Molecular cloning, expression and Insilico 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.

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 abiotic stress responses. The present performed isolation and molecular cloning of full length Myb TF from Eleusine coracana. 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, Insilico analysis was conducted to explore the physiological properties, subcellular localization, potential posttranslational modification sites (phosphorylation and glycosylation sites), and molecular and functional properties of the full-length protein. Overall, the expression profiling and Insilico analysis of EcMyb1 strongly indicated its potential role in abiotic stress response in Eleusine coracana.

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

How to Cite
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 elements in 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 subfamily 1R, R2R3, 3R and 4R MYB 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 coracana. 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 annotate the 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 obtain full-length of EcMyb1 cDNA, two primers,
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°C, 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°C (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°C for 45 min, 55°C for 30 sec and 72°C for 40 min followed by a final extension of 7 min at 72°C. 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 (nBLAST) to find sequence homology with other plant sequences. The isoelectric point (pl), 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...
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). Using ExPASY 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 (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-

---

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

<table>
<thead>
<tr>
<th>S. No.</th>
<th>Name</th>
<th>Accession number</th>
<th>Function</th>
<th>Query cover</th>
<th>E-value</th>
<th>% similarities</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Sateria italica transcription factor myb61</td>
<td>XM_004960378</td>
<td>Involved in regulation in stomatal movement</td>
<td>97%</td>
<td>0.0</td>
<td>89.55%</td>
<td>Liang, 2005</td>
</tr>
<tr>
<td>2</td>
<td>Oryza sativa japonica group myb16</td>
<td>AJ495784</td>
<td>Regulates cell morphogenesis</td>
<td>97%</td>
<td>0.0</td>
<td>86.87%</td>
<td>Katliyar et al., 2012</td>
</tr>
<tr>
<td>3</td>
<td>Saccharumbarbermyb 16</td>
<td>HF546403</td>
<td>Regulates cell morphogenesis</td>
<td>97%</td>
<td>0.0</td>
<td>89.36%</td>
<td>Katliyar et al., 2012</td>
</tr>
<tr>
<td>4</td>
<td>Saccharumundinanecummyb 18</td>
<td>HF546406</td>
<td>control hypocotyl elongation responding to far-red light</td>
<td>97%</td>
<td>0.0</td>
<td>89.35%</td>
<td>Zhang et al., 2018</td>
</tr>
<tr>
<td>5</td>
<td>Sateriavirdismyb transcription factor 61</td>
<td>XM_034728679</td>
<td>involved in regulation in stomatal movement</td>
<td>97%</td>
<td>0.0</td>
<td>89.35%</td>
<td>Baldoni, 2015</td>
</tr>
<tr>
<td>6</td>
<td>Sorghum bicolormyb 61</td>
<td>XR_002447626</td>
<td>involved in regulation in stomatal movement</td>
<td>97%</td>
<td>0.0</td>
<td>89.16%</td>
<td>Baldoni, 2015</td>
</tr>
<tr>
<td>7</td>
<td>Saccharumofficinarummyb 18</td>
<td>HF546401</td>
<td>Plays an important role in drought stress by regulating membrane biosynthesis, antioxidant regulation and osmolyte synthesis</td>
<td>97%</td>
<td>0.0</td>
<td>88.58%</td>
<td>Shingote, 2017</td>
</tr>
<tr>
<td>8</td>
<td>Hordeum vulgare transcription factor MYB6</td>
<td>KAE8806364</td>
<td>Mainly associated with lateral root growth regulation.</td>
<td>97%</td>
<td>0.0</td>
<td>86.67%</td>
<td>Oh, 2011</td>
</tr>
<tr>
<td>9</td>
<td>Oryza sativa Japonica Group transcription factor MYB6</td>
<td>XP_025881490</td>
<td>involved in the regulation of stomatal movement</td>
<td>97%</td>
<td>0.0</td>
<td>86.67%</td>
<td>Baldoni, 2015</td>
</tr>
<tr>
<td>10</td>
<td>Triticum aestivum MYB88</td>
<td>AFH08282</td>
<td>Genes associated with abiotic stress tolerance are positively regulated by MYB 88 TF.</td>
<td>97%</td>
<td>0.0</td>
<td>86.38%</td>
<td>Xie et al., 2010</td>
</tr>
</tbody>
</table>
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 -protparam tool. Physicochemical 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 as 5.18 indicates that the protein is slightly acidic. The protein contains 26 positively charged amino acids (Arginine and Lysine) and 35 negatively charged amino acids (Aspartic acid + Glutamic acid). The result indicates
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 Escherichia coli (in vivo). Half-life of 30 hours 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. The instability 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 Hydropathicity (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 phosphorylation 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 factors by altering their mechanisms, such as DNA binding capacity, stability and interaction with other regulatory proteins (Kirchner et al. 2010). In EcMYB1 protein various phosphorylation 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 Thr126 and Thr131 in Myb75 of Arabidopsis thaliana are sites for MAP kinase phosphorylation. The phosphorylation event activates Myb75 which regulates stress response and secondary metabolism in Arabidopsis (Kreynes et al. 2023).

### Table 2. Physiochemical characterization of EcMYB1 protein using protparam

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

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

<table>
<thead>
<tr>
<th>Begin site</th>
<th>End site</th>
<th>Length</th>
<th>G+C Frequency</th>
<th>CpG o/e ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>707</td>
<td>1266</td>
<td>554</td>
<td>0.5</td>
<td>0.67</td>
</tr>
</tbody>
</table>

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

<table>
<thead>
<tr>
<th>Location</th>
<th>Period size</th>
<th>Copy number</th>
<th>Consensus size</th>
<th>Percent match</th>
<th>Nucleotides</th>
<th>Entropy (0-2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-518</td>
<td>278</td>
<td>1.9</td>
<td>278</td>
<td>98</td>
<td>29 18 28 23</td>
<td>1.98</td>
</tr>
</tbody>
</table>
Other types of posttranslational modifications in proteins include O-linked glycosylation 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 glycosylation (Fig. 4c) were predicted using NetNGlyc server. 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 positions in 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 compact globular protein but a protein containing alpha helix, extended strands and beta turns are turns held together by random coils.

### EcMYB 1 protein 3-D structure prediction

To understand a protein's function, protein it is necessary to collect knowledge about its tertiary or three dimensional 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 sequence which is R2R3 type MYB transcription factor. Ramachandran plot helps visualize energetically favoured protein structure regions structure (Fig 6 b).
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 EcMYB1 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 (Bilas 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
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 PlantPAN 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 (CACGTG), stress-responsive elements (AAGGGG), Myb recognition sites (CAACAG), Myb binding sites (CAACGTG) 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 studies have 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 manifold 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 coracana 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


of drought tolerant plants. LAP Lambert Academic Publishing USA.


