

Provided for non-commercial research and education use.
Not for reproduction, distribution or commercial use.



This article appeared in a journal published by Elsevier. The attached copy is furnished to the author for internal non-commercial research and education use, including for instruction at the authors institution and sharing with colleagues.

Other uses, including reproduction and distribution, or selling or licensing copies, or posting to personal, institutional or third party websites are prohibited.

In most cases authors are permitted to post their version of the article (e.g. in Word or Tex form) to their personal website or institutional repository. Authors requiring further information regarding Elsevier's archiving and manuscript policies are encouraged to visit:

<http://www.elsevier.com/copyright>



Contents lists available at ScienceDirect

Plant Science

journal homepage: www.elsevier.com/locate/plantsci

Isolation and sequence analysis of DREB2A homologues in three cereal and two legume species

Spurthi N. Nayak^{a,b}, Jayashree Balaji^a, Hari D. Upadhyaya^a, C. Tom Hash^a, P.B. Kavi Kishor^b, Debasis Chattopadhyay^c, Lina María Rodríguez^d, Matthew W. Blair^d, Michael Baum^e, Kenneth McNally^f, Dominique This^g, David A. Hoisington^a, Rajeev K. Varshney^{a,h,*}

^a International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), Patancheru 502324, AP, India

^b Department of Genetics, Osmania University, Hyderabad 500007, AP, India

^c National Institute of Plant Genome Research (NIPGR), Aruna Asaf Ali Marg, New Delhi 110067, India

^d International Centre for Tropical Agriculture (CIAT), A.A. 6713, Cali, Colombia

^e International Centre for Agricultural Research in Dry Areas (ICARDA), P.O. Box 5466, Aleppo, Syrian Arab Republic

^f International Rice Research Institute (IRRI), DAPO Box 7777, Metro Manila, Philippines

^g Montpellier SupAgro, 34060 Montpellier Cedex, France

^h Genomics Towards Gene Discovery Sub Programme, Generation Challenge Programme (GCP), c/o CIMMYT, Int APDO Postal 6-641, 06600 Mexico DF, Mexico

ARTICLE INFO

Article history:

Received 4 June 2009

Received in revised form 25 July 2009

Accepted 28 July 2009

Available online 5 August 2009

Keywords:

DREB

Candidate gene

Drought stress

SNP

Haplotype

ABSTRACT

The transcription factor, DREB2A, is one of the promising candidate genes involved in dehydration tolerance in crop plants. In order to isolate DREB2A homologues across cereals (rice, barley and sorghum) and legumes (common bean and chickpea), specific or degenerate primers were used. Gene/phylogenetic trees were constructed using a non-redundant set of 19 DREB1A and 27 DREB2A amino acid sequences and were combined with taxonomic/species tree to prepare reconciled phylogenetic trees. In total, 86 degenerate primers were designed for different clades and 295 degenerate primer combinations were used to amplify DREB homologues in targeted crop species. Successful amplification of DREB2A was obtained in case of sorghum. In parallel, gene-specific primers were used to amplify DREB2A homologues in rice, barley, common bean and chickpea. Seven to eight diverse genotypes from targeted species were used for sequence analysis at DREB2A locus identified/isolated. A maximum of eight SNPs were found in the common bean DREB2A, indicating two distinct haplotypes, three SNPs with five haplotypes were observed in barley whereas a single SNP was observed in rice, sorghum and chickpea. Parsimony based phylogenetic tree revealed distinct clustering of cereals and legumes. Furthermore, alignment of corresponding amino acid sequences showed conservation of AP2 domain across the targeted species.

© 2009 Elsevier Ireland Ltd. All rights reserved.

1. Introduction

Plant growth and development is adversely affected by abiotic stresses such as drought, low temperature and soil salinity. An increasing number of stress-induced genes at transcription level have been identified in several plant species [1–5]. Abiotic stress mediated gene expression has been shown to be regulated by different transcription factors, of which drought responsive element binding (DREB) proteins play a key role. DREB transcription factors bind to dehydration responsive element (DRE) of the genes at the promoter region and regulate the expression of

downstream genes. The DRE containing core sequence A/GCCGAC was identified as a *cis*-acting promoter element, which regulates gene expression in response to drought, high salinity and cold stresses in *Arabidopsis* [6]. On the other hand, C-repeat (CRT) and low temperature responsive element (LTRE) were identified in cold inducible genes [7,8].

The DREB proteins induce a set of abiotic stress responsive genes and maintain water balance in plant systems thus imparting abiotic stress tolerance. DREB proteins are transcription factors that bind to the promoter of genes such as *rd29A*, thereby inducing expression in response to drought, salt and cold [9–14]. The DREB transcription factors have been divided into two classes – DREB1 and DREB2 – based on their involvement in signal transduction pathways under low temperature and dehydration, respectively. They belong to the ERF (ethylene responsive factors) family of transcription factors. ERF proteins are the sub-family of the APETALA (AP2) or ethylene responsive element binding protein

* Corresponding author at: International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), Patancheru 502324, AP, India. Tel.: +91 40 30713071; fax: +91 40 30713074/30713075.

E-mail address: r.k.varshney@cgiar.org (R.K. Varshney).

(EREBP) transcription factors that are unique to plants. DREB1/C-repeat binding factor (CBF) genes are induced by cold stress [15–17], and found to confer tolerance to cold stress in various plant species, e.g. maize [18], rice [9], barley [19–21], wheat [22], soybean [23] and Brassica [24]. The DREB2 genes, induced by drought stress, have been reported to confer drought tolerance in several plant species such as *Arabidopsis* [25–27], rice [9], wheat [28], maize [29], barley DRFs [30] and chickpea CAP2 [31]. While both DREB1/CBF and DREB2 genes share a sequence similarity at AP2 domain and bind to the same DRE sequence, they are either up-regulated by low temperature (DREB1) or by drought/high salt concentration (DREB2).

Water stress; a pressing environmental problem today is likely to worsen in the future due to global warming/climate change, adversely affecting crop production in the developing countries. While linkage mapping based approach has been successful for the identification of genes or quantitative trait loci (QTLs) for traits of interest in several plant species, availability or development of a segregating population for the trait of interest is a pre-requisite for application of this approach [32]. On the other hand linkage disequilibrium (LD) based association mapping approach can be used with a natural population/germplasm collection for trait mapping [33,34]. Either of two approaches, candidate gene sequencing or whole genome scanning, may be used for association mapping [35,36]. While the whole genome scanning approach, requires deployment of a larger number of markers depending on the extent of LD in the species/germplasm, most putative candidate gene(s) involved in controlling the expression of the trait can be used relatively easily for sequencing alleles in a natural population.

DREB2 transcription factor seems to be one of the most promising candidate genes involved in conferring drought tolerance in several crops. The present study to isolate the DREB2 homologues in five crop species including three cereal species, i.e. barley (*Hordeum vulgare*), rice (*Oryza sativa*), sorghum (*Sorghum bicolor*) and two legume species, i.e. chickpea (*Cicer arietinum*) and common bean (*Phaseolus vulgaris*) was undertaken with the assumption that the homologues of DREB2 present in different species will be suitable candidate genes for integrating in genetic maps to associate them with drought tolerance QTLs or undertaking candidate gene sequencing based association mapping. Based on the sequence information available for different type of DREB genes in some model and crop plant species and using a variety of bioinformatics approaches, species-specific as well as degenerate primer pairs were designed to amplify putative DREB2 homologues in the target crop species. Subsequently, an effort was made to identify SNPs in identified DREB2 sequences.

2. Materials and methods

2.1. Plant material

The DNA was isolated from leaf samples from seven/eight diverse genotypes of five crop species, including three cereals, i.e.

rice (*O. sativa*), barley (*H. vulgare*) and sorghum (*S. bicolor*) and two legumes, i.e. chickpea (*C. arietinum*) and common bean (*P. vulgaris*) (Table 1). For understanding the allelic diversity at isolated/identified genes in the species germplasm, these genotypes were selected from the reference collection, a sub-set of global composite collection defined based on molecular diversity (e.g. [37]), for the corresponding species.

2.2. Sequence data

The nucleotide and protein sequences for DREB1A and DREB2A were downloaded from National Centre for Biotechnology Information (NCBI-www.ncbi.nlm.nih.gov), The Institute for Genomic Research (TIGR-<http://compbio.dfci.harvard.edu/tgi/plant.html>) now called J. Craig Venter Institute (JCVI), SWISSPROT (www.expasy.ch/sprot/), The *Arabidopsis* Information Resource (TAIR-www.arabidopsis.org) and Legume Information Systems (LIS-www.comparative-legumes.org) during July 2007. Standalone BLAST was used to check for redundant sequences. Sequence translation was carried out using the EMBOSS Transeq programme (<http://www.ebi.ac.uk/emboss/transeq/index.html>). Gene trees were constructed using the Protpars programme from Phylip V3.65. Tree reconciliation was carried out using the Notung programme [38] with the species tree derived from the NCBI taxonomy (<http://www.ncbi.nlm.nih.gov/Taxonomy/Common-Tree/wwwcmt.cgi>).

2.3. Primer design

Putative homologous consensus region of DREB2 in five crop species were amplified using two approaches. The degenerate primers were designed for consensus sequences of amino acids obtained from reconciled phylogenetic tree using the CODEHOP programme (Consensus DEgenerate Hybrid Oligonucleotide Primers) [39,40]. In the second approach, Primer3 tool was used to design gene-specific primers for the crop species for which sequences were available in the public domain.

2.4. Polymerase chain reaction (PCR)

PCR reactions using degenerate/species-specific gene primers were performed in 20 µl reaction volume in GeneAmp[®] PCR System 9700 thermal cycler (Applied Biosystems USA). The reaction mixture contained final concentration of 10 ng of genomic DNA, 1× PCR buffer, 1.5 mM MgCl₂, 0.1 mM of dNTP mix, 0.2 mM of each primers, and 0.2 U of AmpliTaq Gold *Taq* DNA Polymerase (Applied Biosystems, USA). The temperature profile consisted of an initial denaturation of 15 min at 95 °C followed by first 5 cycles of 94 °C for 20 s, 60 °C for 30 s and 72 °C for 30 s, with 1 °C decrease in annealing temperature per cycle, then 30 cycles of 94 °C for 20 s with constant annealing temperature (56 °C) for 1 min and 72 °C for 1 min 30 s followed by a final extension for 20 min at 72 °C and stored at 4 °C until further use. The amplified products were

Table 1
List of genotypes for five crop species used for assaying allelic diversity.

Sl. No.	Cereals			Legumes	
	Rice	Barley	Sorghum	Chickpea	Common bean
1	AZUCENA	IG 40059	BTX 623	Annigeri	DOR 364
2	N22	IG 138216	B 35	ICCV 2	G 19833
3	Nipponbare	IG 138219	E 36-1	ICC 4958	BAT 477
4	DOM-SOFID	IG 138221	R 16	ICC 283	BAT 93
5	Lijiangxintuanheigu (LTH)	IG 138226	IS 9830	ICC 8261	JaloEEP558
6	IR 64	IG 135258	IS 27761	IC 10029	G 21212
7	Shanhuangzhan-2 (SHZ)	IG 128202	IS 29233	ICC 1882	BAT 881
8		IG 138211	IS 18933	ICC 4411	AFR 298

separated on ethidium bromide stained 1.2% agarose gel and visualized with UV light.

2.5. DNA sequencing and sequence analysis

The PCR amplicons were purified using 1 U each of Exonuclease I and shrimp alkaline phosphatase (SAP) per 200 ng of PCR product. The Exo/SAP added PCR products were incubated for 45 min at 37 °C followed by denaturing at 80 °C for 15 min to deactivate the unused Exonuclease I enzyme. The Exo/SAP treated templates were subjected to BigDye Terminator V 3.1 (Applied Biosystems, USA) sequencing chemistry using respective primers. The cycle sequencing PCR conditions, consisted of an initial denaturation of 96 °C for 30 s, followed by 60 cycles of 96 °C for 10 s, 50 °C for 5 s, and 60 °C for 4 min. Subsequently, PCR products were purified using ethanol–EDTA precipitation and sequenced in both directions with respective primers using an ABI3700 DNA analyzer (Applied Biosystems, USA).

2.6. SNP identification and analysis

The contigs for each genotype were generated using both forward and reverse sequence chromatograms with the help of DNA Baser V 2.9 programme. The contig sequences from seven to eight genotypes of each crop species were used to develop multiple sequence alignment (MSA) using online ClustalW programme (<http://www.ebi.ac.uk/Tools/clustalw2/index.html>). The MSA files were used to analyze the presence of SNP and haplotypes specific to crops. Polymorphic information content (PIC) of each SNP was calculated using

$$PIC = 1 - \sum_{i=1}^n f_i^2$$

where f_i is the frequency of the i th allele.

In order to study the sequence variation at protein level, the nucleotide sequences were translated into six possible amino acid sequences using online EMBOSS Transeq programme (<http://www.ebi.ac.uk/emboss/transeq/index.html>). All six reading

frames were searched for homology with Uniprot database at NCBI. Phylogenetic trees were obtained from protein sequences using MEGA 4.1 programme.

3. Results and discussion

Available nucleotide and protein sequences for DREB in public domain were downloaded by searching various databases, i.e. NCBI, TIGR/JCVI, SWISSPROT, TAIR, LIS. As a result, 64 nucleotide sequences (27 DREB1A and 37 DREB2A) and 31 amino acid sequences (16 DREB1A and 15 DREB2A) were assembled. After translating nucleotide sequences and evaluating redundancies in the entire dataset, a non-redundant dataset consisting of 19 DREB1A and 27 DREB2A protein sequences was used for further analysis.

3.1. Reconciliation of gene tree and species tree

The reconciled trees were generated after rooting the species and gene tree to identify duplication (indicated as red nodes-D) or speciation (blue nodes) (Figs. 1 and 2). Reconciliation is the process of constructing a map between the gene tree and species tree, which resolves disagreement between gene and species trees, by invoking gene duplication and losses to explain topological incongruence. Duplication-loss parsimony is used to infer duplications and losses in the history of the gene family along with the species lineages in which these events occurred.

The reconciled tree of DREB1A (Fig. 1) indicates the clear demarcation of monocots DREB1A from dicots. DREB1A sequences for all dicot species, i.e. *Arabidopsis thaliana* (**BAD43714**, **AAG51606**, **NP_172721**, **DQ372533**, **Q9CAN9**, **AAG51609**, **NP_177887**), *Medicago truncatula* (**Q2XQA6**) and *Nicotiana tabacum* (**ABD65969**) were grouped in Cluster I. Cluster II contained two clades of monocot DREB1A sequences, one cluster (Cluster IIAi) including *O. sativa* (**Q8GVD6**), *H. vulgare* (**Q4ZGK0**), and the other (Cluster IIAii) containing *H. brevisubulatum* (**Q2XUK0**), *Festuca arundinacea* (**CAG30550**) and *O. sativa* (**AAN02486**). *Zea mays* (**AAN76804**) formed a separate group indicated as Cluster IIB. DREB1A from moss *Physcomitrella patens* (**ABA43687**) belonging to

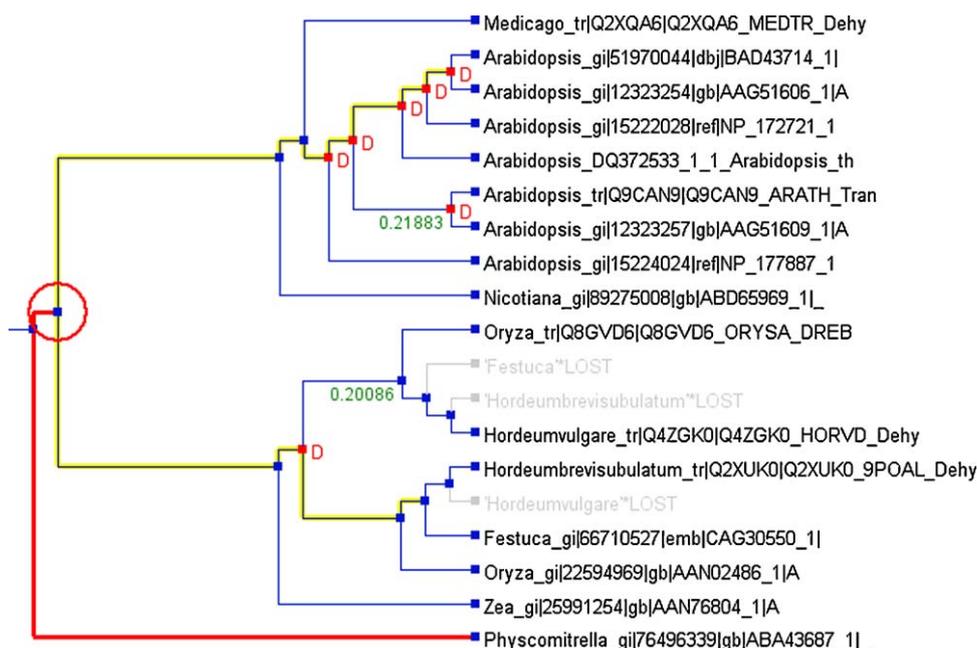


Fig. 1. The reconciled taxonomic (species) tree with DREB1A gene tree. The reconciled tree for DREB1A gene sequences was generated after rooting the species and gene tree to identify duplication (indicated as red nodes-D) or speciation (blue nodes). This tree shows clear grouping of DREB1A sequences of cereals and legumes separately. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)

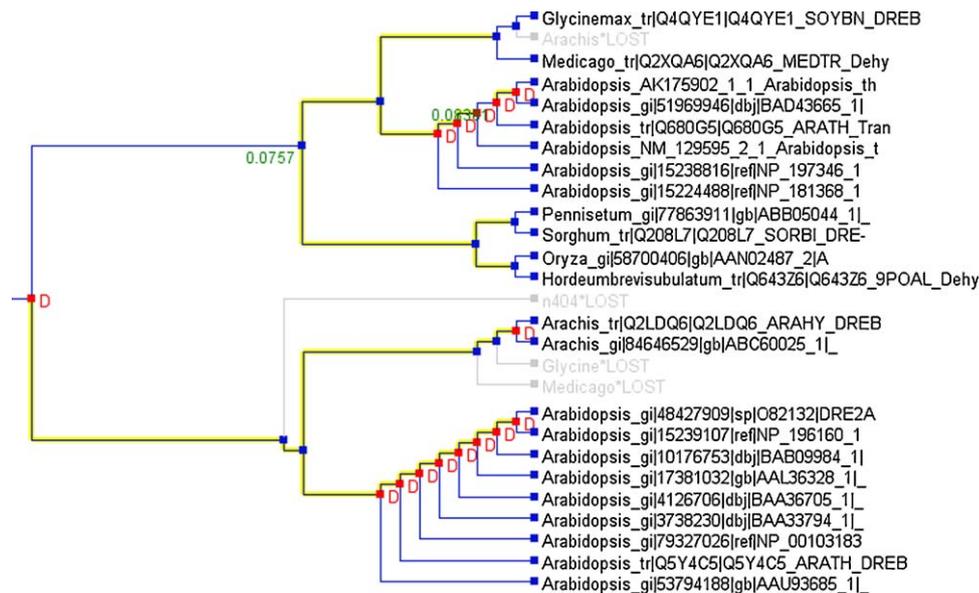


Fig. 2. The reconciled taxonomic (species) tree with DREB2A gene tree. The reconciled tree for DREB2A gene sequences was generated after rooting the species and gene tree to identify duplication (indicated as red nodes-D) or speciation (blue) nodes. DREB2A sequences of cereals and legumes could be grouped in different clades. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)

Bryophyta, as expected, formed an outgroup. As evident from the reconciled tree, the monocot and dicot amino acid sequences of DREB1A separated into distinct groups/clusters. Degenerate primers were designed from the conserved blocks of the orthologous sequences using CODEHOP. This included the sequences from Cluster I (dicots), Cluster IIAi (monocots) and Cluster IIAii (monocots).

Protein sequences of DREB2A (Fig. 2) formed two distinct phylogenetic clusters—Cluster I consisting of DREB2A homologues from dicots and monocot sequences and the other (Cluster II) containing the DREB2 sequences from *Arachis* (Cluster IIA) and DREB2-like sequences from *Arabidopsis* (Cluster IIB). Cluster I was divided into two distinct clades, one containing sequences from dicots (Cluster IA) (that includes *A. thaliana* (**AK175902**), *Glycine max* (**Q4QYE1**), *M. truncatula* (**Q2XQA6**)) and the other containing monocot sequences (Cluster IB) (that includes *Pennisetum glaucum* (**ABB05044**), *S. bicolor* (**Q208L7**), *O. sativa* (**AAN02487**) and *H. brevisubulatum* (**Q643Z6**)). The second dicot Cluster II contained several DREB like sequences from *A. thaliana* (**BAA33794**, **AAU93685**), which were orthologous to *Arachis hypogaea* (**Q2LDQ6**, **ABC60025**). Subsequently, the conserved blocks of the sequences from Cluster IA (dicots), Cluster IB (monocots) and Cluster II (dicot) were considered separately for designing the degenerate primer pairs using CODEHOP in order to amplify DREB2A ortholog in different crop species (Fig ESM 1).

The CODEHOP programme takes into account the codon usage of the target genome in order to design the primers. The degenerate primer sequences are given in ESM Table 1, were checked for amplification across monocot and dicot species targeted in the study. In total 47 degenerate primers for DREB1A and 39 degenerate primers for DREB2A were designed using CODEHOP (ESM Table 1). Around 182 combinations of forward and reverse primers for DREB1A and 113 combinations of DREB2A were tested for amplification. For amplification of DREB1A, 36 primer combinations were tested in case of dicots and 146 combinations for monocots. While for amplification of DREB2A, 47 combinations were tried in monocots and 66 combinations in dicot species. Although 10% of the primer combinations gave amplification across cereals and legumes, only two primer pair combinations could give sequenceable amplification in sorghum and rice.

DREB2A homologue of sorghum was identified by this method. The same primer combinations can be used to isolate DREB2A orthologs from other cereals and millets for which the DREB2A is not yet defined.

The reconciled tree can also help determining the orthologous relationships since it is built by comparing the gene tree (sequence comparison) with the species tree (based on taxonomy) to identify duplication or speciation nodes. Sometimes the gene tree may disagree with the species tree and these differences could be explained by gene duplication in a genome with each copy having its own history. Likewise, the node in the phylogenetic tree can be the result of division of an ancestral species into two others. In the phylogenetic trees so obtained (Figs. 1 and 2), paralogous sequences can be differentiated from orthologs. Two genes are orthologous if the divergence from their last common ancestor results from speciation event while they are paralogous if the divergence results from a duplication event [41]. There is a higher probability for orthologs (homologues that are most closely related through a speciation) to have similar function than for paralogs (homologues that are most closely related through a duplication). It is essential to make the distinction because two paralogous genes are less likely to have preserved the same function as two orthologs [42]. Therefore, the accuracy of gene function prediction by homology between different species depends on whether genes are orthologous or paralogous. Only orthologous sequences were considered for the design of degenerate primers. Using this approach, the DREB2A homologue in sorghum was identified. With appropriate algorithms, it is possible to deduce species history through studying gene sequences.

Species-specific gene primers were designed for DREB2A homologue sequences already available in the public domain such as rice DREB2A (**AF300971**), barley drought responsive factor (**AF521302**), chickpea CAP2 gene (**DQ321719**) and common bean EST sequences resembling the DREB sequences (**CV535836** and **BQ481823**) using Primer3.

3.2. Isolation of DREB2A genes from five crop species

In larger genomes, target genes are often controlled by additional regulators or non-orthologous regulators. The evolution

of major phenotypic differences across organisms may be the effect of transcription factors, which provide regulatory inputs, rather than changes in gene repertoire itself [43]. By using any of two above-mentioned approaches (degenerate primers and species-specific primers), DREB2A homologues were isolated in targeted five crop species by using seven primer pairs mentioned in Table 2. Seven/eight genotypes for the targeted crop species (Table 1), selected from the reference collection, a sub-set of global composite collection defined based on molecular diversity (e.g. [37]) were used for understanding the sequence diversity at DREB2A locus.

The DREB2A homologue (AF300971) from rice is already characterized [9] and the gene-specific primers were used to amplify the corresponding gene from the genomic DNA of seven diverse genotypes of rice. These amplicons were in the range of 1300 bp. Sequence analysis of these amplicons revealed that the sequences of DREB2A from all seven genotypes of rice considerably differed at 5'end from the published sequence (AF300971), although the primers used were taken from the same sequence as mentioned by Dubouzet et al. [9]. The sequence from 1 to 375 bp in the reference sequence does not align with the sequences from all seven genotypes generated in the current study. The multiple sequence alignment of seven sequences revealed the presence of a single nucleotide variant C/A at position 977 of reference sequence (or 644th position from the transcription start site) leading to non-synonymous mutation from glutamate to aspartate amino acid residue. Among seven sequences, three (Azucena, SZH, LTH) had a 'C' variant and the other three (DOM-SOFID, N22, Nipponbare) had an 'A' variant while one genotype (IR 64) had both C/A variants. The PIC value of the SNP was found to be 0.49.

In case of sorghum, the DREB2A homologues were isolated from eight diverse genotypes using degenerate primers (Table 2). The PCR amplicons for eight genotypes were sequenced and these were found identical to recently reported sorghum DREB2 sequence (DQ403725) in NCBI. Further PCR amplicons from the panel of eight diverse genotypes were aligned and analyzed for the presence of SNPs. Only one SNP was observed at 126th position from C to T at the 5' UTR region. The PIC value of this SNP was found to be 0.47 (Table 2).

The DREB2A homologue from barley was isolated using gene-specific primers designed from drought responsive factor (DRF) from barley (AF521302). The DRF sequences from eight diverse genotypes were aligned and multiple sequence alignment of these sequences was used to identify the occurrence of SNPs across the diverse genotypes. As AF521302 sequence was derived from mRNA, 100 bp intronic region is amplified when the genomic DNA from seven diverse genotypes was used as template to amplify corresponding gene. Interestingly, two SNPs were observed in the intronic region of the DRF and a SNP was observed in the exonic region (at 15 bp position) of reference sequence leading to non-synonymous mutation from histidine to glutamine, and contributed five haplotypes (Table 2). SNPs are more prevalent in intronic region than that of exonic regions. In another study Gang-Ping et al. [44] observed that nucleotide polymorphism in the non-coding region in CBF4 locus of Arabidopsis was three times higher than in the coding region and exhibited considerable differences in adaptation to drought tolerance. The two intronic SNPs at positions 13 and 53 bp (calculated according to intronic region) had a PIC of 0.50 and 0.48 respectively, whereas the exonic SNP had a PIC value of 0.49. The average PIC of all the SNPs at barley DRF is found to be 0.49.

In case of chickpea, the CAP2 (DREB2A homologue) promoter and the coding sequence were isolated using the specific primers designed from DQ321719. The sequences obtained in eight diverse genotypes were multiple aligned and analyzed for the presence of SNPs. A single transition nucleotide variant (C/G) was observed at

Table 2
Details of the primers used to isolate DREB2A homologues for crop species.

Species	GenBank accession	Primers	Primer sequences (5'-3')	Length of PCR amplicon (approx)	Sequence read length of contig	Number of SNPs	Average PIC of SNP	Number of haplotypes
Rice	<u>AF300971</u>	Forward Reverse	ATCCGGCCGATGAGCGGGGAGGGGAG GGGGATCTACTCTAATAGGAGAAAAGGCT	1350	1300	1	0.49	3
Barley	<u>AF521302</u>	Forward Reverse	ACATCCTTGCTTGCTT GGTTGCAACTCATAATGC	1300	560	3	0.49	5
Sorghum	Degenerate	Forward Reverse	CGCCGAGACCATCAAGTGrtggaatgarc GAAGCCGTGTCGccytcygrta	850	627	1	0.47	2
Chickpea (CAP2 promoter)	-	Forward	TGTGCTTCAAGTTGCACCTCC	650	580	1	0.5	2
Chickpea (CAP2 gene)	<u>DQ321719</u>	Reverse	ACGGGGTCTTATATACTGCAGA	1000	700	0	0	1
Common bean (DREB2A)	<u>CV535836</u>	Forward Reverse	CGGCTCCCTTCATTCGATCCA AGGCACAAACAAGAATCCA	700	547	8	0.47	2
Common bean (DREB2B)	<u>BQ481823</u>	Forward Reverse	CTAATTCTGCATCTCCCTCAGGTC CAGCTCAGCAGCAGCGTCTACT	650	447	0	0	1

the 124th position in the chickpea CAP2 promoter region and no SNPs were observed in any CAP2 gene (Table 2). The PIC value of the SNP was found to be 0.5. Comparative study of occurrence of SNPs at gene and promoter regions revealed that there are more SNPs in close proximity to transcriptional start sites than in regions further upstream, and the number of SNPs found in the predicted transcription factor binding sites and UTR regions is higher than in non-binding site sequences in eukaryotes including humans [45]. Occurrence of SNP at regulatory region, accounting for the loss of function of seed shattering gene has been already shown in case of rice, which indicates that a single sequence variant can cause a major effect on the function of a gene(s) [46].

Functional regulation of DREB2 type transcription factors seems to be more variable and complicated compared to DREB1/CBF type transcription factors. Expression of intact DREB2A in *Arabidopsis* and rice does not activate downstream genes under normal growth conditions. DREB2A protein requires post-translational modification for its activation [9,25,47]. In contrast, maize ZmDREB2A can function as a transcriptional activator without post-translational modifications [29]. Chickpea DREB2A homologue reported as CAP2 transcription factor is activated by dehydration and high salinity but not by cold, but it is induced by ABA unlike other DREB2A transcription factors. This fact indicates that different or additional pathways may regulate CAP2 as compared to DREB2A in other crops [31].

In the case of common bean, two DREB2 like sequences from EST sequences of common bean were considered for primer design. These were named as putative DREB2A (**CV535836**) and DREB2B (**BQ481823**) based on their expression pattern and comparison to *Arabidopsis* DREB genes (M. Ishitani, personal communication). The isolated DREB sequences were multiple aligned and searched for the presence of SNPs. DREB2A had eight SNPs, belonging to two haplotypes, while DREB2B did not show sequence polymorphism (Table 2). Interestingly, one haplotype corresponds to the cultivars from Andean gene pool (G 19833, AFR 298 and JaloEEP558) and the other to the cultivars from Meso-American (DOR 364, BAT 477, BAT 93, G 21212 and BAT 881) as described by principal component analysis in [48]. The difference in haplotypes agree with variability in morphological trait where common bean cultivars were clustered into two major groups, one having predominantly Meso-American characteristics and the other having Andean characteristics [49]. Several physiological [50] and gene expression [51] studies in common bean reveal that the Andean gene pool were more sensitive to drought stress compared to Meso-American gene pool. The PIC value of all the eight SNPs was 0.47, which corresponds to the average PIC value of SNPs.

The PIC values for identified SNPs in isolated/identified genes for the DREB2A ranged from 0.47 (common bean DREB2A and sorghum DREB2A) to 0.50 (chickpea CAP2 promoter) (Table 2). Based on the PIC value, the SNP locus of CAP2 promoter can be said to be more informative compared to other genes. Less number of SNPs at each locus indicates the conservation of DREB sequence within the crop species.

3.3. Phylogenetic analysis of the sequences

To study DREB2A phylogenetic relationships, the sequences were used to construct a phylogenetic tree using UPGMA based parsimony algorithm in the MEGA 4.1 programme. The level of confidence was estimated using bootstrap of 500 replications. The dendrogram (Fig. 3) shows a clear separation of monocot DREB2 sequences (Cluster IA) from that of dicots, however the CAP2 sequences clustered with the monocots (Cluster IB). The common bean DREB2 sequences clustered separately from other crop sequences (Cluster II). The tree shows a distinct crop-specific

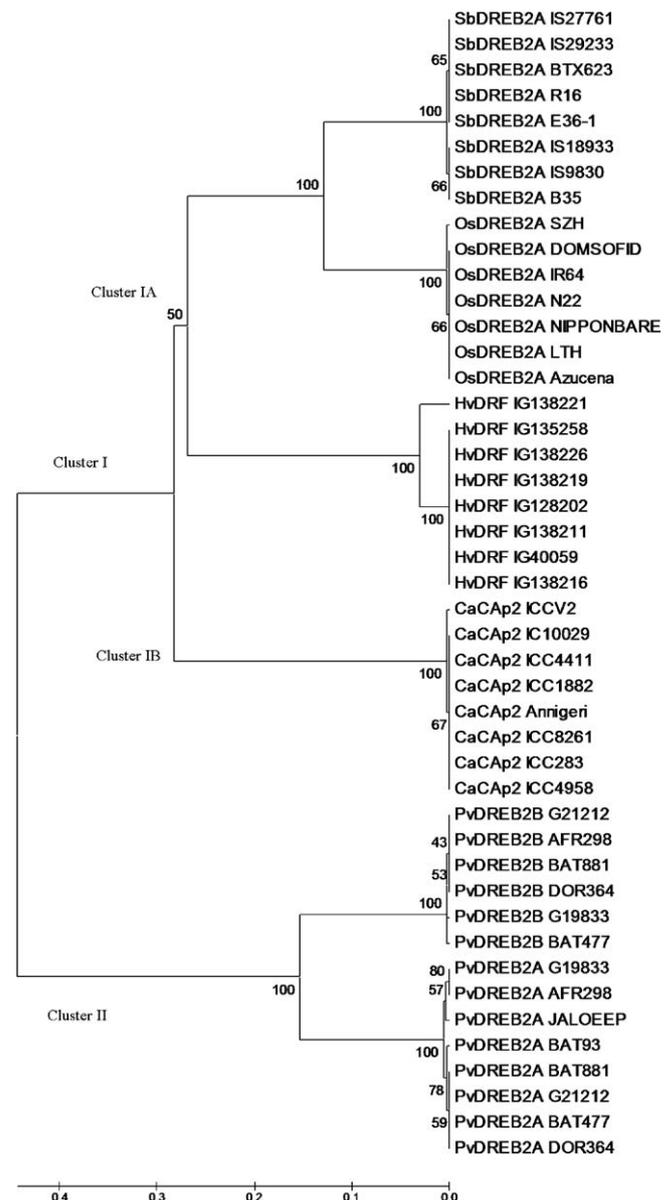


Fig. 3. Phylogenetic analysis of DREB2 homologues from five crop species. The phylogenetic analysis showed clear demarcation of DREB2 homologues into two clusters, Cluster I and Cluster II. Cluster I has two subgroups, Cluster IA containing the DREB2 homologues of three cereals (sorghum, rice and barley) and IB that of chickpea CAP2, whereas Cluster II contains the DREB2 homologues from common bean.

clustering of sequences, indicating clear crop-specific sequence differences.

3.4. Conservation of AP2 domain

The DREB proteins contain ERF/AP2 DNA binding domain unique to plant species and are quite conserved across plant species. In order to assess the conservation of AP2 domain the translated DREB2A nucleotide sequences obtained from seven/eight diverse genotypes along with model crops sequences from *Arabidopsis*, *Medicago* and soybean were aligned to look at the conservation of AP2 domain (Fig. 4). Amino acid alignment of different DREB2 proteins shows high sequence similarity at ERF/AP2 domain, the two amino acids, 14th valine (V) and 19th glutamic acid (E) (indicated as asterisk in Fig. 4) playing a crucial

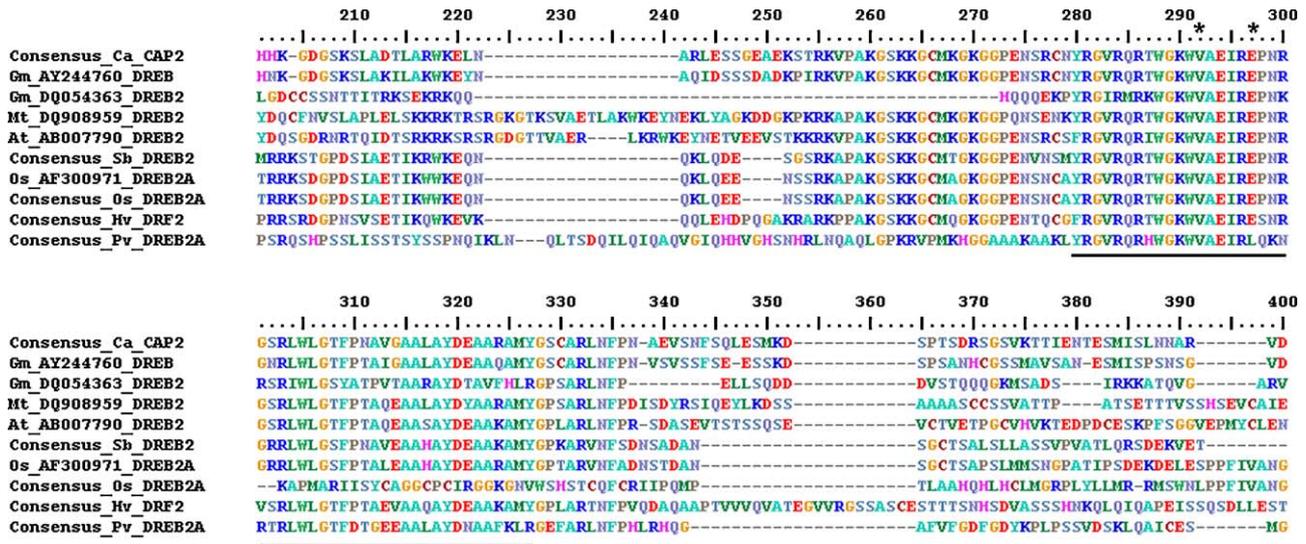


Fig. 4. Comparison of amino acid alignment of DREB2A proteins with that of model species. Part of amino acid sequences (200–400 bp) containing AP2 domain of consensus DREB2A homologues of rice, sorghum, barley, chickpea and common bean are aligned with the model DREB2A from *Arabidopsis* (AB007790), rice (AF300971), *Medicago* (DQ908959) and soybean (AY244760). The ERF/AP2 domain is underlined and the asterisks indicate conserved valine and glutamic acid among DREB proteins.

role in the determination of DNA-binding specificity [25,52]. The amino acid sequences obtained from consensus DREB2A homologue from rice, sorghum, barley, chickpea and common bean indicated that the sequence identity extended throughout the protein, but was greatest at AP2/ERF domain.

4. Conclusion

In the present study, the DREB2A homologues from five crop species were isolated and compared at gene level. It is evident that, DREB2 transcription factor is more conserved within the species, as the SNP polymorphism is comparatively less. The AP2/ERF domain is conserved across the dicot and monocot species. Phylogenetic analysis also showed the clear demarcation of monocot DREBs from those of common bean DREB2 homologues. The SNP in the CAP2 promoter was found to have highest PIC value as compared to other DREB2 homologues. Furthermore, the identified SNPs may be used in linkage mapping to identify, if exist, QTLs associated with DREB2 homologues and/or in candidate gene sequencing based association mapping.

Acknowledgments

Financial support from Generation Challenge Programme (GCP) to the ADOC (Allelic Diversity at Orthologous Candidate Genes in GCP Crops) project is gratefully acknowledged. SNN is thankful to Council of Scientific and Industrial Research (CSIR) for sponsoring fellowship.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.plantsci.2009.07.009.

References

[1] M. Thomashow, Plant cold acclimation: freezing tolerance genes and regulatory mechanisms, *Annu. Rev. Plant Physiol. Plant Mol. Biol.* 50 (1999) 571–599.
 [2] J.K. Zhu, Salt and drought stress signal transduction in plants, *Annu. Rev. Plant Biol.* 53 (2002) 247–273.
 [3] D. Bartels, R. Sunkar, Drought and salt tolerance in plants, *Crit. Rev. Plant Sci.* 24 (2005) 23–58.

[4] H.J. Bohnert, Q. Gong, P. Li, S. Ma, Unravelling abiotic stress tolerance mechanisms—getting genomics going, *Curr. Opin. Plant Biol.* 9 (2006) 180–188.
 [5] K. Yamaguchi-Shinozaki, K. Shinozaki, Transcriptional regulatory networks in cellular responses and tolerance to dehydration and cold stresses, *Annu. Rev. Plant Biol.* 57 (2006) 781–803.
 [6] K. Yamaguchi-Shinozaki, K. Shinozaki, A novel cis-acting element in an *Arabidopsis* gene is involved in responsiveness to drought, low temperature, or high-salt stress, *Plant Cell* 6 (1994) 251–264.
 [7] S.S. Baker, K.S. Wilhelm, M.F. Thomashow, The 5'-region of *Arabidopsis thaliana* cor15a has cis-acting elements that confer cold-, drought- and ABA-regulated gene expression, *Plant Mol. Biol.* 24 (1994) 701–713.
 [8] C. Jiang, B. Lu, J. Singh, Requirement of a CCGAC cis-acting element for cold induction of the BN115 gene from winter *Brassica napus*, *Plant Mol. Biol.* 30 (1996) 679–684.
 [9] J.G. Dubouzet, Y. Sakuma, Y. Ito, M. Kasuga, E.D. Dubouzet, S. Miura, M. Seki, K. Shinozaki, K. Yamaguchi-Shinozaki, OsDREB genes in rice, *Oryza sativa* L., encoded transcription activators that function in drought-, high salt- and cold-responsive gene expression, *Plant J.* 33 (2003) 751–763.
 [10] P.K. Busk, A.B. Jensen, M. Pages, Regulatory elements *in vivo* in the promoter of the abscisic acid responsive gene *rab17* from maize, *Plant J.* 11 (1997) 1285–1295.
 [11] P.K. Busk, M. Pages, Regulation of abscisic acid induced transcription, *Plant Mol. Biol.* 37 (1998) 425–435.
 [12] V. Haake, D. Cook, J.L. Riechmann, O. Pineda, M.F. Thomashow, J.Z. Zhang, Transcription factor CBF4 is a regulator of drought adaptation in *Arabidopsis*, *Plant Physiol.* 130 (2002) 639–648.
 [13] D. Kizis, M. Pages, Maize DRE-binding proteins DBF1 and DBF2 are involved in *rab17* regulation through the drought-responsive element in an ABA-dependent pathway, *Plant J.* 30 (2002) 679–689.
 [14] K. Shinozaki, K. Yamaguchi-Shinozaki, Gene expression and signal transduction in water-stress response, *Plant Physiol.* 115 (1997) 327–334.
 [15] S. Fowler, M.F. Thomashow, *Arabidopsis* transcriptome profiling indicates that multiple regulatory pathways are activated during cold acclimation in addition to the CBF cold response pathway, *Plant Cell* 14 (2002) 1675–1690.
 [16] Y. Ito, K. Katsura, K. Maruyama, T. Taji, M. Kobayashi, M. Seki, K. Shinozaki, K. Yamaguchi-Shinozaki, Functional analysis of rice DREB1/CBF-type transcription factors involved in cold-responsive gene expression in transgenic rice, *Plant Cell Physiol.* 47 (2006) 141–153.
 [17] K. Maruyama, Y. Sakuma, M. Kasuga, Y. Ito, M. Seki, H. Goda, Y. Shimada, S. Yoshida, K. Shinozaki, K. Yamaguchi-Shinozaki, Identification of cold-inducible downstream genes of the *Arabidopsis* DREB/CBF3 transcriptional factor using two microarray systems, *Plant J.* 38 (2004) 982–993.
 [18] F. Qin, Y. Sakuma, J. Li, Q. Liu, Y.Q. Li, K. Shinozaki, K. Yamaguchi-Shinozaki, Cloning and functional analysis of a novel DREB1/CBF transcription factor involved in cold-responsive gene expression in *Zea mays* L., *Plant Cell Physiol.* 45 (2004) 1042–1052.
 [19] D.W. Choi, E.M. Rodriguez, T.J. Close, Barley *cbf3* gene identification, expression pattern and map location, *Plant Physiol.* 129 (2002) 1781–1787.
 [20] J.S. Skinner, J.V. Zitzewitz, P. Szucs, L. Marquez-Cedillo, T. Filichkin, K. Amundsen, E.J. Stockinger, M.F. Thomashow, T.H.H. Chen, P.M. Hayes, Structural, functional and phylogenetic characterization of a large CBF gene family in barley, *Plant Mol. Biol.* 59 (2005) 533–551.
 [21] G.P. Xue, Characterisation of the DNA-binding profile of barley HvCBF1 using an enzymatic method for rapid, quantitative and high-throughput analysis of the DNA-binding activity, *Nucleic Acids Res.* 30 (2002) e77.

- [22] M. Badawi, J. Danyluk, B. Boucho, M. Houde, F. Sarhan, The CBF gene family in hexaploid wheat and its relationship to the phylogenetic complexity of cereal CBFs, *Mol. Genet. Genomics* 277 (2007) 533–554.
- [23] X.P. Li, A.G. Tian, G.Z. Luo, Z.Z. Gong, J.S. Zhang, S.Y. Chen, Soybean DRE-binding transcription factors that are responsive to abiotic stresses, *Theor. Appl. Genet.* 110 (2005) 1355–1362.
- [24] T.J. Zhao, S. Sun, Y. Liu, J.M. Liu, Q. Liu, Y.B. Yan, H.M. Zhou, Regulating the drought-responsive element (DRE)-mediated signaling pathway by synergic functions of trans-active and trans-inactive DRE binding factors in *Brassica napus*, *J. Biol. Chem.* 281 (2006) 10752–10759.
- [25] Q. Liu, M. Kasuga, Y. Sakuma, H. Abe, S. Miura, K. Yamaguchi-Shinozaki, K. Shinozaki, 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 (1998) 1391–1406.
- [26] K. Nakashima, Z.K. Shinwari, Y. Sakuma, M. Seki, S. Miura, K. Shinozaki, K. Yamaguchi-Shinozaki, Organization and expression of two *Arabidopsis* DREB2 genes encoding DRE-binding proteins involved in dehydration- and high-salinity-responsive gene expression, *Plant Mol. Biol.* 42 (2000) 657–665.
- [27] Y. Sakuma, Q. Liu, J.G. Dubouzet, H. Abe, K. Shinozaki, K. Yamaguchi-Shinozaki, DNA-binding specificity of the ERF/AP2 domain of *Arabidopsis* DREBs, transcription factors involved in dehydration- and cold-inducible gene expression, *Biochem. Biophys. Res. Commun.* 290 (2002) 998–1009.
- [28] C. Egawa, F. Kobayashi, M. Ishibashi, T. Nakamura, C. Nakamura, S. Takumi, Differential regulation of transcript accumulation and alternative splicing of a DREB2 homolog under abiotic stress conditions in common wheat, *Gene Genet. Syst.* 81 (2006) 77–91.
- [29] F. Qin, M. Kakimoto, Y. Sakuma, K. Maruyama, Y. Osakabe, L.-S. Phan Tran, K. Shinozaki, K. Yamaguchi-Shinozaki, Regulation and functional analysis of ZmDREB2A in response to drought and heat stresses in *Zea mays* L., *Plant J.* 50 (2007) 54–69.
- [30] G.P. Xue, C.W. Loveridge, HvDRF1 is involved in abscisic acid-mediated gene regulation in barley and produces two forms of AP2 transcriptional activators, interacting preferably with a CT-rich element, *Plant J.* 37 (2004) 326–339.
- [31] R.K. Shukla, S. Raha, V. Tripathi, D. Chattopadhyay, Expression of CAP2, an AP2-family transcription factor from chickpea enhances growth and tolerance to dehydration and salt stress in transgenic tobacco, *Plant Physiol.* 142 (2006) 113–123.
- [32] R.K. Varshney, D.A. Hoisington, A.K. Tyagi, Advances in cereal genomics and applications in crop breeding, *Trends Biotechnol.* 24 (2006) 490–499.
- [33] E.S. Buckler, J.M. Thornsberry, Plant molecular diversity and applications to genomics, *Curr. Opin. Plant Biol.* 5 (2002) 107–111.
- [34] R.K. Varshney, A. Graner, M.E. Sorrells, Genomics-assisted breeding for crop improvement, *Trends Plant Sci.* 10 (2005) 621–630.
- [35] A. Rafalski, Novel genetic mapping tools in plants: SNPs and LD-based approaches, *Plant Sci.* 162 (2002) 329–333.
- [36] A. Rafalski, M. Morgante, Corn and humans: recombination and linkage disequilibrium in two genomes of similar size, *Trends Genet.* 20 (2004) 103–111.
- [37] H.D. Upadhyaya, S.L. Dwivedi, M. Baum, R.K. Varshney, S.M. Udupa, C.L.L. Gowda, D. Hoisington, S. Singh, Genetic structure, diversity and allelic richness in composite collection and reference set in chickpea (*Cicer arietinum* L.), *BMC Plant Biol.* 8 (2008) 106.
- [38] K. Chen, D. Durand, M. Farach-Colton, NOTUNG: a program for dating gene duplications and optimizing gene family trees, *J. Comput. Biol.* 7 (2000) 429–447.
- [39] T. Rose, E. Schultz, J. Henikoff, S. Pietrokovski, C. McCallum, S. Henikoff, Consensus-degenerate hybrid oligonucleotide primers for amplification of distantly related sequences, *Nucleic Acids. Res.* 26 (1998) 1628–1635.
- [40] M. Morant, A. Hehn, D. Werck-Reichhart, Conservation and diversity of gene families explored using the CODEHOP strategy in higher plants, *Plant Biol.* 2 (2002) 7–12.
- [41] J.F. Dufayard, L. Duret, F. Reichenmann, Phylogenetic tree reconciliation: the species/Gene tree problem, *ERCIM News* 43 (2000) 15.
- [42] E.V. Koonin, Orthologs, paralogs, and evolutionary genomics, *Ann. Rev. Gen.* 39 (2005) 309–338.
- [43] M.M. Babu, S.A. Teichmann, L. Aravind, Evolutionary dynamics of prokaryotic transcriptional regulatory networks, *J. Mol. Biol.* 358 (2006) 614–633.
- [44] H. Gang-Ping, W. Zhong-Yi, C. Mao-Sheng, C. Ming-Qing, D. Brunel, G. Pelletier, H. Cong-Lin, Y. Qing, Single nucleotide polymorphisms of CBF4 locus region of *Arabidopsis thaliana* correspond to drought tolerance, *Chin. J. Agric. Biotechnol.* 1 (2004) 181–190.
- [45] Y. Guo, C.D. Jamison, The distribution of SNPs in human gene regulatory regions, *BMC Genomics* 6 (2005) 140.
- [46] S. Konishi, T. Izawa, S.Y. Lin, K. Ebana, Y. Fukuta, T. Sasaki, M. Yano, An SNP caused loss of seed shattering during rice domestication, *Science* 312 (2006) 1392–1396.
- [47] Y. Sakuma, K. Maruyama, K. Osakabe, F. Qin, M. Seki, K. Shinozaki, K. Yamaguchi-Shinozaki, Functional analysis of an *Arabidopsis* transcription factor, DREB2A, involved in drought responsive gene expression, *Plant Cell* 18 (2006) 1292–1309.
- [48] M.W. Blair, M.C. Giraldo, H.F. Buendía, E. Tovar, M.C. Duque, S.E. Beebe, Microsatellite marker diversity in common bean (*Phaseolus vulgaris* L.), *Theor. Appl. Genet.* 113 (2006) 100–109.
- [49] L. Duran, M.W. Blair, M.C. Giraldo, R. Macchiavelli, E. Prophete, J.C. Nin, J.S. Beaver, Morphological and molecular characterization of common bean landraces and cultivars from the Caribbean, *Crop Sci.* 45 (2005) 1320–1328.
- [50] C.G. Munoz-Perea, R. Allen, D.T. Westermann, J.L. Wright, S.P. Singh, Water use efficiency among dry bean landraces and cultivars in drought-stressed and non-stressed environments, *Euphytica* 155 (2007) 393–402.
- [51] T. Kavar, M. Maras, M. Kidric, J. Sustar-Vozlic, V. Meglic, Identification of genes involved in the response of leaves of *Phaseolus vulgaris* to drought stress, *Mol. Breeding* 21 (2008) 159–172.
- [52] P.K. Agarwal, P. Agarwal, M.K. Reddy, S.K. Sopory, Role of DREB transcription factors in abiotic and biotic stress tolerance in plants, *Plant Cell Rep.* 21 (2006) 1263–1274.