Genomic organization and expression profiles of nitrogen assimilation genes in *Glycine max* (#98270)

First submission

Guidance from your Editor

Please submit by **25 Apr 2024** for the benefit of the authors (and your token reward) .

Structure and Criteria
Please read the 'Structure and Criteria' page for general guidance.

Raw data check
Review the raw data.

Image check
Check that figures and images have not been inappropriately manipulated.

If this article is published your review will be made public. You can choose whether to sign your review. If uploading a PDF please remove any identifiable information (if you want to remain anonymous).

Files
Download and review all files from the materials page.

9 Figure file(s)
5 Table file(s)

Structure and Criteria

Structure your review

The review form is divided into 5 sections. Please consider these when composing your review:

1. **BASIC REPORTING**
2. **EXPERIMENTAL DESIGN**
3. **VALIDITY OF THE FINDINGS**
4. General comments
5. Confidential notes to the editor

You can also annotate this PDF and upload it as part of your review

When ready submit online.

Editorial Criteria

Use these criteria points to structure your review. The full detailed editorial criteria is on your guidance page.

BASIC REPORTING

- Clear, unambiguous, professional English language used throughout.
- Intro & background to show context. Literature well referenced & relevant.
- Structure conforms to PeerJ standards, discipline norm, or improved for clarity.
- Figures are relevant, high quality, well labelled & described.
- Raw data supplied (see PeerJ policy).

EXPERIMENTAL DESIGN

- Original primary research within Scope of the journal.
- Research question well defined, relevant & meaningful. It is stated how the research fills an identified knowledge gap.
- Rigorous investigation performed to a high technical & ethical standard.
- Methods described with sufficient detail & information to replicate.

VALIDITY OF THE FINDINGS

- Impact and novelty not assessed. *Meaningful* replication encouraged where rationale & benefit to literature is clearly stated.
- All underlying data have been provided; they are robust, statistically sound, & controlled.
- Conclusions are well stated, linked to original research question & limited to supporting results.

Standout reviewing tips

The best reviewers use these techniques

Tip	*Example*
Support criticisms with evidence from the text or from other sources	*Smith et al (J of Methodology, 2005, V3, pp 123) have shown that the analysis you use in Lines 241-250 is not the most appropriate for this situation. Please explain why you used this method.*
Give specific suggestions on how to improve the manuscript	*Your introduction needs more detail. I suggest that you improve the description at lines 57- 86 to provide more justification for your study (specifically, you should expand upon the knowledge gap being filled).*
Comment on language and grammar issues	*The English language should be improved to ensure that an international audience can clearly understand your text. Some examples where the language could be improved include lines 23, 77, 121, 128 – the current phrasing makes comprehension difficult. I suggest you have a colleague who is proficient in English and familiar with the subject matter review your manuscript, or contact a professional editing service.*
Organize by importance of the issues, and number your points	*1. Your most important issue* *2. The next most important item* *3. …* *4. The least important points*
Please provide constructive criticism, and avoid personal opinions	*I thank you for providing the raw data, however your supplemental files need more descriptive metadata identifiers to be useful to future readers. Although your results are compelling, the data analysis should be improved in the following ways: AA, BB, CC*
Comment on strengths (as well as weaknesses) of the manuscript	*I commend the authors for their extensive data set, compiled over many years of detailed fieldwork. In addition, the manuscript is clearly written in professional, unambiguous language. If there is a weakness, it is in the statistical analysis (as I have noted above) which should be improved upon before Acceptance.*

Genomic organization and expression profiles of nitrogen assimilation genes in *Glycine max*

Hind Abdelmonim Elsanosi [1, 2], **Tiantian Zhu** [1], **Guisheng Zhou** [1], **Li Song** [Corresp. 1]

[1] Joint International Research Laboratory of Agriculture and Agri-Product Safety, the Ministry of Education of China, Jiangsu Key Laboratory of Crop Genomics and Molecular Breeding, Yangzhou University, Yangzhou, Jiangsu, China
[2] Faculty of Agriculture, University of Khartoum, Khartoum, Sudan

Corresponding Author: Li Song
Email address: songli@yzu.edu.cn

Background: Glutamine synthetase (GS), Glutamate synthase (GOGAT), and nitrate reductase (NR) are key enzymes involved in nitrogen assimilation and metabolism in plants .However , the systematic analysis of these gene families lacked reports in soybean (*Glycine max* (L.) Merr.), one of the most important crops worldwide. Methods: In this study, we performed genome-wide identification and characterization of *GS*, *GOGAT*, and *NR* genes in soybean under abiotic and nitrogen stress conditions. Results: We identified a total of 10 *GS* genes, 6 *GOGAT* genes, and 4 *NR* genes in the soybean genome. Phylogenetic analysis revealed the presence of multiple isoforms for each gene family, indicating their functional diversification. The distribution of these genes on soybean chromosomes was uneven, with segmental duplication events contributing to their expansion. Within the NAGs group, there was uniformity in the exon-intron structure and the presence of conserved motifs in NAGs. Furthermore, analysis of cis-elements in NAG promoters indicated complex regulation of their expression. RT-qPCR analysis of seven soybean NAGs under various abiotic stresses, including nitrogen deficiency, drought-nitrogen, and salinity, revealed distinct regulatory patterns. Most NAGs exhibited up-regulation under nitrogen stress, while diverse expression patterns were observed under salt and drought-nitrogen stress, indicating their crucial role in nitrogen assimilation and abiotic stress tolerance. These findings offer valuable insights into the genomic organization and expression profiles of *GS*, *GOGAT*, and *NR* genes in soybean under nitrogen and abiotic stress conditions, with potential applications in developing. stress-resistant soybean varieties through genetic engineering and breeding.

Genomic organization and expression profiles of nitrogen assimilation genes in *Glycine max*

Hind Abdelmonim Elsanosi[1, 2], Tiantian Zhu[1], Guisheng Zhou[1], Li Song[1]*

[1]Joint International Research Laboratory of Agriculture and Agri-Product Safety, the Ministry of Education of China, Jiangsu Key Laboratory of Crop Genomics and Molecular Breeding, Yangzhou University, Yangzhou, Jiangsu 225009, China

[2]University of Khartoum, Faculty of Agriculture, 11115 Khartoum, Sudan.

*Corresponding author

Li Song

Address: International Building, Wenhui Road Campus, Yangzhou University, Yangzhou 225009, Jiangsu Province, China

Email address: songli@yzu.edu.cn

Abstract

Background: Glutamine synthetase (GS), Glutamate synthase (GOGAT), and nitrate reductase (NR) are key enzymes involved in nitrogen assimilation and metabolism in plants .However , the systematic analysis of these gene families lacked reports in soybean (*Glycine max* (L.) Merr.), one of the most important crops worldwide.

Methods: In this study, we performed genome-wide identification and characterization of *GS*, *GOGAT*, and *NR* genes in soybean under abiotic and nitrogen stress conditions.

Results: We identified a total of 10 *GS* genes, 6 *GOGAT* genes, and 4 *NR* genes in the soybean genome. Phylogenetic analysis revealed the presence of multiple isoforms for each gene family, indicating their functional diversification. The distribution of these genes on soybean chromosomes was uneven, with segmental duplication events contributing to their expansion.

Within the NAGs group, there was uniformity in the exon-intron structure and the presence of conserved motifs in NAGs. Furthermore, analysis of cis-elements in NAG promoters indicated complex regulation of their expression. RT-qPCR analysis of seven soybean NAGs under various abiotic stresses, including nitrogen deficiency, drought-nitrogen, and salinity, revealed distinct regulatory patterns. Most NAGs exhibited up-regulation under nitrogen stress, while diverse expression patterns were observed under salt and drought-nitrogen stress, indicating their crucial role in nitrogen assimilation and abiotic stress tolerance. These findings offer valuable insights into the genomic organization and expression profiles of *GS*, *GOGAT*, and *NR* genes in soybean under nitrogen and abiotic stress conditions, with potential applications in developing stress-resistant soybean varieties through genetic engineering and breeding.

Keyword: Soybean; Glutamine synthetase; Glutamate synthase; Nitrate reductase; Abiotic stress

Introduction:

Nitrogen is an essential nutrient for plant growth and development, playing a key role in various physiological processes such as protein synthesis, nucleic acid production, and regulation of enzyme activity. The availability of nitrogen greatly influences crop productivity [1]. Nitrogen use efficiency (NUE) is a vital aspect of crop productivity, relying on the plant's capacity to extract inorganic nitrogen from the soil, assimilate nitrate and ammonium, and recycle organic nitrogen [2].

Glutamine synthetase (GS), Glutamate synthase (GOGAT), and Nitrate reductase (NR) are essential and interrelated genes in plant nitrogen metabolism, playing key roles in nitrogen assimilation, remobilization, storage, reutilization and stress resistance [3, 4]. NR is essential for converting nitrate (NO_3^-) to nitrite (NO_2^-), initiating the process of inorganic nitrogen utilization [5]. GS and GOGAT play a crucial role in converting inorganic ammonium salts into organic nitrogen compounds through the GS/GOGAT cycle [6, 7]. Plants typically have cytoplasmic (GS1) and chloroplast (GS2) forms of GS, along with ferredoxin-dependent GOGAT (Fd-GOGAT) in chloroplasts and plastids, while NADH-dependent GOGAT (NADH-GOGAT) in the cytoplasm[8]. These enzymes facilitate the transfer of an amino group from glutamine to 2-oxoglutarate, producing two molecules of glutamate [9].

Soil salinity, drought, and high temperatures are common environmental stresses that have a significant impact on plant growth and global crop yield [10]. Under stress conditions, plants activate specific nitrogen assimilation genes (NAGs) to improve nitrogen uptake, assimilation, and remobilization. This adaptive response helps plants efficiently utilize nitrogen, allowing them to survive and thrive in challenging conditions. Numerous studies have emphasized the role of NAGs in plant reactions to abiotic and nitrogen stresses. For instance, increasing *GS* genes in rice have been shown to provide resistance to salt, drought and cold stress [11], while introducing pine cytoplasmic glutamine synthetase (GS1) into transgenic poplar has enhanced tolerance to drought stress [12]. Additionally, the induction of *TaGS2* expression by NO_3^- was observed in wheat leaves. In *Zostera marina* L., the expression of the *NR* gene increased in response to NaCl treatment [13].

Soybean (*Glycine max* (L.) Merr.) is a highly valued crop known for its economic and nutritional benefits, including its high oil content (18%) and quality proteins (~40%), as well as positive effects on soil fertility, productivity, and profitability. It is often referred to as a miracle crop [14-16]. Environmental stressors such as drought, salinity, extreme temperatures, and nitrogen stress, can have a significantly impact the growth of soybean plant Growth, leading to reduced yields and cultivation challenges [17, 18]. While the *GS*, *GOGAT*, and *NR* gene families have been studied in other plant species, like rice, pecan, and rapeseed [19-21], a comprehensive genome-wide analysis of these gene families in soybean is still lacking. This study aims to fill this gap by identifying and analyzing these gene families in six legume species (*G. max*, *C. arietinum*, *L. japonicas*, *P. lunatus*, *P. vulgaris*, and *V. unguiculata*), including soybean, using bioinformatics techniques. Various analysis were conducted, including phylogenetic tree construction, gene structure and motif analysis, chromosomal location analysis, gene duplication, synteny analysis, cis-element identification, and protein structure prediction specifically in soybean genes. Additionally, the response of these genes to abiotic and nitrogen stresses was analyzed using RT-qPCR. This research will enhance our understanding of nitrogen utilization and stress tolerance in soybean, providing valuable insights for crop improvement strategies.

provide full form of genera with chromosome numbers of plant species.

Materials and Methods

Identification of NAGs in *Glycine max*

The amino acid sequences of *Arabidopsis GS*, *GOGAT*, and *NR* genes were retrieved from the *Arabidopsis* Information Resource (TAIR) database. These sequences were used as queries for a BLASTP search against the *Glycine max* reference genome (a4.v1 version) to identify members of the soybean *GS*, *GOGAT*, and *NR* gene families. Subsequently, domain analysis tools such as CD search (http://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi), PFAM (http://pfam.xfam.org/), and SMART (http://smart.embl-heidelberg.de/) were employed with default cut-off parameters were used to validate the accuracy of these genes. All candidate genes were aligned with *Arabidopsis* homologous genes. Genomic DNA, cDNA, CDS, and protein sequences of the *GS*, *GOGAT*, and *NR* genes were obtained from Phytozome. The same methods were applied to identify the *GS*, *GOGAT*, and *NR* genes. The confirmed novel NAGs genes were then renamed using a combination of the species abbreviation and the chromosome position.

Analysis of physicochemical properties

Amino acid properties and physicochemical traits, including molecular weight (MW), aliphatic index, instability index (II), and isoelectric point (pI) of *GmGS*, *GmGOGAT*, and *GmNR* proteins were calculated using the ProtParam tool (https://web.expasy.org/protparam). Subcellular localization was predicted using an advanced protein prediction tool WOLF PSORT (https://wolfpsort.hgc.jp/).

Phylogenetic relationship and sequence alignment

The protein sequences of *G.max*, *C. arietinum*, *L. japonicas*, *P. lunatus*, *P. vulgaris*, and *V. unguiculata* were aligned using the MUSCLE tool with default settings. Evolutionary relationships of the NAGs were illustrated through neighbor-joining (NJ) trees for each gene family, constructed with MEGA 11 software using 1000 bootstraps [22]. The resulting trees in Newick format were visualized with iTOL v4 (http://itol.embl.de/) [23].

Analysis of conserved motifs and gene structure

The motif-based sequence analysis tool MEME (https://meme-suite.org/meme/db/motifs) [24] was used to predict the conserved motifs of each protein. Furthermore, details regarding the distribution of exons, introns, and coding sequences were extracted from the GFF3 files of

soybean genome annotation data. The gene architectures were visualized using the TBtools software [25].

Chromosome localization, gene duplication, and syntenic analysis of soybean NAGs

The *GS*, *GOGAT*, and *NR* genes were mapped to specific chromosomes of *G. max* by comparing their physical distances using GFF3 genome files from the Phytozome V13 database (https://phytozome-next.jgi.doe.gov/) [26]. Gene position on the chromosomes was visualized with TBtools software. Collinearity and gene duplication events were examined and presented using the Multiple Collinearity Scan toolkit (MCScanX) with default settings. The collinearity between the homologous gene pairs was visualized using the Circos tool in TBtools. To explore the mechanism behind the amplification of NAGs, gene synteny analysis was conducted between *G. max* and *C. arietinum*, *G. max* and *V. unguiculata*, and *G. max* and *A. thaliana*, and the syntenic relationships were visualized using TBtools software.

Selection pressure and promoter analysis of soybean NAGs

The Ka/Ks ratio was estimated using TBtools software to analyze the selection pressure among soybean NAGs genes within the *G.max* genome. Additionally, we examined the cis-regulatory elements in the promoter regions of NAGs by analyzing the upstream sequences (1500) of NAG proteins downloaded from Phytozome through the PlantCARE database (https://bioinformatics.psb.ugent.be/webtools/plantcare/html/) [27].The distribution of putative cis-elements was visualized using TBtools.

Gene expression pattern of the NAGs in soybean tissues

To examine the expression patterns of NAGs, we analyzed the FPKM values obtained from Phytozome across eight different tissues: root, root tip, lateral root, stem, leaf, shoot tip, open flower, and unopened flower. Further analysis of these expression patterns was carried out using the heatmap function in TBtools.

Protein-protein interaction network

To investigate the interactions among soybean NAGs proteins, researchers utilized STRING V12 (https://string-db.org/) [28]

to analyze the protein sequences. The resulting protein-protein interaction (PPI) networks were then visualized using Cytoscape V3.10.1 [29].

Plant materials and treatments

Williams 82 seeds with uniform size were sterilized using a chlorine gas method [30], and were germinated in a hydroponic system under controlled conditions in a growth chamber with a 16-hour light and 8-hour dark cycle at 25°C. Once the seedlings reached the V1 developmental stage, they were moved to a modified MS liquid medium [31] to assess the impact of different nitrate levels - high (54.3 mM NO_3^--HN), normal (18.81 mM NO_3^--NN), and low (6.27 mM NO_3^--LN) concentrations - for a duration of 7 days. To impose drought-nitrate (D-N) stress, the seedlings were subjected to the specified above nitrate concentrations along with 15% PEG6000 for drought stress over two days. Furthermore, for salt stress experiments, the seedlings were exposed to 150 mM NaCl for 24 and 48 hours. The control treatment exclusively utilized MS medium. Following each treatment, the roots of five plants from three separate biological replicates were harvested and quickly frozen in liquid nitrogen. The samples were ground into powder using a sterilized mortar and pestle in liquid nitrogen. The powdered samples were promptly transferred into 1.5 ml RNase-free micro tubes (Corning Incorporated, Corning, Jiangsu province, China) and stored at -80°C.

is it sufficient for uptake and showing symptoms/ expression of genes? provide reference.

reference?

justify the selection of root only?

specify the time after treatment.

Total RNA extraction and qRT-PCR analysis

Total RNA was isolated from roots using the RNA Pure Plant Kit (DNase1) (Cat#CW0559S, CWBIO, Taizhou, Jiangsu, China). The quality of the RNA samples was assessed for degradation or contamination by 1% agarose gel electrophoresis. Additionally, the purity (A260/A280 ratio) and concentration of the RNA samples were determined using a Nanodrop ND-1000 spectrophotometer (V3.7.9). The primers for quantitative real-time PCR (qRT-PCR) were designed using the IDT online software (https://sg.idtdna.com/) (Table S1). Specificity screening was performed using Phytozome BLAST with the *Glycine max* Wm82.a4.v1 genome as the reference.

To generate cDNA, 1 μg of total RNA was reverse transcribed utilizing HiScript III RT SuperMix for qPCR (Cat# R323-01, Vazyme, Nanjing, Jiangsu, China) following the manufacturer's instructions, reverse transcription reactions are performed at 50°C for 15 min

following by 85°C for 5sec. The qRT-PCR was performed on a CFX96 real-time PCR system (Bio-Rad, USA) using ChamQ SYBR qPCR Master Mix (Cat# Q311, Vazyme Nanjing, Jiangsu, China) which includes SYBR Green1 and other components as specified by the manufacturer. The quantification was performed in triplicate with 10 µl reactions containing 5 µl of 2x ChamQ SYBR qPCR Master mix, 1 µl of primer 10 µM (forward +reversed), 3 µl RNase-free water, and 1 µl cDNA. The PCR conditions involved an initial denaturation at 95°C for 30 s, followed by 40 cycles of 95°C for 5 s, and 60°C for 30 s. A melting curve analysis was conducted by gradually increasing the temperature to 95°C (increment rates of 0.5°C/s) after cooling to 65°C for 5 s.

The raw quantification cycle (Cq) values for each reaction were generated by the Bio-Rad CFX Maestro (version 4.1) as shown in Table S2. The relative expression of the target genes was normalized to the housekeeping gene Actin11 (Glyma18g290800) and calculated using the $2^{-\Delta\Delta Ct}$ method. Ct values were obtained from three biological replicates, each with three technical replicates. Statistically significant differences in gene expression were determined using a t-Test in Excel. Additional qPCR specifics are provided in a MIQE checklist table (Table S3).

Results

Identification and phylogenetic analysis of soybean NAGs

Utilizing bioinformatics techniques, we identified 20, 9, 7, 11, 10, and 10 NAGs in the entire *G. max, C. arietinum*, *L. japonicas*, *P. lunatus*, *P. vulgaris*, and *V. unguiculata* genomes, respectively. ProtParam tool analysis using the ProtParam tool revealed significant differences in molecular weights of GS (ranging from 17.23 to 92.67 kDa), GOGAT (177.20 to 482.52 kDa), and NR (98.27 to 100.08 kDa) proteins in *G. max*. The amino acid sequence length varied from 155 bp to 4395 bp, with a notable diversity in genes encoding GS, GOGAT, and NR. The pI values of all NAGs were less than 7 except *GmGS1*, indicating acidic nature of these proteins. Most soybean NAG proteins exhibited an instability index (II) value below 40, suggesting their stability. Subcellular location predications suggested that the majority of the NAGs are located in the cytoplasm and chloroplast, with a few genes also present in the mitochondria and nucleus, as shown in Table S4.

The phylogenetic relationships of NAGs in six legumes, including soybean, were investigated using the neighbor joining method. The *GS* genes were divided into four groups based on the

phylogenetic tree, with groups 1 and 2 consisting of cytoplasmic *GS* genes, and groups 3 and 4 containing chloroplast *GS* genes (Fig. 1A). The phylogenetic analysis revealed that GOGAT family genes could be classified into three groups with group A consisting 7 *GOGAT* genes, all from the Fd-GOGAT subfamily and groups B and C containing Fd-GOGAT subfamily genes from various plants (Fig. 1B). Additionally, NR genes were categorized into three groups (Fig. 1C).

Chromosomal location and gene duplication analysis of soybean NAGs

The distribution of the 20 NAGs in soybean was uneven across 13 out of the 20 soybean chromosomes. Chromosome14 was notable for containing four genes, while the other chromosomes generally contained one or two genes each. To investigate the role of gene duplication in the expansion of soybean NAGs, we conducted an annotation and analysis of the intraspecific collinearity of these genes. Our analysis revealed that within the *GmGS*, *GmGOGAT*, and *GmNR* genes, there were 14, 4, and 3 segmental duplications, respectively (Fig.2). Additionally, we identified two tandem duplications. One was located on chromosome Gm02 (*GmGS1*/*GmGS2*), and the other on chromosome Gm14 (*GmNIA3*/*GmNIA4*). These tandem duplications involved genes from the same family and were positioned very close to each other on their respective chromosomes (Fig.S1).

Our research was further supported by the observation that the tandem duplication-arranged NAGs, such as *GmNIA3* and *GmNIA4*, clustered together in the phylogenetic tree (Fig. 1), indicating a close evolutionary relationship between these duplicated genes. The presence of segmental and tandem duplications highlights the significance of these mechanisms in molding the genetic landscape of soybean and potentially contributing to its adaptation and functional diversity in nitrogen assimilation processes.

To explore whether selective constraints influenced the duplicated genes, we analyzed the Ka/Ks ratio using the full-length protein sequences of the NAGs. The pairwise comparison revealed a Ka/Ks ratio range of 0.04–0.19, which is notably less than 1, indicating that the soybean NAGs underwent purifying selection pressure with limited functional divergence. Moreover, the average Ka/Ks value for the *GS* gene family members was lower than that of the

GOGAT and *NR* gene families, implying a slower evolution of *GmGS*s and highlighting their higher conservation level. (Table S5).

Structure of genes and conserved motifs in NAGs encoded proteins

The structures of the *GmGS*, *GmGOGAT*, and *GmNR* genes are illustrated in Figure 3 to emphasize their structural diversity. *GmGS* genes exhibit significant variation in the number and length of introns, with *GmGS1* having 5 introns and *GmGS5* containing 17 introns, while other GmGS genes typically have 11–13 introns (Fig. 3A). The arrangement of introns differs among phylogenetic groups, enhancing the structural and functional diversity of *GmGS* genes. Moreover, cytoplasmic *GmGS* genes feature two types of introns (phase-0, phase-1), whereas chloroplastic *GmGS* genes contain phase-0, phase-1, and phase-2 introns, except for *GmGS1*.

In contrast, the distribution of introns in *GmGOGAT* genes reveals that all *GmNADH-GOGAT* genes have 21 introns, whereas *GmFd-GOGAT* genes have 32 introns with three different phase types (Fig. 3B). The *GmNR* genes, on the other hand, have a relatively small number of introns compared to the *GS* and *GOGAT* gene families, typically possessing 3–4 introns with phase-0 and phase-2 introns (Fig. 3C).

In addition, the conserved motifs of the NAGs in soybean were identified based on their amino acid sequences using MEME software. Each family had 10 motifs identified. Genes with close phylogenetic relationships exhibited high similarity conserved motif composition. As shown in Fig3B, all genes in GOAGT family contained the 10 motifs, implies that this family may exhibit highly conserved functions or possibly functional redundancy among its genes, similar to the NR gene family (Fig3C). Notably, the *GS* gene family displayed some variation in motif arrangement, with *GmGS1* and *GmGS5* deviating from the pattern. *GmGS1* had Motif 1 and Motif 3, while *GmGS5* only had Motif 6. These findings suggest functional divergence among nitrogen assimilation genes in soybean.

Synteny analysis of soybean NAGs

To investigate the evolutionary relationships of *GS*, *GOGAT*, and *NR* gene families across different species, we carried out an interspecies collinearity analysis involving *G. max*, *A. thaliana*, *C.arietinum*, and *P. acutifolius* (Fig. 4). The GmGS family members showed the

highest number of collinear pairs with *P. acutifolius*, totaling 16 pairs, indicating a close evolutionary relationship between these two species. Additionally, five collinear gene pairs were identified between the *GmGOGAT* and *AtGOGAT* genes. Conversely, the syntenic relationships of *GmNR* genes with genes from other species were predominantly observed on two or three chromosomes. These results show that the presence of multiple collinear gene pairs among the three species were inferred to be genetic copies with lineage-specific amplification.

Analysis of cis-acting elements in the promoter regions of the soybean NAGs

An analysis was conducted on the 1500 bp upstream of the transcription start site of *GS*, *GOGAT*, and *NR* genes in soybean using the PlantCARE database (Fig. 5). The study revealed that these promoter regions contain three main types of cis-acting elements: light-responsive elements, hormone-responsive elements, and stress-responsive elements. Additionally, five types of hormone-responsive elements were identified, including gibberellin, abscisic acid, auxin, salicylic acid and jasmonic acid-responsive elements. Functional elements related to stress, such as low-temperature responsive elements, were also found. These findings indicate that the *GmGS*, *GmGOGAT*, and *GmNR* genes likely play crucial roles in various physiological processes in soybeans, such as plant growth, development, and responses to different stresses.

Protein-protein interaction network

As shown in Figure 6, NAGs engage in interactions with one another. The most effective interaction was observed between *GmGSs* and *GmNRs*. PPI enrichment p-value <1.0e-16 indicates that the proteins are at least partially biologically connected, as a group. The potential interactions among NAGs could offer valuable insights for investigating their biological roles.

Tissue-specific expression profiles of soybean NAGs

Transcriptomic data from the Phytozome database was utilized to investigate the constant expression of the *GmGS*, *GmGOAGT*, and *GmNR* genes. The analysis involved transcriptome profiles from eight distinct tissue samples of soybean. The resulting expression data of *GmGS*, *GmGOAGT*, and *GmNR* were log-transformed and visualized in a heatmap. Among the genes examined, *GmFd-GOGAT1*, *GmGS6*, *GmGS10*, and *GmNIA1* showed relatively distinct high expression patterns across all eight tissues, suggesting potential involvement in the vegetative

organs of *G. max* (Fig.7). In addition, *GmNADH-GOGAT1* exhibits an unique expression pattern with significant tissue specificity, primarily in root regions such as root tip, lateral root, and flowers (Fig. 7B). In contrast, its expression level is significantly lower in the stem, shoot, and leaf. On the other hand, certain genes like *GmNADH-GOGAT2*, *GmGS7*, *GmGS9*, *GmFd-GOGAT2*, and *GmNIA3* displayed tissues-specific high expression level in the ~~stem or~~ leaf. These results highlight tissue-specific regulation and potential functional roles of these genes in soybean.

Analysis of GS, GOGAT, and NR gene expression under abiotic and nitrogen stresses

The study aimed to investigate the response of seven soybean NAGs to nitrate, salt, and nitrate-drought stress conditions using qRT-PCR. Results showed that *GmNADH-GOGAT3*, *GmGS4*, and *GmGS6* did not exhibit significant changes in transcript levels under nitrate stress, while *GmGS10* was significantly induced under both high and low nitrate treatments. These results suggest that NAGs in soybean may play a potential role in responding to nitrate stress (Fig. 8A). *GmNADH-GOGAT1* and *GmGS4* were significantly up-regulated in response to salt stress over time. In addition, *GmNADH-GOGAT3* showed time-dependent regulation patterns, which were initially significantly up-regulated at 24 hours and later down-regulated at 48 hours, suggesting a possible time-dependent regulation mechanism in response to salt stress (Fig. 8B).

Under drought-nitrate stress treatments (D-HN, D-NN, and D-LN), the expression pattern of these genes was complex (Fig. 8C). For example, *GmNADH-GOGAT1*, *GmGS4, GS6*, and *GmNIA2* were significantly down-regulated in response to all drought-nitrate treatments. Additionally, *GmNADH-GOGAT3* was down-regulated under D-HN treatment but up-regulated under D-NN treatment. Interestingly, *GmFd-GOGAT1* and *GmGS10* were significantly up-regulated after all drought-nitrate treatments compared to the control, suggesting a diverse stress response mechanism among NAG genes in soyabean. Overall, these findings highlight the crucial role of NAGs in nitrogen and abiotic stress responses in soybean.

Discussion:

The *GS*, *GOGAT*, and *NR*, which are among the most crucial NAGs, have been confirmed to be involved in various biological processes, including plant stress tolerance [19, 32]. While these gene families have been extensively studied in several plant species, knowledge of their

functions in soybean remains limited [33, 34]. In this study, 10 *GS*, 6 *GOGAT*, and 4 *NR* genes were identified and characterized through a comprehensive analysis of the soybean genome. We also investigated their phylogeny, duplication patterns, protein sequences, and expression profiles under nitrogen, drought-nitrogen, and salt stress conditions. A comparative phylogenetic analysis revealed a high degree of conservation in NAGs across various legume species including *G.max*, emphasizing the close relationship among nitrogen assimilation genes in diverse crops. The classification of NAGs was consistent with gene structures, motif distributions, and existing literature [20, 21].

Tandem and segmental duplications are thought to have played a significant role in the expansion of gene families over course of evolution [35]. The studies in *B. napus* has shown that the expansion of the NAGs family was mainly driven by segmental duplication [19]. These findings are in line with our current study. A majority of duplicated NAGs pairs were identified as a result of segmental duplication, indicating that it was the primary mechanism driving the expansion of NAGs in soybean during evolution (Figure 2; Table S2). Additional examination of the evolutionary selective pressure revealed that NAGs experienced a strong purifying selection during evolution, implying that their functions may have been conserved over time (Table S2).

The analysis of exon-intron organization and motif patterns within gene families can provide valuable insights into evolutionary relationships [36]. In the current study, the gene structure and motif analysis revealed that genes within the same group tended to have a similar number of introns, similar intron phases, and shared conserved motifs, indicating a pattern of clustering based on these features and strongly supporting the results of the phylogenetic analysis. This observation suggests that these genes share common functions and have evolved from a common ancestor [37, 38].

The interaction between RNA polymerase and the promoter is a crucial event at the onset of transcription, a crucial process in gene expression. The structure of the promoter influences both the binding affinity of the RNA polymerase and gene expression level [37, 39]. The analysis of cis-acting elements in the promoter regions of the *GS*, *GOGAT*, and *NR* genes in soybean provides valuable insights into the potential regulatory mechanisms of these genes. Numerous cis-elements related to light responsiveness were identified, suggesting that these genes might be light-regulated and in plant growth and development processes that are influenced by light

conditions. Previous studies have highlighted the role of light in regulating nitrogen accumulation [34]. Additionally, hormone responses and stress tolerance cis-elements were found in soybean NAGs. This implies a role for the soybean NAGs in modulating hormone response and stress tolerance. Our observations align with previous research on rice and pecan [21, 40].

Moreover, we assessed the expression patterns of *GS*, *GOGAT*, and *NR* genes in different tissues, such as root, stem, and flower, using RNA-seq data. The results revealed that these NAGs have a wide range of expression across these tissues, with *GmGS6* exhibiting the highest expression levels across all tissues. Previous studies have reported the widespread expression of NAG family members in various tissues and organs, indicating their involvement in regulating plant growth and development [33].

Furthermore, the expression patterns of NAGs were validated using RT-qPCR (Fig. 9), indicating that the NAGs might play a crucial role in responding to a wide range of abiotic stresses and contributing to the development of resistance mechanisms, aligning with previous study [20]. The expression of selected *GmGS*, *GmGOGAT*, and *GmNR* genes in response to different nitrate treatments highlights their potential crucial role in soybean plants' nitrate response. Specifically, under high nitrate (HN) treatment conditions, all selected NAGs showed up-regulated expression patterns, aligning with the results of Balotf et al. [41], who observed up regulation of wheat NAG expression in response to high nitrate (50 mM KNO_3) treatment. Similarly, Qiao et al. [21]reported significant up regulation of pecan *GS* genes under high nitrate concentrations. These studies collectively suggest that these genes may play a vital role in the response of plants to high nitrate levels, potentially through their involvement in the assimilation and metabolism of nitrogen compounds. In contrast, the NAGs showed complex expression under low nitrate treatments (6 mM KNO_3); *GmGSs* were up-regulated while *GmGOGAT* s were down-regulated, similar to studies in *B. napus*. The observed different expression patterns of NAGs genes under different nitrogen treatments suggest that these genes have distinct reactions and regulatory mechanisms in response to varying nitrogen treatment conditions. However, the precise mechanisms underlying these differential gene expression patterns and their implications for plant adaptation to nitrate stress require further elucidation. This would contribute to our understanding of plant stress responses and may facilitate the development of strategies for

enhancing crop resilience to nitrate stress. From the above, we can make a hypothesis that up-regulated genes are positively regulated and down-regulated genes are negatively regulated under different stresses.

The analysis of gene expression under salt and drought-nitrate stress conditions reveals interesting patterns and potential mechanisms of plant response to these environmental stresses. The majority of the selected genes exhibited increased expression under salt stress over time, which aligns with previous research highlighting the up regulation of *GS*, *GOGAT*, and *NR* gene in rice and *Arabidopsis* in responses to salt stress [5, 20]. In contrast, the response of NAGs to drought-nitrate stress appears to be more complex. The significant down-regulation of *GmNADH-GOGAT1*, *GmNIA2, GmGS4, and GmGS10* under drought-nitrate treatments is suggestive of a common stress response mechanism. Drought stress can affect the uptake and transport of nitrate in plants. Under drought conditions, the expression of nitrate transporters may be down regulated, reducing the availability of nitrate. This can further contribute to the decrease in GS, NR, and GOGAT activity observed during drought-nitrate interaction treatment. However, the up-regulation of *GmFd-GOGAT1* and *GmGS10* under all drought-nitrate treatments is particularly intriguing. This deviates from the general trend observed in other genes and hints at a unique stress response mechanism. This is a novel finding that has not been reported in previous studies and warrants further investigation.

Conclusions:

The study analyzed the *GS*, *GOGAT*, and *NR* genes in *Glycine max*, including their phylogenetic relationships, chromosomal distribution, gene structure, and tissue-specific expression. The study also explored the interaction of the nitrogen pathway with abiotic stresses using RT-qPCR under nitrogen, salt, and drought-nitrogen stresses. The findings suggest these genes play a crucial role in nitrogen metabolism and are significantly influenced by abiotic stress factors. The differential expression of these genes under different stress conditions suggests their potential role in stress tolerance mechanisms. This research opens new avenues for understanding the complex interplay between nitrogen metabolism and stress response in plants, with further studies needed to understand the precise regulatory mechanisms and explore the potential of these genes in improving stress tolerance in crops. mention specific problem/stress the present study going to address.

List of abbreviations

Fd-GOGAT: Ferredoxin glutamate synthase; **FPKM**: Fragments per kilobase of transcript per Million mapped reads; **GS**: Glutamine synthetase; **GOGAT**: Glutamate synthase ; **MW**: Molecular weight; **MS**: Murashige and Skoog liquid medium; **NADH-GOGAT**: Nicotinamide adenine dinucleotide (NAD) + hydrogen (H) glutamate synthetase; **NAG**: Nitrogen assimilation gene; **NR**: Nitrate reductase; **NUE**: Nitrogen use efficiency; **qRT-PCR:** Quantitative real-time polymerase chain reaction; **II**: Instability index; **pI**: Isoelectric point .

References:

1. Gaudinier A, Rodriguez-Medina J, Zhang L, Olson A, Liseron-Monfils C, Bågman A-M, Foret J, Abbitt S, Tang M, Li BJN: **Transcriptional regulation of nitrogen-associated metabolism and growth**. 2018, **563**(7730):259-264.
2. Liu X, Hu B, Chu CJJoG, Genomics: **Nitrogen assimilation in plants: current status and future prospects**. 2022, **49**(5):394-404.
3. Liu L, Wang J, Han Z, Sun X, Li H, Zhang J, Lu YJAPP: **Molecular analyses of tomato GS, GOGAT and GDH gene families and their response to abiotic stresses**. 2016, **38**:1-14.
4. Sajjad N, Bhat EA, Shah D, Manzoor I, Noor W, Shah S, Hassan S, Ali R: **Nitrogen uptake, assimilation, and mobilization in plants under abiotic stress**. In: *Transporters and Plant Osmotic Stress.* Elsevier; 2021: 215-233.
5. Tang X, Peng Y, Li Z, Guo H, Xia X, Li B, Yin WJIJoMS: **The regulation of nitrate reductases in response to abiotic stress in Arabidopsis**. 2022, **23**(3):1202.
6. García-Gutiérrez Á, Cánovas FM, Ávila CJBg: **Glutamate synthases from conifers: gene structure and phylogenetic studies**. 2018, **19**(1):1-12.
7. Lebedev V, Korobova A, Shendel G, Kudoyarova G, Shestibratov K: **Effect of glutamine synthetase gene overexpression in birch (Betula pubescens) plants on auxin content and rooting in vitro**. In: *Doklady biochemistry and biophysics: 2018*. Springer: 143-145.
8. Yin H, Yang F, He X, Du X, Mu P, Ma WJTCJ: **Advances in the functional study of glutamine synthetase in plant abiotic stress tolerance response**. 2022, **10**(4):917-923.
9. Fortunato S, Nigro D, Lasorella C, Marcotuli I, Gadaleta A, de Pinto MCJB: **The Role of Glutamine Synthetase (GS) and Glutamate Synthase (GOGAT) in the Improvement of Nitrogen Use Efficiency in Cereals**. 2023, **13**(12):1771.
10. Zang X, Geng X, He K, Wang F, Tian X, Xin M, Yao Y, Hu Z, Ni Z, Sun QJFips: **Overexpression of the wheat (Triticum aestivum L.) TaPEPKR2 gene enhances heat and dehydration tolerance in both wheat and Arabidopsis**. 2018, **9**:1710.
11. Cai H, Zhou Y, Xiao J, Li X, Zhang Q, Lian XJPcr: **Overexpressed glutamine synthetase gene modifies nitrogen metabolism and abiotic stress responses in rice**. 2009, **28**:527-537.
12. El-Khatib RT, Hamerlynck EP, Gallardo F, Kirby EGJTp: **Transgenic poplar characterized by ectopic expression of a pine cytosolic glutamine synthetase gene exhibits enhanced tolerance to water stress**. 2004, **24**(7):729-736.
13. Lv X, Yu P, Deng W, Li YJPP, Biochemistry: **Transcriptomic analysis reveals the molecular adaptation to NaCl stress in Zostera marina L**. 2018, **130**:61-68.

14. Sun Y, Wang D, Shi M, Gong Y, Yin S, Jiao Y, Guo SJFiPS: **Genome-wide identification of actin-depolymerizing factor gene family and their expression patterns under various abiotic stresses in soybean (Glycine max)**. 2023, **14**.
15. Zhan Y, Li H, Sui M, Zhao X, Jing Y, Luo J, Teng W, Qiu L, Zheng H, Li WJIC *et al*: **Genome wide association mapping for tocopherol concentration in soybean seeds across multiple environments**. 2020, **154**:112674.
16. Modgil R, Tanwar B, Goyal A, Kumar VJOha, applications f: **Soybean (Glycine max)**. 2021:1-46.
17. Bu M, Li Y, Wang S, Huang C, Wu C, Liu SJBJoB: **Nitrogen stress alters trade-off strategies between reproduction and vegetative growth in soybean**. 2023:1-11.
18. Gavili E, Moosavi AA, Haghighi AAKJIc, products: **Does biochar mitigate the adverse effects of drought on the agronomic traits and yield components of soybean?** 2019, **128**:445-454.
19. He X, Zhang H, Ye X, Hong J, Ding GJP: **Nitrogen assimilation related genes in Brassica napus: Systematic characterization and expression analysis identified hub genes in multiple nutrient stress responses**. 2021, **10**(10):2160.
20. Li Z, Zhu M, Huang J, Jiang S, Xu S, Zhang Z, He W, Huang WJIJoMS: **Genome-Wide Comprehensive Analysis of the Nitrogen Metabolism Toolbox Reveals Its Evolution and Abiotic Stress Responsiveness in Rice (Oryza sativa L.)**. 2022, **24**(1):288.
21. Qiao Z, Chen M, Ma W, Zhu K, Zhao J, Tan P, Peng F: **Genome-Wide Identification and Expression Analysis of Gs and Gogat Gene Family in Pecan (Carya illinoinensis) under Different Nitrogen Forms**. 2023.
22. Tamura K, Stecher G, Kumar SJMb, evolution: **MEGA11: molecular evolutionary genetics analysis version 11**. 2021, **38**(7):3022-3027.
23. Zhou T, Xu K, Zhao F, Liu W, Li L, Hua Z, Zhou XJB: **itol. toolkit accelerates working with iTOL (Interactive Tree of Life) by an automated generation of annotation files**. 2023, **39**(6):btad339.
24. Bailey TL, Johnson J, Grant CE, Noble WSJNar: **The MEME suite**. 2015, **43**(W1):W39-W49.
25. Chen C, Chen H, Zhang Y, Thomas HR, Frank MH, He Y, Xia RJMp: **TBtools: an integrative toolkit developed for interactive analyses of big biological data**. 2020, **13**(8):1194-1202.
26. Goodstein DM, Shu S, Howson R, Neupane R, Hayes RD, Fazo J, Mitros T, Dirks W, Hellsten U, Putnam NJNar: **Phytozome: a comparative platform for green plant genomics**. 2012, **40**(D1):D1178-D1186.
27. Rombauts S, Déhais P, Van Montagu M, Rouzé PJNar: **PlantCARE, a plant cis-acting regulatory element database**. 1999, **27**(1):295-296.
28. Szklarczyk D, Kirsch R, Koutrouli M, Nastou K, Mehryary F, Hachilif R, Gable AL, Fang T, Doncheva NT, Pyysalo SJNar: **The STRING database in 2023: protein–protein association networks and functional enrichment analyses for any sequenced genome of interest**. 2023, **51**(D1):D638-D646.
29. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker TJGr: **Cytoscape: a software environment for integrated models of biomolecular interaction networks**. 2003, **13**(11):2498-2504.
30. Paz MM, Martinez JC, Kalvig AB, Fonger TM, Wang KJPcr: **Improved cotyledonary node method using an alternative explant derived from mature seed for efficient Agrobacterium-mediated soybean transformation**. 2006, **25**:206-213.
31. Guo B, Dai Y, Chen L, Pan Z, Song LJG: **Genome-wide analysis of the soybean root transcriptome reveals the impact of nitrate on alternative splicing**. 2021, **11**(7):jkab162.
32. Li S, Zhou L, Addo-Danso SD, Ding G, Sun M, Wu S, Lin SJSr: **Nitrogen supply enhances the physiological resistance of Chinese fir plantlets under polyethylene glycol (PEG)-induced drought stress**. 2020, **10**(1):7509.

33. Wang Y, Hua Y-p, Zhou T, Huang J-y, Yue C-pJMBR: **Genomic identification of nitrogen assimilation-related genes and transcriptional characterization of their responses to nitrogen in allotetraploid rapeseed**. 2021, **48**:5977-5992.
34. Liang Y, Cossani CM, Sadras VO, Yang Q, Wang ZJFiPS: **The interaction between nitrogen supply and light quality modulates plant growth and resource allocation**. 2022, **13**:864090.
35. Jiang S-Y, González JM, Ramachandran SJPO: **Comparative genomic and transcriptomic analysis of tandemly and segmentally duplicated genes in rice**. 2013, **8**(5):e63551.
36. Liu Q, Wang C, Cui Q, Fan Y, Zhang J, Rao GJA: **Genome-wide analysis of the polyphenol oxidase gene family in Olea europaea provides insights into the mechanism of enzymatic browning in olive fruit**. 2023, **12**(9):1661.
37. Wang T, Long C, Chang M, Wu Y, Su S, Wei J, Jiang S, Wang X, He J, Xing DJSR: **Genome-wide identification of the B3 transcription factor family in pepper (Capsicum annuum) and expression patterns during fruit ripening**. 2024, **14**(1):2226.
38. Li D, Liu P, Yu J, Wang L, Dossa K, Zhang Y, Zhou R, Wei X, Zhang XJBpb: **Genome-wide analysis of WRKY gene family in the sesame genome and identification of the WRKY genes involved in responses to abiotic stresses**. 2017, **17**(1):1-19.
39. Shang H, Li W, Zou C, Yuan YJJoIPB: **Analyses of the NAC transcription factor gene family in Gossypium raimondii Ulbr.: chromosomal location, structure, phylogeny, and expression patterns**. 2013, **55**(7):663-676.
40. Rohilla P, Yadav JPJJoPGR: **Promoter of Nitrate Reductase Gene Behaves Differently Under Salinity Stress in Contrasting Rice Genotype**. 2023:1-11.
41. Balotf S, Kavoosi G, Kholdebarin BJB, biochemistry a: **Nitrate reductase, nitrite reductase, glutamine synthetase, and glutamate synthase expression and activity in response to different nitrogen sources in nitrogen‐starved wheat seedlings**. 2016, **63**(2):220-229.

Figure 1

Unrooted phylogenetic tree of *GS* genes (A) *GOGAT* genes (B), and *NR* genes (C) in *G.max*, *C. arietinum*, *L. japonicas*, *P. lunatus*, *P. vulgaris*, and *V. unguiculata*.

The deduced full-length amino acid sequences were utilized to construct the phylogenetic tree using MEGA 11 software through a neighbor-joining method with 1000 bootstrap replicates. Various groups are distinguished by different colors.

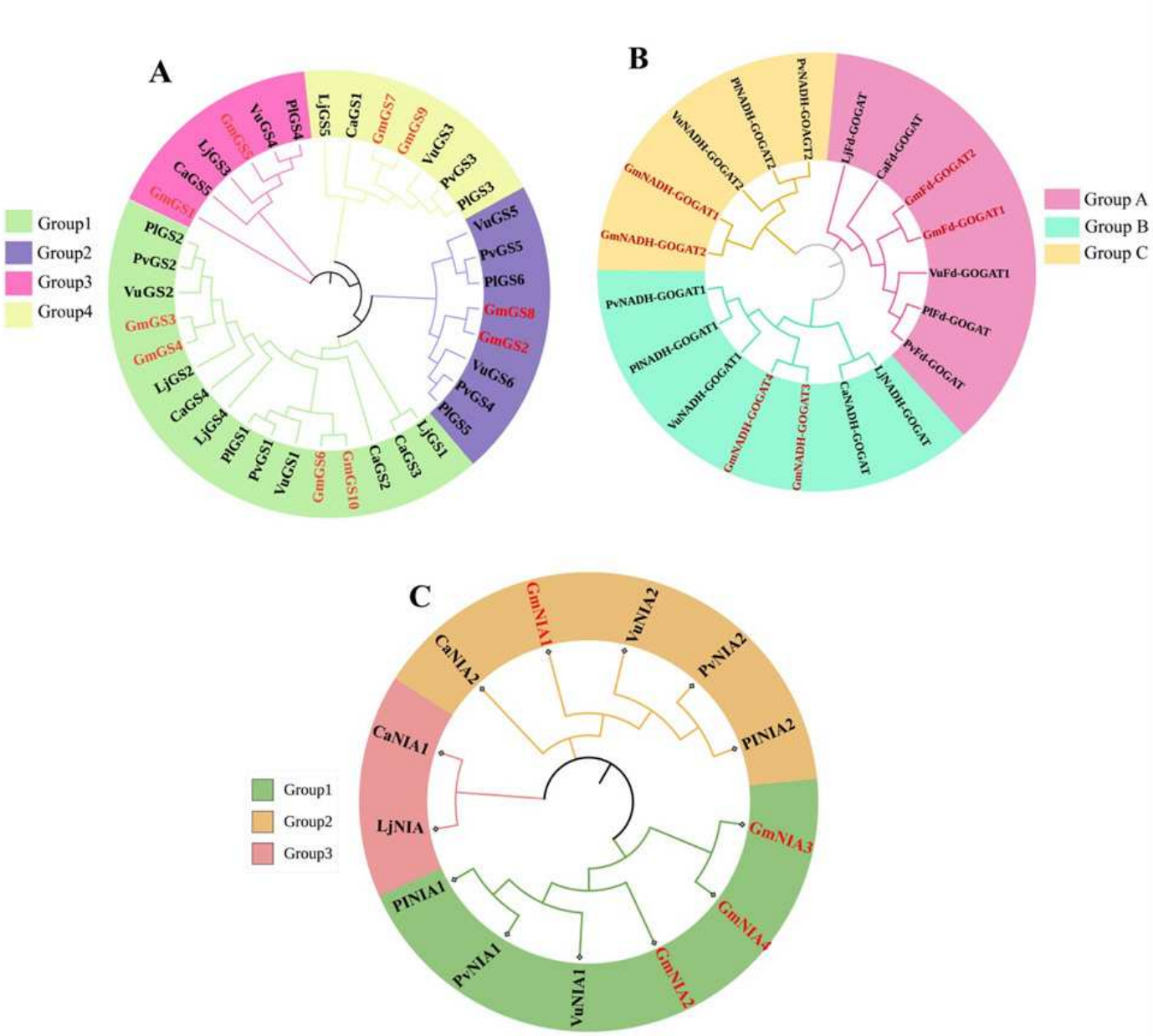

A
Group1
Group2
Group3
Group4
LjGS5
CaGS1
GmGS7
GmGS9
VuGS3
PvGS3
PlGS3
VuGS5
PvGS5
PlGS6
GmGS8
GmGS2
VuGS6
PvGS4
PlGS5
LjGS1
CaGS3
CaGS2
GmGS10
GmGS6
VuGS1
PvGS1
PlGS1
LjGS4
CaGS4
LjGS2
GmGS4
GmGS3
VuGS2
PvGS2
PlGS2
GmGS1
CaGS5
LjGS3
GmGS5
VuGS4
PlGS4
B
Group A
Group B
Group C
LjFd-GOGAT
CaFd-GOGAT
GmFd-GOGAT2
GmFd-GOGAT1
VuFd-GOGAT1
PlFd-GOGAT
PvFd-GOGAT
LjNADH-GOGAT
CaNADH-GOGAT
GmNADH-GOGAT3
GmNADH-GOGAT4
VuNADH-GOGAT1
PlNADH-GOGAT1
PvNADH-GOGAT1
GmNADH-GOGAT2
GmNADH-GOGAT1
VuNADH-GOGAT2
PlNADH-GOAGT2
PvNADH-GOGAT2
C
Group1
Group2
Group3
GmNIA1
VuNIA2
PvNIA2
PlNIA2
GmNIA3
GmNIA4
GmNIA2
VuNIA1
PvNIA1
PlNIA1
LjNIA
CaNIA1
CaNIA2

Figure 2

Genomic distribution and duplication of the *GmGS*, *GmGOGAT*, and *GmNR* genes across 20 chromosomes of soybean.

The colorful lines indicate duplicated *GmGS*, *GmGOGAT*, and *GmNR* genes pairs.

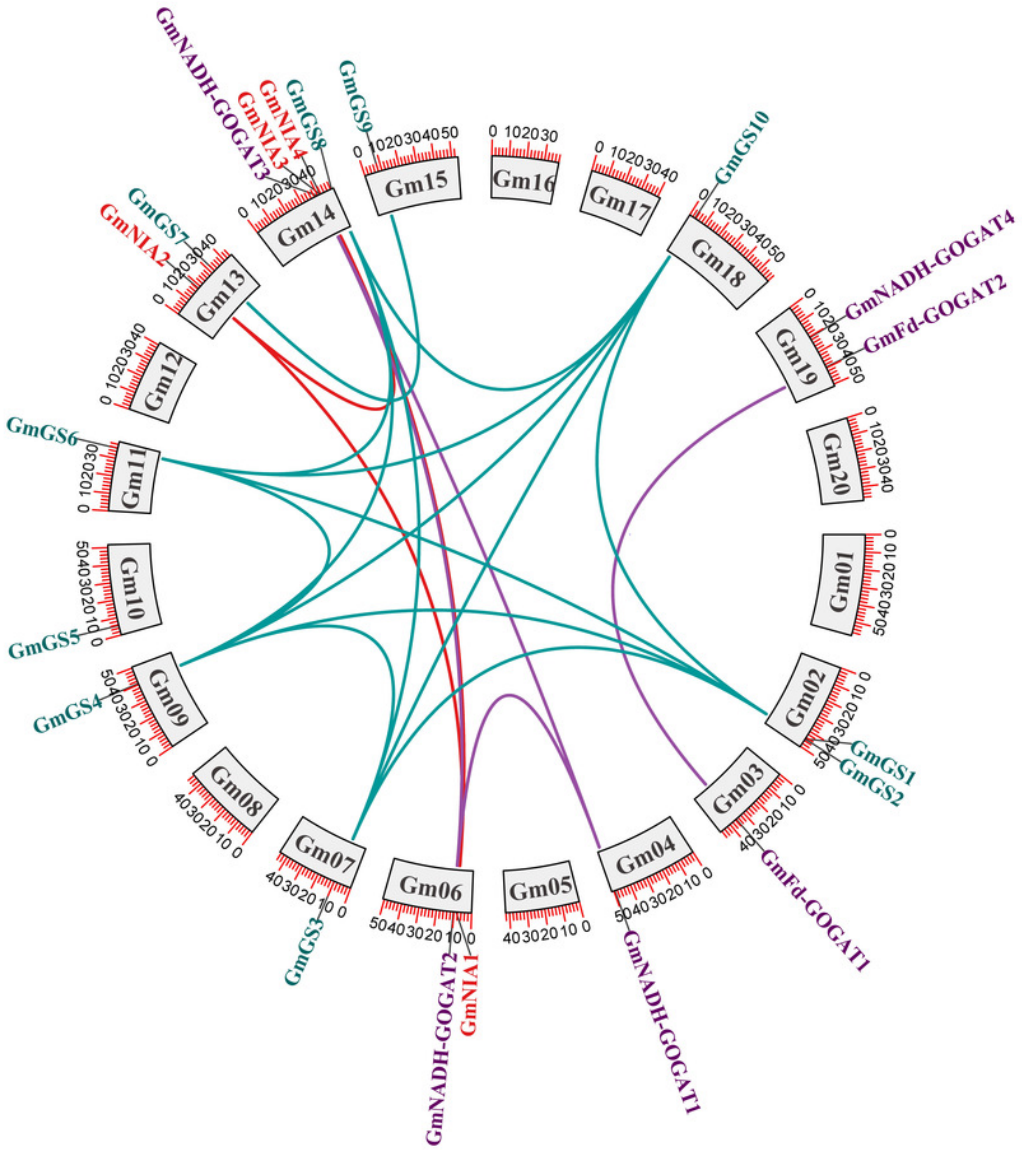


Figure 3

(I) Phylogenetic relationships, (II) Motif compositions, (III) Gene structure of *GmGS* (A), *GmGOGAT* (B), and *GmNR* (C).

Different colored boxes represent different motifs. The green boxes represent UTR and the yellow boxes represent exons. In terms of introns, a phase 0 intron does not disrupt a codon, a phase1 disrupts a codon between the first and second bases, and the phase 2 intron located after second nucleotide of a codon.

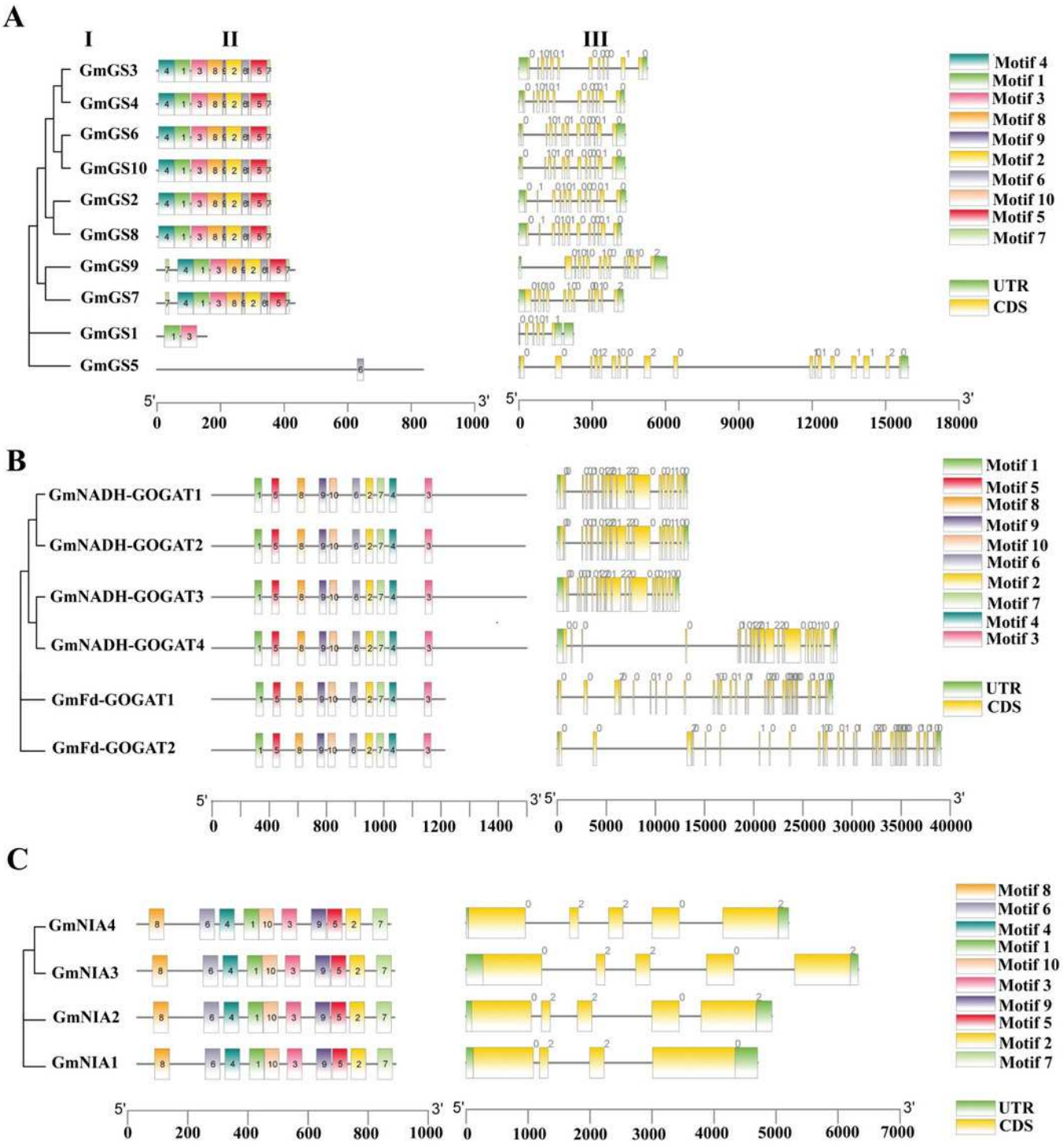

A
I
II
III
GmGS3
GmGS4
GmGS6
GmGS10
GmGS2
GmGS8
GmGS9
GmGS7
GmGS1
GmGS5
Motif 4
Motif 1
Motif 3
Motif 8
Motif 9
Motif 2
Motif 6
Motif 10
Motif 5
Motif 7
UTR
CDS
5'
3'
0
200
400
600
800
1000
0
3000
6000
9000
12000
15000
18000
B
GmNADH-GOGAT1
GmNADH-GOGAT2
GmNADH-GOGAT3
GmNADH-GOGAT4
GmFd-GOGAT1
GmFd-GOGAT2
Motif 1
Motif 5
Motif 8
Motif 9
Motif 10
Motif 6
Motif 2
Motif 7
Motif 4
Motif 3
UTR
CDS
0
400
800
1000
1200
1400
0
5000
10000
15000
20000
25000
30000
35000
40000
C
GmNIA4
GmNIA3
GmNIA2
GmNIA1
Motif 8
Motif 6
Motif 4
Motif 1
Motif 10
Motif 3
Motif 9
Motif 5
Motif 2
Motif 7
UTR
CDS
0
200
400
600
800
1000
0
1000
2000
3000
4000
5000
6000
7000

Figure 4

Synteny analysis of NAGs between *Glycine max* and *A. thaliana*, *C.arietinum*, and *P.acutifolius*.

The gray lines in the background represent the collinear blocks within the soybean and other plant genomes. Pink lines highlight syntenic glutamine synthetase (*GS*) gene pairs, blue lines indicate syntenic glutamate synthase (*GOGAT*) gene pairs, and yellow lines represent syntenic nitrate reductase (*NR*) genes pairs.

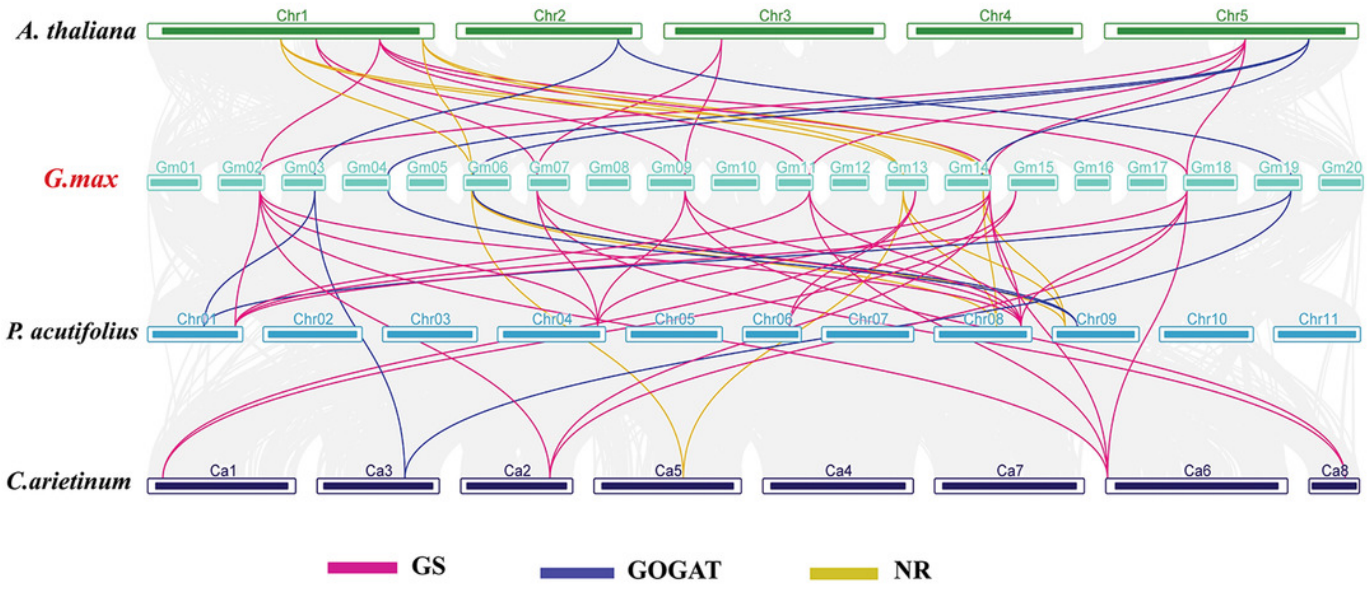


Figure 5

Cis-element analysis on the promoter region of *GmGS* (A) *GmGOGAT* (B), and *GmNR* (C).

The potential cis-regulatory elements in the 1500 bp promoter regions were predicted by PlantCARE software. The elements related to different functional categories were represented by different colors.

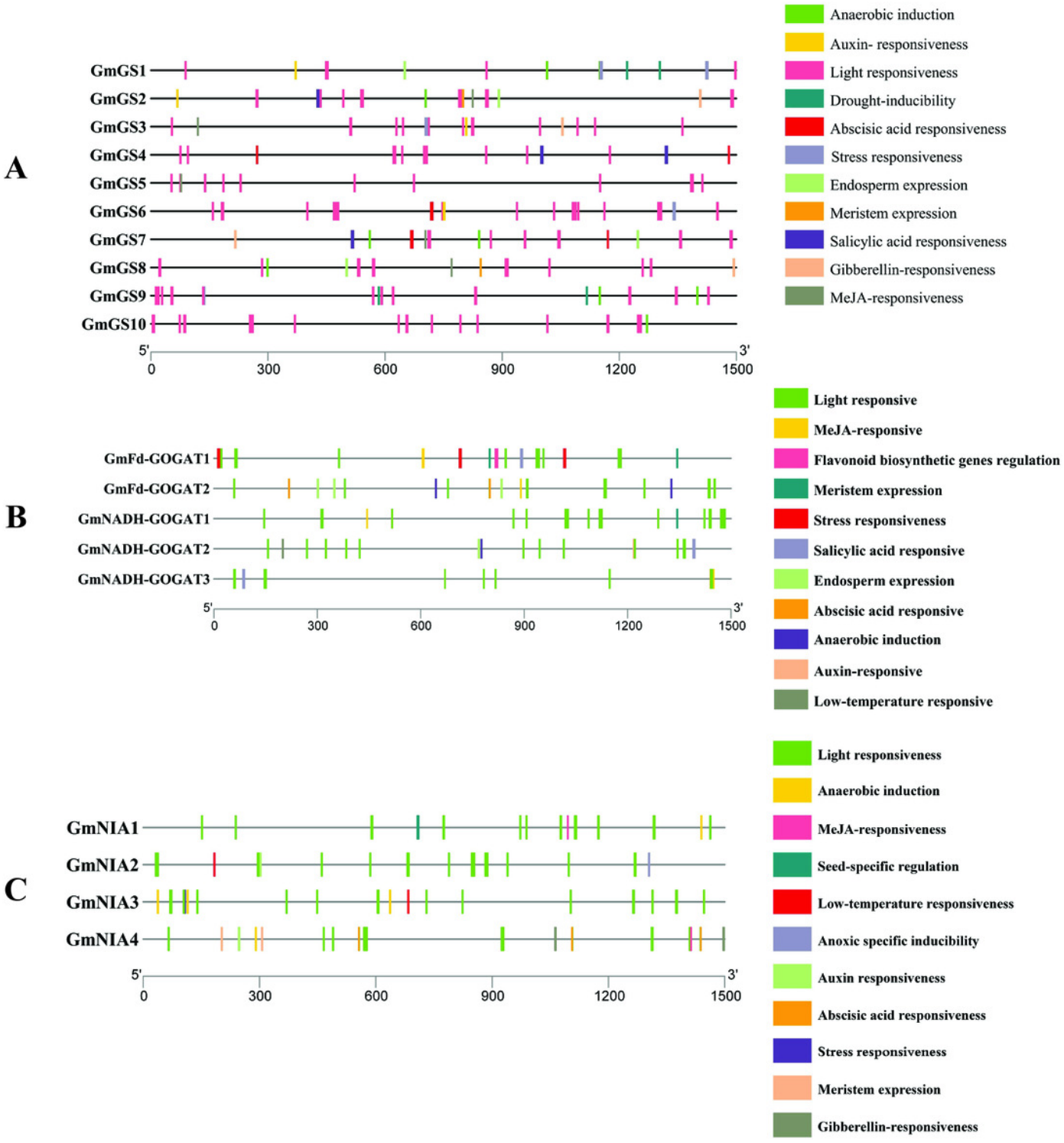

A
GmGS1
GmGS2
GmGS3
GmGS4
GmGS5
GmGS6
GmGS7
GmGS8
GmGS9
GmGS10
5'
3'
0
300
600
900
1200
1500
Anaerobic induction
Auxin- responsiveness
Light responsiveness
Drought-inducibility
Abscisic acid responsiveness
Stress responsiveness
Endosperm expression
Meristem expression
Salicylic acid responsiveness
Gibberellin-responsiveness
MeJA-responsiveness
B
GmFd-GOGAT1
GmFd-GOGAT2
GmNADH-GOGAT1
GmNADH-GOGAT2
GmNADH-GOGAT3
5'
3'
0
300
600
900
1200
1500
Light responsive
MeJA-responsive
Flavonoid biosynthetic genes regulation
Meristem expression
Stress responsiveness
Salicylic acid responsive
Endosperm expression
Abscisic acid responsive
Anaerobic induction
Auxin-responsive
Low-temperature responsive
C
GmNIA1
GmNIA2
GmNIA3
GmNIA4
5'
3'
0
300
600
900
1200
1500
Light responsiveness
Anaerobic induction
MeJA-responsiveness
Seed-specific regulation
Low-temperature responsiveness
Anoxic specific inducibility
Auxin responsiveness
Abscisic acid responsiveness
Stress responsiveness
Meristem expression
Gibberellin-responsiveness

Figure 6

The protein-protein interaction network for soybean NAGs.

Various active interaction sources were indicated by different line colors: blue for databases, pink for experiments, green for neighborhood, yellow for text mining, black for co-expression, and blue for protein homology.

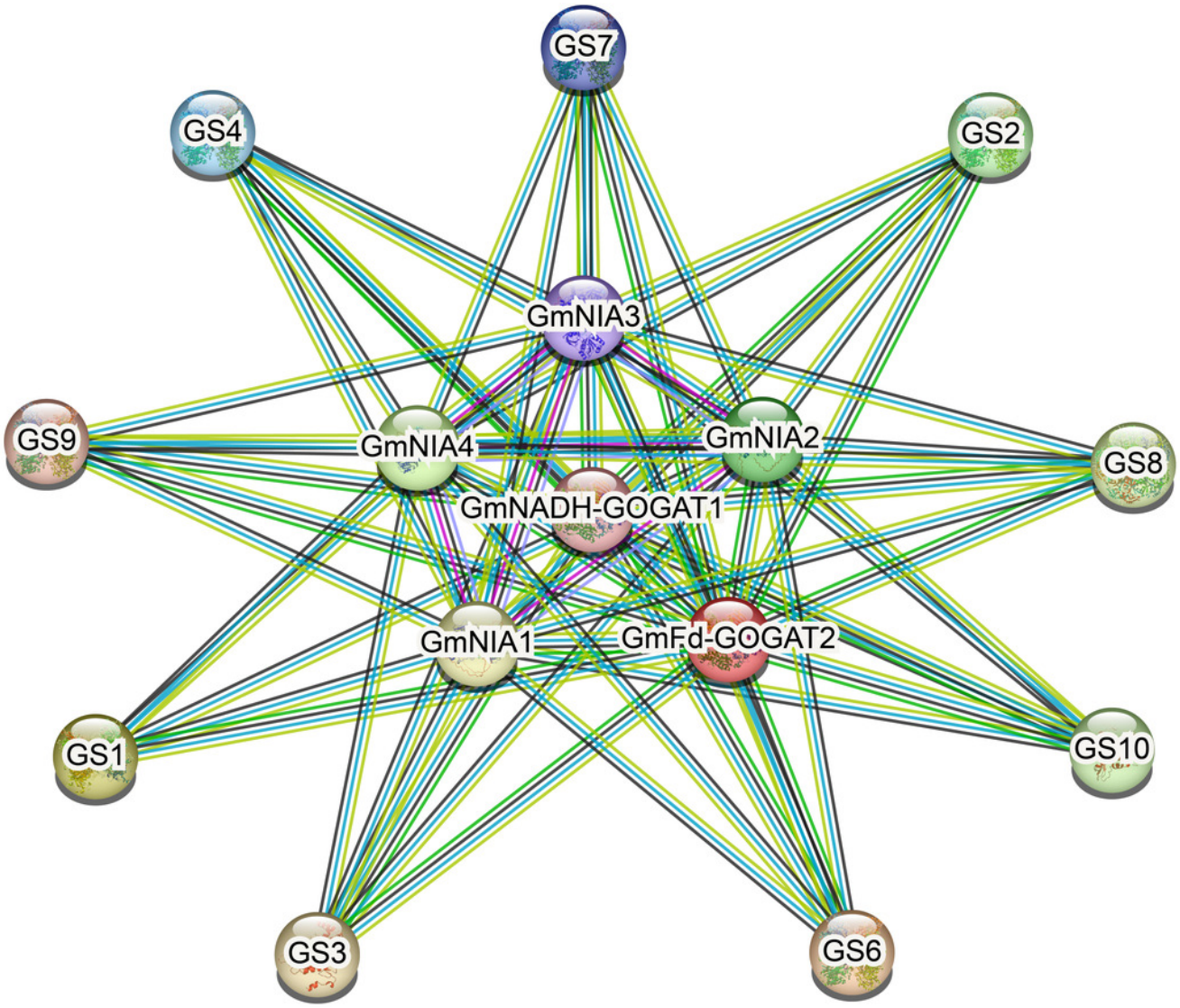


Figure 7

Expression patterns of *GmGS* (A), *GmGOGAT* (B), and *GmNR*(C) genes in eight soybean tissues.

RNA-Seq data were used to construct the expression patterns using the FPKM values. The color scale bars on the right represent the gene expression levels.

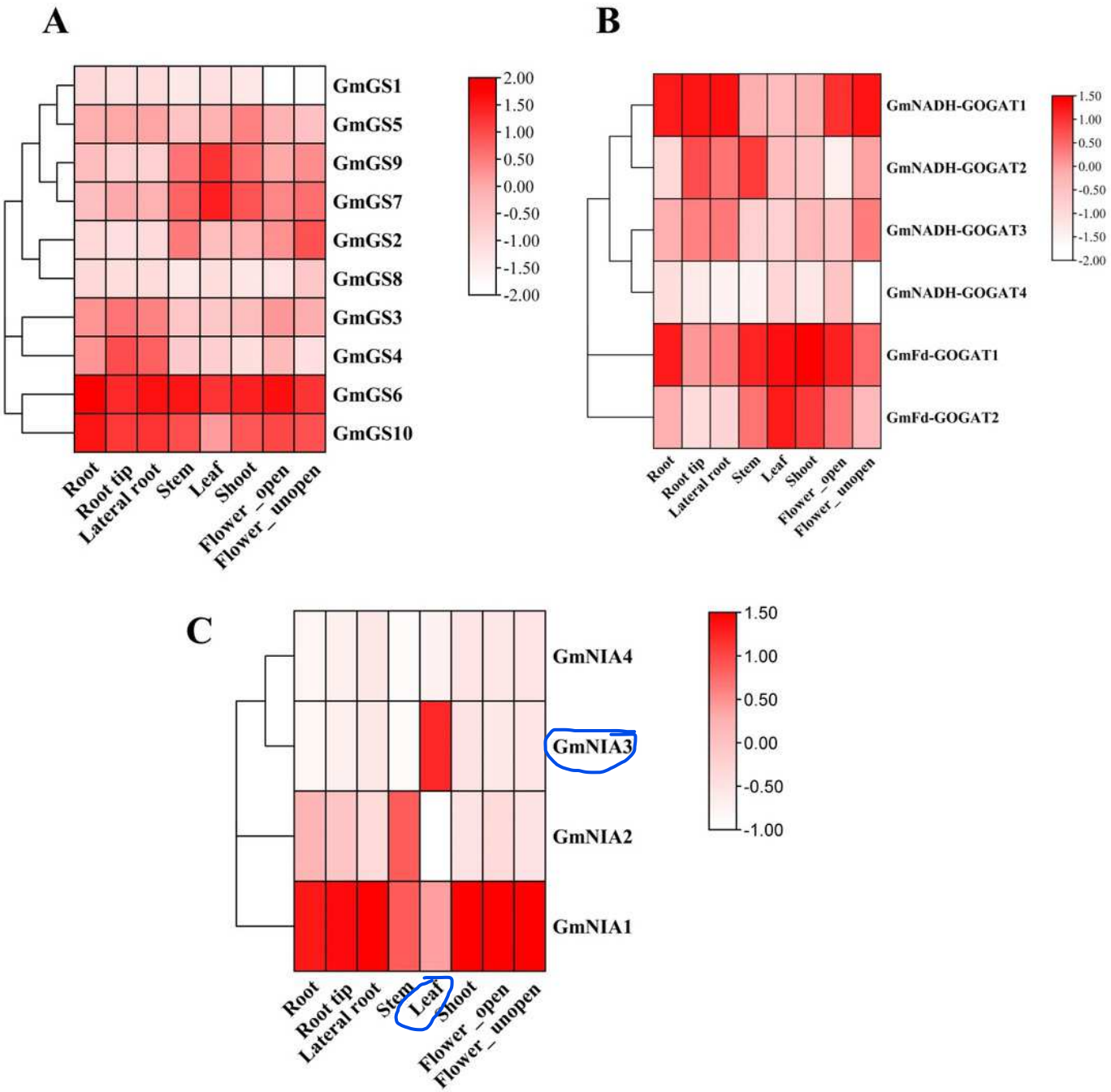
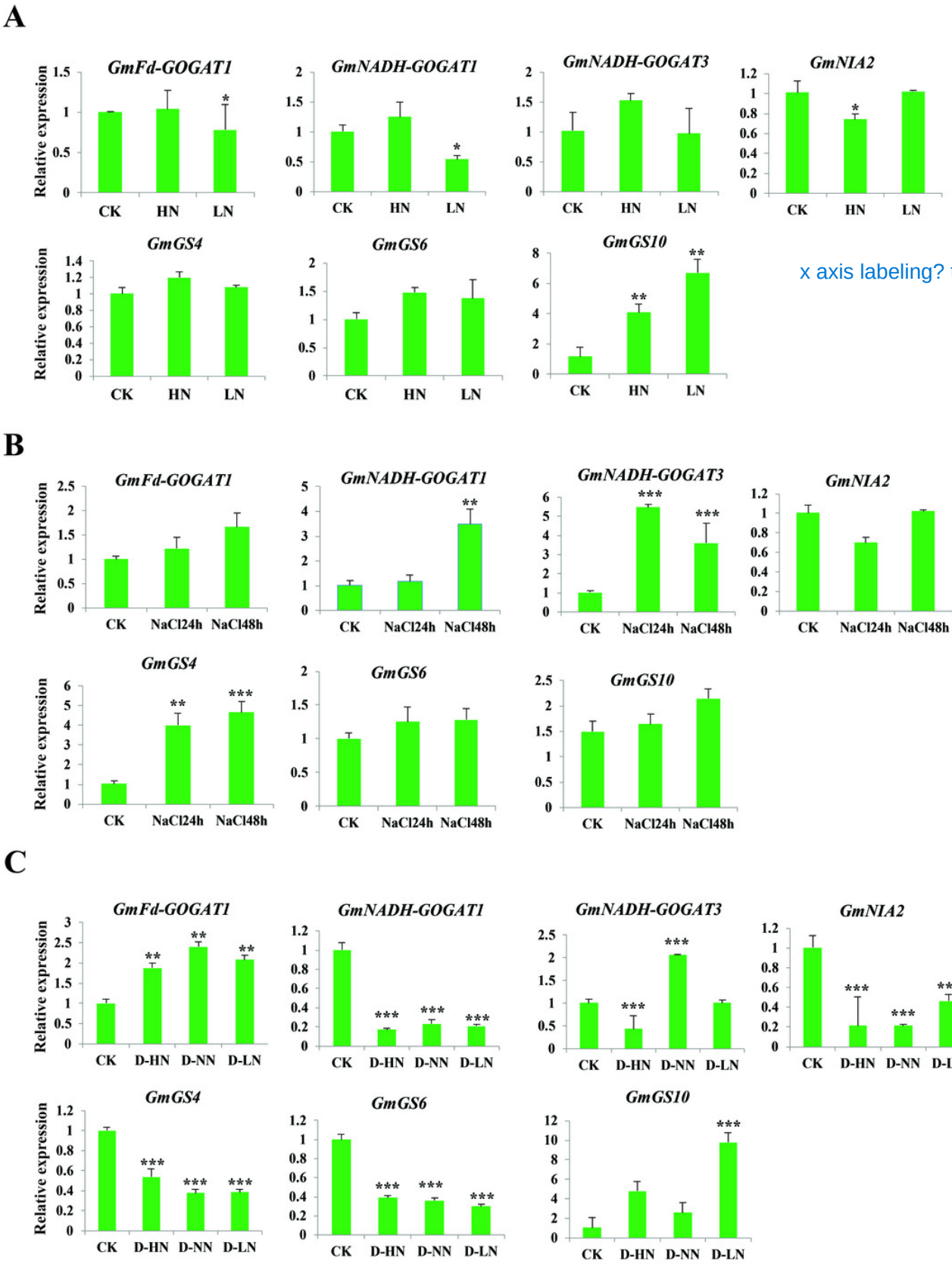

A
GmGS1
GmGS5
GmGS9
GmGS7
GmGS2
GmGS8
GmGS3
GmGS4
GmGS6
GmGS10
2.00
1.50
1.00
0.50
0.00
-0.50
-1.00
-1.50
-2.00
Root
Root tip
Lateral root
Stem
Leaf
Shoot
Flower _ open
Flower _ unopen
B
GmNADH-GOGAT1
GmNADH-GOGAT2
GmNADH-GOGAT3
GmNADH-GOGAT4
GmFd-GOGAT1
GmFd-GOGAT2
1.50
1.00
0.50
0.00
-0.50
-1.00
-1.50
-2.00
Root
Root tip
Lateral root
Stem
Leaf
Shoot
Flower _ open
Flower _ unopen
C
GmNIA4
GmNIA3
GmNIA2
GmNIA1
1.50
1.00
0.50
0.00
-0.50
-1.00
Root
Root tip
Lateral root
Stem
Leaf
Shoot
Flower _ open
Flower _ unopen

Figure 8

Expression profile of seven selected *GmGS, GmGOGAT, and GmNR* genes in response to various stress treatments.

(A) different nitrate concentrations. (B) 24 h and 48 h NaCl stress. (C) drought-nitrate stress. Gene expression levels were determined using qRT-PCR and normalized with soybean *Actin11* (Glyma18g290800) as a reference gene. Statistically significant expression differences were identified using t-tests (P<.01).



x axis labeling? treatments/stress

note: if possible, provide figure showing symptoms of stress or an experimental setup.