Phyton-
International Journal of
Experimental Botany

Tech Science Press

http://doi.org/10.32604/phyton.2026.087409

ARTICLE

Genome-Wide Identification and Expression Analysis of the *DUF506* Gene Family in Sorghum (*Sorghum bicolor* L.) under Abiotic Stresses

Esra Alther[1,2], Hind Abdelmonim Elsanosi[1,2], Ghazi Badawi[2] and Guisheng Zhou[1,*]

[1]The Joint International Research Laboratory of Agriculture and Agri-Product Safety, Ministry of Education, Yangzhou University, Yangzhou, China
[2]Faculty of Agriculture, University of Khartoum, Shambat, Khartoum North, Sudan
*Corresponding Author: Guisheng Zhou. Email: gszhou@yzu.edu.cn

Received: 16 June 2026; Accepted: 25 August 2026; Published: 24 September 2026

ABSTRACT: The domain of unknown function 506 (DUF506) family belongs to the PD-(D/E) XK nuclease superfamily and is related to the regulation of plant growth and development and responses to abiotic stress. To date, no comprehensive analysis of the *DUF506* gene family has been conducted in *Sorghum bicolor*. This study revealed nine *SbDUF506* genes, which have been distributed unevenly across five chromosomes. Gene structure analysis indicated that *SbDUF506* genes composed of one to three exons and introns. Cis-regulatory element analysis indicated that most *SbDUF506* genes are associated with responses to abiotic stress. Collectively, the results suggest that gene duplication events markedly enhanced the evolutionary proliferation of the *SbDUF506* gene family. Ka/Ks analysis indicated that these genes have predominantly evolved under purifying selection. Furthermore, the expression analysis profiles of the *SbDUF506* genes across various tissues and under stress conditions were evaluated via RT-qPCR following heat, salinity, and cold treatments. The results indicated that most *SbDUF506* genes are responsive to cold stress, as well as to heat and salinity. These findings advance the theoretical understanding of the evolutionary organization and stress-responsive regulation of the *DUF506* gene family in sorghum and provide candidate genes for further functional validation and the molecular breeding of Sorghum cultivars with improved tolerance to abiotic stresses.

KEYWORDS: *Sorghum bicolor*; DUF506; gene expression; bioinformatics analysis; abiotic stress

1 Introduction

The Domain of Unknown Function (DUF) encompasses a large group of protein families that are defined by conserved structural domains, despite their biological functions remaining largely uncharacterized [1]. Although these proteins possess conserved motifs, the precise biological functions of many DUF-containing proteins remain largely unresolved [2]. According to the Pfam database (version 35.0), approximately 19,000 protein families have been identified, of which nearly 24% are classified as DUF families. These proteins are broadly distributed across diverse organisms such as plants, animals, fungi, and bacteria [1]. Accumulating studies have reported that DUF domain-containing proteins play critical roles in plant abiotic stress responses. For instance, the DUF569 protein in *Arabidopsis* acts as a positive regulator of drought tolerance, whereas several DUF668 proteins in rice respond to salt and wounding stress [3]. Additionally, DUF proteins influence plant reproductive development. The *Arabidopsis DUF761-1* gene regulates plant morphology and reproductive processes, while mutations in the rice *DUF1668* gene lead to pollen sterility and hybrid incompatibility [4,5]. The *DUF246* family, which participates in pectin biosynthesis, also affects

male fertility in *Arabidopsis* [6]. Furthermore, the *SbDUF4228* family has been associated with the response to salinity in Sorghum [7]. In tobacco, overexpression of *NtDUF868-E5* enhances growth, chlorophyll accumulation, and photosynthetic efficiency [8]. DUF proteins have also been implicated in regulating responses to abiotic stresses. In *Arabidopsis*, knockout of *AtDUF569* confers enhanced disease resistance, likely through the upregulation of salicylic acid–dependent pathogenesis-related (PR) genes [9]. In rice, the stress-inducible DUF1644 protein (*Os-SIDP366*) enhances tolerance to drought and salinity [10], and the *DUF1645*-containing gene *OsSGL* significantly improves drought tolerance in both rice and Arabidopsis by modulating stress-related gene expression [11]. Furthermore, *OsDUF872.2* improves heat tolerance in *Escherichia coli*, and overexpression of *OsDUF6* in rice alters the expression of auxin-related and receptor-like kinase genes under salt stress, functioning as a positive regulator of salinity tolerance [12,13].

The PD-(D/E) XK nuclease superfamily encompasses several DUF domain–containing protein families, including *DUF524*, *DUF506*, and *DUF1626* [14]. These nucleases exhibit considerable structural and functional diversity and are primarily involved in DNA degradation, repair, and recombination processes [15]. As a subfamily within this superfamily, DUF506 proteins display low sequence similarity and distant evolutionary relationships to other PD-(D/E) XK members, making it difficult to predict the specific functions of individual proteins [16]. Recent studies, however, have identified several DUF506 proteins as potential regulators of abiotic stress responses in *Arabidopsis* and rice. For instance, AtRXR3, an *Arabidopsis* DUF506 protein, suppresses root hair elongation under phosphate-deficient conditions [17], while many *DUF506* members in *Oryza sativa* respond to ABA, JA, and low-temperature stimuli [18], highlighting the functional versatility of the *DUF506* family in higher plants.

Sorghum bicolor is a widely cultivated cereal crop known for its exceptional tolerance to abiotic stresses, comprising salinity, temperature, and aridity [19,20]. Sorghum research and breeding initiatives have concentrated on augmenting yield, stress tolerance, and disease resistance to improve productivity and adaptability in various environments. Progress in genomic research and genetic engineering methodologies has enabled the creation of genetically altered Sorghum varieties exhibiting advantageous characteristics, such as greater stress resistance. Despite these advancements, the *DUF506* gene family remains uncharacterized in Sorghum. This study presents a comprehensive genome-wide analysis of the evolutionary features and potential biological roles of *DUF506* genes in the *S. bicolor* genome. We systematically examined their phylogenetic relationships, conserved motifs, gene structures, and expression patterns, as well as cis-acting regulatory elements, nonsynonymous SNP distributions, syntenic relationships, and tissue-specific expression profiles. Additionally, we assessed the expression levels of *SbDUF506* genes under cold, heat, and salinity stresses using qRT-PCR assays. Collectively, our results provide a systematic identification and classification of the *SbDUF506* gene family, deepen insights into their possible functions in abiotic stress responses, and a valuable foundation for future molecular breeding initiatives in Sorghum.

2 Materials and Methods

2.1 Genome-Wide Characterization of the DUF506 Gene Family in Sorghum bicolor

To identify members of the *SbDUF506* gene family in *Sorghum bicolor*, the genome sequence and the corresponding annotation files (version 3.1.1) were downloaded from the Phytozome database (https://phytozome.jgi.doe.gov/pz/portal.html, v13.1, accessed on 23 October 2025), Protein sequences of *Arabidopsis DUF506* genes were also retrieved from Phytozome to serve as reference queries. To comprehensively detect potential DUF506 homologs in *S. bicolor*, the *Arabidopsis* DUF506 protein sequences were used as queries in a BLASTP search against the *Sorghum bicolor* protein database. Candidate sequences were retained using an E-value threshold of $\leq 1 \times 10^{-5}$, query coverage of $\geq 50\%$, and amino acid sequence

identity of ≥30% (http://blast.ncbi.nlm.nih.gov, v13.1, accessed on 24 October 2025). Meanwhile, the conserved DUF506 domain (PF04720; PDDEXK_6) was obtained from the Pfam database (http://pfam.xfam.org) and employed as a query profile to identify DUF506 family members using the HMMER search tool (https://www.ebi.ac.uk/Tools/hmmer/search/hmmscan) using an E-value threshold of $\leq 1 \times 10^{-5}$. All retrieved candidate sequences were subjected to conserved domain validation using the Pfam database (http://pfam.xfam.org/, accessed on 24 October 2025) and SMART (http://smart.embl-heidelberg.de, accessed on 24 October 2025). Proteins lacking the characteristic DUF506 domain were excluded from subsequent analyses. The physicochemical properties, including the amino acid content, theoretical isoelectric point (pI), and molecular weight (MW) of the discovered SbDUF506 proteins were evaluated using the ExPASy ProtParam program (https://web.expasy.org/protparam, accessed on 25 October 2025). Subcellular localization was predicted using the Plant-mPLoc service (http://www.csbio.sjtu.edu.cn/bioinf/plant-multi/, accessed on 27 October 2025).

2.2 Phylogenetic, Gene Structural, and Motif Analysis of SbDUF506 Genes

To analysis of the evolutionary relationships and structural features of the *SbDUF506* gene family, the protein sequences of nine putative *SbDUF506* members were retrieved from the Phytozome v13.1 genome database (accessed on 28 October 2025). Multiple sequence alignment of the identified proteins was performed using the MUSCLE algorithm implemented in MEGA version 11 [21]. An unrooted phylogenetic tree was subsequently constructed in MEGA 11 using the neighbor-joining (NJ) method with the p-distance model, and branch support was evaluated through 1000 bootstrap replicates. The resulting tree was visualized and annotated using the Interactive Tree of Life (iTOL) platform (https://itol.embl.de, accessed on 29 October 2025) [22]. To investigate structural diversity and conserved motif architecture within the *SbDUF506* gene family, exon–intron organization and untranslated regions (UTRs) were analyzed using TBtools. Conserved protein motifs were identified using the MEME Suite (accessed on 14 November 2025) [23]. The XML output generated by MEME was subsequently imported into TBtools to visualize the distribution and organization of conserved motifs across the *SbDUF506* proteins [24].

2.3 Selective Pressure (Ka/Ks) and Conserved Domains Analysis of DUF506 Genes

Non-synonymous (Ka) and synonymous (Ks) substitution rates were calculated to determine the evolutionary divergence of paralogous *DUF506* gene pairs by utilizing the KaKs Calculator 2.0 module embedded in the TBtools bioinformatics platform (accessed on 10 November 2025). Protein sequences harboring the DUF506 domain were obtained from the SMART database (accessed on 12 November 2025) for further analysis. Pairwise Ka and Ks values for the corresponding coding sequences were then calculated using the Nei–Gojobori (NG) method using TBtools to evaluate selective pressures acting on the *DUF506* gene family.

2.4 Chromosomal Localization, Duplication and Synteny Evaluation, and Promoter Cis-Acting Elements

The genomic coordinates and chromosomal positions of the *SbDUF506* genes were retrieved from the *Sorghum bicolor* v3.1.1 (GFF3) annotation file available in the Phytozome database [25]. Gene duplication events were analyzed using TBtools, and Circos plots were produced to illustrate the genomic distribution and structural organization of the *SbDUF506* gene member. Chromosomal distribution was further visualized using TBtools. Comparative synteny analysis was performed among *S. bicolor*, *A. thaliana*, and *O. sativa* using the Multi-Synteny Plot function implemented in TBtools. In addition, 2000-bp upstream promoter

sequences of the *SbDUF506* genes were retrieved from Phytozome v13.1 and subsequently analyzed using the PlantCARE database (accessed on 24 November 2025) to identify potential cis-regulatory elements associated with hormonal signaling and stress responses.

2.5 Tissue-Specific Expression Profiles of SbDUF506 Member in Sorghum

Tissue-specific expression data for *SbDUF506* genes were retrieved from the Phytozome *S. bicolor* expression dataset (accessed on 25 November 2025). The dataset included internode, flag leaf, leaf, root, shoot, and stem. Gene expression levels were transformed using $\log_2$(FPKM + 1), hierarchically clustered, and visualized as a heatmap using TBtools. This analysis revealed distinct tissue-specific expression patterns and provided insights into the potential functional roles of *SbDUF506* family members in different Sorghum tissues.

2.6 Plant Material and Abiotic Stress

The JinZa genotype of Sorghum was chosen for the examination of gene expression. The seeds were obtained from the Joint International Research Laboratory of Agriculture and Agri-Product Safety, Ministry of Education, Yangzhou University, China. Seeds were surface-sterilized and germinated in Petri dishes under controlled environmental conditions at 24°C, 60% relative humidity, and a 16 h light/8 h dark photoperiod for 48 h. Uniformly germinated seedlings were then transferred to a hydroponic culture system containing Hoagland nutrient solution [26]. At the three-leaf stage, seedlings subjected to abiotic stress treatments comprising 150 mM NaCl (salinity stress), 4°C (cold stress), and 42°C (heat stress), with temperature manipulations conducted using a multifunctional incubator. Leaf samples were harvested at 0-, 6-, 12-, and 24-h post-treatment, with three biological replicates for each treatment and time point. Samples were promptly frozen in liquid nitrogen and preserved at −80°C until subsequent utilization.

2.7 RNA Extraction and RT-qPCR Analysis

Total RNA was extracted from leaf tissues using the RNA Pure Plant Kit (DNase I) (Cat. No. CW0559S, CWBIO, Taizhou, China) according to the manufacturer's instructions, with DNase I treatment performed to eliminate residual genomic DNA contamination. First-strand cDNA was synthesized from 1 μg of total RNA using the HiScript III RT SuperMix for qPCR (Cat. #R323-01, Vazyme, Nanjing, China), which includes reverse transcriptase and RNase inhibitor for robust cDNA generation. Primer sequences were designed using the online tools provided by Integrated DNA Technologies (https://sg.idtdna.com/pages/tools), which were subsequently generated by Qingke Biotechnology. The primers were designed from the coding sequences of the target genes. The internal reference for normalization was the *SbPP2A* gene (Table S1), which has been validated as one of the most stable reference genes for qRT-PCR normalization in *S. bicolor* under abiotic stresses [27]. Quantitative reverse transcription polymerase chain reaction (qRT-PCR) was conducted utilizing a CFX96 real-time PCR apparatus (Bio-Rad, Hercules, CA, USA) with SYBR qPCR Master Mix (Catalog #Q711, Vazyme, Nanjing, China). The qPCR conditions were as follows: an initial denaturation step at 95°C for 30 s, followed by 40 cycles of denaturation at 95°C for 10 s and annealing/extension at 60°C for 30 s. Relative gene expression levels (fold change) were calculated using the $2^{-\Delta\Delta Ct}$ method. Three biological replicates were used to determine the cycle threshold (Ct) values.

2.8 Data Analysis

Relative expression was calculated using the $2^{-\Delta\Delta Ct}$ method, with SbPP2A as the internal reference gene and the corresponding 0 h sample as the control. Technical replicates were averaged before analysis,

and three independent biological replicates were included for each treatment and sampling time point. For each gene and stress treatment, differences in relative expression among the 0-, 6-, 12-, and 24-h time points. Statistical analyses were performed using one-way ANOVA followed by Tukey's HSD post hoc test. A probability value of $p < 0.05$ was considered indicative of statistically significant differences.

3 Results

3.1 Comprehensive Analysis and Characterization of the SbDUF506 Gene Family

A comprehensive bioinformatics investigation discovered nine *DUF506* genes in the genome of *Sorghum bicolor*. The indicated genes were systematically labeled *SbDUF506-1* to *SbDUF506-9* according to their chromosomal locations and linear sequence. The nine *SbDUF506* genes were unevenly distributed across five *S. bicolor* chromosomes (Fig. 1). Chromosome 1 contained *SbDUF506-1*, *SbDUF506-2*, and *SbDUF506-3*, whereas chromosome 3 contained *SbDUF506-5*, *SbDUF506-6*, and *SbDUF506-7*. Chromosomes 2, 4, and 9 each contained one gene, namely *SbDUF506-4*, *SbDUF506-8*, and *SbDUF506-9*, respectively. The encoded proteins displayed significant variation in sequence length 281–474 amino acids, with predicted molecular weights of 30.87–50.44 kDa and theoretical isoelectric points (pI) from 6.32 to 9.24, reflecting considerable diversity in their physicochemical properties (Table S2). Predictions of subcellular localization indicated that SbDUF506 proteins are dispersed among various cellular compartments, with three proteins found in the nucleus, five in the chloroplast, and one in the cytoplasm. The distinct subcellular localization patterns suggest potential functional diversification among *SbDUF506* family members, indicating their possible involvement in diverse cellular regulatory processes and responses to abiotic stresses (Table S2).

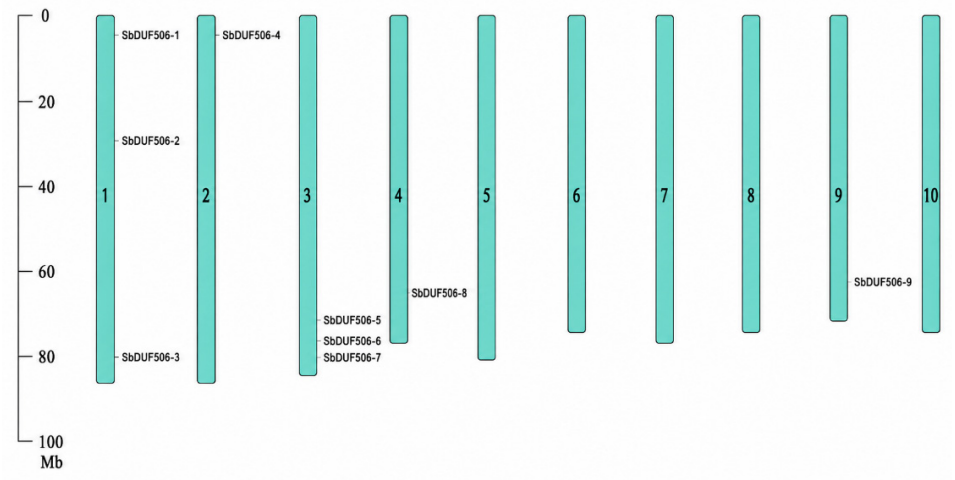


Figure 1: Chromosomal distribution of *SbDUF506* genes in *Sorghum bicolor*. The chromosome numbers and corresponding *SbDUF506* gene names are indicated, while the black scale on the left represents the physical length of each sorghum chromosome.

3.2 Phylogenetic Analysis and Multiple Sequence Alignment of SbDUF506 Genes

The phylogenetic relationships among SbDUF506 proteins and DUF506 homologs from *A. thaliana* were examined using the neighbor-joining method based on the complete protein sequences (Fig. 2). All *SbDUF506* members were categorized into four unique groups based on branch lengths and tree topology. The clustering pattern illustrates the evolutionary relationships among the *DUF506* family members and suggests possible conservation or diversification among the different phylogenetic groups.

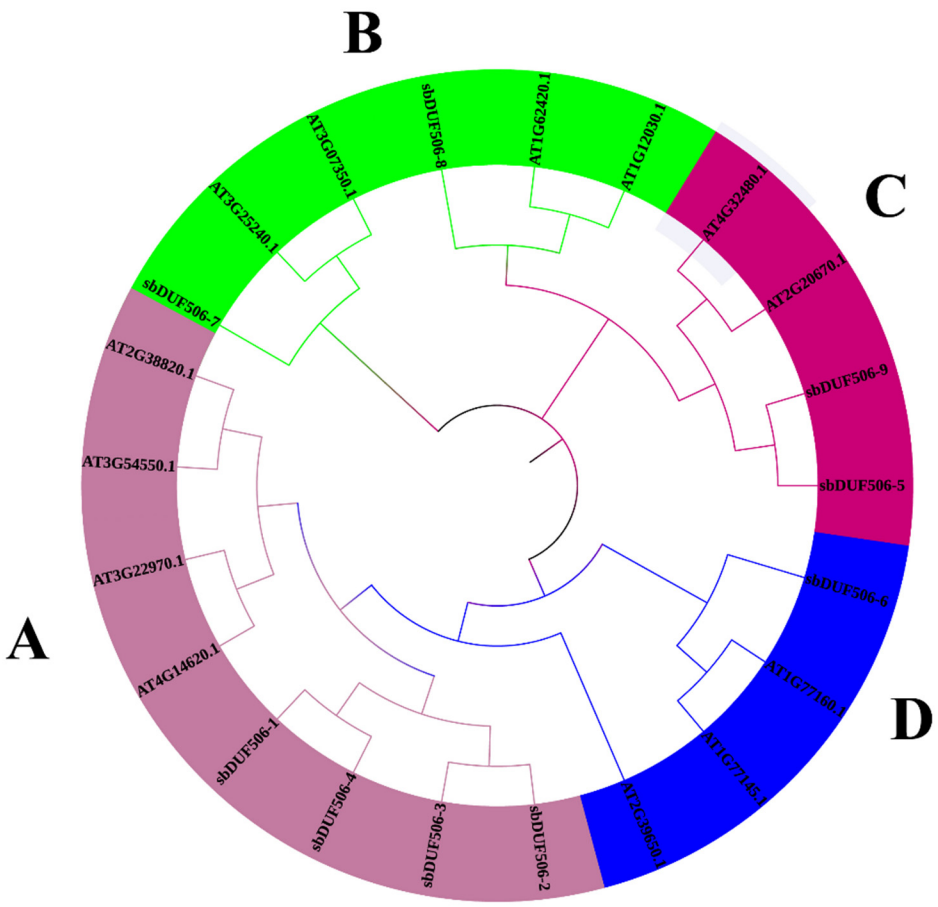


Figure 2: Phylogenetic relationships of DUF506 proteins from *Arabidopsis thaliana* and *Sorghum bicolor*. The phylogenetic tree was constructed using the neighbor-joining method, and the DUF506 proteins were classified into four distinct groups (A–D), highlighted by different colored arcs.

Multiple DUF506 protein sequences from sorghum were aligned, demonstrating significant conservation mostly in the central and C-terminal domains, but the N-terminal and loop regions exhibited substantial diversity (Fig. S1). Essential conserved residues comprise cysteine (C), glycine (G), and charged amino acids (E, K, R), indicating that these areas constitute a stable functional core. Conversely, the variable areas may confer structural flexibility or facilitate species-specific adaptations. The alignment suggests that DUF506 proteins preserve crucial conserved motifs for functionality, but flexible peripheral regions may facilitate various functions in Sorghum development and stress responses.

3.3 Structural Organization and Conserved Characteristics of SbDUF506 Family Members

Using the MEME online platform, ten conserved motifs were found within the SbDUF506 protein family (Fig. 3A). Motifs 1–4 were present in all family members, indicating the same subfamily may have the same biological function. while Motifs 5, 6, and 10 were preserved across most of the proteins (Fig. 3B). To further explore structural diversity, exon–intron organization an important factor of genetic architecture and evolution dynamics was analyzed. Genomic annotation of *S. bicolor* based on published GTF data reveals that *SbDUF506* genes possess between one and three exons and a maximum of two introns. The majority of genes possess both 5′ and 3′ untranslated regions (UTRs), whereas several members possessed a UTR at only one end or lacked detectable UTRs altogether (Fig. 3C). Notably, genes belonging to the

identical phylogenetic grouping exhibited highly similar exon–intron structures, suggesting that paralogous genes sharing conserved domains tend to maintain similar structural configurations.

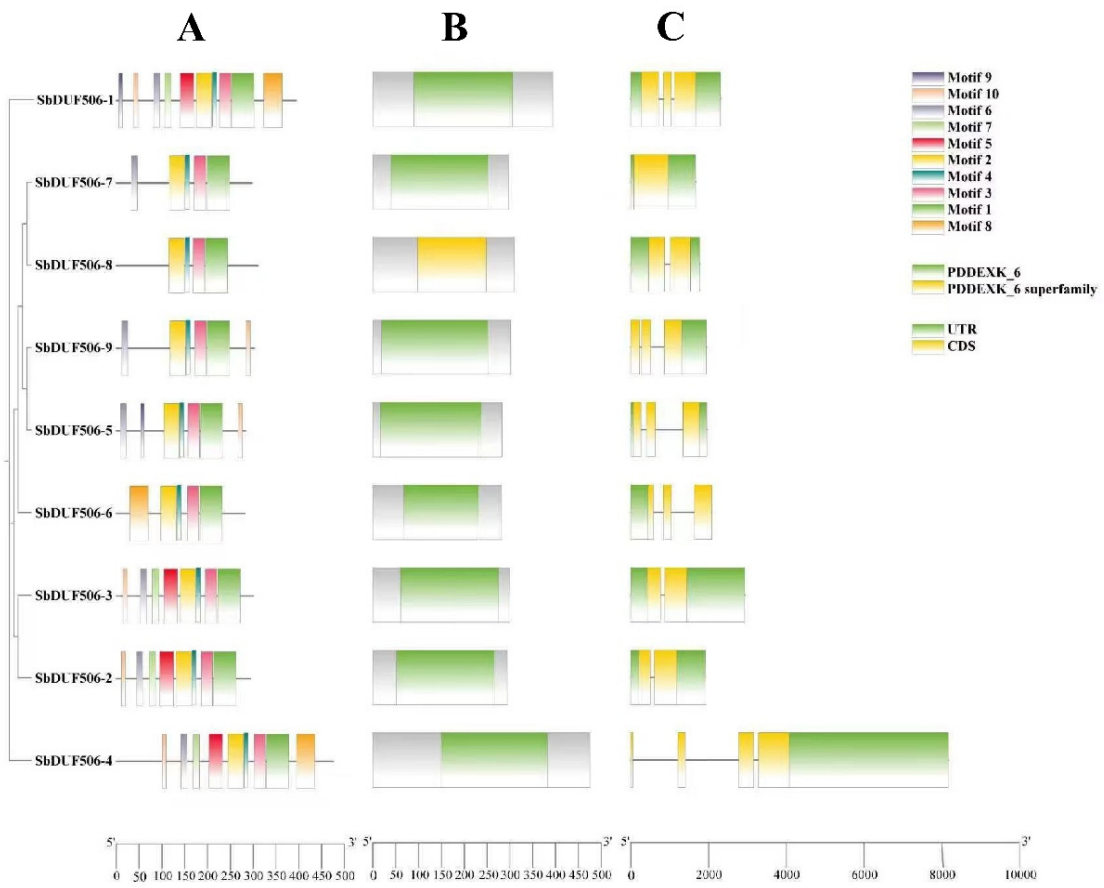


Figure 3: Distribution of conserved motifs, domains, and gene structures of the *SbDUF506* family. (**A**) Conserved motif distribution among *SbDUF506* proteins. (**B**) Identification and visualization of conserved domains within *SbDUF506* proteins. (**C**) Gene structure analysis illustrating exon (yellow boxes), intron (black lines), and untranslated region (UTR; green bars) organization, highlighting structural diversity among *SbDUF506* family members.

3.4 Analysis of Homologous Genes of SbDUF506 and Ka/Ks Evaluation

Homologous genes, which share structural and functional similarities as a consequence of common evolutionary ancestry, highlight the important role of gene duplication in driving evolutionary divergence and functional diversification among gene family members. Several *SbDUF506* paralogous gene pairs were identified within collinear regions located on different chromosomes and were therefore classified as whole-genome duplication (WGD)/segmental duplication pairs (Fig. 4). These findings indicate that the expansion of the *SbDUF506* gene family was driven predominantly by large-scale genomic duplication events. A comparative collinearity study was performed to clarify the evolutionary history and conservation patterns of *DUF506* genes across several plant lineages. The study, involving the monocots *S. bicolor* and *O. sativa* as well as the dicot *A. thaliana*, uncovered lineage-specific duplication events and syntenic interactions. Significantly, twelve collinear gene pairs were found among *S. bicolor* and *O. sativa*, in contrast to merely three pairs between *S. bicolor* and *A. thaliana* (Fig. 5). This pronounced synteny between the two monocot species indicates a higher degree of evolutionary conservation and implies significant functional similarity between their respective *DUF506* gene families. To assess the evolutionary selection pressures acting on

the *SbDUF506* gene family, the ratio of nonsynonymous (Ka) to synonymous (Ks) substitution rates was calculated for each duplicated gene pair. The Ka/Ks ratio is widely used to infer selective constraints on homologous genes: values below 1 indicate purifying selection, values equal to 1 suggest neutral evolution, and values above 1 are indicative of positive selection. The coding sequences of each paralogous gene pair were aligned, and the numbers of nonsynonymous substitutions per nonsynonymous site (Ka) and synonymous substitutions per synonymous site (Ks) were estimated using the Nei–Gojobori method implemented in the KaKs Calculator 2.0 module of TBtools (Table S3). All *SbDUF506* paralogous gene pairs exhibited Ka/Ks ratios below 1, indicating that this gene family has predominantly evolved under purifying selection. This selective constraint likely contributed to the removal of deleterious mutations and the conservation of essential biological functions during the evolutionary history of the *SbDUF506* gene family.

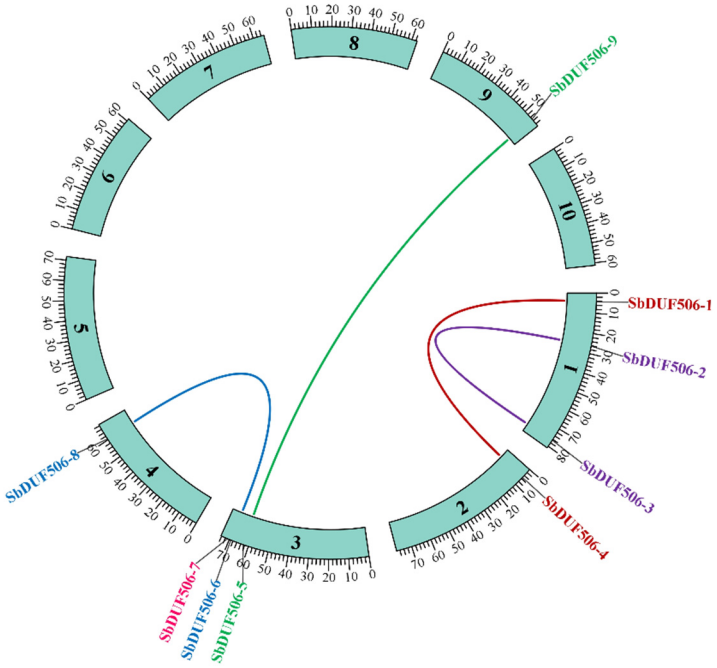


Figure 4: Intraspecies collinearity and duplicated gene pairs of the *SbDUF506* family in the *S. bicolor* genome. Colored curves connect collinear *SbDUF506* gene pairs located on different chromosomes.

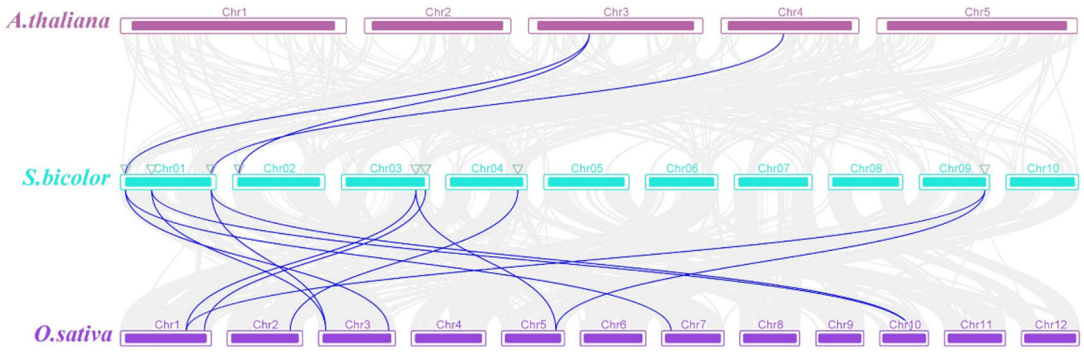


Figure 5: Synteny study of *DUF506* genes across *S. bicolor*, *O. sativa*, and *A. thaliana*; blue lines emphasize syntenic *DUF506* gene pairs.

3.5 Identifying Cis-Acting Regulatory Elements in the DUF506 Gene Family

Cis-regulatory elements, located within promoter regions or nearby regulatory sequences, are essential for influencing transcriptional activity through interactions with transcription factors (TFs), thus influencing gene expression. In this investigation, using the PlantCARE database, we examined the 2-kb upstream promoter regions of the *SbDUF506* gene family to find potential cis-acting regulatory elements. The identified cis-regulatory elements were broadly classified into three major functional categories: (i) light-responsive elements, including Box 4, G-box, TCT-motif, and GATA-motif; (ii) environmental stress-responsive elements associated with drought, light, and other stress conditions; and (iii) transcription factor-binding sites involved in the regulation of gene expression. These results indicate that *SbDUF506* genes are potentially regulated by a complex network of developmental, hormonal, and environmental cues.

The promoter regions of *SbDUF506* genes contain diverse cis-acting regulatory elements, suggesting their potential involvement in a broad range of biological and regulatory processes in *S. bicolor*. Several light-responsive elements, including Box 4, G-box, GATA-motif, and TCT-motif, were identified in the *SbDUF506* promoters. Stress- and hormone-responsive elements, including MYB-, MYC-, ABRE-, and MBS-related motifs, were also detected (Table S4). To explore the potential roles of *SbDUF506* genes in abiotic stress adaptation, we focused on 10 cis-regulatory elements associated with environmental stress responses and stress-related phytohormone signaling pathways. These regulatory elements were consistently detected across the promoter regions of all *SbDUF506* genes, highlighting their likely contribution to stress-responsive regulatory mechanisms (Fig. 6).

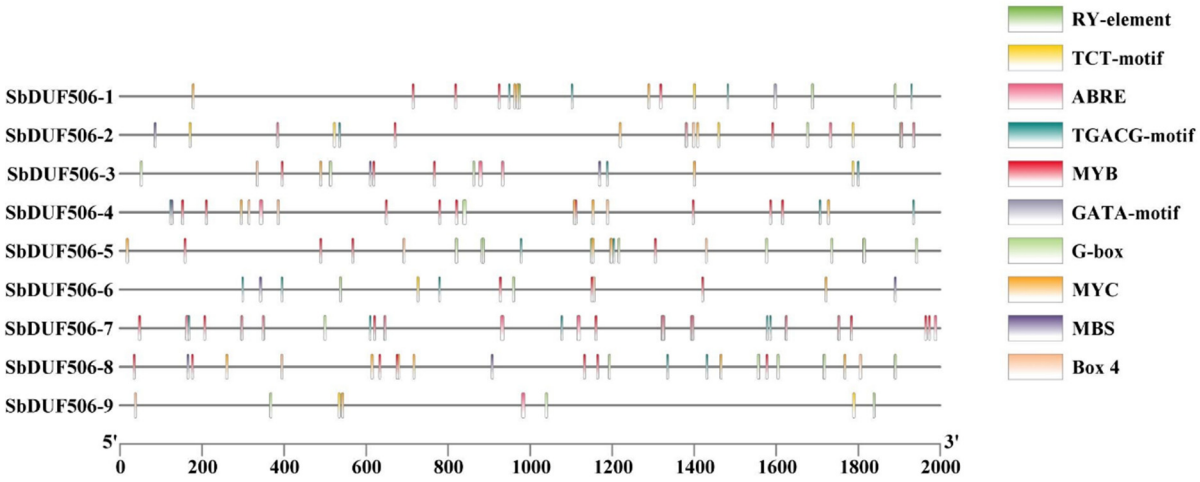


Figure 6: Analysis of cis-regulatory elements in the promoter regions of *SbDUF506* genes. Cis-acting regulatory elements were identified within the 2000-bp upstream promoter sequences of *SbDUF506* genes using the PlantCARE database.

3.6 Expression Profile of SbDUF506 Gene across Various Tissues

Gene expression represents a crucial link between genomic information and the realization of biological function. Genome-wide expression analyses offer important insights into transcriptional patterns and their underlying regulatory mechanisms. To investigate the tissue-specific expression profiles of *SbDUF506* genes, transcriptomic data were retrieved from the Phytozome database (Table S5). The analysis encompassed multiple *S. bicolor* tissues, including internode, leaf, flag leaf, root, shoot, and stem. Raw FPKM values are using for heatmap visualization, the values were transformed as $\log_2$(FPKM + 1) and plotted using TBtools, as illustrated in Fig. 7, most *SbDUF506* genes were expressed in at least one tissue type. Notably, *SbDUF506-1*,

SbDUF506-2, *SbDUF506-3*, *SbDUF506-4*, *SbDUF506-5*, *SbDUF506-7*, and *SbDUF506-9* exhibited broad expression across all examined tissues. Among them, *SbDUF506-1* showed consistently elevated expression levels in all examined tissues (Fig. 7). These findings demonstrate distinct expression patterns among *SbDUF506* genes, indicating potential functional diversification and tissue-specific regulatory functions within the gene family.

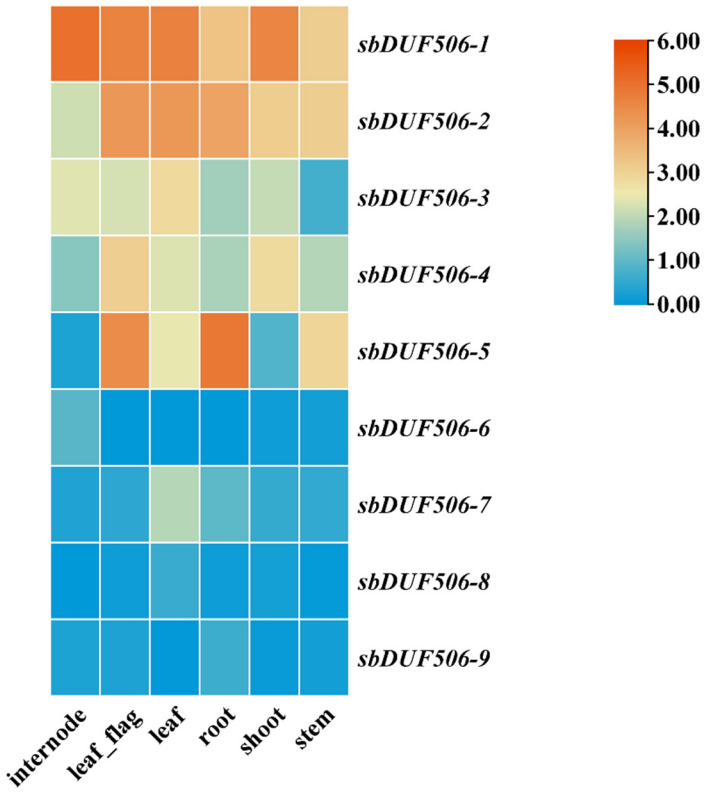
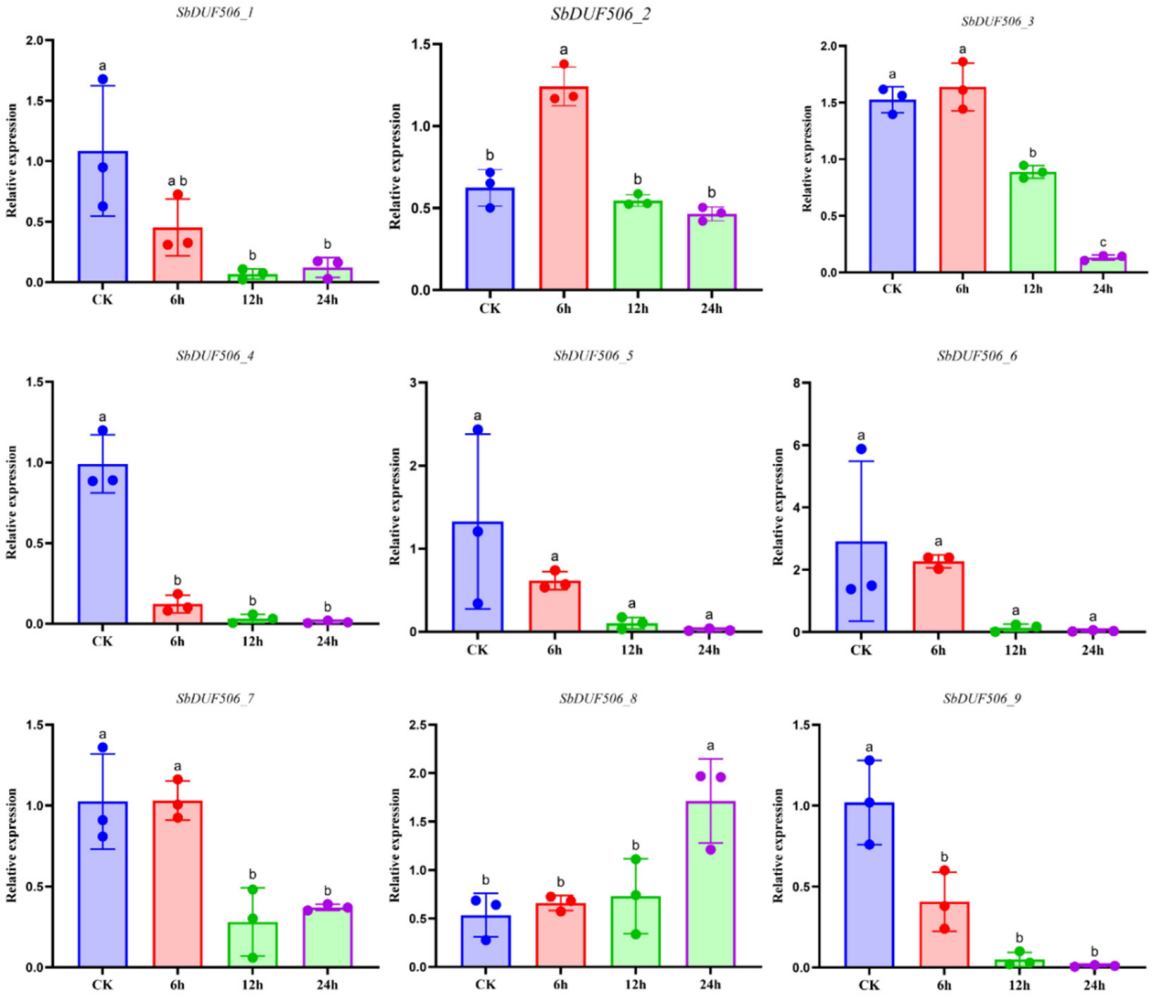


Figure 7: Tissue-specific expression profiles of *SbDUF506* genes in *S. bicolor*. The heatmap illustrates the expression patterns of *SbDUF506* genes across six tissues based on RNA-seq data. Expression values were normalized as fragments per kilobase of transcript per million mapped reads (FPKM) and transformed using $\log_2$(FPKM + 1).

3.7 Expression Profiling of SbDUF506 Genes under Abiotic Stress Conditions

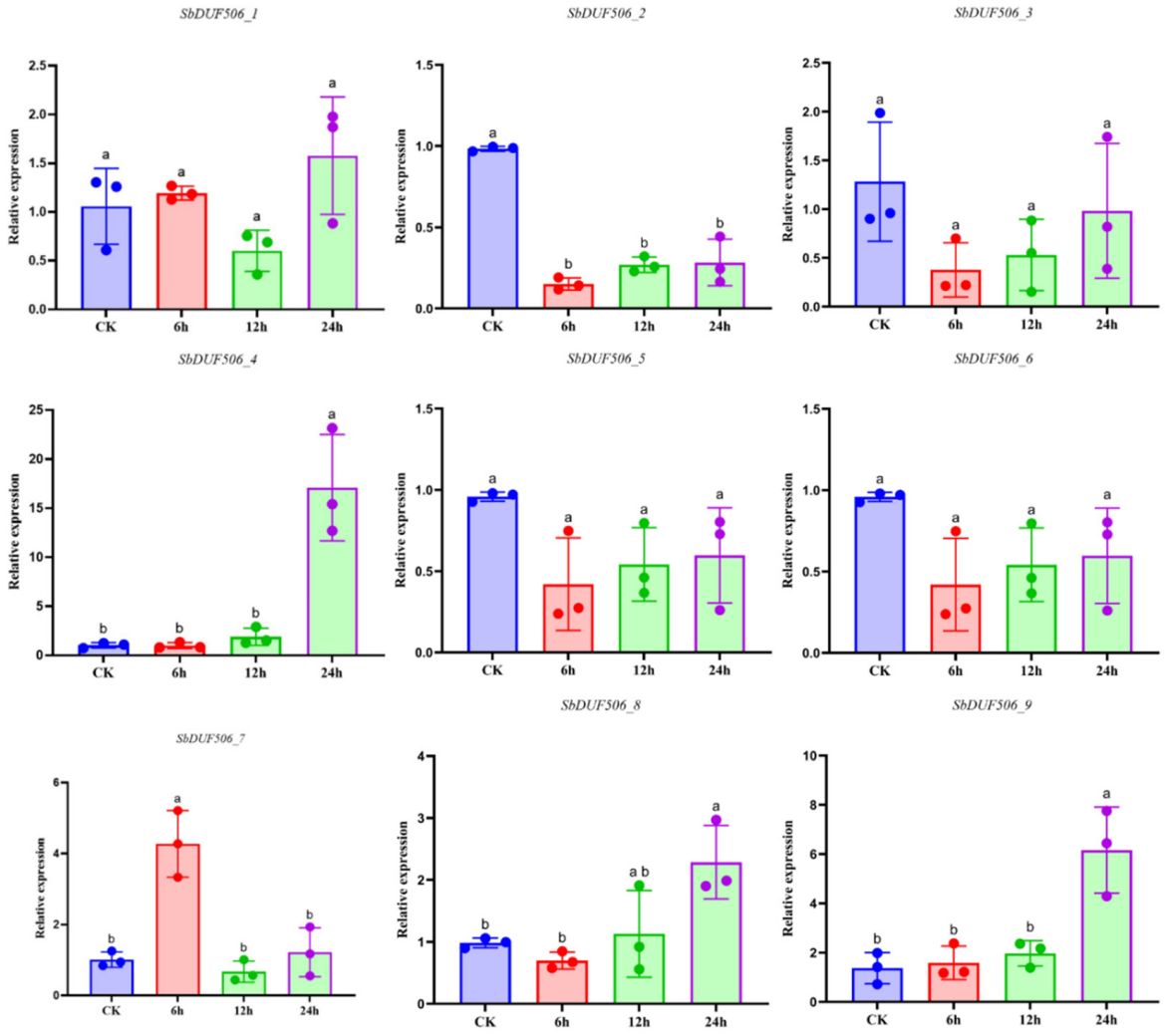
To further elucidate the expression dynamics of *SbDUF506* genes under abiotic stress conditions, qRT-PCR analysis was performed under heat, salinity, and cold treatments (Fig. 8). Samples harvested at 0 h were used as the control (CK). Overall, most *SbDUF506* genes exhibited stress-responsive expression patterns, with significant temporal variation across treatments. Under heat stress, *SbDUF506* genes displayed diverse and time-dependent expression patterns. *SbDUF506-2* showed a rapid and transient induction at 6 h, followed by a gradual decrease at later time points. In contrast, *SbDUF506-8* exhibited a continuous increase in expression, reaching its maximum level at 24 h, suggesting a potential role in the prolonged heat stress response. Conversely, *SbDUF506-1*, *SbDUF506-3*, *SbDUF506-4*, *SbDUF506-7*, and *SbDUF506-9* were progressively downregulated during heat treatment, whereas *SbDUF506-5* and *SbDUF506-6* showed a declining expression trend. Under salinity stress, *SbDUF506* genes exhibited diverse and time-dependent expression patterns. *SbDUF506-7* showed a rapid and strong induction at 6 h, followed by a decrease at 12 h and 24 h, indicating an early response to salt stress. In contrast, *SbDUF506-4*, *SbDUF506-8*, and *SbDUF506-9* displayed substantial upregulation at 24 h, suggesting their potential involvement in the late stages of

salinity adaptation. *SbDUF506-1* also showed a gradual increase in expression, reaching the highest level at 24 h. Conversely, *SbDUF506-2* was significantly downregulated throughout the treatment period, while *SbDUF506-3*, *SbDUF506-5*, and *SbDUF506-6* showed relatively stable expression patterns with no significant changes. Under cold stress, *SbDUF506* genes displayed diverse and time-dependent expression responses. *SbDUF506-1* showed a transient induction, with expression increasing at 6 h and reaching the highest level at 12 h before declining at 24 h. *SbDUF506-2*, *SbDUF506-3*, and *SbDUF506-9* exhibited strong late-stage induction, with their transcript levels markedly increased at 24 h, suggesting their involvement in prolonged cold stress responses. Similarly, *SbDUF506-8* showed an upregulation at 24 h after a slight reduction at early time points. *SbDUF506-4* was initially suppressed at 6 h but gradually increased and reached the highest expression level at 24 h. In contrast, *SbDUF506-5* showed a gradual increase during cold treatment, whereas *SbDUF506-6* and *SbDUF506-7* remained relatively stable with no significant changes. Collectively, these results demonstrate that *SbDUF506* family members respond to salinity stress in a gene-specific and time-dependent manner, suggesting their differential roles in salt stress adaptation.



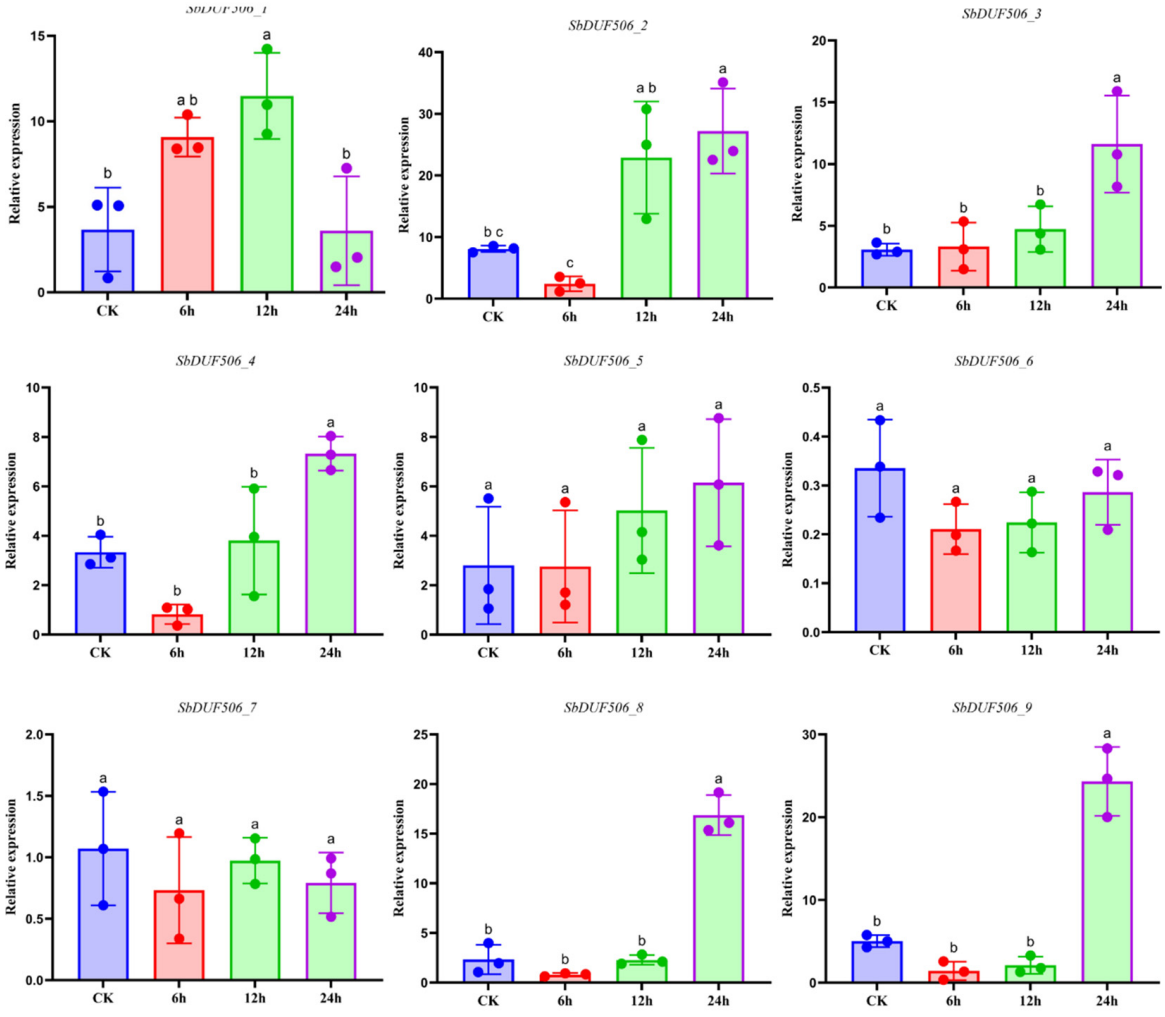
(**A**)

Figure 8: *Cont.*



(**B**)

Figure 8: *Cont.*



(**C**)

Figure 8: Temporal expression profiles of *SbDUF506* genes under heat, salinity, and cold stress. Relative transcript levels of nine *SbDUF506* genes were determined by RT-qPCR at 0 h (CK), 6 h, 12 h, and 24 h under (**A**) heat, (**B**) salinity, and (**C**) cold stress. Relative expression was calculated using the $2^{-\Delta\Delta Ct}$ method, with *SbPP2A* used as the internal reference gene and the corresponding 0 h sample used as the calibrator. Bars represent the mean ± SD of three biological replicates, and individual data points indicate biological replicates. Different lowercase letters indicate significant differences among time points based on one-way ANOVA followed by Tukey's multiple-comparison test ($p < 0.05$).

4 Discussion

The Domain of Unknown Function (DUF) protein families constitute a diverse array of proteins with mainly uncharacterized biological activities; however, they are increasingly associated with plant growth and responses to environmental stressors. Members of these families typically exhibit conserved structural and functional features and are often associated with related physiological processes. The *DUF506* gene family, recently characterized in higher plants, has attracted increasing attention due to its potential roles in regulating plant growth, development, and adaptation to various abiotic stresses [17,28]. Despite these

initial findings, extensive research examining the complete spectrum of roles of *DUF506* genes in plant development and stress responses is scarce and necessitates future thorough exploration.

Sorghum (*Sorghum bicolor*), a cereal crop originating from Africa, ranks as the fifth most important grain crop worldwide. It is well recognized for its remarkable tolerance to a broad range of abiotic stresses, including drought, flooding, salinity, nutrient-poor soils, and high temperatures. As a C4 species, sorghum possesses high photosynthetic efficiency and substantial biomass production capacity, contributing to its considerable agronomic value and ecological significance in sustainable agricultural systems [19,28].

This research presents a thorough genome-wide identification and characterization of the *DUF506* gene family in *S. bicolor* conducted using bioinformatics approaches. A total of nine *SbDUF506* genes were identified from the reference genome (Sorghum v3.1.1). Comparative genomic analysis revealed 13 *DUF506* genes in *A. thaliana*, 10 in *O. sativa*, 18 in *B. rapa*, and 26 in *T. aestivum*. These differences in gene family size across species suggest that gene duplication and gene loss events during evolution have played a key role in shaping the diversification of the *DUF506* gene family [17,28,29]. Phylogenetic analysis of *S. biocolor*, *A. thaliana*, and *O. sativa DUF506* genes enabled their classification into distinct evolutionary clades, offering valuable insights into their potential functional divergence and evolutionary relationships.

The variation in motif composition and abundance within the *S. bicolor DUF506* gene family may contribute to the functional diversification of its members. Notably, motifs 1, 2, 3, and 4 were present in all DUF506 proteins, suggesting that a motif is essential for the discovery of the Sorghum *DUF506* genes. Additionally, our study found that the quantity of exons ranged from 1 to 3, but the quantity of introns varied from 1 to 2. This differs from previous studies, where exon numbers ranged from 1 to 3, in *T. aestivum* containing up to 2 introns [28].

Homologous genes (HGs), which share structural and functional similarities due to a common evolutionary origin either within the same genome or across different species, provide valuable insights into gene function and evolutionary relationships. Gene duplication is a significant catalyst in the evolution and diversity of gene families. The synteny analysis conducted in this study between the *S. bicolor* genome and two other plant genomes revealed that duplicated *DUF506* genes have predominantly undergone purifying selection following gene duplication events. This suggests that the activities of the duplicated *DUF506* genes have likely been preserved via later evolutionary developments. This finding is consistent with previous studies of the *TaDUF506* gene family, which similarly reported duplicated genes subjected to both positive and purifying selection pressures [29]. Cis-acting elements located within gene promoter regions play pivotal roles in regulating transcription through their interactions with specific transcription factors. These regulatory sequences contribute to the precise control of gene expression during various physiological and developmental processes, as well as in response to environmental stresses. Increasing evidence indicates that members of the *DUF506* gene family are involved in the regulation of plant growth and development and contribute to plant adaptation to diverse abiotic stress conditions [20,28]. In this study, cis-acting regulatory elements were systematically analyzed within the 2000-bp upstream promoter regions of *SbDUF506* genes in *S. bicolor*. The cis-elements were classified into three primary functional categories: light-responsive elements, biotic and abiotic stress-responsive elements, and plant hormone-responsive elements. The existence of these varied regulatory patterns indicates that *SbDUF506* genes may play significant roles in governing plant growth, developmental processes, and responses to environmental stressors. A different finding was reported in *O. sativa DUF506* genes, which have cis-acting elements related to hormone response, abiotic stress response, and plant growth and metabolism [18]. Methyl jasmonate (MeJA) and abscisic acid (ABA) are key phytohormones involved in the regulation of diverse physiological and molecular signaling pathways in plants. Both hormones play important roles in plant defense against biotic challenges and in

enhancing tolerance to various abiotic stresses, including drought, salinity, low-temperature stress, and heavy metal toxicity [18,29]. In the present study, a substantial abundance of MeJA- and ABA-responsive cis-acting regulatory elements was detected within the promoter regions of *SbDUF506* genes. This finding is consistent with previous reports describing similar distributions of hormone-responsive cis-elements in the promoters of *OsDUF506* genes in *O. sativa* [18]. Moreover, several additional cis-regulatory elements associated with abiotic stress responses were identified in the *SbDUF506* promoters, further supporting the potential involvement of this gene family in stress adaptation. These observations are also in agreement with previous studies reporting the participation of *DUF506* genes in abiotic stress responses in *O. sativa* [18].

Numerous studies have demonstrated that members of various DUF protein families play important roles in enhancing plant tolerance and resistance to diverse environmental stresses [17,18]. This study further investigated the expression patterns of nine *SbDUF506* genes under heat, salinity, and cold stresses using RT–qPCR analysis. The results revealed distinct stress-specific and time-dependent expression profiles among *SbDUF506* members. Under heat stress, *SbDUF506-2* showed an early transient induction, whereas *SbDUF506-8* exhibited continuous upregulation at later stages. Under salinity stress, *SbDUF506-7* responded rapidly at 6 h, while *SbDUF506-1*, *SbDUF506-4*, *SbDUF506-8*, and *SbDUF506-9* showed significant induction at 24 h. Cold stress induced a broader gene response, with *SbDUF506-2*, *SbDUF506-3*, *SbDUF506-8*, and *SbDUF506-9* strongly upregulated at 24 h, while *SbDUF506-1* showed transient induction and *SbDUF506-4* displayed a biphasic expression pattern. These findings indicate that *SbDUF506* genes are involved in diverse and stage-specific responses to abiotic stresses, providing potential candidates for further functional characterization of stress tolerance mechanisms in Sorghum.

5 Conclusion

The genome-wide analysis of the *DUF506* gene in Sorghum produced significant insights into its putative functions and regulatory mechanisms. We have identified and described nine *DUF506* genes in the Sorghum genome through extensive bioinformatics investigations and expression profiling. Our research indicates that the *DUF506* genes are crucial for multiple biological processes, encompassing plant development and stress responses. The variable expression patterns of *DUF506* genes under various abiotic stressors suggest their role in Sorghum adaptability to such conditions. This study provides a foundation for subsequent research on the specific roles and potential applications of *DUF506* genes in Sorghum breeding and crop improvement.

Acknowledgement: None.

Funding Statement: The research receives funding from the China National Key research and development (R&D) Project (2022 YFE0113400).

Author Contributions: Esra Alther: experimental investigation, analysis of data, and writing the origin article and rewriting editing. Hind Abdelmonim Elsanosi: helpful assistance and discussion. Ghazi Badawi: assistance and reviewing and editing. Guisheng Zhou: project administration and funding acquisition, reviewing and editing and supervision. All authors reviewed and approved the final version of the manuscript.

Availability of Data and Materials: All data included in this work are available in article and Supplementary Materials.

Ethics Approval: No applicable.

Conflicts of Interest: The authors declare no conflicts of interest.

Supplementary Materials: The supplementary material is available online at https://www.techscience.com/doi/10.32604/phyton.2026.087409/s1. Table S1: Description of primers utilized for the real-time PCR (qRT-PCR) assessment; Table S2: List of *SbDUF506* genes, physicochemical properties, and subcellular localization of predicted Sorghum DUF506 proteins; Table S3: Ka/Ks Analysis in *SbDUF506* Protein; Table S4: Details of the cis-elements identified in the *SbDUF506* gene family; Table S5: The *SbDUF506* gene expression profile in different tissues; Figure S1: Protein sequence alignment of the nine SbDUF506 homologs in sorghum.

References

1. Lv P, Wan J, Zhang C, Hina A, Al Amin GM, Begum N, et al. Unraveling the diverse roles of neglected genes containing domains of unknown function (DUFs): Progress and perspective. Int J Mol Sci. 2023;24(4):4187. [CrossRef].
2. Luo C, Akhtar M, Min W, Bai X, Ma T, Liu C. Domain of unknown function (DUF) proteins in plants: Function and perspective. Protoplasma. 2024;261(3):397–410. [CrossRef].
3. Zhong H, Zhang H, Guo R, Wang Q, Huang X, Liao J, et al. Characterization and functional divergence of a novel *DUF668* gene family in rice based on comprehensive expression patterns. Genes. 2019;10(12):980. [CrossRef].
4. Zhang Y, Zhang F, Huang X. Characterization of an *Arabidopsis thaliana* DUF761-containing protein with a potential role in development and defense responses. Theor Exp Plant Physiol. 2019;31(2):303–16. [CrossRef].
5. Sakata M, Takano-Kai N, Miyazaki Y, Kanamori H, Wu J, Matsumoto T, et al. Domain unknown function DUF1668-containing genes in multiple lineages are responsible for F_1 pollen sterility in rice. Front Plant Sci. 2020;11:632420. [CrossRef].
6. Stonebloom S, Ebert B, Xiong G, Pattathil S, Birdseye D, Lao J, et al. A DUF-246 family glycosyltransferase-like gene affects male fertility and the biosynthesis of pectic Arabinogalactans. BMC Plant Biol. 2016;16:90. [CrossRef].
7. Alther E, Abdelmula AA, Elsanosi HA, Mostafa S, Zhou G. Comprehensive genome characterization and expression analysis of the *DUF4228* gene family in Sorghum (*Sorghum bicolor* L.) under salt stress conditions. PeerJ. 2026;14:e21175. [CrossRef].
8. Cao D, Liu C, Zhang W, Zheng C, Zhang S, Jia H, et al. Characterization of the *DUF868* gene family in *Nicotiana* and functional analysis of *NtDUF868-E5* involved in pigment metabolism. Plant Physiol Biochem. 2024;208:108493. [CrossRef].
9. Nabi RBS, Tayade R, Imran QM, Hussain A, Shahid M, Yun BW. Functional insight of nitric-oxide induced DUF genes in *Arabidopsis thaliana*. Front Plant Sci. 2020;11:1041. [CrossRef].
10. Guo C, Luo C, Guo L, Li M, Guo X, Zhang Y, et al. *OsSIDP366*, a DUF1644 gene, positively regulates responses to drought and salt stresses in rice. J Integr Plant Biol. 2016;58(5):492–502. [CrossRef].
11. Cui Y, Wang M, Zhou H, Li M, Huang L, Yin X, et al. *OsSGL*, a novel DUF1645 domain-containing protein, confers enhanced drought tolerance in transgenic rice and *Arabidopsis*. Front Plant Sci. 2016;7:2001. [CrossRef].
12. Li L, Lv M, Li J, Ye T, Li X, He X, et al. Molecular characterisation and function analysis of the rice *OsDUF872* family. Biotechnol Biotechnol Equip. 2018;32(4):881–7. [CrossRef].
13. Ma G, Zhang Y, Li X. Overexpression of *OsDUF6* increases salt stress tolerance in rice. BMC Plant Biol. 2024;24(1):216. [CrossRef].
14. Knizewski L, Kinch LN, Grishin NV, Rychlewski L, Ginalski K. Realm of PD-(D/E)XK nuclease superfamily revisited: Detection of novel families with modified transitive meta profile searches. BMC Struct Biol. 2007;7:40. [CrossRef].
15. Kosinski J, Feder M, Bujnicki JM. The PD-(D/E)XK superfamily revisited: Identification of new members among proteins involved in DNA metabolism and functional predictions for domains of (hitherto) unknown function. BMC Bioinform. 2005;6(1):172. [CrossRef].
16. Steczkiewicz K, Muszewska A, Knizewski L, Rychlewski L, Ginalski K. Sequence, structure and functional diversity of PD-(D/E)XK phosphodiesterase superfamily. Nucleic Acids Res. 2012;40(15):7016–45. [CrossRef].
17. Ying S. Genome-wide identification and transcriptional analysis of *Arabidopsis DUF506* gene family. Int J Mol Sci. 2021;22(21):11442. [CrossRef].
18. Dong W, Tu J, Deng W, Zhang J, Xu Y, Gu A, et al. Genome-wide identification of *DUF506* gene family in *Oryzasativa* and expression profiling under abiotic stresses. PeerJ. 2023;11:e16168. [CrossRef].

19. Huang RD. Research progress on plant tolerance to soil salinity and alkalinity in sorghum. J Integr Agric. 2018;17(4):739–46. [CrossRef].
20. Tu M, Du C, Yu B, Wang G, Deng Y, Wang Y, et al. Current advances in the molecular regulation of abiotic stress tolerance in sorghum via transcriptomic, proteomic, and metabolomic approaches. Front Plant Sci. 2023;14:1147328. [CrossRef].
21. Tamura K, Stecher G, Kumar S. MEGA11: Molecular evolutionary genetics analysis version 11. Mol Biol Evol. 2021;38(7):3022–7. [CrossRef].
22. Zhou T, Xu K, Zhao F, Liu W, Li L, Hua Z, et al. itol.toolkit accelerates working with iTOL (Interactive Tree of Life) by an automated generation of annotation files. Bioinformatics. 2023;39(6):btad339. [CrossRef].
23. Bailey TL, Elkan C. Fitting a mixture model by expectation maximization to discover motifs in biopolymers. Proc Int Conf Intell Syst Mol Biol. 1994;2:28–36.
24. Chen C, Chen H, Zhang Y, Thomas HR, Frank MH, He Y, et al. TBtools: An integrative toolkit developed for interactive analyses of big biological data. Mol Plant. 2020;13(8):1194–202. [CrossRef].
25. Goodstein DM, Shu S, Howson R, Neupane R, Hayes RD, Fazo J, et al. Phytozome: A comparative platform for green plant genomics. Nucleic Acids Res. 2012;40:D1178–86. [CrossRef].
26. Hoagland DR, Arnon DI. The water-culture method for growing plants without soil. Berkeley, CA, USA: College of Agriculture, University of California; 1938.
27. Sudhakar Reddy P, Srinivas Reddy D, Sivasakthi K, Bhatnagar-Mathur P, Vadez V, Sharma KK. 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. 2016;7:529. [CrossRef].
28. Wang Y, Liu S, Lin T, Kan S, Yu P. The *DUF506* gene family in *Triticum aestivum*: Genome-wide identification and expression profiling under salt stress. Agronomy. 2025;15(2):281. [CrossRef].
29. Zhu G, Wang J, He S, Liang K, Zhang R, Huang J, et al. Comprehensive analysis of *BrDUF506* genes across the *Brassica rapa* genome uncovers potential functions in sexual reproduction and abiotic stress tolerance. Int J Mol Sci. 2024;25(20):11087. [CrossRef].