J. BIOL. ENVIRON. SCI.,
2015, 9(25), 1-9

Original Research Article

In Silico Analysis of Cis-Regulatory Elements on Co-Expressed Genes

Parviz Heidari*, Mostafa Ahmadizadeh and Hamid Najafi-Zarrini
Department of Plant Breeding and Biotechnology, Sari Agricultural Sciences and Natural Resources University, Sari, IRAN

Received: 12.02.2015; Accepted: 07.04.2015; Published Online: 13.05.2015

ABSTRACT

Cis-regulatory elements (CREs) are regions of non-coding DNA that regulate the transcription of nearby genes. CREs typically regulate gene transcription by functioning as binding sites for transcription factors. Publicly available database of co-expressed gene sets would be valuable tools for a wide variety of experimental designs, including targeting of genes for functional identification or for regulatory investigation. The study of CREs effect on expression can improve our understanding of co-expression genes and gene networks. In present study we compared the correlation between expression and CREs in co-expression genes with LTP5 by using the Genevestigator database that provides co-expressed genes deduced from microarray data, and SCOPE that uses to find CRDs in 800bp of upstream of the DNA sequences of co-expression genes. The result revealed that three motifs (TGSCAB, ATWTGYMG and CBTATC) of PRISM algorithm, GCCAC motif of BEAM and ATTGNVANNYGG motif of SPACER algorithm distribute in promoter of co-expressed genes and LUX transcription factor was identified by UniPROB database. We present here a new comparison method for detecting key cis-regulatory elements that effect on co-expressed genes to find a relation to clarify the function and regulation of particular genes and gene networks under stress conditions.

Key Words: Co-expression, Stress, Cis elements, Gene networks

Koekspres Genler Üzerindeki Cis-Düzenleyici Öğelerin *In Silico* Analizi

ÖZET

Cis-düzenleyici öğeler (CRE), komşu genlerin transkripsiyonunu düzenleyen kodlanmayan DNA'nın bölgeleridir. CRE'ler tipik olarak transkripsiyon faktörlerinde bağlantı noktaları işlevi görerek gen transkripsiyonunu düzenler. Erişilebilir bir koekspres gen veritabanının bulunması bir çok deneysel tasarım için değerli bir araç olacaktır. CRE'lerin ekspressiyon üzerindeki etkileri ile ilgili yapılan çalışmalar, gen koekspresyonu ve gen ağını anlamamızı kolaylaştırır. Bu çalışmada, ekspresyonlar arasındaki korelasyonlar ile koekspres LTP5'li genlerde bulunan CRE'ler, Genevestigator veritabanı kullanılarak karşılaştırılmıştır. Sonuçlar incelendiğinde PRISM algoritmasının üç motifi (TGSCAB, ATWTGYMG ve CBTATC), BEAM algoritmasının GCCAC motifi ve SPACER algoritmasının ATTGNVANNYGG motifinin koekspres genlerin promotörü içerisinde dağıldığı ve UniPROB veritabanı ile LUX transkripsiyon faktörünün tespit edildiği görülmüştür. Bu çalışma ile koekspres genleri etkileyen önemli cis-düzenleyici öğeler ile bazı genlerin düzenlenmesi ve strest koşulları altındaki gen ağın tespiti üzerine yeni bir yöntem ortaya konmuştur.

Anahtar Kelimeler: Koekspresyon, Stres, Cis öğeleri, Gen ağları

INTRODUCTION

Plants are constantly exposed to a variety of biotic and abiotic stresses and have evolved intricate mechanisms to sense and respond to the adverse conditions. Besides, many stress-associated cis-acting regulatory elements that activate transcription in response to salinity, drought, wounding and pathogen infection have been identified in plants (Ibraheem *et al.* 2010).The short nucleotide motif for this modular code is called the cis-regulatory elements (CREs), which is recognized by factors for transcriptional activation or suppression (Won *et al.* 2009, Wittkop and Kalay 2012). The DNA sequence elements that act as binding sites for transcription factors coordinate the expression of genes in whose regulatory region they appear, and are key to reducing the complexity of the observed expression patterns (Bussemaker *et al.* 2001). Therefore, the study of CREs effect on expression can improve our understanding of co-expression genes and gene networks (Ibraheem *et al.* 2010). The identification of motifs which share in genes promoter, does not necessarily mean that the genes are involved in the same process. Further, the data of microarray expression have notoriously noisy, which affects the ability of motif finders to identify expression patterns under stress (Martyanov and Gross 2010). One of problem to elucidate of functional relationships motifs in co-expressed genes is that some genes have similar expression profiles but they are regulated by different transcription factors (Ahmadizadeh and Heidari 2014,

* Corresponding author: parvizh63@gmail.com

Martyanov and Gross 2010). Allocco *et al*. (2004) believe that similar gene expression patterns are affected by similar regulatory mechanisms.

Computational methods for the identification of cis-regulatory sequences associated with genes have long been sought owing to the strenuous laboratory procedures required to identify them (Wasserman and Sandelin, 2004, Rani 2007). Publicly available database of co-expressed gene sets would be a valuable tool for a wide variety of experimental designs, including targeting of genes for functional identification or for regulatory investigation (Won *et al.* 2009). Web-based databases such as: Genevestigator and Atted 2 professional have provided a convenient and easy way to search for co-expressed genes and cis regulatory element in DNA sequences of *A.thaliana*. PRISM, BEAM and SPACER are three distinct search strategies for motif finding in co-expressed genes that SCOPE (Suite for Computational Identification of Promoter Elements), combines these three methods (Martyanov and Gross 2010).

To date, no work has been published on the transcription factors/cis-acting regulatory elements that regulate the expression of co-expression genes during stresses. In this study, we used the data set of co-expressed genes from miccroarry experimental according high correlation with LTP5 gene under different stresses. Also, we applied the motif finder tools (SCOPE) to identify statistically significant motifs in the upstream regions of the co-expressed genes with *LTP5*, in arabidopsis thaliana.

MATERIALS AND METHODS

Identification of co-expression genes with the LTP5 *gene*

We used Genevestigator database that is a publicly available microarray database coupled with expression-data analysis tools. Co-expressed genes with the *LTP5* gene were obtained from Genevestigator internet database site (https://www.genevestigator.com/gv). All available expression samples of arabidopsis were selected then was entered LTP5 gene on gene selection part. Abiotic and biotic stress conditions with p-value 0.01 and fold change 1.5 were selected by the perturbations tool and created a new sample. The co-expression tool was used to find other genes regulated under similar conditions and the 25 genes with the strongest correlation with the LTP5 gene were selected as shown in table 1. The hierarchical clustering tool was used for grouping and clustering by expression and tissue type.

Table 1. Co-expression genes with the strongest correlation with LTP5 gene.

Gene (Accession No.)	Num. of bp	CDS region	5UTR BeforeATG	Gene Bank	Position on chromosome	Score	Description
AT3G51600	752	129-485	1-128	NM_115019	19138394 - 19139252		lipid transfer protein
AT2G05520	710	56-493	1-55	NM_126575	2026162-2027099	0.93	Glycine-rich protein 3
AT2G34430	1008	63-863	1-62	NM_128995	14524756-14525763	0.92	Light harvesting chlorophyll protein
AT4G17245	641	19-519	1-18	NM_117830	9669365-9670005	0.92	RING/U-box super family protein
AT3G28040	3256	83-3133	1-82	NM_113722	10435057-10438391	0.91	Leucine-rich receptor-like protein kinase family
AT1G79040	720	70-492	1-70	NM_106555	29736016- 29737009	0.91	Photosystem 2 subunit R
AT3G08940	1124	43-906	1-42	NM_111728	2717676 - 2718869	0.90	Light harvesting complex Photosystem 2
AT4G02770	835	73-699	1-72	NM_116511	1229111 - 1229945	0.89	Photosystem 1 subunit D-1
AT1G61520	1198	143-964	1-142	NM_104833	22700010 - 22701383	0.89	Photosystem 1 Light harvesting complex gene
AT1G76100	782	37-552	1-36	NM_106259	28553987 - 28554768	0.89	Plastocyanin 1

AT3G45780	3304	64-3054	1-63	NM_114447	16816866 - 16824210	0.89	Phototropin 1
AT2G45180	678	58-462	1-57	NM_130081	18626320 - 18626997	0.89	Bifunctional inhibitor/lipid-transfer protein
AT1G30380	755	129-521	1-128	NM_102775	10722197 - 10723247	0.88	Photosystem 1 subunit K
AT3G47470	1266	261-1016	1-260	NM_114615	17493372 - 17495033	0.88	Light harvesting chlorophyll protein
AT1G32540	1028	257-721	1-256	NM_102989	11767504 - 11769880	0.88	protein with 3 plant-specific zinc finger domains
AT1G29910	1020	54-857	1-53	NM_102731	10472280 - 10473299	0.88	Chlorophyll A/B binding protein 3
AT1G20340	935	91-594	1-90	NM_101885	7042429 - 7043363	0.88	Cupredoxin superfamily protein
AT2G05070	1063	64-861	1-63	NM_126537	1799234 - 1800392	0.88	Photosystem 2 Light harvesting complex
AT4G05180	1037	74-766	1-73	NM_116757	2671822 - 2673243	0.88	Photosystem 2 subunit Q-2
AT2G15960	380	42-278	1-41	NM_127155	6947309 - 6947688	0.88	Unknown protein
AT3G19850	1965	99-1763	1-98	NM_112875	6898181 - 6901255	0.87	Phototropic-responsive NPH3 family protein
AT1G55670	780	94-576	1-93	NM_104443	20802670 - 20803449	0.87	Photosystem 1 subunit G
AT3G56940	1493	90-1319	1-89	NM_115553	21076505 - 21078443	0.87	Dicarboxylatediiron protein
AT3G54890	1109	182-907	1-181	NM_115346	20339504 - 20341103	0.87	Photosystem 1 Light harvesting complex gene
AT2G03980	1420	173-1276	1-172	NM_126435	1259199 - 1262552	0.87	GDSL-like lipase/Acylhydrolase superfamily protein
AT1G29670	1326	81-1172	1-80	NM_102707	10375763 - 10377871	0.87	GDSL-like lipase/Acylhydrolase superfamily protein

Cis-acting regulatory elements analysis

Co-expressed genes are analyzed by SCOPE (Chakravarty *et al.* 2007, Carlson *et al.* 2007) to generate significant candidate regulatory motifs, which are in 800bp upstream of genes. Statistically significant motifs from SCOPE analysis of gene subsets were used as the input to search for similar motifs and transcription factor binding sites in the UniPROBE database (Newburger and Bulyk 2009). Using database associated search tools, upstream of the 5′ regulatory region of each arabidopsis gene was scanned for the presence of putative cis-acting regulatory elements identical with or similar to the motifs registered in Plant CARE (http://bioinformatics.psb.ugentbe/webtools/plantcare/cgi-bin/CallMatIE55.htpl).

RESULTS AND DISCUSSION

Co-expression analysis

The results of co-expression showed the highest correlation between the *LTP5* gene and AT2G05520 gene (Table 1).AT2G05520 gene is a glycine-rich protein (GRP). GRPs have been found in the cell walls of many higher plants and form a third group of structural protein components of the wall in addition to extensions and proline-rich proteins (Ringlia *et al.* 2001). Ringlia *et al.* (2001) also proposed that GRPs are part of a repair system of the plant during the stretching phase of protoxylem. Pearson correlation between co-expression genes indicated that genes involved in the photosystem have a high correlation with *LTP 5*. RING/U-box superfamily protein (AT4G17245) with 0.92 correlations (Table 1). RING (Really Interesting New Gene)/U-box superfamily protein in plants, like in other eukaryotes, targets numerous intracellular regulators and thus modulates almost every aspect of growth and development (Wang and Deng 2011). Leucine-rich receptor-like proteins (LRR-RLKs) areimportant proteins found in plants under abiotic and biotic stress conditions and the results showed that protein (AT3G28040) had a high correlation (0.91) with *LTP5* (Table 1). LRR-RLKs play important and versatile roles in plant growth, development, hormone perception as well as in the responses of plants to both biotic and abiotic stresses, especially the plant-microbe interactions (Yang *et al.* 2014).

Hierarchical clustering based expression value and tissue type on co-expression genes showed that *LTP5*, At3g28040, At3g45780, At2g03980 and At2g15960 clustered in one group (Figure 1). The result of hierarchical clustering also indicated that location was important on grouping. *LTP5* and most of its co-expression geneshad the highest expression in the shoot, juvenile leaf, adult leaf and rosette stage, whereas these genes were not expressed in seed, root protoplast, cell culture/primary cell and root tissue (Figure 1). Hierarchical clustering one of the most popular methods of cluster analysis, which has two particular characteristics: evaluate relationships between gene expression patterns and color display of gene expression, also hierarchical clustering is based methods component of the neural network.

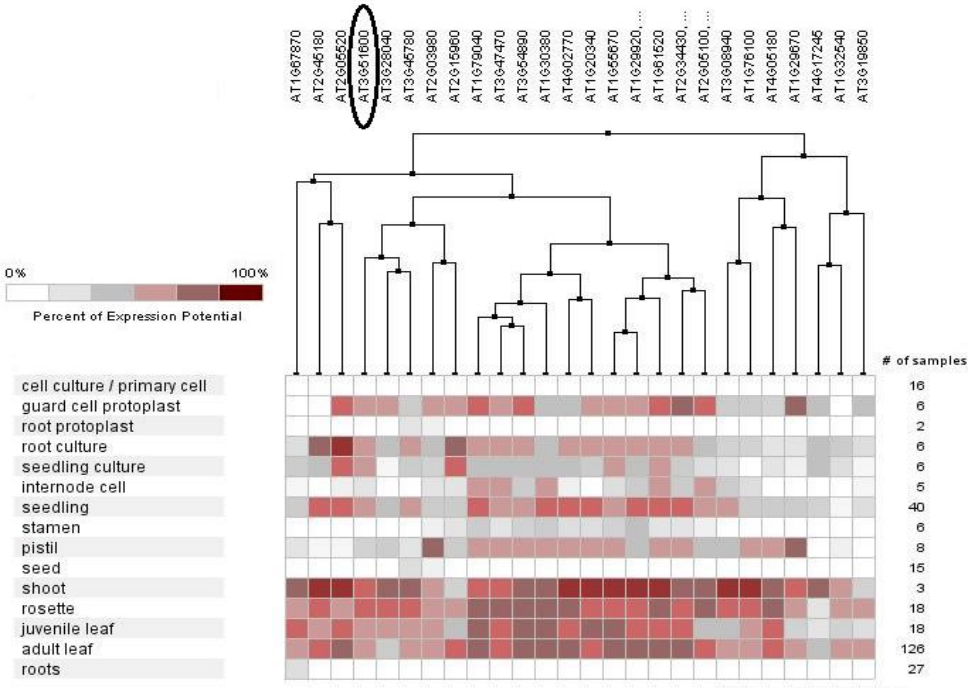


Figure 1. Grouping of gene expression in Arabidopsis tissues based on the expression levels by the hierarchical clustering method.

Cis-acting regulatory elements analysis

Results of the main motif distribution from the SCOPE runs are shown in figure 2 and table 2. The results of SCOPE revealed that motif (light blue in figure 2) with consensus sequence CBTATC covered 81% of upstream

in co-expressed genes. The TGSCAB motif (red in figure 2) and ATWTGYMG motif (orange in figure 2) were positioned in the 20-300 bp and 400-700 bp regions upstream of gene start, respectively (Figure 2). Three motifs (TGSCAB, CBTATC and ATWTGYMG) were found by PRISM algorithm in the promoter region of co-expressed genes (Table 2). PRISM algorithm finds specific kind of motif which is non-degenerate (Carlson *et al.* 2006). The output of UniPROBE showed that consensus sequenceCBTATC and ATWTGYMG motifs were a LUX transcription factor in *A. thaliana* (Table 2). The LUX is necessary for activation of CCA1 (*CIRCADIAN CLOCK ASSOCIATED*) and LHY (*LATE ELONGATED HYPOCOTYL*) expression (Hazen *et al.* 2005). The result of co-expression showed that most genes which had high correlation with LTP5 were involved in photosystem (Table 1), also this result was confirmed by SCOPE and UniPROBE when the LUX was found in upstream regions genes start in co-expressed genes.

Table 2. List of cis-regulatory motifs that were found in upstream of co-expressed genes by SCOPE and comparison with UniPROBE data.

Symbol	SCOPE motifs	Count	match UniPROBE	E-value Motif	coverage	Algorithm
	TGSCAB	35	-	-	73%	PRISM
	GCCAC	17	-	-	54%	BEAM
	ATTGNVANNYGG	12	-	-	39%	SPACER
	ATWTGYMG	17	LUX	0.03	50%	PRISM
	CBTATC	42	LUX	0.0017	81%	PRISM

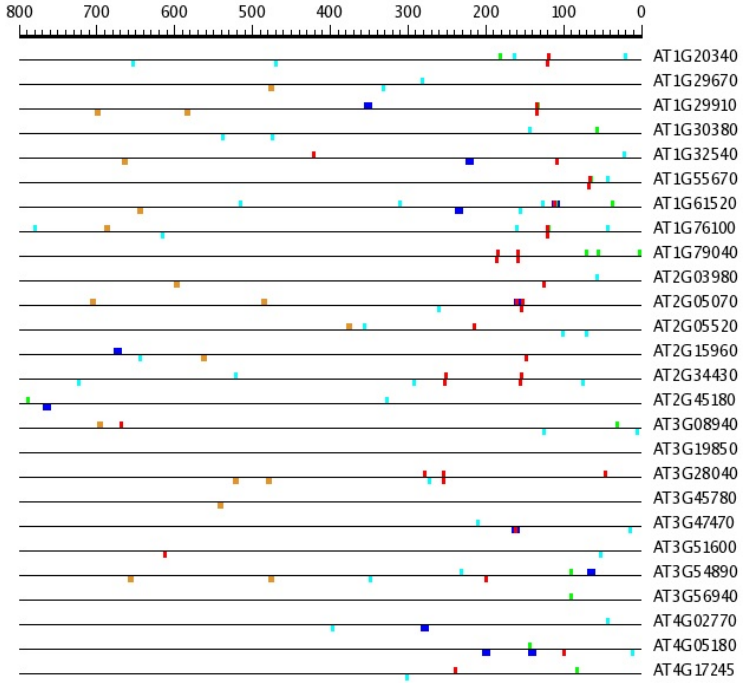


Figure 2. Output of SCOPE for co-expressed genes.

Based on the cis-regulatory elements, in the upstream region of co-expression genes of the LTP5 gene in arabidopsis, the sequences are categorized into different elements. The functions of these predicted cis elements are listed in table 3 and their arrangement on upstream and categorization of cis elements upstream regions is shown in figure 3 and table 3. ATCT, GAG, GTGGC, TCT,2-box, LAMP-element, G-box, Box I, AE-box, G-Box, Box-4 and I-box motif were found at varying frequencies in the regulatory regions of that are associated to the light response (Figure 3 and Table 3). Light is a predominant factor which controls the circadian rhythm of various life processes such as growth, development, nitrate uptake, and stress responses in

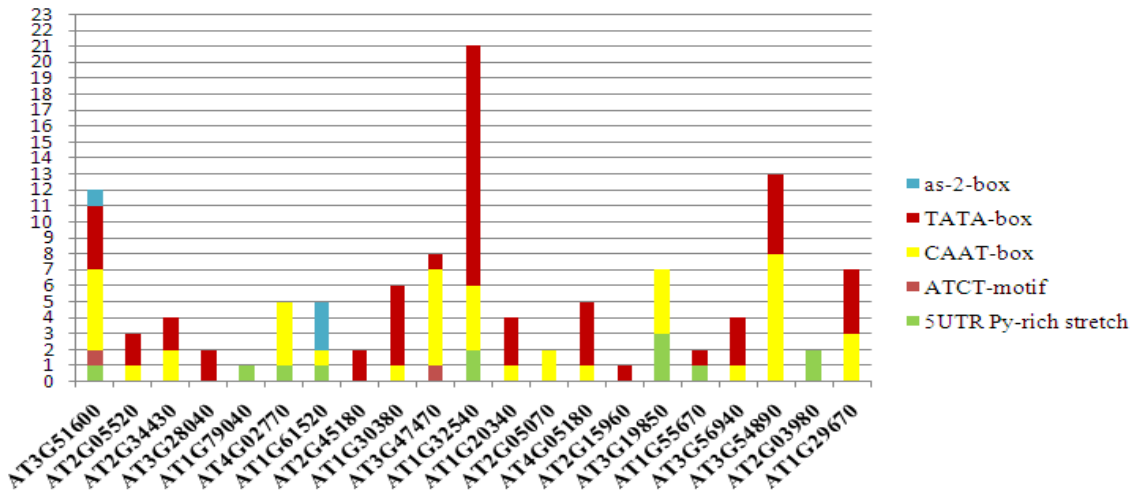
plants. Also, many biotic stress responses in plants are specifically adjusted by light conditions (Genoud *et al.* 2002, Kumar *et al.* 2009, Millar and Kay 1991). These results were confirmed by SCOPE analysis when LUX transcription found in most upstream of co-expressed genes (Figure 2 and Table 2).

Table 3. Potential cis-acting regulatory elements identified in the 5'regulatory sequences co-expression genes of Arabidopsis thaliana. 5' regulatory region as analyzed using Plant CARE databases.

Cis-acting elements	Frequency (%)	Sequence	Function
5UTR Py-rich stretch	30.77	TTTCTTCTCT	cis-acting element conferring high transcription levels
ATCT-motif	7.69	AATCTAATCC	part of a conserved DNA module involved in light responsiveness
Box III	7.69	CATTTACACT	protein binding site
CAAT-box	57.69	CAAAT;CAAT;CCAAT	common cis-acting element in promoter and enhancer regions
TATA-box	57.69	taTATAAAtc;TATAAA;ATATAA;TTTTA;TATA	core promoter element around -30 of transcription start
as-2-box	7.69	GATAatGATG	involved in shoot-specific expression and light responsiveness
circadian	15.38	CAANNNNATC	cis-acting regulatory element involved in circadian control
AE-box	7.69	AGAAACAA	part of a module for light response
Box I	11.54	TTTCAAA	light responsive element
GAG-motif	23.07	AGAGAGT;AGAGATG	part of a light responsive element
AAGAA-motif	7.69	GAAAGAA	Unknown
ABRE	3.85	ACACGTGGC;ACGTGGC;CACGTG;CCTACGTGGC	cis-acting element involved in the abscisic acid responsiveness
G-Box	7.69	CACGTG	cis-acting regulatory element involved in light responsiveness
GARE-motif	3.85	TAACAAG	gibberellin-responsive element
I-box	7.69	GATAAGATT	part of a light responsive element
TCA-element	7.69	GAGAAGAATA	cis-acting element involved in salicylic acid responsiveness
ARE	19.23	TGGTTT	cis-acting regulatory element essential for the anaerobic induction
P-box	7.69	CCTTTTG	gibberellin-responsive element
WUN-motif	3.85	TCATTACGAA	wound-responsive element
TGA-element	7.69	AACGAC	auxin-responsive element
HSE	7.69	AAAAAATTTC	cis-acting element involved in heat stress responsiveness
Box 4	7.69	ATTAAT	part of a conserved DNA module involved in light responsiveness
O2-site	3.85	GATGACATGG	cis-acting regulatory element involved in zein metabolism regulation
TCT-motif	3.85	TCTTAC	part of a light responsive element
LAMP-element	3.85	CCAAAACCA	part of a light responsive element
CCAAT-box	3.85	CAACGG	MYBHv1 binding site
CGTCA-motif	3.85	CGTCA	cis-acting regulatory element involved in the MeJA-responsiveness
GA-motif	7.69	AAAGATGA	part of a light responsive element
TGACG-motif	3.85	TGACG	cis-acting regulatory element involved in the MeJA-responsiveness
TATCCAT/C-	3.85	TATCCAT	Unknown

motif			
MBS	3.85	CAACTG	MYB binding site involved in drought-inducibility
GTGGC-motif	3.85	CATCGTGTGGC	part of a light responsive element
MSA-like	3.85	TCCAACGGT	cis-acting element involved in cell cycle regulation
Skn-1_motif	3.85	GTCAT	cis-acting regulatory element required for endosperm expression
AC-II	3.85	C/T)T(T/C)(C/T)(A/C)(A/C)C(A/C)A(A/C)C(C/A)(C/A)C	Unknown
G-box	3.85	TGACGTGG;CACGTC	cis-acting regulatory element involved in light responsiveness

P-box and GARE-motif were found in the regulatory regions of AT3G19850, AT2G15960 and AT3G45780 genes that are associated to the gibberellin responsive element (Figure 3 and Table 3). A single GARE motif in some promoters is able to direct hormonally regulated transcription at a high level because it cooperates with other cis-acting elements (Rogers *et al.* 1994). The HSE cis-acting regulatory element, involved in heat stress responsiveness, was found in the regulatory regions of AT1G55670 and AT1G29910 genes. AT3G47470 has a TGACG- cis-acting regulatory element, which is involved in MeJA-responsiveness, and an O2-site cis-acting regulatory element, which is involved in zein metabolism regulation (Figure 3 and Table 3). Plant heat shock genes are not only expressed in response to heat stress, but also during zygotic embryogenesis and in other developmental stages in the absence of exogenous stress. Studies of class I sHSP promoters showed that heat shock elements (HSEs), the cis-acting elements necessary for the heat shock response, were also involved in their regulation during zygotic embryogenesis (Carranco *et al.* 1999). TGA-element are related to auxin-responsive element was found in AT3G54890 and AT3G08940 and the ABRE motif was found to be distributed within the regulatory region of the AT3G54890 gene. The ABRE: cis-acting element (TACGGTC) is involved in ABA responsiveness and drought tolerance, MBS: MYB binding site related to ABA signaling and involved in drought inducibility(Figure 3 and Table 3).



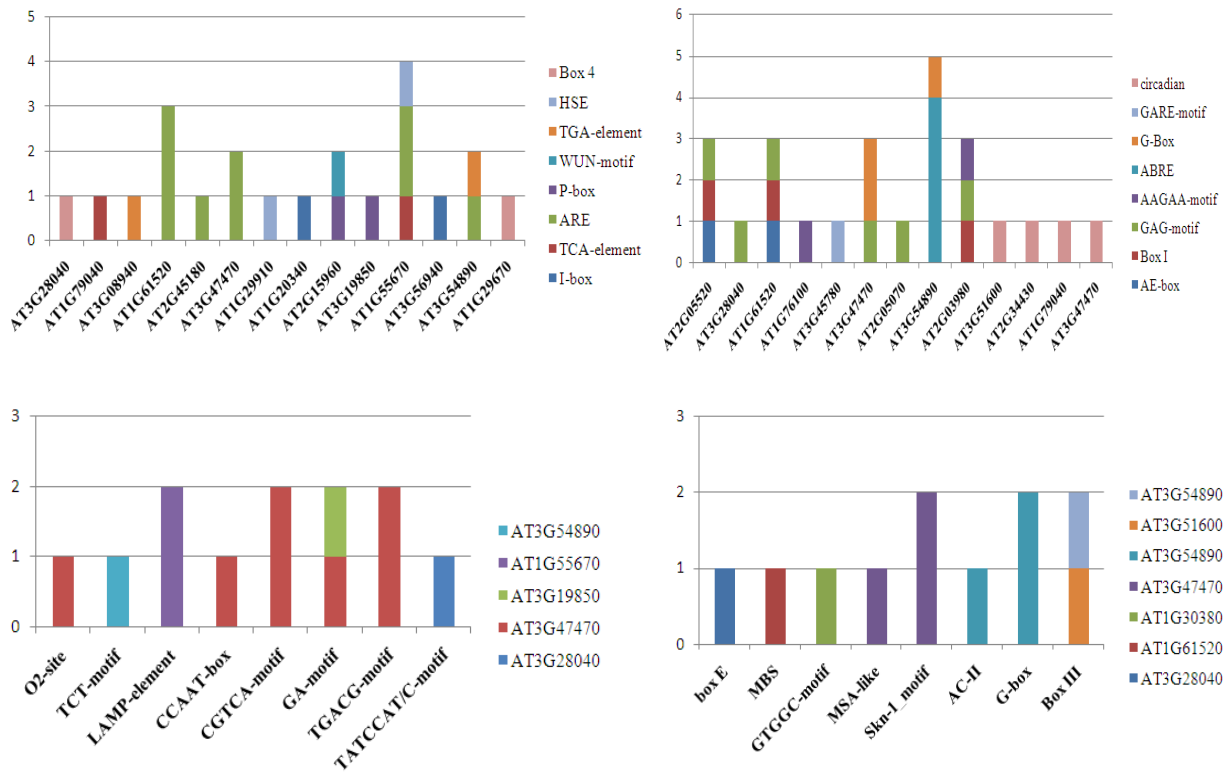


Figure 3. Cis-acting regulatory elements identified in the 5′ regulatory sequences of co-expression genes of arabidopsis.

CONCLUSION

Numerous studies have been reported on the expression of genes during plant development, but little is still known about the regulation of plant genes during stress conditions and interaction between expression pathway and cis regulation. This paper combined gene expression profiles and promoter sequences to introduce a simple way for finding cis-regulatory motifs which is unknown and has the potential to discover new transcription factor. These cis-regulatory motifs could use for prediction of gene expression and subclustering of co-expressed genes. The power of bioinformatics and the availability of whole-genome sequences have enabled a comprehensive description of transcription factor families in several plant species. In addition to the co rebinding sites that seem to be associated with transcription factor families, there may exist a degree of subtlety differences in the cis regulatory elements necessary for transcription factors to facilitate their unique regulatory effects.

REFERENCES

Ahmadizadeh M, Heidari P (2014). Bioinformatics study of transcription factors involved in cold stress. Biharean Biologist 8 (2): 83-86

Allocco DJ, Kohane IS, Butte AJ (2004). Quantifying the relationship between coexpression, co-regulation and gene function. BMC Bioinformatics 5:18.

Arabidopsis Genome Initiative (2000): Arabidopsis Information Resources (TAIR), http://www.arabidopsis.org/in.

Bussemaker HJ, Li H, Siggia ED (2001). Regulatory element detection using correlation with expression. Nature Genetics 27: 167-171.

Carlson JM, Chakravarty A, Khetani RS, Gross RH (2006). Bounded search for *de novo* identification of degenerate cis-regulatory elements. BMC Bioinformatics **7**(1): 254.

Carlson JM, Chakravarty A, DeZiel CE, Gross RH (2007). SCOPE: a web server for practical *de novo* motif discovery J.M. Carlson; A. Chakravarty; C.E. DeZiel; R.H. Gross; Nucleic Acids Research 35:259-264.

Carranco R, Almoguera C, Jordano J (1999). An imperfect heat shock element and different upstream sequences are required for the seed-specific expression of a small heat shock protein gene. Plant Physiology 121: 723–730.

Chakravarty A, Carlson JM, Khetani RS, Gross RH (2007). A novel ensemble learning method for de novo computational identification of DNA binding sites." *BMC Bioinformatics* **8**: 249.

Genoud T, Buchala AJ, Chua NH, Traux MJP (2002). Phytochromesignalling modulates the SA-perceptive pathway in Arabidopsis. Plant Journal 31: 87-95.

Hazen SP, Schultz TF, Pruneda-Paz JL, Borevitz JO, Ecker JR, Kay SA (2005). *LUX ARRHYTHMO* encodes a Myb domain protein essential for circadian rhythms. PNAS 102:10387-10392.

Ibraheem O, Bothab CEJ, Bradley G (2010). In silico analysis of cis-acting regulatory elements in 5′ regulatory regions of sucrose transporter gene families in rice (*Oryza sativa* Japonica) and *Arabidopsis thaliana*. Computational Biology and Chemistry 34: 268-283.

Kumar GM, Mamidala P, Podile AR (2009). Regulation of Polygalacturonase-inhibitory proteins in plants is highly dependent on stress and light responsive elements. Plant Omics Journal 2(6): 238-249.

Martyanov V, Gross RH (2010). Identifying functional relationships within sets of co-expressed genes by combining upstream regulatory motif analysis and gene expression information. BMC Genomics 11: 1-9.

Millar AJ, Kay SA (1991). Circadian control of cab gene transcription and mRNA accumulation in Arabidopsis. The Plant Cell 3: 541-550.

Newburger DE, Bulyk ML (2009). UniPROBE: an online database of protein binding microarray data on protein-DNA interactions. Nucleic Acids Res, 37(Database issue): D77-82.

PlantCARE (2002). A database of plant cis-acting regulatory elements and a portal to tools for in silico analysis of promoter sequences. Lescot, M., P. Déhais, Y. Moreau, B. De Moor, P. Rouzé and S. Rombauts. Nucleic Acids Res. Database issue 30(1): 325-327.

Rani V (2007). Computational methods to dissect cis-regulatory transcriptional networks. J. Biosci 32: 1325–1330.

Ringlia C, Kellera B, Ryser U (2001). Glycine-rich proteins as structural components of plant cell walls. CMLS, Cell. Mol. Life Sci. 58:1430–1441.

Rogers JC, Lanahan MB, Rogers SW (1994). The cis-acting gibberellin response complex in high-pl a-amylase gene promoters. Plant Physiology 105: 151-158.

Wang F, Deng XW (2011). Plant ubiquitin-proteasome pathway and its role in gibberellin signaling. Cell Research 21(9): 1286–1294.

Wasserman WW, Sandelin A (2004). Applied bioinformatics for the identification of regulatory elements. Nature Reviews Genetics 5: 276-287.

Wittkop PJ, Kalay G (2012). Cis-regulatory elements: molecular mechanisms and evolutionary processes underlying divergence. Nature Reviews Genetics 13: 59-69.

Won SK, Lee YJ, Heo YK, Cho M, Chocis HT (2009). Element and transcriptome based screening of root hair-specific genes and their functional characterization in Arabidopsis. Plant Physiology 150: 1459–1473.

Yanga L, Wua K, Gaob P, Liua X, Lic G, Wu Z (2014). GsLRPK, a novel cold-activated leucine-rich repeat receptor-like protein kinase from *Glycine soja*, is a positive regulator to cold stress tolerance. Plant Science 215–216: 19–28.