Plant Omics Journal POJ

POJ 8(4):353-360 (2015) ISSN:1836-3644

In silico genome-wide analysis of the *WRKY* gene family in *Salix arbutifolia*

Guodong Rao[1,3], Jinkai Sui[3], Jianguo Zhang[1,2,3]*

[1]State Key Laboratory of Tree Genetics and Breeding, Research Institute of Forestry, Chinese Academy of Forestry, Beijing 100091, Republic of China
[2]Collaborative Innovation Center of Sustainable Forestry in Southern China, Nanjing Forestry University, Nanjing 210037, Republic of China
[3]Key Laboratory of Tree Breeding and Cultivation, State Forestry Administration, Research Institute of Forestry, Chinese Academy of Forestry, Beijing 100091, Republic of China

***Corresponding author: zhangjg@caf.ac.cn**

Abstract

The *WRKY* genes encode transcriptional regulators that function in response to biotic and abiotic stress in plants. Thus far, no detailed classification and expression profiles of *WRKY* genes are available for willow. In this study, a comprehensive computational analysis identified 89 *WRKY* family genes in willow (*Salix arbutifolia*) by the *in silico* cloning method. Based on the results of phylogenetic analysis and the number of WRKY motifs, the WRKY genes were classified into group I to group III. The expression profile was analyzed by the transcriptome data in different tissues. A comparative analysis of *Salix*, *Populus trichocarpa*, and *Arabidopsis thaliana* were performed, and the *Salix* WRKY family had a similar number as *Populus* and a larger size than the Arabidopsis. A detailed phylogenetic analysis between *Salix* and *Populus* revealed that the WRKY genes had evolved from the same genome duplication event. These results will be useful for future functional analyses of the WRKY family genes.

Keywords: WRKY; phylogenetic analyses; transcriptome; *Salix arbutifolia.*
Abbreviations: TFs_ Transcription factors; MEME_ Motif Elicitation; RPKM_Reads Per kb per Million reads.

Introduction

The genus *Salix* has approximately 400 species, all of which are found in the Northern hemisphere. *Salix* is also a valuable forest resource and an important ecological species due to the commercial significance of willow for renewable energy and multipurpose production (Berlin et al., 2010). The utilization and research of willow for biomass production in short-rotation-intensive culture systems is growing worldwide, which has provided increasing pressure for the breeding of high-yield and resistant clones adapted to different environments (Karp et al., 2011). Transcription factors (TFs) play pivotal roles in signal transduction to activate or suppress defense response genes, which regulate the adaptation of plants, especially when encountering plant face abiotic and biotic stress conditions such as drought and high salinity (Sharma et al., 2010). The WRKY superfamily is a conserved family, which is one of the ten largest transcription factor families in higher plants and extends all over the green lineage (Ulker and Somssich 2004; Wei et al., 2012). WRKY family members contain one or two copies of unique DNA-binding domain with approximately 60 amino acids with the conserved sequence WRKYGXK at the N-terminal followed by a zinc finger like motif (C-X4-5-C-XX22-23-H-X-H or C-X5-8-C-X25-28-H-X1-2-C) at the C-terminus (Ulker and Somssich 2004). WRKY transcription factors play important roles in the plant signaling pathway and modulate many plant processes such as wounding responses (Yoo et al., 2014), pollen development (Guan et al., 2014), stomatal movement (Ding et al., 2014c), flowering time (Cai et al., 2014), cell wall formation (Wang et al., 2010; Yu et al., 2013), terpenoid indole alkaloid biosynthesis (Suttipanta et al., 2011), leaf senescence (Miao et al., 2004), and plant defense different stresses, including biotic stress (Huh et al., 2012; Marchive et al., 2007) and abiotic stress (Ma et al. 2009; Ren et al., 2010; Tang et al., 2013).

To date, many studies have revealed the underlying molecular mechanism of *WRKY* family genes in plant growth and development (Wei et al., 2012; Zhang and Wang, 2005). In *Arabidopsis thaliana*, 72 *WRKY* genes were identified, which were classified into three groups based on the phylogenetic and domain analysis of the proteins sequences. The first group contains two WRKY domains, and the other two groups have only one domain (de Pater et al., 1996; Eulgem et al., 2000). Arabidopsis as a model plant set a good example for the function research of *WRKY* genes in plants, and many Arabidopsis WRKY were functionally characterized. *AtWRKY46* plays dual roles in regulating plant responses to drought and salt stress and light-dependent stomatal opening in guard cells. *AtWRKY41* controls Arabidopsis seed dormancy via direct regulation of ABI3 transcript levels not downstream of ABA (Ding et al., 2014b). *AtWRKY30* is responsive to several stress conditions either from abiotic or biotic origin, suggesting that this gene could have a role in the activation of defense responses at early stages of Arabidopsis growth by binding to W-boxes found in promoters of many stress/developmentally regulated genes (Scarpeci et al., 2013). Defense-related transcription factors AtWRKY70 and AtWRKY54 modulate osmotic stress tolerance by regulating stomatal aperture in Arabidopsis (Li et al., 2013). There are also many *WRKY* genes in various plant species that have been identified apart from *Arabidopsis thaliana*, which includes sweet potatoe (Ishiguro and Nakamura 1994), *Avena fatua*

(Rushton et al., 1995), *Nicotiana tabacum* (Chen and Chen, 2000), *Matricaria chamomilla* (Ashida et al., 2002), *Oryza sativa* (Kim et al., 2000), *Solanum dulcamara* (Huang and Duman 2002), *Triticum aestivum* (Sun et al., 2003), *Vitis vinifera* (Marchive et al., 2007), *Petroselinum crispum* (Cormack et al., 2002), *Hevea brasiliensis* (Li et al., 2014), *Gossypium raimondii* (Dou et al., 2014), *Jatropha curcas* L. (Xiong et al., 2013), Cucumis sativus (Xiong et al., 2013), and *Populus trichocarpa* (He et al., 2012). In this study, we identified 89 *WRKY* genes from *Salix arbutifolia* Pallas (originally named *Chosenia arbutifolia*) which is a riparian dioecious tree species distributed in Northeast Asia (Kadis 2005). These 89 WRKY proteins in *Salix* were categorized into three groups (group I, group II, and group III), Group I contain 42 WRKY members which were classified into two subgroups, group II contain 37 *WRKY* genes which classified into five subgroups, and group III have 10 WRKY members. Phylogenetic trees were constructed by using the WRKY protein amino acid sequences from Arabidopsis, *Populus*, and *Salix*, as well as the motif analysis of the WRKY proteins. Moreover, the transcript level of *WRKY* genes in different tissues of *S. arbutifolia* was analyzed using the transcriptome data. The resulting category of groups and identification of putative functional motifs will be useful in studies on the biological functions of each gene in the *Salix* WRKY families, and the transcriptome reported here on the willow significantly enhance our knowledge of TFs and stress response genes in *Salix*.

Results

Identification and prediction of the WRKY family genes in S. arbutifolia

In this study, a genome-wide analysis was carried out to identify *WRKY* genes in the *S. arbutifolia* genome sequence and putative full-length protein sequences (paper in publish). Initially, 93 partial putative full-length protein sequences of *WRKY* genes were identified with the BLASTP software. Four members were removed because they did not contain the WRKY domains or complete zinc finger motifs from the manual inspection of these putative *WRKY* genes. Eventually, we had a total of 89 members representing the unique *Salix* genes for the next analysis. According to the phylogenetic and motif analysis of the 89 WRKY proteins, the *Salix* WRKY family divided into three groups (group I, group II, and group III). Group I contained two subfamilies (Ia and Ib) and had 42 members that accounted for 47.19% of the total *WRKY* genes. Group II contained four subfamilies (IIa-IIe) and 37 members that accounted for 41.57% of total *WRKY* genes (Table I). Ten genes were in group III accounting for 11.24%. Different numbers and percentages of *WRKY* subfamily genes among *S. arbutifuolia*, *P. trichocarpa*, and *A. thaliana* are shown in Table 1.

Phylogenetic analysis and identification of conserved motifs

In general, transcription factor families contain highly conserved domains or consensus motifs involved in DNA binding (Park et al., 2005). These conserved DNA domains are considered an evolutionary unit with coding sequences that can be duplicated or undergo recombination (Wu et al., 2005). The WRKY superfamily contains 60 amino acid residues, which are the most prominent structural feature of the WRKY proteins (Zhang and Wang 2005). An unrooted phylogenetic tree and a linear distribution map of the conserved motifs in the putative *Salix* WRKY protein were produced using the WRKY amino acid sequences, which provided evolutionary relationships and motif characteristics among the WRKY proteins (Fig 1). The phylogenetic tree showed that 89 WRKY proteins could be divided into three groups and eight subfamilies, containing WRKY members that had the same or similar conserved motif distributions. The eight subfamilies were named I (Ia and Ib), II (IIa-e) and III, based on the 60 WRKY domains and the type of zinc finger in the C-terminal. The gene structure of diverse exon-intron organization of *Salix* WRKY subfamily members was analyzed by the software GSDS (Fig 1). In group I, Ia and Ib subfamily genes showed gene structure divergence. Most of the Ia genes have three exons, while Ib genes are mainly included over four exons. IIa and IIb in group II also showed different gene structures, IIb has a longer DNA sequence compared to IIa, and most of the group II genes have over five exons. In group III, genes in each subfamily have similar gene structures and DNA lengths, and most of the genes have three exons.

The most prominent feature of plant *WRKY* gene family members is the presence of WRKYGQK, which is a highly conserved sequence at the N-terminus. Of the 89 WRKY proteins, 83 have the highly conserved domain, while the other six WRKY proteins have one mismatched amino acid. Group I consisted of 42 members, which were divided into two subgroups (Ia and Ib), Ia subgroup members containing two WRKY domains, while the Ib subgroup members have only one WRKY domain. The other two groups (group II and group III) have 15 and 32 members, respectively, and each member has one WRKY domain (Fig 2.).

Motif analysis of Salix WRKY proteins

The *Salix* WRKY proteins were classified into eight subfamilies and named Ia-b, IIa-e, and III (Fig 1). The eight subfamilies have different conserved motifs, while some of the motifs are shared by all of the WRKY members in *Salix*, such as motif 3. Of the ten motifs, motifs 1, 2, 3, and 9, which represented the distribution of the C- or N-terminal WRKY domains, were contained in the WRKY amino acid residues (Table 2). Group I contained two subgroups (Ia and Ib), and the Ia subgroup consisted of 22 WEKY members, containing two WRKY domains, with most of the members sharing motifs 2 and 3. The Ib subgroup consisted of 20 WRKY members that shared motifs 2, 3, and 4, except for four WRKY proteins (CCG028696, CCG020502, CCG003617, and CCG009663). Group II contained five subgroups (IIa-e) and consisted of 37 WRKY proteins. The IIa subfamily consisted of five members and shared motifs 2 and 3. IIb has ten WRKY members, which have the conserved motif 2 except CCG032743 and CCG008674. IIc and IId have 16 and four WRKY proteins, respectively, and motif 3 is conserved in these two subgroups. Subgroup III contained ten members, which have the conserved motif 1 (Fig 1).

Table 1. Summary of the WRKY family of *S. arbutifolia, P. trichocarpa* and *A. thaliana*. Number and Percent in the table means the number and percentage of sub-family genes in different species.

Plant		*S. arbutifolia*		*P. trichocarpa*		*A. thaliana*	
Group	subgroup	Number	Percent (%)	Number	Percent (%)	Number	Percent (%)
I	Ia	22	24.72	23	22.12	14	19.44
	Ib	20	22.47	27	25.96	18	25.00
	Total	42	47.19	50	48.08	32	44.44
II	IIa	5	5.62	5	4.81	3	4.17
	IIb	10	11.24	9	8.65	8	11.11
	IIc	11	12.36	13	12.50	7	9.72
	IId	6	6.74	13	12.50	7	9.72
	IIe	5	5.50	4	3.85		0.00
	Total	37	41.57	44	42.31	26	36.11
III		10	11.24	10	9.62	14	19.44
Total		89		104		72	

The number of WRKY family from *P. trichocarpa* and *A. thaliana* were according to He et al. (He et al., 2012).

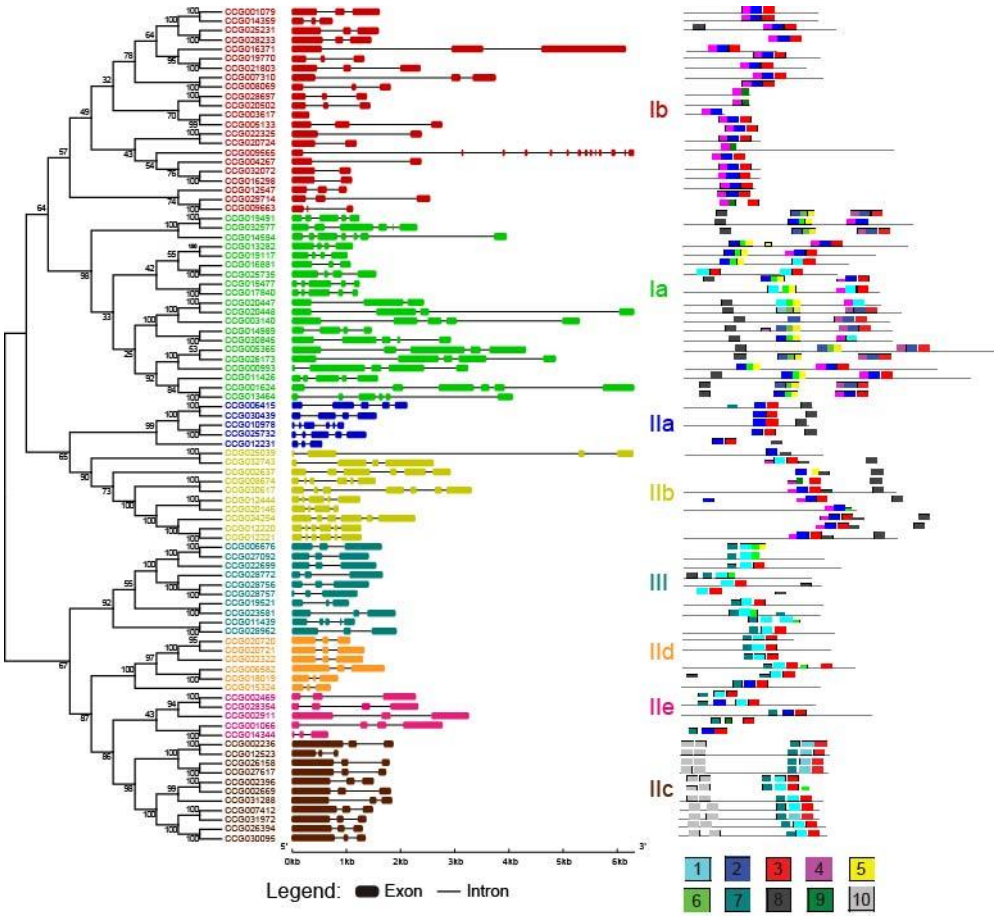

Fig 1. Phylogenetic analysis and conserved domains of WRKY proteins of *Salix*. The unrooted phylogenetic tree is shown on the left side, the gene structure is shown in the middle, and the distribution of conserved motifs of the WRKY protein is shown on the right side of the figure. The different colored boxes at the bottom of the figure indicated different motifs.

Transcripts expression level of WRKY genes in S. arbutifolia

Transcription levels of *WRKY* genes were statistical analyzed from transcriptome sequenced by RNA-seq and de novo assembly. *WRKY* genes displayed differential expressions either in their transcript abundance or expression patterns under normal growth conditions. The expression profiles reveal spatial and temporal variations in the expression of the *WRKY* subfamily genes of different *Salix* tissues (stem, blossom bud, and leaf). In the Ia subfamily, the genes showed most abundant transcript in the stems, and also expressed at moderate levels in the leafs and blossom buds. Seven genes had higher expression level compared to the extremely low transcript level of other genes in Ia subfamily, suggesting the functional redundancy in this family. The 20 members of the Ib subfamily also showed different expression patterns, and genes in this subfamily had the most abundant transcript in stem, and the other half of the family members had very low expression levels in the tissue examined (Fig 3A). In the IIa-IIe subfamilies, the genes in the Ic subfamily showed relative higher expressions level compared to the other subfamilies, and most of the members had abundant transcription in the stems. The ten members of

Table 2. Motif sequences of *Salix* WRKY proteins identified by MEME tolls.

Motif	Width	Consensus sequence	Motif description (Pfam)
Motif 1	25	DGYRWRKYGQKVIKGNPYPRGYYRC	WRKY DNA -binding
Motif 2	25	DGYRWRKYGQKVVKGNPYPRSYYRC	WRKY DNA -binding
Motif 3	25	CPVRKHVERCWEDPTMVITTYEGEH	WRKY DNA -binding
Motif 4	21	HKRIREPRFAFQTRSEVDILD	
Motif 5	15	DGHIFEITYKGEHNH	
Motif 6	15	THPGCPVKKQVERSH	
Motif 7	20	CHCSKRRKHRVKWTVRVPAI	Plant zinc cluster
Motif 8	25	TAAITRDPYFTIPLGLAITSLLDSP	
Motif 9	15	DGYRWRKYGQKVVKN	WRKY DNA -binding
Motif 10	25	NCMEATDAAVFKFKKVISLLCQTQH	

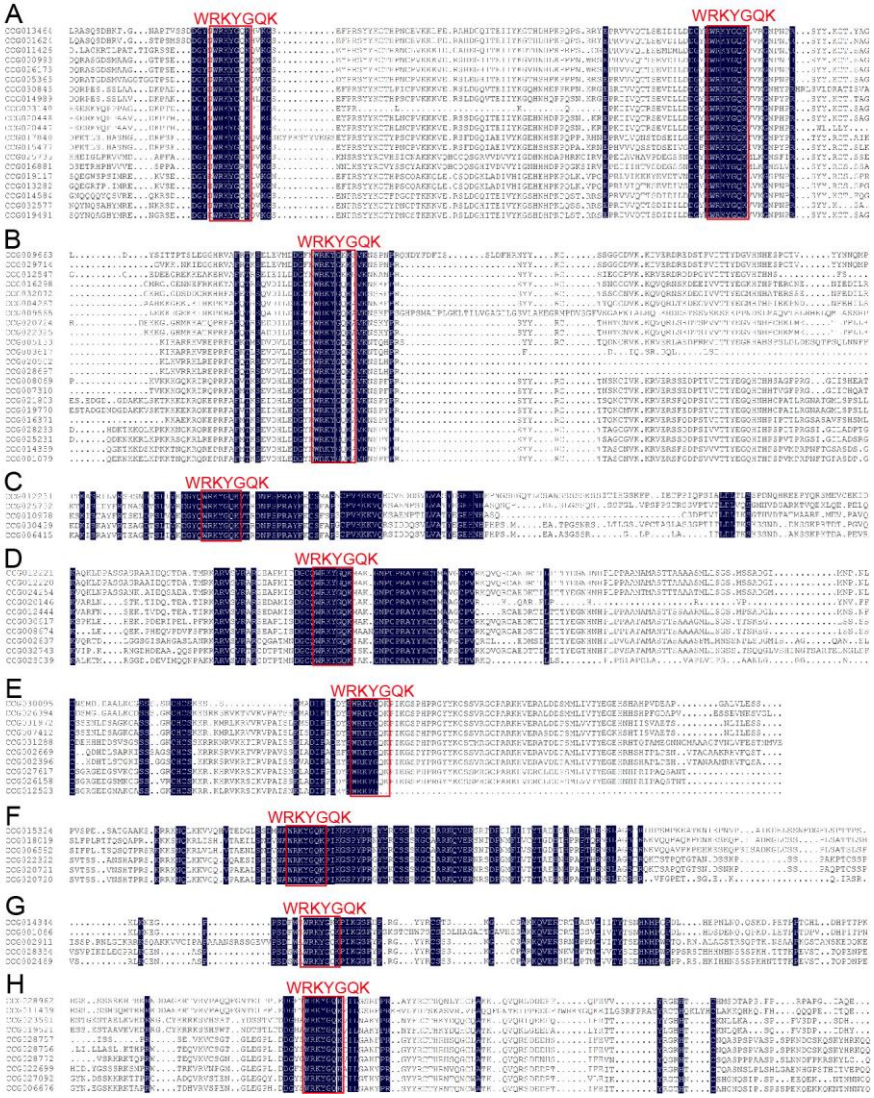

Fig 2. Multiple sequence alignment of WRKY proteins in *Salix*. (A-B) Alignment of Ia and Ib subfamily proteins; (C-G) Alignment of IIa-IIb subfamily proteins; (H) Alignment of III subfamily proteins. The highly conserved WRKYGQK domain was highlighted in dark blue, gaps are marked as dot.

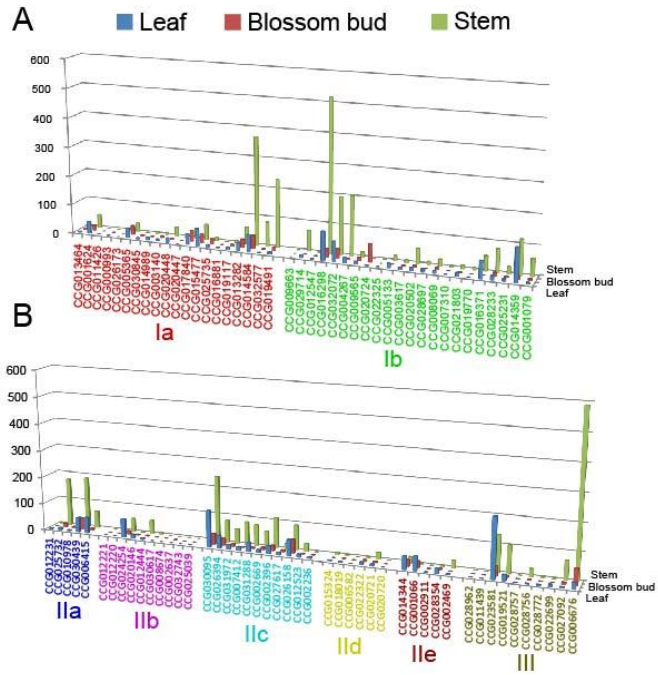

Fig 3. Transcript level of WRKY subfamilies in the blossom buds, stems and leafs. (a) The expression level of Ia and Ib subfamilies; (b) the expression level of IIa-d and III subfamilies.

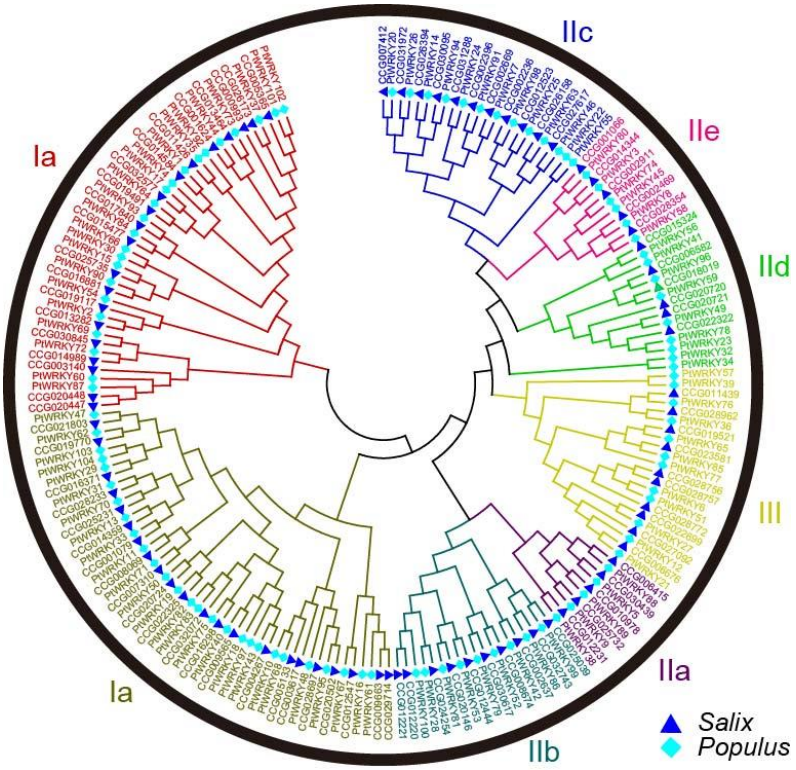

Fig 4. Phylogenetic analysis of WRKY proteins between *Salix* and *Populus*. The different colors indicate different groups. WRKY genes of *Salix* and *Populus* are showed in blue triangle and sky blue diamonds.

the III subfamily showed a variant expression pattern, one of the members (CCG006676) had higher expression levels in all tissues examined, and three of the members had moderate transcription levels (CCG023581, CCG019521, and CCG027092) compared to the low expression level of the remaining genes (Fig 3B).

Discussion

Salix species are trees, shrubs, and subshrubs and members of the Salicaceae family. Previous studies have revealed that many *Salix* species have the capacity to tolerate abiotic and biotic stress. Two Russian willow species, *Salix viminalis* and *Salix*

dasyclados have shown strong genetic control of freezing resistance during autumn (Tsarouhas et al., 2004). *Salix caprea*, *Salix repens* and their F1 hybrids have the ability to tolerate different insects and pathogens (Hjältén 1998). The willow species *S. caprea* and *S. cinera* are grown on nutrient- deficient and industrially-contaminated soils; thus, they serve as biological filters for waste water and sludge disposal (Punshon and Dickinson 1997). Nine different clones of six species of *Salix* (*Salix cordata* Muhlenb. non Michaux, *S. fragilis* L., *S. caprea* L., *S. cinerea* L., *S. burjatica* Nazarov, and *S. viminalis* L.) and one hybrid (*S.* x *calodendron* Wimm.) were exposed to heavy metals in solution culture resulting in reduced phytotoxicity and increased metal resistance (Punshon and Dickinson 1997). Plant transcription factors potentially control downstream gene expression in stress signal transduction pathways through activation and repression of the genes after certain stress treatments (Rao et al., 2014). WRKY play an important role in development and stress responses as one of the largest families of transcription factors in plants. The diverse functions and characteristics of this family have been studied in the model plants, such as Arabidopsis, rice, tobacco, and *Populus*. Survey and characterization of *WRKY* genes in *Salix* would increase our understanding of this superfamily. Previous studies have reported many genome-wide analyses of the *WRKY* gene family members. A genome-wide analysis of the *WRKY* genes family in *Hevea brasiliensis* has recently been reported (Li et al., 2014). These discussed the number, expression profile, and genetic response to jasmonic acid and ethylene, which conclude that the WRKY proteins may be involved in the transcriptional regulation of nature rubber biosynthesis. Two genome-wide analyses of the *WRKY* genes family in cotton were recently reported (Ding et al., 2014a; Dou et al., 2014). They also discussed the gene number, the distribution of *WRKY* genes on chromosomes, the genetic structures, and the motifs in WRKY proteins. A genome-wide analysis of *WRKY* genes family in *Populus trichocarpa* was conducted after the genome sequences was released (He et al., 2012). This study was the first survey of WRKY genes in ligneous species, which researched the number of the *WRKY* gene family, the gene structure, the protein conserved motifs, and the expression profile. In this study, 89 *WRKY* genes were identified and characterized in *S. arbutifolia*, and all of these genes were expressed in one or more tissues according to the transcriptome data. It is the first report of *WRKY* gene family in *Salix*. *WRKY* genes can be divided into three groups based on the protein sequence structure and motifs features. Previous study has revealed the Ib subfamily in group I, with the group II *WRKY* genes likely to have evolved from the Ia subfamily of group I genes. The difference between the two types is that group I WRKY proteins contain the N-terminal WRKY domain (He et al., 2012). Our sequence alignment results also show the same evolution pattern of *Salix*, i.e., members in Ia subfamily have a conserved N-terminal WRKY domain compared to other subfamily WRKY proteins. Comparison analysis of the gene number in Arabidopsis, *Populus*, and *Salix* revealed similar numbers of IIa, IIb, and IId *WRKY* genes among these species indicating that all genes in these subfamily have been fully identified (Table I). As described in *Populus*, the number of genes in group III of *Salix* is also lower than that of the older genotypes such as Arabidopsis and rice, thereby implying reduction of gene size during evolution. A phylogenetic tree based on the WRKY protein sequences between *Salix* and *Populus* was constructed (Fig 4). The tree shows us that over 80% of the WRKY protein exist as sister pairs between *Salix* and *Populus*, which suggest that gene duplication events play a key role in the evolution of *WRKY* gene family. As described previously, the *WRKY* family genes have exhibited tissue specifics in many plants. Genes highly expressed in specific tissue may relate to the development and metabolism of the plant. In Arabidopsis, *AtWRKY44* and *AtWRKY10* were highly expressed in early seed development, *AtWRKY23* is highly expressed in roots, *AtWRKY56* was expressed highly in leafs but to a lower level in root and stem (Xiong et al., 2013). The expression pattern of the stems, leafs and blossom buds in *S. arbutifolia* were analyzed using the transcriptome data. In the 89 *Salix WRKY* family genes, more than half of the genes have higher expression levels in the stem, and the expression level of *WRKY* genes are not restricted to a single organ, with many genes being mainly expressed in the blossom buds, thereby suggestion other functions of *WRKY* family genes. Four genes (CCG012221, CCG012220, CCG025039, and CCG028757) had no detectable expression levels, and another four genes (CCG008674, CCG032743, and CCG020720, and CCG028962) were only expressed in one tissue, which may be due to a very low transcription level in the three tissues and the fact that the RNA-seq method was limited, or that they are expressed only under specific conditions (Ling et al., 2011). These results indicate that the *Salix WRKY* genes participate in various aspects of developmental and metabolic processes. However, more studies are needed to determine the specific function of the WRKYs in *Salix*. The WRKY gene family has been extensively surveyed in many model plant species, such as Arabidopsis, rice, cotton, and poplar, but there is no research on *WRKY* family genes in *Salix*. Here, a total of 89 *WRKY* genes were identified, and their phylogenetic and transcript abundance analyzed in *S. arbutifolia*. This study produced a comprehensive genomics analysis of the *Salix WRKY* gene family, which has provided essential information on the cloning and functional analysis of *WRKY* genes. The comparative analysis of *WRKY* genes between *Salix* and *Populus* can be used in further studies to reveal the functions of the stress response or biosynthesis in trees.

Materials and Methods

Gene isolation and analysis

WRKY genes were isolated from *S. arbutifolia* genome sequence (paper in print) based on annotation and BLAST. First, the sequence indicated to belong to WRKY were isolated from the annotation. Then, using Arabidopsis and *Populus WRKY* genes with the TBALSTN and BLASTP, a batch of related sequences were obtained. All of the sequences that had the WRKY domains were selected as the candidates. To avoid the inclusion of redundant sequences, the characterized *WRKY* genes were assembled again with the CAP3 Sequence Assembly Program (Huang and Madan 1999). To analysis the structure divergence of the *Salix WRKY* gene, motif and gene structure analysis were conducted. The motif identification of WRKY in the encoded protein sequences were carried out using a motif based sequence analysis tool Motif Elicitation (MEME Suite version 4.9.1) (Bailey and Elkan 1994). The optimum width of amino acid sequences was set from 6 to 25, and the maximum number of motifs was set to 10. The gene structure of diverse exon-intron organization of *Salix WRKY* genes were constructed according to the comparison analysis of coding sequences with their corresponding genomic sequence using an online software GSDS (http://gsd.cbi.pku.edu.cn). The protein amino acid alignment was conducted by using the software DNAMAN (http://www.lynnon.com/dnaman.html).

Plant material, Libraries construction and deep sequencing

Stem, leaf and blossom buds from a natural population of *S.*

arbutifolia were harvested in the spring of 2013 in the Niaoning province of China. To ensure the absolute minimal differences in the growth conditions, all materials were sampled within 30×30 m^2. Samples were immediately frozen in liquid nitrogen to extract total RNA. For each group of samples, equal numbers of samples from five individual plants were pooled for the total RNA extraction with the TRIZOL reagent according to the manufacturer' instructions (Invitriogen). Transcriptome sequencing required the total RNA from the different pooled samples. After eliminating traces of genomic DNA through treatment with DNase I, an automated capillary gel electrophoresis was used to assess RNA quality and quantity using a Bioanalyzer 2100 with RNA 6000 Nano Labchips (Agilent Technologies Ireland, Dublin, Ireland). Transcriptome sequencing libraries were prepared using an Illumina Small RNA Sample Prep Kit and an Illumina TruSeq RNA Sample Prep Kit. After the three libraries had been prepared, raw reads generated using the Illumina Hiseq™ 2000 were initially processed to obtain clean reads. Then, all the clean reads were assembled using a *de novo* assembly program Trinity (Haas et al., 2013).

Expression pattern analysis of WRKY genes

An *in silico* analysis was used to compare the relative transcript abundance of the *WRKY* genes in stem, leaf and blossom buds, based on transcriptome data. The gene expression level was normalized to the values of RPKM (Reads Per kb per Million reads) (Wagner et al., 2012).

Phylogenetic Analysis

Phylogenetic and molecular evolutionary analyses were conducted using the sofeware of MEGA version 6. Complete WRKY predicted amino acids sequences was performed with a gap open penalty of 10 and gap extension penalty of 0.2 using ClustalW implemented in MEGA. Neighbour-joining method was used to construct phylogenetic tree by using, and the reliability of the obtained trees was tested using bootstrapping with 1000 replicates.

Conclusion

In the present study, the identification and bioinformatics analysis of willow *WRKY* genes at the whole genome level were conducted. A total of 89 *WRKY* genes were identified. We also investigated the gene structure, phylogenetic relationship, and sequence characteristics. As a conclusion, *in silico* genome wide survey has contributed to the understanding of gene structure and evolutionary relationship of the willow WRKY gene family. The results would be useful for the tree researchers to discovery new *WRKY* gene members in other tress, and useful for the next detailed functional studies.

Acknowledgements

This work was supported by grants from the National Natural Science Foundation of China (31400569), Collaborative Innovation Plan of Jiangsu Higher Education, the Fundamental Research Funds for the Central Non-profit Research Institution of CAF (CAFYBB2014QB028, RIF2013-11), Beijing Co-building Plan for Scientific Research and Postgraduate Education (2014K0151, 2014K0152), and we thank LetPub (www.letpub.com) for its assistance with the manuscript preparation.

References

Ashida Y, Nishimoto M, Matsushima A, Watanabe J, Hirata T (2002) Molecular cloning and mRNA expression of geraniol-inducible genes in cultured shoot primordia of Matricaria chamomilla. Biosci Biotech Bioch. 66 (11):2511-4.

Bailey TL, Elkan C (1994) Fitting a mixture model by expectation maximization to discover motifs in biopolymers. UCSD Technical Report. 2:28-36.

Berlin S, Lagercrantz U, von Arnold S, Ost T, Ronnberg-Wastljung AC (2010) High-density linkage mapping and evolution of paralogs and orthologs in Salix and Populus. BMC Genomics. 11:129.

Cai M, Qiu D, Yuan T, Ding X, Li H, Duan L, Xu C, Li X, Wang S (2008) Identification of novel pathogen-responsive cis-elements and their binding proteins in the promoter of OsWRKY13, a gene regulating rice disease resistance. Plant Cell Environ. 31 (1):86-96.

Cai Y, Chen X, Xie K, Xing Q, Wu Y, Li J, Du C, Sun Z, Guo Z (2014) Dlf1, a WRKY transcription factor, is involved in the control of flowering time and plant height in rice. PLos One. 9 (7):e102529.

Chen C, Chen Z (2000) Isolation and characterization of two pathogen- and salicylic acid-induced genes encoding WRKY DNA-binding proteins from tobacco. Plant Mol BIol. 42 (2):387-96.

Cormack RS, Eulgem T, Rushton PJ, Kochner P, Hahlbrock K, Somssich IE (2002) Leucine zipper-containing WRKY proteins widen the spectrum of immediate early elicitor-induced WRKY transcription factors in parsley. Gene Regul Mech. 1576 (1-2):92-100.

de Pater S, Greco V, Pham K, Memelink J, Kijne J (1996) Characterization of a zinc-dependent transcriptional activator from Arabidopsis. Nucleic Acids Res. 24 (23):4624-31.

Ding M, Chen J, Jiang Y, Lin L, Cao Y, Wang M, Zhang Y, Rong J, Ye W (2014) Genome-wide investigation and transcriptome analysis of the WRKY gene family in Gossypium. Mol Genet Genomics. 209 (1): 151-157.

Ding ZJ, Yan JY, Li GX, Wu ZC, Zhang SQ, Zheng SJ (2014) WRKY41 controls Arabidopsis seed dormancy via direct regulation of ABI3 transcript levels not downstream of ABA. Plant J. 79 (5):810-23.

Dou L, Zhang X, Pang C, Song M, Wei H, Fan S, Yu S (2014) Genome-wide analysis of the WRKY gene family in cotton. Mol Genet Genomics 289 (6):1103-21.

Eulgem T, Rushton PJ, Robatzek S, Somssich IE (2000) The WRKY superfamily of plant transcription factors. Trends Plant Sci. 5 (5):199-206.

Guan Y, Meng X, Khanna R, LaMontagne E, Liu Y, Zhang S (2014) Phosphorylation of a WRKY transcription factor by MAPKs is required for pollen development and function in Arabidopsis. PLos Genet. 10 (5):e1004384.

Haas BJ, Papanicolaou A, Yassour M, Grabherr M, Blood PD, Bowden J, Couger MB, Eccles D, Li B, Lieber M, Macmanes MD, Ott M, Orvis J, Pochet N, Strozzi F, Weeks N, Westerman R, William T, Dewey CN, Henschel R, Leduc RD, Friedman N, Regev A (2013) De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. Nat Protoc. 8 (8):1494-512.

He H, Dong Q, Shao Y, Jiang H, Zhu S, Cheng B, Xiang Y (2012) Genome-wide survey and characterization of the WRKY gene family in Populus trichocarpa. Plant Cell Rep. 31 (7):1199-217.

Hjältén J (1998) An experimental test of hybrid resistance to insects and pathogens using Salix caprea, S. repens and their F_1 hybrids. Oecologia. 177 (1-2):127-132.

Huang T, Duman JG (2002) Cloning and characterization of a thermal hysteresis (antifreeze) protein with DNA-binding activity from winter bittersweet nightshade, Solanum dulcamara. Plant Mol Biol. 48 (4):339-50.

Huang X, Madan A (1999) CAP3: A DNA sequence assembly program. Genome Res. 9 (9):868-77.

Huh SU, Choi LM, Lee GJ, Kim YJ, Paek KH (2012) Capsicum annuum WRKY transcription factor d (CaWRKYd) regulates hypersensitive response and defense response upon Tobacco mosaic virus infection. Plant Sci. 197:50-8.

Ishiguro S, Nakamura K (1994) 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 Gene. 244 (6):563-71.

Karp A, Hanley SJ, Trybush SO, Macalpine W, Pei M (2011) Genetic improvement of willow for bioenergy and biofuels. J Integr Plant Biol. 53: 151-165.

Kadis I (2005) Chosenia: an amazing tree of Northeast Asia. Arnoldia. 63 (3):8-17.

Kim CY, Lee SH, Park HC, Bae CG, Cheong YH, Choi YJ, Han C, Lee SY, Lim CO, Cho MJ (2000) Identification of rice blast fungal elicitor-responsive genes by differential display analysis. Mol Plant Microbe In. 13 (4):470-4.

Li HL, Guo D, Yang ZP, Tang X, Peng SQ (2014) Genome-wide identification and characterization of WRKY gene family in Hevea brasiliensis. Genomics. 104 (1):14-23.

Li J, Besseau S, Toronen P, Sipari N, Kollist H, Holm L, Palva ET (2013) Defense-related transcription factors WRKY70 and WRKY54 modulate osmotic stress tolerance by regulating stomatal aperture in Arabidopsis. New Phytol. 200 (2):457-72.

Ling J, Jiang W, Zhang Y, Yu H, Mao Z, Gu X, Huang S, Xie B (2011) Genome-wide analysis of WRKY gene family in Cucumis sativus. BMC Ggenomics. 12:471.

Ma D, Pu G, Lei C, Ma L, Wang H, Guo Y, Chen J, Du Z, Wang H, Li G, Ye H, Liu B (2009) 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. 50 (12):2146-61.

Marchive C, Mzid R, Deluc L, Barrieu F, Pirrello J, Gauthier A, Corio-Costet MF, 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. J Exp Bot. 58 (8):1999-2010.

Miao Y, Laun T, Zimmermann P & Zentgraf U (2004) Targets of the WRKY53 transcription factor and its role during leaf senescence in Arabidopsis. Plant Mol Biol. 55 (6):853-67.

Park CY, Lee JH, Yoo JH, Moon BC, Choi MS, Kang YH, Lee SM, Kim HS, Kang KY, Chung WS, Lim CO & Cho MJ (2005) WRKY group IId transcription factors interact with calmodulin. FEBS Lett. 579 (6):1545-50. New Phytol. 137 (2):303-314.

Rao G, Sui J, Zeng Y, He C, Duan A & Zhang J (2014) De novo transcriptome and small RNA analysis of two Chinese willow cultivars reveals stress response genes in Salix matsudana. PloS One. 9 (10).

Ren X, Chen Z, Liu Y, Zhang H, Zhang M, Liu Q, Hong X, Zhu JK & Gong Z (2010) ABO3, a WRKY transcription factor, mediates plant responses to abscisic acid and drought tolerance in Arabidopsis. Plant J. 63 (3):417-29.

Rushton PJ, Macdonald H, Huttly AK, Lazarus CM & Hooley R (1995) 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. 29 (4):691-702.

Scarpeci TE, Zanor MI, Mueller-Roeber B & Valle EM (2013) Overexpression of AtWRKY30 enhances abiotic stress tolerance during early growth stages in Arabidopsis thaliana. Plant Mol Biol. 83 (3):265-77.

Sharma MK, Kumar R, Solanke AU, Sharma R, Tyagi AK & Sharma AK (2010) Identification, phylogeny, and transcript profiling of ERF family genes during development and abiotic stress treatments in tomato. Mol Genet Genomics. 284 (6):455-75.

Sun C, Palmqvist S, Olsson H, Boren M, Ahlandsberg S & Jansson C (2003) A novel WRKY transcription factor, SUSIBA2, participates in sugar signaling in barley by binding to the sugar-responsive elements of the iso1 promoter. Plant Cell. 15 (9):2076-92.

Suttipanta N, Pattanaik S, Kulshrestha M, Patra B, Singh SK & Yuan L (2011) The transcription factor CrWRKY1 positively regulates the terpenoid indole alkaloid biosynthesis in Catharanthus roseus. Plant Physiol. 157 (4):2081-93.

Tang J, Wang F, Wang Z, Huang Z, Xiong A, Hou X (2013) Characterization and co-expression analysis of WRKY orthologs involved in responses to multiple abiotic stresses in Pak-choi (Brassica campestris ssp. chinensis). BMC Plant Biol. 13:188.

Tsarouhas V, Gullberg U, Lagercrantz U (2004) Mapping of quantitative trait loci (QTLs) affecting autumn freezing resistance and phenology in Salix. Theor Appl Genet. 108 (7):1335-42.

Ulker B, Somssich IE (2004) WRKY transcription factors: from DNA binding towards biological function. Curr Opin Plant Biol. 7 (5):491-8.

Wagner GP, Kin K, Lynch VJ (2012) Measurement of mRNA abundance using RNA-seq data: RPKM measure is inconsistent among samples. Theor Biosc. 131 (4):281-5.

Wang H, Avci U, Nakashima J, Hahn MG, Chen F, Dixon RA (2010) Mutation of WRKY transcription factors initiates pith secondary wall formation and increases stem biomass in dicotyledonous plants. Proc Natl Acad Sci USA. 107 (51):22338-43.

Wei KF, Chen J, Chen YF, Wu LJ, Xie DX (2012) Molecular phylogenetic and expression analysis of the complete WRKY transcription factor family in maize. DNA Res. 19 (2):153-64.

Wu KL, Guo ZJ, Wang HH, Li J (2005) The WRKY family of transcription factors in rice and Arabidopsis and their origins. DNA Res. 12 (1):9-26.

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 (2):124-32. .

Yoo SJ, Kim SH, Kim MJ, Ryu CM, Kim YC, Cho BH, Yang KY (2014) Involvement of the *OsMKK4-OsMPK1* Cascade and its Downstream Transcription Factor OsWRKY53 in the Wounding Response in Rice. Plant Pathology J. 30 (2):168-77.

Yu Y, Hu R, Wang H, Cao Y, He G, Fu C, Zhou G (2013) MlWRKY12, a novel Miscanthus transcription factor, participates in pith secondary cell wall formation and promotes flowering. Plant Sci. 212:1-9.

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