Journal of Agricultural Science; Vol. 5, No. 8; 2013
ISSN 1916-9752 E-ISSN 1916-9760
Published by Canadian Center of Science and Education

Amplification and Molecular Characterization of DREB1A Transcription Factor Fragment From Finger Millet [(*Eleusine coracana* (L.) Gaertn]

I. C. Mohanty[1], P. S. Gangasagar[1] & S. N. Rath[2]

[1] Department of Agricultural Biotechnology, Orissa University of Agriculture & Technology, Bhubaneswar 751003, Odisha, India

[2] BIF Centre, Department of Bioinformatics, CPGS, Orissa University of Agriculture & Technology, Bhubaneswar 751003, Odisha, India

Correspondence: I. C. Mohanty, Department of Agricultural Biotechnology, College of Agriculture, Orissa University of Agriculture & Technology, Bhubaneswar 751003, Odisha, India. Tel: 91-933-7021-274, E-mail: icmohanty1@rediffmal.com

Received: March 6, 2013 Accepted: June 13, 2013 Online Published: July 15, 2013

doi:10.5539/jas.v5n8p37 URL: http://dx.doi.org/10.5539/jas.v5n8p37

Abstract

Studies have shown that several plant species posses DREB1A and DREB2A (Dehydration-Responsive Element Binding Protein) orthologs. DREB transcription factors, also called C-repeat binding factors (CBFs), are the transacting elements/ transcription factors first identified in Arabidopsis which bind to low-temperature and dehydration responsive element (LTRE/DRE) found in several dehydrin (*Dhn*) promoters as well as in promoters of other cold and drought responsive genes and involved in dehydration-, cold-, and salinity-regulated gene expression. In this study, a fragment of *DREB1A* ortholog named *EcDREB1A* has been amplified from finger millet (*Eleusine coracana*), an important drought-tolerant grain crop with a rich genetic diversity grown in semi-arid tropics. In the current study, a systematic approach has been taken to predict a theoretical primer for *EcDREB1A* based on the cloned *Arabidopsis thaliana AtDREB1A* and other DREB genes for orthologous gene identification from finger millet. Sixteen different but related nucleotide sequences based on *AtDREB1A* gene were retrieved from different databases. A highly conserved region of 287 bp was detected on multiple sequence alignments through claustalw2 program and a set of primers (forward and reverse) was predicted using Primer3plus and Net software on the basis of this conserved region, assuming ideal conditions for primer length, GC content, formation of primer-dimers, hairpin-loops etc. The amplified genomic fragment of *EcDREB1A* was found to be 536 bp long, with possible introns as per translational analysis. Longest detected ORF in the amplified *EcDREB1A* fragment encodes a putative protein of 84 amino acids rich in serine (13.10%) with a predicted molecular mass of 9.29 kDa. Multiple sequence alignment of this *EcDREB1A* fragment with other *DREB* genes revealed presence of 9 highly conserved amino acids. Allele mining of *EcDREB1A* gene fragment across selected 5 finger millet cultivars revealed no variations on nucleotide, probably due to narrow genetic base in the test materials. Identification of novel regulatory genes involved in abiotic stress tolerance and allele mining in a manner similar to this presented herein might lead to a better and quicker solution for improving stress tolerance in crop plants.

Keywords: multiple sequence alignment, abiotic stress, PCR, DNA sequencing.

1. Introduction

Sustainable agricultural productivity requires development of crop plants tolerant to several abiotic stresses and able to grow in adverse environments. Abiotic stresses such as drought, low or high temperature, and salinity have detrimental effects on plant growth and reduce crop yield drastically. The ever increasing pressure put on agricultural land by burgeoning human populations and various human activities like mining, intense agricultural activity to increase crop productivity, the rising land occupation and the heavy and growing industrial activity have resulted in land degradation, drought and salinity across large tracts of agricultural land. Worldwide, it has been estimated that approximately 70% of yield reduction is the direct result of abiotic stresses (Acquaah, 2007). Abiotic stresses are manifested primarily as osmotic stress, resulting in the disruption of homeostasis and ion

distribution in the cell (Zhu, 2001) and may cause denaturation of functional and structural proteins (Smirnoff, 1998). As a consequence, these diverse environmental stresses often activate similar cell signaling pathways (Knight, 2000; Shinozaki & Yamaguchi-Shinozaki, 1996; Zhu, 2001) and cellular responses, such as the production of stress proteins, up-regulation of antioxidants, and accumulation of compatible solutes (Zhu et al., 1997; Cushman & Bohnert, 2000). Most metabolic pathways in plants are subjected to transcriptional regulation, which in turn are often regulated through developmental, environmental, or hormonal processes. Transcription factors modulate the expression of specific groups of genes through sequence-specific DNA binding and protein-protein interaction. They can act as activators or repressors of gene expression, leading to specific effect of genes in terms of cellular response (Latchman, 2003). DREB transcription factors (Dehydration Responsive Element Binding proteins), also called C-repeat binding factors (CBFs), are the transacting elements/ transcription factors first identified in *Arabidopsis thaliana*; they bind to low-temperature and dehydration responsive element (LTRE/DRE) found in several dehydrin (*Dhn*) promoters as well as in promoters of other cold and drought responsive genes and involved in dehydration-, cold-, and salinity-regulated gene expression (Yamaguchi-Shinozaki & Shinozaki, 1994; Stockinger et al., 1997; Liu et al., 1998). DREB transcription factors (genes of *DREB* family) have been grouped into DREBl/CBF and DREB2. DREB1 includes 3 additional novel genes and comprises of DREB1A (CBF3), DREB1B (CBF1), DREB1C (CBF2), and DREB1D (CBF4). DREB2 includes 6 genes (Sakuma et al., 2002). DREB1A and DREB1B share 86% identity; DREB1B and DREB1C have 86% identity; DREB1A and DREB1C, have 87% identity (Liu, et al., 1998). DREB1D (CBF4) gene expression, however, is up-regulated by drought stress , but not by low temperature (Haake et al., 2002).

The DNA sequence information obtained from the model plant *A. thaliana* and several other species has been used towards the identification, structural and functional characterization of genes. In particular, gene sequences from the model plant *A. thaliana* offer a unique opportunity to clone cognate genes and genes from heterologous sources, resulting in discovery of new genes, allele-mining, and large-scale SNP genotyping (Yong-Li et al., 2002; Kirankumar et al., 2001). The Polymerase Chain Reaction (PCR) has proven to be a versatile tool in molecular biology. The use of this technique has generated unprecedented advances in gene discovery and gene expression analysis (Katherine et al., 2008). Optimal primer sequence and appropriate primer concentration are essential for efficiency of the PCR, hence, there is a need of systematic procedure to guide the primer design approach (Binas, 2000).

Finger millet (*Eleusine coracana* (L.) Gaertn; Ragi) is an annual plant widely grown as millet in the arid areas of Africa and Asia. As one of the drought-tolerant grain crops with a rich genetic diversity finger millet poses an excellent model crop of choice for studying the genetic and physiological mechanisms of drought tolerance. It was proposed to isolate dehydration responsive element binding factor from this plant as an variant to be useful in allele mining.

2. Method

2.1 Plant Material

Five finger millet cultivars (1. Dibyasinha, 2. Suvra, 3. Neelachal, 4. Bhairabi, and 5. Chilka) were collected from the department of Plant Breeding & Genetics, College of Agriculture, Bhubaneswar, India for the present investigation.

2.2 Isolation of Genomic DNA

The genomic DNA was isolated from seedlings (3 wks old) of fingermillet cultivars by modified CTAB method (Sambrook et al., 1989). Extraction buffer (100 mM Tris-HCl (pH 8.0), 20 mM EDTA (pH 8.0), 1.4 M NaCl and 2% CTAB (w/v) was used for DNA extraction followed by chloroform: isoamyl alcohol extraction. DNA was precipitated by adding two volumes of cold isopropanol and was stored at -20°C overnight. It was then centrifuged for 10 min at 10,000 rpm. The supernatant was decanted and DNA pellet was washed twice with 70 per cent alcohol and air-dried. The genomic DNA was resuspended in 50µl of T10E1 buffer (Tris 10 mM and EDTA 1.0 mM). The quality and quantity of the DNA was checked spectrophotometrically and was taken for PCR reaction. The DNA was diluted to 1:100 (v/v) with T10E1 and subjected to spectrophotometer using T10E1 buffer as blank. Absorbance at 260 nm and 280 nm was recorded for checking the quality and contamination from protein and RNA. DNA concentration [µg/ml] of original sample was calculated by following the formula: Concentration of DNA [µg/ml] = OD at 260 × 50 × dilution factor.

2.3 In Silico Designing of EcDREB1A Primers

Nucleotide sequences of *DREB1A* were searched in NCBI (http://www.ncbi.nlm.nih.gov/) through Entrez for *A. thaliana* group. From the hits, 13 similar sequences were chosen including four mRNAs, five ESTs, three

complete CDS, and one clone DNA. Furthermore, three genomic DNA sequences were taken from EMBL-EBI database through European Nucleotide Archive (http://www.ebi.ac.uk/ena/) interface. The information related to those sequences, including accession number, sequence type, and sequence length were reported in Table 1. In total, 16 nucleotide sequences belonging to *A. thaliana DREB1A* group were subjected to Multiple Sequence Alignment (MSA) using Clustalw2 program (http://www.ebi.ac.uk/Tools/msa/clustalw2) of European Bioinformatics Institute, Cambridge, United Kingdom (EBI). The conserved region obtained from the MSA of 16 nucleotide sequences was used as input for designing primer using Primer3 Plus tool (http://www.bioinformatics.nl/cgi-bin/primer3plus/primer3plus.cgi). The parameters were adjusted as follows: Melting temperature: $T_{m\,min}$=52ºC, $T_{m\,max}$=58ºC ; Primer size: Min=18 bp, Max=24 bp; GC% contents: Min=50%, Max=60%. The predicted primers were subjected to check various properties; namely, hairpin loops, primer dimers, T_m and GC% using Premier Biosoft's NetPrimer tool (http://www.premierbiosoft.com/netprimer /index.html). Multiple sequence alignment was performed between the selected forward primers along with the 16 nucleotide sequences of *DREB1A* gene homologs belonging to *A. thaliana* group, using ClustalX 2.1 (Thompson et al., 1997; http://www.clustal.org). Again, those 16 nucleotide sequences were subjected for designing reverse primer using GeneFisher2 (http://bibiserv.techfak.unibielefeld.de/genefisher2/submission.html) followed by selection of the most suitable primer using Premier Biosoft's NetPrimer tool. The parameters for primer prediction were adjusted as mentioned above. Multiple sequence alignment was performed between the selected reverse primers along with the 16 nucleotide sequences of *DREB1A* gene homologs belonging to *A. thaliana* group using ClustalX 2.1. The specificity of both forward and reverse primers was checked using Blast program of NCBI database.

2.4 Orthologous Gene Amplification by PCR

Homology-based gene amplification was carried out by PCR using a set of gene-specific primers synthesized by Chromous Biotech, Bangalore. PCR amplification of *EcDREB1A* was optimized by standardizing the annealing temperature and primer concentrations using PCR (Perkin Elmer, ABI system). DNA polymerase, assay buffer, and dNTPs were purchased from Geneaid Biotech Ltd., New Taipei City, 22180 Taiwan (www.geneaid.com). The optimal PCR conditions used for amplification of *EcDREB1A* comprised 45 cycles, one for initial denaturation at 94 ºC for 5 min and 44 cycles of denaturation at 94ºC for 1 min, annealing at 52ºC for 1 min, pre- extension at 72ºC for 2 min, with the final extension at 72ºC for 10 min. The amplification was carried out with gDNA-1µl (50 ng), Forward Primer-2.4 µM, Reverse Primer-2.4 µM, dNTPs-0.4 mM, 10X *Taq* DNA Polymerase Assay Buffer-2.5 µl, *Taq*DNA Polymerase enzyme (2 units/µl)-1.0 µl and water added to make up the total reaction volume: 25 µl.

Table 1. Information regarding 16 *DREB1A* nucleotide sequences belonging to *A. thaliana* group, datamined in NCBI/ Entrez

Sl. No.	Accession No.	Sequence Type	Sequence Length
1	NM_118680.1	mRNA	908
2	AB007787.1	mRNA	908
3	AF074602.1	mRNA	902
4	BT024594.1	mRNA	651
5	CB259241.1	ESTs	613
6	CK119063.1	ESTs	748
7	DR750881.1	ESTs	757
8	DR357280.1	ESTs	425
9	DR357281.1	ESTs	493
10	AY691904.1	Complete cds	805
11	DQ372533.1	Complete cds	651
12	FJ169300.1	Complete cds	651
13	AM992886.1	Clone DNA	651
14	BAA33434.1	Genomic DNA	651
15	AAU93686	Genomic DNA	651
16	AEE85065	Genomic DNA	651

2.5 Electrophoresis and Elution of the Amplified EcDREB1A Fragment

After completion of the thermoprofile, electrophoresis was carried out to separate the amplified products in 1.5% agarose (Himedia Laboratories Pvt. Ltd., Mumbai, India) gel along with 100 bp DNA ladder (Chromous Biotech Pvt. Ltd, Bangalore, India). Large-scale amplification was performed towards elution of the PCR fragment. The single bright amplicon of appx. 0.5 Kbp was eluted using Gel/PCR DNA fragments extraction Kit (Geneaid) following the published protocol (Vogelsteinand & Gillespie, 1979).

2.6 Sequencing of Putative EcDREB1A Gene

The PCR-amplified DNA fragment of ~0.5 Kbp was sequenced by Chromous Biotech, Bangalore, India, following Applied Biosystems fluorescence-based cycle sequencing system (ABI 3500 XL Genetic Analyzer), which is an extension and refinement of Sanger's dideoxy sequencing (Sanger et al., 1974). The sequencing mix composition of 10 µl total contained: 4 µl Ready Reaction Mix, 1 µl (10 ng) Template, 2 µl (10 pmol) Primer, and 3 µl MilliQ Water; PCR Conditions used (Initial Denaturation: 96 ºC for 1 min, and 24 cycles of Denaturation: 96ºC for 10 sec, Hybridization: 50ºC for 5 sec, Elongation: 60ºC for 4 min) were followed for sequencing, using the Big Dye Terminator version 3.1" Cycle sequencing kit and Polymer & Capillary Array (POP_7 polymer 50 cm Capillary Array, Applied Biosystems, 850 Lincoln Centre Drive, Foster City, CA, USA, 94404) and Applied Biosystem Micro Amp Optical 96-Well Reaction plate (Applied Biosystems, USA). Data analysis was carried out using SeqScape_ v 5.2 software (Applied Biosystem) following BDTv3-KB-Denovo_v 5.2 analysis protocol (Applied Biosystem).

2.7 Structural and Phylogenetic Analysis EcDREB1A Genomic Sequence Fragment

The PCR amplicon sequence was subjected for structural analysis. Homology search was done using BLAST algorithm available at http://www.ncbi.nim.nih.gov through Entrez for *A. thaliana* group, *Oryza sativa* Indica group, *Capsella bursa-pastoris*, *Medicago truncatula*, Rosa hybrid cultivar, *Sorghum bicolor*, and *Nicotiana tabacum*. From the hits, nine sequences including four mRNAs, two complete CDS, one clone DNA, and one genomic DNA sequence were taken from EMBL-EBI database through European Nucleotide Archive (http://www.ebi.ac.uk/ena/) interface and were subjected to MSA using Clustalw2 program (http://www.ebi.ac.uk/Tools/msa /clustalw2) of European Bioinformatics Institute (EBI, UK). The alignment parameters were kept default. Multiple sequence alignments were generated with ClustalW at the EBI ClustalW server (http://www.ebi.ac.uk/clustalw/) using default parameters (Thompson et al. 1994). The alignments were visualized using the BioEdit program (version 5.0.9 http://www.mbio.ncsu.edu/BioEdit/bioedit.html). MEGA 3.1 software was used for phylogenetic analysis [version 3.1; http://www.megasoftware.net] (Kumar et al., 2004). The phylogenetic tree was constructed using the neighbor-joining method with bootstrap test.

In silico translation of the obtained genomic DNA sequence of *EcDREB1A* was done using GENSCAN software (genes.mit.edu/GENSCANinfo.html). General features of the predicted protein (molecular weight, Iso-electric point-pI, amino acid composition) were assessed using the Prot Param tool (http://expasy.org/cgi-bin/protparam). All other bioinformatics analyses, such as homology searches etc., were performed using tools that are accessible via different links on the proteomics service site of the Swiss Institute of Bioinformatics (http://www.isb-sib.ch/forms.htm).

3. Results

3.1 In Silico Design of DREB1A Primers

A highly conserved region (100% identity) of 287 bp (Figure 1) obtained through multiple sequence alignment was used for designing of *DREB1A* gene-specific primers. The Primer3Plus program predicted five forward and two reverse primers as per the parameters chosen (Table 2a). Out of these, one forward (F1) primer of 19 bp, 52.63% GC content, T_m of 54.21ºC, and absent of secondary structure was chosen. Comparatively, of the two reverse primers, one (R1) primer was found forming a dimer and other (R2) primer was found to have runs of C's and A's in its sequence. Therefore, reverse primer was not finalized. GeneFisher2-primer designing tool resulted both forward and reverse primers. As forward primer was already selected, only reverse primers were taken into consideration. One backward primer R1 of 18 bp, 50% GC content, T_m of 55.18ºC, and absent of secondary structure was chosen (Table 2b). The primer specificity was checked for both F1 and R1 primer by Blast program, and has shown significant alignment with *AtDREB1A* gene.

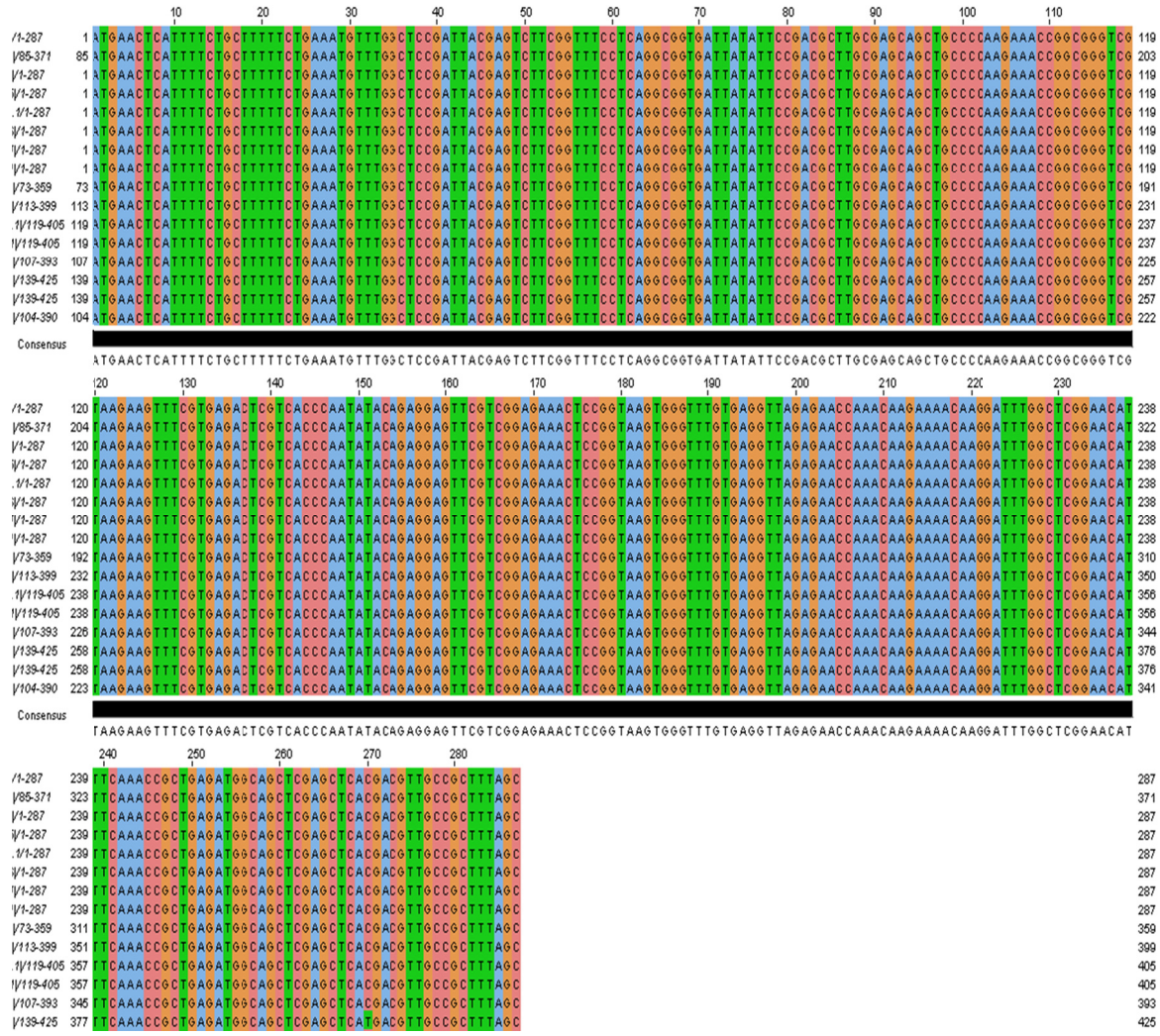

Figure 1. Jalview (www.jalview.org) of Clustalw2 alignment showing consensus sequence of highly conserved (100% identity) domain of 287 bp of *DREB1A* gene sequences from the *A. thaliana* group datamined with NCBI/Entrez

Table 2a. List of primers (Forward & Reverse) predicted by Primer3Plus

Primer Name	Sequence(5'→3')	Length [bp]	Hairpin	Dimer	T_m [°C]	GC%
F1	**GGCTCCGATTACGAGTCTT**	**19**	**0**	**0**	**54.21**	**52.63**
F1	**GGATATATTCCGACGCTG**	**18**	**0**	**0**	**52.2**	**50**
F2	TACGAGTCTTCGGTTTCCTC	20	1	1	54.37	50
F3	TTTGGCTCCGATTACGAG	18	1	1	54.19	50
F4	TGGCTCCGATTACGAGTC	18	1	1	53.56	55.56
F5	GTTTGGCTCCGATTACGA	18	0	0	53.93	50
R1	ACCCACTTACCGGAGTTTCT	20	0	1	55.4	50
R2	ACCTCACAAACCCACTTACC	20	0	0	53.74	50

Table 2b. List of (Reverse) primers predicted by NetPrimer tool

Name of the Primer	Sequence(5'→3')	Length [bp]	Hairpin	Dimer	T_m [°C]	GC%
R1	**ATGTTCCGAGCCAAATCC**	**18**	**0**	**0**	**55.18**	**50**
R2	**GCTGATTGTGATTCCACG**	**18**	**0**	**0**	**52.0**	**50**
R3	GTTCCGAGCCAAATCCTTG	19	1	1	57.85	52.63

3.2 Orthologous Amplification of EcDREB1A Gene Fragment

The template DNA of finger millet cultivars (Section 2.1) produced an amplification of a bright band of ~0.5 kbp (Figure 2). Large-scale amplification of the gene fragment was achieved using the same primer set and was separated in a low melting agarose gel of 1.5%, eluted and purified using Geneaid gel elution kit (www.geneaid.com). The eluted fragment was subjected for DNA sequencing.

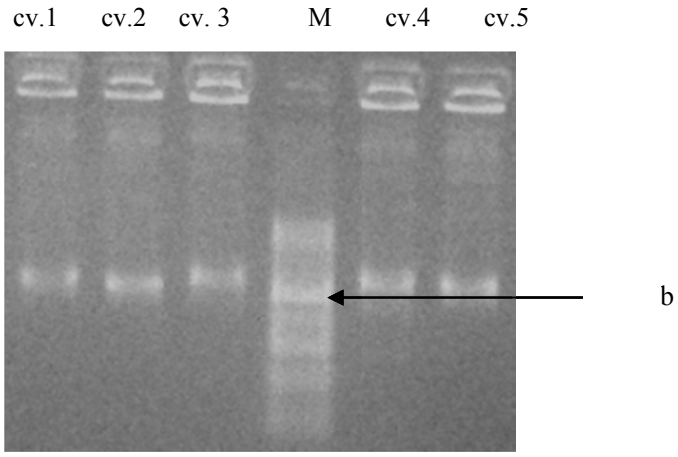

Figure 2. Example of gel electrophoresis of PCR–amplified *EcDREB1A* gene fragment. Shown are ~0.5 Kbp products obtained from 5 different cultivars (Sect.2.1) and DNA ladder marker (M). Position of marker's 500 bp band is indicated by the described arrow

3.3 DNA Sequencing

The amplified *EcDREB1A* gene fragments were sequenced with Applied Biosystems following modified Sanger's Dideoxy method at Chromous Biotech Pvt. Ltd, Bangalore, India and were found to be 536 bp long (Figure 3); it was same case for all the five cultivars of finger millet included in this investigation.

>1.Finger millet (*Eleucine coracana*, cv. Dibyasinha)

CCACATCACTAGGCTGTGATTCTGATCGGACCCCCGCGAGCATCTTCCCCAAGAGACCGG
CGGGGCGTAAGAAGTTACGTGAGACTCGTCGCGAAATTTACAATTGCGCCGCTCGCACAT
ACTACGATAGTGGTGCCTTAGACGTTAGAGACCCAAACACGAAGATGCTCGTTATGCTCG
GTTCCTTTGCAAATCGCTGAGACGACGGACTGGCGCTCAGGACGTTGCCACAAGAGCCTC
TTCGTGGCTGGTCAGCCCGACTCAGTTCAGCAGACTCGTCTTGAAGGCGCCGAATCCTGA
AGTCAACTTCAGATAACAACACTCAATAGGCGTAGGCTGAGTATGGGTAGGAGTGTTGTC
TTGGAATGTCTTAGGCGACGAAGGATCATTTTGATGACATCGAGACCGAAGACTCGTAAG
GGGAGCCACCCGGAGGATCAGAGTAAACTGCGCGTAATAAGCATGATTAGGCGATGGTAG
CGAGTCAGAGACTCCCCGCTTTTATGGCAGCAGGGATGCTTTCGCCGCTTCCGTCC

Figure 3. EcDREB1A gene fragment specific nucleotide sequence derived from amplified and sequenced DNA fragment

3.4 Structural and Phylogenetic Analysis of the Amplified EcDREB1A Gene Fragment

3.4.1 Structural Analysis of Nucleotide Sequence

Table 3. Nucleotide homology of EcDREB1A fragment with orthologous DREB1As datamined by nBLAST analysis of the amplified and sequenced DNA fragment

SlNo.	Accession No.	Sequence Type	Length	Query coverage	e-value(desc.order)	Identity
1	CP002687.1	*Arabidopsis thaliana* chromosome 4, complete sequence	18585056	98%	7e-52	70%
2	FJ169302	*Arabidopsis thaliana* ecotype Nd-1 DRE/CRT-binding factor 3 (CBF3/DREB1a) gene, complete coding DNA sequence	651	98%	7e-52	70%
3	FJ169301	*Arabidopsis thaliana* ecotype Co-1 DRE/CRT-binding factor 3 (CBF3/DREB1a) gene, complete cds	651	98%	7e-52	70%
4	FJ169300	*Arabidopsis thaliana* ecotype Di-G DRE/CRT-binding factor 3 (CBF3/DREB1a) gene, complete cds.	651	98%	7e-52	70%
5	FJ169298	*Arabidopsis thaliana* ecotype Lip-0 DRE/CRT -binding factor 3 (CBF3/DREB1a) gene, complete cds	651	98%	7e-52	70%
6	FJ169297	*Arabidopsis thaliana* ecotype Spr1-2 DRE/CRT-binding factor 3(CBF3/DREB1a) gene, complete cds.	651	98%	7e-52	70%
7	FJ169296	*Arabidopsis thaliana* ecotype Po-0 DRE/CRT-binding factor 3 (CBF3/ DREB1a) gene, complete cds.	651	98%	7e-52	70%
8	FJ169295	*Arabidopsis thaliana* ecotype Gie-0 DRE/CRT -binding factor 3 (CBF3/DREB1a) genes, complete cds.	651	98%	7e-52	70%
9	FJ169294	*Arabidopsis thaliana* ecotype Mt-0 DRE/CRT-binding factor 3 (CBF3/DREB1a) gene, complete cds.	651	98%	7e-52	70%
10	FJ169292	*Arabidopsis thaliana* ecotype Bor-1 DRE/RT-binding factor 3(CBF3 /DREB1a) gene, complete cds.	651	98%	7e-52	70%
11	FJ169291	*Arabidopsis thaliana* ecotype Ta-0 DRE/CRT-binding factor 3 (CBF3/DREB1a) genes, complete cds.	651	98%	7e-52	70%
12	FJ169286	*Arabidopsis thaliana* ecotype Ka-0 DRE/CRT-binding factor 3 (CBF3/DREB1a) genes, complete cds.	651	98%	7e-52	70%
13	FJ169285	*Arabidopsis thaliana* ecotype LP2-2 DRE/CRT-binding factor 3 CBF3/DREB1a) genes, complete cds.	651	98%	7e-52	70%
14	FJ169283	*Arabidopsis thaliana* ecotype Sf-1 DRE/CRT-binding factor 3 (CBF3/DREB1a) genes, complete cds.	651	98%	7e-52	70%
15	FJ169282	*Arabidopsis thaliana* ecotype Ll-0 DRE/CRT-binding factor 3 (CBF3/DREB1a) genes, complete cds.	651	98%	7e-52	70%
16	EF523126	*Arabidopsis thaliana* ecotype Tscha-1 C-repeat binding factor 3(CBF3) mRNA, complete cds	908	98%	7e-52	70%
17	EF523125	*Arabidopsis thaliana* ecotype Sapporo-0 C-repeat binding factor 3 (CBF3) mRNA, complete cds.	908	98%	7e-52	70%
18	EF523124	*Arabidopsis thaliana* ecotype Rsch-0 C-repeat binding factor 3(CBF3) mRNA, complete cds.	908	98%	7e-52	70%
19	EF523122	*Arabidopsis thaliana* ecotype Bur-0C-repeat binding factor 3 (CBF3)mRNA,complete cds	908	98%	7e-52	70%
20	EF523118	*Arabidopsis thaliana* ecotype Litva C-repeat binding factor 3 (CBF3) mRNA, complete cds	908	98%	7e-52	70%
21	EF523115	*Arabidopsis thaliana* ecotype Pog-0 C-repeat binding factor 3 (CBF3)mRNA, complete cds	908	98%	7e-52	70%
22	EF156749	*Capsella bursa-pastoris* DREB1A (DREB1A) gene, complete cds	651	98%	3e-50	70%
23	DQ403814	*Nicotiana tabacum* DREB1A mRNA, complete cds	651	98%	3e-50	70%
24	AY691904	*Arabidopsis thaliana* DREB1A gene, complete cds	805	98%	3e-50	70%
25	XM_002867561	*Arabidopsis lyrata* subsp. lyrata hypothetical protein, mRNA	827	62%	1e-24	70%
26	EF523196	*Arabidopsis lyrata* subsp. petraea C-repeat binding factor 3 (CBF3) mRNA, partial cds	609	69%	1e-24	84%
27	DQ131497	*Iris lactea* var. chinensis CRT/DRE binding factor 1 (CBF1) gene, complete cds.	664	49%	5e-22	72%
28	FJ491243	*Arabis pumila* CBF2 mRNA, complete cds	651	55%	2e-20	70%

Sequence information from the amplified fragment was subjected to homology search using nBLAST algorithm at http://www.ncbi.nlm.nih.gov. The amplified EcDREB1A fragment (536 bp) nucleotide sequence showed 98% query coverage and 70% identity with 651 bp CDSs reported in many of *A. thaliana* sequences (Table 3), with the lowest e-value of 7e-52. This amplified fragment of *EcDREB1A* nucleotide sequence also showed 70% similarity with *Capsella bursa-pastoris* (EF156749), *Nicotiana tabacum* DREB1A sequence (DQ403814) with e-value of 3e-50.The amplified and sequenced nucleotide sequence of partial *EcDREB1A* was submitted to EMBL-ENA which has been accepted with accession No. HF549034 in the said data base.

3.4.2 Structural Analysis of the Predicted *EcDREB1A* Protein Fragment

The obtained amplified genomic sequence was subsequently in silico translated using Expasy translate tool at http://expasy.org/cgi-bin/translate/. The amino acid sequence of *EcDREB1A* gene fragment is shown in Figure 4. The longest detected open reading frame (ORF) of 252 bp is coding for 84 amino acids. The translated *EcDREB1A* fragment encodes a putative protein of 84 amino acids rich in serine (13.10%) with a predicted molecular mass of 9.29 kDa (Figure 4). The amino acid sequence of the obtained *EcDREB1A* fragment showed 88% query coverage and 63% identity with the reported *GhDREB1A* (upland cotton, *Gossypium hirsutum*) DREB sequence, with an e-value of 7e-06. Other cereals, such as *Zea mays* DREBA (53%), *Oryza sativa* DREB1 (55%), *Triticum aestivum* DREB transcription factor 6 (55%), shared varied degrees of identity with the *EcDREB1A* fragment at protein level (Table 5). Multiple sequence alignment of *EcDREB1A* with other amino acid sequences revealed presence of 9 highly conserved amino acids (R, P, G, F, D, R, G, C and N) (Figure 5). However, *EcDREB1A* fragment is phylogenetically separated from other sequences taken for comparison here and forms a distinct orthodox taxonomic unit (OTU) (Figure 6).

acggaagcggcgaaagcatccctgctgccataaaagcggggagtctctgactcgctaccat
R K R R K H P C C H K S G E S L T R Y H
cgcctaatcatgcttattacgcgcagtttactctgatcctccgggtggctccccttacga
R L I M L I T R S L L - S S G W L P L R
gtcttcggtctcgatgtcatcaaaatgatccttcgtcgcctaagacattcccaagacaac
V F G L D V I K M I L R R L R H S Q D N
actcctacccatactcagcctacgcctattgagtgttgttatctgaagttgacttcagga
T P T H T Q P T P I E C C Y L K L T S G
ttcggcgccttcaagacgagtctgctggaactgagtcgggctgaccagccacgaagaggc
F G A F K T S L L E L S R A D Q P R R G
tcttgtggcaacgtcctgagcgccagtccgtcgtcctcagcgatttgcaaaggaaccgag
S C G N V L S A S P S S S A I C K G T E
cataacgagcatcttcgtgtttgggtctctaacgtctaaggcacccactatcgtagtatg
H N E H L R V W V S N V - G T H Y R S M
tgcgagcggcgcaattgtaaatttcgcgacgagtctcacgtaacttcttacgccccgccg
C E R R N C K F R D E S H V T S Y A P P
gtctcttggggaagatgctcgcgggggtccgatcagaatcacagcctagtgatgtpbact
V S W G R C S R G S D Q N H S L V M X T

Figure 4. *In silico* translated regions (red) of the amplified finger millet *DREB1A* fragment resultant from Expasy translate tool (http://expasy.org/cgi-bin/translate/)

Table 5. *In silico* translated protein homology of *EcDREB1A* fragment with orthologous *DREB1A*s

Sl No.	Accession No.	Sequence Type	Length	Query coverage	e-value (desc. order)	Identity
1	gb|ABD72616.1	cold resistance related AP2 transcription factor [*Arabidopsis thaliana*]	150	63%	2e-06	37%
2	gb|ABV82985.1|	cold responsive transcription factor [*Thlaspi arvense*]	214	63%	3e-06	35%
3	emb|CAQ52596.1|	DREB1A [*Arabidopsis thaliana*]	216	63%	4e-06	37%
4	gb|ABD65969.1|	DREB1A [*Nicotiana tabacum*]	216	99%	4e-06	37%
5	gb|AAR26658.1|	Cbcbf [*Capsella bursa-pastoris*]	219	63%	5e-06	36%
6	gb|AAP83936.3|	GhDREB1A (upland cotton) [*Gossypium hirsutum*].	216	88%	7e-06	63%
7	gb|AAS00621.1|	DREB1 [*Eutrema salsugineum*]	216	63%	9e-06	35%
8	gb|AEQ35295.1|	DREB1A-like protein [*Lepidium latifolium*].	213	26%	1e-05	59%
9	gb|ACB45087.1|	CRT binding factor 1 [*Solanum habrochaites*]	219	27%	3e-05	58%
10	dbj|BAE17131.1|	LhCBF1 [*Lycopersicon hirsutum*]	222	27%	3e-05	58%
11	gb|ABX00639.1|	DREB1B [*Brassica juncea*]	214	26%	4e-05	56%
12	ref|NP_001170481.1|	DRE-binding protein 4 [*Zea mays*]	232	29%	1e-04	53%
13	dbj|BAL04971.1|	CBF [*Zoysia japonica*]	212	25%	1e-04	55%
14	gb|AAR88363.1|	DREB-like protein 1 [*Capsicum annuum*]	215	24%	1e-04	57%
15	gb|AAY43213.1|	CRT/DRE binding factor 1 [*Hevea brasiliensis*]	231	36%	1e-04	52%
16	ref|XP_002436386.1|	hypothetical protein SORBIDRAFT_10g001620 [*Sorghum bicolor*]	249	39%	1e-04	47%
17	gb|AAM18961.1|	CBF-like protein CBF17 [*Brassica napus*]	250	24%	1e-04	62%
18	gb|AAX28952.1|	HvCBF5 [*Hordeum vulgare subsp. vulgare*]	214	25%	3e-04	55%
19	ref|NP_001056661.1|	Os06g0127100 [*Oryza sativa* Japonica Group]	214	34%	3e-04	55%
20	gb|ABK55356.1|	CBFII-5.2 [*Triticum aestivum*]	219	26%	3e-04	55%
21	gb|ABF59736.1|	CBF-like transcription factor [*Sabal palmetto*]	210	69%	0.002	48%
22	gb|ABF59744.1|	CBF-like transcription factor [*Dypsis lutescens*]	212	64%	0.002	48%

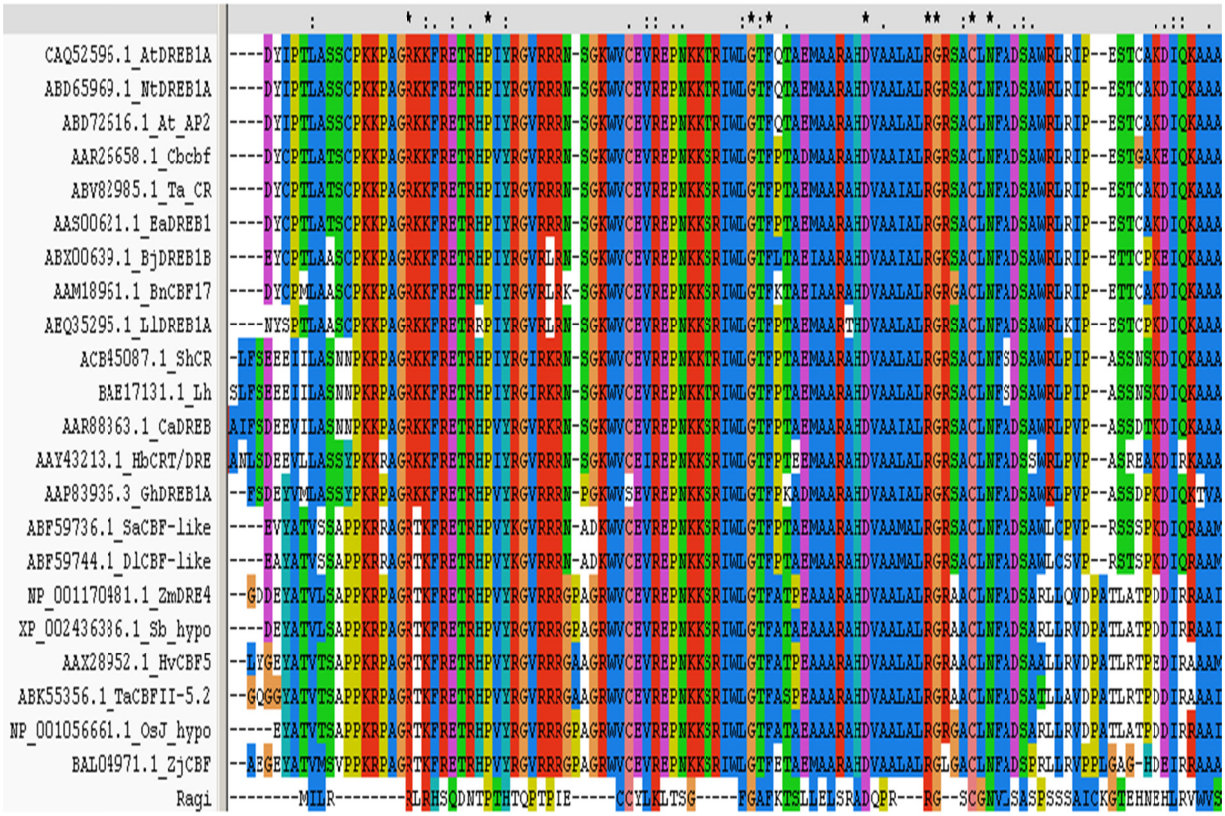

Figure 5. Multiple sequence alignment of 22 orthologous DREB proteins with the obtained EcDREB1A fragment in ClustalX 2.1

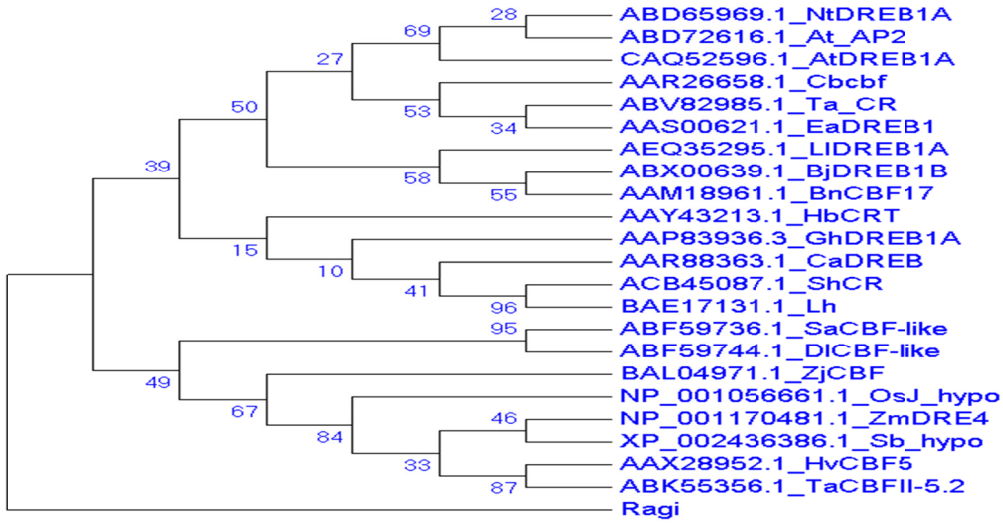

Figure 6. Consensus bootstrap (Maximum Parsimony) tree constructed in MEGA 5.5

4. Discussion

Physiological, biochemical, and genomic tools and techniques have advanced the understanding of plant responses and adaptations to various abiotic stresses like drought, cold, salinity etc. Engineering the stress proteins or the enzymes of the biosynthetic pathways associated with stress responses has been evolving as an encouraging method for improving the abiotic stress tolerance (McCue & Hanson, 1990; Bohnert & Jensen, 1996; Dixon & Arntzen, 1997; Barkla et al., 1999; Blumwald, 2000; Hong et al., 2000).

Dehydration-responsive element (DRE)/C-repeat (CRT) was first identified in *A. thaliana*, as a cis-acting element regulating gene expression in response to dehydration (drought, salinity, & cold stress; Baker et al., 1994; Yamaguchi-Shinozaki & Shinozaki, 1994; Carvallo et al., 2011). Several DRE-binding proteins (DREB)/CRT-binding factors (CBFs) were isolated and identified as key players in dehydration-responsive gene expression (Yamaguchi-Shinozaki & Shinozaki, 1994). *DREB1A* transcription factor seems to be one of the most promising candidate genes involved in conferring the drought tolerance in several crops. Transcriptional regulation is potentially an area for coordinated regulation of genes relevant to stress tolerance, but requires identification of factors limiting the sustained response so that their expression may be manipulated (Xiong & Zhu, 2002). Stress-related genes are found in all plant species. It was hypothesized that genes from hardy plants could have evolved better protection properties and the homologues of *DREB1A* present in different genotypes will be suitable candidate genes for allele mining, integration in the genetic maps to associate them with drought tolerance QTLs, or undertaking candidate gene sequence-based association mapping. Therefore, in the present investigation, afragment of an important transcription factor *EcDREB1A* from finger millet genomic DNA was amplified with species-specific primer pair and characterized based on the sequence information publicly available for different type of *DREB* genes in some model and crop plant species, and using a variety of bioinformatics approaches. Available gene sequences from the model plant *A. thaliana* offer a unique opportunity to clone cognate genes and genes from heterologous sources resulting in discovery of new genes, allele-mining, etc. Primer design and sequence prediction are important steps before synthesizing primers to be used for cognate gene amplification and/or cloning. In the current study, a systematic approach has been made to predict a pair of theoretical primers for the *AtDREB1A* gene for heterologous gene cloning. To achieve this this goal, 16 different but related nucleotide sequences based on *AtDREB1A* gene were retrieved from different public databases. A highly conserved region of 287 bp was detected on multiple alignments through Claustalw2 program and primer set (forward and reverse) was predicted using Primer3plus and NetPrimer software on the basis of the conserved region, considering ideal conditions for primer length, GC content, T_m, formation of primer-dimers, hairpin-loops, etc. The predicted forward and backward primers have shown significant sequence similarity with *AtDREB1A* genes (graphics not shown here) and hence were used for amplification of *EcDREB1A* fragment, possibly leading to discovery of new genes and allele mining.

Finger millet *DREB1* (*EcDREB1A*) genomic sequence fragment was found to be 536 bp long. The similarity search result indicated extensive similarity to other *DREB* genes reported from different plant species suggesting that the amplified fragment belonged to *DREB* class of transcription factors. Significantly high identity of *EcDREB1A* was noticed with *DREB* and its homologs from *A. thaliana, Triticum aestivum*, and *Oryza sativa* (Table 3) Comparative analysis of nucleotide sequences showed 70% similarity and 98% query coverage with the reported *AtDREB2* mRNA, complete CDS, thus indicating the amplified fragment only partially covers the *AtDREB2* sequence. *EcDREB1A* fragment contains a putative ORF of 84 amino acids rich in serine (13.10%) with a predicted molecular mass of 9.29 kDa. Multiple sequence alignment of this *EcDREB1A* ORF with other *DREB* genes revealed presence of 9 highly conserved amino acids (R, P, G, F, D, R, G, C and N).

Allele mining exploits the DNA sequences of one genotype to isolate useful alleles from related genotypes. A significant percentage of allelic polymorphism is known to exist for a given gene and such small differences are implicated in evolutionary relationship between the gene and the trait in question. Such studies could provide the raw materials for plant breeding and translational genomic approaches (Latha et al., 2004; Cavallo et al., 2011). Allele mining of *EcDREB1A* gene across selected five finger millet cultivars revealed no variations on nucleotide, probably due to narrow genetic base among the cultivars taken for current study. Nevertheless, the data generated here indicates the presence of DREB sequence in the plant kingdom. Heterologous gene cloning and allele mining in finger millet with broad genetic base may lead to identification of useful alleles responsible for abiotic stress tolerance.

Acknowledgements

The authors acknowledge the financial support from the Department of Biotechnology, Govt. of India, New Delhi for M.Sc. (Agril. Biotechnology) teaching programme to the university under which the research could be carried out.

References

Acquaah, G. (2007). Principles of plant genetics and breeding. Blackwell, Oxford, UK.

Baker, S. S., Wilhelm, K. S., & Thomshow, M. F. (1994). The 5'-region of *Arabidopsis thaliana cor15a* has *cis*-acting elements that confer cold, drought and ABA-regulatedgene expression. *Plant Molecular Biology, 24*, 701-713. http://dx.doi.org/10.1007/BF00029852

Barkla, B. J., Vera-Estrella, R., & Pantoja, O. (1999). Towards the production of salt tolerant crops. *Advances in Experimental Medicine and Biology, 464*, 77-89. http://dx.doi.org/10.1007/978-1-4615-4729-7_7

Binas, M. (2000). Designing PCR primers on the web. *Biotechniques, 29*, 988-990.

Blumwald, E., Aharon, G. S., & Apse, M. P. (2000). Sodium transport in plant cells. *Biochemical and Biophysical Research Communications, 1465*, 140-151.

Bohnert, H. J., & Jensen, R. G. (1996). Strategies for engineering water stress tolerance in plants. *Trends in Biotechnology, 14*, 89-97. http://dx.doi.org/10.1016/0167-7799(96)80929-2

Cushman, J. C., & Bohnert, H. J. (2000). Genomic approaches to plant stress tolerance. *Current Opinion in Plant Biology, 3*, 117-124. http://dx.doi.org/10.1016/S1369-5266(99)00052-7

Dixon, R. A., & Arntzen, C. J. (1997). Transgenic plant technology is entering the era of metabolic engineering. *Trends in Biotechnology, 15*, 441-444. http://dx.doi.org/10.1016/S0167-7799(97)01115-3

Haake, V., Cook, D., Riechmann, J. L., Pineda, O., Thomashow, M. F., & Zhang, J. Z. (2002). Transcription factor CBF4 is a regulator of drought adaptation in Arabidopsis. *Plant Physiol, 130*(2), 639-648.

Hong, Z., Lakkineni, K., Zhang, Z., & Verma, D. P. S. (2000). Removal of feedback inhibition of 1-pyrroline-5-carboxylate synthetase results in increased proline accumulation and protection of plants from osmotic stress. *Plant Physiology, 122*, 1129-1136. http://dx.doi.org/10.1104/pp.122.4.1129

Katherine, E. V., & Robi, D. M. (2008). Nested Patch PCR enables highly multiplexed mutation discovery in candidate genes. *Genome Res., 18*(11), 1844-1850. http://dx.doi.org/10.1101/gr.078204.108

Kirankumar, S. M., Robert, P.T., & Gregory, B. M. (2001). *Arabidopsis* genome sequence as a tool for functional genomics in tomato. *Genome Biology*.

Knight, H. (2000). Calcium signaling during abiotic stress in plants. *International Review of Cytology, 195*, 269-324. http://dx.doi.org/10.1016/S0074-7696(08)62707-2

Latha, R., Rubia, L., Bennett, J., & Swaminathan, M. S. (2004). Allele mining for stress tolerance genes in *Oryza* species and related germplasm. *Molecular biotechnology, 27*, 101-108. http://dx.doi.org/10.1385/MB:27:2:101

Latchman, D. S. (2003). Eukaryotic transcription factors (4th ed.). San Diego: Academic press.

Liu, Q. K., Sakuma, Y., Abe, H., Miura, S., Yamaguchi-Shinozaki, K., & Shinozaki, K. (1998). Two transcription factors, DREB1 and DREB2, with an EREBP/AP2 DNA bindingdomain separate two cellular signal transduction pathways in drought and low temperature-responsive gene expression, respectively, in *Arabidopsis*. *Plant Cell, 10*, 1391-1406.

McCue, K. F., & Hanson, A. D. (1990). Drought and salt tolerance: towards\ understanding and application. *Trends in Biotechnology, 8*, 358-362. http://dx.doi.org/10.1016/0167-7799(90)90225-M

Sakuma, Y., Liu, Q., Dubouzet, J. G., Abe, H., Shinozaki, K., & Yamaguchi-Shinozaki, K. (2002). DNA-binding specificity of the ERF/AP2 domain of *Arabidopsis* DREB's, transcription factors involved in dehydration- and cold-inducible gene expression. *Biochemical and Biophysics Research Communications, 290*, 998-1009. http://dx.doi.org/10.1006/bbrc.2001.6299

Sambrook, J., Fritsch, E. F., & Maniatis, T. (1989). Molecular cloning-A Laboratory Manual (pp. 51-55, 2nd ed.). New York: Cold Spring Harbor Laboratory Press.

Sambrook, J., & Russel, D. (1989). W. *Molecular cloning: A laboratory manual* (3nd ed.). ColdSpring Harbor Laboratory Press.

Sanger, F., Donelson, J. E., Coulson, A. R., Kössel, H., & Fischer, D. (1973). Use of DNA Polymerase I primed by a synthetic oligonucleotide to determine a nucleotide sequence in Phage f1 DNA. *Proceedings of the National Academy of Sciences USA, 70*(4), 1209-1213. http://dx.doi.org/10.1073/pnas.70.4.1209

Shinozaki, K., & Yamaguchi-Shinozaki, K. (1996). Molecular responses to drought and cold stress. *Current Opinion on Biotechnology, 7*, 161-167. http://dx.doi.org/10.1016/S0958-1669(96)80007-3

Smirnoff, N. (1998). Plant resistance to environmental stress. *Current Opinion in Biotechnology, 9*, 214-219.

Stockinger, E. J., Gilmour, S. J., & Thomashow, M. F. (1997). Arabidopsis thaliana *CBF1* encodes an AP2 domain containing transcription activator that binds to the C repeat DRE, a *cis* acting DNA regulatory

element that stimulates transcription in response to low temperature and eater deficit. *Proceedings of National Academy of Science USA, 94*, 1035-1040. http://dx.doi.org/10.1073/pnas.94.3.1035

Thompson, J. D., Gibson, T. J., Plewniak, F., Jeanmougin, F., & Higgins, D. G. (1997). The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res., 25*, 4876 -4882. http://dx.doi.org/10.1093/nar/25.24.4876

Vogelstein, B., & Gillespie, D. (1979). Preparative and analytical purification of DNA from agarose. *Proceedings of National Academy of Science, 76*(2), 615-619. http://dx.doi.org/10.1073/pnas.76.2.615

Xiong, L., Schumaker, K. S., & Zhu, J. K. (2002). Cell signaling during cold, drought, and salt stress. *Plant Cell (Suppl), 14*, S165-S183.

Yamaguchi-Shinozaki, K., & Shinozaki, K. (1994). A novel *cis*-acting element in an *Arabidopsis* gene is involved in responsiveness to drought, low temperature or high salt stress. *Plant Cell, 6*, 251-264.

Yong-Li, X., Malik, M., Catherine, A. W., Christopher, D. T. (2002). Cloning and Sequencing of cDNAs for Hypothetical Genes from Chromosome 2 of Arabidopsis. *Plant Physiol*. http://dx.doi.org/10.1104/pp.010207

Zhu, J. K. (2001). Cell signaling under salt, water and cold stresses. *Current Opinion in Plant Biology, 4*, 401-406.

Zhu, J. K., Hasegawa, P. M., & Bressan, R. A. (1997). Molecular aspects of osmotic stress in plants. *Critical Reviews in Plant Sciences, 16*, 253-277. http://dx.doi.org/10.1016/S1369-5266(00)00192-8