

Genome-Wide Characterization and Expression Profiling of the LSD Gene Family in Cotton (*Gossypium hirsutum*) Under Drought and Salt Stress

Ayşe Arucan

Istanbul University

Cüneyt Uçarlı

ucarlicu@istanbul.edu.tr

Istanbul University

Research Article

Keywords: Lesion Simulating Disease, abiotic stress, salinity, transcription factors, stress tolerance, zinc finger proteins

Posted Date: May 8th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-9441383/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

Additional Declarations: No competing interests reported.

Abstract

Cotton (*Gossypium hirsutum*) is a major natural fiber crop whose productivity is severely constrained by drought and salinity stress. Understanding the molecular mechanisms underlying stress responses is critical for developing stress-resilient cultivars. The Lesion Simulating Disease (LSD) transcription factor family plays a crucial role in regulating programmed cell death and stress signaling pathways in plants. In this study, a genome-wide analysis identified eleven *GhLSD* genes in *G. hirsutum*, which were unevenly distributed across chromosomes and contained conserved LSD-type zinc finger domains, suggesting their involvement in stress-responsive regulatory networks. Expression profiling of six selected *GhLSD* genes under drought and salinity conditions revealed complex, tissue-specific, and time-dependent expression patterns. Under drought stress, *GhLSD1*, *GhLSD4*, and *GhLSD6* were significantly upregulated, indicating their potential roles as positive regulators of stress tolerance, whereas *GhLSD2* was consistently downregulated, suggesting a possible negative regulatory function. Similarly, under salinity stress, *GhLSD4*, *GhLSD5*, and *GhLSD6* exhibited strong and sustained induction, highlighting their importance in long-term and tissue-specific salt stress responses. In contrast, the persistent downregulation of *GhLSD2* further supports its potential role as a negative regulator in salinity stress adaptation. Overall, these findings demonstrate that the *GhLSD* gene family functions in a coordinated and dynamic manner in response to abiotic stress, with *GhLSD4*, *GhLSD5*, and *GhLSD6* emerging as promising candidate genes for improving drought and salinity tolerance in cotton through molecular breeding approaches.

Introduction

Cotton (*Gossypium spp.*) is an economically important crop that supplies nearly 35% of the world's natural fiber and also serves as a source of vegetable oil and bioenergy (Abdelraheem et al. 2019; Liu et al. 2021). In the 2024–2025 growing season, global cotton production reached around 121.07 million 480-lb bales, with China, India, Brazil, the United States, Australia, Pakistan, and Türkiye as the main producers (USDA, 2025). Türkiye ranks seventh globally, contributing about 3% of world production—mainly from the Southeastern Anatolia and Aegean regions—with an average yield of 1,799 kg/ha between 2020/2021 and 2024/2025 (Najib et al. 2022; USDA, 2025).

Cotton plays an important role in studying polyploid evolution, fiber development, and stress adaptation (Chen et al. 2020). *Gossypium* represents a genus of the Malvaceae family encompasses about 50 species—45 diploid and 5 allotetraploid—with *G. hirsutum* (AD_1) and *G. barbadense* (AD_2) arising from hybridization between *G. arboreum* (A_2) and *G. raimondii* (D_5) followed by chromosome doubling (Li et al. 2014; Chen et al. 2020). *G. hirsutum* accounts for nearly 95% of global production, while *G. arboreum* and *G. raimondii* remain significant genetic resources for plant tolerance in drought and salinity and genomic researches (Zhu et al. 2013; Yan et al. 2016; Shi et al. 2019).

Global climate change has intensified abiotic stresses such as drought and salinity, which negatively affect fiber quality, plant development, and overall cotton productivity (Khan et al. 2025). Although cotton exhibits moderate salt tolerance ($\approx 7.7 \text{ dS m}^{-1}$), severe stress can reduce yield by up to 67%, varying with stress severity and genotype (Maas and Hoffman 1977; Abdelraheem et al. 2021). These conditions impair

photosynthesis, root development, and nutrient uptake, resulting in major agronomic and economic losses (Bano et al. 2022).

Salinity is a critical condition that disrupts plant development by inducing osmotic stress, ionic toxicity, and nutrient imbalance (Ahmad et al. 2022). In cotton, excessive Na⁺ and Cl⁻ accumulation impairs water uptake, damages cellular membranes, and inhibits enzymatic activity, leading to oxidative stress and alleviate fiber characteristics (Lu et al. 2023; Chaudhary et al. 2024). Plants counteract these effects through ion homeostasis, antioxidant activation, and transcriptional regulation of genes associated with stress tolerance (Uçarlı 2020; Zhang et al. 2021).

Drought is a major environmental problem which constrain growth of cotton seeds and fiber quality while promoting soil degradation and erosion (Xiao et al. 2020). Plants employ avoidance, tolerance, escape, and recovery strategies to withstand drought, involving mechanisms such as leaf rolling, stomatal closure, osmotic setting, and metabolic stabilization (Kneebone et al. 1992; Uçarlı 2025). Root system architecture is a key determinant of water uptake and signaling under deficit conditions (Zhang et al. 2024).

At the molecular level, plant adaptation to water scarcity and salinity are governed by complex networks of transcription factors (TFs) that bind cis-regulatory elements to control gene expression (Latchman, 1997; Wärnmark et al. 2003). These TFs contain distinct DNA-binding and activation domains that regulate transcriptional activity and integrate environmental cues into adaptive responses (Nakano et al. 2006; Hrmova et al. 2021; Dhattewal et al. 2024). In adverse condition regulation and secondary metabolism, key TF families are bZIP, MYB, WRKY, NAC, AP2/ERF, and MADS-box (Dubos et al. 2010; Rushton et al. 2010; Puranik et al. 2012; Licausi et al. 2013; Yu et al. 2020; Chowdhary et al. 2023; Gao & Dubos, 2024).

The *Lesion Simulating Disease* (*LSD*) gene encodes a C2C2-type zinc finger protein that functions as a central regulator of programmed cell death (PCD) and stress acclimation (Dietrich et al. 1997; Takatsuji, 1998; Coll et al. 2011). In *Arabidopsis thaliana*, *LSD* negatively regulates PCD, with *lsd1* mutants exhibiting runaway cell death and hypersensitivity to oxidative and pathogen-induced stresses due to disrupted ROS homeostasis (Bernacki et al. 2019). Functioning within the *LSD-EDS1-PAD4* regulatory hub, *LSD* integrates redox and hormonal signaling to coordinate defense activation and stress adaptation under environmental extremes (Mühlenbock et al. 2008; Wituszyńska et al. 2013). However, its molecular functions in cotton remain largely unexplored, and elucidating the evolution and expression of *LSD1* homologs in *Gossypium* could advance understanding of stress tolerance and facilitate breeding of resilient cultivars (Dietrich et al. 1997; Bernacki et al. 2019).

In this study, we aimed to characterize and identify the structural features and conserved domains of *GhLSDs* through in silico analysis; also analyze the expression profile of the *LSD* gene under arid conditions and salinity in upland cotton (*Gossypium hirsutum*). Overall, this study provides new insights into the structural and functional attributes of *GhLSDs* and its potential role in cotton's adaptive response to environmental stresses including salinity and drought.

Materials and Methods

In Silico Analysis of LSD Genes and Proteins in Cotton (*Gossypium hirsutum*)

Members of the LSD (“Lesion Simulating Disease 1”) gene family and their corresponding protein sequences were identified in the *Gossypium hirsutum* genome using the CottonGen database (<https://www.cottongen.org/>), NCBI (<https://www.ncbi.nlm.nih.gov/>), and the Plant Transcription Factor Database (PlantTFDB v4.0 and v5.0; <https://planttfdb.gao-lab.org/>). LSD protein sequences from *A. thaliana*, *G. arboreum*, *G. raimondii*, and *G. hirsutum* were retrieved and confirmed via BLASTP and BLASTN searches (<https://blast.ncbi.nlm.nih.gov/Blast.cgi>). Gene characteristics, including nucleotide and protein length, chromosomal location, and exon-intron organization, were obtained from NCBI. The physicochemical properties of GhLSD proteins—molecular weight, isoelectric point (pI), hydropathicity (GRAVY), and sequence length were computed using ProtParam (<https://web.expasy.org/protparam/>), while subcellular localization was estimated through DeepLoc-2.0 (<https://services.healthtech.dtu.dk/services/DeepLoc-2.0/>). Conserved domains were validated with CDD (<https://www.ncbi.nlm.nih.gov/Structure/cdd/cdd.shtml>) and SMART (<http://smart.embl.de/>), and secondary structure components (α -helix, β -turn, extended strand, random coil) were analyzed with PSIPRED (<https://bioinf.cs.ucl.ac.uk/psipred/>). Tertiary structures of GhLSD proteins were modeled using AlphaFold3 (Abramson et al. 2024). The structural organization of the genes was illustrated via GSDS 2.0 (<http://gsds.cbi.pku.edu.cn/>), and conserved motifs were identified using MEME (<https://meme-suite.org/tools/meme>) with maximum of 10 motifs was allowed, with motif lengths ranging from 6 to 50 amino acids. Cis-acting regulatory elements within 2 kb upstream promoter regions were extracted from NCBI and analyzed using the PlantCARE database (<http://bioinformatics.psb.ugent.be/webtools/plantcare/html/>). Multiple sequence alignments were performed with ClustalW, and a phylogenetic tree was constructed using the neighbor-joining (NJ) method with 1,000 bootstrap replicates in MEGA-X (<https://www.megasoftware.net/>), then visualized and annotated with iTOL (<https://itol.embl.de/>).

Plant Materials, Growth Conditions, and Treatments

Cotton (*Gossypium hirsutum* cv. Harem 2) seeds were obtained from the Republic of Türkiye Ministry of Agriculture and Forestry, Eastern Mediterranean Agricultural Research Institute. Seeds were surface-sterilized with 70% ethanol for 1 min, followed by 10% H₂O₂ for 2 h, and rinsed three times with sterile distilled water. The seed coats were carefully removed, and embryos were transferred onto Murashige and Skoog (MS) medium (Murashige & Skoog, 1962). Germination was carried out at 24°C under controlled growth chamber conditions for 3–4 days.

For salt stress, 10-day-old seedlings were transferred to MS medium supplemented with 200 mM NaCl and sampled after 0, 2, 4, 8, 24, and 48 h. For drought stress, seedlings were transferred to glass jars containing MS medium infused with 25% polyethylene glycol (PEG6000) following the PEG overlay method described by (Pawar & Veena, 2020), and exposed to stress for the same time intervals.

Leaf and root tissues were collected at each time point, immediately frozen in liquid nitrogen, and stored at – 80°C until RNA extraction and gene expression analysis.

Total RNA Isolation and cDNA Synthesis

Total RNA was isolated from the 10 day-old leaf and root samples using the FastPure Universal Plant Total RNA isolation kit (Vazyme- Nanjing, China), which is suitable for plant species rich in polysaccharides or polyphenols.

The concentration and purity of total RNA were measured using a NanoDrop 2000 spectrophotometer (Thermo Fisher Scientific, USA). The A260/A280 ratio between 1.8 and 2.1 indicated acceptable purity, while an A260/A230 ratio above 1.8 confirmed the absence of major contaminants. RNA integrity was verified by 1% agarose gel electrophoresis, where distinct 28S and 18S rRNA bands indicated high-quality, intact RNA suitable for downstream applications.

First-strand complementary DNA (cDNA) was synthesized from 1 µg of total RNA using the qScript cDNA Synthesis Kit (Quanta Biosciences, Germany) according to the manufacturer's protocol. The synthesized cDNA was diluted 1:10 with nuclease-free water and stored at – 20°C until quantitative PCR (qRT-PCR) analysis.

Quantitative Real-Time PCR (qRT-PCR) Analysis

Gene-specific primers for six candidate *GhLSD* genes (*GhLSD1*–*GhLSD6*) were designed using the Primer3 software (<https://primer3.ut.ee/>). The *ubq7* gene was used as an internal reference for normalization (Wang et al., 2013). qRT-PCR reactions were performed in 20 µL volumes containing 1 µL of cDNA (100 ng), 0.5 µL each of forward and reverse primers (10 µM), 10 µL of 2× SYBR Green Master Mix, and nuclease-free water to volume. Amplification was conducted on an A.B.T. (Atlas Biotechnology) real-time PCR system using the following thermal profile: 95°C for 5 min, followed by 40 cycles of 95°C for 15 s, 57°C for 20 s, and 72°C for 20 s, with a final melting curve analysis to verify amplification specificity. Each reaction included three biological replicates and two technical replicates. Relative gene expression was assessed following the $2^{-\Delta\Delta CT}$ method (Livak and Schmittgen, 2001).

Statistical Analyses

The qRT-PCR data were analyzed using one-way ANOVA with the least significant difference Tukey HSD test via the SPSS21 statistical program (IBM). All analyses were performed using three independent biological replicates, each analyzed with two technical replicates.

Results

Identification of GhLSD Family Members

Eleven LSD genes were identified in the tetraploid cotton species *G. hirsutum* genome, and these genes were distributed across the chromosomes of *G. hirsutum* ($2n = 4x = 52$). Eleven *GhLSD* genes were located on 5 out of the 13 chromosomes (chromosomes 5, 6, 8, 10, and 12) in both the A and D subgenomes

(Table 1). In *GhLSD* genes, the coding sequence (CDS) length varied from 294 bp to 513 bp, while the number of exons ranged from four to eight. Notably, a correlation between exon number and CDS length was observed. Among the *GhLSD* genes, five are oriented on the positive strand, whereas six are positioned on the negative strand.

Table 1
Summary of the *GhLSD* genes identified in the cotton (*Gossypium hirsutum* L.) genome.

Gene Name	Accession No	Chromosome	Gene Location (5'-3')	Strand	CDS (bp)	Exon No
GhLSD1	XM_016842666.2	A05	NC_053428.1_16065629–16067143	Plus	294	4
GhLSD2	XM_016896396.2	A06	NC_053429.1_128108232–128112426	Minus	510	8
GhLSD3	XM_016848346.2	A08	NC_053431.1_112468123–112470334	Minus	444	7
GhLSD4	XM_016855484.2	A10	NC_053433.1_426067–430393	Plus	513	8
GhLSD5	XM_041083378.1	A12	NC_053435.1_102589789–102592455	Minus	441	7
GhLSD6	XM_041093586.1	D05	NC_053441.1_9470745–9476686	Plus	504	8
GhLSD7	XM_016832533.2	D05	NC_053441.1_14672500–14674115	Plus	294	4
GhLSD8	XM_041096384.1	D06	NC_053442.1_66551319–66555576	Minus	543	8
GhLSD9	XM_016865838	D08	NC_053444.1_56483411–56485604	Minus	444	6
GhLSD10	XM_016862004.2	D10	NC_053446.1_415235–419837	Plus	513	8
GhLSD11	XM_016881184	D12	NC_053448.1_56748242–56750893	Minus	441	6

Chromosomal localization and Gene structure

Upland cotton *Gossypium hirsutum* is allotetraploid ($2n = 4X = 52$) with subgenomes A and D. All of the A chromosomes are relatively larger than homoeologous D chromosomes. Eleven *GhLSD* genes were found to located on terminal parts of ten chromosomes among 26 chromosomes of *G. hirsutum* (Fig. 1). While 5 *GhLSD* genes were located on chromosomes A (A5, A6, A8, A10, A12), 6 *GhLSD* genes were located on D chromosomes (D5, D6, D8, D10, D12) Among these chromosomes, ChrD5 harbored two *GhLSD* genes (GhLSD6 and GhLSD7).

In order to determine the differences in gene structure, CDS and intron structures of 11 *GhLSD* genes were compared and illustrated. The intron count of *GhLSD* genes varied from 3 to 7 (Fig. 2).

Characterization of cotton LSD1 proteins

The characterization of cotton LSD transcription factors was performed using different bioinformatics tools. The findings obtained are shown in Table 3.2. It was found that the length of the proteins ranged from 97 to 180 amino acids. GHLS1 is the lightest LSD protein with a weight of 10.6 kDa, while the heaviest is GhLSD8 with a weight of 18.7 kDa. It was found that the theoretical isoelectric points (pI) are above 7, and GhLSD proteins have a basic structure. According to GRAVY values, GhLSD1 and GhLSD7 proteins were found to have hydrophilic properties, while other GhLSD proteins were found to have hydrophobic structure (GRAVY > 0). The aliphatic index (AI) values of cotton LSD proteins range from 68.25 (GhLS1) to 82.28 (GhLSD8), which indicates that LSD proteins have a thermostable structure. Subcellular localization analysis predicted that GhLSD proteins are localized in both the nucleus and the cytoplasm. According to the Instability Index (II) analysis results, it was determined that LSD proteins were generally stable in structure, while GhLSD1, GhLSD6 and GhLSD7 proteins (II values > 40) were determined to be unstable in structure (Table 2).

Table 2
Characterization of Cotton LSD1 proteins

Protein Name	Length (aa*)	MW* (Da*)	pI*	GRAVY*	AI*	II*	Charge	KSP*	Subcellular Localization
GhLSD1	97	10608	7.53	-0.161	68.25	61.94	0.7	11	Cyt*, Nuc*
GhLSD2	169	17665	9.01	0.192	80.12	35.68	8.3	20	Cyt, Nuc
GhLSD3	147	15158	8.21	0.287	74.35	38.66	2.3	12	Cyt, Nuc
GhLSD4	170	17929	8.77	0.108	76.24	32.06	6.3	15	Cyt, Nuc
GhLSD5	146	15074	8.21	0.359	79.52	39.87	2.3	10	Cyt, Nuc
GhLSD6	167	17666	8.62	0.141	75.75	41.21	5.3	15	Cyt, Nuc
GhLSD7	97	10618	7.53	-0.169	68.25	64.48	0.7	10	Cyt, Nuc
GhLSD8	180	18776	8.86	0.262	82.28	38.09	7.4	20	Cyt, Nuc
GhLSD9	147	15095	8.21	0.282	75.71	38.08	2.3	12	Cyt, Nuc
GhLSD10	170	178881	8.77	0.116	77.94	31.62	6.3	15	Cyt, Nuc
GhLSD11	146	15074	8.21	0.359	79.52	39.87	2.3	10	Cyt, Nuc

*Molecular Weight (MW), Dalton (Da), Kinase Specific Phosphorylation (KSP) regions, Isoelectric point (pI), Grand Average f Hydropathy (GRAVY), Aliphatic Index (AI), Instability Index (II) and the protein charge was estimated according to the amino acid composition. Cyt (Cytoplasm), Nuc (Nucleus).

Domain and motif analysis

Examination of protein domain structures showed that GhLSD proteins contain two or three conserved LSD-type Zinc Finger (ZF) domains which is defined by CxxCxxRxxLMYxxGASxVxCxxC. While nine of eleven GhLSD proteins contain three LSD-type ZF domains, only GhLSD1 and GhLSD7 proteins contain two LSD-type ZF domains (Fig. 3). The secondary structure predicted for the GhLSD proteins was displayed above the multiple sequence alignments, revealing the presence of one or two α -helices in GhLSD3, GhLSD5, GhLSD8, GhLSD9, and GhLSD11, as well as a variable number of β -strands across all GhLSD proteins. In upland cotton, the conservation analysis of LSD-type ZF domains demonstrated strong sequence conservation across all GhLSD proteins. In particular, cysteine residues at positions 1, 4, 19, and 22 were highly conserved, while leucine (position 8), tyrosine (position 10), and glycine (position 11) were also consistently conserved (Fig. 4).

To explore the diversity of GhLSD proteins, conserved motif analysis was conducted using the MEME online server. Ten distinct motifs (motif 1–10) were identified. Among them, motif 1 and motif 2, corresponding to LSD-type ZF domains, were present in nearly all GhLSD genes, highlighting their conserved functional significance within the gene family (Fig. 5).

Cis-Regulatory Element Analysis

The promoter regions (2 kb upstream) of GhLSD genes were screened for cis-acting regulatory elements using the PlantCARE database. The results revealed that *GhLSD* promoters harbor diverse cis-elements associated with plant growth and development, hormone signaling, stress responses, and light regulation at the transcriptional level (Fig. 6). Development-related motifs included AC-II (in *GhLSD4* and *GhLSD10*), Circadian (in *GhLSD8*), the GCN4 motif (in *GhLSD1*, *GhLSD2*, *GhLSD3*, and *GhLSD8*), and RY elements (in *GhLSD6* and *GhLSD7*). Among hormone-responsive elements, the ERE motif (ethylene-responsive) was notably enriched, appearing four times in the *GhLSD4* promoter, whereas the O₂-site (involved in zein metabolism) was detected only in *GhLSD1*. Numerous stress-responsive elements were identified, including MYB binding sites (present in all *GhLSD* genes except *GhLSD11*) and MYC binding sites, which were detected in varying copy numbers across all members. W-box motifs, recognized by WRKY transcription factors, were restricted to *GhLSD2*, *GhLSD5*, *GhLSD8*, *GhLSD10*, and *GhLSD11*, while the WUN motif (associated with wound responses) was found only in *GhLSD1*, *GhLSD2*, and *GhLSD8*. Light-responsive elements were also abundant, with Box 4 and G-box being the most prominent (Fig. 6).

Multiple Sequence Alignment and Phylogenetic Analysis

To investigate the evolutionary relationships and identify paralogous gene pairs among the eleven *Gossypium hirsutum* GhLSD genes, multiple sequence alignment of their CDS sequences was performed using CLUSTALW. Pairwise sequence identity analysis of the eleven GhLSD genes revealed five nearly identical paralogous pairs (*GhLSD1*–*GhLSD7*, *GhLSD2*–*GhLSD8*, *GhLSD3*–*GhLSD9*, *GhLSD4*–*GhLSD10*, and *GhLSD5*–*GhLSD11*) with > 97% identity, indicating recent duplication events. Other gene pairs exhibited

moderate (80–85%) or low (37–46%) identity, reflecting more divergent evolutionary relationships within the gene family.

Phylogenetic analysis revealed that the eleven LSD transcription factors of upland cotton (*Gossypium hirsutum*) could be classified into three groups (Fig. 7). Group I comprised GhLSD1, GhLSD3, GhLSD5, GhLSD7, GhLSD9, and GhLSD11; Group II included GhLSD2 and GhLSD8; while Group III consisted of GhLSD4, GhLSD6, and GhLSD10. To further elucidate the evolutionary history of the LSD gene family in upland cotton, multiple sequence alignments were performed using ClustalW with LSD proteins from *G. hirsutum* (11 GhLSDs), *Arabidopsis thaliana* (6 AtLSDs), *G. raimondii* (8 GrLSDs), and *G. arboreum* (6 GaLSDs). Phylogenetic analysis of LSD proteins from *Gossypium spp* and *A. thaliana* revealed that the LSD gene family clustered into five distinct clades (Clade I–V) (Fig. 7). Most clades contained members from at three species, while Clade II contained only LSD proteins of *A. thaliana*. In most clades, GhLSDs were grouped closely with homologs from *G. raimondii* and *G. arboreum*, suggesting inheritance from both A- and D-genome progenitors. All GhLSD members of Group I were clustered within Clade V. Within Group III, GhLSD4 and GhLSD10 were clustered in Clade III, while the remaining member GhLSD6 were positioned in Clade I with homologs from *G. raimondii* and *G. Arboreum* (Fig. 7 zein metabolism).

Protein Structure Analysis

AlphaFold3 was used to predict the tertiary structures 11 GhLSD proteins. In all GhLSD proteins, random coil accounted for the largest proportion ranged from 58.50% to 78.35, followed by extended strand (22.65–36.69%). α -Helix structures which are detected on GhLSD3, GhLSD5, GhLSD8, GhLSD9, and GhLSD11 at the low proportions (2.23–6.80%). (Table 3). No Beta turn was found in GhLSD proteins. GhLSD proteins have a secondary structure largely composed of random coils and β -strands, with limited α -helices and no β -turns. These properties, along with their flexible, stable structures, suggest that they are proteins that confer stability (Fig. 8).

Table 3
The secondary structures of GhLSD proteins.

Protein	α -Helix (%)	Extended Strand (%)	Beta Turn (%)	Random Coil (%)
GhLSD1	0	22.68	0	77.32
GhLSD2	0	36.69	0	63.31
GhLSD3	5.44	34.01	0	60.54
GhLSD4	0	31.18	0	68.82
GhLSD5	6.16	33.56	0	60.28
GhLSD6	0	29.94	0	70.06
GhLSD7	0	21.65	0	78.35
GhLSD8	2.23	33.52	0	64.25
GhLSD9	6.80	34.70	0	58.50
GhLSD10	0	31.18	0	68.82
GhLSD11	6.16	33.56	0	60.27

Gene Expression Pattern Analysis Under Abiotic Stress

The *GhLSD* genes in Leaves and roots under drought and salt stresses in the *Gossypium hirsutum* genotype, the expression patterns of 6 GhLSD genes—including *GhLSD1*, *GhLSD2*, *GhLSD3*, *GhLSD4*, *GhLSD5* and *GhLSD6*, were examined using qRT-PCR analysis.

Gene Expression Analysis of LSD genes Under Drought Stress

The expression profiling of *GhLSD* genes revealed diverse and tissue-specific regulatory responses to drought stress. The expression levels of *GhLSD1* showed significant upregulation, with peak expression at 8 h (2.24-fold in leaf, while the expression was decreased to 4.76-fold in leaf tissues at 2h. *GhLSD2* was moderately upregulated in leaves at 8 h (1.48-fold) and 24 h (1.49-fold), but showed only mild induction at 4 h (0.89-fold) and 48 h (1.17-fold). In contrast, it was moderately downregulated in 5leaves at 2 h (1.72-fold). *GhLSD3* exhibited strong upregulation (6.30-fold) in leaves at 8 h under drought stress, followed by moderate induction (2.39-fold) at 24 h. GhLSD4 was markedly upregulated (154.04-fold) in leaves at 8 h of drought exposure. GhLSD5 displayed strong upregulation (3.34-fold) in leaves at 8 h, but was slightly downregulated (1.53-fold) at 4 h under drought stress. GhLSD6 was highly upregulated (4.12-fold and 4.88-fold) in cotton leaves at 8 h and 24h following drought exposure, but displayed only a moderate induction (2.75-fold) under drought stress at 48h respectively (Fig. 9).

The expression levels of GhLSD1 showed significant upregulation, with peak expression at 48 h (40.44-fold) in root, while the expression was decreased to 13.33-fold in root tissues at 2h. GhLSD2 was moderately downregulated in roots at 24 h (1.36-fold). Stronger downregulation was observed in roots at 8

h (2-fold) and 48 h (2.38-fold). In roots, *GhLSD3* was strongly downregulated (9.09-fold and 10-fold) at 2 h and 24 h, respectively. *GhLSD5* was strongly upregulated (27.5-fold) following 48h of drought exposure in roots. In roots, *GhLSD5* showed strong upregulation (27.5-fold) at 48 h, whereas it was highly downregulated (10-fold) at 4 h. *GhLSD6* expression in drought-stressed roots peaked at 48 h, reaching a 19-fold upregulation compared to the control. *GhLSD6* was moderately downregulated (2.77-fold) in roots at 4h under drought stress (Fig. 9).

Gene Expression Analysis of LSD genes Under Salt Stress

The expression analysis of *GhLSD* genes under salt stress revealed distinct temporal and tissue-specific patterns. *GhLSD1* showed moderate upregulation in leaves at 8 h (1.38-fold), but was moderately downregulated at early time points (2 h, 0.44-fold; 4 h, 0.63-fold). *GhLSD2* exhibited strong time-dependent downregulation in leaves (2 h, 0.61-fold; 4 h, 0.51-fold; 24 h, 0.31-fold; 48 h, 0.19-fold). *GhLSD3* displayed dynamic regulation, with strong upregulation in leaves at early time points 2h and 8h but marked downregulation at certain time points, reflecting context-dependent regulatory functions. *GhLSD4* was early induced in leaves at 2 h (1.61-fold), peaking at 4 h (6.2-fold). *GhLSD5* exhibited moderate upregulation in leaves at 2 h (2.42-fold) and 8 h (1.93-fold). Moderate inductions (2.42-fold and 1.93-fold) of *GhLSD5* was observed leaves under salt stress at 2h and 8h respectively. *GhLSD6* showed slight to moderate downregulation in leaves (2 h, 0.61-fold; 4 h, 0.45-fold; 24 h, 0.85-fold; 48 h, 0.56-fold) (Fig. 10).

GhLSD1 was moderately downregulated at early time points (2 h, 0.44-fold; 4 h, 0.63-fold), while in roots it peaked at 24 h (5.68-fold) and was strongly downregulated at 2 h (0.17-fold). *GhLSD2* was strongly downregulated in roots at 2 h (0.20-fold) and 48 h (0.17-fold). *GhLSD3* displayed dynamic regulation, with strong upregulation in roots at later stages 24h and 48h, but marked downregulation at certain time points, reflecting context-dependent regulatory functions. *GhLSD3* was strongly downregulated (0.16-fold) (6.25) and (0.31-fold) (3.22) in roots at 2h and 8h under salt stress. *GhLSD3* was highly downregulated (0.05-fold) (18.18) in roots at 4h under salt stress. In roots, *GhLSD4* showed moderate downregulation at 2 h (0.46-fold) followed by strong upregulation at 4 h (7.39-fold). *GhLSD5* exhibited dramatic upregulation in roots at 24 h (131.20-fold) and 48 h (216.88-fold). *GhLSD5* was strongly upregulated (131.20-fold and 216.88-fold) in roots at 24h and 48h under salt stress. *GhLSD5* showed extreme upregulation in roots at later stages (24–48 h), suggesting a critical role in root-specific salt stress adaptation. *GhLSD6* was strongly downregulated at 8 h (0.62-fold), and highly upregulated at 48 h (5.47-fold). *GhLSD6* was highly upregulated peak in roots at 48 h, highlighting its late-stage role in root stress response (Fig. 10).

Overall, these results indicate that the *GhLSD* gene family operates in a coordinated, tissue- and time-dependent manner under salt stress, with *GhLSD1*, *GhLSD4*, *GhLSD5*, and *GhLSD6* emerging as key positive regulators, whereas *GhLSD2* may act as a repressor (Fig. 10).

Discussion

Gossypium hirsutum (Upland cotton) is an allotetraploid species ($2n = 4x = 52$) composed of two subgenomes, A and D, which originated from the hybridization of *G. arboreum* (A genome) and *G. raimondii* (D genome) (Chen et al. 2020). Its genome size is estimated at 2.25–2.43 Gb and encodes approximately

76,943 genes. The Lesion Simulating Disease (LSD) transcription factor acts as a negative regulator of programmed cell death (PCD) by inhibiting EDS1 and PAD4, and contributes to plant responses to salinity and drought stress (Dietrich et al. 1994). LSD proteins contain a conserved zinc-finger (ZF) motif (CxxCRxxLMYxxGASxVxCxxC) (Dietrich et al. 1997), enabling DNA/protein interactions and integration of ROS, salicylic acid (SA), and ethylene (ET) signaling to modulate cell death and acclimation.

Genome-wide identification and functional characterization of LSD transcription factors have been reported in several model species, including *Arabidopsis thaliana*, *Oryza sativa*, and *Zea mays* (Jiang et al. 2019; Sun et al. 2023). In this study, we identified eleven *GhLSD* genes in the *G. hirsutum* genome and found to be unevenly distributed across 5 out of the 13 chromosomes (chromosomes 5, 6, 8, 10, and 12) in both the A and D subgenomes. The number of LSD genes vary among species. For example, 6 *MsLSD* in *M. Sativa*, 12 *AtLSD* in *A. Thaliana*, 20 *ZmLSD* in *Zea mays* and 18 *GmLSD* in *Glycine max*.

Comprehensive characterization of the *Gossypium hirsutum* LSD (*GhLSD*) gene family provides important insights into their genomic organization, structural variation, and potential functional roles in cotton. The CDS lengths of *GhLSD* genes varied from 294 to 513 bp, and exon numbers ranged from four to eight, suggesting moderate structural diversity within the family. The observed correlation between CDS length and exon number indicates evolutionary constraints that maintain functional domains while allowing flexibility in gene structure. Similar structural variation has been reported in *Arabidopsis thaliana* and soybean (4 to 6 exon and supporting the idea that exon–intron organization contributes to gene diversification and functional specialization of LSD family members (Cabreira et al. 2013). On the other hand, most alpha alpha *MsLSD* genes contain a limited number of exons, one or four. (Sun et al. 2023),

The chromosomal distribution of *GhLSD* genes revealed localization on ten of 26 chromosomes, with an almost balanced presence between the A and D subgenomes. The occurrence of duplicated gene pairs on homoeologous chromosomes, such as *GhLSD1* on chromosome A5 and its paralog *GhLSD7* on chromosome D5, together with the close localization of *GhLSD6* and *GhLSD7* on chromosome D5, indicates that both segmental and tandem duplication events have played a pivotal role in the expansion of the *GhLSD* gene family. The allotetraploid nature of upland cotton further explains the presence of multiple paralogous gene pairs with high sequence identity (> 97%), which likely originated from recent duplication events. Comparable expansion patterns were observed in *Medicago sativa* (6 *MsLSD* genes) and *Glycine max* (18 LSD genes) indicating that polyploidy and duplication events are common mechanisms underlying the expansion of LSD transcription factor families in plants (Cabreira et al. 2013).

At the protein level, *GhLSD* members exhibited lengths of 97–180 amino acids, with predicted molecular weights ranging from 10.6 to 18.7 kDa. Most proteins had basic isoelectric points, consistent with their proposed roles in DNA or protein binding. Variations in hydrophobicity and instability indices suggest that different *GhLSD* proteins may adopt specialized cellular roles, with *GhLSD1*, *GhLSD6*, and *GhLSD7* being less stable and potentially subject to post-translational regulation. Similar physicochemical diversity has been reported in LSD proteins of maize (*Zea mays*), where variation in hydrophilicity was suggested to reflect adaptive subfunctionalization (Jiang et al. 2019). Subcellular localization predictions identified nuclear and cytoplasmic, implying functional diversity in transcriptional regulation and stress signaling.

The nuclear localization is consistent with their role as transcription factors, whereas their cytoplasmic presence suggests potential additional functions, such as acting as scaffold proteins or participating in signaling cascades prior to nuclear translocation. Importantly, consistent with previous findings, (Czarnocka et. al. 2017) reported that LSD1 not only functions as a transcriptional regulator but also acts as a scaffold protein, thereby supporting the notion that LSD proteins exhibit dual roles within the cell. Taken together, these results suggest that GhLSD proteins may function as condition-dependent scaffold proteins in the cytoplasm and as transcriptional regulators in the nucleus, reflecting a versatile role in stress response and developmental regulation in cotton.

Domain analysis confirmed the presence of highly conserved LSD-type zinc finger (ZF) motifs (CxxCRxxLMYxxGASxVxCxxC), also member of a cysteine-2/cysteine-2-class (C2C2), which were maintained across all members, reinforcing their essential role in programmed cell death (PCD) regulation. Nine GhLSD proteins contained three LSD-type zinc finger (ZF) domains, whereas GhLSD1 and GhLSD7 harbored only two, suggesting potential divergence in their DNA-binding affinity or target specificity. Typically, LSD proteins possess three ZF domains; however, the number of domains can vary among different species. For instance, in maize, ZmLSD proteins have been reported to contain between one and three ZF domains, indicating structural diversification within the family across plant lineages (Jiang et al. 2019). Conservation of cysteine residues at positions 1, 4, 19, and 22 across all GhLSD proteins, forming the typical C2C2-type ZF structure, highlights their importance in tertiary structure and structural stability, and protects cell from oxidative damage, redox-mediated regulation as well (Cabreira et al. 2013). In conclusion, GhLSD proteins may play a similar role in managing oxidative stress triggered by drought and salinity.

Motif and cis-regulatory element analyses further supported functional conservation and specialization. The presence of conserved motifs (motif 1 and 2) across almost all GhLSD proteins reflects their core regulatory role, while additional motifs may contribute to lineage-specific adaptation. Promoter analysis revealed abundant stress- and hormone-responsive cis-elements, including MYB, MYC, ERE, and W-box motifs, which are known to regulate transcriptional reprogramming during abiotic stresses. In particular, the enrichment of ethylene-responsive elements (EREs) in GhLSD4 suggests a strong link between LSD proteins and ethylene-mediated signaling pathways, which play central roles in stress adaptation and PCD regulation. These findings are consistent with earlier reports in *Glycine max* and *Medicago sativa*, where LSD genes were shown to integrate reactive oxygen species (ROS), salicylic acid (SA), and ethylene (ET) signaling during stress responses (Cabreira et al. 2013; Sun et al. 2023).

The phylogenetic clustering demonstrates that the LSD gene family is highly conserved across cotton species and *Arabidopsis*, with distinct subgrouping into five clades. The close association of GhLSDs with both *G. arboreum* and *G. raimondii* orthologs indicates that *G. hirsutum* retained most ancestral copies from its diploid progenitors following polyploidization. This conservation suggests that LSD genes play fundamental roles in stress adaptation and development. The absence of *G. hirsutum* members in Clade II likely reflects gene loss, pseudogenization, or significant divergence, highlighting evolutionary flexibility within the family. Furthermore, the distribution of GhLSDs across different clades suggests functional diversification, with some members potentially undergoing subfunctionalization or neofunctionalization to

meet the adaptive demands of upland cotton. These evolutionary patterns are consistent with previous findings that LSD proteins regulate stress signaling and development, implying that different clades may correspond to distinct functional specializations

Protein structural predictions revealed that GhLSD proteins are predominantly composed of random coils and β -strands, with limited α -helices and no β -turns. Such structural flexibility is typical of regulatory proteins that interact with multiple partners within stress signaling networks. The thermostability indicated by aliphatic index values suggests that GhLSD proteins remain stable under variable environmental conditions, a property that has been similarly reported for Alfaalfa and maize LSD proteins (Jiang et al. 2019; Sun et al. 2023).

Taken together, these findings indicate that the LSD transcription factor family in cotton has undergone expansion and diversification through polyploidy and gene duplication while maintaining strong conservation of core domains essential for PCD regulation. The integration of diverse promoter elements, structural variation, and predicted functional divergence highlights the potential of GhLSD genes as key regulators of stress adaptation in cotton. Given the established roles of LSD genes in regulating PCD, ROS signaling, and abiotic stress tolerance in other crops, the GhLSD family represents promising candidates for functional validation studies. Future research focusing on expression profiling under drought and salinity stress, as well as genome-editing approaches such as CRISPR/Cas9, may provide valuable strategies for improving cotton stress tolerance.

The expression profiling of GhLSD genes under drought and salt stresses provides novel insights into their temporal and tissue-specific regulatory functions in *G. hirsutum*. Overall, the results demonstrate that the GhLSD gene family operates in a highly dynamic, coordinated, and context-dependent manner, with distinct members contributing to either activation or repression of stress responses in leaves and roots.

The drought-responsive expression profiles revealed a clear division of regulatory roles among *GhLSD* genes. In leaves, *GhLSD1*, *GhLSD3*, *GhLSD4*, *GhLSD5*, and *GhLSD6* were strongly induced, particularly at 8 h, indicating an early activation of defense and acclimation mechanisms. Notably, *GhLSD4* exhibited a dramatic 154-fold upregulation, positioning it as a potential master regulator of drought-induced signaling pathways in cotton. Similar strong induction of LSD genes under drought stress has been observed in *Arabidopsis* and soybean, where LSD family members were shown to integrate ROS and hormone signals to fine-tune stress responses (Cabreira et al. 2013; Szechyn'ska-Hebda et al. 2016)

In roots, *GhLSD1*, *GhLSD5*, and *GhLSD6* displayed robust late-stage induction (40.44-fold, 27.5-fold, and 19-fold, respectively, at 48 h), suggesting that these genes may contribute to sustained drought adaptation mechanisms, possibly by modulating root growth or enhancing antioxidant defenses. In contrast, *GhLSD2* and *GhLSD3* were consistently downregulated in roots, implying a negative regulatory role, potentially acting to prevent excessive PCD activation under prolonged water deficit. Such contrasting expression patterns between leaves and roots highlight tissue-specific functional diversification, consistent with earlier reports that LSD homologs act as context-dependent regulators balancing survival and cell death in plants

Taken together, *GhLSD1*, *GhLSD4*, and *GhLSD6* appear to function as positive regulators of drought tolerance, while *GhLSD2* may act as a repressor. The strong induction of *GhLSD4* in leaves and *GhLSD6* in roots emphasizes their importance as candidate genes for improving drought resilience in cotton.

Under salt stress, the GhLSD family again exhibited diverse and tissue-dependent regulation. In leaves, *GhLSD4* and *GhLSD5* showed early induction (up to 6.2-fold at 4 h for *GhLSD4*, and 2.42-fold for *GhLSD5* at 2 h), suggesting their involvement in rapid signaling and protective responses against ionic and osmotic stress. *GhLSD3* showed transient upregulation at early time points but was subsequently downregulated, indicating a tightly controlled, context-dependent role. In contrast, *GhLSD2* was consistently downregulated, supporting its putative role as a negative regulator during salt stress.

In roots, striking differences emerged. *GhLSD5* exhibited extremely high induction (131-fold at 24 h and 217-fold at 48 h), suggesting it plays a central role in root-specific salt tolerance, likely through regulation of ion homeostasis, osmotic adjustment, or ROS detoxification. *GhLSD6* also showed late-stage induction (5.47-fold at 48 h), pointing to a supportive role in root acclimation. By contrast, *GhLSD2* remained strongly repressed across time points, further strengthening its proposed role as a negative regulator. Similar patterns of late-stage activation of LSD homologs have been observed in rice, where LSD family members were implicated in long-term adaptive responses to salinity (Sun et al. 2010).

Comparing the two stress conditions reveals both overlapping and divergent roles of GhLSD genes. *GhLSD1*, *GhLSD4*, *GhLSD5*, and *GhLSD6* consistently emerged as positive regulators under both drought and salt stresses, albeit with tissue- and time-dependent differences. *GhLSD4* functioned predominantly in leaves, displaying rapid and strong induction under drought and salt stresses, while *GhLSD5* and *GhLSD6* were highly responsive in roots, particularly under prolonged stress exposure. *GhLSD2*, conversely, was consistently downregulated in both conditions, suggesting it acts as a suppressor of stress responses, possibly to prevent excessive or uncontrolled PCD. The contrasting regulation of *GhLSD3* across tissues and stress conditions suggests a dual, context-dependent role in fine-tuning stress signaling.

These findings highlight that the GhLSD gene family contributes to cotton stress adaptation through a coordinated network of activators and repressors that function in both early and late phases of stress. The strong upregulation of *GhLSD4*, *GhLSD5*, and *GhLSD6* suggests that these genes are promising targets for genetic improvement of drought and salinity tolerance in cotton. Their tissue-specific expression further emphasizes the importance of considering organ-level regulation in stress biology.

Future functional validation using overexpression and CRISPR/Cas9 knockout strategies will be essential to clarify the precise regulatory roles of GhLSD genes and to exploit them for crop improvement. Given the integration of LSD proteins with ROS, SA, and ET signaling, further investigation into their upstream regulation and downstream target genes will provide critical insights into their role in stress resilience mechanisms in cotton.

Conclusion and Recommendation

The comprehensive expression profiling of the *GhLSD* gene family under drought and salinity stress revealed that these transcription factors function in a dynamic, tissue-specific, and time-dependent manner to regulate stress responses in *Gossypium hirsutum*. Among the six genes analyzed, *GhLSD1*, *GhLSD4*, *GhLSD5*, and *GhLSD6* consistently exhibited strong induction under both drought and salt treatments, highlighting their potential roles as positive regulators of abiotic stress tolerance. In particular, *GhLSD4* (in leaves) and *GhLSD5* (in roots) demonstrated exceptionally high expression peaks, suggesting specialized roles in tissue-specific stress adaptation. *GhLSD6* also showed consistent late-stage upregulation, pointing to its involvement in sustained stress responses. Conversely, *GhLSD2* was generally downregulated in both tissues under drought and salinity, suggesting a repressive role in stress signaling pathways. *GhLSD3* displayed variable regulation, alternating between induction and suppression depending on tissue type and exposure duration, indicating a context-dependent function.

Taken together, these findings emphasize that the *GhLSD* family contributes to stress acclimation through a finely tuned regulatory network rather than uniform responses. The strong induction of *GhLSD4*, *GhLSD5*, and *GhLSD6* highlights them as promising candidate genes for further functional studies and potential genetic improvement strategies to enhance drought and salinity tolerance in cotton.

Declarations

Funding: This work was supported by the Scientific Research Projects Coordination Unit of Istanbul University (FYL-2024-40510)

Conflict of interest/Competing interests: The authors declare no conflict of interest.

Ethics approval and consent to participate: Not applicable

Consent for publication: Not applicable

Data availability: Data are included within the article.

Materials availability: Materials are available from the corresponding author upon request.

Code availability: Not applicable

Author contribution: C.U. designed the study, analyzed the data and wrote the manuscript. A. A. performed the experiments and wrote the manuscript. All authors read and approved the final manuscript.

References

1. Abdelraheem A, Esmaeili N, O'Connell M, Zhang J (2019) Progress and perspective on drought and salt stress tolerance in cotton. *Ind Crops Prod* 30:118–129.
<https://doi.org/10.1016/j.indcrop.2018.12.070>
2. Abdelraheem A, Thyssen GN, Fang DD, Jenkins JN, McCarty JC, Wedegaertner T, Zhang J (2021) GWAS reveals consistent QTL for drought and salt tolerance in a MAGIC population of 550 lines

- derived from intermating of 11 upland cotton (*Gossypium hirsutum*) parents. *Mol Genet Genomics* 296:119–129. <https://doi.org/10.1007/s00438-020-01733-2>
3. Abramson J, Adler J, Dunger J, Evans R, Green T, Pritzel A et al (2024) Accurate structure prediction of biomolecular interactions with AlphaFold3. *Nature* 630:493–500. <https://doi.org/10.1038/s41586-024-07487-w>
 4. Ahmad I, Zhu G, Zhou G, Song X, Hussein ME, Ibrahim EG et al (2022) Pivotal role of phytohormones and their responsive genes in plant growth and signaling under salt stress in cotton. *Int J Mol Sci* 23:7339. <https://doi.org/10.3390/ijms23137339>
 5. Bano N, Fakhrah S, Mohanty CS, Bag SK (2022) Transcriptome meta-analysis targeting hub genes and pathways of drought and salt stress responses in cotton (*Gossypium hirsutum*): a network biology approach. *Front Plant Sci* 13:818472. <https://doi.org/10.3389/fpls.2022.818472>
 6. Bernacki MJ, Czarnocka W, Rusaczek A, Witoń D, Kęska S, Czyż J, Szechyńska-Hebda M, Karpiński S (2019) LSD1, EDS1 and PAD4-dependent conditional correlation among salicylic acid, hydrogen peroxide, water use efficiency, and seed yield in *Arabidopsis thaliana*. *Physiol Plant* 165:369–382. <https://doi.org/10.1111/ppl.12863>
 7. Cabreira C, Cagliari A, Bücker-Neto L, Wiebke-Strohm B, de Freitas LB, Marcelino-Guimarães FC et al (2013) The Lesion Simulating Disease (LSD) gene family as a variable in soybean response to *Phakopsora pachyrhizi* infection and dehydration. *Funct Integr Genomics* 13:323–338. <https://doi.org/10.1007/s10142-013-0326-3>
 8. Chaudhary MT, Majeed S, Rana IA et al (2024) Impact of salinity stress on cotton and opportunities for improvement through conventional and biotechnological approaches. *BMC Plant Biol* 24:20. <https://doi.org/10.1186/s12870-023-04558-4>
 9. Chowdhary AA, Mishra S, Mehrotra S, Upadhyay SK, Bagal D, Srivastava V (2023) Plant transcription factors: an overview of their role in plant life. In: *Plant Transcription Factors*. Academic Press, pp 3–20
 10. Chen ZJ, Sreedasyam A, Ando A et al (2020) Genomic diversifications of five *Gossypium* allopolyploid species and their impact on cotton improvement. *Nat Genet* 52:525–533. <https://doi.org/10.1038/s41588-020-0614-5>
 11. Coll NS, Epple P, Dangl JL (2011) Programmed cell death in the plant immune system. *Cell Death Differ* 18:1247–1256. <https://doi.org/10.1038/cdd.2011.37>
 12. Czarnocka W, Van Der Kelen K, Willems P et al (2017) The dual role of LESION SIMULATING DISEASE 1 as a condition-dependent scaffold protein and transcription regulator. *Plant Cell Environ* 40:2644–2662. <https://doi.org/10.1111/pce.12994>
 13. Dhattewal P, Sharma N, Prasad M (2024) Decoding the functionality of plant transcription factors. *J Exp Bot* 75:4745–4759. <https://doi.org/10.1093/jxb/erae231>
 14. Dietrich RA, Delaney TP, Uknes SJ, Ward ER, Ryals JA, Dangl JL (1994) *Arabidopsis* mutants simulating disease resistance response. *Cell* 77:565–577. [https://doi.org/10.1016/0092-8674\(94\)90218-6](https://doi.org/10.1016/0092-8674(94)90218-6)
 15. Dietrich RA et al (1997) A novel zinc finger protein is encoded by the *Arabidopsis* LSD1 gene and functions as a negative regulator of plant cell death. *Cell* 88:685–694. [https://doi.org/10.1016/s0092-8674\(00\)81911-x](https://doi.org/10.1016/s0092-8674(00)81911-x)

16. Dubos C, Stracke R, Grotewold E, Weisshaar B, Martin C, Lepiniec L (2010) MYB transcription factors in *Arabidopsis*. Trends Plant Sci 15:573–581. <https://doi.org/10.1016/j.tplants.2010.06.005>
17. Gao F, Dubos C (2024) The *Arabidopsis* bHLH transcription factor family. Trends Plant Sci 29:668–680. <https://doi.org/10.1016/j.tplants.2023.11.022>
18. Hrmova M, Hussain SS (2021) Plant transcription factors involved in drought and associated stresses. Int J Mol Sci 22:5662. <https://doi.org/10.3390/ijms22115662>
19. Jiang L, Chen D, Yu X, Sun Y, Yu H, Li J, Gao H, Xu H (2019) Bioinformatics analysis of LSD transcription factor in maize and its role in resistance to stem base rot. Intl J Agric Biol 22:1476–1482. <https://doi.org/10.17957/IJAB/15.1225>
20. Khan MA, Anwar S, Abbas M et al (2025) Impacts of climate change on cotton production and genomic approaches for stress resilience enhancement. J Cotton Res 8:17. <https://doi.org/10.1186/s42397-025-00223-3>
21. Kneebone WR, Kopec DM, Mancino CF (1992) Water requirements and irrigation. Turfgrass 32:441–472. <https://doi.org/10.2134/agronmonogr32.c12>
22. Latchman DS (1997) Transcription factors: An overview. Int J Biochem Cell Biol 29:1305–1312. [https://doi.org/10.1016/S1357-2725\(97\)00085-X](https://doi.org/10.1016/S1357-2725(97)00085-X)
23. Li F, Fan G, Wang K et al (2014) Genome sequence of the cultivated cotton *Gossypium arboreum*. Nat Genet 46:567–572. <https://doi.org/10.1038/ng.2987>
24. Liu L, Lindsay PL, Jackson D (2021) Next Generation Cereal Crop Yield Enhancement: From Knowledge of Inflorescence Development to Practical Engineering by Genome Editing. Int J Mol Sci 22(10):5167. <https://doi.org/10.3390/ijms22105167>
25. Livak KJ, Schmittgen TD (2001) Analysis of relative gene expression data using real-time quantitative PCR and the 2 – $\Delta\Delta$ CT method. Methods 25:402–408. <https://doi.org/10.1006/meth.2001.1262>
26. Lu K, Yan L, Riaz M et al (2023) Exogenous boron alleviates salt stress in cotton. Plant Physiol Biochem 201:107858. <https://doi.org/10.1016/j.plaphy.2023.107858>
27. Maas EV, Hoffman GJ (1977) Crop salt tolerance—current assessment. J Irrig Drain Div 103:115–134
28. Murashige T, Skoog F (1962) A revised medium for rapid growth and bio assays with tobacco tissue cultures. Physiol Plant 15:473–497. <https://doi.org/10.1111/j.1399-3054.1962.tb08052.x>
29. Mühlenbock P, Plaszczyc M, Mellerowicz E, Karpinski S (2007) Lysigenous aerenchyma formation in *Arabidopsis* is controlled by LSD1. Plant Cell 19:3819–3830. <https://doi.org/10.1105/tpc.107.052084>
30. Najib DCS, Fei C, Dilanchiev A, Romaric S (2022) Modeling the Impact of Cotton Production on Economic Development in Benin: A Technological Innovation Perspective. Front Environ Sci 10:926350. <https://doi.org/10.3389/fenvs.2022.926350>
31. Nakano T, Suzuki K, Fujimura T, Shinshi H (2006) Genome-wide analysis of the ERF gene family in *Arabidopsis* and rice. Plant Physiol 140:411–432. <https://doi.org/10.1104/pp.105.073783>
32. Pawar KN, Veena VB (2020) Evaluation of Cotton Genotypes for Drought Tolerance using PEG-6000 Water Stress by Slanting Glass Plate Technique. Int J Curr Microbiol Appl Sci 9(11):3203–3212. <https://doi.org/10.20546/ijcmas.2020.911.386>

33. Rushton PJ, Somssich IE, Ringler P, Shen QJ (2010) WRKY transcription factors. Trends Plant Sci 15:247–258. <https://doi.org/10.1016/j.tplants.2010.02.006>
34. Shi Y, Liu A, Li J et al (2019) Dissecting the genetic basis of fiber quality and yield traits in interspecific backcross populations of *Gossypium hirsutum* × *Gossypium barbadense*. Mol Gen Genet 294:1385–1402. <https://doi.org/10.1007/s00438-019-01582-8>
35. Sun SJ, Guo SQ, Yang X, Bao YM, Tang HJ, Sun H, Huang J, Zhang HS (2010) Functional analysis of a novel Cys2/His2-type zinc finger protein involved in salt tolerance in rice. J Exp Bot 61:2807–2818. <https://doi.org/10.1093/jxb/erq120>
36. Sun S, Ma W, Jia Z, Ou C, Li M, Mao P (2023) Genomic Identification and Expression Profiling of Lesion Simulating Disease Genes in Alfalfa (*Medicago sativa*) Elucidate Their Responsiveness to Seed Vigor. Antioxidants 12:1768. <https://doi.org/10.3390/antiox12091768>
37. Szechyńska-Hebda M, Czarnocka W, Hebda M, Bernacki MJ, Karpiński (2016) PAD4, LSD1 and EDS1 regulate drought tolerance, plant biomass production, and cell wall properties. Plant Cell Rep 35:527–539. <https://doi.org/10.1007/s00299-015-1901-y>
38. Takatsuji H (1998) Zinc-finger transcription factors in plants. Cell Mol Life Sci 54:582–596. <https://doi.org/10.1007/s000180050186>
39. Uçarlı C (2020) Effects of Salinity on Seed Germination and Early Seedling Stage. In: Abiotic Stress in Plant. Fahad S, Saud S, Chen Y, Wu C, Wang D (eds). IntechOpen London, UK. <https://doi.org/10.5772/intechopen.93647>
40. Uçarlı C (2025) Drought Stress and the Role of NAC Transcription Factors in Drought Response. In: Chaudhry UK, Öztürk ZN, Gökçe AF (eds) Drought Stress. Springer, Cham. https://doi.org/10.1007/978-3-031-80610-0_11
41. USDA (2025) <https://www.fas.usda.gov/data/commodities/cotton>, Visited date: 18.12.2025
42. Wang M, Wang Q, Zhang B (2013) Evaluation and selection of reliable reference genes for gene expression under abiotic stress in cotton (*Gossypium hirsutum* L). Gene 530:44–50. <https://doi.org/10.1016/j.gene.2013.07.084>
43. Wärnmark A, Treuter E, Wright APH, Gustafsson JÅ (2003) Activation functions 1 and 2 of nuclear receptors: Molecular strategies for transcriptional activation. Mol Endocrinol 17:1901–1909. <https://doi.org/10.1210/me.2002-0384>
44. Wituszynska W, Slesak I, Vanderauwera S et al (2013) Lesion simulating disease1, enhanced disease susceptibility1, and phytoalexin deficient4 conditionally regulate cellular signaling homeostasis, photosynthesis, water use efficiency, and seed yield in Arabidopsis. Plant Physiol 161:1795–1805. <https://doi.org/10.1104/pp.112.208116>
45. Xiao S, Liu L, Zhang Y et al (2020) Fine root and root hair morphology of cotton under drought stress revealed with RhizoPot. J Agron Crop Sci 206:679–693. <https://doi.org/10.1111/jac.12429>
46. Yan R, Liang C, Meng Z, Malik W, Zhu T, Zong X, Guo S, Zhang R (2016) Progress in genome sequencing will accelerate molecular breeding in cotton (*Gossypium spp*). 3 Biotech 6(2):217. <https://doi.org/10.1007/s13205-016-0534-3>

47. Yu Y, Qian Y, Jiang M, Xu J, Yang J, Zhang T, Gou L, Pi E (2020) Regulation Mechanisms of Plant Basic Leucine Zippers to Various Abiotic Stresses. *Front Plant Sci* 11:01258. <https://doi.org/10.3389/fpls.2020.01258>

48. Zhang J, Zhang P, Huo X, Gao Y, Chen Y, Song Z, Wang F, Zhang J (2021) Comparative phenotypic and transcriptomic analysis reveals key responses of upland cotton to salinity stress during postgermination. *Front Plant Sci* 12:639104. <https://doi.org/10.3389/fpls.2021.639104>

49. Zhang Q, Jiang G (2024) The root system architecture of wheat: Core mechanisms of drought resistance and research prospects. *Geogr Res Bull* 3:377–380. https://doi.org/10.50908/grb.3.0_377

50. Zhu YX, Li FG (2013) The *Gossypium raimondii* genome, a huge leap forward in cotton genomics. *J Integr Plant Biol* 55:570–571. <https://doi.org/10.1111/jipb.12076>

Figures

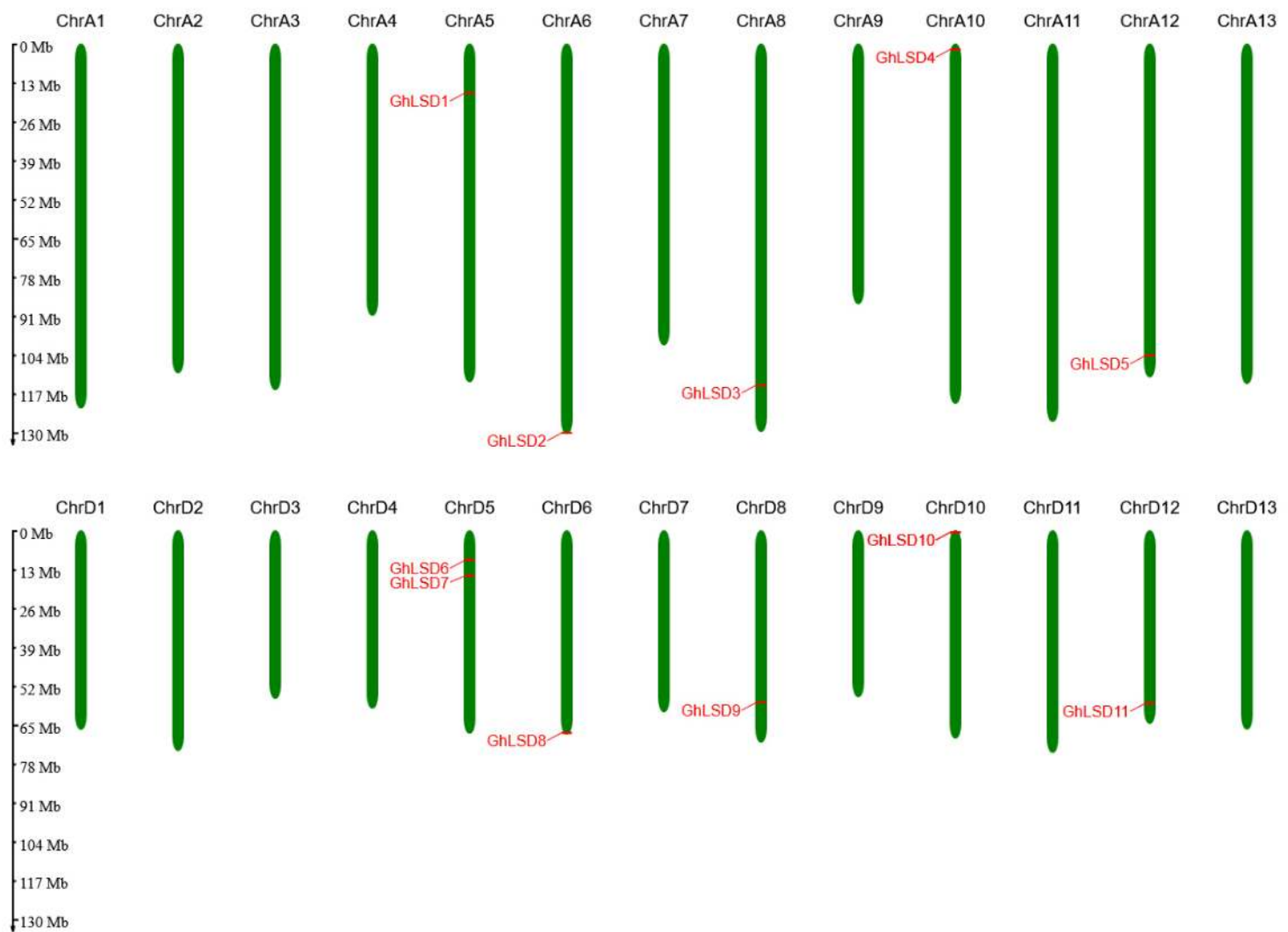


Figure 1

Chromosomal distributions of GhLSD genes on sub-genomes A and D of *G. hirsutum*.

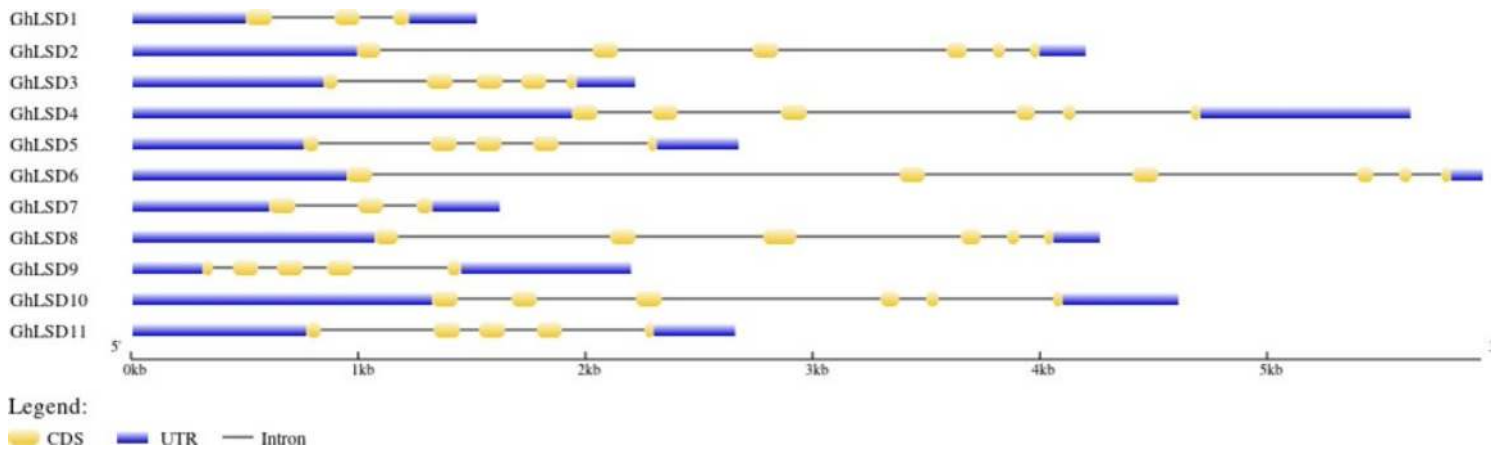


Figure 2

CDS-intron structure analysis of GhLSD genes. The blue boxes show UTRs, the yellow boxes show CDS and the black lines indicate introns.

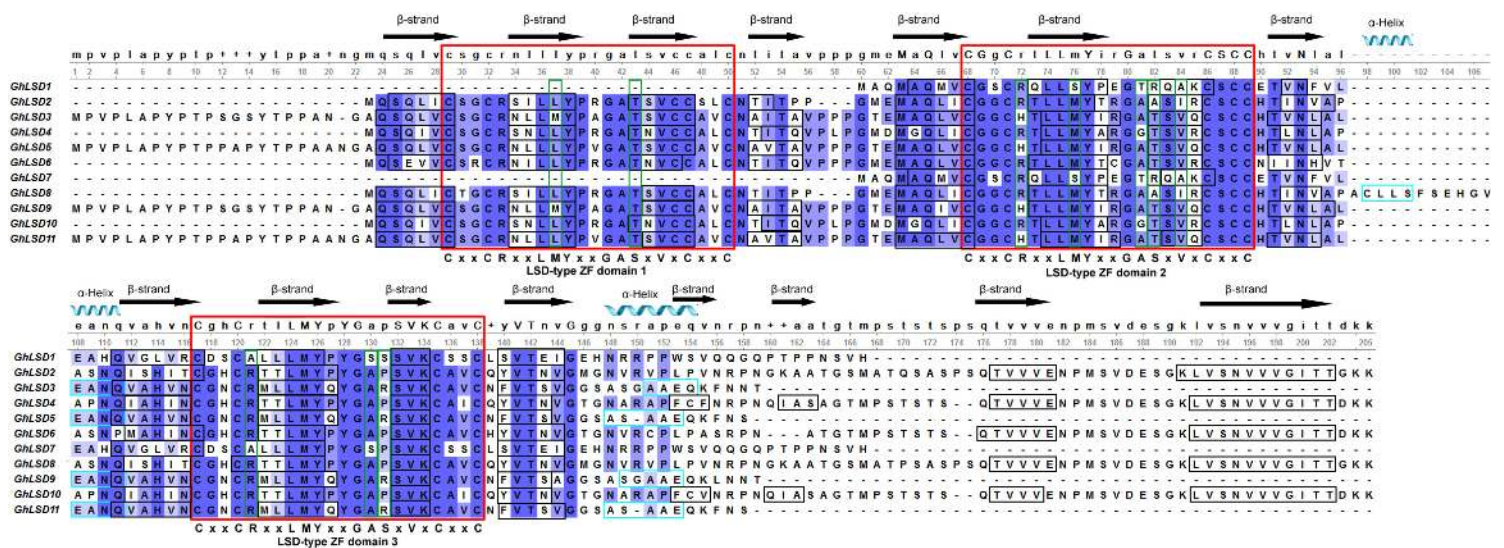


Figure 3

Multiple sequence alignment of the GhLSD proteins. The secondary structural elements, α -helix and β -strand, predicted of GhLSD and LSD-type ZF domains are shown above. Red box, LSD-type ZF domains; blue frame represents similarity across groups. β -strands are shown as black arrows and black frames. α -helices are shown as blue solenoids and blue frames.

LSD-type ZF domain	C x x C R x x L M Y x x G A S x V x C x x C																					
	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50
GhLSD1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
GhLSD2	C	S	G	C	R	S	I	L	L	Y	P	R	G	A	T	S	V	C	C	S	L	C
GhLSD3	C	S	G	C	R	N	L	L	M	Y	P	A	G	A	T	S	V	C	C	A	V	C
GhLSD4	C	S	G	C	R	S	N	L	L	Y	P	R	G	A	T	N	V	C	C	A	L	C
GhLSD5	C	S	G	C	R	N	L	L	L	Y	P	V	G	A	T	S	V	C	C	A	V	C
GhLSD6	C	S	R	C	R	N	I	L	L	Y	P	R	G	A	T	N	V	C	C	A	L	C
GhLSD7	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
GhLSD8	C	T	G	C	R	S	I	L	L	Y	P	R	G	A	T	S	V	C	C	A	L	C
GhLSD9	C	S	G	C	R	N	L	L	M	Y	P	A	G	A	T	S	V	C	C	A	V	C
GhLSD10	C	S	G	C	R	S	N	L	L	Y	P	R	G	A	T	N	V	C	C	A	L	C
GhLSD11	C	S	G	C	R	N	L	L	L	Y	P	V	G	A	T	S	V	C	C	A	V	C

LSD-type ZF domain	C x x C R x x L M Y x x G A S x V x C x x C																					
	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89
GhLSD1	C	G	S	C	R	Q	L	L	S	Y	P	E	G	T	R	Q	A	K	C	S	C	C
GhLSD2	C	G	G	C	R	T	L	L	M	Y	T	R	G	A	A	S	I	R	C	S	C	C
GhLSD3	C	G	G	C	H	T	L	L	M	Y	I	R	G	A	T	S	V	Q	C	S	C	C
GhLSD4	C	G	G	C	R	T	L	L	M	Y	A	R	G	G	T	S	V	R	C	S	C	C
GhLSD5	C	G	G	C	H	T	L	L	M	Y	I	R	G	A	T	S	V	Q	C	S	C	C
GhLSD6	C	G	G	C	R	T	L	L	M	Y	T	C	G	A	T	S	V	R	C	S	C	C
GhLSD7	C	G	S	C	R	Q	L	L	S	Y	P	E	G	T	R	Q	A	K	C	S	C	C
GhLSD8	C	G	G	C	R	T	L	L	M	Y	T	R	G	A	A	S	I	R	C	S	C	C
GhLSD9	C	G	G	C	H	T	L	L	M	Y	I	R	G	A	T	S	V	Q	C	S	C	C
GhLSD10	C	G	G	C	R	T	L	L	M	Y	A	R	G	G	T	S	V	R	C	S	C	C
GhLSD11	C	G	G	C	H	T	L	L	M	Y	I	R	G	A	T	S	V	Q	C	S	C	C

LSD-type ZF domain	C x x C R x x L M Y x x G A S x V x C x x C																					
	117	120	122	124	126	128	130	132	134	136	138											
GhLSD1	C	D	S	C	A	L	L	L	M	Y	P	Y	G	S	S	S	V	K	C	S	S	C
GhLSD2	C	G	H	C	R	T	T	L	M	Y	P	Y	G	A	P	S	V	K	C	A	V	C
GhLSD3	C	G	N	C	R	M	L	L	M	Y	Q	Y	G	A	R	S	V	K	C	A	V	C
GhLSD4	C	G	H	C	R	T	T	L	M	Y	P	Y	G	A	P	S	V	K	C	A	I	C
GhLSD5	C	G	N	C	R	M	L	L	M	Y	Q	Y	G	A	R	S	V	K	C	A	V	C
GhLSD6	C	G	H	C	R	T	T	L	M	Y	P	Y	G	A	P	S	V	K	C	A	V	C
GhLSD7	C	D	S	C	A	L	L	L	M	Y	P	Y	G	S	P	S	V	K	C	S	S	C
GhLSD8	C	G	H	C	R	T	T	L	M	Y	P	Y	G	A	P	S	V	K	C	A	V	C
GhLSD9	C	G	N	C	R	M	L	L	M	Y	Q	Y	G	A	R	S	V	K	C	A	V	C
GhLSD10	C	G	H	C	R	T	T	L	M	Y	P	Y	G	A	P	S	V	K	C	A	I	C
GhLSD11	C	G	N	C	R	M	L	L	M	Y	Q	Y	G	A	R	S	V	K	C	A	V	C

Figure 4

Multiple sequence alignment of LSD-type ZFN domains of GhLSD proteins.

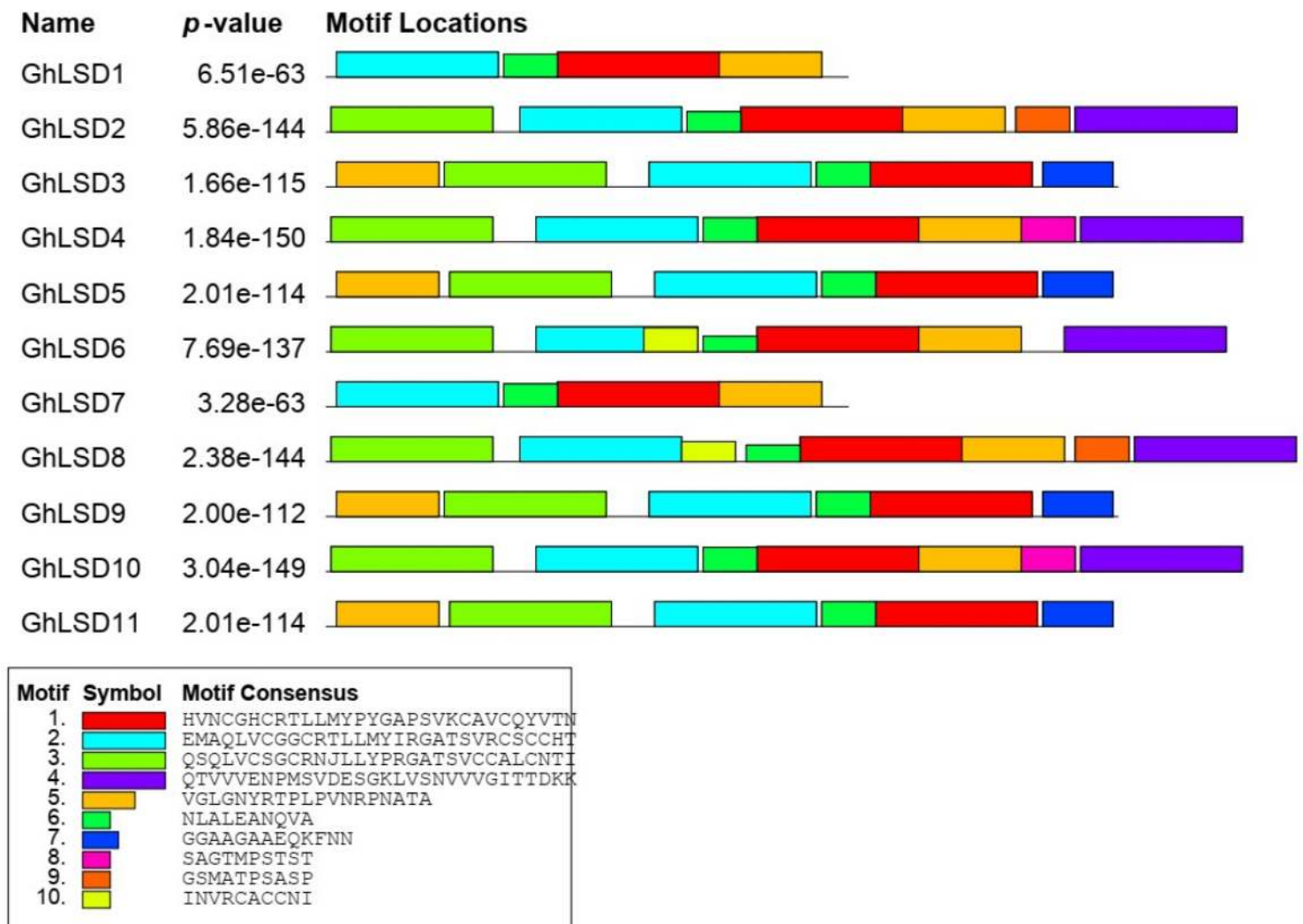
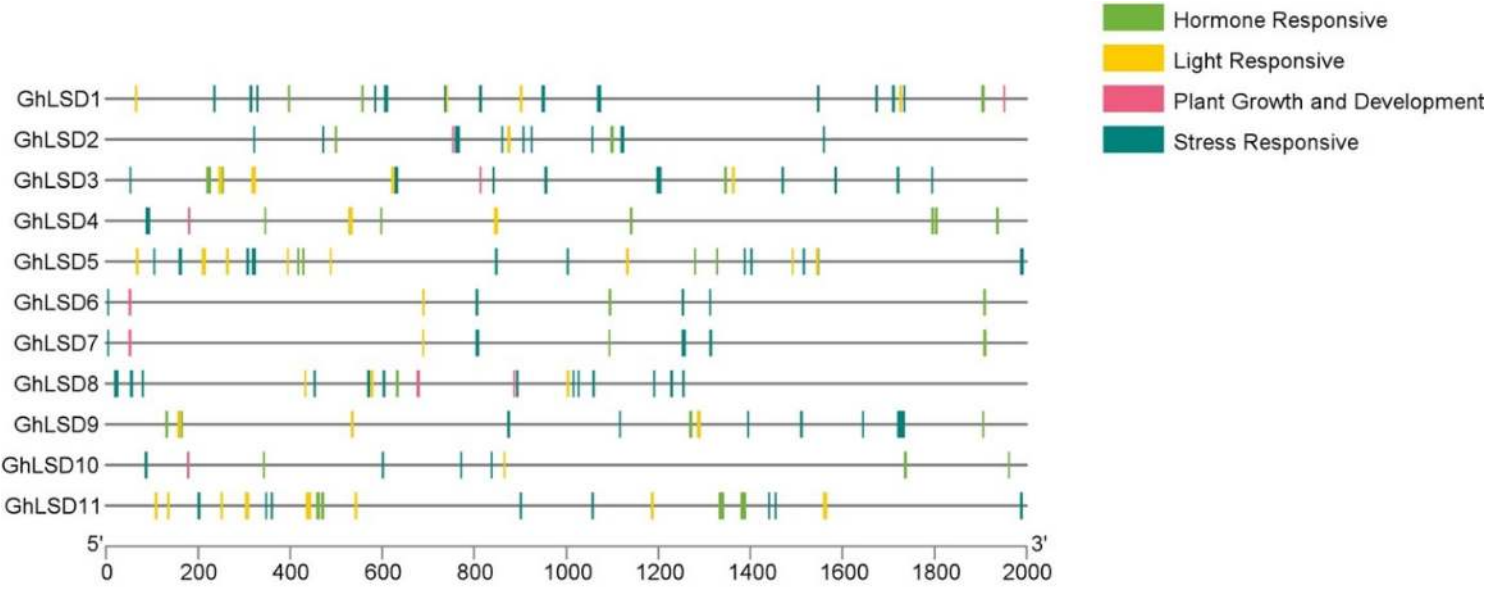


Figure 5

Conserved motifs of LSD proteins in *Gossypium hirsutum*. Motifs 1–10, represented by colored boxes, illustrate the conserved sequence patterns identified within the protein family.



	Plant Growth				Hormone Responsive										Stress responsive										Core														
	AC-II	Circadian	GCN4-motif	RY-element	AAGAA-motif	ABRE	as-1	CGTCA-motif	ERE	GARE-motif	O2-site	P-box	TCA-element	TGA-element	TGACG-motif	ARE	AT-rich element	LTR	MRE	MYB	MYC	STRE	TG-rich repeats	TCT-motif	W-box	WRE3	WWR-motif	Box 4	chs-CMA1a	G-box	GA-motif	GATA-motif	GT1-motif	I-box	TCCC-motif	CAAT-box	TATA-box		
GhLSD1			1		1	1			1		1			1		2				4	3	1	1	1		1			1	1	1				2				
GhLSD2			1					1						1		1				1	4				1		1		1										
GhLSD3			1					1					1	1		2				1	1	4		1				2			1					1			
GhLSD4	1					2			4			2								1	1										2								
GhLSD5						1	2			1			1		2		1	1	1	1	2		2		1			2		1	2	2	1		1				
GhLSD6				1	1								1							1	1	1	1			1				1									
GhLSD7				1	1								1							1	1	1	1							1									
GhLSD8		1	1											1		2				4	4				1	1	1	1						2					
GhLSD9								1		1			1	1		1				1	1	4					1	1		1						1			
GhLSD10	1								2			1								1	1				1	1			1		1								
GhLSD11							2			1			1		2		1		1		3		2		1				2			1	2	2	1	1			

Figure 6

Distribution of cis-acting elements related to stress responses, hormonal regulation, and plant growth in the 2 kb upstream promoter regions of *GhLSD* genes.

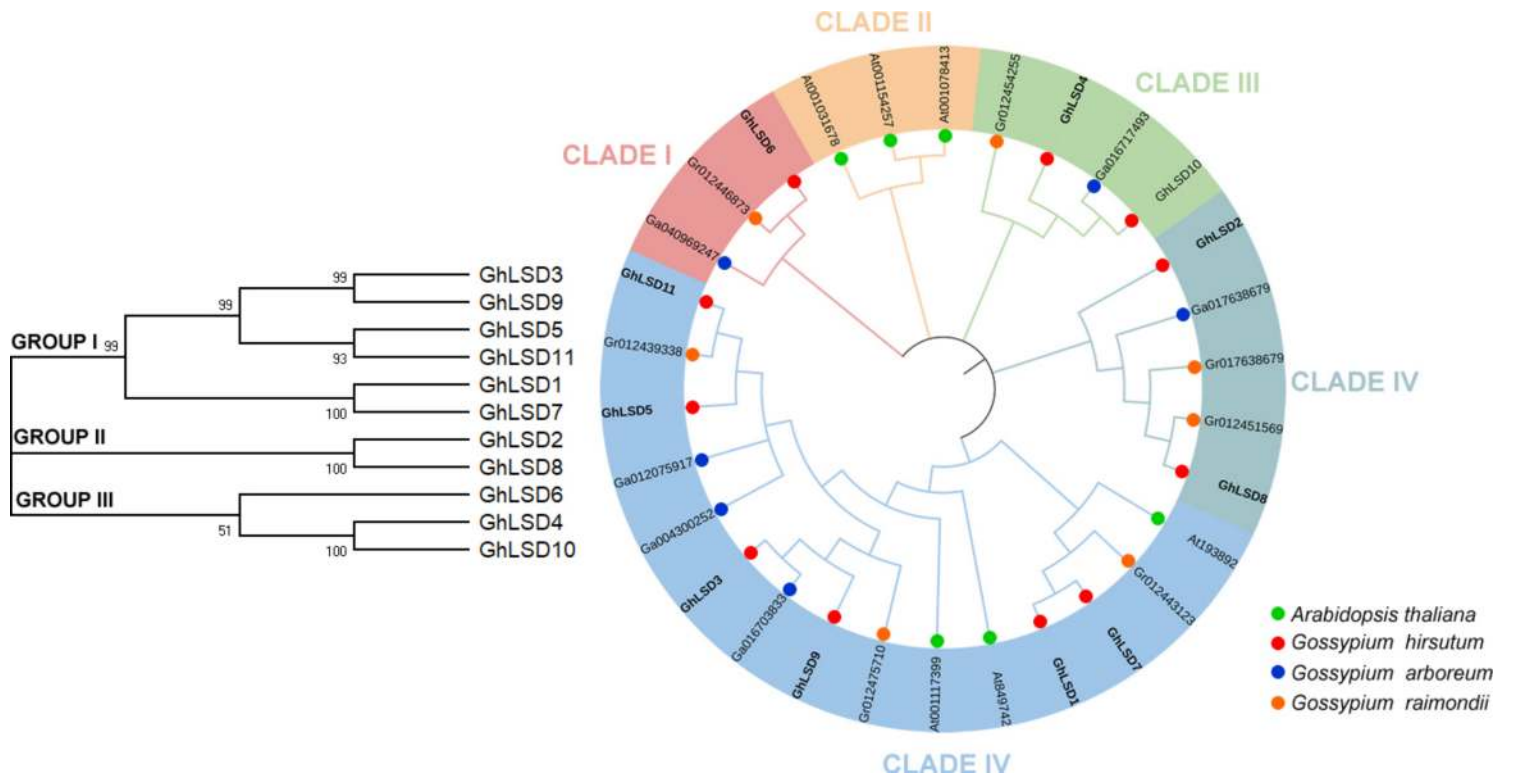


Figure 7

Phylogenetic analysis of LSD genes among *H. hirsutum*, *Arabiopsis thaliana*, *G. arboretum*, *G. raimondii*.

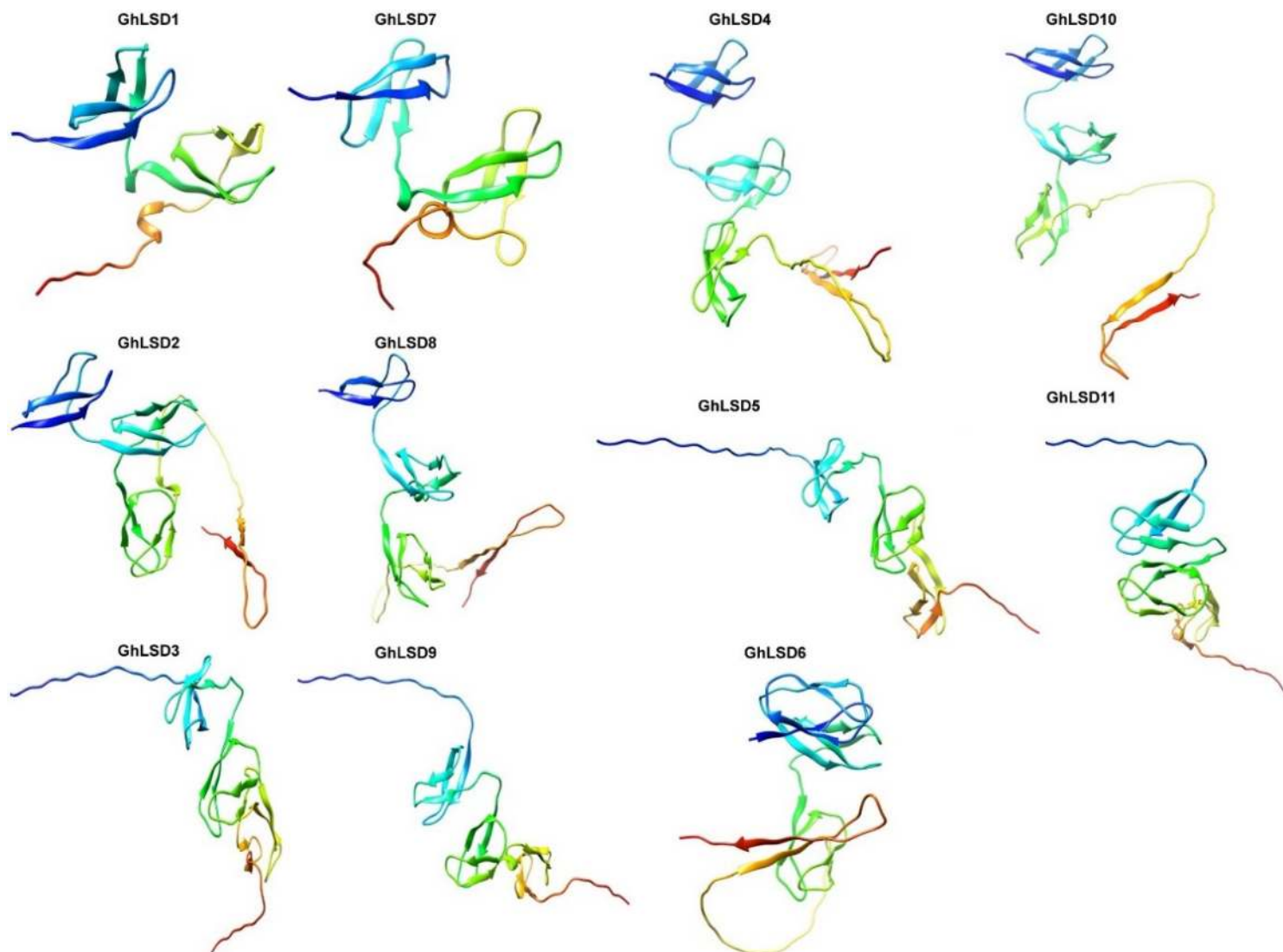


Figure 8

Tertiary (3°) structure of eleven GhLSD Proteins.

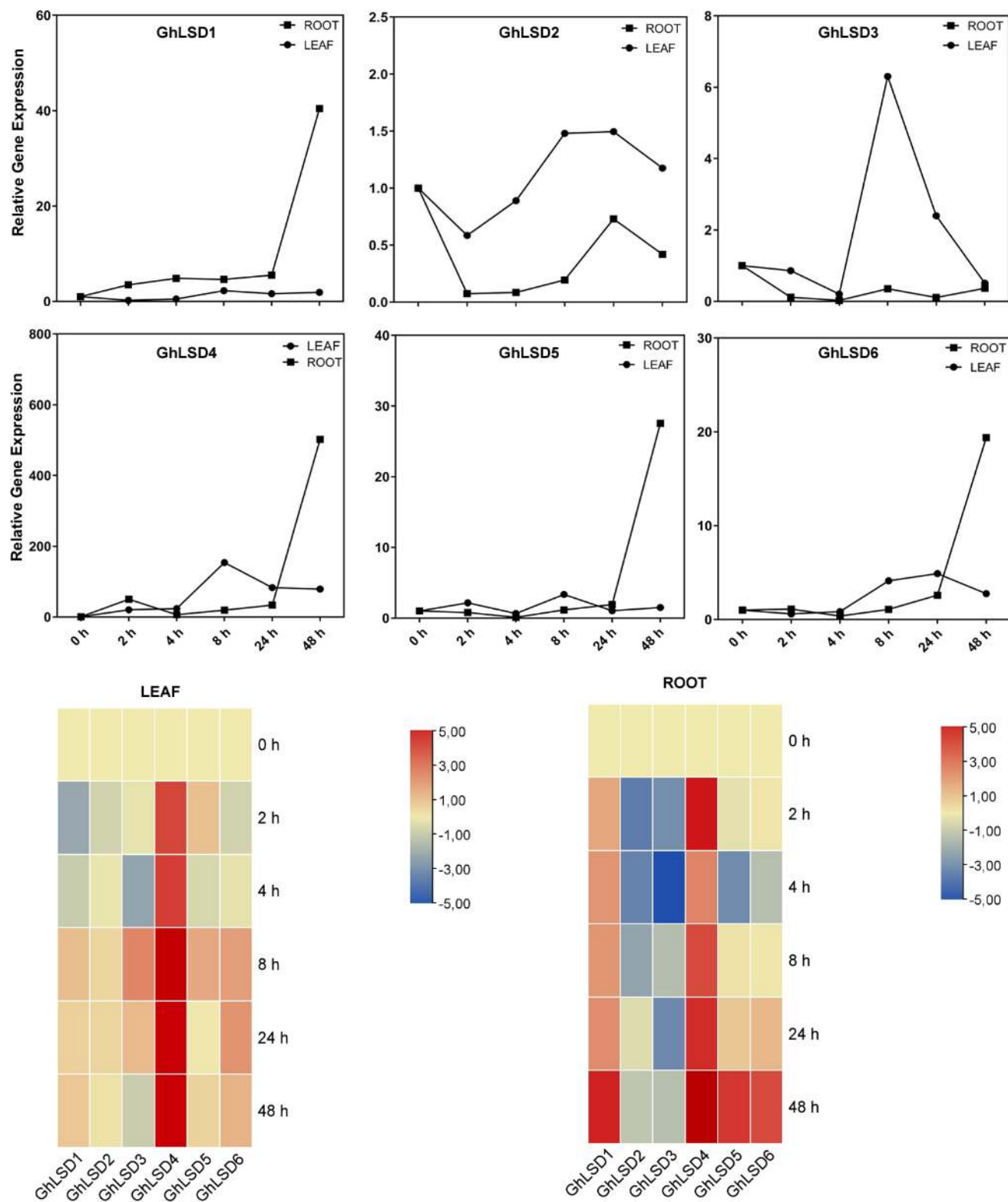


Figure 9

Relative expression of GhLSD genes in leaves and roots of *Gossypium hirsutum* under drought stress at 0, 2, 4, 8, 24 and 48 h. Heat map of GhLSD gene expression patterns in leaves and roots of *Gossypium hirsutum* under drought stress at 0, 2, 4, 8, 24, and 48 h. Downregulated genes are shown in blue, whereas upregulated genes are shown in red.

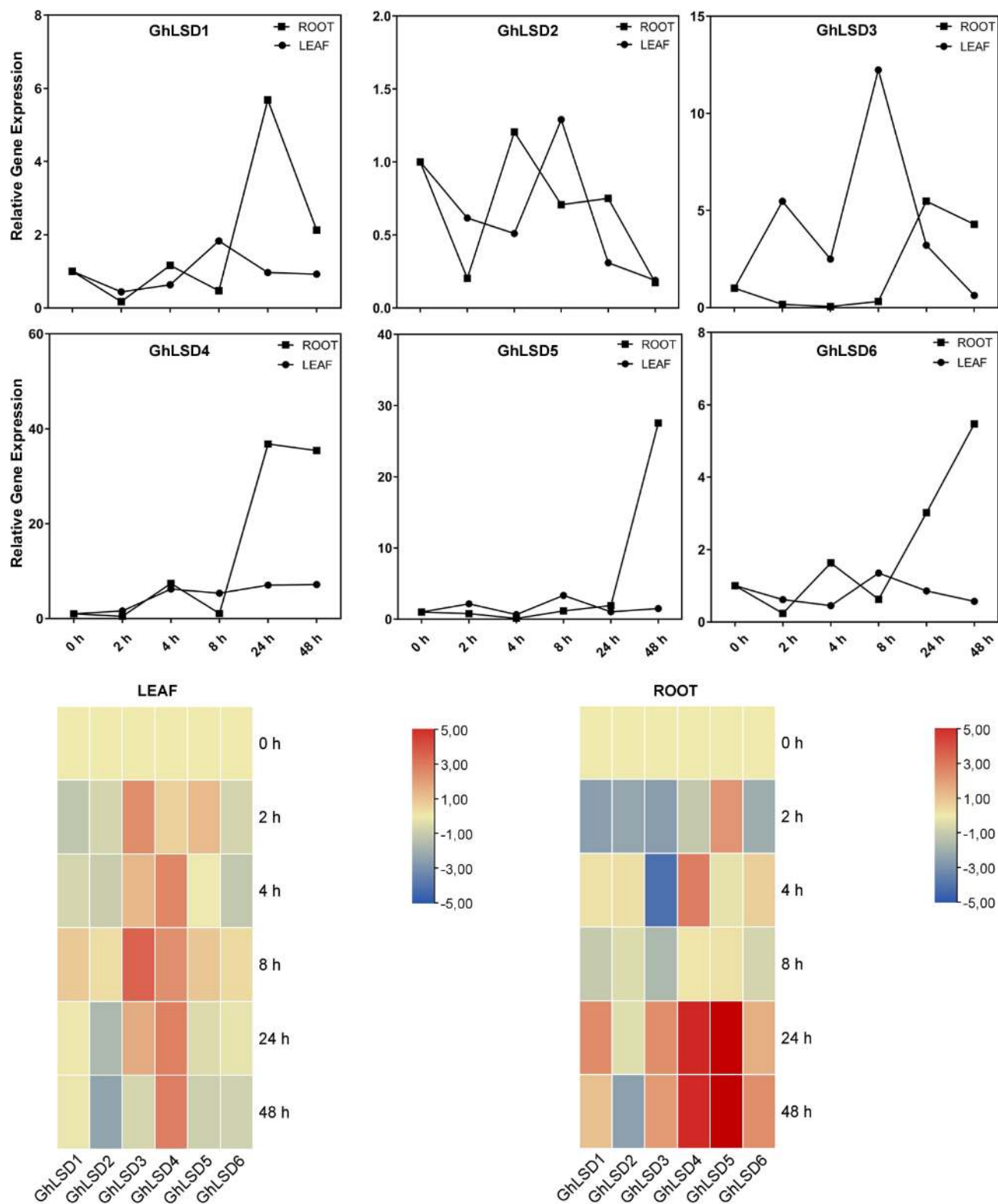


Figure 10

Relative expression of GhLSD genes in leaves and roots of *Gossypium hirsutum* under salt stress at 0, 2, 4, 8, 24 and 48 h. Heat map of *GhLSD* gene expression patterns in leaves and roots of *Gossypium hirsutum* under salt stress at 0, 2, 4, 8, 24, and 48 h. Downregulated genes are shown in blue, whereas upregulated genes are shown in red.