academicJournals

Vol. 12(23), pp. 3636-3642, 5 June, 2013
DOI: 10.5897/AJB12.2781
ISSN 1684-5315 ©2013 Academic Journals
http://www.academicjournals.org/AJB

African Journal of Biotechnology

Full Length Research Paper

Discovery AP2/ERF family genes *in silico* in *Medicago truncatula*

Zhifei Zhang*, Qian Zhou, Zhijian Yang and Jingpeng Jiang

College of Agronomy, Hunan Agricultural University, Furong District, Changsha, Hunan Province 410128, P.R. China.

Accepted 27 May, 2013

Medicago truncatula **is a legume model plant due to its small genome and it has been used to study the molecular events of legume biology. As a crucial plant-specific gene family, AP2/EREBP transcription factors (TFs) are important for plant development and biotic and abiotic stress responses. The purpose of the work was to determine AP2/ERF family genes *in silico* of *M. truncatula*, and also sheds light on molecular mechanism of stress responses of AP2/EREBPs. We investigated AP2/ERF family genes of *M. truncatula* using BLAST search. Thirty-seven (37) AP2/ERF family genes were identified and sorted into the corresponding subfamily or subgroup, with sequences alignment and phylogenetic analysis of the AP2-like TFs proteins between in Arabidopsis and in *M. truncatula*, and expression patterns of putative 35 AP2/ERF family genes in *M. truncatula* were revealed. Identification of AP2/ERF family genes would make them easier to clone and position those functional genes, and which also would open new opportunities for the study of molecular regulatory network of stress resistance in *M. truncatula*.**

Key words: *Medicago truncatula*, transcription factor, AP2/ERF.

INTRODUCTION

As leguminous model plant, *Medicago truncatula* is featured by small genome (500 to 550 Mbp) (Young and Udvardi, 2009), high genetic transformation efficiency, self-pollination, nitrogen fixation, etc., which facilitate to study legumes biology and genomics. Transcription factors (TFs) are a class of proteins regulating the gene transcription and expression by recognizing and binding to *cis*-acting elements in certain gene promoters. APETALA2/ethylene-responsive element binding protein (AP2/EREBP) belongs to one of important plant-specific transcription factor families. In the model plant *Arabidopsis thaliana*, AP2/EREBPs are related to plant development, drought and salt stress responses and so on. They are composed of five subfamilies containing AP2 (APETALA 2), ERF (ethylene-responsive transcription factor), DREB/CBF (dehydration-responsive element-binding protein/C-repeat binding factor), RAV (related to ABI3/VP1) (Feng et al., 2005) and Solosit. Compared to the model plant Arabidopsis, few study is focused on AP2/EREBPs of *M. truncatula* and several AP2/EREBPs have been identified, including Mt*CBF*1, Mt*CBF*2, Mt*DREB*1C/*CBF*3 (Pennycooke et al., 2008), Mt*CBF*4, *WXP*1, *WXP*2 (Zhang et al., 2005; Zhang et al., 2007) and Mt*DREB*2A (Chen et al., 2009).

Expressed sequence tags (EST) technology has been widely used to acquire plant genetic information. Recently, BLAST analysis in EST database with AP2 domain sequence probe was exploited to identify putative AP2/ERF family TFs in *Triticum aestivum* (Zhuang et al., 2011a), *Hordeum vulgare*, soybean (Mochida et al., 2009), *Brassica napus* L. HuYou15 etc. (Zhuang et al., 2011b), moreover, based on these discoveries of AP2-like genes, GmERF4 (Zhang et al., 2010) and *BnaRAV-1-HY15* gene (Zhuang et al., 2011b), were isolated successfully from *B. napus* L. cv HuYou15, which suggested that this method was a simple and effective

*Corresponding author. E-mail: zzf0917@yahoo.com.cn. Tel: (+81) 0731 84618684.

way to study the AP2-like genes in plants.

To investigate the roles of AP2/EREBP in *M. truncatula*, we isolated AP2/ERF family genes using the conserved sequence of *Arabidopsis* AP2/ERF transcription factors as the electron probe to screen NCBI Unigene database of *M. truncatula*. The deduced protein sequences, domain structure, function and expression pattern have been predicted and analyzed by bioinformatics method.

MATERIALS AND METHODS

Isolation of AP2/ERF family genes

A search for probes specific to AP2/ERF family was performed according to **Zhuang et al. (2011a, 2011b)** and Zhuang et al. (2008). The EST databases were downloaded for *M. truncatula* from the NCBI site (ftp://ftp.ncbi.nih.gov/repository/UniGene/Medicago_truncatula/, the data were released on 26 July 2010) and *A. thaliana* from DATF site (http://datf.cbi.pku.edu.cn/) to a local computer. The NCBI BLAST was available at ftp://ftp.ncbi.nlm.nih.gov/blast /executables/release/2.2.25/ (the installers were released on 1 January 2011), the BLAST program was used for searching AP2-like EST sequence fragments from the databases with the query sequences (the amino acid sequence of AtCBF1, AT4G25490.1), then the searched UniGenes were assembled using the CAP3 assembly program (http://pbil.univ-lyon1.fr/cap3.php), the candidate coatings with highly significant hits to known genes were selected using NCBI blast.

Sequence analysis of the AP2-like genes

The open reading frames (ORFs) were identified using ORF finder (http://www.ncbi.nlm. nih.gov/gorf/gorf.html) and GeneMark. HMM (http://opal.biology.gatech.edu/GeneMark/ gmhmm2_prok.cgi). The domain feature and functional properties of cDNA-deduced amino acid sequence were revealed with UniProt and NCBI Blast. The ClustalW2 online server (http://www.ebi.ac.uk/ Tools/msa/ clustalw2/) were used to determine sequence alignment and similarity between the AP2-like TFs in *Arabidopsis* and *M. truncatula*. The neighbor joining (NJ) phylogenetic tree of the aligned protein sequences was built and edited using MEGA 5.05.

Expression profile of AP2-like genes

The digital expression profiles derived from EST counts of UniGene (http://www.ncbi.nlm.nih.gov/UniGene/) were analyzed for the putative AP2-like genes in *M. truncatula*.

RESULTS

Isolation of AP2-like genes from *M. truncatula*

The numbers of AP2-like genes from *Arabidopsis* were consistent with Sakuma et al. (2002), which proved the reliability of the BLAST research methods. A total of 37 AP2-like genes containing a single AP2-domain were identified from *M. truncatula* database containing 231,739 sequences with the query sequence. Based on the structure analysis and sequence alignment of their amino acid sequences, 15 AP2-like genes encoding ERF-like proteins and 22 AP2-like genes encoding DREB-like proteins were identified as ERF subfamily and DREB subfamily respectively (Table 1).

Phylogenetic analysis between the AP2-like TFs in Arabidopsis and *M. truncatula*

Phylogenetic trees were inferred from alignments of *Arabidopsis* AP2/ERF family TFs and the AP2-like protein sequences of *M. truncatula*. 37 AP2-like genes were subdivided into ERF and DREB subfamily, which were distributed in B1, B2, B3, B4, B6 ERF subgroup or A1, A2, A4, A5, A6 DREB subgroup, respectively (Figure 1 and Table 1).

Comparative analysis between the AP2-like TFs in Arabidopsis and *M. truncatula*

Comparison of deduced amino acid sequence from AP2-like *TF* genes showed that most of the DREB and ERF subfamily proteins contained the YRG element and the WLG motif, which were conserved among all DREB and ERF subfamily proteins in both *Arabidopsis* and *M. truncatula* (Figure 2). Sequence alignment indicated that amino acid residues outside AP2 domain exhibit greater variation than those inside AP2 domain in Arabidopsis and *M. truncatula* (data not shown); however, AP2 subfamily gene containing duplicated AP2 domains, DREB-A3 and ERF-B5 subgroup gene were not found in *M. truncatula*.

Expression profile of *AP2-like* genes in *M. truncatula*

The *AP2-like* genes in *M. truncatula* were found in seven kinds of tissue with different distribution (seed 9.09%, stem 13.64% and flower 9.09%, glandular trichome 6.06%, leaf 15.15%, pod 7.58% and root 53.03%). The highest and lowest expression levels of AP2-like TFs were found in root and glandular trichome, respectively. Expression pattern of each in putative 35 AP2/ERF family of *M. truncatula* was revealed, except MtrERF-B6-1 and MtrDREB-A6-4 (Table 1).

The gene expression levels of most AP2-like genes are tissue specific, which were found in some tissues but not in others, only 11 genes were root specific (Figure 3). Thirty-three (33) AP2/ERF family genes with different expressive level were detected in root, except MtrDREB-A4-5 and MtrERF-B2-5; how-ever, only 4 AP2/ERF family genes were detected in glandular trichome; MtrDREB-A5-1 and MtrERF-B3-3 had the highest expression level in seed and pod of *M. truncatula*, respectively, only MtrERF-B2-1 was presented in six examined tissues (Figure 3).

Table 1. The AP2/ERF family member expression profiles suggested by analysis of EST counts based UniGene (Transcripts per millon, TPM). Database: Mtr.seq.all 231,739 sequences.

This research	UniGene number	Flower	Glandular trichome	Leaf	Pod	Root	Seed	Stem
MtrERF-B1-1	Mtr.2504	0	0	0	0	151	347	0
MtrERF-B1-2	Mtr.14925	0	0	59	0	58	0	0
MtrERF-B1-3	Mtr.7533	0	0	0	0	69	0	0
MtrERF-B2-1	Mtr.2743	439	0	148	614	570	695	76
MtrERF-B2-2	Mtr.9472	146	0	0	307	127	0	76
MtrERF-B2-3	Mtr.15517	0	224	0	0	221	115	0
MtrERF-B2-4	Mtr.16629	146	0	0	0	46	0	0
MtrERF-B2-5	Mtr.14998	16	0	0	29	0	58	0
MtrERF-B3-1	Mtr.6697	146	0	148	307	360	0	230
MtrERF-B3-2	Mtr.11998	0	0	118	0	139	0	0
MtrERF-B3-3	Mtr.16440	0	224	118	921	186	0	384
MtrERF-B3-4	Mtr.18262	0	0	0	0	58	0	0
MtrERF-B4-1	Mtr.22373	0	0	0	0	46	0	0
MtrERF-B6-1	Mtr.21747	/	/	/	/	/	/	/
MtrERF-B6-2	Mtr.2833	0	0	0	0	104	0	0
MtrDREB-A1-1	Mtr.6959	0	0	0	0	34	115	0
MtrDREB-A1-2	Mtr.11886	0	224	0	0	34	0	0
MtrDREB-A1-3	Mtr.12121	0	0	0	0	11	0	0
MtrDREB-A1-4	Mtr.22246	0	0	0	0	11	0	0
MtrDREB-A1-5	Mtr.10770	0	0	29	0	23	0	0
MtrDREB-A1-6	Mtr.23141	0	0	0	0	11	0	0
MtrDREB-A2-1	Mtr.20682	0	0	0	0	23	0	0
MtrDREB-A2-2	Mtr.13132	0	0	0	0	58	0	0
MtrDREB-A2-3	Mtr.14309	0	0	0	307	58	0	0
MtrDREB-A4-1	Mtr.5745	0	0	0	0	23	0	0
MtrDREB-A4-2	Mtr.12113	0	0	0	0	46	0	153
MtrDREB-A4-3	Mtr.12074	0	0	59	0	46	0	0
MtrDREB-A4-4	Mtr.8969	0	0	0	0	34	0	0
MtrDREB-A4-5	Mtr.22444	0	0	0	0	0	0	76
MtrDREB-A5-1	Mtr.8574	0	0	0	0	46	347	76
MtrDREB-A5-2	Mtr.8670	0	0	29	0	23	0	0
MtrDREB-A5-3	Mtr.12015	0	0	89	0	69	0	0
MtrDREB-A5-4	Mtr.17392	0	0	0	0	34	0	0
MtrDREB-A6-1	Mtr.6075	146	0	0	0	46	0	76
MtrDREB-A6-2	Mtr.8520	0	0	59	0	46	0	231
MtrDREB-A6-3	Mtr.9850	146	0	0	0	23	0	0
MtrDREB-A6-4	Mtr.23658	/	/	/	/	/	/	/

DISCUSSION

Legumes are one of the most important sources of animal feed, edible oil and industrial fuel. To improve legume abiotic and biotic stresses tolerance, researches have been focused on the physiological and molecular mechanisms involved in the response and tolerance to these stresses. An important goal in understanding the molecular basis and providing the genetic resources required for the legume genetic improvement is to isolate and identify members of the legume AP2/ERF family TFs. *M. truncatula* is a valuable legume model plant because of features such as relatively simple and small genomes, short life cycle and appropriate developmental responses for environmental stress enable the use of this plant in studying how a legume plant interacts with its environment (Young and Udvardi, 2009; Ray, 2008). The authors have used a bio-informatics-based approach to identify and characterize sequences of the AP2/EREBP family of transcription factors from the genome of *M. truncatula*, establish a phylogenetic comparison of these sequences, and have presented data on deduced protein

Figure 1. Phylogenetic tree of AP2-like TF proteins in *M. truncatula*. After alignment of amino acid sequences from the *M. truncatula* (MtrDREB and MtrERF) and Arabidopsis (ATDREB and ATERF), the phylogenetic tree inferred from comparative analyses of AP2 TF protein sequences in *M. truncatula* and Arabidopsis, MtrDREB and MtrERF are classified to B1, B2, B3, B4, B6 ERF subgroup and A1, A2, A4, A5 and A6 DREB subgroups, respectively (Table 1).

sequences, domain structure of this family of transcription factors and tissue-specific gene expression profiles.

Thirty seven (37) genes in *M. truncatula* were identified as possibly encoding putative AP2/ERF family factor. The number of total AP2/ERF family genes from *M. truncatula* (37) was lower 3.9-fold than those genes from *A. thaliana* (147) (Nakano et al., 2006), which might be due to a small genomes, or lack of the more divergent ESTs to identify the unknown members in *M. truncatula* AP2/ERF family TFs. Twenty two (22) genes were identified as possibly encoding DREB subfamily, fifteen genes were predicted to encode ERF subfamily. The interesting EST

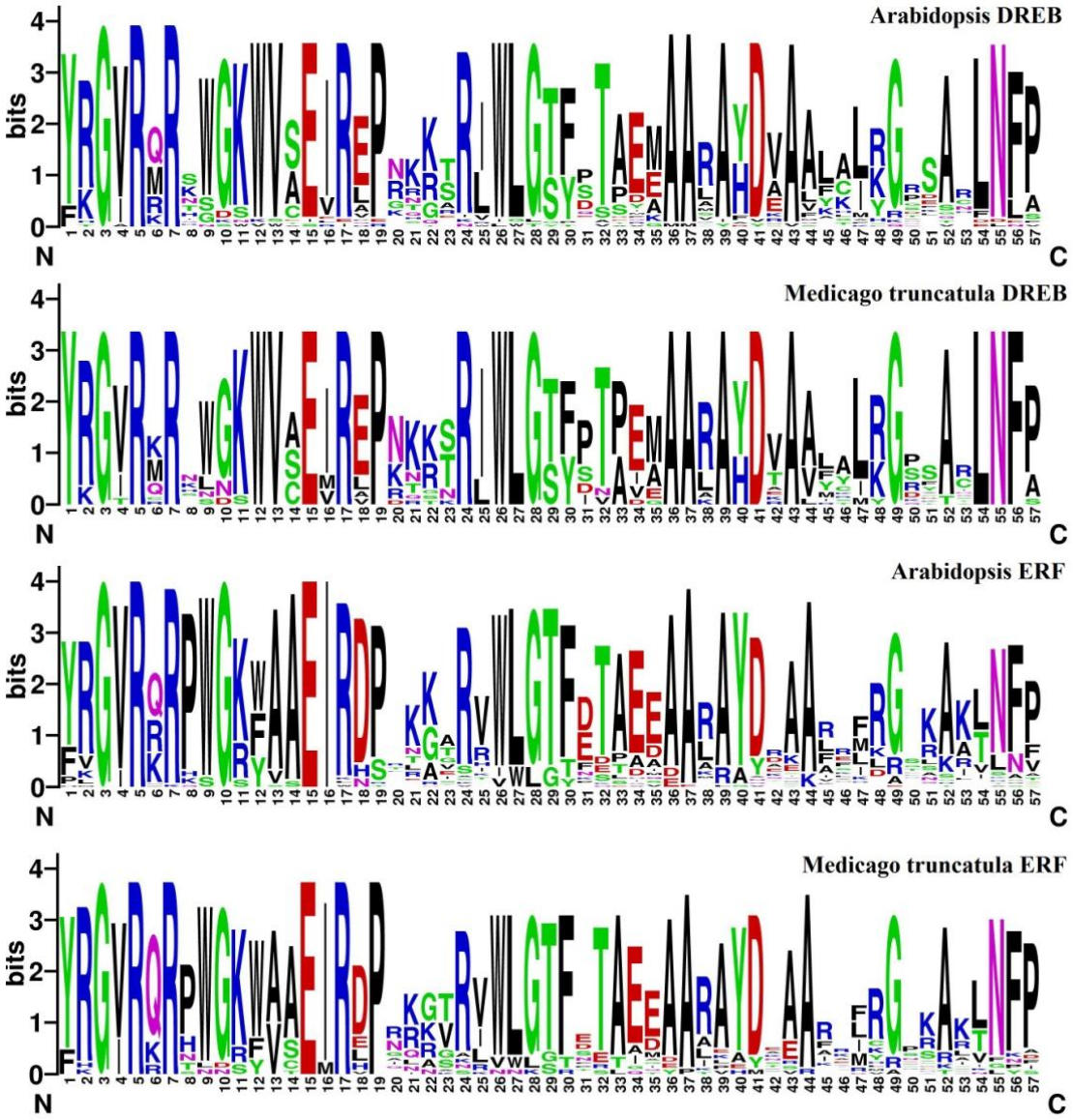

Figure 2. Web logo representation for AP2 TF protein sequences in *M. truncatula* and Arabidopsis. The height of each letter symbols reflects the sequence conservation (measured in bits), the relative frequency of the corresponding amino acid at sequence conservation at that position.

data in our study are useful resources for researchers to understand the regulatory function and evolution of *M. truncatula* AP2/ERF TFs. More recently, MtDREB1C containing an AP2 domain of 57 amino acids was isolated and characterized a *DREB* gene was a cold-acclimation-specific gene, exhibiting the highest homology to Arabidopsis *DREB1C* gene, and MtDREB1C from *M. truncatula* enhances freezing tolerance in transgenic *M. truncatula* and China Rose (Rosa chinensis Jacq.) (Pennycooke et al., 2008).

Most AP2/ERF TFs identified in the present study showed relatively high homology with the corresponding AP2/ERF TFs in Arabidopsis by sequence alignment. MtrERF-B3-2 (Mtr.11998) was ERF1A; MtrDREB-A1-3, MtrDREB-A1-5 and MtrDREB-A2-3 were DREB1A, DREB1C and DREB2A, respectively, based on Unigene name annotation in NCBI, which suggest that the four genes might belong to known genes, function in drought or salt stress resistance. Most *M. truncatula* AP2-like genes were determined to be transcribed locus, similar to AP2/ERF-domain transcription factor or predicted proteins in *Physcomitrella* patens subsp. patens, *A. thaliana* or *Populus trichocarpa*, respectively. MtrDREB-A4-5 (Mtr.22444), MtrERF-B3-4 (Mtr.18262), MtrERF-B3-3 (Mtr.16440), MtrERF-B3-1 (Mtr.6697), MtrERF-B2-5 (Mtr.15517) and MtrERF-B2-3 (Mtr.15517) were deter-

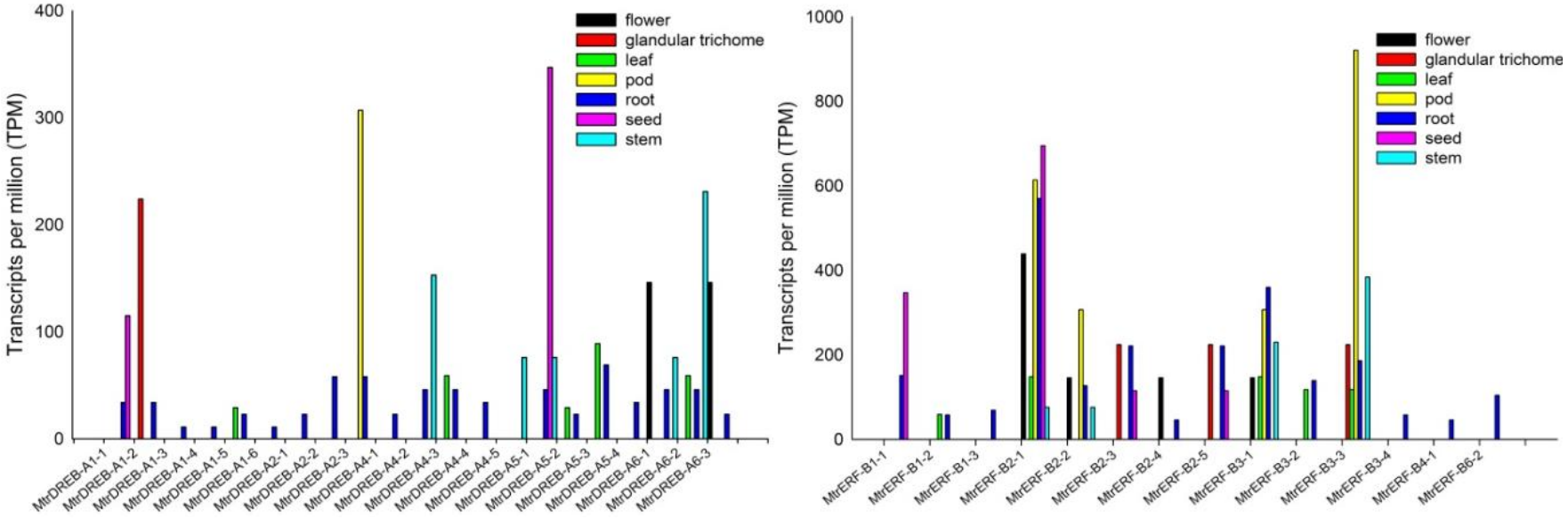

Figure 3. Tissue-specific expression of *M. truncatula* AP2-like genes. Gene expression of DREB and ERF subfamilies was represented as transcripts per million (TPM). AP2-like genes expression in 7 tissues (seed, stem, flower, glandular trichome, leaf, pod and root) that was indicated by different colors exhibited different tissue-specific expression profiles for *M. truncatula*.

mined to unknown MRNA without their putative function identified. These genes are to be worked out in future functional studies, since they may represent new genes not described and unique to *M. truncatula*.

In this research, to detect the pattern of expression of the 37 *M. truncatula* AP2/ERF family genes, the expression of these genes in seven tissues were analyzed according to the annotation of ESTs. The seven tissues have different abundance (about 6.06 to 53%). Recently, several AP2 genes (*MtCBF1*, *MtCBF2*, *MtDREB1C*/*CBF3*, *MtCBF4* and *MtDREB2A*) have been isolated from *M. truncatula*, which functioned in stress responses to high salt or drought. The gene expression analysis on tissue specificity showed that MtCBF1, MtCBF3 and MtCBF4s were expressed at relatively high levels in root (Benedito et al., 2008). Consistent with this gene expression pattern, our results revealed 33 MtAP2/ERF family genes with different expression level were detected in roots; moreover, 11 genes had root tissue-specificity. Very recently, a detailed transcriptomic analysis of the salt stress response of *M. truncatula* Jemalong A17 roots led to the identification of an AP2-EREBP TF able to enhance salt tolerance in Arabidopsis and *M. truncatula* roots. A specific AP2-EREBP TF, CBF4 was shown to play a significant role in most abiotic stresses, including drought, cold and salt using overexpression both in Arabidopsis and *M. truncatula* roots (Li et al., 2011).

Plants can cope with salt and osmotic stresses by modifying their root architecture in order to keep root growth under the stresses (Malamy, 2005). The AP2-EREBP TF expression in root might support the stress response. The gene expression is related to those stress tolerance and response; some genes are induced only by water stress, low temperature and others, some genes are induced by a combination of this stress (Nakashima et al., 2009; Yamaguchi and Shinozaki, 2006). Accurate gene expression is mainly due to TFs at the transcriptional level of control. These expression data results will be helpful in studying the AP2-like genes in *M. truncatula*. EST database contains a wealth of molecular biology information, which showed clear advantages in cloning AP2/ERF from

EST database. Making full use of the EST data resources should lay the foundation for studies of comparative genomics and functional genomics of *M. truncatula*. The information will guide our future study, including isolation and characterization of candidate AP2/ERF family genes in *M. truncatula* using PCR and real-time PCR.

ACKNOWLEDGEMENTS

This work was supported by National Nature Science Foundation of China (No. 31101761) and the Talent Fund of Hunan Agricultural University (No. 09WD17).

REFERENCES

Benedito VA, Torres-Jerez I, Murray JD, Andriankaja A, Allen S, Kakar K, Wandrey M, Verdier J, Zuber H, Ott T (2008). A gene expression atlas of the model legume *Medicago truncatula*. Plant J. 55: 504-513

Chen JR, Lu JJ, Wang TX, Chen SY, Wang HF (2009). Activation of a DRE-binding transcription factor from *Medicago truncatula* by deleting a Ser/Thr-rich region. In Vitro Cell Dev. Biol. Plant 45(1): 1-11.

Feng JX, Liu D, Pan Y, Gong W, Ma LG, Luo JC, Deng XW, Zhu YX (2005). An annotation update via cDNA sequence analysis and comprehensive profiling of developmental, hormonal or environmental responsiveness of the Arabidopsis AP2/EREBP transcription factor gene family. Plant Mol. Biol. 59:853-868

Li D, Zhang Y, Hu X, Shen X, Ma L, Su Z, Wang T, Dong J (2011). Transcriptional profiling of Medicago truncatula under salt stress identified a novel CBF transcription factor MtCBF4 that plays an important role in abiotic stress responses. BMC Plant Biol. 11:109

Malamy JE (2005) Intrinsic and environmental response pathways that regulate root system architecture. Plant Cell Environ. 28 67-77

Mochida K, Yoshida T, Sakurai T, Yamaguchi-Shinozaki K, Shinozaki K, Tran LSP (2009). In silico analysis of transcription factor repertoire and prediction of stress responsive transcription factors in soybean. DNA Res. 16:353-369

Nakano T, Suzuki K, Fujimura T, Shinshi H (2006). Genome-Wide Analysis of the ERF Gene Family in Arabidopsis and Rice. Plant Physiol. 40(2):411-32

Nakashima K, Ito Y, Yamaguchi-Shinozaki K (2009) Transcriptional regulatory networks in response to abiotic stresses in Arabidopsis and grasses. Plant Physiol. 149:88-95

Pennycooke JC, Cheng H, Stockinger EJ (2008). Comparative genomic sequence and expression analyses of *Medicago truncatula* and Alfalfa subspecies falcata COLD- ACCLIMATIONSPECIFIC genes. Plant Physiol. 146:1242-1254.

Ray JR (2008). Medicago truncatula as a model for understanding plant interactions with other organisms, plant development and stress biology: past, present and future. Funct. Plant Biol. 35: 253-264.

Sakuma Y, Liu Q, Dubouzet JG, Abe H, Shinozaki K, Yamaguchi-Shinozaki K (2002). DNA-binding specificity of the ERF/AP2 domain of Arabidopsis DREBs, transcription factors involved in dehydration- and cold-inducible gene expression. Biochem. Biophys. Res. Commun. 290:998-1009.

Yamaguchi SK, Shinozaki K (2006) Transcriptional regulatory networks in cellular responses and tolerance to dehydration and cold stresses. Annu. Rev. Plant Biol. 57:781-803.

Young ND, Udvardi M (2009) Translating Medicago truncatula genomics to crop legumes. Curr. Opin. Plant Biol. 12: 193-201.

Zhang G, Chen M, Chen X, Xu Z, Li L, Guo J, Ma Y (2010) Isolation and characterization of a novel EAR-motif-containing gene GmERF4 from soybean (*Glycine max* L.). Mol. Biol. Rep. 37(2): 809-818.

Zhang JY, Broeckling CD, Blancaflor EB, Sledge MK, Sumner LW, Wang ZY (2005) Overexpression of *WXP1*, a putative *Medicago truncatula* AP2 domain-containing transcription factor gene, increases cuticular wax accumulation and enhances drought tolerance in transgenic alfalfa (*Medicago sativa*). Plant J. 42(5):689-707.

Zhang JY, Broeckling CD, Sumner LW, Wang ZY (2007) Heterologous expression of two *Medicago truncatula* putative ERF transcription factor genes, WXP1 and WXP2, in Arabidopsis led to increased leaf wax accumulation and improved drought tolerance, but differential response in freezing tolerance. Plant Mol. Biol. 64(3):265-278.

Zhuang J, Cai B, Peng RH, Zhu B, Jin XF, Xue Y, Gao F, Fu XY, Tian YS, Zhao W, Qiao YS, Zhang Z, Xiong AS, Yao QH (2008) Genome-wide analysis of the AP2/ERF gene family in *Populus trichocarpa*. Biochem. Biophys. Res. Commun. 371(3):468-474.

Zhuang J, Chen JM, Yao QH, Xiong F, Sun CC, Zhou XR, Zhang J, Xiong AS (2011a). Discovery and expression profile analysis of AP2/ERF family genes from *Triticum aestivum*. Mol. Biol. Rep. 38(2): 745-753.

Zhuang J, Sun CC, Zhou XR, Xiong AS, Zhang J (2011b). Isolation and characterization of an AP2/ERF-RAV transcription factor BnaRAV-1-HY15 in *Brassica napus* L. HuYou15. Mol. Biol. Rep. 38(6):3921-3928.

.