Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Genome-wide identification of the GRAS transcription factor family and regulation of metabolites under drought and salt stress in Isatis indigotica

Jiajia Chen
Qiqihar Medical University
Keyong Zhang
Qiqihar Medical University
Hui Li
Qiqihar Medical University
Zun Lv
Qiqihar Medical University
Ying Yu
Qiqihar Medical University
Dan Ha
Xinjiang Hongtao Agricultural Development Limited Liability Company
Jicheng Liu
Qiqihar Medical University
Yanshi Ma
mayanshi19900827@163.com
Qiqihar Medical University
Ming Jiang
Qiqihar Medical University

Research Article

Keywords: Isatis indigotica, GRAS family, IiSCL subfamily, salt stress, drought stress, transcriptomic and metabolomic integration

Posted Date: April 13th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-9241824/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

Background

The GRAS transcription factor family plays essential roles in plant development and adaptation to environmental stress. However, its genomic organization and regulatory mechanisms remain largely unknown in medicinal plants.

Result

Through genome-wide analysis, 39 *IiGRAS* genes were identified in *I. indigotica*. Phylogenetic, structural, and synteny analyses indicated that family expansion was driven by tandem and segmental duplications under strong purifying selection. Promoter analysis revealed stress-responsive cis-elements related to abscisic acid, jasmonic acid, and drought signaling. Integrated transcriptomic and metabolomic analyses further showed that *IiSCL* genes are co-expressed with transcription factors such as AP2, bZIP, WRKY, and HLH, and correlate with metabolites including trehalose-6-phosphate, lysophosphatidylcholine (LPC 18:1), and N-lactoyl-phenylalanine. These associations suggest *IiSCL* genes participate in coordinated transcriptional and metabolic regulation, contributing to osmotic balance, membrane stability, and oxidative defense under stress.

Conclusion

This study establishes a molecular framework for understanding SCL-mediated regulatory networks in *I. indigotica* and identifies potential genetic targets for improving stress tolerance, thereby supporting the sustainable cultivation of medicinal plants.

Introduction

Isatis indigotica Fort. (I. indigotica), a perennial herb from the *genus Isatis* within the *Brassicaceae family*, is a significant traditional medicinal and food homology plant in China. Its dried roots, known as “Ban-lan-gen”, and its leaves, referred to as “Daqing Ye” are widely used in traditional Chinese medicine for their properties of clearing heat, detoxifying, cooling blood, and alleviating sore throat. They are applied in treating viral and inflammatory diseases and modulating immune responses [1–3]. *I. indigotica* is mainly grown in the northern, eastern, and northwestern areas of China, and has also been introduced and cultivated in various countries throughout East Asia and Europe. Documented initially in the Shennong Bencaojing, *I. indigotica* has been used in Chinese medicine for over two thousand years, emphasizing its long-standing medicinal value [4]. Recent years have seen increasing demand for the prevention and treatment of influenza and emerging viral diseases, which has boosted sustained growth in the market for Banlangen-based traditional Chinese medicines and led to a significant expansion in *I. indigotica* cultivation. This increasing trend demonstrates trust in traditional Chinese medicine and

underscores the antiviral capabilities of *I. indigotica*. Numerous studies have demonstrated that key bioactive constituents in *I. indigotica*, such as clemastanin B epigoitrin, phenylpropanoids, and their complexes with alkaloids and organic acids exert potent inhibitory effects against influenza virus strain FM1 [5]. However, similar to many medicinal plants, the cultivation and industrial development of *I. indigotica* face multiple challenges, including germplasm degradation, limited stress tolerance, variability in medicinal quality, and inconsistent yields, all of which severely hinder the standardization and sustainable production of authentic “Ban-lan-gen” herbal materials [6].

Abiotic stress is a major environmental factor limiting plant growth, development, and distribution [7]. With the increasing demand for medicinal herbs, the cultivation area of *I. indigotica* has expanded significantly. However, its growth is frequently subjected to abiotic stresses, leading to substantial fluctuations in both yield and quality of the herbal products [8]. Among these stresses, drought and salt stresses have emerged as key limiting factors for stable production in major authentic growing regions of *I. indigotica* [9]. This species is predominantly cultivated in semi-arid and salt-affected soils, where water deficit and excessive accumulation of salt ions induce osmotic stress, ionic toxicity, and oxidative damage, thereby impairing photosynthetic efficiency, nutrient uptake, and the biosynthesis and accumulation of secondary metabolites [10, 11]. Under drought and salt stress, *I. indigotica* activates complex physiological defenses, particularly ROS scavenging and the upregulation of antioxidant systems, which form a key regulatory mechanism [12]. Saline-alkali-induced ionic stress not only suppresses plant growth but also disrupts the homeostasis of secondary metabolic pathways, consequently altering the profile of pharmacologically active compounds. Notably, alkaloids such as indiforine A and B have been isolated from *I. indigotica* leaves and exhibit significant neuroprotective activity, suggesting that these secondary metabolites may play a regulatory role in the plant’s adaptation to environmental challenges [13]. A systematic investigation of the molecular response mechanisms underlying drought and salt stress tolerance in *I. indigotica* holds considerable theoretical value and practical implications for elucidating its stress adaptation strategies, optimizing cultivation practices, and enabling standardized production of high-quality medicinal materials.

The GRAS gene family, named after GAI (gibberellic-acid insensitive), RGA (repressor of gai-3) and SCR (scarecrow), comprises plant-specific transcription factors first identified for roles in gibberellin signaling and root development [14]. Proteins belonging to this family typically feature a highly conserved C-terminal GRAS domain that mediates protein-protein interactions and participates in intracellular signaling, whereas the N-terminal region exhibits substantial sequence divergence, potentially contributing to functional diversification [15]. The GRAS family can be phylogenetically classified into several distinct subfamilies, including DELLA, SCR, SHR, LAS, PAT1, and SCL. The *SCL* subfamily is central to abiotic stress responses and participates in abscisic acid (ABA), drought, and salt-related signaling pathways [16]. Advances in genomics and transcriptomics have enabled systematic characterization of the *GRAS* family across model and crop species, revealing both functional diversity and evolutionary conservation. Genome-wide analysis identified 188 *TaGRAS* genes in wheat, classified into 12 phylogenetic subfamilies. Comparative genomic studies reveal that tandem and segmental duplications have predominantly driven the expansion of the *TaGRAS* gene family in wheat, a process

likely contributing to its adaptability under diverse environmental conditions [17, 18]. Functional analyses across multiple crops confirm that GRAS transcription factors regulate growth and stress responses. Overexpression of *GmGRAS37* enhances drought and salt tolerance in soybean, emphasizing its function in abiotic stress responses [19, 20]. In cotton, silencing of *GhGRAS55* accelerates germination indicating its potential involvement in regulating early developmental traits [21]. GRAS genes of rapeseed exhibit elevated expression during root formation and under stress exposure, implying their role in root architecture and adaptation [22]. Moreover, gene family expansion in lychee through tandem and segmental duplication has led to tissue-specific expression divergence, suggesting functional specialization in development and metabolism [23]. Members of the *GRAS* family in pineapple are also implicated in floral organ differentiation, fruit ripening, and cold stress responses, reflecting the evolutionary diversification of this family across plant species [24]. These findings show that *GRAS* genes regulate multiple aspects of plant growth, development, and stress adaptation, especially under drought and salinity. Despite these advances, *GRAS* family studies in medicinal plants remain scarce, particularly regarding their roles in abiotic stress responses. Therefore, comprehensive identification of the *GRAS* gene family in *I. indigotica* and detailed analysis of their expression profiles under drought and salt stress conditions are essential for elucidating the molecular basis of stress tolerance in medicinal plants.

We systematically identified and characterized GRAS transcription factors in *I. indigotica* at the genome-wide level. Thirty-nine *IiGRAS* genes were identified and classified, providing the first comprehensive overview of this family in *I. indigotica.* Particular attention was given to the *IiSCL* subfamily, whose expression patterns and regulatory roles were examined under drought and salt stress. Integration of transcriptomic and non-targeted metabolomic data revealed a co-expression network connecting *IiSCL* genes, transcription factors, and stress-related metabolites. This integration uncovered coordinated transcriptional and metabolic regulation during stress adaptation. These findings elucidate the molecular basis of stress resilience in *I. indigotica* and highlight genetic targets for improving stress tolerance in medicinal plants, supporting the sustainable cultivation and molecular breeding of *I. indigotica*.

Materials and Methods

Plant Materials

The seeds of *Isatis indigotica Fort.* (*I. indigotica*) employed in this research were obtained from the medicinal plant garden of Qiqihar Medical University in Heilongjiang Province, China. Their identity was verified by Associate Researcher Hui Li of the Astragalus Industry Research Institute. A voucher specimen (No. QMU-Ii-2026-007) has been archived in the public herbarium of the School of Pharmacy at the same institution (Qiqihar, Heilongjiang 161006, China).Experimental plants were grown in a controlled-environment growth chamber at Qiqihar Medical University under a 16-h light (24°C)/8-h dark (22°C) photoperiod, with relative humidity maintained at 55% and a photosynthetic photon flux density of 10,000 Lux. After 60 days of cultivation, plants were subjected to salt and drought stress treatments. Salt

stress was imposed by irrigation with 200 mM NaCl solution to simulate salt stress conditions, while drought stress was simulated using 20% (w/v) polyethylene glycol (PEG-6000) solution. Control plants were irrigated with an equivalent volume of distilled water. Sampling was conducted at 0 h (control), 24 h, and 48 h post-treatment. Throughout the experimental period, environmental conditions in the growth chamber were kept constant to minimize non-treatment-related variability. Leaf tissues were collected as the experimental material. At each sampling time point, biological replicates were obtained from three independently grown plants exhibiting uniform growth status. Collected samples were immediately labeled, flash-frozen in liquid nitrogen, and stored at -80°C until further analysis.

Identification and Characterization of the *IiGRAS* Gene Family and Subcellular Localization Prediction

The genome assembly and corresponding annotation files of *I. indigotica* used in this study were retrieved from Figshare (https://doi.org/10.6084/m9.figshare.8940383.v1) [25]. All protein sequences were extracted from the genome using TBtools-II (v2.263) and subsequently submitted to the Plant Transcription Factor Database (PlantTFDB, http://planttfdb.ga-lab.org/index.php) for transcription factor (TF) classification [26]. Candidate GRAS family members were identified based on homology to annotated TFs in *Arabidopsis thaliana* (*A. thaliana*), with functional annotations guiding the selection of best-hit sequences. To ensure sequence reliability, conserved domain analysis was performed using both the CD-Search tool (https://www.ncbi.nlm.nih.gov/Structure/cdd/wpsb.cgi) and the Pfam database (accession: PF03514). Sequences lacking the characteristic GRAS domain or exhibiting redundancy were excluded. The remaining high-confidence sequences were designated as putative IiGRAS gene family members [27]. Basic physicochemical properties of these proteins, including amino acid length, molecular weight, theoretical isoelectric point (pI), instability index, aliphatic index, and grand average of hydropathicity (GRAVY) were calculated using TBtools-II. Subcellular localization predictions were conducted using the WoLF PSORT platform (https://wolfpsort.hgc.jp/), which integrates multiple sorting signals to infer protein localization.

Chromosomal Localization and Gene Duplication Analysis

Chromosomal position information for the *IiGRAS* genes was extracted from the genome annotation file, including chromosome number, chromosome length, gene orientation, start position, and end position. The chromosomal distribution of *IiGRAS* genes were visualized using TBtools-II, with consistent chromosome nomenclature applied according to the genomic annotation. To investigate the potential mechanisms underlying gene family expansion, gene duplication events were analyzed using the “One Step MCScanX-Super Fast” module in TBtools-II under a stringent E-value threshold of $\leq 1\times10^{-10}$. Tandem, segmental, and other types of duplicated gene pairs were identified, and their genomic positions were integrated into a synteny map to visually represent the duplication patterns and homologous relationships among *IiGRAS* genes.

Phylogenetic Analysis and Interspecies Collinearity Analysis

GRAS protein sequences from *A. thaliana* and Brassica oleracea (B. oleracea) were retrieved from the PlantTFDB database (https://planttfdb.gao-lab.org/). Multiple sequence alignment of GRAS amino acid sequences from *I. indigotica*, *A. thaliana*, and *B. oleracea* was performed using MEGA-X. A phylogenetic tree was constructed using the Neighbor-Joining (NJ) method with 1,000 bootstrap replicates and all other parameters were set to their default values. The resulting tree was visualized and annotated using the Evolview platform (https://www.evolgenius.info/evolview). Genome sequences and annotation files for *A. thaliana* and *B. oleracea* were downloaded from Ensembl Plants (http://plants.ensembl.org). Using these data in combination with the previously obtained genome and annotation resources for Isatis indigotica, collinearity analysis of the *GRAS* gene family was conducted via the MCScanX toolkit implemented in TBtools-II. Syntenic relationships between *I. indigotica* and the two reference species were visualized using the Dual Synteny Plot function, enabling the identification of evolutionary conservation and chromosomal segmental duplication events within the *GRAS* gene family.

Gene Structure, Conserved Motif, and Cis-Regulatory Element Analysis

The gene structures of the *IiGRAS* family members, including untranslated regions (UTRs), coding sequences (CDSs), introns, and exons, were extracted and visualized using TBtools-II based on the *I. indigotica* genome and its annotation file. To confirm the presence of the characteristic *GRAS* domain, protein sequences were subjected to conserved domain analysis via the Batch CD-Search tool (https://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi). Conserved motifs were predicted using the MEME Suite online platform (https://meme-suite.org/meme/tools/meme) with the number of motifs set to six and all other parameters maintained at default settings. For promoter analysis, 2000 bp upstream sequences from the transcription start site of each *IiGRAS* gene were retrieved as putative promoter regions. These sequences were analyzed using the PlantCARE database (http://bioinformatics.psb.ugent.be/webtools/plantcare/) to identify putative cis-regulatory elements. The resulting element distribution was visualized as a heat map using TBtools-II, facilitating the comparative interpretation of potential transcriptional regulatory differences among *IiGRAS* genes.

Expression Profiling and qRT-PCR Validation under Stress

FPKM expression values of the *IiGRAS* gene family were extracted from RNA sequencing (RNA-seq) data obtained from *I. indigotica* leaves subjected to salt and drought stress conditions. The expression profiles were visualized using TBtools-II, with all FPKM values subjected to row-wise normalization to facilitate cross-gene and cross-sample comparison. To validate the reliability of the RNA-seq results, quantitative real-time PCR (qRT-PCR) was performed to assess the expression patterns of selected *GRAS* genes in *I. indigotica* leaves under both salt and drought stress. The experiment included five treatment groups (control, 24-hour salt stress, 48-hour salt stress, 24-hour drought stress, and 48-hour drought stress). Total RNA was extracted from leaf tissues following stress treatments, and first-strand cDNA was synthesized via reverse transcription for use as the PCR template. qRT-PCR was conducted using the SYBR Green method on a real-time PCR system, with the following thermal cycling parameters: initial denaturation at 95°C for 2 min, followed by 40 cycles of 95°C for 10 s (denaturation) and 60°C for 30 s

(annealing and extension). The IiMUB gene was selected as the reference gene in this study. Based on transcriptome analysis results, six significantly differentially expressed *GRAS* genes were identified under salt stress treatment and four under drought stress treatment, which were subsequently validated using qRT-PCR. Three technical replicates were performed for each sample to guarantee result reliability. All primers were synthesized by Sangon Biotech (Shanghai) Co., Ltd., and the corresponding primer sequences (Supplementary Table 1) are provided .

Non-Targeted Metabolomic Analysis under Stress

To systematically characterize the metabolic responses of *I. indigotica* to salinity and drought stress, a non-targeted metabolomic analysis was performed on leaf samples from five treatment groups, control (normal growth), 24-hour salt stress, 48-hour salt stress, 24-hour drought stress, and 48-hour drought stress. Each treatment included three biological replicates to ensure experimental reliability and statistical robustness. Metabolite profiling was conducted using the non-targeted metabolomics platform at Metware Biotechnology Co., Ltd. (Wuhan), based on liquid chromatography-tandem mass spectrometry (LC-MS/MS) technology, following an optimized standard operating procedure. The workflow encompassed sample extraction, chromatographic separation, and mass spectrometric detection. High-resolution MS data were acquired and subjected to qualitative identification and relative quantification of metabolites through comparison with an in-house database. The relative abundance of each metabolite was determined by its chromatographic peak area, enabling a comprehensive assessment of dynamic metabolic changes in *I. indigotica* across different stress conditions.

Statistical Analysis

Data processing and basic statistical analysis were performed using Microsoft Excel. GraphPad Prism (version 9.0) was employed for significance testing and correlation analysis. Intergroup differences were assessed by one-way analysis of variance (ANOVA), followed by Duncan's multiple range test for post-hoc comparisons at a significance level of $p < 0.05$. The co-expression network was constructed and visualized using Cytoscape software (version 3.9.1).

Results

Identification and Characterization of the *IiGRAS* Gene Family

Following domain integrity validation and phylogenetic-based redundancy elimination, a total of 39 *IiGRAS* genes were identified in this study (Supplementary Table 2) and designated as *IiGRAS01* to *IiGRAS*39 according to their chromosomal positions. The corresponding protein sequences encoded by these genes are summarized in Supplementary Table 3. The encoded proteins vary in length from 252 to 781 amino acids, with predicted molecular weights ranging from 28.6 to 86.6 kDa. The theoretical isoelectric points (*pI*) span a broad range of 4.69 to 8.28, reflecting considerable diversity in the physicochemical properties among family members. Instability index analysis indicates that the majority of IiGRAS proteins (35 out of 39, 89.7%) are classified as unstable, which may imply roles in dynamic

regulatory processes. All members exhibit high aliphatic indices (67.87-96.00), suggesting potential thermal stability. Furthermore, the grand average of hydropathy (GRAVY) values are uniformly negative (ranging from -0.533 to -0.029), confirming the hydrophilic nature of all IiGRAS proteins.

The subcellular localization predictions (Supplementary Table 4) indicate that the IiGRAS proteins are predominantly localized in the nucleus (23 members, 58.97%), with additional distributions in the cytoplasm (9 members, 23.08%) and chloroplasts (7 members, 17.95%). The nuclear enrichment of most IiGRAS proteins is highly consistent with the established role of GRAS family proteins as transcriptional regulators involved in nuclear gene expression control. The presence of a subset of proteins in the cytoplasm and chloroplasts suggests potential involvement in diverse biological processes, including cell signal transduction and the regulation of photosynthetic metabolism.

Chromosomal Localization and Replication Event Analysis

The chromosomal distribution of the 39 identified *IiGRAS* genes was uneven across 10 chromosomes (Figure 1A). Chr03 and Chr07 harbored the highest gene densities, with 7 and 6 members, respectively. Several genes were arranged in clusters, such as *IiGRAS25* and *IiGRAS04*, *IiGRAS26* and *IiGRAS12* on Chr03, indicating that tandem duplication events likely played a significant role in the expansion of the *IiGRAS* gene family. Synteny analysis revealed 15 duplicated gene pairs (Figure 1B). Further Ka/Ks analysis (Supplementary Table 5) showed that all duplicated pairs had Ka/Ks ratios less than 1, ranging from 0.148 to 0.496, suggesting strong purifying selection during evolution and high conservation of both amino acid sequences and functional domains. Collectively, these duplication events including tandem and segmental duplications have contributed substantially to the expansion and functional preservation of the *IiGRAS* gene family in *I. indigotica*.

Phylogenetic Analysis and Interspecies Collinearity Assessment

A phylogenetic tree was constructed based on GRAS protein sequences from *I. indigotica*, the model plant *A. thaliana*, and the closely related species *B. oleracea*. (Figure 2A; Supplementary Table 6), revealing six distinct evolutionary clades (Groups I-VI). The *IiGRAS* family members were distributed across all clades and clustered closely with their orthologs in *A. thaliana* and *B. oleracea*, indicating strong evolutionary conservation of the *GRAS* gene family within the *Brassicaceae*. Furthermore, interspecies collinearity analysis identified numerous conserved homologous gene pairs among *I. indigotica*, *A. thaliana*, and *B. oleracea* (Figure 2B). The highest degree of synteny was observed between *I. indigotica* and *B. oleracea*, consistent with their closer phylogenetic relationship, thereby reinforcing the notion that the structure and function of the *GRAS* gene family have been largely conserved during evolution in cruciferous plants.

Gene Structure and Motif Analysis of the *IiGRAS* Gene Family

Based on phylogenetic relationships, conserved domain architecture, and gene structure visualization, the 39 *IiGRAS* genes from *I. indigotica* were classified into multiple evolutionary subfamilies, including

SCL, PAT1, and DELLA (Figure 3A), indicating substantial functional divergence and evolutionary diversification within this gene family. Conserved motif analysis revealed that all IiGRAS proteins harbor the characteristic GRAS domain (Figure 3B, 3E). Notably, Motifs 1-6 are highly conserved across subfamilies and exhibit consistent amino acid sequence features, suggesting their critical roles in maintaining protein structural integrity and mediating functional activity.

Gene structure analysis uncovered marked variation in exon-intron organization among different subfamilies (Figure 3C, 3D). Most *IiGRAS* genes possess long and structurally complex introns, while exon numbers vary considerably across subgroups. These structural characteristics may contribute to the evolutionary adaptation of *IiGRAS* genes by enabling intricate transcriptional regulation and facilitating functional specialization.

Cis-Acting Regulatory Element Analysis of the *IiGRAS* Gene Family

A systematic analysis of cis-acting elements in the promoter regions of the 39 *IiGRAS* genes in *I. indigotica* (Figure 4, Supplementary Table 7) revealed that all promoters harbor multiple core regulatory elements associated with stress responses, hormone signaling, and developmental processes. The ubiquitous presence of fundamental promoter motifs such as the TATA-box and CAAT-box indicates a conserved basal transcriptional machinery maintained throughout the evolution of the *IiGRAS* gene family. Further inspection identified abscisic acid-responsive elements (ABRE), methyl jasmonate-responsive motifs (CGTCA-motif), and drought-responsive elements (MBS) in the majority of promoters, suggesting that *IiGRAS* genes are likely involved in abscisic acid and jasmonic acid signaling pathways as well as in responses to drought stress. Additionally, several promoters contain gibberellin-responsive elements (P-box) and light-responsive elements (GATA-motif), which are implicated in developmental regulation, indicating potential roles of *IiGRAS* genes in integrating light signals and hormonal cues to coordinate plant growth and development.

Expression Profiling and qRT-PCR Validation of SCL Subfamily Genes under Drought and Salt Stress

Previous studies have systematically identified and classified the *IiGRAS* gene family, revealing that the majority of its members belong to the *SCL* subfamily (Supplementary Table 8). Given the dominance of this subfamily and its potential role in stress response regulation, this study integrated transcriptome sequencing data with quantitative real-time PCR (qRT-PCR) to systematically analyze the expression patterns of *IiSCL* genes under drought and salt stress. Under drought stress, multiple *IiSCL* genes exhibited significant differential expression at 24 h and 48 h post-treatment compared to the control (CK) (Figure 5A; Supplementary Table 9-1). *IiSCL13*, *IiSCL27*, and *IiSCL31* were markedly upregulated, with several genes reaching peak expression at 24 h and sustaining high transcript levels at 48 h. In contrast, *IiSCL33* and *IiSCL34* showed significant down-regulation throughout the treatment period. *IiSCL3* displayed a transient induction pattern, with strong up-regulation at 24 h followed by a decline to near-control levels at 48 h. Under salt stress, *IiSCL* genes also demonstrated distinct stress-responsive expression dynamics (Figure 5B; Supplementary Table 9-2). *IiSCL6* and *IiSCL30* exhibited progressive upregulation, peaking at 48 h, whereas *IiSCL5* and *IiSCL14* showed gradual downregulation, with

significantly reduced expression levels at 48 h relative to both the control and the 24 h time point. To validate the reliability of the RNA-seq data, a subset of differentially expressed genes was subjected to qRT-PCR analysis (Figure 5C). The results confirmed a high degree of concordance between qRT-PCR and transcriptomic data under both drought and salt stress, supporting the accuracy and reproducibility of the transcriptome profiling.

Co-expression Network Construction

Correlation Analysis between *IiSCL* Genes and TFs

To investigate the potential role of the *IiSCL* subfamily in the transcriptional regulatory network of *I. indigotica* within the GRAS gene family, a genome-wide correlation analysis of TFs expression profiles were performed based on RNA-seq data using a custom Python script. Gene pairs with a Pearson correlation coefficient exceeding 0.94 ($r > 0.94$) were considered significantly co-expressed and used to construct the *IiSCL-TFs* co-expression network (Supplementary Table 10).

Thirteen *IiSCL* genes exhibited significant positive correlations with members of diverse TF families under drought stress (Figure 6). *IiSCL*14 showed strong co-expression with genes encoding G-patch, OTU, R3H, and P450 domain-containing proteins, suggesting its potential centrality in the drought-responsive signaling network. *IiSCL31* displayed the highest correlation with C2 domain-containing proteins, while *IiSCL27* was significantly co-expressed with E1-E2-ATR and AP2 family members. *IiSCL3* showed high correlation with WRKY transcription factors. These findings suggest that distinct *IiSCL* genes may contribute to drought adaptation through involvement in multiple biological processes, including signal transduction, energy metabolism regulation , and activation of stress responses, thereby collectively enhancing the abiotic stress resilience of *I. indigotica*.

Under salt stress, 11 *IiSCL* genes were significantly co-expressed with various TF families, particularly HLH and WRKY (Figure 7). *IiSCL3* exhibited a strong positive correlation with HLH family members, whereas *IiSCL6* showed tight association with DnaJ-domain proteins. WRKY transcription factors were frequently linked to multiple *IiSCL* genes ($r > 0.94$), underscoring their coordinated regulatory role. In light of TPR1-mediated protein homeostasis regulation, these *IiSCL* members may integrate phosphorylation signaling via Pkinase-Tyr, WRKY/HLH-regulated stress response pathways, and DnaJ/TPR1-involved molecular chaperone systems to form a core regulatory network underlying salt tolerance in *I. indigotica*. This highlights the critical role of the *IiSCL* subfamily in orchestrating stress adaptation and signal transduction in this species.

Correlation Analysis between *IiSCL* Genes and Metabolites

To further investigate the metabolic regulatory roles of the *IiSCL* subfamily genes under salt and drought stress in *I. indigotica*, a correlation matrix linking *IiSCL* gene expression with upregulated metabolites was constructed by integrating RNA-seq and non-targeted metabolomics data. The analysis revealed that, under both stress conditions, the majority of *IiSCL* genes exhibited significant positive correlations

with specific metabolites (r > 0.7; Supplementary Table 11), suggesting their potential involvement in plant stress responses through the modulation of key metabolic pathways.

Under drought stress, *IiSCL14* showed the strongest correlation with trehalose-6-phosphate (r = 0.93), indicating its likely role in maintaining osmotic homeostasis and carbon metabolic balance (Figure 8A). *IiSCL30*, *IiSCL3*, and *IiSCL34* were significantly correlated with 17(S)-hydroperoxydocosahexaenoic acid (17(S)-HpDHA), a lipid peroxidation intermediate (r > 0.85), implying their potential participation in mitigating oxidative stress and regulating lipid-mediated signaling. *IiSCL1* was strongly associated with N-lactoyl-phenylalanine and multiple phenylpropanoid derivatives, including (2E)-3-(3,4-dimethoxyphenyl)prop-2-enic acid and gramine (r > 0.7), suggesting that it may enhance drought tolerance by modulating the biosynthesis of aromatic amino acid-derived compounds, promoting the accumulation of antioxidant or signaling molecules.

IiSCL13 displayed the most prominent correlation with valepotriate (r = 0.94), implying a potential role in enhancing stress resilience through the regulation of secondary metabolite biosynthesis under salt stress (Figure 8B). *IiSCL3*3, *IiSCL34*, and *IiSCL1* were highly correlated with lysophosphatidylcholine (LPC 18:1) and lysophosphatidylinositol (LPI 18:3), indicating their possible involvement in membrane lipid remodeling and signal transduction via lipid mediators, thus contributing to ion homeostasis and cellular membrane integrity.

Collectively, the robust positive correlations between *IiSCL* subfamily members and diverse metabolites highlight the functional significance of this gene family in coordinating complex physiological adaptations in *I. indigotica* under abiotic stress, potentially through the integrated regulation of carbohydrate metabolism, lipid metabolism, and secondary metabolic networks.

Integrated Analysis of Differentially Expressed *IiSCL* Genes with TFs and Metabolites

To systematically elucidate the comprehensive regulatory mechanisms underlying the response of the *IiSCL* subfamily genes in *I. indigotica* to drought and salt stress, an integrative correlation analysis was performed using differentially expressed genes (DEGs), TFs, and significantly altered metabolites (Supplementary Table 12). The results demonstrate that *IiSCL* genes form tightly coordinated regulatory modules with multiple TFs and metabolites during stress responses, constituting a complex transcription, metabolism interaction network that plays a central role in mediating abiotic stress adaptation in *I. indigotica*.

Under drought stress conditions, *IiSCL14* exhibited highly significant positive correlations with several TFs (r > 0.98), including the PHD associated with epigenetic regulation (r = 0.992), the bZIP-1 involved in DNA binding and transcriptional activation (r = 0.989), and the ACT linked to cytoskeletal dynamics (r = 0.988) (Figure 9A). These associations suggest that *IiSCL14* may contribute to drought signal transduction and adaptive responses through the modulation of chromatin remodeling, transcriptional regulation, and cytoskeletal reorganization. *IiSCL14* showed a strong positive correlation with trehalose-6-phosphate (r = 0.930), underscoring its potential role in maintaining cellular osmotic homeostasis and

balancing carbon metabolism. *IiSCL1* was significantly co-expressed with the CTP-synth-N domain-containing TF (r = 0.988) and positively correlated with N-lactoyl-phenylalanine (r = 0.892) and the membrane lipid metabolite LPC 18:1 (r = 0.762), indicating its involvement in nucleotide biosynthesis and amino acid derivative metabolism, which may enhance drought tolerance by promoting membrane integrity and signal transduction.

Under salt stress conditions (Figure 9B), *IiSCL6* displayed significant positive correlations with stress-responsive TFs, including the molecular chaperone DnaJ (r = 0.979), the proteasome component Mov34 (r = 0.976), and the NAM domain implicated in hormone signaling (r = 0.969), suggesting a role for *IiSCL6* in regulating protein folding, degradation, and hormonal responses under salt stress. *IiSCL6* was significantly associated with the lipid metabolite LPC 18:1 (r = 0.809) and the benzophenanthridine alkaloid stylopine (r = 0.789), highlighting its potential function in membrane lipid remodeling and the accumulation of defense-related secondary metabolites. *IiSCL13* exhibited high correlation with ubiquitination-related TFs (r = 0.960) and the pathogenesis-related protein thaumatin (r = 0.958), implying its involvement in targeted protein degradation and stress signal transduction pathways.

These findings indicate that *IiSCL* subfamily genes serve as key regulatory hubs in the transcriptional and metabolic networks governing stress responses in *I. indigotica*. Through their integration with diverse TFs and metabolites, *IiSCL* genes coordinate critical physiological processes under adverse environmental conditions, providing a robust theoretical foundation for further investigation into their molecular regulatory mechanisms.

Discussion

This study provides a genome-wide overview of the GRAS transcription factor family in *I. indigotica*, clarifying its composition, structure, and evolution. The analysis focused on the expression dynamics and metabolic regulation of the *IiSCL* subfamily under drought and salt stress. The results reveal molecular mechanisms by which *GRAS* genes contribute to abiotic stress adaptation, providing a basis for identifying stress-resistance genes and guiding molecular breeding in medicinal plants.

A total of 39 *IiGRAS* genes were identified and distributed across 10 chromosomes, exhibiting an uneven genomic distribution, with the highest gene density observed on Chr03 and Chr07. Gene duplication, particularly tandem and segmental events, largely contributed to the expansion of the *IiGRAS* family. All duplicated gene pairs had Ka/Ks ratios < 1, indicating strong purifying selection and high structural and functional conservation. This finding is consistent with previous studies on the GRAS family in species such as *Secale cereale* [28], *Glycine max* [20], and *Brassica napus* [22], suggesting a conserved pattern of gene duplication and evolutionary constraint within cruciferous plants. The most pronounced synteny was observed between *I. indigotica* and *B. oleracea*, further supporting their close phylogenetic relationship and providing compelling evidence for the evolutionary origin of the *IiGRAS* gene family [29]. Gene structure and motif analyses confirmed that all GRAS proteins contain the conserved GRAS domain. Motifs 1–6 are highly conserved across subfamilies and are likely critical for protein stability

and transcriptional regulatory activity. Marked differences in exon-intron structures among subfamilies suggest adaptive evolution in transcriptional regulation and functional diversification within the gene family. Cis-acting element analysis demonstrated that the promoters of most *IiGRAS* genes are enriched with stress-responsive elements, including abscisic acid-responsive elements (ABRE), methyl jasmonate-responsive elements (CGTCA-motif), and drought-responsive elements (MBS), indicating that these genes may mediate diverse abiotic stress responses through integration of hormonal signaling pathways. This agrees with studies on *OsGRAS23* and *AtSCL14*, which act via ABA-dependent pathways, suggesting that *IiGRAS* genes may share similar regulatory functions [30, 31].

The *IiSCL* subfamily displayed distinct temporal and stress-specific expression patterns under drought and salt stresses. Under drought stress, *IiSCL13*, *IiSCL27*, and *IiSCL31* were significantly upregulated, peaking at 24 hours post-treatment before stabilizing, suggesting their potential involvement in early signal perception and transcriptional activation. In contrast, *IiSCL3* displayed a transient upregulation followed by a return to baseline levels, indicating high sensitivity during the initial phase of stress response, a pattern consistent with the rapid induction of *OsSCL30* in rice within 12 hours of drought exposure [32]. Under salt stress, *IiSCL6* and *IiSCL30* showed sustained upregulation, likely contributing to longterm maintenance of ion homeostasis and osmotic balance; *IiSCL5* and *IiSCL14* were downregulated, implying a potential role in negative feedback regulation. Previous studies have demonstrated that *TaSCL14* in wheat is induced under photo-oxidative stress, and its silencing results in reduced photosynthetic efficiency and impaired antioxidant capacity [33], further underscoring the central role of GRAS transcription factors in oxidative stress responses. *IiSCL14* responded to both drought and salt stresses and correlated with antioxidant metabolites, indicating possible dual roles in stress tolerance and metabolic regulation.

Integrated transcriptomic and metabolomic analyses revealed that the *IiSCL* subfamily coordinates multiple transcriptional and metabolic responses under stress. Under drought stress, *IiSCL14*, *IiSCL30*, *IiSCL3*, and *IiSCL34* play crucial roles in antioxidant defense, osmotic adjustment, and lipid metabolism. *IiSCL14* exhibits a strong positive correlation with trehalose-6-phosphate (T6P) ($r = 0.93$) and shows co-expression with TFs such as AP2, bZIP, C2H2, and G-patch domain-containing proteins, suggesting its involvement in a potential signaling cascade of “ABA/JA-AP2/bZIP→SCL14→ T6P/sucrose metabolism genes”, thereby modulating energy allocation and osmotic homeostasis to improve drought resilience. This finding aligns with studies in poplar demonstrating that overexpression of T6P synthase PagTPS1/10 elevates T6P levels and improves drought resistance [34], further supporting the conserved role of T6P as a central metabolic signal in plant stress adaptation. *IiSCL30*, *IiSCL3*, and *IiSCL34* exhibit significant correlations with the lipid peroxidation intermediate 17(S)-HpDHA ($r > 0.85$) and show co-expression with WRKY family transcription factors, suggesting their potential involvement in enhancing antioxidant defense via the lipoxygenase (LOX)/oxylipin (JA) signaling pathway and reactive oxygen species (ROS) scavenging mechanisms [35]. *IiSCL1* is significantly associated with the aromatic amino acid derivative N-lactoyl-phenylalanine (Lac-Phe) and the signaling lipid lysophosphatidylcholine (LPC) 18:1. Lac-Phe, an N-lactoyl-amino acid formed by CNDP2-catalyzed condensation of lactate and phenylalanine, is well studied in animals but remains poorly understood in plants [36]. The LPC/PC cycle

contributes to membrane lipid remodeling and stress responses through LPCAT-mediated acyl editing, a mechanism that has been experimentally validated in *A. thaliana* and sunflower [37]. In conjunction with previous reports demonstrating that GRAS family members upregulate the LOX/JA pathway [38], it is proposed that *IiSCL1* may coordinately maintain cellular homeostasis through an integrated regulatory axis involving “aromatic amino acid-derived metabolism and nucleotide metabolism→membrane lipid remodeling (LPC/PC cycle)→oxylipin/ROS signaling”, thereby contributing to enhanced stress tolerance.

Under salt stress *IiSCL6*, *IiSCL13*, *IiSCL33*, and *IiSCL34* showed strong associations with membrane lipid metabolism, protein homeostasis, and secondary metabolic processes. *IiSCL33* and *IiSCL34* correlated positively with LPC 18:1 and LPI 18:3 and co-expressed with HLH, DnaJ, and NAM TFs, suggesting their involvement in the HLH/DnaJ/NAM→SCL33/34→phospholipid enzyme/membrane transport pathway, which may contribute to maintaining membrane integrity and ion homeostasis [39, 40]. *IiSCL6* shows strong correlation with the chaperone protein DnaJ and the proteasome subunit Mov34, indicating its potential role in the coordinated network of protein folding and degradation, thereby supporting cellular protein quality control [41]. Overexpression of *LaSCL6* in *A. thaliana* inhibits root elongation, delays flowering, and extends lifespan, demonstrating that SCL6 plays a regulatory role in balancing vegetative growth and reproductive development [42]. *IiSCL13* is significantly co-expressed with the terpenoid metabolite valepotriate as well as ubiquitin and thaumatin-like proteins, suggesting it may serve as a molecular link between ubiquitin-mediated defense responses and secondary metabolism regulation. Its expression profile aligns with findings in *A. thaliana*, where *SCL13* participates in the jasmonic acid (JA)/salicylic acid (SA)-co-induced metabolic network [43].

In summary, the *IiSCL* subfamily in *I. indigotica* cooperates with diverse transcription factor modules under drought and salt stress to regulate multiple metabolic networks, including sugar signaling and osmotic adjustment, oxylipin and reactive oxygen species (ROS) signaling, membrane lipid remodeling (LPC/LPI), and the biosynthesis of secondary metabolites such as valepotriate and N-lactoyl-phenylalanine (Lac-Phe). This regulatory network coordinates energy balance, membrane stability, and defense signaling, forming the molecular basis for *I. indigotica* adaptation and pharmacological trait development under stress.

Conclusions

This study systematically characterized the GRAS transcription factor family in *I. indigotica*, identifying 39 *IiGRAS* genes with conserved structures and evidence of expansion under purifying selection. Members of the SCL subfamily showed distinct expression profiles under drought and salt stresses, suggesting their roles in adaptive signaling. Integrated transcriptomic and metabolomic analyses revealed that *IiSCL* genes co-regulate with transcription factors such as AP2, bZIP, and WRKY, and associate with metabolites including trehalose-6-phosphate and lysophosphatidylcholine. The results suggest that *IiSCL* genes enhance stress tolerance by linking transcriptional and metabolic regulation. These findings advance understanding of SCL-mediated stress responses in medicinal plants and

provide a molecular foundation for breeding stress-resilient *I. indigotica* cultivars to support sustainable production.

Declarations

Ethical approval

This publication does not report any research involving human participants or animals conducted by the authors.

Authors' contributions

All authors contributed to the conception and design of the study. Material preparation, data collection and analysis were performed by HL and JC. ZL, YY designed the experiments and performed the data analysis. The manuscript was drafted by MJ, HL, YM. Formal analysis was performed by KY, JL, DH. Supervision, Project administration, Methodology were conducted by MJ and KZ. All authors comment on previous versions of the manuscript. All authors read and approved the manuscript.

Funding

This work was supported by Natural Science Foundation of Heilongjiang Province of China (Grant No. ZL2024H018), Program for Young Talents of Basic Research in Universities of Heilongjiang Province (no. YQJH2024283) and Qiqihar Academy of Medical Sciences project (no. QMSI2023B-17).

Data Availability Statement

RNA sequencing (RNA-seq) data from the three *I. indigotica* leaf samples were acquired and uploaded to the NCBI Sequence Read Archive (accession no. SUB11718505) and to BioProject (ID: PRJNA854335).

Ethics approval and consent to participate

The seeds of *I. indigotica* used in the experiment were authenticated by Associate Professor Hui Li from Qiqihar Medical University.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Acknowledgments

During the preparation of this work the authors used Google AI Studio in order to improve language. After using this tool, the authors reviewed and edited the content as needed and take full responsibility for the content of the publication.

References

1. Özbek, Ö. Breeding Strategies of Isatis indigotica (Chinese Woad). In Biodiversity and Genetic Improvement of Medicinal and Aromatic Plants I; Springer: 2025; pp. 137-158. https://doi.org/10.1007/978-3-031-81846-2_6
2. Chen, Q.; Lan, H.-Y.; Peng, W.; Rahman, K.; Liu, Q.-C.; Luan, X.; Zhang, H. Isatis indigotica: a review of phytochemistry, pharmacological activities and clinical applications. Journal of Pharmacy and Pharmacology 2021, 73, 1137-1150. https://doi.org/10.1093/jpp/rgab014
3. Wang, T.; Wang, X.; Zhuo, Y.; Si, C.; Yang, L.; Meng, L.; Zhu, B. Antiviral activity of a polysaccharide from Radix Isatidis (Isatis indigotica Fortune) against hepatitis B virus (HBV) in vitro via activation of JAK/STAT signal pathway. Journal of Ethnopharmacology 2020, 257, 112782. https://doi.org/10.1016/j.jep.2020.112782
4. Bai, N.; Wang, W.; Li, Q.; Guo, Y.; Meng, X.; Chen, Y. Study and Application of Radix Isatidis Polysaccharide. Botanical research 2018, 7, 268-276. https://doi.org/10.34038/jpss.2019.01.004
5. Ping, X.; Weiyang, Y.; Jianwei, C.; Xiang, L. Antiviral activities against influenza virus (FM1) of bioactive fractions and representative compounds extracted from Banlangen (Radix Isatidis). Journal of Traditional Chinese Medicine 2016, 36, 369-376. https://doi.org/10.1016/s0254-6272(16)30051-6
6. Li, X.; Chen, Y.; Lai, Y.; Yang, Q.; Hu, H.; Wang, Y. Sustainable utilization of traditional Chinese medicine resources: Systematic evaluation on different production modes. Evidence-Based Complementary and Alternative Medicine 2015, 218901. https://doi.org/10.1155/2015/218901
7. Kijne, J.W. Abiotic stress and water scarcity: identifying and resolving conflicts from plant level to global level. Field Crops Research 2006, 97, 3-18. https://doi.org/10.1016/j.fcr.2005.08.011
8. Wang, Z.; Cao, B.; Du, J.; Deng, T.; Wang, R.; Chen, Y.; Li, X.; Fu, J.; Pang, J.; Yang, M. Transcriptomic and proteomic-based analysis of the mechanisms by which drought and salt stresses affect the quality of Isatidis Folium. BMC Plant Biology 2025, 25, 332. https://doi.org/10.1186/s12870-025-06309-z
9. Wang, Z.; Cao, B.; Ma, Y.; Xu, W.; Fu, J.; Zhang, Z.; Du, J.; Deng, T.; Pang, J.; Yang, M. Exploration of biophoton characteristics of fresh Isatis indigotica fort leaves under salt and drought stresses and the feasibility analysis for the quality prediction of Isatidis Folium. Frontiers in Plant Science 2025, 16, 1523636. https://doi.org/10.3389/fpls.2025.1523636
10. Arif, Y.; Singh, P.; Siddiqui, H.; Bajguz, A.; Hayat, S. Salinity induced physiological and biochemical changes in plants: An omic approach towards salt stress tolerance. Plant Physiology and Biochemistry 2020, 156, 64-77. https://doi.org/10.1016/j.plaphy.2020.08.042

11. Munns, R. Comparative physiology of salt and water stress. Plant, cell & environment 2002, 25, 239-250. https://doi.org/10.1046/j.0016-8025.2001.00808.x
12. Souri, Z.; Karimi, N.; de Oliveira, L.M. Antioxidant enzymes responses in shoots of arsenic hyperaccumulator, Isatis cappadocica Desv., under interaction of arsenate and phosphate. Environmental technology 2018, 39, 1316-1327. https://doi.org/10.1080/09593330.2017.1329349
13. Liu, S.-F.; Zhang, Y.-Y.; Zhou, L.; Lin, B.; Huang, X.-X.; Wang, X.-B.; Song, S.-J. Alkaloids with neuroprotective effects from the leaves of Isatis indigotica collected in the Anhui Province, China. Phytochemistry 2018, 149, 132-139. https://doi.org/10.1016/j.phytochem.2018.02.016
14. Waseem, M.; Nkurikiyimfura, O.; Niyitanga, S.; Jakada, B.H.; Shaheen, I.; Aslam, M.M. GRAS transcription factors emerging regulator in plants growth, development, and multiple stresses. Molecular Biology Reports 2022, 49, 9673-9685. https://doi.org/10.1007/s11033-022-07425-x
15. Tian, C.; Wan, P.; Sun, S.; Li, J.; Chen, M. Genome-wide analysis of the GRAS gene family in rice and Arabidopsis. Plant molecular biology 2004, 54, 519-532. https://doi.org/10.1023/b:plan.0000038256.89809.57
16. Grimplet, J.; Agudelo-Romero, P.; Teixeira, R.T.; Martinez-Zapater, J.M.; Fortes, A.M. Structural and functional analysis of the GRAS gene family in grapevine indicates a role of GRAS proteins in the control of development and stress responses. Frontiers in plant science 2016, 7, 353. https://doi.org/10.3389/fpls.2016.00353
17. Liu, Y.; Wang, W. Characterization of the GRAS gene family reveals their contribution to the high adaptability of wheat. PeerJ 2021, 9, e10811. https://doi.org/10.7717/peerj.10811
18. Kumar, B.; Bhalothia, P. Evolutionary analysis of GRAS gene family for functional and structural insights into hexaploid bread wheat (Triticum aestivum). Journal of Biosciences 2021, 46, 45. https://doi.org/10.1007/s12038-021-00163-5
19. Wang, T.-T.; Yu, T.-F.; Fu, J.-D.; Su, H.-G.; Chen, J.; Zhou, Y.-B.; Chen, M.; Guo, J.; Ma, Y.-Z.; Wei, W.-L. Genome-wide analysis of the GRAS gene family and functional identification of GmGRAS37 in drought and salt tolerance. Frontiers in Plant Science 2020, 11, 604690. https://doi.org/10.3389/fpls.2020.604690
20. Wang, L.; Ding, X.; Gao, Y.; Yang, S. Genome-wide identification and characterization of GRAS genes in soybean (Glycine max). BMC Plant Biology 2020, 20, 415. https://doi.org/10.1186/s12870-020-02636-5
21. Xie, Z.; Yang, D.; Zhou, Z.; Li, K.; Yi, P.; Liu, A.; Zhou, Z.; Tu, X. A genome-wide analysis of the GRAS gene family in upland cotton and a functional study of the role of the GhGRAS55 gene in regulating early maturity in cotton. Biotechnology Journal 2023, 18, 2300201. https://doi.org/10.1002/biot.202300201
22. Guo, P.; Wen, J.; Yang, J.; Ke, Y.; Wang, M.; Liu, M.; Ran, F.; Wu, Y.; Li, P.; Li, J. Genome-wide survey and expression analyses of the GRAS gene family in Brassica napus reveals their roles in root development and stress response. Planta 2019, 250, 1051-1072. https://doi.org/10.1007/s00425-019-03199-y

23. Chen, J.; Yan, Q.; Li, J.; Feng, L.; Zhang, Y.; Xu, J.; Xia, R.; Zeng, Z.; Liu, Y. The GRAS gene family and its roles in seed development in litchi (Litchi chinensis Sonn). BMC Plant Biology 2021, 21, 423. https://doi.org/10.1186/s12870-021-03193-1
24. Lin, J.; Wu, J.; Zhang, D.; Cai, X.; Du, L.; Lu, L.; Liu, C.; Chen, S.; Yao, Q.; Xie, S. The GRAS gene family and its roles in pineapple (Ananas comosus L.) developmental regulation and cold tolerance. BMC Plant Biology 2024, 24, 1204. https://doi.org/10.1186/s12870-024-05913-9
25. Wang, Z.; Zhang, Z.; Wang, P.; Qin, C.; He, L.; Kong, L.; Ren, W.; Liu, X.; Ma, W. Genome-wide identification of the NAC transcription factors family and regulation of metabolites under salt stress in Isatis indigotica. International Journal of Biological Macromolecules 2023, 240, 124436. https://doi.org/10.1016/j.ijbiomac.2023.124436
26. Chen, C.; Wu, Y.; Li, J.; Wang, X.; Zeng, Z.; Xu, J.; Liu, Y.; Feng, J.; Chen, H.; He, Y. TBtools-II: A “one for all, all for one” bioinformatics platform for biological big-data mining. Molecular plant 2023, 16, 1733-1742. https://doi.org/10.1016/j.molp.2023.09.010
27. Wang, J.; Chitsaz, F.; Derbyshire, M.K.; Gonzales, N.R.; Gwadz, M.; Lu, S.; Marchler, G.H.; Song, J.S.; Thanki, N.; Yamashita, R.A. The conserved domain database in 2023. Nucleic acids research 2023, 51, D384-D388. https://doi.org/10.1093/nar/gkac1096
28. Fan, Y.; Wan, X.; Zhang, X.; Zhang, J.; Zheng, C.; Yang, Q.; Yang, L.; Li, X.; Feng, L.; Zou, L. GRAS gene family in rye (Secale cereale L.): genome-wide identification, phylogeny, evolutionary expansion and expression analyses. BMC Plant Biology 2024, 24, 46. https://doi.org/10.1186/s12870-023-04674-1
29. Yu, J.; Hu, F.; Dossa, K.; Wang, Z.; Ke, T. Genome-wide analysis of UDP-glycosyltransferase super family in Brassica rapa and Brassica oleracea reveals its evolutionary history and functional characterization. Bmc Genomics 2017, 18, 474. https://doi.org/10.1186/s12864-017-3844-x
30. Dutta, M.; Saha, A.; Moin, M.; Kirti, P.B. Genome-wide identification, transcript profiling and bioinformatic analyses of GRAS transcription factor genes in rice. Frontiers in Plant Science 2021, 12, 777285. https://doi.org/10.3389/fpls.2021.777285
31. Wang, J.; Zhang, Q.; Yu, Q.; Peng, L.; Wang, J.; Dai, Q.; Yang, Y.; Li, X. CARK6 is involved in abscisic acid to regulate stress responses in Arabidopsis thaliana. Biochemical and Biophysical Research Communications 2019, 513, 460-464. https://doi.org/10.1016/j.bbrc.2019.03.180
32. Zhang, J.; Sun, Y.; Zhou, Z.; Zhang, Y.; Yang, Y.; Zan, X.; Li, X.; Wan, J.; Gao, X.; Chen, R. OsSCL30 overexpression reduces the tolerance of rice seedlings to low temperature, drought and salt. Scientific Reports 2022, 12, 8385. https://doi.org/10.1038/s41598-022-12438-4
33. Chen, K.; Li, H.; Chen, Y.; Zheng, Q.; Li, B.; Li, Z. TaSCL14, a novel wheat (Triticum aestivum L.) GRAS gene, regulates plant growth, photosynthesis, tolerance to photooxidative stress, and senescence. Journal of Genetics and Genomics 2015, 42, 21-32. https://doi.org/10.1016/j.jgg.2014.11.002
34. Yang, Y.; Wang, C.; Liang, Y.; Xiao, D.; Fu, T.; Yang, X.; Liu, J.; Wang, S.; Wang, Y. PagTPS1 and PagTPS10, the trehalose-6-phosphate synthase genes, increase trehalose content and enhance drought tolerance. International Journal of Biological Macromolecules 2024, 279, 135518. https://doi.org/10.1016/j.ijbiomac.2024.135518

35. Sugimoto, K.; Allmann, S.; Kolomiets, M.V. Oxylipins: the front line of plant interactions. 2022, 13, 878765. https://doi.org/10.3389/fpls.2022.878765
36. Jansen, R.S.; Addie, R.; Merkx, R.; Fish, A.; Mahakena, S.; Bleijerveld, O.B.; Altelaar, M.; IJlst, L.; Wanders, R.J.; Borst, P. N-lactoyl-amino acids are ubiquitous metabolites that originate from CNDP2-mediated reverse proteolysis of lactate and amino acids. Proceedings of the National Academy of Sciences 2015, 112, 6601-6606. https://doi.org/10.1073/pnas.1424638112v
37. Mapelli-Brahm, A.; Sánchez, R.; Pan, X.; Moreno-Pérez, A.J.; Garcés, R.; Martínez-Force, E.; Weselake, R.J.; Salas, J.J.; Venegas-Calerón, M. Functional characterization of lysophosphatidylcholine: Acyl-CoA acyltransferase genes from sunflower (Helianthus annuus L.). Frontiers in Plant Science 2020, 11, 403. https://doi.org/10.3389/fpls.2020.00403
38. Wang, Z.; Wong, D.C.J.; Wang, Y.; Xu, G.; Ren, C.; Liu, Y.; Kuang, Y.; Fan, P.; Li, S.; Xin, H. GRAS-domain transcription factor PAT1 regulates jasmonic acid biosynthesis in grape cold stress response. Plant Physiology 2021, 186, 1660-1678. https://doi.org/10.1093/plphys/kiab142
39. Wang, L.; Kazachkov, M.; Shen, W.; Bai, M.; Wu, H.; Zou, J. Deciphering the roles of Arabidopsis LPCAT and PAH in phosphatidylcholine homeostasis and pathway coordination for chloroplast lipid synthesis. The Plant Journal 2014, 80, 965-976. https://doi.org/10.1111/tpj.12683
40. Rajappa, S.; Krishnamurthy, P.; Kumar, P.P. Regulation of AtKUP2 expression by bHLH and WRKY transcription factors helps to confer increased salt tolerance to Arabidopsis thaliana plants. Frontiers in plant science 2020, 11, 1311. https://doi.org/10.3389/fpls.2020.01311
41. Wang, L.; Mai, Y.-X.; Zhang, Y.-C.; Luo, Q.; Yang, H.-Q. MicroRNA171c-targeted SCL6-II, SCL6-III, and SCL6-IV genes regulate shoot branching in Arabidopsis. Molecular plant 2010, 3, 794-806. https://doi.org/10.1093/mp/ssq042
42. Xing, J.-X.; Zang, Q.-L.; Ye, Z.-L.; Qi, L.-W.; Yang, L.; Li, W.-F. Overexpression of Larch SCL6 Inhibits Transitions from Vegetative Meristem to Inflorescence and Flower Meristem in Arabidopsis thaliana (L.) Heynh. Plants 2024, 13, 1232. https://doi.org/10.3390/plants13091232
43. Zhang, N.; Zhou, S.; Yang, D.; Fan, Z. Revealing shared and distinct genes responding to JA and SA signaling in Arabidopsis by meta-analysis. Frontiers in plant science 2020, 11, 908. https://doi.org/10.3389/fpls.2020.00908

Figures

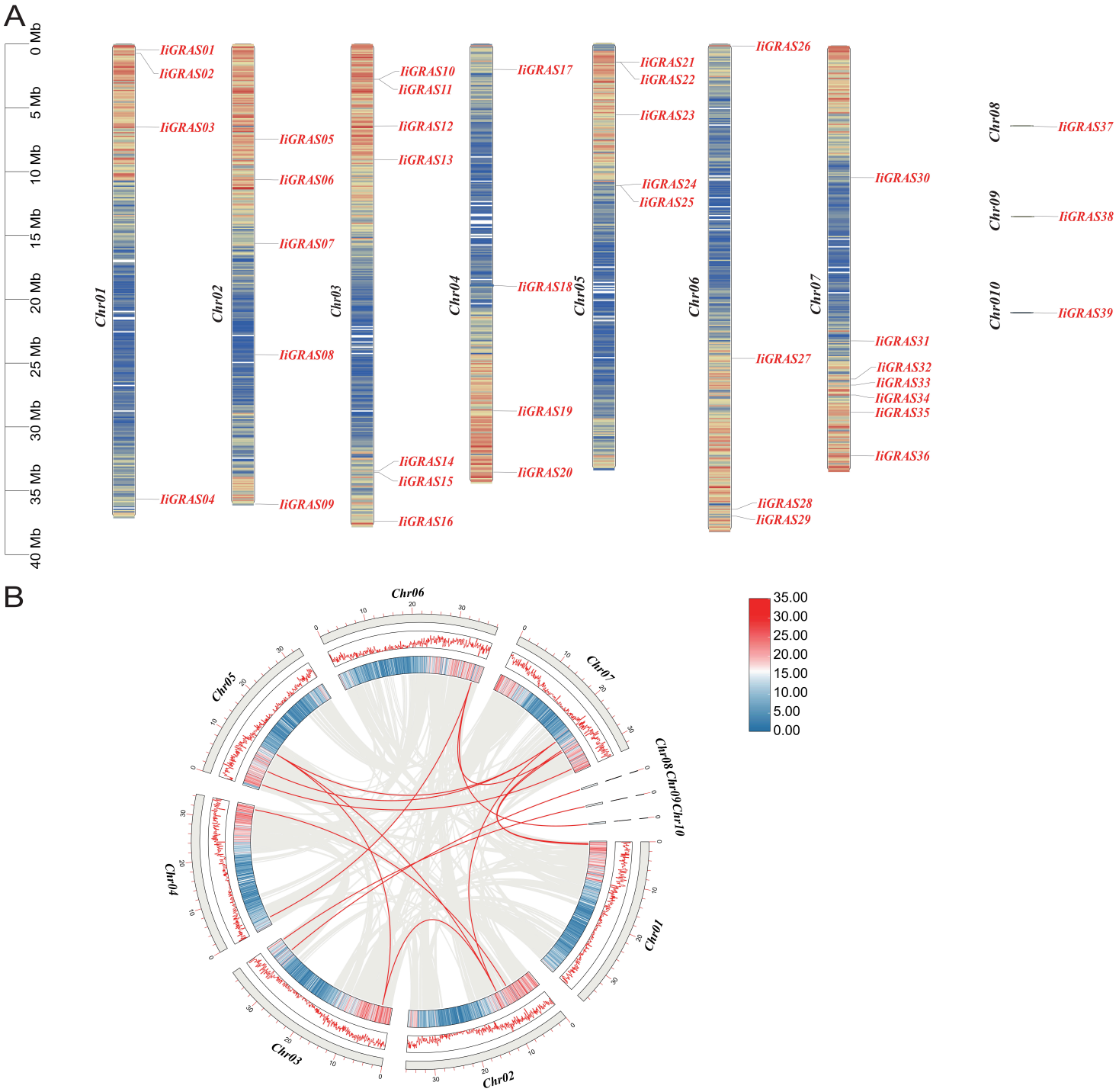


Figure 1

Chromosomal distribution and duplication events of the *GRAS* gene family in *Isatis indigotica*. (A) Chromosomal localization of *GRAS* genes across 10 chromosomes of *I. indigotica*. Red labels indicate gene names, and the color gradient represents gene density (red = high, blue = low). (B) Genome-wide synteny analysis of *I. indigotica GRAS* genes. Red lines represent tandem or segmental duplication relationships among *IiGRAS*genes, while gray lines indicate other collinear gene pairs across the genome. The inner rings depict gene density, where red denotes high density and blue denotes low density.

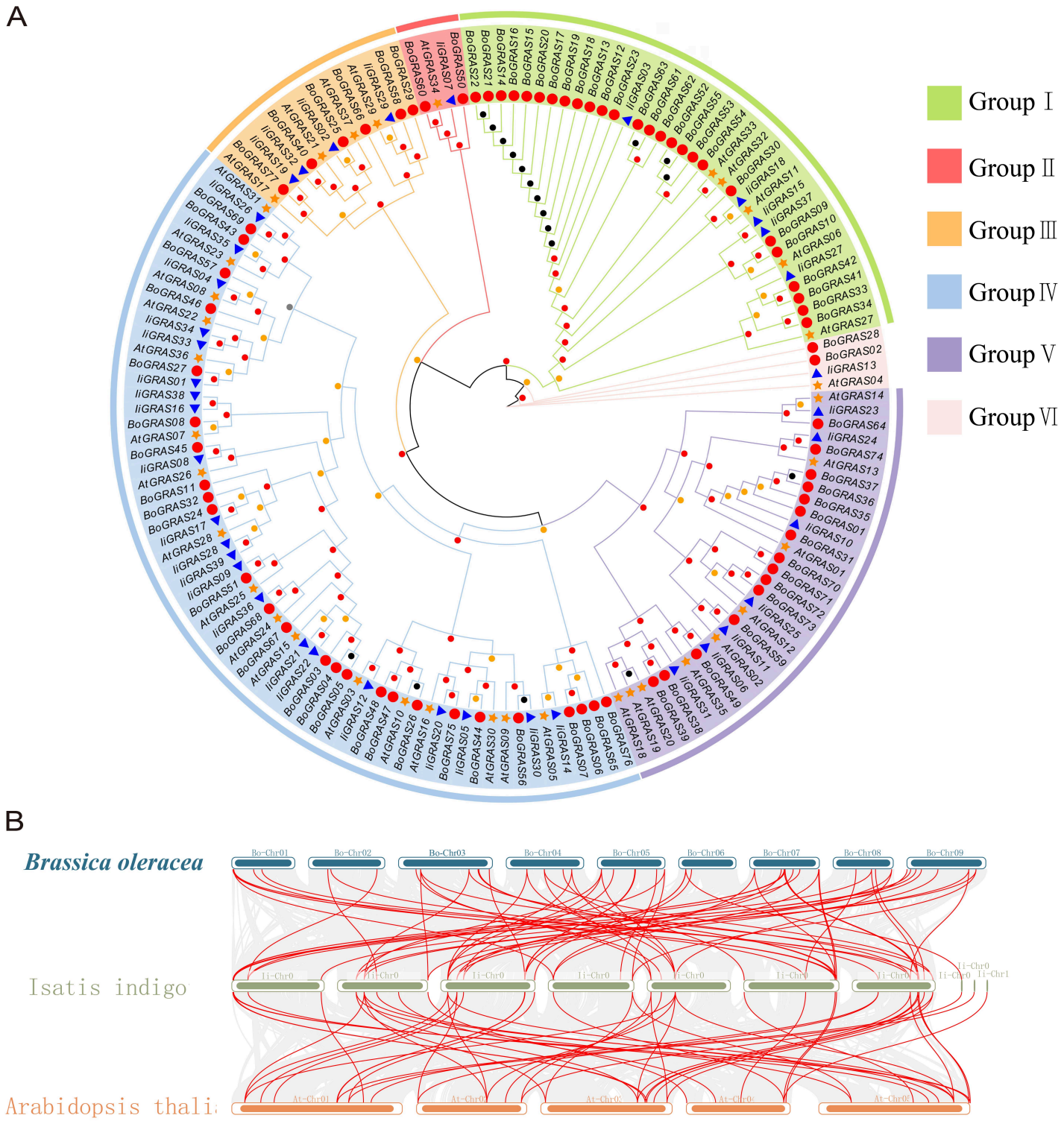


Figure 2

Phylogenetic relationships and synteny analysis of the *GRAS* gene family among *I. indigotica*, *A. thaliana*, and *B. oleracea*. (A) Phylogenetic tree of GRAS proteins from *I. indigotica*, *A. thaliana*, and *B. oleracea*. (B) The Neighbor-Joining (NJ) method was used to construct the phylogenetic tree in MEGA-X with 1000 bootstrap replicates. Different colors represent distinct *GRAS* subfamilies. Circles, triangles, and pentagrams denote GRAS proteins from *B. oleracea*, *I. indigotica*, and *A. thaliana*, respectively. (C)

Synteny relationships of GRAS genes among *I. indigotica*, *A. thaliana*, and *B. oleracea*. Red lines indicate conserved homologous gene pairs between species.

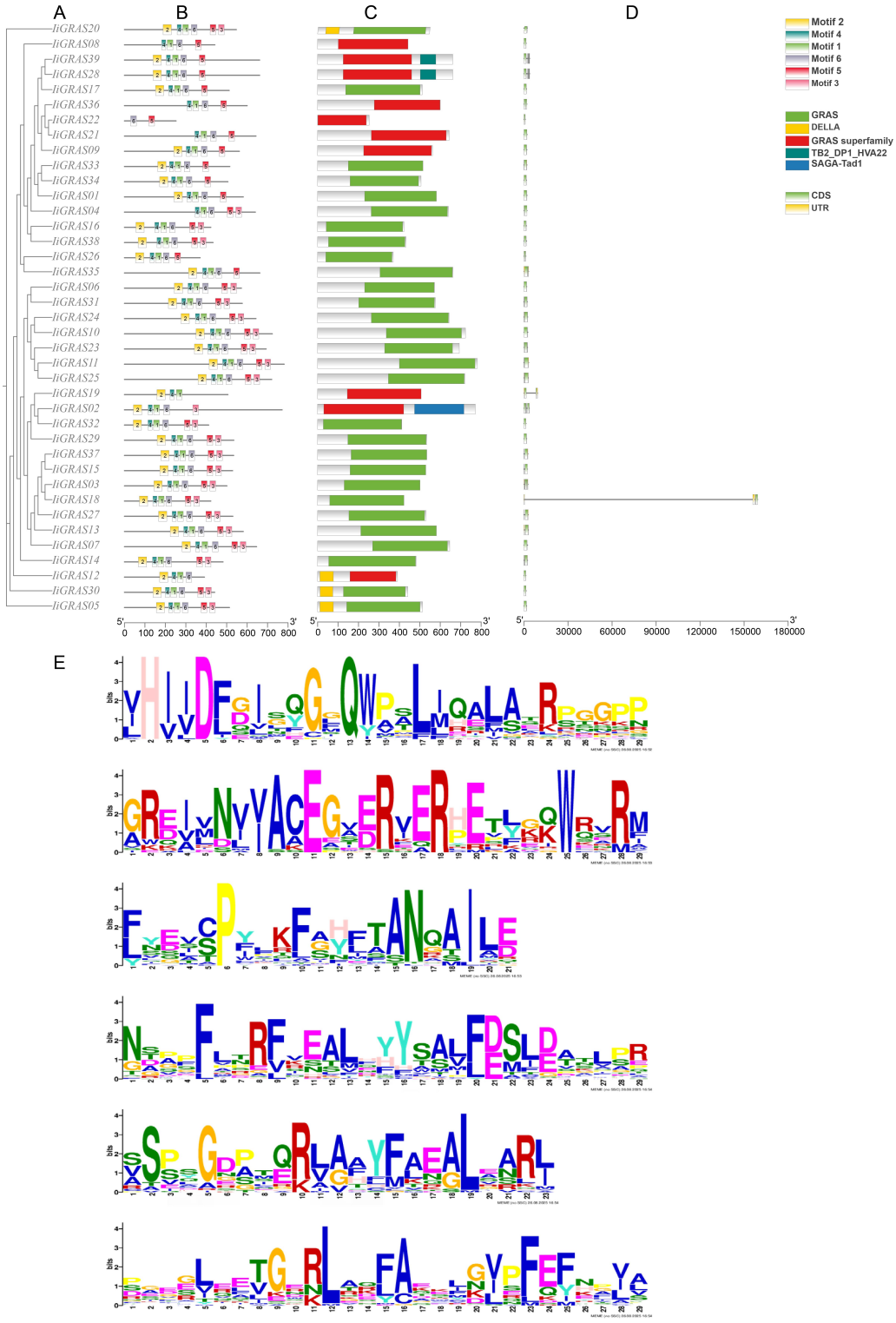


Figure 3

Phylogenetic relationships, conserved motifs, gene structures, and motif logos of the *GRAS*gene family in *I. indigotica.* (A) Phylogenetic tree of 39 *IiGRAS*proteins constructed using the Neighbor-Joining (NJ)

method in MEGA-X with 1000 bootstrap replicates, showing the evolutionary relationships among *IiGRAS*members. (B) Distribution of conserved motifs in *IiGRAS* proteins. Different colored boxes represent distinct motifs (Motif 1-6) predicted by MEME, illustrating the motif composition and arrangement of each protein. (C) Gene structures of *IiGRAS* members. Green boxes indicate coding sequences (CDS), yellow boxes represent untranslated regions (UTRs), and black lines denote introns. Structural variations in exon-intron organization suggest potential functional divergence among subfamilies. (D) Gene length distribution of *IiGRAS* genes, with base pairs (bp) on the x-axis and gene IDs on the y-axis. (E) Conserved motif sequence logos of the GRAS protein family generated by MEME. The six major motifs correspond to conserved amino acid sites that form the core functional domains of GRAS proteins.

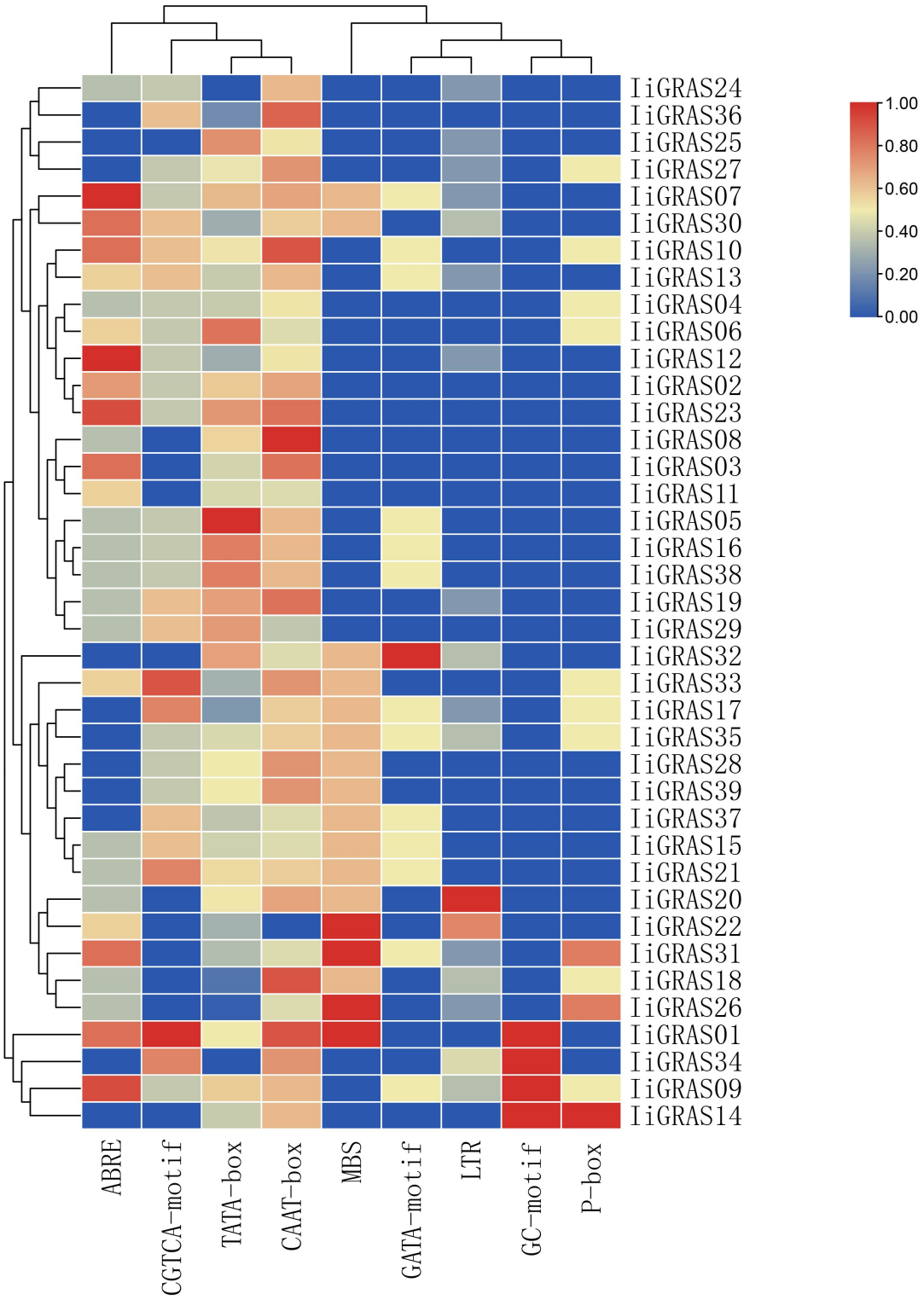


Figure 4

Distribution and enrichment analysis of cis-acting elements in the promoters of *IiGRAS* genes in *I. indigotica*.

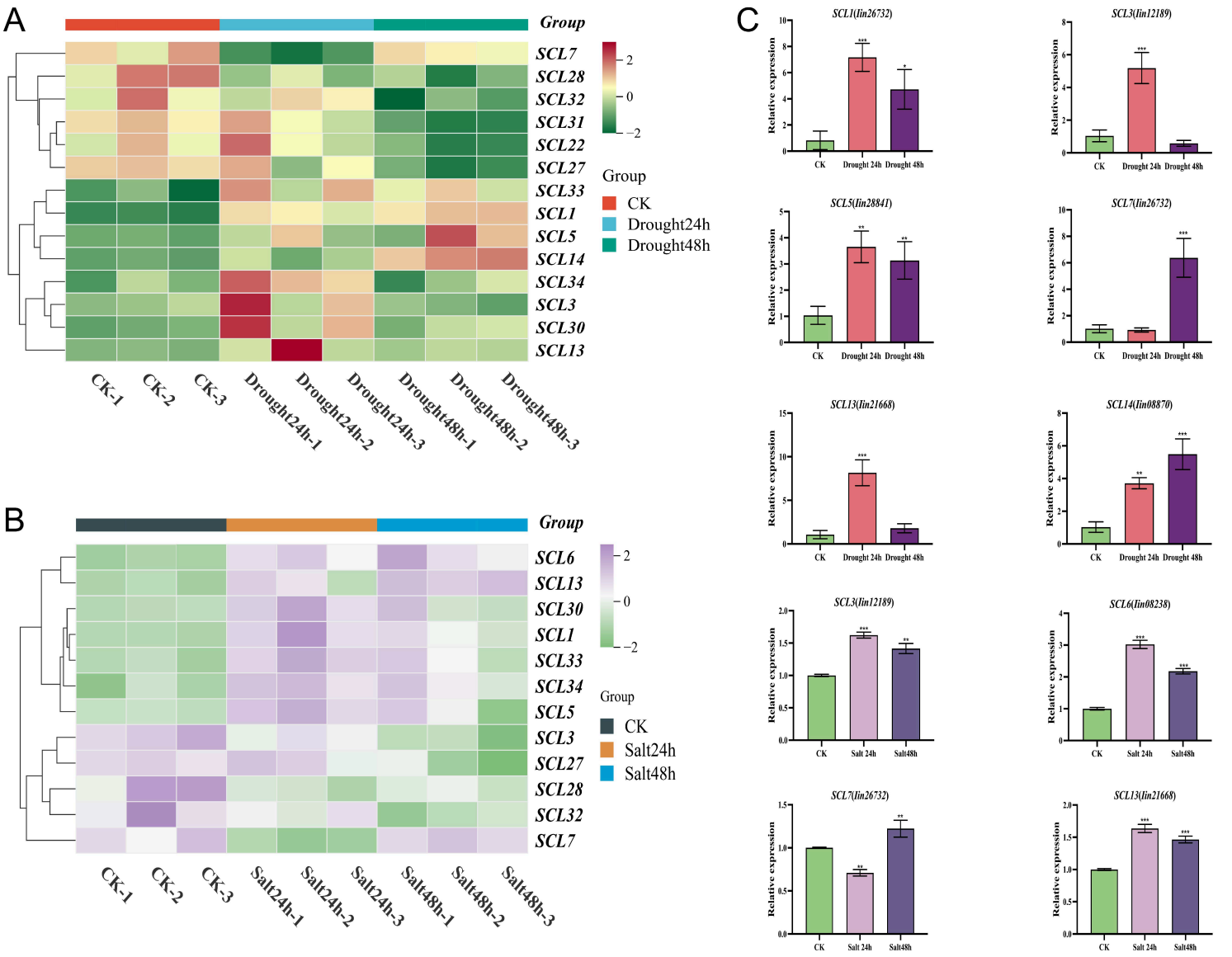
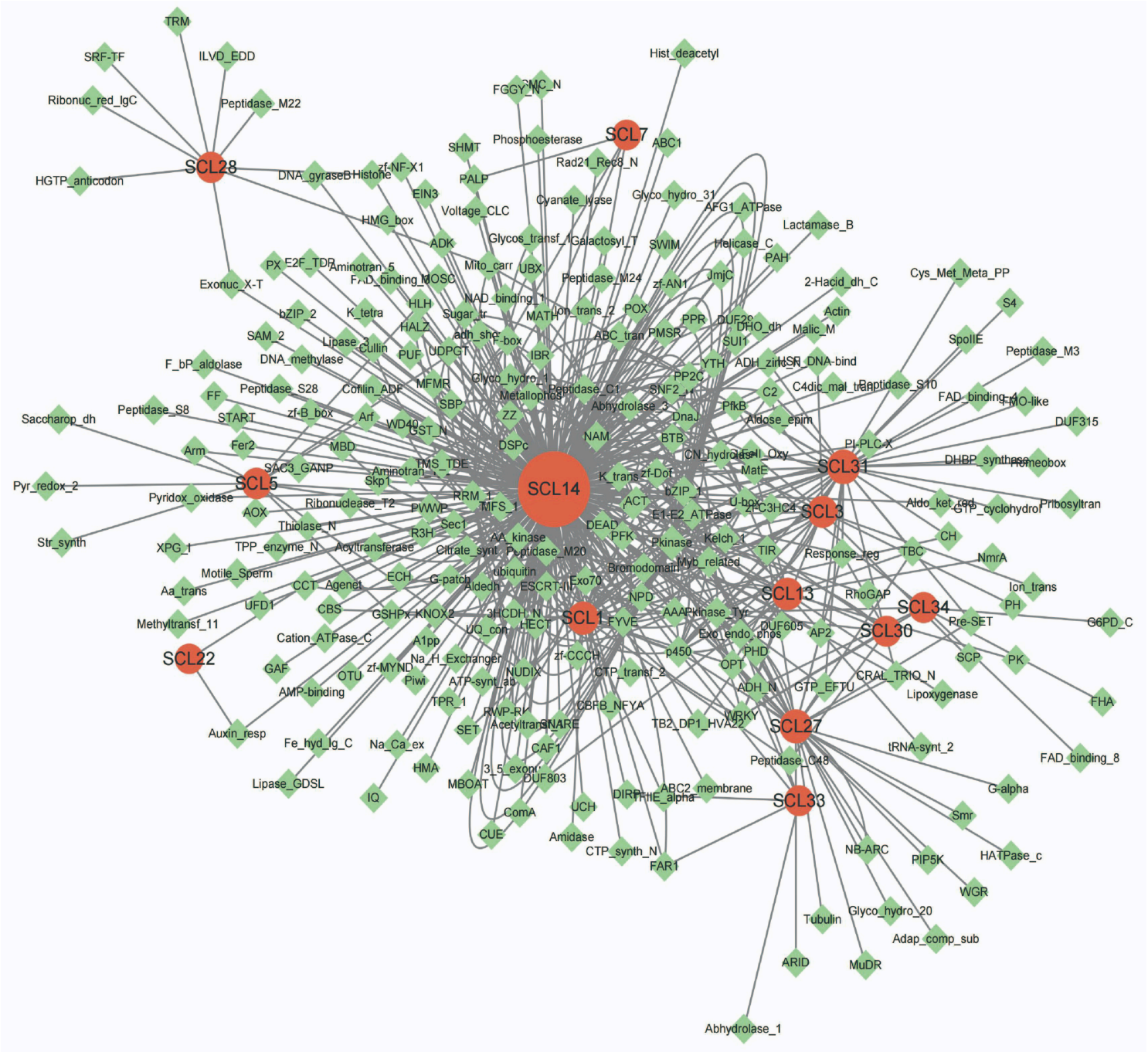


Figure 5

Expression patterns and qRT-PCR validation of *IiSCL* subfamily genes in *I. indigotica* under drought and salt stresses. (A) Heatmap of transcriptome expression profiles of *IiSCL* subfamily genes under drought stress (0 h, 24 h, and 48 h). The color gradient from green to red represents low to high expression levels. (B) Heatmap showing the expression profiles of *IiSCL*subfamily genes under salt stress (0 h, 24 h, and 48 h). The color gradient from green to purple indicates the variation in expression levels. (C) qRT-PCR validation of selected *IiSCL* genes. Bar charts indicate relative expression levels under drought and salt treatments, with error bars representing the standard error (±SE) of three technical replicates. Different letters indicate statistically significant differences at *($p < 0.05$).


SCL14
SCL28
SCL7
SCL5
SCL31
SCL3
SCL13
SCL34
SCL30
SCL1
SCL22
SCL27
SCL33
TRM
SRF-TF
ILVD_EDD
Ribonuc_red_lgC
Peptidase_M22
HGTP_anticodon
Hist_deacetyl
FGGY_N
Phosphoesterase
SHMT
Rad21_Rec8_N
ABC1
PALP
Cyanate_lyase
Glyco_hydro_31
AFG1_ATPase
Lactamase_B
Helicase_C
PAH
Cys_Met_Meta_PP
2-Hacid_dh_C
S4
SpoIIE
Peptidase_M3
Peptidase_S10
FAD_binding_4
FMO-like
DUF315
Homeobox
DHBP_synthase
GTP_cyclohydrol
PRibosyltran
Aldo_ket_red
CH
NmrA
Ion_trans
PH
G6PD_C
Pre-SET
SCP
PK
FHA
Lipoxygenase
CRAL_TRIO_N
GTP_EFTU
ADH_N
tRNA-synt_2
FAD_binding_8
G-alpha
Smr
HATPase_c
WGR
Adap_comp_sub
MuDR
ARID
Tubulin
Glyco_hydro_20
NB-ARC
PIP5K
Abhydrolase_1
FAR1
CTP_synth_N
Amidase
UCH
ComA
CUE
DUF803
MBOAT
3_5_exonu
HMA
IQ
Lipase_GDSL
Auxin_resp
Fe_hyd_lg_C
Na_Ca_ex
SET
Acetyltransf
SNARE
CAF1
TPR_1
RWP-RK
CBFB_NFYA
TB2_DP1_HVA22
WRKY
Peptidase_C48
ABC2_membrane
PHE_alpha
DIRP
Piwi
ATP-synt_ab
NUDIX
zf-CCCH
CTP_transf_2
OPT
p450
PHD
Exo_endo_phos
DUF605
AP2
RhoGAP
AAA
Pkinase_Tyr
Bromodomain
Myb_related
Response_reg
TIR
TBC
Kelch_1
Pkinase
PFK
DEAD
E1-E2_ATPase
U-box
ZFC3HC4
ACT
bZIP_1
MatE
CN_hydrolase
Asn_synthase
Oxy
PI-PLC-X
DnaJ
Aldose_epim
PfkB
C4dic_mal_tran
C2
PP2C
SNF2_N
Peptidase_C1
Metallophos
Abhydrolase_3
YTH
ADH_zinc_N
DNA-bind
SUI1
DHO_dh
Malic_M
Actin
PPR
DUF26
PMSR
ABC_tran
Ion_trans_2
POX
MATH
Glyco_hydro_1
SBP
ZZ
NAM
BTB
DSPc
IBR
F-box
UDPGT
PUF
Cullin
Lipase_3
SAM_2
bZIP_2
K_tetra
HALZ
HLH
Sugar_tr
NAD_binding_1
adh_short
Mito_carr
UBX
Peptidase_M24
zf-AN1
JmjC
MOSC
FAD_binding_4
Aminotran_5
PX
E2F_TDP
Exonuc_X-T
ADK
Glycos_transf_1
Galactosyl_T
SWIM
PAH
HMG_box
Voltage_CLC
EIN3
DNA_gyraseB
Histone
zf-NF-X1
HGTP_anticodon
Ribonuc_red_lgC
Peptidase_M22
SRF-TF
ILVD_EDD
TRM
Saccharop_dh
Peptidase_S8
FF
Peptidase_S28
Cofilin_ADF
MFMR
F_bP_aldolase
DNA_methylase
Lipase_3
Cullin
PUF
UDPGT

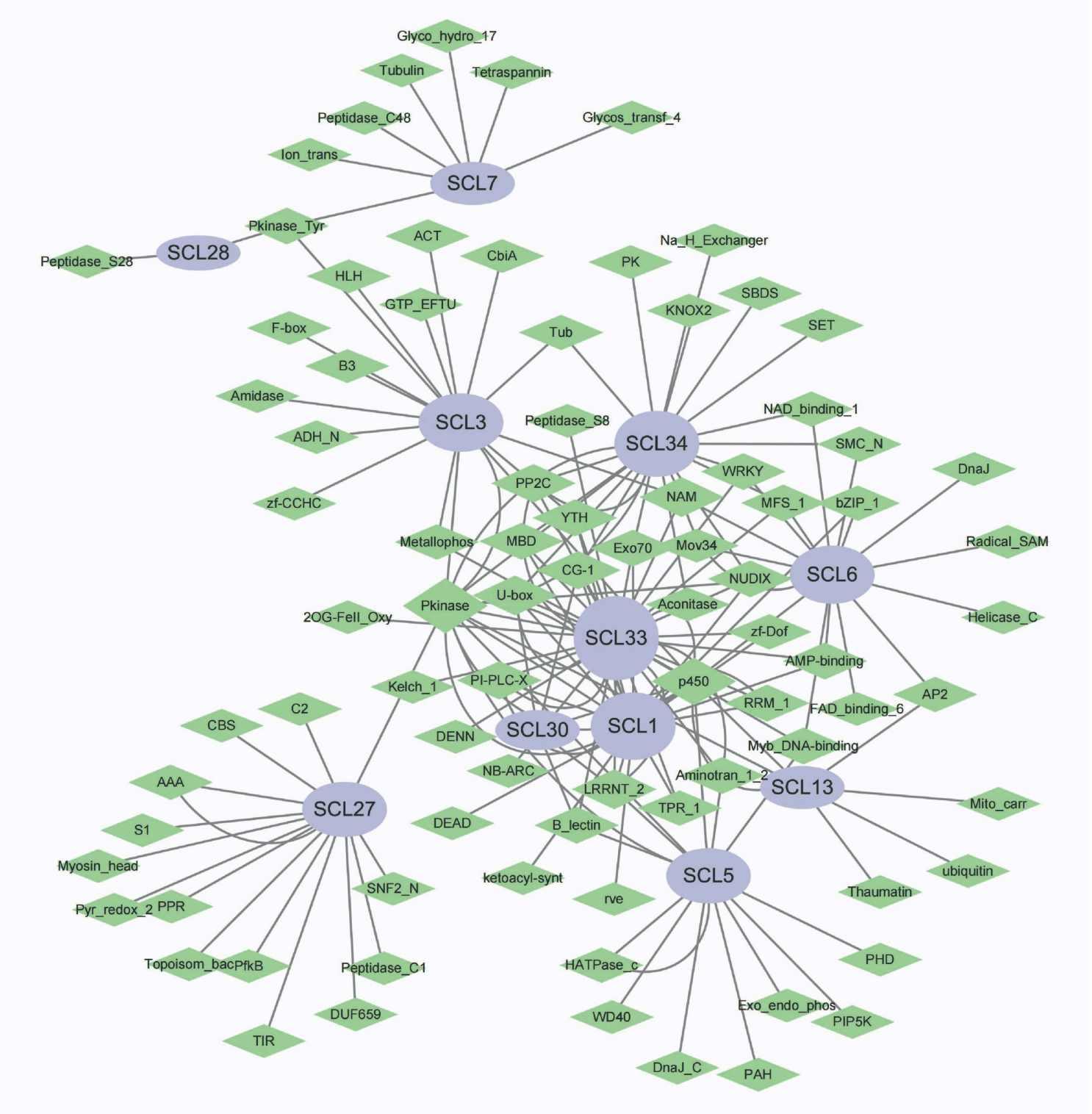


Figure 7

Co-expression network between *IiSCL* genes and TFs in *I. indigotica* under salt stress. Purple circular nodes represent *IiSCL* genes, while green diamond-shaped nodes represent transcription factors or functional protein domains. Node size corresponds to the strength of correlation, and gray edges indicate significant positive associations.

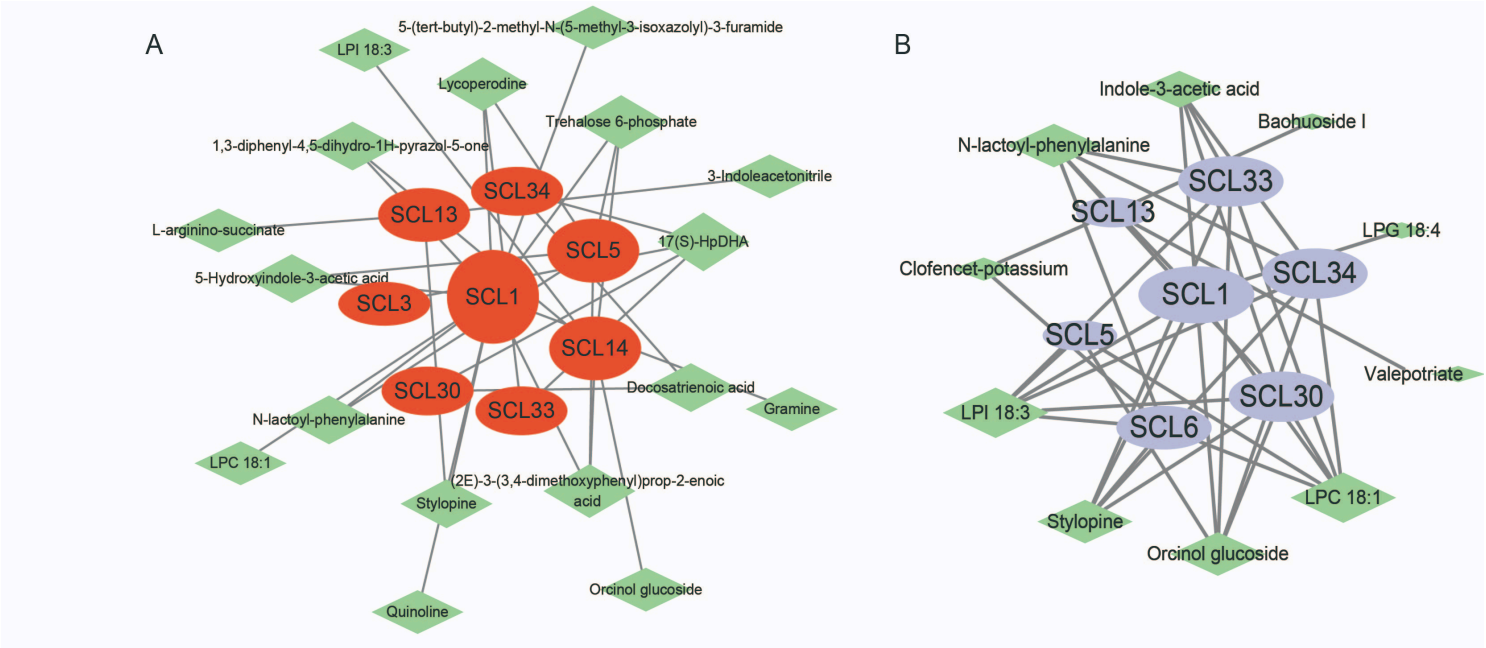


Figure 8

Correlation network between *IiSCL* subfamily genes and metabolites in *I. indigotica*. (A) Correlation network between *IiSCL* genes and significantly upregulated metabolites under drought stress. (B) Correlation network between *IiSCL* genes and significantly upregulated metabolites under salt stress. Circular nodes represent *IiSCL* genes, while diamond-shaped nodes indicate metabolites. Node size is proportional to the correlation coefficient (|r|), and edges represent significant positive correlations ($r > 0.7$).

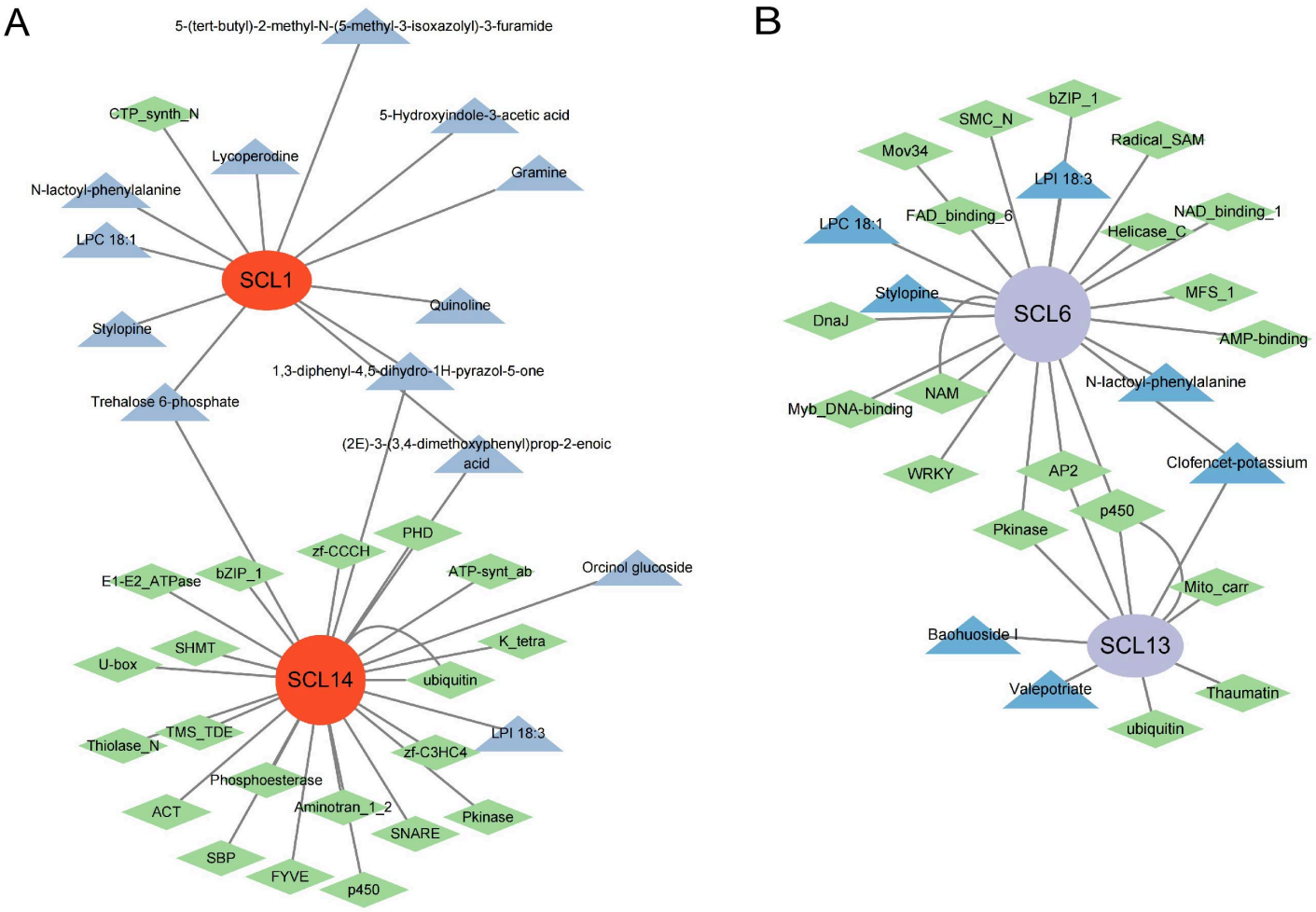


Figure 9

Integrated correlation network of *IiSCL* subfamily genes, TFs, and metabolites in *I. indigotica*. (A) Integrated co-expression network of differentially expressed *IiSCL* genes with significantly correlated TFs and metabolites under drought stress. (B) Integrated co-expression network of differentially expressed *IiSCL* genes with TFs and metabolites under salt stress. Circular nodes represent *IiSCL* genes, triangular nodes denote TFs, and diamond-shaped nodes represent metabolites. Node size is proportional to the correlation coefficient (|*r*|), and edges indicate significant correlations ($r > 0.7$).

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Supplementarytable.zip