

Full Length Research Paper

Molecular cloning of WRKY transcription factor sequences in wild emmer wheat (*Triticum dicoccoides*)

Pan Dong^{1,2}, Guo-Yue Chen¹, Ya-Xi Liu¹, Yu-Ming Wei¹, Qian-Tao Jiang¹, Wei Li³, Eviatar Nevo⁴, Yong-Sheng Liu² and You-Liang Zheng^{1*}

¹Triticeae Research Institute, Sichuan Agricultural University, Wenjiang, Chengdu, Sichuan 611130, China.

²School of Life Sciences, Chongqing University, Chongqing 400030, China.

³Agronomy College, Sichuan Agricultural University, Wenjiang, Chengdu, Sichuan 611130, China.

⁴Institute of Evolution, University of Haifa, Mt. Carmel, Haifa 31905, Israel.

Accepted 16 October, 2012

WRKY transcription factors are one of the largest families of transcriptional regulators in plants and the WRKY gene family is involved in several diverse pathways. WRKY genes contain one or two highly conserved DNA binding domains called WRKY domain interrupted by an intron. In this study, 20 sequence fragments with WRKY genes/proteins were identified by a pair of degenerate primer (WRKY 1 FP + WRKY 2 RP) from wild emmer wheat (*Triticum dicoccoides*), which were represented fragments of 7 unique WRKY loci. The *T. dicoccoides* WRKY (TdWRKY) putative loci ranged in size from 207 bp (TdWRKY-7) to 562 bp (TdWRKY-2). Size differences among the 7 TdWRKY putative loci were primarily due to variations in the length of the intron contained within the WRKY domain. These introns ranged in size from 97 bp (TdWRKY-7) to 449 bp (TdWRKY-2). According to the group assignments of 7 TdWRKY domains with 38 AtWRKY domains and the position of the putative intron for the TdWRKY loci, the 5 TdWRKY loci (TdWRKY-1, TdWRKY-3, TdWRKY-4, TdWRKY-5 and TdWRKY-7) belonged to WRKY group I domains; and the rest 2 loci's domains (TdWRKY-2 and TdWRKY-6) were clustered into WRKY group IIb. This research provide some useful information for studying the WRKY transcriptional factors families in the genus *Triticum* L.

Key words: WRKY transcription factor; Wild emmer wheat, phylogenetic analysis

INTRODUCTION

Wild emmer wheat (*Triticum dicoccoides*, AABB, 2n=28) is the tetraploid progenitor of tetraploid and hexaploid cultivated wheat, which could be crossed with durum wheat (*Triticum durum* L.) easily, because of the same genome formula (Zohary, 1970), and also with common wheat (*Triticum aestivum* L.) (Nevo et al., 2002). It naturally grows in the Middle-east nations, such as Israel, Syria, Jordan, Lebanon, northern Iraq, and western Iran (Nevo et al., 2002), also in some European countries, such as south-east Turkey, Italy, Spain, and in Asian country India, too (Konvalina et al., 2011). And the centre of distribution of *T. dicoccoides* was found in the

catchment area of the Upper Jordan Valley (East Galilee, Golan, Gilead, Hauran, Mt. Hermon) and vicinity (Nevo and Beiles, 1989). In this centre of diversity, *T. dicoccoides* harbors rich genetic resources for wheat improvement, including abiotic stress tolerances (salt, drought and heat), biotic stress tolerances (powdery mildew, rusts, and Fusarium head blight), grain protein quality and quantity, and micronutrient concentrations (Zn, Fe, and Mn) (Xie and Nevo, 2008).

The WRKY proteins, one of the largest families of transcriptional regulators in plants, have a well-conserved amino acid sequence, the WRKY domain, from which their name originated (Eulgem et al., 2000; Zhang and Wang, 2005; Rushton et al., 2010). Most of the WRKY domains contain a signature WRKYGQK motif followed by a distinctive zinc-finger-like motif.

All recognized WRKY proteins contain either one or two

*Corresponding author. E-mail: ylzheng@sicau.edu.cn. Tel: (+86 835) 2882 007. Fax: (+86 835) 2883 153.

WRKY domain. They could be classified into three main groups (I, II, and III) on the basis of both the number of WRKY domains and the features of their zinc-finger-like motif (Eulgem et al., 2000). Group II is further subdivided into five subgroups (a to e) based on the additional amino acid sequence present outside the WRKY domain (Eulgem et al., 2000; Rushton et al., 2010; Agarwal et al., 2011). A common element of WRKY genes is the interruption of the coding region of the C-terminal WRKY domain of group I and the single WRKY domain of groups II and III genes by an intron (Eulgem et al., 2000). The size and sequence of the intron vary in each gene, but its position is highly conserved, being localized after the codon encoding arginine that is N terminal to the zinc-finger-like motif, and aids in identifying the group/subgroup to which each gene belongs (Eulgem et al., 2000; Borrone et al., 2004). Since the first WRKY cDNAs was identified in the sweet potato (*Ipomoea batatas*) in 1994 (Ishiguro and Nakamura, 1994), lots of WRKY genes have been isolated from many plants, such as *Arabidopsis thaliana*, rice (*Oryza sativa* L.), cotton (*Gossypium* spp.), wheat and tomato (*Lycopersicon esculentum* M.) (Ross et al., 2007; Dong et al., 2003; Xu et al., 2004; Wu et al., 2008; Hofmann et al., 2008). Till August 2012, there were more than 1131 WRKY genes registered in NCBI (National Centre for Biotechnology Information) website, including 83 from *Arabidopsis thaliana*, 110 from rice (*Oryza Sativa*), 67 from corn (*Zea Mays*) and 13 from wheat. The WRKY proteins are involved in the regulation of various physiological programs, including regulating seed dormancy, germination and development, senescence, development, and plant responses to both abiotic and biotic stresses (Guo et al., 2004; Ulker and Somssich, 2004; Xie et al., 2005; Rushton et al., 2010; Zhao et al., 2012). While both the gene sequences and function of WRKY family of wheat are not well understood, some single WRKY genes and their functions have made big progress, such as TaWRKY1 tolerant to freezing, TaWRKY2 tolerant to drought and salt, and TaWRKY19 tolerant to drought, salt and freezing, and so on (Houde et al., 2006; Gregersen and Holm, 2007; Wu et al., 2008; Qin, 2009; Zhao et al., 2012).

Up to now, only one WRKY sequence from *T. dicoccoides* was released through the isolation of differentially expressed cDNA (Ergen and Budak, 2009). It has been reported that the WRKY genes could be obtained from degenerate primer pairs (Borrone et al., 2004). In this study, we reported that the isolation and analysis of WRKY gene fragments of *T. dicoccoides*.

MATERIALS AND METHODS

Plant material

All the four materials (TD88, TD89 from population J'aba, TD98 from population Amirim and TD129 from population Dalliya) were provided by the cereal gene bank of the Institute of Evolution,

University of Haifa.

DNA isolation

The seed of all genotypes were germinated under darkness at 23°C for 1 week. Young leaves were harvested and crushed into powder with the aid of liquid nitrogen. The genomic DNA was extracted by a CTAB method (Murray and Thompson, 1980) and diluted to 50 ng/μl for PCR.

PCR amplification

The degenerate primer pair WRKY 1 FP and WRKY 2 RP was released by Borrone et al. (2004). The detailed primer information was listed in Table 1 and Figure 1. WRKY analysis was conducted according to previously established protocols with minor modifications (Borrone et al., 2004). Each 20 μl PCR reaction mixture consisted of 2.0 μl of 10 × PCR buffer (2 mM of MgCl₂), 300 μmol dNTPs, 0.3 μmol primers, 100 ng genomic DNA, and 1U Taq polymerase. The PCR reaction procedure was: 94°C for 5 min, followed by 35 cycles of three steps: 30 s denaturing at 94°C, 45 s annealing at 55°C, and 30 s elongation at 72°C, with a final elongation step at 72°C for 5 min. The PCR products were separated on 1.2% agarose gels, and then the targeted DNA fragments were recovered and cloned into the pGEMT-Easy vector (Promega). The ligated products were transformed into *Escherichia coli* (DH5α) cells and the recombinated plasmids were screened as a sequencing template.

Bioinformatic analysis

Individual sequences and the consensus sequence of each gene fragment were putatively identified using BLASTN, BLASTX and TBLASTS searches (Altschul et al., 1997) of GenBank at the National Center for Biotechnology and Information (NCBI). The ORFs were translated into amino acids sequences using the ORF finder program at the NCBI network service (<http://www.ncbi.nlm.nih.gov/gorf/gorf.html>). The nucleotide sequences encoding WRKY domain region with or without the intron, and those of the putative open reading frames (ORFs) were aligned using softwares DNAMAN 5.2.2 (<http://www.lynnon.com>) and CLUSTAL W 1.81 (Thompson et al., 1994). The phylogenetic tree was generated based on the NJ (neighbour-joining) sequences distance method (Saitou and Nei, 1987) and depicted and edited by MEGA 3.1 program (Kumar et al., 2004). The used distance for the NJ grouping was 0.05 and the bootstrap value was estimated based on 500 replications.

RESULTS

Five distinctive bands ranging in size from 100 to 700 bp were observed in each of 4 tested *T. dicoccoides* accessions using the degenerate primer WRKY 1 FP + WRKY 2 RP. Of the random 54 colonies sequenced, 20 (37%) identified with WRKY genes/proteins using BLASTX and BLASTN or with the feature of WRKY genes/proteins. The vast majority (15) of the identifying WRKY sequences (20) ranged in size from 207 bp to 341 bp.

On the basis of multi-sequence alignment based on the WRKY domain, the 20 sequenced clones represented

Table 1. WRKY degenerate primers.

Primer ^a	Sequence ^b	Deg ^c	Corresponding amino acids ^b	Targeted sequences
WRKY 1 FP	TGGMGIAARTAYGGNCARA	64	WRKYGQ	All groups, all WRKY domains
WRKY 2 RP	TGRBYRTGYTTICCYTCRTAIGTD GT	576	TTYEGKH(T/N/S/G/A/D)(H/Q)	Group I, C-terminal WRKY domain and Group II, subgroup a-c

^aFP, Forward primer; RP, reverse primer; ^bThe nucleotide sequence is given in the 5'-3' direction using the standard IUB code where M = A or C; R = A or G; Y = C or T; B = T, C, or G; N = A, T, C, or G; I = Inosine. Amino acids are given in the standard one-letter code; ^cDeg, overall degeneracy of the primer omitting inosine.

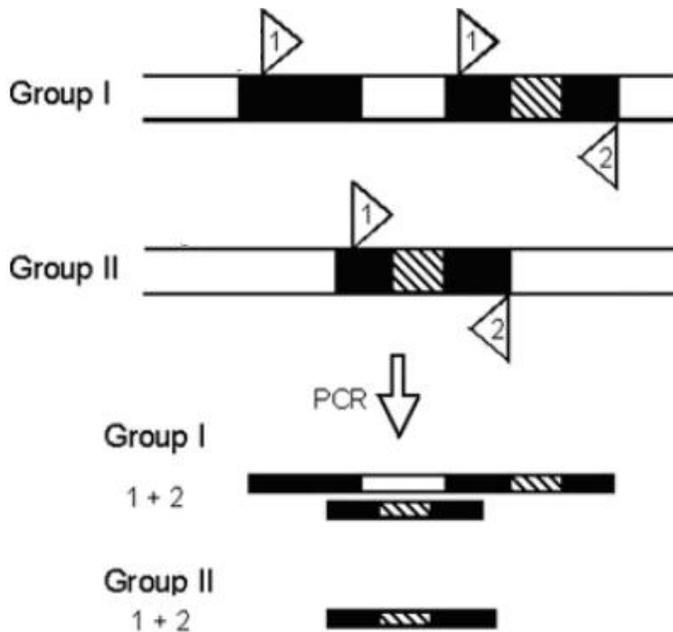


Figure 1. A scheme depicting the organization of group 1 and group II, a-c, WRKY proteins and the expected results from PCR with WRKY 1 FP + WRKY 2 RP based only upon the potential binding sites for the primer. Arrows indicate the approximate location of potential binding sites for the degenerate primers. 1, WRKY 1 FP, 2, WRKY 2 RP. The WRKY domain is indicated by the black box, and the intron interrupting the WRKY domain is indicated by diagonal lines. The locations of other introns are not indicated. The detailed information is from Borrone et al. (2004).

fragments of 7 unique WRKY loci (Table 2, Figure 2). For all the 7 individual loci, the translated amino acid sequences of each locus were completely identical. When the coding regions of WRKY domains of multiple clones representing a single TdWRKY putative locus were compared, the nucleotide sequences of the overlapping segments were more than 96% identical. While the results of comparing the multiple clones including the introns representing a single TdWRKY putative locus showed the nucleotide sequences of the overlapping segments were more than 94% identical, except TdWRKY-5, the identity is 73.33%. The 7 putative loci were different with one another based on the comparison of the nucleotide or

amino acid sequence of the coding region of the WRKY domain. The closest nucleotide percentage identity between two TdWRKY putative loci for the WRKY domain coding region was 85% for TdWRKY-1 and TdWRKY-4. As the DNA-binding WRKY domain was expected to be the most conserved region of the gene, it was concluded that each identified TdWRKY putative locus represented an individual gene locus. Among the 7 TdWRKY putative loci, four (TdWRKY-1, TdWRKY-4, TdWRKY-5 and TdWRKY-6) were decided by 2 clones, while the rest three were represented by 5, 4 and 3 clones, respectively. The TdWRKY putative loci ranged in size from 207 bp (TdWRKY-7) to 562 bp (TdWRKY-2). The putative 7 loci contained only the C-terminal WRKY domain for each of these genes.

The difference of the size among the 7 TdWRKY putative loci was primarily due to variations in the length of the intron contained within the WRKY domain. Alignments of all the 20 nucleotide sequences of the 7 TdWRKY putative loci or of the WRKY domains were not possible, when the introns existed, due to their extreme variation in length and sequence. These introns ranged in size from 97 bp (TdWRKY-7) to 449 bp (TdWRKY-2). The introns within the domains were at positions identical to those described in *Arabidopsis* for each group and subgroup (Eulgem et al., 2000). According to the feature of the degenerate primers, all the putative loci should belong to Group I or Group II a, b or c. The position of the putative intron for two of the TdWRKY loci (-2, -7) following the conserved CX₅C motif of the zinc-finger identified these as belonging to group II, subgroups a or b, WRKY genes. The position of the putative intron for the other five of the TdWRKY loci followed by the conserved CX₄C motif and the whole zinc-finger identified these as belonging to Group I or Group II c, WRKY genes.

Group assignments were conducted by aligning the translated amino acid sequences of 7 TdWRKY domains with 38 AtWRKY domains, including Group I or Group II a, b or c, obtained from Eulgem et al. (2000) (Figure 3). The WRKY domains separated into distinct clusters representing each WRKY group and subgroup. The tree generated was consistent with trees from previous analyses (Eulgem et al., 2000). The association of individual TdWRKY domains with one another received high bootstrap support (>60%). For example, TdWRKY-7, TdWRKY-1 and TdWRKY-4 were clustered together, with

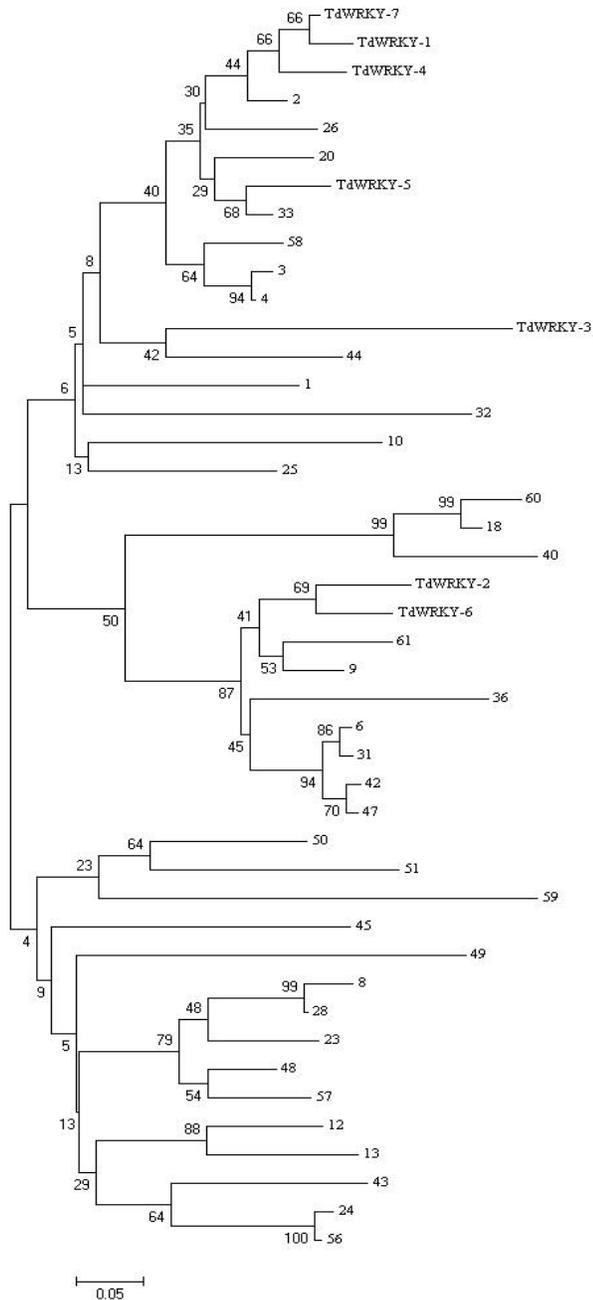


Figure 3. Phylograms depicting the relationship of TdWRKY domains with *Arabidopsis thaliana* (AtWRKY) domains. AtWRKY domains are indicated by number only, that is, 1 = AtWRKY1. Bootstrap support is given as a percentage of 500 datasets at each node, and groups/subgroups are designated as has been previously described for AtWRKY sequences (Eulgem et al., 2000). The complete tree and a list of AtWRKY accession numbers used are provided in the ESM.

the 66% of the times in bootstrap analysis; and TdWRKY-2 and TdWRKY-6 were obtained 69% of the times to get together. The domains of 5 TdWRKY loci (TdWRKY-1, TdWRKY-3, TdWRKY-4, TdWRKY-5 and

TdWRKY-7) grouped with AtWRKY group I domains. And the rest 2 loci's domains (TdWRKY-2 and TdWRKY-6) were placed with AtWRKY group IIb.

DISCUSSION

WRKY genes have been identified by three methods: sequence analysis of the entire genomes of *A. thaliana* (The Arabidopsis Genome Initiative, 2000) and *O. sativa* (Goff et al., 2002; Yu et al., 2002), or in the EST database of *T. aestivum* (Qin, 2009; Niu et al., 2012); the isolation of differentially expressed cDNA (Alexandrova and Conger, 2002; Hara et al., 2000; Hinderhofer and Zentgraf, 2001; Huang and Duman, 2002) and the isolation on the basis of the degenerate PCR primers (Chen and Chen, 2000; Trognitz et al., 2002; Borrone et al., 2004; Mauro-Herrera et al., 2006). It is an effective, cheap, easy and fast way to get WRKY genes from the plant with huge and un-sequenced genome using the method of PCR cloning. In this research, 20 nucleotide sequences with WRKY domain were successfully isolated from *T. dicoccoides* using one degenerate PCR primer: WRKY 1 FP + WRKY 2 RP. However, up to now, in the NCBI database only one WRKY nucleotide sequence from *T. dicoccoides* was reported through the isolation of differentially expressed cDNA (Ergen and Budak, 2009). The size and sequence variability of the intron, the position of the intron within the WRKY domain and the conserved amino acid motifs within the WRKY domain were used to classify the cloned PCR products into 7 groups. For the nucleotide and amino acid alignments of the most conserved portion of the TdWRKY loci, the DNA-binding WRKY domain, the TdWRKY loci were distinct with one another.

According to the BLAST among 7 putative TdWRKY loci and other WRKY sequences in the NCBI database one by one, 7 other WRKY sequences (TaWRKY1, BnWRKY6-1, AtWRKY44, BnWRKY2, BnWRKY33-1, BnWRKY72 and HvWRKY6) got the highest similarity to TdWRKY sequences (from TdWRKY1 to TdWRKY7) (Table 2). The BnWRKY genes were response to fungal pathogens and hormone treatments (Yang et al., 2009). So we could presume the TdWRKY transcriptional factors should also be related to some function. Further research should be made to get the full lengths of 7 putative loci from *T. dicoccoides* and to analyze the function of them. Furthermore, some single WRKY genes and their functions in wheat have made big progress, which will provide useful information for studying TdWRKY genes and functions.

According to the feature of the primers, the WRKY loci obtained should belong to group I and group II a, b or c. All the four types WRKY putative loci were gotten from *T. cacao* (Borrone et al., 2004), and the three quarters are group I and IIb. In our *T. dicoccoides* study, two types: group I and group IIb were obtained successfully. So we can presume this primer pair (WRKY 1 FP + WRKY 2 RP) has the privilege to get group I (especially C-terminal

WRKY domain) and IIb in *T. cacao* and *T. dicoccoides* genomes. In previous research, there was no group IIb after assigning 61 and 43 WRKY genes from *T. aestivum* cDNA database (Qin, 2009) and wheat EST database (Niu et al., 2012), and none of the 15 genes encoding WRKY transcription factor in wheat belonged to group IIb (Wu et al., 2008). While in our research, 2 of 7 TdWRKY putative loci belonged to group IIb. It must provide a very effective method for isolation group IIb from *T. aestivum*. However, because of the two potential binding sites for the forward degenerate WRKY 1FP (Figure1), theoretically the nucleotide sequences both N-terminal and C-terminal could be obtained from the primer pair. Only the C-terminal WRKY domains in group I were obtained in *T. dicoccoides*, and only 6 nucleotide sequenced out of 324 contained in *T. cacao* (Borrone et al., 2004). So we can predict that this primer pair (WRKY 1 FP + WRKY 2 RP) has the privilege to get the N-terminal WRKY domains. The reverse primer WRKY 2 RP determined the specificity of the WRKY 1FP + WRKY 2 RP combination for group I and group II, subgroups a-c, WRKY genes. WRKY domains of group II, subgroups d and e, and group III WRKY proteins differ in their amino acid sequence at the C-terminal portion of the zinc-finger motif (Eulgem et al., 2000). It was expected that WRKY genes belonging to these groups would not be amplified. None were identified in *T. cacao* (Borrone et al., 2004) and also in *T. dicoccoides*. This suggests that degenerate primers could be designed to specifically amplify WRKY genes from each group or subgroup.

ACKNOWLEDGEMENTS

This work was supported by the National Basic Research Program of China (973 Program 2010CB134400), China Transgenic Research Program (2011ZX08002) and Chongqing Natural Science Funds (CSTC, 2011BB1102).

REFERENCES

- Agarwal P, Reddy MP, Chikara J (2011). WRKY: its structure, evolutionary relationship, DNA-binding selectivity, role in stress tolerance and development of plants. *Mol. Biol. Rep.* 38:3883-3896.
- Alexandrova KS, Conger BV (2002). Isolation of two somatic embryogenesis-related genes from orchard grass (*Dactylis glomerata*). *Plant Sci.* 162:301-307.
- Altschul SF, Madden TL, Schaffer AA, Zhang J, Z Zhang, Miller W, Lipman DJ (1997). GappedBLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* 25:3389-3402.
- Borrone JW, Kuhn DN, Schnell RJ (2004). Isolation, characterization, and development of WRKY genes as useful genetic markers in *Theobroma cacao*. *Theor. Appl. Genet.* 109:495-507.
- Chen CH, Chen ZX (2000). Isolation and characterization of two pathogen- and salicylic acid-induced genes encoding WRKY DNA-binding proteins from tobacco. *Plant Mol. Biol.* 42:387-396.
- Dong JX, Chen CH, Chen ZX (2003). Expression profiles of the *Arabidopsis* WRKY gene superfamily during plant defense response. *Plant Mol. Biol.* 51:21-37.
- Ergen NZ, Budak H (2009). Sequencing over 13000 expressed sequence tags from six subtractive cDNA libraries of wild and modern wheats following slow drought stress. *Plant Cell Environ.* 32: 220-236.
- Eulgem T, Rushton PJ, Robatzek S, Somssich IE (2000). The WRKY superfamily of plant transcription factors. *Trends plant Sci.* 5:199-206.
- Goff SA, Ricke D, Lan TH, Presting G, Wang R, Dunn M, Glazebrook J, Sessions A, Oeller P, Varma H, Hadley D, Hutchison D, Martin C, Katagiri F, Lange BM, Moughamer T, Xia Y, Budworth P, Zhong J, Miguel T, Paszkowski U, Zhang S, Colbert M, Sun WL, Chen L, Cooper B, Park S, Wood TC, Mao L, Quail P, Wing R, Dean R, Yu Y, Zharkikh A, Shen R, Sahasrabudhe S, Thomas A, Cannings R, Gutin A, Pruss D, Reid J, Tavtigian S, Mitchell J, Eldredge G, Scholl T, Miller RM, Bhatnagar S, Adey N, Rubano T, Tusneem N, Robinson R, Feldhaus J, Macalma T, Oliphant A, Briggs S (2002). A draft sequence of the rice genome (*Oryza sativa* L. ssp. *Japonica*). *Science* 296:92-100.
- Gregersen PL, Holm PB (2007). Transcriptome analysis of senescence in the flag leaf of wheat (*Triticum aestivum* L.). *Plant Biotechnol. J.* 5:192-206.
- Guo Y, Cai Z, Gan S (2004). Transcriptome of *Arabidopsis* leaf senescence. *Plant Cell Environ.* 27:521-549.
- Hara K, Yagi M, Kusano T, Sano H (2000). Rapid systemic accumulation of transcripts encoding a tobacco WRKY transcription factor upon wounding. *Mol. Gen. Genet.* 263:30-37.
- Hinderhofer K, Zentgraf U (2001). Identification of a transcription factor specifically expressed at the onset of leaf senescence. *Planta* 213:469-473.
- Hofmann MG, Sinha AK, Proels RK, Roitsch T (2008). Cloning and characterization of a novel LpWRKY1 transcription factor in tomato. *Plant Physiol. Biochem.* 46:533-540.
- Houde M, Belcaid M, Ouellet F (2006). Wheat EST resources for functional genomics of abiotic stress. *BMC Genomics*, 7: 149.
- 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.* 48:339-350.
- Ishiguro S, Nakamura K (1994). Characterization of a c 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. General Genet.* 244:563-571.
- Konvalina P, Capouchová I, Stehno Z, Moudrý J jr, Moudrý J (2011). Spike productivity in relation to yield as criterion for emmer wheat breeding. *Romanian. Agric. Res.* 28:49-56.
- Kumar S, Tamura K, Jakobsen IB, Nei M (2004). MEGA3: integrated software for molecular evolutionary genetics analysis and sequence alignment. *Brief Bioinform.* 5: 150-163.
- Mauro-Herrera, Meerow AW, Borrone JW, Kuhn DN, Schnell RJ (2006). Ten informative markers developed from WRKY sequences in coconut (*Cocos nucifera*). *Mol. Ecol. Notes* 6:904-906.
- Murray M, Thompson WF (1980). Rapid isolation of high molecular weight plant DNA. *Nucleic Acids Res.* 8:4321-4325.
- Nevo E, Beiles A (1989). Genetic diversity of wild emmer wheat in Israel and Turkey: Structure, evolution and application in

- breeding. *Theor. Appl. Genet.* 77:421-455.
- Nevo E, Korol AB, Beiles A, Fahima T (2002). Evolution of Wild Emmer and Wheat Improvement. Population Genetics, Genetic Resources, and Genome Organization of Wheat's, *Triticum dicoccoides*. Springer-Verlag. p. 364.
- Niu CF, Wei W, Zhou QY, Tian AG, Hao YJ, Zhang WK, Ma B, Lin Q, Zhang ZB, Zhang JS, Chen SY (2012). Wheat WRKY genes TaWRKY2 and TaWRKY19 regulate abiotic stress tolerance in transgenic Arabidopsis plants. *Plant Cell. Environ.* 35:1156-1170.
- Qin W (2009). Sequence analysis of WRKY gene family and their response to powdery mildew fungus in wheat. Qufu Normal University, thesis.
- Ross CA, Liu Y, Shen QJ (2007). The WRKY gene family in Rice (*Oryza sativa*). *J. Integrat. Plant Biol.* 49:827-842.
- Rushton PJ, Somssich IE, Ringler P, Shen QJ (2010). WRKY transcription factors. *Trends Plant Sci.* 15:247-258.
- Saitou N, Nei M (1987). The neighbor-joining method: A new method of reconstructing phylogenetic trees. *Mol. Biol. Evol.* 4:406-425.
- The Arabidopsis Genome Initiative (2000). Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. *Nature* 408:796-815.
- Thompson JD, Higgins DG, Gibson TJ (1994). Clustal-W improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* 22:4673-4680.
- Trognitz F, Manosalva P, Gysin R, Ninio-Liu D, Simon R, Herrera MR, Trognitz B, Ghislain M, Nelson R (2002). Plant defense genes associated with quantitative resistance to potato late blight in *Solanum phureja* x dihaploid *S. Tuberosum* hybrids. *Mol. Plant Microbe. Interact.* 15:587-597.
- Ulker B, Somssich IE (2004). WRKY transcription factors: from DNA binding towards biological function. *Curr. Opin. Plant Biol.* 7:491-498.
- Wu HL, Ni ZF, Yao YY, Guo GG, Sun QX (2008). Cloning and expression profiles of 15 genes encoding WRKY transcription factor in wheat (*Triticum aestivum* L.). *Prog. Nat. Sci.* 18:697-705.
- Xie Z, Zhang ZL, Zou X, Huang J, Ruas P, Thompson D, Shen QJ (2005). Annotation and functional analyses of the rice WRKY gene superfamily reveal positive and negative regulators of abscisic acid signalling in aleurone cells. *Plant Physiol.* 137:176-189.
- Xie WL, Nevo E (2008). Wild emmer: genetic resources, gene mapping and potential for wheat improvement. *Euphytica* 164:603-614.
- Xu YH, Wang JW, Wang S, Wang JY, Chen XY (2004). Characterization of GaWRKY1, a cotton transcription factor that regulates the sesquiterpene synthase gene (+)-delta-cadinene synthase-A. *Plant Physiol.* 135:507-515.
- Yang B, Jiang YQ, Rahman MH, Deyholos MK, Kav NNV (2009). Identification and expression analysis of WRKY transcription factor genes in canola (*Brassica napus* L.) in response to fungal pathogens and hormone treatments. *BMC Plant Biol.* 9:68
- Yu J, Hu S, Wang J, Wong GK, Li S, Liu B, Deng Y, Dai L, Zhou Y, Zhang X (2002). A draft sequence of the rice genome (*Oryza sativa* L. ssp. *Indica*). *Science* 296:79-92.
- Zhang Y, Wang L (2005). The WRKY transcription factor superfamily: its origin in eukaryotes and expansion in plants. *BMC Evol. Biol.* 5:1. doi:10.1186/1471-2148-5-1.
- Zhao MY, Zhang ZB, Chen SY, Zhang JS, Shao HB (2012). WRKY transcription factor superfamily: structure, origin and function. *Afr. J. Biotechnol.* 11(32):8051-8059.
- Zohary D (1970). Centers of diversity and centers of origin. In: Frankel OH, Bennet E (eds) *Genetic Resources in Plants-their exploration and conservation*, Blackwell: Oxford. pp. 33-42