

Molecular phylogenetic characterization and analysis of the WRKY transcription factor family responsive to *Rhizoctonia solani* in maize

Jian Gao^{1†}, Hua Peng^{1,3†}, Xiujing He¹, Mao Luo², Zhe Chen¹, Haijian Lin¹, Haiping Ding¹, Guangtang Pan^{1*}, Zhiming Zhang^{1*}

¹Institute of maize Research, Key Laboratory of Biology and Genetic Improvement of maize in Southwest Region, Ministry of Agriculture, Sichuan Agricultural University, Wenjiang, 631000, Sichuan, People's Republic of China

²Drug Discovery Research Center of Luzhou Medical College, Luzhou Medical College, Luzhou, 656000, Sichuan, People's Republic of China

³Sichuan Tourism University, Chengdu, 610000, Sichuan, People's Republic of China

†These authors contributed equally to this work

*Corresponding author: E-mail: pangt1956@163.com, zzmmaize@gmail.com

Abstract

In this study we have identified, based on the maize genome, 85 WRKY genes that were phylogenetically clustered into three families formed by 8 distinct subfamilies. The exon/intron structures and motif compositions of these WRKY genes were highly conserved in each subfamily suggesting their functional conservation. Moreover, based on qTeller analyses, the majority of these WRKY genes showed a specific temporal and spatial expression pattern. These WRKY genes, within the same group, manifested a distinct expression, indicating a similar function in their expression during the evolutionary process; this is reflected by their sub-functionalizations in their expression pattern concerning leaf developmental gradient, while mature bundle sheath, and mesophyll cells had a similar cellular localization and modality of expression. This study has also provided evidence of the presence of a subset of WRKY genes exhibiting a putative functional role in leaf sheath when infected with *Rhizoctonia solani*. This finding appears helpful for future functional investigations to unravel the roles of WRKY genes in plant pathogen resistance. Interestingly, in this study we have identified three WRKY genes that are predicted to be potential targets of miR160 and miR171b families. Therefore, this finding appears relevant in elucidating the biological functions of these transcription factors to clarify the molecular mechanisms affecting leaf sheath growth and development during fungal infection and plant resistance.

Keywords: WRKY transcription factor, maize, phylogenetic analysis, expression profile, *R solani*

Introduction

Transcription factors (TFs) are proteins that regulate gene expression by binding to specific cis-acting promoter elements in controlling many important cellular processes during plants growth and development, including cellular morphogenesis, signal transduction, and environmental stress responses (Riechmann et al, 2000; Wray et al, 2003). In this context, based on bioinformatic analyses, fifty different families of TFs were identified (Riechmann et al, 2000; Riaño-Pachón et al, 2007; Pérez-Rodríguez et al, 2010). In plants, the WRKY TFs, are an integral component of the signaling networks that modulate many plant processes, and formed one of the largest families of transcriptional regulators represented by approximately 81 genes in *Arabidopsis*, 99 genes in rice, and 88 genes in *Fragaria vesca*. (Miao et al, 2012; Eulgem and Somssich, 2007; Ross et al, 2007). It was also found that the WRKY TFs are plant-specific and are characterized by N-terminal ends, containing a conserved WRKYGQR amino acids motif, formed by nearly 60 amino acid residues followed by a novel

zinc-finger-like motif such as C2H2 (C-X4-5-C-X22-23-H-X-H) or C2HC (C-X7- -X23-H-X-C). Further studies have provided evidence that the WRKY TFs regulate their target genes through the W-box elements, located in the promoter regions, by specifically binding to the (T)(T)TGAC(C/T) sequence (Eulgem et al, 2000a) Moreover, it was reported that they act in concert with other components of the transcriptional machinery and direct the temporal- and spatial-specific expression of the designated genes, thereby ensuring a proper cellular response to both internal and external stimuli. (Eulgem et al, 1999; Somssich, 2004; Ciolkowski et al, 2008). Further research have also indicated that the WRKY TFs are probably involved in plant defense mechanism responses following pathogen infection (Tian et al, 2006). In this respect, a number of plants after pathogen infection or treatment with pathogen elicitors or salicylic acid (SA) were found to induce a rapid expression of several WRKY genes (Eulgem et al, 2000b; Dellagi et al, 2000; Dong et al, 2003; Hara et al, 2000; Kim and Zhang, 2004). Additionally, it was shown that a number of defense-related genes, including the well-studied

pathogen related (PR) genes, containing W-box elements in their promoter regions (Eulgem et al, 1999; Du and Chen, 2000; Yang et al, 1999), are specifically recognized by WRKY proteins that appear necessary for the inducible expression of these genes. Thus, it can be argued that a single WRKY TF might provide activity on both abiotic and biotic stresses and cross talks with different signal transduction pathways. For instance, the rice WRKY45 (OsWRKY45) gene expression is markedly induced in response to abscissic acid (ABA treatments, various abiotic stress factors (e.g. NaCl and dehydration), and by pathogen attack attributable to *Pyricularia oryzae* Cav and *Xanthomonas oryzae* pv oryzae (Chen et al, 2012; Tao et al, 2009).

Maize (*Zea mays* L) is one of the most important agronomic crops in the world. The kernel provides feed, food, and a resource for many unique industrial and commercial products. However, this plant is frequently infected by various pathogenic fungi, such as *Rhizoctonia solani* Kühn (Peters et al, 2001), especially in Asia and southwest of China (Sharma and Saxena, 2002). These fungi cause the Banded leaf and sheath blight (BLSB) symptoms, that negatively affect both yield quality and quantity. To date, research in this field are focused to reveal the underlying molecular mechanism of the plant WRKY TFs in responses to pathogen infection, as well as to elucidate interactions in the pathways involved in response to pathogen attacks. The maize B73 genome sequence, recently published (Schnable et al, 2009), provides a good opportunity to study the WRKY genes. Therefore, in this research we have identified in the maize genome 85 WRKY genes, which were phylogenetic clustered into three families including eight distinct subfamilies. The exon/intron structure and motif compositions of WRKY genes were highly conserved in each subfamily, indicating their functional conservation. Based on qTeller analyses, the majority of WRKY genes herein identified showed a specific temporal and spatial expression pattern. These distinctive expression patterns, within the same group, are suggestive of a similar function in their expression during the evolutionary process. Their sub-functionalizations are reflected by a differential expression in leaf developmental gradient, while mature bundle sheath, and mesophyll cells showed similar cellular localization and expression modality. Finally, in the current study we have identified a subset of maize WRKY genes with putative functional roles in leaf sheaths infected by *R. solani*. Collectively the results herein presented appears helpful for future functional studies directed to unravel the roles of the WRKY genes involved in biotic resistance of plants to contrast pathogen infection.

Materials and Methods

Database search and sequence retrieval

Hidden Markov Model (HMM) profile of WRKY

domain (PF03106) downloaded from Protein family (Pfam; <http://pfam.sanger.ac.uk/>) was exploited for the identification of the WRKY genes from maize genome with HMMER (v 2.3.2) (Eddy, 2001). Conserved sequences of WRKY was extracted from the HMM profile by the HMMER software (Eddy, 2008), and then was adopted to query the B73 maize sequencing database (<http://www.maizesequence.org/index.html>). Searching parameters were followings: BLASTp, data-base-filtered gene sets (release 4a.53), E=1e⁻¹⁰, and other parameters were defaulted. All non-redundant hits with expected values less than 1.0 were collected and then compared with the WRKY family in PlnTFDB (<http://plntfdb.bio.unipotsdam.de/v3.0/>; Pérez-Rodríguez et al, 2010) and PlantTFDB (<http://plantfdb.cbi.pku.edu.cn>; Zhang et al, 2011). The re-annotated sequences were further manually analyzed to confirm the presence of WRKY domain with the SMART (<http://smart.embl-heidelberg.de/>), CDD (<http://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi>) and Inter-ProScA (<http://www.ebi.ac.uk/Tools/InterProScan/>) to examine domains of obtained sequences (Letunic et al, 2004; Marchler-Bauer et al, 2005; Hunter et al, 2009). Sequences of Arabidopsis and rice WRKY domain proteins were downloaded from the Arabidopsis genome TAIR 9.0 release (<http://www.Arabidopsis.org>) and rice genome annotation database (<http://rice.plantbiology.msu.edu/>, release 5.0), respectively.

Phylogenetic analysis

For phylogenetic analysis multiple sequence alignments of the full-length protein sequences, including the highly conserved N-terminal WRKY domain and the more divergent C-terminal activation domain, were performed by Clustal X (version 1.83) program (Larkin et al, 2007). The un-rooted phylogenetic trees were constructed with MEGA 4.0 by the Neighbor-Joining (NJ), Minimal Evolution (ME) and Maximum Parsimony (MP) methods, which carried out with 1000 iterations (Tamura et al, 2007). The protein sequences of Arabidopsis and rice WRKY transcription factors were obtained from the TIGR database, phylogenetic analysis was performed with MRBAYES 3.1.2 program (Ronquist and Huelsenbeck, 2003) by Bayesian method (Huelsenbeck et al, 2001), and the bootstrap test was carried out with 1,000,000 iterations.

Mapping WRKY genes on maize chromosomes

The maize databases were used in BLAST-based databases for the search of the entire maize genomic sequence to confirm the physical locations of all WRKY genes. The Genome Pixelizer software was used for a graphical display of the WRKY loci in each pair of corresponding maize chromosomes (<http://atgc.org/GenomePixelizer/41>).

Genomic structure and chromosomal location

Gene structure display server (GSDS) program (Guo et al, 2007) was used to illustrate exon/intron organization for individual WRKY genes by comparing

the cDNAs with their corresponding genomic DNA sequences from maize sequences (<http://www.maizesequence.org/>). Blocks of the same color represent the homologous chromosome segments. The tandem gene duplications in maize were identified according to the same criteria described in rice (Ouyang et al, 2007). Genes separated by five or fewer gene loci in a range of 100 kb distance were considered to be tandem duplicates.

Identification of conserved motifs

The program MEME, version 4.3.0, was used for the elucidation of motifs in 85 deduced Zea mays WRKY protein sequences (<http://meme.sdsc.edu>; Bailey et al, 2006). MEME was run locally with the following parameters: number of repetitions - any; maximum number of motifs - 20; and the optimum motif widths were constrained to between 6 and 200 residues. Structural motif annotation was performed with SMART (<http://smart.embl-heidelberg.de>; Letunic et al, 2004) and Pfam (<http://pfam.sanger.ac.uk> databases; Finn et al, 2006).

Prediction of potential targets for small RNA

Putative small RNA target sites were searched by using the miRanda software, which is an algorithm for finding genomic targets for microRNAs.

RNA profiling and depth analysis of leaf sheath responsive to *R. solani* in maize

For the expression analysis of ZmWRKY genes, our private available Genome expression profile data of the maize sheath infected by *R. solani* was used (data unpublished). Single gene expression was predicted online by qTeller (http://qteller.com/qteller3/generate_figures.php; Buell et al, 2011), which is a

tool to turning genome wide RNA-seq datasets into expression data on your favorite gene or genes. In addition, their differentiated expression in leaf developmental gradient, and in mature bundle sheath and mesophyll cells were predicted by maize maize eFP Browser (http://bar.utoronto.ca/efp_maize/cgi-bin/efpWeb.cgi; Li et al, 2010)

Results and Discussion

Identification and chromosome localization of maize WRKY genes

A key search against the NCBI and UniProt protein sequence databases have previously identified 34 and 31 of annotated maize WRKY protein sequences, respectively, while the searching for protein sequences in Phytozome have identified 32 annotated maize WRKY proteins. In this respect, we have greatly appreciated the foregoing efforts made by Dao-Xin Xie and coworkers (Wei et al, 2012), who have located these sequences in chromosome and in the association group orders. However, in this analysis it was only predicted the protein sequence based on maize genome, without identifying the WRKY genes from the maize genome sequences. Therefore, we have extended this analysis by Hidden Markov Model (HMM) to profile of WRKY domain (PF03106), moreover, we have examining domains of obtained sequences through the SMART, CDD and Inter-ProScan to get further clear data. A useful filter strategy was subsequently applied to avoid unclear results. Thus, in this investigation the WRKY genes were identified and validated through four steps: i) conserved sequences of WRKY proteins extracted from the Pfam database were firstly used to query the

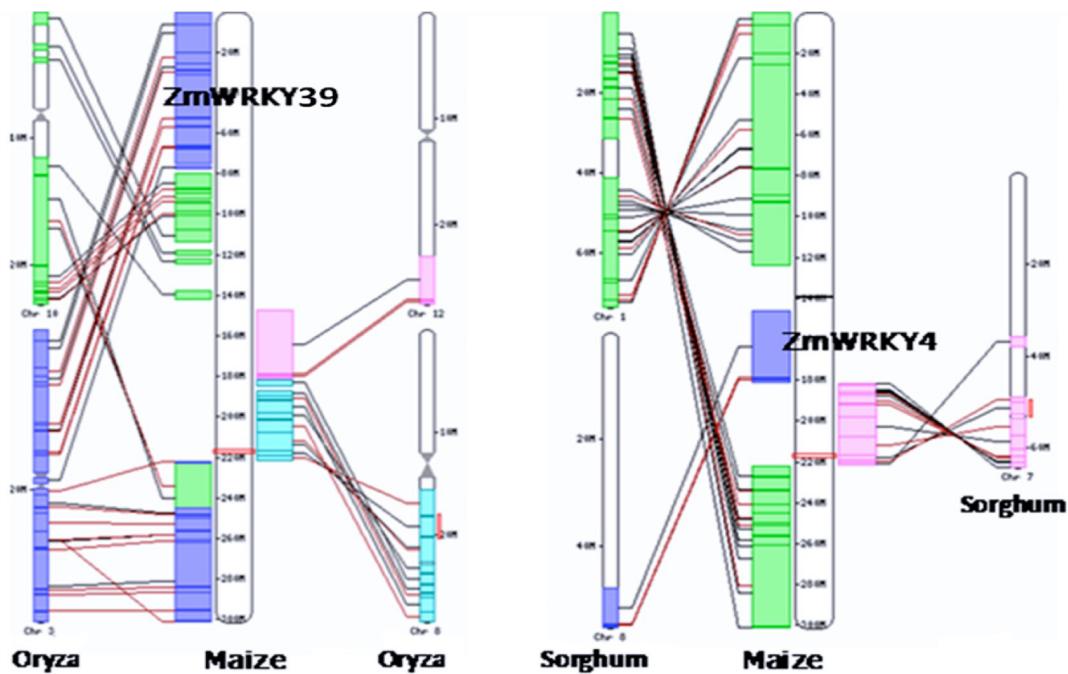


Figure 1 - Micro-colinearity of maize and sorghum, Oryza duplicate chromosome blocks containing WRKY.

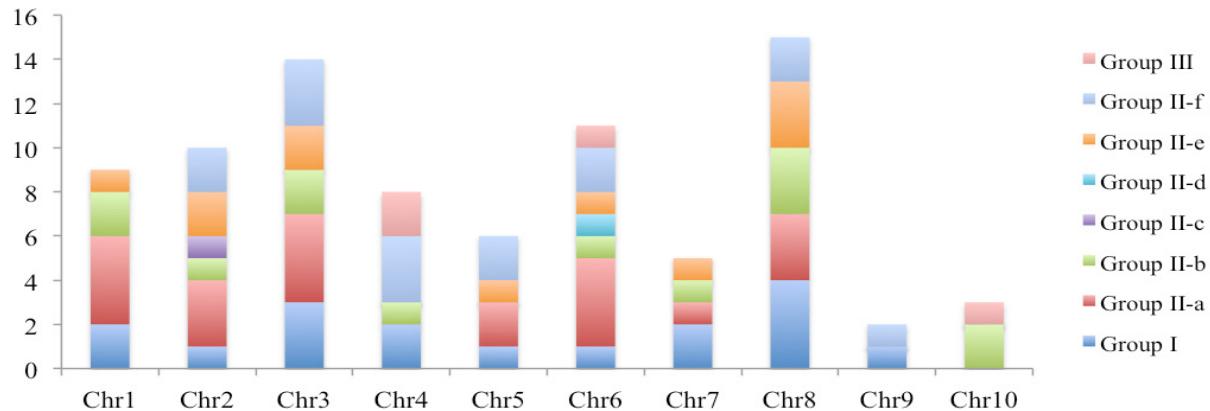


Figure 2 - Histogram showing the number and distribution of WRKY genes of eight subgroups on 10 chromosomes.

B73 maize-sequencing database with low stringency ($E=1.0$). Moreover, the NCBI database was searched for additional maize WRKY genes that were missed; ii) domains of obtained sequences were employed to develop the maize-specific HMM profile, which was adopted for the following data mining; iii) four WRKY genes that represented the genetic diversity of maize WRKY family were used to search for other maize WRKY genes; iv) all obtained sequences were examined on domains of WRKY proteins. Additionally, sequences that are not presented in the B73 maize filtered gene sets (*Zea mays* (AGPv2) release 5b.60) were eliminated in our analysis. In this study, only the most conserved transcripts, namely the transcript with the lowest e-value of domains confirmation was selected. We filtered and renamed the 85 genes sequences that we have identified through an overall search of the complete genome sequence from Zm-WRKY1 to ZmWRKY85 based on the exact position of their corresponding genes on chromosomes 1 to 10 from top to bottom, including the variant proteins produced from the same locus. (Supplementary Table 1)

As a good plant for experiments, sorghum, together with monocotyledonous plants rice was used to assign orthology for maize genes, which are valuable tools for functional analysis of the WRKY family members in maize (Supplementary Table 2). The putative orthologues were identified according to their e-value (under $1e^{-20}$) and the topology of phylogenetic tree. In addition, to screen the putative WRKY homologs, we specially investigated the micro-colinearity of maize/sorghum and maize/rice duplicate chromosomal blocks containing the WRKY homologs identified in this study. Results showed that genomic regions in sorghum chromosome 6, containing WRKY gene Sb01g000696, were syntenic with two maize genomic blocks: one in maize chromosomes 6 containing Sb01g000696 ortholog named Zm-WRKY4, the other in maize chromosomes 9 containing Sb01g000696 ortholog to ZmARF39 (Figure 1).

To map the 85 WRKY genes to the maize chromosomes, the physical location of each gene was required. The physical distribution of WRKY genes in the maize B73 genome illustrates the genetic events that result in the diversity and complexity of this gene family (Figure 1). The study of chromosome map, along with the histogram, suggests that the WRKY genes are dispersedly distributed across all the chromosomes in the maize genome. It is evident that the highest number of WRKY genes is present on Chromosome 8 (14 genes), representing 17.71% of the total. The least number of WRKY genes is located on chromosome 9 which contain only two genes belonging to eight different groups, accounting for just 1.265% of the total. (Figure 2)

Phylogenetic analysis and genomic structure of ZmWRKYs

Bayesian phylogenetic analysis was also performed and the 85 ZmWRKY proteins were classified into three classes: class I, class II and class III, including eight subgroups: Group I, Group II-a- II f, Group III which belong to the class I, class II and class III, respectively. Specifically class I contains two conserved WRKY amino acid signature, class II is formed of WRKY amino acid signature and C2H2 Zinc-finger, which was further divided into 5 groups, on the bases of the their construction of MEME, class III shares with WRKY amino acid signature and C2HC Zinc-finger (Supplementary Figure 1)

It is worthy to note that, in the joint phylogenetic tree, most of the ZmWRKY proteins fell as related sister pairs, such as in Group I, ZmWRKY63 and 17, ZmWRKY77 and 49, or triple (ZmWRKY24, 73 and 58) and quadruplets in the case of ZmWRKY44, 48, 71, and 80; similar condition has been noted in the other group (Figure 3). The result of phylogenetic analysis implies that domain gain and loss is a divergent force for expansion of the WRKY gene family. All previously isolated WRKY proteins contain the WRKYGQK sequence in their DNA-binding zinc motifs (Ishiguro and Nakamura, 1994; Hara et al, 2000; Dong et al, 2003).

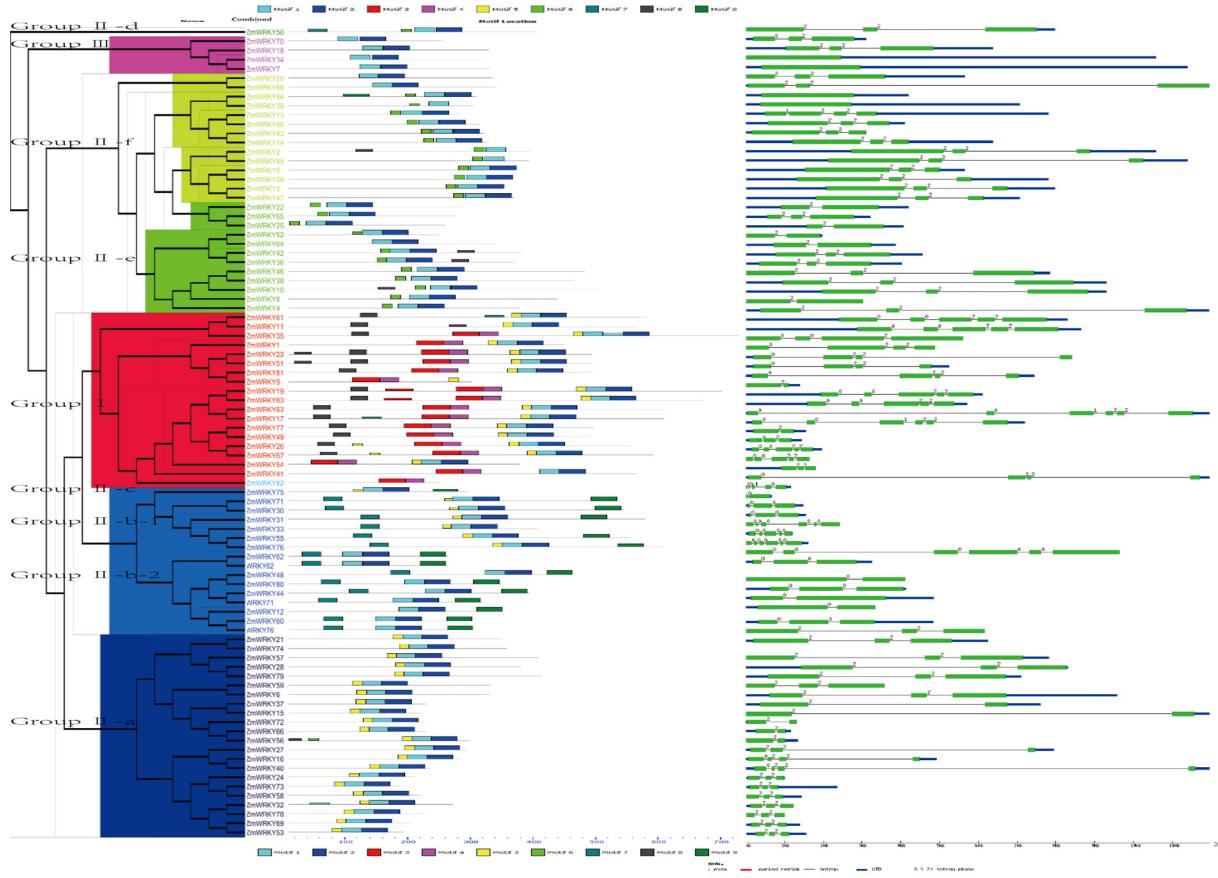


Figure 3 - Phylogenetic relationships, gene structure and motif compositions of maize WRKY genes. A. Multiple alignments of 85 full-length amino acids of WRKY genes from Maize were executed by Clustal X 1.83 and the phylogenetic tree was constructed using MEGA 4.0 by the Neighbor-Joining (NJ) method with 1,000 bootstrap replicates. The percentage bootstrap scores higher than 50% are indicated on the nodes. The six major phylogenetic families are marked with different color backgrounds. B. Schematic representation of the conserved motifs in the WRKY proteins from Maize elucidated by MEME. A number in the colored box represents each motif. The black lines represent the non-conserved sequences. C. Exon/intron structures of WRKY genes from Maize. Green boxes and black lines represent exons and introns, respectively. The sizes of exons and introns can be estimated using the scale at bottom.

Although the WRKYGQK peptide is highly conserved, nine variants with one or two amino acids substitution were observed in 7 domains belong to Groups IIa (Figure 3). While WRKYGKK is the only common variant shared by seven (all in Group II a) domains, it was also found that class II and class III had a single WRKY protein. It has been reported that WRKY transcription factors have their evolutionary origin in ancient eukaryotes with the most basal WRKY genes identified in the unicellular protist, *Guardia lamblia*, and in the slime mold *Dictyostelium discoideum* (Wu et al, 2005). Additionally, it was demonstrated that the proteins of this group might have evolve early and represent the ancestral form because of the two WRKY domains. Moreover, a large number of WRKY proteins exist in the three above mentioned species, suggesting that these proteins play a crucial role in plant developmental and physiological processes. (Babu et al, 2006). These evidence suggested that the rapid duplication of WRKY genes occurred before the divergence of monocots and dicots. (Wu

et al, 2005). In the current study, in class III, eight WRKYS of *Arabidopsis* were found to divergence in comparison with 24 WRKYS of rice. This implies that monocots have been subjected to a rapid duplication of the WRKY genes that occurred before their divergence of dicots. In addition, we found that rice has owned the special group (group II-g) and contain 9 WRKYS with zf-BED superfamily (PF02892), which is an about 50 to 60 amino acid residues domain that contains a characteristic motif with two highly conserved aromatic positions, as well as a shared pattern of cysteines and histidines that is predicted to form a zinc finger.

The intron/exon structures of ZmWRKY genes were determined by alignment of cDNA to genomic sequences. This sequence analysis revealed that introns were found in coding sequences of all the WRKY genes. Of which, the number of exons varied from 2 to 12. As expected, most ZmWRKY genes in the same sister pair or triplets showed similar distribution of intron/exon, whereas the others were more

divergent in genomic structure, showing that these sister pair genes lies in duplicated genomic regions. To further reveal the diversification of WRKY genes in maize, putative motifs were predicted in the MEME (Multiple Expectation Maximization for Motif Elicitation) program. In addition, most of the closely related members in the phylogenetic tree shared common motif compositions, indicating functional similarities among the WRKY proteins within the same subfamily (Figure 3).

The annotated WRKY gene family in Arabidopsis and rice enabled us to determine the phylogenetic relationship between dicot and monocot WRKY proteins. A phylogenetic tree construction, using the protein sequences of 85 ZmWRKYs, 99 OsWRKYs and 81 AtWRKYs, respectively, which depicted altogether 265 WRKY proteins, were also divided into three classes (Figure 4). Moreover, the similar gene structures and conserved motifs of WRKY genes in the same subfamilies may provide additional supports to the phylogenetic analysis. On the other hand, the differences among gene organizations and the divergences in motif compositions among different subfamilies may also indicate that maize WRKYs are

functionally diversified. However, the biological significance of most of the putative motifs remains to be elucidated because they do not have homologs when searching against Pfam and SMART (Simple Modular Architecture Research Tool) databases.

Prediction of potential targets for small RNA

Previous studies have revealed that WRKY TFs are a large family of regulatory proteins forming a network in defense signaling.(Eulgem and Somsich, 2007) that is involved in various plant processes and most notably in coping with diverse biotic and abiotic stresses by acting on targets through small RNAs. Predicted targets for several miRNAs are encoded WRKY factors (Zhang et al, 2008; Pandey and Somsich, 2009), suggesting smRNA-mediated regulation of WRKY TFs. It was also reported that osa-miRNA396-like is induced and its target genes were predicted to encode a WRKY domain protein in maize. However, the direct predicted targets of several miRNAs encoding WRKY factors are poorly known. In this study, using the miRanda software searched putative small RNA target sites. Out of these, only three WRKY genes were predicted to be the potential targets of small RNA and the number of target

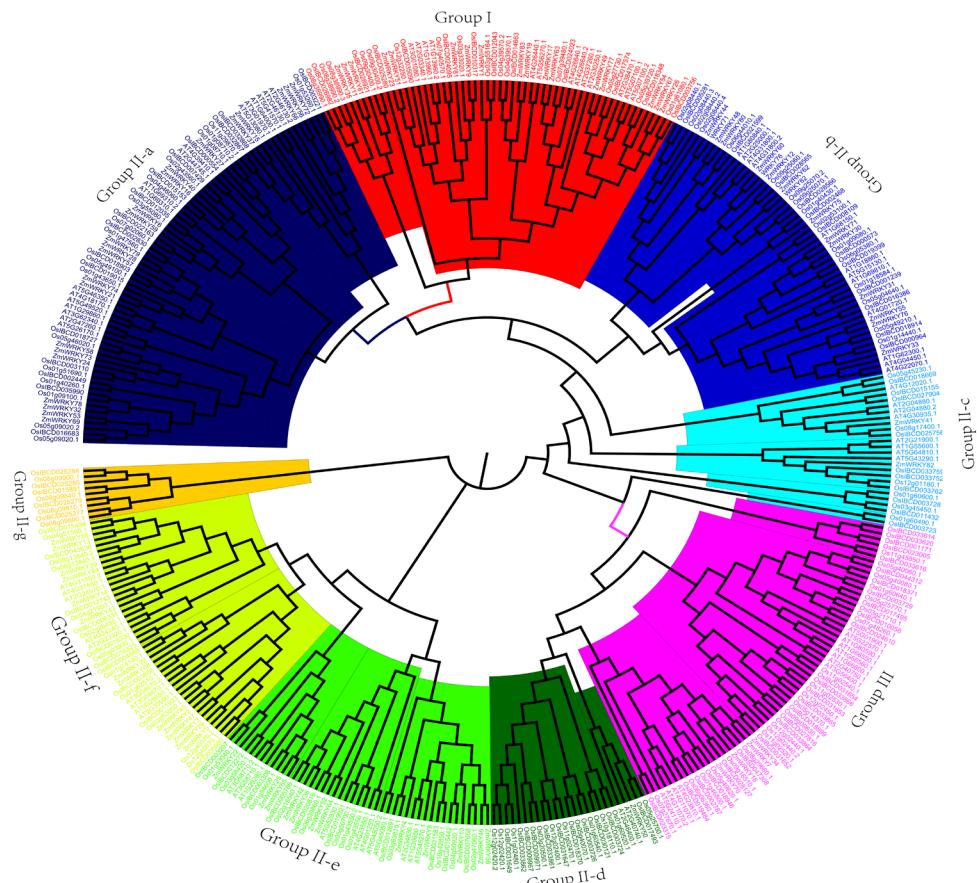


Figure 4 - Joined phylogenetic trees of WRKY domain-containing proteins from Maize, Arabidopsis and rice. The deduced full-length amino acid sequences of 85 maize, 81 Arabidopsis and 99 rice WRKY genes were aligned by Clustal X 1.83 and the phylogenetic tree was constructed using MEGA 4.0 by the Neighbor-Joining (NJ) method with 1,000 bootstrap replicates. Each WRKY subfamily was indicated in a specific color.

Table 1 - Prediction of potential targets for small RNA by miRanda.

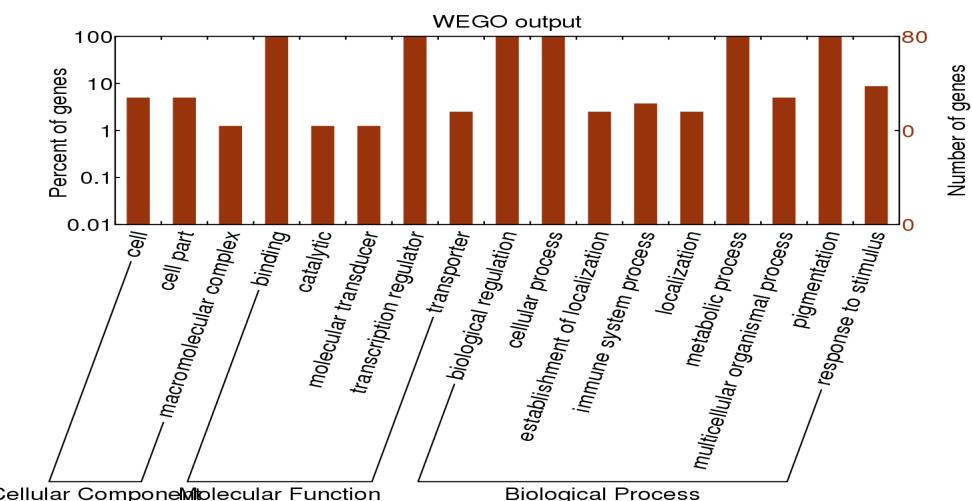
Zma-miR171b		
ZmWRKY13:	1855bp	TTGAGGCGTGAGAGTCATCA 1874bp
Zma-miR171b:		: : :
		3'-UUGAGCCGUGCCAAU-AUCA-5'
Zma-miR160 family		
ZmWRKY60:	9884bp	TGCCTGGCGCCGTCACGCCG 9904bp
ZmWRKY80:	7015bp	TGGCTTGGCCCCCTGAAGGGGC 7033bp
Zma-miR160c:		: : :
Zma-miR160b:		UGCCUGGCUCCCUGUAUGCCA
Zma-miR160f:		UGCCUGGCUCCCUGUAUGCCA
Zma-miR160a:		UGCCUGGCUCCCUGUAUGCCA
Zma-miR160d:		UGCCUGGCUCCCUGUAUGCCA

genes for miR160 family and miR171b was 2 and 1, respectively (Table 1). The importance of smRNAs is increasingly becoming important in plant processes to response to abiotic stresses and the endogenous plant-derived smRNAs probably have broad implications in post-transcriptional regulating plant responses to pathogen attack (Navarro et al, 2006; Brodersen et al, 2008; Pandey and Baldwin, 2008). For example, in rice phytohormone treatments were shown to induce the expression of several miRNAs (Tao et al, 2009); moreover several miRNAs have been predicted for encoding WRKY factors, suggesting smRNA-mediated regulation of WRKY TFs. Conversely, it was found that several miRNA gene promoters are highly abundant in W box sequences, implicating WRKY TFs in their activation/repression(Tao et al, 2009).

Expression analysis of WRKY factors in global transcriptome at different developmental stages and in specific organs

Research has revealed multiple roles of WRKY factors in response to abiotic stresses, including drought and salt, which are regarded as ancestral roles of WRKY proteins (Singh et al, 2002) Additionally, they also play multiple roles in response to biotic stresses such as bacteria and fungi. RNA-seq is one of the useful global transcriptome analysis technolo-

gies, which provides us an opportunity to understand the patterns of gene expression. In this study, qTeller has collected all the RNA-seq data currently available, which provided valuable resources for gene discovery and functional characterization. In a previous study (unpublished data of this laboratory), we have found 54 genes associated with BLSB-inoculated plants Those data were used herein to mine gene expression data in 30 specific organs in maize during BLSB infection as shown in **Supplementary Figure 2**. It can be noted from the heat map that all of the 54 detected transcripts are involved in many biological processes and are expressed in all tissues, although to different extent. In fact most of the genes appear to be invariable and lowly expressed among all tissues. Interestingly, ZmWRKY63 and ZmWRKY45 were the most stable and are highly expressed across maize organs. Additionally, structural studies indicated that this domain is a four-stranded beta-sheet with a zinc-binding pocket, which forms a novel zinc and DNA binding structure. The WRKYGQK residues correspond to the most N-terminal beta-strand, which enables extensive hydrophobic interactions, contributing to the structural stability of the beta-sheet. The stable gene expression across all tissues can be regarded as constitutive expression. From this we can infer that many maize WRKY genes were expressed at low level, which may work synergistically with other family of proteins during plant growth and development. Gene ontology annotation of differentially expressed genes (DEG) showed that these DEGs are involved in biological process, such as response to stimulus, signal transmission, and molecular function, such as catalytic and transcriptional regulation (Figure 5). In our study, we have mainly focused our analysis on the spatial and temporal specific expression patterns of maize WRKY genes to identify their roles in developmental regulation, as the expression

**Figure 5** - GO annotation of differentially expressed genes association with BLSB treatment.

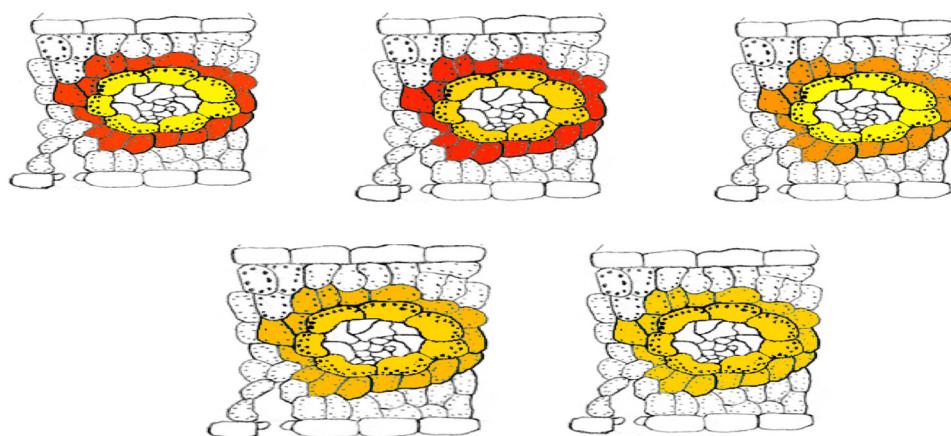


Figure 6 - The spatial and temporal specific expression patterns of gradual increasing BLSB responsive genes.

of the WRKY genes was detected in a wide variety of plant species and is involved in plant growth and development. Thus in this study the expression patterns of maize WRKY genes were investigated with SBS data. The results showed that most of ZmWRKY genes are transcribed in checked tissues and organs (Supplementary Table 3). The remaining three transcripts with no detectable expression signal were ZmWRKY85, ZmWRKY37, and ZmWRKY59.

Finally we have mined RNA-seq data, which recorded the gene expression levels of 30 distinct tissues representing 11 major organs and various developmental stages of the maize plant (Supplementary Table 4). It was shown that the majority of WRKY genes exhibited a specific temporal and spatial expression pattern. These distinct expressions within the same group suggested a similar function in their expression during the evolutionary process. The spatial and temporal specific expression patterns of gradual increasing BLSB responsive genes were analysis by maize eFP Browser (http://bar.utoronto.ca/efp_maize/cgi-bin/efpWeb.cgi). It was observed that most of the genes were differentially expressed along the developmental gradient and between bundle sheath and mesophyll cells, respectively. It is interesting that the increasing expression genes showed distinct spatial and temporal expression patterns. Altogether, these findings highlight the importance of WRKY factors in transcriptionally reprogramming plant responses toward different invading pathogens. Some appear to positively influence the outcome of such plant-pathogen interactions, while others appear to act negatively (Figure 6). This negative influence may be due to active targeting of the WRKY genes/factors, or products under their control, by certain pathogens. Manipulation of WRKY proteins by pathogen effectors may partly explain the existence of redundancy within the WRKY TF family as a reinforcement measure for essential regulatory functions. Coordinated modulation of positive- and negative-acting factors could also enable the proper amplitude and duration of the plant response during

pathogen attack.

Conclusion

In present study, a comprehensive analysis of WRKY gene family in maize was performed, including phylogeny, chromosomal location, gene structure, conserved motifs, and expression profiling. We identified a sum of 85 WRKY genes phylogenetically clustered into three families with 24 distinct subfamilies in maize genome. The exon/intron structure and motif compositions of WRKYs were highly conserved in each subfamily, indicating their functional conservation. A majority of these WRKY genes showed specific temporal and spatial expression patterns based on qTeller analyses, their distinct expression in the same group, indicating a similar function association with their expression during the evolutionary process. Further, the differentiated expression of sub-functionalization in leaf developmental gradient, mature bundle sheath and mesophyll cells, indicating a similar cellular localization and expression mode. By means of the RNA-seq based data mining and homologous analysis, we can obtain much useful information about the putative functions of the WRKYs in maize sheath infected by *R. solani*. In addition, three WRKY genes were predicted to be the potential targets of miR160 and miR171 families. It will be of great importance to elucidate the biological functions of these WRKY TFs, which could provide us deeper understanding in molecular mechanisms of sheath leaf infected growth and development and resistance response.

References

- Babu MM, Iyer LM, Balaji S, Aravind L. 2006. The natural history of the WRKY-GCM1 zinc fingers and the relationship between transcription factors and transposons. *Nucleic Acids Res* 34: 6505-6520
- Bailey TL, Williams N, Misleh C, Li WW. 2006. MEME: discovering and analyzing DNA and protein sequence motifs. *Nucleic Acids Res* 34: W369-W373

Brodersen P, Sakvarelidze-Achard L, Bruun-Rasmussen M, Dunoyer P, Yamamoto YY, Sieburth L, Voinnet O, 2008. Widespread translational inhibition by plant miRNAs and siRNAs. *Science's STKE* 320: 1185

Buell CR, Kaeppler SM, de Leon N, Vaillancourt B, Sekhon RS, Childs KL, Hansey CN, Davidson RM, Jiang N, Lin H, 2011. Utility of RNA sequencing for analysis of maize reproductive transcriptomes. *The Plant Genome* 4: 191-203

Chen L, Song Y, Li S, Zhang L, Zou C, Yu D, 2012. The role of WRKY transcription factors in plant abiotic stresses. *Biochimica et Biophysica Acta (BBA)-Gene Regulatory Mechanisms* 1819(2): 120-128

Ciolkowski I, Wanke D, Birkenbihl RP, Somssich IE, 2008. Studies on DNA-binding selectivity of WRKY transcription factors lend structural clues into WRKY-domain function. *Plant Mol Biol* 68: 81-92

Dellagi A, Heilbronn J, Avrova AO, Montesano M, Palva ET, Stewart HE, Toth IK, Cooke DEL, Lyon GD, Birch PRJ, 2000. A potato gene encoding a WRKY-like transcription factor is induced in interactions with *Erwinia carotovora* subsp. *atroseptica* and *Phytophthora infestans* and is coregulated with class I endochitinase expression. *Molecular plant-microbe interactions* 13: 1092-1101

Dong J, Chen C, Chen Z, 2003. Expression profiles of the Arabidopsis WRKY gene superfamily during plant defense response. *Plant Mol Biol* 51: 21-37

Du L, Chen Z, 2000. Identification of genes encoding receptor like protein kinases as possible targets of pathogen and salicylic acid induced WRKY DNA binding proteins in Arabidopsis. *Plant J* 24: 837-847

Eddy SR, 2001. {HMMER: Profile hidden Markov models for biological sequence analysis}.

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

Eulgem T, Rushton PJ, Schmelzer E, Hahlbrock K, Somssich IE, 1999. Early nuclear events in plant defence signalling: rapid gene activation by WRKY transcription factors. *EMBO J* 18: 4689-4699

Eulgem T, Somssich IE, 2007. Networks of WRKY transcription factors in defense signaling. *Curr Opin Plant Biol* 10: 366-371

Finn RD, Mistry J, Schuster-Beckler B, Griffiths-Jones S, Hollich V, Lassmann T, Moxon S, Marshall M, Khanna A, Durbin R, 2006. Pfam: clans, web tools and services. *Nucleic Acids Res* 34: D247-D251

Guo AY, Zhu QH, Chen X, Luo JC, 2007. GSDB: a gene structure display server]. Yi chuan= Hereditas/Zhongguo yi chuan xue hui bian ji 29: 1023

Hara K, Yagi M, Kusano T, Sano H, 2000. Rapid systemic accumulation of transcripts encoding a tobacco WRKY transcription factor upon wounding. *Mol Gen Genetics* 263: 30-37

Huelsenbeck JP, Ronquist F, Nielsen R, Bollback JP, 2001. Bayesian inference of phylogeny and its impact on evolutionary biology. *Science* 294: 2310-2314

Hunter S, Apweiler R, Attwood TK, Bairoch A, Bateman A, Binns D, Bork P, Das U, Daugherty L, Duquerre L, 2009. InterPro: the integrative protein signature database. *Nucleic Acids Res* 37: D211-D215

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 Genet* 244: 563-571

Kim CY, Zhang S, 2004. Activation of a mitogen-activated protein kinase cascade induces WRKY family of transcription factors and defense genes in tobacco. *Plant J* 38: 142-151

Larkin M, Blackshields G, Brown N, Chenna R, McGettigan P, McWilliam H, Valentin F, Wallace I, Wilm A, Lopez R, 2007. Clustal W and Clustal X version 2.0. *Bioinformatics* 23: 2947-2948

Letunic I, Copley RR, Schmidt S, Ciccarelli FD, Doerks T, Schultz J, Ponting CP, Bork P, 2004. SMART 4.0: towards genomic data integration. *Nucleic Acids Res* 32: D142-D144

Li P, Ponnala L, Gandotra N, Wang L, Si Y, Tausta SL, Kebrom TH, Provart N, Patel R, Myers CR, 2010. The developmental dynamics of the maize leaf transcriptome. *Nature Genetics* 42: 1060-1067

Marchler-Bauer A, Anderson JB, Cherukuri PF, DeWeese-Scott C, Geer LY, Gwadz M, He S, Hurwitz DI, Jackson JD, Ke Z, 2005. CDD: a Conserved Domain Database for protein classification. *Nucleic Acids Res* 33: D192-D196

Miao L-X, Zhang Y-C, Yang X-F, Jiang G-H, 2012. Genome-wide identification and analysis of WRKY transcription factors In *Fragaria vesca*. *Journal of Nuclear Agricultural Sciences* 8: 1124-1131

Navarro L, Dunoyer P, Jay F, Arnold B, Dharmasiri N, Estelle M, Voinnet O, Jones JDG, 2006. A plant miRNA contributes to antibacterial resistance by repressing auxin signaling. *Science's STKE* 312: 436

Ouyang S, Zhu W, Hamilton J, Lin H, Campbell M, Childs K, Thibaud-Nissen F, Malek RL, Lee Y, Zheng L, 2007. The TIGR rice genome annotation resource: improvements and new features. *Nucleic Acids Res* 35: D883-D887

Pandey SP, Baldwin IT, 2008. Silencing RNA directed RNA polymerase 2 increases the susceptibility of *Nicotiana attenuata* to UV in the field and in the glasshouse. *Plant J* 54: 845-862

Pandey SP, Somssich IE, 2009. The role of WRKY transcription factors in plant immunity. *Plant Physiol* 150: 1648-1655

Pérez-Rodríguez P, Riaño-Pachón DM, Corrêa LGG, Rensing SA, Kersten B, Mueller-Roeber B, 2010.

PInTFDB: updated content and new features of the plant transcription factor database. *Nucleic Acids Res* 38: D822-D827

Peters FR, Datnoff L, Korndörfer G, Seibold K, Rush M, 2001. Effect of silicon and host resistance on sheath blight development in rice. *Plant Dis* 85: 827-832

Riaño-Pachón DM, Ruzicic S, Dreyer I, Mueller-Roeber B, 2007. PInTFDB: an integrative plant transcription factor database. *BMC Bioinformatics* 8: 42

Riechmann J, Heard J, Martin G, Reuber L, Keddie J, Adam L, Pineda O, Ratcliffe O, Samaha R, Creelman R, 2000. Arabidopsis transcription factors: genome-wide comparative analysis among eukaryotes. *Science* 290: 2105-2110

Ronquist F, Huelsenbeck JP, 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* 19: 1572-1574

Ross CA, Liu Y, Shen QJ, 2007. The WRKY gene family in rice (*Oryza sativa*). *J Integr Plant Biol* 49: 827-842

Schnable PS, Ware D, Fulton RS, Stein JC, Wei F, Pasternak S, Liang C, Zhang J, Fulton L, Graves TA, 2009. The B73 maize genome: complexity, diversity, and dynamics. *Science* 326: 1112-1115

Sharma G, Saxena S, 2002. Integrated management of banded leaf and sheath blight of maize (*Zea mays* L) caused by *Rhizoctonia solani* (Kuhn). *Curr Adv Plant Sci* 15: 107-114

Singh KB, Foley RC, Oñate-Sánchez L, 2002. Transcription factors in plant defense and stress responses. *Curr Opin Plant Biol* 5: 430-436

Somssich IE, 2004. WRKY transcription factors: from DNA binding towards biological function. *Curr Opin Plant Biol* 7: 491-498

Tamura K, Dudley J, Nei M, Kumar S, 2007. MEGA4: molecular evolutionary genetics analysis (MEGA) software version 4.0. *Mol Biol Evol* 24: 1596-1599

Tao Z, Liu H, Qiu D, Zhou Y, Li X, Xu C, Wang S, 2009. A pair of allelic WRKY genes play opposite roles in rice-bacteria interactions. *Plant Physiol* 151: 936-948

Tian Y, Lu XY, Peng LS, Fang J, 2006. The structure and function of plant WRKY transcription factors]. *Yi chuan= Hereditas/Zhongguo yi chuan xue hui bian ji* 28: 1607

Wei K-F, Chen J, Chen Y-F, Wu L-J, Xie D-X, 2012. Molecular phylogenetic and expression analysis of the Complete WRKY transcription factor family in maize. *DNA Res* 19(2): 153-64

Wray GA, Hahn MW, Abouheif E, Balhoff JP, Pizer M, Rockman MV, Romano LA, 2003. The evolution of transcriptional regulation in eukaryotes. *Mol Biol Evol* 20: 1377-1419

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: 9-26

Yang P, Chen C, Wang Z, Fan B, Chen Z, 1999. A pathogen and salicylic acid induced WRKY DNA binding activity recognizes the elicitor response element of the tobacco class I chitinase gene promoter. *Plant J* 18: 141-149

Zhang H, Jin J, Tang L, Zhao Y, Gu X, Gao G, Luo J, 2011. PlantTFDB 2.0: update and improvement of the comprehensive plant transcription factor database. *Nucleic Acids Res* 39: D1114-D1117

Zhang Z, Wei L, Zou X, Tao Y, Liu Z, Zheng Y, 2008. Submergence-responsive microRNAs are potentially involved in the regulation of morphological and metabolic adaptations in maize root cells. *Ann Bot* 102: 509-519

Table S1: List and annotations of WRKY proteins in maize

Name	Chromosome	RefSeq	Ensemble Acc	Protein name in this article	Description
GRMZM2G127064	6	#/N/A	GRMZM2G127064_P01	Group II-f	Transcription factor activity
GRMZM2G063880	8	#/N/A	GRMZM2G063880_P01	Group I	Transcription factor activity
GRMZM2G163418	2	NP_001147748	GRMZM2G163418_P01	Group II-b	WRKY74
AC205562.3_FG002	4	#/N/A	AC205562.3_FGP002	Group II-f	Transcription factor activity
GRMZM2G324999	1	NP_001150829	GRMZM2G324999_P01	Group I	WRKY69
AC165171.2_FG002	3	#/N/A	AC165171.2_FGP002	Group I	Hypothetical protein LOC100191206
GRMZM2G045560	8	#/N/A	GRMZM2G045560_P01	Group II-f	Transcription factor activity
GRMZM2G083717	1	NP_001149833	GRMZM2G083717_P01	Group II-a	WRKY transcription factor 14
GRMZM2G383594	1	#/N/A	GRMZM2G383594_P01	Group II-b	Transcription factor activity
GRMZM2G024898	2	#/N/A	GRMZM2G024898_P01	Group II-f	Transcription factor activity
GRMZM2G138683	4	NP_001167966	GRMZM2G138683_P01	Group II-f	Hypothetical protein LOC100381682
GRMZM5G823157	5	#/N/A	GRMZM5G823157_P01	Group I	Transcription factor activity
GRMZM2G038158	4	NP_001130964	GRMZM2G038158_P01	Group III	Hypothetical protein LOC100192069
GRMZM2G048450	5	NP_001151984	GRMZM2G048450_P01	Group II-a	WRKY39v2
GRMZM2G006497	8	#/N/A	GRMZM2G006497_P01	Group II-f	Transcription factor activity
GRMZM2G156529	6	#/N/A	GRMZM2G156529_P01	Group II-a	Transcription factor activity
GRMZM2G090594	10	NP_001151889	GRMZM2G090594_P01	Group II-b	WRKY25
GRMZM2G148561	4	NP_001151725	GRMZM2G148561_P01	Group I	Transcription factor activity
GRMZM2G071907	2	NP_001140970	GRMZM2G071907_P03	Group II-a	Transcription factor activity
GRMZM2G091331	10	C4J3W2	GRMZM2G091331_P01	Group III	Transcription factor activity
GRMZM2G073272	5	#/N/A	GRMZM2G073272_P01	Group II-f	Transcription factor activity
GRMZM2G102583	2	NP_001142073	GRMZM2G102583_P02	Group II-e	LOC100274230
GRMZM2G018487	1	Q32SG4	GRMZM2G018487_P01	Group II-e	WRKY1
GRMZM2G147880	5	NP_001150830	GRMZM2G147880_P01	Group II-e	WRKY transcription factor 21
GRMZM2G130374	1	NP_001136596	GRMZM2G130374_P01	Group II-b	Hypothetical protein LOC100216719
GRMZM2G173680	3	NP_001130531	GRMZM2G173680_P01	Group II-e	Hypothetical protein LOC100191630
GRMZM2G070211	1	NP_001147091	GRMZM2G070211_P01	Group II-a	WRKY transcription factor 21
GRMZM5G880069	5	NP_001131554	GRMZM5G880069_P02	Group II-f	Hypothetical protein LOC100192894
GRMZM2G040298	3	NP_001132768	GRMZM2G040298_P01	Group I	Hypothetical protein LOC100194255
GRMZM2G013391	8	C4J8W2	GRMZM2G013391_P01	Group II-e	Transcription factor activity
GRMZM2G141299	3	#/N/A	GRMZM2G141299_P02	Group II-b	Transcription factor activity
GRMZM2G027972	4	#/N/A	GRMZM2G027972_P01	Group II-f	Transcription factor activity
GRMZM2G052671	2	NP_001170182	GRMZM2G052671_P01	Group II-e	Hypothetical protein LOC100384128
GRMZM2G151407	7	NP_001143143	GRMZM2G151407_P02	Group II-b	Hypothetical protein LOC100275623
GRMZM2G008029	1	#/N/A	GRMZM2G008029_P01	Group II-a	Transcription factor activity
GRMZM2G076657	3	NP_001147897	GRMZM2G076657_P01	Group II-b	LOC100281507
GRMZM2G143765	6	NP_001130833	GRMZM2G143765_P01	Group II-d	Hypothetical protein LOC100191937
GRMZM2G171428	9	#/N/A	GRMZM2G171428_P01	Group I	Transcription factor activity
GRMZM2G2452430	1	#/N/A	GRMZM2G2452430_P01	Group II-a	Transcription factor activity
GRMZM2G020254	10	#/N/A	GRMZM2G020254_P03	Group II-b	Transcription factor activity
GRMZM2G549512	4	NP_001151453	GRMZM2G549512_P01	Group II-b	WRKY transcription factor 4
GRMZM2G169966	6	#/N/A	GRMZM2G169966_P01	Group II-b	Transcription factor activity
GRMZM2G036703	8	#/N/A	GRMZM2G036703_P01	Group I	Transcription factor activity
GRMZM2G148087	3	NP_001152322	GRMZM2G148087_P01	Group II-a	Transcription factor activity
GRMZM2G012724	6	NP_001147949	GRMZM2G012724_P01	Group III	WRKY53
GRMZM2G449681	8	NP_001147551	GRMZM2G449681_P01	Group II-b	Transcription factor activity
GRMZM2G130854	2	NP_001168562	GRMZM2G130854_P01	Group II-f	Hypothetical protein LOC100382344
GRMZM2G398506	7	NP_001146223	GRMZM2G398506_P02	Group II-a	LOC100279793
GRMZM2G031963	8	NP_001147816	GRMZM2G031963_P03	Group II-a	WRKY22
GRMZM5G816457	2	NP_001169214	GRMZM5G816457_P02	Group II-c	Hypothetical protein LOC100383070
GRMZM2G304573	8	#/N/A	GRMZM2G304573_P01	Group II-a	Transcription factor activity
GRMZM2G083350	8	#/N/A	GRMZM2G083350_P01	Group II-b	Transcription factor activity
GRMZM2G176489	3	#/N/A	GRMZM2G176489_P01	Group II-e	Transcription factor activity
GRMZM2G327349	3	#/N/A	GRMZM2G327349_P01	Group II-f	Transcription factor activity
GRMZM5G871347	3	#/N/A	GRMZM5G871347_P02	Group II-f	Transcription factor activity
GRMZM2G366795	6	#/N/A	GRMZM2G366795_P02	Group II-a	Transcription factor activity
GRMZM2G448605	8	#/N/A	GRMZM2G448605_P01	Group II-b	Transcription factor activity
GRMZM2G169149	7	NP_001147635	GRMZM2G169149_P01	Group I	WRKY62
AC209050.3_FG003	6	#/N/A	AC209050.3_FGP003	Group II-e	Transcription factor activity
GRMZM2G111711	9	#/N/A	GRMZM2G111711_P01	Group II-f	Transcription factor activity
GRMZM2G120320	5	NP_001146039	GRMZM2G120320_P01	Group II-a	Hypothetical protein LOC100279570
GRMZM2G057011	2	#/N/A	GRMZM2G057011_P01	Group II-a	Transcription factor activity
GRMZM2G125653	7	NP_001120723	GRMZM2G125653_P01	Group II-e	WRKY DNA-binding protein
AC198725.4_FG009	3	#/N/A	AC198725.4_FGP009	Group II-f	WRKY11
GRMZM2G145554	8	NM_001153177	GRMZM2G145554_P01	Group I	Transcription factor activity
GRMZM2G453571	6	#/N/A	GRMZM2G453571_P01	Group I	Transcription factor activity
GRMZM2G151763	3	NP_001169830	GRMZM2G151763_P01	Group II-a	Hypothetical protein LOC100383722
GRMZM5G812272	8	NP_001145926	GRMZM5G812272_P02	Group I	Transcription factor activity
GRMZM2G018721	7	#/N/A	GRMZM2G018721_P01	Group I	Transcription factor activity
GRMZM2G143204	1	NP_001130077	GRMZM2G143204_P01	Group I	Hypothetical protein LOC100191170
GRMZM2G054125	4	#/N/A	GRMZM2G054125_P01	Group I	Transcription factor activity
GRMZM2G106560	2	#/N/A	GRMZM2G106560_P01	Group II-a	Transcription factor activity
GRMZM2G111354	8	#/N/A	GRMZM2G111354_P01	Group II-e	Transcription factor activity
GRMZM2G015433	8	NP_001148624	GRMZM2G015433_P01	Group I	WRKY23
GRMZM2G401521	6	#/N/A	GRMZM2G401521_P01	Group II-a	Transcription factor activity
GRMZM2G151444	3	NP_001132878	GRMZM2G151444_P01	Group II-a	Hypothetical protein LOC100194371
GRMZM2G123387	2	NP_001151912	GRMZM2G123387_P01	Group I	WRKY36
GRMZM2G377217	4	#/N/A	GRMZM2G377217_P01	Group III	Transcription factor activity
GRMZM2G101405	3	#/N/A	GRMZM2G101405_P01	Group I	Transcription factor activity
GRMZM2G137802	8	NP_001148337	GRMZM2G137802_P01	Group II-e	WRKY7
GRMZM5G863420	6	#/N/A	GRMZM5G863420_P01	Group II-f	Transcription factor activity
GRMZM2G475984	3	#/N/A	GRMZM2G475984_P01	Group II-a	Transcription factor activity
GRMZM2G516301	8	COP8K1	GRMZM2G516301_P01	Group II-a	Transcription factor activity

GRMZM2G057116	8	NP_001148599	GRMZM2G057116_P01	Group II-e	WRKY67
GRMZM2G163054	6	C4J6I0	GRMZM2G163054_P02	Group II-a	Transcription factor activity

Table S2. Ortholog of WRKY between maize, sorghum and rice

Gene name	Chr	Start	Sorghum	Rice	Protein name in this article	Subgroup
GRMZM2G324999	Chr1	216580897	Sb07g019400	13108.m03064	ZmWRKY7	Group III
GRMZM2G083717	Chr1	299282977	Sb01g000696	13103.m07081	ZmWRKY4	Group II-e
GRMZM2G383594	Chr1	105951470	Sb01g027770	13110.m04048	ZmWRKY8	Group II-e
GRMZM2G018487	Chr1	274880474	Sb01g008550	13103.m05789	ZmWRKY2	Group II-f
GRMZM2G130374	Chr1	167896338	Sb08g020270	13112.m04223	ZmWRKY5	Group II-f
GRMZM2G070211	Chr1	253502280	Sb01g014180		ZmWRKY3	Group II-f
GRMZM2G008029	Chr1	278602439		13103.m06001	ZmWRKY1	Group I
GRMZM2G143204	Chr1	278162171		13103.m05984	ZmWRKY6	Group II-a
GRMZM2G425430	Chr1	75169610	Sb01g032120		ZmWRKY9	Group I
GRMZM2G163418	Chr2	175752568	Sb02g022290		ZmWRKY18	Group III
GRMZM2G024898	Chr2	12799342	Sb06g027290	13104.m05176	ZmWRKY10	Group II-e
GRMZM2G071907	Chr2	11753531	Sb06g027710	13104.m05262	ZmWRKY13	Group II-f
GRMZM2G102583	Chr2	79282357			ZmWRKY14	Group II-f
GRMZM2G052671	Chr2	189310197	Sb02g027950	13109.m03036	ZmWRKY11	Group I
GRMZM2G123387	Chr2	21500615	Sb06g024220	13104.m04629	ZmWRKY16	Group II-a
GRMZM2G057011	Chr2	180621363	Sb02g024760	13109.m02392	ZmWRKY12	Group II-b
GRMZM2G106560	Chr2	221753525	Sb05g017250		ZmWRKY15	Group II-a
GRMZM2G130854	Chr2	207841392	Sb02g037660	13107.m04089	ZmWRKY17	Group I
GRMZM5G816457	Chr2	38760033	Sb06g019710	13104.m03838	ZmWRKY19	Group I
AC165171.2_FG002	Chr3	144743920	Sb03g047350	13101.m08100	ZmWRKY20	Group II-f
GRMZM2G173680	Chr3	116024859	Sb08g020270	13112.m04223	ZmWRKY29	Group II-f
GRMZM2G040298	Chr3	201013567	Sb03g033640	13101.m05556	ZmWRKY22	Group II-e
GRMZM2G141299	Chr3	198214591	Sb03g034670	13101.m05741	ZmWRKY25	Group II-e
GRMZM2G076657	Chr3	134915466	Sb08g016240	13112.m03291	ZmWRKY23	Group I
GRMZM2G475984	Chr3	8213140			ZmWRKY32	Group II-a
GRMZM2G101405	Chr3	203083255	Sb03g032800	13101.m05395	ZmWRKY24	Group II-a
GRMZM2G151444	Chr3	9234113	Sb03g003640	13101.m00922	ZmWRKY27	Group II-a
AC198725.4_FG009	Chr3	217212884	Sb03g028530	13101.m04516	ZmWRKY21	Group II-a
GRMZM2G151763	Chr3	210977929	Sb03g030480	13101.m04918	ZmWRKY28	Group II-a
GRMZM2G148087	Chr3	183888339	Sb03g038510	13101.m06501	ZmWRKY26	Group I
GRMZM2G176489	Chr3	8217987	Sb03g003370	13101.m00961	ZmWRKY30	Group II-b
GRMZM2G327349	Chr3	49562769	Sb03g011800	13101.m02015	ZmWRKY31	Group II-b
GRMZM5G871347	Chr3	1260877	Sb03g000240	13101.m01589	ZmWRKY33	Group II-b
AC205562.3_FG002	Chr4	73366185	Sb07g019400	13108.m03064	ZmWRKY34	Group III
GRMZM2G138683	Chr4	158861929	Sb04g030930	13102.m05292	ZmWRKY38	Group II-e
GRMZM2G038158	Chr4	223519261	Sb04g009800	13102.m01849	ZmWRKY36	Group II-e
GRMZM2G148561	Chr4	69920108		13108.m01482	ZmWRKY39	Group II-f
GRMZM2G027972	Chr4	198130795	Sb07g028430	13108.m04150	ZmWRKY35	Group I
GRMZM2G377217	Chr4	151070766	Sb04g033240	13102.m04838	ZmWRKY40	Group II-a
GRMZM2G054125	Chr4	17852745		13111.m02887	ZmWRKY37	Group II-a
GRMZM2G054125	Chr4	17852745		13111.m02887	ZmWRKY37	Group II-a
GRMZM2G549512	Chr4	58331160			ZmWRKY41	Group I
GRMZM2G090594	Chr10	68725531	Sb07g006980		ZmWRKY84	Group II-f
GRMZM2G091331	Chr10	141270670	Sb06g027710	13104.m05262	ZmWRKY85	Group II-f
GRMZM2G031963	Chr10	124659347	Sb06g019710	13104.m03838	ZmWRKY83	Group I
GRMZM2G020254	Chr10	65416247	Sb07g006230		ZmWRKY82	Group II-c

Table S3: Expression patterns investigation of ZmWRKY using the maize SBS WGS database

Gene ID	Smop1	Swt	Seeding	Shoots	Seeding Roots	Ear	Pollen	Roots	Seedlings	Tassel
ZmWRKY43	0	0	4		2	11	0	33	0	0
ZmWRKY26	1	1	0		0	0	0	0	0	0
ZmWRKY60	50	19	17		20	25	34	57	24	13
ZmWRKY3	0	0	1		2	0	0	0	0	0
ZmWRKY67	0	0	1		0	0	0	0	0	0
ZmWRKY1	0	0	0		0	0	0	0	0	0
ZmWRKY20	0	2	5		4	30	0	51	22	8
ZmWRKY32	0	1	9		4	15	0	22	2	2
ZmWRKY71	0	0	0		0	0	3	0	0	0
ZmWRKY13	0	2	1		2	1	0	4	0	0
ZmWRKY47	0	0	3		2	0	0	4	0	0
ZmWRKY83	0	0	0		0	0	0	0	0	0
ZmWRKY18	0	2	3		0	0	0	2	0	0
ZmWRKY21	0	1	3		2	0	0	0	0	0
ZmWRKY5	1	0	0		0	0	0	0	0	0
ZmWRKY58	0	0	0		0	0	0	0	0	0
ZmWRKY33	0	0	1		0	0	0	0	0	0
ZmWRKY54	0	0	1		0	1	0	0	0	0
ZmWRKY28	0	0	2		4	1	0	1	0	0
ZmWRKY34	3	1	2		4	1	0	5	0	0
ZmWRKY29	0	0	0		0	1	0	0	0	0
ZmWRKY36	1	3	1		0	1	0	2	0	0
ZmWRKY10	6	1	2		2	6	0	28	2	10
ZmWRKY52	1	7	18		6	28	0	41	4	2
ZmWRKY44	3	0	0		0	1	0	1	0	0
ZmWRKY64	2	1	0		0	1	0	0	0	0
ZmWRKY27	7	1	0		0	1	0	0	0	0
ZmWRKY86	0	0	0		0	0	0	0	0	0
ZmWRKY19	0	0	3		2	0	0	2	0	0
ZmWRKY8	3	0	1		0	0	0	2	0	0
ZmWRKY48	2	0	0		2	0	0	1	0	0
ZmWRKY14	0	0	0		0	0	0	0	0	0
ZmWRKY22	0	0	3		4	1	0	3	0	0
ZmWRKY55	0	0	1		0	3	0	1	0	0
ZmWRKY6	0	0	0		0	0	0	1	0	0
ZmWRKY30	11	5	1		2	7	0	6	0	0
ZmWRKY50	0	1	1		0	0	0	3	0	0
ZmWRKY63	5	0	0		24	0	0	0	0	0
ZmWRKY74	1	0	0		0	0	0	0	0	0
ZmWRKY12	0	0	0		0	0	0	0	0	0
ZmWRKY80	1	0	0		0	0	0	0	0	0
ZmWRKY62	3	28	26		14	13	0	26	14	11
ZmWRKY17	4	0	1		0	1	0	2	0	0
ZmWRKY53	4	5	0		4	0	0	2	0	0
ZmWRKY7	7	1	4		0	0	0	3	0	0
ZmWRKY76	11	1	4		2	1	0	7	0	0
ZmWRKY45	3	1	1		6	3	1	8	0	0
ZmWRKY72	12	28	23		10	25	1	37	3	8
ZmWRKY15	4	1	1		0	0	0	0	0	0
ZmWRKY82	0	0	0		0	0	0	0	0	0
ZmWRKY66	0	0	1		2	0	0	0	0	0
ZmWRKY31	0	0	2		0	1	0	3	0	0
ZmWRKY65	0	0	2		0	1	0	1	0	0
ZmWRKY68	0	0	2		0	0	0	0	0	0
ZmWRKY85	0	0	0		0	0	0	0	0	0
ZmWRKY69	0	0	0		0	0	0	0	0	0
ZmWRKY75	0	5	0		2	8	0	17	0	2
ZmWRKY61	0	0	0		0	0	0	0	0	0
ZmWRKY4	1	0	0		2	0	0	0	0	0
ZmWRKY39	0	0	0		0	0	0	0	0	0
ZmWRKY40	1	0	4		4	0	0	0	0	0
ZmWRKY24	0	0	0		0	0	0	0	0	0
ZmWRKY42	0	0	0		0	0	0	1	0	0
ZmWRKY2	0	0	7		6	13	0	16	0	1
ZmWRKY51	0	1	0		0	0	0	0	0	0
ZmWRKY77	2	0	0		0	0	0	0	0	0
ZmWRKY57	2	0	0		0	0	0	0	0	0
ZmWRKY81	0	0	0		0	0	0	0	0	0
ZmWRKY11	1	0	0		0	0	0	0	0	0
ZmWRKY49	1	0	1		0	0	0	1	0	0
ZmWRKY23	1	0	0		0	0	0	0	0	0
ZmWRKY37	0	0	0		0	0	0	0	0	0
ZmWRKY38	0	0	1		0	0	0	0	0	0
ZmWRKY9	0	0	0		0	0	0	1	0	0
ZmWRKY73	0	0	2		0	0	0	2	0	0
ZmWRKY56	0	3	15		8	25	0	66	15	14
ZmWRKY44	5	6	5		4	14	0	23	0	0
ZmWRKY70	0	0	0		0	0	0	0	0	0

Table S4: Expression profile of the candidate WRKY genes in 30 distinct tissues

gene_name	Aliases	Developing_s1Embryo	Embryo_20DAS	Endosperm	Ovule2	Pollen2	PPTassel	Seed_15DAP	Seed_10DAS	Tassel	Unpollinated	SAM_mcp1	SAM_wt	Bundle_shoot	Developing_s1Nature	Leaf	Mesophyl	Embryo_sac	Ovule	Pollen	Seeding_c10Seeding_sh1Embryo	14D_Endosperm	seedlings20d_seeding20d_seeding20d	seedlings20d_seeding20d_seeding20d			
ZmWRKY26	0.17923 0.23684	0.609687	0.102299	1.10655	0	1.76968	0.299103	0.16803	0	0	0.79078	4.46163	2.74316	0	0	3.55688	0.168964	0	0	1.42361	1.41718	0.725491	0	2.75033	2.14101	0.236157	0.418846
ZmWRKY43	0.10000 0.00000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0	0.00000	0.179104	0.180170	0	0	0.203300	0.247474	0	0	0.156702	0	0	0	0.165462	0.115993	0.0256185	0.179404
ZmWRKY47	0.10000 0.00000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0	0.00000	0.179104	0.180170	0	0	0.203300	0.247474	0	0	0.156702	0	0	0	0.165462	0.115993	0.0256185	0.179404
ZmWRKY5	0.16651 0.23604	0.4971735 0.844154	0.000000	0.000000	0.000000	0.000000	0.000000	1.19869 0.15324	0.180169	0	0	0.82164	1.02041	0	0	0.203300	0.247474	0	0	0.156702	0	0	0	0.165462	0.115993	0.0256185	0.179404
ZmWRKY14	0.10000 0.00000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0	0.00000	0.179104	0.180170	0	0	0.203300	0.247474	0	0	0.156702	0	0	0	0.165462	0.115993	0.0256185	0.179404
ZmWRKY21	0.117771 0.210597	0.276382 0.092206	0.291578	0	0.487556	2.18724	1.82049	1.64323	0.690452	0	0	0.81742	0.202401	0	0	0.000000	0.000000	0	0	0.020481	0.059776	0.939653	0.305446	0.654559	0.622222	0.80139	0.026601
ZmWRKY22	0.117771 0.210597	0.276382 0.092206	0.291578	0	0.487556	2.18724	1.82049	1.64323	0.690452	0	0	0.81742	0.202401	0	0	0.000000	0.000000	0	0	0.020481	0.059776	0.939653	0.305446	0.654559	0.622222	0.80139	0.026601
ZmWRKY23	0.117771 0.210597	0.276382 0.092206	0.291578	0	0.487556	2.18724	1.82049	1.64323	0.690452	0	0	0.81742	0.202401	0	0	0.000000	0.000000	0	0	0.020481	0.059776	0.939653	0.305446	0.654559	0.622222	0.80139	0.026601
ZmWRKY24	0.080887 0.191169	21.0333 24.6005	9.6191	0	0.32884	5.66819	5.3108	1.30818	3.21379	11.0038	6.30275	1.75000	0	0	0.000000	0.000000	0	0	0.020481	0.059776	0.939653	0.305446	0.654559	0.622222	0.80139	0.026601	
ZmWRKY25	0.080887 0.191169	21.0333 24.6005	9.6191	0	0.32884	5.66819	5.3108	1.30818	3.21379	11.0038	6.30275	1.75000	0	0	0.000000	0.000000	0	0	0.020481	0.059776	0.939653	0.305446	0.654559	0.622222	0.80139	0.026601	
ZmWRKY26	0.080887 0.191169	21.0333 24.6005	9.6191	0	0.32884	5.66819	5.3108	1.30818	3.21379	11.0038	6.30275	1.75000	0	0	0.000000	0.000000	0	0	0.020481	0.059776	0.939653	0.305446	0.654559	0.622222	0.80139	0.026601	
ZmWRKY27	0.15789 22.4997	6.87529 7.60702	11.6761	0	0.40078	4.18943	2.8574	1.38803	13.3452	20.3741	1.59106	1.82308	0	0	0.62945	3.30781	0	0	0.020481	0.059776	0.939653	0.305446	0.654559	0.622222	0.80139	0.026601	
ZmWRKY28	0.15789 22.4997	6.87529 7.60702	11.6761	0	0.40078	4.18943	2.8574	1.38803	13.3452	20.3741	1.59106	1.82308	0	0	0.62945	3.30781	0	0	0.020481	0.059776	0.939653	0.305446	0.654559	0.622222	0.80139	0.026601	
ZmWRKY29	0.15789 22.4997	6.87529 7.60702	11.6761	0	0.40078	4.18943	2.8574	1.38803	13.3452	20.3741	1.59106	1.82308	0	0	0.62945	3.30781	0	0	0.020481	0.059776	0.939653	0.305446	0.654559	0.622222	0.80139	0.026601	
ZmWRKY30	0.15789 22.4997	6.87529 7.60702	11.6761	0	0.40078	4.18943	2.8574	1.38803	13.3452	20.3741	1.59106	1.82308	0	0	0.62945	3.30781	0	0	0.020481	0.059776	0.939653	0.305446	0.654559	0.622222	0.80139	0.026601	
ZmWRKY31	0.15789 22.4997	6.87529 7.60702	11.6761	0	0.40078	4.18943	2.8574	1.38803	13.3452	20.3741	1.59106	1.82308	0	0	0.62945	3.30781	0	0	0.020481	0.059776	0.939653	0.305446	0.654559	0.622222	0.80139	0.026601	
ZmWRKY32	0.15789 22.4997	6.87529 7.60702	11.6761	0	0.40078	4.18943	2.8574	1.38803	13.3452	20.3741	1.59106	1.82308	0	0	0.62945	3.30781	0	0	0.020481	0.059776	0.939653	0.305446	0.654559	0.622222	0.80139	0.026601	
ZmWRKY33	0.15789 22.4997	6.87529 7.60702	11.6761	0	0.40078	4.18943	2.8574	1.38803	13.3452	20.3741	1.59106	1.82308	0	0	0.62945	3.30781	0	0	0.020481	0.059776	0.939653	0.305446	0.654559	0.622222	0.80139	0.026601	
ZmWRKY34	0.15789 22.4997	6.87529 7.60702	11.6761	0	0.40078	4.18943	2.8574	1.38803	13.3452	20.3741	1.59106	1.82308	0	0	0.62945	3.30781	0	0	0.020481	0.059776	0.939653	0.305446	0.654559	0.622222	0.80139	0.026601	
ZmWRKY35	0.15789 22.4997	6.87529 7.60702	11.6761	0	0.40078	4.18943	2.8574	1.38803	13.3452	20.3741	1.59106	1.82308	0	0	0.62945	3.30781	0	0	0.020481	0.059776	0.939653	0.305446	0.654559	0.622222	0.80139	0.026601	
ZmWRKY36	0.15789 22.4997	6.87529 7.60702	11.6761	0	0.40078	4.18943	2.8574	1.38803	13.3452	20.3741	1.59106	1.82308	0	0	0.62945	3.30781	0	0	0.020481	0.059776	0.939653	0.305446	0.654559	0.622222	0.80139	0.026601	
ZmWRKY37	0.15789 22.4997	6.87529 7.60702	11.6761	0	0.40078	4.18943	2.8574	1.38803	13.3452	20.3741	1.59106	1.82308	0	0	0.62945	3.30781	0	0	0.020481	0.059776	0.939653	0.305446	0.654559	0.622222	0.80139	0.026601	
ZmWRKY38	0.15789 22.4997	6.87529 7.60702	11.6761	0	0.40078	4.18943	2.8574	1.38803	13.3452	20.3741	1.59106	1.82308	0	0	0.62945	3.30781	0	0	0.020481	0.059776	0.939653	0.305446	0.654559	0.622222	0.80139	0.026601	
ZmWRKY39	0.15789 22.4997	6.87529 7.60702	11.6761	0	0.40078	4.18943	2.8574	1.38803	13.3452	20.3741	1.59106	1.82308	0	0	0.62945	3.30781	0	0	0.020481	0.059776	0.939653	0.305446	0.654559	0.622222	0.80139	0.026601	
ZmWRKY40	0.15789 22.4997	6.87529 7.60702	11.6761	0	0.40078	4.18943	2.8574	1.38803	13.3452	20.3741	1.59106	1.82308	0	0	0.62945	3.30781	0	0	0.020481	0.059776	0.939653	0.305446	0.654559	0.622222	0.80139	0.026601	
ZmWRKY41	0.15789 22.4997	6.87529 7.60702	11.6761	0	0.40078	4.18943	2.8574	1.38803	13.3452	20.3741	1.59106	1.82308	0	0	0.62945	3.30781	0	0	0.020481	0.059776	0.939653	0.305446	0.654559	0.622222	0.80139	0.026601	
ZmWRKY42	0.15789 22.4997	6.87529 7.60702	11.6761	0	0.40078	4.18943	2.8574	1.38803	13.3452	20.3741	1.59106	1.82308	0	0	0.62945	3.30781	0	0	0.020481	0.059776	0.939653	0.305446	0.654559	0.622222	0.80139	0.026601	
ZmWRKY43	0.15789 22.4997	6.87529 7.60702	11.6761	0	0.40078	4.18943	2.8574	1.38803	13.3452	20.3741	1.59106	1.82308	0	0	0.62945	3.30781	0	0	0.020481	0.059776	0.939653	0.305446	0.654559	0.622222	0.80139	0.026601	
ZmWRKY44	0.15789 22.4997	6.87529 7.60702	11.6761	0	0.40078	4.18943	2.8574	1.38803	13.3452	20.3741	1.59106	1.82308	0	0	0.62945	3.30781	0	0	0.020481	0.059776	0.939653	0.305446	0.654559	0.622222	0.80139	0.026601	
ZmWRKY45	0.15789 22.4997	6.87529 7.60702	11.6761	0	0.40078	4.18943	2.8574	1.38803	13.3452	20.3741	1.59106	1.82308	0	0	0.62945	3.30781	0	0	0.020481	0.059776	0.939653	0.305446	0.654559	0.622222	0.80139	0.026601	
ZmWRKY46	0.15789 22.4997	6.87529 7.60702	11.6761	0	0.40078	4.18943	2.8574	1.38803	13.3452	20.3741	1.59106	1.82308	0	0	0.62945	3.30781	0	0	0.020481	0.059776	0.939653	0.305446	0.654559	0.622222	0.80139	0.026601	
ZmWRKY47	0.15789 22.4997	6.87529 7.60702	11.6761	0	0.40078	4.18943	2.8574	1.38803	13.3452	20.3741	1.59106	1.82308	0	0	0.62945	3.30781	0	0	0.020481	0.059776	0.939653	0.305446	0.654559	0.622222	0.80139	0.026601	
ZmWRKY48	0.15789 22.4997	6.87529 7.60702	11.6761	0	0.40078	4.18943	2.8574	1.38803	13.3452	20.3741	1.59106	1.82308	0	0	0.62945	3.30781	0	0	0								

