Frontiers in Bioscience Landmark

Front. Biosci. (Landmark Ed) **2026**; 31(9): 54457
https://doi.org/10.31083/FBL54457

Original Research

Genome-Wide Characterization, Expression Profiling and Source-Sink Nitrogen Responses of Rice *Glutamine Synthetase* Gene Family

Haiting Hu[1], Huihui Zhang[1], Zixin Xiang[1], Yuelin Wu[1], Shuai Fu[1], Zhuocheng Liu[2,*], Han Yang[1,*]

[1]Anhui Province Key Laboratory of Plant Resources and Biology, Anhui Engineering Research Center for Green Production Technology of Drought Grain Crops, College of Life Sciences, Huaibei Normal University, 235000 Huaibei, Anhui, China
[2]School of Biological and Food Engineering, Suzhou University, 234000 Suzhou, Anhui, China
*Correspondence: lzcdzsyx@163.com (Zhuocheng Liu); yanghan@chnu.edu.cn (Han Yang)
Academic Editor: Jen-Tsung Chen
Submitted: 2 June 2026 Revised: 23 July 2026 Accepted: 11 August 2026 Published: 31 August 2026

Abstract

Background: Nitrogen is the most abundant mineral nutrient required by rice (*Oryza sativa*) and the primary limiting factor for grain yield. Glutamine synthetase (GS) is a key enzyme involved in nitrogen assimilation in rice, and is closely associated with nitrogen use efficiency (NUE) and source–sink nitrogen dynamics. **Methods:** We performed a genome-wide characterization of the *GS* gene family across 12 plant species. Haplotype analysis of four rice *GS* genes was conducted using resequencing data from approximately 2000 rice accessions. A reproductive-stage nitrogen-gradient experimental system (0N, 1N, 3N) was established to simultaneously profile *GS* transcript levels and enzyme activity in source tissues (roots and flag leaves) and sink tissues (young embryos at different developing stages). **Results:** A total of 86 non-redundant GS proteins were identified, which clustered into two evolutionarily distinct subfamilies (GS1 and GS2). In rice, *OsGS1-1* exhibited pronounced indica–japonica differentiation and a pattern consistent with a domestication bottleneck, while other members showed high sequence conservation. Transcript abundance and enzymatic activity data revealed that *OsGS1-3* among *GS* isoforms exhibited its highest transcript level at 5 days after pollination (DAP), while total GS activity peaked at 5 DAP in developing embryos, suggesting this stage may represent a critical window for active nitrogen assimilation during embryogenesis. Distinct tissue-specific expression patterns and nitrogen-responsive profiles were also observed among *OsGS* isoforms. **Conclusions:** These findings provide fundamental insights into the evolutionary and functional diversification of the rice *GS* gene family. The tissue-specific and nitrogen-dependent transcriptional and enzymatic responses of *GS* isoforms indicate their coordinated involvement in nitrogen assimilation processes, which may further influence source–sink nitrogen utilization in rice under varying nitrogen conditions.

Keywords: *Oryza sativa*; *glutamine synthetase* (*GS*); haplotype; nitrogen; gene expression profiling; enzyme activity

1. Introduction

Nitrogen is the most abundant mineral nutrient required by rice and the primary limiting factor for grain yield [1,2]. As an essential macronutrient, it participates in protein biosynthesis, nucleic acid formation, chlorophyll synthesis, and photosynthetic electron transport, directly determining crop biomass accumulation and reproductive development [3,4,5]. Globally, nitrogen fertilizer application has increased more than 10-fold over the past 60 years to meet the food demands of a growing population, with annual consumption exceeding 110 million tons [2,6]. Rice (*Oryza sativa*) serves as a primary food source for over half of the global population, and nitrogen fertilizer application is the cornerstone of high-yield rice production, accounting for approximately 30% of global nitrogen fertilizer consumption [7]. However, excessive nitrogen fertilization has led to persistently low nitrogen use efficiency (NUE) in rice, with only 30–40% of applied nitrogen being effectively utilized by crops [8]. The remaining nitrogen is lost to the environment through leaching, volatilization, and denitrification, causing severe environmental problems including water eutrophication, greenhouse gas emissions, and soil acidification [9,10]. Therefore, dissecting the molecular mechanisms underlying nitrogen metabolism and improving crop NUE has become a critical scientific issue for sustainable agricultural development and global food security.

Glutamine synthetase (GS, EC 6.3.1.2) is one of the key enzymes in plant nitrogen assimilation, catalyzing the ATP-dependent conversion of inorganic NH_4^+ into organic glutamine [11,12]. The GS/GOGAT cycle constitutes the core pathway for nitrogen assimilation in higher plants, operating throughout the entire growth period from root uptake to grain translocation [13,14]. Across higher plant lineages, the *GS* gene family is categorized into two major groups according to their subcellular localization: cytosolic *GS1* and chloroplastic *GS2* [15]. *GS1* is mainly involved in the primary assimilation of nitrogen taken up from the soil and the reassimilation of NH_4^+ released during protein degradation, while *GS2* is predominantly responsible for reassimilating NH_4^+ released from photorespiration [13,16,17]. In maize, a landmark study by Martin et al. (2006) [18] provided definitive genetic evidence for the

non-redundant functions of GS isoforms in cereal yield formation. Using knockout mutants and overexpression lines, they demonstrated that *GS1-3* specifically controls kernel number by regulating nitrogen supply during ear development, while *GS1-4* determines kernel size by mediating nitrogen remobilization from senescing leaves [18]. In rice, recent functional studies using knockout, knockdown, and overexpression lines have systematically characterized the roles of GS isoforms in NUE and yield traits: *OsGS1;1* maintains basal nitrogen assimilation and pleiotropically regulates grain yield under fluctuating nitrogen conditions; its alternative splicing further modulates amylose content [19,20]. *OsGS1;2* specifically modulates root NH_4^+ assimilation and tiller outgrowth [21,22,23]. *OsGS1;3* is specifically expressed in the endosperm during grain filling, participating in nitrogen accumulation [24]. *OsGS2* dominates photorespiratory NH_4^+ reassimilation and affects leaf nitrogen status; its overexpression enhances nitrogen assimilation capacity under abiotic stresses [11,25]. In soybean, a recent genome-wide analysis by Elsanosi et al. (2024) [26] identified 10 *GS* family members and revealed their distinct expression patterns in response to nitrogen deficiency, drought, and salt stress, highlighting the conserved roles of *GS* genes in nitrogen metabolism and abiotic stress tolerance across plant species. These findings collectively establish that *GS* genes are central regulators of crop productivity and stress adaptation, warranting systematic investigation in rice.

Previous studies on rice *GS* genes have focused on single-gene functional validation in vegetative tissues under normal growth conditions [25,27,28,29], while genome-wide evolutionary analysis and comprehensive investigation of their expression dynamics in source-sink tissues during the critical reproductive growth stage are still lacking. At the genome-wide level, several *GS* family identification and characterization studies have been reported in rice and other Poaceae crops [30,31], but most were limited to basic bioinformatic annotation and vegetative tissue expression profiling. Critically, these studies lacked integration with large-scale population genetics data and biochemical assays under controlled nitrogen gradients. In rice, the reproductive growth stage is the most critical period for yield formation, during which a highly efficient and tightly regulated source-sink nitrogen transport and allocation system is established [32]. Root tips and fully expanded flag leaves are the most important nitrogen source tissues: root tips absorb inorganic nitrogen from the soil, while flag leaves dominate photosynthetic nitrogen assimilation and transport organic nitrogen to reproductive organs via vascular bundles [33,34]. In sharp contrast, developing young embryos are the primary terminal nitrogen sink tissues throughout the reproductive period, and efficient nitrogen allocation from source tissues directly determines grain yield and quality [13,35]. However, how *GS* family genes coordinate nitrogen assimilation, translocation, and utilization in this source-sink system under both low and high nitrogen stress conditions remains poorly understood.

In this study, we performed genome-wide identification and characterization of the rice *GS* gene family, including phylogenetic analysis across 12 plant species, gene structure characterization, conserved motif analysis, and haplotype diversity analysis using resequencing data from nearly 2000 rice accessions. Furthermore, we established a reproductive-stage nitrogen gradient experimental system and systematically investigated the transcriptional profiles and enzyme activities of *GS* genes in source tissues (roots and flag leaves) and sink tissues (developing embryos at four consecutive stages) under low, normal, and high nitrogen conditions. Our findings provide fundamental insights into the evolutionary and functional diversification of the rice *GS* family, and identify valuable genetic resources for molecular breeding of nitrogen-efficient rice varieties.

2. Materials and Methods

2.1 Genome-Wide Screening and Chromosomal Positioning of OsGS Genes

Genome-wide protein sequences and chromosomal annotation files of 12 plant species, including *Oryza sativa*, *Triticum aestivum*, *Zea mays*, *Sorghum bicolor*, *Hordeum vulgare*, *Setaria italica*, *Arabidopsis thaliana*, *Chenopodium quinoa*, *Solanum lycopersicum*, *Glycine max*, *Arachis hypogaea* and *Brassica napus*, were accessed from the Ensembl Plants database (https://plants.ensembl.org/index.html) and the NCBI Genome database (https://www.ncbi.nlm.nih.gov/genome/) (**Supplementary Table 1**). The six full-length protein sequences of *Arabidopsis* AtGS used for homologous search were retrieved from The *Arabidopsis* Information Resource (TAIR) repository (https://www.arabidopsis.org/) [36].

To identify *GS* gene family members across the whole genome, we adopted a combined approach integrating BLASTP homology alignment and hidden Markov model (HMM) search. First, complete amino acid sequences of *Arabidopsis* GS proteins served as reference sequences to conduct BLASTP comparison with the genome-wide full protein sequences of the 12 species using TBtools (v2.085. E-value cutoff: $<1 \times 10^{-5}$). Duplicate sequences were filtered out to obtain preliminary candidate genes. Next, HMM profiles corresponding to the conserved GS domains Gln-synt_N (PF00120) and Gln-synt_C (PF03951) were downloaded from the Pfam database (https://pfam.xfam.org/), and HMMER 3.0 was used for further screening. E-value cutoff: $<1 \times 10^{-4}$ [37,38].

Candidate sequences from two complementary analytical strategies were merged, and redundant or incomplete sequences were removed using TBtools. The remaining sequences were uploaded to the SMART platform (https://smart.embl.de/), Pfam, and NCBI Conserved Domain Database (CDD, https://www.ncbi.nlm.nih.gov/cdd/) for verification. Only sequences containing both conserved

GS domains were retained for final identification [37,38]. According to the chromosomal annotation of *Oryza sativa*, chromosomal distribution patterns of the four *OsGS* genes were analyzed and visualized using TBtools.

2.2 Phylogenetic Tree Construction of the GS Gene Family

The 86 full-length protein sequences of the *GS* gene family used for phylogenetic analysis were obtained from our previous genome-wide identification, covering 12 representative plant species. Multiple sequence alignment of the GS full-length protein sequences from the 12 species was carried out using the CLUSTALW tool embedded in MEGA 11 software (http://www.megasoftware.net/) with default parameters, yielding a standardized sequence alignment matrix for phylogenetic analysis. Based on the alignment result, the phylogenetic tree of the *GS* gene family was inferred with the FastTree software (http://www.microbesonline.org/fasttree/) using the Maximum Likelihood (ML) method under the default Jones-Taylor-Thornton (JTT) amino acid substitution model, with the CAT approximation to model rate heterogeneity across sites. The reliability of branch nodes in the phylogenetic tree was assessed by bootstrap test with 1000 replicates [39].

Visualization, refinement and annotation of the phylogenetic tree were completed via the iTOL web server (https://itol.embl.de/). Subfamily classification and evolutionary characteristic analysis of the GS family proteins were conducted according to the topological structure and clustering outcomes of the phylogenetic tree [40].

2.3 Cross-Species Synteny Analysis of the GS Gene Family in Rice

The MCScanX tool embedded in TBtools software (https://github.com/CJ-Chen/TBtools-II) was applied to assess the collinearity data of *GS* genes, and the collinear homologous gene pairs of the *GS* family between rice and other plant species were identified and counted [41].

2.4 Gene Structure Analysis of the GS Gene Family in Rice

According to the whole-genome annotation information of rice, the positional data of the coding sequence (CDS), 5′ and 3′ untranslated regions (UTRs), exons and introns of all *OsGS* genes were extracted. The exon-intron structural features of *OsGS* genes were systematically dissected, and core structural parameters including full gene length, exon number and intron length were statistically analyzed. TBtools software was utilized to visualize the structural characteristics of *OsGS* genes [42].

2.5 Multiple Sequence Alignment of GS Proteins From Different Plant Species

Multiple sequence alignment of the full-length GS protein sequences from six plant species was conducted using CLUSTALW (v2.1, https://www.ebi.ac.uk/Tools/msa/clustalw2/), with all alignment parameters set to the program default values [43]. Based on the alignment results, the online tool ESPript 3.0 (https://espript.ibcp.fr/ESPript/ESPript/) was used to visualize the sequence alignment, and the conserved amino acid sites as well as the corresponding positions of the two characteristic functional domains (Gln-synt_N and Gln-synt_C) of GS proteins were annotated.

2.6 Prediction and Visualization Analysis of Cis-Acting Elements in the Promoter Regions of GS Genes

The 2000-bp promoter sequences upstream of the ATG start codon of *GS* genes were submitted to the PlantCARE online database (http://bioinformatics.psb.ugent.be/webtools/plantcare/html/) to predict and annotate the cis-acting elements in their promoter regions [44,45]. TBtools software (https://github.com/CJ-Chen/TBtools) was employed to visualize the distribution and quantitative characteristics of the cis-acting elements.

2.7 Conserved Motif Analysis of GS Proteins

Conserved motifs of GS proteins from rice, wheat, *Arabidopsis* and soybean were characterized via the MEME web server (https://meme-suite.org/meme/tools/meme), and the upper limit of predicted conserved domains was defined as 15 [46]. TBtools software (https://github.com/CJ-Chen/TBtools) was employed to visualize the number, type and distribution of conserved motifs [47]. Meanwhile, HMM sequence logos of key conserved motifs were constructed, and the distribution of subfamily-specific motifs and their correspondence with the Gln-synt_N and Gln-synt_C functional domains were analyzed in combination with the phylogenetic clustering results.

2.8 Structural Prediction and Functional Analysis of Interaction Networks of Rice GS Proteins

The secondary structures of OsGS proteins were predicted using the SOPMA online tool (https://npsa-prabi.ibcp.fr/cgi-bin/npsa_automat.pl?page=npsa_sopma.html). The tertiary structures of OsGS monomer proteins were predicted via AlphaFold2 (https://alphafold.ebi.ac.uk/). Three-dimensional (3D) structural models of four OsGS proteins were performed using SWISS-MODEL (https://swissmodel.expasy.org/) to obtain high-confidence 3D structural models of these proteins. The interaction network of rice GS proteins was constructed via the STRING database (https://string-db.org/), and interacting proteins were screened for functional correlation analysis [48]. Gene Ontology (GO) enrichment analysis was conducted for the interacting proteins to elucidate the key biological processes in which GS proteins participate [49].

2.9 Ka/Ks Analysis of Rice GS Genes

TBtools software (https://github.com/CJ-Chen/TBtools) was applied for the computation of the Ka, Ks and Ka/Ks values of *GS* homologous gene pairs between rice and six other species, including maize, wheat, sorghum,

Arabidopsis, soybean, and rapeseed [50]. The distribution of Ka/Ks ratios across different species combinations was visualized using box plots to dissect the evolutionary selection pattern of the *GS* gene family.

2.10 Haplotype Analysis

The 1500-bp promoter and full-length coding sequence of *GS1-1*, *GS1-2*, *GS1-3* and *GS2* in 2003, 2002, 1981 and 1982 rice accessions were obtained from the MBKbase rice repository (https://mbkbase.org). The cultivated panel comprised approximately 2,005 _*Oryza sativa*_ accessions aggregated from nine pre-release projects within MBKbase (HBWild461, XD248, HB950, Japan176, MBK1287, HXH67, Rice3k, ISG575, HAU533). Subpopulation assignments (Indica, Basmati/Sadri, Aus/Boro, Temperate Japonica, Tropical Japonica) were obtained from the rice 3K project (RFGB, http://www.rmbreeding.cn/Index/s) [51] and collapsed into the legend categories Ind, Bas, Aus, Tem-Jap, and Tro-Jap, respectively. The wild rice panel comprised 19 _Oryza rufipogon_ accessions from the HBWild461 project, drawn from the Or-I, Or-II, Or-III and Wild ecotype groups and merged into a single Wild category. The Nipponbare reference genome was used as the coordinate reference throughout. Nucleotide diversity profiles were generated with a 2000 bp window and a 100 bp step size. To assess the evolutionary patterns of *GS* genes, all polymorphic sites with a minor allele frequency (MAF) ≥0.005 within this locus were used to construct phylogenetic trees and haplotype networks. The phylogenetic trees were built using the UPGMA method in MEGA 11, and the resulting trees were visualized in EvolView [52], whereas the haplotype network was established with the pegas package [53]. All parameter settings and detailed nucleotide polymorphism statistics for the haplotype analysis of the four *GS* genes are provided in **Supplementary Table 2**.

2.11 Plant Materials and Hydroponic Treatments

Rice plant materials used in this study were obtained from the indica rice cultivar Zhongxian 3037. Seedling hydroponic growth was conducted inside an artificial climate incubator under 23–35 °C conditions. The growth settings included a 10-hour light and 14-hour dark photoperiod, 70% ambient humidity, and a light intensity of 200 μmol m^{-2} s^{-1}. Plants were planted in porous ceramic (PROFILE Products LLC) cultivated with Kimura B nutrient solution which was defined as normal N condition (**Supplementary Table 3**). The nutrient solution was renewed every three days. For gradients treatment with different N concentrations, $(NH_4)_2SO_4$ and KNO_3 were replaced by NH_4NO_3, and $Ca(NO_3)_2$ was replaced by $CaCl_2 \cdot 2H_2O$. Three gradient contents were set for hydroponic culture, including 0 mM, 0.9 mM and 2.7 mM NH_4NO_3. Compared with the N concentration of Kimura B solution (1.7 mM), 0 mM NH_4NO_3 was defined as 0N condition (low N), 0.9 mM NH_4NO_3 was defined as 1N (the moderate N condition), and 2.7 mM NH_4NO_3 was defined as excessive N condition (high N) in this study.

2.12 Sample Collection and RNA Isolation

Collection of flag leaves and root tips and total RNA extraction: Apical 5 cm of flag leaves and 5 cm root tips from WT Zhongxian 3037 under three different N treatments were collected. Three independent biological replicates were set for each genotype-treatment group, with each replicate containing panicles from at least 15 uniform plants. Following snap-freezing of samples in pre-chilled liquid nitrogen, total RNA isolation was conducted following routine protocols with TRIZOL reagent. RNA-seq analysis was performed by Metware, Wuhan, China (https://www.metware.cn/) (**Supplementary Table 4**).

Rice young embryo collection and total RNA extraction: Total RNA was isolated from rice embryos of four developing stages in WT grown under three different N gradients. Stage 1 (1 DAP): Zygote polarization, no organ differentiation, active cell division. Stage 2 (2 DAP): Globular proembryo, axial polarity established, rapid cell proliferation. Stage 3 (3 DAP): Organ primordia initiation, tissue differentiation, functional transition. Stage 4 (5 DAP): Complete embryonic organs, active nutrient storage and nitrogen metabolism and rapid accumulation of storage products. After snap freezing rice young embryos in liquid nitrogen, total RNA extraction was conducted following conventional protocols with a magnetic bead-based high-efficiency total RNA extraction kit for polysaccharide- and polyphenol-rich plants (TIANGEN, DP772). RNA purity and concentration were detected by a NanoDrop 2000 spectrophotometer (Thermo Fisher Scientific, USA), and RNA integrity was evaluated using an Agilent 2100 Bioanalyzer (Agilent Technologies, USA). Samples with RNA integrity number (RIN) ≥7.0 and $OD_{260/280}$ ratio between 1.8 and 2.1 were used for subsequent library construction. mRNA was enriched via oligo (dT) magnetic beads, fragmented into short fragments, and reversely transcribed into first-strand cDNA. Second-strand cDNA was synthesized, followed by end repair, A-tailing, adapter ligation and PCR amplification to construct paired-end cDNA libraries. All libraries were sequenced on the Illumina NovaSeq 6000 platform (Illumina, USA) with a 150 bp paired-end (PE150) read strategy, and approximately 6 Gb of clean data was generated for each sample. RNA-seq analysis was performed by Metware, Wuhan, China (https://www.metware.cn/) (**Supplementary Table 5**).

2.13 RNA Sequencing and Transcriptome Analysis

Read Processing and Gene Expression Quantification: Raw sequencing reads (raw reads) were first subjected to quality control using FastQC (v0.11.9, Babraham Bioinformatics, Cambridge, UK) to assess base quality, GC content and adapter contamination. Trim Galore (v0.6.7, Babraham Bioinformatics, Cambridge, UK) was then applied to

remove adapter sequences, low-quality bases (Phred quality score <20) and reads shorter than 50 bp, yielding high-quality clean reads. The *Oryza sativa* Nipponbare IRGSP-1.0 reference genome and corresponding gene structure annotation file (GTF format) were downloaded from the Rice Genome Annotation Project (RGAP, Michigan State University, East Lansing, MI, USA) database. Clean reads were aligned to the reference genome using HISAT2 (v2.2.1, Johns Hopkins University, Baltimore, MD, USA) with default parameters. The number of reads mapped to each gene was counted using featureCounts (v2.0.3, Walter and Eliza Hall Institute of Medical Research, Melbourne, Victoria, Australia) based on the genome annotation, and the raw read count matrix of all genes was generated for subsequent differential expression analysis.

Quality Control of Transcriptome Data: Prior to differential expression analysis, the raw read count matrix was preprocessed and subjected to systematic quality control. Genes with consistently low expression (raw count <1 in more than half of all samples) were filtered out to reduce statistical noise and improve the reliability of downstream statistical tests. The raw counts were then subjected to regularized logarithm (rlog) transformation via the DESeq2 package (v1.42.0, Bioconductor project, implemented in R statistical software, R Core Team, Vienna, Austria), which corrects for sequencing depth across samples and stabilizes the variance of low-abundance genes, making the data suitable for distance-based quality assessment. To evaluate the reproducibility of biological replicates, pairwise Pearson correlation coefficients of genome-wide expression profiles were calculated between all samples based on rlog-normalized values, and the results were visualized as a clustered heatmap using the pheatmap R package (v1.0.12, CRAN, implemented in R statistical software, R Core Team, Vienna, Austria). Principal component analysis (PCA) was further performed using the prcomp function in R (R Core Team, Vienna, Austria) to capture the major sources of transcriptomic variation across all samples. The first two principal components were plotted to inspect the overall separation between different genotypes and nitrogen treatments, as well as the clustering pattern of biological replicates within each group. Samples that deviated markedly from their corresponding group were regarded as potential outliers and would be excluded from further analysis.

Differential Expression and Functional Enrichment Analysis: Differential expression analysis between groups was performed using the DESeq2 R package (v1.42.0, Bioconductor project, implemented in R statistical software, R Core Team, Vienna, Austria). Genes with adjusted p-value < 0.05 and $|\log_2(\text{fold change})| \geq 1$ were defined as differentially expressed genes (DEGs). Principal component analysis (PCA) and sample correlation analysis were conducted to evaluate the reproducibility among biological replicates and the overall separation between groups. Functional enrichment analysis of DEGs was carried out using the clusterProfiler R package (v4.6.2, Bioconductor project, implemented in R statistical software, R Core Team, Vienna, Austria), including GO annotation covering biological process, cellular component and molecular function categories, and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment. Enriched terms with adjusted p-value < 0.05 were considered statistically significant.

2.14 GS Enzyme Activity Determination

The apical 10 cm of flag leaves, the apical 5 cm of root tips, and developing embryos at three developmental stages (1, 5 and 10 DAP) were collected. All samples were immediately frozen in liquid nitrogen and stored at –80 °C until further analysis. GS enzyme activity assays were performed by the Technical Service Department of Beijing Solarbio Science & Technology Co., Ltd. (https://solarbio.bioon.com.cn/) via the GS Activity Detection Kit BC0915 (**Supplementary Table 6**).

2.15 Statistical Analysis

Each assay contained three independent biological replicates, and data statistics were processed via GraphPad Prism 10.1.2. For all statistical comparisons, we carried out one-way ANOVA followed by Tukey's HSD post-hoc test for multiple pairwise comparisons, with a significance level set at $p < 0.05$. For letter marking: within the same gene, treatments under different nitrogen conditions sharing identical letters indicate no significant difference ($p > 0.05$), while distinct letters represent statistically significant differences ($p < 0.05$). Only intra-gene comparisons were conducted, and cross-gene comparison was not performed.

3. Results

3.1 Identification of GS Gene Family Members

To comprehensively characterize *GS* gene family members across multiple plant species, we employed two complementary bioinformatics approaches. First, six *Arabidopsis thaliana* GS (AtGS) protein sequences were employed as queries to conduct BLASTP searches against the protein databases of 12 plant species, including *Oryza sativa* (rice), *Triticum aestivum* (wheat), *Zea mays* (maize), *Sorghum bicolor* (sorghum), *Hordeum vulgare* (barley), *Setaria italica* (foxtail), *Arabidopsis thaliana* (*Arabidopsis*), *Chenopodium quinoa* (quinoa), *Solanum lycopersicum* (tomato), *Glycine max* (soybean), *Arachis hypogaea* (peanut), and *Brassica napus* (rapeseed), with E-value cutoff $<1 \times 10^{-5}$. Second, the hidden Markov model (HMM) profiles corresponding to the conserved GS domains (PF00120 and PF03951) were extracted from the Pfam database, and HMMER 3.0 software was used to screen the same protein databases with a more stringent E-value cutoff $<1 \times 10^{-4}$.

After removing redundant sequences, all candidate proteins were submitted to the SMART online server for

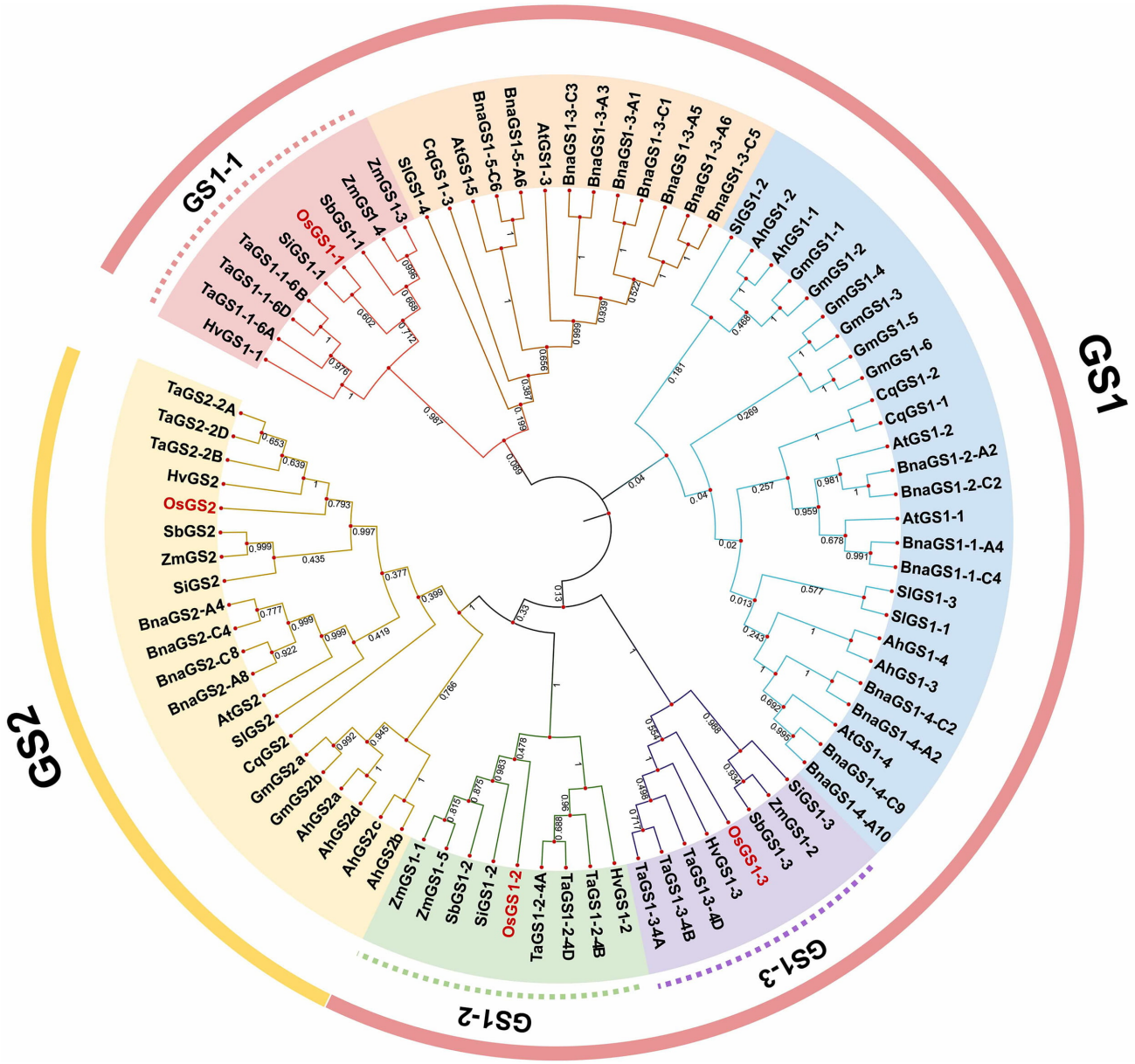

Fig. 1. Phylogenetic tree of glutamine synthetase (GS) family proteins from 12 plant species. The tree is divided into two major evolutionarily distinct subfamilies: cytosolic GS1 and plastidic GS2. The GS1 clade is further classified into three well-supported subgroups (GS1-1, GS1-2, GS1-3) in Poaceae species according to the rice GS nomenclature. Different colored regions represent distinct subfamily or subgroup classifications, and bootstrap support values are indicated at each node.

domain validation. Only sequences containing both the N-terminal (Gln-synt_N, PF00120) and C-terminal (Gln-synt_C, PF03951) conserved domains of GS were subjected to further analysis. This rigorous screening process yielded a total of 86 non-redundant GS protein sequences across the 12 species. Specifically, the numbers of GS proteins identified were: 4 in rice (OsGS), 12 in wheat (TaGS), 6 in maize (ZmGS), 4 in sorghum (SbGS), 4 in barley (HvGS), 4 in foxtail millet (SiGS), 6 in *Arabidopsis* (AtGS), 4 in quinoa (CqGS), 5 in tomato (SlGS), 8 in soybean (GmGS), 8 in peanut (AhGS), and 21 in rapeseed (BnaGS) (**Supplementary Table 1**). Corresponding coding gene sequences were then retrieved for all identified GS proteins, and subsequent analyses were focused on the four *OsGS* genes.

Subsequently, chromosomal positioning analysis of the four *OsGS* genes was conducted via TBtools software (**Supplementary Fig. 1**). The results revealed that the four *OsGS* genes were localized to three different rice chromosomes: *OsGS1-1* (*Os02g0735200*) was located on chromosome 2; *OsGS1-2* (*Os03g0223400*) and *OsGS1-3* (*Os03g0712800*) were tandemly arranged on chromosome 3; and *OsGS2* (*Os04g0659100*) was mapped to chromosome 4.

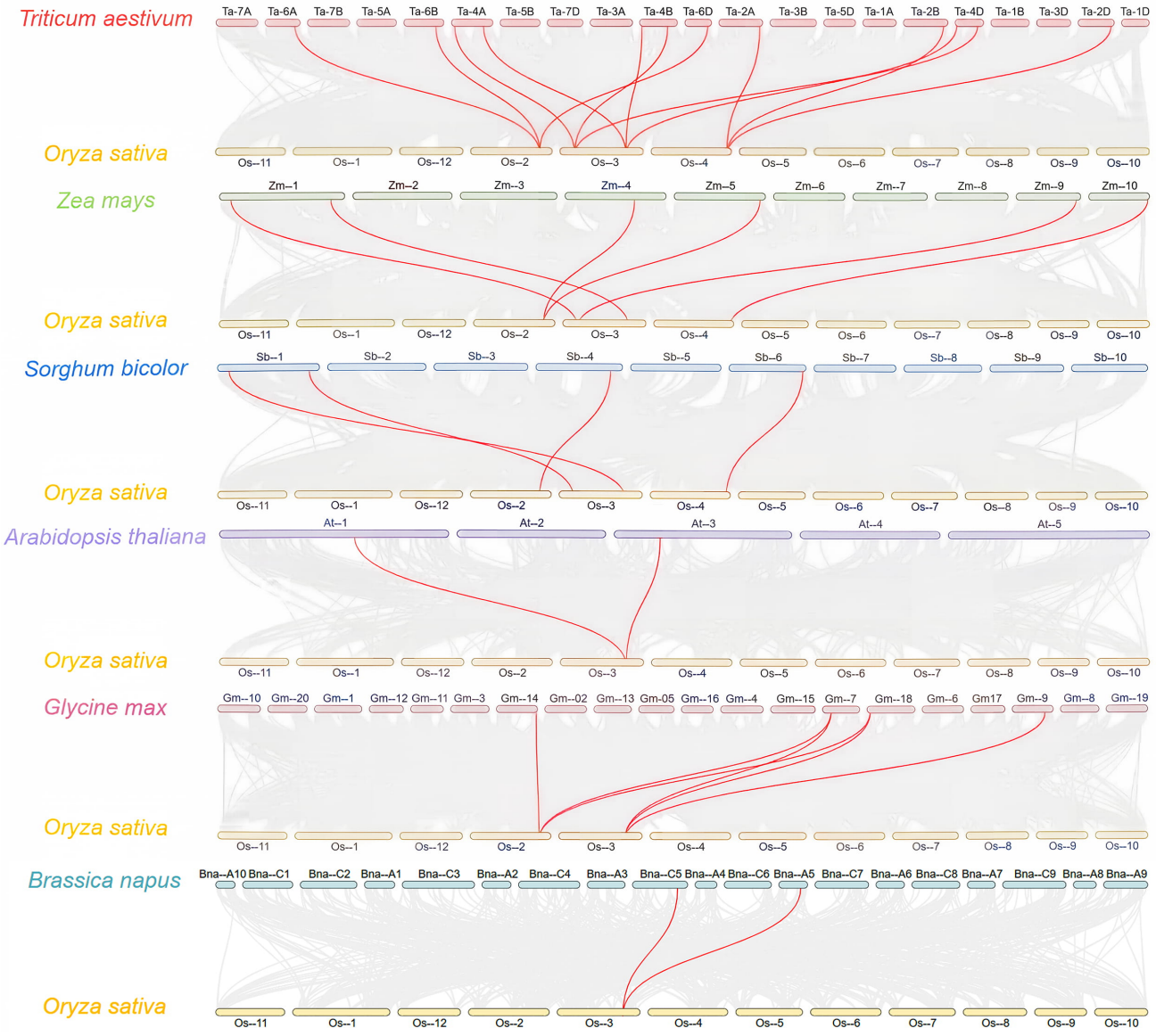

Fig. 2. Genome-wide synteny analysis of *GS* genes between rice and six representative plant species. Syntenic relationships between rice chromosomes and the genomes of wheat, maize, sorghum, *Arabidopsis*, soybean, and rapeseed are shown. Collinear blocks between the paired genomes are indicated by gray background lines, while orthologous collinear pairs of *GS* genes are highlighted with red lines.

3.2 Phylogenetic Analysis of GS Gene Family Members

To elucidate the phylogenetic affinities of the *GS* gene family, we constructed an unrooted maximum likelihood phylogenetic tree based on the complete amino acid sequences of the 86 non-redundant GS proteins identified from 12 representative plant species (Fig. 1). All GS proteins were clearly clustered into two evolutionarily distinct major clades corresponding to the cytosolic GS1 and plastidic GS2 subfamilies, consistent with their subcellular localization. Notably, the GS1 clade in Poaceae (grass) species was further resolved into three well-supported monophyletic subgroups, designated as GS1-1, GS1-2, and GS1-3 following the standardized rice GS nomenclature system. Phylogenetic clustering was highly consistent with the species evolutionary relationships: all GS proteins from monocot grasses and dicotyledonous plants formed separate, independent clades with no cross-grouping.

A notable species-specific proliferation of *GS* homologs was detected in polyploid lineages, particularly rapeseed (21 members) and wheat (12 members), with expansion detected in both the *GS1* and *GS2* subfamilies. This pattern indicates that gene duplication events, likely including whole-genome duplication and segmental duplication, are important contributors to the lineage-specific expansion of the *GS* family in these species, and these duplication events may serve as key drivers of functional diversification of the family. In contrast, the *GS* gene family remained evolutionarily conserved in diploid grass species such as rice, sorghum, and foxtail millet, each containing exactly 4 members.

3.3 Synteny Analysis of GS Genes Between Rice and Six Representative Plant Species

To further explore evolutionary mechanisms underlying the phylogenetic diversification of the *GS* gene family, genome-wide synteny maps were constructed between rice and six representative plant species (Fig. 2). These species included three monocot grasses (wheat, maize, sorghum) and three dicotyledonous plants (*Arabidopsis*, soybean, rapeseed).

The counts of orthologous collinear *GS* gene couples identified between rice and wheat, maize, sorghum, *Arabidopsis*, soybean, and rapeseed were 12, 6, 4, 2, 6, and 2, respectively. Abundant syntenic *GS* gene pairs were observed between rice and the Poaceae species (wheat, maize, sorghum), demonstrating the highly conserved syntenic relationships of *GS* gene loci across grass genomes during evolution. In sharp contrast, only bits of collinear *GS* gene pairs were detected between rice and the two Brassicaceae species (*Arabidopsis* and rapeseed), reflecting the substantial erosion of *GS* gene synteny following the ancient monocot-dicot divergence.

Notably, the collinear *GS* gene pairs between soybean and rice displayed a clear expansion signature, implying that species-specific duplication events have driven the evolution and functional diversification of the *GS* gene family.

3.4 Gene Structure Analysis of GS Genes in Rice

Gene structural divergence is a key driver of adaptive evolution and functional diversification in gene families [38]. Based on the rice reference genome annotation files, we analyzed the exon-intron structures of the four *OsGS* genes and visualized the results using TBtools software (**Supplementary Fig. 2**). All four *OsGS* genes contained complete 5′-UTR, 3′-UTR and CDS regions, and exhibited the characteristic multi-exon structure of the *GS* gene family. Notably, the three *GS1* subfamily members (*OsGS1-1*, *OsGS1-2* and *OsGS1-3*) shared nearly identical gene architectures, with exactly 10 exons and highly conserved exon lengths and arrangements, reflecting strong evolutionary conservation within the cytosolic GS1 clade. This high structural uniformity aligns with their conserved cytosolic localization and overlapping roles in basal nitrogen assimilation, reflecting strong evolutionary constraint on the core catalytic function of cytosolic GS isoforms.

In stark contrast, the plastidic *OsGS2* gene displayed a distinct structural profile: it had a significantly longer genomic length than the *GS1* genes, primarily due to the presence of multiple large introns, and contained 12 exons. The acquisition of additional exons and expanded introns during evolution likely underlies the functional specialization of *GS2*: the extra exon sequences may encode the chloroplast transit peptide that determines its plastid localization, while the enlarged intronic regions may harbor cis-regulatory elements that drive its light-responsive and photosynthetic tissue-specific expression. This marked structural divergence between the GS1 and GS2 subfamilies suggests that structural variation has accompanied the functional specialization of these two GS isoforms during rice evolution.

3.5 Amino Acid Sequence Alignment of GS Proteins Across Representative Plant Species

To systematically investigate the sequence conservation characteristics of the *GS* gene family, multiple sequence alignments were generated with CLUSTALW using GS protein sequences from six representative plant species: rice, wheat, maize, *Arabidopsis*, soybean, and rapeseed (**Supplementary Fig. 3**). The alignment results showed that all GS family members shared extensive highly conserved amino acid regions across different species. All GS proteins contained the two characteristic core functional domains of the family: Gln-synt_N (PF00120) and Gln-synt_C (PF03951), which exhibited remarkably high sequence conservation. Notably, multiple core functional sites, including the NH_4^+-binding active sites within the Gln-synt_N domain and ATP-binding active sites located in the Gln-synt_C domain, were completely identical in both monocots (rice, wheat, maize) and dicots (*Arabidopsis*, soybean, rapeseed). This finding reflects the strong evolutionary constraint on the *GS* gene family and the essential structural and functional stability of this crucial nitrogen assimilation enzyme. Further structural feature prediction revealed that rice GS proteins are soluble intracellular proteins lacking transmembrane helices and signal peptides, and are not directed to the canonical secretory pathway (**Supplementary Fig. 4**).

3.6 Cis-Acting Element Prediction of GS Genes Across Four Representative Plant Species

Gene expression is predominantly modulated by cis-acting elements that are embedded within the promoter sequences of the corresponding genes [44]. To dissect the transcriptional control mechanisms and circuits underlying the *GS* gene family, we predicted and annotated the cis-acting elements within the upstream promoter regions of *GS* genes across four representative plant species (rice, wheat, *Arabidopsis*, and soybean) (Fig. 3).

The results showed that four major classes of predicted cis-elements were detected in the promoter regions: hormone-responsive elements, stress-responsive elements, light-responsive elements, as well as tissue-specific cis-regulatory elements. Within these predicted cis-element classes, light-responsive elements were the most widely distributed and had the highest occurrence frequency in *GS* genes across the four species at the sequence level. ABA-responsive elements, defense and stress-responsive elements, and anaerobic induction elements were universally detected in the predicted cis-elements of *GS* genes across all four species, which suggests potential evolutionarily conserved regulatory roles of the *GS* family in environmental adaptation and primary nitrogen metabolism.

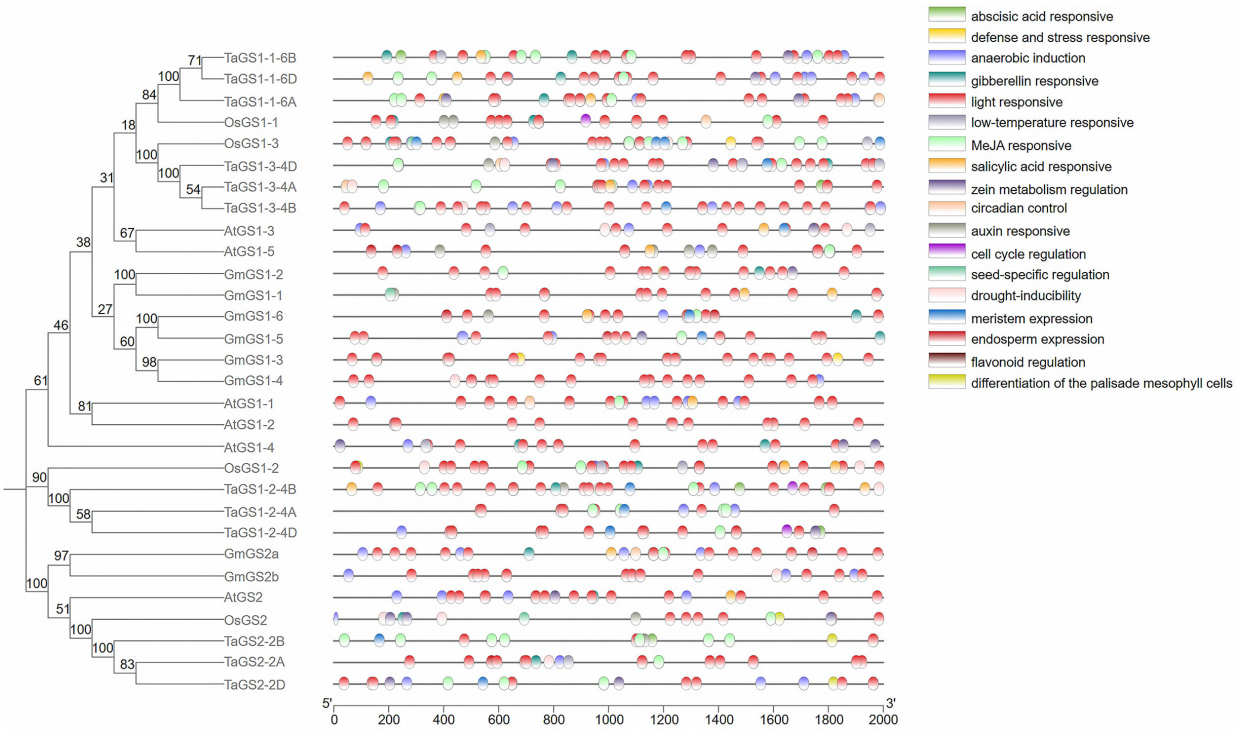

Fig. 3. Prediction of cis-acting elements in *GS* genes across diverse plant species. The left panel presents an unrooted maximum likelihood phylogenetic tree inferred with complete GS amino acid sequences of rice, wheat, *Arabidopsis*, and soybean, with bootstrap support values (1000 replicates) marked at each node. The right panel shows the distribution of cis-regulatory elements within the 2000-bp promoter regions upstream of the translation initiation codon (ATG) of the corresponding *GS* genes. Color-coded circles denote distinct cis-acting regulatory motifs grouped by their biological functions as specified in the legend. The bottom scale bar denotes the nucleotide positions of the promoter regions from the 5′ to 3′ end.

At the species level, rice and wheat *GS* genes contained more meristem-specific expression elements (4 and 7, respectively) than *Arabidopsis* (1) and soybean (2). Seed-specific regulatory elements were exclusively detected in rice (*OsGS1-3*, *OsGS2*) and soybean (*GmGS1*), but absent in wheat and *Arabidopsis*. Low-temperature-responsive elements were widely distributed in rice, wheat and *Arabidopsis*, but were not significantly enriched in soybean. These sequence-level findings suggest that the promoter regions of *GS* genes have conserved motif distribution features across species, while species-specific motif enrichment patterns may contribute to divergent spatiotemporal expression regulation and adaptive evolution of *GS* genes in different plant lineages.

3.7 Conserved Motif Analysis of GS Proteins Across Four Representative Plant Species

To characterize the conserved structural features of GS proteins, conserved motif analysis was performed on GS proteins from rice, wheat, *Arabidopsis*, and soybean (Fig. 4). In total, 15 conserved motifs were characterized, and their composition and linear arrangement were highly consistent across all analyzed GS proteins. Among them, the core motifs (Motif 1, Motif 2 and Motif 4) were universally conserved in all GS members. Motif 1 and Motif 2 corresponded to the Gln-synt_C domain, while Motif 4 matched the Gln-synt_N domain (Fig. 5).

The motif composition, number, and distribution were highly consistent among proteins belonging to an identical GS subfamily, whereas marked divergence was observed between different subfamilies. This pattern was in excellent agreement with the phylogenetic clustering patterns of our phylogenetic tree. Specifically, Motifs 1 to 12 are largely conserved across GS proteins, although a few proteins such as AtGS1-5 lack individual motifs, while Motif 14 and Motif 15 were restricted to the GS2 subfamily. Collectively, the conserved motif patterns of GS proteins not only underscore the evolutionary conservation of core family functions, but also reveal the molecular basis for functional diversification among different GS isoforms conferred by subfamily-specific motif distribution patterns.

3.8 Secondary and Tertiary Structural Characterization of OsGS Family Proteins

The secondary structures of OsGS proteins were predicted using the SOPMA online tool, and three-dimensional (3D) structural models were constructed via AlphaFold2 (**Supplementary Fig. 5**). All four OsGS proteins were predominantly composed of random coils and α-helices, with extended strands as the secondary component. Struc-

tural analysis revealed that all four OsGS proteins shared highly conserved spatial conformations, each consisting of two core functional modules: an N-terminal Gln-synt_N ammonia-binding domain and a C-terminal catalytic Gln-synt_C domain. The assembly of α-helices and β-sheets in these two core domains formed a stable βαβ sandwich-like fold, which constitutes the structural foundation for substrate binding and glutamine biosynthesis—the core catalytic reaction mediating plant primary nitrogen assimilation.

The composition of α-helices and β-sheets in the core catalytic domains was completely conserved among the four OsGS proteins, whereas random coils were mainly distributed in the interdomain linker regions and protein surfaces, providing structural flexibility for conformational changes during substrate binding and protein-protein interactions—a feature consistent with the conserved catalytic function and diversified regulatory patterns of GS isoforms in coordinating source-sink nitrogen allocation. The predicted secondary structure features were highly consistent with the topological characteristics of the 3D models (**Supplementary Fig. 6**).

3.9 Prediction of Interacting Proteins and Functional Network Analysis

Protein–protein interaction (PPI) networks contribute to the prediction of functionally homologous proteins among sequence-based homology groups, which is critical for investigating functional gene interactions and regulatory networks [54]. Therefore, we used the STRING database to predict proteins interacting with rice GS proteins. The results revealed that the four OsGS proteins formed a compact interaction network with NH_4^+ transporters (AMT family), vitamin transporters (VITH family), iron/manganese ion homeostasis-related proteins, and pyridoxal metabolism-related proteins (PDX family), suggesting that GS proteins are not only involved in glutamine biosynthesis but also widely implicated in biological processes including nitrogen uptake, ion homeostasis, and vitamin metabolism (Fig. 6A).

Furthermore, GO functional enrichment analysis (Fig. 6B) demonstrated that the interacting proteins were significantly overrepresented in manganese ion homeostasis, intracellular iron ion sequestration, inorganic cation transmembrane transport, and pyridoxal phosphate biosynthesis, as well as glutamine biosynthesis itself. These findings indicate that rice GS proteins function as pivotal regulators in integrating core physiological processes such as nitrogen metabolism, ion balance, and cofactor synthesis through synergistic interactions with diverse functional proteins.

3.10 Ka/Ks Analysis of GS Genes in Rice

To deeply dissect the phylogenetic trajectory of the *GS* family genes, TBtools software was utilized to analyze the Ka/Ks ratios of *GS* homologous gene pairs between

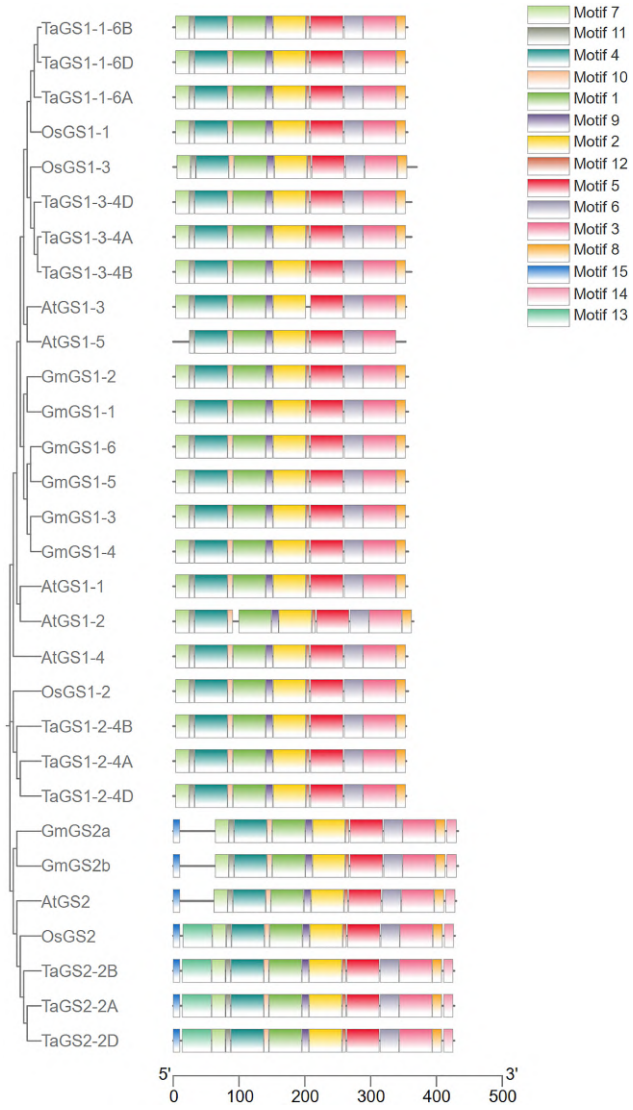

Fig. 4. Conserved motif analysis of GS proteins from four plant species. Conserved motif composition and arrangement are shown for GS proteins from rice, wheat, *Arabidopsis*, and soybean. A total of 15 conserved motifs (Motif 1–15) were detected, with distinct colored boxes representing different motifs. The amino acid length of the proteins is indicated by the scale bar shown at the bottom.

rice and other plants. The results revealed that Ka/Ks ratios of all species combinations were significantly below unity (**Supplementary Fig. 7**), suggesting that *GS* genes underwent intense purifying selection during the evolution of rice and other analyzed plant species, which reflects the functional conservation of GS as a core enzyme in nitrogen metabolism. Overall, the purifying selection pattern of *GS* genes is consistent across different species, with only minor differences in evolutionary rates. These results further validate the central and conserved role played by the GS family in the regulation of plant nitrogen metabolism.

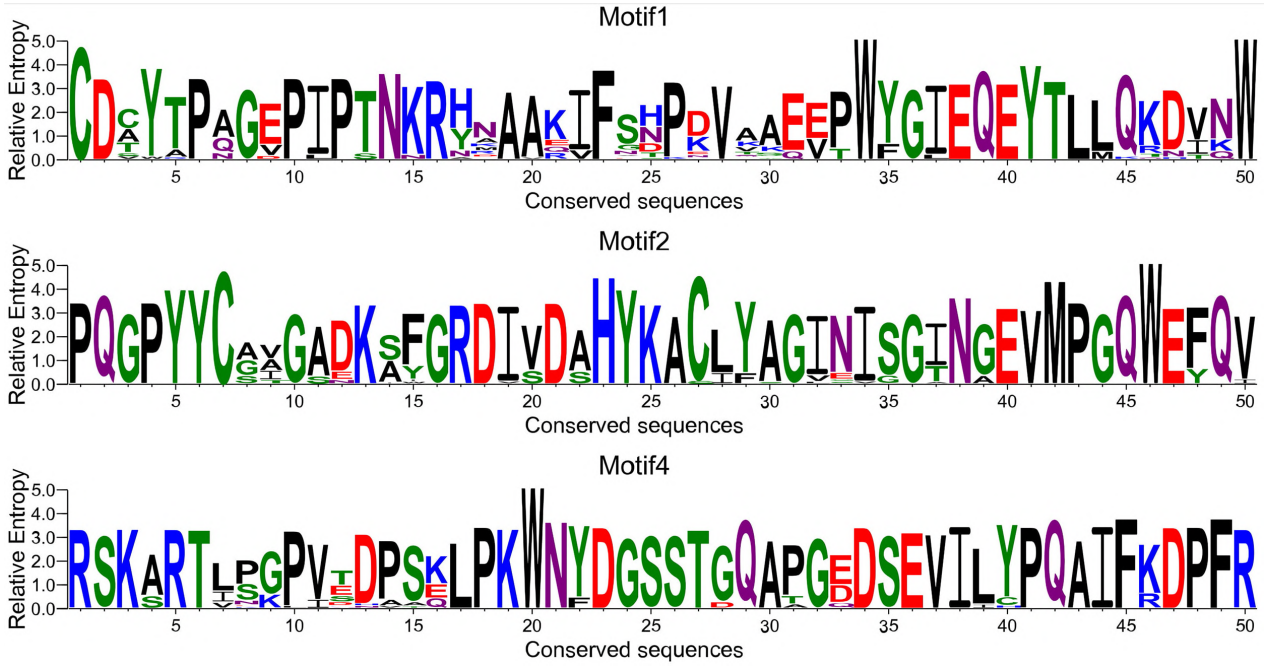

Fig. 5. Hidden Markov model (HMM) logos of conserved motifs in GS proteins. Sequence logos of three conserved motifs (Motif 1, Motif 2, and Motif 4) from GS proteins were generated to visualize the amino acid conservation at each position. The X-axis represents the consensus amino acid sequences of each motif, while the Y-axis denotes the relative entropy values, which reflects the conservation rate of each residue. Motif 1 and Motif 2 correspond to the Gln-synt_C domain, and Motif 4 corresponds to the Gln-synt_N domain of GS proteins.

3.11 Haplotype Diversity and Evolutionary Divergence Analysis of OsGS Genes in Rice Germplasm Resources

To comprehensively investigate the genetic variation and evolutionary characteristics of the entire rice *GS* gene family, we performed comprehensive haplotype analysis of four members: *OsGS1-1* (*Os02g0735200*), *OsGS1-2* (*Os03g0223400*), *OsGS1-3* (*Os03g0712800*) and *OsGS2* (*Os04g0659100*). For each gene, we analyzed the 1.5 kb upstream promoter region and complete CDS using resequencing data from 1975–2003 rice accessions, including approximately 1975 cultivated rice (*Oryza sativa*) and 14–19 wild rice (*Oryza rufipogon*) accessions obtained from the MBKbase database (**Supplementary Table 2**).

OsGS1-1 exhibited the most pronounced evolutionary dynamics among all *GS* family members. After strict quality filtering, a total of 10 single nucleotide variants (SNVs) were identified, defining 48 distinct haplotypes (Fig. 7A). *OsGS1-1* displayed a classic domestication bottleneck signature, though a few haplotypes were shared between cultivated and wild rice. Two dominant haplotypes accounted for 74.8% of all cultivated rice accessions: Hap1 (n = 851, 42.9%), which was almost exclusively restricted to the indica subspecies (90.1% indica), and Hap2 (n = 632, 31.9%), which was predominantly enriched in the japonica subspecies (78.2% japonica). This extreme complementary distribution pattern suggests that *OsGS1-1* may have experienced divergent selection during the independent domestication of indica and japonica rice. Phylogenetic analysis (Fig. 7B) and haplotype network (Fig. 7C) revealed three distinct evolutionary clades corresponding to indica, japonica and wild rice, with 5–7 mutational steps separating cultivated and wild haplotypes. Notably, 8 of the 10 SNVs were located in the promoter region, and 2 SNVs in CDS suggested that functional divergence could potentially be primarily intended to be mediated by transcriptional regulation.

For *OsGS1-2*, only one high-quality SNV was identified in the promoter region, defining only 3 haplotypes. Other single nucleotide polymorphisms (SNPs) were located in introns. Similar to *OsGS1-1*, *OsGS1-2* exhibited a domestication bottleneck with 14 wild-rice accessions not present in cultivated rice. Two dominant haplotypes accounted for 98.1% of cultivated rice accessions: Hap1 (n = 1439, 72.0%) and Hap2 (n = 524, 26.2%). Hap1 was predominantly found in indica accessions (81.2% indica), while Hap2 was enriched in japonica accessions (84.5% japonica), indicating a moderate level of indica-japonica subspecies differentiation, although less pronounced than that observed for *OsGS1-1*. Non-synonymous mutations were not detected in the CDS, confirming the high conservation of the OsGS1-2 protein sequence (**Supplementary Fig. 8**).

A total of 11 SNVs were identified in *OsGS1-3*, including 10 SNVs in the promoter and 1 in the 3′UTR, defining 31 haplotypes. A single dominant haplotype, Hap1 (n

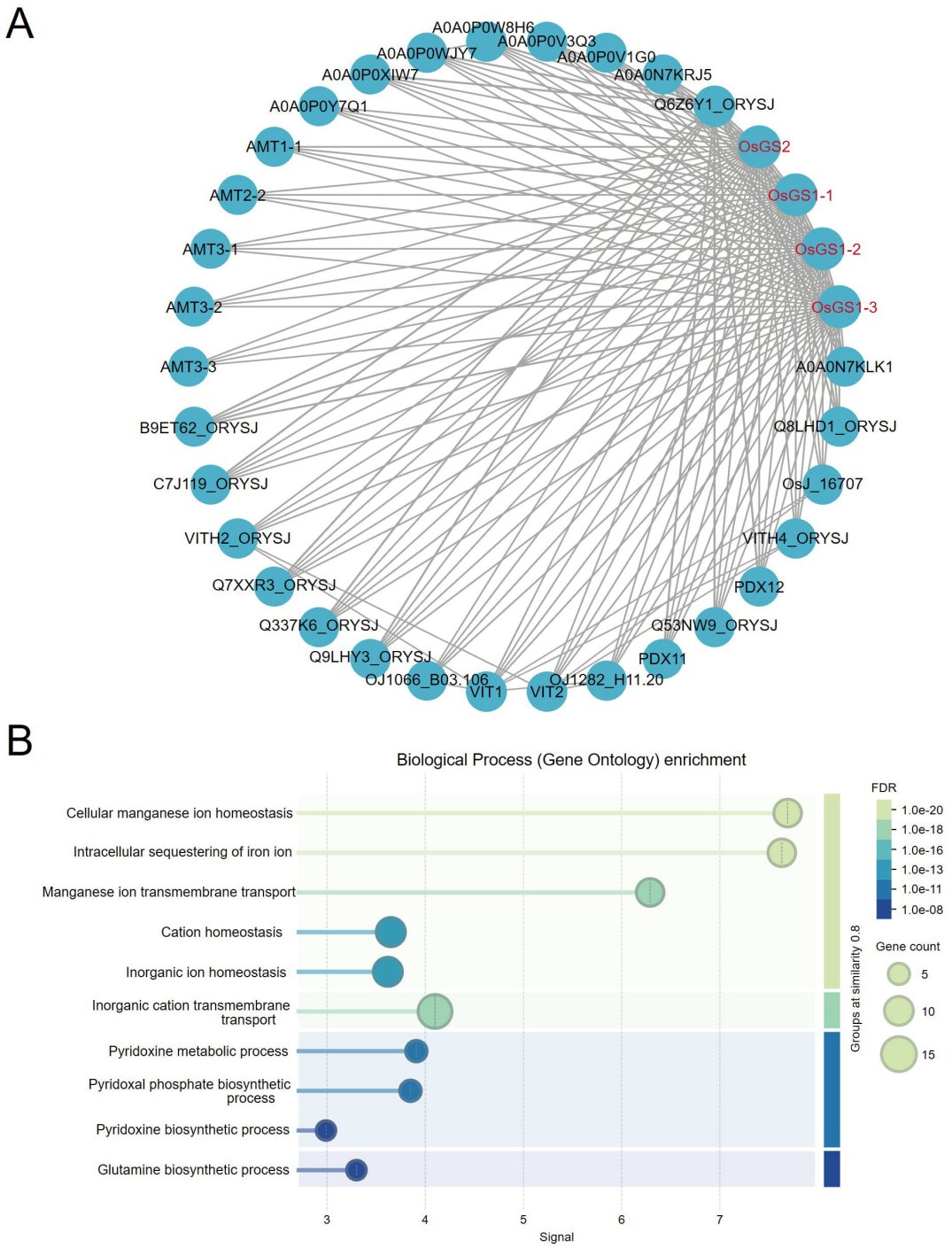

Fig. 6. Protein–protein interaction (PPI) network and Gene Ontology (GO) enrichment analysis of rice OsGS proteins. (A) Protein-protein interaction network of four OsGS proteins (OsGS1-1, OsGS1-2, OsGS1-3, and OsGS2, highlighted in red). Gray lines denote pairwise protein interactions. (B) GO biological process enrichment analysis of the interacting proteins was performed, with the x-axis indicating the signal value, the y-axis listing the enriched biological processes, the size of the circles representing the gene count, and the color gradient representing the false discovery rate (FDR).

= 1363, 68.8%), was present at high frequency across all rice subpopulations: 777 indica, 293 temperate japonica and 239 tropical japonica accessions. The remaining 30 haplotypes were all low-frequency variants with no significant subpopulation enrichment. Notably, no unique wild rice haplotypes were detected, indicating that *OsGS1-3* has not experienced a strong genetic bottleneck during domestication. All mutations in the CDS region were synonymous substitutions, further supporting the functional conservation of this gene (**Supplementary Fig. 9**).

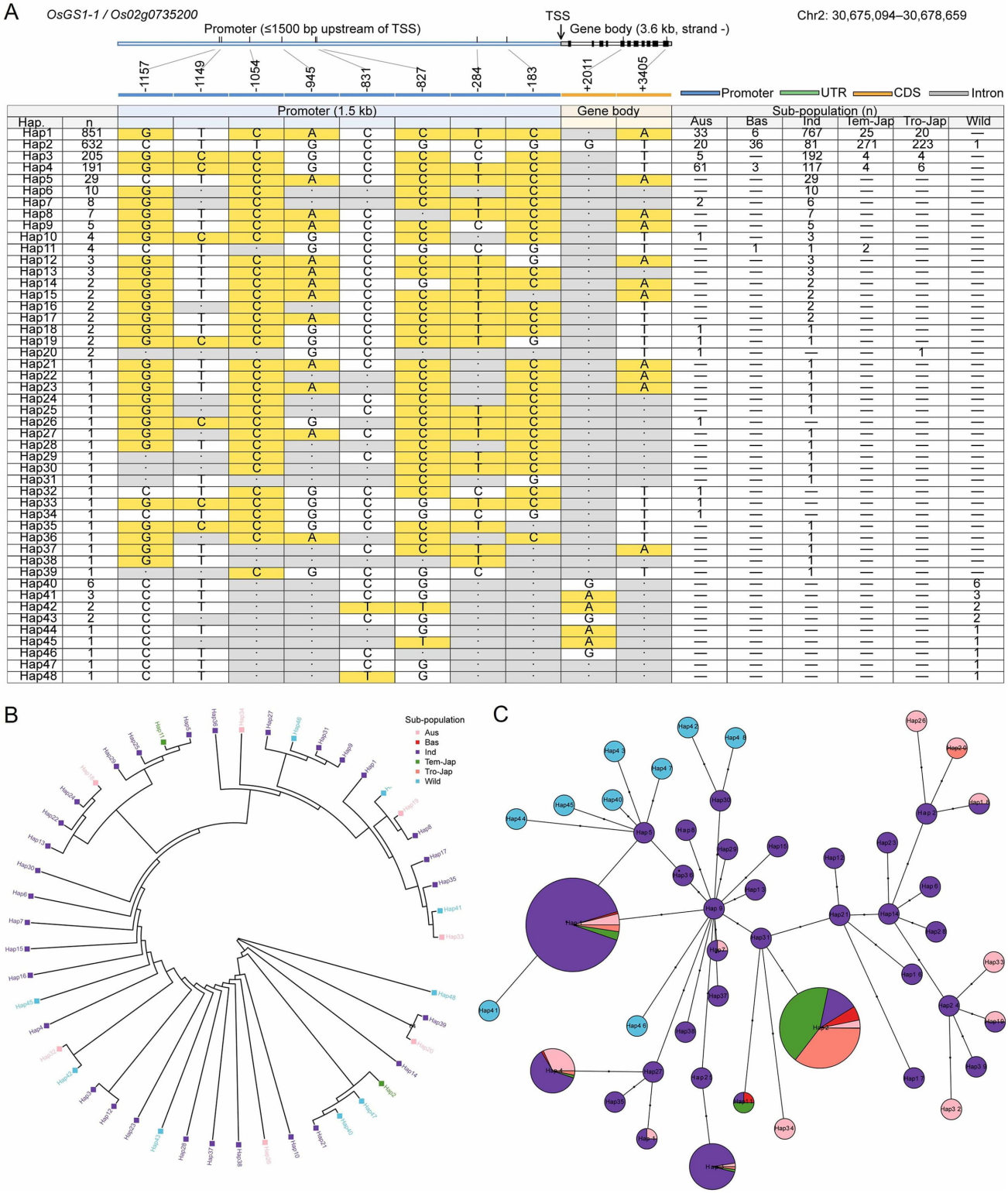

Fig. 7. Haplotype analysis of *OsGS1-1*. (A) Haplotype polymorphism matrix of *OsGS1-1* promoter and coding regions from 2003 rice accessions. Yellow-highlighted boxes indicate different nucleotides, and white boxes within letters indicate non-polymorphic nucleotide sites, and white boxes within black dots indicate no sequencing coverage. Ind, indica population. Tem-Jap, temperate japonica population. Tro-Jap, tropical japonica population. Bas, Basmati population. (B) Phylogenetic tree of the 48 haplotypes from 1984 cultivated and 19 wild rice accessions. Text on nodes = bootstrap %. Branches: cladogram (equal lengths). (C) *OsGS1-1* haplotype network. The size of each circle is proportional to the sample count per haplotype. Black dots on the connecting lines denote mutational steps between different haplotypes.

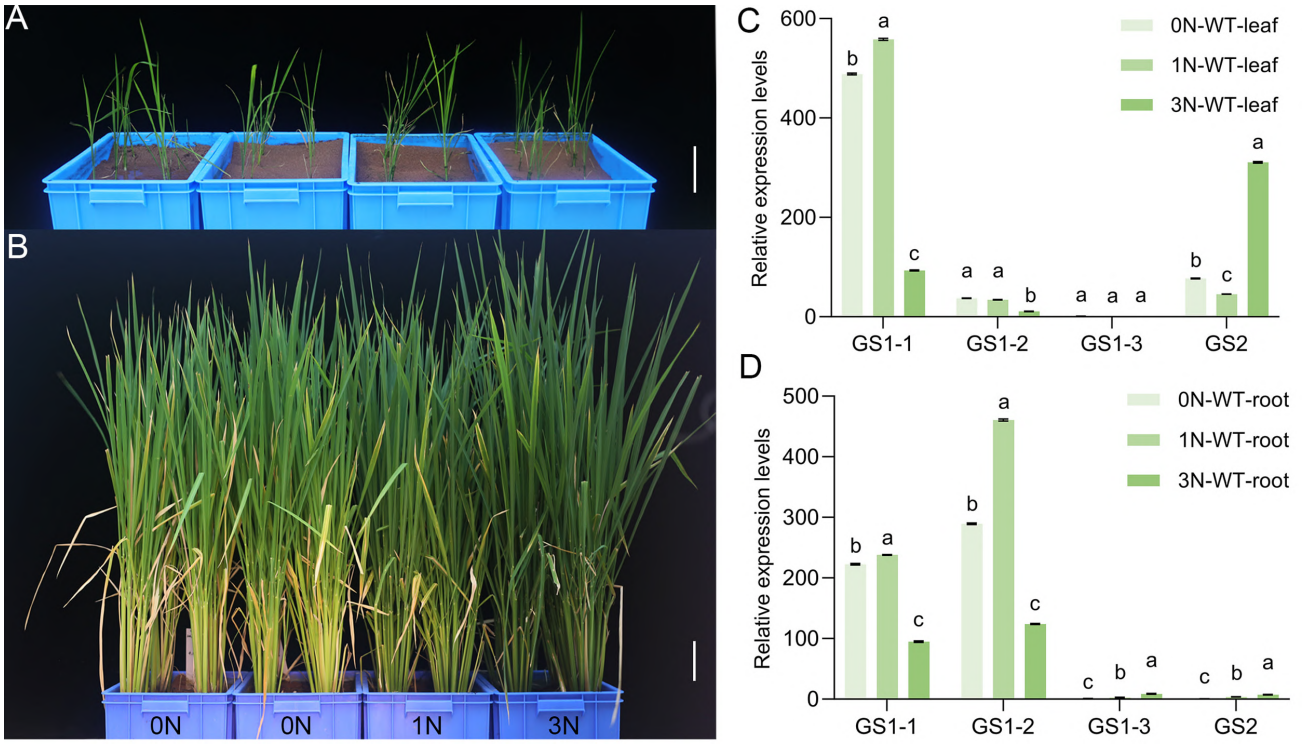

Fig. 8. Transcriptional profiles of *GS* genes in rice leaf and root tissues under three nitrogen gradient treatments. (A,B) All rice materials used in this study were of the Zhongxian 3037 genetic background. Rice seedlings were transplanted into porous ceramic supports and cultivated with Kimura B complete nutrient solution. (A) Whole plants at seedling stage. Bar = 10 cm. (B) Whole plants after 50 days for vegetative growth treated with three nitrogen nutritional conditions (0N, 1N and 3N). To meet the sampling quantity requirements for plant material collection, two pots of rice plants treated with 0N (low-nitrogen condition) were prepared. Bar = 10 cm. (C) Expression dynamics of four *GS* genes during reproductive development (August–September) in leaf tissues via RNA-seq. Results are shown as mean ± SD ($n = 3$). Distinct lowercase letters indicate statistically significant differences among different nitrogen treatment groups for the same gene ($p < 0.05$); groups with the same letter show no significant difference. (D) Expression dynamics of four *GS* genes during reproductive development (August–September) in root tissues obtained by RNA-seq analysis. FPKM values are shown as relative expression levels. Values are presented as mean ± SD ($n = 3$). Distinct lowercase letters indicate statistically significant differences among different nitrogen treatment groups for the same gene ($p < 0.05$); groups with the same letter show no significant difference.

Coincidentally, only one high-quality SNV was identified in the promoter region of *OsGS2* like *OsGS1-2*, defining 3 haplotypes. For *OsGS2*, one SNV was identified in the promoter region and one in the CDS region, while the remaining three SNPs were all located in intronic regions. *OsGS2* exhibited a moderate domestication bottleneck with 14 wild-rice accessions carrying private haplotypes absent in cultivated rice. Two dominant haplotypes accounted for 97.2% of all cultivated rice accessions: Hap1 (n = 1050, 52.5%) and Hap2 (n = 895, 44.7%). Hap1 was more prevalent in indica accessions (83.3% indica), while Hap2 showed weak enrichment in japonica accessions (44.1% japonica), indicating a very low level of subspecies differentiation. Non-synonymous mutations were not detected in the CDS region, with only one SNV located in the promoter region, consistent with the essential and highly conserved function of OsGS2 in photosynthetic nitrogen assimilation (**Supplementary Fig. 10**).

3.12 Expression Profiles of OsGS Family Genes in Source and Sink Tissues Under Different Nitrogen Regimes

To investigate the transcriptional responses of *GS* family genes to nitrogen availability and reveal the molecular mechanism underlying nitrogen remobilization, reallocation and efficient source-sink communication in rice, we performed high-throughput transcriptome sequencing to systematically analyze the expression patterns of four *OsGS* members (*OsGS1-1*, *OsGS1-2*, *OsGS1-3*, and *OsGS2*) in major nitrogen source and sink tissues (**Supplementary Table 4**).

In this study, *Oryza sativa* cv. Zhongxian 3037, an elite indica rice cultivar with stable genetic background, was hydroponically cultured with modified Kimura B nutrient solution under three strictly controlled and biologically representative nitrogen regimes: 0N (low-nitrogen stress), 1N (normal nitrogen condition, standard control), and 3N (high-nitrogen treatment). This graded nitrogen experi-

mental design possesses profound agronomic and scientific significance: 0N accurately simulates the nitrogen-limited soil conditions that widely restrict crop growth in low-input agricultural systems, 1N corresponds to the optimal nitrogen supply that maintains normal vegetative and reproductive growth of rice in conventional cultivation, and 3N replicates the excessive nitrogen fertilization practice commonly adopted in intensive rice production, which causes low nitrogen use efficiency and serious environmental pollution [4,55]. Collectively, these three nitrogen gradients establish a rigorous comparative system to dissect the adaptive transcriptional regulation of nitrogen metabolic genes, and to clarify how external nitrogen levels reshape the source-sink nitrogen distribution network in rice.

In leaf tissues, the chloroplastic isoform *OsGS2* was highly expressed across four *OsGS* genes (Fig. 8A–C), which is highly consistent with its well-documented specialized function in reassimilating the large amount of NH_4^+ released from photorespiration in photosynthetic green organs [56]. The transcript abundance of *OsGS1-1* and *OsGS1-2* was significantly and negatively regulated by external nitrogen supply levels (Fig. 8C): their transcripts were significantly reduced under 3N conditions but strongly induced under 0N stress in leaves, while *OsGS2* exhibited an opposite regulatory pattern with the highest expression under 3N conditions. In contrast, the grain-specific cytosolic isoform *OsGS1-3* showed extremely low basal expression in leaf tissues and was not sensitive to changes in external nitrogen availability, reflecting its strict tissue-specific expression pattern in reproductive tissues.

In root tissues, the root-predominant *OsGS1-2*, an essential functional gene responsible for primary nitrogen assimilation of absorbed NH_4^+ in rice root system, displayed the highest expression level among all family members (Fig. 8D). Consistent with the expression trend in leaf tissues, compared with 1N conditions, nitrogen deficiency led to a significant reduction in the transcript abundance of cytosolic *OsGS1* isoforms (*OsGS1-1* and *OsGS1-2*), whereas excess nitrogen supply also repressed their transcription in root tissues, with both genes reaching peak transcript levels under 1N. And although *OsGS2* and *OsGS1-3* were up-regulated along the nitrogen gradient, their transcript levels remained comparatively low relative to *OsGS1-1* and *OsGS1-2*. These results imply that external nitrogen availability strongly and differentially modulates the transcriptional activity of *OsGS* family genes in vegetative source tissues, and the distinct tissue-specific expression pattern of each *OsGS* member are consistent with their proposed non-redundant biological function in nitrogen uptake, primary assimilation and metabolic recycling.

We further profiled the detailed spatiotemporal expression dynamics of *OsGS* family genes in developing young embryos at four consecutive and critical developmental stages (1, 2, 3, and 5 DAP) (Fig. 9A) (**Supplementary Table 5**). Morphological and cytological observations confirmed that rice young embryos undergo a highly ordered, precise and sequential developmental program after fertilization: zygote polarity establishment and initial asymmetric cell division at 1 DAP, compact globular proembryo formation and vigorous cell proliferation at 2 DAP, active organ primordium initiation and tissue regional patterning at 3 DAP, and complete embryonic organ maturation coupled with the most vigorous starch biosynthesis, amyloplast proliferation and massive storage reserve deposition at 5 DAP [35]. Transcriptome data showed that mRNA of four *OsGS* genes were constitutively detected throughout the entire embryonic developmental process, with clear and distinct stage-specific expression characteristics. Under 1N conditions, the transcriptional abundance of most *OsGS* genes increased gradually and continuously with embryonic development, with *OsGS1-3* reaching its prominent maximum expression peak at 5 DAP, which was highly synchronized with the rapid accumulation of starch and storage reserves in late-stage embryos (Fig. 9B). Different nitrogen supply treatments significantly and differentially altered the transcriptional abundance of *OsGS* genes in young embryos: 0N stress tended to severely repress most OsGS family members at certain developmental stages, whereas 3N treatment remarkably and continuously induced transcription for a subset of these genes, with the most significant regulatory effect observed at the key storage accumulation stage of 5 DAP (Fig. 9C–F). Notably, the grain-specific *OsGS1-3* showed a dramatic and specific upregulation trend during late embryogenesis, suggesting that it may play an important role in nitrogen remobilization, translocation and storage protein synthesis in developing rice embryos (Fig. 9E). Collectively, these comprehensive findings demonstrated that *OsGS* family genes exhibit typical tissue-specific, developmental stage-specific and nitrogen-responsive expression patterns in rice, and their coordinated transcriptional regulation at the critical 5 DAP stage plays a pivotal regulatory role in coupling nitrogen assimilation with carbon storage metabolism, thereby fine-tuning source-sink nitrogen utilization efficiency and normal embryonic development in rice.

3.13 Tissue-Specific Regulation of GS Activity by Exogenous Inorganic Nitrogen in Rice

We determined the GS enzyme activity in flag leaves, roots and developing embryos of wild-type rice under three nitrogen regimes (0N, 1N and 3N) (**Supplementary Table 6**). The results showed that GS activity exhibited distinct tissue-specific regulatory patterns in response to exogenous inorganic nitrogen levels (Fig. 10).

In flag leaves and developing embryos (at 1, 5 and 10 days after pollination), GS activity was significantly decreased under high nitrogen (3N) treatment compared with the normal nitrogen (1N) control, while no statistically significant difference was detected between nitrogen-free (0N) and normal nitrogen conditions. In contrast, as the most

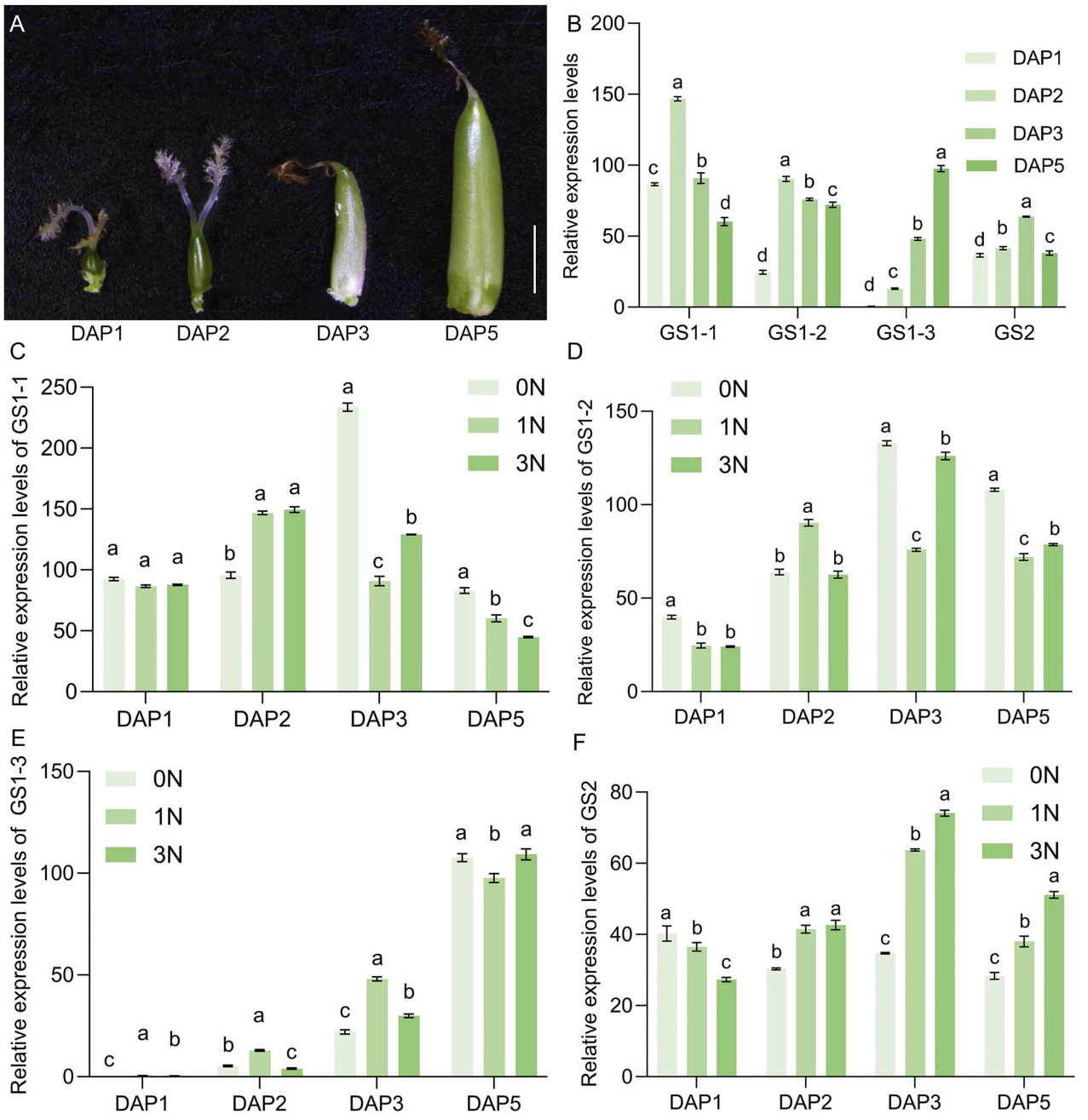

Fig. 9. Expression pattern analysis of *GS* genes in rice young embryos under three nitrogen gradient treatments. (A) Morphology of rice young embryos at four stages (1, 2, 3 and 5 days after pollination). Bar = 3 mm. (B) Expression dynamics of four *GS* genes during reproductive development (August–September) in young embryos of rice treated with normal nitrogen nutrition condition (1N) via RNA-seq. Bars represent mean ± SD (n = 3). Distinct lowercase letters indicate statistically significant differences among different developmental stages for the same gene under the 1N condition ($p < 0.05$); groups with the same letter show no significant difference. (C–F) Expression dynamics of four *GS* genes in young embryos of rice treated with three nitrogen nutrition conditions (0N, 1N and 3N) via RNA-seq. Bars represent mean ± SD (n = 3). Distinct lowercase letters indicate statistically significant differences among the three nitrogen treatments (0N, 1N, 3N) for the same gene at the same developmental stage ($p < 0.05$); groups sharing the same letter are not significantly different.

direct nitrogen source tissue, roots showed a different response pattern: both high nitrogen (3N) and nitrogen-free (0N) treatments led to a significant reduction in GS activity relative to the normal nitrogen control.

Importantly, we observed that GS activity displayed an identical developmental trajectory in early rice embryos regardless of exogenous nitrogen supply. Specifically, GS activity increased from 1 DAP to 5 DAP, reached its peak

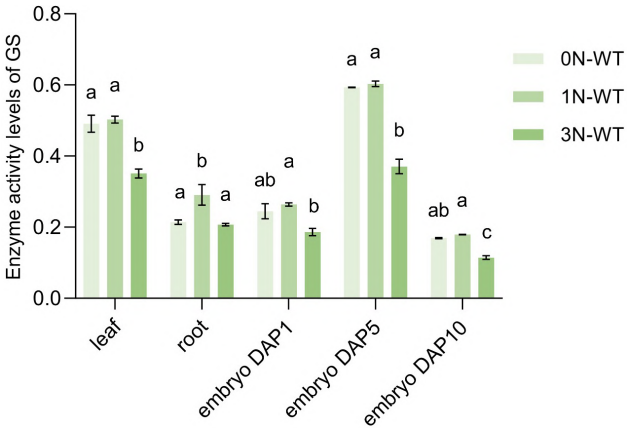

Fig. 10. GS enzyme activity in source and sink tissues of wild-type rice under different nitrogen concentrations. Source tissues: leaf and root; sink tissues: embryo at 1, 5 and 10 DAP. 0N, nitrogen-free treatment; 1N, normal nitrogen treatment (control); 3N, high nitrogen treatment. Enzyme activity assays were performed with three independent biological replicates, each containing three technical replicates. Values are expressed as mean ± standard error (SE). Bars with different lowercase letters indicate statistically significant differences among the three nitrogen treatments (0N, 1N, 3N) within the same tissue or developmental stage ($p < 0.05$), as determined by one-way ANOVA followed by Tukey's HSD post-hoc test; bars sharing the same letter show no significant difference.

at 5 DAP, and subsequently decreased by 10 DAP. Given that GS is the only enzyme catalyzing glutamine formation and plays a central role in nitrogen remobilization to sink organs, this temporal pattern is highly consistent with the metabolic demand of developing embryos. In rice, most nitrogenous compounds (amino acids and proteins) are deposited in the embryo and aleurone layer during seed development. Therefore, our findings provide direct biochemical evidence that 5 DAP is the metabolically most active stage, corresponding to the completion of aleurone layer differentiation and the onset of rapid starch accumulation in the endosperm.

4. Discussion

4.1 Divergent Evolutionary Trajectories of Rice GS Family Members Are Tightly Coupled to Their Functional Specialization

Our comprehensive haplotype analysis of the four rice *GS* genes revealed strikingly distinct evolutionary patterns that directly reflect their non-redundant roles in nitrogen metabolism. *OsGS1-1* emerged as the most evolutionarily dynamic member, exhibiting extreme indica-japonica differentiation and a pattern consistent with a domestication bottleneck. This pattern is consistent with its central role in root primary nitrogen assimilation [13,30], a process subject to intense divergent selection during the independent domestication of indica and japonica rice in distinct ecological environments. In contrast, *OsGS1-3* showed minimal subspecies differentiation and no domestication bottleneck, likely due to its specialized function in reproductive tissues that are less sensitive to environmental nitrogen variation. Notably, all four *GS* genes displayed strong purifying selection on their coding regions, with functional variation restricted almost exclusively to promoter regions. This finding provides compelling evidence that the catalytic function of GS enzymes is evolutionarily immutable, while transcriptional regulation has been the primary driver of adaptive evolution in the rice GS family.

4.2 A Reproductive-Stage Source-Sink Experimental System Uncovers Novel Nitrogen Response Mechanisms

A major innovation of this study is the establishment of a porous ceramic-based hydroponic system that enabled precise nitrogen gradient treatments (0N, 1N, 3N) during the critical reproductive growth stage of rice. Unlike previous studies that focused primarily on vegetative tissues [33,57], we simultaneously profiled *GS* expression and enzyme activity in both source tissues (roots and flag leaves) and sink tissues (developing embryos at four consecutive stages). This integrated approach revealed that nitrogen availability exerts tissue-specific regulatory effects on *GS* family members: while root GS activity was suppressed under both low and high nitrogen stress, leaf and embryo GS activity was only reduced under high nitrogen conditions. Furthermore, we identified 5 DAP as the metabolically most active stage for nitrogen assimilation in developing embryos, where *GS* expression and activity peaked in synchrony with rapid storage reserve accumulation. These findings provide a previously unrecognized temporal and spatial resolution of nitrogen metabolism during rice reproductive development.

4.3 Transcriptional and Biochemical Evidence Reveals Functional Division and Coordination of GS Isoforms in Source-Sink Nitrogen Allocation

Despite high sequence conservation at the amino acid level, the four rice GS isoforms exhibited distinct tissue-specific expression patterns and nitrogen response profiles. *OsGS1-1* showed constitutive abundant expression in all tested tissues, consistent with its housekeeping function in basal nitrogen assimilation [19]. *OsGS1-2* was predominantly expressed in roots, reflecting its specialized function in primary nitrogen uptake [21]. *OsGS1-3* displayed strict reproductive tissue specificity, being highly expressed only in developing embryos, which is consistent with a potential role in nitrogen remobilization to seeds. *OsGS2* was almost exclusively expressed in photosynthetic tissues, consistent with its function in photorespiratory NH_4^+ reassimilation. Importantly, these expression patterns were mirrored at the enzyme activity level, suggesting that transcriptional regulation is the primary mechanism controlling GS function in

different tissues. Collectively, our results support a working model of *GS* family function in rice, where distinct isoforms act in a coordinated manner to mediate efficient nitrogen assimilation, translocation, and storage across source and sink tissues.

5. Limitations and Future Perspectives

This study provides a comprehensive multi-omics atlas of the rice GS family, establishing a foundational framework for understanding its roles in source-sink nitrogen allocation. While our integrative analyses offer robust evolutionary and transcriptional insights, certain aspects warrant further investigation in future studies.

Evolutionary Mechanisms: Our cross-species synteny analysis revealed lineage-specific expansion patterns, supporting the role of duplication events in functional diversification. To precisely date these events, subsequent research should integrate intraspecific collinearity analysis with Ks calculations, calibrated against established ancient polyploidization events in rice. This will refine the evolutionary timeline of the *GS* family and clarify how specific duplication types contributed to its expansion and functional partitioning.

Validation of Computational Predictions: Key findings regarding cis-regulatory elements (PlantCARE), protein structures (AlphaFold2/SOPMA), and interaction networks (STRING) are derived from *in silico* predictions. These predictions provide high-value hypotheses for downstream validation. For instance, predicted structural divergences between GS1 and GS2 isoforms offer a plausible basis for their functional specialization, while predicted interaction networks highlight candidate regulatory modules. Future experimental work, employing reporter assays, ChIP, X-ray crystallography, and Co-IP, will be essential to transition these bioinformatic inferences into empirical evidence.

Functional Correlation and Genetic Context: We identified tissue-specific and nitrogen-responsive patterns at both transcriptional and enzymatic levels. However, the quantitative relationship between transcript abundance and enzymatic activity, along with the direct genetic validation of isoform function, remains to be elucidated. While the use of a single elite *indica* cultivar (Zhongxian 3037) minimized genetic background noise, expanding this work to diverse germplasm will be crucial to confirm the broader applicability of our findings. Furthermore, integrating CRISPR-Cas9-mediated gene editing and targeted qRT-PCR will help establish causal links between gene expression, enzyme activity, and phenotypic outcomes.

Physiological and Population Genetics Extensions: The absence of direct tissue nitrogen measurements limits the precision of linking GS activity to *in planta* nitrogen assimilation efficiency. Future physiological experiments incorporating total nitrogen quantification will address this gap. Additionally, while our haplotype analysis suggests strong selection on *OsGS1-1*, formal population genetic tests—such as Tajima's *D* and F_{ST} scans—are required to confirm selection signals and intensities. Functional characterization of promoter polymorphisms via dual-luciferase assays will also be vital to determine their impact on gene expression and nitrogen use efficiency.

In conclusion, this study constructs a comprehensive theoretical framework for the rice GS family. Future integration of quantitative physiology, structural biology, and population genomics will be pivotal in translating these correlative insights into causal mechanisms, thereby providing robust theoretical support for molecular breeding aimed at improving nitrogen use efficiency in rice.

6. Conclusions

This study presents comprehensive characterization of the rice *GS* gene family to date, integrating evolutionary genomics, transcriptomics, and biochemical approaches to elucidate their roles in source-sink nitrogen allocation under different nitrogen regimes. We identified 86 *GS* genes across 12 plant species, revealing that the GS family has undergone strong purifying selection during evolution, with significant species-specific expansion in polyploid crops. In rice, the four *GS* members exhibited distinct evolutionary trajectories: *OsGS1-1* showed extreme indica-japonica differentiation and a pattern consistent with a domestication bottleneck, while *OsGS1-3* displayed minimal subspecies differentiation, reflecting their divergent functional roles.

A key innovation of this study is the establishment of a reproductive-stage nitrogen gradient experimental system, which enabled simultaneous profiling of *GS* expression and enzyme activity in both source tissues (roots and flag leaves) and sink tissues (developing embryos at four consecutive stages). We identified 5 days after pollination as the critical period for nitrogen assimilation in developing embryos, where *GS* expression and activity peaked in synchrony with rapid storage reserve accumulation. Furthermore, we demonstrated that despite high sequence conservation, the four rice GS isoforms exhibit distinct tissue-specific expression patterns and nitrogen response profiles, indicating distinct transcriptional response patterns that are consistent with functional division and coordination in nitrogen metabolism.

Collectively, our findings provide fundamental insights into the evolutionary and functional diversification of the plant *GS* family. The divergent haplotypes of *OsGS1-1* between indica and japonica rice represent valuable genetic resources for molecular breeding to improve nitrogen use efficiency in rice, addressing the pressing global challenges of food security and environmental sustainability.

Abbreviations

ABA, Abscisic acid; AMT, Ammonium transporter; ANOVA, Analysis of Variance; ATP, Adenosine triphosphate; BLASTP, Basic Local Alignment Search Tool

for Proteins; CAT, Category; CDD, Conserved Domain Database; CDS, Coding sequence; ChIP, Chromatin Immunoprecipitation; Co-IP, Co-Immunoprecipitation; DAP, Days after pollination; FDR, False discovery rate; Glu, Glutamate; Gln, Glutamine; GO, Gene Ontology; GOGAT, Glutamate synthase; GS, Glutamine synthetase; Hap, haplotype; HISAT2, Hierarchical Indexing for Spliced Alignment of Transcripts 2; HMM, Hidden Markov model; HSD, Honestly Significant Difference; Ka, Non-synonymous substitution rate; KEGG, Kyoto Encyclopedia of Genes and Genomes; Ks, Synonymous substitution rate; ML, Maximum likelihood; NCBI, National Center for Biotechnology Information; N, Nitrogen; NUE, Nitrogen use efficiency; PCA, Principal Component Analysis; PDX, Pyridoxal metabolism-related protein; PEG, Polyethylene Glycol; Pfam, Protein Family Database; PPI, Protein–protein interaction; qRT-PCR, quantitative Reverse Transcription Polymerase Chain Reaction; RFGB, Rice Functional Genomics and Breeding database; RNA-seq, RNA sequencing; rlog, Regularized Logarithm; SMART, Simple Modular Architecture Research Tool; SNPs, single nucleotide polymorphisms; SNVs, single nucleotide variants; SOPMA, Self-Optimized Prediction Method with Alignment; STRING, Search Tool for the Retrieval of Interacting Genes/Proteins; TAIR, The *Arabidopsis* Information Resource; TBtools, Toolbox for Biologists; UPGMA, Unweighted Pair Group Method with Arithmetic Mean; UTR, Untranslated region; VITH, Vitamin transporter; WT, Wild Type.

Availability of Data and Materials

The original contributions presented in this study are included in the article/supplementary material. The raw data for RNA-seq of source and sink tissue under different nitrogen treatment can be accessed from NCBI SRA database (accession number: PRJNA1473341). Further inquiries can be directed to the corresponding authors.

Author Contributions

Data curation, formal analysis, investigation, writing original draft, HH; analysis of RNA-seq, data validation, software, SF; plant material cultivation, maintenance and treatment, HZ; biological sample collection for RNA-seq, ZX and YW; conceptualization, data validation, supervision, ZL; experiment designing, funding acquisition, methodology, manuscript revision, HY. All authors contributed to editorial changes in the manuscript. All authors read and approved the final manuscript. All authors have participated sufficiently in the work and agreed to be accountable for all aspects of the work.

Ethics Approval and Consent to Participate

Rice plants (Zhongxian 3037) used in this study were provided by Dr. Yang Han, and the original materials were obtained from the Institute of Genetics and Developmental Biology, Chinese Academy of Sciences (CAS, Beijing, China).

Acknowledgment

We gratefully acknowledge the assistance and instruction from Professor Lifu Qian and Professor Jiyuan Wang from College of Life Sciences of Huaibei Normal University for helping to revise the manuscript.

Funding

This work was supported by grants from the National Natural Science Foundation of China (grants 32401891), the Project of Anhui Provincial Association for Science and Technology (RCTJ202406), and the Natural Science Foundation of Anhui Province (2308085QC81), the Scientific Research Project of Anhui Provincial Universities (2022AH030058), Anhui Provincial Department of Education Scientific Research Project (2025AHGXZK40140), Open Project of School-level Laboratories of Huaibei Normal University (2025syskf030), and Anhui Province College Students' Innovation and Entrepreneurship Training Program Project (X202610373027).

Conflicts of Interest

The authors declare no conflicts of interest.

Declaration of AI and AI-Assisted Technologies in the Writing Process

During the preparation of this work, the authors used Doubao AI (v10.1.0) in order to check spelling and grammar. After using this tool, the authors reviewed and edited the content as needed and take full responsibility for the content of the publication.

Supplementary Material

Supplementary material associated with this article can be found, in the online version, at https://doi.org/10.31083/FBL54457.

References

[1] Siqueira JA, da Fonseca-Pereira P, Fernie AR, Nunes-Nesi A, Araújo WL. Recycling amino acids ensures meiosis and seed development. Trends in Plant Science. 2022; 27: 1084–1086. https://doi.org/10.1016/j.tplants.2022.07.010

[2] Gaudinier A, Rodriguez-Medina J, Zhang L, Olson A, Liseron-Monfils C, Bågman AM, et al. Transcriptional regulation of nitrogen-associated metabolism and growth. Nature. 2018; 563: 259–264. https://doi.org/10.1038/s41586-018-0656-3

[3] Luo L, Zhang Y, Xu G. How does nitrogen shape plant architecture? Journal of Experimental Botany. 2020; 71: 4415–4427. https://doi.org/10.1093/jxb/eraa187

[4] Yang H, Li Y, Cao Y, Shi W, Xie E, Mu N, et al. Nitrogen nutrition contributes to plant fertility by affecting meiosis initiation. Nature Communications. 2022; 13: 485. https://doi.org/10.1038/s41467-022-28173-3

[5] Yang H, Fu S, Hu HT, Chen C, Xu XX, Li GP. An Overview of Molecular Mechanisms Underlying Nitrogen-Mediated Regulation of Plant Fertility Establishment. Journal of Huaibei Normal University (Natural Science). 2026; 47: 25–30. (In Chinese)
[6] Wen A, Havens KL, Bloch SE, Shah N, Higgins DA, Davis-Richardson AG, et al. Enabling Biological Nitrogen Fixation for Cereal Crops in Fertilized Fields. ACS Synthetic Biology. 2021; 10: 3264–3277. https://doi.org/10.1021/acssynbio.1c00049
[7] Swamy BPM, Kumar A. Genomics-based precision breeding approaches to improve drought tolerance in rice. Biotechnology Advances. 2013; 31: 1308–1318. https://doi.org/10.1016/j.biotechadv.2013.05.004
[8] Hou M, Yu M, Li Z, Ai Z, Chen J. Molecular Regulatory Networks for Improving Nitrogen Use Efficiency in Rice. International Journal of Molecular Sciences. 2021; 22: 9040. https://doi.org/10.3390/ijms22169040
[9] Guo JH, Liu XJ, Zhang Y, Shen JL, Han WX, Zhang WF, et al. Significant acidification in major Chinese croplands. Science (New York, N.Y.). 2010; 327: 1008–1010. https://doi.org/10.1126/science.1182570
[10] Dimkpa CO, Fugice J, Singh U, Lewis TD. Development of fertilizers for enhanced nitrogen use efficiency - Trends and perspectives. The Science of the Total Environment. 2020; 731: 139113. https://doi.org/10.1016/j.scitotenv.2020.139113
[11] Bernard SM, Habash DZ. The importance of cytosolic glutamine synthetase in nitrogen assimilation and recycling. The New Phytologist. 2009; 182: 608–620. https://doi.org/10.1111/j.1469-8137.2009.02823.x
[12] Moreira E, Coimbra S, Melo P. Glutamine synthetase: an unlikely case of functional redundancy in Arabidopsis thaliana. Plant Biology (Stuttgart, Germany). 2022; 24: 713–720. https://doi.org/10.1111/plb.13408
[13] Fortunato S, Nigro D, Lasorella C, Marcotuli I, Gadaleta A, de Pinto MC. The Role of Glutamine Synthetase (GS) and Glutamate Synthase (GOGAT) in the Improvement of Nitrogen Use Efficiency in Cereals. Biomolecules. 2023; 13: 1771. https://doi.org/10.3390/biom13121771
[14] Liu X, Hu B, Chu C. Nitrogen assimilation in plants: current status and future prospects. Journal of Genetics and Genomics. 2022; 49: 394–404. https://doi.org/10.1016/j.jgg.2021.12.006
[15] Yin H, Yang F, He X, Du X, Mu P, Ma W. Advances in the functional study of glutamine synthetase in plant abiotic stress tolerance response. The Crop Journal. 2022; 10: 917–923. https://doi.org/10.1016/j.Cj.2022.01.003
[16] Lothier J, Gaufichon L, Sormani R, Lemaître T, Azzopardi M, Morin H, et al. The cytosolic glutamine synthetase GLN1;2 plays a role in the control of plant growth and ammonium homeostasis in Arabidopsis rosettes when nitrate supply is not limiting. Journal of Experimental Botany. 2011; 62: 1375–1390. https://doi.org/10.1093/jxb/erq299
[17] Thomsen HC, Eriksson D, Møller IS, Schjoerring JK. Cytosolic glutamine synthetase: a target for improvement of crop nitrogen use efficiency? Trends in Plant Science. 2014; 19: 656–663. https://doi.org/10.1016/j.tplants.2014.06.002
[18] Martin A, Lee J, Kichey T, Gerentes D, Zivy M, Tatout C, et al. Two cytosolic glutamine synthetase isoforms of maize are specifically involved in the control of grain production. The Plant Cell. 2006; 18: 3252–3274. https://doi.org/10.1105/tpc.106.042689
[19] Liu X, Tian Y, Chi W, Zhang H, Yu J, Chen G, et al. Alternative splicing of OsGS1;1 affects nitrogen-use efficiency, grain development, and amylose content in rice. The Plant Journal: for Cell and Molecular Biology. 2022; 110: 1751–1762. https://doi.org/10.1111/tpj.15768
[20] Bao A, Zhao Z, Ding G, Shi L, Xu F, Cai H. Accumulated expression level of cytosolic glutamine synthetase 1 gene (OsGS1;1 or OsGS1;2) alter plant development and the carbon-nitrogen metabolic status in rice. PloS One. 2014; 9: e95581. https://doi.org/10.1371/journal.pone.0095581
[21] Kusano M, Fukushima A, Tabuchi-Kobayashi M, Funayama K, Kojima S, Maruyama K, et al. Cytosolic GLUTAMINE SYNTHETASE1;1 Modulates Metabolism and Chloroplast Development in Roots. Plant Physiology. 2020; 182: 1894–1909. https://doi.org/10.1104/pp.19.01118
[22] Funayama K, Kojima S, Tabuchi-Kobayashi M, Sawa Y, Nakayama Y, Hayakawa T, et al. Cytosolic glutamine synthetase1;2 is responsible for the primary assimilation of ammonium in rice roots. Plant & Cell Physiology. 2013; 54: 934–943. https://doi.org/10.1093/pcp/pct046
[23] Ohashi M, Ishiyama K, Kusano M, Fukushima A, Kojima S, Hanada A, et al. Lack of cytosolic glutamine synthetase1;2 in vascular tissues of axillary buds causes severe reduction in their outgrowth and disorder of metabolic balance in rice seedlings. The Plant Journal: for Cell and Molecular Biology. 2015; 81: 347–356. https://doi.org/10.1111/tpj.12731
[24] Tabuchi M, Sugiyama K, Ishiyama K, Inoue E, Sato T, Takahashi H, et al. Severe reduction in growth rate and grain filling of rice mutants lacking OsGS1;1, a cytosolic glutamine synthetase1;1. The Plant Journal: for Cell and Molecular Biology. 2005; 42: 641–651. https://doi.org/10.1111/j.1365-313X.2005.02406.x
[25] James D, Borphukan B, Fartyal D, Ram B, Singh J, Manna M, et al. Concurrent Overexpression of *OsGS1;1* and *OsGS2* Genes in Transgenic Rice (*Oryza sativa* L.): Impact on Tolerance to Abiotic Stresses. Frontiers in Plant Science. 2018; 9: 786. https://doi.org/10.3389/fpls.2018.00786
[26] Elsanosi HA, Zhu T, Zhou G, Song L. Genomic organization and expression profiles of nitrogen assimilation genes in *Glycine max*. PeerJ. 2024; 12: e17590. https://doi.org/10.7717/peerj.17590
[27] Lee HJ, Abdula SE, Jang DW, Park SH, Yoon UH, Jung YJ, et al. Overexpression of the glutamine synthetase gene modulates oxidative stress response in rice after exposure to cadmium stress. Plant Cell Reports. 2013; 32: 1521–1529. https://doi.org/10.1007/s00299-013-1464-8
[28] Park SI, Kim YS, Kim JJ, Mok JE, Kim YH, Park HM, et al. Improved stress tolerance and productivity in transgenic rice plants constitutively expressing the Oryza sativa glutathione synthetase OsGS under paddy field conditions. Journal of Plant Physiology. 2017; 215: 39–47. https://doi.org/10.1016/j.jplph.2017.05.006
[29] Xia JQ, Liang QY, He DY, Zhang ZY, Wu J, Zhang J, et al. Knockout of OsSPL10 confers enhanced glufosinate resistance in rice. Plant Communications. 2024; 5: 100731. https://doi.org/10.1016/j.xplc.2023.100731
[30] Li Z, Zhu M, Huang J, Jiang S, Xu S, Zhang Z, et al. Genome-Wide Comprehensive Analysis of the Nitrogen Metabolism Toolbox Reveals Its Evolution and Abiotic Stress Responsiveness in Rice (*Oryza sativa* L.). International Journal of Molecular Sciences. 2022; 24: 288. https://doi.org/10.3390/ijms24010288
[31] Yin H, Sun Q, Lu X, Zhang L, Yuan Y, Gong C, et al. Identification of the glutamine synthetase (GS) gene family in four wheat species and functional analysis of Ta4D.GSe in Arabidopsis thaliana. Plant Molecular Biology. 2022; 110: 93–106. https://doi.org/10.1007/s11103-022-01287-4
[32] Tegeder M, Masclaux-Daubresse C. Source and sink mechanisms of nitrogen transport and use. The New Phytologist. 2018; 217: 35–53. https://doi.org/10.1111/nph.14876
[33] Huang Y, Ji Z, Tao Y, Wei S, Jiao W, Fang Y, et al. Improving rice nitrogen-use efficiency by modulating a novel monouniquitination machinery for optimal root plasticity response to nitrogen. Nature Plants. 2023; 9: 1902–1914. https://doi.org/10.1038/s41477-023-01533-7

[34] Zhang S, Zhang Y, Li K, Yan M, Zhang J, Yu M, et al. Nitrogen Mediates Flowering Time and Nitrogen Use Efficiency via Floral Regulators in Rice. Current Biology: CB. 2021; 31: 671–683.e5. https://doi.org/10.1016/j.cub.2020.10.095

[35] Wu X, Liu J, Li D, Liu CM. Rice caryopsis development II: Dynamic changes in the endosperm. Journal of Integrative Plant Biology. 2016; 58: 786–798. https://doi.org/10.1111/jipb.12488

[36] Huang R, Zhang X, Luo K, Tembrock LR, Li S, Wu Z. The Identification of Auxin Response Factors and Expression Analyses of Different Floral Development Stages in Roses. Genes. 2025; 16: 41. https://doi.org/10.3390/genes16010041

[37] Xiong W, Zhao Y, Gao H, Li Y, Tang W, Ma L, et al. Genomic characterization and expression analysis of *TCP* transcription factors in *Setaria italica* and *Setaria viridis*. Plant Signaling & Behavior. 2022; 17: 2075158. https://doi.org/10.1080/15592324.2022.2075158

[38] Dai WS, Peng T, Wang M, Liu JH. Genome-wide identification and comparative expression profiling of the WRKY transcription factor family in two Citrus species with different Candidatus Liberibacter asiaticus susceptibility. BMC Plant Biology. 2023; 23: 159. https://doi.org/10.1186/s12870-023-04156-4

[39] Yan X, Huang W, Liu C, Hao X, Gao C, Deng M, et al. Genome-Wide Identification and Expression Analysis of the *AP2/ERF* Transcription Factor Gene Family in Hybrid Tea Rose Under Drought Stress. International Journal of Molecular Sciences. 2024; 25: 12849. https://doi.org/10.3390/ijms252312849

[40] Yang Y, Li Y, Guo Z, Zhao Y, Zhou X, Han Y, et al. Identification of DREB gene family in foxtail millet (*Setaria italica*) and analysis of its expression pattern in response to abiotic stress. Frontiers in Plant Science. 2025; 16: 1552120. https://doi.org/10.3389/fpls.2025.1552120

[41] Yuan Y, Wang L, Zhao Q, Liu C, Fu X, Li X, et al. Genome-Wide Identification, Expression Analysis and Functional Study of *DELLA* Genes in Chinese Cabbage (*Brassica rapa* L. ssp. *pekinensis*). Frontiers in Bioscience (Landmark Edition). 2024; 29: 198. https://doi.org/10.31083/j.fbl2905198

[42] Hamid R, Panahi B, Ghorbanzadeh Z, Jacob F, Zeinalabedini M, Ghaffari MR. Genome-wide identification and characterization of DUF789 genes in cotton: implications for fibre development. BMC Plant Biology. 2025; 25: 1192. https://doi.org/10.1186/s12870-025-07258-3

[43] Zhang D, Zhao C, Liu X, Wang H, Zhu B, Zhao G, et al. Genome-Wide Analysis of the AT-Hook Gene Family in *Malus sieversii* and Functional Characterization of MsAHL13. Plants (Basel, Switzerland). 2025; 14: 2625. https://doi.org/10.3390/plants14172625

[44] Wu L, Fan J, Su X, Peng D, Xing S. Genome-Wide Identification of R2R3-MYB Family Genes and Their Response to Stress in *Dendrobium nobile*. Frontiers in Bioscience (Landmark Edition). 2024; 29: 1. https://doi.org/10.31083/j.fbl2901001

[45] Ali S, Wang W, Zhang Z, Xie L, Boer DR, Khan N. Genome-Wide Identification, Expression and Interaction Analysis of ARF and AUX/IAA Gene Family in Soybean. Frontiers in Bioscience (Landmark Edition). 2022; 27: 251. https://doi.org/10.31083/j.fbl2708251

[46] Ma W, Liu X, Chen K, Yu X, Ji D. Genome-Wide Re-Identification and Analysis of CrRLK1Ls in Tomato. International Journal of Molecular Sciences. 2023; 24: 3142. https://doi.org/10.3390/ijms24043142

[47] Yang Y, Sun F, Wang P, Yusuyin M, Kuerban W, Lai C, et al. Genome-Wide Identification and Preliminary Functional Analysis of *BAM* (β-Amylase) Gene Family in Upland Cotton. Genes. 2023; 14: 2077. https://doi.org/10.3390/genes14112077

[48] Chen L, Wu X, Zhang M, Yang L, Ji Z, Chen R, et al. Genome-Wide Identification of *BrCMF* Genes in *Brassica rapa* and Their Expression Analysis under Abiotic Stresses. Plants (Basel, Switzerland). 2024; 13: 1118. https://doi.org/10.3390/plants13081118

[49] Gao L, Shi A, Wang M, Tian H, Zhang Y. Genome-Wide Identification and Expression Profiling of the GRAS Gene Family in *Taxus cuspidate*. Genes. 2025; 16: 1345. https://doi.org/10.3390/genes16111345

[50] Shi G, Zhang Z, Li J. Genome-Wide Identification of Basic Helix-Loop-Helix (*bHLH*) Family in Peanut: Potential Regulatory Roles in Iron Homeostasis. International Journal of Molecular Sciences. 2024; 25: 12057. https://doi.org/10.3390/ijms252212057

[51] Zhang H, Li Z, Li J, Ruan J, Yu H, Xu J, et al. Rice functional genomics and breeding database (RFGB)-3K-rice SNP and InDel sub-database. Chinese Science Bulletin. 2015; 60: 367–371. https://doi.org/10.1360/n972014-01231

[52] Zhang H, Gao S, Lercher MJ, Hu S, Chen WH. EvolView, an online tool for visualizing, annotating and managing phylogenetic trees. Nucleic Acids Research. 2012; 40: W569–W572. https://doi.org/10.1093/nar/gks576

[53] Paradis E. pegas: an R package for population genetics with an integrated-modular approach. Bioinformatics (Oxford, England). 2010; 26: 419–420. https://doi.org/10.1093/bioinformatics/btp696

[54] Li J, Fan M, Yang S, Huo H, Fan W, Lü S, et al. Genome-wide identification of castor bean class III peroxidase genes and analysis of expression patterns under abiotic stresses. BMC Plant Biology. 2025; 25: 900. https://doi.org/10.1186/s12870-025-06945-5

[55] Fukai C, Tanabata T, Nishizawa T, Koizumi M, Kutsuwada K, Kusano M. A developed system to extract specific responses of increment length in rice shoots under gradient changes in nitrogen concentration regimes. Plant Biotechnology (Tokyo, Japan). 2023; 40: 1–8. https://doi.org/10.5511/plantbiotechnology.22.1107a

[56] Parry MAJ, Andralojc PJ, Scales JC, Salvucci ME, Carmo-Silva AE, Alonso H, et al. Rubisco activity and regulation as targets for crop improvement. Journal of Experimental Botany. 2013; 64: 717–730. https://doi.org/10.1093/jxb/ers336

[57] Pérez-Tienda J, Corrêa A, Azcón-Aguilar C, Ferrol N. Transcriptional regulation of host NH4+ transporters and GS/GOGAT pathway in arbuscular mycorrhizal rice roots. Plant Physiology and Biochemistry: PPB. 2014; 75: 1–8. https://doi.org/10.1016/j.plaphy.2013.11.029