Journal of Medicinal Plants Research Vol. 6(6), pp. 950-958, 16 February, 2012
Available online at http://www.academicjournals.org/JMPR
DOI: 10.5897/JMPR11.240
ISSN 1996-0875 ©2012 Academic Journals

Full Length Research Paper

Transcriptional regulatory networks in response to salt and drought stress in *Arabidopsis thaliana*

ZHANG Li[1], LI Qiaoying[2], SHEN Jie[1], XUE Jinai[1] and HAN Yuanhuai[1]*

[1]Agricultural Bioengineering Centre, School of Agriculture, Shanxi Agricultural University, Taigu, Shanxi 030801, China.
[2]Department of Agriculture of Shanxi Province, Taiyuan, Shanxi, 030006, China.

Accepted 27 December, 2011

Many transcription factors are involved in the progress of stress-inducible regulation. Transcription factors play an important role not only in stress tolerance but also in stress response. Some of the transcription factors and their target genes, whether in tolerance or in response, take part in the pathway metabolism. To explain the relationship between the transcription factor and target gene in *Arabidopsis thaliana*, gene regulation networks under salt and drought stresses were constructed. The regulation network showed that the transcription factor WRKY53 played a key role in the regulatory network through regulating ATWRKY18 and GBF3. The overlap of salt and drought stress regulation networks showed that degradation-related pathways were repressed, while alpha-linolenic acid and phenylpropanoid pathways were activated .

Key words: Regulation network, *Arabidopsis thaliana*, stress, pathway.

INTRODUCTION

Salinity and drought stress

Environmental stresses, such as drought and high salinity, have adverse effects on plant growth and production. Plants respond and adapt to these stresses through various biochemical and physiological processes, thereby acquiring stress tolerance. Salinity is a major environmental stress and a substantial constraint to crop production. High salinity causes both hyperionic and hyperosmotic stress and can lead plant to demise. High salinity depositions in the soil generate a low water potential zone in the soil making it increasingly difficult for the plant to acquire both water and nutrients. Therefore, salt stress essentially results in a water deficit condition in the plant and takes the form of a physiological drought (Mahajan and Tuteja, 2005). Compared to salt stress, the problem of drought is even more pervasive and economically damaging. Most studies on water stress signaling have focused on salt stress primarily because plant's responses to salt and drought are closely related and the mechanisms overlap. From a practical point, salt stress can be imposed more easily and precisely in laboratory settings. In drought stress responses, guard cell signaling is of critical importance because it is a key denominator within the plant water budget. Much effort has been justifiably dedicated to guard cell signaling and substantial advances have been made. Several hundred genes which respond to these stresses at the transcriptional level have been identified, so did the products of these genes' function under stress (Kreps et al., 2002; Seki et al., 2002; Xiong et al., 2002; Zhu, 2002; Shinozaki et al., 2003; Matsui et al., 2008), and several cis- and trans-acting factors family, such as DRE/CRT, ABRE and MYCRS/ MYBRS, involve in stress-response (Mahajan and Tuteja 2005). Salinity and drought elicit many common and interactive downstream effects

*Corresponding author. Email: swgctd@163.com. Tel: +86-0354-6286-399.

Abbreviation: ABA, Abscisic acid; **TFs,** Transcription factors; **AGRIS,** Arabidopsis Gene Regulatory Information Server; **DEGs,** Differentially expressed genes; **PCC,** Pearson correlation coefficient; **DREB2,** dehydration response element binding factor 2; **ERF,** ethylene response factor; **AP2,** APETALA2; **PPI,** protein protein interaction; **CK,** Cytokinin; **GSH,** glutathione; **CORN,** co-regualtion network under salt and drought stress; **WARGNW,** ATWRKY18 (AT4G31800) and AT5G01380 regulatory network.

Table 1. Data source of microarray data, pathway data and regulation data.

Source	Description	Link
KEGG	130 pathway involved 2287 proteins	http://www.genome.jp/kegg/
AthaMap	Involved 86 TF and more than 5000 target genes	http://www.AthaMap.de/
AGRIS	Involved 85 TF and 10255 target genes, total 13266pairs (contain our collected)	http://arabidopsis.med.ohio-state.edu/
Microarray	ATH1 platform, total 148 chips	http://arabidopsis.org/portals/expression/microarray/ ATGenExpress.jsp http://www.ncbi.nlm.nih.gov/geo/

(Shinozaki et al., 2003). For example, drought and salt stresses activate dehydration response element binding factor 2 (DREB2), members of the ethylene response factor (ERF)/APETALA2 (AP2) transcription factors family. DREB2 binds CRT/DRE promoter elements in stress response genes (Gosti et al., 1995; Yamaguchi-Shinozaki and Shinozaki, 2006). Microarray technology employing cDNAs or oligonucleotides is a powerful tool for analysing gene expression profiles of plants exposed to abiotic stresses, such as drought, high salinity (Narusaka et al., 2001), and it has been developed to infer gene regulatory network (Wang et al., 2007). For example, MADS-box and WRKY transcription factor families were found that they could be induced by salt treatment in tomato upon the microarray analysis (Zhou et al., 2007). Regulatory network analysis revealed that the interactions between different transcription machineries function proved that cross-talk could be exited between different stress signaling pathways (Shinozaki et al., 2003) and between different stresses (Tran et al., 2007).

DATAS AND METHODS

Affymetrix microarray data

Three transcription profiles of Arabidopsis under salt and drought stress were obtained from AtGenExpress (http://arabidopsis.org/portals/expression/microarray/ATGenExpress.jsp) and GEO (http://www.ncbi.nlm.nih.gov/geo/). They all are the Affymetrix ATH1 platform data.

The dataset ME00328 is extracted from AtGenExpress. This experiment studies the effects of continuous salt stress on gene expression of Arabidopsis treated with 150 mM NaCl. Total 52 slides are divided into 26 sets, and in each set, two samples serve as replicates. The dataset ME00338 is also extracted from AtGenExpress.

This experiment studies the effects of 15 min drought stress on gene expression of Arabidopsis. Total 60 slides are divided into 30 sets, and in each set, two samples serve as replicates. The dataset GSE7641 is extracted from GEO. This experiment analyzes the root cell-types after treated with 140 mM NaCl for 1 h. Total 36 slides are divided into 12 sets, and in each set, three samples serve as replicates.

Pathway data

Kyoto Encyclopedia of Genes and Genomes (KEGG) is a collection of online databases dealing with genomes, enzymatic pathways, and biological chemicals. The pathway database records networks of molecular interactions in the cells, and variants of them specific to particular organisms (http://www.genome.jp/kegg/). Total 130 pathways, involving 2287 genes, were collected from KEGG.

Regulation data

There are approximately 1,770 transcription factors in the Arabidopsis. These transcription factors are grouped into 50 families, based on the presence of conserved DNA-binding domains. More than 5000 target genes supposed to have at least one binding site of total 86 transcription factors (TFs) (which can be mapped to the AGI locus number of TAIR), including total 11,137 interactions, were collected from the AthaMap (http://www.AthaMap.de/) database (Bulow et al., 2009). 13041 interactions between 69 transcription factors (TFs) and 9423 target genes (Palaniswamy et al., 2006) were collected from Arabidopsis Gene Regulatory Information Server (AGRIS) (http://arabidopsis.med.ohio-state.edu/). 7758 pairs of regulatory relationship were collected manually. Combined the two regulation datasets, total 13266 regulatory relationships between 85 TFs and 10255 target genes were collected (Table 1).

Differentially expressed genes (DEGs) analysis

For the ME00328 and ME00338 datasets, we directly used the differentially expressed genes (DEGs) list proposed by the authors (Ma and Bohnert, 2007).

For the GSE7641 dataset, the limma method (Smyth, 2004) was used to identify DEGs. The original expression datasets from all conditions were processed into expression estimates using the RMA method with the default settings implemented in Bioconductor, and then construct the linear model. The DEGs with the fold change value larger than 2 were selected, and 505 pathway-related genes were kept.

Co-expression analysis

To demonstrate the potential regulatory relationship, the Pearson correlation coefficient (PCC) was calculated for all pair-wise comparisons of gene-expression values between TFs and the DEGs. The regulatory relationships whose absolute PCC are larger

than 0.7 were considered significant.

Regulation network construction

Base on the significant relationships (PCC > 0.7 or PCC < -0.7) between TFs and its target genes, which are pathway-related DEGs, we predicted the possible binding sites on the target gene promoter regions with AthaMap. Using these criteria, 2015 putative regulatory relationships were predicted between 70 TFs on 382 target genes.

Using the AGRIS regulation data that have been experimentally verified, we collected the relationships between differentially expressed TFs and its differentially expressed target genes. Base on the aforementioned two regulation datasets and the pathway regulationship of the target genes, we build the regulation networks by Cytoscape (Shannon et al., 2003).

Gene ontology analysis

The BiNGO analysis (Maere et al., 2005) was used to identify over-represented Several gene ontology (GO) categories in biological process.

Significance analysis of pathway

First, the DEGs were mapped to the pathways. Then differential pathways were selected through the Fisher exact test (Francesconi et al., 2008). The Fisher exact test computes the probability p* by using the hypergeometric distribution with parameters (M, n, N):

$$p^* = p(X = \alpha \mid M, n, N) = \frac{\binom{M}{m}\binom{N-M}{n-m}}{\binom{N}{n}} \quad (1)$$

m= number of significant genes in one pathway; M= number of significant genes in the array; N= total number of measured genes and n= number of genes in one pathway. The p value to reject the null hypothesis is given by the sum of the probabilities of all the probability lower than p*, that is:

$$p-value = \sum_{\substack{m'=1,n \\ p^* < p}} \frac{\binom{M}{m'}\binom{N-M}{n-m'}}{\binom{N}{n}} \quad (2)$$

If p-value ≤ 0.05, the pathway is considered as significantly expressed.

RESULTS

Regulation network construction under salt stress

To get pathway-related DEGs of Arabidopsis thaliana under high salinity stress, we obtained publicly available microarray data sets GSE7641 (from GEO) and ME00328 (from AtGenExpress). After microarray analysis, we got 2 groups of DEGs, that is, 2244 genes from GSE7641 and 3373 from ME00328. To get the regulatory relationships, we selected the co-expressed value (PCC ≥0.7) as the threshold. To further verify the relationships, we mapped all of the TFs' binding sites to the DEGs' promoter region by using AthaMap.

Finally, we got 24 regulatory relationships between different expressed TFs and its' differently expressed target genes (269) (Figure 1A). The differently expressed target genes were mapped to pathways (232). By integrating the regulatory relationships in the foregoing, a salt stress regulation network was build between TFs and its target genes (Figure 1A). To further investigate the regulatory relationships between TFs and pathways, we mapped DEGs to pathways and got a regulation network between TFs and pathways (Figure 1B). A displays the regulation between TF and their target genes. B displays the regulation between TF and pathway. These two graphs show that the regulation network of certain TFs to target genes or pathways in salt stress. In the graph, triangle spot denotes TF, round spot denotes target gene, square spot denotes pathway; red edge denotes activated target genes or pathways, vice versa, green edge denotes repression.

Several gene ontology (GO) analysis of the regulation network under salt stress

Several gene ontology (GO) categories were enriched among these genes in the regulatory network, including response to stimulus, response to stress and metabolic processes (Table 2).

Regulation network construction under drought stress

To get pathway-related DEGs of Arabidopsis thaliana under high salinity stress, we obtained publicly available microarray data ME00338 (from AtGenExpress). After microarray analysis, to get the regulatory relationships, we selected the co-expressed (PCC ≥ 0.7) TFs and DEGs. To further verify the relationships, we mapped all of the TFs' binding sites to the DEGs' promoter region (using AthaMap).

By integrating the pathway DEGs and the regulatory relationships in the foregoing, we build a drought stress regulation network between TFs and its target genes (Figure 2A). To further investigate the regulatory relationships between TFs and pathways, we mapped DEGs to pathways and got a regulation network of TFs to pathways (Figure 2 B). Figure 2A displays the regulation between TF and their target genes. Figure 2B displays the regulation between TF and pathway. These two graphs show that the regulation network of certain TFs to target genes or pathways in drought stress. In the graph, triangle spot denotes TF, round spot denotes target

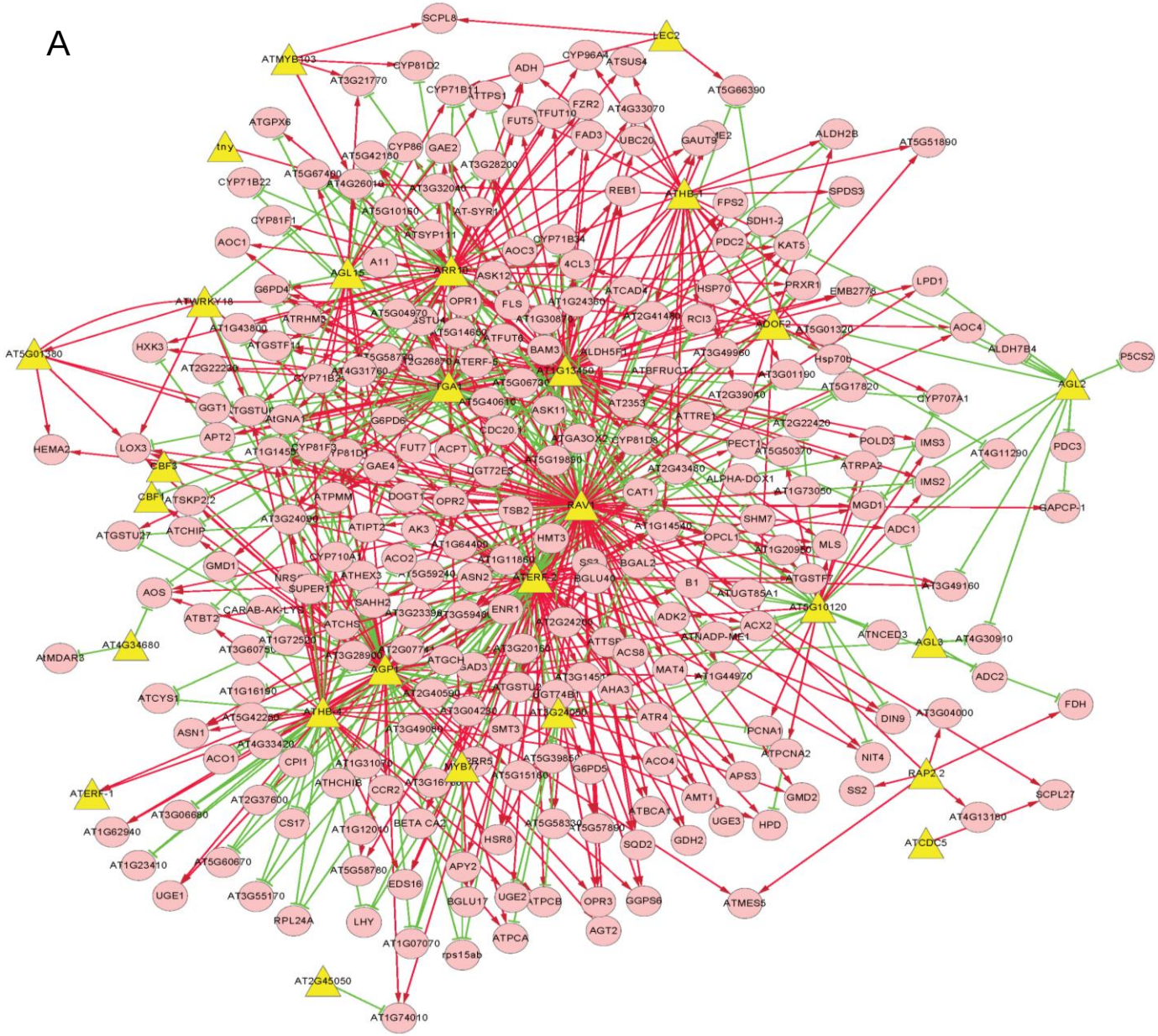

Figure 1. Regulation network analysis under high salinity stress.

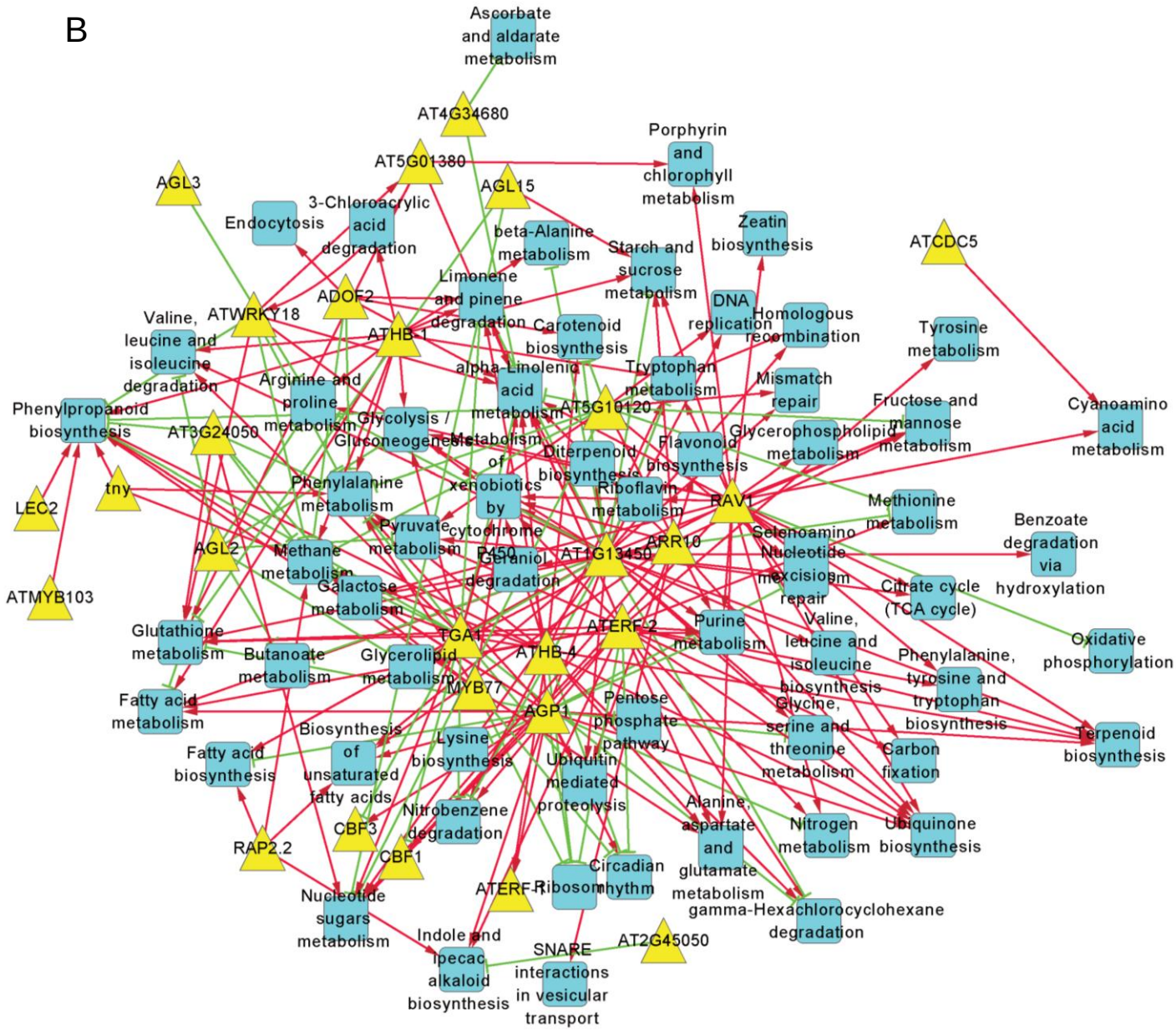

Figure 1. Contd.

Table 2. GO analysis of regulation network genes under salt stress.

GO-ID	GO category	p-value	Description
51869	BP	1.24E-33	response to stimulus
19752	BP	1.43E-33	carboxylic acid metabolic process
6082	BP	1.43E-33	organic acid metabolic process
6950	BP	3.04E-33	response to stress
6979	BP	2.72E-30	response to oxidative stress
9058	BP	1.91E-28	biosynthetic process
8152	BP	2.57E-28	metabolic process
44237	BP	2.66E-28	cellular metabolic process
44255	BP	5.10E-28	cellular lipid metabolic process
42221	BP	1.58E-27	response to chemical stimulus

Note: BP means biology process.

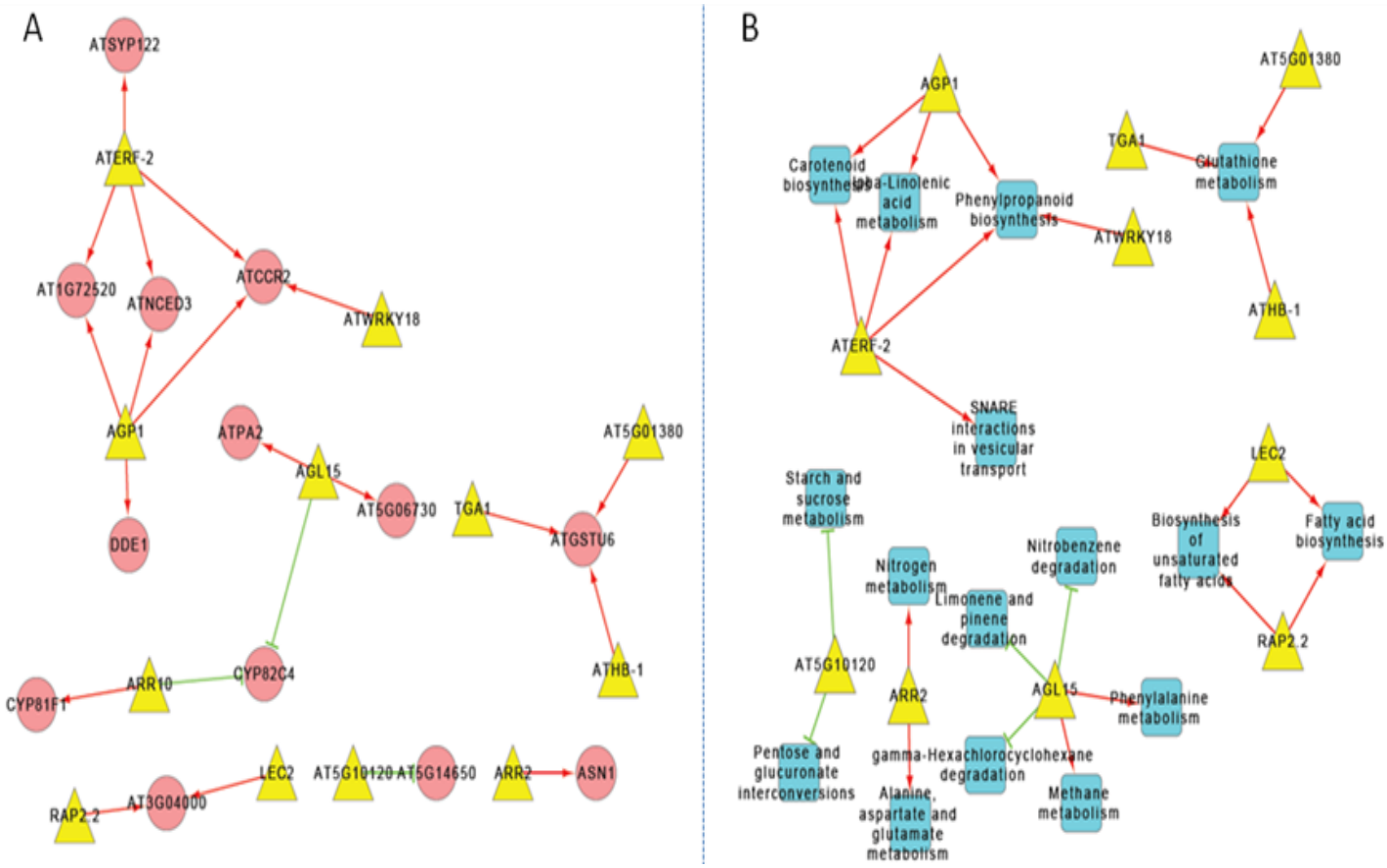

Figure 2. Regulation network construction under drought stress.

gene, square spot denotes pathway; red edge denotes activated target genes or pathways, vice versa, green edge denotes repression. Several gene ontology (GO) categories of drought stress were similar to salt stress (Table 3).

The conserved regulatory network under salt and drought stresses

Co-regualtion network under salt and drought stress (CORN) is the regulation network of salt and drought

Table 3. GO analysis of regulation network genes under drought stress.

GO_ID	GO Category	p-value	Description
GO:0006950	BP	0.000232	response to stress
GO:0008152	BP	0.00452	metabolic process
GO:0055114	BP	0.00475	oxidation-reduction process

Note: BP means biology process.

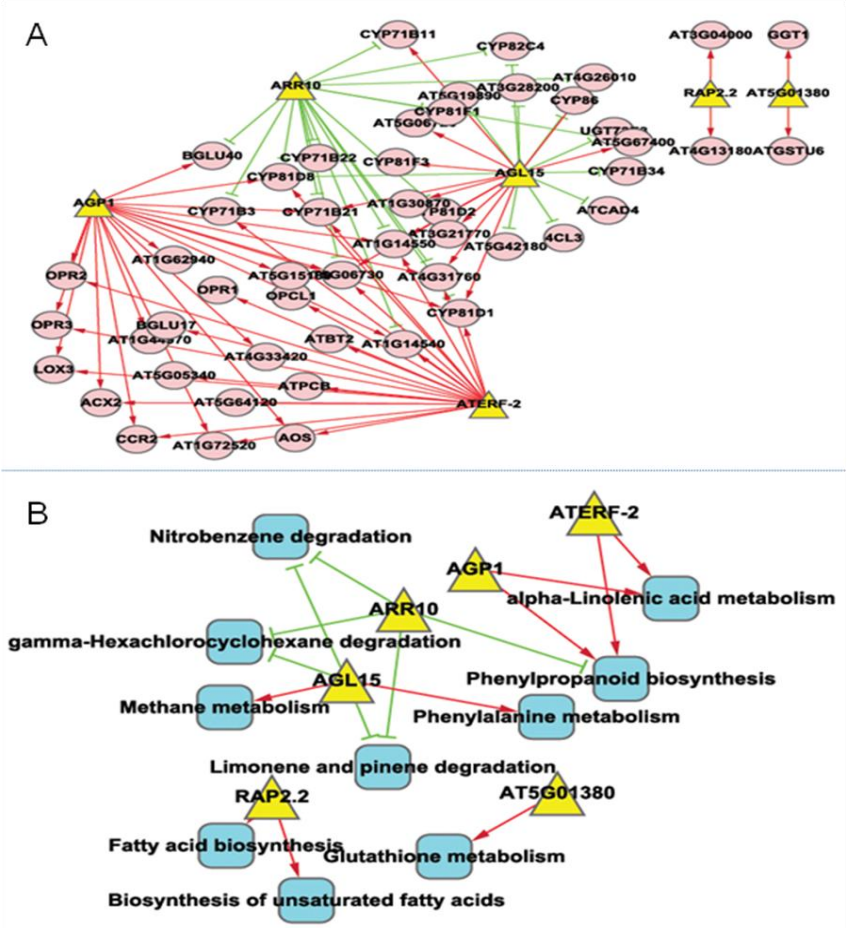

Figure 3. Co-regualtion network under salt and drought stress (CORN).

stresses. A displays the regulatory relationships of TFs and target genes; B displays the regulatory relationships of TFs and pathways. Triangle denotes TF, circle denotes target gene, square denotes pathway, red edges denotes activation and green edges denotes repression (Figure 3).

Significant analysis of pathway

We calculated the significance of each pathway by using the P value by the formula (1) and (2). Tables 4 and 5 show the result of pathway significant analysis.

Through the pathway shown in Tables 4 and 5 significant analyses, we found that 5 (ath00190, ath00230, ath00592, ath00901, and ath00980) pathways under salt stress in our salt stress regulation network are all stress related.

DISCUSSION

Drought and salt stresses all stimulate the accumulation of compatible osmolytes and antioxidants in plants

Table 4. Significant pathways under the salt stress.

Pathway	Pathway names	Genes	DEGs ME00328	P value ME00328	DEGs GSE7641	P value GSE7641
ath00190	Oxidative phosphorylation	141	4	3.00E-07	6	0.022432
ath00195	Photosynthesis	74	2	0.0004	1	0.009348
ath00230	Purine metabolism	105	9	0.0249	4	0.032424
ath00592	alpha-Linolenic acid metabolism	29	15	1.00E-05	11	5.37E-05
ath00901	Indole and ipecac alkaloid biosynthesis	8	5	0.0045	5	0.000403
ath00980	Metabolism of xenobiotics by cytochrome P450	20	9	0.0027	6	0.010497

Note: Table 5 List the overlap pathway of GSE7641 and ME00328 of pathway significant analysis. The pathways which are bold also exited in the regulation network in the salt stress based on the AthaMap.

Table 5. ME00338 pathway significant analysis.

Pathway	Pathway names	Genes	DEGs ME00338	P value ME00338
ath00190	Oxidative phosphorylation	141	1	0.01102
ath00592	alpha-Linolenic acid metabolism	29	11	7.38E-09
ath00901	Indole and ipecac alkaloid biosynthesis	8	5	1.75E-05
ath00980	Metabolism of xenobiotics by cytochrome P450	20	5	0.002885

Note: The bold pathways exit in the co-regulation network in the overlap of salt and drought stress based on the AthaMap.

(Hasegawa et al., 2000). Salinity and drought are among the major stresses, which adversely affect plants growth and productivity (Mahajan and Tuteja, 2005). Both drought and salt stress ultimately result in dehydration of the cell and osmotic imbalance. Virtually, every aspect of plants, from physiology to cellular metabolism is affected by salt and drought stresses (Mahajan and Tuteja, 2005). Function of stress response gene and the relationship between TF and its target gene can be explained in regulatory network. By integrating the salt and drought regulatory relationship data, we constructed the consensus regulatory network (Figure 3). Our network displays consistent regulatory relationships in both salt and drought stresses. In the Figure 3B, there are10 pathways: ath00940 (Phenylpropanoid), ath01040 (Biosynthesis of unsaturated fatty acids)，ath00903 (Limonene and pinene degradation), ath00626 (Nitrobenzene degradation), ath00592 (alpha-Linolenic acid metabolism), ath00680 (Methane metabolism), ath00361 (gamma- Hexachlorocyclohexane degradation), ath00480 (Glutathione metabolism), ath00061 (Fatty acid biosynthesis), and ath00360 (Phenylalanine metabolism)), and 6 TFs: AT3G14230 (RAP2.2), AT4G31920 (ARR10), AT5G01380, AT5G13790 (AGL15), AT5G47220 (ATERF-2), AT5G64310 (AGP1) are in the Figure 3B. AP2/EREBP family TF ATERF2 (AT5G47220), ethylene response element, and RAP2.2 (AT3G14230) are induced by dehydration in the drought stress (Liu et al., 1998). GATA transcription factors family, to which AGP1 (AT5G64310) belong, is related to stress. The family GATA target genes respond to stress in tobacco (Sugimoto et al., 2003). AGL15 (AT5G13790) of MADS family, related to plant growth and development, resopnses to cold stress (Arora et al., 2007; Tardif et al., 2007). ARR10 (AT4G31920) is related to cytokinin (CK) (Yokoyama et al., 2007), and CK response to stress (Tran et al., 2010). AGP1 and ATERF-2 co-activate the alpha-Linolenic acid metabolism and Phenylpropanoid pathway. AGL15 and ARR10 co-repress Limonene and pinene degradation, Nitrobenzene degradation and gamma-Hexachlorocyclohexane degradation pathway, and AGL15 activate methane and phenylalanine pathways. RAP2.2 activate fatty acid biosynthesis and unsaturated fatty acids pathways. AT5G01380 activate glutathione pathway. GO function analysis of CORN shows that, besides some basic metabolism, most of genes in the network are enriched in response to stress and stimulus. Some TFs co-regulated one or more pathways, which indicate that the network has the crosstalk between pathways (Seo and Koshiba, 2002; Shinozaki et al., 2003).

According to the regulatory network discussed previously, we could clearly see that ATWRKY53 plays a critical role by regulating ATWRKY18 and GBF3 in response to salt stress. Through regulating the other 3 TFs, ATWRKY53 participate in regulating all the downstream pathways in the regulatory network. The plant is a complex organism, and pathways responding to stress, which are regulated by TFs, may be more complicated. The pathway of glutathione, alpha-linolenic acid and cytochrome P450 related directly or indirectly are regulated by TFs to response to stress. In our

research of the salt stress, Trihelix family and WRKY family have the same regulatory relationship and activate each other. They both response to downstream genes ATGSTU6 and LOX3.

AGP1 and ATERF-2 co-actiavte alpha-Linolenic acid pathway and phenylalanine pathway, and these two pathways response to stress (Dixon and Paiva, 1995; Katsoulieris et al., 2009). In our research, we predicated that the Garp family, which belongs to ARR10, response to stress by influencing the CK metabolism. Recognizing the cross-talk between different pathways will provide information useful to elucidate unknown regulation networks (Ma et al., 2006). Stress imposes injury on cellular physiology and result in metabolic dysfunction. This injury imposes a negative influence on cell division and growth of a plant. This is an indirect advantage to the plant as reduction of leaf expansion reduces the surface area of leaves exposed for transpiration and thereby reducing water loss. That causes the osmotic imbalance. Stress injury and ROS that are generated in response to stress also triggers a detoxification signaling by activating genes responsible for damage control and repair mechanism therefore leading to stress tolerance (Mahajan and Tuteja, 2005). Meanwhile, some important pathways are not in our networks. One reason is the limitation of our collected data, including pathways, regulatory relationships and microarrays data. More than 10,000 regulatory relationships are in AGRIS, but only 85 TFs were involved. In AthaMap database, there are only 86 TFs involved (could match to AGI numbers). Till now, there are approximately 1,770 transcription factors in Arabidopsis. These transcription factors are grouped into 50 families, based on the presence of conserved DNA-binding domains (Palaniswamy et al., 2006). But the intersection of these two databases is only the 12 TFs. In this paper, TFs are researched less than 1%. Results that lots of classical pathways and others important information do not belong to our research areas. The networks were constructed on the intersection of different datasets, which can help to screen out a lot of useful information. But the intersection of data could be more robust in the stress related regulation networks. We found that some pathways do not emerge in our regulation network, one reason is that the regulation network we selected was the overlap of 2 or 3 groups of microarrays. Different data sources, different DEGs that cause DEGs in the same pathway may be different.

The basic understanding of the mechanisms underlying the functioning of stress genes is important for the development of transgenic plants. Each stress is a multigenic trait and therefore their manipulation may result in alteration of a large number of genes as well as their products. A deeper understanding of the transcription factors regulating these genes, the products of the major stress responsive genes and cross talk between different signaling components should remain an area of intense research activity in future (Mahajan and Tuteja, 2005).

Our regulation network is useful in investigating the complex interacting mechanisms of cellular metabolic pathways in response to stresses. Some TFs interactively regulate downstream genes to respond to salinity and drought. We predicted that pathways induced by certain stress have cross-talk through the similar TFs. In the future, a combination of molecular, genomic and genetic analyses will be used to elucidate the complex systems that regulate the responses of gene expression to stresses.

ACKNOWLEDGEMENT

This paper was supported by “The Research Project for the Returned Overseas Chinese Scholars of Shanxi (2011-045)”.

REFERENCES

Bulow L, Engelmann S, Schindler M, Hehl R (2009). AthaMap, integrating transcriptional and post-transcriptional data. Nucl. Acids Res., 37(Database issue): D983-986.

Francesconi M, Remondini D, Neretti N, Sedivy JM, Cooper LN, Verondini E, Milanesi L, Castellani G (2008). Reconstructing networks of pathways via significance analysis of their intersections. B.M.C. Bioinformatics, 9 Suppl 4: S9.

Ma S, Bohnert HJ (2007). Integration of Arabidopsis thaliana stress-related transcript profiles, promoter structures, and cell-specific expression. Genome Biol., 8(4): R49.

Ma S, Gong Q, Bohnert HJ (2006). Dissecting salt stress pathways. J. Exp. Bot., 57(5): 1097-1107.

Maere S, Heymans K, Kuiper M (2005). BiNGO: a Cytoscape plugin to assess overrepresentation of gene ontology categories in biological networks. Bioinformatics, 21(16): 3448-3449.

Mahajan S, Tuteja N (2005). Cold, salinity and drought stresses: an overview. Arch. Biochem. Biophys., 444(2): 139-158.

Palaniswamy SK, James S, Sun H, Lamb RS, Davuluri RV, Grotewold E (2006). AGRIS and AtRegNet. a platform to link cis-regulatory elements and transcription factors into regulatory networks. Plant Physiol., 140(3): 818-829.

Seki M, Narusaka M, Abe H, Kasuga M, Yamaguchi-Shinozaki K, Carninci P, Hayashizaki Y, Shinozaki K (2001). Monitoring the expression pattern of 1300 Arabidopsis genes under drought and cold stresses by using a full-length cDNA microarray. Plant Cell, 13(1): 61-72.

Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker T (2003). Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res., 13(11): 2498-2504.

Shinozaki K, Yamaguchi-Shinozaki K, Seki M (2003). Regulatory network of gene expression in the drought and cold stress responses. Curr. Opin. Plant Biol., 6(5): 410-417.

Smyth GK (2004). Linear models and empirical bayes methods for assessing differential expression in microarray experiments. Stat. Appl. Genet. Mol. Biol., 3: 3.

Tran LS, Nakashima K, Shinozaki K, Yamaguchi-Shinozaki K (2007). Plant gene networks in osmotic stress response: from genes to regulatory networks. Methods Enzymol., 428: 109-128.

Tran LS, Shinozaki K, Yamaguchi-Shinozaki K (2010). Role of cytokinin responsive two-component system in ABA and osmotic stress signalings. Plant Signal Behav., 5(2) :148-150.

Zhou S, Wei S, Boone B, Levy S (2007). Microarray analysis of genes affected by salt stress in tomato. Afr. J. Environ.l Sci. Technol., 1(2): 014-026.