Nature Communications

https://doi.org/10.1038/s41467-026-74688-4

Article in Press

A rhizobium-induced *FT-FD* module locally activates nodule stem cell gene to promote nodulation

Received: 6 March 2025

Accepted: 11 June 2026

Cite this article as: Sun, H., Zhang, X., Ma, C. *et al.* A rhizobium-induced *FT-FD* module locally activates nodule stem cell gene to promote nodulation. *Nat Commun* (2026). https://doi.org/10.1038/s41467-026-74688-4

Huiyu Sun, Xiaomei Zhang, Chao Ma, Penghui Huang, Qi Sun, Guolong Yu, Limin Liu, Mingyang Lu, Lu Han, Kexin Qu, Xianzhong Feng, Yuchen Miao, Qingshan Chen, Jingjuan Yu, Dawei Xin & Yong-Fu Fu

A rhizobium-induced *FT-FD* module locally activates nodule stem cell gene to promote nodulation

Huiyu Sun[1,2,§], Xiaomei Zhang[1,§], Chao Ma[3], Penghui Huang[1], Qi Sun[1], Guolong Yu[3], Limin Liu[3], Mingyang Lu[1], Lu Han[3], Kexin Qu[3], Xianzhong Feng[4], Yuchen Miao[5], Qingshan Chen[3], Jingjuan Yu[2], Dawei Xin[3], Yong-Fu Fu[1,*]

[1] State Key Laboratory of Crop Gene Resources and Breeding, National Key Facility for Crop Gene Resources and Genetic Improvement, MARA Key Laboratory of Soybean Biology (Beijing), Institute of Crop Sciences, Chinese Academy of Agricultural Sciences, Beijing 100081, China

[2] State Key Laboratory of Plant Environmental Resilience, College of Biological Sciences, China Agricultural University, Beijing 100193, China

[3] Key Laboratory of Soybean Biology, Ministry of Education/College of Agriculture, Northeast Agricultural University, Harbin,150030, China

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

[5] State Key Laboratory of Crop Stress Adaptation and Improvement, School of Life Sciences, Henan University, Kaifeng 475004, China

[§] These authors contributed equally.

[*]Corresponding author: fuyongfu@caas.cn

Abstract

Functional divergence of gene homologs enables species to evolve unique physiological processes, such as nitrogen fixation of nodules in legumes. Nodulation is initiated via rhizobia-induced reactivation of root cortical cells into stem cells, yet the molecular mechanism underlying this process has remained elusive. There are eight florigen homologs in the soybean genome, including *GmFT2a* and *GmFT5a*, which are produced in leaves and possibly translocated to roots to promote nodulation. Here, we report the identification of a distinct florigen homolog, *GmFT5b*, and its partner *GmFDL23*, both of which are locally upregulated in the root cortex upon rhizobia infection. This *GmFT5b-GmFDL23* module directly suppresses *GmCLV3-1* expression in the root cortex, thereby alleviating its repression of the stem cell regulator *GmWUS1*, which further drives cortical cell reprogramming into nodule stem cells. Our findings delineate a florigen-mediated pathway that spatiotemporally controls stem cells in cortex to drive nodule formation.

Introduction

Nodulation is a unique biological process through which leguminous plants to fix inorganic nitrogen into organic nitrogen, benefiting not only themselves but also the entire ecosystem. Until now, nearly two hundred genes have been found to form a complex regulatory network that functions in different processes of nodulation across different plant species [1-3]. The early events of signaling communication between plants and rhizobia result in the simultaneous formation of infection threads and reactivation of vascular bundle sheath/pericycle, endodermis, and root cortex cells into stem cells capable of cell division [1, 4]. The growth of infection threads facilitates rhizobia entry into daughter cells derived from *de novo* stem cell divisions. However, the mechanism underling the reactivation of nodule stem cells from root somatic cells has remained unknown until now.

Post-embryonic organogenesis depends on the progressive reprogramming of somatic cells into stem cells [5]. During this process, the stem cell identity gene *WUS* plays a central role [6-9], and high *WUS* activity globally enhances organogenesis [10, 11]. *WUS* not only is a direct target of *CLV3*, but also directly inhibits *CLV3* expression [12-15]; such regulatory dynamics of this loop drives *de novo* stem cell formation to initiate a meristem or anlagen [5, 16].

Florigen is well known as a systemic and mobile signal produced in leaves and then transported not only to shoot apices to initiate flower formation [17, 18] or regulate bud dormancy and burst [19], but also to underground organs to enhance tuberization [20] or nodulation [21]. The whole-genome duplication promotes evolutionary innovation and gene duplications provide an opportunity for the assignment of additional functions to homologous genes, allowing them adapt to unique environmental stresses and developmental programs in animals [22] and plants [23-25]. Legume plants have evolved multiple florigen homolog genes (Supplementary Fig. 1). Soybean is a palaeopolyploid plant. At least eight florigen homologous genes have been retained in the soybean genome (Supplementary Fig. 1) through the evolutionary process and they encode proteins sharing similar flowering-promoting functions [26-28]. However, their biological functions have diverged during soybean evolution, partially due to their different expression patterns [26]. Recently, we found that one soybean florigen homolog, *GmFT5b* (*Glyma.19G108200*), is expressed in roots and nodules (Supplementary Fig. 2), suggesting that it may locally function in these organs. In stark

contrast to *GmFT5b* expression, analyses of expression database indicate that other *GmFT* members are not detected in roots or nodules and are exclusively expressed in aerial tissues. (Supplementary Fig. 2d). Among them, *GmFT2a* (*Glyma.16G150700*) and *GmFT5a* (*Glyma.16G044100*) as systemic signals: both are produced in leaves, and then move upward to shoot apices to initiate floral primordia [28] or downward to roots to regulate nodulation [21]. We here show that GmFT5b, together with its interacting partner GmFDL23, indirectly induces the expression of stem cell gene *GmWUS1* to enhance nodule formation.

Results

GmFT5b is induced in root cortex by rhizobia

To test our hypothesis, we first investigated the expression pattern of *GmFT5b* promoter (*GmFT5b*$_{pro}$), approximately 10 kb of sequence upstream of the ATG start codon was analyzed, as this region may contain potential *cis*-elements - consistent with the finding that *cis*-regulatory elements are present in over 8 kb of sequence upstream of the ATG start codon in the *Arabidopsis* florigen *FT* promoter [29, 30]. Histochemical staining of *GmFT5b*$_{pro}$*:GUS* transgenic plants and hairy roots revealed strong GUS activity in vascular tissues of roots, nodules, and leaves (Supplementary Fig. 2a, 2b). Consistent with this, RT-qPCR detected *GmFT5b* transcripts in all examined organs, which contrasts with publicly available transcriptome data (Supplementary Fig. 2c, 2d). The expression activity of *GmFT5b*$_{pro}$ in roots suggested that *GmFT5b* could be locally produced in roots. Interestingly, following rhizobia infection, GUS signals progressively expanded into the endodermis and cortex (Fig. 1a), potential sites for nodule primordia initiation [1, 4] at 12 hours after infection (HAI). At 5 days after infection (DAI), strong GUS signals were observed in the cortex and particularly pronounced in recently divided cells (or dividing cells) (Fig. 1a, 1b). RT-qPCR results supported that *GmFT5b* expression was induced by rhizobium infection (Fig. 1c). In contrast, no strong GUS activity was detected in the nodule primordium itself (Supplementary Fig. 2b). In nodule latter developmental stages, vascular tissues also exhibited high GUS signals, while the nitrogen fixation zone showed weak signals (Supplementary Fig. 2b). The results suggest that *GmFT5b* functions predominantly during two stages of nodulation: nodule initiation and later nodule developmental stage (vascular tissues).

Next, we generated CRISPR/Cas9-edited mutants of the *GmFT5b* gene, *Gmft5b-1* and *Gmft5b-9*, which contained premature stop codons in exon 1 and 4, respectively (Supplementary Fig. 3a). Both *Gmft5b-1* and *Gmft5b-9* mutants displayed few nodules - especially small nodules - compared with wild type plants (Fig. 1d, e and Supplementary Fig. 3b), while overexpressing *GmFT5b* increased nodule number and rescued the nodulation phenotypes of *Gmft5b-1* (Fig. 1f and Supplementary Fig. 3c-3e), indicating that *GmFT5b* may be involved in both nodule initiation and development. Furthermore, we found that *GmFT5b* may affect rhizobial infection, as the number of infection threads was significantly reduced in *Gmft5b* mutants (Fig. 1g, 1h). Since GmFT2a and GmFT5a proteins act as systemic signals that are transported from shoots to roots to enhance nodulation in soybean [21], we hypothesized that other members of *GmFTs* family may be involved in nodulation (Supplementary Data 1). We then analyzed the function of *FT*-like genes in nodulation using *GmFTs* family RNAi transgenic lines (RNAi-*GmFTs-1* and *-3*) [27], and found that the expression of all *GmFTs* genes was knocked down in two lines (Supplementary Fig. 3f). Silencing the *GmFT* family slightly increased nodule number-especially for small nodules-(Supplementary Fig. 3g-3i), indicating that there may exist antagonistic effects among different members on nodulation. In *Gmft5b-1* and *Gmft5b-9* mutants, the expression of other *GmFTs* genes remained unaltered (Supplementary Fig. 3f), indicating mutation of *GmFT5b* gene does not affect the expression of other *GmFT* genes. To clarify whether *GmFT5b* functioned as a local signal in roots, we performed grafting experiments. The results showed that it was *Gmft5b-1* rootstock rather than *Gmft5b-1* scion that determined the nodule number of the graft chimeras (Fig. 1i,1j and Supplementary Fig. 3j). Taken together, these results demonstrate that *GmFT5b* acts locally in the roots to promote nodulation in soybean.

GmFT5b partner GmFDL23 is induced by rhizobia

FT regulates flower anlagen initiation through its interaction with FD to activate floral identity gene, such as *AP1* [31-33]. We wondered whether *GmFT5b* functioned in soybean nodulation through a similar molecular mechanism. Firstly, we checked if GmFT5b proteins interact with GmFDL23 or GmFDL20 proteins, both of which are highly expressed in roots (Supplementary Fig. 4a, 4b and Supplementary Data 1). GmFT5b proteins distributed throughout cells, while GmFDL23 or GmFDL20 proteins localized in the nucleus (Fig. 2a, Supplementary Fig. 4c, 4d). Their interactions in the nucleus were verified by BiFC

(bimolecular fluorescence complementation), Y2H (yeast-two-hybrid), and Co-IP (co-immunoprecipitation) experiments (Fig. 2b-2d and Supplementary Fig. 4e-4h). *GmFDL23* also showed rhizobium-induced expression patterns in roots and nodules (Fig. 1c), but *GmFDL20* expression was decreased after rhizobium infection, although it increased at DAI 5 (Supplementary Fig. 4i), indicating *GmFDL20* is not involved in nodule initiation. Consistent with the RT-qPCR data, histochemical GUS analysis revealed that *GmFDL23* was only weakly expressed in the root stele under non-infected conditions. Upon infection, however, *GmFDL23* was strongly upregulated in the stele and was additionally detected in some pericycle or endodermis cells (Fig. 2e, 2f). Strong GUS signals were also present in unmature nodules (Supplementary Fig. 4j). However, unlike *GmFT5b*, *GmFDL23* may not participate in rhizobial infection, as the number of infection threads was not changed in the *Gmfdl23* mutants compared to wild type plants (Supplementary Fig. 5). These results suggest that GmFDL23 may collaborate with GmFT5b to initiate of cell division during nodule formation.

To test this hypothesis, we first analyzed nodule phenotype of *GmFDL23*-knockout hairy roots. Phenotypic analysis of nodules demonstrated that loss of function of *GmFDL23* led to the reduced number of nodules, while overexpression of *GmFDL23* increased nodule number (Fig. 2g and Supplementary Fig. 6, 7a, 7b). Additionally, the expression of early nodulation marker genes in both *Gmft5b* and *Gmfdl23* mutants was similarly decreased (Fig. 2h and Supplementary Data 1). Stem cells and cell division genes (*GmSCR1*, *GmSHR4*, *GmCYCD6;1-6*, *GmPLT2*) [34, 35], indispensable regulators for primordium initiation, also displayed lower expression levels in *Gmft5b* mutants than in wild-type plants (Supplementary Fig. 8 and Supplementary Data 1). To elucidate the genetic relationship between *GmFT5b* and *GmFDL23*, we generated genetically modified hairy roots in which *GmFDL23* was either knocked out or constitutively overexpressed in the *Gmft5b-1* background. Compared with single mutant roots, the double mutant roots exhibited a similar number of nodules but slightly reduced density (Fig. 2i, 2j and Supplementary Fig.7c). The *GmFT5b* mutation partially rescued the hypernodulation phenotype of *GmFDL23*-overexpressing roots, restoring nodulation to near wild-type levels (Fig. 2i, 2j and Supplementary Fig. 7d). These results demonstrate that *GmFDL23* contributes to nodule formation partially dependent on *GmFT5b*. However, we found that *GmFT* and *GmFDL* did not significantly activate *GmAP1* promoter (Supplementary Fig. 9a

and Supplementary Data 1), indicating the *GmFT-GmFDL* module may regulate nodulation through another signaling pathway, independent of *GmAP1* gene.

GmFT5b-GmFDL23 module inhibits *GmCLV3-1* expression in root cortex

The results above indicated that the *GmFT5b-GmFDL23* module is involved in nodule formation by regulating the expression of stem cells. The WUS-CLV3 loop plays pivotal roles in *de novo* stem cell formation [6-9]. Thus, we aimed to investigate the WUS-CLV3 loop when screening targets of the FT-FD module. Bioinformatic analysis identified binding motifs of the bZIP transcription factor (FD) in the promoters of *GmCLV3-1* and *GmWUS1* (AGCT [37], Fig. 3a, Supplementary Fig. 9b, and Supplementary Data 1). Therefore, we hypothesized that *GmFT5b* and *GmFDL* directly act on a WUS-CLV3 feedback loop. ChIP-qPCR confirmed that GmFDL23 enriched specific fragments of the *GmCLV3-1* promoter, but not of *GmWUS1* (Fig. 3a, Supplementary Fig. 9b), and EMSA test confirmed that GmFDL23 proteins physically interacted with the corresponding fragments of the *GmCLV3-1* promoter (Fig. 3b). *GmFDL23* and *GmFT5b* synergistically inhibited the promoter activity of *GmCLV3-1* (Fig. 3c). In the*Gmft5b* mutants, *GmCLV3-1*$_{pro}$ promoter activities were significantly enhanced (Fig. 3d). Furthermore, the mutation of either *GmFDL23* or *GmFT5b* genes resulted in increased *GmCLV3-1* expression (Fig. 3e), while overexpression of either *GmFT5b* or *GmFDL23* repressed *GmCLV3-1* expression (Fig. 3f). However, we found *GmFT5b* did not inhibit *GmCLV3-2* expression (Supplementary Fig. 9d and Supplementary Data 1). To further clarify how the interaction between GmFDL23 and GmFT5b regulates *GmCLV3-1 in vivo*, we used *Gmft5b* mutants to examine the binding ability and regulatory function of GmFDL23 on the *GmCLV3-1* gene. The results showed that the *GmFT5b* mutation did not disrupt the binding of GmFDL23 to the *GmCLV3-1* promoter but significantly attenuated transcriptional repression (Fig. 3a and 3g). This suggests that GmFT5b may not act as a DNA-binding factor but rather functions as a transcriptional co-effector with *GmFDL23* as *Arabidopsis FT* does in the *FT-FD* module [38].

To investigate whether *GmCLV3-1* participates in regulation of nodulation, we analyzed the spatiotemporal expression of this gene via histochemical staining using a promoter-GUS reporter system. Rhizobium infection strongly inhibited the activity of the *GmCLV3-1* promoter in roots: at 3 DAI, GUS

signals in root cortex were barely detectable (Fig. 3h). Consistent with this observation, RT-qPCR results showed that rhizobium infection significantly reduced the transcriptional level of *GmCLV3-1* (Fig. 1c). Overexpression of *GmCLV3-1* led to a decrease in nodule number (Fig 3I and Supplementary Fig. 10b). In contrast, loss of *GmCLV3-1* function not only increased nodule number and nodule density per plant, but also rescued the nodulation-related phenotypes of the *Gmft5b* mutant (Fig 3i, 4j and Supplementary Fig. 10a, 11). Additionally, the *Gmfdl23 Gmclv3-1* double mutant exhibited a hypernodulation phenotype (Fig 4i, 4j and Supplementary Fig. 10c, 12). Collectively, these results suggest that the *GmFT5b-GmFDL23* module promotes nodulation by directly suppressing *GmCLV3-1* activity in the root cortex.

GmWUS1 is released from _GmCLV3-1_ repression to trigger nodulation

Further evidence showed that *GmCLV3-1* inhibited *GmWUS1* activity, similar to its homologs do in shoot apices [12-15]. This is because overexpression of *GmCLV3-1* suppressed the expression of *GmWUS1* gene and reduced the activity of its promoter (Fig. 4a and 4b). Although GmFDL23 proteins do not directly bind to the fragments of the *GmWUS1* promoter in planta (Supplementary Fig. 9b) and both GmFT5b and GmFDL23 had no effect on $GmWUS1_{pro}$ (Supplementary Fig. 9c), *GmWUS1* expression was down-regulated when *GmFT5b* or *GmFDL23* genes were knocked out (Fig. 4c). During early nodule primordium formation, the tissue-specific expression of the *GmWUS1* promoter was induced by rhizobia induction in pericycle and endodermis cells (Fig. 4d and 4e) - a pattern consistent with the spatial expression of the *GmFT5b-GmFDL23* module (Fig. 1a, 1b, 2e, and 2f). Furthermore, analysis of nodule phenotypes confirmed *GmWUS1* acted as a positive regulator of nodulation: mutation of *GmWUS1* inhibited nodulation, while overexpression of *GmWUS1* enhanced nodulation (Fig. 4f, and Supplementary Fig. 10d, 13).

Then, to investigate whether the *GmFT5b–GmFDL23* module regulates nodulation through *GmWUS1*, we first overexpressed *GmWUS1* in the *Gmft5b-1* background. The *Gmft5b-1 35S:GmWUS1* hairy roots exhibited a hypernodulation phenotype (Fig. 4g, 4h and Supplementary Fig. 10e), confirming the genetic epistasis of *GmWUS1* to the *GmFT5b–GmFDL23* module. Additionally, using a multiplex sgRNA system, we generated *Gmclv3-1 Gmwus1* double mutant roots. Knockout of *GmWUS1* partially rescued the

hypernodulation phenotype of the *Gmclv3-1*, restoring nodule number and density to near wild-type levels (Fig. 4i, 4j and Supplementary Fig. 10c, 14). Soybean nodules are determinate nodules, which are lack of a meristem-like tissue [4]. However, we found that the *GmWUS1* promoter was active in vascular tissues of nodule primordia (Supplementary Fig.15), suggesting that *GmWUS1* functions in nodule development.

Previous reports showed *WUSCHEL-RELATEDHOMEOBOX 5* (*WOX5*), a homolog of *GmWUS1*, might control cell proliferation in nodule meristems [39, 40] (Supplementary Data 1). In our experiment, although *GmFT5b* enhanced the activity of *GmWOX5* promoter, the *GmFT5b-GmFDL23* complex did not significantly regulate *GmWOX5*. (Fig. 4k). Mutation of *GmFT5b* reduced *GmWOX5* expression (Fig. 4l), while mutation of *GmFDL23* had no significant effect on *GmWOX5* expression (Fig. 4m), suggesting that *GmFDL23* is not required for the expression of *GmWOX5* gene. These results suggest that both *GmWUS1* and *GmWOX5* may play roles in nodulation through different pathways. However, *GmWUS1*, but not *GmWOX5*, is dependent on the *GmFT5b-GmFDL23* module.

The type III secretion system (T3SS) is a conserved apparatus employed by rhizobia [41, 42] to deliver the type III effectors (T3Es), also named Nop (nodulation outer protein), into the host cell to activate the symbiosis. ErnA, a T3E from *Bradyrhizobium* ORS3257 that targets the plant nucleus, is responsible for nodulation of nodulate *Aeschynomene indica* (Supplementary Data 1). Ectopic expression of *ErnA* induces formation of nodule-like structures in *A. indica* roots [43]. It is hypothesized that ErnA most likely triggers nodulation by targeting to key regulators in the nodulation signaling pathway, such as NIN and NF-Y [44]. As previously reported, ErnA is derived from *Bradyrhizobium* strain ORS3257, which does not induce nodulation in soybean; however, when another T3 effector, Bel2-5, derived from the strain *B. elkanii* USDA61 is introduced into ORS3257, the resulting strain induces nodulation in soybean [45]. Given the above results, we wondered whether the *GmFT5b–GmFDL23* module could trigger nodulation in the absence of rhizobia. Thus, we overexpressed *GmFT5b*, *GmFDL23*, and *GmWUS1* individually and found nodule-like structures on their hairy roots under rhizobia-free conditions (Supplementary Fig. 16a-c). Similarly, *Gmclv3-1* mutant roots also exhibited comparable spontaneous organogenesis (Supplementary Fig. 16a-c). Spontaneous nodules exhibited aberrant morphology and failed to form well-differentiated

vascular tissues, as compared with normal nodules. These results further confirm that *GmFT5b*, *GmFDL23*, *GmWUS1*, and *GmCLV3-1* genes all participate in regulation of the process of nodule organogenesis. Then, we examined whether ErnA directly affects the *GmFT5b-GmFDL23* module after rhizobia HH103 infection. Interestingly, overexpression of the soybean codon-optimized *ErnA* gene enhanced soybean nodulation partially dependent on *GmFT5b* (Fig. 5a and Supplementary Fig. 16d, 16e). Additionally, overexpression of *ErnA* upregulated the expression of *GmFT5b*, *GmFDL23* and nodulation marker genes, with the additional finding that its upregulation of the marker genes occurred in a partially *GmFT5b*-dependent manner. (Fig. 5b).

To further investigate the regulation of the *GmFT5b–GmFDL23* module by type III effectors, we inoculated hairy roots with the *NopL*-deficient mutant of HH103 (*nopL*) [46] (we choose *NopL* because its protein shows the highest similarity to ErnA, Supplementary Fig. 17, and Supplementary Data 1 and 2). We first examined the expression of early nodulation-related genes in hairy roots inoculated with *nopL* and found that their expression was downregulated, except for *GmNFR1* (Fig. 5c), consistent with previous reports [46]. We also observed that disruption of *NopL* suppressed the expression of *GmFT5b*, *GmFDL23*, and *GmWUS1*, while upregulating *GmCLV3-1* (Fig. 5d). Moreover, histochemical assays revealed that after inoculation with the *nopL*, $GmCLV3\text{-}1_{pro}$ activity was increased in the root cortex, whereas $GmWUS1_{pro}$ activity was reduced (Fig. 5e). Finally, we found that either overexpression of *GmFT5b*, *GmFDL23* and *GmWUS1*, or knockout of *GmCLV3-1*, could partially rescue the reduced-nodule phenotype induced by inoculation with the *nopL* mutant (Fig. 5f and 5g). The results suggest that rhizobium T3Es probably trigger nodulation through activating the *GmFT5b-GmFDL23* module.

Discussion

Based on our results, we proposed a working model (Fig. 5h). Under uninfected conditions, the expression of *GmFT5b* and *GmFDL23* is restricted to stele cells in soybean roots, while *GmCLV3* is predominantly expressed in root cortical cells, and thereby inhibiting the expression of the stem cell gene GmWUS1 in these cells. The analysis results from two single-cell sequencing databases [47, 48] are both consistent with

the expression patterns of the model-related genes identified in this study (Supplementary Fig. 18). Upon rhizobial infection, rhizobial signals (possibly T3E and other factors) activate the expression of *GmFT5b* and *GmFDL23* in root cortical cells. These two proteins then locally inhibit *GmCLV3* expression, which in turn releases the repression of *GmWUS1* by *GmCLV3-1*. Increased GmWUS1 activity triggers *de-novo* stem cell formation, thereby initiating nodule primordia. Our work elucidates a signaling pathway for the *de novo* activation of stem cell genes in the root cortex during soybean nodulation and supports the conserved function of WUS (re-activating somatic cells) in soybean nodulation as observed in other developmental processes [6-11]. This model is not involved in the formation of infection thread because *GmFDL23* and *GmCLV3-1* does not participate in this process (Supplementary Fig. 5). From this perspective, there is a gap between rhizobia and the *GmFT5b-GmFDL23* module. We also uncovered a distinct working mode of action for florigens: a local function in roots, in contrast to their canonical role as systemic signals [21]. Thus, florigens participate in nodulation through at least in two distinct processes, acting as different signals: GmFT5b (a local signal, this study) promotes stem cell formation, while GmFT2a and GmFT5a (systemic signals from leaves, previous reports [21]) regulate nodule primordium initiation or development. This pathway for *de novo* stem cell formation in soybean roots may be independent of autoregulation of nodulation (AON) pathway [49, 50], as the function of the *GmFT5b-GmFDL23* module mediates nodulation without requiring shoot-derived signals (graft experiments, Fig.1iand1j) and we also did not detect an interaction between GmCLV3-1 and GmNARK (GmCLV1-1) proteins (Supplementary Fig. 19). This signaling pathway may also explain the mechanism of *de novo* organ formation in plants, such as flowers, tubes, onion bulbs, and sugar beets, because FT-FD [18, 20, 28, 31, 32, 51, 52] and WUS-CLV3 [12-14] modules share conserved molecular functions across plant kingdom.

In our study, we revealed a pathway for the *de novo* activation of stem cell genes in the root cortex of soybean, mediated by the *GmFT5b-GmFDL23* module and induced by rhizobia. This pathway is likely an early and essential step for various types of *de novo* organogenesis. However, several important questions remain unanswered. First, we did not identify the direct activators of the *GmFT5b-GmFDL23* module in soybean of rhizobia. Second, *de novo* stem cell formation is a progressive process that begin with a transient regulatory network in a small group of cells [5], but we failed to identify this specific cell

population or capture precise spatiotemporal-dynamics of *GmWUS1* activity in the root cortex, as has been happen in the apex [12-15]. Third, we do not know the role of *GmWUS1* in nodule cell proliferation, which is different from stem cell proliferation in shoot apices [4]. Fourth, it would be interesting to explore the relationship between *GmWUS1* and its homolog *WOX5* in nodulation. Additionally, *CLV3* is not the only gene that regulates *WUS* expression [53]; other signals also modulate *WUS* activity. For example, nitrate, an important signal for nodulation, can modulate *WUS* expression in shoot meristems via cytokinins [10]. Furthermore, a WUS-independent pathway has been identified that regulates stem cell proliferation [54-56], and the effect of interaction between *GmFT5b* and its two homologs (*GmFT5a* and *GmFT2a*) on nodulation is unknown. These gaps indicate the mechanism of nodulation is complex and requires further investigation.

Methods

Plant materials and growth conditions

The genetic background of *Gmft5b* mutants and *GmFT5bPro:GUS* transgenic plants is soybean (*Glycine max* (Merr.) L.) Williams 82 (WS82), and the genetic background of *GmFTs* RNAi lines is soybean Tianlong 1 (TL1). Hairy root transformation experiment was performed with WS82 and relative transgenic plants. All materials were grown in a growth chamber set to 25°C under a 16 h light / 8 h dark cycle and a light intensity of 400 μmol m^{-2} s^{-1}, using vermiculite as the growth medium. During growth, nutrient solution was applied as 1 mM KNO_3, 1 mM $CaCl_2$, 10 μM Fe-citrate, 0.25 μM $MgSO_4$, 0.25 μM K_2SO_4, 1 μM $MnSO_4$, 2 μM H_3BO_3, 0.5 μM $ZnSO_4$, 0.2 μM $CuSO_4$, 0.1 μM $CoSO_4$, 0.1 μM Na_2MoO_4, 5 μM KH_2PO_4. Rhizobia (*Sinorhizobium fredii* HH103) were applied to soils after three days of transplanting. HH103 was previously cultured in TY liquid medium (3 g/L yeast extract, 5 g/L tryptone, 0.4 g/L $CaCl_2$) at 28°C for 2-3 days.

Vector construction

For promoter activity and transcriptional activity analysis, the upstream sequences of the start codon ATG (serving as the gene promoters) were cloned from WS82 into the Fu76 entry vector [57] via enzyme digestion and ligation. The length of the cloned promoters were as follows: *GmFT5b* (10 kb), *GmFDL23*

(3 kb), *GmAP1-2* (2.6 kb), *GmAP1-3* (2.5 kb), *GmCLV3-1* (3 kb), *GmCLV3-2* (3 kb), *GmWUS1* (3 kb), *GmWOX5* (3 kb). Subsequently, these promoter fragments, together with the Fu79-GUS or Fu79-LUC entry vectors, were transferred into the pSoy10 binary vector via LR recombination reactions.

To construct vectors for overexpressing target genes in hairy roots and analyzing transcriptional activity in *Nicotiana benthamiana* leaves, the coding sequences (CDS) of *GmFT5b*, *GmFDL9*, *GmFDL16*, *GmFDL20*, *GmFDL23*, *GmCLV3-1*, and *GmWUS1* were cloned from WS82, while the CDS of *ErnA* with soybean optimized codons was synthesized by Genecreate company. All CDS sequences were ligated into the Fu28 vector through enzymatic digestion [57]. Subsequently, these CDS fragments, together with the Fu76-$35S_{pro}$ vector were transferred to the pSoy10 binary vector via LR recombination reactions.

For gene editing, two sgRNA sequences were designed using CRISPR-P 2.0 software (http://crispr.hzau.edu.cn/CRISPR2/) and ligated into the Fu79 entry vector containing the U6 promoter at the *Bsp*Q I and *Bsa* I sites. These sgRNA-containing Fu79 vectors, together with the Fu76-Cas9 vector, were then transferred to the pSoy10 or pSoy13 vector via LR recombination reactions.

Hairy root transformation

WS82 and different mutants were used for hairy root transformation. The constructs were transformed into the *Agrobacterium rhizogenesis* strain K599 and cultured in 50 mL of liquid LB medium containing kanamycin (selection marker) until OD_{600} = 0.8. Root transformation was performed following a previously reported method [58]. Briefly, sterilized soybean seeds were soaked and germinated, and the elongated radicles were excised. The explants were infected in the resuspension solution for 30 minutes and then transferred to co-cultivation medium (1/10× Gamborg B5 salts, 30 g/L sucrose, 3.9 g/L MES, 4.25 g/L agar, pH 5.4, supplemented with 400 mg/L cysteine, 154.2 mg/L dithiothrietol, and 40 mg/L acetosyringone after sterilization). The explants were maintained in the dark for 3 days and then transferred to hairy root induction medium under 12 h light / 12 h dark cycle. After 3 days, the explants were rolled up in moistened germination paper and cultured for an additional 10 days until roots emerged. Transgenic hairy roots were selected by detecting the DsRed signal under a fluorescent stereomicroscope,

and non-transgenic hairy roots were removed. Composite plants with positive hairy roots were then transplanted into pots (7.5×7.5 cm) containing vermiculite with nutrient solution and grown in a growth chamber at 25°C under a 16 h light / 8 h dark cycle. Five days later, the plants were inoculated with a suspension of rhizobia strain HH103 suspension (OD_{600} = 0.1) for nodule phenotype analysis and GUS histochemical staining.

GUS histochemical staining and microscopic observation

Transgenic hairy roots and nodules were harvested at different stages after rhizobial inoculation, along with leaves of stable transgenic lines. These samples were immersed in GUS staining solution (50 mM sodium phosphate buffer, pH 7.0, 0.2% Triton-X-100, 5 mM $K_4Fe(CN)_6$, 5 mM $K_3Fe(CN)_6$, and 1~2 mM X-gluc). The samples were vacuum-infiltrated for 15 minutes and then incubated in dark at 37°C for 12 hours. Subsequently, the samples were rinsed three times with 70% ethanol and observed under an Olympus stereo microscope (SZX7). For sectioning, the samples were embedded in 4% agarose and cut into 80-100 μm slices using a Leica vibrating microtome (VT1000 S). Finally, images were taken using an Olympus optical microscope (IX71).

The proportion of hairy roots with GUS signals in the stele was as follows: *GmFT5b*$_{pro}$*:GUS* (10/10), *GmFDL23*$_{pro}$*:GUS* (7/10), *GmCLV3-1*$_{pro}$*:GUS* (10/10), and *GmWUS1*$_{pro}$*:GUS* (8/10). Concurrently, the empty vector (EV) control was closely examined, and no GUS signals were detected in the EV control (0/6).

DNA extraction

Plant or hairy root samples were harvested for DNA extraction using the SDS method [59]. Briefly, plant tissues were ground in lysis buffer (2% sodium dodecyl sulfate (SDS), 50 mM Tris-HCl (pH 8.0), 10 mM ethylenediaminetetraacetic acid (EDTA), 200 mM NaCl). The homogenate was incubated at 65°C for 10 minutes, then placed on ice, and one-half volume of 3 M potassium acetate (KAc) was added. After incubation on ice, DNA was precipitated with an equal volume of isopropanol and washed with 70% ethanol. The extracted DNA was further used for PCR amplification.

RNA extraction and RT-qPCR

The plant tissues were ground into a fine powder in liquid nitrogen, and total RNA was extracted using HiPure Total RNA Mini Kit (Magen). Genomic DNA contamination was removed, and first strand of cDNA was synthesized using the PrimeScript RT Reagent Kit with gDNA Eraser (TaKaRa). RT-qPCR was performed using the ChamQTM Universal SYBR qPCR Master Mix (Vazyme) on a qPCR instrument. Each RT-qPCR experiment included at least three technical replicates. *GmUKN1* was used as the internal reference gene [60]. The RT-qPCR primers used for RT-qPCR are listed in Supplementary Data 3.

Grafting experiments

Seeds of WS82 and Gmft5b-1 were sown in vermiculite with nutrient solution for germination (16 h light/8 h dark, 25℃). The hypocotyls of 7-day-old seedlings were split into two halves, and the scion was then gently tapered and grafted onto the reciprocal rootstock. Grafting clips and sealing film were used to secure the graft junction. Grafted plants were cultured in dark, warm, and humid environment for 5 days. Successfully grafted plants were inoculated with rhizobium strain HH103, and nodule phenotypes were evaluated at 21 days after inoculation.

Subcellular localization and bimolecular fluorescence complementation (BiFC)

The coding sequences (CDSs) of *GmFT5b*, *GmFDL20*, and *GmFDL23* were cloned into the pGWC entry vector [61], and then recombined into the pGWB5-GFP expression vector for subcellular localization experiments using LR recombination reactions [61]. Additionally, pGWC-*GmFDL20* and pGWC-*GmFDL23* were recombined into the pEARLYGAYE201-YN vector by LR recombination reactions, while pGWC-*GmFT5b* was recombined into the pEARLYGAYE202-YC vector for BiFC experiments. These constructs were transformed into *Agrobacterium tumefaciens* strain EHA105. When the bacterial culture reached an OD_{600} of 1.0, the cells were harvested and resuspended in infiltration buffer (10 mM MES, 150 μM acetosyringone, 10 mM $MgCl_2$). The bacterial suspension was used to infiltrate *Nicotiana benthamiana* leaves. Fluorescent signals were observed and recorded using a Zeiss LSM980 confocal laser scanning microscope.

Co-immunoprecipitation

Via LR recombination reactions, pGWC-*GmFDL20* and pGWC-*GmFDL23* were recombined into the pGWB20-Myc vector. The resulting binary vectors were transformed into *Agrobacterium tumefaciens* strain EHA105. These strains were co-infiltrated together with EHA105 containing pGWB5-*GmFT5b:GFP* constructs into *Nicotiana benthamiana* leaves. Two days after infiltration, proteins were extracted from the *Nicotiana benthamiana* leaves. Immunoprecipitation was performed at 4°C for 1 hour using anti-Myc magnetic beads (M047, Anti-Myc-tag mAb-Magnetic Agarose, MBL). Western blot analysis was then performed using anti-Myc (M192-3, MBL) and anti-GFP (M048-3, MBL) antibodies.

Yeast two-hybrid

The coding sequences (CDSs) were cloned into the pGADT7/Pxgy17 vector and the pGBKT7/ Pxgy18 vector via restriction enzyme digestion and ligation. These constructs were co-transformed into *Saccharomyces cerevisiae* strain AH109. Initially, yeasts were grown on SD-Leu/-Trp medium. After colony formation, the colonies were transferred to SD-Leu/-Trp/-His/-Ade medium and cultured for 3-4 days. The interaction between the proteins was determined based on the growth of yeasts on the selective conditions.

Transcriptional activity assay

pSoy10-*AP1-2*$_{pro}$*:LUC*, pSoy10-*AP1-3*$_{pro}$*:LUC*, pSoy10-*GmCLV3-1*$_{pro}$*:LUC*, pSoy10-*GmWUS1*$_{pro}$*:LUC*, and pSoy10-*GmWOX5*$_{pro}$*:LUC* were used as reporter vectors. pSoy10-*35S:GmFT5b*, pSoy10-*35S:GmFDL9*, pSoy10-*35S:GmFDL16*, pSoy10-*35S:GmFDL20*, pSoy10-*35S:GmFDL23*, and pSoy10-*35S:GmCLV3-1* were used as effector vectors. All constructs were transformed into *Agrobacterium tumefaciens* strain EHA105 and co-infiltrated into *Nicotiana benthamiana* leaves. Two days later, the leaves were sprayed with 3 mM D-luciferin substrate (Provenand Published, LUCK-1G). Luciferase signals were detected using a Tanon 5200 imaging system, and signal intensity was quantified using Image J.

ChIP-qPCR assay

Chromatin Immunoprecipitation (ChIP) was performed following the previously reported method [62]. Briefly, four grams of *35S:GmFDL23:GFP* transgenic roots were harvested 3 days after rhizobium inoculation and minced. The minced tissue was crosslinked with 1% formaldehyde for 10 minutes, followed by quenching with 0.4 M glycine to terminate the crosslinking reaction. The tissues were then ground into a fine powder in liquid nitrogen for chromatin extraction. Chromatin was sonicated using a Bioruptor UCD-200 to generate DNA fragments of 200 bp to 1000 bp in length. The sonicated chromatin was incubated with anti-GFP antibodies and IgG (as a negative control) at 4°C overnight, followed by immunoprecipitation using a 1:1 volume mixture of protein A and protein G beads. Cross-links were reversed by incubation with 215 mM NaCl and 1% SDS at 65°C for 4 hours. DNA was purified using the QIAquick PCR Purification Kit (QIAGEN). Quantitative PCR (qPCR) was then performed to analyze the precipitated DNA, and the primers used listed in Supplementary Data 3.

EMSA

The full-length coding sequence (CDS) of *GmFDL23* was cloned into the pET28a-His expression vector using *Bam*H I and *Sal* I restriction sites. The construct was then transformed into *Escherichia coli* BL21 competent cells for protein expression and purification. Biotin-labeled and unlabeled probes were synthesized by BGI and their sequences are listed in Supplementary Data 3. Electrophoretic Mobility Shift Assays (EMSAs) were performed following the protocol of the Light-Shift Chemiluminescent EMSA Kit (Thermo Scientific). For DNA binding, purified recombinant protein was mixed with 50 fMol of the labeled probe in a reaction buffer (2.5% (v/v) glycerol, 5 mM $MgCl_2$, 50 ng/mL poly (dI, dC), 0.05% (v/v) Nonidet P-40). For competition assay, at least 50-folds of unlabeled or mutated probe were added to the reaction mixture. The samples were incubated at room temperature for 20 minutes and then separated by non-denaturing 8% polyacrylamide gel electrophoresis (PAGE).

Identification of CRISPR/Cas9 hairy roots

Positive CRISPR/Cas9-edited hairy roots were identified following previously reported methods [63, 64]. Briefly, PCR amplification was performed using genomic DNA from positive transgenic hairy root as a

template, with primers designed within 400 bp of both sides of the two sgRNA target sites. The amplified DNA fragments were then sequenced. Heterozygous biallelic mutations - characterized by double peaks at the two target sites - were analyzed using CRISPR-GE (http://skl.scau.edu.cn/) to decode the mutation types. All sequencing data of CRISPR/Cas9-edited hairy roots are provided in Supplementary Data 4.

Statistical analysis

All statistical analyses were performed using GraphPad Prism 9.0.0 (GraphPad Software). Data were visualized as figures using the same software. Individual data points were represented by dots in the figures. Statistical significance was examined using One-way ANOVA with Tukey's post hoc test or two-tailed Student's *t*-test. A p-values < 0.05 was considered as statistically significant.

Data availability

All materials generated in this study are available upon request. Data supporting the findings of this work are available within the paper and its Supplementary Information files. A reporting summary for this Article is available as a Supplementary Information file. Source data are provided with this paper.

References

1. Roy S*, et al.* Celebrating 20 Years of Genetic Discoveries in Legume Nodulation and Symbiotic Nitrogen Fixation. *Plant Cell* **32**, 15-41 (2020).
2. Yang J, Lan L, Jin Y, Yu N, Wang D, Wang E. Mechanisms underlying legume-rhizobium symbioses. *Journal of Integrative Plant Biology* **64**, 244-267 (2022).
3. Desbrosses GJ, Stougaard J. Root nodulation: a paradigm for how plant-microbe symbiosis influences host developmental pathways. *Cell Host Microbe* **10**, 348-358 (2011).
4. Hirsch AM. Developmental biology of legume nodulation. *New Phytol* **122**, 211-237 (1992).
5. Nicolas A, Laufs P. Meristem Initiation and *de novo* Stem Cell Formation. *Frontiers in Plant Science* **13**, 891228 (2022).
6. Shimotohno A. Illuminating the molecular mechanisms underlying shoot apical meristem

homeostasis in plants. *Plant Biotechnol* **39**, 19-28 (2022).

7. Lopes FL, Galvan-Ampudia C, Landrein B. WUSCHEL in the shoot apical meristem: old player, new tricks. *J Exp Bot* **72**, 1527-1535 (2021).
8. Lindsay P, Swentowsky KW, Jackson D. Cultivating potential: Harnessing plant stem cells for agricultural crop improvement. *Molecular Plant* **17**, 50-74 (2024).
9. van der Graaff E, Laux T, Rensing SA. The WUS homeobox-containing (WOX) protein family. *Genome Biol* **10**, 248 (2009).
10. Landrein B*, et al.* Nitrate modulates stem cell dynamics in Arabidopsis shoot meristems through cytokinins. *Proc Natl Acad Sci U S A* **115**, 1382-1387 (2018).
11. Schoof H, Lenhard M, Haecker A, Mayer KF, Jurgens G, Laux T. The stem cell population of Arabidopsis shoot meristems in maintained by a regulatory loop between the CLAVATA and WUSCHEL genes. *Cell* **100**, 635-644 (2000).
12. Muller R, Borghi L, Kwiatkowska D, Laufs P, Simon R. Dynamic and compensatory responses of Arabidopsis shoot and floral meristems to CLV3 signaling. *Plant Cell* **18**, 1188-1198 (2006).
13. Hirakawa Y. CLAVATA3, a plant peptide controlling stem cell fate in the meristem. *Peptides* **142**, 170579 (2021).
14. Fletcher JC, Brand U, Running MP, Simon R, Meyerowitz EM. Signaling of cell fate decisions by CLAVATA3 in Arabidopsis shoot meristems. *Science* **283**, 1911-1914 (1999).
15. Brand U, Fletcher JC, Hobe M, Meyerowitz EM, Simon R. Dependence of stem cell fate in Arabidopsis on a feedback loop regulated by CLV3 activity. *Science* **289**, 617-619 (2000).
16. Birnbaum KD, Roudier F. Epigenetic memory and cell fate reprogramming in plants. *Regeneration* **4**, 15-20 (2017).
17. Tamaki S, Matsuo S, Wong HL, Yokoi S, Shimamoto K. Hd3a protein is a mobile flowering signal in rice. *Science* **316**, 1033-1036 (2007).
18. Corbesier L*, et al.* FT protein movement contributes to long-distance signaling in floral induction of Arabidopsis. *Science* **316**, 1030-1033 (2007).
19. Wigge PA. FT, a mobile developmental signal in plants. *Curr Biol* **21**, R374-378 (2011).
20. Navarro C*, et al.* Control of flowering and storage organ formation in potato by *FLOWERING*

LOCUS T. *Nature* **478**, 119-122 (2011).

21. Wang T*, et al.* Light-induced mobile factors from shoots regulate rhizobium-triggered soybean root nodulation. *Science* **374**, 65-71 (2021).
22. Storz JF, Opazo JC, Hoffmann FG. Gene duplication, genome duplication, and the functional diversification of vertebrate globins. *Mol Phylogenet Evol* **66**, 469-478 (2013).
23. Ohno S, Wolf U, Atkin NB. Evolution from fish to mammals by gene duplication. *Hereditas* **59**, 169-187 (1968).
24. de Bruijn S, Angenent GC, Kaufmann K. Plant 'evo-devo' goes genomic: from candidate genes to regulatory networks. *Trends Plant Sci* **17**, 441-447 (2012).
25. Schmutz J*, et al.* Genome sequence of the palaeopolyploid soybean. *Nature* **463**, 178-183 (2010).
26. Fan C*, et al.* Conserved CO-FT regulons contribute to the photoperiod flowering control in soybean. *BMC Plant Biol* **14**, 9 (2014).
27. Xu K*, et al.* Fine-Tuning Florigen Increases Field Yield Through Improving Photosynthesis in Soybean. *Frontiers in Plant Science* **12**, 710754 (2021).
28. Kong F*, et al.* Two coordinately regulated homologs of FLOWERING LOCUS T are involved in the control of photoperiodic flowering in soybean. *Plant Physiol* **154**, 1220-1231 (2010).
29. Takada S, Goto K. Terminal flower2, an Arabidopsis homolog of heterochromatin protein1, counteracts the activation of flowering locus T by constans in the vascular tissues of leaves to regulate flowering time. *Plant Cell* **15**, 2856-2865 (2003).
30. Adrian J, Farrona S, Reimer JJ, Albani MC, Coupland G, Turck F. cis-Regulatory elements and chromatin state coordinately control temporal and spatial expression of FLOWERING LOCUS T in Arabidopsis. *Plant Cell* **22**, 1425-1440 (2010).
31. Taoka K*, et al.* 14-3-3 proteins act as intracellular receptors for rice Hd3a florigen. *Nature* **476**, 332-U397 (2011).
32. Abe M*, et al.* Transient activity of the florigen complex during the floral transition in *Arabidopsis thaliana*. *Development* **146**, dev171504 (2019).
33. Yamaguchi N. LEAFY, a Pioneer Transcription Factor in Plants: A Mini-Review. *Frontiers in Plant Science* **12**, 701406 (2021).

34. Schiessl K*, et al.* NODULE INCEPTION Recruits the Lateral Root Developmental Program for Symbiotic Nodule Organogenesis in *Medicago truncatula*. *Current Biology* **29**, 3657-3668 (2019).

35. Roux B*, et al.* An integrated analysis of plant and bacterial gene expression in symbiotic root nodules using laser-capture microdissection coupled to RNA sequencing. *Plant J* **77**, 817-837 (2014).

36. Xiao C, Chen F, Yu X, Lin C, Fu YF. Over-expression of an AT-hook gene, AHL22, delays flowering and inhibits the elongation of the hypocotyl in Arabidopsis thaliana. *Plant Mol Biol* **71**, 39-50 (2009).

37. Jakoby M*, et al.* bZIP transcription factors in Arabidopsis. *Trends Plant Sci* **7**, 106-111. (2002).

38. Abe M*, et al.* FD, a bZIP protein mediating signals from the floral pathway integrator FT at the shoot apex. *Science* **309**, 1052-1056 (2005).

39. Osipova MA*, et al.* Wuschel-related homeobox5 gene expression and interaction of CLE peptides with components of the systemic control add two pieces to the puzzle of autoregulation of nodulation. *Plant Physiol* **158**, 1329-1341 (2012).

40. Wang C*, et al.* SHORT-ROOT paralogs mediate feedforward regulation of D-type cyclin to promote nodule formation in soybean. *Proc Natl Acad Sci U S A* **119**, (2022).

41. Deakin WJ, Broughton WJ. Symbiotic use of pathogenic strategies: rhizobial protein secretion systems. *Nat Rev Microbiol* **7**, 312-320 (2009).

42. Tampakaki AP. Commonalities and differences of T3SSs in rhizobia and plant pathogenic bacteria. *Frontiers in Plant Science* **5**, 114 (2014).

43. Teulet A*, et al.* The rhizobial type III effector ErnA confers the ability to form nodules in legumes. *Proc Natl Acad Sci U S A* **116**, 21758-21768 (2019).

44. Teulet A, Camuel A, Perret X, Giraud E. The Versatile Roles of Type III Secretion Systems in Rhizobium-Legume Symbioses. *Annu Rev Microbiol* **76**, 45-65 (2022).

45. Ratu STN*, et al.* Rhizobia use a pathogenic-like effector to hijack leguminous nodulation signalling. *Scientific Reports* **11**, 2034 (2021).

46. Ma C*, et al.* The type III effector NopL interacts with GmREM1a and GmNFR5 to promote symbiosis in soybean. *Nature Communications* **15**, 5852 (2024).

47. Fan J*, et al.* A large-scale integrated transcriptomic atlas for soybean organ development. *Molecular Plant* **18**, 669-689 (2025).

48. Liu Z*, et al.* Integrated single-nucleus and spatial transcriptomics captures transitional states in soybean nodule maturation. *Nat Plants* **9**, 515-524 (2023).

49. Caetano-Anolles G, Gresshoff PM. Plant genetic control of nodulation. *Annu Rev Microbiol* **45**, 345-382 (1991).

50. Ferguson BJ*, et al.* Legume nodulation: The host controls the party. *Plant Cell and Environment* **42**, 41-51 (2019).

51. Lee R, Baldwin S, Kenel F, McCallum J, Macknight R. FLOWERING LOCUS T genes control onion bulb formation and flowering. *Nature Communications* **4**, 2884 (2013).

52. Pin PA*, et al.* An antagonistic pair of FT homologs mediates the control of flowering time in sugar beet. *Science* **330**, 1397-1400 (2010).

53. Baurle I, Laux T. Regulation of WUSCHEL transcription in the stem cell niche of the Arabidopsis shoot meristem. *Plant Cell* **17**, 2271-2280 (2005).

54. Huang W*, et al.* ALTERED MERISTEM PROGRAM1 suppresses ectopic stem cell niche formation in the shoot apical meristem in a largely cytokinin-independent manner. *Plant Physiol* **167**, 1471-1486 (2015).

55. Kimura Y, Tasaka M, Torii KU, Uchida N. ERECTA-family genes coordinate stem cell functions between the epidermal and internal layers of the shoot apical meristem. *Development* **145**, dev156380 (2018).

56. Lee C, Clark SE. A WUSCHEL-Independent Stem Cell Specification Pathway Is Repressed by PHB, PHV and CNA in Arabidopsis. *PloS one* **10**, e0126006 (2015).

57. Wang X, Fan C, Zhang X, Zhu J, Fu YF. BioVector, a flexible system for gene specific-expression in plants. *BMC Plant Biology* **13**, 198 (2013).

58. Huang P*, et al.* An Efficient Agrobacterium rhizogenes-Mediated Hairy Root Transformation Method in a Soybean Root Biology Study. *International Journal of Molecular Sciences* **23**, 12261 (2022).

59. Ahmed I, Islam M, Arshad W, Mannan A, Ahmad W, Mirza B. High-quality plant DNA extraction

for PCR: an easy approach. *Journal of Applied Genetics* **50**, 105-107 (2009).

60. Huang G, Ma J, Han Y, Chen X, Fu YFJPMBR. Cloning and Expression Analysis of the Soybean CO-Like Gene *GmCOL9*. *Plant Molecular Biology Reporter* **29**, 352-359 (2011).
61. Chen QJ, Zhou HM, Chen J, Wang XC. Using a modified TA cloning method to create entry clones. *Anal Biochem* **358**, 120-125 (2006).
62. Malapeira J, Mas P. ChIP-seq analysis of histone modifications at the core of the Arabidopsis circadian clock. *Methods Mol Biol* **1158**, 57-69 (2014).
63. Liu W, Xie X, Ma X, Li J, Chen J, Liu YG. DSDecode: A Web-Based Tool for Decoding of Sequencing Chromatograms for Genotyping of Targeted Mutations. *Molecular Plant* **8**, 1431-1433 (2015).
64. Xie X, Ma X, Zhu Q, Zeng D, Li G, Liu YG. CRISPR-GE: A Convenient Software Toolkit for CRISPR-Based Genome Editing. *Molecular Plant* **10**, 1246-1249 (2017).

Acknowledgements

We thank Dr. Eric Giraud for providing helpful suggestions on rhizobia and T3E.

Funding

This work was supported by the National Key Research and Development Program of China (2021YFF1001202 to Y.F., H.S, X.Z., P.H., Q.S., X.F., Y.M., and Q.C.), the National Natural Science Foundation of China (32372028 and 31771714 to Y.F, G.Y., L.L., M. L., L.H., K.Q. and 32360797 to M. L.), and the Innovation Program of Chinese Academy of Agricultural Sciences (to Y.F and X.Z.).

Author contributions

Y.F. and X.Z conceived, supervised the project and drafted the manuscript; H.S. conducted most of the experiments, including root nodule phenotype analysis, grafting experiments, vector construction, mutant identification, RT-qPCR, GUS histochemical staining, co-immunoprecipitation, subcellular localization, BiFC, transcriptional activation assays, EMSA, and ChIP-qPCR; X.Z. conducted soybean transformation

and hairy root transformation, phenotype investigation, and project and lab management; P.H. conducted some root transformation; Q.S. performed the yeast two-hybrid experiments; G.Y., L.L., M.L., L.H., and K.Q. cloned some gene and promoter and constructed related vectors; C. M. and D. X. designed and performed infection thread experiments and provided the sequences of Nop effectors of rhizobia HH103. Y.F. and H.S. analyzed the data and prepared the figures and tables; X.F, Y.M., Q.C., and J.Y. analyzed some data and reviewed the manuscript; all authors contributed to data analysis and manuscript preparation.

Competing interest

The authors declare no competing interests.

Figure legends

Fig. 1. *GmFT5b* enhances nodulation in soybean. a *GmFT5b*$_{pro}$*:GUS* activity in roots. Top: Stereomicrograms of roots. Middle: Cross-sectional views. Bottom: Magnified views of the boxed regions in the cross-sections. Violet arrows indicate pericycle cells, blue arrows indicate endodermis cells, and black triangles identify cortical cells exhibiting GUS staining. The experiments were repeated three times with similar results. **b** *GmFT5b*$_{pro}$*:GUS* activity in nascent nodule primordia. Red dashed lines indicate cells undergoing division. The experiments were repeated three times with similar results. **c** Time-course expression of *GmFT5b*, *GmFDL23*, *GmWUS1* (the left Y-axis), and *GmCLV3-1* (the right Y-axis) genes in roots after rhizobia inoculation. *GmUKN1* was used as the reference gene. Data are normalized to those in the uninfected roots. **d** and **e** Images and the number of large (≧ 2 mm) or small (< 2 mm) nodules per plant of WT and *Gmft5b* mutants at 21 DAI. **f** Nodule number of hairy roots expressing EV and *35S:GmFT5b*. **g** and **h** Images and the number of infection threads of WT and *Gmft5b* mutants at 1 DAI. **i** and **j** Nodule phenotypes in grafting experiments. Stock indicates Root stock. Scale bar: 100 μm (**a**, **b**), 5 mm (**d**, **i**) or 50 μm (**g**). HAI, hours after inoculation; DAI, days after inoculation; WT, wild-type; EV, empty vector. Data are presented as mean ± s.d. with $n = 3$ (**c**), 11 (**e**), 13 (**f**), 14 (**j**) or $n = 14$ (**h**) biological replicates; Two-way ANOVA (**e**, **j**), different letters indicate $p < 0.05$; or two-tailed Student's *t*-test (**f**, **h**),

** $p < 0.01$. Box plots display the distribution of data with boxes representing the first and third quartiles, a central black line indicating the median, whiskers extending to the minimum and maximum values, and individual measurements are shown as black circles (**e**, **f, h, j**). Source data are provided as a Source Data file.

Fig. 2. GmFT5b interacts with GmFDL23 to enhance nodulation. a Subcellular localization of GmFT5b-GFP and GmFDL23-GFP proteins in *N. benthamiana* leaves. The experiments were repeated three times with similar results. **b** BiFC assay for GmFDL23-nYFP and GmFT5b-cYFP in *N. benthamiana* leaves. AHL22-mCherry served as the nuclear marker [36]. The experiments were repeated three times with similar results. **c** Y2H assay for the interaction between GmFT5b and GmFDL23. **d** Co-IP assay for GmFDL23-Myc and GmFT5b-GFP. Proteins were immunoprecipitated with anti-Myc beads and probed by anti-Myc or anti-GFP antibodies. Green arrows indicate GmFDL23-Myc, red arrows indicate negative control GmCLV3-1-Myc. The experiments were repeated three times with similar results. **e** *GmFDL23$_{pro}$:GUS* activity in roots after inoculation with rhizobia HH103. Violet arrows indicate pericycle or endodermis cells, and red arrow indicates cortical cells. The experiments were repeated three times with similar results. **f** *GmFDL23$_{pro}$:GUS* activity in nascent nodule primordia. Red dashed lines indicate cells undergoing division. The experiments were repeated three times with similar results. **g** Nodule number per plant in EV and *GmFDL23*-knockout hairy roots or in EV and *35S:GmFDL23*-overexpressing hairy roots at 21 DAI. **h** Relative expression of nodulation marker genes in different mutant roots at 5 DAI. *GmUKN1* was used as the reference gene. Data are normalized to those in WT or EV. **i** and **j** Nodule number (**I**) and nodule density (**j**) of different mutants and overexpressing lines in different backgrounds at 21 DAI. DAI, days after inoculation; WT, wild-type; EV, empty vector. Scale bar: 50 μm (**a**, **b**, **f**), 100 μm (**e**). Data are presented as mean ± s.d. with n = 10 and 13 (**g**), 3 (**h**) or 12 (**i**, **j**) biological replicates; Two-tailed Student's t-test (**g**, **h**), *$p < 0.05$, **$p < 0.01$; or One-way ANOVA (**i**, **j**), different letters indicate $p < 0.05$. Box plots display the distribution of data with boxes representing the first and third quartiles, a central black line indicating the median, whiskers extending to the minimum and maximum values, and individual measurements are shown as black circles (**g, i, j**). Source data are provided as a Source Data file.

Fig. 3. *GmFT5b* and *GmFDL23* synergistically inhibit *GmCLV3-1* expression. a ChIP–qPCR assay to verify binding sites (P1-P4, sharing AGCT *cis*-elements) of GmFDL23 on *GmCLV3-1*$_{pro}$ in WT or *Gmft5b-1*. **b** EMSA assays. Probe localizes P1, P2 and P4. The experiments were repeated three times with similar results. **c** *GmCLV3-1*$_{pro}$*:LUC* activity in *N. benthamiana* leaves. **d** GUS staining (Top) and RT-qPCR analysis (Bottom) of *GmCLV3-1*$_{pro}$*:GUS* hairy roots in WT or the *Gmft5b-1* mutant**. e**-**g** Relative expression level of *GmCLV3-1* in the *Gmft5b* and *Gmfdl23* single mutants (**e**), *35S:GmFT5b* and *35:GmFDL23* hairy roots (**f**), or *35:GmFDL23 Gmft5b-1* roots (**g**). **h** *GmCLV3-1*$_{pro}$*:GUS* activity in hairy roots at 3 DAI. Violet arrows indicate pericycle or endodermis cells, and red arrow indicates cortical cells. **i** Nodule number of *GmCLV3-1* knockout or overexpressing roots in *Gmft5b-1* background at 21 DAI. CK, *35S:GFP*; WT, wild-type; EV, empty vector. Scale bar: 100 μm (**d**, **h**). *GmUKN1* was used as the reference gene (**d**, **e**, **f**, **g** Data are normalized to those in CK in luciferase activity analysis (**c**), or in WT or EV in RT-qPCR analysis (**d**, **e**, **f**, **g**). Data are presented as mean ± s.d. with $n = 3$ (**a**, **d**, **e**, **f**, **g**), 10 (**c**), or 11 (**i**) biological replicates; One-way ANOVA (**a**, **c**, **g**, **i**), different letters indicate $p < 0.05$; two-tailed Student's *t*-test (**d**, **e**, **f**), $^{**}p < 0.01$. Box plots display the distribution of data with boxes representing the first and third quartiles, a central black line indicating the median, whiskers extending to the minimum and maximum values, and individual measurements are shown as black circles (**i**). Source data are provided as a Source Data file.

Fig. 4. GmCLV3-1 represses *GmWUS1* to negatively regulate nodulation. a Expression of *GmWUS1* in the *Gmclv3-1* mutant and *35S:CLV3-1* hairy roots. **b** Analysis of *GmWUS1*$_{pro}$*:LUC* activity. **c** Expression of *GmWUS1* in the *Gmft5b-1* and *Gmfdl23* mutant. **d** *GmWUS1:GUS* activity with or without rhizobial induction. Violet arrows indicate pericycle or endodermis cells, and red arrow indicates cortical cells. The experiments were repeated three times with similar results. **e** *GmWUS1*$_{pro}$*:GUS* activity in nascent nodule primordia. Red dashed lines indicate cells undergoing division. The experiments were repeated three times with similar results. **f**-**h** Nodule number and density of different genetic modified lines related to *GmWUS1* gene at 21 DAI. **i** and **j** Nodule number and nodule density analysis in *Gmfdl23 Gmclv3-1* and *Gmwus1 Gmclv3-1* double mutant roots at 21 DAI. **k**-**m** Analysis of the effect of the

GmFT5b-GmFDL23 module on *GmWOX5*. DAI, days after inoculation; WT, wild-type; EV, empty vector; CK, *35S:GFP*. Scale bar: 100 μm (**d**, **e**). *GmUKN1* was used as the reference gene (**a**, **c**, **l**, **m**). Data are normalized to those in CK in luciferase activity analysis (**b**, **k**) and in WT or EV in RT-qPCR analysis (**a**, **c**, **l**, **m**); Data are presented as mean ± s.d. with $n = 3$ (**a**, **c**, **l**, **m**), 10 (**b**, **f**), 12 (**g**, **h**, **i**, **j**), or 13 (**k**) biological replicates; One-way ANOVA (**g**, **h**, **i**, **j**, **k**), different letters indicate $p < 0.05$; two-tailed Student's *t*-test (**a**, **b**, **c**, **f**, **l**, **m**) , $^{**}p < 0.01$. Box plots display the distribution of data with boxes representing the first and third quartiles, a central black line indicating the median, whiskers extending to the minimum and maximum values, and individual measurements are shown as black circles (**f-j**). Source data are provided as a Source Data file.

Fig. 5. ErnA and NopL effector induce the *GmFT5b-GmFDL23* module to trigger nodulation. a *35S:ErnA* enhances nodulation partially dependent on *GmFT5b*. **b** Expression of *GmFT5b*, *GmFDL23*, and nodulation marker genes in *35S:ErnA* hairy roots at 7 DAI. **c**-**e**, Effect of *NopL* effector on expression of nodulation marker genes (**c**), *GmFT5b-GmFDL23* module-related genes (**d**), and *GmCLV3-1* and *GmWUS1* promoters (**e**). **f** and **g** Nodule number (**f**) and nodule density (**g**) analysis in *35S:GmFT5b, 35S:GmFDL23*, *Gmclv3-1* and *35S:GmWUS1* at 21 DAI inoculated with the *nopl* mutant. **h** A proposed model of the *GmFT5b-GmFDL23* module promoting nodule formation. In uninfected roots, *GmFT5b* and *GmFDL23* are specifically expressed in stele cells, whereas cortical GmCLV3 represses *GmWUS1*. Following rhizobial infection, rhizobial signals (e.g., ErnA and NopL) activate GmFT5b and GmFDL23 in cortical cells where they locally inhibit *GmCLV3-1*, thereby relieving the repression of *GmWUS1* and facilitating nodule formation. DAI, days after inoculation; WT, wild-type; EV, empty vector. Scale bar: 100 μm (**e**). *GmUKN1* was used as the reference gene (**b**, **c**, **d**). Data are normalized to those in WT-EV (**a**, **b**) or HH103 (**c**, **d**) in RT-qPCR analysis; Data are presented as mean ± s.d. with $n = 10$ (**a**), 3 (**b**, **c**, **d**) or 12 (**f**, **g**) biological replicates; One-way ANOVA (**a**, **b, f, g**), different letters indicate $p < 0.05$; two-tailed Student's *t*-test (**c**, **d**) , $^{**}p < 0.01$. Box plots display the distribution of data with boxes representing the first and third quartiles, a central black line indicating the median, whiskers extending to the minimum and maximum values, and individual measurements are shown as black circles (**a**). Source data are provided as a Source Data file.