All sequences have been submitted to GenBank (Supplementary Table S1). Primers used for cloning these five genes are listed in Table 1.

### Bioinformatics analysis

Heatmap was generated using HemI software (<http://hemi.biocuckoo.org/down.php>) [32]. The ORF analysis and amino acid sequence comparison were carried out using ORF finder (<http://www.ncbi.nlm.nih.gov/projects/gorf/>) and BLASTP tools (<http://www.ncbi.nlm.nih.gov/>), respectively. The theoretical iso-electric point (PI), molecular weight (MW), grand average of hydropathicity (GRAVY), and other physicochemical properties were determined using ExPASy server (<http://web.expasy.org/protparam/>). Phylogenetic tree and multiple sequence alignment were carried out with MEGA 4.0 using NJ algorithm with 100 bootstrap trials and DNAMAN software, respectively. DNA-binding domain was predicted using NCBI Conserved Domains Finder (<http://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi>). Transmembrane domain prediction was analyzed using the TMHMM Server V.2.0 (<http://www.cbs.dtu.dk/services/TMHMM-2.0/>). The secondary structure prediction for *PgWRKY1* protein was performed with NPS@ ([https://npsa-prabi.ibcp.fr/cgi-bin/npsa\\_automat.pl?page=npsa\\_gor4.html](https://npsa-prabi.ibcp.fr/cgi-bin/npsa_automat.pl?page=npsa_gor4.html)). The 3D structure predictions were accomplished using Phyre2 (<http://www.sbg.bio.ic.ac.uk/phyre2/html/page.cgi?id=index>) and Swiss-Model tool (<http://www.swissmodel.expasy.org/>) of ExPASy server. Protein function predictions were performed using ProtFun procedure

(<http://www.cbs.dtu.dk/services/ProtFun/>). Primers were designed using Primer Premier 5.0 software (<http://www.idtdna.com/Primerquest/Home/Index>).

### Subcellular localization analysis

To examine the subcellular localization of the *PgWRKY1* protein, the longest ORF of the *PgWRKY1* cDNA without termination codon was amplified by PCR with primer1 containing *Bgl*II site (underlined) at 5'-end: 5'-GAAGATCTAATGACTGTTGACCTGA-3' and primer2 containing *Spe*I site (underlined) at 5'-end: 5'-GGACTAGTCGAGCCGTCCGATTAG-3' using 2 × *Taq* Master Mix kit. The PCR product was introduced into the 5' side of the *GFP* gene of the pCAMBIA1302 vector (Supplementary Fig. S1) to generate the *CaMV35S::PgWRKY1:GFP* construct which was verified by sequencing. The *CaMV35S::GFP* construct was used as the control. The recombinant vector (pCAMBIA1302-*PgWRKY1*) and the control (pCAMBIA1302) were transformed into onion and tobacco epidermal cells with injector via *Agrobacterium tumefaciens* GV3101. After the transformed onion and tobacco epidermis cells were incubated for 48 h at 28°C in the dark, green fluorescent protein (GFP) and DAPI (4',6-diamidino-2-phenylindole) were visualized under a fluorescent microscope (Nikon Eclipse 80i; Nikon, Tokyo, Japan). GFP excitation was detected between 460 and 500 nm as blue color and emission from 515 to 550 nm as green color, while DAPI excitation was observed as ultraviolet (UV) light from 340 to 500 nm and emission from 435 to 485 nm in blue color. Finally, all the images were captured with a ×40 objective lens.

### Stress and hormone treatments

To examine the responses of *PgWRKY1* gene to salt and hormone treatments, adventitious roots that had been cultured for 2 weeks were transferred in Murashige and Skoog media containing 200 µM MeJA, 200 µM SA, 20 µM ABA, and 50 µM NaCl, and then kept shaking (110 rpm) at 28°C. The adventitious root samples at 0, 3, 6, 12, 24, 36, and 48 h post-treatment were collected and frozen in liquid nitrogen. The harvested samples were then stored at -80°C for RNA extraction.

### Expression analysis by quantitative real-time PCR (qRT-PCR)

Total RNA was extracted using E.Z.N.A plant RNA kit from the above harvested adventitious root samples according to the manufacturer's procedure. The RNA samples were then reverse transcribed into cDNA using the RevertAid First-Strand cDNA Synthesis kit (Thermo Fisher Scientific). The following gene-specific primers were used in the reaction: 5'-CACCAACAGCATCAACAGCG-3' (forward primer) and 5'-TTCGGGACCCATTAGACATCA-3' (reverse primer). A housekeeping gene *β-actin* (GenBank accession number KF699319.1) was used as standard control using the following primers: 5'-TGCCGGTTGAGGTGGTGGGT-3' (forward primer) and 5'-GAGCGTCC TACTAACAAGGCAGAG-3' (reverse primer). The cDNA was amplified with the fluorescent dye 2 × UltraSYBR Mixture (CW BIO) using Mastercycler ep realplex (Eppendorf, Hamburg, Germany) under the following conditions: 95°C for 10 min; 95°C for 15 s, 60°C for 30 s, 40 cycles. The reaction mixture (50 µl) contained 25 µl of 2 × UltraSYBR Mixture (with ROX), 1 µl of each primer (10 µM), 1.5 µl of cDNA, and 21.5 µl of ddH<sub>2</sub>O. All the qRT-PCRs, including non-template controls, were performed in triplicate with three independent biological replicates. At the end of the amplification cycle, the melting curve of amplified products was generated for each reaction to ensure specific amplification.

The relative expression level of *PgWRKY1* was calculated using the  $2^{-\Delta\Delta t}$  method.

### Statistical analysis

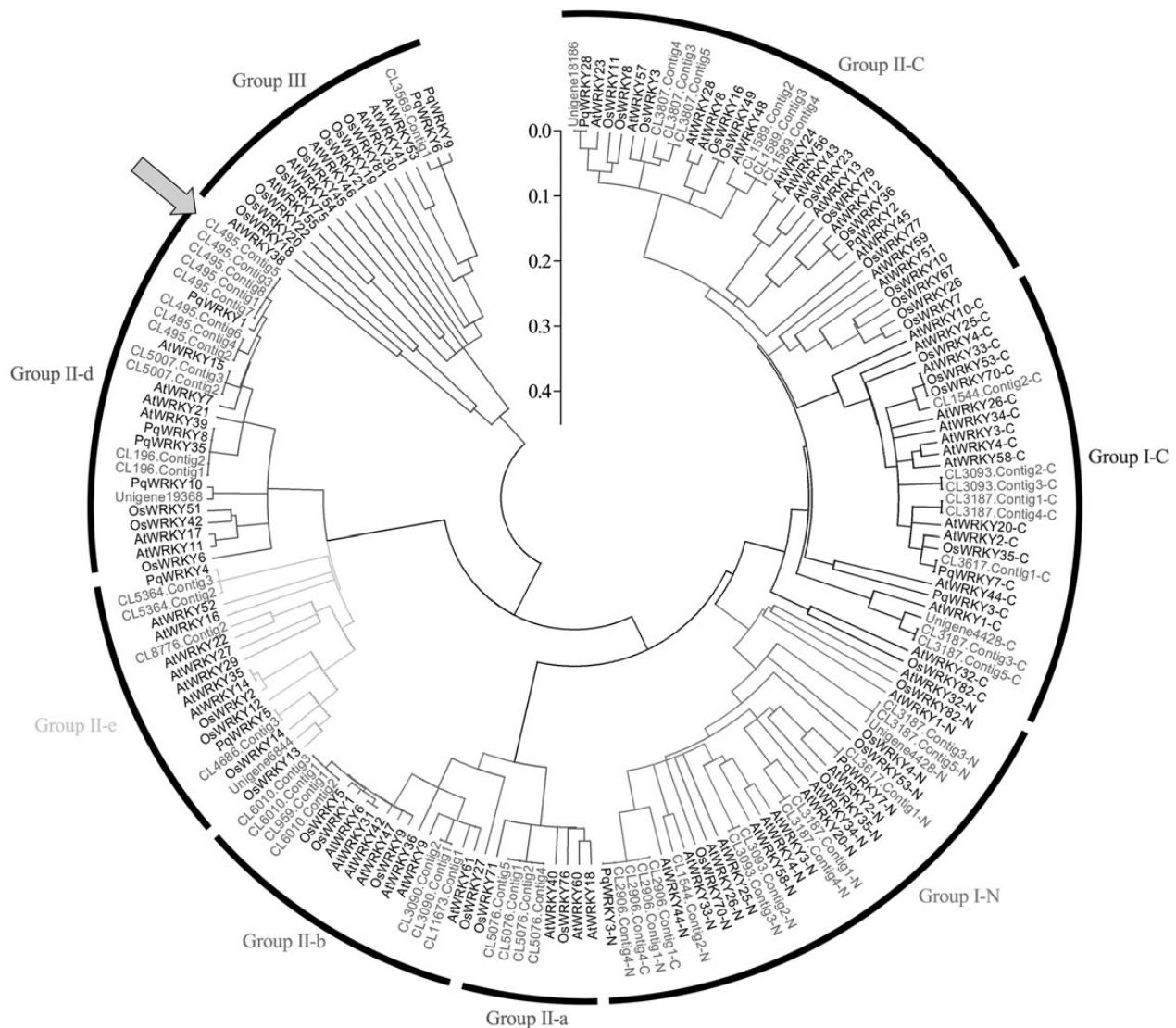
Results were presented as the mean  $\pm$  the standard deviation (SD) from three independent experiments, and each of those was calculated with three replicates. Statistical analysis was conducted using the analysis of variance. Comparisons of mean were carried out using the least significant difference test at  $P = 0.05$ . The difference was considered statistically significant when  $P < 0.05$ .

## Results

### Identification of *PgWRKY1* from the Illumina/Solexa dataset

A total of 71,095 unigenes from MeJA-treated adventitious roots of *P. ginseng* were identified via Illumina/Solexa platform sequencing

in our previous study [29]. Among them, 186 unigenes based on NCBI Nr, NCBI Nt, Swiss-prot, and KEGG databases were found to be annotated as putative WRKY genes. After excluding sequences that were wrong or incomplete ORFs, a total of 48 promising *PgWRKY* unigenes were identified, which contain nucleotide sequences between 1123 and 2543 bp with the average length of 1608 bp (Supplementary Table S2). These WRKY genes were translated into proteins through the MEGA 4.0 software [30] and the WRKY domains containing the WRKYGQK sequence motif and the zinc-finger motif ( $C_2H_2$  or  $C_2HC$ ) were further checked to avoid duplicate results. The 48 WRKY domains with complete N-terminal and C-terminal WRKY ends were then classified into three groups (Groups I, II, and III) using unrooted phylogenetic analysis (Fig. 1) with the already classified AtWRKYs from *Arabidopsis*, OsWRKYs from *O. sativa*, and PqWRKYs from *P. quinquefolius* [1,9,21] as references. Eleven *PgWRKY*s with two WRKY domains were assigned to Group I. Their zinc-finger structure is  $C_2H_2$  with a pattern of

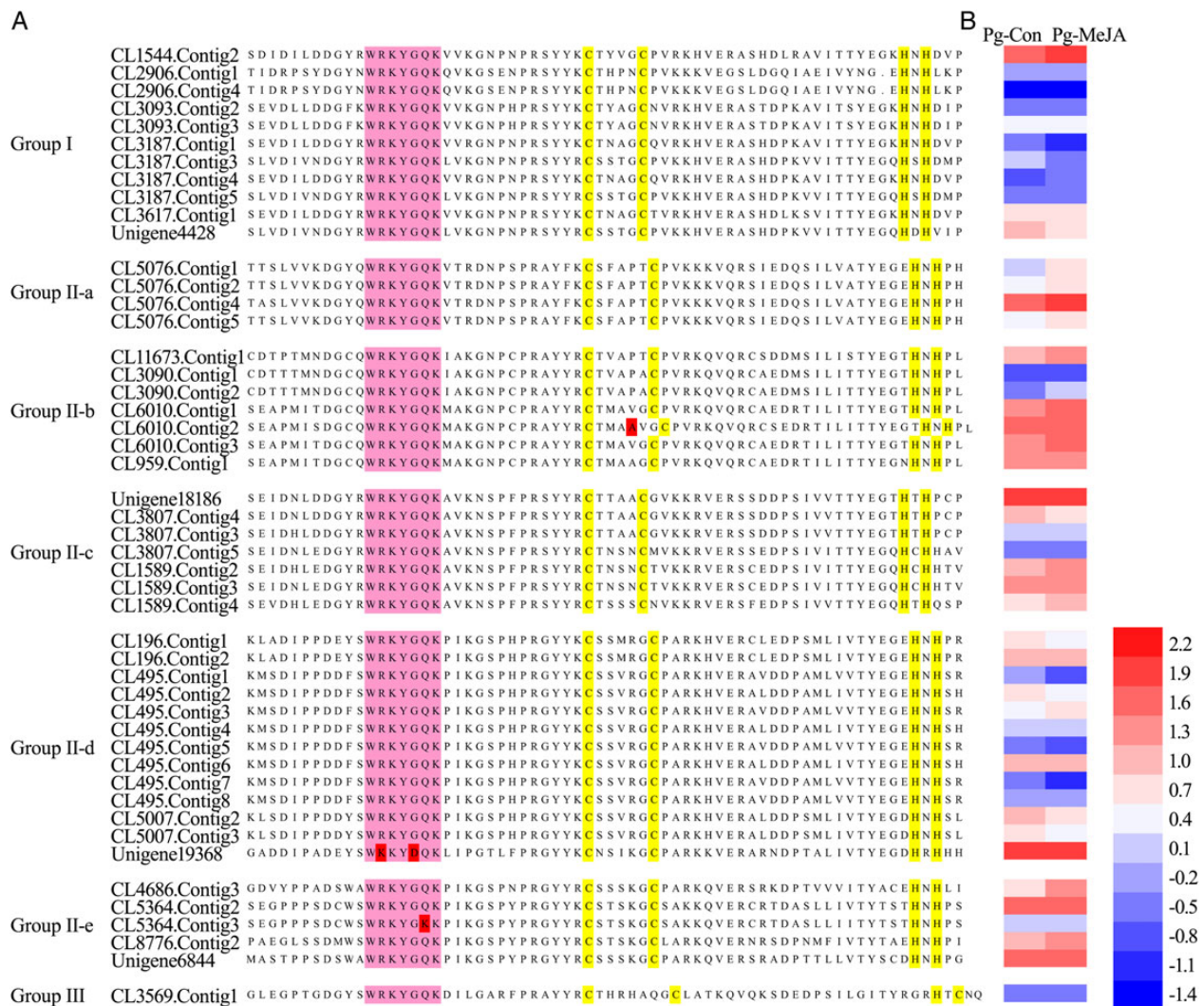


**Figure 1. Unrooted phylogenetic tree analysis of 48 *PgWRKY* domains in three groups** The *PgWRKY*s are highlighted in red and other plant WRKYs in black. Group I-N and Group I-C denote the N-terminal and C-terminal WRKY domain of Group I members, respectively. The alignment of amino acid sequences of the WRKY domains was performed with MEGA 4.0 software and the phylogenetic tree was based on *A. thaliana*, *O. sativa*, and *P. quinquefolius* WRKY classification constructed using NJ method. The arrow indicates the position of the *PgWRKY1* protein in the phylogenetic tree.

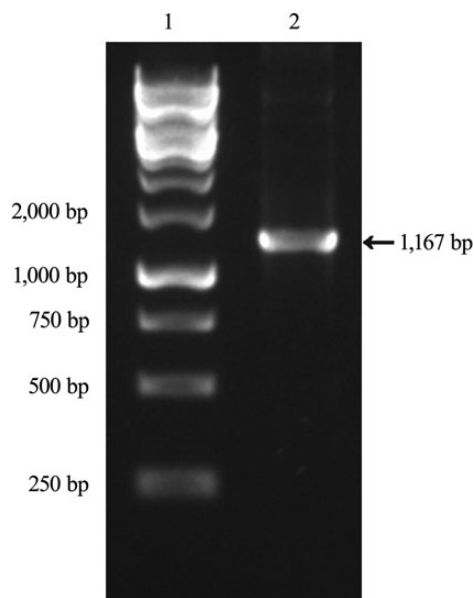
C-X<sub>5</sub>-C-X<sub>23</sub>-H-X-H. Since the N- and C-terminal domains produce distinct clusters, we designated them as I-N and I-C, respectively. Thirty-six *PgWRKYs* with single WRKY domain containing C<sub>2</sub>H<sub>2</sub> zinc-finger (C-X<sub>5</sub>-C-X<sub>23</sub>-H-X-H) were assigned to Group II. They were further assembled into five subgroups: Group II-a (4 proteins), Group II-b (7 proteins), Group II-c (7 proteins), Group II-d (13 proteins), and Group II-e (5 proteins). The remaining one protein with single WRKY domain was assigned into Group III and its zinc-finger structure was C<sub>2</sub>HC (C-X<sub>7</sub>-C-X<sub>23</sub>-H-C).

Moreover, the amino acid sequences of the 48 *PgWRKY* proteins were also aligned and only the regions containing the WRKY motif and the zinc-finger motif were used for comparison (Fig. 2A). In addition, although the two WRKY domains of Group I members were found to be functionally distinct, only the second WRKY domains near the C-terminal ends were compared as they were proved to have sequence-specific binding ability to the target DNA sequence

[2,33]. The sequences in the WRKY motif and the zinc-finger motif were found to be highly conserved with the exception of three variants including CL6010.Contig2 (Group II-b), Unigene 19368 (Group II-d), and CL5364.Contig3 (Group II-e). The ‘CL6010.Contig2’ WRKY protein showed a little variation in the sequences of zinc-finger motif compared with other members in the group. Moreover, most of the 48 *PgWRKY* proteins were identified to contain the conserved WRKYGKQ residues with the exception of two proteins with the alternative WRKY motif sequences: WKKYDQK (Unigene 19368) and WRKYGKK (CL5364. Contig3). These two proteins were grouped into the subgroup II-d and II-e, respectively. It is noteworthy that WKKYDQK sequence of ‘Unigene 19368’ has not previously been identified in any other species. This sequence variation within the conserved domain suggested that multiple WRKY genes should widely represent *P. ginseng* genome and may play a species-specific biological role in the growth and development of this species.



**Figure 2. Amino acid sequences of 48 *PgWRKY* domains along with expression levels of their corresponding unigenes** (A) Comparison of the amino acid sequences of WRKY domains. WRKY motifs within each of the groups are marked in pink and cysteine (C) and histidine (H) residues forming zinc-finger motif are highlighted in yellow. The amino acids that are variants are highlighted in red. Only the C-terminal WRKY domain for the members with two WRKY domains was used for comparison. (B) Heatmap diagram showing expression levels of *PgWRKY* unigenes. The levels of expression were determined based on comparison of the intensity of two colors (red and blue) from two samples (Pg-Con and Pg-MeJA). The more intensity of red is, the more expression occurs, while the phenomenon is vice versa for blue. The values representing intensity were measured using log<sub>10</sub>(FPKM). Log<sub>10</sub>(FPKM) values are presented with the ‘color-scale’ shown at the right-bottom. The expression map was generated using Heml program with log<sub>10</sub>(FPKM) values.



**Figure 3. Full-length cDNA of *PgWRKY1* generated from PCR amplification**  
Lane 1: marker (1 kb) and lane 2: the amplified products of the full-length cDNA (1167 bp, marked by arrowhead).

To get a full picture in terms of the MeJA-treated expression levels of the above-mentioned 48 *PgWRKY* candidates, a heatmap diagram (Fig. 2B) was generated on the basis of the  $\log_{10}$ (FPKM) values of Pg-Con (control) and Pg-MeJA (MeJA treated), which showed an average expression of each unigene. Among them, 39.58% (19 out of 48 unigenes) showed upregulated trends, 27.08% (13 unigenes) showed downregulated trends, and 33.33% (16 unigenes) did not show differential expression levels (Supplementary Table S2), implying their putative roles in secondary metabolism. It is worth mentioning that the unigene ‘CL495.Contig5’ selected for cloning in this study was found to be downregulated in our MeJA-treated Solexa/Illumina dataset.

### Cloning and phylogeny of *PgWRKY1*

Based on the current data, one annotated unigene (CL495.Contig5) corresponding to MeJA in our Illumina/Solexa-based expression dataset was randomly selected. The primers were then designed, and the selected WRKY candidate named *PgWRKY1* (GenBank accession number KR060074) was successfully cloned. The full-length cDNA of the *PgWRKY1* is 1167 bp with 5'-UTR (35 bp) and 3'-UTR (55 bp) sequences (Figs. 3 and 4). The ORF encodes a protein with 358 amino acid residues with a calculated MW of 39.68 kDa and a PI of 9.54.

Comparison with the sequence alignment of other WRKY group II members from different plants, the *PgWRKY1* protein was found to have a single WRKY domain consisting of one WRKYGQK sequence motif located at the N-terminal and one C<sub>2</sub>H<sub>2</sub> (C-X<sub>5</sub>-C-X<sub>23</sub>-H-X-H) type zinc-finger motif at the C-terminal (Fig. 5), indicating that it is a subgroup II-d protein according to the classification of the WRKY superfamily [9,31]. According to phylogenetic tree, *PgWRKY1* was found to be grouped closely with *PqWRKY1*. Although *PqWRKY1* was assigned to subgroup II-c [21], it was really found to be included within the II-d subgroup when compared with the entire *Arabidopsis* and *O. sativa* WRKY family (Fig. 1). In addition, the rooted phylogenetic analysis (Fig. 6) showed that WRKY proteins along with

*PgWRKY1* formed one branch in the tree, even if they were derived from different plant species. In particular, *PqWRKY1*, *PcWRKY3*, *IbWRKY1*, *NbWRKY17*, and *StWRKY12* showed very high similarity to *PgWRKY1*, with 98.60%, 59.84%, 57.41%, 62.80% and 67.12% similarity, respectively. All these observations indicated that *PgWRKY1* is a member of WRKY gene family.

### Structure analysis of *PgWRKY1* protein

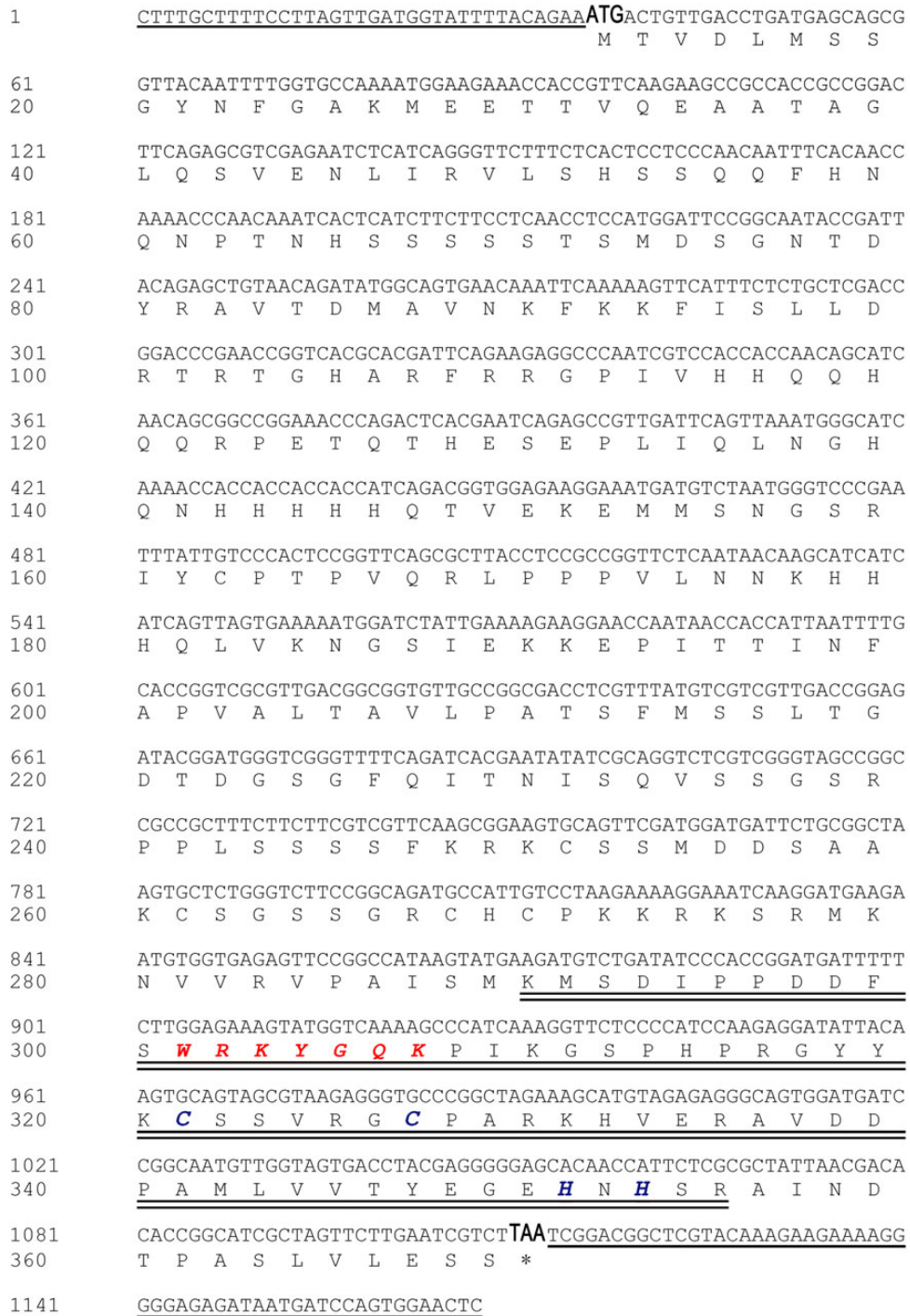
The single WRKY-DBD of *PgWRKY1* was detected via NCBI Conserved Domains Finder (CDF) database (Fig. 7A). The instability index (II) of the protein was determined to be 47.36, indicating that it is an unstable (stability factor >40) protein. The protein GRAVY value was computed to be -0.737, implying that the *PgWRKY1* was a hydrophilic protein in nature. GRAVY values generally range between  $\pm 2$ , where the positive values show the hydrophobic nature of proteins and the negative values indicate their hydrophilic nature. Transmembrane helices are necessary for membrane-related proteins associated with cell signaling and substrate transport [34,35]. Transmembrane prediction with TMHMM server V.2.0 showed that such transmembrane helix was not found in *PgWRKY1* protein, indicating that it is not a transmembrane protein (Fig. 7B). The predicted secondary structure of *PgWRKY1* protein is shown in Fig. 7C. It has 16.48% (59 amino acid residues) alpha helix (Hh), 18.44% (66 amino acid residues) extended strand (Ee), and 65.08% (233 amino acid residues) random coil (Cc). Based on amino acid sequence (Fig. 8A), the predicted three-dimensional structure of *PgWRKY1* protein using Phyre2 [37] and Swiss-Model software [38] was shown in Fig. 8B-E and F,G, respectively. The results of *PgWRKY1* protein function prediction with ProtFun procedure [39] listed in Table 2 showed that the key function of *PgWRKY1* was translation, and its gene ontology category indicated transcription regulation.

### Subcellular localization of *PgWRKY1* protein

All the discovered WRKY proteins in various plant species have been shown to be present in the nucleus [6,22]. To demonstrate that *PgWRKY1* is indeed nucleus-localized, 35S::*PgWRKY1*:GFP fusion and the 35S::GFP control gene were injected into onion (*Allium cepa*) and tobacco (*Nicotiana tabacum*) epidermal cells by *Agrobacterium rhizogenes* injector, and then cells were observed under a fluorescent microscope. Both the *PgWRKY1*:GFP and GFP genes were expressed at a high level in the injected cells under the control of highly powerful cauliflower mosaic virus (CaMV) 35S promoter. Fluorescent microscope imaging showed that *PgWRKY1*:GFP fusion protein was prominently present in the nucleus of onion and tobacco epidermal cells in a transient expression assay (Fig. 9). In contrast, the GFP control protein alone showed GFP signal both in the entire cytoplasm and in the nucleus. These results suggested that the *PgWRKY1* protein is a nuclear-localized protein that is consistent with its predicted function as a TF.

### Expression pattern of *PgWRKY1*

More evidence showed that plant WRKY genes have various expression patterns in different plant organs. For example, *OsWRKY82* was found to be expressed differentially in all investigated tissues, including stems, leaves, flowers, and grains [40]. *CtWRKY1* in *Catharanthus roseus* was shown to be preferentially expressed in roots, indicating that it is a key factor in determining the root-specific accumulation of serpentine in *C. roseus* plants [28]. In this study, the expression pattern of the *PgWRKY1* gene in *P. ginseng* tissue was investigated using qRT-PCR analysis. As shown in Fig. 10, *PgWRKY1* was expressed

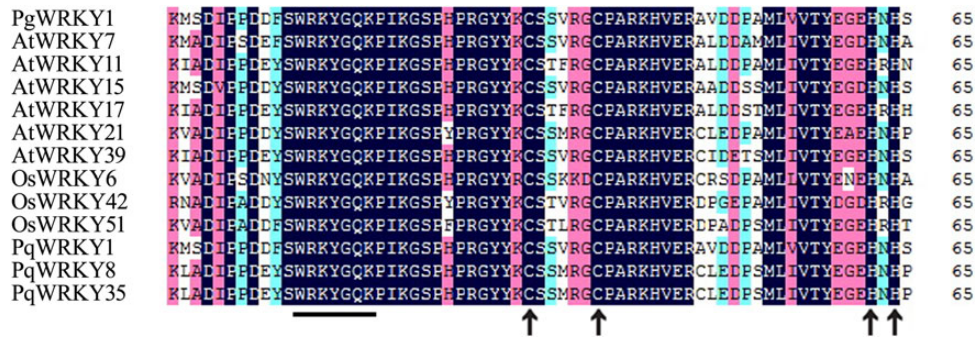


**Figure 4. Alignment of nucleotide and amino acid sequences of *PgWRKY1*** The nucleotides are shown in the upper line and the amino acid residues are presented below. Bold letters in the upper line indicate translation initiation (ATG) and stop (TAA) codon. 5' and 3' UTRs are singly and the WRKY domain is doubly underlined. The WRKY motif is marked in red and cysteine (C) and histidine (H) residues in the putative zinc-finger motif are marked in blue.

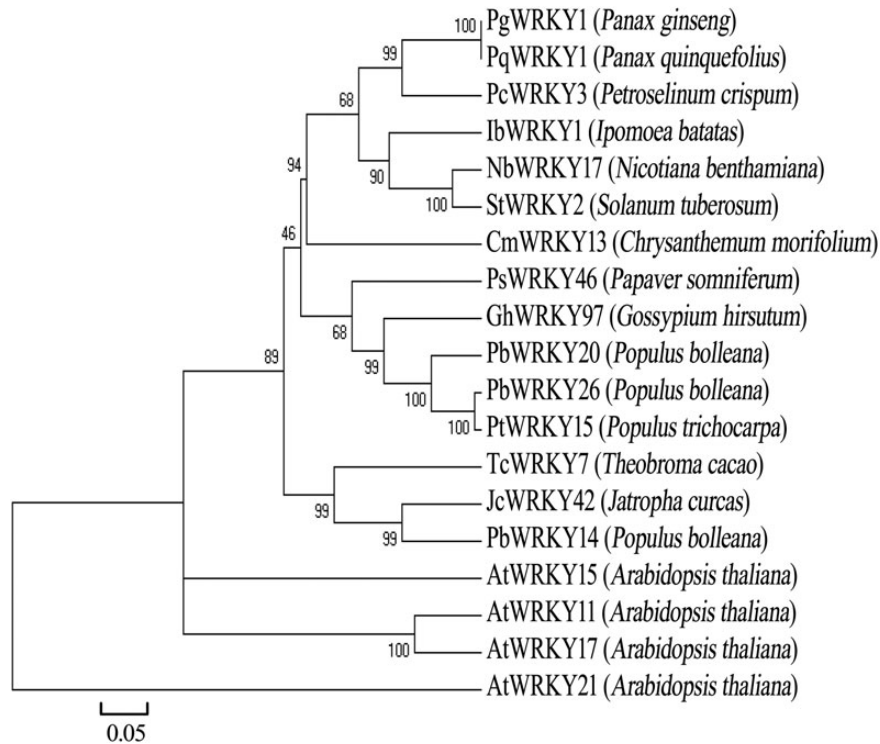
differentially in leaves, roots, lateral roots, stems, and seeds. The expression of *PgWRKY1* was high in lateral roots (5.07-fold) and roots (4.53-fold) compared with that in leaves (1-fold), indicating an increased accumulation of *PgWRKY1* mRNA in *P. ginseng* roots. Likewise, a relative decrease of *PgWRKY1* transcripts was observed in stems (0.97-fold) and seeds (0.55-fold) compared with that in leaves. These results indicated a constitutive basal expression of *PgWRKY1* in *P. ginseng* in a tissue- and development-dependent manner.

**Expression profiles of *PgWRKY1* in response to hormones and abiotic stress**

The WRKY genes have been demonstrated to respond to MeJA, ABA, SA, and salt treatments associated with many life processes, including stress resistance and natural products regulation [21,41,42]. Therefore, to identify the possible involvement of *PgWRKY1* gene in hormone-dependent defense signaling pathways and in improving stress tolerance, its expression profiles at different time points after



**Figure 5. WRKY domain comparison of *PgWRKY1* with other members of WRKY group II** Alignment analysis of *PgWRKY1* with *A. thaliana* AtWRKY7 (AAK28440.1), AtWRKY11 (AAK96194.1), AtWRKY15 (AEC07442.1), AtWRKY17 (AAL13049.1), AtWRKY21 (AEC08412.1), and AtWRKY39 (AAK96198.1), *O. sativa* OsWRKY6 (AAQ20906.1), OsWRKY42 (AAT84154.1), and OsWRKY51 (DAA05115.1), and *P. quinquefolius* PqWRKY1 (AEQ29014.1), PqWRKY8 (AEQ29021.1), and PqWRKY35 (AEQ29025.1) performed using DNAMAN software. The fully conserved WRKY motif (WRKYGQK) is underlined and cysteines (C) and histidines (H) in the zinc-finger motif are indicated by arrowheads. Identical amino acids are shaded in deep blue and similar residues are in other colors.



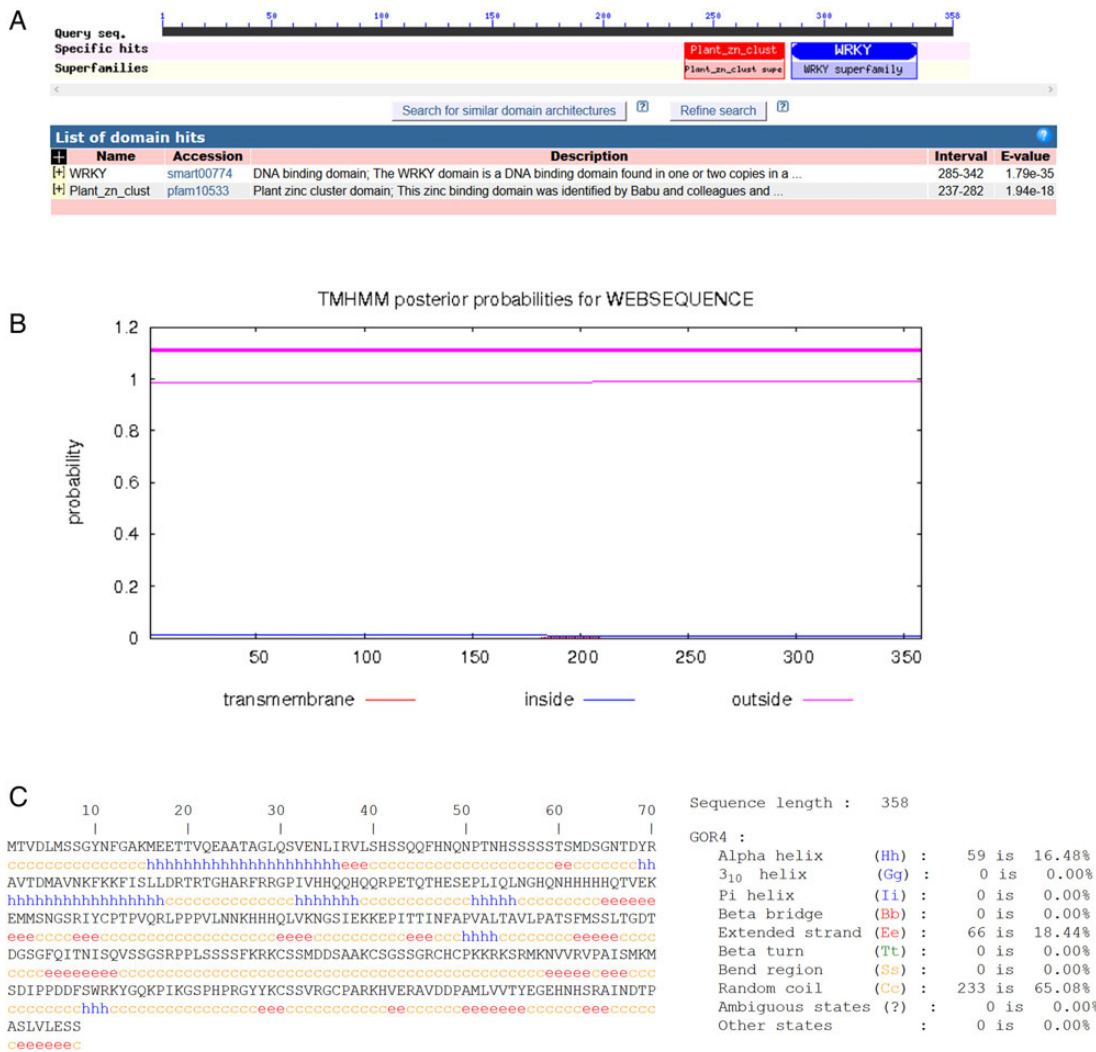
**Figure 6. A neighbor-joining phylogenetic tree of *PgWRKY1* protein with other closely related WRKY proteins** The other WRKY amino acid sequences including their accession number used to prepare the phylogenetic tree are as follows: PqWRKY1 (AEQ29014.1), PcWRKY3 (AAC49528.1), IbWRKY1 (ACT55331.1), NbWRKY17 (AIR74899.1), StWRKY2 (NP\_001275001.1), CmWRKY13 (AHC54618.1), PsWRKY46 (AFU81789.1), GhWRKY97 (AGV75972.1), PbWRKY20 (ACV92022.1), PbWRKY26 (ACV92028.1), PtWRKY15 (XP\_002310122.1), TcWRKY7 (XP\_007017155.1), JcWRKY42 (AGQ04232.1), PbWRKY14 (ACV92016.1), AtWRKY15 (NP\_179913.1), AtWRKY11 (NP\_849559.1), AtWRKY17 (NP\_565574.1), and AtWRKY21 (NP\_565703.1). Bootstrap values calculated for 100 replicates are shown on the nodes. The name of the respective plant species of the above WRKY proteins are indicated in the phylogenetic tree.

MeJA, SA, ABA, and NaCl treatments of *P. ginseng* adventitious roots were monitored.

MeJA is an important and potent elicitor of specialized metabolites formation including ginsenosides in *P. ginseng* and other medicinal plants [28,43]. MeJA treatment was demonstrated to regulate the expression of *AtWRKY38* and *AtWRKY70* in *A. thaliana* in a time-dependent manner [44,45]. Moreover, expression analysis of *P. quinquefolius* treated with MeJA at different time points showed that *PqWRKY1* was induced within 24 h [21]. The candidate ‘CL495.

Contig5’ selected for cloning (*PgWRKY1*) in this study was found to be downregulated in our MeJA-treated Solexa/Illumina dataset (Fig. 2B). Based on the analysis of the full-length *PgWRKY1* protein, it was concluded that the protein structure of *PgWRKY1* is closely related to the structure of *PqWRKY1* with 98.60% similarity. Therefore, the expression profiles of *PgWRKY1* in the adventitious roots after treatment with MeJA were analyzed at different time points. Results showed that MeJA did not induce *PgWRKY1* expression (Fig. 11A), rather it significantly reduced *PgWRKY1* expression,



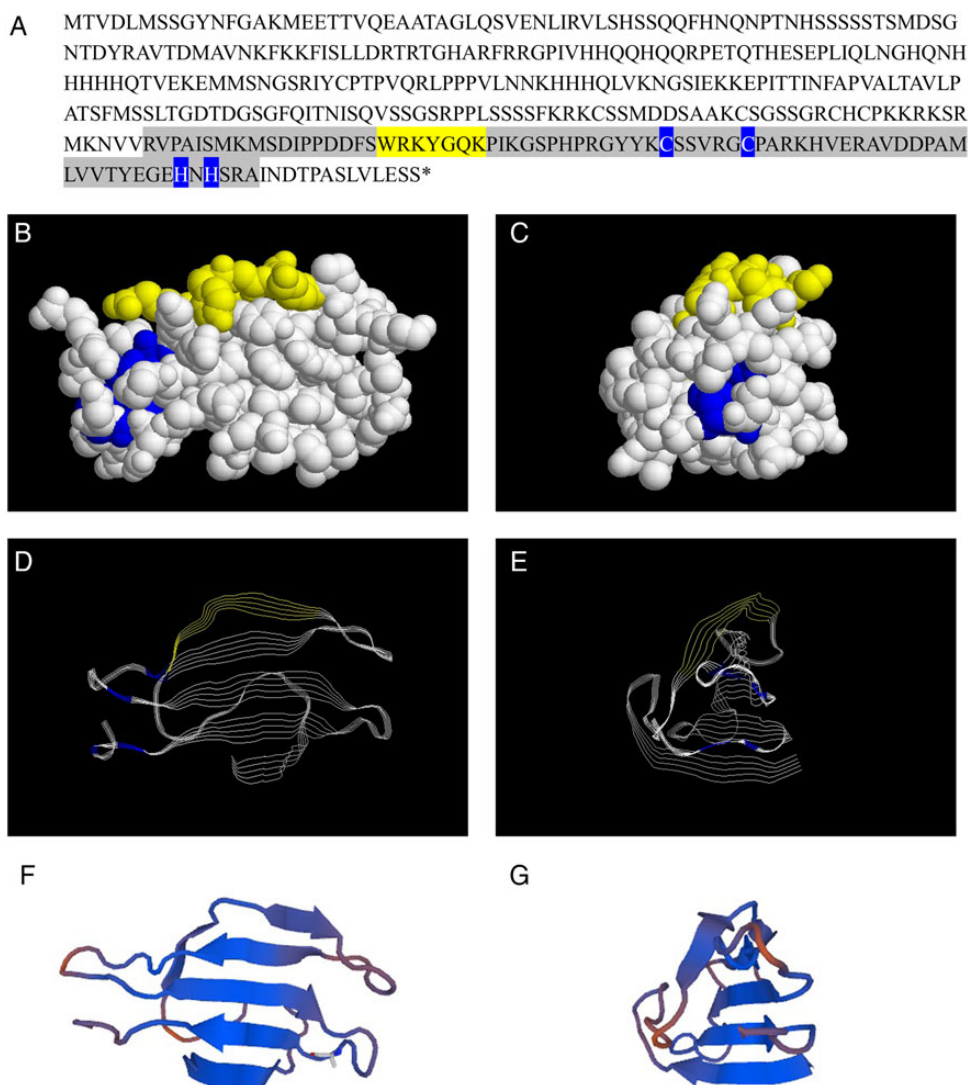


**Figure 7. Results of conserved domain analysis, transmembrane helices, and secondary structure of *PgWRKY1* protein** (A) Conserved domains analysis indicating one conserved WRKY domain. (B) Transmembrane helices prediction via TMHMM server (version 2.0). Red bars show transmembrane, blue lines indicate inside, and magenta lines indicate outside. (C) The predicted secondary structure. Protein sequence is presented at the left side and the secondary structure is shown at the right side. Alpha helix (Hh) is marked in blue, extended strand (Ee) in red, and random coil (Cc) in yellow.

which validates our MeJA-treated transcriptome expression data, but differs from the *PqWRKY1* expression data. *PqWRKY1* induction has been identified in American ginseng leaves, but not in adventitious roots that may respond to MeJA differently [46]. Our qRT-PCR analysis (Fig. 11A) showed a remarkable decrease in the transcripts of *PgWRKY1* gene at the first 3 h of MeJA treatment. The decreased transcripts were observed at all time points of treatment and the expression level reached the lowest level at 48 h post-treatment compared with the control. The statistical analysis indicated that the change in the expression level of *PgWRKY1* was extremely significant ( $P < 0.01$ ). Previous reports also showed that *CrWRKY18*, *CrWRKY21*, *CrWRKY41*, *CrWRKY45*, and *CrWRKY48* in *C. roseus* were all significantly downregulated at all time points after MeJA treatment. *CrWRKY8* and *CrWRKY5* were upregulated after 1 h of MeJA treatment and downregulated thereafter [47], and this phenomenon was without exception for *AaWRKY1* in *A. annua* leaves [20].

In model plants, WRKY TF family is well recognized in SA and defense signaling pathways. The major part of *Arabidopsis* WRKY TFs are induced upon SA treatment [48,49]. Previously, MeJA-regulated

*WRKY* genes were found to be regulated by SA as well [42,50]. To support this idea, the expression analyses of *PgWRKY1* in response to SA were investigated. After this treatment (Fig. 11B), the *PgWRKY1* expression was induced by SA, but in different patterns. The expression of *PgWRKY1* was decreased at 3 h and further down-regulated at 6 h after treatment. However, the expression increased from 12 h and reached the highest level (~2.69-fold) at 36 h post-treatment compared with the control. Finally, the transcripts of *PgWRKY1* were remarkably decreased during the next 48 h after treatment ( $P < 0.01$ ). The SA induction pattern of *PgWRKY1* is similar to that of WRKYs from other plants. For instance, SA treatment exhibited the best induction of *OsWRKY71* in rice leaves at 30 min and significantly decreased thereafter [51]. In some cases, SA also generated responses in abiotic stresses. The *MusaWRKY71* gene of *Musa* spp. cv. *Karibale Monthan* was shown to be induced rapidly in both roots and leaves in response to cold, salt, and dehydration stresses [22]. Together with these reports, our results indicated an interaction between MeJA and SA signaling pathways and signified the importance of the *PgWRKY1* gene in plant defense responses.



**Figure 8. Different views of 3D structural model of the *PgWRKY1* protein** (A) Amino acid sequence of *PgWRKY1*. Grey (WRKY domain), yellow (WRKY motif), and blue (zinc-finger motif) marked portion were used in 3D structure. (B,C) Front view and side view of spacefill model, respectively, prepared by Phyre2 software. (D,E) Front view and side view of strands model, respectively, designed by Phyre2 software. (F,G) Front view and side view of cartoons model, respectively, produced by Swiss-Model software. The models were produced on the basis of the solution structure [36].

ABA is a stress hormone that also mediates plant responses to abiotic stresses, including high salinity, drought, and low temperature. Evidence showed that MeJA positively regulates *AtWRKY18* and *AtWRKY40*, which negatively regulate ABA response [52]. *LtWRKY21* transcripts in vegetative tissues of *Larrea tridentate* is induced upon ABA treatment and in response to various environment stresses such as water deficit, wounding, drought, and salt stresses [53]. The transcripts of *GbWRKY 3/14/20/24/33/34* in cotton and *OsWRKY24/51/71/72* in rice were previously reported to be upregulated in roots by ABA treatment [54,55]. In this study, *PgWRKY1* expression was found to be induced rapidly and robustly at the beginning of treatment with ABA (Fig. 11C). The transcript level of *PgWRKY1* increased rapidly and peaked within 3 h (6.68-fold more than the control level). Then, the expression decreased ~558% at 6 h and 201% at 12 h post-treatment. The transcript level of *PgWRKY1* increased slightly at 24 h and began to decline again and returned to the control level within 36 h. Expressions of *WRKY* genes that may contribute to secondary metabolites regulation

are predicted to be altered before early and mid-stages of the metabolites biosynthesis pathway [47].

Salinity stress in soil or water is one of the most serious abiotic stresses especially in arid and semi-arid regions that can severely limit plant growth and crop yield, as well as affect the gene expression in plants [56]. In addition, the salinity stresses are closely associated with the ABA signaling network. Besides various hormones, salt was also implicated in regulating *WRKY* gene expression. Both ABA and NaCl have previously been reported to regulate the expression of *MusaWRKY71* in cotton and *OsWRKYs* in rice [22,42]. Taking this view into account, *PgWRKY1* expression analysis was carried out upon NaCl treatment. Results showed that the induction pattern was almost the same as that of the *MusaWRKY71* and *OsWRKY* genes expression. After this treatment (Fig. 11D), *PgWRKY1* expression began to rise immediately at 3 h (2.34-fold) and maintained at this level up to 6 h. Then, the *PgWRKY1* expression level continually declined and this trend lasted up to 48 h post-treatment until the

**Table 2. Prediction of *PgWRKY1* protein function**

Functional category	Prob	Odds	Gene ontology category	Prob	Odds
Amino acid biosynthesis	0.011	0.500	Signal transducer	0.272	1.271
Co-factors biosynthesis	0.210	2.917	Receptor	0.008	0.047
Cell envelope	0.033	0.541	Hormone	0.114	17.538
Cellular processes	0.030	0.411	Structural Protein	0.003	0.107
Central intermediary metabolism	0.048	0.762	Transporter	0.043	0.394
Energy metabolism	0.035	–	Ion Channel	0.079	1.386
Fatty acid metabolism	0.017	1.308	Voltage-gated ion channel	0.003	0.136
Proteins and pyrimidines	0.331	1.362	Cation channel	0.010	0.217
Regulatory functions	0.034	0.211	Transcription	0.031	0.242
Replication and transcription	0.020	0.075	–	–	–
Translation	≥0.071	1.614	Transcription regulation	≥0.031	0.248
Translation and binding	0.773	1.885	Stress response	0.223	2.534
			Immune response	0.123	1.447
			Growth factor	0.040	2.857
			Metal ion transport	0.026	0.056

original level was reached. The statistical analysis showed that the expression change was extremely significant ( $P < 0.01$ ) upon NaCl treatment. Based on these results, *PgWRKY1* gene can be presumed as part of the complex signal networks regulating salt stress response in *P. ginseng*.

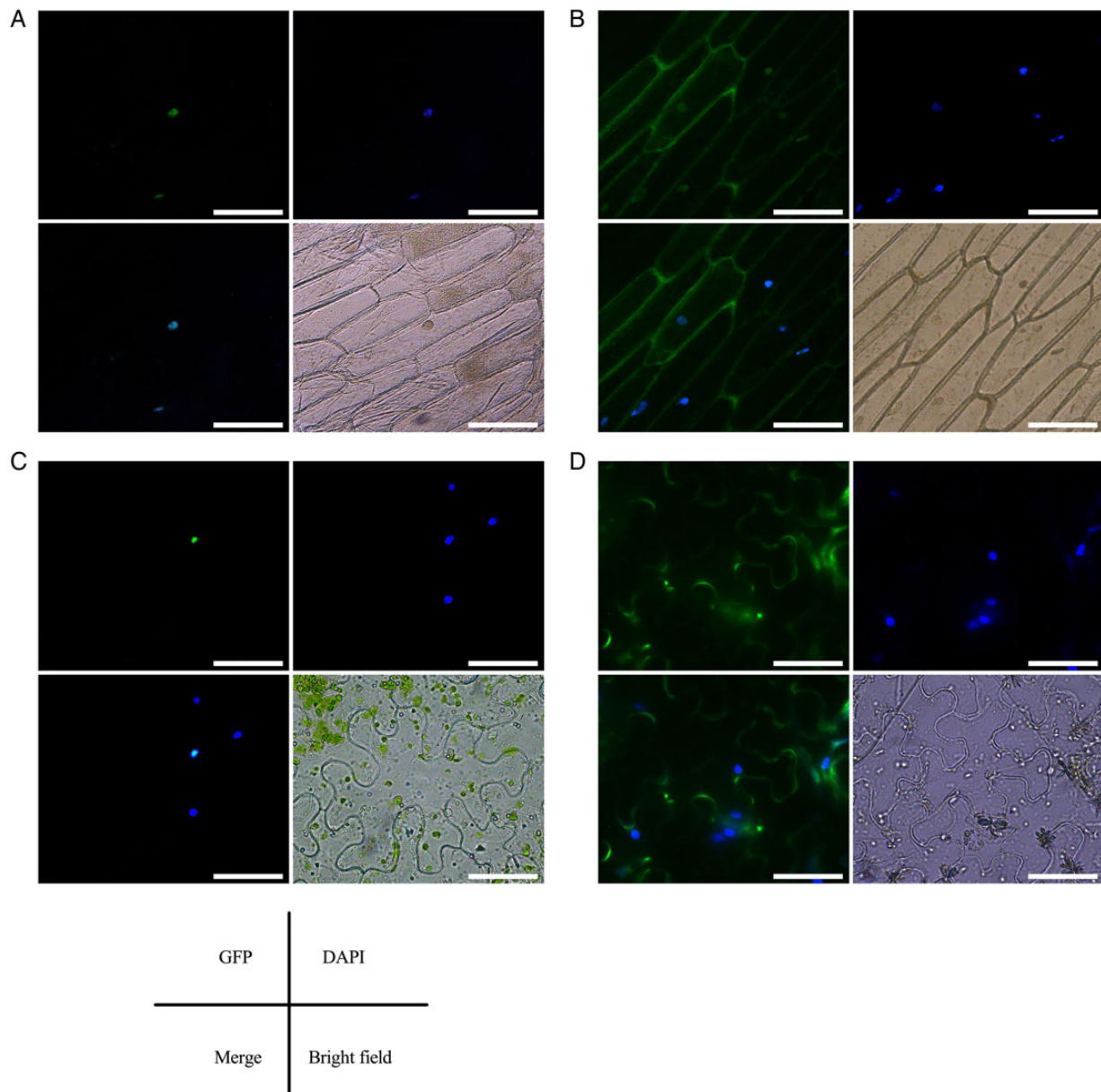
## Discussion

Based on the different DBDs, the TFs have been categorized into different families or superfamilies, where the WRKY proteins are plant-specific proteins involved in multiple functions of plant growth and development [1,5,10]. In the past, however, a large number of WRKY members and their novel functions and significances have been identified mostly in model plants like *A. thaliana* and *O. sativa* [9,10,57]. Moreover, in recent years, many WRKY genes have been isolated and functionally characterized in medicinal plants such as *Coptis japonica*, *C. roseus*, and *A. annua* [19,20,28]. The *PgWRKY1* from *P. quinquefolius* represents the first discovery of WRKY TF in *Panax* species [21]. In spite of pharmaceutical and industrial importance associated with *P. ginseng* since ancient times, no study has been carried out in the characterization and expression of any WRKY in *P. ginseng*. Moreover, public databases for WRKY gene discovery are not currently available for this species with the exception of a few expressed sequence tags using the traditional Sanger sequencing method. In addition, the large-sized tetraploid ( $2n = 4x = 48$ ) and highly complex genome of ~3.2 Gb of this species [58] are big challenges for obtaining the whole genomic sequence using this traditional method. As a result, the lack of genomic knowledge complicates the analysis of WRKY gene family in *P. ginseng*, in particular the gene structures and expression profiles. The completion of our transcriptome sequence on this species based on the NGS technology [29] has enabled comparative genomic studies as well as the identification of new *PgWRKY* genes.

In this study, our Illumina/Solexa dataset on MeJA-treated *P. ginseng* transcriptome was comprehensively searched for WRKY unigenes. A total of 492 unigenes encoding TF proteins were identified, which represent three major TF families, namely WRKY, MYB, and MYC. The WRKY family represents a large family in *P. ginseng* with high number of unigenes (186), indicating that WRKYs might have essential functions in many life processes of this species. Therefore, by further analysis, a total of 48 promising transcripts encoding *PgWRKY* proteins along with their expression levels upon MeJA

treatment were identified (Fig. 2B and Supplementary Table S2). Our results indicated that the *P. ginseng* WRKY family is similar in size to *Cucumis sativus*, *Fragaria vesca*, *Jatropha curcas*, *Carica papaya*, and *C. roseus* with 55, 56, 58, 66 and 52 WRKYs, respectively [7,47,59–61]. *Panax quinquefolius* is closely related to *P. ginseng* and also produces pharmaceutically valuable ginsenosides. Forty-five WRKY transcripts (16 Group I, 24 Group II, and 5 Group III) were found in *P. quinquefolius* [21], close to the 48 WRKY transcripts (11 Group I, 36 Group II, and 1 Group III) identified in *P. ginseng*. Using unrooted phylogenetic analysis of the WRKY families in *P. ginseng*, *A. thaliana*, and other plant species, these *PgWRKY* genes were assigned into three groups (Fig. 1), which facilitated the identification of specific subgroups and expansion process of the WRKY genes as well. Group III WRKY TFs are believed to have dramatically expanded during the evolution of angiosperms and can be classified into different subgroups depending on the species [62,49]. *Panax quinquefolius* contains five group III WRKY TFs that are in subgroup III-a [21]. Only one group III WRKY TF was identified in *P. ginseng*. Proportionally, the number of group III *PgWRKY* TFs is smaller than that of *P. quinquefolius*, suggesting that this group has not undergone significant expansion as occurred in American ginseng, rice, and *Arabidopsis* [21,62]. The aligned amino acid sequences of 48 *PgWRKYs* revealed that WRKYGQK sequence motif across the WRKY domains is conserved with the exception of ‘Unigene 19368 (WKKYDQK)’ and ‘CL5364.Contig3 (WRKYGKK)’ (Fig. 2A). More evidence indicated that several plants hold variants of the highly conserved WRKYGQK motif including WRKYGEK, WRKYGKK, WRKYGSK, and WRKYGRK [1,47]. Variation in this motif can reduce, alter, or eliminate DNA-binding activity of WRKY TFs. Previously, mutagenesis of the conserved glutamine was shown to reduce but not eliminate DNA-binding [63]. Recently, AtWRKY50 was reported to bind to the GAC core of the W-box with less preference for 5′ or 3′ bases [64]. Therefore, both the WKKYDQK and WRKYGKK variants identified in this study are anticipated to bind with DNA. These results will be a potential resource to provide useful *PgWRKY* candidates in support of genetic improvement of medicinal traits and stress tolerance in *P. ginseng* and probably other ginseng species.

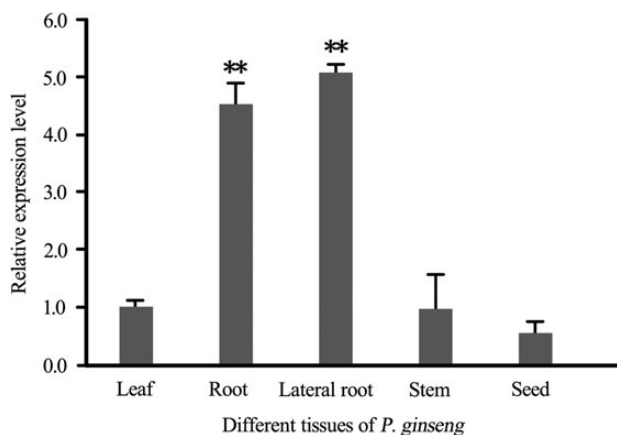
Importantly, the *PgWRKY1* (CL495.Contig5) containing the complete coding sequence (CDS) and responding to MeJA in the Illumina/Solexa dataset was selected for cloning, which provides more insights into the expression patterns under SA, ABA, NaCl, or MeJA



**Figure 9. Subcellular localization of PgWRKY1 protein** The *35S::PgWRKY1::GFP* gene and *35S::GFP* control gene were introduced into onion and tobacco epidermal cells by *Agrobacterium rhizogenes* injector. (A) Fluorescence image of *35S::PgWRKY1::GFP* fusion protein from onion epidermal cells. Scale bar = 200  $\mu$ m. (B) Fluorescence image of *35S::GFP* control construct from onion epidermal cells. Scale bar = 200  $\mu$ m. (C) Fluorescence image of *35S::PgWRKY1::GFP* fusion protein from tobacco epidermal cells. Scale bar = 50  $\mu$ m. (D) Fluorescence image of *35S::GFP* control construct from tobacco epidermal cells. Scale bar = 50  $\mu$ m. GFP, DAPI, bright field, and merge (GFP and DAPI) in each fluorescence image shown as clock wise motion are demonstrated in the 'plus icon' at the bottom.

treatments. Recently, the WRKY domains thought to be functional units of WRKY proteins were shown to play an important role in the regulation of various physiological processes of wheat in response to various environmental stimuli [65]. Moreover, analysis on the 3D structure of the C-terminal domain of AtWRKY4 has shown that the conserved WRKYGQK sequences are strongly and directly involved in DNA binding [36] and the substitution of each of the amino acids from WRKYGQK residues with alanine decreased the DNA-binding activity [66]. Thus, the PgWRKY1 protein sequence possesses one conserved WRKY domain which includes a fully established invariant WRKY motif (WRKYGQK) and a zinc-finger motif  $C_2H_2$  in the center of polypeptide (Fig. 4), indicating that PgWRKY1 might have high

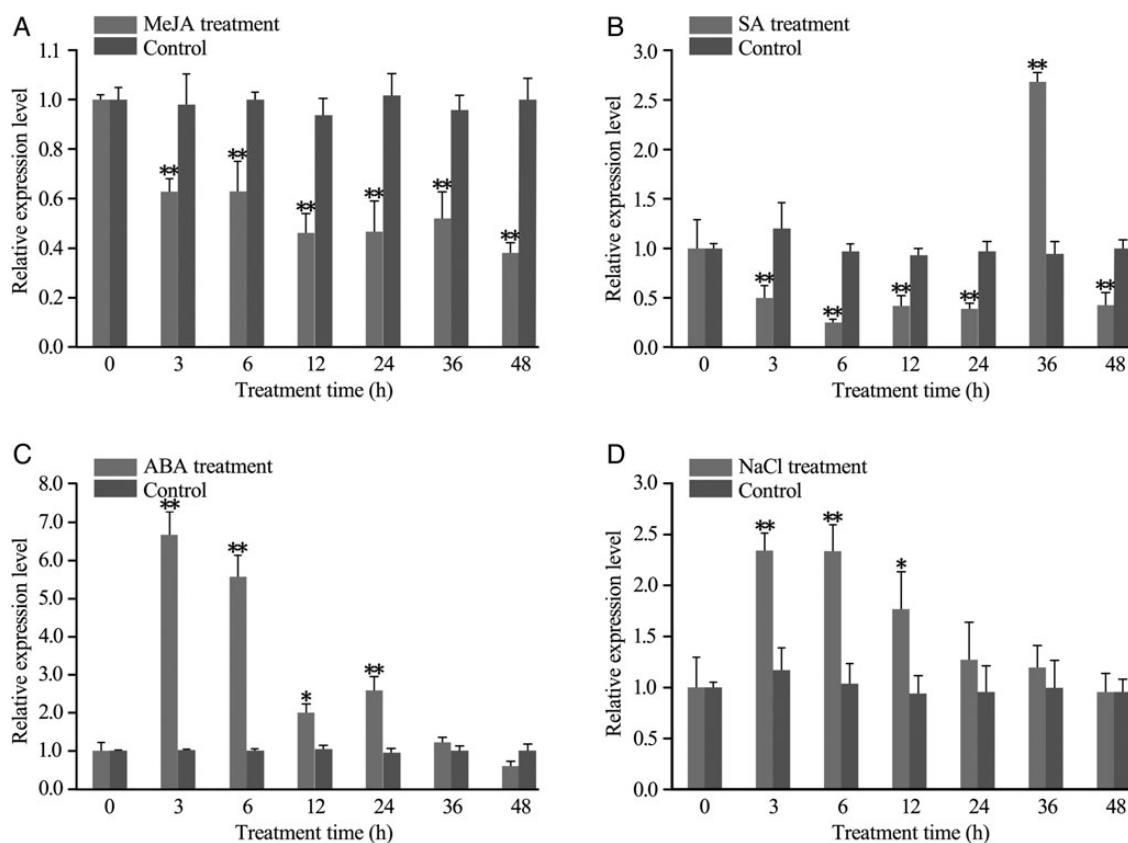
DNA-binding activity like other functional plant WRKY TFs. Furthermore, domain prediction of the deduced full-length protein also clearly showed that PgWRKY1 contains one conserved WRKY-DBD (Fig. 7A). To clarify the relationships among the WRKY-DBD regions, we performed multiple alignment analysis of WRKY subgroup II-d protein members from *A. thaliana*, *O. sativa*, and *P. quinquefolius* (Fig. 5). The high sequence identities indicate that the amino acid residues are highly conserved among WRKY proteins. Our phylogenetic analysis (Fig. 6) showed that in the WRKY family, PgWRKY1, PqWRKY1, and PcWRKY3 are highly homologous and show highly similar features, indicating a possible functional similarity among them.



**Figure 10. Tissue-specific expression of *PgWRKY1* in different organs of *P. ginseng* analyzed by qRT-PCR** Error bars indicate standard deviation (SD) of means from three independent experiments. The relative expression was calculated using the  $2^{-\Delta\Delta t}$  method based on the corresponding gene expression in leaf. The differences between the transcript level of lateral roots, roots, leaves, stems, and seeds are statistically significant (\* $P < 0.05$ , \*\* $P < 0.01$ ) according to independent samples *t*-test.

WRKY TFs have previously been reported to be a nucleus-targeted protein and tissue-specific [28,33]. Similar to those plant WRKY factors, *PgWRKY1* was also found to be localized in the nucleus (Fig. 9). Expression analysis of *PgWRKY1* in various tissues (Fig. 10) under normal condition reveals that it is preferentially expressed in lateral roots and roots compared with leaves, stems, and seeds, indicating its possible importance for high valuable roots of *P. ginseng*. Moreover, the wide expressions of *PgWRKY1* in various organs suggested that it may play regulatory roles in multiple developmental stages of *P. ginseng* tissues.

Gene expression analysis is essential for understanding the possible biological functions of a gene. Hormones have been proved to affect the expression of many genes and regulate a wide range of developmental processes, as well as responses of plants to stresses. Several WRKY genes from a number of plant species have been confirmed to be upregulated by one or more phytohormones, including SA, MeJA, ABA, and abiotic stress NaCl [40,55]. About 109 WRKY proteins have been identified in rice [42,51], which mostly were involved in responses to various hormones, biotic, and abiotic stresses. Besides, the recently discovered *PgWRKY1* in American ginseng was shown to be up-regulated rapidly during MeJA treatments [21]. Here, we tried to demonstrate whether the expression profiles of *PgWRKY1* are related



**Figure 11. Expression profiles of *PgWRKY1* in *P. ginseng* adventitious roots measured by qRT-PCR** (A) Treatment with 200  $\mu$ M MeJA. (B) Treatment with 200  $\mu$ M SA. (C) Treatment with 20  $\mu$ M ABA. (D) Treatment with 50  $\mu$ M NaCl. The adventitious roots treated with equal amount solvent of the corresponding chemicals were used as a control (0, 3, 6, 12, 24, 36, and 48 h). The levels of *PgWRKY1* transcript at different time points (0, 3, 6, 12, 24, 36, and 48 h) were analyzed by qRT-PCR. The gene expression level in the treated samples was calculated based on the corresponding gene expression at 0 h using the  $2^{-\Delta\Delta t}$  method. The error bars represent the standard deviation (SD) from three independent experiments. The differences between the transcript level of treated adventitious roots and control adventitious roots at 0 h are statistically significant (\* $P < 0.05$ , \*\* $P < 0.01$ ) according to independent samples *t*-test.

to various phytohormone treatments and NaCl stress. According to our results, *PgWRKY1* transcript was significantly upregulated by the treatment with SA, ABA, and salt, but declined by MeJA treatment (Fig. 11). The expression of *PgWRKY1* upon MeJA treatment seemed not similar to that of the *PqWRKY1* expression, upon SA, ABA, and salt treatments, respectively. However, it is noteworthy that *PgWRKY1* expression result is consistent with the result of our MeJA-treated adventitious root transcriptome expression dataset (Fig. 2B). Previous studies reported that ABA, SA, and MeJA treatment resulted in different trends of expression patterns. For example, GaWRKY1 from cotton and AtWRKY40 from *Arabidopsis* were found to negatively regulate ABA response, but positively regulate MeJA responses [52,67]. Moreover, it is also evident that the genes which were significantly upregulated by SA were not shown to be induced upon MeJA treatment and vice versa, which is consistent with the known antagonistic effect of MeJA and SA on gene expression [68]. The *PgWRKY1* in this study was found to be upregulated after treatment with SA, but downregulated upon MeJA treatment, which is also in agreement with the above result. *PgWRKY1* is the first TF gene that has been cloned and preliminarily identified in *P. ginseng* as a multiple stress-inducible gene in response to salt stress and several phytohormones. Moreover, this work provides a fundamental knowledge that will help to further elucidate the molecular mechanisms controlling the biosynthesis of pharmaceutically valuable ginsenosides in *P. ginseng* and its closely related species.

## Supplementary Data

Supplementary Data is available at ABBS online.

## Acknowledgements

We would like to thank Dr. Ru Zhang (College of Chemistry and Chemical Engineering, Hunan Institute of Engineering, Xiangtan, China) for his valuable suggestions and help in revision of this manuscript.

## Funding

The work was supported by the grants from the National Natural Science Foundation of China (Nos. 30470189, 81071821, 81250110086, and 812111392), the Key Hunan Provincial Natural Science Foundation of China (Nos. 13JJ2016 and 07JJ5096), the Science and Technology Program of Hunan Province of China (No. 2014SK4050), Changsha National High Tech Zone Innovation Leading Talent Program (No. 20143701), and the Fundamental Research Funds for the Central Universities of Central South University (No. 721500011).

## References

- Zhang Y, Wang L. The WRKY transcription factor superfamily: its origin in eukaryotes and expansion in plants. *BMC Evol Biol* 2005, 5: 1–12.
- Ishiguro S, Nakamura K. Characterization of a cDNA encoding a novel DNA-binding protein, SPF1, that recognizes SP8 sequences in the 5' upstream regions of genes coding for sporamin and beta-amylase from sweet potato. *Mol Gen Genet* 1994, 244: 563–571.
- Rushton PJ, Macdonald H, Huttly AK, Lazarus CM, Hooley R. Members of a new family of DNA-binding proteins bind to a conserved cis-element in the promoters of alpha-Amy2 genes. *Plant Mol Biol* 1995, 29: 691–702.
- Rushton PJ, Torres JT, Parniske M, Wernert P, Hahlbrock K, Somssich IE. Interaction of elicitor-induced DNA-binding proteins with elicitor response

- elements in the promoters of parsley *PR1* genes. *EMBO J* 1996, 15: 5690–5700.
- Ulker B, Somssich IE. WRKY transcription factors: from DNA binding towards biological function. *Curr Opin Plant Biol* 2004, 7: 491–498.
- Ross CA, Liu Y, Shen QJ. The WRKY gene family in rice (*Oryza sativa*). *J Integr Plant Biol* 2007, 49: 827–842.
- Ling J, Jiang W, Zhang Y, Yu H, Mao Z, Gu X, Huang S, et al. Genome-wide analysis of WRKY gene family in *Cucumis sativus*. *BMC Genomics* 2011, 12: 471.
- He H, Dong Q, Shao Y, Jiang H, Zhu S, Cheng B, Xiang Y. Genome-wide survey and characterization of the WRKY gene family in *Populus trichocarpa*. *Plant Cell Rep* 2012, 31: 1199–1217.
- Eulgem T, Rushton PJ, Robatzek S, Somssich IE. The WRKY superfamily of plant transcription factors. *Trends Plant Sci* 2000, 5: 199–206.
- Cheong YH, Chang HS, Gupta R, Wang X, Zhu T, Luan S. Transcriptional profiling reveals novel interactions between wounding, pathogen, abiotic stress, and hormonal responses in *Arabidopsis*. *Plant Physiol* 2002, 129: 661–677.
- Rizhsky L, Liang H, Mittler R. The combined effect of drought stress and heat shock on gene expression in tobacco. *Plant Physiol* 2002, 130: 1143–1151.
- Huang T, Duman JG. Cloning and characterization of a thermal hysteresis (antifreeze) protein with DNA-binding activity from winter bittersweet nightshade, *Solanum dulcamara*. *Plant Mol Biol* 2002, 48: 339–350.
- Pnueli L, Hallak-Herr E, Rozenberg M, Cohen M, Goloubinoff P, Kaplan A, Mittler R. Molecular and biochemical mechanisms associated with dormancy and drought tolerance in the desert legume *Retama raetam*. *Plant J* 2002, 31: 319–330.
- Johnson CS, Kolevski B, Smyth DR. TRANSPARENT TESTA GLABRA2, a trichome and seed coat development gene of *Arabidopsis*, encodes a WRKY transcription factor. *Plant Cell* 2002, 14: 1359–1375.
- Lagace M, Matton DP. Characterization of a WRKY transcription factor expressed in late torpedo-stage embryos of *Solanum chacoense*. *Planta* 2004, 219: 185–189.
- Hinderhofer K, Zentgraf U. Identification of a transcription factor specifically expressed at the onset of leaf senescence. *Planta* 2001, 213: 469–473.
- Robatzek S, Somssich IE. Targets of AtWRKY6 regulation during plant senescence and pathogen defense. *Genes Dev* 2002, 16: 1139–1149.
- Mao G, Meng X, Liu Y, Zheng Z, Chen Z, Zhang S. Phosphorylation of a WRKY transcription factor by two pathogen-responsive MAPKs drives Phytoalexin biosynthesis in *Arabidopsis*. *Plant Cell* 2011, 23: 1639–1653.
- Kato N, Dubouzet E, Kokabu Y, Yoshida S, Taniguchi Y, Dubouzet JG, Yazaki K, et al. Identification of a WRKY protein as a transcriptional regulator of benzyloquinoline alkaloid biosynthesis in *Coptis japonica*. *Plant Cell Physiol* 2007, 48: 8–18.
- Ma D, Pu G, Lei C, Ma L, Wang H, Guo Y, Chen J, et al. Isolation and characterization of AaWRKY1, an *Artemisia annua* transcription factor that regulates the Amorpha-4,11-diene synthase gene, a key gene of artemisinin biosynthesis. *Plant Cell Physiol* 2009, 50: 2146–2161.
- Sun Y, Niu Y, Xu J, Li Y, Luo H, Zhu Y, Liu M, et al. Discovery of WRKY transcription factors through transcriptome analysis and characterization of a novel methyl jasmonate-inducible *PqWRKY1* gene from *Panax quinquefolius*. *Plant Cell Tiss Org Cult* 2013, 114: 269–277.
- Shekhawat UKS, Ganapathi TR, Srinivas L. Cloning and characterization of a novel stress-responsive WRKY transcription factor gene (*MusaWRKY71*) from *Musa* spp. cv. *Karibale Monthan* (ABB group) using transformed banana cells. *Mol Biol Rep* 2011, 38: 4023–4035.
- Wen J, Zimmer EA. Phylogeny and biogeography of *Panax* L. (the ginseng genus, araliaceae): inferences from ITS sequences of nuclear ribosomal DNA. *Mol Phylogenet Evol* 1996, 6: 167–177.
- Kim KH, Choi I, Lee YW, Cho CK, Yoo HS, Lee SB, Choi SH, et al. Target genes involved in antiproliferative effect of modified ginseng extracts in lung cancer A549 cells. *Acta Biochim Biophys Sin* 2014, 46: 441–449.
- Yang N, Chen P, Tao Z, Zhou N, Gong X, Xu Z, Zhang M, et al. Beneficial effects of ginsenoside-Rg1 on ischemia-induced angiogenesis in diabetic mice. *Acta Biochim Biophys Sin* 2012, 44: 999–1005.

26. Chu S, Zhang J. New achievements in ginseng research and its future prospects. *Chin J Integr Med* 2009, 15: 403–408.
27. Afrin S, Zhu J, Cao H, Huang J, Xiu H, Luo T, Luo Z. Molecular cloning and expression profile of an abiotic stress and hormone responsive MYB transcription factor gene from *Panax ginseng*. *Acta Biochim Biophys Sin* 2015, 47: 267–277.
28. Suttipanta N, Pattanaik S, Kulshrestha M, Patra B, Singh SK, Yuan L. The transcription factor CrWRKY1 positively regulates the terpenoid indole alkaloid biosynthesis in *Catharanthus roseus*. *Plant Physiol* 2011, 157: 2081–2093.
29. Cao H, Nuruzzaman M, Xiu H, Huang J, Wu K, Chen X, Li J, et al. Transcriptome analysis of methyl jasmonate-elicited *Panax ginseng* adventitious roots to discover putative ginsenoside biosynthesis and transport genes. *Int J Mol Sci* 2015, 16: 3035–3057.
30. Tamura K, Dudley J, Nei M, Kumar S. MEGA4: Molecular evolutionary genetics analysis (MEGA) software version 4.0. *Mol Biol Evol* 2007, 24: 1596–1599.
31. Rushton PJ, Somssich IE, Ringler P, Shen QJ. WRKY transcription factors. *Trends in Plant Sci* 2010, 15: 247–258.
32. Deng W, Wang Y, Liu Z, Cheng H, Xue Y. Heml: a tool kit for illustrating heatmaps. *Plos One* 2014, 9: e111988.
33. de Pater S, Greco V, Pham K, Memelink J, Kijne J. Characterization of a zinc-dependent transcriptional activator from *Arabidopsis*. *Nucleic Acids Res* 1996, 24: 4624–4631.
34. Nuruzzaman M, Zhang R, Cao HZ, Luo ZY. Plant pleiotropic drug resistance transporters: transport mechanism, gene expression, and function. *J Integr Plant Biol* 2014, 56: 729–740.
35. Zhang R, Huang J, Zhu J, Xie X, Tang Q, Chen X, Luo J, et al. Isolation and characterization of a novel PDR-type ABC transporter gene *PgPDR3* from *Panax ginseng* C.A. Meyer induced by methyl jasmonate. *Mol Biol Rep* 2013, 40: 6195–6204.
36. Yamasaki K, Kigawa T, Inoue M, Tateno M, Yamasaki T, Yabuki T, Aoki M, et al. Solution structure of an *Arabidopsis* WRKY DNA binding domain. *Plant Cell* 2005, 17: 944–956.
37. Kelley LA, Sternberg MJE. Protein structure prediction on the Web: a case study using the Phyre server. *Nat Protoc* 2009, 4: 363–371.
38. Biasini M, Bienert S, Waterhouse A, Arnold K, Studer G, Schmidt T, Kiefer F, et al. SWISS-MODEL: modelling protein tertiary and quaternary structure using evolutionary information. *Nucleic Acids Res* 2014, 42: W252–W258.
39. Jensen LJ, Gupta R, Blom N, Devos D, Tamames J, Kesmir C, Nielsen H, et al. Prediction of human protein function from post-translational modifications and localization features. *J Mol Biol* 2002, 319: 1257–1265.
40. Peng XX, Tang XK, Zhou PL, Hu YJ, Deng XB, He Y, Wang HH. Isolation and expression patterns of rice WRKY82 transcription factor gene responsive to both biotic and abiotic stresses. *Agri Sci China* 2011, 10: 893–901.
41. Kim YJ, Shim JS, Lee JH, Jung DY, Sun H, In JG, Yang DC. Isolation and characterization of a novel short-chain alcohol dehydrogenase gene from *Panax ginseng*. *BMB Rep* 2009, 42: 673–678.
42. Ramamoorthy R, Jiang SY, Kumar N, Venkatesh PN, Ramachandran S. A comprehensive transcriptional profiling of the WRKY gene family in rice under various abiotic and phytohormone treatments. *Plant Cell Physiol* 2008, 49: 865–879.
43. Zhang R, Zhu J, Cao HZ, An YR, Huang JJ, Chen XH, Mohammed N, et al. Molecular cloning and expression analysis of PDR1-like gene in ginseng subjected to salt and cold stresses or hormonal treatment. *Plant Physiol Biochem* 2013, 71: 203–211.
44. Sheard LB, Tan X, Mao HB, Withers J, Ben-Nissan G, Hinds TR, Kobayashi Y, et al. Jasmonate perception by inositol-phosphate-potentiated COI1-JAZ co-receptor. *Nature* 2010, 468: 400–405.
45. Xie DX, Feys BF, James S, Nieto-Rostro M, Turner JG. COI1: an *Arabidopsis* gene required for jasmonate-regulated defense and fertility. *Science* 1998, 280: 1091–1094.
46. Pan Q, Chen Y, Wang Q, Yuan F, Xing S, Tian Y, Zhao J, et al. Effect of plant growth regulators on the biosynthesis of vinblastine, vindoline and catharanthine in *Catharanthus roseus*. *Plant Growth Regul* 2010, 60: 133–141.
47. Schluttenhofer C, Pattanaik S, Patra B, Yuan L. Analyses of *Catharanthus roseus* and *Arabidopsis thaliana* WRKY transcription factors reveal involvement in jasmonate signaling. *BMC Genomics* 2014, 15: 502.
48. Dong J, Chen C, Chen Z. Expression profiles of the *Arabidopsis* WRKY gene superfamily during plant defense response. *Plant Mol Biol* 2003, 51: 21–37.
49. Kalde M, Barth M, Somssich IE, Lippok B. Members of the *Arabidopsis* WRKY group III transcription factors are part of different plant defense signaling pathways. *Mol Plant Microbe Interact* 2003, 16: 295–305.
50. Qiu D, Xiao J, Ding X, Xiong M, Cai M, Cao C, Li X, et al. OsWRKY13 mediates rice disease resistance by regulating defense-related genes in salicylate- and jasmonate-dependent signaling. *Mol Plant Microbe Interact* 2007, 20: 492–499.
51. Liu X, Bai X, Wang X, Chu C. OsWRKY71, a rice transcription factor, is involved in rice defense response. *J Plant Physiol* 2007, 164: 969–979.
52. Pandey SP, Roccaro M, Schon M, Logemann E, Somssich IE. Transcriptional reprogramming regulated by WRKY18 and WRKY40 facilitates powdery mildew infection of *Arabidopsis*. *Plant J* 2010, 64: 912–923.
53. Zou X, Shen QJ, Neuman D. An ABA inducible WRKY gene integrates responses of creosote bush (*Larrea tridentata*) to elevated CO<sub>2</sub> and abiotic stresses. *Plant Sci* 2007, 172: 997–1004.
54. Zhou L, Wang NN, Kong L, Gong SY, Li Y, Li XB. Molecular characterization of 26 cotton WRKY genes that are expressed differentially in tissues and are induced in seedlings under high salinity and osmotic stress. *Plant Cell Tiss Organ Cult* 2014, 119: 141–156.
55. Xie Z, Zhang ZL, Zou X, Yang G, Komatsu S, Shen QJ. Interactions of two abscisic-acid induced WRKY genes in repressing gibberellin signaling in aleurone cells. *Plant J* 2006, 46: 231–242.
56. Zhu JK. Cell signaling under salt, water, and cold stresses. *Curr Opin Plant Biol* 2001, 4: 401–406.
57. Song Y, Ai CR, Jing SJ, Yu DQ. Research progress on functional analysis of rice WRKY genes. *Rice Sci* 2010, 17: 60–72.
58. Chen SL, Sun YZ, Xu J, Luo HM, Sun C, He L, Cheng XL, et al. Strategies of the study on herb genome program. *Acta Pharmacol Sin* 2010, 45: 807–812.
59. Miao LX, Zhang YC, Yang XF, Jiang GH. Genome-wide identification and analysis of WRKY transcription factors in *Fragaria vesca*. *J Nucl Agriculture Sci* 2012, 8: 1124–1131.
60. Pandey SP, Somssich IE. The role of WRKY transcription factors in plant immunity. *Plant Physiol* 2009, 150: 1648–1655.
61. Xiong W, Xu X, Zhang L, Wu P, Chen Y, Li M, Jiang H, et al. Genome-wide analysis of the WRKY gene family in physic nut (*Jatropha curcas* L.). *Gene* 2013, 524: 124–132.
62. Huang S, Gao Y, Liu J, Peng X, Niu X, Fei Z, Cao S, et al. Genome-wide analysis of WRKY transcription factors in *Solanum lycopersicum*. *Mol Genet Genomics* 2012, 287: 495–513.
63. Duan MR, Nan J, Liang YH, Mao P, Lu L, Li L, Wei C, et al. DNA binding mechanism revealed by high resolution crystal structure of *Arabidopsis thaliana* WRKY1 protein. *Nucleic Acids Res* 2007, 35: 1145–1154.
64. Brand LH, Fischer NM, Harter K, Kohlbacher O, Wanke D. Elucidating the evolutionary conserved DNA-binding specificities of WRKY transcription factors by molecular dynamics and *in vitro* binding assays. *Nucleic Acids Res* 2013, 41: 9764–9778.
65. Satapathy L, Singh D, Ranjan P, Kumar D, Kumar M, Prabhu KV, Mukhopadhyay K. Transcriptome-wide analysis of WRKY transcription factors in wheat and their leaf rust responsive expression profiling. *Mol Gene Genomics* 2014, 289: 1289–1306.
66. Mao K, Hayashi S, Kojima-Suzuki H, Morikami A, Nakamura K. Role of conserved residues of the WRKY domain in the DNA-binding of tobacco WRKY family proteins. *Biosci Biotech Bioch* 2001, 65: 2428–2436.
67. Chen H, Lai Z, Shi J, Xiao Y, Chen Z, Xu X. Roles of *Arabidopsis* WRKY18, WRKY40 and WRKY60 transcription factors in plant responses to abscisic acid and abiotic stress. *BMC Plant Biol* 2010, 10: 281.
68. Moons A. Transcriptional profiling of the PDR gene family in rice roots in response to plant growth regulators, redox perturbations and weak organic acid stresses. *Planta* 2008, 229: 53–71.