TÜBİTAK

Turkish Journal of Biology

http://journals.tubitak.gov.tr/biology/

Research Article

Turk J Biol
(2015) 39: 49-59
© TÜBİTAK
doi:10.3906/biy-1403-9

Genome-wide distribution of superoxide dismutase (SOD) gene families in *Sorghum bicolor*

Ertuğrul FİLİZ[1,*], Hüseyin TOMBULOĞLU[2]

[1]Department of Crop and Animal Production, Çilimli Vocational School, Düzce University, Çilimli, Düzce, Turkey
[2]Department of Biology, Faculty of Science and Arts, Fatih University, Büyükçekmece, İstanbul, Turkey

Received: 05.03.2014 • **Accepted:** 24.06.2014 • **Published Online:** 02.01.2015 • **Printed:** 30.01.2015

Abstract: Superoxide dismutases (SODs) are critical enzymes protecting cells against toxic superoxide radicals. To date, 3 types of SODs have been identified: Cu-ZnSODs, Fe-MnSODs, and NiSODs. In this study, a genome-wide analysis was performed in *Sorghum bicolor* to characterize *SOD* genes and proteins. Using several bioinformatics tools, we characterized a total of 8 *SOD* genes from the *Sorghum* genome. Gene structure, chromosomal distribution, tissue specific expression, conserved domain, and phylogenetic analyses of *SOD* genes were carried out. Additionally, 3-dimensional structures were determined and compared within each SbSOD protein. Chromosomal distributions revealed that the highest number of *SOD* genes was on chromosomes 1 and 10, with 2 members on each. Single segmental gene duplication was observed between the genes Sb*SOD2* and Sb*SOD5*. Intron numbers of Sb*SOD* genes ranged from 5 to 7. Motif analyses showed that SbSODs included 2 and 3 common motifs in Cu-ZnSOD and Fe-MnSODs, respectively. In addition, 3 functional domains were identified in SbSODs: 1) copper-zinc domain (Pfam: 00080) in Cu-ZnSOD; and 2) iron/manganese superoxide dismutases alpha-hairpin domain (Pfam: 00081) and 3) iron/manganese superoxide dismutases, C-terminal domain (Pfam: 02777) in Fe-MnSODs. Gene Ontology term enrichment analysis showed that 8 *SOD* genes have similar molecular functions and biological processes, and variable cellular components. Phylogenetic analysis revealed that Cu-ZnSODs (92%) and Fe-MnSODs (100%) were separated by high bootstrap values. Additionally, predicted motif structures and critical binding sites of SbSODs were found to be similar within each SOD group. The results of this study contribute to a better understanding of *SOD* genes and proteins in plants, especially in *Sorghum* taxa.

Key words: Superoxide dismutase, SOD, *Sorghum bicolor*, genome-wide analysis, in silico analysis

1. Introduction

Environmental stressors, including heavy metals, herbicides, salinity, drought, and pathogens, are important factors affecting plant metabolism and productivity (Mittler and Blumwald, 2010). Reactive oxygen species (ROS) such as super oxide ($O_2^{\cdot-}$), perhydroxyl radical ($HO_2^{\cdot}$), hydrogen peroxide (H_2O_2), hydroxyl radical ($OH^{\cdot}$), alkoxy radical ($RO^{\cdot}$), peroxy radical ($ROO^{\cdot}$), singlet oxygen (O_2), and organic hydroperoxide (ROOH) are some of the common responses in both biotic and abiotic stress (Miller et al., 2008; Bhattacharjee, 2012). ROS are produced in the chloroplast, mitochondria, plasma membrane, apoplast, endoplasmic reticulum, peroxisomes, and cell walls (Sharma et al., 2012). Major ROS-scavenging enzymes in plants are superoxide dismutase (SOD), ascorbate peroxidase (APX), catalase (CAT), glutathione peroxidase (GPX), and peroxiredoxin (PrxR) (Mittler et al., 2004).

The SODs (EC 1.15.1.1) catalyze the dismutation of superoxide ($O_2^{\cdot-}$) into molecular oxygen (O_2) and H_2O_2 [$2O_2^{\cdot-} + 2H^+ \rightarrow O_2 + H_2O_2$] (Qu et al., 2010). Three classes of SODs have been identified based on protein folds and catalytic metal ions: Cu-ZnSODs, Fe-MnSODs, and NiSODs (Perry et al., 2010). Plants include Cu-ZnSODs in cytoplasm, chloroplast, and peroxisomes. Chloroplastic and cytosolic Cu-ZnSODs are differentiated in terms of their numbers and positions of introns (Bueno et al., 1995; Fink and Scandalios, 2002; Miller, 2012). FeSODs are localized in chloroplasts and also have been found in cytoplasm in cowpea (Moran et al., 2003; Miller, 2012). MnSODs have been observed in mitochondria, which are essential for the survival of aerobic life (Carlioz and Touati, 1986; Miller, 2012). MnSODs may play an important role in protecting ROS in mitochondria, and hence it is expected that plant genomes contain at least one copy of MnSODs (Møller, 2001). NiSODs were first identified in *Streptomyces* (Youn et al., 1996) and are not found in plants (Wolfe-Simon et al., 2005). To date, 7 *SOD* genes were identified in *Arabidopsis*, including 1 MnSOD (*MSD1*),

* Correspondence: ertugrulfiliz@gmail.com

3 FeSOD isoforms (*FSD1*, *FSD2*, and *FSD3*), and 3 Cu-ZnSODs (*CSD1*, *CSD2*, and *CSD3*). CSD2 and FeSOD proteins were also observed in *Arabidopsis* chloroplast (Kliebenstein et al., 1998).

Sorghum bicolor L. (hereafter *Sorghum*), sorghum, is one of the important grain crops in the world. The genus *Sorghum* contains 25 recognized species with various chromosome numbers (2n = 10, 20, 30, or 40) (Lazarides et al., 1991; Price et al., 2005). Sorghum has a small genome size (~730 Mb) that makes it a good candidate for functional genomics. Among the 34,496 sorghum gene models, ~27,640 bona fide protein-coding genes were found. Approximately 24% of the genes are grass-specific and 7% are sorghum-specific (Paterson et al., 2010). In this study, we characterized a comprehensive and nonredundant set of *SOD* genes at a genome-wide scale in *Sorghum* by using bioinformatics tools. Furthermore, predicted 3D modeling, subcellular localization, and physiochemical properties of *Sorghum* SOD proteins were evaluated.

2. Materials and methods

2.1. Characterization of SOD genes in *Sorghum*

We retrieved *Arabidopsis* (accession nos. AAC24834, AEE28355, AEE28354, and AAO39917), maize (BAI50563 and BAI50561), and rice (EEC79969 and EEE60811) *SOD* protein sequences from the National Center for Biotechnology Information (NCBI) protein database (http://www.ncbi.nlm.nih.gov/protein/) to be used as query sequences. We then used the Basic Local Alignment Search Tool for Proteins (BLASTP search) for the *Sorghum* genome at the Joint Genome Institute (JGI) (http://www.phytozome.net). For every BLAST search, BLAST default settings were used, and BLAST hits were accepted significant with E-value of $\leq e^{-10}$. The Pfam server (http://pfam.sanger.ac.uk) (Sonnhammer et al., 1997) was used to confirm the inclusion of the SOD domain in each predicted SOD protein sequence. Physicochemical characteristics of SOD proteins were calculated by using the ProtParam tool (http://www.expasy.org/tools/protparam.html), containing the number of amino acids, molecular weight, and theoretical isoelectric point (*pI*).

2.2. Gene structure, duplication, and expression pattern analysis of *Sorghum SOD* genes

Sorghum SOD (Sb*SOD*) gene data, including accession number, chromosomal location, and open reading frame (ORF) length were collected from the JGI database (http://www.phytozome.net). Exon and intron structures of Sb*SOD* genes were determined by using the Gene Structure Display Server (GSDS) (http://gsds.cbi.pku.edu.cn/) (Guo et al., 2007). Transcript levels of Sb*SOD* genes were analyzed in the NCBI EST database (http://www.ncbi.nlm.nih.gov/dbEST/) by using the MEGABLAST tool. Parameters of searching were adopted as: maximum identity > 95%, length > 200 bp, and E value < 10^{-10}. In addition, we identified gene duplication based on the following criteria: 1) the length of alignable sequence cover > 80% of the longer gene; and 2) the similarity of the aligned regions > 80% (Gu et al., 2002; Yang et al., 2008).

2.3. Conserved motif and phylogenetic analyses of SbSOD proteins

The conserved motif analysis of SbSOD proteins was performed using the Multiple Expectation Maximization for Motif Elicitation (MEME) program (http://meme.nbcr.net/meme/) (Bailey et al., 2009). Alignments of SOD protein sequences were performed using the ClustalW program (Higgins et al., 1996). Phylogenetic analysis was conducted with MEGA version 5.1 (Tamura et al., 2011) by a neighbor-joining (NJ) tree method, including the following parameters: Poisson correction, pairwise deletion, and bootstrap analysis with 1000 replicates (Filiz et al., 2014).

2.4. Putative functional analysis of *Sorghum* SODs

Functional annotations of SbSOD proteins were analyzed by using the Gene Ontology (GO) term analysis tool (www.amigo.geneontology.org) (Ashburner et al., 2000). Eight *Sorghum* SOD proteins were assessed based on their molecular functions, biological processes, and cellular localizations.

2.5. Predicted 3D structure of SbSODs

The predicted 3D structures of SbSODs were generated by using the 3D LigandSite server (http://www.sbg.bio.ic.ac.uk/3dligandsite/) (Wass et al., 2010). Structural evaluation and stereochemical analyses were assessed by using RAMPAGE Ramachandran plot analysis (http://mordred.bioc.cam.ac.uk/~rapper/rampage.php) (Lovell et al., 2003).

3. Results and discussion

3.1. Genome-wide characterization of *SOD* genes in *Sorghum*

To characterize *SOD* genes, *Arabidopsis*, rice, and maize SOD protein sequences were used as queries. We detected 8 *SOD* genes in *Sorghum* by using BLASTP searching in the Phytozome v9.1 database (Table 1). The 8 SbSOD protein sequences were then used for BLASTP analysis in the NCBI database. It was observed that 3 SOD proteins (SbSOD2, -3, and -5) did not match with any *Sorghum* protein sequences, whereas 5 SOD proteins (SbSOD1, -4, -6, -7, and -8) matched with *Sorghum* hypothetical proteins. The sequences were analyzed by using the Pfam database for the presence of a SOD domain. Based on the domain analysis, 5 protein sequences had a copper-zinc domain (Pfam: 00080) and 3 SOD proteins had an iron/manganese superoxide dismutases alpha-hairpin

domain (Pfam: 00081) and an iron/manganese superoxide dismutases, C-terminal domain (Pfam: 02777) (Table 1). In *Arabidopsis*, 7 *SOD* genes were determined, including 3 Cu-ZnSODs, 3 Fe-MnSODs, and 1 MnSOD (Kliebenstein et al., 1998). In this study, the number of *Sorghum SOD*s (totally 8) was found to be the same as that of *Arabidopsis SOD*s. Physicochemical analysis of SbSOD proteins revealed that the length, molecular weight, and *pI* values of SbSOD proteins varied between 151 and 392 amino acids, 15.1 and 43.5 kDa, and 5.13 and 7.14, respectively (Table 1). All Cu-ZnSODs were acidic in character, whereas Fe-MnSODs indicated variable *pI* values: 1 member was slightly basic (SbSOD8) and 2 members (SbSOD6 and SbSOD7) were acidic in character. In rice, *pI* values (isoelectric points) were found to be acidic for Cu-ZnSOD and MnSOD proteins and basic for FeSOD proteins (Dehury et al., 2012). These findings are similar to our results. The prediction of subcellular localizations of SbSOD proteins showed that 3 members of Cu-ZnSODs (SOD1, SOD2, and SOD5) were localized in the cytoplasm, and 2 members of Cu-ZnSODs (SOD3 and SOD4) were localized in chloroplasts. Furthermore, 2 members (SOD6 and SOD8) and 1 (SbSOD7) member of Fe-MnSODs were localized in mitochondria and chloroplasts, respectively (Table 1).

3.2. Conserved motif analysis of SbSOD proteins

The MEME server was used for motif analysis and 5 types of motifs were identified as shown in the combined block diagram (Figure 1). Furthermore, motif sequences and their domain patterns are shown in Table 2. It was

Table 1. The details of *Sorghum SOD* genes and proteins, containing physiochemical, structural, and sequence properties.

Gene name	Sequence ID	Chr	ORF length (bp)	Intron number	Length (aa)	MW (kDa)	*pI*	Predicted Pfam domain	Subcellular prediction by CELLO	Subcellular prediction by PSORT
Sb*SOD1*	Sobic.001G453800.1	1	492	6	163	16.7	6.82	CZ	C	C
Sb*SOD2*	Sobic.002G407900.3	2	456	5	151	15.1	5.60	CZ	C	C
Sb*SOD3*	Sobic.006G185700.1	6	939	5	312	32.2	5.65	CZ, HMA	Ch	Ch
Sb*SOD4*	Sobic.007G166600.1	7	621	7	206	20.7	5.31	CZ	Ch	Ch
Sb*SOD5*	Sobic.001G371900.1	1	459	6	152	15.1	5.65	CZ	C	C
Sb*SOD6*	Sobic.009G093200.1	9	702	5	233	25.3	6.71	IMA, IMC	M	M
Sb*SOD7*	Sobic.010G012900.1	10	1179	7	392	43.5	5.13	IMA, IMC	Ch	Ch
Sb*SOD8*	Sobic.010G033000.1	10	786	7	261	29.9	7.14	IMA, IMC	M	M

Sb: *Sorghum bicolor*, CZ: copper/zinc superoxide dismutase (SODC), HMA: heavy-metal-associated domain; IMA: iron/manganese superoxide dismutases alpha-hairpin domain, IMC: iron/manganese superoxide dismutases, C-terminal domain; C: cytoplasm, Ch: chloroplast, M: mitochondrion.

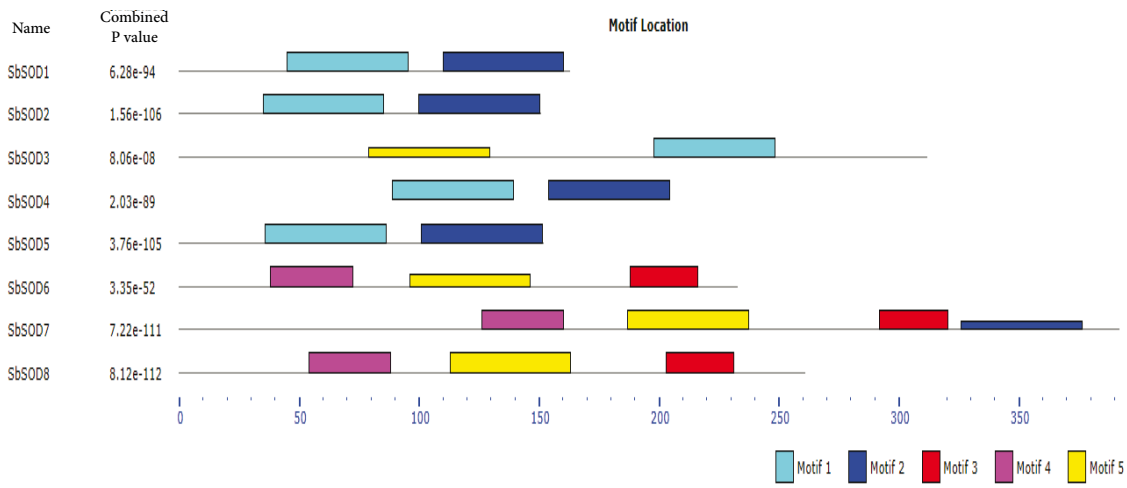

Figure 1. The most conserved protein motifs in SbSODs (motif I, motif II, motif III, motif IV, and motif V, respectively) Each motif is represented in boxes with different colors: motif 1, cyan; motif 2, blue; motif 3, red; motif 4, purple; and motif 5, yellow.

Table 2. Five different motifs commonly observed in SbSOD protein sequences with the best possible matched amino acid sequences using MEME server.

Motif number	Width sequence	Protein sequences	Pfam domain
1	50	LKPGLHGFHIHAFGDTTNGCMSTGPHFNPHNKEHGAPEDENRHAGDLGNI	Copper/zinc superoxide dismutase
2	50	QIPLSGPHSIIGRAVVVHADPDDLGKGGHELSKSTGNAGGRIACGIIGLQ	Copper/zinc superoxide dismutase
3	28	PLLGIDVWEHAYYLDYKDDRPDYISNIW	Iron/manganese superoxide dismutases C-terminal domain
4	34	PYPLDALEPYISKETMRLHWGKHHQIYVDNYNKQ	Iron/manganese superoxide dismutases, alpha-hairpin domaindomain
5	50	PPFFNAAQVWNHDFFWESMQPEGGGKPPERILQFIEKDFGSFENFREQFM	Iron/manganese superoxide dismutases, C-terminal and alpha-hairpin

observed that motif I and motif II were related to the copper/zinc superoxide dismutase domain, whereas motifs III, IV, and V were related to the iron/manganese superoxide dismutase domain. Based on motif signatures, Cu-ZnSODs and Fe-MnSODs were separated distinctively so that Cu-ZnSODs (SbSOD1, -2, -4, and -5) contained motif I and motif II (except for SbSOD3, which included motif I and motif V), while Fe-MnSODs (SbSOD5, -6, and -7) had motifs III, IV, and V. Interestingly, SbSOD7 had motif II, while SbSOD3 had motif V; these motifs were not related to the type of SODs. Most of the motif signatures, however, were supported by the classifications of SbSODs obtained from this study (Table 1). The unexpected motif signatures in SbSOD3 (putative chloroplastic Cu-ZnSOD) and SbSOD7 (putative chloroplastic FeSOD) could be explained by the enzymatic activities in response to different stress conditions of SOD types in chloroplasts.

3.3. Chromosomal locations, gene structures, and duplications of Sb*SOD* genes

Chromosomal distributions of 8 Sb*SOD* genes were determined. Accordingly, the genes were segregated among the 10 chromosomes as follows: 2 genes, 1 gene,

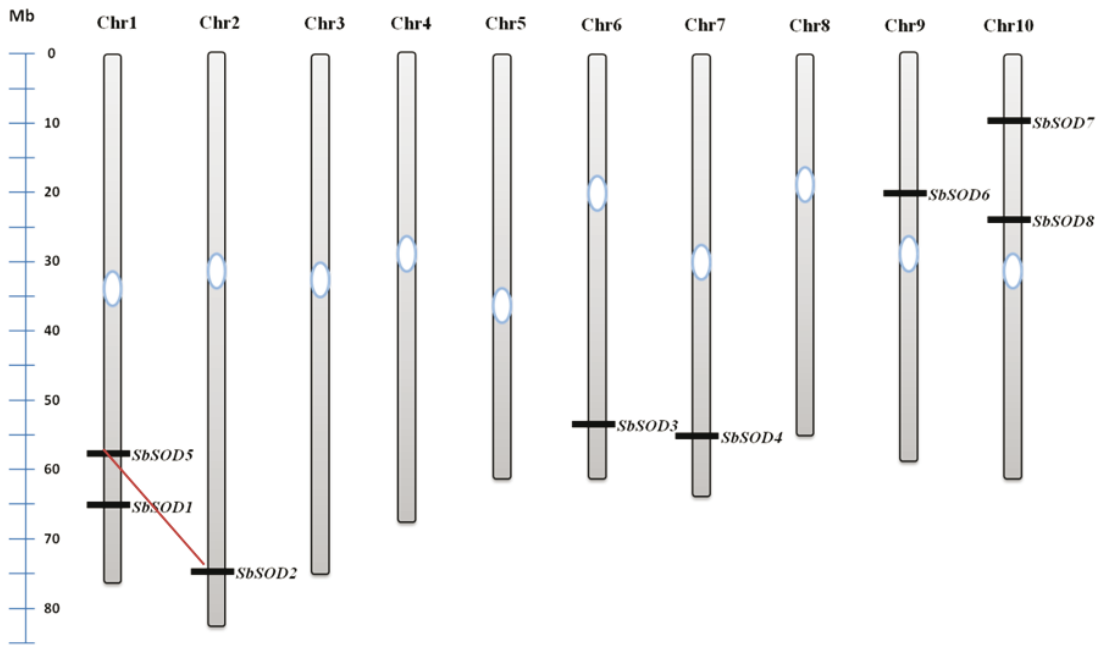

Figure 2. Chromosomal localization of 8 Sb*SOD* genes on 10 chromosomes of *Sorghum*. Red line shows segmental duplication between *SOD2* and *SOD5*. The chromosome numbers are indicated on top of chromosomes, and size of chromosome is represented using a vertical scale.

1 gene, 1 gene, 1 gene, and 2 genes on chr1, chr2, chr6, chr7, chr9, and chr10, respectively (Figure 2). The highest number of *SOD* genes was determined on chr2 and chr10 with 2 genes on each, while the others had 1 gene member. Additionally, chromosomes 3, 4, 5, and 8 do not contain any *SOD* genes.

Gene structure analysis revealed that ORF length of *SOD* genes ranged from 459 to 1176 bp (Table 1). Schematic structures of *SOD* genes were generated by the GSDS server as shown in Figure 3. Intron numbers ranged from 5 to 7 and the highest number of intron was found in Sb*SOD4*, -7, and -8 with 7 introns; in contrast, the lowest intron number was found in Sb*SOD2*, -3, and -6 with 5 introns (Table 1). Fink and Scandalios (2002) reported that plant *SOD*s have highly conserved intron patterns and all cytosolic and chloroplastic *SOD*s contain 7 introns except for 1 member. In our analysis, only 2 Sb*SOD* genes (Sb*SOD4*, chloroplastic Cu-ZnSOD and Sb*SOD7*, chloroplastic FeSOD) contained 7 introns and the rest varied. Thus, our findings did not corroborate the findings of Fink and Scandalios (2002). Divergences of exon-intron structures are shaped by 3 main mechanisms: exon/intron gain/loss, exonization/pseudoexonization, and insertion/deletion (Xu et al., 2011). The structural divergences of Sb*SOD* genes were affected by these mechanisms in the *SOD* gene evolution of *Sorghum*. In addition, these structural divergences may be related to enzyme function that responds to various biotic and abiotic stress conditions with expression pattern divergences.

Based on gene duplication analysis, one segmental duplication event was identified between Sb*SOD2* and Sb*SOD5* (Figure 2). Gene duplication could be a crucial factor for diversification. Duplicated genes appear by regional genomic events or genome-wide events (polyploidization) (Lawton-Rauh, 2003). Functional divergences in duplicated genes also contribute to molecular innovation in higher organisms (Ganko et al., 2007). It can be suggested that segmental duplications have played an important role in the expansion of *SOD* genes in *Sorghum* genome. This segmental duplication event may provide support for finer regulation of SOD activities by functional divergences in various stress conditions.

3.4. Digital expression analysis of Sb*SOD*s

Digital expression analyses of Sb*SOD* genes were performed by using the NCBI-EST database and mixed EST data were not taken into account (Table 3). Tissue-specific expressions of Sb*SOD* genes were evaluated in 4 different organs: leaf, root, flower, and stem. Exceptionally,

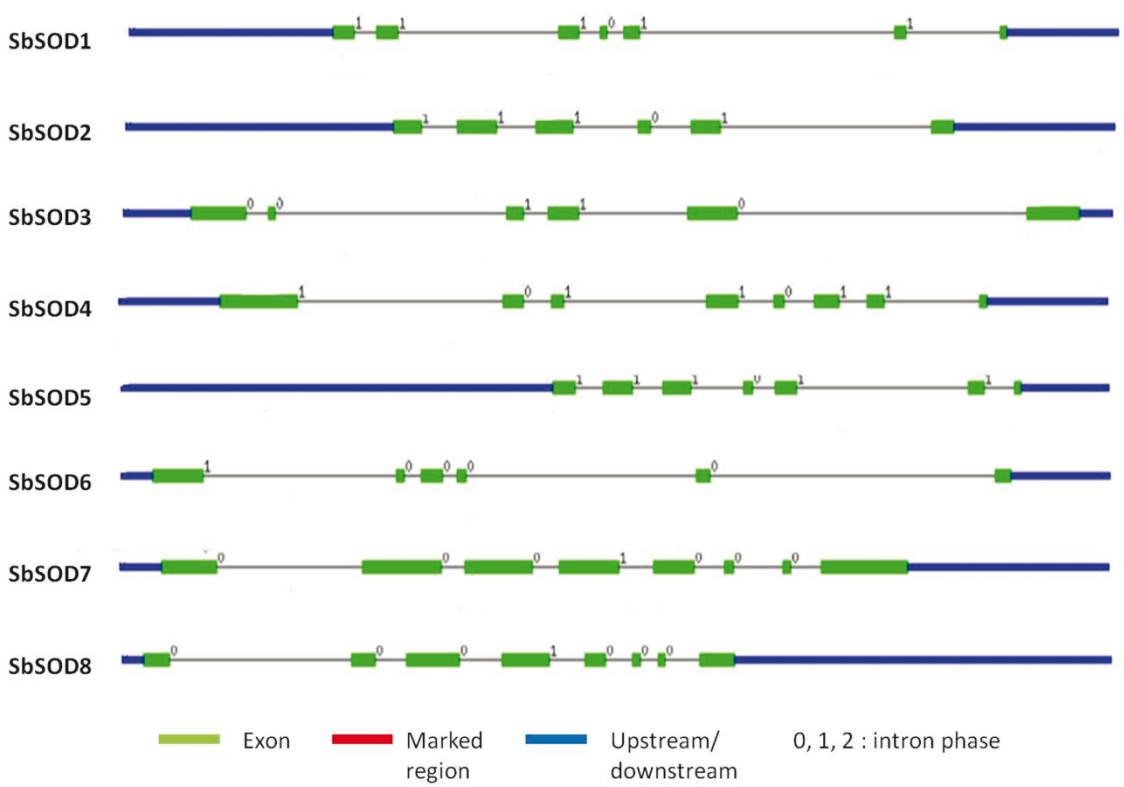

Figure 3. Exon-intron structures of *SOD* genes in *Sorghum*. Exons and introns are depicted by filled green boxes and single lines, respectively. Intron phases 0, 1, and 2 are indicated by numbers 0, 1, and 2. Untranslated regions are displayed by thick blue lines at both 5' and 3' ends of each gene.

Table 3. Digital expression analysis of Sb*SOD* genes.

Gene	Leaf	Tissue and organ type (NCBI)			
		Root	Flower	Stem	Mixed
Sb*SOD1*					
Sb*SOD2*		+	+		+
Sb*SOD3*	+	+		+	
Sb*SOD4*		+			
Sb*SOD5*		+	+		
Sb*SOD6*		+	+		+
Sb*SOD7*	+	+			
Sb*SOD8*				+	

+: Expressed; blank: not expressed.

Sb*SOD1* expression was not observed in the tissues. It can be proposed that the Sb*SOD1* gene has temporal and spatial expression pattern. Interestingly, the Sb*SOD8* gene was only expressed in stems, while the Sb*SOD4* gene was only expressed in roots. Six of the 8 *SOD* genes were found to be expressed in roots. Duplicated gene pairs (Sb*SOD2* and Sb*SOD5*) shared the same expression patterns in roots and flowers and this could indicate tissue-specific expression patterns.

3.5. GO annotations of SODs

Biological processes, molecular function, and cellular components of genes are characteristics of genes or gene products and these characteristics let us understand diverse molecular functions of proteins (Ashburner et al., 2000). GO annotations of Sb*SOD* genes showed that all Cu-ZnSODs were involved in "copper ion binding" (GO:0005507), "superoxide dismutase activity" (GO:0004784), and "zinc ion binding" (GO:0008270); in contrast, Fe-MnSODs had only "superoxide dismutase activity" (GO:0004784). According to biological processes, Cu-ZnSODs and MnSODs belonged to the "removal of superoxide radicals" (GO:0019430) group, and FeSODs were in the "oxidation-reduction process" (GO:0055114) group. The cellular components of SbSOD proteins contained "cytosol" (GO:0005829), "chloroplast" (GO:0009507), and "mitochondrion" (GO:0005759). The cellular component data were not in agreement with the subcellular prediction of some SODs (Table 1). These results may be related to protein sequence similarity caused by genomic events. Particularly, SbSOD3 contained distinctive molecular functions (superoxide dismutase copper chaperone activity) and biological processes (intracellular copper ion transport). These data revealed that SbSOD proteins generally have similar molecular functions and biological processes in GO classifications with regard to cofactor types. Consequently, functions of 8 Sb*SOD* genes were predicted by considering orthology and/or homology of *Arabidopsis* and *Oryza SOD* genes.

3.6. Phylogenetic analysis of SODs

Phylogenetic analysis was performed on 41 SOD protein sequences from 23 higher plant species, including 7 grass species (Figure 4). Phylogenetic analysis revealed that 2 major groups were obtained with Cu-ZnSODs and Fe-MnSODs. The Cu-ZnSODs cluster had 3 subgroups (A, B, and C), whereas the Fe-MnSODs cluster had 2 subgroups (D and E). Predicted *Sorghum* and other plant cytosolic SODs (SbSOD1, -2, and -5) were clustered in subgroup A with 95% and 96% bootstrap values. Subgroup B and C contained predicted chloroplastic SbSODs and other chloroplastic SODs with 98%. All Cu-ZnSODs of *Sorghum* (SbSOD1, -2, -4, and -5) were grouped with grass species except for SbSOD3. Notably, SbSOD3 was separated from the other Cu-ZnSODs and clustered in subgroup C with *G. max*, at 100% bootstrap value. GO data did not support the predicted subcellular localization of SbSOD3, which was localized in cytosol (Table 4). This SbSOD3 protein contained a copper-zinc domain (Pfam: 00080) and also had a heavy-metal-associated domain (Pfam: 00403). GO annotations revealed that SbSOD3 had superoxide dismutase copper chaperone activity (Table 4). It can be

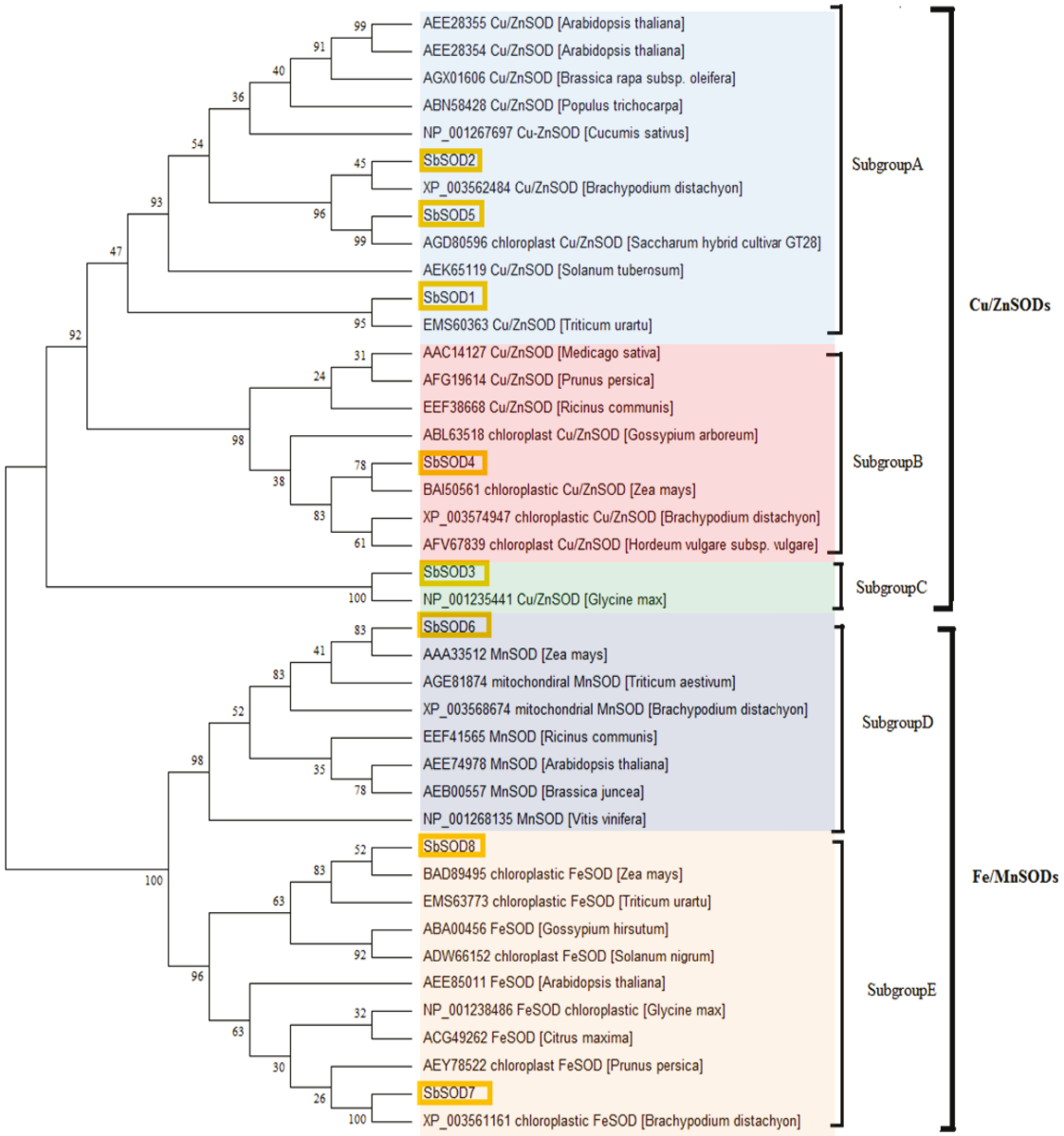

Figure 4. The phylogenetic tree of SOD protein sequences from 23 higher plant species. The rooted phylogenetic tree was conducted with MEGA5.1 by the NJ method. Bootstrap values from 1000 replicates are indicated at each branch. SOD proteins of *Sorghum* are shown as yellow boxes.

suggested that the combined domain structure of SbSOD3 caused this cluster with the highest bootstrap value (100%) in subgroup C. Plant Cu-ZnSODs contain highly homologous groups, but cytosolic and chloroplastic Cu-ZnSODs show some distinctive features (Miller, 2012). In eukaryotic and prokaryotic SOD analyses, plant cytosolic and chloroplastic SODs were separated in a joined phylogenetic tree (Fink and Scandalios, 2002). Our findings that cytosolic (subgroup A) and chloroplastic (subgroup B) SODs of *Sorghum* were separated clearly support these results.

Plant FeSODs and MnSODs were separated by the highest bootstrap value (100%), and MnSODs (subgroup D) and FeSODs (subgroup E) were grouped in 98% and 96% bootstrap values, respectively. Interestingly, predicted subcellular localization data did not support the phylogenetic data regarding SbSOD8. We expected to find it in MnSODs, but SbSOD8 was located in FeSODs in subgroup E. However, GO annotations supported the chloroplastic activities of SbSOD8 (Table 4). SbSOD6 (accepted as MnSOD) and SbSOD7 (accepted as FeSOD)

Table 4. Putative functions and cellular localizations of SOD proteins in *Sorghum*.

Gene name	Sequence ID	GO: Molecular function	GO: Biological process	GO: Cellular component
Sb*SOD1*	Sobic.001G453800.1	Copper ion binding, superoxide dismutase activity, zinc ion binding	Removal of superoxide radicals	Cytosol
Sb*SOD2*	Sobic.002G407900.3	Copper ion binding, superoxide dismutase activity, zinc ion binding	Removal of superoxide radicals	Cytosol
Sb*SOD3*	Sobic.006G185700.1	Copper ion binding, superoxide dismutase activity, ion binding , superoxide dismutase copper chaperone activity	Intracellular copper ion transport, removal of superoxide radicals	Cytosol
Sb*SOD4*	Sobic.007G166600.1	Copper ion binding, superoxide dismutase activity, zinc ion binding	Removal of superoxide radicals	Chloroplast
Sb*SOD5*	Sobic.001G371900.1	Copper ion binding, superoxide dismutase activity, zinc ion binding	Removal of superoxide radicals	Cytosol
Sb*SOD6*	Sobic.009G093200.1	Superoxide dismutase activity	Removal of superoxide radicals	Mitochondrion
Sb*SOD7*	Sobic.010G012900.1	Superoxide dismutase activity	Oxidation-reduction process	Chloroplast
Sb*SOD8*	Sobic.010G033000.1	Superoxide dismutase activity	Oxidation-reduction process	Chloroplast

were also in agreement with predicted subcellular localization data. Fink and Scandalios (2002) reported that FeSODs and MnSODs were separated in plants. Plant MnSODs have 70% homology but are not so similar to plant FeSODs, suggesting different ancestral gene origination (Miller, 2012). These data support our results that MnSOD and FeSODs of *Sorghum* were diverged with 100% bootstrap value.

3.7. Predicted 3D structures of Sb*SODs*

Eukaryotic Cu-ZnSODs contain a highly conserved quaternary structure with 2 identical subunits, including a β-barrel consisting of 8 antiparallel β-strands in a Greek key motif. The Cu and Zn sites are located outside the β-barrel in the active site channel (Richardson, 1977; Perry et al., 2010). In *Sorghum* Cu-ZnSODs it was observed that the β-barrel structures and predicted binding sites were positioned outside the β-barrel (Figures 5A–5E). The conserved disulfide bonds (between Cys-Cys) were observed in the active site channel of the Cu-ZnSODs (Perry et al., 2010). The disulfide residues were determined in *Sorghum* Cu-ZnSODs, including Cys65/Cys154, Cys55/Cys144, Cys100/Cys301, Cys109/Cys198, and Cys56/Cys145 in SbSOD1, SbSOD2, SbSOD3, SbSOD4, and SbSOD5, respectively. It can be stated that these cysteine residues contribute to the stability of Cu-ZnSODs. In Fe-MnSODs of rice, 3 antiparallel β-sheets dominated by α-helices were identified (Dehury et al., 2012). Similar β-sheets and α-helices were observed in soybean Fe-MnSODs (Gopavajhula et al., 2013). The analysis of Fe-MnSODs of *Sorghum* revealed that the dominant α-helices and β-sheets structures were observed as rice and soybean (Figures 5F–5H).

In model validation, the Ramachandran plot analysis using the RAMPAGE server showed that 98.7%, 93.3%, 85.9%, 86.2%, 95.3%, 93.5%, 93.9%, and 94.6% were in the favored region; 0.7%, 5.4%, 10.9%, 8.6%, 4.7%, 4.5%, 4.5%, and 7.9% in the allowed region; and 0.7%, 1.3%, 3.2%, 5.3%, 0.0%, 2%, 1.5%, and 1.5% in the outlier region in SbSOD1, SbSOD2, SbSOD3, SbSOD4, SbSOD5, SbSOD6, SbSOD7, and SbSOD8, respectively (Figure 6), indicating that the 3D models were fairly good in quality.

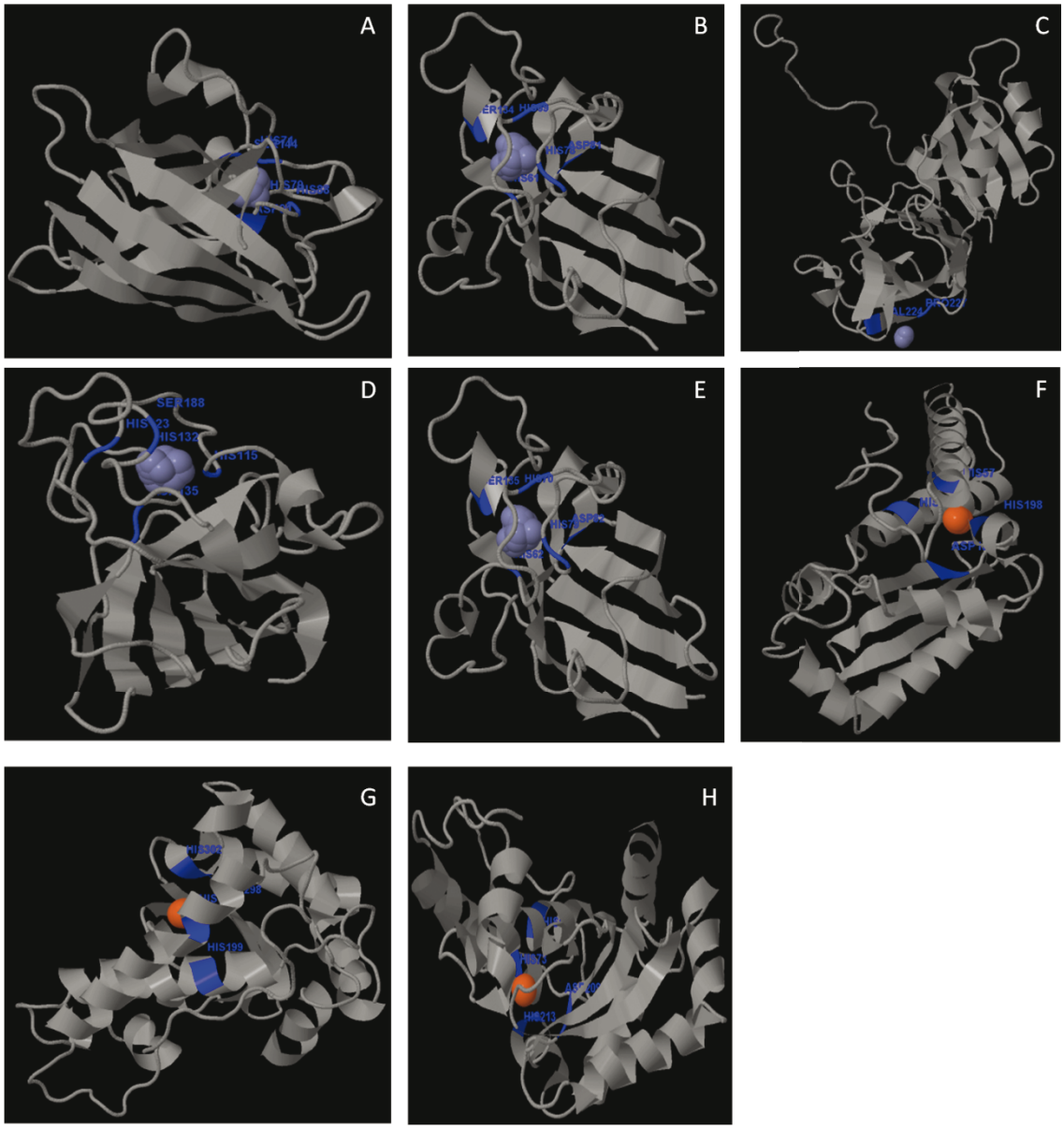

Figure 5. Predicted 3D structures and binding sites of *Sorghum* SODs by using the 3D LigandSite server (**A**: SbSOD1, Cu-ZnSOD cytosolic; **B**: SbSOD2, Cu-ZnSOD cytosolic; **C**: SbSOD3, Cu-ZnSOD cytosolic; **D**: SbSOD4, Cu/ZnSOD chloroplastic; **E**: SbSOD5, Cu/ZnSOD cytosolic; **F**: SbSOD6, Fe-MnSOD mitochondrial; **G**: SbSOD7, Fe-MnSOD chloroplastic; **H**: SbSOD8, Fe-MnSOD chloroplastic).

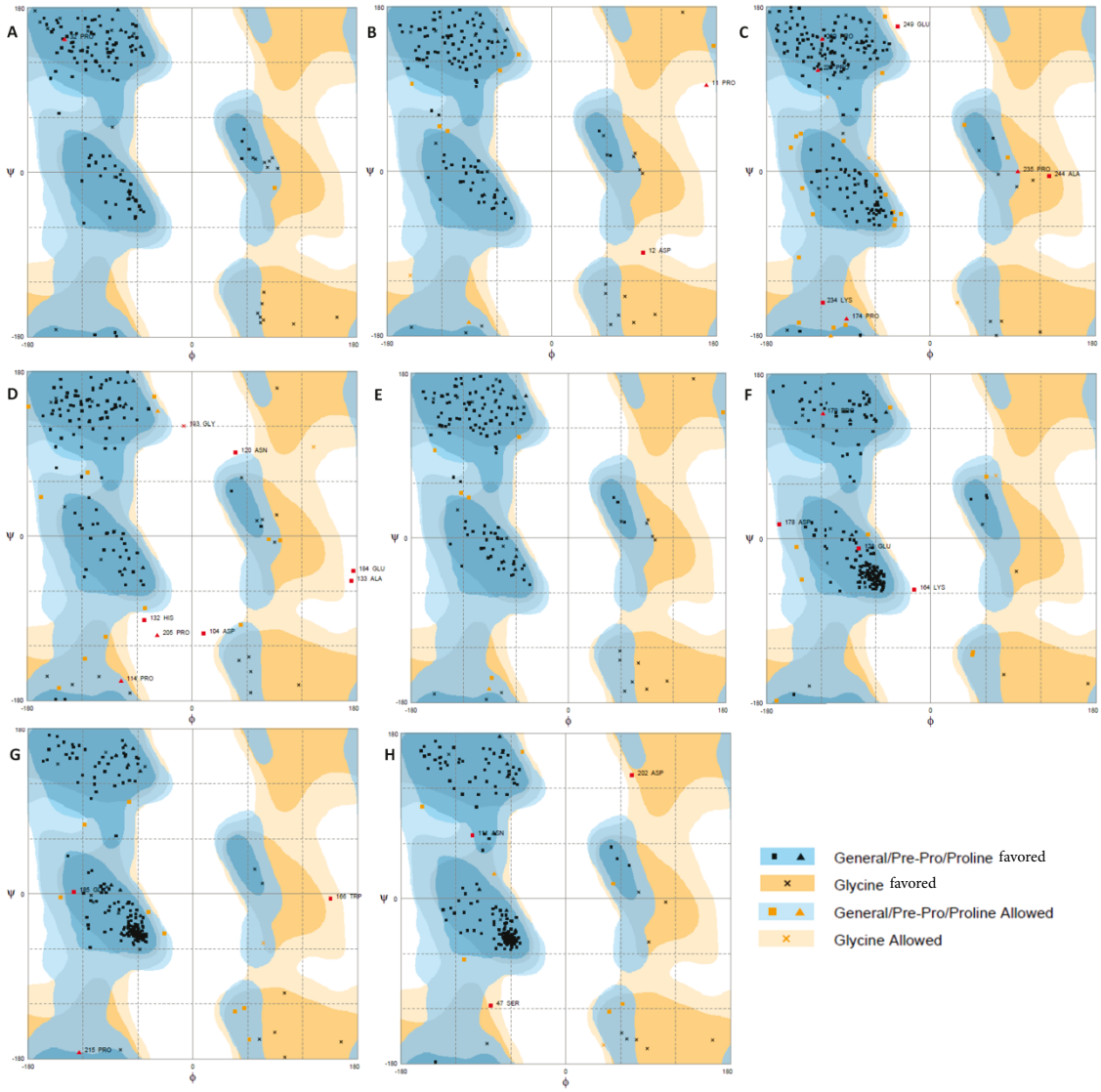

Figure 6. Ramachandran plots of *Sorghum* Cu-ZnSODs (**A**, **B**, **C**, **D**, and **E**) and Fe-MnSODs (**F**, **G**, and **H**).

In conclusion, we have analyzed the *SOD* gene family in *Sorghum*, including genome-wide characterization, chromosomal location, gene structure, digital expression and phylogenetic analysis, and 3D modeling. Our findings can contribute to the understanding of *SOD* genes in *Sorghum* and other grasses. More wet-lab studies are needed to confirm the functional analyses of Sb*SOD* genes in various stress conditions.

References

Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT et al. (2000). Gene Ontology: tool for the unification of biology. Nat Genet 25: 25–29.

Bailey TL, Boden M, Buske FA, Frith M, Grant C, Clementi L, Ren J, Li WW, Noble WS (2009). MEME SUITE: tools for motif discovery and searching. Nucl Acids Res 37: 202–208.

Bhattacharjee S (2012). The language of reactive oxygen species signaling in plants. Journal of Botany 2012: 985298.

Bueno P, Varela J, Giménez-Gallego G, Del Río LA (1995). Peroxisomal copper, zinc superoxide dismutase-characterization of the isoenzyme from watermelon cotyledons. Plant Physiol 108: 1151–1160.

Carlioz A, Touati D (1986). Isolation of superoxide dismutase mutants in *Escherichia coli*: is superoxide dismutase necessary for aerobic life? EMBO J 5: 623–630.

Dehury B, Sarma K, Sarmah R, Sahu J, Sahoo S, Sahu M, Sen P, Modi KM, Barooah M (2012). In silico analyses of superoxide dismutases (SODs) of rice (*Oryza sativa* L.). J Plant Biochem Biotechnol 22: 150–156.

Filiz E, Koç I, Tombuloğlu H (2014). Genome-wide identification and analysis of growth regulating factor genes in *Brachypodium distachyon*: in silico approaches. Turk J Biol 38: 296–306.

Fink RC, Scandalios JG (2002). Molecular evolution and structure-function relationships of the superoxide dismutase gene families in angiosperms and their relationship to other eukaryotic and prokaryotic superoxide dismutases. Arch Biochem Biophys 399: 19–36.

Ganko EW, Meyers BC, Vision TJ (2007). Divergence in expression between duplicated genes in *Arabidopsis*. Mol Biol Evol 24: 2298–2309.

Gu Z, Cavalcanti A, Chen FC, Bouman P, Li WH (2002). Extent of gene duplication in the genomes of *Drosophila*, nematode, and yeast. Mol Biol Evol 19: 256–262.

Guo AY, Zhu QH, Chen X, Luo JC (2007). GSDS: a gene structure display server. Yi Chuan 29: 1023–1026.

Higgins DG, Thompson JD, Gibson TJ (1996). Using CLUSTAL for multiple sequence alignments. Methods Enzymol 266: 383–402.

Kliebenstein DJ, Monde RA, Last RL (1998). Superoxide dismutase in *Arabidopsis*: an eclectic enzyme family with disparate regulation and protein localization. Plant Physiol 118: 637–650.

Lawton-Rauh A (2003). Evolutionary dynamics of duplicated genes in plants. Mol Phylogenet Evol 29: 396–409.

Lazarides M, Hacker JB, Andrew MH (1991). Taxonomy, cytology and ecology of indigenous Australian sorghums (*Sorghum* Moench: Andropogoneae: Poaceae). Aust Syst Bot 4: 591–635.

Lovell SC, Davis IW, Bryan Arendall W, Bakker PIW, Word JM, Prisant MG, Richardson JS, Richardson DC (2003). Structure validation by C_α geometry: φ, ψ, and C_β deviation. Proteins 50: 437–450.

Miller AF (2012). Superoxide dismutases: ancient enzymes and new insights. FEBS Lett 586: 585–595.

Miller G, Shulaev V, Mittler R (2008). Reactive oxygen signaling and abiotic stress. Physiol Plantarum 133: 481–489.

Mittler R, Blumwald E (2010). Genetic engineering form modern agriculture: challenges and perspectives. Annu Rev Plant Biol 61: 443–462.

Mittler R, Vanderauwera S, Gollery M, Breusegem FN (2004). Reactive oxygen gene network of plants. Trends Plant Sci 9: 490–498.

Møller IM (2001). Plant mitochondria and oxidative stress: electron transport, NADPH turnover, and metabolism of reactive oxygen species. Annu Rev Plant Physiol Plant Mol Biol 52: 561–91.

Moran JF, James EK, Rubio MC, Sarath G, Klucas RV, Becana M (2003). Functional characterization and expression of a cytosolic iron superoxide dismutase from cowpeak root nodules. Plant Physiol 133: 773–782.

Paterson AH, Bowers JE, Bruggmann R, Dubchak I, Grimwood J, Gundlach H, Haberer G, Hellsten U, Mitros T, Poliakov A et al. (2009). The *Sorghum bicolor* genome and the diversification of grasses. Nature 457: 551–556.

Perry JJP, Shin DS, Getzoff ED, Tainer JA (2010). The structural biochemistry of the superoxide dismutases. Biochim Biophys Acta 1804: 245–262.

Price HJ, Dillon SN, Hodnett G, Rooney WL, Ross L, Johnston JS (2005). Genome evolution in the genus *Sorghum* (Poaceae). Ann Bot 95: 219–227.

Qu CP, Xu ZR, Liu GJ, Liu C, Li Y, Wei ZG, Liu GF (2010). Differential expression of copper-zinc superoxide dismutase gene of *Polygonum sibiricum* leaves, stems, and underground stems, subjected to high-salt stress. Int J Mol Sci 11: 5234–5245.

Sharma P, Jha AB, Dubey RS, Pessarakli M (2012). Reactive oxygen species, oxidative damage, and antioxidative defense mechanism in plants under stressful conditions. Journal of Botany 2012: 217037.

Sonnhammer EL, Eddy SR, Durbin R (1997). Pfam: a comprehensive database of protein domain families based on seed alignments. Proteins 28: 405–420.

Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S (2011). MEGA5: Molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol Biol Evol 28: 2731–2739.

Wass MN, Kelley LA, Sternberg MJ (2010). 3DLigandSite: predicting ligand-binding sites using similar structures. Nucl Acid Res 38 (Suppl.): W469–473.

Wolfe-Simon F, Schofield DGO, Falkowski PG (2005). The role and evolution of superoxide dismutase in algae. J Phycol 41: 453–465.

Xu G, Guo C, Shan H, Kong H (2012). Divergence of duplicate genes in exon–intron structure. P Natl Acad Sci USA 109: 1187–1192.

Yang S, Zhang X, Yue JX, Tian D, Chen JQ (2008). Recent duplications dominate NBS-encoding gene expansion in two woody species. Mol Genet Genomics 280: 187–198.

Youn HD, EJ Kim, JH Roe, YC Hah, SO Kang (1996). A novel nickel-containing superoxide dismutase from *Streptomyces* spp. Biochem J 318: 889–896.