

Original Research Article**DOI: 10.26479/2022.0804.01**

COMPUTATIONAL ANALYSIS OF CARE ELEMENTS IN MYB TRANSCRIPTION FACTOR GENES OF SUGARCANE

M. Babu, S. Geethalakshmi*

Department of Biotechnology, Sree Narayana Guru College,
KG Chavadi, Coimbatore, Tamilnadu, India.

ABSTRACT: Plant transcription factors are involved in practically all of the plant's functions. *MYB* transcription factor genes implicated in abiotic stress response have been examined extensively with the available EST data in our research on sugarcane, a crop that has lately acquired relevance due to its usage in biofuel production. However, little is known about the promoter organization and its role in the abiotic stress response. Due to the polyploidy of sugarcane, the entire gene sequence and regulatory sequences have yet to be determined. *Sorghum bicolor* *MYB* gene sequences were compared with *Saccharum* species utilizing computational techniques in an attempt to address this deficiency in a closely related plant species. As a result, the findings of this study provide information on the CARE ELEMENTS distribution, which can be used to investigate MYB regulatory elements in upstream sequences.

Keywords: cis-acting regulatory elements, promoter, transcription factor, MYB, Sugarcane.

Article History: Received: June 28, 2022; Revised: July 15, 2022; Accepted: July 29, 2022.

Corresponding Author: Dr. S. Geethalakshmi* Ph.D.

Department of Biotechnology, Sree Narayana Guru College,
KG Chavadi, Coimbatore, Tamilnadu, India. Email Address: s.geethalakshmi@gmail.com

1. INTRODUCTION

Plants are exposed to a variety of stress stimuli, which cause coordinated changes in gene expression at the transcriptional level and, as a result, transcription factors. Some stress-responsive transcription factors from the NAC, AP2/ERF, MYB, and WRKY families that mediate plant response and tolerance to abiotic stress have been found through extensive research. These findings suggest that determining the mechanisms underlying plant stress responses and tolerance to abiotic stress

requires transcriptional regulation of stress-responsive genes and that these transcription factors may be important targets for the development of crops with improved abiotic stress tolerance [1]. MYB transcription factor regulates stress responses as well as the activation of drought-related genes, which aids in the improvement of drought stress tolerance in crops. *MYB* proteins are thought to be involved in controlling plant responses to abiotic stress, according to research [2], [3], [4], [5]. Transcription factors bind to certain regions in DNA called cis-acting regulatory elements (CAREs), which then mediate transcription initiation. CAREs are short, conserved motifs that are found near the promoter at the 5' end and are usually 5-20 nucleotides long. Promoters, enhancers, insulators, and silencers are all part of this system. Cis sequences and trans factors are two types of regulators, according to their structure. Linear nucleotide segments of non-coding DNA make up cis-regulatory sequences. Their position and direction regarding genes and activity are different [6]. When the plant is subjected to a stimulus, these components help to differentially expressed transcription factors. Under varied environmental settings, the interaction of CAREs with transcription factors determines and regulates transcription start events [7], [8]. Depending on the gene, the Cis sequence can have different copy counts, distances, and orientations [6], [9]. Depending on the gene, the Cis sequence can have different copy counts, distances, and orientations [6]. The PlantCARE database and currently available analysis tools make it feasible to learn about plant promoters and how they affect gene expression [10], [11]. These motifs are difficult to identify since they appear at random sites and actual regulatory elements are difficult to separate from misleading elements [12]. Sugarcane is one of the important tropical crops that is the major source of sugar worldwide. Recent interest in sugarcane has been increasing because of its use in biofuel production [13]. The most common abiotic stress encountered by sugarcane is drought and salinity since it is a tropical crop [14]. In this study, for the identification of CAREs in the promoters of genes involved in stress/stimuli response, the promoter regions of putative sugarcane MYB genes were analyzed to determine the motifs present that function as CAREs. Because there is no whole genome for plants like sugarcane and the existing EST data is insufficient for such research, the process becomes much more difficult. This study aims at identifying the CARE elements present in the upstream regions of the MYB transcription factor genes involved in abiotic stress response using computational analysis in sugarcane.

2. MATERIALS AND METHODS

Sequences of 52 MYB *Saccharum* genes (HF679407.1 to HF679441.1) were identified in our previous research investigation [15] and were blasted with SbGDB (*Sorghum bicolor*), and MaizeGDB. *Sorghum* and maize are evolutionarily connected to *Saccharum* species, according to studies of genetic diversity correlated with chromosome positions [16], [17]. The *Sorghum* hits with more than 90% match and start site were collected. The >90% match sorghum hit genes upstream sequence of -1200 details were collected. These upstream sequences were analyzed using

PLANTCARE, PlantPAN, and PLACE databases to check for the distribution of CARE elements. These are the databases that contain data about the Plant CARE elements. This analysis yielded a total of 26 *Saccharum* MYB genes implicated in stress/stimuli response. The entire cis-elements located in the upstream regions were consolidated and represented in the form of a figure [10], [11].

3. RESULTS AND DISCUSSION

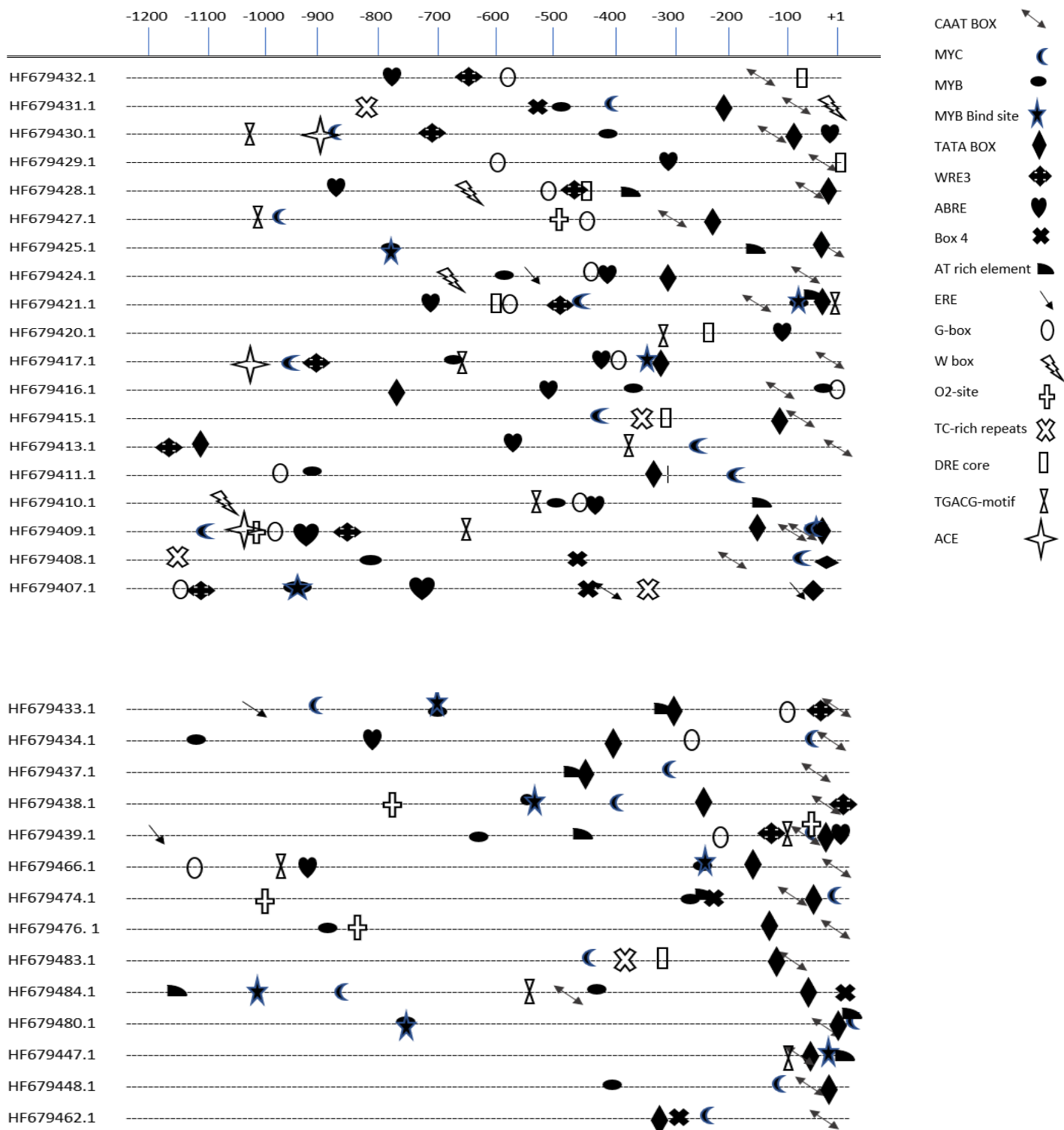
Studies of genetic diversity correlated with chromosome locations have shown that Sorghum and maize are evolutionarily related to *Saccharum* species [16], [17]. Hence for the determination of putative promoter regions for the identified *Saccharum* MYB genes, the data available in SbGDB and MaizeGDB were used. The retrieved 1200 bases prior to the translation start site of the sugarcane MYB genes, when subjected to TSSP (Transcription Start Site Prediction Tool), showed that the CAP site in all the genes showed the pattern pyApypy which correlates with the pattern proposed by [18]. Moreover, the distance between TSS and +1 site was found to be on average, ranging from 50 bp to 450 bp in length [19]. After the TSS determination, the upstream sequences were subjected to CARE analysis using Plant CARE [10], PLANTPAN [20] and PLACE [21] tools. This was done to determine the consensus domains present in the promoter regions. Various domains were found to be present in the upstream region regulating the function of the MYB genes. The TATA box showed the consensus TATA, TATAA, and TTTTA. The position of the TATA box ranged approximately from -25 to -50 [23]. The CAAT box was widely distributed with no specific location range and had the consensus of CAAT, CAAAT, and CCAAT. But in all the cases, the CAAT box was upstream of the TATA box [23]. The Plant CARE analysis of the putative promoter regions showed various motifs tabulated in Table I.

Table I: CARE Elements and their functionality:

Cis Element	Function	Sequence
CAP SITE	Transcription start site	pyApypy
TATA BOX	Core promoter element around - 30 of transcription start site	TATA,TATAA, TTTTA
CATT BOX	Common in promoter and enhancer region	CAAT, CAAAT, CCAAT
5'UTR Py-RICH REPEAT	high transcription levels	TTTCTTCTCT
ACE	light responsiveness	GCGACGTACC
ABRE	abscisic acid responsiveness	TACGGTC
AT Rich Element	Binding site of DNA Binding Proteins	ATAGAAATCAA
BOX-4	light responsiveness	ATTAAT
DRE Core	Dehydration Response Element	GCCGAC
ERE	Ethylene response	ATTTCAAA
G BOX	light responsiveness	CACGTT
MYC	Plant growth and Defense Response	TCTCTTA
MYB	Hormone signal transduction and abiotic stress response	CAACAG
MBS	MYB binding site involved in drought stress response	TAACTG
O2 SITE	zein metabolism	GATGACATGG
TC RICH REPEATS	defense and stress response	ATTTTCTTCA
TGACG MOTIF	Me-JA responsiveness	TGACG
W Box	Fungal elicitor, belongs to WRKY Family	TTGACC
WRE3	Involved in signaling during anaerobic conditions	CCACCT

The domains were found to have the functions like a hormonal response, stress response, regulatory elements, tissue-specific elements, and light responsiveness. The domains present in the promoter region of the 52 sequences are shown in Figure 1.

Key



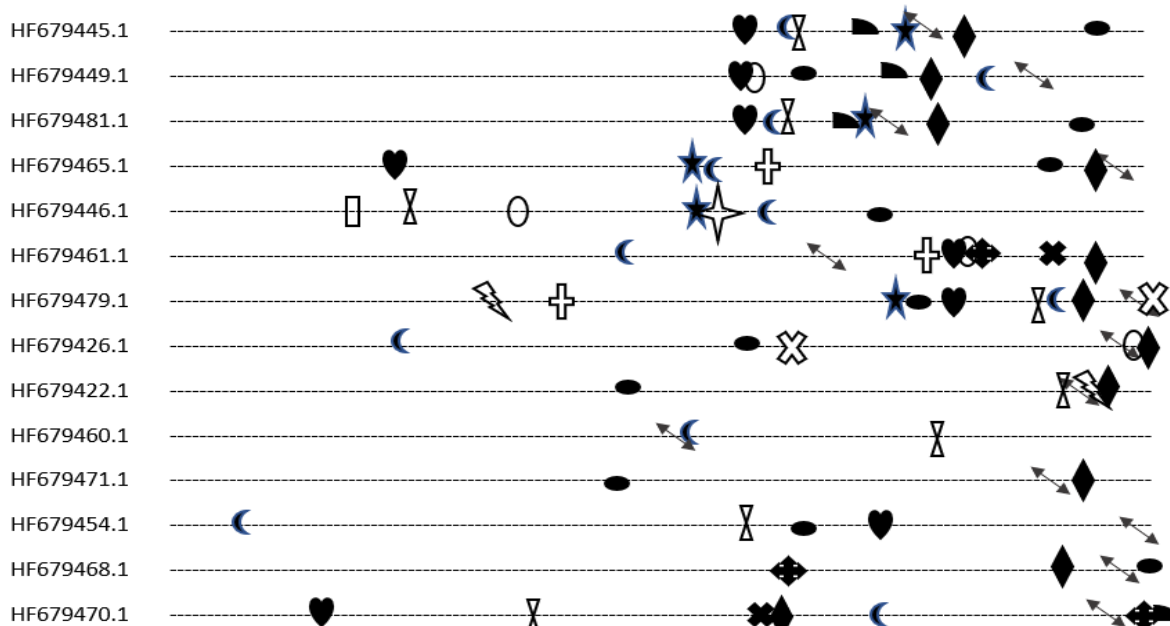


Figure I: Distribution of CARE elements in the upstream region

The percentage occurrence of each category motif from the above sequence is shown in Figure 2. Further study is going to focus on these MYB domains, hence the sequences that contain prominent MYB and MYC domains very close to (approximately within -400 to -450 bp) TATA and CAAT boxes were selected for further study [24]. This is because these domains have a high probability of recognition during gene expression. Hence, from the pool of 52 sequences, 33 were shortlisted based on this parameter.

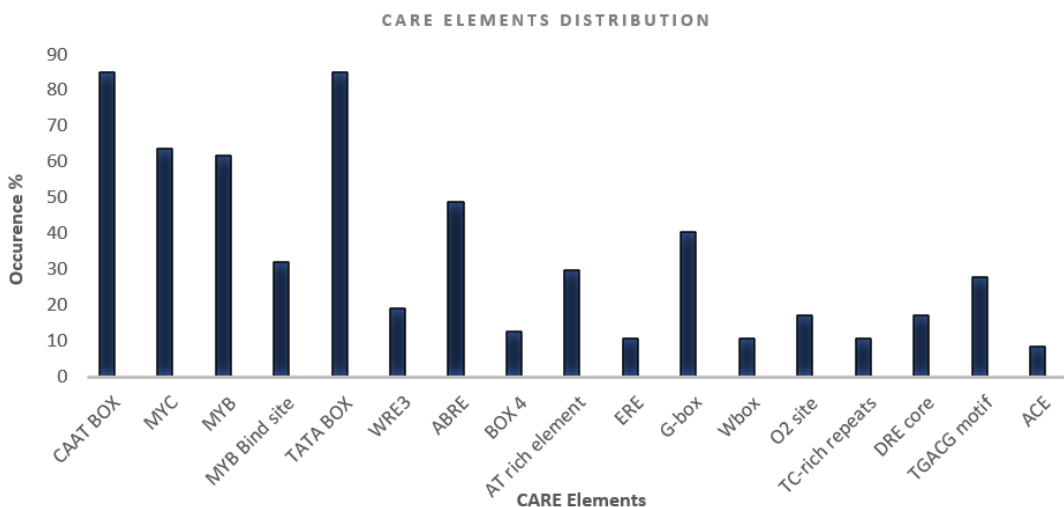


Figure II: CARE elements distribution:

The location distribution of these domains was analyzed and found that in the case of stress-responsive elements, MYB and MYC [22] occur commonly (70% approx.) in most of the MYB promoters. The hormone response elements ABRE (Abscisic acid response element) [25], ERE element (Ethylene Response element) [26], and TGACG (Methyl jasmonate response element) [27] occurs more in the promoter regions (25%, 18.25%, 25% respectively). The element involved in stress and defense response like ACE [28], G-Box [29] and Box 4 [30] (Light response elements) are specific only to MYB gene promoters and occur in all the sequences [31]. The abiotic stress response elements like MYB, MYC, MBS, DRE (Dehydration response element) and TC-rich repeats are widespread among the promoters and are seen to occur in the region close to the transcription start site and polymerase binding site. The W Box (fungal response element) is found in very few sequences (10%) analyzed [32]. The tissue-specific element which regulates the expression of tissues like WRE3, anaerobic metabolism domain, and O₂ site, involved in zein metabolism are less distributed in the MYB genes. The motif which regulates high transcription level, the 5' UTR-Py rich repeat and AT-rich element are more pronounced in all the sequences [33], [34].

4. CONCLUSION

Sugarcane represents one of the important biofuel crops not only involved in sugar production but has various industrial applications including energy cogeneration, dietary fiber, and many more. In the present study, genome-wide searches for full-length cDNA sequences yielded a considerable number of TSS of MYB genes involved in abiotic stress. This study indicates that structural features of the promoters involved in stress resistance govern the function of the gene and help them in expressing specific proteins required for the specific situation. Also, the putative cis-regulatory elements in the promoter region exhibit a high degree of conservation when analyzed in two species, maize, and sorghum. These MYB transcription factor genes when incorporated into the gene regulatory networks can give rise to stress-tolerant plant species showing a high level of resistance in different parts of the plant. Further studies are required for analyzing the precise function of the CAREs which will pave way for the complete analysis of gene networks in an industrial crop like sugarcane.

ETHICS APPROVAL AND CONSENT TO PARTICIPATE

Not applicable.

HUMAN AND ANIMAL RIGHTS

No Animals/Humans were used for studies that are base of this research.

CONSENT FOR PUBLICATION

Not applicable.

AVAILABILITY OF DATA AND MATERIALS

The author confirms that the data supporting the findings of this research are available within the

article.

FUNDING

None.

CONFLICT OF INTEREST

No conflict of interest.

REFERENCES

1. Yoon Y, Seo DH, Shin H, Kim HJ, Kim CM, Jang G. The role of stress-responsive transcription factors in modulating abiotic stress tolerance in plants. *Agron.* 2020 Jun;10(6):788-811.
2. Lv Y, Yang M, Hu D, Yang Z, Ma S, Li X, Xiong L. The OsMYB30 transcription factor suppresses cold tolerance by interacting with a JAZ protein and suppressing β -amylase expression. *Plant physiol.* 2017 Feb;173(2):1475-91.
3. Wang L, Li Z, Lu M, Wang Y. ThNAC13, a NAC transcription factor from *Tamarix hispida*, confers salt and osmotic stress tolerance to transgenic *Tamarix* and *Arabidopsis*. *Front. Plant Sci.* 2017 Apr 26;8:635.
4. Wei Q, Luo Q, Wang R, Zhang F, He Y, Zhang Y, Qiu D, Li K, Chang J, Yang G, He G. A wheat R2R3-type MYB transcription factor TaODORANT1 positively regulates drought and salt stress responses in transgenic tobacco plants. *Front. Plant Sci.* 2017 Aug 8;8:1374.
5. Prabu G, Theertha Prasad D. Functional characterization of sugarcane MYB transcription factor gene promoter (PScMYBAS1) in response to abiotic stresses and hormones. *Plant Cell Rep.* 2012 Apr;31(4):661-9.
6. Venter M, Botha FC. Synthetic promoter engineering. In *Plant Developmental Biology- Biotechnological Perspectives 2010* (pp. 393-414). Springer, Berlin, Heidelberg.
7. Yamaguchi-Shinozaki K, Shinozaki K. Organization of cis-acting regulatory elements in osmotic-and cold-stress-responsive promoters. *Trends Plant Sci.* 2005 Feb 1;10(2):88-94.
8. Sharma N, Russell SD, Bhalla PL, Singh MB. Putative cis-regulatory elements in genes highly expressed in rice sperm cells. *BMC Res. Notes.* 2011 Dec;4(1):1-0.
9. Shamloo-Dashtpajardi R, Razi H, Aliakbari M, Lindlöf A, Ebrahimi M, Ebrahimie E. A novel pairwise comparison method for in silico discovery of statistically significant cis-regulatory elements in eukaryotic promoter regions: Application to *Arabidopsis*. *J. Theor. Biol.* 2015 Jan 7;364:364-76.
10. Rombauts S, Déhais P, Van Montagu M, Rouzé P. PlantCARE, a plant cis-acting regulatory element database. *Nucleic Acids Res.* 1999 Jan 1;27(1):295-6.
11. Lescot M, Déhais P, Thijs G, Marchal K, Moreau Y, Van de Peer Y, Rouzé P, Rombauts S. PlantCARE, a database of plant cis-acting regulatory elements and a portal to tools for in silico analysis of promoter sequences. *Nucleic Acids Res.* 2002 Jan 1;30(1):325-7.
12. Rombauts S, Florquin K, Lescot M, Marchal K, Rouzé P, Van de Peer Y. Computational

- approaches to identify promoters and cis-regulatory elements in plant genomes. *Plant Physiol.* 2003 Jul;132(3):1162-76.
13. Hotta CT, Lembke CG, Domingues DS, Ochoa EA, Cruz GM, Melotto-Passarini DM, Marconi TG, Santos MO, Mollinari M, Margarido GR, Crivellari AC. The biotechnology roadmap for sugarcane improvement. *Trop. Plant Biol.* 2010 Jun;3(2):75-87.
 14. Mahajan S, Tuteja N. Cold, salinity and drought stresses: an overview. *Arch. Biochem. Biophys.* 2005 Dec 15;444(2):139-58.
 15. Geethalakshmi S, Barathkumar S, Prabu G. The MYB transcription factor family genes in sugarcane (*Saccharum sp.*). *Plant Mol. Biol. Rep.* 2015 Jun;33(3):512-31.
 16. Dufour P, Deu M, Grivet L, D'Hont A, Paulet F, Bouet A, Lanaud C, Glaszmann JC, Hamon P. Construction of a composite sorghum genome map and comparison with sugarcane, a related complex polyploid. *Theor. Appl. Genet.* 1997 Mar;94(3):409-18.
 17. Gaut BS, Le Thierry d'Ennequin M, Peek AS, Sawkins MC. Maize as a model for the evolution of plant nuclear genomes. *Proc. Natl. Acad. Sci.* 2000 Jun 20;97(13):7008-15.
 18. Breathnach R, Chambon P. Organization and expression of eucaryotic split genes coding for proteins. *Annu. Rev. Biochem.* 1981 Jul;50(1):349-83.
 19. Shahmuradov IA, Umarov RK, Solovyev VV. TSSPlant: a new tool for prediction of plant Pol II promoters. *Nucleic Acids Res.* 2017 May 5;45(8):e65-.
 20. Chang WC, Lee TY, Huang HD, Huang HY, Pan RL. PlantPAN: Plant promoter analysis navigator, for identifying combinatorial cis-regulatory elements with distance constraint in plant gene groups. *BMC Genom.* 2008 Dec;9(1):1-4.
 21. Higo K, Ugawa Y, Iwamoto M, Korenaga T. Plant cis-acting regulatory DNA elements (PLACE) database: 1999. *Nucleic Acids Res.* 1999 Jan 1;27(1):297-300.
 22. Urao T, Yamaguchi-Shinozaki K, Urao S, Shinozaki K. An Arabidopsis myb homolog is induced by dehydration stress and its gene product binds to the conserved MYB recognition sequence. *Plant Cell.* 1993 Nov;5(11):1529-39.
 23. Joshi CP. An inspection of the domain between putative TATA box and translation start site in 79 plant genes. *Nucleic Acids Res.* 1987 Aug 25;15(16):6643-53.
 24. Porto MS, Pinheiro MP, Batista VG, dos Santos RC, de Albuquerque Melo Filho P, de Lima LM. Plant promoters: an approach of structure and function. *Mol. Biotechnol.* 2014 Jan;56(1):38-49.
 25. Yamaguchi-Shinozaki K, Shinozaki K. Transcriptional regulatory networks in cellular responses and tolerance to dehydration and cold stresses. *Annu Rev Plant Biol.* 2006 Jun 2;57(1):781-803.
 26. Liang H, Lu Y, Liu H, Wang F, Xin Z, Zhang Z. A novel activator-type ERF of *Thinopyrum intermedium*, TiERF1, positively regulates defence responses. *J. Exp. Bot.* 2008 Aug 1;59(11):3111-20.
 27. Rushton PJ, Somssich IE. Transcriptional control of plant genes responsive to pathogens.

Curr. Opin. Plant Biol. 1998 Aug 1;1(4):311-5.

28. Logemann E, Hahlbrock K. Crosstalk among stress responses in plants: pathogen defense overrides UV protection through an inversely regulated ACE/ACE type of light-responsive gene promoter unit. *Proc. Natl. Acad. Sci.* 2002 Feb 19;99(4):2428-32.
29. Menkens AE, Schindler U, Cashmore AR. The G-box: a ubiquitous regulatory DNA element in plants bound by the GBF family of bZIP proteins. *Trends Biochem. Sci.* 1995 Dec 1;20(12):506-10.
30. Argüello-Astorga G, Herrera-Estrella L. Evolution of light-regulated plant promoters. *Annu Rev Plant Biol.* 1998 Jun;49(1):525-55.
31. Chattopadhyay S, Puente P, Deng XW, Wei N. Combinatorial interaction of light-responsive elements plays a critical role in determining the response characteristics of light-regulated promoters in *Arabidopsis*. *Plant J.* 1998;15(1):69-77.
32. Fukuda Y, Shinshi H. Characterization of a novel cis-acting element that is responsive to a fungal elicitor in the promoter of a tobacco class I chitinase gene. *Plant Mol. Biol.* 1994 Feb;24(3):485-93.
33. Jang CS, Kamps TL, Skinner DN, Schulze SR, Vencill WK, Paterson AH. Functional classification, genomic organization, putatively cis-acting regulatory elements, and relationship to quantitative trait loci, of sorghum genes with rhizome-enriched expression. *Plant Physiol.* 2006 Nov;142(3):1148-59.
34. Diaz-De-Leon F, Klotz KL, Lagrimini LM. Nucleotide sequence of the tobacco (*Nicotiana tabacum*) anionic peroxidase gene. *Plant Physiol.* 1993 Mar;101(3):1117.