Journal of Applied and Natural Science
15 (4), 1407 - 1420 (2023)
ISSN : 0974-9411 (Print), 2231-5209 (Online)
journals.ansfoundation.org

Research Article

Molecular cloning, expression and *Insilco* analysis of drought stress inducible MYB transcription factor encoding gene from C4 plant *Eleusine coracana*

Megha Bhatt*
Department of Molecular Biology & Genetic Engineering, College of Basic Science and Humanities, G. B. Pant University of Agriculture & Technology, Pantnagar (Uttarakhand), India
Prafull Salvi
Agri-Biotechnology Department, National Agri-Food Biotechnology Institute, Mohali, (Punjab), India
Ankita Chauhan
Department of Molecular Biology & Genetic Engineering, College of Basic Science and Humanities, G. B. Pant University of Agriculture & Technology, Pantnagar (Uttarakhand), India
Mihir Joshi
Faculty of Management Studies, WISDOM, Banasthali Vidyapith University (Rajasthan) India
Pushpa Lohani
Department of Molecular Biology & Genetic Engineering, College of Basic Science and Humanities, G. B. Pant University of Agriculture & Technology, Pantnagar (Uttarakhand), India

*Corresponding author. E-mail: Megha4bhatt@gmail.com.

Article Info
https://doi.org/10.31018/jans.v15i4.4891
Received: July 5, 2023
Revised: November 17, 2023
Accepted: November 25, 2023

How to Cite
Bhatt, M. *et al.* (2023). Molecular cloning, expression and *Insilco* analysis of drought stress inducible MYB transcription factor encoding gene from C4 plant *Eleusine coracana*. *Journal of Applied and Natural Science*, 15 (4), 1407 - 1420. https://doi.org/10.31018/jans.v15i4.4891

Abstract
Drought is one of the key abiotic stresses that critically influences the crops by restraining their growth and yield potential. Being sessile, plant tackle the detrimental effects of drought stress by modulating the cellular state by changing the gene expression. The transcriptional syndicate essentially drives such alteration of gene expression. Transcription factors (TF) are the key regulatory protein that controls the expression of their target gene by binding to the cis-regulatory elements present in the promoter region. Myb-TF, ubiquitously present in all eukaryotes belong to one of the largest TF family, and play a wide array of biological functions in plants, including anthocyanin biosynthesis, vasculature system, cell signalling, seed maturation and abiotc stress responses. The present performed isolation and molecular cloning of full length Myb TF from *Eleusine corocana*. The isolated full-length coding sequence has 1053 bp and 350 aa was submitted to NCBI (Accession number MT312253). The transcript level of EcMYB increases under different abiotic stress treatments including dehydration, salinity, and high-temperature stress. The promoter region of *EcMyb1* was found to be enriched in stress-responsive cis-regulatory elements such as DRE, HSE, ABRE etc. In phylogenetic analysis, *EcMyb1* appeared to have high homology with its monocot orthologs particularly *Sateria italica, Hordeum vulgare, Saccharum barberi and Oryza sativa*. The three-dimension protein structure was generated based on homology modeling and structural aspects were discussed. Further, *Insilco* analysis was conducted to explore the physiological properties, subcellular localization, potential posttranslational modification sites (phosphorylation and glycosylation sites), and molecular and biological function of the full-length protein. Overall, the expression profiling and Insilco analysis of *EcMyb1* strongly indicated its potential role in abiotic stress response in *Eleusine corocana*.

Keywords: Abiotic stress, cis-regulatory elements, finger millet , *In- silico* analysis, MYB-transcription factor, Transcriptional regulation

INTRODUCTION

Drought stress is one of the major challenges encountered by plants due to the dramatic increase in climate change. The sharp rise in population and the anthropogenic activities further aggravated the situation. Greenhouse emission rise has led to erratic precipitation, increase in arid land area, desertification and ultimately diminution of crop productivity (Yang *et al*. 2010). Due to their molecular plasticity, plants can adapt and survive these changes in their environment. The molecular plasticity of plants is driven by the cell signalling cascade that follows the stress sensing, signal perception and adequate response towards it. The signalling pathway is initiated by signalling molecules that detect the stress and relay the signalling pathway by activating secondary signalling molecules, which further amplify the signalling and trigger the stress response. Among different secondary signalling molecules, calcium, cAMP, ROS, NO3, phosphorylation cascade etc. play an important role. Moreover, stress signalling is largely governed by phytohormonal and transcriptional regulation. The Interplay of phytohormones holds a remarkable role in the activation of TF. Activated transcription factors regulate the stress-responsive gene expression by limiting stress-induced cellular damage. TF, being a master regulator of gene expression are gaining more attention/ importance in context of target of crop improvement. In Arabidopsis, about 6% (or more than 1500 genes) of total number of genes are assigned to encode TF's only. TF binds to the cis-regulatory elementsin the gene's promoter region and modulates its transcription. MYB (Myeloblastosis), is one of the largest TF family, found ubiquitously in eukaryotes with varying functional role in plants. Like all TF, MYB-TF is also characterized by a highly conserved DNA-binding domain. Majority of MYB proteins contain variable numbers of N-terminus conserved MYB repeats (R).

The MYB proteins categorization is carried out on the basis of repeat numbers present in sequences that may range from 1 to 4.A composition of 52 amino acids forms each repeats ultimately creates 3 α-helices.Of the three α-helices, the second and third helices in every repeat create a helix turn helix (HTH) domain (Ogata *et al*. 1996). Further, MYB protein family is categorized into four different subfamily1R, R2R3, 3R and 4RMYB proteins, respectively on the basis of MYB domain (Dubos C *et al*. 2010). Plants regulate several responses through MYB proteins associated with R2R3 MYB subfamily (Lippold *et al*. 2009, Segarra et al.2009). Large number of MYB transcription factors has been identified and have been reported to function in many important physiological and biochemical processes involving anthocyanin accumulation, cell cycle and cell development, metabolism, hormone biosynthesis and signal transduction as well as drought stress responses (Allan *et al*. 2008, Ambawat *et al*. 2013, Dubos *et al*. 2010). Previously we have reported a drought stress induced MYB transcription factor from *Eleusine coroca-na.* The expression level increases *EcMyb1* significantly enhanced with the progression of severity of drought stress (Salvi *et al.* 2012, Jadhav *et al.* 2018). Henceforth, it is essential to research drought-responsive gene Myb besides their protein structure and promoter. Finger millet, a member of Poaceae, is a hardy crop and shows tolerance towards abiotic stresses specifically drought (Gupta *et al*., 2017).Finger millet can survive in harsh environmental stress conditions due to their potent alleles, which show drought-resistant characteristics. Such crops will be of great significance in meeting challenges. Understanding the mechanisms involved in the response of plants to adverse environmental conditions will help to generate crops with high tolerance to these stresses (Sanchita *et al*., 2013). Therefore, isolate the full-length coding sequence of Myb gene from drought tolerant variety of *Eleusine coracana* (PRM 6107), sequencing and cloning. Also, *Insilco* analysis was conducted to annotatethe sequence-structure-function relationship using bioinformatics tools and transcript analysis through real-time PCR during abiotic stress like drought, heat and salt stress.

MATERIALS AND METHODS

Plant material and growth conditions

Seeds of *Eleusine coracana* (finger millet) genotype - 6107, drought tolerant genotype, were obtained from the Molecular Biology Department, GB Pant University. Seeds of finger millet were sown in pots filled with soil, peat moss and vermiculite in a 3:1:1 proportion in polyhouse and allowed to grow to the seedling stage. Further plants were subjected to drought stress by withholding watering for 11 days.

RNA Isolation and cDNA preparation

After 11 days of drought treatment leaf samples of finger millets were collected and total RNA was extracted using the RNA-Xpress™ Reagent (Himedia) as described in the manufacturer's instructions. Using RNA as template, the first-strand cDNA synthesis was performed with oligo dT primers and avian myeloblastosis virus reverse transcriptase enzyme using revert first strand cDNA synthesis kit (Thermo Scientific India Pvt. Ltd., Mumbai) according to manual instructions.

PCR Amplification, Cloning and sequencing of *EcMyb1* cDNA

To the obtain full-length of *EcMyb1* cDNA, two primers,

5′TCAACTAATG GTAGCCCTTCC CTCT-3) ′sense) and 5 -′GATATT CTCAAAA GACAGTTGC ATTCT -3) ′ antisense) were designed based on the partial sequence of *EcMyb1* gene which has already been submitted with Genbank accession number JN107890.1 BLAST After attaining full length sequence of *EcMyb1* gene, the sequence was submitted to NCBI GenBank with accession number MT312253(Bhatt *et al.* 2021).

Differential expression analysis by quantitative real -time reverse transcription–PCR of *EcMyb1* gene for different Abiotic stresses Stress treatment

Finger millet genotype PRM 6107 was used for stress treatment. Seeds were surface sterilized and germinated in half-strength MS media. For germination, bottles were kept in dark for 2 days. After 2 days bottles were transferred to growth chamber under controlled conditions (light/dark regime of 18/6 h at 25℃, and light intensity of 200 µmol photons m–2 s–1. The seedlings were allowed to grow for 15 days. After 15 days seedlings were subjected to different abiotic stresses. For water stress seedlings were transferred to bottles containing MS media with three different concentration of PEG-6000. Three concentrations were 100ml of 5% (P1), 10% (P2), 15% (P3).

For salt stress, seedlings were allowed to grow on MS media with 100mM, 150mM and 200mM NaCl. Additionally, heat stress was also given to the seedling. For heat stress germination bottles were kept in an incubator at 35° C (H1), 40° C (H2) and 45℃ (H3) temperature. All these treatments were carried out with control (in water) for 6 hours.

Analysis of *EcMyb1* gene through Real-time PCR

After 6 hours total RNA was extracted from the treated and control leaves samples using the RNA-Xpress™ Reagent (Himedia) as described in the manufacturer's instructions and converted to cDNA using revert first strand cDNA synthesis kit (Thermo Scientific India Pvt. Ltd.) according to manual instructions. Real-time PCR (Realplex; Applied Biosystem) was performed in duplicates by applying SYBR green (PowerUp™ SYBR™ Green; Applied Biosystems, Thermo Fisher Scientific) using gene-specific primers. Cycle conditions were 94° C for 3 min, 40 cycle of denaturation at 94℃ for 45 min, 55° C for 30 sec and 72℃ for 40 min followed by a final extension of 7 min at 72℃. To check the relative fold in the expression of *EcMyb1* gene for all stress treatments compared to control were calculated through ΔΔCT method. Two different housekeeping genes actin and tubulin were used as endogenous control.

Statistical analysis

An independent sample t-test was applied for statistical analysis, and mean values represented the measurements. The t-test results showed a significant effect of abiotic stressors on the relative gene expression at $P<0.05$.

Sequence analysis of cloned DNA fragment using bioinformatics approach

For the sequence analysis, the cloned DNA fragment's nucleotide sequence was first converted to amino acid sequence using bioinformatics translate tool, ExPASy (Expert Protein Analysis System). This tool allows the translation of a nucleotide (DNA/RNA) sequence to an amino acid sequence (Gasteiger *et al*. 2005). Further the deduced amino acid sequences were subjected to protein blast (pBLAST) to find sequence homology with other plant sequences. The isoelectric point (pI), molecular weight, , total number of positive and negative residues, extinction coefficient, instability index, aliphatic index (AI) and grand average hydropathy (GRAVY) parameters were predicted by using protparam software (https://web.expasy.org/compute_pi/) (Bjellqvist *et al* 1993). The protein sequence of Myb domain was annotated by pfam (http://pfam.xfam.org/search/sequence) (Finn *et al.*2015). NetPhos2.0 and NetNGly 1.0 server was used for computing potential posttranslational modification sites. The domains in the functional region that are responsible for the activity of *EcMyb1* gene were identified using PROSITE (http://expasy.hcuge.ch/sprot/ prosite.html)(Sigrist *et al*. 2012). The Location of Motifs was identified using the MEME suite (Motif-based sequence analysis tools) (Bailey *et al.* 2011). Phylogenetic tree was constructed using software named Neighbor joining BIONJ program of MEGA version 7.0 which follows Maximum Likelihood method for evolutionary analysis (Kumar *et al.,* 2016). Secondary structure of *EcMyb1* protein, such as α-helix, β-sheet, and turn, were predicted using self-optimized prediction method with alignment (SOPMA) (Geourjon *et al.*1995). 3-D structures *EcMyb1* protein was characterized by Swiss model web server (https://swissmodel.expasy.org/), through homology template approach. These are bioinformatics tools method for predicting three-dimensional structure model of protein molecules from amino acid sequences (Yang *et al.,* 2015).

Analysis of promoter and cis-regulatory elements

Nucleotide sequence of cloned DNA fragment further BLAST with *Eleusine coracana* whole genome sequence (GenBank: LXGH01099917.1) using bioinformatics tool nucleotide blast (nBLAST). Upstream sequences in the vicinity of the gene were selected and subjected to promoter analysis. Web-based bioinformatics tools such as, PLACE (http://www.dna.affrc.go.jp/htdocs/PLACE/) (Higo *et al.* 1999), PlantCARE (http://bioinformatics.psb .ugent.be/webtools/ plantcare/html/(Lescot *et al.* 2002) and PlantPAN (http://plantpan2.itps.ncku.edu.tw/promoter.php

(Chow *et al.* 2015) have been used for the analysis of the cis-regulatory elements present in the upstream region of *EcMyb1* gene.

RESULTS AND DISCUSSION

Isolation and cloning of full-length gene encoding *EcMyb1*

Our previous study identified and isolated a partial gene (CDS) encoding MYB-TF from Eleusine coracana, which was induced after drought stress exposure. To further understand the molecular role and regulation of the *EcMyb1*, we sought to clone the gene's full length sequence.To clone the full-length gene, we isolated the RNA from the from *Eleusine coracana* leaves and reverse-transcribed it to cDNA. The cDNA was used to amplify the full-length gene amplicon of *EcMyb1* using gene specific- primers, and subsequently cloned in pGEM-T easy vector. The cloned amplicon was confirmed by sequencing.

Expression of *EcMyb1* gene in response to abiotic stresses in seedling stage

Drought stress

For imposing drought stress finger millet seedlings were grown in PEG medium at three different concentrations. In the present study PEG-6000 was used, molecular weight of more than 3000 cannot enter the cell wall space therefore, it does not show any harmful effects on plants growth and creates significant water stress **(**Meher *et al*., 2018). It was observed that as the concentration of PEG increases relative expression fold also increases as compared to control plant. The expression fold increases from 12 to 28 fold as compare to control (Fig 1).

Salt stress

In plants NaCl causes osmotic imbalance due to generation of reactive oxygen species. Present study indicated that as the salt stress increases the expression of *EcMyb1* gene also increases. In higher salt concentrations (200mM), the expression fold increases 29 fold as compared to minimum salt concentration (100mM) (Fig 1).

Heat stress

Heat stress causes detrimental effects on plant activities, including seed germination, growth development, photosynthesis and reproduction. Heat stress causes the direct accumulation of toxic compounds such as reactive oxygen species and it also disturbs cellular homeostasis. *EcMyb1* gene also shows enhanced expression during heat stress (Fig. 1). Relative expression of *EcMyb1* gene enhanced from 19 to 46 fold. This indicates that *EcMyb1* gene was induced in response to all stresses. In other words gene expression is induced with the onset of drought or other abiotic stresses.

Characterizing *EcMyb1* sequence by homology search

The extensively used bioinformatics tool for homology search is nBLAST for nucleotide sequence and pBLAST for protein sequence. The nBLAST and pBLAST tool were used to find out the similarity of *EcMyb1* gene with myb gene sequences of other members of the poaceae family.

It was observed that *EcMyb1* gene had 86.87 % sequence similarity with *Oryza sativa japonica* group myb16, 89.55% sequence similarity with *Sataria italica* transcription factor myb 61.Additionally, EcMYB 1 protein had 86.67% similarity with *Hordeum vulgare* transcription factor MYB86, 88.63% similarity with *Oryza sativa Japonica* Group transcription factor MYB61 and 86 % sequence similarity with *Triticum aestivum* MYB88 protein sequence (Table1). This result indicates that *EcMyb1* gene shows good similarity with Myb genes of different poaceae family members, which

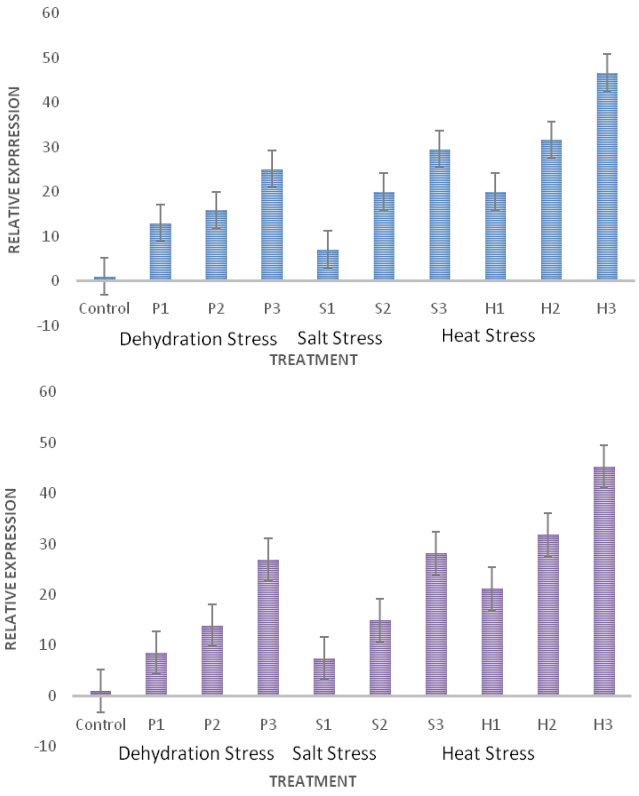

Fig. 1. *Quantitative real-time PCR analyses showing relative expressionof EcMyb1 gene under control and different abiotic stress conditions. P1-5% PEG, P2-10%PEG, P3-15%PEG, S1-100mM, S2- 150mM, S3-200mM NaCl solution and H1-35℃, H2-40℃ and H3-45° C temperature. Expression of EcMyb1 mRNA was normalized usingtwo endogenous controls, Tubulin (a) and actin (b) and calculated using the ΔΔCT method. Two replicates for each sample analyzed by real-time PCR.*

Table 1. Homology of *EcMyb1* gene with the other nucleotide and protein sequences using nBLAST and pBLAST tool

S. No.	Name	Accession number	Function	Query cover	E-value	% similarities	References
1	*Satariaitalica* transcription factor myb61	XM_004960378	Involved in regulation in stomatal movement	97%	0.0	89.55%	Liang, 2005
2	*Oryza sativa japonica* group myb16	AJ495784	Regulates cell morphogenesis	97%	0.0	86.87 %	Katiyar *et al.,* 2012
3	*Saccharumbarberi*myb 16	HF546403	Regulates cell morphogenesis	97%	0.0	89.36 %	Katiyar *et al.,* 2012
4	*Saccharumarundinaceum*myb 18	HF546406	control hypocotyl elongation responding to far-red light	97%	0.0	89.35 %	Zhang *et al.,* 2018
5	*Satariaviridis*myb transcription factor 61	XM_034728679	involved in regulation in stomatal movement	97%	0.0	89.35%	Baldoni, 2015
6	*Sorghum bicolor*myb 61	XR_002447626	involved in regulation in stomatal movement	97%	0.0	89.16 %	Baldoni, 2015
7	*Saccharumofficinarum*myb 18	HF546401	Plays an important role in drought stress by regulating membrane biosynthesis, antioxidant regulation and osmolyte synthesis	97%	0.0	88.58	Shingote, 2017
8	*Hordeumvulgare* transcription factor MYB86	KAE8806364	Mainly associated with lateral root growth regulation.	97%	0.0	86.67	Oh, 2011
9	*Oryza sativa Japonica* Group transcription factor MYB61	XP_025881490	involved in the regulation of stomatal movement	97%	0.0	88.63	Baldoni, 2015
10	*Triticum aestivum* MYB88	AFH08282	Genes associated with abiotic stress tolarence are positively regulated by MYB 88 TF.	97%	0.0	86.38	Xie *et al*., 2010

directly or indirectly contribute to the in generation of drought tolerance in plants (Jadhav *et al.* 2018).

Phylogenetic analysis

A phylogenetic tree using MEGA version 7.0, which is based on Neighbor joining BIONJ program was prepared with *EcMyb1* nucleotide and protein sequence as well as other similar sequences were obtained by nBLAST and pBLAST (Fig 2 a and b). The bootstrapping value shows the number of times the same branch was displayed during the repetition of phylogenetic reformation on data re-sampling for phylogenetic tree. For *EcMyb1* gene 1000 bootstrap was used. EcMyb1 nucleotide sequence was predicted to be adjacent to *Triticum aestivum* followed by *Sorghum bicolor, Oryza sativa,Zea mays* and *Sateria italica.* Further, EcMYB1 protein was nearest to *Sateria italica* preceded by *Zea mays*, *Oryza sativa*,*Sorghum bicolor* and others, as they show more recent common ancestors.

Motif analysis

Motifs play a functional or structural role that helps in recognizing new likely members of extant family or superfamily (Voet and Voet 2006).UsingExPASY PROSITE motif search program it was observed that in EcMYB1 protein HTH conserved motif is present. It is a major characteristic of MYB transcription factor protein (Fig 2a) (Peters *et al.* 1987; Biedenkapp *et al.* 1988). Comparison of the expression of Myb genes in *Arabidopsis* and *Oryza* by genome wide comparative analysis revealed that HTH domain plays a vital role in response to stress (Ogata *et al.,*1996). The MYB tran-

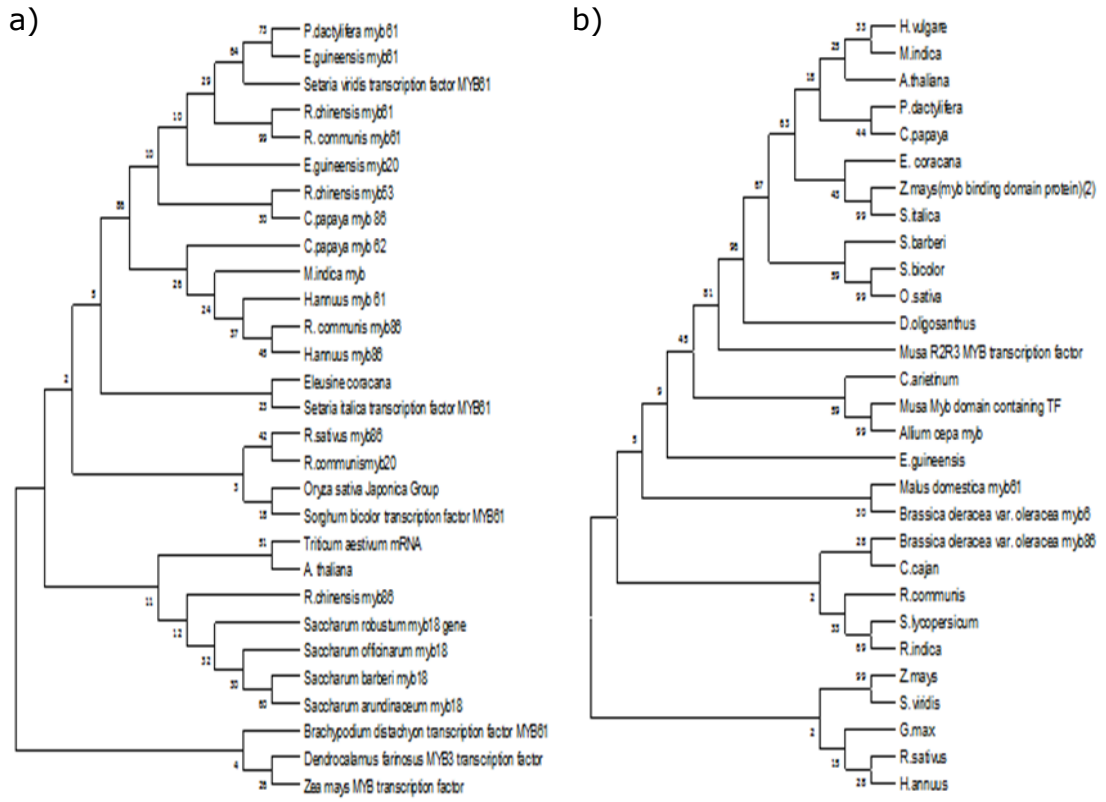

Fig. 2. ***a)*** *Phylogenetic tree of EcMyb1gene using Neighbor joining BIONJ program of MEGA version 7.0.with sequences obtained from nBLAST. Arabidopsis thaliana mybtranscription factor (TF),Carica papaya mybTF62, Carica papaya TFmyb86, Elaeisguineensismyb 20, Elaeisguineensismyb TF 61, Helianthus annuusTF myb61, Helianthus annuusmyb TF 86, SetariaitalicaTF MYB61, Zea mays MYB TF, Saccharumbarberimyb TF 18, Magniferaindicamyb TF, Phoenix dactyliferamyb TF 61, Ricinuscommunismyb TF 61, Ricinuscommunismyb TF 86, Rosa chinensismyb TF53, Rosa chinensismyb TF 61, Rosa chinensismyb TF86, Setariaviridis TF MYB61, Saccharumarundinaceummyb TF 18, Sorghum bicolor transcription factor MYB TF 61, Ricinuscommunismyb TF 20, Raphanussativusmyb TF 86Saccharumofficinarummyb TF18, Oryza sativa Japonica Group myb TF, Triticumaestivum mRNA, Brachypodiumdistachyon TF MYB61, Saccharumrobustummyb TF 18 gene, Dendrocalamusfarinosus MYB3 TF.* ***b)*** *Phylogenetic tree of putative EcMYB1 protein with different MYB Transcription factors of Dichantheliumoligosanthes, Zea mays, Oryza sativa Japonica Group, Hordeumvulgare, Arabidopsis thaliana, Sorghum bicolor, Allium cepa, Setariaitalica, Saccharumbarberi, Elaeisguineensis, Phoenix dactylifera, Setariaviridis, Carica papaya, Raphanus sativus, Rosa indica, Solanumlycopersicum, Helianthus annuus, Ricinus communis, Brassica oleracea var. oleraceamyb 6, Cicerarietinum, Brassica oleracea var. oleraceamyb 86, Malusdomestica myb61, Malusdomesticamyb 86, MusaMyb domain containing TF, Musa R2R3 type Myb TF, Glycine max, Cajanuscajan, Mangifera indica and Putative MYB DNA-binding domain superfamily protein Zeamayswas constructed using Neighbor joining BIONJ program of MEGA version 7.0.*

scription factors recognize and bind specific sequence of DNA by HTH domain (Fig 3a). Moreover, MEME suite (Motif-based sequence analysis) was used for identification of location of motif for EcMYB1 protein (Fig 3b). It was predicted that EcMYB1 protein shows a similar location of motifs as shown by other MYB transcription factor proteins of *Zea mays*, *Sateria italica* or *Saccharum barberi.*

Physiochemical properties of EcMYB1 protein

Various physiochemical properties of a protein such as molecular weight, isoelectrical point, total number of negatively and positively charged amino acids, atomic composition, extinction coefficients, estimated half life, instability index, aliphatic index and Grand average of hydropathicity (GRAVY) can be analyzed using Expasy -protpraram tool. Physiochemical characterization of EcMYB1 protein was carried out using this tool and results are presented in Table 2.

The molecular mass of EcMYB1 protein was found to be 38.26 KDa. The value of extinction coefficient for EcMYB1 protein range from 47440 M^{-1} cm^{-1}wavelength (where all Cys residues get reduced), to 47815 M^{-1} cm^{-1} (where all pairs of Cys residue are in cystines form) was reported by the software during analysis. The isoelectric point of EcMYB1 protein reported as 5.18 indicates that the protein is slightly acidicThe protein contains 26 positively charged amino acids (Arginine and Lysine) and 35 negatively charged amino acids (Aspartic acid + Glutamic acid). The result indicates

Table 2. Physiochemical characterization of EcMYB1 protein using protparam

S.No.	Physiochemical properties	Values
1	Molecular weight	38262.53 or 38.26 kDa
2	Isoelectrical point,	5.18
3	Total number of negatively charged amino acids of a protein,	(Asp + Glu) : 35
4	Total number of positively charged amino acids of a protein,	(Arg + Lys) : 26
4	Atomic composition,	$C_{1661}H_{2576}N_{464}O_{542}S_{17}$
5	Extinction coefficients,	$47815M^{-1}$ cm^{-1}
6	Estimated half life,	30 hours (mammalian reticulocytes, *in vitro*). >20 hours (yeast, *in vivo*). >10 hours *(Escherichia coli*, *in vivo*).
7	Instability index,	57.76
8	Aliphatic index	61.94
9	Grand average of hydropathicity (GRAVY).	-0.586

Table 3. Location of CpG islands in the promoter region of *EcMyb1* gene

Begin site	End site	Length	G+C Frequency	CpG o/e ratio
707	1266	554	0.5	0.67

Table 4. Location of Tandem repeats in the promoter region of *EcMyb1* gene

Location	Period size	Copy number	Consensus size	Percent match	Nucleotides A T C G	Entropy (0-2)
1-518	278	1.9	278	98	29 18 28 23	1.98

that the overall charge on EcMYB1 protein is negative. Using ProtParam tool half life for a protein can be predicted by observing the N-terminal amino acid sequence of a protein. This tool gives its prediction based on three human organisms (*in vitro*), yeast and *E.coli. (in vivo).* Half-life of 30hours in mammalian reticulocytes (*in vitro*), >20 hours for yeast (*in vivo*) and >10 hours for *Escherichia coli* (*in vivo*) was predicted for EcMYB1 protein. Instability index for a protein gives an approximate idea of its stability in a test tube. Theinstability index for EcMYB1 protein was predicted to be 57.76. The aliphatic index of EcMYB1 protein was computed to be 61.94. Characteristically, higher the aliphatic index, a protein's thermostability. For a protein or a peptide, the Grand Average of Hydropathy (GRAVY) is the total hydropathy of all amino acids which is divided by a total number of residues present in the sequence. A negative score of -0.586 was recorded for EcMYB1 protein, which predicts that protein is soluble and is an important characteristic of all transcription factors (Kyte *et al.* 1982; Katiyar *et al.* 2012).

Post Translational modification sites

When DNA gets transcribed into RNA and translated into protein,some chemical alteration may occur in the amino acid chain of proteins known as post translational modifications or PTMs. This modification includes covalent joining of specific chemical groups like phosphates, sulphates, etc. or lipid or carbohydrate moieties on proteins. (Burkle 2013; Gui *et al.*2019). Mostly phoshorylation occurs on serine (86.4%), threonine (11.8%) and is lowest in tyrosine (1.8%) (Ardito *et al*.2017).

Phosphorylation and protein-protein interactions can regulate transcription factorsby altering their mechanisms, such as DNA binding capacity, stability and interaction with other regulatory proteins (Kirchler *et al*. 2010). In EcMYB1 protein various phoshorylation sites were observed using NetPhos server (Fig.4a). Total 42 sites for serine, 21 for threonine, and 6 for tyrosine were computed. The sites for serine phosphorylation are mostly located at position 56-70, 225-250 and 290-300 in aminoacid sequence, while the sites for threonine and tyrosine phosphorylation could be found at positions 26,50, 60-65, 124,180 and 220-224, 236, 300 -305 respectively. It has been reported that Thr_{126} and Thr_{131} in Myb75of *Arabidopsis thaliana* are sites for MAP kinase phosphorylation. The phosphorylation event activates Myb75 which regulates stress response and secondary metabolism in *Arabidopsis* (Kreynes

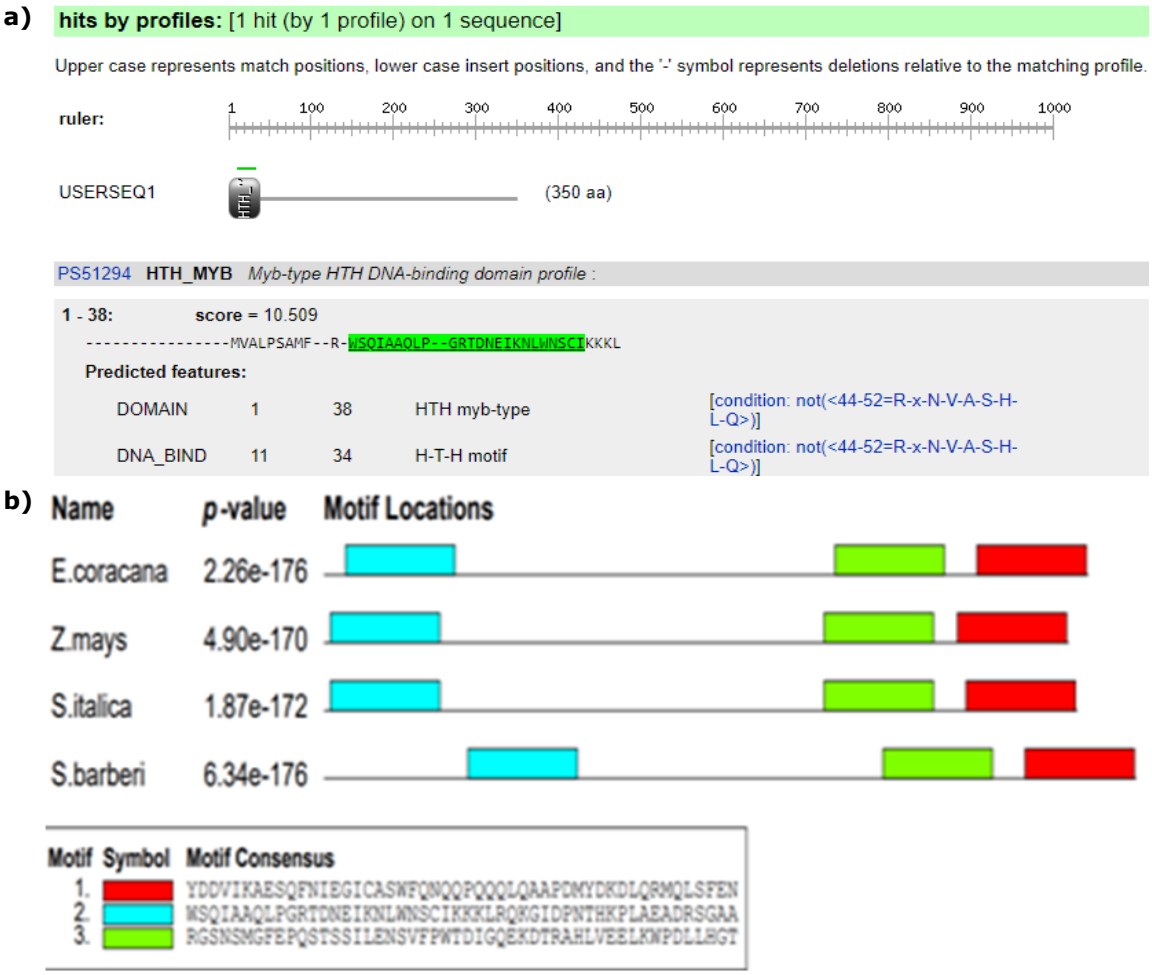

Fig. 3. ***a)*** *ExPASY PROSITEmotif search program showed the presence of Myb type HTH (Helix turn Helix) DNA binding domain in EcMYB1 protein.* ***b)*** *Location ofMotifs were identified using the MEME suite (Motif-based sequence analysis tools) observed that the translated sequences of MYB protein of E.coracana, Zea mays, Sateriaitalica and Saccharum-barbericonsisted of similar motifs location.*

2018, Morse *et.al.* 2009).Other types of posttranslation-al modifications in proteins include O-linked glycosyla-tion and N-linked glycosylation. (Steen 2008; Aebi 2013). In EcMYB1 protein, 20 O-linked glycosylation sites (Fig. 4b) and 4 potential sites for N- linked glyco-sylation (Fig. 4c) were predicted using NetNGlyc serv-er. It has been reported that O –linked glycosylation of transcription factor plays an important role in regulating RNA polymerase II activity (Jackson *et al*. 1988). It may be thought that o-linked glycosylation at different posi-tionsin EcMYB1 protein may help modulate the activity of RNA polymerase II involved in the transcription of drought-responsive genes.

Secondary Structure Annotation of EcMYB1 protein

When backbone atoms of a polypeptide chain interact with each other and form local folded structure, it leads to the formation of a secondary structure of a protein. To predict of EcMYB1 protein secondary structure SOPMA (self-optimized prediction method) was used (Geourjon *et al.,* 1995). It was observed that EcMYB1 protein contains 28% alpha helix, 5.43% extended strand, 2.57% beta-turn and 64% random coil (Fig. 5). This result suggests that EcMYB1 protein is not a com-pact globular protein but a protein containing alpha he-lix, extended strands and beta turns are turns held to-gether by random coils.

EcMYB 1 protein 3-D structure prediction

To understand a protein's function, protein it is neces-sary to collect knowledge about its tertiary or three di-mensional structure (3-D). The 3D structure of EcMYB1 protein was characterized by the online tool SWISS-MODEL, which is based on the homology modeling approach. SWISS-MODEL gives protein structure based on similarity in sequence between template and query. In EcMYB1 protein 3D structures Myb type helix turn helix domain is seen (Fig 6a). EcMYB1 protein shows 52.78 % sequence similarity with template se-quence which is R2R3 type MYB transcription factor. Ramachandran plot helps visualize energetically fa-voured protein structure regions structure (Fig 6 b). The

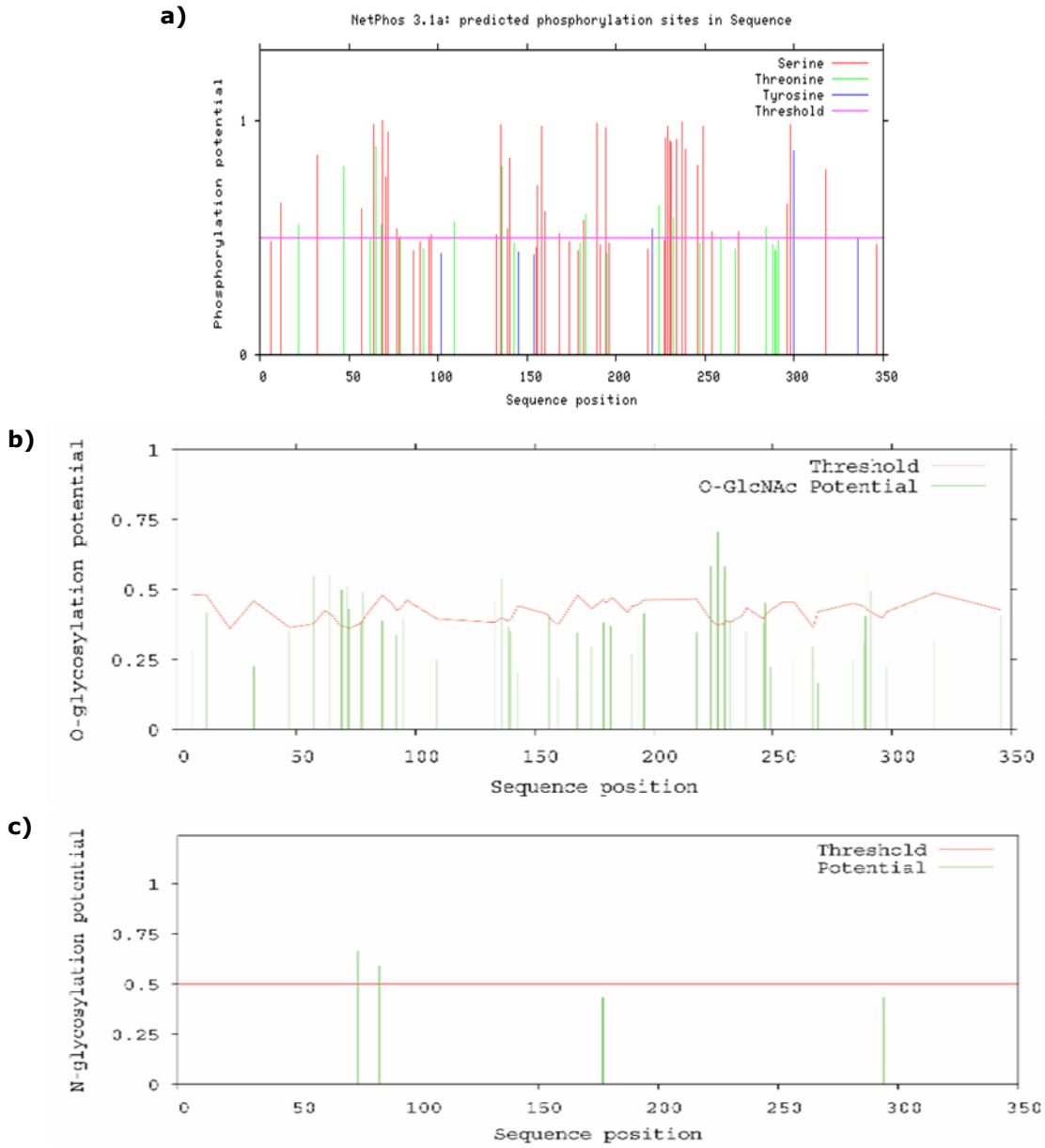

Fig. 4. *Prediction of post translational sites for EcMYB1 protein using NetPhos server and NetNGlyc 1.0 server (A) Phosphorylation sites (B) O-linked Glycosylation sites (C) N-linked Glycosylation sites*

feasibility of the secondary structure of EcMYB1 protein was predicted using SWISS MODEL Mol Probity tool. Many studies have reported that a good protein model will usually contain 90% of its amino acid residues in the favorable regions of a Ramachandran plot (Laskowski 1993; Pramanik *et al*. 2018; Chen *et al.* 2010).The molprobity score for EcMYB1 protein was 1.05. Mol Probity score, which should be as low as possible. The clash score generated due to overlapping of any two non-bonding atoms in the protein structure was zero. (Fig. 6).It can be concluded that EcMYB1 protein lies in an energetically favored region and, therefore could exist in nature.

The QMEAN Z-score gives a good estimate of the "degree of nativeness" of the structural features observed in the model on a global scale (Benkert *et al*.2011). An approximate QMEAN- Z score of zero stipulate the higher quality agreement between the experimental and modeled structure. However, the score of −4.0 or below shows a low-quality model. The QMEAN Z -score value for EcMYB 1 protein showed -0.38 (fig 6c) indicating that the proposed homology model has good reliability, shows adequate fit,and is acceptable.

Promoter analysis of *EcMyb1* gene

Promoter of a gene is located upstream of the coding regions of the gene and facilitates optimum expression of the gene by enabling the binding of various transcription factors with their *cis*- acting regulatory sequences (Biłas *et al*. 2016). The *cis*-regulatory sequences of the promoter need to be identified and studied to understand the expression of the gene. The present study

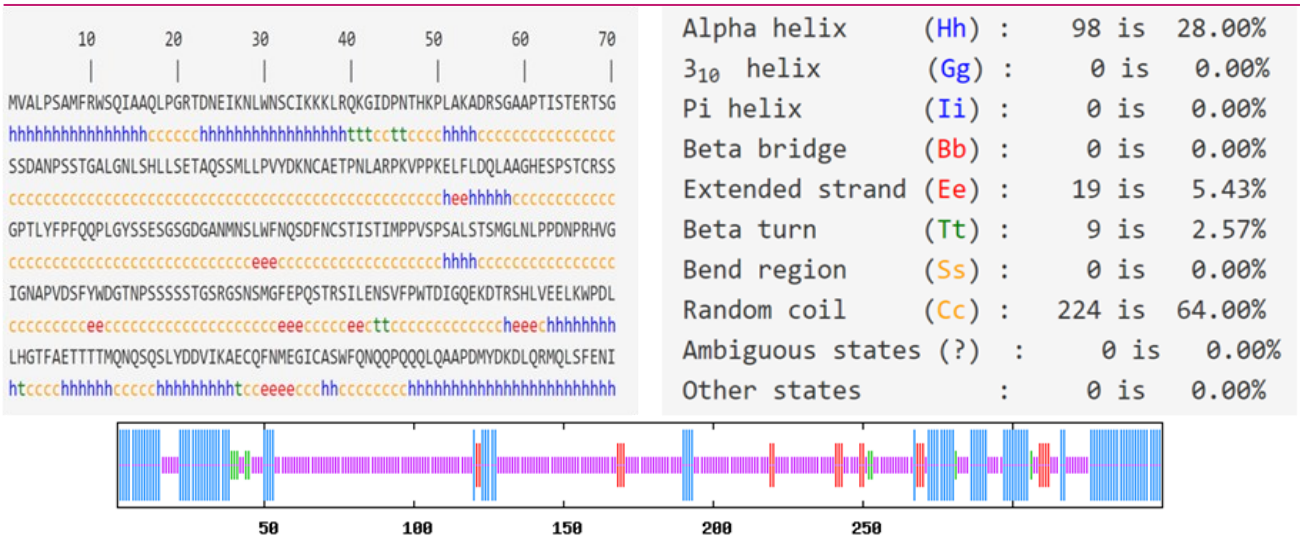

Fig.5. *Secondary structure of EcMYB1 protein computed using SOPMA web-based tool*

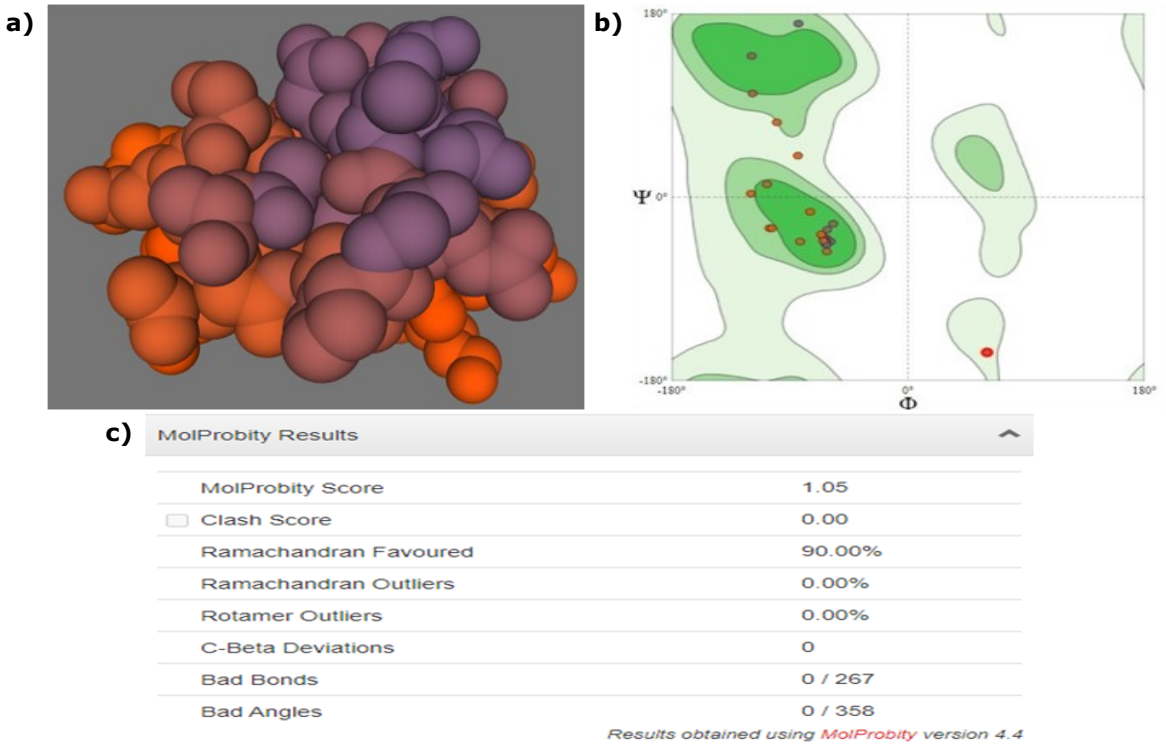

Fig. 6. ***(a)*** *3D structure of EcMYB1 protein by SWISS model* ***b)*** *Ramchandran plot of EcMYB1 protein indicating 90% residues in favored region* ***c)*** *Quality estimate of EcMYB1 protein based upon QMEAN value quality score*

determines and designs an expression cassette for exogenous gene expression.

Analysis of *cis*-regulatory elements

The upstream region of the*EcMyb1*sequence was selected and subjected to promoter analysis. Nearly 2 kb nucleotides were scanned using various to determine the promoter region's regulatory elements**.**

Analysis of CpG/CpNpG island and Tandem repeats by Plant PAN

Plant PAN is an important platform for plant promoters analysis and helps create transcriptional regulatory networks. CpG islands are non-methylated DNA sequences in plant genome are rich in G and C nucleotides. The characteristic feature of CpG island is that most of them are transcription initiation sites(Deaton *et al*. 2011; Elango *et al*. 2011). PlantPAN employs following criteria to characterize CpG islands:

(i) GC content should be above 50%

(ii) Length of CpG/CpNpG region should be greater than 200 bp

(iii) Ratio of observed to expected dinucleotide number should be above 0.6.

Table 3 shows the CpG/CpNpG analysis by the Plant-PAN tool that shows the presence of CpG/CpNpG is-

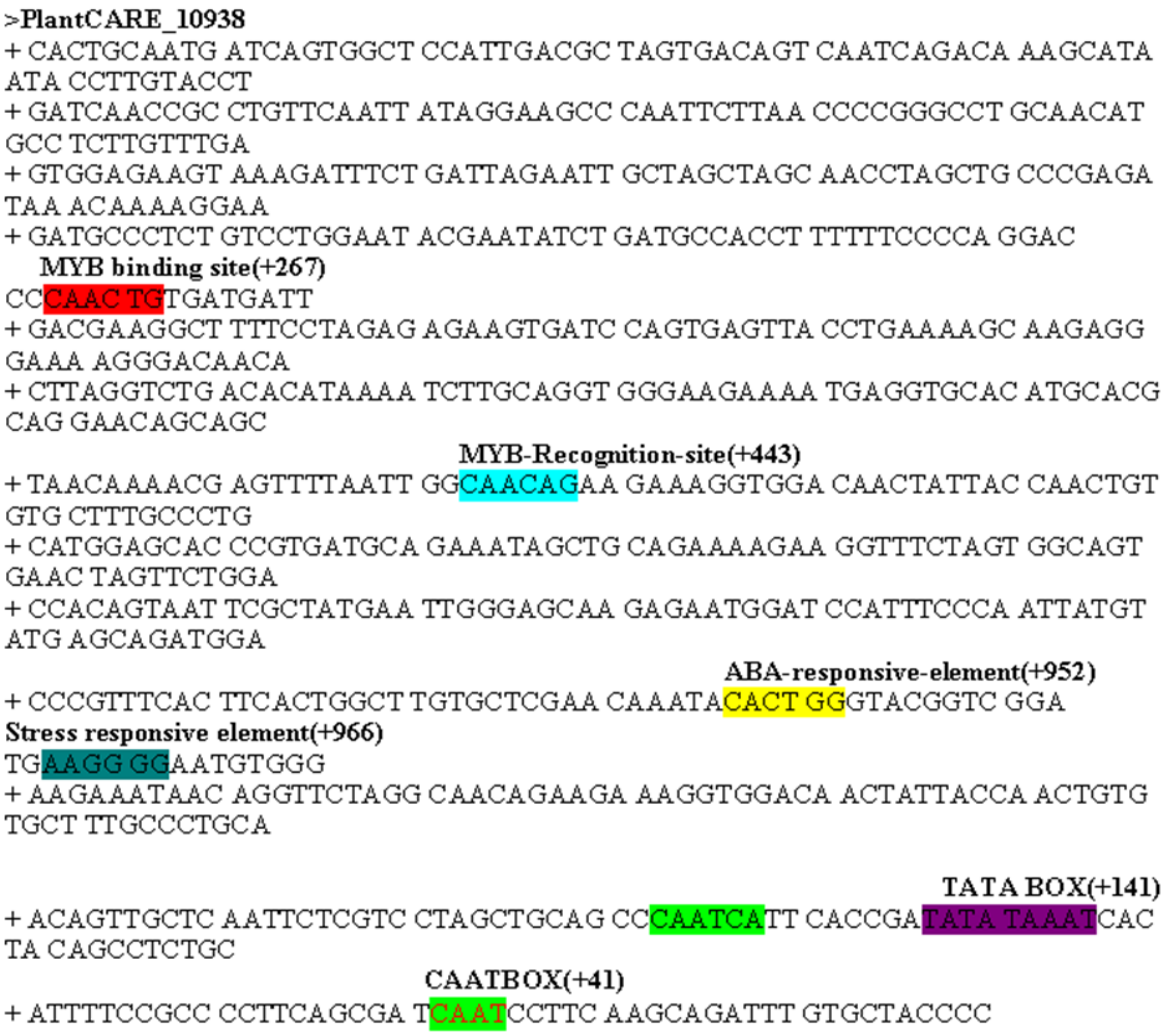

Fig. 7. *EcMyb1 gene promoter sequence analysis by PlantCARE*

lands in *EcMyb1* gene towards the 3' end in a distal promoter region.Presence of CpG island in the promoter region of *EcMyb1* gene indicates that the gene is transcriptionally active and the level of expression is not intermediate.

Tandem repeats occur in DNA when a pattern of one or more nucleotides is repeated and the repetitions are directly adjacent to each other(Mehrotra *et al*. 2014). Tandem repeat in the promoter region of a gene is associated with transcriptional regulation of the gene (Richard *et al.* 2009; Tian *et al*. 2017).

As represented by Table 4, the presence of DNA tandem repeats is shown in the upstream promoter region of *EcMyb1* gene. The repeat is 278 nucleotides in length and repeats twice in the promoter region. It may play a role in regulating the conditional expression of the *EcMyb1* gene.

Analysis of regulatory elements and ABA responsive elements

The transcription of a gene is controlled by the *cis* regulatory elements which are non-coding regions of the DNA. These elements function as molecular switches by being present in promoter sequence and controlling the transcriptional regulation (Banu *et al*. 2014). Further, the program PlantCare and PLACE was utilized to scan the promoter sequence upto 2 Kbp upstream from the translation commencement site of *EcMyb1* gene of *Eleusine coracana*. The length of *cis*-acting regulatory elements varied from 4- 10 bp in *EcMyb1* gene. This scanning led to the identification of various *cis*-acting elements such as ABA-responsive elements (CACT GG), stress-responsive elements (AAGG GG), Myb recognition sites (CAACAG), Myb binding sites (CAAC TG) etc.

Accumulation of abscisic acid (ABA) in plant cells is one of the rapid responses of drought stress (Hsiao, 1973), which modulates ABA inducible gene expression (Yamaguchi-Shinozaki *et al.* 2006) and stomatal closure for shrinking water loss due to transpiration (Schroeder *et al.* 2001**)**.Abscisic acid binding site in the promoter region of a gene indicates that the gene is expressed in response to the ABA signalling (Fernando *et al.* 2016). Promoter analysis of *EcMyb1* gene by

PlantCARE (Fig 7) revealed the presence of ABA-Responsive Elements (ABRE) in the promoter. The presence of ABRE in *EcMyb1* promoter indicates that some protein of ABA signalling pathway binds to this element and upregulates the expression of *EcMyb1* gene at drought conditions. This is supported by the study in which the expression of *EcMyb1* gene takes place not only in the drought-tolerant genotype but also in the drought-sensitive genotype of *Eleusine coracana* on exogenous application of abscisic acid(Kumari *et al.* 2017).

Conclusion

Abiotic stress tolerance is a multi-genic quantitative trait involving complex genetic control. None of the studieshave been successfully carried out that develop tolerant cultivars by targeting a single gene. Therefore,the present study holds high significance and adds to the literature on master regulators that modulate the expression of related genes. The master regulators are the transcription factors induced during stress conditions and bind specifically to the promoters of downstream genes associated with stress tolerance. They build strong associations with general transcription factors at the promoter of target gene and regulate their expression. The association of transcription factors developed in response to intracellular signals leads to activation or repression of target genes. In this study, we have cloned a full-length coding sequence of abiotic stress-inducible gene encoding Myb- transcription factor and submitted it to NCBI (Accession number MT312253). The expression of the gene increased manifolds during abiotic stress treatment, indicating it to be a stress-responsive gene. The stress- inducible expression of the gene was well correlated with the presence of different stress-responsive *cis*-regulatory elements in its promoter region. Expression profiling and Insilco analysis of *Ecmyb1*gene strongly indicated its role in abiotic stress response in *Eleusine corocana* and recommended it as a potential candidate gene for producing abiotic stress tolerance in plants.

Conflict of interest

The authors declare that they have no conflict of interest.

REFERENCES

1. Abe, H., Urao, T., Ito, T., Sek,i M., Shinozaki, K. & Yamaguchi-Shinozaki, K. (2003). Arabidopsis AtMYC2 (bHLH) and AtMYB2 (MYB) function as transcriptional activators in abscisic acid signalling. *Plant Cell*, 15, 63-78.doi: 10.1105/tpc.006130.
2. Abe, H., Yamaguchi-Shinozaki, K., Urao, T., Iwasaki, T., Hosokawa, D. & Shinozaki, K. (1997). Role of Arabidopsis MYC and MYB homologs in drought- and abscisic acid-regulated gene expression. *Plant Cell*, 9, 1859-1868.doi: 10.1105/tpc.9.10.1859.
3. Allan, AC, Hellens, RP & Laing, WA (2008).*MYB* transcription factors that colour our fruit. *Trends in Plant Science*, 13, 99–102. doi: 10.1016/j.tplants.2007.11.012.
4. Ambawat S, Sharma P, Yadav NR & Yadav RC (2013).*MYB* transcription factor genes as regulators for plant responses: an overview. *Physiology and Molecular Biology of Plants*, 19 ,3307-21.doi: 10.1007/s12298-013-0179-1.
5. Ardito F, Giuliani M, Perrone D, Troiano G& Muzio LL (2017). The crucial role of protein phosphorylation in cell signaling and its use as targeted therapy. *International Journal of Molecular Medicine*, 40**,** 271-280.doi: 10.3892/ijmm.2017.3036.
6. Bailey TL (2011). DREME: motif discovery in transcription factor ChIP-seq data. *Bioinformatics*, 27(12),1653-1659. doi: 10.1093/bioinformatics/btr261.
7. Baldoni, E., Genga, A. & Cominelli, E. (2015). Plant MYB transcription factors: their role in drought response mechanisms. *International Journal of Molecular Sciences,* 16 (7),15811-15851.doi: 10.3390/ijms160715811.
8. Banu SA, Huda KM, & Tuteja N. (2014). Isolation and functional characterization of the promoter of a DEAD-box helicase Psp68 using Agrpobacterium-mediated transient assay. *Plant Signalling & amp: Behavior,* 91, (1)8-9.doi: 10.4161/psb.28992.
9. Benkert P, Biasini M & Schwede T. (2011). Toward the estimation of the absolute quality of individualprotein structure models. *Bioinformatics*., 27(3) , 343-350.DOI: 10.1093/bioinformatics/btq662.
10. Bhatt, M., Salvi, P.& Lohani, P. (2021). Gateway cloning and in-planta transformation of drought stress responsive Ecmyb1 gene isolated from Eleusine coracana var.PRM 6107. *Environment Conservation Journal*, *22*(1&2), 205–211. https://doi.org/10.36953/ECJ.2021.221229.
11. Biedenkapp H, Borgmeyer U, Sippel AE & Klempnauer KH. (1988). Viralmyb oncogene encodes a sequence-specific DNA-binding activity. *Nature,* 335,835-837.doi: 10.1093/nar/29.17.3546.
12. BiłasR, Szafran K, Hnatuszko-Konka K &Kononowicz AK. (2016). Cis-regulatory elements used to control gene expression in plants. *Plant Cell, Tissue and Organ Culture (PCTOC)*, 127, 2269-87.
13. Bjellqvist B, Hughes GJ, Pasquali C, Paquet N, Ravier F & Sanchez JC. (1993). The focusing positions of polypeptides in immobilized pH gradients can be predicted from their amino acid sequences. *Electrophoresis*, 14, 1023-1031.doi: 10.1002/elps.11501401163.
14. Burkle A & Uversky VN. (2013). Posttranslational Modification. *Elsevier,* 3,425-430.doi: 10.1021/pr060394e.
15. Chen VB, Arendall WB, Headd JJ, Keedy DA, Immormino RM, Kapral GJ, Murray LW, Richardson JS& Richardson DC. (2010). MolProbity: all-atom structure validation for macromolecular crystallography. *ActaCrystallographica Section D: Biological Crystallography*, 66(1), 12-21.doi: 10.1107/S0907444909042073.
16. Chinnusamy V, Schumaker K & Zhu JK. (2004). Molecular genetic perspectives on cross-talk and specificity in abiotic stress signalling in plants. *Journal of Experimental Botany*, 55, 225-236.https://doi.org/10.1093/jxb/erh005.

17. Chow C, Zheng H, Wu N, Chien C, Huang H., Tzong-Yi& Lee T. (2015). PlantPAN 2.0: an update of plant promoter analysis navigator for reconstructing transcriptional regulatory networks in plants. *Nucl Acids Res*, 44,1154–1160.doi: 10.1093/nar/gkv1035.
18. Deaton AM & Bird A. (2011). CpG islands and the regulation of transcription. *Genes & Development*, 25,101010-22.doi: 10.1101/gad.2037511.
19. Dubos C, Stracke R, Grotewold E, Weisshaar B, Martin C & Lepiniec L. (2010). MYBtranscription factors in ArabidopsisTrends in Plant. *Science*,15,10573-81.doi: 10.1016/j.tplants.2010.06.005.
20. Elango N &Soojin VY. (2011). Functional relevance of CpG island length for regulation of gene expression.*Genetics,* 187,4 1077-83. doi: 10.1534/genetics.110.126094.
21. Fernando, V.D. & Schroeder, D.F., 2016.Role of ABA in Arabidopsis salt, drought and desiccation tolerance.In *Abiotic and biotic stress in plants-recent advances and future perspectives*.I*ntechOpen*, 507-524.
22. Finn RD, Coggill P, Eberhardt RY, Eddy SR., Mistry J& Mitchell AL. (2015). The Pfam protein families database: towards a more sustainable future. *Nucleic Acids Res*, 44, D279-D285.DOI: 10.1093/nar/gkv1344.
23. Geourjon C & Deleage G (1995) SOPMA: Significant improvements in proteinsecondary structure prediction by consensusprediction from multiple alignments. Oxford University Press 11, 681-684.doi: 10.1093/bioinformatics/11.6.681.
24. Gui J, Luo L, Zhong Y, Sun J, UmezawaT & Li L. (2019). Phosphorylation of LTF1, an MYB Transcription Factor in Populus, Acts as a Sensory Switch Regulating Lignin Biosynthesis in Wood Cells. *Molecular Plant*, 12, 1325–1337. DOI: 10.1016/j.molp.2019.05.008.
25. Gupta SM, Arora S, Mirza N, Pande A, Lata C, Puranik S, Kumar J & Kumar A. (2017). Finger Millet: A“Certain” Crop for an “Uncertain” Future and a Solution to Food Insecurity and Hidden Hunger under Stressful Environments. *Frontiers in plant science*, 8,643. https://doi.org/10.3389/fpls.2017.00643.
26. Higo K, Ugawa Y, Iwamoto M & KorenagaT.(1999). Plant cis-acting regulatory DNA elements (PLACE) database. *Nucl Acids Res,* 27,1297-300.doi: 10.1093/nar/27.1.297.
27. Hsiao, T.C., 1973. Plant responses to water stress. *Annual Review of Plant Physiology*. *24*(1),519-570.
28. Jackson SP & Tjian R. (1988). O-glycosylation of eukaryotic transcription factors: Implications for mechanisms of transcriptional regulation. *Cell*, 55,125-133. doi: 10.1016/0092-8674(88)90015-3.
29. JadhavP, Salvi P, Bhatt M & Lohani P. (2018). Expression ofEcMYB Transcription Factor Gene Under Different Abiotic Stress Conditions in *Eleusine coracana*. *International Journal of Agriculture,Environment and Biotechnology*, 11,5 799-806.doi: 10.30954/0974-1712.10.2018.12.
30. Katiyar A, Smita S, Lenka SK, Rajwanshi R, Chinnusamy V & Bansal K. (2012). Genome-wide classification and expression analysis of MYB transcription factor families in rice and Arabidopsis. *BMC genomics*, 13(1), 544.
31. Kirchler T, Briesemeister S, Singer M, Schutze K, Keinath M, Kohlbacher O, Carbajosa JV, Teige M, Harter K & Chaban C. (2010). The role of phosphorylatable serine residues in the DNA-binding domain of Arabidopsis bZIP transcription factors. *European Journal of Cell Biology,* 89,175–183.doi: 10.1016/j.ejcb.2009.11.023.
32. Kreynes AE. (2018). Phosphorylation of MYB75 Transcription Factor by MAP Kinases in *Arabidopsis thaliana*. PhD Thesis. The University of British Columbia.
33. Kumar S, StecherG&Tamura K. (2016). MEGA7: Molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Molecular BiologyEvolution*, 33,1870-1874.doi: 10.1093/molbev/msw054.
34. Kyte J, Russell F & Doolittle A. (1982). Simple Method for Displaying the Hydropathic Character of a Protein. *Journal of Molecular Biology*, 157, 105-132.doi: 10.1016/0022-2836(82)90515-0.
35. Laskowski RA, MacArthur MW, Moss DS & Thornton JM. (1993). PROCHECK: a program to check the stereochemical quality of protein structures. *Journal of applied crystallography*, 26(2), 283-291. https://doi.org/10.1107/S0021889892009944.
36. Les-cot M, Dhais P, Thijs G, Marchal K & Moreau Y. (2002). PlantCARE a database of plant cis-acting regulatory elements and a portal to tools for in silico analysis of promoter sequences. *NucleicAcids Res*, 30,1325-327.doi: 10.1093/nar/30.1.325.
37. Liang, YK., Dubos, C., Dodd, IC., Holroyd, GH., Hetherington, AM. & Campbell, MM. (2005). AtMYB61, an R2R3-MYB transcription factor controlling stomatal aperture in *Arabidopsis thaliana*. *Current Biology*, 15(13),1201-1206. doi: 10.1016/j.cub.2005.06.041.
38. Lippold F, Sanchez DH, Musialak M, Schlereth A, Scheible WR, Hincha DK & Udvardi MK. (2009). AtMyb41 regulates transcriptional and metabolic responses to osmotic stress in Arabidopsis. *Plant Physiol*, 149,1761-72. doi: 10.1104/pp.108.134874.
39. Mehrotra S, Goel S, Raina SN & Rajpal VR. (2014). Significance of satellite DNA revealed by conservation of a widespread repeat DNA sequence among angiosperms. *Applied Biochemistry and Biotechnology*, 173, 71790-801.doi: 10.1007/s12010-014-0966-3.
40. Ogata K, Kanei-Ishii C, Sasaki M, Hatanaka H, Nagadoi A, EnariM,Nakamura H, Nishimura Y, Ishii S & Sarai A. (1996). The cavity in the hydrophobic core of Myb DNA binding domain is reserved for DNA recognition and transactivation. *Nat StructBiol*, 3,178-818.doi: 10.1038/nsb0296-178.
41. Oh, J.E., Kwon, Y., Kim, J.H., Noh, H., Hong, S.W. & Lee, H. (2011). A dual role for MYB60 in stomatal regulation and root growth of Arabidopsis thaliana under drought stress. *Plant Molecular Biology*, 77(1-2),91-103. doi: 10.1007/s11103-011-9796-7.
42. Peters CWB, Sippel AE Vingron & Klempnauer (1987). MKHDrosophila and vertebrate myb proteins share two conserved regions, one of which functions as a DNA-binding domain. *EMBO J*, 6, 3085-3090.doi: 10.1002/j.1460-2075.1987.tb02616.x.
43. Pramanik K, Kundu S, Banerjee S, Ghosh PK & Maiti TK. (2018). Computational-based structural, functional and phylogenetic analysis of Enterobacterphytases. *3 Biotech*, 8(6), 262.doi: 10.1007/s13205-018-1287-y.
44. Richard, V., Brendolise, EC.&Chagné D., 2009. Multiple Repeats of a Promoter Segment Causes Transcription Factor Autoregulation in Red Apples.*The Plant Cell, 2*(1),168-183.
45. Salvi,P , Arora A & Lohani P. (2012). Expression of MYB Transcription Factor in Eleusinecoracana: towards making

of drought tolerant plants.LAP Lambert Academic Publishing USA.

46. Sanchita, Blessy BM & Sharma A. (2013). In silico analysis of putative transcription factor binding sites in differentially expressed genes: Study of the turnover of TFBSs under salt stress responsiveness in solanaceae family. *Plant Omics Journal*, 6,4278-285.
47. Schroeder JI, KwakJM & Allen GJ. (2001). Guard cell abscisic acid signalling and engineering drought hardiness in plants.*Nature*, 410, 6826327. doi: 10.1038/35066500.
48. Segarra G, Vander Ent. S, Trillas I & Pieterse CMJ (2009). MYB72, a node of convergence in induced systemic resistance triggered by a fungal and a bacterial beneficial Microbe. *Plant Biol*, 11,90-6. doi: 10.1111/j.1438-8677.2008.00162.x.
49. Shingote, P.R., Kawar, P.G., Pagariya, M.C., Rathod, P.R. & Kharte, S.B (2017). Ectopic Expression of SsMYB18, a Novel MYB Transcription Factor from *Saccharum spontaneum* Augments Salt and Cold Tolerance in Tobacco. *Sugar Tech*, 19(3), 270-282.
50. Sigrist CJ, De Castro E, Cerutti L, Cuche BA, Hulo N & Bridge A. (2012). New and continuing developments at PROSITE. *Nucleic Acids Res.,* 41,D344-D347.https://doi.org/10.1093/nar/gks1067.
51. Steen PVD, Rudd PM, Dwek RA & Opdenakker G. (2008). Concepts and Principles of O-Linked Glycosylation. *Critical Reviews in Biochemistry and Molecular Biology*, 33(3), 151-208.doi: 10.1080/10409239891204198.
52. Tian SL & Li Z. *(2017).*Analysis of tandem repeat units of the promoter of capsanthin/capsorubin synthase (Ccs) gene in pepper fruit.*PhysiolMolBiol Plants,*23, 3685–691. doi: 10.1007/s12298-017-0437-8.
53. Voet, D. & Voet, J.G. (2006). Protein Structural Motifs: Identification, Annotation and Use in Function Prediction. *Via and Tramontano*, 145-171.
54. Xie, Z., Li, D., Wang, L., Sack, F.D. & Grotewold, E. (2010). Role of the stomatal development regulators FLP/MYB88 in abiotic stress responses. *The Plant Journal*, 64 (5), 731-739. doi: 10.1111/j.1365-313X.2010.04364.x.
55. Yamaguchi-Shinozaki K & Shinozaki K. (2006). Transcriptional regulatory networks in cellular responses and tolerance to dehydration and cold stresses. *Annual Review of Plant Biology*, 57, 781–803.doi: 10.1146/annurev.arplant.57.032905.105444.
56. Yang J, Yan R, Roy A, Xu D, Poisson J & Zhang Y. (2015). The I-TASSER Suite: protein structure and function prediction. *Nature methods*, 12(1),7-8.doi: 10.1038/nmeth.3213.
57. Yang S, Vanderbeld B, Wan J &Huang Y. (2010). Narrowing down the targets: towards successful genetic engineering of drought-tolerant crops.*Mol Plant*, 3469–490.doi: 10.1093/mp/ssq016.
58. Zhang, T., Zhao, Y., Wang, Y., Liu, Z. &Gao, C. (2018). Comprehensive analysis of MYB gene family and their expressions under abiotic stresses and hormone treatments in *Tamarix hispida*. *Frontiers in Plant Science*, 9,1303. https://doi.org/10.3389/fpls.2018.01303