BIOLOGIA (PAKISTAN)
December, 2018, 64 (II), 275-280

PKISSN 0006 – 3096 (Print)
ISSN 2313 – 206X (On-Line)

Cis-acting regulatory elements and transcription factors as a key regulator in plant gene expression

HIRA MUBEEN[1, 2*], AMMARA NASEEM[1], AMMARA MASOOD[2], SHAHID RAZA[3] & NAUREEN NAEEM[3]

[1]*Government College University, Faisalabad, Pakistan*
[2]*University of Central Punjab, Lahore, Pakistan*
[3]*Lahore Garrison University, Lahore, Pakistan*

ARTICLE INFORMAION
Received: 02-08-2018
Received in revised form: 31-12-2018
Accepted: 30-01-2019

***Corresponding Author:**

Hira Mubeen:
hira_sh@hotmail.com

A Review Article

ABSTRACT

Gene expression regulation is an important molecular process for monitoring overall expression level of genes in plants. The expression of genes is regulated by certain regulatory elements including, *Cis*-acting regulatory elements (CRE), transcription factors (TFs) and promoters. Cis-acting elements are actually specific class of DNA binding proteins that act at particular site of DNA. Promoters are part of DNA fragment essential for transcriptional regulation of genes with certain transcription factors. These transcription factors could be useful in developmental regulation, interpretation and validation of candidate genes. This review highlights the importance of C*is*-acting regulatory elements and transcription factors for regulation of gene expression. Furthermore, the use of bioinformatics approach for identification of transcription factors and putative motifs within the promoter region has gained much success for studying DNA-protein interactions. These findings promote the importance of CRE and TF inactivation of transcriptional networks for enhanced gene expression studies in plants.

Keywords: promoter, TF, CRE, transcriptional network

INTRODUCTION

Understanding molecular mechanisms associated with regulatory control parameters is an important key to success. Gene expression is a useful tool to study genetic regulation and process of transcription. Expression of genes in plants is highly controlled by transcriptional regulators. Modern technologies, which are specific for genetic engineering of plants and improvement of crops have been discovered. The development of high throughput methods for stable expression of genes in plants is valuable for many viewpoints. One important approach for plants is to cope up with all the environmental changes due to stress conditions (Baena *et al*, 2010; Lauria *et al*, 2011). Growth of a plant is completely dependent on variety of such factors, which are responsible for controlling mechanism of expression and regulation.

Various molecular processes, in all living organisms, control gene regulatory networks. The binding of DNA elements with their recognition protein factors is also a controlled process under umbrella of specific genes. In plants, the overall control of genes in developmental processes has been studied previously (Ahmad *et al.,* 2010). Genetic engineering of plants includes insertion of foreign genes in other plants by modifying traits as desired. Production of such plants with desired traits can be achieved by specialized transformation methods. Several trials are in practice for stable transformation of genes in higher plants. The stable integration of exogenous DNA in the plastid genome of a unicellular alga, Chlamydomonasital reinhardtii have been reported (Boynton *et al*., 1988; Blowers *et al*., 1989; Boynton *et al.,* 1990).

Expression and regulation of plants are dependent on several motifs and transcription factors, which play an important role as its machinery. The presence of these major elements in plant genome contributes in control of gene expression level by recording overall interactions among regulatory proteins. The function of genes can be specified and predicted by expression and action of specialized regulatory elements or motifs. The key regulators involved in the whole process of regulation are classified according to their particular structures. All structural motifs allow binding to specific sequences which are called as DNA binding domains. One of the transcription factor

Author's Contribution: H.M., Performed Insilco analysis, analyzed the data, and wrote manuscript; **A.M.**, Helped for analyzing data; **A.N., & N.N.**, Performed critical reading and editing of manuscript; **S.R.,** Compared and compiled findings

from *Arabidopsis thaliana* is AGL3, which encodes for MADS domain. These are well known member of MADS box family and plays an important role as transcription factors (Mubeen *et al.,* 2018). Furthermore, these C*is*-acting regulatory elements have linear structure comprising of nucleotide fragments of non-coding DNA (Venter & Botha*,* 2010).

To measure the rate of transcription and gene regulation at wide scale, several elements of promoter including transcription factors, transcription start sites, and motif plays an important role. Promoters consists of two main parts: The first part represents "core promoter", lies within 100-250 bp around the transcription start site. The second part represents a distal part, which contains the element that regulates the spatio-temporal expression (Mubeen *et al*., 2016). The, identification of these factors and regulatory elements will be useful for studying the function of conserved motif within the promoter sequence. Still, due to lack of modern technologies for validation of motifs leads towards the path of limitation. The regulation of genes is also considered as an important regulatory step for building relationship among gene networks (Mazzucotelli *et al.,* 2008).

Transcription activation and regulation of genes is completely dependent on interaction of promoters and transcription factors. Transcription factors binds at specific sites on promoter sequences known as transcription factor binding sites, and also consists of vast amount of *cis*-regulatory elements. These elements are helpful to understand the spatial and temporal features of promoters. Promoters are important regulatory elements essential for transcription of all genes. These are responsible for initiation of transcription and gene regulation control process. The activity of promoter is truly dependent on its size, copy number and position.

In this review, we have highlighted the core mechanism of gene expression and transcriptional regulation for identification of C*is*-regulatory elements and transcription factors by using bioinformatics approach.

Finding transcription factors and *Cis*-regulatory elements using bioinformatics approach

The identification of *cis*-regulatory elements (CRE) enhance our understanding of gene regulation and expression process (Cai *et al*., 2010). CRE consists of several transcription factor binding sites which allows specific transcription factors (TFs) to recognize their best fit and start the regulatory process. Uptil now, studies related to CRE have shown only few combinations of transcription factors. However, this is really complex mechanism to understand as to how different plants share TFs and how these TFs are actively involved in regulation of plant genes. We have identified few Cis-regulatory elements in the promoter sequence of Sucrose Phosphate synthase SPS) promoter, their TFs, TFBs and some of the DNA specific binding domains.

Currently, the use of computational methods for prediction of regulatory elements of promoter region is a common approach. The promoter region of sucrose phosphate synthase gene was matched for finding transcription factors and C*is*-acting regulatory motif using high throughput genome sequencing. Sucrose phosphate synthase is a key enzyme involved in conversion of fructose-6-phosphate and UDP-glucose into sucrose-6-phosphate. This is useful in biosynthesis of plants (Winter & Huber, 2003). Sucrose phosphate synthase catalyzes the first step in the synthesis of sucrose in photosynthetic tissues. The activity of SPS has been shown to be highly regulated at different transcriptional and post-transcriptional levels (Rubab *et al*., 2017).One of the motif from family (bHLH) of transcription factors found was as ACE (ACGTGGA) and plays an important role as a transcriptional activator involved in cell elongation. The findings are shown in Table 1 below: The presence of ACE motif was useful for understanding the expression regulation of genes involved in cell expansion and binding to G-box motif. Whereas, the bHLH plays a vitol role as one of the important protein structural motif characterizing large families of transcription factors (Massar *et al.,* 2000). Moreover, the bHLH transcription factors are important in development of cell activity. The transcription factor (AT1G68920)for ACE motif and (AT5G5187) was searched in Plant PAN database for AE-box for its relationship with other regulatory elements. The resulting transcription factor binding sequence was found at position 0619 with the following tandem repeat (CACGTG) for ACE and at position 0470 with tandem repeat (TTCCAAATGGAA) as shown in Figure 1 and 2 below.

Table 1: *Cis*-regulatory motifs in *SPS* promoter.

***Cis*-regulatory element**	**Organism**	**Sequence**	**TF family**	**TF ID**	**Function**
ACE	*Arabidopsis thaliana*	ACGTGGA	bHLH	AT1G68920	*cis*-acting element involved in light responsiveness
AE-Box	*Arabidopsis thaliana*	AGAAACAT	MADS box	AT5G5187	Active part of a module for light response
ATE	*Arabidopsis thaliana*	CGGTCAAC	WRKY	AT1G13960	Act as a sequence specific DNA binding protein
ATE	*Arabidopsis thaliana*	GGTCAA	WRKY	AT1G55600	DNA binding transcription factor binding protein activity

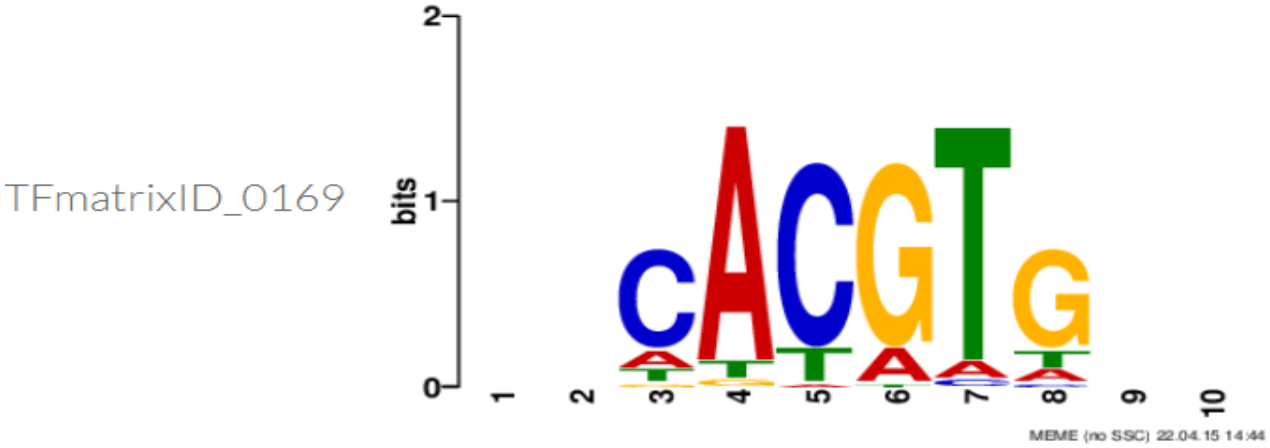

Fig. 1: Shows the TF binding sequence of ACE (AT1G68920) obtained from PlantPAN 2.0.

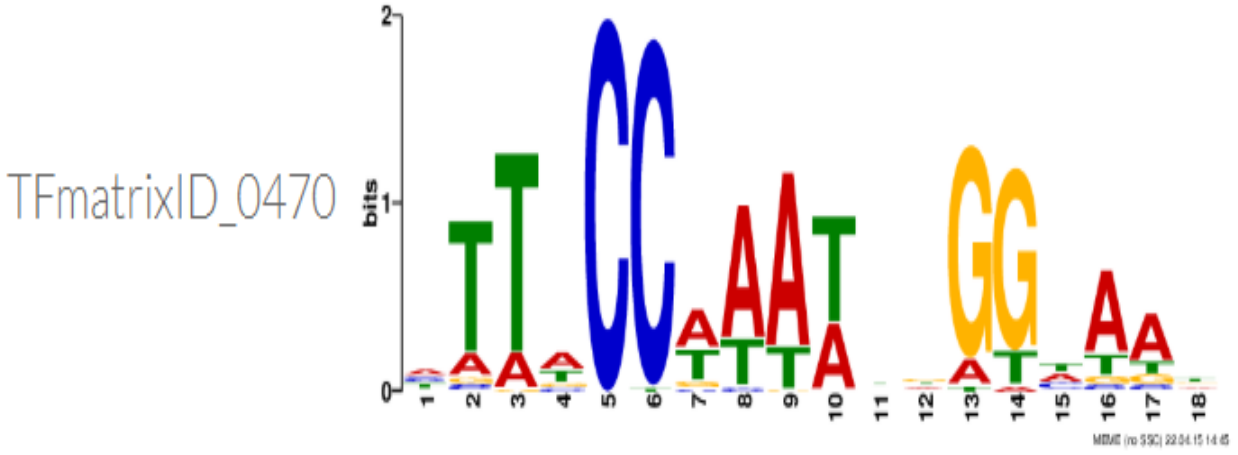

Fig. 2: Shows the TF binding sequence of AE (AT5G5187) obtained from PlantPAN 2.0.

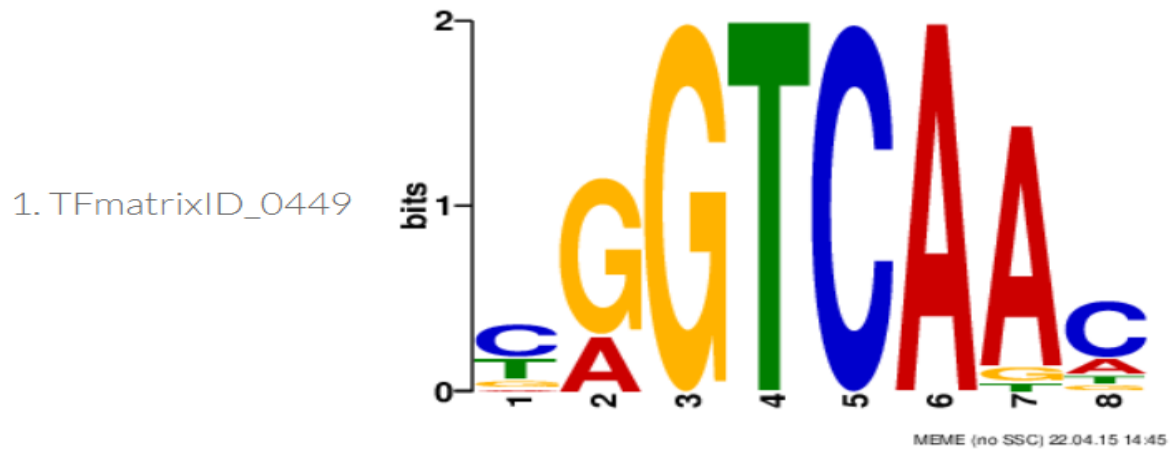

Fig. 3: Shows TF binding sequence of ATE (AT1G13960)obtained from PlantPAN 2.0.

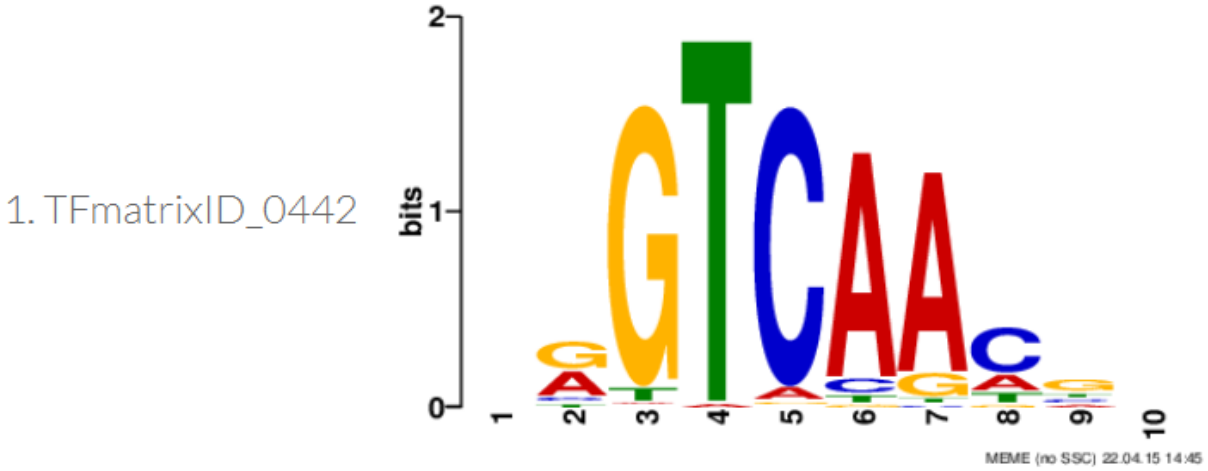

Fig. 4: Shows TF binding sequence of ATE (AT1G55600)obtained from PlantPAN 2.0.

Fig. 5: Shows 3 domains of specific size at different locations of WRKY .

WRKY DNA binding domains

The WRKY domain consists of 60 amino acids having conserved sequence WRKYGQK at N-terminal along with a zinc-finger motif. Many of the plant transcription factors are rich with WRKY domains, which are involved in regulation of various functions including DNA binding specificity, host defense, biosynthesis of secondary metabolites. The WRKYGQK residues give strength to N-terminal and also enables hydrophobic interactions and also provides structural stability to beta-sheet.

Table 2: Shows DNA binding domains in TF of *SPS* promoter

Domain	TF Family	Motif	Position: Start-End
NB-ARC	WRKY	ATE	157-388
LRR-3	WRKY	ATE	577-596
WRKY	WRKY	ATE	1210-1270

Promoter: as a transcriptional control unit

The use of promoters and transcription factors predicted specific transcription factor binding sites appears as one of the key fact in controlling expression profiles and regulatory networks. All target genes are regulated by specialized promoters attached to the upstream of gene. However, this regulation depends on spatial and temporal expression patterns. Interaction among regulatory proteins will help to control the plant gene expression level by cis-acting regulatory elements. However, the plant regulatory elements are present in the transcribed DNA strand or can be added during posttranslational modifications (Vaughn *et al*., 2012). The network of regulatory elements making up a complex of gene structures results after interaction among regulatory proteins called trans elements and *cis*-acting elements.

The level of gene expression can be active or non-active or can be slow or it can shows spatial pattern. The network of transcription factors are key regulators of overall expression profiling of genes and depends on protein transport and efficiency (Li *et al.* 2012).

The use of promoters is useful for understanding the events within transcriptional regulatory networks. Promoters are categorised in different types: constitutive, inducible, tissue specific and synthetic. However, one of the fourth type is synthetic promoters. These are more valuable with desired characteristics for controlling gene expression. Synthetic promoters can be designed according to structure and organization of regulatory elements. (Bhullar *et al*., 2003; Mehrotra *et al*.,2011).The number of cis- regulatory elements and transcription factors in model organisms can be identified, by various bioinformatics tools (Thomas and Chiang, 2006). Further, for analysis of sequence specific motifs by transcription factors, the first and far more important task is to identify transcription factor binding site (TFB) **(**Carey *et al*., 2009**)**.

Future directions

Plant biotechnology involves variety of signaling pathways being regulated under the umbrella of transcriptional regulatory networks modulated by transcription factors and cis acting regulatory elements for controlling the expression of target genes. Many of the TFs are involved in regulation of diverse genes and expression patterns in *Arabidopsis.* Understanding the diverse expression patterns will help to explore more deeper functional studies on a transcriptional scale. Moreover, use of high-throughput data analysis techniques, it is now more easy to predict possible TFBs and TFs within a promoter sequence. With the use of modern technology approach and bioinformatics, the interaction among all transcriptional network members and regulatory proteins can be found easily. This will help to understand more details of plant molecular genomics.

REFERENCES

Baena-González, E. (2010). Energy signaling in the regulation of gene expression during stress. *Molecular plant*, *3*(2), 300-313.

Lauria, M., & Rossi, V. (2011). Epigenetic control of gene regulation in plants. *Biochimicaet Biophysica Acta (BBA)-Gene Regulatory Mechanisms*, *1809*(8), 369-378.

Ahmad, A., Zhang, Y., & Cao, X. F. (2010). Decoding the epigenetic language of plant development. *Molecular plant*, *3*(4), 719-728.

Boynton, J. E., Gillham, N. W., Harris, E. H., Hosler, J. P., Johnson, A. M., Jones, A. R., & Shark, K. B. (1988). Chloroplast transformation in Chlamydomonas with high velocity micro-projectiles. *Science*, *240*(4858), 1534-1538.

Blowers, A. D., Bogorad, L., Shark, K. B., & Sanford, J. C. (1989). Studies on Chlamydomonas chloroplast

transformation: foreign DNA can be stably maintained in the chromosome. *The Plant Cell*, *1*(1), 123-132.

Boynton, J. E., Gillham, N. W., Harris, E. H., Newman, S. M., Randolph-Anderson, B. L., Johnson, A. M., & Jones, A. R. (1990). Manipulating the chloroplast genome of Chlamydomonas molecular genetics and transformation. In *Current research in photosynthesis* (pp. 2415-2422). Springer, Dordrecht.

Mubeen, H., Masood, A., Bashir, A., Raza, S., (2018).Identification and Analysis of DNA binding specific transcription factor binding sites in sucrose synthase promoter. *Pak. J. Bot*, *49*(4), 1105-1112.

Venter, M., & Botha, F. C. (2010). Synthetic promoter engineering. In *Plant Developmental Biology-Biotechnological Perspectives* (pp. 393-414). Springer, Berlin, Heidelberg.

Mazzucotelli, E., Mastrangelo, A. M., Crosatti, C., Guerra, D., Stanca, A. M., &Cattivelli, L. (2008). Abiotic stress response in plants: when post-transcriptional and post-translational regulations control transcription. *Plant Science*, *174*(4), 420-431.

Cai, X., Hou, L., Su, N., Hu, H., Deng, M., & Li, X. (2010). Systematic identification of conserved motif modules in the human genome. *BMC genomics*, *11*(1), 567.

Winter, H., & Huber, S. C. (2000). Regulation of sucrose metabolism in higher plants: localization and regulation of activity of key enzymes. *Critical Reviews in plant sciences*, *19*(1), 31-67.

Lunn, J. E., &MacRae, E. (2003). New complexities in the synthesis of sucrose. *Current opinion in plant biology*, *6*(3), 208-214.

Massari, M. E., &Murre, C. (2000). Helix-loop-helix proteins: regulators of transcription in eucaryotic organisms. *Molecular and cellular biology*, *20*(2), 429-440.

Vaughn, J. N., Ellingson, S. R., Mignone, F., & von Arnim, A. (2012). Known and novel post-transcriptional regulatory sequences are conserved across plant families. *RNA*, *18*(3), 368-384.

Li, W. J., Dai, L. L., Chai, Z. J., Yin, Z. J., & Qu, L. Q. (2012). Evaluation of seed storage protein gene 3′-untranslated regions in enhancing gene expression in transgenic rice seed. *Transgenic research*, *21*(3), 545-553.

Bhullar, S., Chakravarthy, S., Advani, S., Datta, S., Pental, D., & Burma, P. K. (2003). Strategies for development of functionally equivalent promoters with minimum sequence homology for transgene expression in plants: cis-elements in a novel DNA context versus domain swapping. *Plant Physiology*, *132*(2), 988-998.

Mehrotra, R., Gupta, G., Sethi, R., Bhalothia, P., Kumar, N., & Mehrotra, S. (2011). Designer promoter: an artwork of cis engineering. *Plant molecular biology*, *75*(6), 527-536.

Carey, M. F., Peterson, C. L., & Smale, S. T. (2009). Chromatin immune-precipitation (chip). *Cold Spring Harbor Protocols*, *2009*(9), pdb-prot5279.

Thomas, M. C., & Chiang, C. M. (2006). The general transcription machinery and general cofactors. *Critical reviews in biochemistry and molecular biology*, *41*(3), 105-178.