



# *DREB2* (dehydration-responsive element-binding protein 2) type transcription factor in sorghum (*Sorghum bicolor*): genome-wide identification, characterization and expression profiles under cadmium and salt stresses

M. Aydın Akbudak<sup>1</sup> · Ertugrul Filiz<sup>2</sup> · Kubra Kontbay<sup>1</sup>

Received: 25 July 2018 / Accepted: 24 September 2018 / Published online: 28 September 2018  
© Springer-Verlag GmbH Germany, part of Springer Nature 2018

## Abstract

Biotic and abiotic stresses negatively affect fitness, biomass production, and crop yield in plants. The dehydration-responsive element-binding proteins (*DREB*) are important transcription factors (TFs), and are induced by abiotic and biotic stresses. In this study, genome-wide identification, in silico sequence, and phylogenetic analyses and expression analyses of *DREB2* genes under cadmium (Cd) and salt (NaCl) stresses in sorghum (*Sorghum bicolor*, *Sb*) were performed. Six putative *SbDREB2* genes were identified in sorghum genome and all contained AP2 domain (PF00847). Nucleotide diversities in *SbDREB2* genes were calculated as  $\pi$ : 0.53 and  $\theta$ : 0.39, respectively. While exon numbers of them were either one or two, length of *SbDREB2* proteins ranged from 238 to 388 amino acid residues. Fifty-six *cis*-acting regulatory elements, which are tissue specific, light, hormone, and stress responsive, were identified in the promotor regions of *SbDREB2* genes. Analyses on digital expression data indicated that *SbDREB2A* and *SbDREB2B* are more expressed genes than other *SbDREB* genes in sorghum. Under Cd and NaCl stresses, expressions of *SbDREB2* genes were induced at different levels. All *SbDREB2* genes in root were up-regulated under salt stress. In case of Cd stress, *SbDREB2D* gene was particularly up-regulated in leaves and roots. Co-expression analyses revealed four of TFs in co-expression network, indicating that they have roles in transcriptional cascade. Furthermore, five miRNA target regions were identified for four *SbDREB2* genes, indicating their roles in post-transcriptional regulation. The predicted 3D structure of *SbDREB2* proteins showed some structural divergences and structure overlap between rice and sorghum varied at between 26.58 and 50%. Finally, obtained data could be used in breeding of stress-tolerant plants, particularly genetically engineered *DREB2* expressing plants. Findings in this study would also contribute to the understanding of *DREB2* genes in plants, especially in sorghum.

**Keywords** Sorghum · Abiotic stress · Gene expression · Genome-wide analyses

## Introduction

Plant growth is adversely affected by various abiotic stresses including high temperature, drought, salinity, and low temperature (Sakuma et al. 2006a). Many stress-inducible genes

have been determined in plants, and they can be divided into two groups: stress tolerance proteins and regulatory proteins (Khan et al. 2017). Transcription factors (TFs) are essential for regulating expression of functional protein genes in the genome. *AP2/EREBP* (*APETALA2*/ethylene-responsive element-binding protein) represents a large family of TF genes in plants and contains highly conserved AP2/ERF DNA-binding domain (Riechmann and Meyerowitz 1998; Zhou et al. 2010; Sharoni et al. 2011; Yan et al. 2013; Vatansever et al. 2017). *AP2/EREBP* gene family can be divided into four subfamilies, including AP2, RAV (related to ABI3/VP1), dehydration-responsive element-binding protein (DREB), and ERF (Sakuma et al. 2002). The AP2/EREBP supergene family is characterized by the AP2/ERF domain

✉ M. Aydın Akbudak  
akbudak@akdeniz.edu.tr

✉ Ertugrul Filiz  
ertugrulfiliz@gmail.com

<sup>1</sup> Department of Agricultural Biotechnology, Akdeniz University, Antalya, Turkey

<sup>2</sup> Department of Crop and Animal Production, Cilimli Vocational School, Duzce University, Cilimli, Duzce, Turkey

consisting of 50–60 amino acids, and these proteins play important roles in various regulatory mechanisms in plant metabolism (Liu et al. 1998). Ethylene-responsive factor (ERF) proteins are one of the transcription factors involved in biotic and abiotic stress responses. The *DREBs* (dehydration-responsive element binding) are members of *ERF* family of transcription factors and belonged to ABA-independent signal transduction pathway. There are two subclasses of *DREBs*, *DREB1/CBF* and *DREB2*, which are induced by cold and dehydration, respectively, and there is a cross-talk between them (Agarwal et al. 2006; Lata and Prasad 2011; Mizoi et al. 2012; Vatansever et al. 2017). *DREB* genes can be divided into six small groups named as A-1–A-6 based on their sequence similarity, which play multiple roles in plants (Liu et al. 1998; Sakuma et al. 2002). Dehydration-responsive element (DRE)/C-repeat (CRT, A/GCCGAC) core sequence is presented in many promoters involved in environmental stress response (Yamaguchi-Shinozaki and Shinozaki 1994; Stockinger et al. 1997; Sakuma et al. 2006b). The *DREB2* subgroup is composed of eight members in *Arabidopsis* (Sakuma et al. 2002) and five in rice (Matsukura et al. 2010). In *Arabidopsis*, *DREB2A* has various roles in transcriptional cascade of heat-shock responses. *DREB2A* is rapidly induced by heat shock, which results in the expression of many heat shock-responsive gene-encoding transcription factors and molecular chaperones (Sakuma et al. 2006b). In rice, transgenic expression of sorghum *DREB2* gene improves tolerance and yield under water limitation. Particularly, drought-stressed *rd29A::SbDREB2* transgenic plants indicated a significantly higher number of panicles compared to wild-type rice plants (Bihani et al. 2011). *ZmDREB2A* was induced by cold, dehydration, salt, and heat stresses in maize seedlings. In addition, overexpression of *ZmDREB2A* improved thermo-tolerance in transgenic *Arabidopsis* plants (Qin et al. 2007). Mizoi et al. (2013) reported that a new soybean *DREB2* gene, *GmDREB2A;2*, was highly expressed not only under the conditions of dehydration and heat but also low temperature. In another study, transcription of *DREB2C* was induced in *Arabidopsis* by a superoxide anion propagator, methyl viologen (MV). The *DREB2C*-overexpressing transgenic plants showed more oxidative stress tolerance than wild-type plants (Hwang et al. 2012). In tomato, *SIDREB2* was up-regulated in roots and young leaves under NaCl exposure. In addition, it was shown that *SIDREB2* was induced by KCl and drought (Hichri et al. 2016). Agarwal et al. (2010) reported that transgenic tobacco plants transformed with *Pennisetum glaucum DREB2A* gene (*PgDREB2A*) showed the abiotic stress tolerance and induced higher expression of downstream stress-responsive genes. This study reveals genome-wide identification and comparison of *DREB2* genes in sorghum genome with *Arabidopsis* and rice genomes as well as their expression profiles under Cd and NaCl stresses.

## Materials and methods

### Identifications of *SbDREB2* genes/proteins

Amino acid sequences of eight *Arabidopsis* *DREB2* proteins—*DREB2A* (AT5G05410.1), 2B (AT3G11020.1), 2C (AT2G40340.1), 2D (AT1G75490.1), 2E (AT2G38340.1), 2F (AT3G57600.1), 2G (AT5G18450.1), and 2H (AT2G40350.1)—and five rice *DREB2* proteins—*DREB2A* (LOC\_Os01g07120.1), 2B (LOC\_Os05g27930.1), 2C (LOC\_Os08g45110.1), 2D (LOC\_Os05g39590.1), and 2E (LOC\_Os03g07830.1)—were obtained from UniProt (<http://www.uniprot.org/>) database. These amino acid sequences were queried in sorghum genome in Phytozome database with  $E^{-20}$  value (<http://phytozome.jgi.doe.gov/pz/portal.html>) (Goodstein et al. 2012). Later, Hidden Markov Model (HMM) was performed in putative *SbDREB2* protein sequences by Pfam (<http://pfam.sanger.ac.uk>) (Punta et al. 2012) to confirm the presence of AP2 domain (PF00847). In addition, coding sequence (CDS) and peptide sequences of identified *DREB2* genes were collected from Phytozome database for further analyses.

### Sequence analyses of *DREB2* genes/proteins

First, physico-chemical properties of *DREB2* proteins were analyzed using ProtParam tool (<https://web.expasy.org/protparam/>) (Gasteiger et al. 2005). Sub-cellular localizations were predicted by WoLF PSORT (<https://wolfsort.hgc.jp/>) (Horton et al. 2007). Multiple sequence alignment and amino acid composition of *DREB2* proteins were performed by Bioedit v7.2.5 (Hall 1999). The pairwise distance analyses were evaluated by the MEGA7 software (Kumar et al. 2016). In addition, the coding sequences of *DREB2* genes were analyzed using DnaSP 5.1 (Librado and Rozas 2009), including the estimates of genetic diversity such as  $\pi$  (Tajima 1989) and  $\theta$  (Watterson 1975), and segregating (polymorphic) sites (S). Exon/intron organization of *DREB2* genes was obtained from Phytozome database.

### Annotations, conserved motif, and phylogenetic analyses

The ontology annotations of *DREB2* genes were made using PhytoMine in Phytozome (<https://phytozome.jgi.doe.gov/phytomine/begin.do>) (Goodstein et al. 2012) and AmiGO (<http://amigo.geneontology.org/amigo/landing>) (Carbon et al. 2009) databases. Conserved motifs in *DREB2* protein sequences were predicted using MEME (Multiple Em for Motif Elicitation) suites (<http://meme-suite.org/>) with parameters; maximum motif number to find: 5 and

minimum–maximum motif width to find: 6–50 (Bailey et al. 2015). A phylogenetic tree was constructed by the MEGA7 software using with 1000 bootstraps (Kumar et al. 2016). The evolutionary history was inferred using the maximum-likelihood method based on the JTT matrix-based model (Jones et al. 1992). All positions containing gaps and missing data were eliminated. There were a total of 78 positions in the final data set.

### Promotor sequence, digital expression, and miRNA target analyses of *SbDREB2* genes

For promoter analysis, the 1000 bp upstream regions of *SbDREB2* genes were retrieved from Phytozome database and searched in PlantCARE database (<http://bioinformatics.psb.ugent.be/webtools/plantcare/html/>) (Lescot et al. 2002). The digital expression data of *DREB2* genes at anatomical and developmental levels in sorghum were obtained from the Genevestigator database for plant biology ([https://genevestigator.com/gv/doc/intro\\_plant.jsp](https://genevestigator.com/gv/doc/intro_plant.jsp)) (Hruz et al. 2008). In addition, expression analyses of putative *SbDREB2* genes were analyzed by MOROKOSHI transcriptome database (<http://sorghum.riken.jp/morokoshi/Home.html>) based on approximately 40,000 full-length cDNA (FL-cDNAs) and expression profiles from RNA-Seq analysis (Makita et al. 2015). The co-expression network of *SbDREB2A* gene was generated by MOROKOSHI transcriptome database. To identify putative *SbDREB2*-targeted miRNAs, *SbDREB2*-coding sequences were searched against the published sorghum miRNAs in psRNATarget database (<http://plantgrn.noble.org/psRNATarget/analysis?function=2>) (Dai and Zhao 2011).

### Secondary and tertiary structure analysis

First, secondary structures of *SbDREB2* proteins were evaluated using the SOPMA server ([https://npsa-prabi.ibcp.fr/cgi-bin/npsa\\_automat.pl?page=/NPSA/npsa\\_sopma.html](https://npsa-prabi.ibcp.fr/cgi-bin/npsa_automat.pl?page=/NPSA/npsa_sopma.html)) (Geourjon and Deleage 1995). 3D models of them were predicted at intensive mode using the Phyre2 server (<http://sbj.bio.ic.ac.uk/phyre2/>) (Kelley and Sternberg 2009). Model

quality was checked using Ramachandran plot analysis by Vadar 1.8 server (<http://vadar.wishartlab.com/>) (Willard et al. 2003).

### Plant material and growth condition

Sorghum (*S. bicolor* L. cv. Aldari) seeds were individually planted in 10-cm square-top pots containing perlite or peat pellet, and grown at 28 °C under LD 12:12 regime (170  $\mu\text{mol m}^{-2} \text{s}^{-1}$ —12 h light period/day), and 70% relative humidity as described previously (Akbadak et al. 2018). Plants were watered with 30 ml half-strength Hoagland solution every other day. On the 14th day of plant growth, plants were watered with 30 ml Hoagland solution-containing 200  $\mu\text{M CdCl}_2$ . No  $\text{CdCl}_2$  was added into Hoagland solution used for watering control plants (Panda et al. 2011). After 6 h and 24 h of  $\text{CdCl}_2$  treatments, leaves and roots of plants were separately harvested for RNA isolation. To induce salinity stress, 2-week-old plants were watered with half strength Hoagland solution-containing supplemental sodium chloride (NaCl). Salt concentration was increased gradually (50 mM, 100 mM, 150 mM, and 200 mM) every other day intervals to avoid osmotic shock (Swami et al. 2011). After 24 h treatment with 200 mM NaCl, leaves and roots were separately harvested for RNA isolation.

### RNA extraction and gene expression analyses

RNA was separately isolated from leaves and roots of each plant using RNA Plant Mini Kit (Qiagen, Cat no.: 74904) following the manufacturer's instructions. RNA amounts in the samples were determined with BioDrop  $\mu\text{LITE}$  (Biodrop, UK). RNA samples were treated with RQ1 RNase-Free DNase (Promega, WI, USA). The intactness of the RNA and its contamination by DNA were checked by gel electrophoresis. Real-time quantitative PCR (RT-qPCR) was carried out on a Light Cycler 96 System (Roche). Expressions of putative *DREB2* genes were quantified in 10 ng DNAase-treated RNA samples using EvoScript RNA SYBR Green I Master Kit (One-step) (Roche). The forward and reverse primers (Table 1) designed for *SbDREB2* genes were used in

**Table 1** Primers used in RT-qPCR analysis of *SbDREB2* genes

Genes	Forward primer sequences	Reverse primer sequences	Size (bp)
<i>SbDREB2A</i>	5'-AAGGCACGTGTCAACTTCTC-3'	5'-TAGATGCCAGCAACGAAAGC-3'	79
<i>SbDREB2B</i>	5'-CAGCCCGGAAGCAGGAAGAA-3'	5'-CTGCCTCACTCCACGATAACC-3'	222
<i>SbDREB2C1</i>	5'-GATGATGCAGTACTCCGCCT-3'	5'-CATCTGGAACGTCTGGTGGT-3'	151
<i>SbDREB2C2</i>	5'-TCCGGCTCCTTTTCCAACACTAC-3'	5'-GTGGAATGTCTGGTGGTTGTTG-3'	99
<i>SbDREB2D</i>	5'-GATGCGGTATCCGACATGGT-3'	5'-GGCGTCAATACCTGATGCT-3'	75
<i>SbDREB2E</i>	5'-TCAGGGAGTTCTTGCAGCAG-3'	5'-CTGAAAACCTCCGTTGACGC-3'	209
<i>SbPPP2A</i>	5'-AACCCGCAAAACCCAGACTA-3'	5'-TACAGGTCGGGCTCATGGAAC-3'	138

RT-qPCR analysis. Gene expression was determined using the  $\Delta\Delta C_T$  method (Livak and Schmittgen 2001). *S. bicolor* serine/threonine-protein phosphatase (*SbPPP2A*) gene served as a reference gene in the analysis (Reddy et al. 2016). For gene expression analysis, the  $C_T$  values of reference genes (*PPP2A*) were subtracted from the  $C_T$  of the target genes to normalize each sample. For example:

$$\Delta C_T = C_T(\text{Target Gene}) - C_T(\text{SbPPP2A}).$$

$\Delta\Delta C_T$  was obtained by comparing the  $\Delta C_T$  value of Cd/NaCl-stressed plants to control plants:

$$\Delta\Delta C_T = \Delta C_T(\text{Cd/NaCl-stressed plant}) - \Delta C_T(\text{Control}).$$

The fold change for each line based on its corresponding reference (Control) was calculated using the following equation:

$$\text{Fold Change} = 2^{-(\Delta\Delta C_T)}.$$

Average  $C_T$  values for leaf and root tissues were obtained from at least three biological and three technical replicates for each gene.

## Results and discussion

### Nucleotide and protein sequence variations of *DREB2* genes/proteins

To identify *DREB2* genes in the sorghum genome, BLASTP analyses were performed using *Arabidopsis* (eight members) and rice (five members) orthologues obtained from UniProt database. As a result, we identified six putative *SbDREB2* genes and named them based on phylogenetic relationships to rice and MOROKOSHI transcriptome database (Table 2). In addition, to get more insights about *DREB2* genes, *Arabidopsis* and rice orthologues were subjected to comparative analyses (Table 2). Gene structure analyses revealed that exon numbers of *DREB2* genes were between one and two in sorghum. The lengths of the coding sequences were ranged from 717 bp (*SbDREB2D*) to 1167 bp (*SbDREB2C1*). Furthermore, G+C (guanine + cytosine) content was found between 53.91 and 72.13%, and overall pairwise average distance was identified as 0.796, suggesting that a genetic variation was between *DREB2* genes in sorghum. Variable (polymorphic) sites, ~61% (364/593) of which were parsimony informative sites and ~39% (229/593) of which were singleton variable sites, were also identified in *SbDREB2* genes. In addition, 63 sites were identified as invariable

**Table 2** Details of *DREB2* genes in sorghum, rice, and *Arabidopsis*

Gene ID (phytozome)	Uniprot ID	Gene name	Protein domain family	Exon no.	Protein length (aa)	MW (kDa)	<i>pI</i> value	Localization (WoLF PSORT)
AT5G05410.1	O82132	AtDREB2A	PF00847	1	335	37.70	5.17	Nucleus
AT3G11020.1	O82133	AtDREB2B	PF00847	1	330	37.12	5.02	Nucleus
AT2G40340.1	Q8LFR2	AtDREB2C	PF00847	2	341	37.83	4.67	Nucleus
AT1G75490.1	Q9LQZ2	AtDREB2D	PF00847	1	206	22.58	6.10	Nucleus
AT2G38340.1	O80917	AtDREB2E	PF00847	1	244	27.37	8.39	Nucleus
AT3G57600.1	Q9SVX5	AtDREB2F	PF00847	1	277	31.57	5.33	Nucleus
AT5G18450.1	P61827	AtDREB2G	PF00847	1	307	34.23	6.32	Nucleus
AT2G40350.1	Q9SIZ0	AtDREB2H	PF00847	2	157	17.79	10.02	Nucleus
LOC_Os01g07120.1	Q0JQF7	OsDREB2A	PF00847	1	274	30.66	5.77	Nucleus
LOC_Os05g27930.1	Q5W6R4	OsDREB2B	PF00847	2	373	40.47	4.76	Nucleus
LOC_Os08g45110.1	Q84ZA1	OsDREB2C	PF00847	1	230	24.95	5.57	Nucleus
LOC_Os05g39590.1	Q65WX1	OsDREB2D	PF00847	2	261	27.69	5.06	Nucleus
LOC_Os03g07830.1	Q10R18	OsDREB2E	PF00847	1	318	33.46	5.91	Nucleus
Sobic.003G058200.1	–	SbDREB2A	PF00847	2	262	28.64	5.52	Nucleus
Sobic.009G101400.1	–	SbDREB2B	PF00847	2	336	36.17	5.07	Nucleus
Sobic.007G162700.1	–	SbDREB2C1	PF00847	1	388	41.19	6.67	Nucleus
Sobic.001G378900.1	–	SbDREB2C2	PF00847	2	351	38.62	8.67	Nucleus
Sobic.009G170800.1	–	SbDREB2D	PF00847	2	238	25.43	6.82	Nucleus
Sobic.001G486800.1	–	SbDREB2E	PF00847	1	316	33.27	5.93	Nucleus

PF00847: AP2 domain

(monomorphic). While Tajima's *D* value was found to be 2.28, nucleotide diversities were calculated as  $\pi$ : 0.53 and  $\theta$ : 0.39. The Tajima's *D* indicates the frequency spectrum for sequence polymorphism, including  $\pi$  from the average number of pairwise nucleotide differences (Tajima 1989) and  $\theta$  derived from number of segregating sites (Watterson 1975). Because a positive value of Tajima's *D* statistic indicates a balanced selection maintaining advantageous genetic diversity within populations (Yamasaki et al. 2007; Delph and Kelly 2014), *DREB2* genes in the sorghum genome

exhibit positive selection. All sorghum *DREB2* proteins contain AP2 domain structure (PF00847) (Fig. 1). Agarwal et al. (2006) stated that the DREB proteins have an ERF/AP2 DNA-binding domain which is quite conserved. This conserved domain structure (AP2) was identified in all DREBs (Fig. 1). The protein lengths were ranged from 238 to 388 amino acids, and *pI* values varied between 5.07 and 8.67. Similar results were observed in *Arabidopsis* and rice. ZmDREB2A was found to be in the length of 274 amino acid residues in maize (Qin et al. 2007), while Mizoi et al.



**Fig. 1** Comparison of amino acid alignments of the DREB2 proteins in *Arabidopsis*, rice, and sorghum, respectively. The yellow and red rectangles indicate the nuclear localization site (NLS) and AP2

domain structures, respectively. Sequences were aligned by ClustalW; identical and similar residues were shaded as black and grey, respectively, with 100% threshold

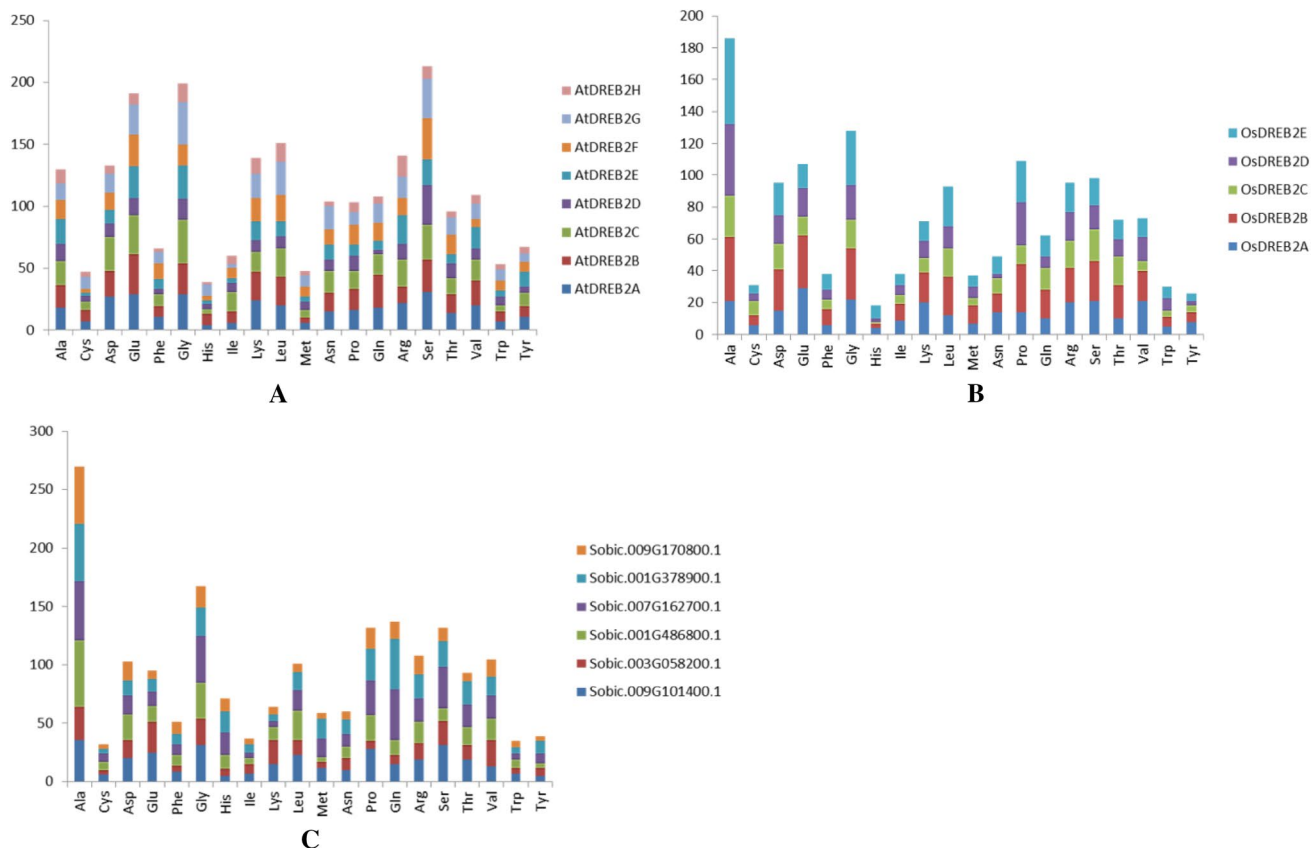
(2013) reported that GmDREB2A;2 was found in the length of 393 amino acid residues in soybean. In 19 DREB proteins of 14 grass species, protein lengths were found between 213 and 394 amino acid residues (Filiz and Tombuloglu 2014), which is also consistent with our data. Moreover, all DREB proteins found to be localized in the nucleus, suggesting TF functions of these proteins.

The amino acid pairwise distance analyses were performed among 19 DREB2 protein sequences and overall average was found as 0.42 by MEGA7 software. Filiz and Tombuloglu (2014) reported that pairwise distance of DREB proteins was 0.588 in 14 grass species. Their result was higher than our finding. In genetic comparison of rice and sorghum DREB proteins, the lowest distance values were found as 0.013 between SbDREB2E and OsDREB2E, followed by 0.214 between OsDREB2B and SbDREB2B, 0.214 between OsDREB2B and SbDREB2C1, and 0.279 between OsDREB2C and SbDREB2B. For *Arabidopsis* and sorghum, the lowest distance value was found as 0.167 between AtDREB2F and SbDREB2E, followed by 0.246 between AtDREB2A and SbDREB2A and 0.279 between AtDREB2A and SbDREB2B and AtDREB2B and SbDREB2B. According to these data, it can be proposed that *DREB2* genes showed variations at amino acid level.

The amino acid composition of DREB2A proteins in *Arabidopsis* (Fig. 2a), rice (Fig. 2b) and sorghum (Fig. 2c) were analyzed. The highest number of amino acid residue was found as serine (Ser) and alanine (Ala) in *Arabidopsis* and rice, respectively. In sorghum, the highest number of amino acid residue was identified as alanine (Ala), followed by glycine (Gly) and glutamine (Gln). Remarkably, Gly residue was found to be abundant in DREB proteins of all three plant species. In conserved motif analyses, motif I related with AP2 domain structure contained seven Gly residues (Table 2). Filiz and Tombuloglu (2014) determined the amino acid compositions of DREB proteins (%) in 14 grass species. It was found that glutamic acid (10%) was the most abundant amino acid followed by serine (9.3%), alanine (8.2%), and glycine (8%). Their result is consistent with our findings.

### Annotations of sorghum *DREB2* genes

The ontology annotations of *DREB2* genes in sorghum were performed using PhytoMine in Phytozome and AmiGO databases. All DREB2 proteins contained PF00847 (AP2 domain), SM00380 (DNA-binding domain in plant proteins such as APETALA2 and EREBPs), and PR00367



**Fig. 2** Amino acid distributions of DREB2 proteins from *Arabidopsis* (a), rice (b), and sorghum (c)

(ethylene-responsive element-binding protein signature) domain structures based on Pfam, SMART, and PRINTS databases, respectively. GO annotations were found as GO:0003700 (transcription factor activity and sequence-specific DNA binding), GO:0043565 (sequence-specific DNA binding), GO:0044212 (transcription regulatory region DNA binding), and GO:0003677 (DNA binding) for molecular function; GO:0006355 (regulation of transcription, DNA-templated) and GO:0045893 (positive regulation of transcription, DNA-templated) were found for biological process, and GO:0005634 (nucleus) was identified for cellular component.

### Analyses of conserved motifs

In the analysis, five conserved motif sequences were detected and motif-domain relationships were investigated (Table 3). The motif width and detected sites were found between 11 and 41 amino acid residues, and between 7 and 19 sites, respectively. In all sequences, motif 1 (GKGGPENAQCSYRGVQRQTWGWKVAEIREPNRGARLWLGTF) was the only one found to be related with AP2 domain structure. Based on these analyses, it can be suggested that AP2 domain structure is well conserved in all DREB2 proteins.

### Phylogenetic analyses

To analyze the evolutionary relationships among DREB2 proteins, we have performed phylogenetic analyses with 19 DREB2 protein sequences (Fig. 3). The two major groups were identified as groups A (subgroup A1 and A2) and B in phylogenetic tree. While the three plant species were clustered together in group A with monocots and dicot, group B was only consisted of monocot species with rice and sorghum. Particularly, rice DREB2C was isolated from the others and clustered with sorghum Sobic.007G162700.1 and Sobic.001G378900.1 in group B. Phylogenetic analyses also revealed that sorghum DREB2 proteins were continually clustered with rice orthologues, indicating conservation of *DREB2* genes in grasses. In addition, a clear distinction

could be observed between *Arabidopsis* and rice and sorghum DREB2 proteins in different subgroups. Mizoi et al. (2013) stated that monocot–dicot of DREB2 proteins separated subgroups in phylogenetic tree of *Arabidopsis*, rice, and soybean. These data are in agreement with our findings.

### Analyses of promoter sequences

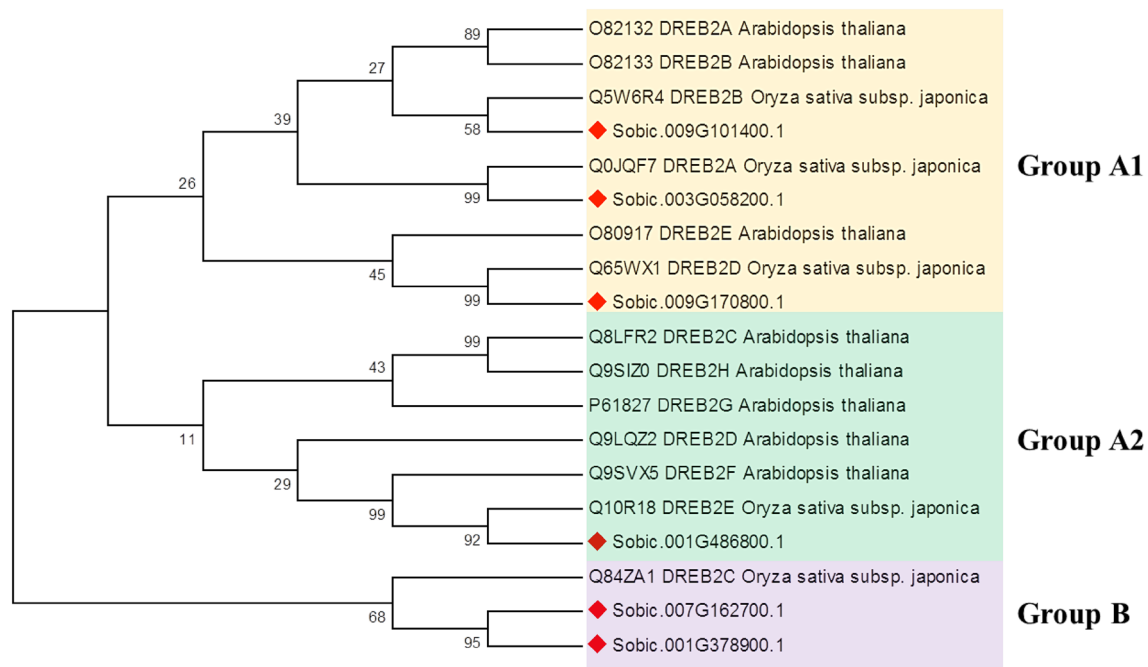
The gene expression levels through interactions with DNA or regulatory proteins were adjusted by many elements located in plant genome at processing of genetic information. *Cis*-regulatory sequences are linear nucleotide regions of non-coding DNA and their localizations change due to genes activities. There are two main types of *cis* sequences, containing enhancers and silencers (Venter and Botha 2010). In silico analysis of promoter regions of sorghum *DREB2* genes were performed in upstream of 1000 bp regions and 56 *cis*-element elements were identified by PlantCare database (Fig. 4). *Cis*-regulatory elements whose functions are unknown were removed from the heat map and present and absent elements were shown in red and green colors, respectively. In *DREB2* genes, 19 type *cis*-elements were found to be related with light-responsive elements, including ATCT-motif, Box 4, CGTCA-motif, MBS, Skn-1\_motif, Sp1, 3-AF1 binding site, AE-box, CATT-motif, GT1-motif, TCT-motif, TGG-motif, as-2-box, box II, GA-motif, GAG-motif, AT1-motif, and I-box. On the other hand, hormone responsive *cis*-elements such as abscisic acid responsiveness (ABRE), auxin-responsive (AuxRR-core and TGA-element), gibberellin-responsive element (MNF1), and salicylic acid responsiveness (TCA-element) were identified as 11 types, while three types of tissue-specific *cis*-elements were identified such as meristem (OCT), endosperm (P-box), and seed specific (RY-element). The sorghum *DREB2* genes also contained low-temperature responsiveness (LTR), heat stress responsiveness (HSE), anaerobic induction (ARE), zein metabolism regulation (O2-site), fungal elicitor responsive elements (Box-W1), anoxic-specific inducibility (G-Box), core promoter (TATA and CAAT boxes), MYB-binding site (GC-motif), and cell cycle regulation (MSA-like). Bonetta

**Table 3** Most conserved five motifs in DREB2 proteins in sorghum, rice, and *Arabidopsis*

Motif no.	Motif width	Detected sites	Motif sequence	Motif-domain association <sup>a</sup>
1	41	19	GKGGPENAQCSYRGVQRQTWGWKVAEIREPN-RGARLWLGTF	AP2 domain (PF00847)
2	28	18	PTAEEAALAYDEAARALYGPDARLNLPH	Not found
3	11	14	PAKGSKKGCMK	Not found
4	21	12	RKRNGPTSVAEIJRWKEYNE	Not found
5	14	7	DDCFDIBELLEMJN	Not found

A set of 19 DREB2 sequences were analyzed using MEME server

<sup>a</sup>Motifs are searched in Pfam database to figure out their associations with protein domains



**Fig. 3** Phylogenetic tree of 19 DREB2A proteins in three plant species, including Arabidopsis, rice, and sorghum. Tree was constructed by MEGA 7 using maximum-likelihood (ML) method with 1000 bootstraps. The sorghum DREB2 proteins were indicated as red diamonds

and McCourt (1998) stated that many stress-reacting promoters contain at least two common *cis*-elements, one of which was the ACGTG(G/T)C sequence, called ABA-responsive element (ABRE). In the present study, it was identified that *SbDREB2A*, *SbDREB2B*, *SbDREB2C1*, and *SbDREB2D* harbor ABRE *cis*-regulatory elements, indicating that they have major roles in response to stress conditions in sorghum. Consequently, it can be proposed that the various functions of *cis*-regulatory elements of *DREB2* genes in metabolic pathways have exhibited a complicated system which regulates many cell metabolisms in sorghum.

### Digital expression analyses of *DREB2* genes

The expression profiles of sorghum *DREB2* genes were investigated at different developmental stages and anatomical parts including seedling, stem elongation, booting, flowering, and dough stages (Fig. 5). The expression data of four *DREB2* genes, *SbDREB2A*, *SbDREB2B*, *SbDREB2C1*, and *SbDREB2D* were identified. The high expression levels were found for *SbDREB2A* and *SbDREB2B* genes, while *SbDREB2C1* and *SbDREB2D* genes showed the medium-level expression. In addition, *SbDREB2B* has more variable expression in these different stages/components, whereas *SbDREB2A* is more consistent in the expression levels.

The expression level of *SbDREB2A*, *SbDREB2B*, *SbDREB2C1*, and *SbDREB2D* were investigated at six different anatomical parts, containing pith, rind, shoot, roots,

leaf, and internode (Fig. 6). High expression levels were identified for *SbDREB2A* and *SbDREB2B* genes, while *SbDREB2C1* and *SbDREB2D* genes showed medium level of expression. With regard to the expression data, it can be proposed that expression levels of *DREB2* genes show variations and complex regulations based on gene type, anatomical part, and developmental stage.

To get deep insights about function of *SbDREB2* genes, MOROKOSHI transcriptome database was used, and expression data of four *SbDREB2* genes were obtained (Fig. 7). When expression profiles of *SbDREB2* genes were analyzed, highest FPKM (fragments per kilobase million) values were found for *SbDREB2B* genes (Fig. 7b). Particularly, *SbDREB2B* gene was highly expressed in leaf and stem, while its expression was the lowest in embryo. Furthermore, infections induced the expression level of *SbDREB2B* gene. While the *SbDREB2C1* gene was highly expressed in embryo and endosperm tissues (Fig. 7c), *SbDREB2A* was highly expressed in pistil, embryo, and endosperm tissues (Fig. 7a). Expression of *SbDREB2D* gene was high in endosperm, seed, and anther (Fig. 7d). In addition, *SbDREB2D* gene was up-regulated by nitrogen, PEG (polyethylene glycol), and ABA (abscisic acid) treatments in roots. In general, it can be suggested that *SbDREB2A* and *SbDREB2B* genes reported higher expression levels under various conditions and in tissues of sorghum.



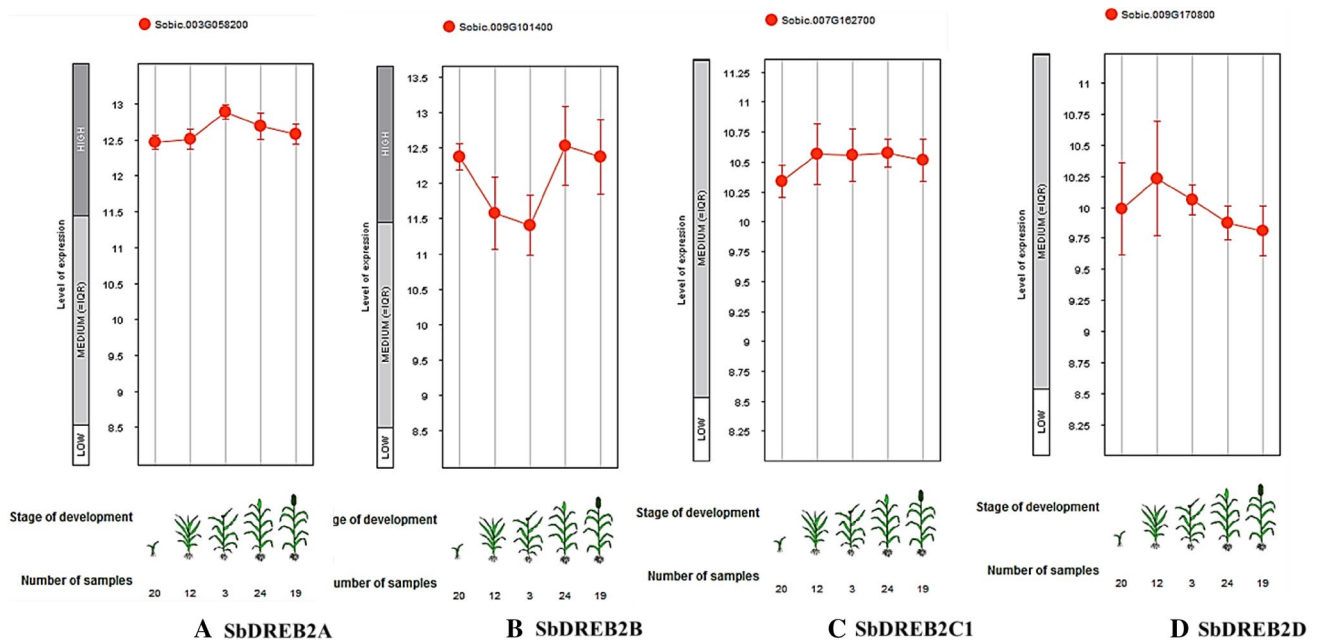
Cis-elements	Function	<i>Sobic.009G101400</i>	<i>Sobic.003G058200</i>	<i>Sobic.001G486800</i>	<i>Sobic.007G162700</i>	<i>Sobic.001G378900</i>	<i>Sobic.009G170800</i>
A-box	regulatory element						
ABRE	abscisic acid responsiveness						
ACE	light responsive element						
ARE	anaerobic induction						
ATCT-motif	light responsive element						
AuxRR-core	auxin-responsive element						
Box 4	light responsive element						
box S	core promoter						
CAAT-box	meristem specific activation						
CCGTCC-box	MeJA-responsiveness						
CGTCA-motif	light responsive element						
G-Box	anoxic specific inducibility						
GC-motif	MYB binding site						
MBS	light responsive element						
MNF1	gibberellin-responsive element						
P-box	endosperm expression						
Skn-1_motif	light responsive element						
Sp1	light responsive element						
TATA-box	auxin-responsive element						
TGA-box	MeJA-responsiveness						
TGACG-motif	regulatory element						
3-AF1 binding site	light responsive element						
AAGAA-motif	regulatory element						
AE-box	light responsive element						
CATT-motif	light responsive element						
GT1-motif	light responsive element						
LTR	low-temperature responsiveness						
MSA-like	cell cycle regulation						
O2-site	zein metabolism regulation						
TC-rich repeats	stress responsiveness						
TCA-element	salicylic acid responsiveness						
TCT-motif	light responsive element						
TGA-element	auxin-responsive element						
TGG-motif	light responsive element						
as-2-box	light responsive element						
box II	light responsive element						
GA-motif	light responsive element						
GAG-motif	light responsive element						
HSE	heat stress responsiveness						
RY-element	seed-specific regulation						
TCCACCT-motif	regulatory element						
AT1-motif	light responsive element						
I-box	light responsive element						
JERE	regulatory element						
OCT	meristem specific activation						
motif IIb	abscisic acid responsive element						
AC-I	regulatory element						
CAT-box	meristem expression						
CCAAT-box	MYBHv1 binding site						
GCN4_motif	endosperm expression						
TATCCAT/C-motif	regulatory element						
Box-W1	fungal elicitor responsive element						
CE3	ABA responsiveness						
GARE-motif	gibberellin-responsive element						
GCC box	regulatory element						
W box	regulatory element						

**Fig. 4** Heat map of *cis*-regulatory elements in *DREB2* genes in sorghum (green: absent and red: present). The *cis*-regulatory elements were investigated using 1000-bp upstream regions of *DREB2* genes

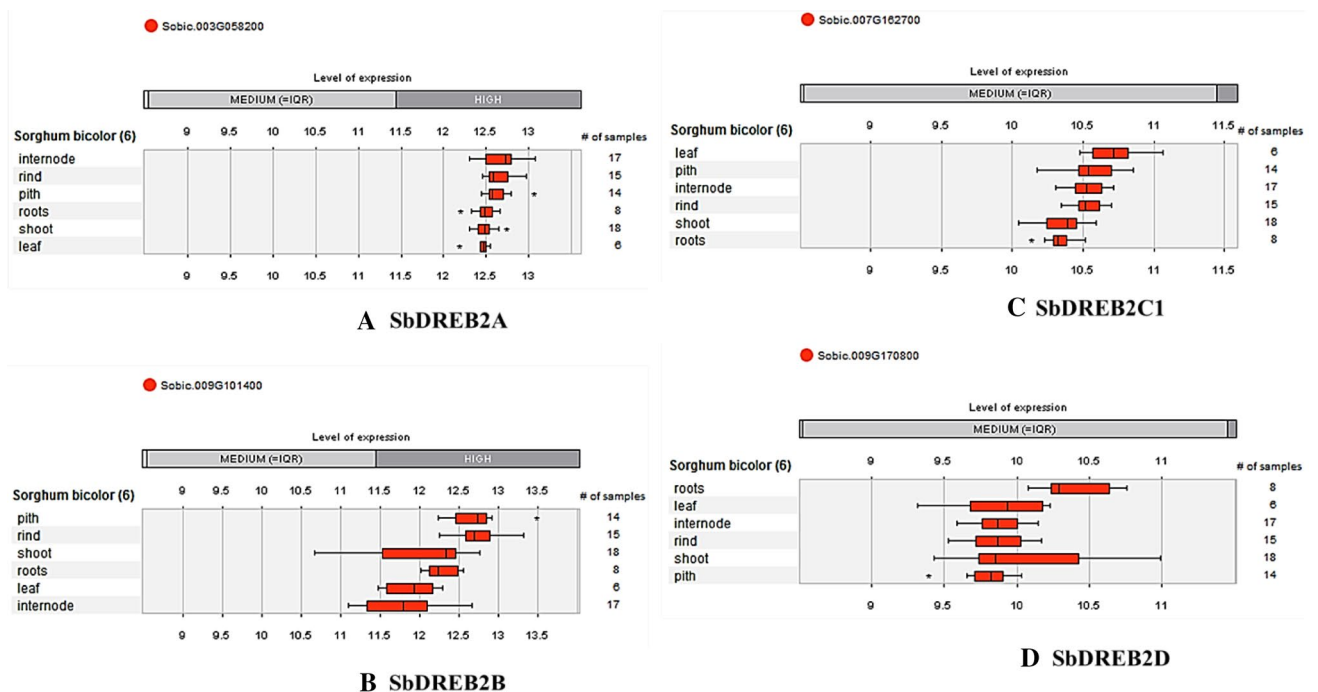
### The co-expression network of *SbDREB2A* gene

To understand roles of *SbDREB2A* gene in metabolic pathways, a co-expression network of *SbDREB2A*

(*Sobic.003G058200*) was constructed using MOROKO-SHI *Sorghum* transcriptome database (Fig. 8). The six first neighbors were identified for *SbDREB2A* gene, including *Sb03g033020.1* (transcription factor GT-2 and related



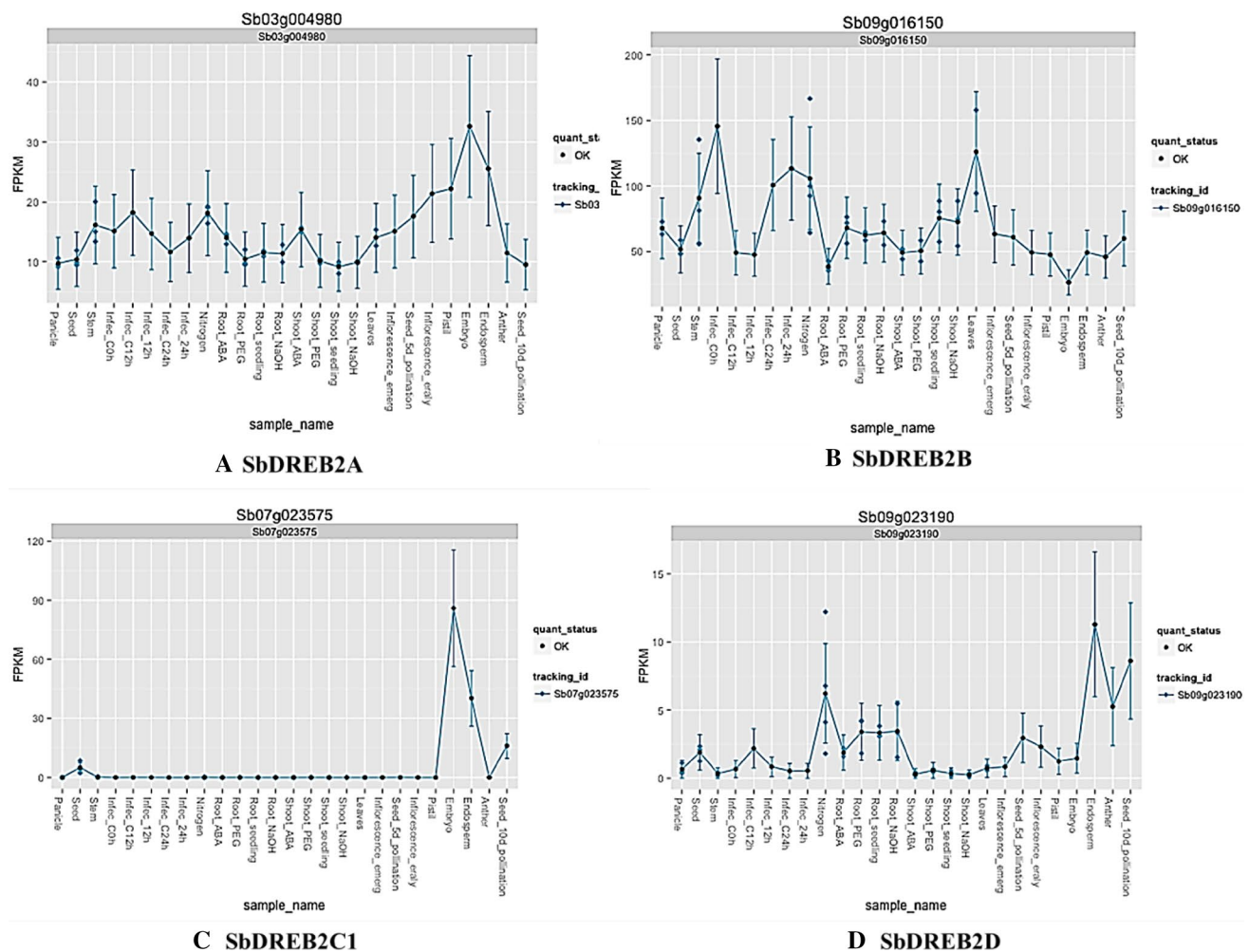
**Fig. 5** Expression levels of *DREB2* genes in different five development stages of sorghum, such as seedling, stem elongation, booting, flowering, and dough stages by Genevestigator database



**Fig. 6** Expression levels of *DREB2* genes in different six anatomical parts of sorghum, such as pith, rind, shoot, roots, leaf, and internode by Genevestigator database

proteins, contains trihelix DNA-binding/SANT domain), Sb01g010570.1 (transcription regulator), Sb03g023230.1 (unknown protein), Sb03g005080.1 (DNA-binding protein, putative), Sb01g000750.1 (RNA recognition motif

(RRM)-containing protein), and Sb01g014700.1 (F-box family protein). In addition, two transcription factors were found as Sb02g029870.1 (*bZIP* transcription factor, putative *bZIP69*) and Sb04g003450.1 (pathogenesis related



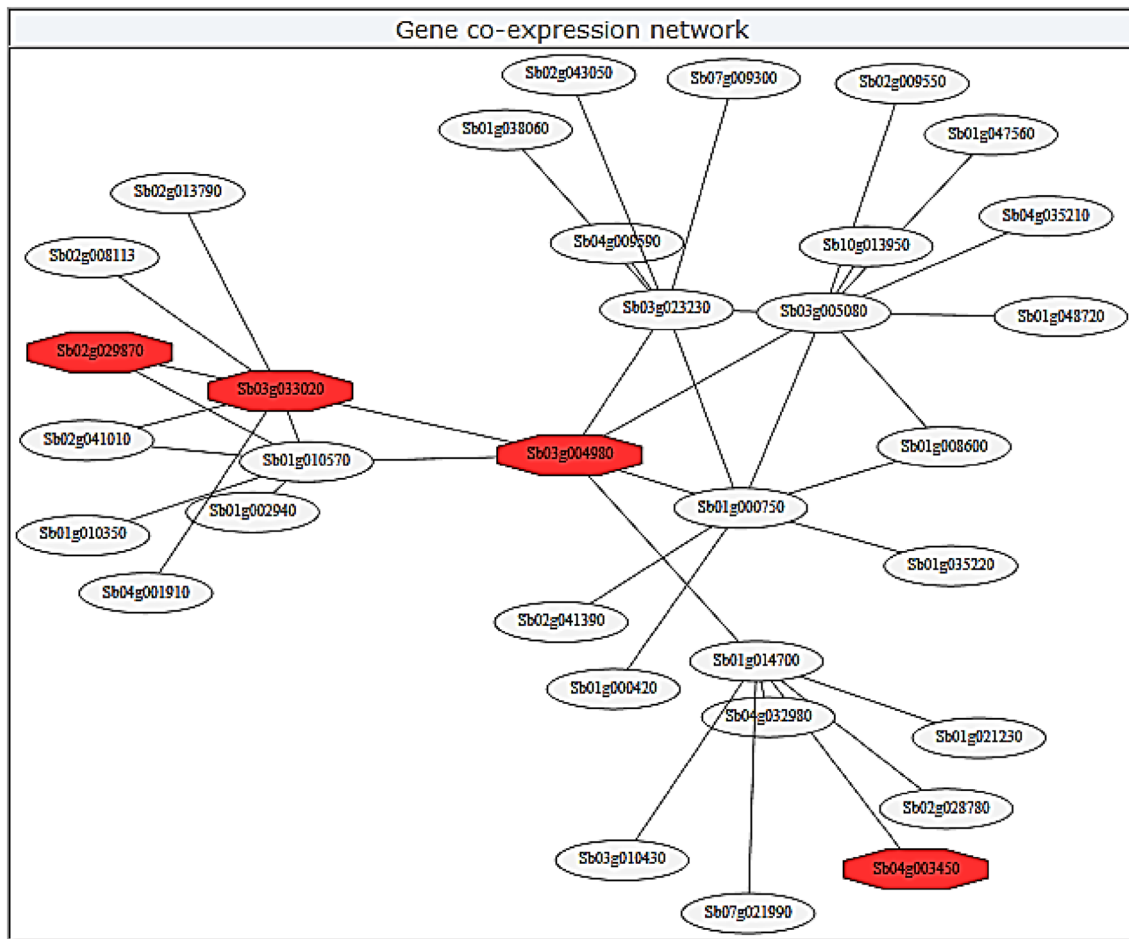
**Fig. 7** Expression profiles of *SbDREB2A*, *B*, *C1*, and *D* genes in various development stages, anatomical parts, tissues, and stress conditions using MOROKOSHI transcriptome database. The legends on the figures show the synonym name of *SbDREB2* genes

homeodomain protein A, PRHA) in other neighbors. The F-box is a protein motif consisting of nearly 50 amino acids that function as a site of protein–protein interaction (Kipreos and Pagano 2000). F-box proteins play roles in various cellular processes such as cell cycle transition, transcriptional regulation, and signal transduction with Skp1p–cullin–F-box protein (SCF) complexes or non-SCF complexes (Kuroda et al. 2002). Trihelix family transcription factor (known as *GT factor*) is one of the plant-specific TF families, and they function in development and stress tolerance in various plants (Li et al. 2017). The GT factors bind to GT elements present in the promoters for many plant genes (Ayadi et al. 2004). Members of the RNA-editing factor interacting protein (RIP) family and Organelle RNA Recognition Motif-containing (ORRM) family are vital components for editosome in *Arabidopsis* (Shi et al. 2016). The basic leucine zipper (*bZIP*) TF family regulates essential processes in all eukaryotes. The *bZIPs* are regulators of many metabolic processes such as leaf and seed formation, photomorphogenesis,

energy homeostasis, and abiotic and biotic stress responses in plants (Corrêa et al. 2008). On the basis of the above given data, it can be proposed that *SbDREB2A* gene may play roles in the regulation of various metabolic processes in sorghum. TFs regulate gene expressions depending on intrinsic and extrinsic signal; thus, there are regulatory relationships among TFs. When primary transcription factors are expressed, they induce another one or another target gene. This process repeats and as known transcriptional cascade (Jothi et al. 2009). In this study, four TFs were found in co-expression network, suggesting that *SbDREB2A* gene may be a member of transcriptional cascade in metabolic pathways of sorghum.

### Analyses of miRNAs targeting

MicroRNAs (miRNAs) are endogenous small RNA class with mostly consisting of 20–24 nucleotide controlling gene expression. In plants, miRNAs regulate gene expression,



**Fig. 8** Co-expression networks of *SbDREB2A* gene by MOROKOSHI Sorghum transcriptome database. Sb03g004980 shows the synonym of *SbDREB2A* (Sobic.003G058200). In addition, the red filled octagons show the transcription factors

including transcription factors, stress response proteins, and proteins affecting development, growth, and physiology of plants (Ding et al. 2012; Rogers and Chen 2013). miRNAs function in two ways: suppressing translation of a target gene or degrading target mRNAs post-transcriptionally (Carrington and Ambros 2003). In sorghum, 113 conserved miRNA homologs belonging to 31 distinct miRNA families were found by processing the small RNA reads (Zhang et al. 2011). Under drought stress, a significant de-regulation was identified with miR396, miR393, miR397-5p, miR166,

miR167, and miR168 in 11 elite sorghum genotypes (Hamza et al. 2016). In the present study, six *SbDREB2* genes were searched against psRNATarget database using 241 known miRNAs for sorghum to identify miRNAs that play roles in regulation of *SbDREB2* genes. Finally, five types of miRNAs were identified for *SbDREB2A*, *SbDREB2B*, *SbDREB2C1*, and *SbDREB2C2* (Table 4). Katiyar et al. (2015) reported that induced expression of miR156b, miR396b-c, miR396d-e, miR396f, and miR5385 was enhanced under drought stress in drought tolerant genotype in sorghum. Zhang et al.

**Table 4** Details of predicted miRNAs for *SbDREB2* genes using psRNATarget database

Target gene	miRNA acc.	Expectation	miRNA sequence	Inhibition type
Sobic.009G101400.1	sbi-miR6232a-5p	4.5	UACAUGGUUUUUUCAGUUUCGUG	Cleavage
Sobic.009G101400.1	sbi-miR437x-5p	5	UUUCAACUGAAUCCUGUUGAGAU	Cleavage
Sobic.003G058200.1	sbi-miR5568e-3p	5	UGCAAAAUCGAAAAGAUCUAU	Translation
Sobic.007G162700.1	sbi-miR5385	5	CUCUUCGCCACCCCAACCACCA	Translation
Sobic.001G378900.1	sbi-miR5386	4.5	GUCGCGCGCUGUCGUGC	Cleavage

reported that the least-abundant expression was found for miR393, miR437, miR1126, and miR1436 in sweet Sorghum (M81E). In addition, miR5568-TAS derived tasiRNA (trans-acting short-interfering RNA) was detected in drought-stressed library of a drought-sensitive genotype targeting a universal stress protein (USP) (Katiyar et al. 2015). In the present study, sbi-miR5385 was identified for *SbDREB2C1* gene by translation inhibition, suggesting that this gene may play roles in response to drought stress.

### Expression profiles of *SbDREB2* genes under Cd and drought stresses

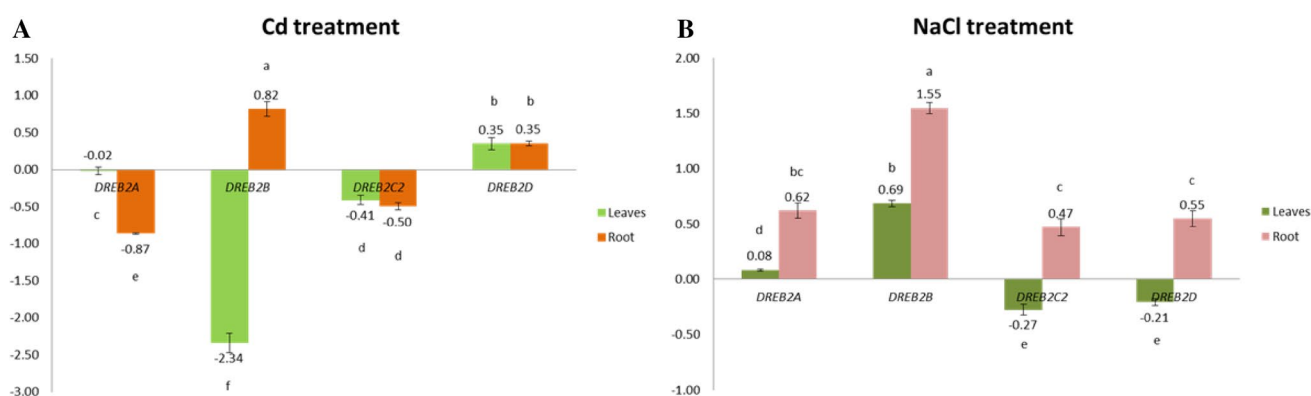
DREBs are important transcription factors that are induced by abiotic stress-related genes. They can be separated into two groups, *DREB1* and *DREB2*, members of which are involved in two different signal transduction pathways under low temperature and dehydration, respectively (Agarwal et al. 2006). In the present study, to elucidate their roles in stress conditions, expressions of *SbDREB2* genes in sorghum root and leaf tissues under NaCl and Cd stresses were quantified using RT-qPCR. In the analysis, the expressions of four *DREB2* genes (*SbDREB2A*, *2B*, *2C2*, and *2D*) were detected, but *SbDREB2C1* and *2E* (Fig. 9). Under Cd stress, *SbDREB2A* and *2C2* were down-regulated in both root and leaf tissues, while *SbDREB2B* was down-regulated only in leaves (Fig. 9a). Conversely, *SbDREB2D* in both leaves and roots was up-regulated, but *SbDREB2B* only in roots. Mizoi et al. (2013) reported that *GmDREB2A;2* was expressed not only by dehydration and heat but also by low temperature in soybean. In a similar study, expression level of *ZmDREB2A* increased by cold, dehydration, salt and heat stresses in maize seedlings (Qin et al. 2007). Hwang et al. (2012) proposed that *Arabidopsis DREB2C* gene was expressed by a superoxide anion propagator, methyl viologen (MV),

suggesting that *Arabidopsis DREB2C* is an oxidative stress-responsive gene. The mentioned data establish that *DREB2* TFs are induced under various abiotic stresses, explaining the induction of *SbDREB2* genes under Cd stress.

Under NaCl stress, all *SbDREB2* genes were up-regulated in roots, while *SbDREB2A* and *2B* gene were up-regulated only in leaves (Fig. 9b). On the other hand, *SbDREB2C2* and *2D* were down-regulated in leaves. It was reported that *DREB2A* and *DREB2B* transcription factors in *Arabidopsis* were induced by dehydration stress and enhanced expression of various genes related to drought tolerance (Liu et al. 1998; Nakashima et al. 2014). In pearl millet (*P. glaucum*), *PgDREB2A* gene was up-regulated under drought stress within 1 h of treatment, whereas the expression was delayed when plants were exposed to cold and salinity stresses (Agarwal et al. 2007). Herath (2016) showed that *OsDREB2B* was highly expressed under three stress conditions (drought, cold, and salt). Highest *OsDREB2B* expression was estimated under cold stress *in silico* expression analysis of *OsDREB2s* genes. In addition, both *OsDREB2A* and *OsDREB2C* indicated highest expression under drought stress. In tomato, *SIDREB2* was up-regulated in roots and young leaves under NaCl exposure, and it was also induced by KCl and drought (Hichri et al. 2016). In foxtail millet (*Setaria italica*), a significant up-regulation of *SiDREB2* was detected under dehydration (polyethylene glycol) and salinity (NaCl) stresses, while its expression was less affected by other stresses (Lata et al. 2011). Finally, the results of these reported studies were similar to our findings.

### Secondary and tertiary structures analyses of *DREB2* proteins

To better understand functions of *SbDREB2* proteins, we first performed secondary and tertiary structure analyses

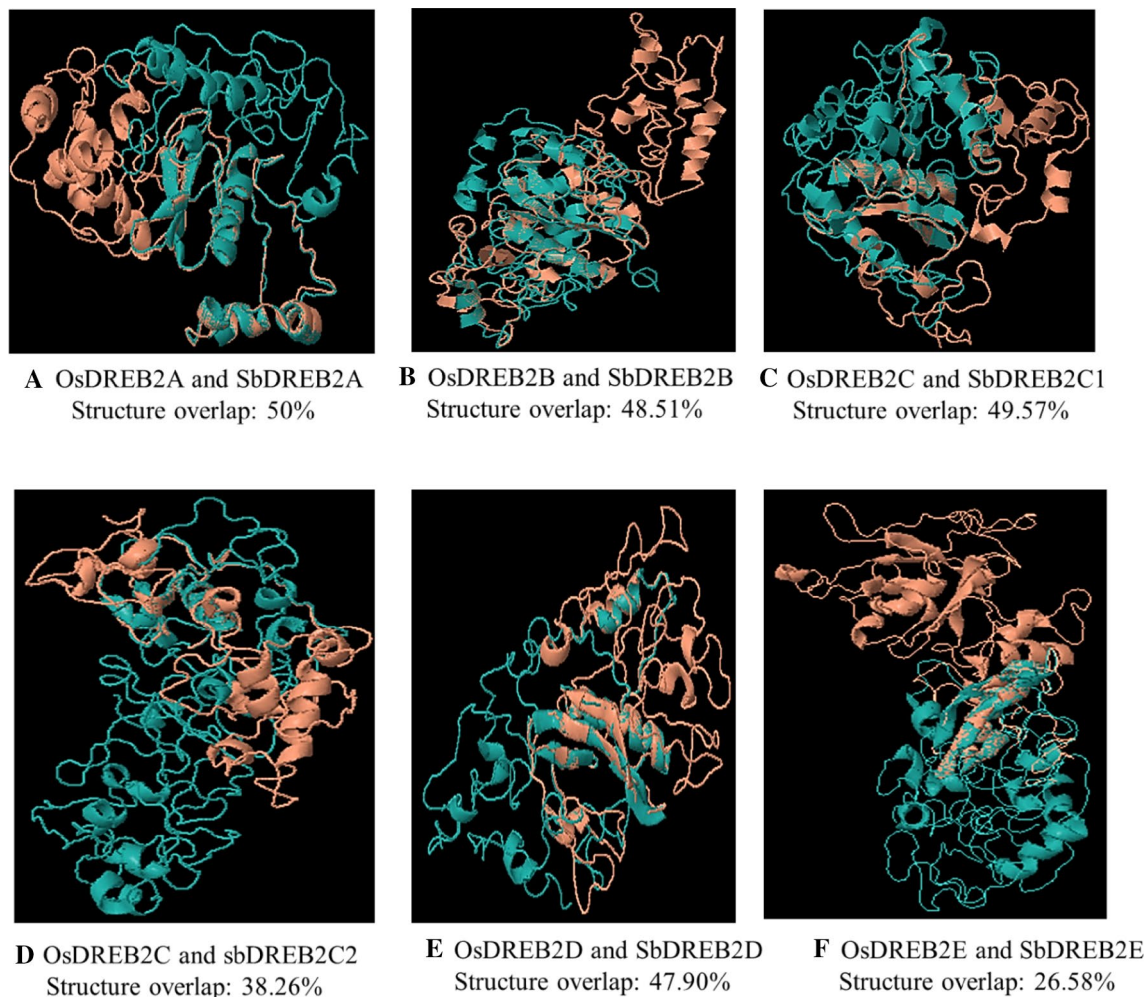


**Fig. 9** Expression of *SbDREB2* genes under Cd (a) and NaCl (b) stresses in sorghum leaves and roots detected by RT-qPCR. Bars above/below x-axis demonstrate up- and down-regulation of genes, respectively. Expression values are indicated as log<sub>2</sub> scale to bet-

ter represent the symmetry of magnitude for up and down-regulated genes. Error bars depict standard deviation of the mean (sdom;  $n = 3$ ). Values with identical lowercase letters are not significantly different ( $P < 0.01$ ) as compared by Student's *t* test

(Fig. 10). According to the results of secondary structure analysis, percentage (%) of alpha helix ( $\alpha$ -helix), extended strand (E-strands), beta-turn ( $\beta$  turn), and random coil were ranged from 30.15 to 40.08, from 10.76 to 16.79, from 3.35 to 9.16, and from 33.97 to 53.09, respectively. Ji and Li (2010) reported that the secondary structures play important roles in protein structure and folding. These results proved that some structural variations observed in SbDREB2 proteins may be related with functional flexibilities of SbDREB2 proteins. Subsequently, all SbDREB2 proteins were modelled using Phyre2 server at intensive mode. The Ramachandran plot analysis was used for model validation by Vadar server. The results showed that more than 90% of residues were present in core and allowed regions, indicating the reliability of 3D models. In addition, topology independent comparison of the 3D structures was performed by CLICK server comparing sorghum and rice DREB2 proteins (Fig. 8). While the highest structure overlap was found as

50% between OsDREB2A and SbDREB2A (Fig. 7a), the lowest overlap found as 26.58% between rice OsDREB2E and SbDREB2E (Fig. 7f). In phylogenetic tree, OsDREB2A and SbDREB2A and OsDREB2E and SbDREB2E were grouped together with 99% and 92% bootstrap values, respectively. While OsDREB2B and SbDREB2B were clustered as 58% bootstrap value which is the lowest value in phylogenetic tree, structure overlap for this group was found as 48.51%, suggesting that phylogenetic relationship may not be directly connected with 3D structure similarity for all *SbDREB2* genes. Upon having compared 3D models of AP2 domain structures of DREB proteins in *Oryza sativa* and *Avena sativa*, some structural divergences were detected at Pro59, Gly60, and Arg62 amino acid residues contributing to  $\beta$ -sheet formation in rice, which were not presented in their *A. sativa* complements (Filiz and Tombuloglu 2014). In the present study, some structural divergences were detected on 3D structures of SbDREB2 proteins; these variations



**Fig. 10** Structure overlaps of DREB2 proteins at superimposition mode by CLICK server. The orange and blue 3D structures show the rice and sorghum, respectively. In addition, structure overlaps values (%) were indicated in figure legends

could support the functional diversities of *SbDREB2* genes in cell metabolism.

In conclusion, *SbDREB2* genes and proteins were analyzed at two steps: bioinformatics approaches and transcript levels in this study. The bioinformatics analyses showed that *SbDREB2* genes indicated structural variations and differences at expression levels. In addition, *SbDREB2* proteins exhibited some structural differences in secondary and tertiary structures. According to the expression profiles under Cd and NaCl stresses, expression of *SbDREB2* genes showed variations with regard to stress and tissue types. Particularly, *SbDREB2A* and *2B* under NaCl stress and *SbDREB2D* under Cd stress showed more dynamic expression profiles over the others. Finally, results of this study will contribute to understandings of *DREB2* genes in grasses, particularly sorghum under abiotic stress conditions.

**Author contribution** EF and MAA conceived the study; MAA, EF, and KK conducted the experiments. MAA and EF wrote the manuscript; all authors read, edited, and approved the manuscript.

## Compliance with ethical standards

**Conflict of interest** The authors declare no conflict of interest.

## References

- Agarwal PK, Agarwal P, Reddy MK, Sopory SK (2006) Role of *DREB* transcription factors in abiotic and biotic stress tolerance in plants. *Plant Cell Rep* 25:1263–1274
- Agarwal P, Agarwal PK, Nair S, Sopory SK, Reddy MK (2007) Stress-inducible *DREB2A* transcription factor from *Pennisetum glaucum* is a phosphoprotein and its phosphorylation negatively regulates its DNA-binding activity. *Mol Genet Genom* 277:189–198
- Agarwal P, Agarwal PK, Joshi AJ, Sopory SK, Reddy MK (2010) Overexpression of *PgDREB2A* transcription factor enhances abiotic stress tolerance and activates downstream stress-responsive genes. *Mol Biol Rep* 37:1125–1135
- Akbudak MA, Filiz E, Kontbay K (2018) Genome-wide identification and cadmium induced expression profiling of sulfate transporter (*SULTR*) genes in sorghum (*Sorghum bicolor* L.). *Biometals* 31:91
- Ayadi M, Delaporte V, Li YF, Zhou DX (2004) Analysis of *GT-3a* identifies a distinct subgroup of trihelix DNA-binding transcription factors in *Arabidopsis*. *FEBS Lett* 562:147–154
- Bailey TL, Johnson J, Grant CE, Noble WS (2015) The MEME suite. *Nucleic Acids Res* 43(W1):39–49
- Bihani P, Char B, Bhargava S (2011) Transgenic expression of sorghum *DREB2* in rice improves tolerance and yield under water limitation. *J Agric Sci* 149:95–101
- Bonetta D, McCourt P (1998) Genetic analysis of ABA signal transduction pathways. *Trends Plant Sci* 3:231–235
- Carbon S, Ireland A, Mungall CJ, Shu S, Marshall B et al (2009) AmiGO: online access to ontology and annotation data. *Bioinformatics* 25:288–289
- Carrington JC, Ambros V (2003) Role of microRNAs in plant and animal development. *Science* 301:336–338
- Corrêa LGG, Riaño-Pachón DM, Schrago CG, Vicentini dos Santos R, Mueller-Roeber B, Vincenz M (2008) The role of bZIP transcription factors in green plant evolution: adaptive features emerging from four founder genes. *PLoS One* 3(8):e2944
- Dai X, Zhao PX (2011) psRNATarget: a plant small RNA target analysis server. *Nucleic Acids Res* 39:W155–W159
- Delph LF, Kelly JK (2014) On the importance of balancing selection in plants. *New Phytol* 201:45–56
- Ding J, Li D, Ohler U, Guan J, Zhou S (2012) Genomewide search for miRNA-target interactions in *Arabidopsis thaliana* with an integrated approach. *BMC Genom* 13(Suppl 3):S3
- Filiz E, Tombuloğlu H (2014) In silico analysis of *DREB* transcription factor genes and proteins in grasses. *Appl Biochem Biotechnol* 174:1272–1285
- Gasteiger E, Hoogland C, Gattiker A, Duvaud S, Wilkins MR, Appel RD et al (2005) Protein identification and analysis tools on the ExPASy server. In: Walker JM (ed) *The proteomics protocols handbook*. Humana, Louisville, pp 571–607
- Geourjon C, Deleage G (1995) SOPMA: significant improvements in protein secondary structure prediction by consensus prediction from multiple alignments. *Comput Appl Biosci* 11:681–684
- Goodstein DM, Shu S, Howson R, Neupane R, Hayes RD, Fazo J et al (2012) Phytozome: a comparative platform for green plant genomics. *Nucleic Acids Res* 40:1178–1186
- Hall TA (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symp Ser* 41:95–98
- Hamza NB, Sharma N, Tripathi A, Sanan-Mishra N (2016) MicroRNA expression profiles in response to drought stress in *Sorghum bicolor*. *Gene Expr Patterns* 20:88–98
- Herath V (2016) Small family, big impact: in silico analysis of *DREB2* transcription factor family in rice. *Comput Biol Chem* 65:128–139
- Hichri I, Muhovski Y, Clippe A, Žižková E, Dobrev PI, Motyka V, Lutts S (2016) *SIDREB2*, a tomato dehydration-responsive element-binding 2 transcription factor, mediates salt stress tolerance in tomato and *Arabidopsis*. *Plant Cell Environ* 39:62–79
- Horton P, Park KJ, Obayashi T, Fujita N, Harada H, Adams-Collier CJ, Nakai K (2007) WoLF PSORT: protein localization predictor. *Nucleic Acids Res* 35:W585–W587
- Hruz T, Laule O, Szabo G, Wessendorf F, Bleuler S, Oertle L, Widmayer P, Gruissem W, Zimmermann P (2008) Genevestigator v3: a reference expression database for the meta-analysis of transcriptomes. *Adv Bioinform* 2008:420747
- Hwang JE, Lim CJ, Chen H, Je J, Song C, Lim CO (2012) Overexpression of *Arabidopsis* dehydration responsive element-binding protein 2C confers tolerance to oxidative stress. *Mol Cells* 33:135–140
- Ji YY, Li YQ (2010) The role of secondary structure in protein structure selection. *Eur Phys J E* 32:103–107
- Jones DT, Taylor WR, Thornton JM (1992) The rapid generation of mutation data matrices from protein sequences. *Comput Appl Biosci* 8:275–282
- Jothi R, Balaji S, Wuster A, Grochow AJ, Gsponer J, Przytycka MT, Aravind L, Babu MM (2009) Genomic analysis reveals a tight link between transcription factor dynamics and regulatory network architecture. *Mol Syst Biol* 5:294
- Katiyar A, Smita S, Muthusamy SK, Chinnusamy V, Pandey DM, Bansal KC (2015) Identification of novel drought-responsive microRNAs and trans-acting siRNAs from *Sorghum bicolor* (L.) Moench by high-throughput sequencing analysis. *Front Plant Sci* 6:506
- Kelley LA, Sternberg MJE (2009) Protein structure prediction on the Web: a case study using the Phyre server. *Nat Protoc* 4:363–371
- Khan ZH, Kumar B, Dhatteval P, Mehrotra S, Mehrotra R (2017) Transcriptional regulatory network of cis-regulatory elements (Cres) and transcription factors (TFs) in plants during abiotic stress. *Int J Plant Biol Res* 5(2):1064

- Kipreos ET, Pagano M (2000) The F-box protein family. *Genome Biol* 1(5):reviews3002.1–3002
- Kumar S, Stecher G, Tamura K (2016) MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Mol Biol Evol* 33:1870–1874
- Kuroda H, Takahashi N, Shimada H, Seki M, Shinozaki K, Matsui M (2002) Classification and expression analysis of *Arabidopsis* F-box-containing protein genes. *Plant Cell Physiol* 43:1073–1085
- Lata C, Prasad M (2011) Role of DREBs in regulation of abiotic stress responses in plants. *J Exp Bot* 62:4731–4748
- Lata C, Bhutty S, Bahadur RP, Majee M, Prasad M (2011) Association of a SNP in a novel DREB2-like gene SiDREB2 with stress tolerance in foxtail millet [*Setaria italica* (L.)]. *J Exp Bot* 62:3387–3401
- Lescot M, De 'hais P, Moreau Y, De Moor B, Rouze' P, Rombauts S (2002) PlantCARE: a database of plant cis-acting regulatory elements and a portal to tools for in silico analysis of promoter sequences. *Nucleic Acids Res* 30:325–327
- Li H, Huang W, Liu ZW, Wu ZJ, Zhuang J (2017) Trihelix family transcription factors in tea plant (*Camellia sinensis*): identification, classification, and expression profiles response to abiotic stress. *Acta Physiol Plant* 39:217
- Librado P, Rozas J (2009) DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* 25:1451–1452
- Liu Q, Kasuga M, Sahana Y, Abe H, Miura S, Yamaguchi-Shinozaki K, Shinozaki K (1998) Two transcription factors, DREB1 and DREB2, with an EREBP/AP2 DNA binding domain separate two cellular signal, respectively, in *Arabidopsis*. *Plant Cell* 10:1391–1406
- Livak KJ, Schmittgen TD (2001) Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) method. *Methods* 25:402–408
- Makita Y, Shimada S, Kawashima M, Kondou-Kuriyama T, Toyoda T, Matsui M (2015) MOROKOSHI: transcriptome database in *Sorghum bicolor*. *Plant Cell Physiol* 56:e6
- Matsukura S, Mizoi J, Yoshida T, Todaka D, Ito Y, Maruyama K, Shinozaki K, Yamaguchi-Shinozaki K (2010) Comprehensive analysis of rice DREB2-type genes that encode transcription factors involved in the expression of abiotic stress-responsive genes. *Mol Genet Genom* 283:185–196
- Mizoi J, Shinozaki K, Yamaguchi-Shinozaki K (2012) AP2/ERF family transcription factors in plant abiotic stress responses. *Biochem Biophys Acta* 1819:86–96
- Mizoi J, Ohori T, Moriwaki T, Kidokoro S, Todaka D, Maruyama K, Kusakabe K, Osakabe Y, Shinozaki K, Yamaguchi-Shinozaki K (2013) GmDREB2A;2, a canonical dehydration-responsive element-binding protein 2-type transcription factor in soybean, is posttranslationally regulated and mediates dehydration-responsive element-dependent gene expression. *Plant Physiol* 161:346–361
- Nakashima K, Yamaguchi-Shinozaki K, Shinozaki K (2014) The transcriptional regulatory network in the drought response and its crosstalk in abiotic stress responses including drought, cold, and heat. *Front Plant Sci* 5:170
- Panda P, Nath S, Chanu TT, Sharma GD, Panda SK (2011) Cadmium stress-induced oxidative stress and role of nitric oxide in rice (*Oryza sativa* L.). *Acta Physiol Plant* 33:1737–1747
- Punta M, Coghill PC, Eberhardt RY, Mistry J, Tate J, Bournsnel C et al (2012) The Pfam protein families database. *Nucleic Acids Res* 40:D290–D301. <https://doi.org/10.1093/nar/gkr1065>
- Qin F, Kakimoto M, Sakuma Y, Maruyama K, Osakabe Y, Tran LS, Shinozaki K, Yamaguchi-Shinozaki K (2007) Regulation and functional analysis of ZmDREB2A in response to drought and heat stresses in *Zea mays* L.. *Plant J* 50:54–69
- Reddy DS, Sivasakthi K, Bhatnagar-Mathur P, Vadez V, Sharma KK (2016) Evaluation of sorghum [*Sorghum bicolor* (L.)] reference genes in various tissues and under abiotic stress conditions for quantitative real-time PCR data normalization. *Front Plant Sci* 7:529
- Riechmann JL, Meyerowitz EM (1998) The AP2/EREBP family of plant transcription factors. *Biol Chem* 379:633–646
- Rogers K, Chen X (2013) Biogenesis, turnover, and mode of action of plant microRNAs. *Plant Cell* 25:2383–2399
- Sakuma Y, Liu Q, Dubouzet JG, Abe H, Shinozaki K, Yamaguchi-Shinozaki K (2002) DNA-binding specificity of the AP2/ERF domain of *Arabidopsis* DREBs, transcription factors involved in dehydration- and cold-inducible gene expression. *Biochem Biophys Res Commun* 290:998–1009
- Sakuma Y, Maruyama K, Osakabe Y, Qin F, Seki M, Shinozaki K, Yamaguchi-Shinozaki K (2006a) Functional analysis of an *Arabidopsis* transcription factor, *DREB2A*, involved in drought-responsive gene expression. *Plant Cell* 18:1292–1309
- Sakuma Y, Maruyama K, Qin F, Osakabe Y, Shinozaki K, Yamaguchi-Shinozaki K (2006b) Dual function of an *Arabidopsis* transcription factor *DREB2A* in water-stress responsive and heat-stress-responsive gene expression. *Proc Natl Acad Sci USA* 103:18822–18827
- Sharoni AM, Nuruzzaman M, Satoh K, Shimizu T, Kondoh H, Sasaya T, Choi IR, Omura T, Kikuchi S (2011) Gene structures, classification and expression models of the AP2/EREBP transcription factor family in rice. *Plant Cell Physiol* 52:344–360
- Shi X, Germain A, Hanson MR, Bentolila S (2016) RNA recognition motif-containing protein ORRM4 broadly affects mitochondrial RNA editing and impacts plant development and flowering. *Plant Physiol* 170:294–309
- Stockinger EJ, Gilmour SJ, Thomashow MF (1997) *Arabidopsis thaliana* CBF1 encodes an AP2 domain-containing transcriptional activator that binds to the C-repeat/DRE, a cis-acting DNA regulatory element that stimulates transcription in response to low temperature and water deficit. *Proc Natl Acad Sci USA* 94:1035–1040
- Swami AK, Alam SI, Sengupta N, Sarin R (2011) Differential proteomic analysis of salt stress response in *Sorghum bicolor* leaves. *Environ Exp Bot* 71:321–328
- Tajima F (1989) Statistical methods to test for nucleotide mutation hypothesis by DNA polymorphism. *Genetics* 123:585–595
- Vatansever R, Uras ME, Sen U, Ozyigit II, Filiz E (2017) Isolation of a transcription factor *DREB1A* gene from *Phaseolus vulgaris* and computational insights into its characterization: protein modeling, docking and mutagenesis. *J Biomol Struct Dyn* 35:3107–3118
- Venter M, Botha FC (2010) Synthetic promoter engineering. In: Pua EC, Davey MR (eds) *Plant developmental biology—biotechnological perspectives*. Springer, Berlin, pp 393–414
- Watterson GA (1975) On the number of segregating sites in genetical models without recombination. *Theor Popul Biol* 7:188–193
- Willard L, Ranjan A, Zhang H, Monzavi H, Boyko RF, Sykes BD, Wishart DS (2003) VADAR: a web server for quantitative evaluation of protein structure quality. *Nucleic Acids Res* 31:3316–3319
- 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
- Yamasaki M, Wright SI, McMullen MD (2007) Genomic screening for artificial selection during domestication and improvement in maize. *Ann Bot* 100:967–973
- Yan HW, Hong L, Zhou YQ, Jiang HY, Zhu SW, Fan J, Cheng BJ (2013) A genome-wide analysis of the ERF gene family in *Sorghum*. *Genet Mol Res* 12:2038–2055
- Zhang L, Zheng Y, Jagadeeswaran G, Li Y, Gowdu K, Sunkar R (2011) Identification and temporal expression analysis of conserved and novel microRNAs in *Sorghum*. *Genomics* 98:460–468
- Zhou M, Ma J, Pang J, Zhang Z, Tang Y, Wu Y (2010) Regulation of plant stress response by dehydration responsive element binding (DREB) transcription factors. *Afr J Biotech* 9:9255–9279