Advances in Life Science and Technology
ISSN 2224-7181 (Paper) ISSN 2225-062X (Online)
Vol 3, 2012

Evaluation of the coexpressed gene network of DREB2A

Santosh Kumar Mehar*

Department of Botany, College of Sciences, Sri Venkateswara University, Tirupati, Andhra Pradesh, India

*Corresponding author email: santoshkumar.1@rediffmail.com

Abstract

Different proteins in the cell perform their function in a unified way, interacting with the other proteins involved in completion of the particular metabolic pathway. The information about a single protein estimated from an organism subjected to a particular treatment will reveal only part of the process. To understand the process completely it is very essential to have an assessment of the different proteins linked together in different pathways and also the level of expression of different genes under different conditions. The coexpressed genes reveal the part of the genome which is intricately involved in carrying the process being studied.

The network of coexpressed genes with DREB2A was analyzed. DREB2A is involved in different processes related to drought and salt stress. Out of the genes coexpressed with it, At1g16030 is Hsp70b/ Heat Shock Protein 70B, and is reported to be involved in virus attack and in protective processes under heat stress. Similarly At2g26150 was also observed to be closely networked with DREB2A, and is a member of heat stress transcription factor. The information from the coexpressed network of DREB2A shows that the interacting proteins play important role in protection of plants from pathogen attack and also in protecting the protein repertoire of the plant from denaturation when it is subjected to salt and heat stress.

Keywords: DREB2A, Heat shock protein, transcription factor, coexpressed genes

1. Introduction

DNA microarray analysis produces information on relative expression levels for thousands of genes simultaneously. Microarray data contain information about concerted changes in transcript levels in these datasets beyond the original purpose of each dataset. Such microarray data have been collected in several general repositories like ArrayExpress (Parkinson et al., 2005), Gene Expression Omnibus (GEO) (Barrett et al., 2005) and the Center for Information Biology Gene Expression Database (CIBEX) (Ikeo et al., 2003). The Arabidopsis Information Resource (TAIR) (Rhee et al., 2003) and the Nottingham Arabidopsis Stock Centre Arrays (NASC Arrays) (Craigon et al., 2004) are species-specific repositories for microarray data. From these information resources it is possible to collect information about individual genes or samples. However, knowledge about the expression of genes at the individual scale tells only part of the story, finding interaction partners for the particular gene can reveal its function in greater detail. The importance of such information can be best judged by the report (Venkatesan et al., 2009), suggesting that at any time a human cell may contain about 130,000 binary interactions between proteins. So far, a mere 33,943 unique human protein–protein interactions have been listed (http://thebiogrid.org).

Realizing the importance of such information, several databases have been created as secondary database for microarray data. The comprehensive systems-biology database (CSB.DB) (Steinhauser et al., 2004), Botany Array Resource (BAR) (Toufighi et al., 2005), Arabidopsis Co-expression Tool (ACT) (Manfield et al., 2006) and Genevestigator (Zimmermann et al., 2005), provide co-expressed gene relationships calculated from the array data stored in TAIR and/or the NASC Arrays. Such gene relationship information is valuable for predicting gene function, because co-expressed genes are often involved in the same or related pathways. ATTED-II is a valuable addition to this list. It stands for A.thaliana trans-factor and cis-element prediction database (ATTED-II) for retrieving gene-to-gene relationships like other databases for co-expressed genes. Besides the valuable attributes of the earlier listed databases, ATTED-II has some advancements like visualization of network of co-expressed gene relationships, facilitating understanding of the structural basis of gene co-expression, pre calculated results for cis- element prediction linked to every gene and functional category and graphical representation of the gene expression patterns.

Present study utilizes the resources available at ATTED-II, to understand the molecular interaction and coexpressed genes with DREB2A (Dehydration-Responsive Element Binding Protein2A) coded by At5g05410. Transcription factor DREB2A interacts with a cis-acting dehydration-responsive element (DRE) sequence to activate the expression of downstream genes that are involved in drought- and salt-stress response in Arabidopsis thaliana. In its intact form DREB2A expression does not activate downstream genes under normal growth conditions. However, a negative regulatory domain that exists in the central region of DREB2A, when deleted transforms DREB2A to a constitutive active form (DREB2A CA). Overexpression of DREB2A CA induces not only drought- and salt-responsive genes but also heat-shock (HS)-related genes. Besides the transient induction of the DREB2A occurs rapidly by HS stress, and the sGFP-DREB2A protein accumulates in nuclei of HS-stressed cells (Sakuma et al., 2006).

2. Methodology

ATTED-II database (Obayashi et al., 2009) was used to retrieve the significant inforamation about the gene interactions and coregulated gene network of DREB2A. The detailed information and graphical details were obtained from ATTED-II and the other online databases such as TAIR (The Arabidopsis Information Resource, Lamesch et al., 2011) and KEGG (Kyoto Encyclopedia of Genes and Genomes, Kanehisa et al., 2011).

3. Results and discussion

DREB2A has extensive gene interaction network (Fig.1). The list of top 50 coexpressed gees is given in table1. Some of them like At5g10695, At3g27880, At5g03210 (Table 1), do not have estabished proven annotations. However, the other genes coexpressed with it have diverse and important roles during the development of the plant and/or response to certain kind of stresses. The important functions of some of them are given below;

At1g01720 belongs to a large family of putative transcriptional activators with NAC domain. Their transcript levels increases in response to wounding and abscisic acid. ATAF1 attentuates ABA signaling and sythesis (Collinge and Boller, 2001), and is one of the directly connected gene on the network.

At1g71000 is involved in protein folding and its identified locations are cytoplasm and nucleus. It is expressed in different plant organs like carpel, flower, pedicel, plant sperm cell, pollen, root and stamen (Schmid et al., 2005, Inzé et al., 2011).

At4g15420 is Ubiquitin fusion degradation UFD1 family protein. It has peptidase activity and also involved in proteolysis, ubiquitin-dependent protein catabolic process and zinc ion binding (Schmid et al., 2005, and InterPro to GO annotation (AnalysisReference:501739686).

At4g12410 is involved in response to auxin stimulus and is expressed in carpel, collective leaf structure, flower, pedicel, petal, plant embryo, pollen, sepal, stamen (Schmid et al., 2005).

At1g16030 is Hsp70b/ HEAT SHOCK PROTEIN 70B. It is involved in processes related to response to virus attack and heat stress (Sung et al., 2001, Aparicio et al., 2005).

At2g26150 is also directly connected on the network with DREB2A. It is member of Heat Stress Transcription Factor (Hsf) family. It is involved in response to misfolded protein accumulation in the cytosol. It is regulated by alternative splicing and non-sense-mediated decay. It is involved in regulatoin of transcription and its annotated location is endomembrane system and nucleus (Riechmann et al., 2000, Sugio et al., 2009, Banti et al., 2010).

Predicted cis elements for At5g05410 and the sequence information are given in fig. 2 and table 2 respectively. The correlation of expression between At5g05410 and the four directly connected genes on the coexpression network during different development stages is diplayed in fig. 3 and detailed in table 3, As it is clear from the figure that the epression profiles of these grouped genes fluctuates with the stage of development and under specific conditions to which the plant is subjected. They can have all variations in correlations ranging from strong negative to strong positive.

The study of the interaction between the different genes in an organism is an evolving science and its utility is undisputed. The knowledge about the groups of genes which are always expressed together give an indication of the response of the organism to the specific conditions. As it has been seen in case of DREB2A, it is strongly coexpressed with genes involved in responses to various kinds of stresses and has protective roles to play such as the expression of HSPs which assist the denaturing proteins during heat stress in folding and also play the role. Since each of the coexpressed genes has its own coexpressed gene network, this network involves from hundreds to thousands of genes together. Any disturbance in the expression of genes forming the network shifts the balance in favour to the network portion which will be most useful under the new condition emerging from the disturbance from the previous condition to help the plant (or any other organism) in maintaining the equilibrium.

References

Aparicio, F., Thomas, C.L., Lederer, C., Niu, Y., Wang, D. & Maule, A.J. (2005). Virus induction of heat shock protein 70 reflects a general response to protein accumulation in the plant cytosol. *Plant Physiology*, 138, 529-36

Banti, V., Mafessoni, F., Loreti, E., Alpi, a. & Perata, P. (2010).The heat-inducible transcription factor HsfA2 enhances anoxia tolerance in Arabidopsis. *Plant Physiology*, 152, 1471-1483.

Barrett, T., Suzek, T.O., Troup, D.B., Wilhite, S.E., Ngau, W., Ledoux, P., Rudnev, D., Lash. A.E., Fujibuchi. W. & Edgar, R. (2004). NCBI GEO: mining millions of expression profiles--database and tools. *Nucleic Acids Research*, 33, D562-D566.

Collinge, M. & Boller, T. (2001). Differential induction of two potato genes, Stprx2 and StNAC, in response to infection by Phytophthora infestans and to wounding. *Plant Molecular Biology*, 46, 521-529.

Craigon, D. J. (2004). NASCArrays: a repository for microarray data generated by NASC's transcriptomics service. *Nucleic Acids Research*, 32, D575D-577.

Ikeo, K., Ishi-I, J., Tamura, T., Gojobori, T. & Tateno, Y. (2003). CIBEX: center for information biology gene expression database. *Comptes Rendus Biologies*, 326, 1079-1082.

Inzé, A., Vanderauwera, S., Hoeberichts, F.A., Vandorpe, M., VAN Gaever, T. & VAN Breusegem, F. (2011) A subcellular localization compendium of hydrogen peroxide-induced proteins. *Plant, Cell & Environment*, doi: 10.1111/j.1365-3040.2011.02323.x

Kanehisa, M., Goto, S., Sato, Y., Furumichi, M. & Tanabe, M. (2011). KEGG for integration and interpretation of large-scale molecular data sets. *Nucleic Acids Research*, doi:10.1093/nar/gkr988

Lamesch, P., Berardini, T.Z., Li, D., Swarbreck, D., Wilks, C., Sasidharan, R., Muller, R., Dreher, K., Alexander, D.L., Garcia-Hernandez, M., Karthikeyan, A.S., Lee, C.H.,Nelson, W.D., Ploetz, L., Singh, S., Wensel, A & Huala, E. (2011). The Arabidopsis Information Resource (TAIR): improved gene annotation and new tools. *Nucleic Acids Research*, doi:10.1093/nar/gkr1090

Manfield, I.W., Jen, C., Pinney, J.W., Michalopoulos, I., Bradford, J.R., Gilmartin, P.M. &. Westhead, D.R. (2006). Arabidopsis Co-expression Tool (ACT): web server tools for microarray-based gene expression analysis. *Nucleic Acids Research*, 34, W504-W509.

Obayashi, T., Hayashi, S., Saeki, M., Ohta, H. & Kinoshita, K. (2009). ATTED-II provides coexpressed gene networks for Arabidopsis. *Nucleic Acids Research*, 37, D987-991.

Parkinson, H (2004). ArrayExpress--a public repository for microarray gene expression data at the EBI. *Nucleic Acids Research*, 33, D553-D555.

Rhee, S.E., Beavis, W. Berardini, T. Z. Chen,G., Dixon, D., Doyle, A., Garcia-Hernandez, M., Huala, E., Lander, G., Montoya, M., Miller, N., Mueller, L.A., Mundodi, S., Reiser, L., Tacklind, J., Weems, D.C., Wu, Y., Xu, I., Yoo, D., Yoon, J. & Zhang, P. (2003). The Arabidopsis Information Resource (TAIR): a model organism database providing a centralized, curated gateway to Arabidopsis biology, research materials and community. *Nucleic Acids Research*, 31, 224 -228.

Riechmann, J.L., Heard, J., Martin, G., Reuber, L., Jiang, C., Keddie, J., Adam, L., Pineda, O., Ratcliffe, O.J., Samaha, R.R., Creelman, R., Pilgrim, M., Broun, P., Zhang, J.Z., Ghandehari, D., Sherman, B.K. & Yu, G. (2000). Arabidopsis transcription factors: genome-wide comparative analysis among eukaryotes. *Science*, 290, 2105-2110.

Sakuma, Y., Maruyama, M., Osakabe, Y., Qin, F., Seki, M., Shinozaki, K., & Yamaguchi-Shinozaki, K. (2006). Functional Analysis of an Arabidopsis Transcription Factor, DREB2A, Involved in Drought-Responsive Gene Expression. *The Plant Cell Online*, 18, 1292 -1309.

Schmid, M., Davison, T.S., Henz, S.R., Pape, U.J., Demar, M., Vingron, M., Schölkopf, B., Weigel, D. & Lohmann, J.U. (2005). A gene expression map of Arabidopsis thaliana development. *Nature Genetics*, 37, 501-506.

Steinhauser, D., Usadel, B., Luedemann, A., Thimm, O. & Kopka, J (2004). CSB.DB: a comprehensive systems-biology database. *Bioinformatics*, 20, 3647 -3651.

Sugio, A., Dreos, R., Aparicio, F. & Maule, A.J. (2009). The cytosolic protein response as a subcomponent of the wider heat shock response in Arabidopsis. The Plant Cell, 21, 642-654.

Sung, D.Y. Vierling, E. & Guy, C.L. (2001). Comprehensive expression profile analysis of the Arabidopsis Hsp70 gene family. *Plant Physiology*, 126, 789-800.

Toufighi, K., Brady, S.M., Austin, R., Ly, E. & Provart, N.J. (2005). The Botany Array Resource: e-Northerns, Expression Angling, and promoter analyses. *The Plant Journal: For Cell and Molecular Biology*, 43, 153-163.

Venkatesan, K., Rual, J.F., Vazquez, A., Stelzl, U., Lemmens, I., Hirozane-Kishikawa, T., Hao, T., Zenkner, M., Xin, X., Goh, K.I., Yildirim, M.A., Simonis, N., Heinzmann, K., Gebreab, F., Sahalie, J.M., Cevik, S., Simon, C., de Smet, A.S., Dann, E., Smolyar, A., Vinayagam, A., Yu, H., Szeto, D., Borick, H., Dricot, A., Klitgord, N., Murray, R.R., Lin, C., Lalowski, M., Timm, J., Rau, K., Boone, C., Braun, P., Cusick, M.E., Roth, F.P., Hill, D.E., Tavernier, J., Wanker, E.E., Barabási, A.L. & Vidal, M. (2009). An empirical framework for binary interactome mapping. *Nature Methods*, 6, 83-90

Zimmermann, P., Hennig, L., & Gruissem, W. (2005). Gene-expression analysis and network discovery using Genevestigator. *Trends in Plant Science*, 10, 407-409.

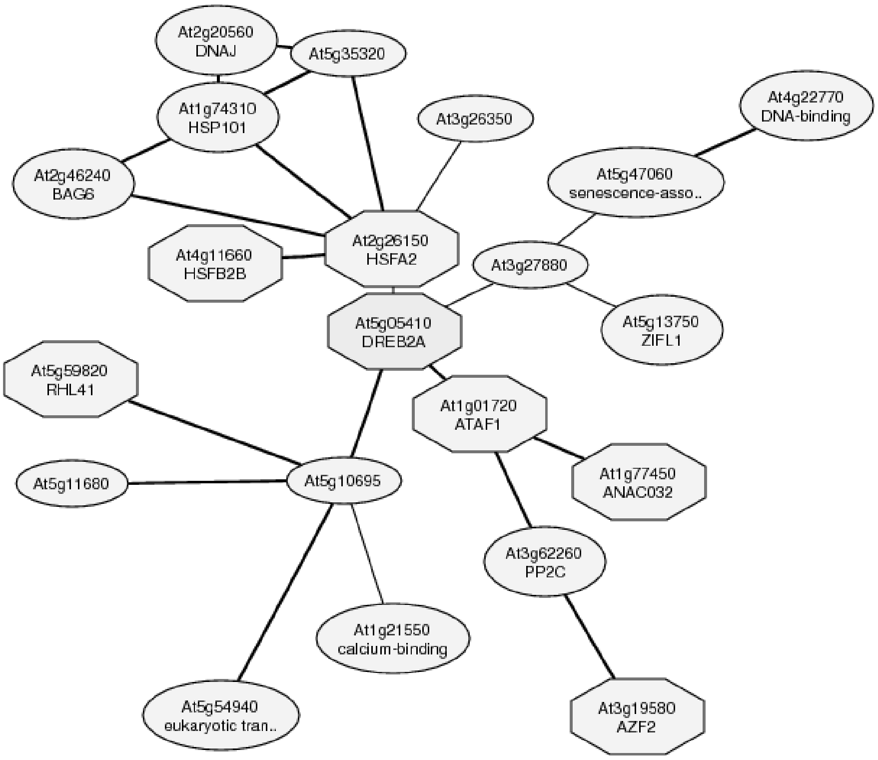


Fig.1. Coexpressed gene network around At5g05410

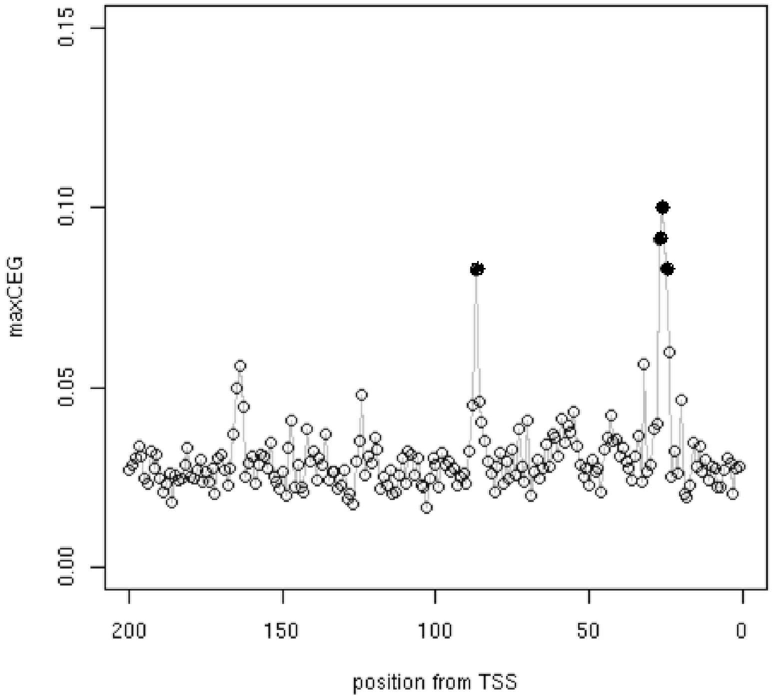
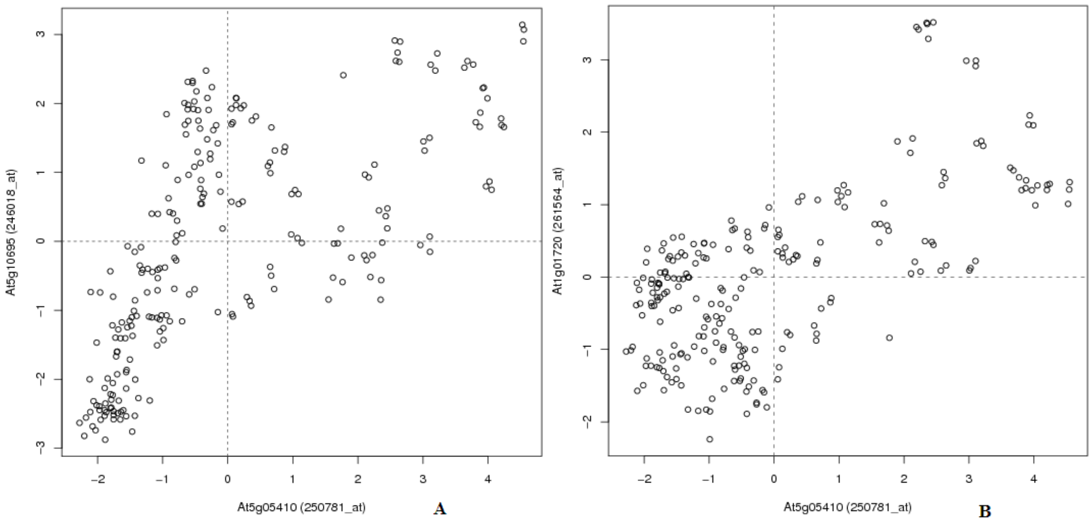


Fig. 2. Predicted cis elements for At5g05410 (those with high CEG are filled black)



A

B

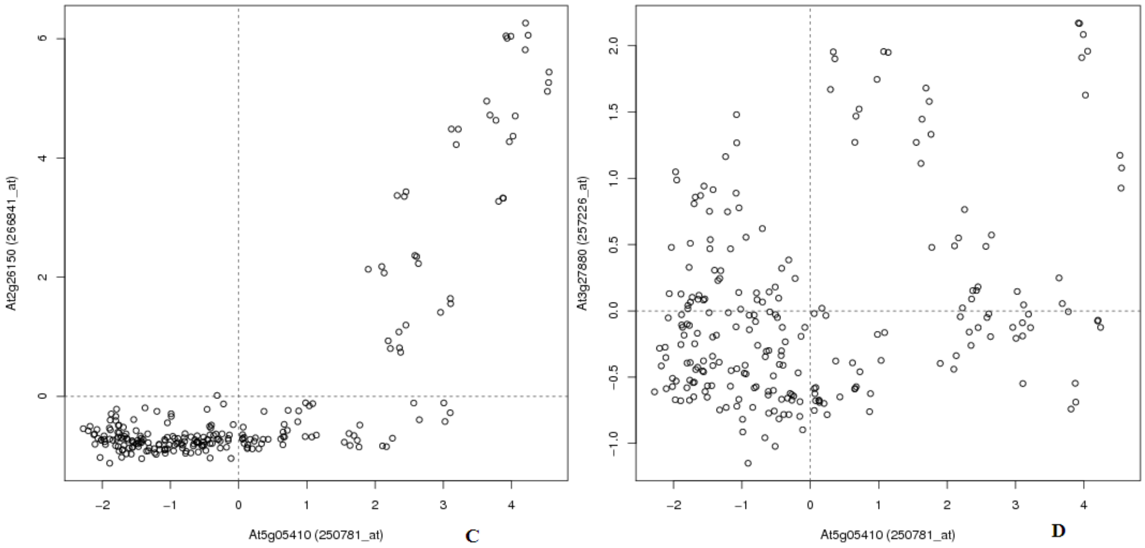


Fig. 3. Correlation of expression between At5g05410 and other genes directly connected with it on the network during development, (a) At5g10695, (b) At1g01720, (c) At2g26150 and (d) At3g27880.

Table 1. List of coexpressed genes with At5g05410.

S.No.	**Locus**	**Alias**	**Mutual Rank**	**S.No.**	**Locus**	**Alias**	**Mutual Rank**
1	At5g10695	-	1.4	**26**	At5g12030	HSP17.6A	28.3
2	At1g01720	ATAF1	4.5	**27**	255155_at	-	28.4
3	At2g26150	HSFA2	8.5	**28**	At4g11660	HSFB2B	32.9
4	At3g11020	DREB2B	9.4	**29**	At3g09350	armadillo/beta-catenin repeat	32.9
5	At3g27880	-	9.6	**30**	At5g37670	HSP15.7-CI	34
6	At2g46240	BAG6	10.7	**31**	At3g23170	-	34.1
7	At5g62020	HSFB2A	11	**32**	At4g12410	auxin-responsive	34.7
8	At1g14200	zinc finger	13	**33**	At2g29500	HSP17.6B-CI	35.3
9	At3g62260	PP2C	13.4	**34**	At3g10020	-	36.7
10	At2g20560	DNAJ	14.3	**35**	256356_s_at	-	37.1
11	At5g59820	RHL41	15.2	**36**	At1g53540	HSP17.6C-CI	40.4
12	At5g03210	-	16.3	**37**	At1g55530	zinc finger	41.1
13	At1g74310	HSP101	16.8	**38**	At3g08970	ERDJ3A	42.5
14	At4g15420	PRLI-interacting factor K	19	**39**	At4g27657	-	43.4
15	At4g25380	SAP10	20.9	**40**	At1g16030	Hsp70b	44.2
16	262911_s_at	-	21.2	**41**	At3g63350	HSFA7B	44.3
17	At3g14200	N-terminal	22.1	**42**	At2g32120	HSP70T-2	45.5
18	At1g26800	zinc finger	22.4	**43**	At4g26080	ABI1	46
19	At3g46230	HSP17.4	22.8	**44**	At3g48070	protein binding	46.5
20	At5g04340	ZAT6	23.8	**45**	At1g60190	U-box	46.7
21	263823_s_at	-	24.4	**46**	At5g12020	HSP17.6II	47.3
22	At1g71000	heat shock binding	25.4	**47**	At1g52560	HSP26.5-P	48.7
23	At3g24500	MBF1C	26.5	**48**	At5g11680	-	49.5
24	At5g35320	-	27.9	**49**	At3g17611	RBL14	50.7
25	At3g17110	-	28	**50**	At3g12580	HSP70	51.4

Table 2. Cis element with high CEG (correlation between expression and a defined group of genes) value

Position	**Cis element**	**CEG**
25	ATAAATA	0.08
26	TATAAAT	0.10
27	CTATAAA	0.09
87	GACACGT	0.08

Table 3. Genes directly connected with At5g05410 on the network

Mutual Rank	**Correlation**	**Locus**	**Function**
1.4	0.69	At5g10695	unknown protein
4.5	0.67	At1g01720	ATAF1; transcription activator/ transcription factor
8.5	0.69	At2g26150	ATHSFA2; DNA binding / transcription factor
9.6	0.49	At3g27880	unknown protein

This academic article was published by The International Institute for Science, Technology and Education (IISTE). The IISTE is a pioneer in the Open Access Publishing service based in the U.S. and Europe. The aim of the institute is Accelerating Global Knowledge Sharing.

More information about the publisher can be found in the IISTE's homepage:
http://www.iiste.org

The IISTE is currently hosting more than 30 peer-reviewed academic journals and collaborating with academic institutions around the world. **Prospective authors of IISTE journals can find the submission instruction on the following page:** http://www.iiste.org/Journals/

The IISTE editorial team promises to the review and publish all the qualified submissions in a fast manner. All the journals articles are available online to the readers all over the world without financial, legal, or technical barriers other than those inseparable from gaining access to the internet itself. Printed version of the journals is also available upon request of readers and authors.

IISTE Knowledge Sharing Partners

EBSCO, Index Copernicus, Ulrich's Periodicals Directory, JournalTOCS, PKP Open Archives Harvester, Bielefeld Academic Search Engine, Elektronische Zeitschriftenbibliothek EZB, Open J-Gate, OCLC WorldCat, Universe Digtial Library , NewJour, Google Scholar