Expression Network Analysis of Abiotic Stress Responsive Myb in Rice

Shuchi Smita [1, 2] , Amit Katiyar[1, 2], Dev Mani Pandey[2], Viswanathan Chinnusamy[3], Kailash Chander Bansal[1]*[1]

[1]National Research Centre on Plant Biotechnology, Indian Agricultural Research Institute Campus, New Delhi-110012, India

[2]Department of Biotechnology, B.I.T. Mesra, Ranchi, Jharkhand, India

[3]Department of Botany and Plant Sciences, 2150 Batchelor Hall, University of California, Riverside, USA

Abstract - *In post-genomic era, bioinformatic tools allow us to explore and reconstruct the precise gene interaction network. To deduce the function of uncharacterized gene, genetic network by co-regulatory analysis from an expression data is a foremost approach. In this study, we report comprehensive identification of co-expressed MYB gene modules in rice. MYB transcription factor family is involved in phenylpropanoid and flavonoid biosynthesis and various other metabolic and developmental processes. By a reiterative database exploration, 249 potential OsMYB genes were retrieved. Computational analysis has shown the presence of several other functional domains including WD domain, G-beta repeat, response regulator receiver domain, BTB/POZ domain, SWIRM/Zinc finger domain and many more. Several studies have pointed out their involvement in a range of biological processes, revealing that a large number of MYB genes are transcriptionally regulated under conditions of biotic and/or abiotic stress. To investigate the existence of MYB co-regulatory network, a whole genome MYB expression study was carried out in rice. We identified the existence of co-expression clusters comprising phylogenetically related MYB genes, suggesting that specific sets of MYB genes might act in co-regulatory network. Thus, the co-expression networks identified in this study illustrate gene cooperation pathways that have not been identified by classical genetic.*

Keywords*: MYB gene, clusters analysis, Oryza Sativa*

1 Introduction

These are Plant growth and development are regulated by the coordinated expression of thousand of genes. To infer the function of uncharacterized genes, coexpression analysis of gene-to-gene is a useful approach. Regulation of gene expression is highly complex process, influenced by genotype environment interactions. The huge biological information available publically forms a foundation for system biology study nowadays [2]. System networks are often analyzed using visualization and analysis of network to deduce gene function, pathway components and links between and genes [1]. Network can be analyzed by direct and module based methods as in graph [8]. On the basis of gene-to-gene correlation coefficient derived from microarray hybridization data, cluster-based analysis give the idea of co-expressed gene or connections between genes that respond simultaneously to various stimuli [7]. For network study, expression profiling data are seems as highly useful resource. Microarray gene expression data is analyzed by a variety of bioinformatics techniques. In addition to commonly studied gene-specific expression patterns, gene expression analysis can be used to elucidate module and system-level organization of the transcriptome. Gene clustering method for module detection based on similarity of expression levels in different set of condition (gene co-expression networks) were used in many studies [9]. Several clustering algorithms have been developed for this purpose. Here, we report comprehensive identification of coexpression gene modules of *MYB* genes in rice. Several studies have indicated that *MYB* significantly involved in stress induced responses in *Arabidopsis thaliana* and other plants also. Several studies have pointed out their involvement in a range of biological processes, revealing that a large number of *MYB* genes are transcriptionally regulated under biotic and/or abiotic stresses [6]. In computational biology, use of network has greatly changed the analytical ability of researcher. In the present studies an attempt has been made to study this relationship of *MYB* genes in rice under abiotic stresses conditions. In our study, analysis of the *MYB* genes expression helped us to understand the cellular process where they involved, their interaction with other genes and their products.

2 Materials and Methods

2.1 *MYB* domain identification and Phylogenetic Analysis

Myb domain was retrieved by searching for PFAM ID PF00249 as a query in Rice at TIGR (http://rice.plantbiology.msu.edu/). Only the longest one was saved, when more than one alternative splicing sequence was found for the same locus. Phylogenetic tree for *MYB* proteins were constructed by iTOL (http://itol.embl.de/) to know the conserved pattern between rice *MYB* genes.

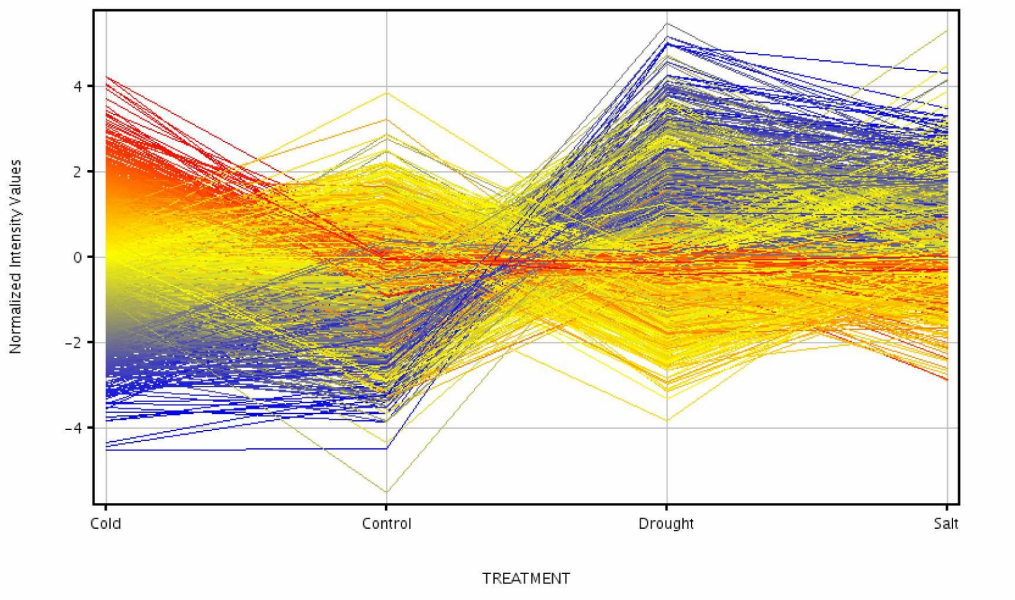

Figure 1. Profile plot for the differentially expressed genes found in our study

2.2 Expression Profiling of *MYB* and cluster analysis

To analyze the genome-wide expression profiles of rice *OsMYB* genes, microarray analysis was carried out using Affymetrix rice whole genome array. Expression data of MYB expression under abiotic stress were extracted from result of 12 hybridization experiment GSE6901 retrieved from GEO Database. .CEL files were downloaded and subjected to Genespring GX 10 (Agilent Technologies Inc, Santa Clara CA) and normalized with the PLIER16 algorithm (3) for further analysis. Obtained expression value were log2 transformed, probes having two fold up - down regulation were taken. Hierarchical clustering was performed by average linkage and Euclidean distance algorithm using GeneSpring GX 10.

3 Results

3.1 Identification of *MYB* genes and Phylogenetic analysis

By a reiterative database exploration 249 potential *OsMYB* domains in rice were retrieved. Non-redundant dataset for *MYB* genes in rice genome were used as input for further analysis. Computational analysis of 249 identified MYB has shown the presence of several other functional domains including WD domain, G-beta repeat, response regulator receiver domain, BTB/POZ domain, SWIRM/Zinc finger domain and many more. Phylogenetic analysis performed with the Maximum Likelihood method using all 249 proteins containing a single or double MYB domain, divided the genes into 3 main phylogenetic groups. Other subgroups and smaller clades were identified within each group, based upon bootstrap values.

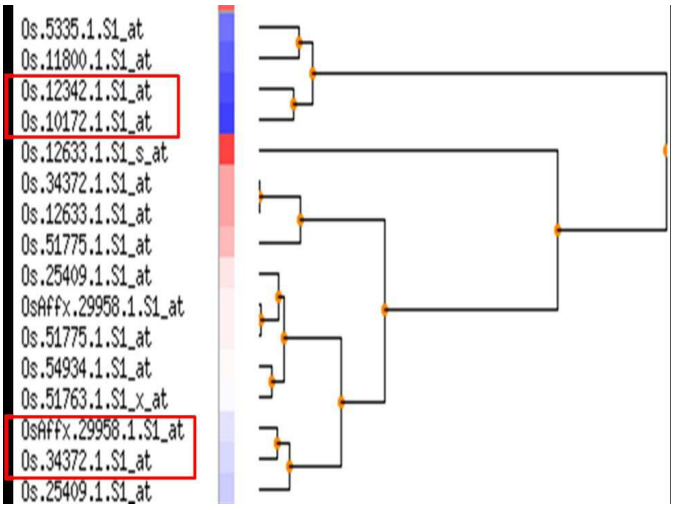

Figure 2. Clustering of upregulated OsMYB genes. Red boxes highlight the presence of co-expressed *MYB* gene clusters.

Table 1. Some up regulated *MYB* genes in three abiotic stress conditions.

Probes IDs (Cold stress)	**Locus ID**	**2 Fold Upregulation**
Os.57355.1.S1_x_at	LOC_Os01g58550	4.7924976
Os.5335.1.S1_at	LOC_Os04g49450	3.9795542
Os.11800.1.S1_at	LOC_Os01g50100	3.7111404
Os.12342.1.S1_at	LOC_Os12g04204	3.4641871
Os.10172.1.S1_at	LOC_Os02g41510	3.2356162
Probes IDs (Drought stress)	**Locus ID**	**2 Fold Upregulation**
Os.12633.1.S1_s_at	LOC_Os11g26790	7.9102745
Os.34372.1.S1_at	LOC_Os06g48300	6.630616
Os.51775.1.S1_at	LOC_Os12g05210	6.3568006
Os.25409.1.S1_at	LOC_Os06g45184	5.8992395
OsAffx.29958.1.S1_at	LOC_Os09g21180	5.7759666
Os.54934.1.S1_at	LOC_Os05g37060	5.6580715
Os.51763.1.S1_x_at	LOC_Os01g12690	5.57427
Probes IDs (Salt stress)	**Locus ID**	**2 Fold Upregulation**
Os.12633.1.S1_at	LOC_Os11g26790	6.636118
Os.51775.1.S1_at	LOC_Os12g05210	5.7613506
OsAffx.29958.1.S1_at	LOC_Os09g21180	5.308271
Os.34372.1.S1_at	LOC_Os06g48300	5.186286
Os.25409.1.S1_at	LOC_Os06g45184	5.0420284

3.2 Expression profiling and Clustering analysis result

Microarray data analysis was done by employing .CEL file to GeneSpring. Profile plot of all 2866 up and down regulated genes crossed all statistical test was made (Figure 1).

Differentially expressed genes were analyzed to extract *MYB* genes showing expression in drought, salt and cold stress conditions. We found 158 *MYB* genes out of 249 showing differential expression. Drought, salt and cold stresses upregulated (≥ 2 fold) 102, 72 and 16 *MYB* genes, respectively. Table 1 shows the up-regulated *MYB* genes found in our study. Clustering analysis of the upregulated *MYB* genes indentified in this study was performed to pinpoint genes with similar expression profiles between different stress conditions. Understanding of this functional network structure of MYB genes, such as gene regulatory and biochemical networks, systems biology is the area that has to be explored and the area that we believe to be the main stream in biological sciences in this century [4].

4 Conclusions

Our approach has identified co-regulated MYB gene networks that have potential role in abiotic stress response of rice. This will contribute to illustrate the functions of gene cooperation pathways not yet identified by classical genetic analyses. We defined the existence of Os*MYB* gene clusters comprising both phylogenetically related and unrelated genes that were significantly co-expressed, suggesting that specific sets of *MYB* genes might act in co-regulatory networks.

5 References

[1] Daniele Merico, David Gfeller & Gary D Bader (2009) **How to visually interpret biological data using networks,** *nature biotechnology* volume 27 number 10 october 2009.

[2] H. Kitano, (2002) **Systems biology: A brief overview**, *Science* 295, 1662–1664.

[3] Hubbell E, Liu WM and Mei R (2002) **Robust estimators for expression analysis**.*Bioinformatics* 18:1585-1592.

[4] Michael B. Eisen, Paul T. Spellman,Patrick O. Brown and David Botstein, **Cluster analysis and display of genome-wide expression patterns**". Department of Genetics and, Department of Biochemistry and Howard Hughes Medical Institute, Stanford University School of Medicine, 300 Pasteur Avenue, Stanford, CA 94305

[5] Michael TP, Mockler TC, Breton G, McEntee C, Byer A, et al. (2008) **Network discovery pipeline elucidates conserved time-of-day–specific cis-regulatory modules.** *PLoS Genet* 4(2): e14.

[6] Nafees A. Khan and Sarvajeet Singh, I.K. (2008) **Abiotic Stress and Plant Responses** 300 p, 65, 81-89866-95-2,

[7] Nicholas A. Heard, Christopher C. Holmes, David A. Stephens, David J. Handand George Dimopoulos, **Bayesian Co-clustering of Anopheles Gene Expression Time Series: A Study of Immune Defense Responses to Multiple Experimental Challenges.**

[8] Sharan R, Ulitsky I, Shamir R: **Network-based prediction of protein function.** *Mol Syst Biol* 2007, **3:**88.

[9] Wille A, Zimmermann P, Vranova E, Furholz A, Laule O, Bleuler S, Hennig L, Prelic A, von Rohr P, Thiele L, *et al.*: **Sparse graphical Gaussian modeling of the isoprenoid gene network in Arabidopsis thaliana.** *Genome Biol* 2004, 5(11):R92.