the plant journal

SEB SOCIETY FOR EXPERIMENTAL BIOLOGY

*The Plant Journal* (2025) **123,** e70406 doi: 10.1111/tpj.70406

# Identification and functional verification of key iron homeostasis genes in soybean roots and nodules through integrated transcriptome and proteome analysis

**Jie Li[1,2,†], Qihui Kong[1,†], Jiamei Zhu[1], Lin Li[1,2], Shoudong Wang[3], James Whelan[1,2] and Huixia Shou[1,2,*]**

[1]*State Key Laboratory of Plant Environmental Resilience, College of Life Sciences, Zhejiang University, Hangzhou, Zhejiang 310058, P.R. China,*

[2]*The Provincial International Science and Technology Cooperation Base on Engineering Biology, International Campus of Zhejiang University, Haining, Zhejiang 314400, China, and*

[3]*Key Laboratory of Soybean Molecular Design Breeding, Northeast Institute of Geography and Agroecology, Chinese Academy of Sciences, Changchun 130102, China*

Received 20 July 2022; revised 14 July 2025; accepted 26 July 2025.
*For correspondence (e-mail huixia@zju.edu.cn).
†These authors contributed equally to this work.

## SUMMARY

**Iron (Fe) is an essential nutrient for soybean [*Glycine max* (L.) Merr.] growth and symbiotic nitrogen fixation. However, the mechanisms underlying Fe homeostasis in nodules remain poorly understood. In this study, we conducted integrated transcriptome and proteome analyses of soybean roots and nodules under Fe deficiency to identify distinct Fe regulatory networks. Notably, nodules retained 42% of Fe levels under Fe-deficient conditions, despite severe depletion in roots (85% loss) and leaves (71% loss), suggesting a prioritized Fe allocation mechanism. Transcriptome and proteome sequencing of roots and nodules under Fe-sufficient and -deficient conditions revealed significant differences, confirming distinct expression profiles in nodules compared with roots. Among the differentially expressed genes, those encoding vacuolar Fe transporter-like protein 1a (*GmVTL1a*), yellow-strip like 7 (*GmYSL7*), and natural resistance-associated macrophage protein 3a (*GmNRAMP3a*) were highly expressed in nodules, emerging as key candidates. Transgenic soybeans expressing promoter:*GUS* fusion constructs for *GmVTL1a*, *GmYSL7*, and *GmNRAMP3a* confirmed their expression in nodules. Functional studies demonstrated that GmVTL1a mediates Fe transport across the symbiosome membrane, while GmYSL7 is critical for nodule development. Knockout of either gene impaired nitrogen fixation and ureide synthesis. Co-expression analysis of *GmVTL1a* and *GmVTL1b* identified 19 putative transcription factors potentially regulating GmVTL1a. An immunoprecipitation-mass spectrometry assay on nodule protein extracts from the *pGmVTL1a-3Flag-gGmVTL1a* plants yielded 55 candidate interactors, including 26 nodule-expressed proteins and 17 that overlapped with known symbiosome membrane proteins. Taken together, our study reveals nodule-specific adaptations in Fe homeostasis, highlighting GmVTL1a and GmYSL7 as central players.**

**Keywords: soybean, iron, transcriptome, proteome, nodule, symbiosome.**

## INTRODUCTION

Legumes form a symbiosis with nitrogen-fixing soil bacteria that convert atmospheric nitrogen into amides or ureides (Masson-Boivin et al., 2009). This symbiosis is a key component of the biological nitrogen cycle. The process of symbiotic nitrogen fixation (SNF) begins through intricate chemical and physical interactions between rhizobia and legume roots. Both nodulation and nitrogen fixation are tightly regulated processes that depend on substantial host-derived resources, including photosynthates, macronutrients (e.g., phosphate and sulfate), and essential metal micronutrients (Ke et al., 2022; Roy et al., 2020). Among these, metal ions are vital for the development of nodules and the nitrogen fixation process (González-Guerrero et al., 2014). Plants have evolved complex mechanisms to coordinate transporters for the absorption and distribution of these essential nutrients (Li, Liu, et al., 2024). Iron (Fe) is a crucial element for various ferroproteins, including leghemoglobin, nitrogenase, cytochromes, ferredoxin, and hydrogenase, which are essential

during SNF (Tang et al., 1992). Despite its abundance in soil, Fe is often a growth-limiting nutrient for plants due to its low solubility (Guerinot & Yi, 1994). Fe deficiency not only directly impacts the growth and development of the plant host but also significantly affects SNF (Brear et al., 2013). While the mechanisms of Fe acquisition by roots have been extensively studied, knowledge about Fe uptake in determinate nodules remains less explored. Recent work revealed that Fe status acts as a key signal for nodulation, with the Fe sensor BRUTUS A integrating Fe and nodulation cues (Ren et al., 2025).

Fe is likely imported into the nodules as Fe(III)-citrate via the xylem and then crosses several cell layers to reach the infected cells (Brear et al., 2013; Day & Smith, 2021; Rodriguez-Haas et al., 2013). Both apoplastic and symplastic pathways are utilized for this transport (Rodriguez-Haas et al., 2013). MtNRAMP1, a member of the natural resistance-associated macrophage protein (NRAMP) family in *Medicago truncatula*, was reported to transport Fe(II) from the apoplast into infected cells (Tejada-Jimenez et al., 2015). It acts in concert with multidrug and toxic compound extrusion protein MtMATE67 to facilitate Fe uptake by infected cells (Kryvoruchko et al., 2018). In contrast, symplastic Fe transport requires the reduction of Fe(III) to Fe(II) by membrane-localized ferric chelate reductase (FRO, Slatni et al., 2009). The resulting Fe(II) is then either imported into bacteroids via the FeoAB transporter in *B. japonicum* (Sankari & O'Brian, 2016), or acquired by infected host cells through the host divalent metal transporter DMT1, which is a member of the NRAMP family (Kaiser et al., 2003). A recent study showed that both DMT1 and GmNRAMP2b localize to the tonoplast of uninfected nodule cells and mediate Fe transport from non-infected cells to infected cells (Zhou et al., 2024). Alternatively, Fe(II) can be chelated to nicotianamine (NA) and then imported into cells. In *Medicago truncatula*, the genes involved in the NA-mediated Fe(II) transport include *MtNAS2* for NA biosynthesis, *MtYSL3* for NA-Fe(II) transport, and *MtFPN2* encoding Ferroportin2, a putative Fe efflux transporter (Castro-Rodríguez et al., 2020; Escudero, Abreu, del Sastre, et al., 2020; Escudero, Abreu, Tejada-Jimenez, et al., 2020).

*Saccharomyces cerevisiae* CCC1 is the first identified member of the vacuolar iron transporter (VIT) family. It mediates Fe(II) transport from the cytosol into the vacuole (Li et al., 2001). In plants, Arabidopsis VITs, VIT-like (VTL), and rice orthologs *OsVIT1/OsVIT2* have been shown to catalyze vacuolar Fe transport (Gollhofer et al., 2014; Kim et al., 2006; Zhang et al., 2012). In infected cells of legume nodules, vacuoles are replaced by symbiosomes (Gavrin et al., 2014). Given the functional similarity between these compartments, VIT family transporters were predicted to facilitate Fe delivery into symbiosomes. Indeed, in *Medicago truncatula*, VTL4 and VTL8 mediate cytosolic Fe transport to the symbiosome, ensuring Fe supply for nitrogen fixation (Walton et al., 2020). In *Lotus japonicus*, the VIT family gene *LjSEN1* is specifically expressed in infected nodule cells and is essential for nitrogen fixation activity and symbiosome/bacteroid differentiation in legume nodules (Hakoyama et al., 2012; Suganuma et al., 2003). The soybean GmVTL1a, a homolog of LjSEN1, functions as a symbiosome membrane (SM) Fe transporter. It mediates Fe delivery from the infected root cell cytosol to the symbiosome, playing a crucial role in the SNF process in soybean (Liu et al., 2020). GmVTL1a can restore nitrogen fixation when expressed in the *Ljsen1* mutant (Brear et al., 2020). There are two VTL paralogs in the soybean genome, GmVTL1a and GmVTL1b. Complementation with *GmVTL1a* fully restores SNF and nitrogenase activity in the *vtl1-1* mutant (*vtl1a/vtl1b* double mutant), demonstrating that soybean relies predominantly on GmVTL1a for Fe delivery to bacteroids (Liu et al., 2020).

Multi-omics studies have been widely used across various plant species to investigate the specific mechanisms and potential genes behind regulatory phenomena (Lardi & Pessi, 2018; Li et al., 2021; Mahmood et al., 2022; McLoughlin et al., 2018). In this study, we analyzed the transcriptome and proteome profiles of soybean roots and nodules under Fe-sufficient and Fe-deficient conditions. Comparative profiling revealed tissue-specific expression patterns of key Fe-related genes, including YSL, FRO, NRAMP, VTL, and FER. Notably, we identified nodule-enriched transporters GmVTL1a, GmYSL7, and GmNRAMP3a as critical for Fe allocation or nodule development. Furthermore, we uncovered a putative regulatory network comprising 19 transcription factors and 26 interacting proteins associated with GmVTL1a, providing mechanistic insight into nodule-specific Fe adaptation. This work establishes a foundation for targeting these genes to optimize Fe homeostasis and enhance SNF in soybean.

## RESULTS

### Effect of Fe deficiency on the growth of soybean plants and nodule formation

Five-day-old soybean seedlings, initially germinated in vermiculite, were transferred to hydroponic systems and grown under either Fe-sufficient (+Fe) or -deficient (−Fe) conditions. After 30 days of growth under Fe deficiency, soybean seedlings showed leaf chlorosis and a significant decrease in the number of nodules (Figure 1a,b). Plant height was significantly reduced by 30%; shoot fresh weight by 37%; root fresh weight by 23%; nodule fresh weight by 50%; and nodule number by 42% (Figure 1c). The Fe content in leaves and roots of −Fe treated plants decreased by over 71%. In contrast, nodule Fe content decreased by 42%. The relative expression levels of *GmIRT1* (iron-regulated transporter) and *GmFRO2* (ferric

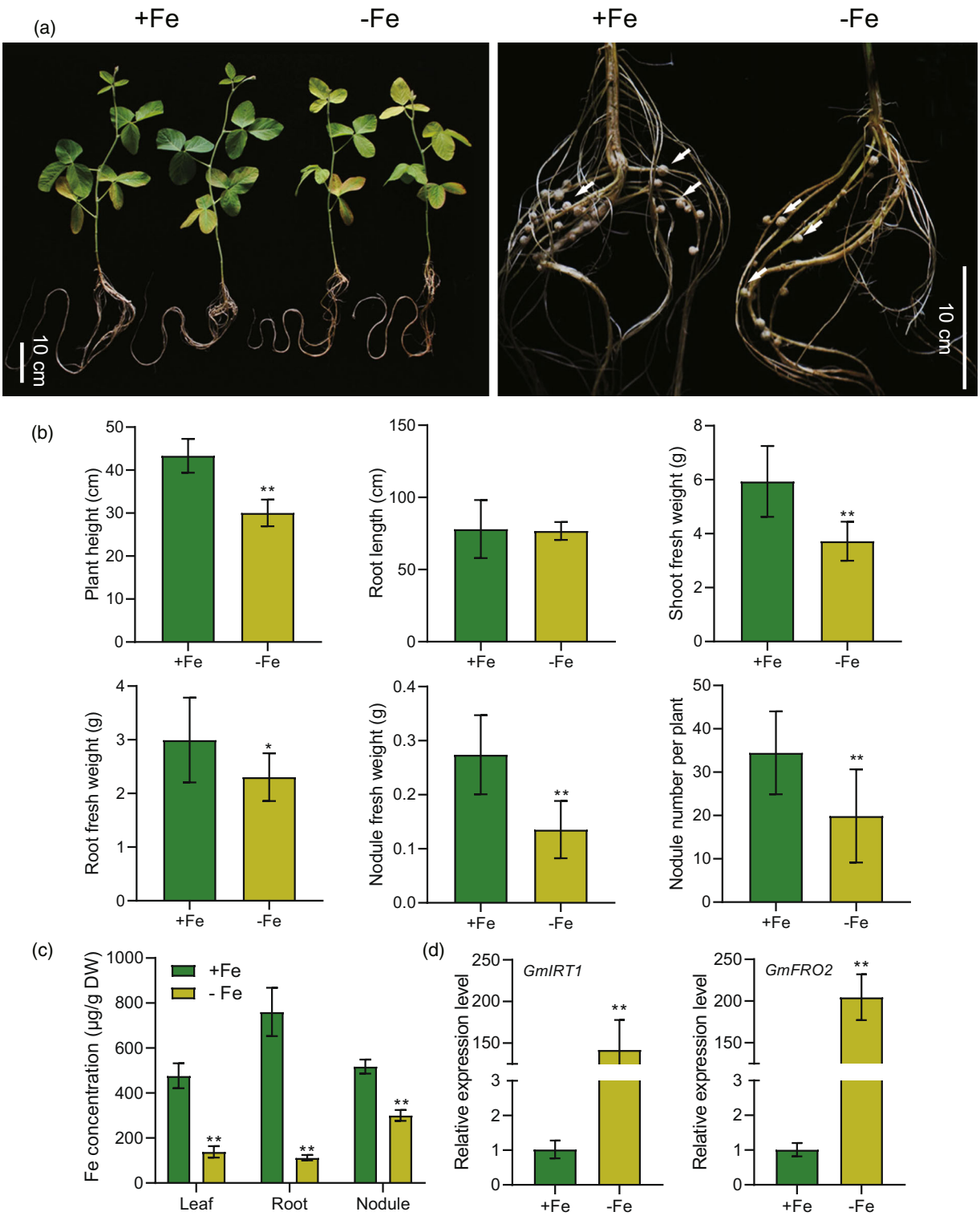


**Figure 1.** Effect of Fe supply on soybean shoot, root, and nodule development.
(a) Phenotype of wild-type (WT) soybean under Fe-sufficient (+Fe, 100 μM) and −deficient (−Fe, 0 μM) conditions. Seedlings were grown on nutrient solution for 30 days. Scale bar, 10 cm.
(b) Quantification of plant height, root length, shoot fresh weight, root fresh weight, nodule fresh weight, and nodule number.
(c) Fe concentrations (ICP-MS) in leaves, roots, and nodules.
(d) Transcript levels of Fe-responsive genes *GmIRT1* and *GmFRO2* in roots (normalized to *GmTefs1*). Significance of differences is indicated by asterisks (LSD's ANOVA test; $*P < 0.05$, $**P < 0.01$).

reduction oxidase), both of which are responsive to Fe deficiency, were significantly increased in the roots under the −Fe condition, by 140- and 200-fold, respectively (Figure 1d). The reduced Fe contents and increased expression of Fe-deficiency marker genes confirmed the efficacy of the Fe deficiency treatment.

### Changes of transcriptome and proteome in Fe deficiency

RNA-sequencing (RNA-seq) and tandem mass tag (TMT) quantitative proteomic analysis were conducted on whole roots or nodules without dissection from soybean plants grown under both +Fe and −Fe conditions. RNA-seq analysis generated 532.4 million reads, totaling 73 Gb of raw data for roots and nodules, with an average mapping rate of 95% to the Williams82.v2 reference genome (Tables S1 and S2). The quality of the transcriptome data was validated through multiple orthogonal assessments: (i) fragment per kilobase million (FPKM) distributions confirmed consistent expression profiles across replicates (Figure S1a); (ii) principal component analysis (PCA) demonstrated clear separation by experimental conditions in PCA1 accounting for 77.07% variance (Figure S1b); and (iii) high inter-replicate Pearson correlation coefficients (>0.95) supported robust reproducibility (Figure S1c). These comprehensive quality metrics (Figure S1; Tables S1 and S2) confirm the suitability of the dataset for downstream analyses. A total of 37 040 differentially expressed genes (DEGs) were identified across comparison groups, with fold change ≥2 or ≤0.25 and *P*-adj ≤0.05.

TMT-based quantitative proteomic analysis of soybean roots and nodules identified 8813 proteins from 26,858 unique peptides, corresponding to an average of three unique peptides per protein (Table S3). Protein identification confidence was robust, with 79.05% of peptides exceeding the significance threshold (IonScore >20) and a median Ion Score of 32.28 (Figure S2a). The peptide length, molecular weight, peptide count, protein sequence coverage, and isoelectric point distributions confirmed both the reliability of protein identifications and key physicochemical properties of the detected peptides and proteins (Figure S2b–f). These metrics collectively support that the quality of the proteomics data is reliable. For protein quantification, relative abundances were calculated as the ratio of each protein's signal in the sample group to its corresponding internal reference signal (Table S3). A total of 7033 differentially expressed proteins (DEPs) (fold change ≥1.2 or ≤0.83, and $P \leq 0.05$) were identified by TMT labeling. Although the proteomic analysis was conducted on total protein extracts without specific enrichment of membrane proteins, several transporter proteins, including members of the VIT family, were identified (Tables S4 and S5).

Venn and Upset plot diagrams were used to visualize DEGs and proteins (DEPs) across the different comparisons: 17 309 DEGs and 3142 DEPs in nodules (N) versus roots (R) under +Fe condition (N+Fe/R+Fe, Table S4); 17 807 DEGs and 3029 DEPs in nodules versus roots under −Fe condition (N−Fe/R−Fe, Table S5); 451 DEGs and 81 DEPs in −Fe versus +Fe in nodules (N−Fe/N+Fe, Table S6); and 1473 DEGs and 781 DEPs −Fe versus +Fe in roots (R−Fe/R+Fe, Table S7) were detected (Figure 2).

To explore more information, gene ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analysis were performed in these DEGs and DEPs. GO functions are mainly enriched in multiple processes related to Fe ion transport, binding, and homeostasis (Figure S3; Tables S8 and S9). KEGG enriched the main functions of the DEGs and DEPs in different comparisons into starch and sucrose metabolism, plant–pathogen interaction, nitrogen metabolism, hormone signal transduction, and fatty acid biosynthesis (Figure S4; Tables S10 and S11).

### Fe-related DEGs and DEPs that are highly expressed in roots or nodules

Proper development and function of nodules require an adequate supply of Fe. Thus, under Fe-deficient conditions, the expression of genes associated with Fe homeostasis is likely to be altered. To identify genes that are significantly expressed in soybean nodules under either Fe-sufficient or -deficient conditions, the DEGs in these N versus R comparison groups were screened and selected based on the criteria of $|\log_2\text{Fold Change (FC)}| \geq 1$ and $P\text{-adj} \leq 0.05$. The DEPs in the N versus R comparison groups were screened and selected based on the criteria of $|\log_2\text{FC}| \geq 0.26$ and $P \leq 0.05$. A total of 14 329 DEGs (1st + 4th + 7th + 9th bars in Figure 2a,b; Table S12) and 2269 DEPs (1st + 4th + 8th + 12th bars in Figure 2c,d; Table S13) were identified in the N versus R comparison in both Fe-sufficient and -deficient conditions. Among these, 5663 genes and 1114 proteins were expressed in nodules, while 8650 genes and 1152 proteins were expressed in roots (Figure S5a,b; Tables S12 and S13). The resulting DEGs and DEPs (nodules versus roots) were further screened by GO enrichment analysis for categories related to Fe ion homeostasis (GO:0055072), Fe ion transport (GO:0006826), Fe ion binding (GO:0005506), Fe ion transmembrane transporter activity (GO:0005381), and so on (Figure 3a,b; Tables S8 and S9). To prioritize biologically meaningful changes, we applied stringent thresholds for identifying DEGs and proteins (DEPs) in the root nodule comparison. DEGs were required to exhibit $|\log_2\text{FC}| \geq 3$ (corresponding to an eightfold change), while DEPs were selected at $|\log_2\text{FC}| \geq 0.5$ (1.4-fold change). These genes and proteins were highlighted in green and orange box for root and nodule-expressed genes, respectively (Figure 3c, d). Of these, three vacuolar Fe transporter (VIT) genes,

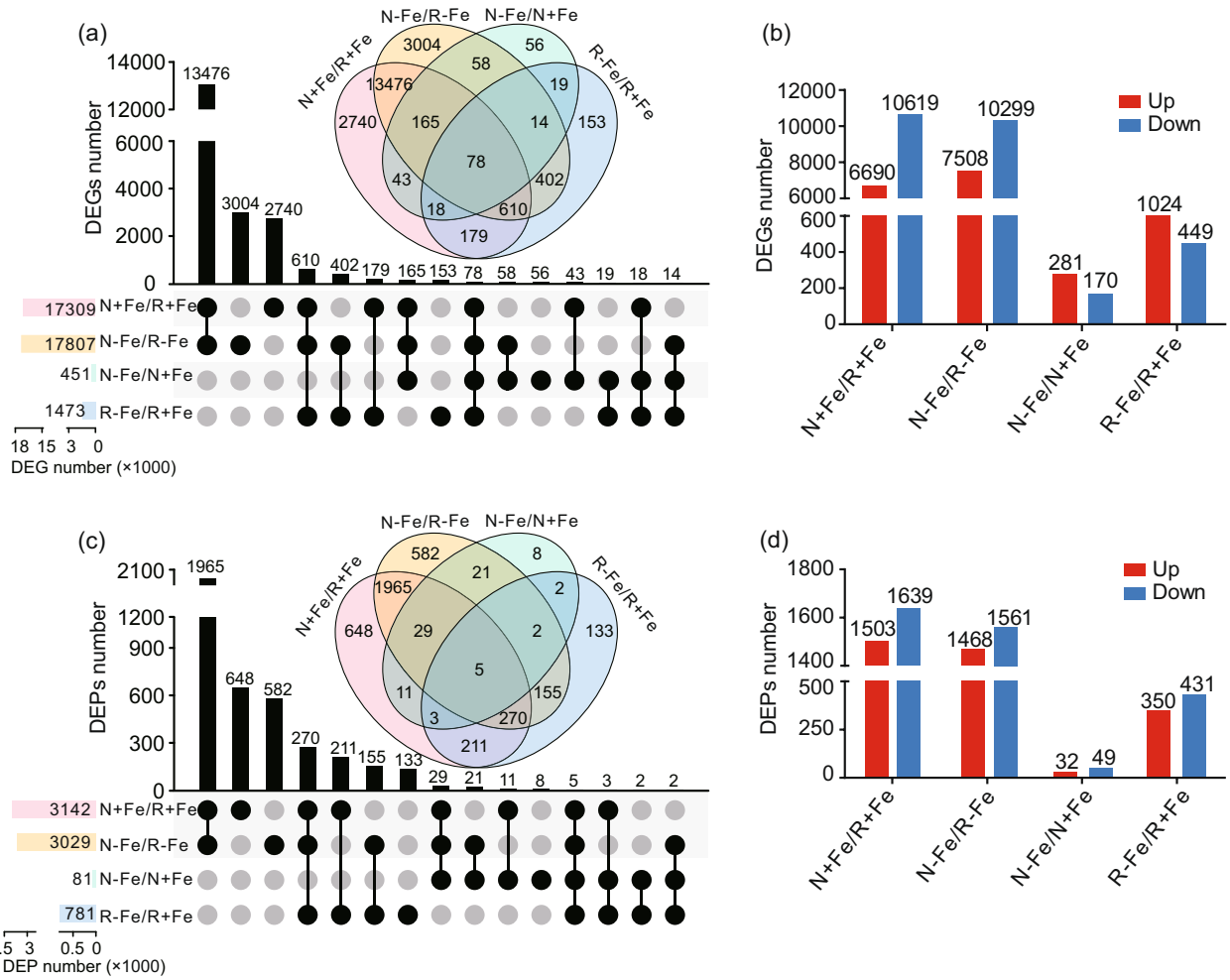


**Figure 2.** Distinct transcriptomic and proteomic responses (DEGs and DEPs) to Fe deficiency in roots and nodules.
(a, b) Venn diagrams and histograms of DEGs ($|\log_2 FC| \geq 1$, $P \leq 0.05$).
(c, d) Venn diagrams and histograms of DEPs (≥1.2-fold or ≤0.8-fold, $P \leq 0.05$). The comparisons were named as N+Fe/R+Fe (nodule versus root, +Fe), N−Fe/R−Fe (nodule versus root, −Fe), N−Fe/N+Fe (−Fe versus +Fe, nodule), R−Fe/R+Fe (−Fe versus +Fe, root). All experimental groups were repeated in triplicate.

*02G082500*, *08G076000*, and *08G181900*, were specifically expressed in roots, while three other VIT-related genes, *05G121500*, *05G121600*, and *08G076300*, were significantly expressed in nodules (Figure 3c). The protein encoded by *05G121600* was also identified in nodules by proteomic data (Figure 3d), which is consistent with the DEGs results. The genes encoding Ferretin 1 (*FER*, *01G124500* and *03G050100*) and Ferroportin (*FPN2*, *03G042400*) were highly expressed in roots. Notably, the transcript abundances of Fe-regulated 1 (*IREG1*, *01G128300*) and ferric reduction oxidase 2 (*FRO2*, *16G033800*) known to respond to Fe deficiency, were 38- and 1580-fold higher in nodules than in roots (Peiffer et al., 2012). Similar preferential expression in nodules was also found in Fe-deficient conditions, in which their expression in nodules was 258- and 827-fold higher than in roots (Figure 3c; Table S12). The proteins encoding Ferretin 4 (14G056800) and Ferredoxin 3 (11G183300) were highly expressed in nodules (Figure 3d). Additionally, numerous cytochrome P450 family genes were differentially expressed between roots and nodules.

### DEGs and DEPs in response to Fe deficiency in roots or nodules

Roots exhibited a stronger response to Fe deficiency than nodules at both the transcriptomic and proteomic levels, as evidenced by comparing N−Fe versus N+Fe and R−Fe versus R+Fe. A total of 129 Fe-responsive DEGs and 12 DEPs were associated with both nodules and roots, while a large number of Fe-responsive DEGs and DEPs were specific to roots or nodules (Figure 4a,b; Tables S14 and S15), indicating significant differences in root and nodule responses to Fe deficiency. Among the co-regulated DEGs and DEPs, 75 DEGs and 4 DEPs were co-upregulated by Fe deficiency, and 53 genes and 6 proteins were co-downregulated, indicating that these genes or proteins responded consistently to Fe deficiency in roots and

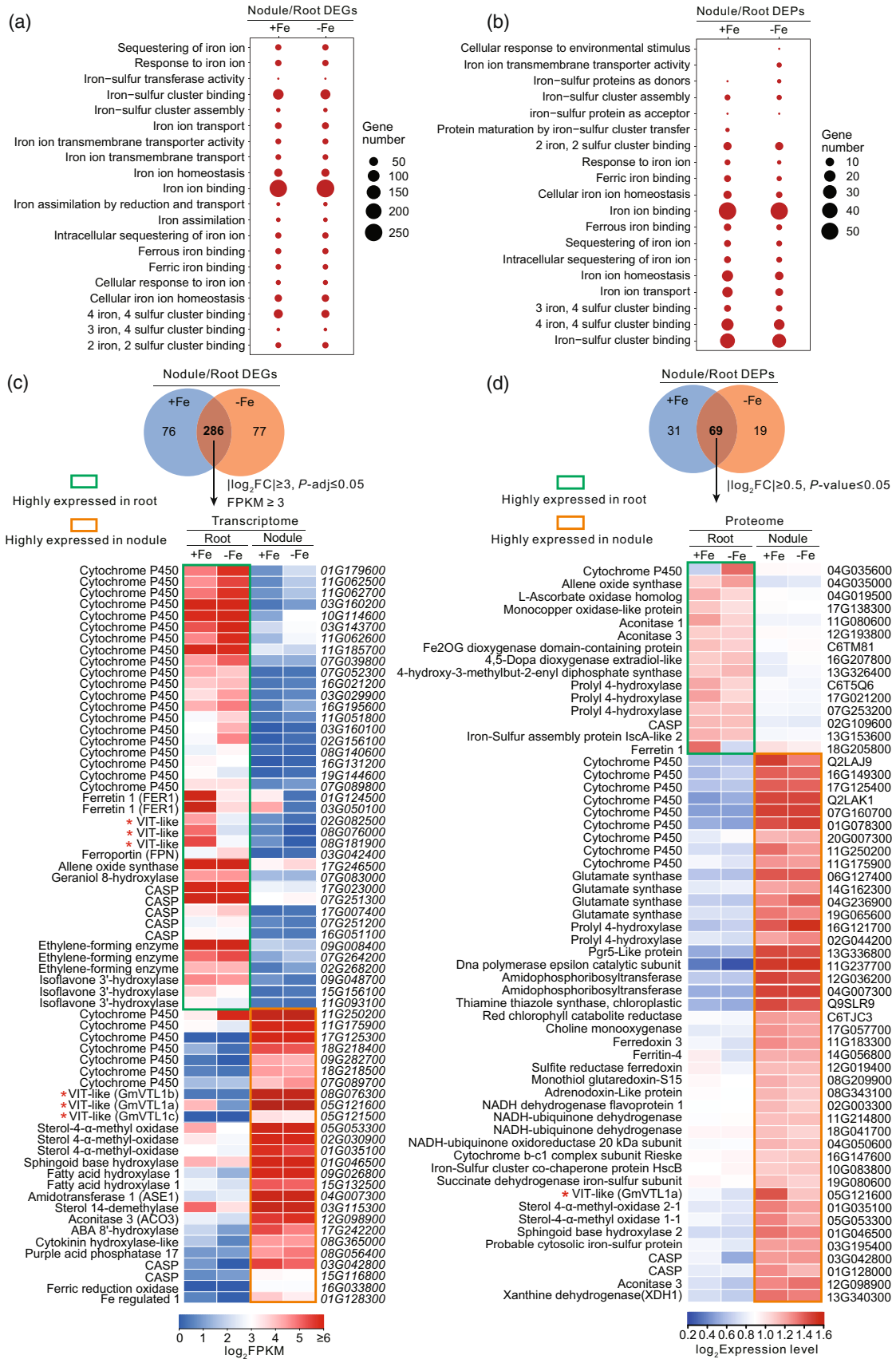

**Figure 3.** Nodule-enriched Fe homeostasis genes and proteins.
(a, b) GO terms for Fe-related DEGs and DEPs in nodules versus roots.
(c) Heatmap of nodule versus root DEGs ($|\log_2 FC| \geq 3$, FPKM $\geq 3$, $P \leq 0.05$).
(d) Heatmap of nodule versus root DEPs ($|\log_2 FC| \geq 0.5$, $P \leq 0.05$). The red asterisk indicates the VIT-like protein. Green and orange boxes indicate genes or proteins enriched in roots and nodules, respectively.

nodules (Tables S6c,d; Figures S14 and S15). However, one gene encoding alanine-glyoxylate aminotransferase (*18G210500*) showed an antagonistic tendency. The expression of *18G210500* in nodules was induced by Fe deficiency, while that in roots was suppressed (Figure S5c; Table S14). Representative genes or proteins that responded to Fe deficiency in both roots and nodules are shown in Figure 4(c,d). In the Fe homeostasis category, four *GmFER* genes were significantly inhibited by Fe deficiency in roots and nodules (Figure 4c), which is consistent with the fact that *GmFER* gene expression reflects tissue Fe content (Lönnerdal, 2009). In addition, two *GmVTL* genes, *08G181900* and *16G168200*, were also suppressed by Fe deficiency. Fe deficiency in roots and nodules induced the expression of many Fe homeostasis genes, including the Fe(II) transporter *NRAMP2b* (*17G165200*, also known as *GmDMT1*) and its homolog *NRAMP2a* (*05G101700*) (Figure 4d) (Kaiser et al., 2003; Zhou et al., 2024).

### Genes and proteins highly expressed in nodules and significantly responding to Fe deficiency

To identify nodule-specific responses to Fe deficiency, we compared the DEGs and DEPs between Fe-sufficient and -deficient conditions. A total of 322 DEGs and 69 DEPs expressed in nodules significantly responded to Fe deficiency (Tables S16 and S17). Among them, *GmFRO2* (*18G235900*) exhibited the highest level of induction under Fe deficiency, up to 60-fold (Figure 4e; Table S16). *MtYSL3* has been reported to be involved in Fe delivery to the SNF complex (Castro-Rodríguez et al., 2020). In this study, *GmYSL3* (*17G191300*), the homolog of *MtYSL3*, was significantly inhibited by Fe deficiency. However, the expression level of this particular gene is relatively low in nodules. For the proteome analysis, DEPs in the N–Fe versus N+Fe comparison were screened and selected based on the criteria of $|\log_2 FC| \geq 0.4$ and $P \leq 0.05$. A total of 30 DEPs were identified in nodules. The GmVTL family protein (*05G121600*) and two Fe superoxide dismutase proteins (*20G196900* and *10G193500*) were suppressed by Fe deficiency in nodules (Figure 4f; Table S17).

### Transcriptional changes of genes involved in Fe homeostasis

A variety of gene families are known to be involved in Fe homeostasis, including FER, FRO, NRAMP, YSL, and VTL. In this study, we systematically identified members of these gene families in the soybean genome and characterized their expression profiles in roots and nodules. Ferritin is a class of Fe storage proteins, whose expression is positively regulated by cellular Fe status. Among the 12 *FER* genes in the soybean genome (Figure S6a), six *FER* genes were highly expressed in roots under Fe-sufficient conditions; only one was detected in nodules. This is likely reflecting the high Fe demand in nodules that reduces Fe storage (Figure 5a). The proteins corresponding to the *FER* genes, *03G050100* and *02G262500*, were identified in proteomic data with protein IDs I1JL80 and I1JIE1, respectively (Figure 5a). Both FER proteins were significantly suppressed by Fe deficiency in roots but not in nodules (Table S18).

Using the conserved domains of ferric reductase-like transmembrane, FAD-binding, and NAD-binding domains as features of FRO proteins, a blastp search identified 25 *FRO*-like genes (Figure S6b). While most *GmFRO* genes showed similar expression under Fe-sufficient and -deficient conditions, *07G067700* was specifically expressed in Fe-deficient roots (Figure 5b; Table S18). Proteomic analysis showed that two FROs, 10G152200 and 06G162300, were induced in roots and suppressed in nodules by Fe deficiency (Table S18).

This study further investigated the *NRAMP*, *YSL*, and *VTL* gene families known to transport Fe or other metals. *MtNRAMP1* is known to transport Fe from the apoplast to rhizobia-containing cells (Tejada-Jimenez et al., 2015). Thirteen *NRAMP* genes were identified in the soybean genome (Figure S6c). Among these, *GmNRAMP3a* (*06G044200*) and *GmNRAMP3b* (*04G044000*) were expressed at higher levels in nodules than in roots, and their expression was not affected by Fe supply conditions (Figure 5c; Table S18). *GmNRAMP2b* (*17G165200*), *GmNRAMP2a* (*05G101700*) and *GmNRAMP7* (*06G115800*) were induced by Fe deficiency in both roots and nodules, while *GmNRAMP5b* (*08G218200*) was induced only in roots (Figure 5c; Table S18). Proteomic data only detected the GmNRAMP7 protein, I1KAC9, which was significantly induced by Fe deficiency in roots (Table S18). A recent study demonstrated that GmYSL7, an oligopeptide transporter previously characterized by Gavrin et al. (2021), plays a key role in regulating Fe translocation from roots to nodules and its intra-nodular distribution in soybean. This process directly impacts nodule development and nitrogenase activity (Wu et al., 2023). Similarly, MtYSL3 was shown to mediate the vascular transport of both Fe and zinc essential for SNF

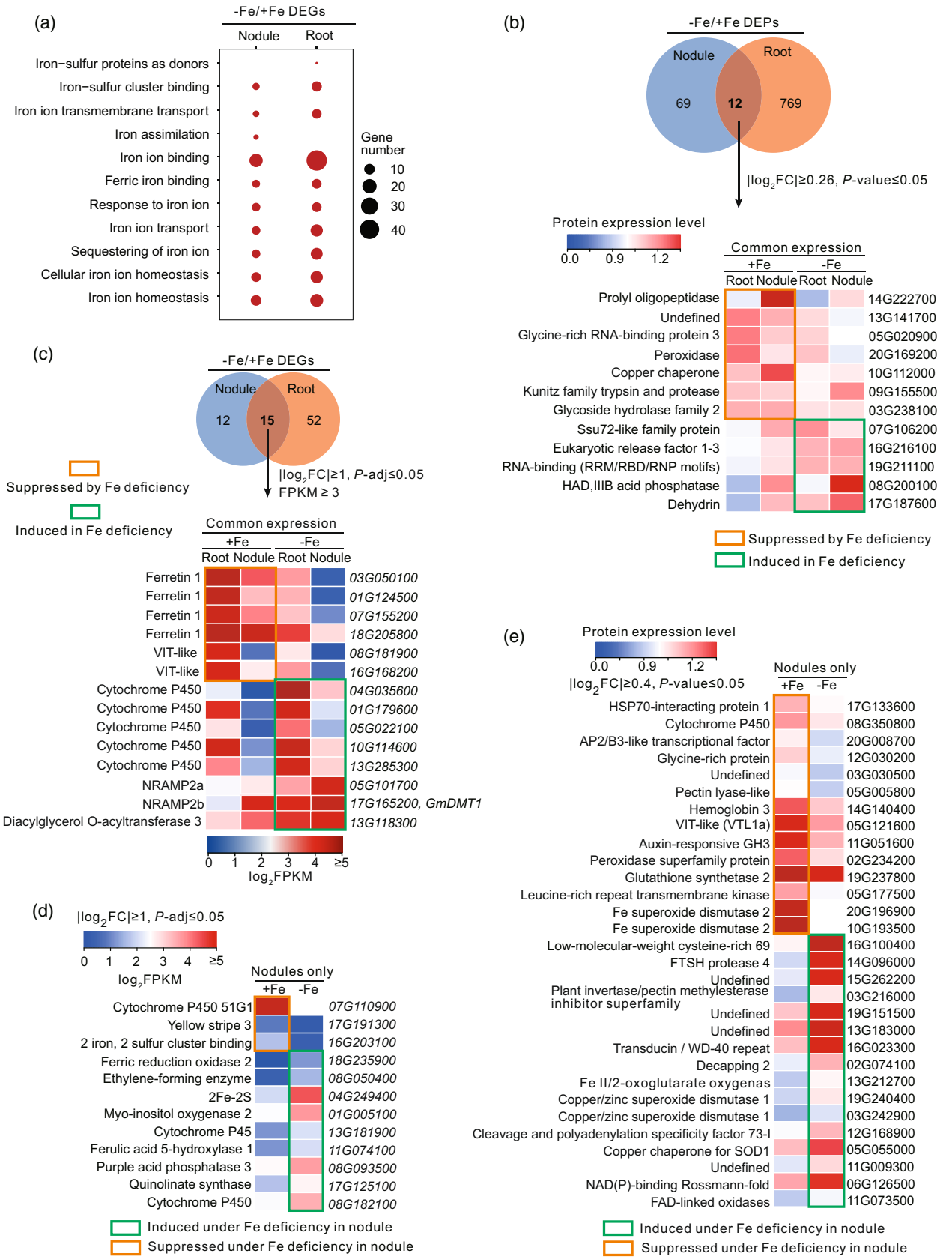

**Figure 4.** Tissue-specific Fe homeostasis related genes at mRNA and protein levels.
(a) GO terms of Fe-related DEGs.
(b) Venn diagrams of DEPs significantly induced or suppressed in Fe deficiency ($|\log_2 FC| \geq 0.26$, $P \leq 0.05$).
(c) DEGs in Fe homeostasis significantly induced or suppressed by Fe deficiency ($|\log_2 FC| \geq 1$, $P \leq 0.05$, FPKM $\geq 3$).
(d) Nodule-specific DEGs under Fe deficiency ($|\log_2 FC| \geq 1$, $P \leq 0.05$).
(e) Nodule-specific DEPs under Fe deficiency ($|\log_2 FC| \geq 0.4$, $P \leq 0.05$). Green and orange boxes indicate genes or proteins that were induced and suppressed by Fe deficiency, respectively.

(Castro-Rodríguez et al., 2020). These findings collectively underscore the importance of YSL family transporters in maintaining metal nutrient balance during rhizobial symbiosis. A total of 15 *YSL* genes encoding OPT domain-containing transporters were identified (Figure S6d). *YSL* genes *16G054200*, *19G094800*, and *16G212900* showed similar expression levels in roots and nodules, whereas *06G133000*, *GmYSL7* (*11G203400*), and *04G231900* were predominantly expressed in nodules (Figure 5d; Table S18). The I1MQH0 (16G212900) protein was detected in the proteomic data in all samples, indicating high expression in both roots and nodules regardless of Fe supply conditions (Table S18). A total of 20 *GmVTL* genes encoding VIT domain-containing proteins were identified in the soybean genome (Figure S6e). Eight *GmVTL* genes showed low expression levels in both roots and nodule tissues (Figure 5e). Three *GmVTLs* (*GmVTL1a/b/c*) were expressed with a nodule to root FC values greater than 5, and were defined as nodule-preferential expression genes (Figure 5e; Figure S7a; Table S18). Among them, *GmVTL1a* had the highest expression level, with expression 2.62 or 2.37-fold higher than that of *GmVTL1b* under Fe-sufficient or deficient conditions, respectively, while *GmVTL1c* showed substantially lower expression than *GmVTL1a* and *GmVTL1b* (Figure 5e; Figure S7A). Quantitative RT-PCR analysis confirmed that *GmVTL1a* and *GmVTL1b* are preferentially expressed in nodules, with *GmVTL1a* being expressed at a significantly higher level than *GmVTL1b* (Figure S8). The results are consistent with single-nucleus and spatial transcriptome data (Figure S9a). The published single-nucleus and spatial transcriptome data also showed that both *GmVTL1a* and *GmVTL1b* were mainly expressed in the infected zone, with *GmVTL1a* exhibiting higher expression levels than *GmVTL1b* (Figure S9).

Overall, the study found that several Fe transport genes in the *NRAMP*, *YSL*, and *VTL* families were preferentially expressed in nodules, implying their roles in soybean nodule development. These nodule-preferential genes include *GmNRAMP3a/3b*, *GmYSL7*, and *GmVTL1a/b*, which were further validated.

### Verification of the nodule-preferential expression of *GmVTL1a*, *GmYSL7*, and *GmNRAMP3a*, and their roles in nodules

To further clarify the expression patterns of these nodule-preferential expression genes, transgenic soybean plants expressing constructs of the *GmVTL1a*, *GmYSL7*, or *GmNRAMP3a* promoters fused to a *GUS* ($\beta$-glucuronidase) reporter gene were generated. Transgenic seedlings were grown under nitrogen-deprivation conditions and inoculated with *Bradyrhizobium japonicum* strain USDA110 to investigate the expression of these genes in nodules. Tissues of the corresponding transgenic plants after GUS staining were imaged by stereomicroscope. While *GmVTL1a* and *GmYSL7* were specifically expressed in nodules, *GmNRAMP3a* was expressed in multiple tissues, including stems, leaves, flowers, pods, and nodules (Figure 6a–c). This implies that *GmVTL1a* and *GmYSL7* play roles primarily in nodules, while Gm*NRAMP3a* functions not only in nodules but also in other tissues. When infected by rhizobium, the expression of all three genes was significantly induced at 21–42 days post-inoculation (dpi) (Figure 6d).

To determine the biological functions of the two nodule-preferential expression genes, *GmVTL1a* and *GmYSL7*, two CRISPR/Cas9-mediated knockout mutant lines for each gene were obtained using the CRISPR/Cas9 gene editing system, named *vtl1a-1*, *vtl1a-2*, *ysl7-1*, and *ysl7-2* respectively (Figure S10). No significant differences in growth phenotypes were observed between the wild-type (WT) and *vtl1a* or *ysl7* mutants when grown under high N conditions, where SNF is largely blocked. As both genes are exclusively expressed in nodules, we further investigated their effect on symbiotic interactions. After 28 dpi, the *vtl1a* and *ysl7* mutants exhibited shorter plant height, lighter shoot and root fresh weight than those of the WT (Figure 7a,e). In addition, the nodule number and fresh weight of the *vtl1a* and *ysl7* mutant lines were significantly lower than the WT. The number of nodules decreased by 31.6 and 24.7%, and the fresh weight decreased by 19.4 and 33.4% in both *vtl1a* lines, respectively (Figure 7a,b). Similar to *GmVTL1a*, the nodule number and weight, plant height, shoot and root weight of the two *ysl7* mutant lines were also significantly lower than in WT (Figure 7e,f). The fixed nitrogen in the nodules of leguminous plants is converted into ureides, which serve as a stored form of nitrogen for transport to other parts. Compared with the WT, *vtl1a-1* and *vtl1a-2* showed a significant reduction in nitrogenase activity compared with the WT (Figure 7c). Similarly, there was a significant decrease in nfiH expression in *ysl7-1 and ysl7-2* compared with the WT (Figure 7g). To investigate the effects of GmVTL1a and

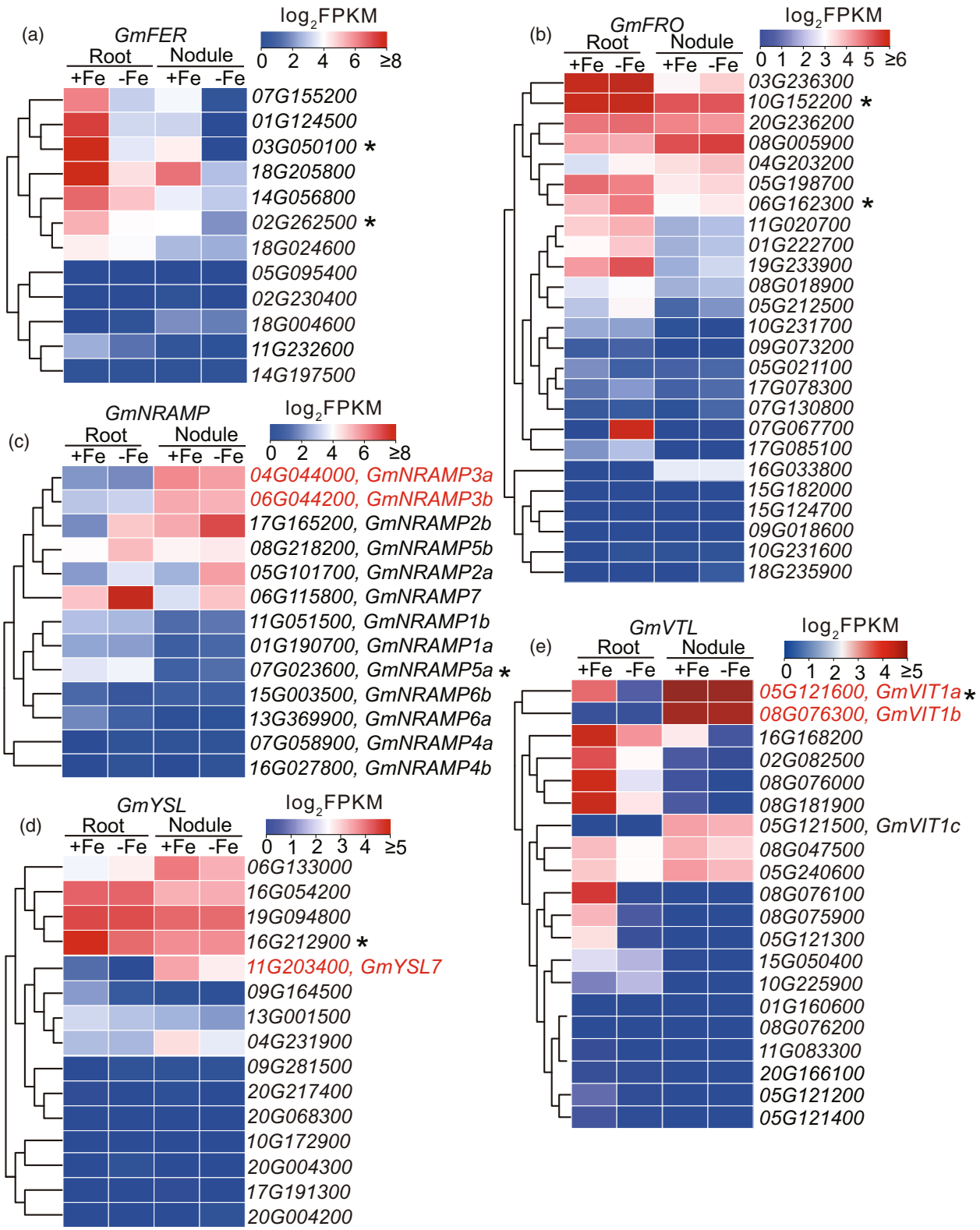


**Figure 5.** Transcriptional changes of Fe homeostasis genes in roots and nodules under different Fe supply conditions.
Hierarchical clustering heatmap analysis of soybean ferritin (*FER*) (a), ferric reduction oxidase (*FRO*) (b), natural resistance-associated macrophage protein (*NRAMP*) (c), yellow stripe-like (*YSL*) (d) and vacuolar iron transporter-like proteins (*VTL*) (e). The value represents $log_2$FPKM. Asterisks indicate that the proteins of the corresponding genes were detected in proteomic data.

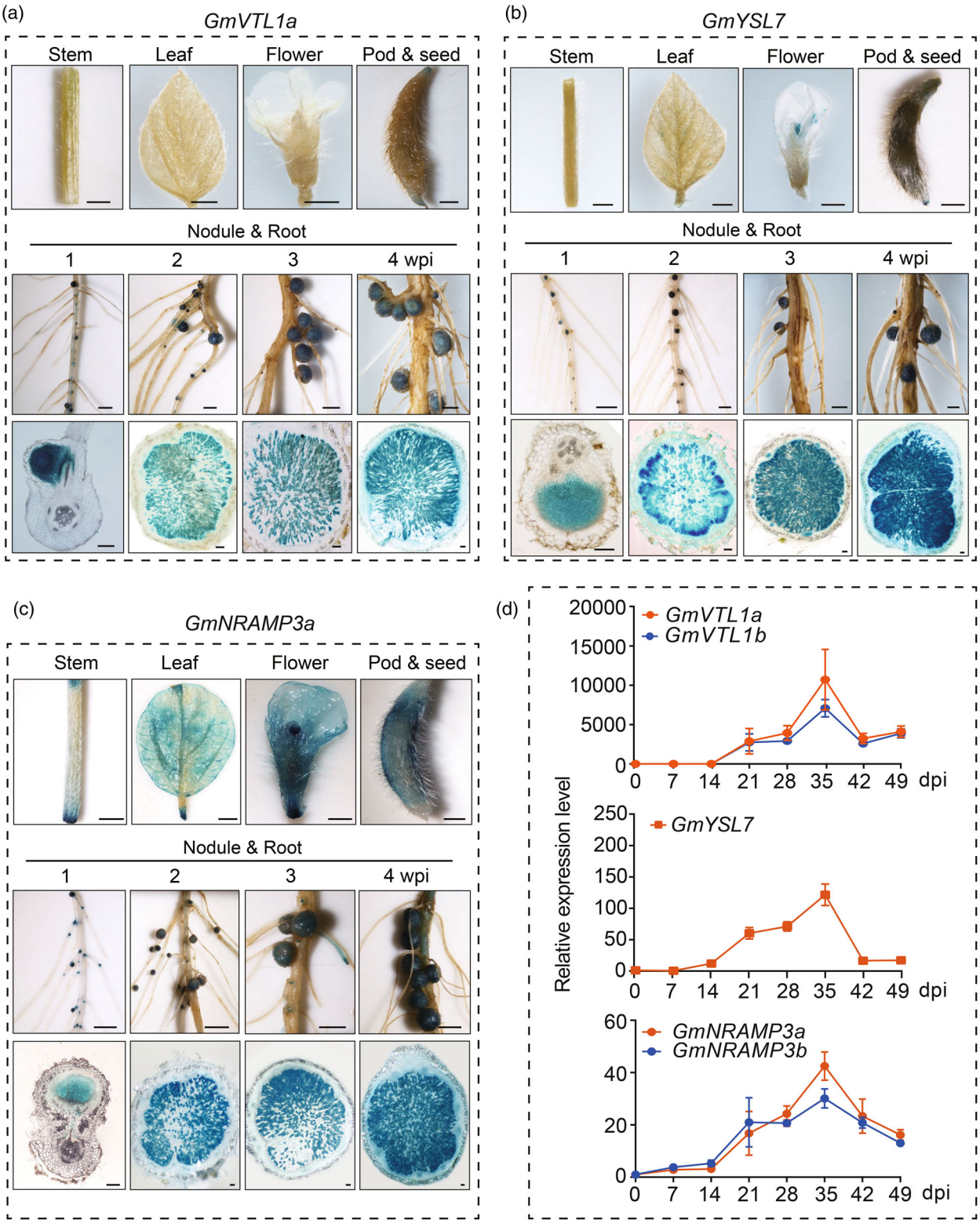


**Figure 6.** Expression patterns of *GmVTL1a*, *GmYSL7*, and *GmNRAMP3*.
(a–c) Promoter-GUS activity in different tissues. wpi, weeks post-inoculation; Scale bar, 100 μm for nodule section and 2 mm for other tissues.
(d) qPCR analysis of *GmVTL1*, *GmYSL7*, and *GmNRAMP3* during nodule development. dpi, days post-inoculation.

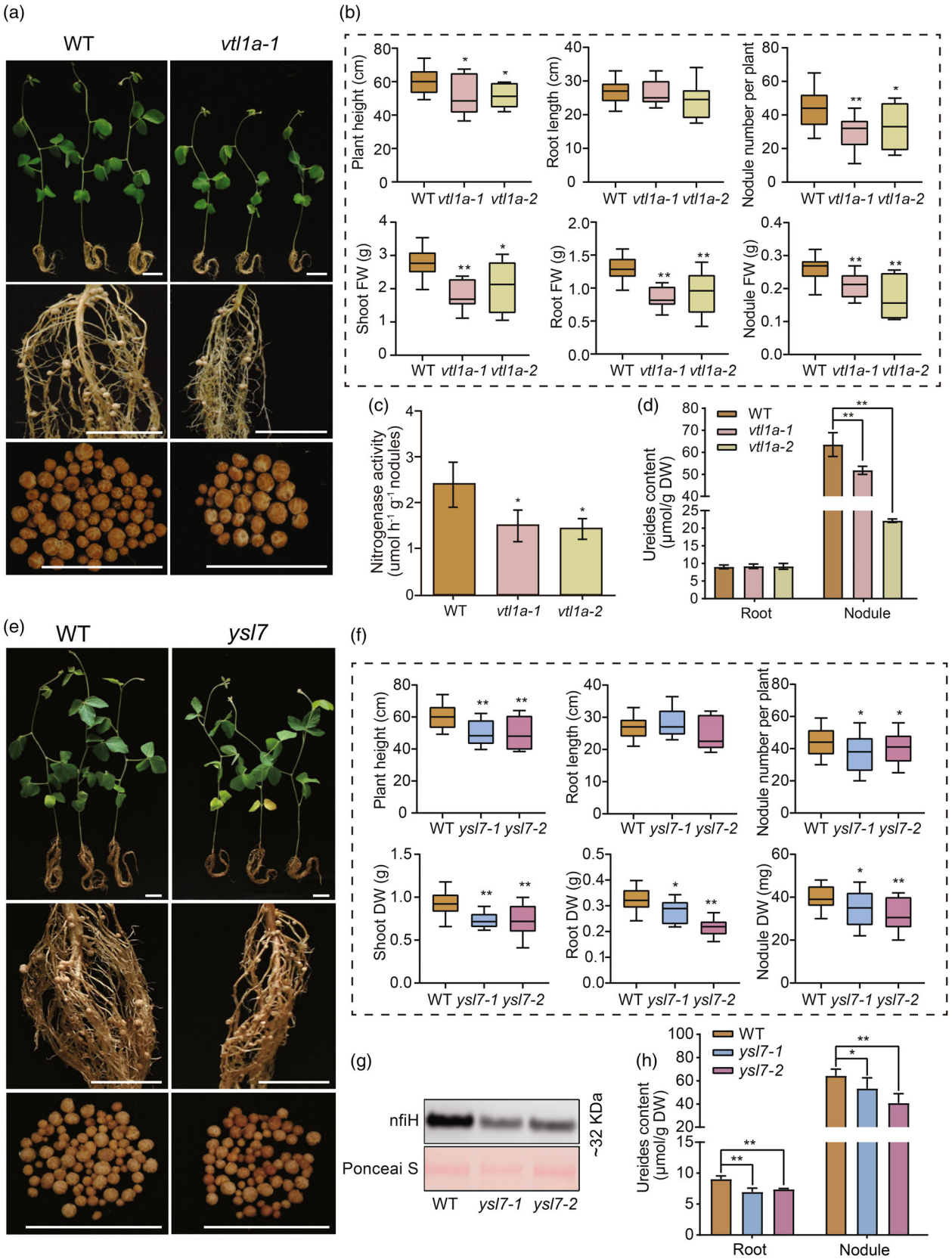

**Figure 7.** GmVTL1a and GmYSL7 are essential for nodule function.
(a) Phenotype of wild-type and *vtl1a*-1 mutant. Five-day-old seedlings were inoculated with *B. japonicum* strain USDA110 in vermiculite. Plants were grown under N-free condition for 4 weeks. Scale bar, 3 cm.
(b) Statistical analysis of plant height, root length, shoot fresh weight (FW), root FW, nodule FW, and nodule number of WT and two vtl1a mutant lines.
(c) Nitrogenase activity of nodules of WT and *vtl1a* mutants at 28 dpi.
(d) Ureide content in different tissues of WT and *vtl1a* mutant plants.
(e) Phenotype of wild-type and *ysl7* mutant.
(f) Statistical analysis of plant height, root length, shoot fresh weight (DW), root DW, nodule DW, and nodule number of WT and two *ysl7* mutant lines.
(g) Western blot analysis of nifH in WT and *ysl7* mutant nodules.
(h) Ureide content in different tissues of WT and *ysl7* mutant plants. Data are mean ± SD. Significance of differences is indicated by asterisks (LSD's ANOVA test; $*P < 0.05$, $**P < 0.01$).

GmYSL7 on nitrogen fixation, colorimetric analysis of glyoxylate derivatives was used to determine the ureide content in roots and nodules. The ureide content in the vtl1a and ysl7 mutants was significantly reduced compared with the WT (Figure 7d,h).

The study investigated the effects of *GmVTL1a* gene overexpression by generating overexpression lines *GmVTL1a*-OE-2, -5, and-11 using the $P_{GmVTL1a}$*::3Flag::gDNA* construct (Figure S11a). The *GmVTL1a*-OE lines showed a 5.7-, 3.7-, and 3.2-fold increase in *GmVTL1a* gene expression, and the Flag-tagged transgene protein could be detected (Figure S11b,c). However, compared with the WT, the *GmVTL1a*-overexpressing lines exhibited no significant differences in nodule number per plant, nodule weight, shoot weight, root weight, or overall plant growth (Figure S12). The lack of significant phenotypic differences may be attributed to the high expression levels of *GmVTL1a* (FPKM = 1500) under normal conditions, which are not markedly altered by overexpression. The results suggest that endogenous expression of *GmVTL1a* might be sufficient to achieve vacuole Fe storage to a saturation level.

### Transcription factors co-expressed with *GmVTL1a/b* and potential interacting proteins of GmVTL1a

To explore the regulatory factors of the *GmVTL* family genes, weighted correlation network analysis (WGCNA) was conducted on the nodule-preferentially expressed genes, *GmVTL1a*, *GmVTL1b*, and *GmVTL1c*. Analysis identified 749 genes that co-express with GmVTL1a/b in nodules (Figure 8a; Table S19), of which 58 are transcription factors. Of these 58 co-expressed transcription factors (TF), 19 were highly expressed in nodules (FPKM in nodules >10), including those in TF families of bHLH (08G063900), NF-YA (10G082800), and NIN (06G000400, 11G125500, 01G159200), as specifically or highly expressed in nodules (Figure 8b). Studies showed that NF-YA and NIN are essential TFs in nodule development (Bu et al., 2020; Wang et al., 2021).

Among the 749 co-expressed genes, 648 and 664 (Table S19) were detected in infection cells at 12 and 21 dpi, respectively, based on single-nucleus RNA-seq (snRNA-seq) data (Liu et al., 2023). To cross-reference the 19 co-expressed TFs with snRNA-seq data, we compiled their expression profiles and presented them in Figure S16. Further analysis identified 17 of the 19 TFs co-expressed with *GmVTL1a* and *GmVTL1b* in infected cells (highlighted in red, Figure 8b; Figure S16). Among them, 07G128700 (HRT-like) exhibited significantly high expression in the infection zone (Figure S16), indicating its potential role as an important candidate TF in regulating the expression of *GmVTL1a* and *GmVTL1b*. The potential regulatory roles and mechanisms of these TFs need to be further validated and characterized.

To identify potential proteins that interact with GmVTL1a, immunoprecipitation-mass spectrometry (IP-MS) was performed on nodule protein extracts from *pGmVTL1a-3Flag-gGmVTL1a* transgenic plants. IP-MS yielded 55 candidate interactors with scores greater than 30 (Table S20). These potential interacting proteins identified by IP-MS were cross-referenced with the soybean SM proteome atlas (Luo et al., 2023). Seventeen of the 26 candidate GmVTL1a-interacting proteins overlapped with the SM proteins (highlighted in red, Figure 8c).

## DISCUSSION

### Distinct mechanisms for maintaining Fe homeostasis in nodules and roots

Due to their ability to fix nitrogen through associations with symbiotic bacteria, legume plants play a crucial role in achieving sustainable agriculture. Although legumes can be productive without the addition of nitrogen fertilizer, the symbiotic relationship between legumes and rhizobium bacteria necessitates an adequate Fe supply (Brear et al., 2013; Kryvoruchko et al., 2018). To understand the molecular responses to Fe deficiency in nodules and roots, we conducted transcriptomic and proteomic analyses on nodules and roots collected from plants grown under Fe-sufficient (100 μM) and -deficient conditions (0 μM). Both the number of nodules, fresh weight, and Fe concentration significantly decreased under Fe-deficient conditions (Figure 1). Comparisons of DEGs and DEPs revealed significant differences in gene expression within nodules under both Fe-sufficient and -deficient conditions compared with roots (Figure 4; Tables S6 and S7). These differences exhibit two

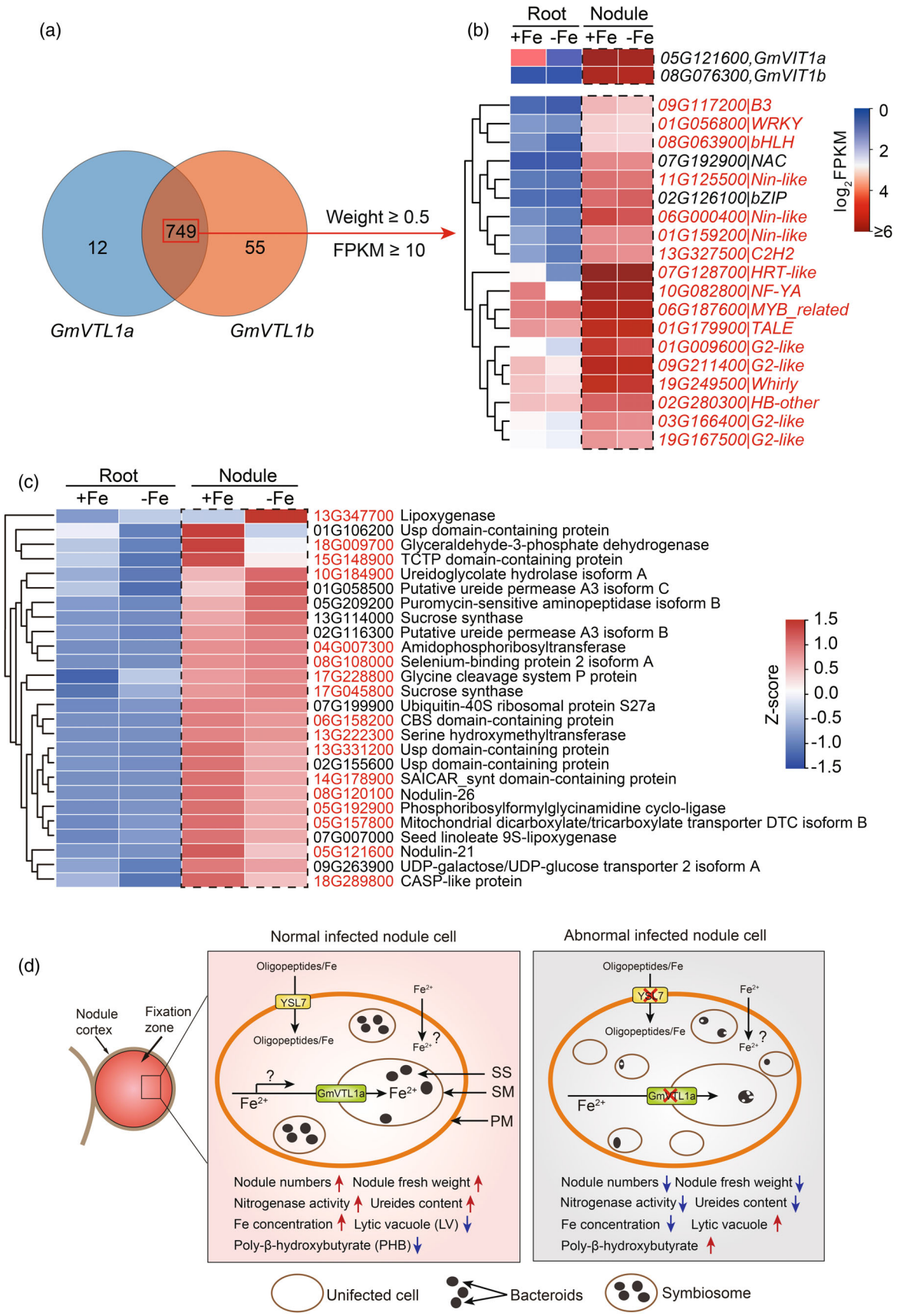

**Figure 8.** Putative regulatory and interaction networks of GmVTL1a.
(a) Co-expressed genes with *GmVTL1a* and *GmVTL1b*.
(b) Transcription factors co-expressed with *GmVTL1a* and *GmVTL1b* in nodules (FPKM > 10, weight > 0.5). The transcription factors expressed in infected cells were highlighted in red.
(c) Interacting proteins of GmVTL1a identified by IP-MS. The list of 26 potential interaction proteins was obtained by filtering out these proteins with scores below 30 and those with low nodular expression. The 17 interactors that overlapped with the SM proteins were highlighted in red.
(d) Model of Fe absorption and transport by GmYSL7 and GmVTL1a in soybean. GmYSL7 can transport oligopeptides or Fe into the nodule from the plasma membrane (PM). The GmVTL1a transports ferrous Fe from the symbiosome space (SS) into the interior of the symbiosome via the symbiosome membrane (SM) to support nodule development. The red upward arrows indicate maintenance of normal or increased levels, while the blue downward arrows indicate reduced levels.

main characteristics. First, numerous genes involved in Fe transport and storage are activated in both roots and nodules following Fe deficiency (Figure 4). Secondly, the response to Fe deficiency is largely characterized by the induction of different isogenes in nodules compared with roots (Figure 3; Tables S4 and S5). Thus, the regulatory network for Fe deficiency in roots and nodules is significantly different. Through transcriptome and proteome analysis, it was found that most of the DEGs and DEPs were associated with Fe homeostasis or transport, nitrogen metabolism, plant–pathogen interaction, starch and sucrose metabolism, fatty acid metabolism or biosynthesis, and plant hormone signal transduction (Figures S3 and S4; Tables S8–S11).

Multiple omics studies have uncovered significant differences in the numbers of genes or proteins in roots and nodules (Figure 2). Notably, genes that respond to Fe limitation exhibit a particularly large discrepancy between these two tissues, prompting the question: why is there such a significant difference in their response to Fe limitation? This difference may be attributed to the aerobic environment of roots and the low $O_2$ concentration in the infected region of nodules, which ensures that nitrogenase activity is not inhibited (Bergersen, 1997; Schulte et al., 2021). One of the primary routes for Fe toxicity or cell damage is through the Fenton reaction, leading to cellular damage of lipids, proteins, and DNA (Ensing et al., 2003). As reactive oxygen species play significant signaling roles in maintaining Fe homeostasis, this can cause differential gene expression between nodules and roots. An additional factor contributing to the DEGs involved in Fe homeostasis between nodules and roots may include differences in Fe demand and uptake pathways. The Fe requirement in rhizobia-infected nodules is substantially higher than in regular roots, as efficient N fixation by bacteroids relies on the expression of numerous Fe-dependent proteins. These include nitrogenases, ferritins, hydrogenases, and cytochromes, which are essential for respiration under hypoxic conditions (Brear et al., 2013). The observation that vacuolar Fe storage protein (*GmVTL1a/b*) expression levels are higher in nodules compared with roots under Fe-replete conditions confirms the higher Fe demand in nodules (Figure 5e). LeVier et al. (1996) proposed that Fe(III)-citrate is the primary form of Fe in the cytosol of nodule cells and that a ferric chelate reductase likely reduces Fe(III) to Fe(II) before uptake into symbiosomes. This aligns with the presence of VTL1a, which is likely a ferrous Fe transporter on the SM. From the data of the study, it can be seen that the bacteroids and the root cells likely use different Fe transporters.

### Case studies of the nodule-preferential expressed Fe transporter genes

NRAMP, YSL, and VTL transporter families are known to play important roles in Fe or other metal transportation in non-legume plants (Gavrin et al., 2021; Kim et al., 2006; Kobayashi & Nishizawa, 2012; Zhang et al., 2012). The functional diversification of these transporter families within legume nodules is particularly noteworthy. Key nodule-enriched genes, including *GmNRAMP3a/b*, *GmYSL7*, and *GmVTL1a/b*, were identified (Figures 5c–e and 7d) and their spatial expression patterns were confirmed through promoter-driven *GUS* assays in transgenic soybean (Figure 6a–c). Mutants of *ysl7* and *vtl1a* demonstrated that these two genes are critical for nodule development and function (Figure 7). Although other research groups published papers on these genes recently (Brear et al., 2020; Liu et al., 2020; Wu et al., 2023), the phenotypic data obtained from stable transgenic soybeans in the study not only prove the reliability of transcriptomic analysis results, but also provide independent support for peer research data.

Among the 20 members of the soybean VIT family, 10 genes form two tandem gene duplication clusters on chromosomes 5 (05G121200–05G121600) and 8 (08G075900–08G076300) (Figure S13a,b). A phylogenetic tree showed the members of the VTL clusters on Chr 5 and Chr 8 are paired with each other (Figure S7b), implying that a possible gene duplication event likely occurs during the evolutionary process. Analysis of *GmVTL1a* homologous sequences in different varieties (SoyOD website, Li, Ni, et al., 2024) revealed that variations in the VIT domains of GmVTL1a among 26 genomes resulted in different transmembrane domains (Figure S13c). Among the five pairs of *VTL* genes, the pair of *GmVTL1a* and *GmVTL1b* exhibits the highest expression in nodules (Figure 5e; Figure S7a). Comparison of the expression of *GmVTL1a* and *GmVTL1b* showed that *GmVTL1a* is expressed at 2.6

times higher levels than *GmVTL1b* (Figure 5e; Figure S7). Similar results were observed in the single-nucleus and spatial transcriptome data, where the number of cells expressing *GmVTL1a* in infected areas was significantly higher than those expressing *GmVTL1b* (Liu et al., 2023, Figure S8). Phenotypically, while the *vtl1a/1b* double mutant exhibits severe defects in nodule development and SNF (<20% activity, Liu et al., 2020), the *vtl1a* single mutant retains approximately 50% of nitrogenase activity (Figure 7c), suggesting partial compensation by *GmVTL1b*. This partial functional compensation by GmVTL1b is supported by the following evidence: (i) GmVTL1b exhibits weak but detectable iron transport activity in yeast Δccc1 complementation (Figure S14; and Brear et al., 2020); (ii) GmVTL1b was identified in purified SM and symbiosome-enriched membrane samples (Brear et al., 2020); (iii) *GmVTL1b* co-expresses with *GmVTL1a* in nodule cells in the spatial transcriptomic data (Liu et al., 2023; Figure S9). In the future, comparing *vtl1a*, *vtl1b* single, and *vtl1a/vtl1b* double mutants would clarify the contribution of GmVTL1b.

In this study, we systematically compared the spatial expression patterns of the three nodule-expressed transporter genes, *GmNRAMP3a*, *GmYSL7*, and *GmVTL1a*, using the *promoter-GUS* transgenic plants. While all three transporters are highly expressed in nodules, only *GmNRAMP3a* shows broad activity in stems, leaves, flowers, and pods (Figure 6c). This suggests a functional divergence between *GmNRAMP3a*, which is specialized for nodule Fe homeostasis, and GmVTL1a/GmYSL7 (which are specialized for nodule Fe homeostasis). Complementation assays on yeast mutants of the *fet3fet4* [DEY1453 (Eide et al., 1996)] and ▵*ycf1* mutant strain DTY167 strains (Li et al., 1996) revealed that GmNRAMP3 does not transport Fe and manganese but does transport cadmium (Figure S15). The function and mechanism of the *GmNRAMP3* gene in soybean nodules are worth studying in the future.

### Regulatory network of GmVTL1a

TFs that regulate plant VIT/VTLs have not been previously reported. In this study, a group of *GmVTL* genes was found to be preferentially expressed in nodules (Figure 3c). Co-expression analysis was conducted to identify genes co-expressed with these *GmVTL* genes. As a result, 19 nodules preferentially expressed TFs were found to be co-expressed with these *GmVTL* genes (Figure 8b). In addition, the potential interacting proteins of GmVTL1a were identified using IP-MS experiments. Among these putative GmVTL1a-interacting proteins, many are associated with the SNF and nodule development, including 04G007300, 08G108000, 06G158200, 08G120100, and 18G289800 (Figure 8c). Genes encoding these proteins exhibited significantly higher expression levels in the nodules compared with the roots (Table S20). We hypothesize that GmVTL1a either directly or indirectly interacts with these proteins in nodules, thereby affecting nodule development. Although the regulatory mechanisms and functions of these TFs and IP-MS candidates are yet fully clear, the list includes key factors that may be important for nodule function and warrant further investigation in the future.

Taken together, our results provide a novel transcriptomic and proteomic analysis, identifying multiple DEGs and DEPs under Fe-deficient conditions, and confirming distinct expression profiles in nodules compared with roots. Meanwhile, we analyzed various genes and proteins with unique expression profiles in roots and nodules that are involved in regulating Fe homeostasis and transport. This study takes the *GmVTL*, *GmYSL*, and *GmNRAMP* genes as case studies. Through stable soybean transgenic materials and spatial transcriptome data, we confirm the specific expression of *GmVTL1a* in the infected zone of nodules and the roles of *GmVTL1a* and *GmYSL7* in nodule Fe transport, nodule development, and nitrogen fixation (Figure 8d). Through WGCNA and IP-MS analyses, we investigated potential interactions among TFs and proteins, providing valuable insights for further study on SNF in soybeans. This research deepens our understanding of the regulatory mechanisms of Fe homeostasis in soybean nodules and Fe uptake, thereby establishing a theoretical framework to improve the efficiency of SNF in nodules.

## EXPERIMENTAL PROCEDURES

### Plant materials, rhizobium strain, and growth conditions

Soybean (*Glycine max* cv. Williams82 and Tianlong1) and *Bradyrhizobium japonicum* strain USDA110 were used in this study. Hydroponic growth conditions have been described previously (Li et al., 2018). Soybean seeds were surface-sterilized using chlorine gas for 10–12 h and then germinated in moist, sterilized vermiculite. Five-day-old seedlings were inoculated with rhizobium suspension ($OD_{600}$ = 0.8–1.0) for 1 h and transplanted into a low-N solution supplemented with 100 μM Fe(II)-EDTA (Fe-sufficiency, +Fe) or 0 μM Fe(II)-EDTA (Fe-deficiency, −Fe). The low-N nutrient solution (modified 1/2 Hoagland) was prepared with the following components: 0.1 mM $NH_4NO_3$, 1.5 mM $K_2SO_4$, 2.5 mM $CaCl_2 \cdot 2H_2O$, 1.0 mM $MgSO_4 \cdot 7H_2O$, $3.8 \times 10^{-4}$ mM $ZnSO_4 \cdot 7H_2O$, $1.57 \times 10^{-3}$ mM $CuSO_4 \cdot 5H_2O$, $0.09 \times 10^{-3}$ mM $(NH_4)_6Mo_7O_{24} \cdot 4H_2O$, 0.02313 mM $H_3BO_3$, $0.1 \times 10^{-3}$ mM $CoSO_4 \cdot 7H_2O$, $4.57 \times 10^{-3}$ mM $MnCl_2 \cdot 4H_2O$, 0.25 mM $KH_2PO_4$, 0.1 mM Fe(II)-EDTA (formulated with $FeSO_4 \cdot 7H_2O$ and $Na_2EDTA$). Nutrient solution was refreshed every 3 days, and pH was adjusted to 5.8. Roots were continuously aerated. For transgenic soybean plants culture, plants were grown in vermiculite and fertilized with an N-free solution once a week. Soybean plants were grown in a growth chamber under condition of 16 h light at 28°C and 8 h dark at 24°C. The light intensity was 216 μmol $m^{-2}$ $sec^{-1}$ and the humidity maintained at 55–65%.

### Metal content measurement

Metal content in different soybean tissues was determined according to a modified procedure from Wang et al. (2013). Leaves, roots, and nodules were oven-dried at 65°C for 3 days. Samples

were weighed and digested with 5 ml of 14 M $HNO_3$ and 1 ml 30% (v/v) $H_2O_2$ in microwave digestion apparatus (CEM, MARS6). The metal content of different tissues was determined by inductively coupled plasma-mass spectrometry (ICP-MS).

For measuring yeast vacuole Fe concentration, vacuoles were isolated according to Li et al. (2001) with modifications. Briefly, yeast was cultured to an $OD_{600}$ of 0.6–0.8. Cells were harvested by centrifugation at 3000 ***g*** for 5 min at 24°C, washed with a buffer containing 100 mM Tris–HCl and 10 mM DTT (pH 9.4) for 10 min, and recentrifuged. The pellet was resuspended in spheroplasting buffer (1.2 M sorbitol, 20 mM potassium phosphate, pH 7.4) and incubated with 20 μg $ml^{-1}$ oxalyticase for 30 min at 30°C. Spheroplasts were collected by centrifugation at 3500 ***g*** for 5 min at 4°C and resuspended in 2.5 ml of 15% (w/v) Ficoll, 200 mM sorbitol, 10 mM K-PIPES, pH 6.8. DEAE-Dextran at a final concentration of 90 μg/100$OD_{600}$ units of cells was added and incubated on ice for 5 min, followed by a 2 min heat shock at 30°C. The vacuoles were obtained from a four-tiered gradient of Ficoll: 6 ml of 15% (w/v) Ficoll, 2.5 ml of 8% Ficoll, and 2.5 ml of 4% (w/v) Ficoll, and the tube was filled with 200 mM sorbitol, 10 mM K-PIPES, pH 6.8. The gradient was centrifuged at 175 000 ***g*** for 90 min at 4°C. The vacuoles were recovered on the 0–4% interphase by centrifugation for 30 min at 110 000 ***g*** for 30 min at 4°C. The protein content and Fe content in the vacuoles were determined by the Bradford (1976) method and ICP-MS, respectively.

### Illumina library preparation and RNA-seq

Each individual nodule was labeled at its initiation stage with strings to ensure the sampled nodules that were at their maturation stage, which occurs 3 weeks after the nodule initiation. A total of 12 samples were collected for three independent biological replicates for RNA extraction, mRNA library construction, and sequencing. Total RNA was isolated using Trizol Reagent (Invitrogen, Carlsbad, CA, USA) following the manufacturer's protocol. The quantity and purity of total RNA were checked using the RNA 6000 Nano kit (Agilent Technologies, Santa Clara, CA, USA). Three micrograms of total RNA were subjected to mRNA enrichment using oligo (dT) magnetic beads (Invitrogen). The mRNA was then fragmented into smaller pieces using an RNA fragmentation buffer. Sequencing was performed on the Illumina Novaseq 6000. Raw sequence reads were aligned to the JGI Williams82.a2 soybean assembly (*Glycine max*, cv. Williams82) using HISAT2 software (v2.1.0) (Kim et al., 2015). Gene expression levels were normalized using fragments per kilobase million mapped reads (FPKM) method (Trapnell et al., 2010). DESeq2 (Love et al., 2014) was used to identify differential gene expression between different comparison groups. The DEGs with threshold were set at $\log_2$ |Fold Change| ≥ 1 and $P \leq 0.05$.

### TMT labeled quantitative proteome sequencing and analysis

The protein samples were extracted from whole soybean roots or nodules without dissection and subjected to TMT-based quantitative proteomics analysis using the following workflow (McAlister et al., 2012): Proteins were extracted with SDT lysis buffer (4% SDS, 100 mM Tris/HCl pH 7.6, 0.1 M DTT), quantified via bicinchoninic acid assay, and digested into peptides using the filter aided sample preparation (FASP) method. Peptides (100 μg per sample) were labeled using the TMT labeling kit (Thermo Fisher Scientific, Waltham, MA, USA) following the manufacturer's protocol. Labeled peptides were fractionated using the high-pH reversed-phase peptide fractionation kit (Thermo Fisher Scientific) according to the manufacturer's instructions before being analyzed by LC–MS/MS analysis on a Q-exactive mass spectrometer coupled with an EASY-nLC 1000 liquid chromatograph instrument (Thermo Fisher Scientific). MS/MS spectra data were searched against the UniProt database using MASCOT (v2.2; Matrix Science, London, UK) within Proteome Discoverer 1.4 for peptide identification and quantification. Peptides were filtered at a 1% false discovery rate, and IonScore distributions were generated from aggregated peptide matches. Differentially expressed proteins were defined as those with |fold change| ≥ 1.2 and $P \leq 0.05$. The experiment included three biological replicates to ensure statistical reliability.

### Gene ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment

GO term enrichment analysis was performed using Blast2GO. Pathway annotation was conducted using the KEGG database. The raw RNA-seq data generated in this study were deposited in the Sequence Read Archive (SRA) of the National Center for Biotechnology Information (NCBI) under accession number SRP364857.

### Gene family identification and hierarchical clustering heatmap

The genome annotations and sequences of *Glycine max* Williams82.a2 were downloaded from phytozome (https://phytozome-next.jgi.doe.gov/), and a local BLASTP database was built using the BioEdit tool. The comprehensive identification of *GmVTL*, *GmYSL*, *GmFRO*, *GmNRAMP*, and *GmFER* gene family members was achieved using identified protein sequence. The gene family members were identified by a BLASTP search. Each new query parameter was set to an *E*-value ≤$10^{-2}$ to avoid the loss and addition of orthologs. The domains of obtained proteins were also further verified using the NCBI-Conserved Domain database (https://www.ncbi.nlm.nih.gov/Structure/) search program and SMART databases (https://smart.emblheidelberg.de/). Proteins lacking common domains were removed from further analysis. For hierarchical clustering, $\log_2$ FPKM values were normalized and then clustered using average linkage based on Euclidean distance. The SoyOD web service at https://bis.zju.edu.cn/soyod/home/ was utilized to identify homologous genes across different varieties (Li, Ni, et al., 2024). The phylogenetic tree was constructed using the maximum likelihood method and was validated with a bootstrap analysis of 1000 replicates using MEGA X (Kumar et al., 2018). The transmembrane domain was predicted using TMHMM (v2.0, https://services.healthtech.dtu.dk/services/TMHMM-2.0/). Identity and similarity were obtained by protein sequence alignment (https://www.ebi.ac.uk/jdispatcher/psa/emboss_needle). Soybean single-nucleus and spatial transcriptomic data were analyzed using the atlas of soybean nodule website (http://119.45.35.29:3569/).

### Vector construction and soybean transformation

For yeast expression, full-length open reading frames of Gm*VTL1a*, Gm*VTL1b*, At*VIT1*, *GmYSL7*, and *GmNRAMP3a* were ligated into pAG426GPD Gateway destination vector. To investigate tissue-specific expression of Gm*VTL1a*, a 2 kb promoter region upstream from the start codon of Gm*VTL1a*, *GmYSL7*, and *GmNRAMP3a* were amplified and constructed into pBI101.3 to generate $P_{GmVTL1a}$::*GUS*, $P_{GmYSL7}$::*GUS*, and $P_{GmNRAMP3a}$::*GUS* constructs. The CRISPR/Cas9 system was used to generate *vtl1a and gmysl7* mutants as described (Wang et al., 2021). Briefly, the first exon of *GmVTL1a* and *GmYSL7* was selected to design the guide RNA using the online tools CRISPRP v2.0 (http://crispr.hzau.edu.cn/cgi-bin/CRISPR2/CRISPR). For overexpression of

Gm*VTL1a*, the endogenous promoter (2 kb upstream from the start codon of *GmVTL1a*) was utilized to drive Gm*VTL1a* genomic DNA (935 bp). These fragments were cloned and ligated into a modified pLM-B001 (with Flag tag) to generate the overexpression constructs ($P_{GmVTL1a}$*::3Flag::gDNA*). The primers used for amplifying these fragments are listed in Table S21. After verification by DNA sequencing, constructs were introduced into *Agrobacterium tumefaciens* strain LBA4404 for soybean transformation. T2 and T3 homozygous seeds were used for phenotypic analysis.

### Histochemical GUS staining

Using the soybean stable transformation method, *pVTL1a*-GUS and *pYSL7*-GUS were transferred to the Williams 82, and *pNRAMP3a*-GUS was transferred to the Tianlong1. Transgenic soybean tissues were incubated overnight at 37°C in GUS staining buffer (100 mM sodium phosphate, pH 7.0, 10 mM $Na_2EDTA$, 1 mM 5-bromo-4-chloro-3-indolyl-b-D-glucuronidase, 0.5% TritonX-100 and 20% methanol, with freshly added potassium ferrocyanide and potassium ferricyanide to a final concentration of 2.5 mM). After staining, samples were washed with 75% (v/v) ethanol. The sectioning process of nodule tissue involved encapsulating the samples in 5% agarose, cutting approximately 50 μm thick sections using a vibrating microtome (VT 1000S; Leica, Bensheim, Germany), and immersing the sections in GUS staining solution for incubation. Images were captured under a stereomicroscope (SMZ745T; Nikon, Tokyo, Japan).

### Quantitative RT-PCR

Samples were collected from different tissues, and total RNA was isolated using the RNA-easy isolation reagent (R701; Vazyme, Nanjing, China) according to the manufacturer's instructions. cDNA was synthesized from the total RNA using a cDNA synthesis kit (AG Biology, Hunan, China), and RT-qPCR was performed on a LightCycler® 480 machine (Roche, Basel, Switzerland) with SYBR Green Supermix (CWBIO, Beijing, China). The housekeeping gene *GmTefs1* (*Glyma.17G186600*) was used as an internal control for sample normalization (Jian et al., 2008). The qRT-PCR primers are listed in Table S21. All amplification reactions were conducted with three biological and two technical replicates. Normalized relative expression was calculated using the $2^{-\Delta\Delta Ct}$ method.

### Transport activity assay in yeast

To test for the transport of Fe(II), *GmVTL1a*, *GmVTL1b*, and *AtVIT1* were constructed into the vector of pAG426GPD and then introduced into yeast Δ*ccc1* mutant (Li et al., 2001). The empty vector was transformed into both WT and mutant strains as the controls. Yeast transformants were grown on SD-Ura medium for 3 days and resuspended in sterile water. Yeast cells were collected by centrifugation, serially diluted, and spotted on either SD−Ura plates supplied with 0, 2.5, 5, and 7.5 mM $FeSO_4$ and incubated at 28°C for 7 days in the dark. To test for the transport of Fe(III) and $CdCl_2$, *GmNRAMP3a*, *GmNRAMP3b*, and *AtIRT1* were constructed into the vector of pAG426GPD and then introduced into yeast *DDY4* (*fet3fet4*) and Δ*ycf1* mutant. Yeast cells were collected by centrifugation, serially diluted, and spotted on either SD−Ura plates supplemented with $FeCl_3$ (0, 2, 20, and 100 μM) and $CdCl_2$ (0, 10, 20, and 30 μM). The plates were then incubated at 28°C for 7 days in the dark.

### Co-expression analysis

A weighted gene co-expression network analysis (WGCNA) was performed using FPKM values (Langfelder & Horvath, 2008). The optimal soft threshold was selected to convert the correlation matrix into an adjacency matrix, followed by the generation of a topological overlap matrix. To enhance reliability, weak connections (weight <0.25) were removed. Candidate genes co-expressed with *GmVTL1a* and *GmVTL1b* were identified and cross-referenced with single-nucleus RNA-seq data at infection time points (12 and 24 dpi) (Liu et al., 2023). Finally, nodule-specific co-expressed TFs were filtered based on FPKM >10 and correlation weight >0.5.

### Immunoprecipitation-mass spectrometry

An overexpression construct *pGmVTL1a-3Flag-gGmVTL1a* was generated by cloning the genomic sequence of *GmVTL1a*, along with its native promoter, into the pLM-B001 vector. The resulting construct carried a C-terminal 3× Flag tag for protein detection. Stable transgenic soybean lines were produced via cotyledon node transformation. Wild-type plants were included as controls to account for background signal. After 4 weeks of growth, nodules were harvested for protein extraction and downstream analysis.

Protein extraction was performed using a protein buffer solution (50 mM Tris–HCl pH 7.5, 150 mM NaCl, 10% [v/v] glycerol, 0.1% [v/v] Nonidet P-40, 1 mM PMSF, and protease inhibitor cocktail [Sigma-Aldrich, St. Louis, MO, USA]). Following centrifugation, the supernatant from the lysate buffer was incubated with anti-Flag® M2 magnetic beads (Sigma; #M8823) at 4°C overnight and eluted according to the manufacturer's instructions. The bands were then obtained by Western blot analysis.

The in-gel tryptic digestion process involved destaining the gel pieces with a solution of 50 mM $NH_4HCO_3$ and 50% acetonitrile until they were clear. Subsequently, the gel pieces were dehydrated with 100 μl of 100% acetonitrile for 5 min; the liquid was discarded, and the gel pieces were rehydrated in 10 mM dithiothreitol and incubated at 56°C for 60 min. After another dehydration step with 100% acetonitrile, the gel pieces were rehydrated with 55 mM iodoacetamide and incubated at room temperature in the dark for 45 min. Following this, the gel pieces were washed with 50 mM $NH_4HCO_3$, dehydrated again with 100% acetonitrile, and then rehydrated with 10 ng $\mu l^{-1}$ trypsin in 50 mM $NH_4HCO_3$ on ice for 1 h. Excess liquid was removed, and the gel pieces were digested with trypsin at 37°C overnight. Peptides were extracted using 50% acetonitrile or 5% formic acid, followed by 100% acetonitrile, dried completely, and resuspended in 2% acetonitrile or 0.1% formic acid. For LC–MS/MS analysis, the tryptic peptides were dissolved in 0.1% formic acid and loaded onto a reversed-phase analytical column. Peptides were ionized using an NSI source and analyzed by tandem mass spectrometry (MS/MS) on a Q ExactiveTM Plus (Thermo) instrument coupled online to the UPLC.

The identified proteins were systematically matched against their corresponding protein dataset to generate a list of candidates interacting proteins. Candidate interactors were filtered in three steps: (i) exclusion of proteins with IP-MS scores <30; (ii) removal of background contaminants identified in control samples; (iii) elimination of low-abundance nodule proteins (FPKM < 10). The resulting proteins were cross-referenced with the SM proteome to identify shared proteins (Luo et al., 2023).

### Determination of ureide content

Ureide concentrations were measured using the Lescano (2020) method, with corresponding improvements made to the procedure. Soybean tissues were collected and dried in an oven at 65°C, then ground into a fine powder. For extraction,

approximately 50 mg of the powdered tissue was weighed and mixed with 500 μl of extraction solution (25 mM potassium phosphate buffer). The mixture was heated in a 100°C water bath for 20 min. Subsequently, it was centrifuged at 13 523 ***g*** for 10 min at 4°C. The supernatant was transferred to a 1.5 ml centrifuge tube and kept on ice for subsequent analysis. Take 200 μl of the sample solution, add 100 μl of 0.5 M NaOH, mix, and heat in a boiling water bath for 15 min. Cool to room temperature, add 100 μl of 0.65 M HCl and 100 μl of 22.8 mM phenylhydrazine hydrochloride solution, mix, and heat in a boiling water bath for 4 min. Immediately place the sample on ice and cool for 15 min. Add 400 μl of pre-chilled concentrated HCl and 100 μl of 0.05 M potassium ferricyanide to the sample. Incubate at room temperature for 10 min, then measure the absorbance at 520 nm using a spectrophotometer.

### Acetylene reduction assay

The acetylene reduction assay was used to measure nitrogenase activity. The nodules from WT and vtl1a plants were isolated and placed in airtight glass bottles. Acetylene gas (1 ml) was injected into each bottle after the same volume of air was pumped out and incubated for 2 h at 28°C. After injecting sodium hydroxide to terminate the reaction, 1 ml gas samples were taken for ethylene determination by gas chromatography and mass spectrometry. The amount of ethylene produced was determined by measuring the height of the ethylene peak relative to the background. Subsequently, the weights of the nodules were measured.

### Nodule ultrastructure observation

The ultrastructure of nodules was observed by transmission electron microscope. Fresh nodules were cut with a razor blade, and the central part with red color was fixed with 2.5% glutaraldehyde in PBS (0.1 M, pH 7.0) overnight at 4°C, followed by post-fixation with 1% $OsO_4$ in PBS for 2 h. The fixed samples were first dehydrated using a graded ethanol series (30, 50, 70, and 80%) for 15 min each, then further dehydrated using a graded acetone series (90, and 95%) for 15 min each, and finally subjected to absolute acetone dehydration twice for 20 min each. The samples were placed in a 1:1 mixture of absolute acetone and Spurr resin for 1 h at room temperature, then transferred to a 1:3 mixture of absolute acetone and Spurr resin for 3 h, and finally embedded in pure Spurr resin overnight. The specimens were heated at 70°C, and sections were cut using a Leica EM UC7 ultratome. The sections were stained with uranyl acetate and alkaline lead citrate for 5–10 min each and observed using a Hitachi H-7650 transmission electron microscope (Hitachi, Tokyo, Japan).

## AUTHOR CONTRIBUTIONS

HS, JW, and JL conceived and designed the experiments. QK, JL, LL, and JZ carried out the experiments. JL analyzed the transcriptome and proteome data, visualized the data presentation, and drafted the manuscript. HS and JW reviewed and edited the manuscript. All authors read and approved the final manuscript.

## ACKNOWLEDGMENTS

This work was supported by the Ministry of Science and Technology of China (2021YFF1001204), Key Research and Development Project of Zhejiang Province (2021C02057), the National Natural Science Foundation of China (32572127) and the 111 project of the Ministry of Education (B14027). We also thank Prof. Jiming Xu (Zhejiang University) and Dr. Shelong Zhang (Zhejiang University) for technical assistance, and Weimin Zhang of the Agricultural Experiment Station of Zhejiang University for plant propagation and field test.

## CONFLICT OF INTEREST

The authors declare that they have no competing interests.

## DATA AVAILABILITY STATEMENT

The raw RNA-seq data used in this study were deposited in the National Center for Biotechnology Information (NCBI) Sequence Read Archive (SRA) database (https://www.ncbi.nlm.nih.gov/sra/), under the accession number PRJNA817859. The proteomic data generated in this study have been deposited into the OMIX, China National Center for Bioinformation/Beijing Institute of Genomics, Chinese Academy of Sciences (OMIX ID: OMIX009667), and are publicly accessible at https://ngdc.cncb.ac.cn/omix/release/OMIX009667.

## SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article.

**Figure S1.** Distribution of total expression (a) and principal component analysis (b) of the RNA-seq data.

**Figure S2.** Quality control of the proteome and characterization of peptides.

**Figure S3.** Gene ontology (GO) of the DEGs and DEPs.

**Figure S4.** Kyoto Encyclopedia of Genes and Genomes (KEGG) classification of the DEGs (a) and DEPs (b).

**Figure S5.** Venn plot showing the interaction of different comparison groups.

**Figure S6.** Domain analysis of *GmFER*, *GmNRAMP*, *GmYSL*, *GmFRO*, and *GmVTL* gene family.

**Figure S7.** Cluster, phylogenetic tree, expression level of the GmVIT family, and amino acid sequence comparison of *GmVTL1a/b* with *AtVIT1* and *LjSEN1*.

**Figure S8.** Expression profiling of *GmVTL1a* and *GmVTL1b* in soybean plants.

**Figure S9.** Expression analysis of *GmVTL1a* and *GmVTL1b* in soybean nodule sections using data achieved from the single-nucleus and spatial transcriptomic data (Liu et al., 2023).

**Figure S10.** Identification of CRISPR/Cas9-mediated knockout *vtl1a* and *ysl7* mutants.

**Figure S11.** Identification of *GmVTL1a* overexpression lines.

**Figure S12.** Effect of overexpression of *GmVTL1a* on soybean plant growth and nodules.

**Figure S13.** Chromosome distribution, gene cluster, and phylogenetic tree of the *GmVTL* gene family.

**Figure S14.** Analysis of Fe(II) transport activity of GmVTL1a and GmVTL1b expressed in yeast and ultrastructure of infected cells in WT and vtl1a mutant nodules.

**Figure S15.** Metal transport activities of GmNRAMP3a/3b.

**Figure S16.** Cross-referencing with single-nucleus expression data was performed on the 19 screened transcription factors.

**Table S1.** Quality assessment and sample information of raw RNA-seq data.

**Table S2.** Reads compared with reference genome in RNA-seq data.

**Table S3.** Protein identification list based on TMT quantification.

**Table S4.** Fold changes of transcripts or proteins significantly expressed in N+Fe versus R+Fe.

**Table S5.** Fold changes of transcripts or proteins significantly expressed in N−Fe versus R−Fe.

**Table S6.** Fold changes of transcripts or proteins significantly expressed in N−Fe versus N+Fe.

**Table S7.** Fold changes of transcripts or proteins significantly expressed in R−Fe versus R+Fe.

**Table S8.** GO analysis of DEGs in the different comparison groups.

**Table S9.** GO analysis of DEPs in the different comparison groups.

**Table S10.** KEGG analysis of DEGs in the different comparison groups.

**Table S11.** KEGG analysis of DEPs in the different comparison groups.

**Table S12.** Fold changes of the common DEGs (nodules versus roots) concurrently associated with Fe-sufficient or -deficient conditions.

**Table S13.** Fold changes of the common DEPs (nodules versus roots) concurrently associated with Fe-sufficient or -deficient conditions.

**Table S14.** Fold changes of the common DEGs (−Fe versus +Fe) concurrently associated with nodules and roots.

**Table S15.** Fold changes of the common DEPs (−Fe versus +Fe) concurrently associated with nodules and roots.

**Table S16.** Fold changes of the DEGs (−Fe versus +Fe) only expressed in nodules.

**Table S17.** Fold changes of the DEPs (−Fe versus +Fe) only expressed in nodules.

**Table S18.** Expression levels (FPKM) of the *GmFER*, *GmNRAMP*, *GmYSL*, *GmFRO*, and *GmVTL* gene families.

**Table S19.** Co-expression network of nodule higher expression genes GmVTL1a, GmVTL1b, and GmVTL1c.

**Table S20.** Proteins of GmVTL1a associated with co-immunoprecipitation are identified by LC–MS/MS analysis.

**Table S21.** Primers for vector construction and qRT-PCR.

## REFERENCES


**Bergersen, F.J.** (1997) Regulation of nitrogen fixation in infected cells of leguminous root nodules in relation to $O_2$ supply. *Plant and Soil*, **191**, 189–203.

**Bradford, M.** (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Analytical Biochemistry*, **72**, 248–254.

**Brear, E.M., Bedon, F., Gavrin, A., Kryvoruchko, I.S., Torres-Jerez, I., Udvardi, M.K. *et al.*** (2020) GmVTL1a is an iron transporter on the symbiosome membrane of soybean with an important role in nitrogen fixation. *New Phytologist*, **228**, 667–681.

**Brear, E.M., Day, D.A. & Smith, P.M.C.** (2013) Iron: an essential micronutrient for the legume-rhizobium symbiosis. *Frontiers in Plant Science*, **4**, 359.

**Bu, F.J., Rutten, L., Roswanjaya, Y.P., Kulikova, O., Rodriguez-Franco, M., Ott, T. *et al.*** (2020) Mutant analysis in the nonlegume *Parasponia andersonii* identifies NIN and NF-YA1 transcription factors as a core genetic network in nitrogen-fixing nodule symbioses. *New Phytologist*, **226**, 541–554.

**Castro-Rodríguez, R., Abreu, I., Reguera, M., Novoa-Aponte, L., Mijovilovich, A., Escudero, V. *et al.*** (2020) The *Medicago truncatula* yellow Stripe1-Like3 gene is involved in vascular delivery of transition metals to root nodules. *Journal of Experimental Botany*, **71**, 7257–7269.

**Day, D.A. & Smith, P.M.C.** (2021) Iron transport across symbiotic membranes of nitrogen-fixing legumes. *International Journal of Molecular Sciences*, **22**, 432.

**Eide, D., Broderius, M., Fett, J. & Guerinot, M.L.** (1996) A novel iron-regulated metal transporter from plants identified by functional expression in yeast. *Proceedings of the National Academy of Sciences of the United States of America*, **93**, 5624–5628.

**Ensing, B., Buda, F. & Baerends, E.J.** (2003) Fenton-like chemistry in water: oxidation catalysis by Fe(III) and $H_2O_2$. *The Journal of Physical Chemistry A*, **107**, 5722–5731.

**Escudero, V., Abreu, I., del Sastre, E., Tejada-Jiménez, M., Larue, C., Novoa-Aponte, L. *et al.*** (2020) Nicotianamine synthase 2 is required for symbiotic nitrogen fixation in *Medicago truncatula* nodules. *Frontiers in Plant Science*, **10**, 1780.

**Escudero, V., Abreu, I., Tejada-Jimenez, M., Rosa-Núñez, E., Quintana, J., Prieto, R.I. *et al.*** (2020) Ferroportin2 mediates iron import into nodule symbiosomes. *New Phytologist*, **228**, 194–209.

**Gavrin, A., Kaiser, B.N., Geiger, D., Tyerman, S.D., Wen, Z.Y., Bisseling, T. *et al.*** (2014) Adjustment of host cells for accommodation of symbiotic bacteria: vacuole defunctionalization, HOPS suppression, and TIP1g retargeting in *Medicago*. *The Plant Cell*, **26**, 3809–3822.

**Gavrin, A., Loughlin, P.C., Brear, E., Griffith, O.W., Bedon, F., Suter Grotemeyer, M. *et al.*** (2021) Soybean yellow stripe-like 7 is a symbiosome membrane peptide transporter important for nitrogen fixation. *Plant Physiology*, **186**, 581–598.

**Gollhofer, J., Timofeev, R., Lan, P., Schmidt, W. & Buckhout, T.J.** (2014) Vacuolar-iron-transporter1-like proteins mediate iron homeostasis in *Arabidopsis*. *PLoS One*, **9**, e110468.

**González-Guerrero, M., Matthiadis, A., Sáez, Á. & Long, T.A.** (2014) Fixating on metals: new insights into the role of metals in nodulation and symbiotic nitrogen fixation. *Frontiers in Plant Science*, **5**, 45.

**Guerinot, M.L. & Yi, Y.** (1994) Iron: nutritious, noxious, and not readily available. *Plant Physiology*, **104**, 815–820.

**Hakoyama, T., Niimi, K., Yamamoto, T., Isobe, S., Sato, S., Nakamura, Y. *et al.*** (2012) The integral membrane protein SEN1 is required for symbiotic nitrogen fixation in *Lotus japonicus* nodules. *Plant and Cell Physiology*, **53**, 225–236.

**Jian, B., Liu, B., Bi, Y., Hou, W., Wu, C. & Han, T.** (2008) Validation of internal control for gene expression study in soybean by quantitative real-time PCR. *BMC Molecular Biology*, **9**, 59.

**Kaiser, B.N., Moreau, S., Castelli, J., Thomson, R., Lambert, A., Bogliolo, S. *et al.*** (2003) The soybean NRAMP homologue, GmDMT1, is a symbiotic divalent metal transporter capable of ferrous iron transport. *Plant Journal*, **35**, 295–304.

**Ke, X.L., Xiao, H., Peng, Y.Q., Wang, J., Lv, Q. & Wang, X.L.** (2022) Phosphoenolpyruvate reallocation links nitrogen fixation rates to root nodule energy state. *Science*, **378**, 971–977.

**Kim, D., Langmead, B. & Salzberg, S.L.** (2015) HISAT: a fast spliced aligner with low memory requirements. *Nature Methods*, **12**, 357–360.

**Kim, S.A., Punshon, T., Lanzirotti, A., Li, L., Alonso, J.M., Ecker, J.R. *et al.*** (2006) Localization of iron in Arabidopsis seed requires the vacuolar membrane transporter VIT1. *Science*, **314**, 1295–1298.

**Kobayashi, T. & Nishizawa, N.K.** (2012) Iron uptake, translocation, and regulation in higher plants. *Annual Review of Plant Biology*, **63**, 131–152.

**Kryvoruchko, I.S., Routray, P., Sinharoy, S., Torres-Jerez, I., Tejada-Jiménez, M., Finney, L.A. *et al.*** (2018) An iron-activated citrate transporter, MtMATE67, is required for symbiotic nitrogen fixation. *Plant Physiology*, **176**, 2315–2329.

**Kumar, S., Stecher, G., Li, M., Knyaz, C. & Tamura, K.** (2018) MEGA X: molecular evolutionary genetics analysis across computing platforms. *Molecular Biology and Evolution*, **35**, 1547–1549.

**Langfelder, P. & Horvath, S.** (2008) WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics*, **9**, 559.

**Lardi, M. & Pessi, G.** (2018) Functional genomics approaches to studying symbioses between legumes and nitrogen-fixing rhizobia. *HighThroughput*, **7**, 15.

**Lescano, I.** (2020) Determination of ureides content in plant tissues. *BioProtocol*, **10**, e3642.

**LeVier, K., Day, D.A. & Guerinot, M.L.** (1996) Iron uptake by symbiosomes from soybean root nodules. *Plant Physiology*, **111**, 893–900.

**Li, J.**, **Lu, Z.Y.**, **Yang, Y.**, **Hou, J.F.**, **Yuan, L.Y.**, **Chen, G.H.** ***et al.*** (2021) Transcriptome analysis reveals the symbiotic mechanism of *Ustilago esculenta*-induced gall formation of *Zizania latifolia*. *Molecular Plant-Microbe Interactions*, **34**, 168–185.

**Li, J.**, **Ni, Q.Y.**, **He, G.Q.**, **Huang, J.L.**, **Chao, H.Y.**, **Li, S.D.** ***et al.*** (2024) SoyOD: an integrated soybean multi-omics database for mining genes and biological research. *Genomics, Proteomics & Bioinformatics*, **22**, qzae080.

**Li, L.**, **Gao, W.W.**, **Peng, Q.**, **Zhou, B.**, **Kong, Q.H.**, **Ying, Y.H.** ***et al.*** (2018) Two soybean bHLH factors regulate response to iron deficiency. *Journal of Integrative Plant Biology*, **60**, 608–622.

**Li, L.T.**, **Chen, O.S.**, **Ward, D.M.** & **Kaplan, J.** (2001) CCC1 is a transporter that mediates vacuolar iron storage in yeast. *Journal of Biological Chemistry*, **276**, 29515–29519.

**Li, Y.**, **Liu, Q.**, **Zhang, D.X.**, **Zhang, Z.Y.**, **Xu, A.**, **Jiang, Y.L.** ***et al.*** (2024) Metal nutrition and transport in the process of symbiotic nitrogen fixation. *Plant Communications*, **5**, 100829.

**Li, Z.S.**, **Szczypka, M.**, **Lu, Y.P.**, **Thiele, D.J.** & **Rea, P.A.** (1996) The yeast cadmium factor protein (YCF1) is a vacuolar glutathione S-conjugate pump. *Journal of Biological Chemistry*, **271**, 6509–6517.

**Liu, S.**, **Liao, L.L.**, **Nie, M.M.**, **Peng, W.T.**, **Zhang, M.S.**, **Lei, J.N.** ***et al.*** (2020) A VIT-like transporter facilitates iron transport into nodule symbiosomes for nitrogen fixation in soybean. *New Phytologist*, **226**, 1413–1428.

**Liu, Z.**, **Kong, X.**, **Long, Y.**, **Liu, S.**, **Zhang, H.**, **Jia, J.** ***et al.*** (2023) Integrated single-nucleus and spatial transcriptomics captures transitional states in soybean nodule maturation. *Nature Plants*, **9**, 515–524.

**Lönnerdal, B.** (2009) Soybean ferritin: implications for iron status of vegetarians. *The American Journal of Clinical Nutrition*, **89**, 1680S–1685S.

**Love, M.I.**, **Huber, W.** & **Anders, S.** (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology*, **15**, 550.

**Luo, Y.**, **Liu, W.**, **Sun, J.**, **Zhang, Z.R.** & **Yang, W.C.** (2023) Quantitative proteomics reveals key pathways in the symbiotic interface and the likely extracellular property of soybean symbiosome. *Journal of Genetics and Genomics*, **50**, 7–19.

**Mahmood, U.**, **Li, X.**, **Fan, Y.**, **Chang, W.**, **Niu, Y.**, **Li, J.** ***et al.*** (2022) Multi-omics revolution to promote plant breeding efficiency. *Frontiers in Plant Science*, **13**, 1062952.

**Masson-Boivin, C.**, **Giraud, E.**, **Perret, X.** & **Batut, J.** (2009) Establishing nitrogen-fixing symbiosis with legumes: how many rhizobium recipes? *Trends in Microbiology*, **17**, 458–466.

**McAlister, G.C.**, **Huttlin, E.L.**, **Haas, W.**, **Ting, L.**, **Jedrychowski, M.P.**, **Rogers, J.C.** ***et al.*** (2012) Increasing the multiplexing capacity of TMTs using reporter ion isotopologues with isobaric masses. *Analytical Chemistry*, **84**, 7469–7478.

**McLoughlin, F.**, **Augustine, R.C.**, **Marshall, R.S.**, **Li, F.**, **Kirkpatrick, L.D.**, **Otegui, M.S.** ***et al.*** (2018) Maize multi-omics reveal roles for autophagic recycling in proteome remodelling and lipid turnover. *Nature Plants*, **4**, 1056–1070.

**Peiffer, G.A.**, **King, K.E.**, **Severin, A.J.**, **May, G.D.**, **Cianzio, S.R.**, **Lin, S.F.** ***et al.*** (2012) Identification of candidate genes underlying an iron efficiency quantitative trait *locus* in soybean. *Plant Physiology*, **158**, 1745–1754.

**Ren, Z.**, **Zhang, L.**, **Li, H.**, **Yang, M.**, **Wu, X.**, **Hu, R.** ***et al.*** (2025) The BRUTUS iron sensor and E3 ligase facilitates soybean root nodulation by monoubiquitination of NSP1. *Nature Plants*, **11**, 595–611.

**Rodriguez-Haas, B.**, **Finney, L.**, **Vogt, S.**, **Gonzalez-Melendi, P.**, **Imperial, J.** & **Gonzalez-Guerrero, M.** (2013) Iron distribution through the developmental stages of *Medicago truncatula* nodules. *Metallomics*, **5**, 1247–1253.

**Roy, S.**, **Liu, W.**, **Nandety, R.S.**, **Crook, A.**, **Mysore, K.S.**, **Pislariu, C.I.** ***et al.*** (2020) Celebrating 20 years of genetic discoveries in legume nodulation and symbiotic nitrogen fixation. *Plant Cell*, **32**, 15–41.

**Sankari, S.** & **O'Brian, M.R.** (2016) The *Bradyrhizobium japonicum* ferrous iron transporter FeoAB is required for ferric iron utilization in free living aerobic cells and for symbiosis. *Journal of Biological Chemistry*, **291**, 15653–15662.

**Schulte, C.C.M.**, **Borah, K.**, **Wheatley, R.M.**, **Terpolilli, J.J.**, **Saalbach, G.**, **Crang, N.** ***et al.*** (2021) Metabolic control of nitrogen fixation in rhizobium-legume symbioses. *Science Advances*, **7**, eabh2433.

**Slatni, T.**, **Krouma, A.**, **Gouia, H.** & **Abdelly, C.** (2009) Importance of ferric chelate reductase activity and acidification capacity in root nodules of $N_2$-fixing common bean (*Phaseolus vulgaris* L.) subjected to iron deficiency. *Symbiosis*, **47**, 35–42.

**Suganuma, N.**, **Nakamura, Y.**, **Yamamoto, M.**, **Ohta, T.**, **Koiwa, H.**, **Akao, S.** ***et al.*** (2003) The *Lotus japonicus Sen1* gene controls rhizobial differentiation into nitrogen-fixing bacteroids in nodules. *Molecular Genetics and Genomics*, **269**, 312–320.

**Tang, C.X.**, **Robson, A.D.** & **Dilworth, M.J.** (1992) The role of iron in the (Brady) rhizobium legume symbiosis. *Journal of Plant Nutrition*, **15**, 2235–2252.

**Tejada-Jimenez, M.**, **Castro-Rodríguez, R.**, **Kryvoruchko, I.**, **Lucas, M.M.**, **Udvardi, M.**, **Imperial, J.** ***et al.*** (2015) *Medicago truncatula* natural resistance-associated macrophage protein1 is required for iron uptake by rhizobia-infected nodule cells. *Plant Physiology*, **168**, 258–272.

**Trapnell, C.**, **Williams, B.A.**, **Pertea, G.**, **Mortazavi, A.**, **Kwan, G.**, **van Baren, M.J.** ***et al.*** (2010) Transcript assembly and quantification by RNA-seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nature Biotechnology*, **28**, 511–515.

**Walton, J.H.**, **Kontra-Kováts, G.**, **Green, R.T.**, **Domonkos, Á.**, **Horváth, B.**, **Brear, E.M.** ***et al.*** (2020) The *Medicago truncatula* vacuolar iron transporter-like proteins VTL4 and VTL8 deliver iron to symbiotic bacteria at different stages of the infection process. *New Phytologist*, **228**, 651–666.

**Wang, L.**, **Ying, Y.H.**, **Narsai, R.**, **Ye, L.X.**, **Zheng, L.Q.**, **Tian, J.L.** ***et al.*** (2013) Identification of OsbHLH133 as a regulator of iron distribution between roots and shoots in. *Plant, Cell & Environment*, **36**, 224–236.

**Wang, T.**, **Guo, J.**, **Peng, Y.**, **Lyu, X.**, **Liu, B.**, **Sun, S.** ***et al.*** (2021) Light-induced mobile factors from shoots regulate rhizobium-triggered soybean root nodulation. *Science*, **374**, 65–71.

**Wu, X.**, **Wang, Y.**, **Ni, Q.**, **Li, H.**, **Wu, X.**, **Yuan, Z.** ***et al.*** (2023) GmYSL7 controls iron uptake, allocation, and cellular response of nodules in soybean. *Journal of Integrative Plant Biology*, **65**, 167–187.

**Zhang, Y.**, **Xu, Y.H.**, **Yi, H.Y.** & **Gong, J.M.** (2012) Vacuolar membrane transporters OsVIT1 and OsVIT2 modulate iron translocation between flag leaves and seeds in rice. *The Plant Journal*, **72**, 400–410.

**Zhou, M.**, **Li, Y.**, **Yao, X.L.**, **Zhang, J.**, **Liu, S.**, **Cao, H.R.** ***et al.*** (2024) Inorganic nitrogen inhibits symbiotic nitrogen fixation through blocking NRAMP2-mediated iron delivery in soybean nodules. *Nature Communications*, **15**, 8946.