Plant Science Horizons

Research Article

Open-Access

Plant Science Horizons 2026, 1(2): e0015
https://doi.org/10.66835/psh.2026.e0015

Check for updates

WGCNA Reveals Conserved Molecular Responses to Osmotic Stress in *Gossypium hirsutum* and *Gossypium barbadense*

Haiying Li[1] and Pengbo Hao[1*]

[1]*College of Life Sciences, Zaozhuang University, Zaozhuang, Shandong, 277160, China*
***Correspondence:** hao_pengbo@163.com (Hao P)

Abstract

Cotton is an important natural fiber crop worldwide. Drought and osmotic stress severely impair cotton growth, fiber quality, and yield. Although *Gossypium hirsutum* and *G. barbadense* exhibit different phenotypic and physiological changes under water deficit, whether there are conserved molecular mechanisms underlie their responses to osmotic stress remains unclear. Here, we compared the transcriptome responses of *G. hirsutum* and *G. barbadense* to polyethylene glycol (PEG)-induced osmotic stress using weighted gene co-expression network analysis (WGCNA). Eight and nine gene co-expression modules were identified in *G. hirsutum* and *G. barbadense*, respectively. Among these modules, the *G. hirsutum* black module and the *G. barbadense* purple module exhibited similar expression patterns, a high proportion of orthologous genes and strong module preservation, indicating that the two modules are highly conserved. In addition, the two modules were jointly enriched in 181 gene ontology (GO) terms, mainly involving water transport, photosynthesis, metabolism, and response to stimuli. Seven pairs of candidate water transport-related orthologous genes were identified, such as genes encoding ABC transporters and plasma membrane intrinsic proteins (PIPs). Furthermore, both modules were enriched in established stress-responsive TF families including ERF, bHLH, MYB, NAC, WRKY, bZIP and HSF. Hub genes of the two modules were jointly enriched in 21 GO terms related to chloroplast organization, circadian rhythm, light and hormone responses, and abiotic stress responses. These results reveal conserved molecular responses to PEG-induced osmotic stress in *G. hirsutum* and *G. barbadense* and provide important candidate genes and theoretical basis for genetic improvement of drought resistance in cotton.

Keywords: *G. hirsutum*; *G. barbadense*; Osmotic stress; WGCNA; Conserved module

Citation: Li h., Hao P. 2026. WGCNA Reveals Conserved Molecular Responses to Osmotic Stress in *Gossypium hirsutum* and *Gossypium barbadense*. Plant Science Horizons 1(2): e0015 https://doi.org/10.66835/psh.2026.e0015

Introduction

Climate change has led to frequent extreme weather events, among which drought stress has become one of the primary abiotic factors limiting agricultural productivity. As the world's most important natural fiber crop, cotton is highly sensitive to water deficiency[1]. Cotton exhibits indeterminate growth, and the formation of new nodes depends highly on water availability[2]. Water deficit directly inhibits carbon assimilation, reduces biomass accumulation, and causes depletion of stored carbohydrates due to ongoing respiration, ultimately impairing the transport of photosynthates to reproductive organs. This leads to boll shedding and reduced boll weight, resulting in substantial yield losses ranging from 50% to 73%[3, 4].

© The Author(s)

The response of plants to drought stress is a complex regulatory process involving numerous genes. Transcription factors (TFs) serve as central orchestrators of plant drought stress responses, acting as molecular switches that transduce stress signals into downstream transcriptional reprogramming. Under drought conditions, plants activate both ABA-dependent and ABA-independent signaling pathways, in which multiple TF families play pivotal regulatory roles. In the ABA-dependent pathway, bZIP TFs such as AREB/ABF bind to ABRE cis-elements to activate stress-responsive genes, while MYB and MYC TFs regulate target genes through MYBRS and MYCRS elements, respectively[5]. In the ABA-independent pathway, AP2/ERF family TFs, particularly the DREB/CBF subfamily, recognize DRE/CRT sequences to induce dehydration-responsive genes[6]. Additionally, NAC, WRKY, bHLH, and HSF TFs have been identified as key positive or negative regulators of drought tolerance, modulating stomatal closure, osmolyte accumulation, antioxidant defense, and photosynthetic adjustment. These TF families have been functionally characterized across diverse plant species, establishing them as fundamental components of drought adaptation networks[7].

In cotton, several genes involved in drought response have been identified. The phosphorylation feedback loop composed of *GhMAPKKK15-GhMKK4-GhMPK6-GhWRKY59* regulates cotton responses to drought through an ABA-independent pathway[8]. The *GhCIPK6* gene is induced under drought, salt, and ABA treatment, and its overexpression can enhance cotton drought tolerance[9]. Overexpression of *GhABF2* increases the activities of catalase (CAT) and superoxide dismutase (SOD) in cotton, improving yield of transgenic plants[10]. *GbMYB5* positively regulates drought adaptation by reducing water loss, increasing proline content, and enhancing antioxidant enzyme activities[11]. *GhWRKY41* enhances drought tolerance by increasing antioxidant enzyme activities, reducing malondialdehyde content, and inducing stomatal closure[12]. Overexpression of the receptor-like kinase gene *GbRLK* can enhance drought tolerance by activating the ABA signaling pathway, which helps reduce water loss and improve stress resistance[13].

Transcriptome analysis has been widely applied to elucidate the molecular response mechanisms of different cotton varieties and tissues under various drought conditions. Zheng used transcriptome sequencing to identify a set of candidate genes associated with drought tolerance in *G. hirsutum*. Among them, *GhGABA-T* participates in cotton drought response by regulating the concentration of gamma-aminobutyric acid, while *GhAO* functions by affecting NAD synthesis[14]. In addition, transcriptome analysis has also revealed that genes such as *WRKY70*, *CIPK6*, *SnRK2.6*, and *NET1A* are significantly induced under drought stress, suggesting that these genes may act as key signaling components involved in the drought response[15]. In terms of hormone signal transduction, transcriptome studies further revealed the synergistic action of multiple signaling pathways. Han and colleagues found that the MAPK signaling pathway and endoplasmic reticulum stress-related heat shock proteins cooperatively regulate photosynthesis, drought signal transduction, and water transport processes, thereby inducing higher drought tolerance in cotton[16]. Integrated transcriptomic and metabolomic analyses have revealed that phenylpropanoid biosynthesis and lipid metabolism are key pathways differentially activated between drought-tolerant and sensitive cotton genotypes under PEG-induced drought stress[17]. Phylogeny-based comparative analysis of gene expression modulation under drought stress

across three cotton diploids has demonstrated that the transcriptional responses to drought varied significantly among cotton species, with only a small subset of genes exhibiting conserved regulation patterns[18]. WGCNA of contrasting varieties has identified core gene modules and hub genes, such as *WRKY22*, that coordinate proline biosynthesis and jasmonate signaling to enhance drought adaptation[19].

G. hirsutum (upland cotton, 2n = 4x = 52, AADD) and *G. barbadense* (sea-island cotton, 2n = 4x = 52, AADD) are two major cultivated allotetraploid cotton species that originated from a common polyploidization event approximately 1–1.5 million years ago[20]. Under various abiotic stresses, *G. hirsutum* generally exhibits stronger tolerance. Transcriptome results show that more genes are induced in *G. hirsutum* under low or high temperatures, especially genes related to ethylene and ABA signaling pathways, which are more actively involved in regulating *G. hirsutum*'s responses to cold and heat stress[20]. Under drought stress, *G. hirsutum* better maintains plant height, leaf area, and dry matter accumulation, displaying greater growth stability[21], and accumulates more proline, soluble sugars, and proteins, while the activity of antioxidant enzymes (such as SOD) is more significantly upregulated, effectively preserving chloroplast integrity and reducing membrane lipid peroxidation damage[22].

Currently, research on the molecular mechanisms of drought response in *G. hirsutum* and *G. barbadense* mostly focuses on identifying drought response genes or transcriptome analyses in individual cotton species, However, comparative WGCNA between *G. hirsutum* and *G. barbadense* under osmotic stress has not been reported. Specifically, whether conserved molecular mechanisms of osmotic stress response exist between *G. hirsutum* and *G. barbadense* remains unknown. In this study, using WGCNA, we compared the transcriptome changes of *G. hirsutum* and *G. barbadense* under PEG-induced osmotic stress, identified conserved gene co-expression modules, and performed GO enrichment and TF family analyses. Our aim was to reveal the conservation of osmotic stress response between *G. hirsutum* and *G. barbadense* and to identify candidate genes for further functional studies.

Materials and Methods

Acquisition and Quantitative Analysis of Transcriptome Data of *G. hirsutum* and *G. barbadense*

Raw transcriptome data of PEG-treated and control samples of *G. hirsutum* and *G. barbadense* were downloaded from the NCBI SRA database (Project ID: PRJNA490626). The original experiment was conducted by Hu et al.[20]. In brief, four-week-old seedlings were subjected to osmotic stress with PEG (200 g/liter), and the third true leaves were harvested at 0 h (control), 1 h, 3 h, 6 h, 12 h, and 24 h after treatment. Three biological replicates were collected for each time point and treatment. Illumina paired-end RNA sequencing was performed to generate transcriptome data. The raw data were quality-controlled using fastp (v1.1.0)[23], and then Hisat2 (v2.2.1)[24], Samtools (v1.22)[25], and Stringtie (v2.2.3)[26] were used for sequence alignment, sorting of alignment results, and transcript assembly and quantification, respectively, following the standard RNA-seq data processing pipeline as described in a recent transcriptomic study[27]. The genome annotations of *G. hirsutum* (AD1_T2T_JZU) and *G. barbadense* (H7124_ZJU) were downloaded from the COTTONGEN website (https://www.cottongen.org/data/download).

Weighted Gene Co-expression Network Analysis of *G. hirsutum* and *G. barbadense* Transcriptomes

The hclust function in R (v4.3.2) was used to perform sample clustering on the transcriptome data of *G. hirsutum* and *G. barbadense*, and outlier samples were removed based on the clustering results. The top 10,000 genes with the largest median absolute deviation (MAD) across all samples were selected for WGCNA to reduce computational noise while retaining the most variable genes. The pickSoftThreshold function in the WGCNA package (v1.72)[28] was used to calculate the scale-free topology fitting index (R^2). A soft-thresholding power (β) that achieved $R^2 > 0.8$ was selected to obtain an approximate scale-free network topology. The blockwiseModules function was used to construct the gene co-expression network in one step with network type, correlation method and minimum module size set as unsigned, pearson and 50, respectively. Then the cutHeight was set to 0.5 to merge similar modules.

Analysis of the Expression Patterns of Module Eigengenes in *G. hirsutum* and *G. barbadense*

The expression of module eigengenes (MEs) in *G. hirsutum* and *G. barbadense* contained multiple biological replicates. The average expression was calculated and visualized using heatmap with expression patterns clustered.

Analysis of Orthologous Genes Between *G. hirsutum* and *G. barbadense* Co-expression Modules

Using Diamond (v0.39)[29], all protein sequences of *G. hirsutum* and *G. barbadense* were aligned against each other with an e-value threshold of 1×10^{-5} and a minimum sequence coverage of 80%. Reciprocal best hits (RBH) between the two species were identified as orthologous gene pairs. Then, orthologous gene pairs specifically between co-expression modules of *G. hirsutum* and *G. barbadense* were extracted. The significance of 914 orthologous gene pairs in the *G. hirsutum* black module (2343) and the *G. barbadense* purple module (2191) was evaluated against a genomic background using a hypergeometric test, with all annotated genes in both genomes as the background set (*G. hirsutum*: 79642, *G. barbadense*: 75071, orthologous genes: 44316).

Module Preservation and Connectivity Correlation Analysis

The modulePreservation function in the WGCNA package (v1.72) was used to perform module preservation analysis. Gene IDs in the expression data of *G. barbadense* were replaced with their orthologous genes of *G. hirsutum*. The expression data and module assignments of *G. hirsutum* were used as the reference network to evaluate the preservation of its modules in *G. barbadense*, and vice versa. For each module, a preservation statistic Zsummary was calculated based on multiple density- and connectivity-based preservation statistics. The co-expression network of the *G. hirsutum* black module and the *G. barbadense* purple module were exported, then input to the igraph package to calculate the weighted degree of all orthologous gene pairs between the two modules. Pearson correlation analysis was performed to assess the correlation of weighted degree between orthologous genes across the two species.

GO Enrichment of Genes in *G. hirsutum* Black Module and *G. barbadense* Purple Module

All the protein sequences of *G. hirsutum* and *G. barbadense* were submitted to the eggNOG-mapper website (http://eggnog-mapper.embl.de/) for GO

annotation. The GO annotation and the gene lists in the *G. hirsutum* black module and the *G. barbadense* purple module were input into the GO Enrichment function in TBtools (v2.475) for GO enrichment[30], in which hypergeometric test was performed, background gene set comprised all genes with GO annotations in the respective genome, adjusted p-value (FDR, Benjamini–Hochberg method) < 0.05 was applied. The enriched GO terms from the two modules were compared using a Venn diagram to identify jointly enriched GO terms. These jointly enriched GO terms were submitted to the Revigo website (http://revigo.irb.hr/) and clustered based on the semantic similarity of GO terms[31].

Identification of Water Transport-Related Orthologous Genes in *G. hirsutum* Black Module and *G. barbadense* Purple Module

GO terms related to water transport were selected, and the orthologous genes annotated to these GO terms were extracted. The protein sequences of these genes to were submitted to the TAIR website (https://www.arabidopsis.org/) to find the *Arabidopsis* homologous genes and their functional annotations.

Identification of Transcription Factors in *G. hirsutum* Black Module and *G. barbadense* Purple Module

The protein sequences of genes in the *G. hirsutum* black module and the *G. barbadense* purple module were submitted to the PlantTFDB website (https://planttfdb.gao-lab.org/prediction.php) to predict the transcription factors[32].

Identification and Functional Analysis of Hub Genes in *G. hirsutum* Black Module and *G. barbadense* Purple Module

The top 5% of genes with the highest weighted degree from the *G. hirsutum* black module and the *G. barbadense* purple module were selected as hub genes in the two modules. Orthologous relationships between hub genes of the two modules were extracted from previous analysis of orthologous genes. GO enrichment analysis of hub genes in each module was performed using the eggNOG-mapper annotations and the GO Enrichment function in TBtools (v2.475) as described previously. The enriched GO terms from the two modules were compared using a Venn diagram to identify jointly enriched functional categories. The significance of the jointly enriched GO terms was visualized using a bar plot.

Results

Identification of Gene Co-expression Modules in *G. hirsutum* and *G. barbadense* under PEG Treatment

Gene co-expression modules were identified through weighted gene co-expression network analysis (WGCNA). First, hierarchical clustering was performed on PEG-treated and control samples of *G. hirsutum* and *G. barbadense*, and no outlier samples were detected in either species (Figure 1). Therefore, the expression matrices of the top 10,000 genes with the highest median absolute deviations in all the samples of *G. hirsutum* and *G. barbadense* were used for WGCNA. According to the scale-free topology model, the co-expression network constructed by WGCNA approximates a scale-free network when $R^2 > 0.8$. Therefore, the soft-thresholding power was set to 6 for constructing the co-expression networks of *G. hirsutum* and *G. barbadense* (Figure 2). A total of 15 and 21 gene co-expression modules were initially identified in *G. hirsutum* and *G. barbadense*, respectively. After merging similar modules, eight and nine gene co-expression modules remained in *G. hirsutum* and *G.*

barbadense, respectively (Figure 3A and 3B). The number of genes in *G. hirsutum* co-expression modules ranged from 48 (grey module) to 4,239 (greenyellow module), while in *G. barbadense* co-expression modules, it ranged from 24 (grey module) to 3,173 (turquoise module) (Figure 3C and 3D).

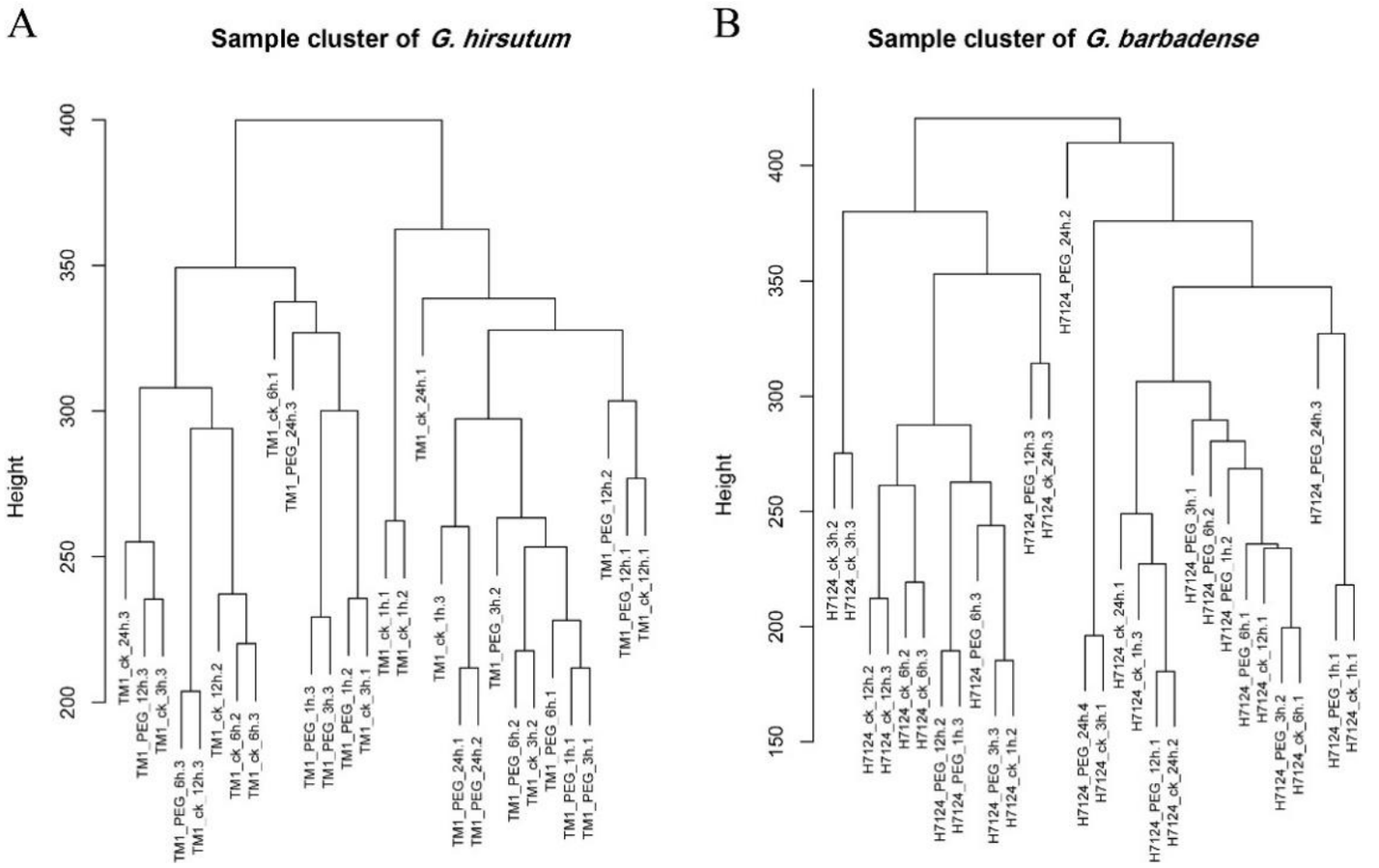

Figure 1. Sample cluster of *G. hirsutum* and *G. barbadense*. (A) Sample cluster of *G. hirsutum* (TM-1) under PEG treatment (PEG) and control (ck) conditions. (B) Sample cluster of *G. barbadense* (Hai7124) under PEG treatment (PEG) and control (ck) conditions. The y-axis indicates the clustering height (dissimilarity).

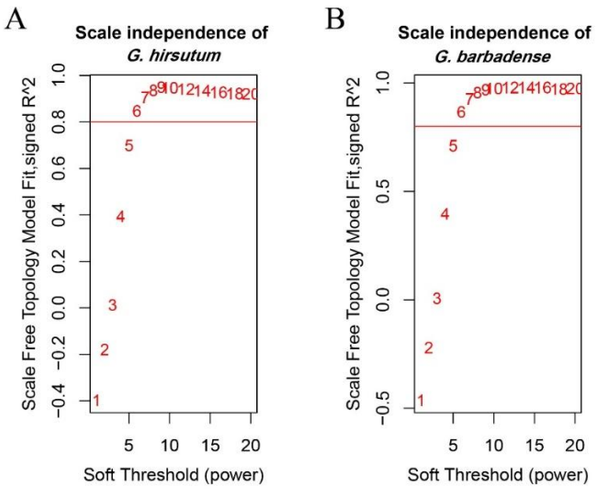

Figure 2. Determination of soft threshold for network construction. (A) Scale independence of *G. hirsutum*. (B) Scale independence of *G. barbadense*. The red horizontal line marks the threshold of $R^2 = 0.8$.

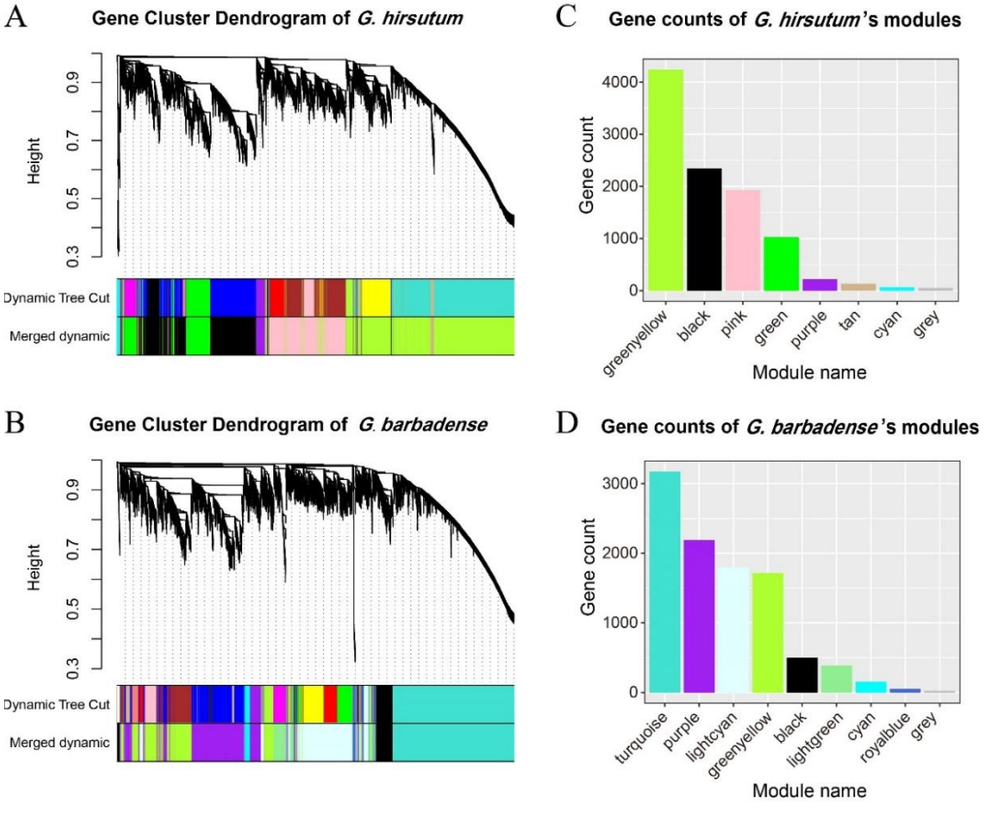

Figure 3. Gene co-expression modules in *G. hirsutum* and *G. barbadense*. (A, B) Gene cluster dendrograms with module assignments for *G. hirsutum* (A) and *G. barbadense* (B). The color rows beneath each dendrogram show the modules identified by dynamic tree cut (Dynamic Tree Cut) and the modules after merging similar modules at cutHeight = 0.5 (Merged dynamic). (C, D) Gene counts of the co-expression modules in *G. hirsutum* (C) and *G. barbadense* (D).

Correlation Analysis of Gene Co-expression Modules in *G. hirsutum* and *G. barbadense*

The expression patterns of 8 module eigengenes in *G. hirsutum* and 9 module eigengenes in *G. barbadense* were clustered. The results showed that, except for two gray modules (genes that could not be classified into co-expression modules were placed in the gray modules), three pairs of *G. hirsutum* and *G. barbadense* co-expression modules exhibited similar expression patterns (Figure 4). The expression of the *G. hirsutum* black module and the *G. barbadense* purple module first decreased and then increased during PEG treatment, while their expression levels remained low in control samples, except that the *G. hirsutum* black module showed higher expression at 1 hour and 3 hours in control samples (Figure 4). *G. hirsutum* cyan module and *G. barbadense* lightcyan module maintained low expression during PEG treatment, whereas they maintained high expression in control samples, except at 3 hours (Figure 4). *G. hirsutum* pink module and *G. barbadense* cyan module showed increased expression levels at 12 and 24 hours of PEG treatment, while their expression in control samples fluctuated at high levels (Figure 4). Notably, the expression changes of the *G. hirsutum* black and pink modules were more pronounced than those of the *G. barbadense* purple and cyan modules, suggesting that *G. hirsutum* may respond to osmotic stress by altering gene expression to a greater extent.

Orthologous Genes in *G. hirsutum* and *G. barbadense* Co-expression Modules with Similar Expression Patterns

The *G. hirsutum* black module and the *G. barbadense* purple module contained 39% (914/2343) and 41.7% (914/2191) orthologous genes, respectively (Figure 5A), suggesting that the genes in these two modules may play similar

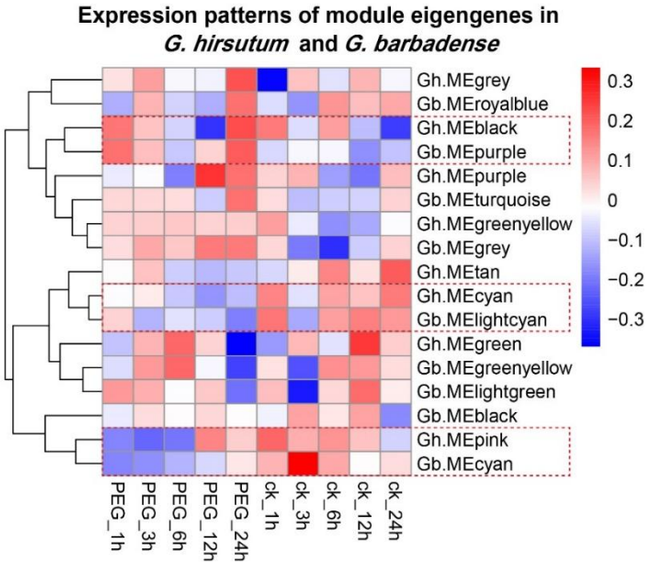

Figure 4 Expression patterns of module eigengenes in *G. hirsutum* and *G. barbadense*. Rows represent module eigengenes, labeled with the species prefix (Gh., *G. hirsutum*; Gb., *G. barbadense*) followed by the module name. For each ME, the mean expression across biological replicates was used. The color scale indicates the relative expression level of each ME, with red and blue representing higher and lower values, respectively, and white indicating values close to zero. Three pairs of *G. hirsutum* and *G. barbadense* modules showing similar expression patterns are marked by dashed boxes.

functions in osmotic response of *G. hirsutum* and *G. barbadense*. A hypergeometric test confirmed that this overlap was highly significant compared to the genomic background ($p \approx 0$). *G. hirsutum* cyan module and *G. barbadense* lightcyan module did not contain orthologous genes (Figure 5B). *G. hirsutum* pink module and *G. barbadense* cyan module contained 0.6% (12/1929) and 7.7% (12/156) orthologous genes, respectively (Figure 5C), indicating that the functions of genes in *G. hirsutum* pink module may be more diverse than those in *G. barbadense* cyan module.

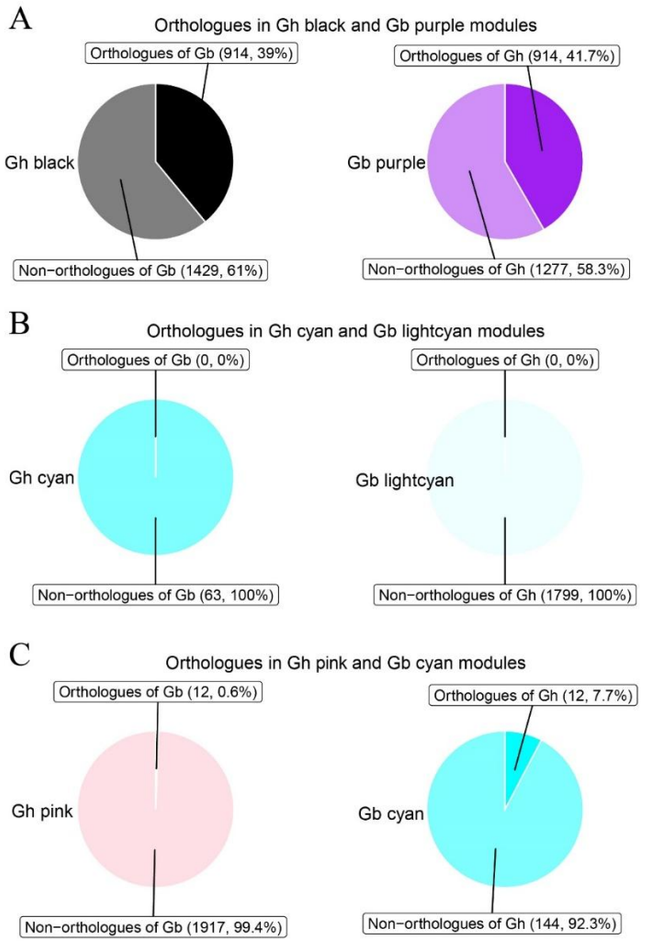

Figure 5. Proportion of orthologous genes in the *G. hirsutum* and *G. barbadense* co-expression modules with similar expression patterns. (A) *G. hirsutum* black module (left) and *G. barbadense* purple module (right). (B) *G. hirsutum* cyan module (left) and *G. barbadense* lightcyan module (right). (C) *G. hirsutum* pink module (left) and *G. barbadense* cyan module (right).

Quantitative Assessment of Module Conservation between *G. hirsutum* and *G. barbadense*

To further validate the conservation of the *G. hirsutum* black module and of the *G. barbadense* purple module beyond similar expression patterns and a high proportion of orthologous genes, we performed quantitative module preservation analysis. when *G. hirsutum* was used as the reference network, the *G. hirsutum* black module showed the highest Zsummary score (Figure 6A), confirming strong preservation of the black module in *G. barbadense*; when *G. barbadense* was used as the reference network, the *G. barbadense* purple module showed the highest Zsummary score (Figure 6B), confirming strong preservation of the purple module in *G. hirsutum*. In addition, we compared the weighted degree of the 914 orthologous gene pairs between the *G. hirsutum* black module and the *G. barbadense* purple module. The results revealed a highly significant positive correlation between the connectivity of orthologous genes in the two modules (Pearson's $R = 0.85$, $p < 2.2e-16$; Figure 6C), indicating that orthologous genes occupy similar topological positions and maintain comparable connectivity within their respective modules. Collectively, these expression pattern, orthologous genes, module preservation and connectivity correlation analyses indicate that the *G. hirsutum* black module and the *G. barbadense* purple module represent conserved functional modules in the osmotic stress response of the two cotton species.

GO Enrichment of Genes in *G. hirsutum* Black Module and *G. barbadense* Purple Module

Genes in the *G. hirsutum* black module and the *G. barbadense* purple module were jointly enriched in 181 GO terms, accounting for 64.9% of the 279 GO terms enriched in the *G. hirsutum* black module and 78.4% of the 231 GO terms enriched in the *G. barbadense* purple module (Figure 7A), indicating that the two modules share largely overlapping biological functions in the osmotic stress response. In addition, the 181 GO terms jointly enriched in the *G. hirsutum* black module and the *G. barbadense* purple module could be categorized into four categories, including GO terms related to transport, photosynthesis, metabolism, and stimulus response. The transport-related GO terms included

multiple biological processes (BP) and molecular functions (MF) such as water transport, water channel activity, and ion transport (Figure 7B). The photosynthesis-related GO terms included multiple biological processes and cellular components (CC) such as regulation of photosynthesis, regulation of circadian rhythm, and chloroplast (Figure 7C). The metabolism-related GO terms included multiple biological processes of carbohydrate biosynthetic and metabolic processes (Figure 7D). The stimulus response-related GO terms included multiple biological processes of response to stimulus (Figure 7E).

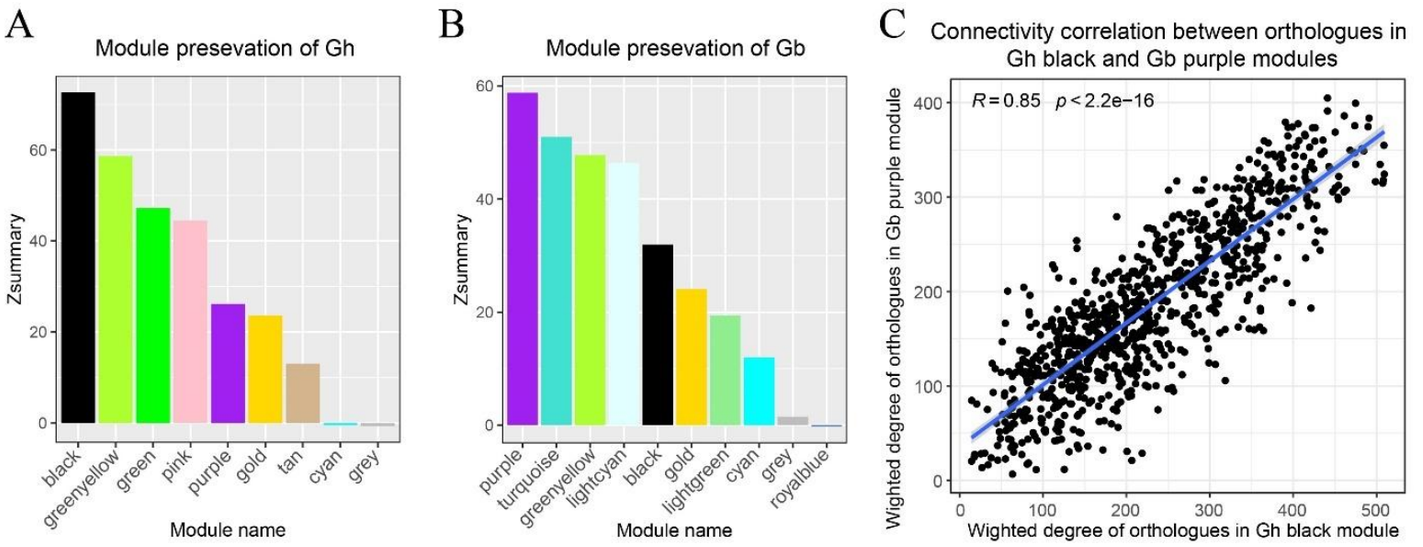

Figure 6. Module preservation of the *G. hirsutum* black module and the *G. barbadense* purple module. (A) Preservation of *G. hirsutum* (Gh) modules in *G. barbadense*, using the *G. hirsutum* network as the reference. (B) Preservation of *G. barbadense* (Gb) modules in *G. hirsutum*, using the *G. barbadense* network as the reference. (C) Correlation of gene connectivity (weighted degree) between the 914 orthologous gene pairs of the *G. hirsutum* black module and the *G. barbadense* purple module.

Candidate Water Transport-Related Genes in the *G. hirsutum* Black Module and *G. barbadense* Purple Module

Since the *G. hirsutum* black module and the *G. barbadense* purple module were jointly enriched in multiple water transport-related GO terms, we identified candidate water transport-related orthologous genes in the two modules. A total of 7 pairs of candidate water transport-related orthologous genes were identified (Table 1). Among them, *GhChrA01G2499* and *GB_A01G2272*, as well as *GhChrD01G2401* and *GB_D01G2355* were homologous to *AT5G06530* (*ABCG22*) in *Arabidopsis*, which encodes an ABC transporter protein. Mutation of this gene leads to increased water transpiration and decreased drought resistance (Table 1). *GhChrD09G1713* and *GB_D09G1633*, *GhChrD10G2823* and *GB_D10G2701*, *GhChrD10G2824* and *GB_D10G2702*, *GhChrA10G2921* and *GB_A10G2740*, and *GhChrA06G2271* and *GB_A06G1934* belonged to the plasma membrane intrinsic protein (PIP) family and are putatively involved in water transport based on their *Arabidopsis* homologs (Table 1).

Transcription Factors in *G. hirsutum* Black Module and *G. barbadense* Purple Module

Transcription factors (TFs) play a key role in regulating the expression of stress-responsive genes. To characterize the transcriptional regulatory components within the conserved modules, we identified candidate TFs in the *G. hirsutum* black module and the *G. barbadense* purple module. In the *G. hirsutum* black module, 7.9% (186/2343) of genes were predicted as TFs, while in the *G. barbadense* purple module, 9.6% (210/2191) of genes were predicted as TFs (Figure 8A). Notably,

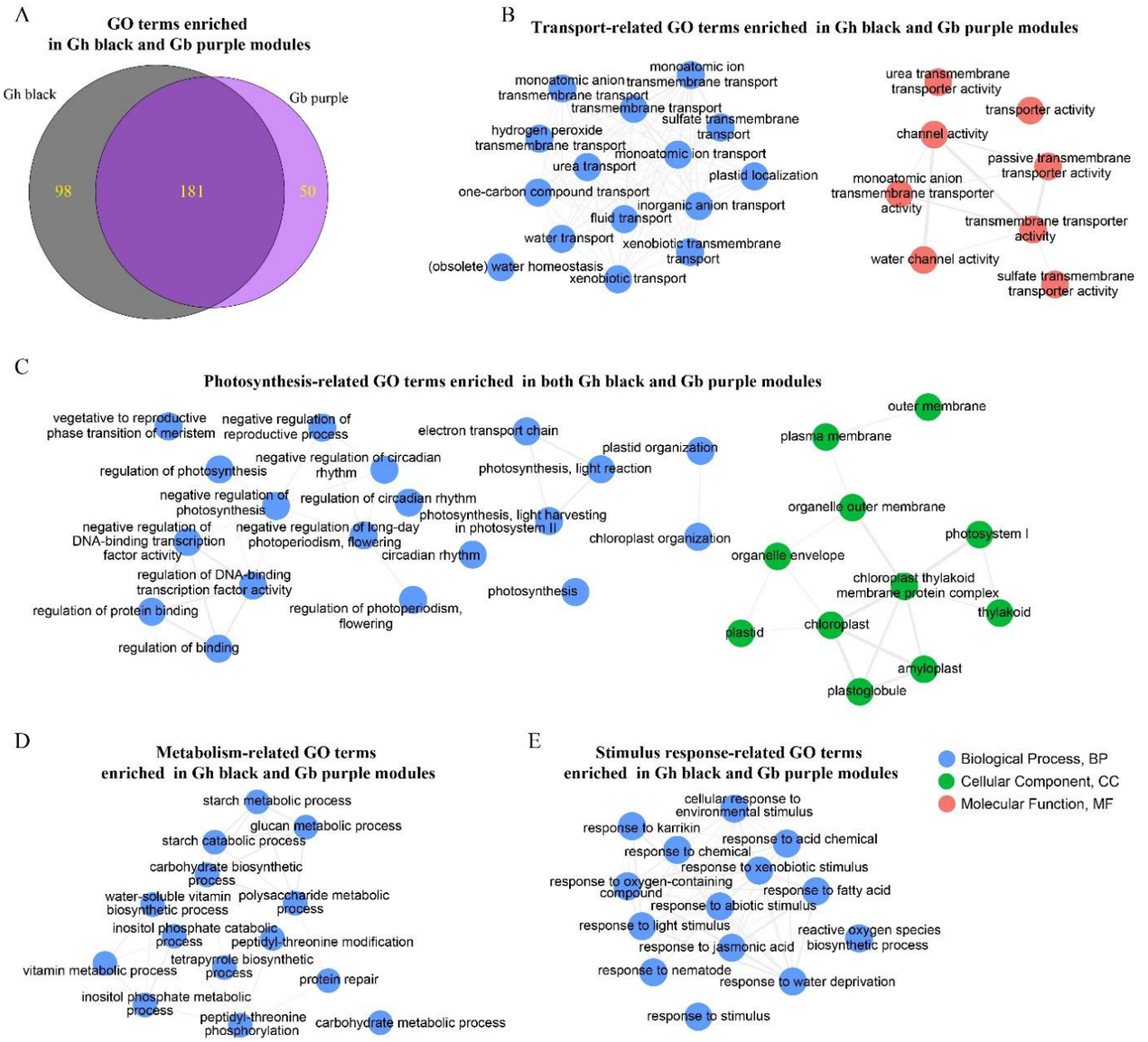

Figure 7. GO terms enriched in the *G. hirsutum* module black and the *G. barbadense* module purple. (A) Venn diagram comparing the enriched GO terms of the *G. hirsutum* black module and the *G. barbadense* purple module. (B–E) The 181 jointly enriched GO terms were submitted to REVIGO and clustered based on semantic similarity into four functional groups. edges indicate semantic similarity between terms. (B) Transport-related GO terms. (C) Photosynthesis-related GO terms. (D) Metabolism-related GO terms. (E) Stimulus-response-related GO terms.

Table 1. Candidate water transport-related orthologous genes in the *G. hirsutum* black module and the *G. barbadense* purple module

Gene ID of Gh	Gene ID of Gb	Homolog in *Arabidopsis*	Putative gene function
GhChrA01G2499	*GB_A01G2272*	*AT5G06530 (ABCG22)*	ABC transporter, Mutation results in increased water transpiration and drought susceptibility
GhChrD01G2401	*GB_D01G2355*	*AT5G06530 (ABCG22)*	ABC transporter, Mutation results in increased water transpiration and drought susceptibility
GhChrD09G1713	*GB_D09G1633*	*AT2G37170 (PIP2B)*	Plasma membrane intrinsic protein, exhibiting water transport activity in Xenopus oocyte
GhChrD10G2823 *GhChrD10G2824* *GhChrA10G2921*	*GB_D10G2701* *GB_D10G2702* *GB_A10G2740*	*AT2G45960 (PIP1B)*	Plasma membrane intrinsic protein, exhibiting water transport activity in Xenopus oocyte
GhChrA06G2271	*GB_A06G1934*	*AT4G00430 (PIP1;4)*	Plasma membrane intrinsic protein, involved in water transport

both modules were enriched in members of the ERF, bHLH, MYB, NAC, WRKY, bZIP, and HSF families (Figure 8B), all of which have been extensively documented as key regulators of drought and osmotic stress responses in plants. In addition, both modules contained TFs belonging to the Dof, G2-like, CO-like, and MIKC_MADS families (Figure 8B), which have also been implicated in stress adaptation and developmental reprogramming under water deficit. Although the relative abundance of individual TF families varied between the two modules, the co-occurrence of multiple stress-responsive TF families in both the *G. hirsutum* black module and the *G. barbadense* purple module indicates that both species deploy a broad repertoire of established drought-related transcriptional regulators in these conserved modules.

Hub Genes and Their Functional Conservation in *G. hirsutum* Black Module and *G. barbadense* Purple Module

To elucidate the core regulatory genes within the conserved modules, we identified hub genes in the *G. hirsutum* black module and the *G. barbadense* purple module based on weighted degree. The top 5% of genes with the highest weighted degree were selected as hub genes, yielding 117 genes from the *G. hirsutum* black module and 110 genes from the *G. barbadense* purple module. Among these hub genes, 48 genes (41.0%) from the *G. hirsutum* black module had orthologous genes in the *G. barbadense* purple module, and conversely, 48 genes (43.6%) from the *G. barbadense* purple module had orthologous genes in the *G. hirsutum* black module (Figure 9A). GO enrichment analysis of hub genes revealed that 21 GO terms were jointly enriched in both modules (Figure 9B), including functions related to chloroplast and plastid, circadian rhythm and rhythmic process, response to light stimulus and radiation, response to hormones, response to abiotic stimulus, and negative regulation of transcription (Figure 9C). Interestingly, the enrichment significance of these 21 GO terms varied

between the two species. Generally, GO terms exhibited higher enrichment significance in the *G. hirsutum* black module than in the *G. barbadense* purple module (Figure 9C).

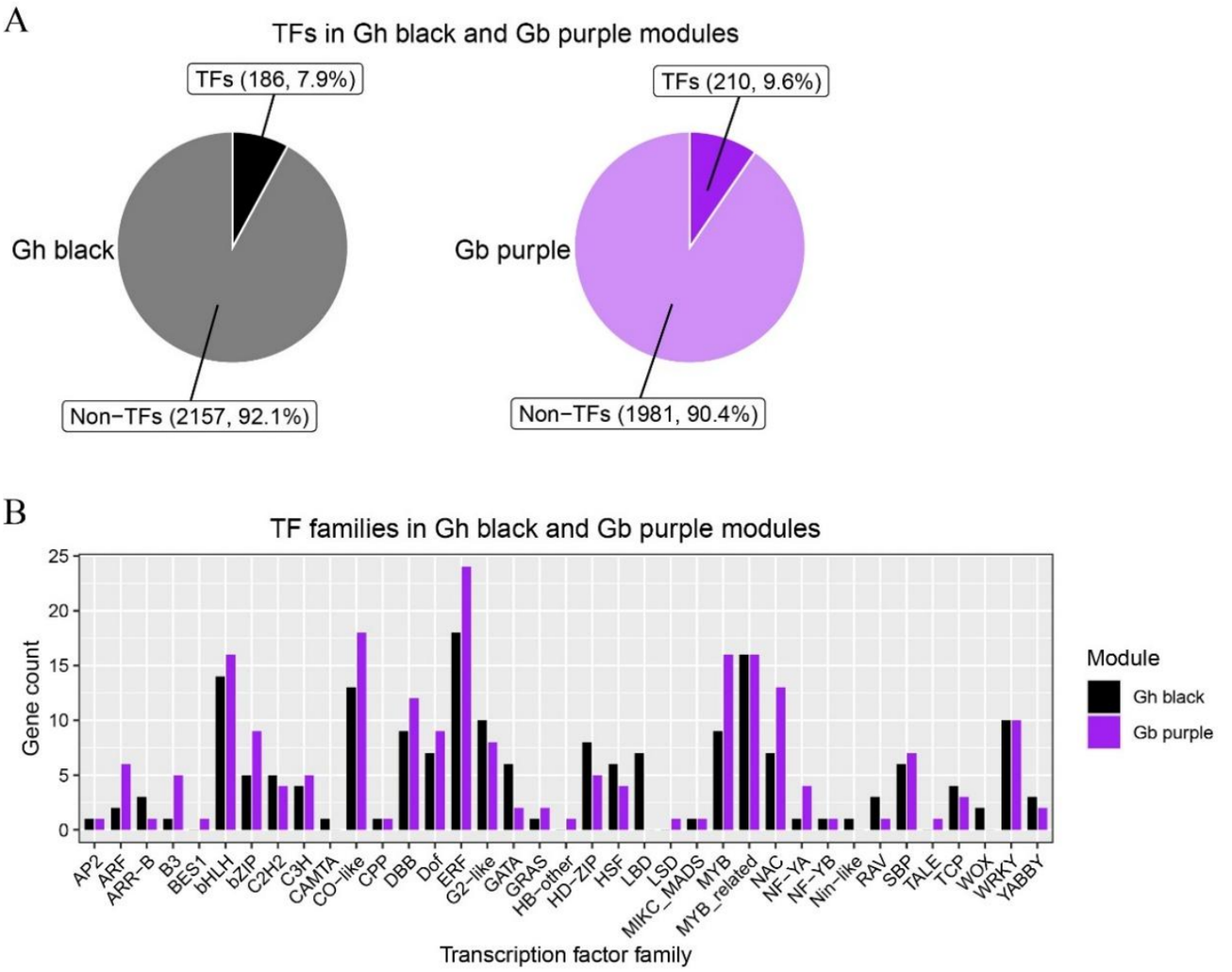

Figure 8. Transcription factors in the *G. hirsutum* module black and the *G. barbadense* module purple. (A) Pie charts showing the numbers and percentages of predicted transcription factors (TFs) and non-TFs in each module. (B) Numbers of TF genes in each transcription factor family of the *G. hirsutum* black module (black bars) and the *G. barbadense* purple module (purple bars).

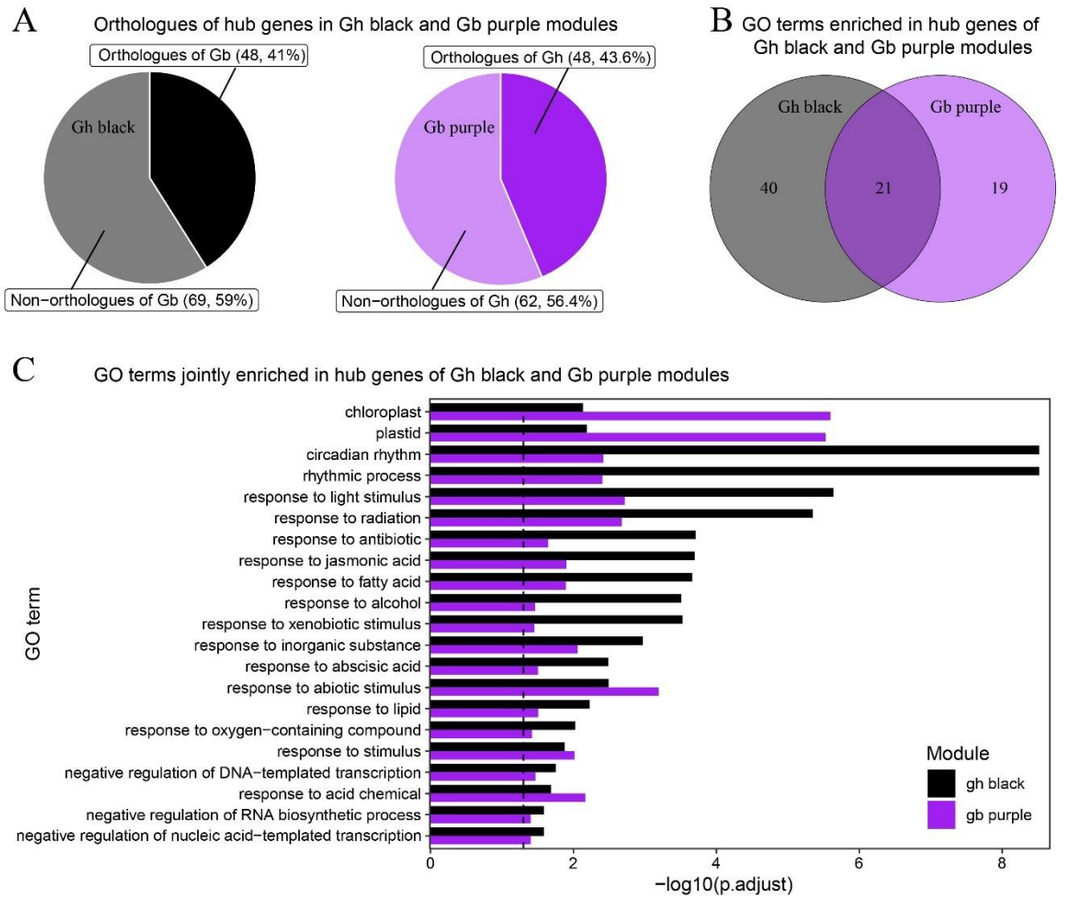

Figure 9. Functional conservation of hub genes in the *G. hirsutum* black module and the *G. barbadense* purple module. (A) Pie charts showing the numbers and percentages of hub genes in each module that have orthologous genes in the other module. (B) Venn diagram comparing the GO terms enriched among the hub genes of the two modules. (C) Bar plot of the 21 jointly enriched GO terms.

Discussion

This study applied the WGCNA method to identify multiple gene co-expression modules with conserved expression patterns in *G. hirsutum* and *G. barbadense* under drought stress. Among them, the *G. hirsutum* black module and the *G. barbadense* purple module showed a consistent 'first decreasing then increasing' expression pattern, with 39% and 41.7% of the genes in each module being orthologous, respectively (Figure 5A). The 181 jointly enriched GO terms of the two modules suggest that they may perform similar functions in the drought response of *G. hirsutum* and *G. barbadense*. Furthermore, TF families in the two modules were primarily composed of stress-responsive ERF, bHLH, MYB, NAC, WRKY, bZIP and HSF families (Figure 8). These results demonstrate that although *G. hirsutum* and *G. barbadense* exhibit different phenotypic and physiological changes under drought stress[20-22], conserved molecular mechanisms for drought response still exist.

The *G. hirsutum* black module and *G. barbadense* purple module were jointly enriched in four major GO functional categories related to transport, photosynthesis, metabolism, and stimulus

response (Figure 7). These GO functional categories are highly consistent with the known mechanisms of plant drought response[33], which involve two phases of transcriptional reprogramming: during the early stage of drought stress (1–6 hours), due to a sudden increase in ABA signaling, stomatal closure, and inhibition of photosynthesis, growth- and assimilation-related genes are passively downregulated to reduce energy consumption; as drought stress continues (12–24 hours), plants activate defense-recovery programs through transcription factors such as *bHLH*/*ERF*/*MYB*/*WRKY*/*NAC*, leading to the expression of genes related to water channel proteins, photosystem repair components, osmotic regulation, and antioxidants rebounding[34]. Among the 181 GO terms enriched in the *G. hirsutum* black module and *G. barbadense* purple module, enrichment of functions related to water transport is particularly prominent. Seven pairs of candidate water transport-related orthologous genes were identified, including genes encoding ABC transporters and plasma membrane intrinsic proteins (PIPs) (Table 1), directly pointing to the core processes of water absorption and transport. For example, *ABCG22* encodes an ABC transporter in *Arabidopsis*, and mutation of this gene leads to increased transpiration and reduced drought resistance[35], indicating its key negative regulatory role in drought response. PIPs are the main channels facilitating transmembrane water flow, and play key roles in drought response[36, 37].

Drought stress directly inhibits carbon assimilation, leading to reduced photosynthetic efficiency and biomass accumulation. Photosystem II (PSII), as a sensitive target in plants responding to stress, suffers photoinhibition damage under drought stress, ultimately resulting in decreased yield. Cotton varieties with strong drought resistance can mitigate drought-induced damage by maintaining the stability of PSII[38, 39]. This study found that GO terms related to photosynthesis regulation, light reaction, chloroplast organization, and circadian regulation were enriched in the *G. hirsutum* black module and *G. barbadense* purple module (Figure 7C), indicating that after perceiving osmotic signals, plants can initiate protection of the photosynthetic system, such as optimizing energy metabolism by adjusting the circadian clock or slowing reactive oxygen species production through negative regulation of photosynthesis. This is consistent with previous report showing widespread change in photosynthesis-related genes in cotton transcriptomes under drought stress[16], further confirming that fine-tuning of photosynthetic efficiency is an important component of cotton's drought adaptation.

Although the *G. hirsutum* black module and *G. barbadense* purple module exhibited similar expression patterns and contained a high proportion of orthologous genes, other gene co-expression modules of *G. hirsutum* and *G. barbadense* exhibited different expression dynamics or contained few orthologous genes, indicating that *G. hirsutum* and *G. barbadense* also have species-specific osmotic stress response pathways. In addition, the expression changes of the *G. hirsutum* black and pink modules were more dramatic than those of the *G. barbadense* purple and cyan modules (Figure 4). At the hub gene level, the higher enrichment significance of GO terms in the *G. hirsutum* black module compared to *G. barbadense* purple module (Figure 9C) suggests that *G. hirsutum* may possess more robust osmotic stress response. This is consistent with previous findings that drought-tolerant cotton varieties exhibit more pronounced activation of hub transcription factors, such as *WRKY22*, which coordinates osmotic regulation and antioxidant defense[19].

Conclusion

This study used WGCNA method to identify 8 and 9 gene co-expression modules in *G. hirsutum* and *G. barbadense*, respectively. Among them, the *G. hirsutum* black module and *G. barbadense* purple module showed highly similar expression patterns under osmotic stress and contained a high proportion of orthologous genes. GO enrichment analysis indicated that the two modules are both involved in drought response-related functions such as water transport, photosynthesis regulation, metabolism, and stimulus response. Additionally, seven pairs of candidate water transport-related orthologous genes were identified in *G. hirsutum* and *G. barbadense*. Furthermore, in the two modules, the TFs families are mainly stress-responsive and hub genes are enriched in stress response-related functions. These results reveal that *G. hirsutum* and *G. barbadense* possess conserved molecular mechanisms for osmotic stress response and provide a theoretical basis for in-depth understanding of the molecular basis of cotton drought resistance. The functions of *ABCG22* and *PIPs* in cotton drought response can be further verified through cotton genetic transformation, providing direct targets for molecular breeding of drought-resistant cotton.

Data availability

All data supporting the findings of this study are included in the article. The datasets generated and/or analyzed during the current study are available in the NCBI SRA repository, PRJNA490626, https://www.ncbi.nlm.nih.gov/bioproject/PRJNA490626

Acknowledgments

Not applicable

Conflict of interest

The authors declare that they have no conflict of interest.

Author Contribution

The authors confirm contribution to the paper as follows: study conception and design, data analysis, draft manuscript preparation, manuscript revision: HL and PH. All authors have read and approved the final version of the manuscript and agree to be accountable for their contributions to the work.

Funding

Not applicable

Ethics Statement

Not applicable

Publisher's note

The views expressed in this article are solely those of the authors and do not necessarily reflect those of their institutions, the publisher, editors, or reviewers. Any product mentioned is neither guaranteed nor endorsed by the publisher.

Generative AI Statement

The authors declare that generative artificial intelligence (AI) was not used in the scientific content creation, analysis, or interpretation of this manuscript. Any AI-assisted language editing or formatting support, where applicable, was carefully reviewed and verified by the authors.

Dates

Received 09 August 2026; Revised 17 November 2026; Accepted 21 November 2026; Published online 27 November 2026

References

[1] A. Khan, D. K. Y. Tan, M. Z. Afridi, et al. Nitrogen fertility and abiotic stresses management in cotton crop: a review. Environmental Science and Pollution Research, vol. 24, no. 17, pp. 14551-14566, 2017.

[2] R. Zahoor, W. Zhao, M. Abid, et al. Potassium application regulates nitrogen metabolism and osmotic adjustment in cotton (Gossypium hirsutum L.) functional leaf under drought stress. Journal of Plant Physiology, vol. 215, pp. 30-38, 2017.

[3] W. T. Pettigrew. Physiological Consequences of Moisture Deficit Stress in Cotton. Crop Science, vol. 44, no. 4, pp. 1265-1272, 2004.

[4] P. Berry, J. Ramirez-Villegas, H. Bramley, et al. Regional impacts of climate change on agriculture and the role of adaptation. CABI, pp. 78–97, 2014.

[5] X. Wang, H. Wei, K. Wang, et al. MYB transcription factors: Acting as molecular switches to regulate different signaling pathways to modulate plant responses to drought stress. Industrial Crops and Products, vol. 226, pp. 120676, 2025.

[6] P. K. Agarwal, P. Agarwal, M. K. Reddy, et al. Role of DREB transcription factors in abiotic and biotic stress tolerance in plants. Plant cell reports, vol. 25, no. 12, pp. 1263-1274, 2006.

[7] K. Shinozaki, K. Yamaguchi-Shinozaki. Gene networks involved in drought stress response and tolerance. Journal of experimental botany, vol. 58, no. 2, pp. 221-227, 2007.

[8] F. Li, M. Li, P. Wang, et al. Regulation of cotton (Gossypium hirsutum) drought responses by mitogen-activated protein (MAP) kinase cascade-mediated phosphorylation of GhWRKY59. New Phytologist, vol. 215, no. 4, pp. 1462-1475, 2017.

[9] L. He, X. Yang, L. Wang, et al. Molecular cloning and functional characterization of a novel cotton CBL-interacting protein kinase gene (GhCIPK6) reveals its involvement in multiple abiotic stress tolerance in transgenic plants. Biochemical and Biophysical Research Communications, vol. 435, no. 2, pp. 209-215, 2013.

[10] C. Liang, Z. Meng, Z. Meng, et al. GhABF2, a bZIP transcription factor, confers drought and salinity tolerance in cotton (Gossypium hirsutum L.). Scientific reports, vol. 6, no. 1, pp. 35040, 2016.

[11] T. Chen, W. Li, X. Hu, et al. A Cotton MYB Transcription Factor, GbMYB5, is Positively Involved in Plant Adaptive Response to Drought Stress. Plant and Cell Physiology, vol. 56, no. 5, pp. 917-929, 2015.

[12] X. Chu, C. Wang, X. Chen, et al. The Cotton WRKY Gene GhWRKY41 Positively Regulates Salt and Drought Stress Tolerance in Transgenic Nicotiana benthamiana. PloS one, vol. 10, no. 11, pp. e0143022, 2015.

[13] J. Zhao, Y. Gao, Z. Zhang, et al. A receptor-like kinase gene (GbRLK) from Gossypium barbadense enhances salinity and drought-stress tolerance in Arabidopsis. BMC plant biology, vol. 13, no. 1, pp. 110, 2013.

[14] J. Zheng, Z. Zhang, Y. Liang, et al. Whole Transcriptome Sequencing Reveals Drought Resistance-Related Genes in Upland Cotton. Genes (Basel), vol. 13, no. 7, pp. 1159, 2022.

[15] H. M. Li, S. D. Liu, C. W. Ge, et al. Analysis of Drought Tolerance and Associated Traits in Upland Cotton at the Seedling Stage. International Journal of Molecular Sciences, vol. 20, no. 16, pp. 3888, 2019.

[16] B. Han, F. Wang, Z. Liu, et al. Transcriptome and metabolome profiling of interspecific CSSLs reveals general and specific mechanisms of drought resistance in cotton. Theoretical and Applied Genetics, vol. 135, no. 10, pp. 3375-3391, 2022.

[17] F. Sun, J. Wang, S. Geng, et al. Comprehensive transcriptomic and metabolomic analysis revealed drought tolerance regulatory pathways in upland cotton. Frontiers in plant science, vol. 16, pp. 1571944, 2025.

[18] T. Huan, X. Zhang, M. Lv, et al. Phylogeny-based comparative analysis of gene expression modulation upon drought stress across three cotton diploids. vol. 25, pp. 31, 2025.

[19] H. Zhang, W. Zhang, Y. Tang, et al. Mining Key Drought-Resistant Genes of Upland Cotton Based on RNA-Seq and WGCNA Analysis. Plants (Basel), vol. 14, no. 10, pp. 1407, 2025.

[20] Y. Hu, J. Chen, L. Fang, et al. Gossypium barbadense and Gossypium hirsutum genomes provide insights into the origin and evolution of allotetraploid cotton. Nature genetics, vol. 51, no. 4, pp. 739-748, 2019.

[21] V. Hejnák, Ö. Tatar, G. D. Atasoy, et al. Growth and photosynthesis of Upland and Pima cotton: response to drought and heat stress. Plant, Soil and Environment, vol. 61, no. 11, pp. 507-514, 2015.

[22] M. M.-U. Hasan, F. Ma, Z. H. Prodhan, et al. Molecular and Physio-Biochemical Characterization of Cotton Species for Assessing Drought Stress Tolerance. International Journal of Molecular Sciences, vol. 19, no. 9, pp. 2636, 2018.

[23] S. Chen, Y. Zhou, Y. Chen, et al. fastp: an ultra-fast all-in-one FASTQ preprocessor. Bioinformatics, vol. 34, no. 17, pp. i884-i890, 2018.

[24] D. Kim, B. Langmead, S. L. Salzberg. HISAT: a fast spliced aligner with low memory requirements. Nature methods, vol. 12, no. 4, pp. 357-360, 2015.

[25] P. Danecek, J. K. Bonfield, J. Liddle, et al. Twelve years of SAMtools and BCFtools. GigaScience, vol. 10, no. 2, 2021.

[26] M. Pertea, G. M. Pertea, C. M. Antonescu, et al. StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. Nature biotechnology, vol. 33, no. 3, pp. 290-295, 2015.

[27] L. Wei, J. Jiao, S. Shi, et al. Metabolomic-transcriptomic analysis unravels flavonoids accumulation mechanism in 'Tainong 1' mango pulp under enhanced UV-B radiation. Horticultural Plant Journal, vol. 12, no. 3, pp. 624-638, 2026.

[28] P. Langfelder, S. Horvath. WGCNA: an R package for weighted correlation network analysis. Bmc Bioinformatics, vol. 9, no. 1, pp. 559, 2008.

[29] B. Buchfink, K. Reuter, H. G. Drost. Sensitive protein alignments at tree-of-life scale using DIAMOND. Nature methods, vol. 18, pp. 366-368, 2021.

[30] C. Chen, H. Chen, Y. Zhang, et al. TBtools - an integrative toolkit developed for interactive analyses of big biological data. Molecular plant, vol. 13, no. 8, pp. 1194-1202, 2020.

[31] N. Škunca, T. Šmuc, F. Supek. REViGO: Redundancy Elimination and Visualization of Gene Ontology Term Lists. Proceedings of the Adriatic Meeting on Computational Solutions in the Life Sciences, 2009.

[32] J. Jin, F. Tian, D.-C. Yang, et al. PlantTFDB 4.0: toward a central hub for transcription factors and regulatory interactions in plants. Nucleic acids research, vol. 45, no. D1, pp. D1040-D1045, 2017.

[33] J. Wang, Z. Xu. Plant response to and recovery from drought. Current Biology, vol. 36, no. 8, pp. r343-r362, 2026.

[34] T. Mahmood, S. He, D. Zhu, et al. Transcriptomic insights into the stress signaling and drought tolerance mechanisms in sea-island cotton (Gossypium barbadense). Environmental and Experimental Botany, vol. 228, pp. 106048, 2024.

[35] T. Kuromori, E. Sugimoto, K. Shinozaki. Arabidopsis mutants of AtABCG22, an ABC transporter gene, increase water transpiration and drought susceptibility. the Plant Journal, vol. 67, no. 5, pp. 885-894, 2011.

[36] I. Pawłowicz, K. Masajada. Aquaporins as a link between water relations and photosynthetic pathway in abiotic stress tolerance in plants. Gene, vol. 687, pp. 166-172, 2019.

[37] Q. Chen, R. Liu, Y. Wu, et al. ERAD-related E2 and E3 enzymes modulate the drought response by regulating the stability of PIP2 aquaporins. The Plant cell, vol. 33, no. 8, pp. 2883-2898, 2021.

[38] B. Genty, J.-M. Briantais, J. B. V. Da Silva. Effects of Drought on Primary Photosynthetic Processes of Cotton Leaves. Plant physiology, vol. 83, no. 2, pp. 360-364, 1987.

[39] X.-P. Yi, Y.-L. Zhang, H.-S. Yao, et al. Rapid recovery of photosynthetic rate following soil water deficit and re-watering in cotton plants (Gossypium herbaceum L.) is related to the stability of the photosystems. Journal of plant physiology, vol. 194, pp. 23-34, 2016.

Copyright: © 2026 by the author(s). Published by Inspire Science and Tech Publisher, Pakistan. This article is an open access article distributed under Creative Commons Attribution License (CC BY 4.0), visit https:// creativecommons.org/licenses/by/4.0.