Pak. J. Bot., 53(3): 905-914, 2021. DOI: http://dx.doi.org/10.30848/PJB2021-3(37)

IDENTIFICATION AND EXPRESSION ANALYSIS OF THE AUXIN RESPONSE FACTOR (*ARF*) GENE FAMILY IN SORGHUM [*SORGHUM BICOLOR* (L.) MOENCH.]

LING ZHENG, JI ZHOU, LI YI AND XIAOJUN ZHOU*

Department of Biology, Luoyang Normal University, Luoyang, NC 471934, China
**Corresponding author's email: lynubio@126.com*

Abstract

Many ARF genes have been identified in plants, yet there remains a limitation of comprehensive analyses of this gene family in sorghum [*Sorghum bicolor* (L.) Moench]. In this study, we identified a total of 46 *S. bicolor ARF (SbARF)* genes by bioinformatics methods. Then, we analyzed this gene family in terms of conserved domain, phylogenetic relationship, chromosome location, gene structure, and expression pattern. The results showed that All *SbARF* genes could be divided into four subfamilies (classes I–IV) according to their relationship in *Arabidopsis thaliana*. Each of the *SbARF* genes consisted of one to 13 introns, with three highly conserved regions of the ARF, B3, and auxin/indole-3-acetic acid (Aux/IAA) domains, respectively. Class III genes were mainly expressed in 10 different tissues of sorghum, indicating important roles in growth of leaves, flowers and seeds. Expression levels of some *SbARF* genes were significantly upregulated or downregulated in roots or shoots of sorghum under stress induced by exogenous abscisic acid or polyethylene glycol, indicating critical roles in abiotic stress responses. This study provides insight into the role of the *ARF* gene family in the growth, development, and stress responses of sorghum.

Key words: Auxin response factor, Abiotic stress response, Abscisic acid, Polyethylene- glycol, Sorghum.

Introduction

Auxin is a plant hormone that regulates a vast array of physiological processes in higher plants including apical dominance, vascular tissue differentiation, lateral root development, embryo formation, flowering, and fruit ripening (Wu *et al.*, 2011). To date, auxin has been widely applied as a master regulator in plant tissue culture and agricultural planting production. ARFs accept auxin signals and thereby initiate or inhibit specific expression of downstream genes (Feng *et al.*, 2018).

Generally, the molecular weight of ARF proteins vary from 67 kDa to129 kDa (Guan *et al.*, 2016). Most ARF proteins contain three regions: an N-terminal DNA-binding domain, a middle region that functions as an activation or repression domain, and a C-terminal interaction domain that dimerizes with the auxin/indole-3-acetic acid (Aux/IAA) gene family (Tiwari *et al.*, 2003). The activation domain is rich in serine, glutamine, and leucine residues, while the repression domain is abundant in proline, serine, glycine, and leucine residues (Guilfoyle *et al.*, 2007).

The model plant *Arabidopsis thaliana* is used by the earliest study on the biological functions of ARFs. For example, *AtARF1* and *AtARF2* regulate flowering time, seed volume and leaf senescence (Ellis *et al.*, 2005); *AtARF3* participates in leaf polarity through the control of KANAD protein synthesis (Ellis *et al.*, 2005); *AtARF5* affects hypocotyl formation as well as vascular tissue growth and development (Nagpal *et al.*, 2005); *AtARF7* and *AtARF19* are transcriptional activators that mediate lateral root formation (Okushima *et al.*, 2005); and *AtARF6* and *AtARF8* are involved in the biosynthesis of jasmonic acid associated with regulation of stress responses (Schruff *et al.*, 2006; Wu *et al.*, 2006). In rice (*Oryza sativa*), *OsARF23* plays a critical role in vegetative organ growth and seed development (Wang *et al.*, 2007; Attia *et al.*, 2009). In tomato (*Lycopersicon esculentum* Mill.), *SlARF4* modulates carbohydrate metabolism during fruit development, while *SlARF7* negatively regulates fruit initiation and controls the auxin response during fruit growth (De *et al.*, 2009).

In 2005, a total of 23 *ARF* genes were identified in *Arabidopsis* (Okushima *et al.*, 2005). Subsequently, many *ARF* gene family members were also identified in other plants; 18 were identified in cucumber (*Cucumis sativus* L.) (Sheng *et al.*, 2014) , 11 in papaya (*Carica papaya* L.) (Liu *et al.*, 2015), 29 in apple (*Malus domestica*) (Luo *et al.*, 2014), 20 in pepper (*Capsicum annuum* L.) (Wei *et al.*, 2017) and 47 in switchgrass (*Panicum virgatum* L.) (Wang *et al.*, 2018). Sorghum [*Sorghum bicolor* (L.) Moench] is a special species in the family Poaceae, which is tolerant to barren, drought, and saline-alkali conditions (Zhang *et al.*, 2014); it is of great significance to future bioenergy production. The completion of the sequencing of the sorghum genome has provided the potential to identify and analyze *ARF* genes within it, yet no such study has been reported.

In our study, we conducted a comprehensive analysis of the *ARF* gene family in sorghum. Based on available studies of *ARF* gene functions in *Arabidopsis* (Ellis *et al.*, 2005; Okushima *et al.*, 2005; Schruff *et al.*, 2006; Nagpal *et al.*, 2005), we proposed the functions of some sorghum *ARF* genes. Moreover, we analyzed the expression patterns of sorghum *ARF* genes in different tissues and in response to abiotic stresses. This work is useful to study the role of *ARF* genes in the growth, development, and stress responses of sorghum.

Materials and Methods

Sequence retrieval and gene identification: The conserved ARF domain (PF06507) was obtained from the Pfam database (http://pfam.sanger.ac.uk/). Sequence PF06507 was employed to search against the sorghum genome database in Phytozome (http://phytozome.net/) to download all *SbARFs* (E-value $< 1 \times 10^{-10}$). The downloaded data were aligned with sequences retrieved from the PlantTFDB database

(http://planttfdb.cbi.pku.edu.cn/) and candidate genes were obtained after removing redundancy. The candidate genes were validated online using InterPro (http://www.ebi.ac.uk/interpro/) to identify whether they contained the ARF domain. Finnaly, selected amino acid sequences were analyzed online using ExPASy (http//www.expasy.org/) to determine the molecular weight, isoelectric point, subcellular localization, instability index and grand average of hydropathicity (GRAVY) of the ARF proteins. Information on subcellular localization was obtained using the Plant-PLoc online software (http://www.csbio.sjtu.edu.cn/bioinf/plant/).

Conserved domain analysis: To identify conserved domains of ARF proteins, multiple sequence alignments were performed using ClustalX 2.0 and visualized using GENEDOC (Nicholas *et al*., 1997).

Phylogenetic analysis: A total of 68 ARF amino acid sequences from *Arabidopsis* and sorghum were subjected to multiple sequence alignments using ClustalW in MEGA 6.0 (Tamura *et al.*, 2013). Phylogenetic trees were constructed using the neighbor-joining method in MEGA 6.0.

Gene structure and conserved motif analysis: Gene sequences, cDNA sequences, and coding sequences of sorghum *ARF*s were downloaded from the Phytozome database. Intron gene structure was displayed using the GSDS 2.0 online software (http://gsds.cbi.pku.edu.cn/). Motif analysis was performed via the MEME website (http://meme-suite.org/tools/meme), and the number of motifs that MEME should find was set to 10.

Chromosome localization analysis: The physical locations of *ARF* genes on sorghum chromosomes were obtained from the Phytozome database and visualized using MapInspect (Zhang *et al.*, 2013).

Expression pattern analysis: Transcriptional levels of sorghum *ARF* genes in different tissues and in response to different abiotic stresses, induced by treatment with 20 µm abscisic acid (ABA) or 20% polyethylene glycol (PEG) for eight days, were retrieved from the qTeller database (http://qteller.com/) and visualized using MEV4 (Saeed *et al.*, 2006).

Results

Identification of *ARF* genes in sorghum: A total of 46 *S. bicolor ARF* genes (*SbARF1* to *SbARF46*) were identified. We found that the length of SbARF proteins were 518 (SbARF32) to 1159 (SbARF45) amino acids (Table 1). The predicted molecular weights of which ranged from 56284.76 kDa (SbARF32) to 128163.34 kDa (SbARF45). The isoelectric point of the deduced proteins varied from 5.51 (SbARF19) to 9.16 (SbARF41), which was >7 (alkaline) for seven proteins and <7 (acidic) for the remaining 39 proteins. All deduced proteins had an instability index of >40 and thus were unstable. The deduced SbARF17 protein had a positive GRAVY value, representing a hydrophilic protein; all the others were found to be hydrophobic proteins. With regards to subcellular localization, 28 SbARFs were found to be located in the nucleus and 18 on the chloroplast.

Conserved protein domains of SbARFs: Based on the conserved domain analysis, we found that all 46 SbARFs contained a B3 domain at the N-terminus and an ARF domain in the middle region. In addition, 21 SbARFs contained an Aux/IAA domain at the C-terminus (Figs. 1-3). These three domains were generally conserved in sorghum. In the ARF domain, five amino acids were completely conserved in all sequences, covering 5.4% of the total length. In the B3 domain, eight amino acids were completely conserved in all sequences, covering 7.6% of the total length. In the Aux/IAA domain, 10 amino acids were completely conserved in the 21 sequences, covering 25% of the total length.

Phylogenetic relationships of SbARFs: We generated a neighbor-joining tree based on 46 SbARFs and 22 AtARFs (Fig. 4). The 46 SbARFs could be assigned into four subfamilies alongside AtARFs, with 17 SbARFs and two AtARFs in class I, six SbARFs and three AtARFs in class II, nine SbARFs and 12 AtARFs in class III, and 14 SbARFs and five AtARFs in class IV. Many sorghum ARFs were clustered together with their homologs from *Arabidopsis*, indicating close phylogenetic relationships.

Gene structures and conserved motifs of SbARFs: We found at least one intron in each of the 46 *SbARF* genes; the intron number varied from one to 13 (Fig. 5, Table 2). The intron number was smallest in class II genes, at only one or two. The largest intron numbers appeared in class III and class IV genes, at between 11 and 13. Class I genes contained nine to 11 introns each. Additionally, a total of 10 conserved motifs were identified (Table 3), designated motif1 to motif10. These motifs were located in proteins SbARF10 to SbARF46, with a width of 16–50 amino acids. There were considerable differences in the type and number of motifs found across the four subfamilies of the 46 SbARFs (Fig. 5). All protein sequences contained motif1, motif2 and motif10. However, motif8 and motif9 were only present in nine SbARFs of class II, while all members of this subfamily lacked motif7. Motif4 and motif10 belonged to the ARF domain, motif2 belonged to the B3 domain and motif7 belonged to the Aux/IAA domain. Other motifs are unknown proteins.

Chromosome locations of *SbARF*s: As shown in Fig. 6, the *SbARF* genes were unevenly distributed across 10 chromosomes, with the highest concentration of genes being 15 on chromosome 3. Chromosome 6 had the second highest concentration of *SbARF* genes, carrying eight. Chromosomes 4, 9 and 10 carried five *SbARF* genes each. Fewer *SbARF* genes appeared on chromosome 5 (two only), while the fewest genes were found on chromosomes 1, 2 and 7 (one each). Excluding a small number of *SbARF* genes located at the upper end of chromosomes 3, 4 and 10, all the remaining genes were mainly found at the lower end of the chromosome. In addition, there were prominent gene clusters on chromosomes 3, 4, 5, 6, 9, and 10.

Table 1. The auxin response factor (*ARF*) gene family in sorghum.

Gene name	Gene symbol	Protein length (amino acids)	Molecular weight (Da)	Isoelectric point	Subcellular localization	Instability index	GRAVY
SbARF1	Sobic.001G217300.1	689	75404.92	6.49	Chloroplast	45.51	-0.275
SbARF2	Sobic.002G290600.1	663	72693.19	8.22	Chloroplast	64.46	-0.459
SbARF3	Sobic.003G003800.1	688	76895.05	5.97	Nucleus	51.72	-0.456
SbARF4	Sobic.003G251700.1	702	77003.66	5.99	Nucleus	57.76	-0.415
SbARF5	Sobic.003G251700.2	673	73888.33	6.11	Chloroplast	57.04	-0.411
SbARF6	Sobic.003G298600.1	622	67843.54	6.25	Chloroplast	55.46	-0.395
SbARF7	Sobic.003G298600.1	685	74377.98	6.75	Chloroplast	54.26	-0.364
SbARF8	Sobic.003G298600.2	682	74143.72	6.75	Chloroplast	54.66	-0.370
SbARF9	Sobic.003G298600.3	651	70835.04	6.50	Chloroplast	54.99	-0.380
SbARF10	Sobic.003G298600.4	654	71069.30	6.50	Chloroplast	54.57	-0.374
SbARF11	Sobic.003G298600.5	654	71069.30	6.50	Chloroplast	54.57	-0.374
SbARF12	Sobic.003G298600.6	651	70835.04	6.50	54.99		-0.380
SbARF13	Sobic.003G298600.7	625	68077.79	6.25	Chloroplast	55.01	-0.389
SbARF14	Sobic.003G298600.8	625	68077.79	6.25	Chloroplast	55.01	-0.389
SbARF15	Sobic.003G298600.9	622	67843.54	6.25	Chloroplast	55.46	-0.395
SbARF16	Sobic.003G411900.1	810	90774.62	6.04	Nucleus	60.81	-0.650
SbARF17	Sobic.003G411900.2	704	78702.14	6.96	Nucleus	65.31	0.643
SbARF18	Sobic.004G037800.1	1070	119870.15	6.01	Nucleus	57.99	-0.546
SbARF19	Sobic.004G051900.1	911	100417.56	5.51	Nucleus	68.82	-0.432
SbARF20	Sobic.004G051900.3	844	93419.80	5.58	Nucleus	67.65	-0.430
SbARF21	Sobic.004G178500.1	672	74930.18	5.87	Nucleus	60.63	-0.579
SbARF22	Sobic.004G221400.1	708	76087.88	6.63	Chloroplast	50.93	-0.270
SbARF23	Sobic.005G132000.1	815	90621.08	6.32	Nucleus	53.59	-0.603
SbARF24	Sobic.005G132000.2	814	90492.95	6.32	Nucleus	53.03	-0.599
SbARF25	Sobic.006G089500.1	661	73201.76	5.80	Nucleus	58.48	-0.423
SbARF26	Sobic.006G149600.1	722	78108.18	8.27	Chloroplast	60.15	-0.357
SbARF27	Sobic.006G255300.1	946	103749.51	5.85	Nucleus	60.89	-0.462
SbARF28	Sobic.006G255300.2	945	103692.45	5.85	Nucleus	61.24	-0.462
SbARF29	Sobic.006G255300.3	946	103749.51	5.85	Nucleus	60.89	-0.462
SbARF30	Sobic.006G255300.4	945	103692.45	5.85	Nucleus	61.24	-0.462
SbARF31	Sobic.006G262100.1	819	91051.22	5.94	Nucleus	66.51	-0.472
SbARF32	Sobic.006G278900.1	518	56284.76	5.70	Chloroplast	47.25	-0.305
SbARF33	Sobic.007G203500.1	1095	121248.57	6.11	Nucleus	58.34	-0.455
SbARF34	Sobic.008G096000.1	839	92454.03	6.27	Nucleus	60.06	-0.648
SbARF35	Sobic.008G096000.2	838	92325.85	6.21	Nucleus	60.12	-0.644
SbARF36	Sobic.008G169400.1	895	98609.50	5.65	Nucleus	65.99	-0.409
SbARF37	Sobic.009G196900.1	676	73463.61	6.23	Chloroplast	43.93	-0.364
SbARF38	Sobic.009G196900.2	575	63169.94	7.25	Nucleus	42.72	-0.467
SbARF39	Sobic.009G196900.3	575	63169.94	7.25	Nucleus	42.72	-0.467
SbARF40	Sobic.009G231800.1	739	80854.09	7.58	Nucleus	50.82	-0.478
SbARF41	Sobic.009G231800.2	573	63788.17	9.16	Nucleus	47.43	-0.539
SbARF42	Sobic.010G073600.1	1053	116455.26	6.16	Nucleus	70.43	-0.502
SbARF43	Sobic.010G229000.1	919	101692.05	5.95	Nucleus	70.19	-0.469
SbARF44	Sobic.010G236300.1	709	76700.16	7.37	Chloroplast	46.51	-0.307
SbARF45	Sobic.010G253300.1	1159	128163.34	6.27	Nucleus	64.18	-0.551
SbARF46	Sobic.010G253300.2	1075	119669.69	6.26	Nucleus	61.92	-0.574

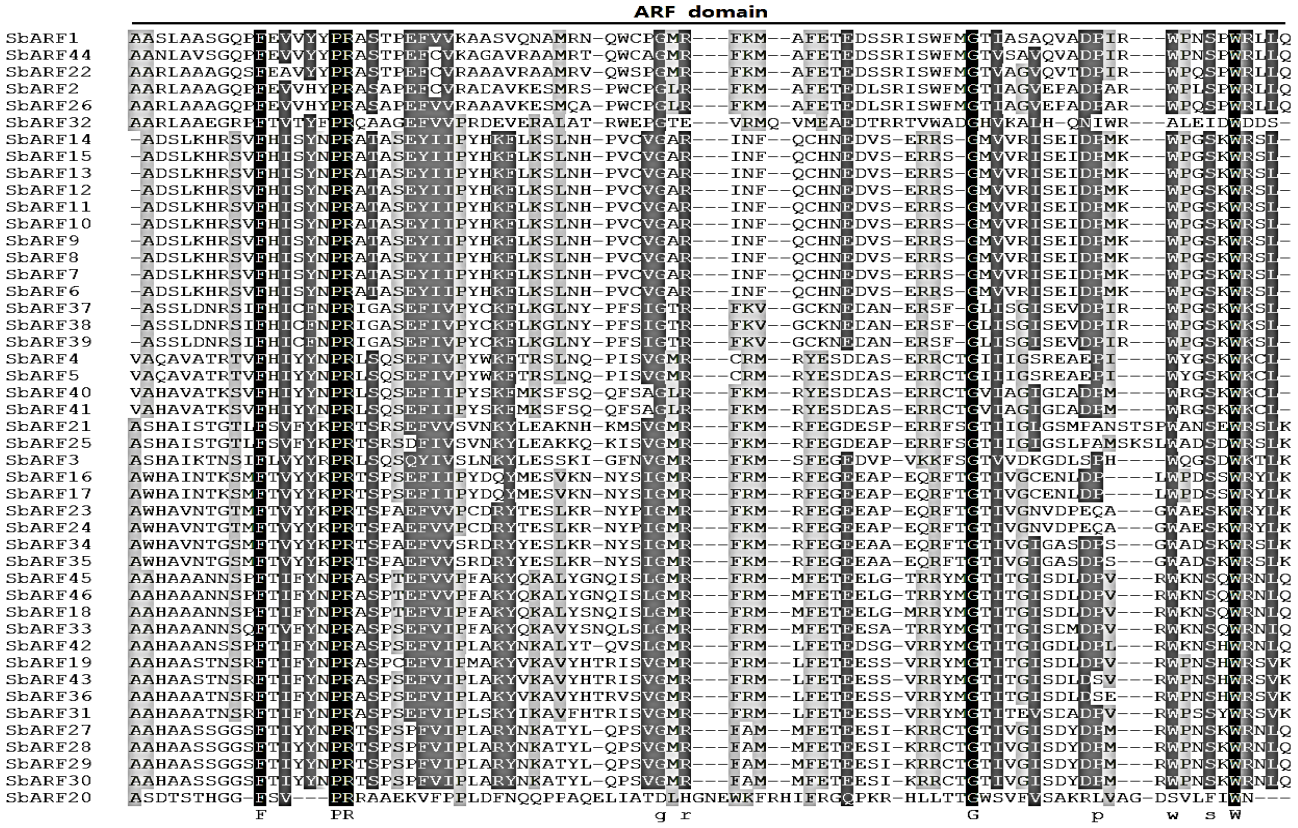

Fig. 1. Sequence alignments of the ARF domain of ARF proteins in sorghum.

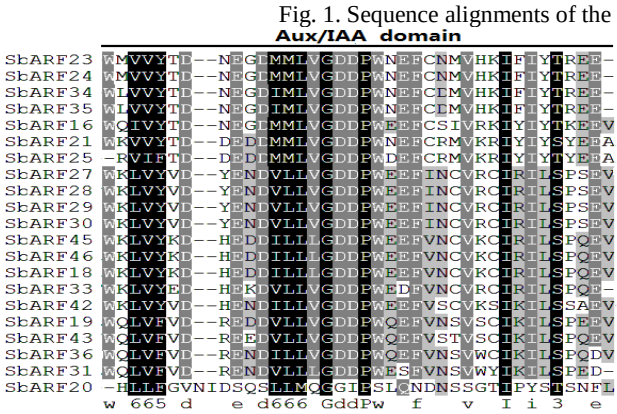

Fig. 2. Sequence alignments of the Aux/IAA domain of ARF proteins in sorghum.

Expression patterns of *SbARF*s in different plant tissues and their responses to abiotic stresses: Transcriptional profiles showed that several *SbARF* genes (including *SbARF16/17/21/23 /24/25/33*) were widely expressed in 10 different tissues of sorghum, namely the roots, shoots, leaves, emerging inflorescences, seeds, early inflorescences, pistils, embryos, endosperms and anthers (Fig. 7). In total, 35 genes were expressed in early inflorescences, followed by 32 in the emerging inflorescences; the number of genes expressed in endosperms was smallest at 19. Generally, class IV genes were not expressed in leaves or endosperms, whereas *SbARF18/33/31/36/43/19/20* were highly expressed in emerging inflorescences, seeds, early inflorescences, pistils and embryos. Most class III genes exhibited high expression levels in all 10 tissues, excluding *SbARF3*, which was highly expressed in roots, inflorescences and pistils only. By contrast, *SbARF34/35* were hardly expressed in roots. Most class I genes were expressed at high levels in emerging inflorescences, early inflorescences, seeds, and embryos, with low expression in leaves, pistils, endosperms and anthers. In class II, only *SbARF1/22/44* displayed high expression levels in seeds, early inflorescences, and pistils, whereas the remaining subfamily members were not generally expressed in any tissues.

In plants, interactions exist between relevant genes in the hormone pathways. Thus, exogenous ABA stress and PEG-induced osmotic stress affect the biosynthesis pathways of various hormones. The resulting differential expression within amino acid metabolic pathways in turn regulates the signaling pathways to different degrees and many pathways involve both upregulation and downregulation of genes (Okushima *et al.*, 2007). The transcriptional profiles of *SbARF* genes under ABA and PEG treatments are shown in Fig. 7. Comparing the ABA treatment group with the NaOH control group, it was found that the majority of class I genes (*SbARF6/7/8/9/10/11/12/13/14/15/40/41*) and the *SbARF36* gene were significantly upregulated in the roots, whereas *SbARF3/45/46* were significantly downregulated in this tissue. In addition, the expressions of *SbARF33/34/35/44* were significantly increased in the shoots, while no genes presented expression decreases in this tissue. When comparing the PEG treatment group with the H_2O control group, *SbARF45* and *SbARF46* were upregulated in the roots, but *SbARF3* and *SbARF36* were downregulated in this tissue. Meanwhile, the expressions of *SbARF21/34/35* were markedly upregulated in the shoots and *SbARF*33 was the only gene that was considerably downregulated in this tissue.

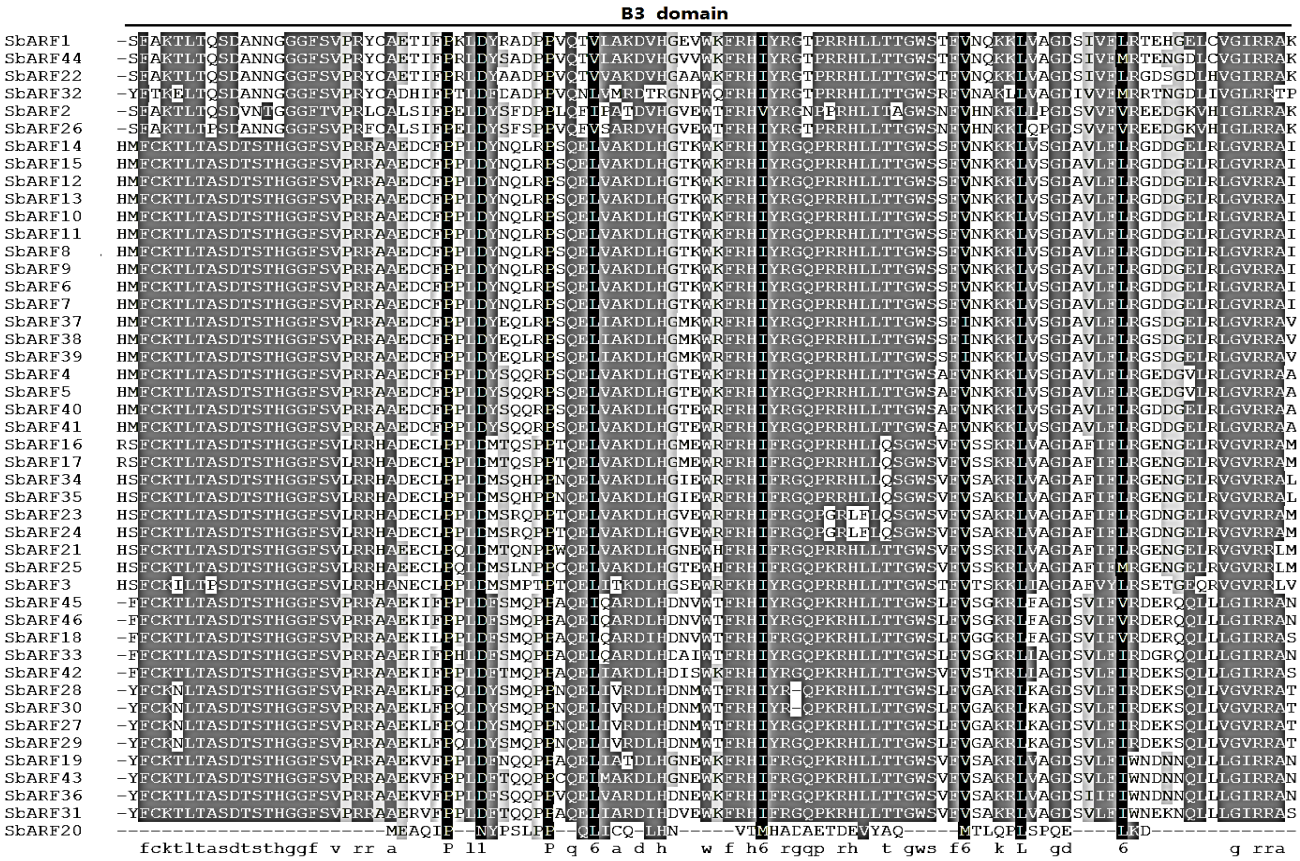

Fig. 3. Sequence alignments of the B3 domain of ARF proteins in sorghum.

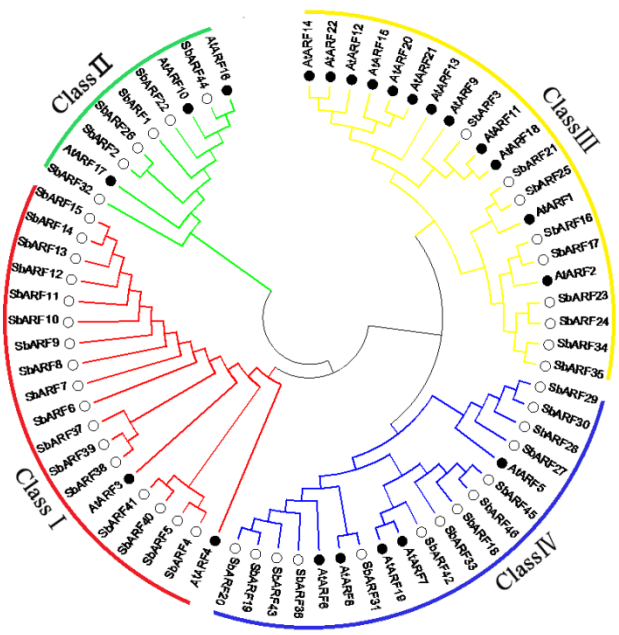

Fig. 4. Phylogenetic relationship of ARF proteins in sorghum (SbARFs) and *Arabidopsis* (AtARFs).

Discussion

Most ARF proteins are known to contain three conserved regions, a unique B3 domain at the N-terminus, a repression/activation ARF domain in the middle region and a dimerization domain Aux/IAA at the C-terminus (Guilfoyle *et al.*, 2007). Here we identified 46 *SbARF* genes in sorghum, all of which encoded proteins containing the B3 and ARF domains; however, only 21 SbARF proteins contained the Aux/IAA domain. In *Arabidopsis*, only one of the 23 AtARFs lacks the B3 domain, and with the exception of AtARF3/13/17, all AtARF family members contain the C-terminal domain (Schruff *et al.*, 2006). In rice, OsARF20 contains two B3 domains, whereas most of the other OsARFs contain one B3 domain and an ARF domain (Jain *et al.*, 2009). These findings indicate that like other plant ARFs, SbARF proteins contain the Aux/IAA, B3, and ARF domains, but the number of domains is higher than those in other well-studied plants.

In the present study, we selected and analyzed 46 *SbARF* genes at the genome level, assigning them into four subfamilies. Since the localization pattern of introns can provide important evidence for the phylogenetic relationship of the genome (Wang *et al.*, 2012), we combined the structural arrangements of introns with the phylogenetic tree of *SbARF* genes. We found that each *SbARF* gene contained one to 13 introns. There were no major differences in the intron number per gene across classes I, III, and IV; however, genes of class II had a much smaller intron number compared with the other three subfamilies. In higher plants, such differences in the number of introns per gene may be associated with diverse gene functions between different genetic lineages (Attia *et al.*, 2009). Similar results have been observed in tomato (De *et al.*, 2009), Arabidopsis (Okushima *et al.*, 2005), and rice (Jain *et al.*, 2009).

Table 2. Introns in *ARF* genes in sorghum.

Gene name	Intron number	Gene name	Intron number	Gene name	Intron number	Gene name	Intron number
SbARF1	2	SbARF13	11	SbARF25	13	SbARF37	9
SbARF2	2	SbARF14	11	SbARF26	2	SbARF38	9
SbARF3	12	SbARF15	11	SbARF27	12	SbARF39	9
SbARF4	9	SbARF16	13	SbARF28	12	SbARF40	10
SbARF5	9	SbARF17	11	SbARF29	12	SbARF41	10
SbARF6	11	SbARF18	11	SbARF30	12	SbARF42	12
SbARF7	9	SbARF19	13	SbARF31	13	SbARF43	13
SbARF8	9	SbARF20	13	SbARF32	1	SbARF44	2
SbARF9	10	SbARF21	13	SbARF33	12	SbARF45	13
SbARF10	10	SbARF22	2	SbARF34	13	SbARF46	11
SbARF11	10	SbARF23	12	SbARF35	13		
SbARF12	10	SbARF24	12	SbARF36	13		

Table 3. Motifs in ARF proteins in sorghum.

Motif	Isoelectric point	Chromosome location	Width	Motif sequence
1	6.2e-1898	46	50	DTSTHGGFSVPRRAAEDCFPPLDYSQ QPPSQELVAKDLHGTEWKFRHIYR
2	2.9e-1596	46	50	QPRRHLLTTGWSVFVNKKKLVAGDA VLFLRGEBGZLRLGVRRAIRLKNEA
3	1.0e-921	45	33	WPGSKWRSLKVRWDEGAEGERPDR VSPWEIEIEPA
4	1.7e-1218	40	50	SSVLSSDSMHLGVLAAAAHAAKTRS VFTIYYNPRASPSEFIIPYAKYLKS
5	3.8e-979	40	41	VNRELWHACAGPLVALPRRGSLVVY FPQGHLEQVGASTVAA
6	3.9e-781	43	29	LPPKVLCRVABVELHADAETDEVYA QLTL
7	3.3e-642	22	50	EDPGRSGWKLVYVDNENDVLLVGD DPWEEFVNCVRCIRILSPZEVQQMSL
8	8.2e-409	10	50	GQEISRAVPMFQGMMSEACSLKGGY GLHSYMHTPVAANGLSAPAQECCLT
9	2.8e-403	10	50	DNIFNRTVVPQLGLASKFGGGGTNGQ QSGPFDRRREIWTKPQHETPDQMN
10	1.8e-390	45	16	PVSVGMRFKMRFETED

Based on the phylogenetic analysis of sorghum and *Arabidopsis ARF* genes, we found that *SbARF16/17/21/25* were most closely related to *AtARF1* and *AtARF2* (Ellis *et al.*, 2005). We suspect that these four *SbARF* genes are involved in sorghum leaf, flower and seed growth processes. In addition, *SbARF27/28/29/30* displayed a close phylogenetic relationship with *AtARF5* (Nagpal *et al.*, 2005); these four *SbARF* genes may play important roles in sorghum hypocotyl formation and microtubule growth. Moreover, *SbARF42* was located on the same branch as *AtARF7* and *AtARF19 (*Okushima *et al.*, 2005), and hence it may mediate lateral root formation. Furthermore, *SbARF19/20/31/36/43* were most closely related to *AtARF6* and *AtARF8* (Schruff *et al.*, 2006; Wu *et al.*, 2006), suggesting that these five *SbARF* genes may participate in the biosynthesis of hormones associated with stress resistance. Our observation of gene clusters on multiple chromosomes indicates that those clustered genes are likely structural genes that encode enzymes to catalyze different steps of the same metabolic pathway (Feng *et al.*, 2018).

At the transcriptional level, some *SbARF* genes were widely expressed in different tissues of sorghum. Similar expression patterns have also been found in tea [*Camellia sinensis* (L.) O. Ktze.] (Xu *et al.*, 2016), apple (Luo *et al.*, 2014), and tomato (De *et al.*, 2009). For example, 13 of 15 *CsARF* genes are expressed in roots, stems, leaves, flowers, and fruits of tea; eight of 31 *MdARF* genes are expressed in stems, leaves, flowers, and fruits of apple; and 17 *SlARF* genes are expressed in roots, stems, leaves, flower buds, and ovaries of tomato. Expression levels of the *SbARF* genes were considerably different across various tissues. Most class III genes were highly expressed in leaves, seeds and flowers, which is in accordance with the reported functions of two *Arabidopsis* homologs, *AtARF1* and *AtARF2* (Ellis *et al.*, 2005). In addition, *AtARF1* and *AtARF2* were assigned to class III based on the phylogenetic analysis. This corroborates from two perspectives that genes in class III play critical roles in sorghum leaf, flower and seed growth processes (Ellis *et al.*, 2005). Most genes in class I regulated the development of emerging inflorescences, early inflorescences, seeds, and embryos. However, expression of *SbARF2/26/32* rarely occurred in any tissues, suggesting that these genes may not function during the development of these tissues, or that their functions have been lost during evolution.

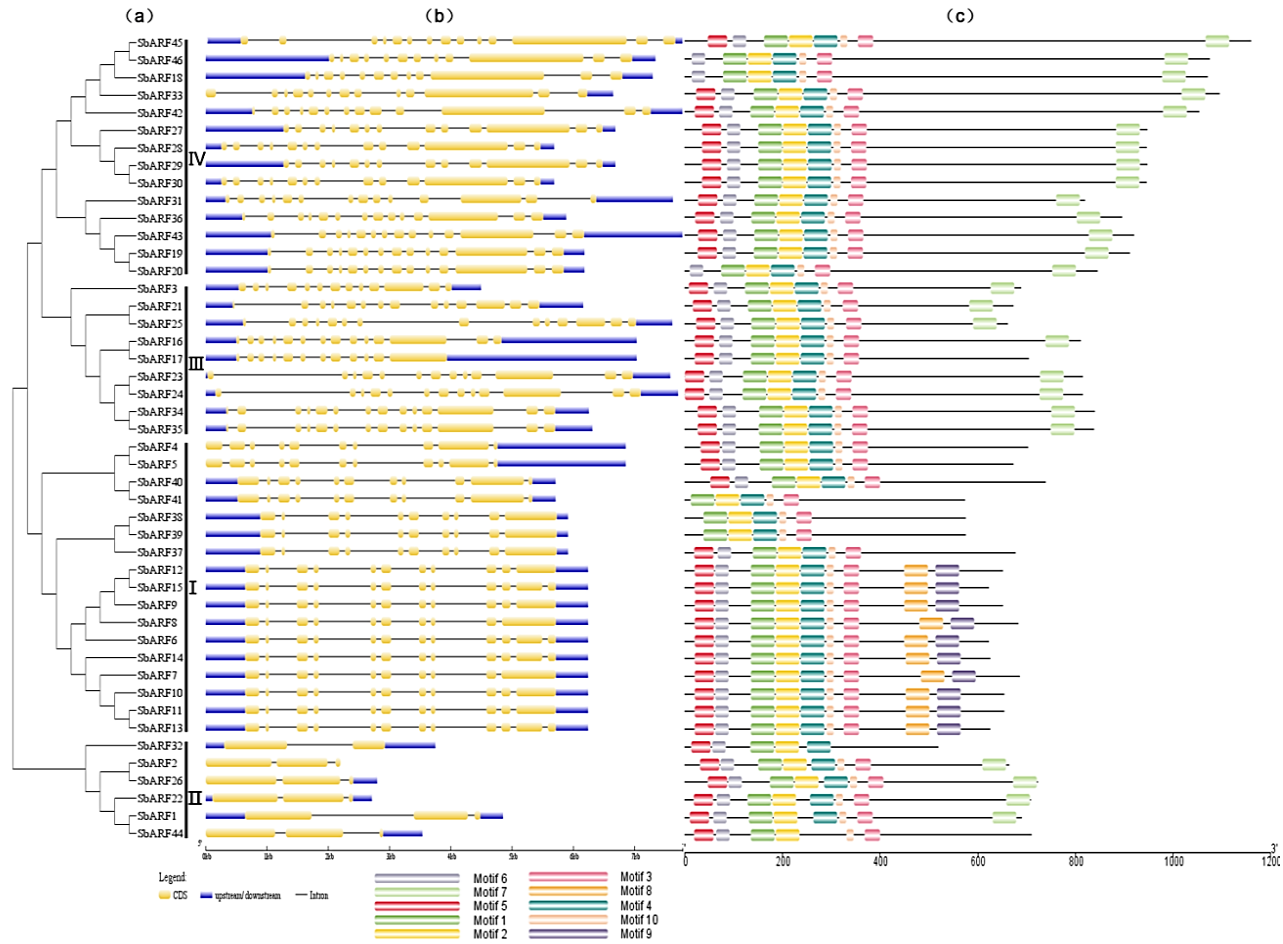

Fig. 5. Phylogeny and gene structure of *ARF* genes in sorghum. (a) Phylogeny of *SbARF* genes. (b) Intron arrangements of *SbARF* genes. (c) Schematic representation of conserved motifs in the SbARF proteins.

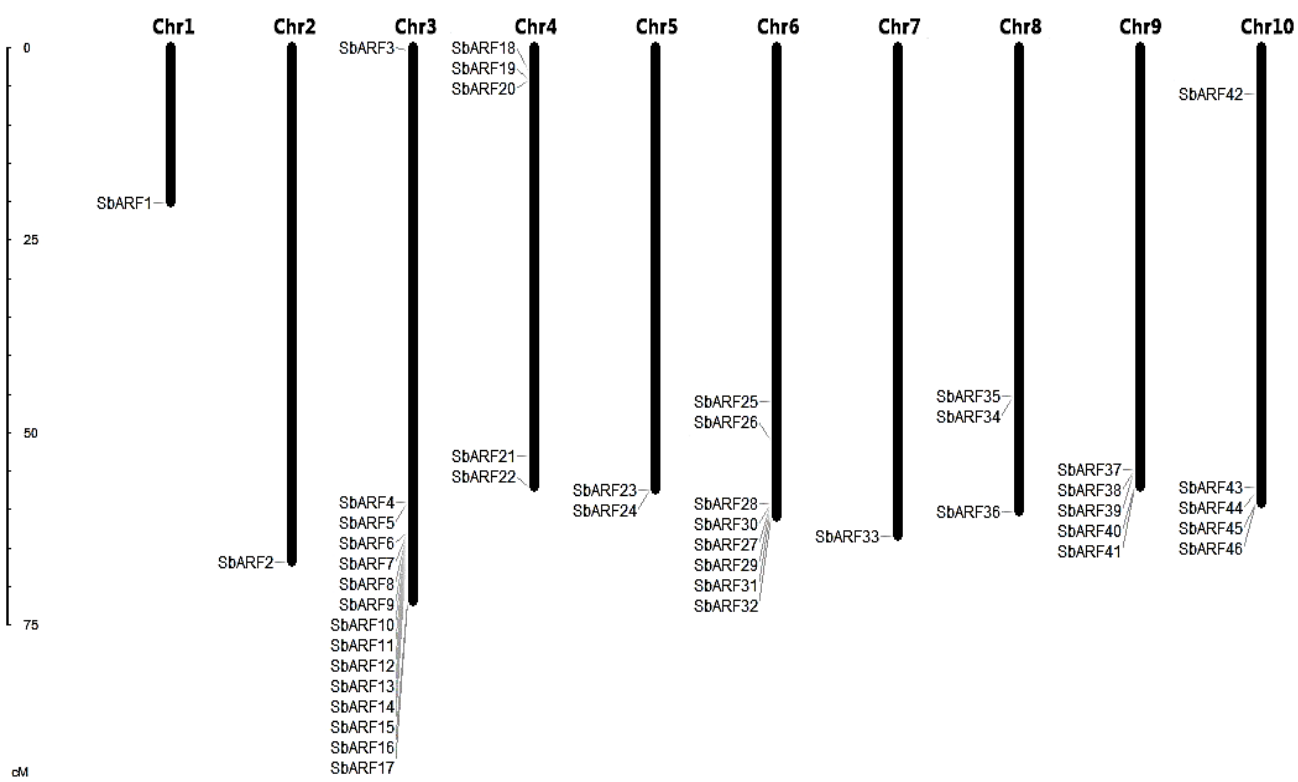

Fig. 6. Chromosome locations of *ARF* genes in sorghum.

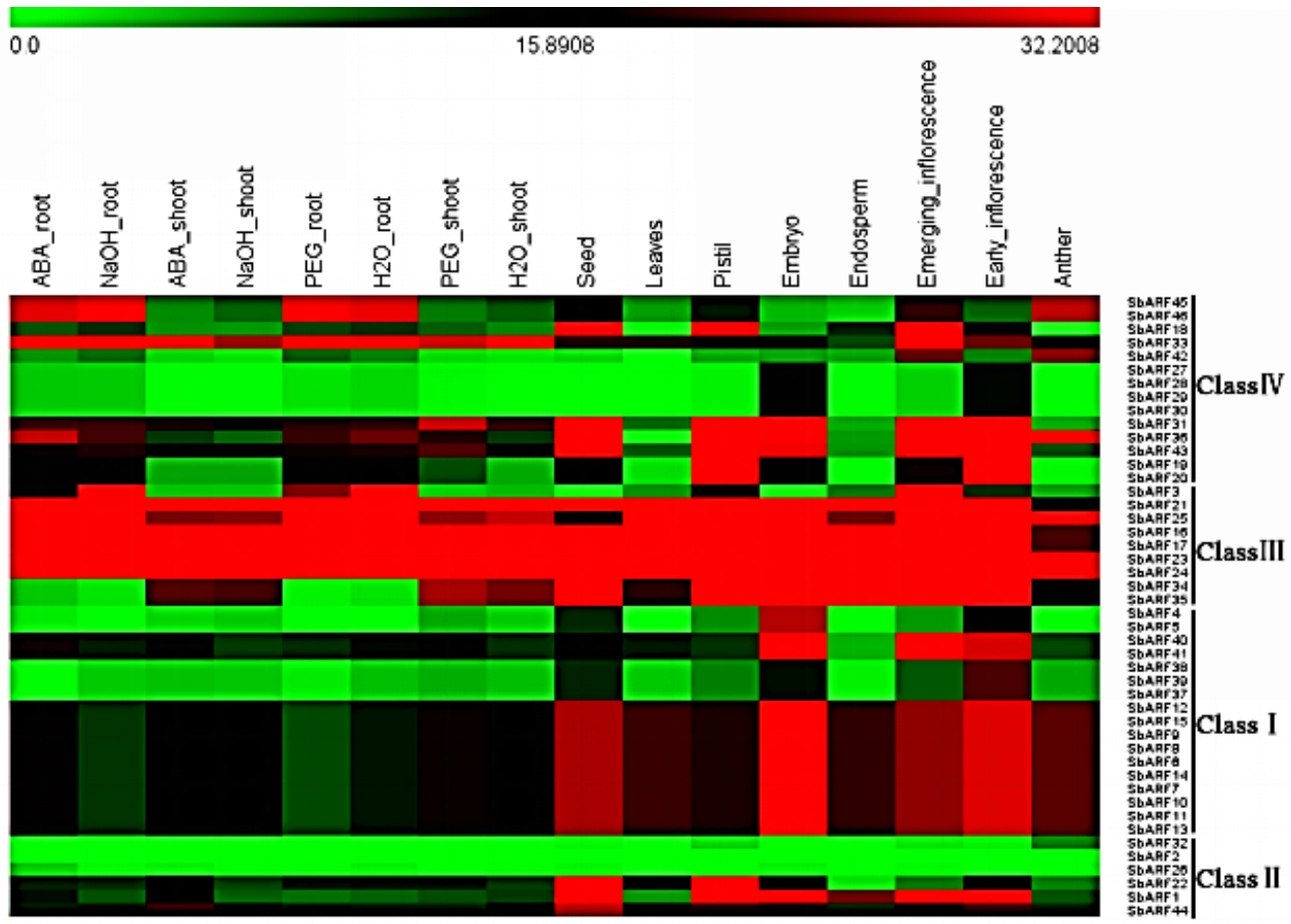

Fig. 7. Transcriptional profiling of *ARF* genes in sorghum in relation to different tissues and stresses induced by abscisic acid (ABA, with NaOH as a control) or polyethylene glycol (PEG, with H_2O as a control).

ARFs are transcription factors that mediate the expression of auxin-responsive genes, while exogenous hormones play a main role in the regulatory of *ARF* gene expression. For example, in *Arabidopsis*, expression levels of *AtARF5/10/16/19* are markedly changed under exogenous auxin treatment; *AtARF7* and *AtARF19* also perform important functions in ethylene signal transduction (Li *et al.*, 2006; Okushima *et al.*, 2007). In switchgrass, expression levels of *PvARF5/6/7/8/11/12/15/16/29/30/35/36/37/38/40/41/42/43/44/45* are downregulated by exogenous auxin (1-naphthalene acetic acid, NAA) (Wang *et al.*, 2018); in contrast, expression levels of *PvARF3/4/23/24/25/26* are upregulated by NAA, suggesting that these genes are potential major auxin-responsive genes. In rice (Attia *et al.*, 2009) and maize (Wang *et al.*, 2012), *OsARF1/23* and *ZmARF3/8/13/15/21/27/30* appear to be upregulated under exogenous auxin treatment, whereas *OsARF5/14* and *ZmARF5/18* are downregulated; these genes are therefore considered to play crucial roles in regulating plant growth and development under exogenous auxin stress conditions.

ABA, a hormone essential for plant growth and development, is extensively involved in response to abiotic stresses such as drought, low temperature and osmotic stress (Dalal *et al.*, 2009). Here we investigated the responses of *SbARF* genes after treating sorghum plants with exogenous ABA or PEG. At least 15 *SbARF* genes in the roots and four genes in the shoots were responsive to exogenous ABA treatment. Following PEG treatment, four genes in the roots and four genes in the shoots were responsive. Interestingly, these responsive genes exhibited distinct expression patterns. A total of five genes were downregulated in roots and shoots of sorghum under ABA or PEG stress; meanwhile, 19 genes were upregulated in the roots and shoots, suggesting that they may be potentially important ABA-responsive genes. Previous study has shown that ABA can induce massive accumulation of proline as an osmoticum in plants, while increasing the activity of related protective enzymes (Scandaliol, 1993). Similarly, corresponding changes may also occur in the content of proline in plants under PEG-induced osmotic stress (Zhou *et al.*, 2014). Therefore, we suspect that the above-mentioned *SbARF* genes play a major role in ABA-dependent stress response and PEG-induced osmotic stress through increasing the proline content in sorghum plants. In summary, these *SbARF* genes are likely to play a critical role in sorghum's drought tolerance and thus may be considerate as candidate genes for the research about plant tolerance to salt and drought. However, the specific functions and mechanisms of these genes need to be further elucidated via experimentation.

The *ARF* gene family can mediate plant hormone metabolism, thereby affecting plant growth and

development (Dalal *et al.*, 2009). By using the sorghum genome sequence as the genetic background, this study explored the conserved domain, phylogenetic relationship, gene structure, chromosome location, and expression pattern of the *ARF* gene family in sorghum. Our work provides evidence for the regulation mechanisms of the *ARF* gene family involved in plant growth and development of sorghum. However, further investigation is required to determine how the sorghum *ARF* genes respond to hormonal signals and thereby mediate the role of auxin.

Conclusion

In conclusion, our study provided comprehensive information on the ARF family in sorghum, including gene structures, chromosome locations, phylogenetic relationships and expression patterns. SbARF genes were highly expressed in leaves, flowers and seeds, which were significantly upregulated or downregulated in roots or shoots of sorghum under stress induced by exogenous abscisic acid or polyethylene glycol. These results suggested that *SbARF* genes played a key role in growth, development and stress responses in sorghum.

Acknowledgement

This study was supported from Luoyang Normal University horizontal research project (No. 190307) and National Natural Science Foundation of China under Grant (No.31870697), 2020 training plan for key teachers of Luoyang Normal University.

References

Attia, K.A., A.F. Abdelkhalik, M.H. Ammar, C. Wei, J. Yang, D.A. Lightfoot, W.M. E1-Sayed and H.A. E1-Shemy. 2009. Antisense phenotypes reveal a functional expression of OsARF 1, an anxin response factor, in transgenic rice. *Curr. Issu. Mol. Biol*., 11: 29-34.

Dalal, M., D. Tayal, V. Chinnusamy and K.C. Bansal. 2009. Abiotic stress and ABA-inducible Group 4 LEA from *Brassica napus* plays a key role in salt and drought tolerance. *Biotechnology*, 139: 137-145.

De, M., M. Wolters-Art, R. Feron, C. Mariani and W.H. Vriezen. 2009. The *Solanum lycopersicum* auxin response Factor 7 (SlARF7) regulates auxin singnaling during tomato fruit set and development. *Plant J.*, 57: 160-170.

Ellis, C.M., P. Nagpal, J.C. Young, T.J. Guilfoy and J.W. Reed. 2005. Auxin response factor 1 and factor 2 regulate senescence and foral organ abscission in *Arabidopsis thaliana*. *Development (Cambridge, England)*, 132: 4563-4574.

Feng, T., G.Y. Si., G.H. Li, N. Guo and Y. Cai. 2018. Gome-wide identification and analysis of ARF gene family in *Gossypium arboretum* L. *China J. Bioinform.*, 16: 22-28.

Guan, X.C., H.J. Hu, Y.N. Wu and X.Y. Lu. 2016. Research progress on auxin response factor. *Mol. Plant Bre. (in Chinese)*, 14: 1892-1897.

Guilfoyle, T.J. and G. Hagen. 2007. Auxin response factor. *Curr. Opin. Plant Biol.*, 10: 453-460.

Jain, M. and J.P. Khurana. 2009. Transcript profiling reveals diverse roles of auxin-responsive genes during reproductive development and abiotic stress in rice. *FEBS J.*, 276(11): 3148-3162.

Li, J., X. Dai and Y. Zhao. 2006. A Role for Auxin Response Factor 19 in Auxin and Ethylene Signaling in Arabidopsis. *Plant Physiol.*, 140: 899-908.

Liu, K.D., C.C. Yuan, H.L. Li, W.H. Lin, Y.J. Yang, C.J. Shen and X.L. Zheng. 2015. Genome-wide identification and characterization of auxin response factor (ARF) family genes related to flower and fruit development in papaya (*Carica papaya* L.). *BMC Genom.*, 16: 901-912.

Luo, X.C., M.H. Sun, R.R. Xu, H.R. Shu, J.W. Wang and S.Z. Zhang. 2014. Genomewide identification and expression analysis of the ARF gene family in apple. *J. Genet*, 93: 785-797.

Nagpal, P., C.M. Ellis, H. Weber, S.E. Ploense, L.S. Barkawi, T.J. Guilfoyle, G. Hagen, J.M. Alonso, J.D. Cohen, E.E. Farmer, J.R. Ecker and J.W. Reed. 2005. Auxin response factors ARF 6 and ARF 8 promote jasmonic acidproduction and flower maturation. *Development (Cambridge, England)*, 132: 4107-4118.

Nicholas, K.B., H.B. Nicholas and D.W. Deerfield. 1997. GeneDoc: Analysis and visualization of genetic variation. *Plos One*, 4: 1-14.

Okushima, Y., H. Fukaki, M. Onoda, A. Theologis and M. Tasaka. 2007. ARF7 and ARF19 regulate lateral root formation via direct activation of LBD/ASL genes in *Arabidopsis*. *Plant Cell*, 19: 118-130.

Okushima, Y., P.J. Overvoorde, K. Arima , J.M. Alonso, A. Chan, C. Chang, J.R. Ecker, B. Hughes, A. Lui, D. Nguyen, C. Onodera, H. Quach, A. Smith, G. Yu and A. Theologis. 2005. Functional genomic analysis of the auxin response factor gene family members in *Arabidopsis thaliana*: Unique and overlapping functions of ARF7 and ARF19. *Plant Cell*, 17: 444-463.

Saeed, A.I., N.K. Hagabati, J.C. Braisted, W. Liang, V. Sharov, E.A. Howe, J. Li, M. Thiagarajan, J.A. White and J. Quackenbush. 2006. TM4 microarray software suite. *Methods Enzymol*., 411: 134-193.

Scandaliol, J.G. 1993. Oxygen stress and super oxide dimutases. *Plant Physiol*, 101: 7-12.

Schruff, M.C., M. Spielmanlman, S. Tiwari, S. Adams, N. Fenby and R.J. Scott. 2006. The auxin response factor 2 gene of Arabidopsis links auxin signaling, cell division, and the size of seeds and other organs. *Development (Cambridge, England)*, 133: 251-261.

Sheng, H., Z.W. Qin and W.B. Li. 2014. Cucumber Auxin Response Factor (ARF) family identification and expression specificity analysis. *Chinese Agricultural Science (in Chinese)*, 47: 1985-1994.

Tamura, K., G. Stecher, D. Peterson, A. Filipski and S. Kumar. 2013. MEGA6: Molecular evolutionary genetics analysis version 6.0. *Mol. Biol. Evol.*, 30: 2725-2729.

Tiwari, S.B., G. Hagen and T. Guilfoyle. 2003. The roles of auxin response factor domains in auxin responsive transcription. *Plant cell*, 15: 533-543.

Wang Y.J., D.X. Deng, Y.T. Shi, N. Miao, Y.L. Bian and Z.T. Yin . 2012. Diversification, phylogeny and evolution of auxin response factor (ARF) family: insights gained from analyzing maize ARF genes. *Mol. Biol. Rep.*, 39: 2401-2415.

Wang, D., K. Pei, Y. Fu, Z .Sun, H. Liu, K. Tang, B. Han and Y. Tao. 2007. Genome-wide analysis of the anxin response factors (ARF) gene family in rice (*Oryza sative*). *Gene*, 394: 13-24.

Wang, J.L., Z.Y. Wu, Z.B. Shen, Z.T. Bai, P. Zhong, L.C. Ma, D.F. Pan, R.B. Zhang, D.M. Li, H.L. Zhang, C.X. Fu, G.Q. Han and C.H. Guo. 2018. Genome-wide identification, phylogeny, and expression analysis of ARF genes involved in vegetative organs development in switch grass. *Int. J. Genom.*, 29: 1-13.

Wei, R.M., L.L. Xie, X. Ouyang, Y.L. Zhang, X.Z. Dai and F. Liu. 2017. Identification and expression analysis of ARF gene family in pepper. *J. Agr. Sci. Tech.,* 37: 1047-1058.

Wu, J., F.Y. Wang, L. Cheng, F.L. Kong, Z. Peng, S.Y. Liu, X.L. Yu and G. Lu. 2011. Identification, isolation and expression analysis of auxin response factor (ARF) genes in *Solanum lycopersi*cum. *Plant Cell Rep*., 30: 2059-2073.

Wu, M.F., Q. Tian and J.W. Reed. 2006. Arabidopsis microRNA167 controls patterns of ARF6 and ARF8 expression, and regulates both female and male reproduction. *Development (Cambridge, England)*, 1332: 4211-4218.

Xu, Y., J. Mao, W. Chen, T.T. Qian, S.C. Liu, W.J. Hao, C.F. Li and L. Chen. 2016. Identification and expression profiling of the auxin response factors (ARFs) in the tea plant (*Camellia sinensis* (L.) O. Kuntze) under various abiotic stresses. *Plant Physiol. Bioch.*, 98: 46-56.

Zhang, S.J., O. Amerjan, X.Z. Xue, X. Zhang, X.F. Guo and L.Q. Chen. 2014. Quality analysis on different sweet sorghum silages in Southern Xinjiang compared with a corn silage. *Acta Prataculturae Sinica (in Chinese)*. 23: 232-240.

Zhang, S.P., M.M. Liu, H. Miao, S.Q. Zhang, Y.H. Yang, B.Y. Xie, T.C. Wehner and X.F. Gu. 2013. Chromosomal mapping and QTL analysis of resistance to downy mildew in *Cucumis sativus*. *Plant Disease (in Chinese)*, 97: 245-251.

Zhou, J.J., P. Liu, J.J. Yin and X.Z. Xie. 2014. Involvement of proline metabolic pathway in regulating drought tolerance in rice *phyB* mutant. *Shandong Agri. Sci., (in Chinese)*, 46: 1-4.

(Received for publication 25 June 2019)