Hereditary Genetics

Jadhao et al., Hereditary Genet 2014, 3:3
DOI: 10.4172/2161-1041.1000133

Research Article Open Access

Studies on Molecular Characterization of DREB Gene in Indica Rice (*Oryza sativa L.*)

Jadhao KR[2], Samal KC[2], Pradhan SK[1] and Rout G[2]*

[1]*Department of Bioinformatics, Centre for Post-Graduate Studies, Orissa University of Agriculture & Technology, Bhubaneswar-751003, Odisha, India*
[2]*Department of Agricultural Biotechnology, College of Agriculture, Orissa University of Agriculture & Technology, Bhubaneswar-751003, Odisha, India*

Abstract

Number of potential candidate genes (CGs) has been identified by transcriptome and transgenic approaches involved in the adaptive responses to drought in cereals. One of these genes conferring tolerance to drought is *DREB* gene family. The exploration of genetic/allelic variation within the *DREB* gene family is an important prerequisite towards a better understanding of stress tolerance. The present investigation aims to study allelic variation in nucleotide sequence within *DREB* gene of different rice cultivars need for crop improvement programme. Ten upland as well as lowland rice cultivars was used for molecular characterization and allele mining. The results showed that the *DREB* gene was found in all the tested germplasm and submitted to NCBI GenBank (accession numbers KF 545561 to KF 545569). Bioinformatics analysis showed that 100 percent identity with AP2 DNA binding domain of 59 amino acids which showed conserved three sheets with 14th valine and 19th glutamic acid conserved residues. AP2 domain of putative DREB protein were found rich in alanine (17.6%) and arginine (15.7%) amino acids with a predicted molecular mass of 5.89 kDa and iso-elecric point (pI) 10.38. It is observed that *DREB* gene nucleotide sequences were verified with DNA polymorphism analysis, out of which 196 invariable (monomorphic) and 8 variable (polymorphic) i.e. segregating sites were identified including 9 number of mutation and 5 haplotypes. The haplotypes (gene) diversity was 0.756; the variance and standard deviation diversity were 0.01678 and 0.130 respectively. The result showed that the variety 'Khandagiri' showed 97.5% similarity with all accessions *DREB* gene of rice and 0.036% evolutionary divergence with AK121956 accession present in the database. This might have enhanced the capability of *DREB* gene as a transcription factor and that lead to provide better adoption and tolerance capability of the cultivar during drought.

Keywords: Rice; DREB Transcription factor; Computational Sequence data analysis; AP2 domain

Introduction

Rice (*Oryza sativa L.*) is the second most important staple food for the largest part of the world. It belongs to family graminae (Poaceae) and a model system for cereal biology with the smallest genome consisting of 430 Mb across 12 chromosomes. Abiotic stress like drought and salinity are the most limiting factor of crop productivity and it is estimated to be more than 50% decline in the average yields of major crops worldwide [1]. However, drought tolerance is a genetically complex trait that involves multiple genes [2]. Abiotic stresses solely associated with physiological and developmental changes in plants, which are due to changes in plant genes expression [3]. There are some transcription factor(s) that regulate the expression of several genes related to stress. *DREB* (Dehydration responsive element binding factor) play key roles in plant stress signalling transduction pathway, they can specifically bind to DRE/CRT element (G/ACCGAC) and activate the expression of many stress inducible genes. Each *DREB* protein contains a basic N-terminal region that might function as a nuclear localization signal and acidic C-terminal region that might act as an activator domain for transcription. *DREB* genes family has been grouped into *DREB* l/CBF and *DREB*2. *DREB*1 includes 3 novel genes *viz*., *DREB*1A (CBF3), *DREB*1B (CBF1), *DREB*1C (CBF2) and *DREB*1D (CBF4) [4]. Five *DREB* homologs were identified in rice which includes Os DREB1A, Os DREB1B, OsDREB1C, OsDREB1D and OsDREB2A [4]. Analysis of the genomic sequences related to rice ERF and *DREB* gene families are useful in identification of new DREB genes that could play a major role in drought tolerance and their phylogenetic analyses [5]. The present study indicates the isolation and characterization of *DREB* gene present in indica rice varieties which are linked to drought tolerance.

Material and Method

Genotypes used

Ten promising upland as well as lowland cultivars of rice (*Oryza sativa* L.) viz. 'SwarnaSub-1', 'Udaygiri', 'Lalat', 'Bandana', 'RGL', 'Jagannath', 'Daya', 'Pary', 'Mahalaxmi', and 'Khandagiri were selected for the present study and were collected from the Rice Research Station, OUAT, Bhubaneswar and Central Rice Research Institute (CRRI), Cuttack.

Primer design

Nucleotide sequences of DREB gene family were retrieved from Genbank (http://www.ncbi.nlm.nih.gov/Genbank) database of NCBI and used for primer designing. Overlapping oligos of DREB gene sequences were designed using Primer3 tool (http://primer3.wi.mit.edu/ (Table 1) The predicted primers were subjected to check for various properties namely hairpin loops, primer dimer, Tm (temperature), GC% by using Premier Biosoft's Net Primer tool (http://www.premierbiosoft.com/netprimer/index.html). The specificity of both forward and reverse primers as well as product size was checked using primer Blast program of NCBI database (http://blast.ncbi.nlm.nih.gov/Blast.cgi) against our retrieved sequence.

DNA extraction and PCR analysis

The genomic DNA of ten rice cultivars were extracted using modified CTAB method [6] and purified DNA of each cultivars were subjected for PCR amplification using gene-specific left 5'–CCTCATTGGGTCAGGAAGAA-3' and right

***Corresponding author:** Rout GR, Department of Agricultural Biotechnology,, Orissa University of Agriculture & Technology, Bhubaneswar-751003, Odisha, India, Tel: 09437308014; E-mail: grrout@rediffmail.com

Received May 20, 2014; **Accepted** August 29, 2014; **Published** September 03, 2014

Citation: Jadhao KR, Samal KC, Pradhan SK, Rout GR (2014) Studies on Molecular Characterization of DREB Gene in Indica Rice (*Oryza sativa L.*). Hereditary Genet 3: 133. doi:10.4172/2161-1041.1000133

5'-GGATCTCAGCCACCCACTTA-3'primers (Merck Bioscience, India.). PCR amplification was carried out with the template DNA 25- 50 ng, 2.5 µl 10X PCR assay buffer (Merck Bioscience, India), 1.5 µl each of 10 mM dNTPs (M/S Merck Bioscience, India), 1 µl of 5 µM forward primer, 1 µl of 5µM primer and 1µl of 1U high fidelity *Pfu* DNA polymerase (Merck Bioscience). M/s Peqlab, 96 universal gradient thermal cycler was used for the PCR amplification consisted of a total of 35 cycles of initial denaturation (94°C for 5 min) followed by denaturation (94°C for 1 min), annealing (59°C for 2 min), elongation (72°C for 2 min) and final elongation (72°C for 10 min). The 1X Tris-acetate-ethylenediamine–tetra acetic acid (TAE) buffer was used to electrophoreses PCR amplified products in 2.5% (w/v) agarose gel (Merck Bioscience, India) along with 100 bp DNA ladder (Himedia Laboratories Pvt. Ltd., Mumbai, India). The gel image was documented using gel documentation system (UVITECH, Cambridge, UK).

PCR product purification and sequencing

The single bright amplicon of approximately 0.24 Kbp were eluted and purified from each ten samples using small DNA fragments extraction Kit (Gene Aid) following the published protocol [7]. The PCR amplified fragments of ~0.24 Kbp were sequenced with 96 capillary high throughput sequencer; ABI 3730 XL genetic analyserfollowing modified Sanger's dideoxy method [8] at Xcelris Genomics Ltd., Ahmedabad, India.

Nucleotide sequence data analysis

After sequencing all ten sequenced amplicons of different cultivars were subjected for Multiple Sequence Alignment (MSA) using ClustalW program at the EBI ClustalW server () and muscle (www. ebi.ac.uk/Tools/msa/muscle [9]. MEGA (Molecular Evolutionary genetic analysis 5.5) tool was used for phylogenetic tree construction for three different groups i.e. 'Khandagiri' with other nine cultivars, 'Khandagiri' with other rice cultivars and 'Khandagiri' with other crop plants (http://www.megasoftware.net) [10]. The third datasets were constructed by searching the homologus of 'Khandagiri'. DnaSP (DNA sequence polymorphism, version 5.10; http://www.ub.es/dnasp) [11] was used DNA polymorphism, InDel (insertion/deletion) polymorphism, nucleotide diversity, polymorphic sites etc.

Sl. No.	Accession Number	Gene	Sequence Type	Sequence length in base pairs (bp)
1	JN561151.1	DREB1A	complete cds	717
2	JQ885955.1	DREB1A	complete cds	717
3	JQ885956.1	DREB1A	complete cds	717
4	HM807364.1	DREB1	complete cds	969
5	JQ341059.1	DREB2A	mRNA	846
6	JQ885965.1	DREB1B	5'UTR	540
7	JQ885957.1	DREB1B	complete cds	657
8	JQ885966.1	DREB1B	5'UTR	531
9	JQ885958.1	DREB1B	complete cds	657
10	AY785895.1	DREB1D	complete cds	762
11	JF915844.1	DREB2B	mRNA	1602
12	JF915842.1	DREB2B	Complete cds	3299
13	JF915845.1	DREB2B	Complete cds	1684

Table 1: Nucleotide sequence information of DREB gene family of Rice (*Oryza sativa*) retrieved from NCBI.

Computational protein analysis

These ten sequenced of DREB gene were verified at protein level by using Blast-x (http://blast.ncbi.nlm.nih.gov/blast.cgi) programme from NCBI and Expasy (http://web.expasy.org/cgi-bin/translate/dna_aa) server. General features such as molecular weight, iso-electric point (pI), amino acid composition etc. of the predicted protein were assessed by using the ProtParam tool (http://expasy.org/cgi-bin/protparam).The consensus sequences from Blast-x for analysis and characterization of DREB gene in Plant TFDB (Transcription Factor Database, version 3.0) (http://planttfdb.cbi.pku.edu.cn/) for AP2 domain in rice.The structure prediction of AP2/ERF domain was done by using the Phyre2 () and model validations of the predicted structure was done by using Structural Analysis and Verification Server ().

Result

Out of nine primer pairs, three having 18 bp to 20 bp, 55 to 60% GC content, Tm value ranged from 59 to 62°C showed PCR amplification (Table 2). The primer specificity was checked for both forward and reverse primer by primer blast program which has shown significant alignment with DREB1 gene. The DREB1 gene fragment (0.24 Kbs) isolated from template DNA of different rice cultivars were amplified and presented in Figure 1. The 10 eluted fragments of rice cultivars were sequenced having 193 bp to 215 bp long nucleotides. On the basis of computational genomic analysis, the conserved region was identified in all the varieties (Figure 2) and found very little changes in the conserved region (Table 3). All the ten nucleotide sequences were used further for DNA polymorphism analysis in which 188 invariable (monomorphic) and 2 variable (polymorphic) sites were identified including total two

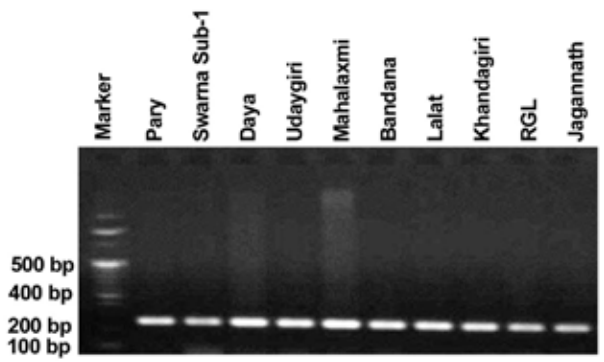

Figure 1: Amplification of 10 Rice cultivars employing DREB1 gene specific primer. M=Low range DNA markers.

Sl. No.	Primer	Length (bp)	Tm (°C)	Ta (°C)	GC (%)	Sequence (5' → 3')	Position of amplicon (nt)	
							Start	End
1	Left primer	20	60.04	59.00	50.00	5'CCTCATTGGGTCAGGAAGAA 3'	94	113
	Right primer	20	60.07	59.00	55.00	5'GGATCTCAGCCACCCACTTA 3'	334	315
2	Left primer	20	59.03	59.00	55.00	5'GACCAAGTTCAGGGAGACGA 3'	120	139
	Right primer	20	59.83	59.00	55.00	5'CAAGCTCGCGTAGTACAGGT 3'	627	608
3	Left primer	20	59.55	59.00	55.00	5'GGATCAAGCAGGAGATGAGC 3'	08	27
	Right primer	20	59.83	59.00	55.00	5'AAGCTCGCGTAGTACAGGTC 3'	626	607

Table 2: List of primers (forward and reverse) predicted by Primer-Blast.

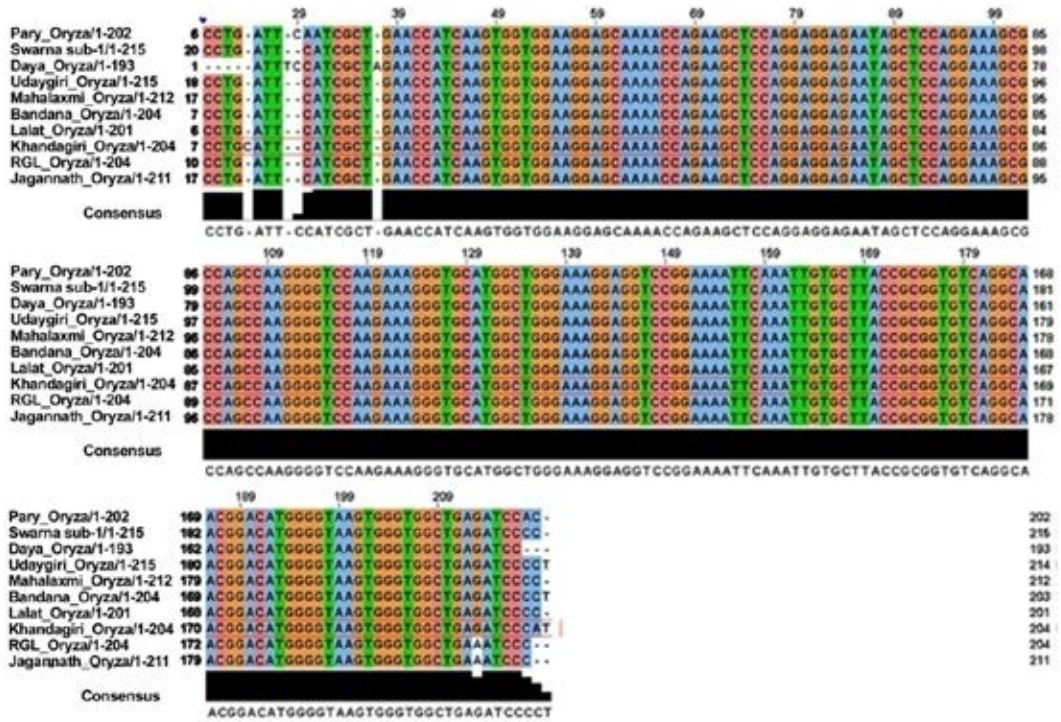

Figure 2: Jalview of muscle multiple sequence alignment tool (www.ebi.ac.uk Toolsmsa muscle) based on nucleotide sequences of ten varieties of rice.

Cultivars	Lowland/ Upland	Tolerant/ Susceptible	Position of Nucleotide InDel/substitution					
			24	28	29	30	37	213
Daya	Upland	Tolerant	-	T	C	-	A	-
Khandagiri	Upland	Tolerant	C	-	-	-	-	-
Pary	Upland	Tolerant	-	-	C	C → A	-	-
RGL	Upland	Moderate	-	-	-	-	-	G → A
Jagannath	Lowland	Susceptible	-	-	-	-	-	G → A

Table 3: Nucleotide changes due to insertion/deletion (InDel)/ substitution in DNA binding region of AP2 domain according to position.

	Pary	Swarna Sub-1	Daya	Udaygiri	Mahalaxmi	Bandana	Lalat	Khandagiri	RGL	Jagannath
Pary	-	**0.005**	**0.005**	**0.010**	**0.005**	**0.010**	**0.010**	**0.010**	**0.011**	**0.011**
Swarna Sub-1	97.51	**-**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.005**	**0.005**
Daya	99.48	100.0	-	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.005**	**0.005**
Udaygiri	97.01	94.37	100.0	**-**	**0.000**	**0.000**	**0.000**	**0.000**	**0.005**	**0.005**
Mahalaxmi	97.01	93.87	100.0	98.11	-	**0.000**	**0.000**	**0.000**	**0.005**	**0.005**
Bandana	97.51	99.01	100.0	97.55	97.52	**-**	**0.000**	**0.000**	**0.005**	**0.005**
Lalat	98.01	99.00	100.0	98.51	98.51	99.50	-	**0.000**	**0.005**	**0.005**
Khandagiri	96.52	97.52	100.0	97.04	97.03	98.52	98.51	**-**	**0.005**	**0.005**
RGL	96.50	96.57	99.47	96.08	95.59	98.51	98.00	98.01	-	**0.000**
Jagannath	97.00	93.84	99.47	97.16	96.68	97.01	98.00	97.01	96.08	-

N.B. Figures (normal font) represents of percent Similarity, Figures (bold font) represents of percent Diversity

Table 4: Percent Similarity and Diversity Matrix of sequenced rice cultivars.

number of mutation and three haplotypes. The haplotypes (gene) diversity was 0.51; the variance and standard deviation of haplotypes diversity were determined 0.027 and 0.164 respectively. On the basis of mismatch distribution and segregating sites, the frequency spectrum was predicted through graphical in which an average number of pair wise differences for observed value was estimated at 0.556 (Figure 3A and B). However, the variance and Coefficient of variance (CV) for observed value were 0.3434 and 1.0812 respectively as compared to expected value 0.244 and 0.8863.

The percent of similarity and evolutionary distance matrixes were generated from ten nucleotide sequences in the dataset by Clustal 2.1 in muscle tool and MEGA 5.5 (Table 4). The result showed that the variety 'Daya' showed 100% similarity with 'Udaygiri', 'Mahalaxmi',

'Bandana', 'Lalat' and 'Khandagiri' whereas 'Pary', 'RGL' and 'Jagannath' showed diversity ranges from 0.0053 to 0.011. Phylogenetic tree was construced along with control by using maximum parsimony method with bootstrap test in MEGA 5.5. There were three distinct clusters were formed in which the variety 'Khandagiri' and 'Daya' placed in separate cluster whereas another 8 varieties along with control were placed lie in third cluster (Figure 1A). The sequence information for DREB alleles have been deposited in the NCBI nucleotide database with accession numbers KF545561, KF545562, KF545563, KF545564, KF545565, KF545566, KF545567, KF545568 and KF545569.

Molecular and Phylogenetic analysis of DREB1 gene of variety 'Khandagiri' with different accessions of indica rice

DREB 1 gene sequence information of 'Khandagiri' was further analysed with the information available in NCBI database showed that a single nucleotide 'T' was substituted by 'A' (Figure 4). All *DREB* gene nucleotide sequences were further verified with DNA polymorphism analysis, out of which 196 invariable (monomorphic) and 8 variable (polymorphic) i.e segregating sites were identified including 9 number of mutation and 5 haplotypes. The haplotypes (gene) diversity was 0.756; the variance and standard deviation diversity were 0.01678 and 0.130 respectively. The percent of similarity and evolutionary distance matrixes were generated from second dataset by using the variety 'Khandagiri' with other possible DREB gene nucleotides of rice by Clustal 2.1 in muscle tool and MEGA 5.5 (Table 5). The result showed that the variety 'Khandagiri' showed 97.5% similarity with all accessions *DREB* gene of rice and 0.036% evolutionary divergence with AK121956 accession present in the database. It was further noted that there were three distinct clusters and the variety 'Khandagiri' placed in separate clustal but rest

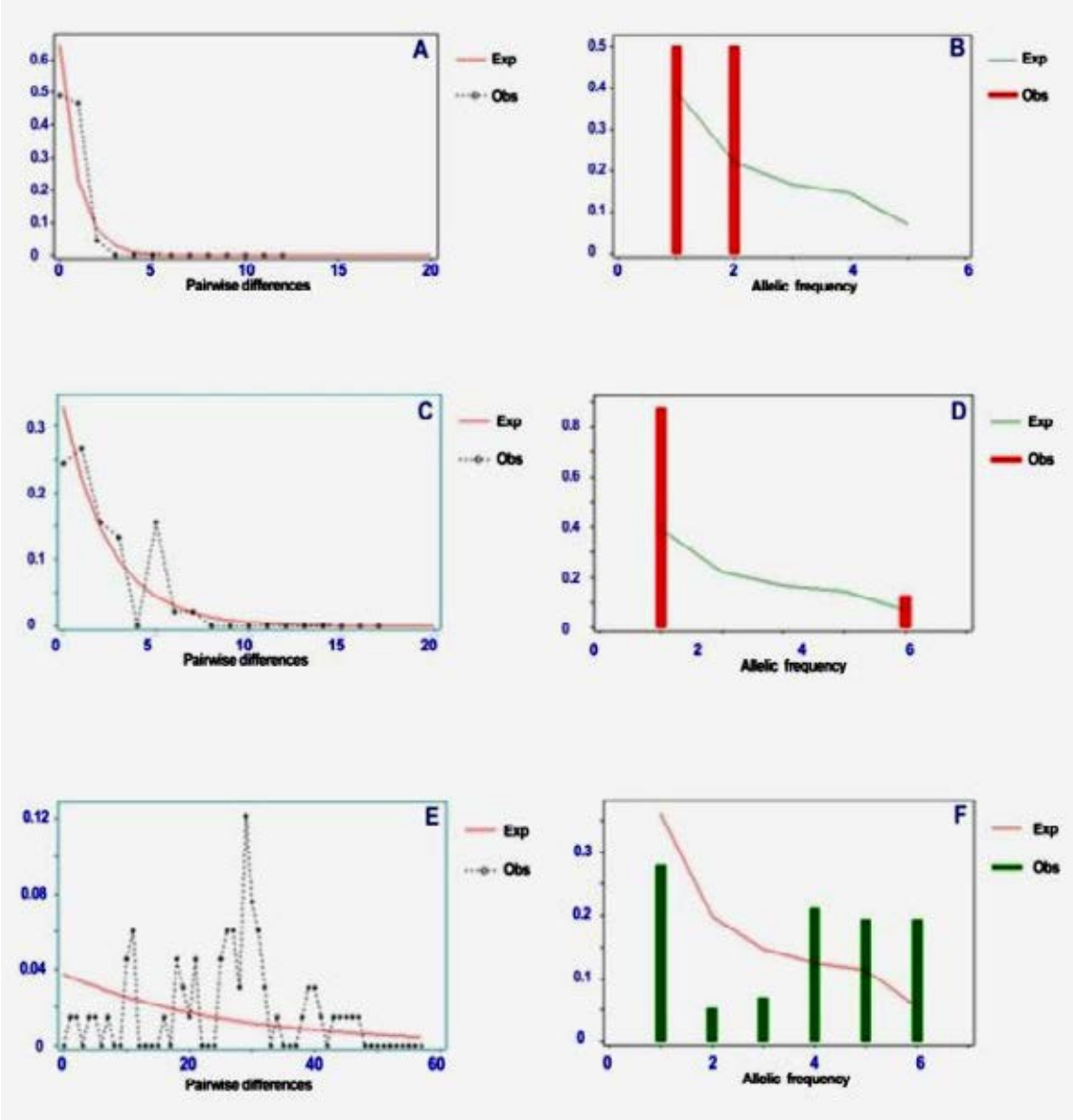

Figure 3: Graphical representation of (mismatch distribution) and frequency spectrum of 10 rice varieties (A and B), Variety 'Khabndagiri with other rice accessions (C and D), variety 'Khandagiri' with other monocotyledonds and dicotyledons (E and F) by using DnaSP software's.

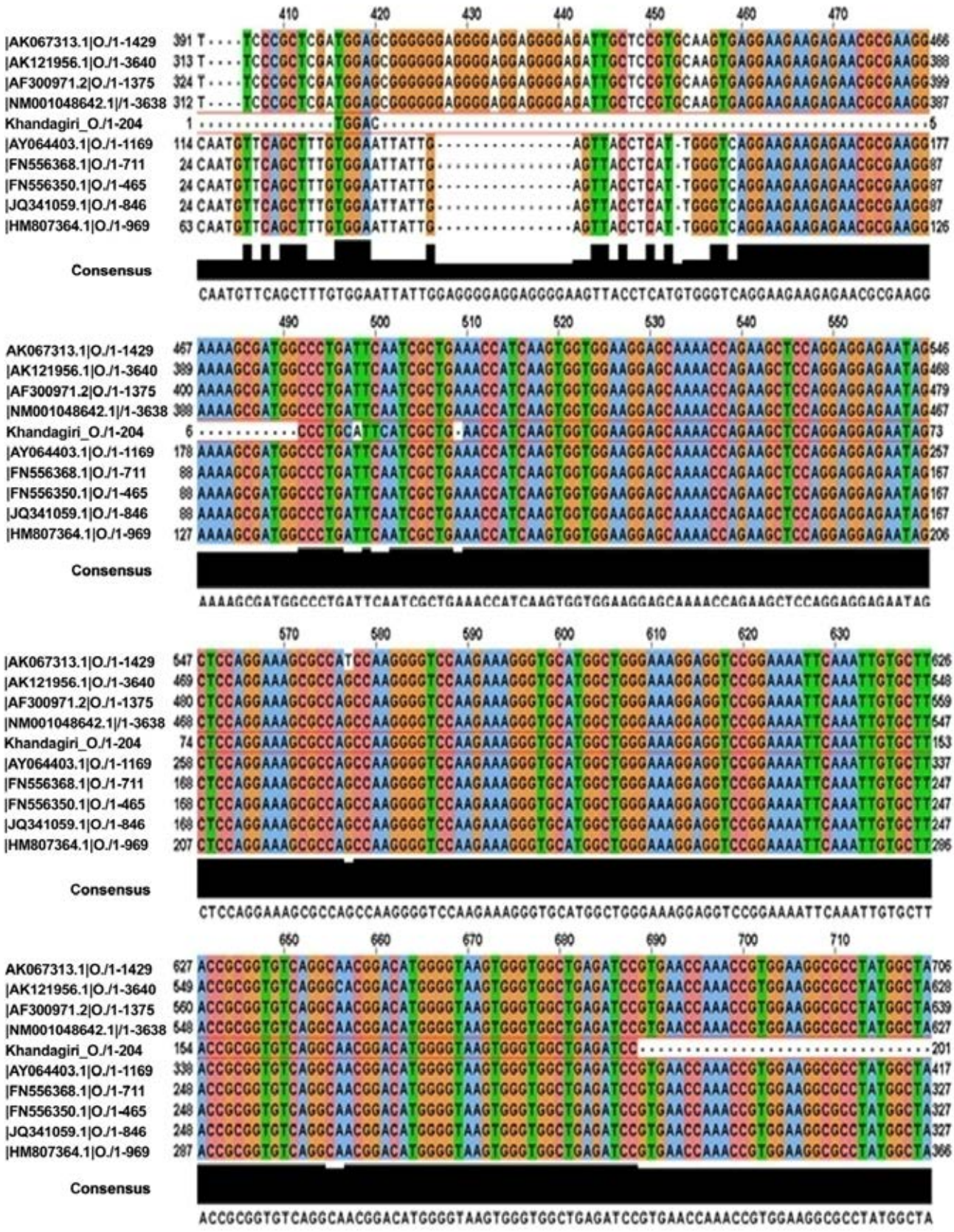

Figure 4: Jalview of Multiple Sequence Alignment of variety 'Khandagiri' with different accession of rice available in NCBI.

	AK067313.1	AK121956.1	AF300971.2	NM001048642.1	Khandagiri	AY064403.1	FN556368.1	FN556350.1	JQ341059.1	HM807364.1
AK067313.1	-	**0.15**	**0.005**	**0.005**	**0.031**	**0.01**	**0.01**	**0.01**	**0.01**	**0.01**
AK121956.1	99.7	**-**	**0.01**	**0.01**	**0.036**	**0.015**	**0.015**	**0.015**	**0.015**	**0.015**
AF300971.2	99.93	99.78	-	**0**	**0.025**	**0.005**	**0.005**	**0.005**	**0.005**	**0.005**
NM001048642.1	99.93	99.95	100	-	**0.025**	**0.005**	**0.005**	**0.005**	**0.005**	**0.005**
Khandagiri	97.06	96.57	97.55	97.55	-	**0.005**	**0.005**	**0.005**	**0.005**	**0.005**
AY064403.1	92.36	92.28	92.78	92.54	97.55	-	**0.001**	**0.001**	**0.002**	**0.003**
FN556368.1	95.76	95.62	95.9	95.9	97.55	99.86	-	**0.002**	**0.002**	**0.003**
FN556350.1	93.71	93.49	93.93	93.93	97.55	100	100	-	**0.001**	**0.002**
JQ341059.1	96.44	96.32	96.56	96.56	97.55	99.88	99.72	100	-	**0.001**
HM807364.1	93.89	93.78	93.99	93.99	97.55	99.9	99.72	100	100	-

N.B. Figures (normal font) represents of per cent Similarity, Figures (bold font) represents of per cent Diversity

Table 5: Per cent Similarity and Diversity Matrix of 'Khandagiri' with DREB rice accessions.

Accession number	Conserved consensus nucleotide positions in DREB gene											
	699	722	728	731	732	746	755	756	758	759	762	771
|AY920495.1|*Gossypium hirsutum*	A	T	G	G	T	T	A	A	T	G	T	T
|EU727155.1|*Nicotiana tabacum*	A	T	G	G	T	T	A	A	T	G	T	T
|FJ169307.1|*Arabidopsis thaliana*	A	T	G	G	T	T	A	A	T	G	T	T
|AF448789.1|*Zea mays*	A	T	G	G	T	T	A	A	T	G	T	T
|HE647689.1|*Glycine max*	A	T	G	G	T	T	A	A	T	G	T	T
|AY642596.1|*Glycine soja*	A	T	G	G	T	T	A	A	T	G	T	T
|AF500011.1|*Lycopersicon esculentum*	A	T	G	G	T	T	A	A	T	G	T	T
Khandagiri_*O.sativa*_DREB1	A	T	G	G	T	T	A	A	T	G	T	T
|HM807364.1|*Oryza sativa Indica*	A	T	G	G	T	T	A	A	T	G	T	T
|JN191708.1|*Triticum aestivum*	A	T	G	G	T	T	A	A	T	G	T	T
|DQ012941.1|*Hordeum vulgare*	A	T	G	G	T	T	A	A	T	G	T	T
|JN107537.1|*Hordeum brevisubulatum*	A	T	G	G	T	T	A	A	T	G	T	T

Table 6: Conserved consensus nucleotide positions in DREB gene family.

of the accesions represents two different phylogeny with bootstrap value (Figure 1B).

Molecular and Phylogenetic analysis of DREB1 gene of 'Khandagiri' with other crops

On the basis of comparison with other monocotyledons, dicotyledons and model plants available in NCBI database revealed that the conserved nucleotide sequences were changed the position as indicated in Figure 5. 'A' at position number 699, consecutively at 755 and 756. 'T' at position number 722, 732, 746, 758, 762, and 771. Likewise, 'G' at position number 728, 731 and 759 (Table 6). The results indicate that 'C' nucleotide was not conserved in *DREB* gene sequences. All the *DREB* gene nucleotide sequences were further analysed with DNA polymorphism, in which 13 invariable (monomorphic) and 57 variable (polymorphic) (i.e segregating sites) were identified having 94 number of mutation and 12 haplotypes. The haplotypes (gene) diversity was 1.00 and the variance, standard deviation of haplotypes diversity was 0.00116 and 0.034. On the basis of constant population size, the mismatch distribution and frequency spectrum was determined through DnaSP (Figure 3C-F), in which an average pair wise differences was 25.625. The variance and coefficient of variance were 122.5382 and 0.4405 respectively as observe value. However, the expected value of variance and coefficient of variance (no recombination) were 146.912 and 0.4725. On the basis of multiple sequence alignment for DREB1 gene using Neighbour joining method with 1000 replications in Bootstrap test in MEGA 5.5, the variety 'Khandagiri' was grouped with *Oryza sativa* indica group with bootstrap value 62 and *Arabidopsis thaliana*, *Nicotiana tabacum* with maximum bootsrtap value 90 (Figure 1A-C). The monocotyledon like *Zea mays* was more similar with *Arabidopsis thaliana* and other dicotyledons having bootsrap value 86.

Computational Structural analysis of DREB1 protein

The computational analysis showed that the amino acid sequence of DREB1 of varity Khandagiri (Figure 6) showed 100 percent similarity with rice germplasm *DREB* amino acid sequence of AP2 (apetala) domain available in the NCBI database with maximum score 122 and lowest E- value of 2e-33 (Table 7). All 68 amino acids present in ten tested varities of indica rice were further characterized through plant transcription factor database (http://planttfdb.cbi.pku.edu.cn/) for confirmation. The results revealed that out of 68 amino acids,59 amino acids were conserved to DNA binding region of AP2 domain. The amino acid sequences for AP2 domain were retrieved from NCBI database based on homology and used for multiple sequence alignment along with AP2/ERF domain available in variety 'Khandagiri'. There was very negligible differences with regard to DREB protein in conserved region of aminoacid sequences (Figure 7). The conserved consensus sequence of AP2 DNA binding domain of variety 'Khandagiri' was further used for computational biochemical anlysis which signifies the molecular weight of 5.89 kDa. with therotical isoelectric point (pI) 10.38. Among 22

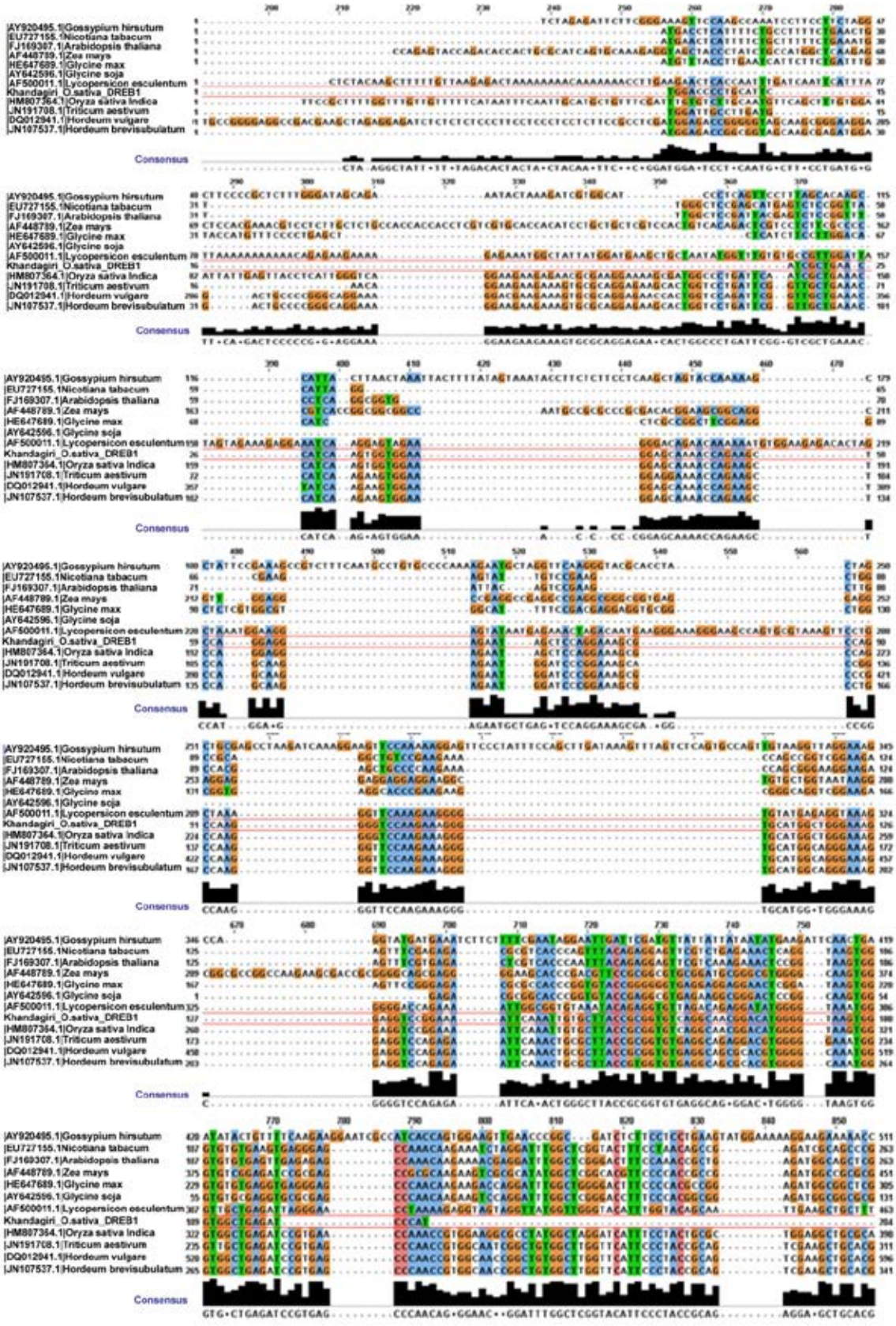

Figure 5: Jalview of different monocotyledon and dicotyledons DREB gene with the variety 'Khandagiri'.

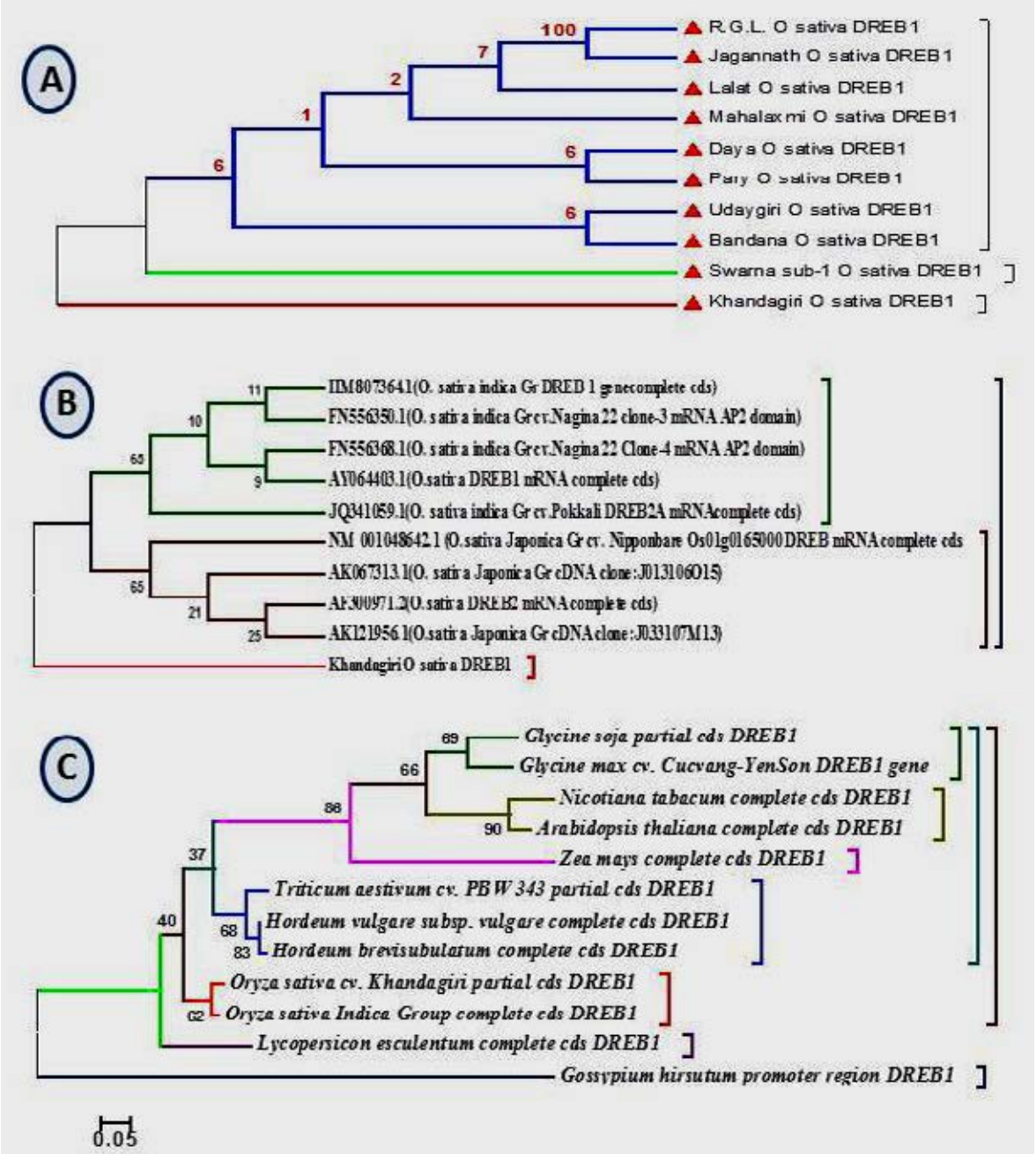

Figure 6: Phylogenetic tree based on nucleotide sequences of A) 10 cultivars constructed by maximum parsimony method. B) 'Khandagiri' cultivar with different accession of rice by maximum parsimony method. C) 'Khandagiri' cultivar with different crop plants by neighbour joining method.

Sl. No.	Accession number	Size of Amino acid	Amino acid sequences obtained by BLASTx tool	Per cent homology	E-value	Score
1	CBH19549.1	154	AP2 domain containing protein (*O. sativa indica* group)	100	2.00E-33	122
2	CBH19547.1	168	AP2 domain containing protein (*O. sativa indica* group)	100	2.00E-33	122
3	A2WL19.2	274	Dehydration-responsive element binding protein 2A short-protein DREB2A	100	2.00E-33	122
4	NP001042107.1	274	[OS01g0165600] *O. sativa japonica* group	100	2.00E-33	122
5	ADT9548.1	281	DREB like protein *O. sativa indica* group	100	2.00E-33	122
6	CBH19548.1	204	AP2 domain containing protein (*O. sativa indica* group)	100	2.00E-33	122
7	AAL40870.1	281	DREB1(*Oryza sativa*)	100	2.00E-33	122
8	CBH19567.1	236	AP2 domain containing protein (*O. sativa indica* group)	100	2.00E-33	122

Table 7: Amino acid homology of DREB1 by using BLASTX tool.

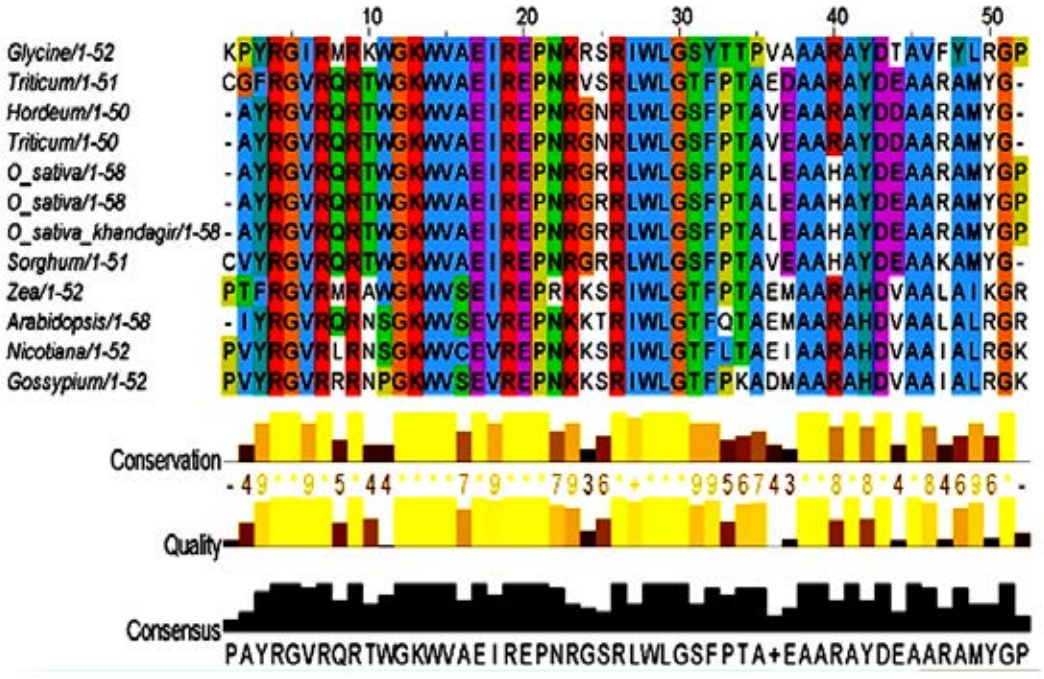

Figure 7: Comparison of Multiple sequence alignment of AP2 domain amino acids of different monocotyledon, dicotyledons and Arabidosis model plants available in NCBI with Khandagiri.

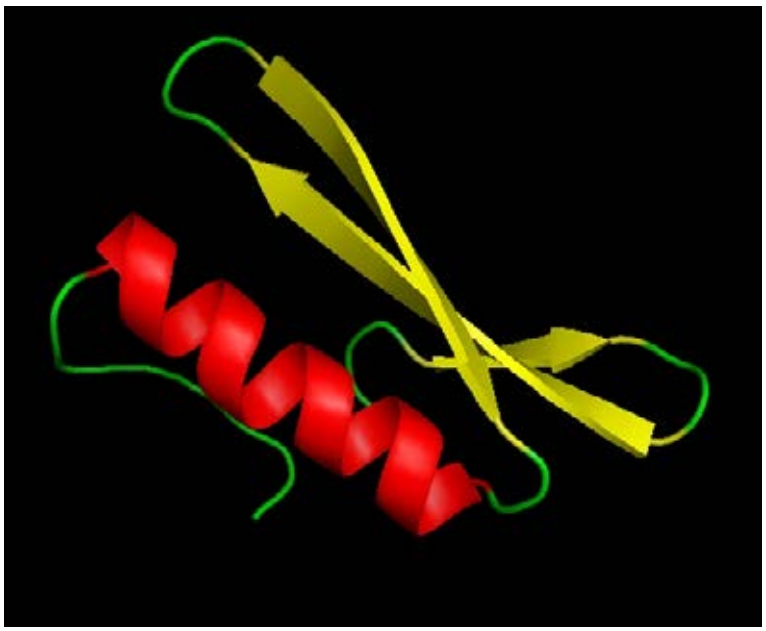

Figure 8: Structure of AP2 domain of variety 'Khandagiri' by using Phyre2 (httpwww.sbg.bio.ic.ac.ukphyre2).

essential amino acids, AP2 domain comprises maximum percentage of alanine (A) and arginine (R) with 17.6% and 15.7% respectively. Through computational observation, about 819 of atoms were encompasses in AP2 domain protein with their atomic composition and formula $C_{264}H_{401}N_{83}O_{70}S_1$ including total number of negatively charged (Asp + Glu) and positively charged (Arg + Lys) residues 5 and 9 respectively. The computed instability index was 34.21 which indicate the stability of AP2 domain of DREB 1 protein. It was observed that though there were differences at the sequence level but structurely they were similar in all other crops. The results revealed that there was three strand anti-parallel β-sheets towards N- terminal region and α-helix towards C-terminal end of the AP2 domain (Figure 8) with the conserved position of 14 th and 19 th valine and glutamic acid residues. The results showed that 89.1% of amino acid with 41residues were present in most favoured region of variety 'Khandagiri' whereas 10.9% of amino acid with 5 residues in additional allowed region (Figure 9).

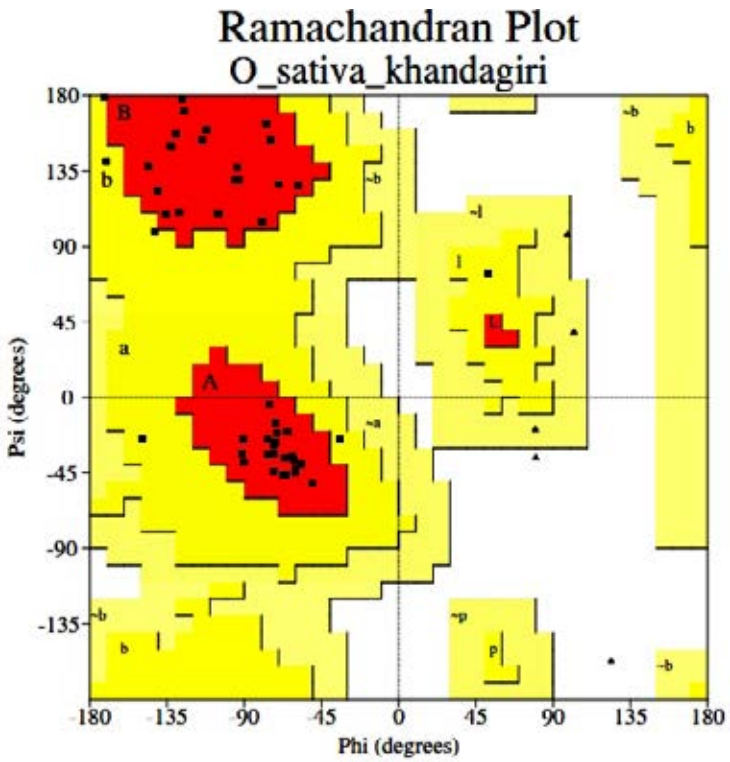

Figure 9: Structure validation (Ramchandran Plot) of AP2 domain through Structural Analysis and Verification Server. (http://nihserver. mbi.ucla.edu SAVES).

Discussion

Crop production is seriously affected by both biotic and abiotic stresses. Abiotic stresses such as drought lead to physiological and developmental changes in plant genes expression beginning with stress perception, which initiates a signal transduction pathway and is manifested in changes at the cellular, physiological and developmental levels [4]. Numbers of genes and their products respond to these stresses at transcriptional and translational level [12,13]. The present study indicates that a group of genes that regulate the gene expression and signal transduction during drought stress. These include transcription factors like CBF/DRE-binding protein or CBF/DREB [14,15].

In the present study, 241 base pairs long amplicon fragment was isolated from ten indica rice varieties and sequenced by their respective validated allele specific primers predicted by using bioinformatics software. The nucleotide sequences derived from ten varieties were subjected to multiple sequence alignment and the conserved region was identified with minor changes in the conserved region that might be resulted due to substitution, insertion and deletion of nucleotide(s) in the DNA binding region of AP2 domain during evolution process. The nucleotide sequence of variety 'Khandagiri' was further subjected to multiple sequence alignment along with rice *DREB* gene sequences retrieved from NCBI using Muscle tool. Jelview showed that the nucleotide sequences of all DREB genes exhibited the changes at initial sequence region but in rest of the region conserved consensus sequences which probably comprises AP2 domain. The variety 'Khandagiri' and cDNA clone: J013106O15 of Oryza *sativa* japonica Group showed substitution of nucleotide in conserved region in which 'T' was substituted by 'A' and 'T' was substituted by 'G'. Similar observation was obtained in number of monocotyledon and dicotyledonous crops.

All ten nucleotide sequences derived from indica rice were used for DNA polymorphism analysis in which the nucleotide diversity was 0.00292 with average number of nucleotide differences 0.556. The variance and coefficient of variance for observed value were 0.3434 and 1.0812 respectively.. The polymorphism analysis of DREB gene nucleotide sequences of rice revealed 196 invariable (monomorphic) and 8 variable (polymorphic) i.e segregating sites including total 9 number of mutation and 5 haplotypes. So the nucleotide diversity was estimated 0.01002 with average number of nucleotide differences 0.01002. Similarly, DNA polymorphism was also observed in *DREB* gene nucleotide sequences of other monocotyledons and dicotyledons which contain total 94 number of mutation and 12 haplotypes. The haplotypes (gene) diversity was found 1.0; the variance and standard deviation of heliotypes diversity were determined 0.00116 and 0.034. Similar results were reported in maize [16]. They reported that ideal polymorphism occur every 309 bp and SNPs occur every 79 bp in randomly selected sequences in the maize inbred B73 relative to inbred Mo17. Nucleotide variation (SNPs/InDels) was reported to underpin R-gene evolution and its function. Presence and absence of polymorphism have reported in 18.8% of total R-genes in *Arabidopsis* and 22.2% in rice [17,18]. The sequenced based phylogentic tree were constructed which revealed that the three distinct clusters in which 'Khandagiri' was found to be most diverged one and placed in the third cluster whereas the cultivar 'Swarna sub-1' was placed in the 2nd cluster and rest of the varieties were placed in first cluster. The phylogenetic analysis based on MSA against reported accesions of DREB gene in rice retrieved from database that the variety 'Khandagiri' placed in separate cluster but rest of the varieties represents two different phylogeny with high bootstrap value. The phylogenetic tree of variety 'Khandagiri' was constructed with other monocotyledon and dicotyledons, it was observed that *Gossypium hirsutum* was placed separately in a distinct cluster while the variety 'Khandagiri' grouped with *Oryza sativa* indica group with maximum bootstrap value 62. The model plant *Arabidopsis thaliana* formed a cluster with *Nicotiana*

tabacum showed more similarity with maximum bootsrap value 90. Similar observations were reported by Thakur et al. [19,20]. They reported that phylogentic analysis of all the *Pi-ta* alleles isolated from Indian land races shared identity with their wild progenotors and cultivated species except Lri_2 and Lri_15.

The nucleotide sequences were further translated to amino acid using BlastX programme (). The amino acid sequence of *DREB*1 showed 100 per cent similarity with reported rice *DREB* amino acid sequence of AP2 (apetala) domain of rice with maximum score 122 and lowest E- value of 2e-33. All the 68 amino acids from ten varieties were further investgated in Plant TFDB (http:// planttfdb.cbi.pku.edu.cn/) and it was noted that out of 68 amino acids, 59 amino acids were conserved to DNA binding region of AP2 domain. The result showed that AP2 domain of the sequences have three β-sheets, one α-helix with valine and glutamic acid residues are typical characteristics of AP2/EREBP domain of DREB protein [21-23]. The N-terminal end of AP2/ERF domain of rice has the conserved YRG basic and hydrophilic amino acid residues, essential for DNA binding property of DREB transcription factor. Similarly, towards the carboxylic end of AP2/ERF domain has conserved RAYD amino acid residues which involved in forming C-terminal amphipathic α-helix structure, which was considered as a vital for protein-protein interaction. After the characterization of AP2 domain, it was estimated that valine and glutamic acid in the 14^{th} and 19^{th} positions were conserved. These amino acid residues may play vital roles in recognition of the DNA-binding sequence [22,23], however, further research showed that E_{19} might not be as important as V_{14} for the recognition of the DNA-binding sequence [4]. Similarly while the characterization of ZmDREB1A, a homolog of *Arabidopsis* DREB1 in maize proteins precisely bound to DRE which includes highly conserved Valine at the 14^{th} residue in the ERF/AP2 domain. DNA binding domain of DREB protein was a play a crucial role to determine the specific interaction between this protein and the DRE consensus sequence. This study showed that ZmDREB1A has functional similarity to DREB1/CBFs in *Arabidopsis* [24]. Similar observation was obtained in cultivar 'Khandagiri' in which the AP2 domain of DREB1 putative protein also comprises three β-sheets and one α-helix structure which includes conserved valine and glutamic acid residues on $14^{th}$and 19^{th} position respectively.

From the study it was perceived that very minor modification in nucleotide sequence of *DREB* genes have been taken place during evolution process. The nucleotide sequences of all *DREB* genes are highly conserved but little insertion or deletion of nucleotide are occurred at the initial region of the sequences. The conserved consensus sequences in all *DREB* gene except in variety 'Khandagiri' in which 'T' was substituted by 'A' and one deletion mutation for 'A' has been occurred in the DNA binding region of AP2 domain of *DREB* protein. *DREB* (Dehydration responsive element binding factor) play key roles in plant stress signalling transduction pathway, they can specifically bind to DRE/ CRT element (G/ACCGAC) and activate the expression of many stress inducible genes. The insertion and deletion of nucleotide at the initial region of nucleotide sequence of DREB gene of the 'Khandagiri' cultivar might have enhanced the capability of *DREB* gene as a transcription factor and that lead to provide better adoption and tolerance capability of the cultivar during drought.

Acknowledgement

The authors wish to acknowledge to Department of Biotechnology, Government of India, New Delhi for providing financial assistance under HRD Post Graduate teaching program.

References

1. Bray EA, Bailey-Serres J, Weretilnyk E (2000) Response to abiotic stresses of Ca^{+2} dependent phospolipid-binding proteins in higher plant cells. FEBS Lett 244: 456-460.
2. Shinozaki K, Yamaguchi-Shinozaki K (2006) Gene networks involved in drought stress response and tolerance. Jour Exp Bot 56: 1-7.
3. Kobayashi F, Takumi S, Nakata M, Ohno R, Nakamura T, et al. (2004) Comparative study of the expression profiles of the Cor/Lea gene family in two wheat cultivars with contrasting levels of freezing tolerance. Plant Physiol 120: 585-594.
4. Sakuma Y, Liu Q, Dubouzet JG, Ab H, Shinozaki K, et al. (2002) DNA-binding specificity of the ERF/AP2 domain of Arabidopsis DREBs, transcription factors involved in dehydration- and cold-inducible gene expression. Biochem. Biophys. Res Comm 290: 998-1009.
5. Hemalatha N, Rajesh MK, Narayanan NK (2012) Genome-Wide Analysis of Putative ERF and DREB GENE Families in Indica Rice (*O. sativa* L.subsp. Indica) Int Journal of Machine Learning and Computing 2: 188.
6. Edwards A, Civitello A, Hammond HA, Casey CT (1991) DNA typing and genetic mapping with trimeric and tetrameric tendem repeats. Am Jour Hum Genet 49: 746-756.
7. Vogelstein B, Gillespie D (1979) Preparative and analytical purification of DNA from agarose. Proc National Acad Sci 76: 615-619.
8. Sanger F, Nicklen S, Coulson AR (1977) DNA sequencing with chain-terminating inhibitors. Proc Natl Acad Sci USA 74: 5463-5467.
9. Thompson JD, Gibson TJ, Plewniak, F, Jeanmougin F, Higgins DG, et al. (1997) The CLUSTAL-X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res 25: 4876-4882.
10. Tamura K, Peterson D, Peterson N, Stecher G, Nei M, et al. (2011) MEGA5: Molecular Evolutionary Genetics Analysis using Maximum Likelihood, Evolutionary Distance, and Maximum Parsimony Methods. Mol. Biology and Evol 28: 2731-2739.
11. Rozas J, Sánchez-DelBarrio JC, Messeguer X, Rozas R (2003) DnaSP, DNA polymorphism analyses by the coalescent and other methods. Bioinformatics 19: 2496-2497.
12. Yamaguchi-Shinozaki K, Shinozaki K (2005) Organization of cis-acting regulatory elements in osmotic- and cold-stress-responsive promoters. Trends in Plant Science.10: 88-94.
13. Umezawa T, Fujita M, Fujita Y, Yamaguchi-Shinozak K, Shinozaki K, et al. (2006) Engineering drought tolerance in plants: discovering and tailoring genes unlock the future. Current Opinion in Biotech 17: 113-122.
14. Shen YG, Zhang WK, He SJ, Zhang JS, Liu QS, et al. (2003) An EREBP/ AP2-.type protein in *Triticum aestivum* was a DRE- binding transcription factor induced by cold dehydration and ABA stress. Theor Appl Genet 106: 923-930.
15. Garg N, Pundhir S, Prakash A, Kumar A (2008) Primer designing for DREB1A, A cold induced gene. Jour Proteomics Bioinformatics 1: 37-46.
16. Vroh BI, McMullen MD, Villeda HS, Schroeder S, Gardiner J, et al. (2006) Single nucleotide polymorphisms and insertion deletions for genetic markers and anchoring the maize fingerprint counting physical map. Crop Sci 46: 12-21.
17. Meyers BC, Kozik A, Griego A, Kuang H, Michelmore RW, at al. (2003) Genome-wide analysis of NBS-LRR-encoding genes in *Arabidopsis*. Plant Cell 15: 809-834.
18. Shen JH, Araki L, Chen JQ, Chen TD (2006) Unique evolutionary mechanism in R-genes under the presence/absence of polymorphism in *Arabidopsis thaliana*. Genetics 172: 1243-1250.
19. Thakur S, Singh PK, Rathour R, Variar M, Prashanthi SK, et al. (2012) Positive selection pressure on rice blast resistance allele Piz-t makes it divergent in Indian landraces. Jour Plant Interact 1: 34-44
20. Thakur S, Gupta YK, Singh PK, Rathour R, Variar M, et al. (2013) Molecular diversity in rice blast resistance gene Pi-ta makes it highly effective against dynamic population of Magnaportheoryzae. Funct. Integr Genomics 13: 309-322.
21. Medina J, Bargues M, Terol J, Perez-Alonso M, Salinas J, et al. (1999) The Arabidopsis CBF genefamily is composed of three genes encoding AP2domain-containing proteins whose expression isregulated by low temperature but not by abscisic acid or dehydration. Plant Physiol 119: 463-469.
22. Okamuro JK, Caster B, Villarroe IV, Van Montagu M, Jofuku KD, et al. (1997) The AP2 domain of APETALA2 defines a large new family of DNA binding

proteins in Arabidopsis. Proc Natl Acad Sci USA 94: 7076-7081.

23. Liu Q, Kasuga M, Sakuma Y, Abe H, Miura S, et al. (1998) Two transcription factors, DREB1 and DREB2, with an EREBP/AP2 DNA binding domain separate two cellular signal transduction pathways in drought-and low-temperature-responsive gene expression, respectively, in *Arabidopsis*. Plant Cell 10: 1391-1406.

24. Qin F, Sakuma Y, Li J, Liu Q, Li Y, et al. (2004) Cloning and functional analysis of novel DREB1c/CBF transcription factor involved in cold-responsive gene expression in *Zea mays*. Plant and Cell Physiol 45: 1042-1052.