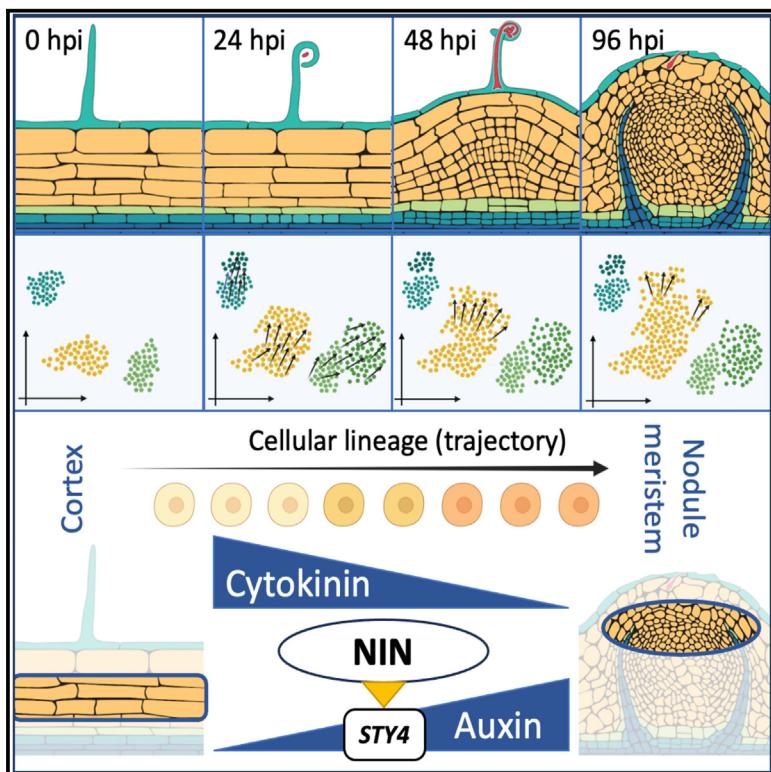


The single-cell transcriptome program of nodule development cellular lineages in *Medicago truncatula*

Graphical abstract



Authors

Wendell J. Pereira, Jade Boyd,
Daniel Conde, ..., Jean-Michel Ané,
Julia Frugoli, Matias Kirst

Correspondence

mkirst@ufl.edu

In brief

Plant nitrogen fixation by rhizobia is agriculture's most important symbiotic association, but only a few species develop root nodules to house the bacteria. Using single-cell analysis, Pereira et al. dissect the transcriptional program necessary for nodule formation, an essential step for the future introduction of this symbiotic relationship into crops.

Highlights

- Single-cell sequencing of a *Medicago* mutant captures rare cell response to rhizobia
- Root hair and stele cells differentiate to support infection and nodulation
- Cortex-derived cellular lineages modulate phytohormone activity in their trajectory
- High-dimension gene regulatory network analysis identifies RNS regulators



Resource

The single-cell transcriptome program of nodule development cellular lineages in *Medicago truncatula*

Wendell J. Pereira,¹ Jade Boyd,¹ Daniel Conde,^{1,2} Paolo M. Triozi,^{1,12} Kelly M. Balmant,^{1,3} Christopher Dervinis,¹ Henry W. Schmidt,¹ Carolina Boaventura-Novaes,¹ Sanhita Chakraborty,⁴ Sara A. Knaack,⁵ Yueyao Gao,^{6,7} Frank Alexander Feltus,^{6,8,9} Sushmita Roy,^{5,10,11,13} Jean-Michel Ané,^{4,13} Julia Frugoli,^{6,13} and Matias Kirst^{1,13,14,*}

¹School of Forest, Fisheries, and Geomatics Sciences, University of Florida, Gainesville, FL 32611, USA

²Centro de Biotecnología y Genómica de Plantas, Universidad Politécnica de Madrid – Instituto Nacional de Investigación y Tecnología Agraria y Alimentaria, 28223 Madrid, Spain

³Horticultural Sciences Department, University of Florida, Gainesville, FL 32611, USA

⁴Department of Bacteriology, University of Wisconsin – Madison, Madison, WI 53706, USA

⁵Wisconsin Institute for Discovery, University of Wisconsin, Madison, WI 53715, USA

⁶Department of Genetics & Biochemistry, Clemson University, Clemson, SC 29634, USA

⁷Broad Institute of MIT and Harvard, Cambridge, MA 02142, USA

⁸Biomedical Data Science and Informatics Program, Clemson University, Clemson, SC, USA

⁹Clemson Center for Human Genetics, Clemson University, Greenwood, SC 29646, USA

¹⁰Department of Biostatistics and Medical Informatics, University of Wisconsin, Madison, WI 53726, USA

¹¹Department of Computer Sciences, University of Wisconsin, Madison, WI 53706, USA

¹²PlantLab, Center of Plant Sciences, Sant'Anna School of Advanced Studies, 56010 Pisa, Italy

¹³Senior author

¹⁴Lead contact

*Correspondence: mkirst@ufl.edu

<https://doi.org/10.1016/j.celrep.2024.113747>

SUMMARY

Legumes establish a symbiotic relationship with nitrogen-fixing rhizobia by developing nodules. Nodules are modified lateral roots that undergo changes in their cellular development in response to bacteria, but the transcriptional reprogramming that occurs in these root cells remains largely uncharacterized. Here, we describe the cell-type-specific transcriptome response of *Medicago truncatula* roots to rhizobia during early nodule development in the wild-type genotype Jemalong A17, complemented with a hypernodulating mutant (*sunn-4*) to expand the cell population responding to infection and subsequent biological inferences. The analysis identifies epidermal root hair and stele sub-cell types associated with a symbiotic response to infection and regulation of nodule proliferation. Trajectory inference shows cortex-derived cell lineages differentiating to form the nodule primordia and, posteriorly, its meristem, while modulating the regulation of phytohormone-related genes. Gene regulatory analysis of the cell transcriptomes identifies new regulators of nodulation, including *STYLISH 4*, for which the function is validated.

INTRODUCTION

Nitrogen is an essential nutrient for plant growth. To acquire nitrogen, plants of the legume family establish a symbiosis with rhizobia bacteria that leads to the development of root nodules. Nodules are *de novo* plant organs that provide an optimal environment for the rhizobia to fix nitrogen in exchange for photosynthates. The establishment of root nodule symbioses (RNSs) requires the coordination of two processes: bacterial infection and nodule organogenesis. In the model legume *Medicago truncatula* responding to the rhizobium *Sinorhizobium meliloti*, the nucleus of the epidermal root hair cells increases in size, and genes involved in defense responses are transiently activated, together with bacterial

lipo-chito-oligosaccharides (LCOs) signaling within the first 24 h of interaction.^{1,2} In parallel, cell divisions occur in the pericycle, initiating nodule organogenesis. This first mitotic event is followed by the division of cortex cells that will become the nodule primordium.^{2,3} Together with the endodermis, the pericycle-derived cells differentiate into the uninfected cells at the base of nodules, while inner-cortex cells originate the other nodule compartments, in a process that involves a large number of genes.⁴

Nodules are modified lateral roots.^{5,6} Their development is regulated by the interplay between the phytohormones auxin and cytokinin (CK), synthesized and transported across root cells.⁷ Rhizobia induces the biosynthesis of CK by *LONELY GUY 3* (*MtLOG3*) in the epidermis,⁸ which is translocated to inner



root cells by the ABCG CK transporter *MtABCG56*.⁹ Cytokinin signaling in the root cortex and pericycle activates the expression of transcription factors such as *NODULE INCEPTION* (*MtNIN*) and the biosynthesis and accumulation of auxin through the induction of *LOB-DOMAIN PROTEIN 16* (*MtLBD16*) by *MtNIN*.⁶ Because of the metabolic cost to the plant, the host regulates nodule development through the autoregulation of nodulation (AON, reviewed by Ferguson et al.¹⁰). After rhizobial infection, peptides of the CLE (CLAVATA3 [CL3]/endosperm-surrounding region) family are induced and transported to the shoot. These peptides are perceived by a leucine-rich repeat receptor-like kinase (LRR-RLK), triggering the AON pathway by releasing a mobile signal. The LRR-RLK is encoded by *MtSUNN* (*SUPER NUMERIC NODULES*) in *M. truncatula*.¹¹ A null mutation in *MtSUNN* (*sunn-4*) results in a significant increase (>5-fold) in the number of nodules.¹²

The molecular mechanisms involved in the regulation of gene expression during nodule development have been examined by transcriptome analysis of whole nodulating roots (reviewed by Mergaert et al.¹³). In these studies, the response of each cell is masked by the tissue RNA profile, a limitation considering that distinct root cell types assume specific roles in nodulation.³ Approaches such as single-cell RNA sequencing (scRNA-seq) allow for uncovering cell transcriptional heterogeneity within a tissue or organ.¹⁴ Furthermore, methods such as pseudotime analysis add a developmental and temporal dimension to single-cell genomics data, allowing for cell lineages to be inferred and genes that control the process to be identified.¹⁵ Consequently, putative regulators of developmental processes such as shoot apex,^{16,17} leaf,^{18,19} and wood differentiation^{20,21} were identified and, in some cases, validated. Similar strategies could uncover the molecular regulators involved in the differentiation of root cell types into nodule components. This requires data to be collected across the developmental process, a limitation of existing studies.^{22–24}

Cells responding to rhizobia represent a small fraction of the whole root tissue, a limitation for single-cell genomic analysis of nodule development. Here we present a high-resolution and cell-type-specific gene expression map of the roots of the legume *M. truncatula* during early nodule development. To capture a significant fraction of cells responding to the rhizobia, we complemented the wild-type genotype Jemalong A17 analysis with data collected from the *sunn-4* mutant,²⁵ a regulator of the AON pathway. By disrupting the AON, the developmental program of nodules remains unchanged while affecting an extensive fraction of the root cortex and pericycle cells.²⁵ Using this approach, we identified the infected epidermal root hair cells and characterized their transcriptional differences from the uninfected cells. Moreover, we detected a subset of cells of the stele responding to infection, including the expression of RNS and AON regulators. Finally, we reconstructed the developmental trajectory of cortical cells differentiating to generate the nodule meristem, identifying genes that participate in this process and that may be required for engineering nodule organogenesis and N-fixation in crop plants. A web application has been developed²⁶ to allow the exploration of the expression profile of all *M. truncatula* genes during the rhizobial infection.

RESULTS

Single-cell analysis identifies root and nodule cell types

Before identifying the cell populations present in the sampled tissue, we removed entries from empty droplets and doublets (Table S1). We performed clustering after combining the single-cell transcriptomes obtained in four time points (0, 24, 48, and 96 h post inoculation [hpi]) for each of the two genotypes (WT and *sunn-4*), individually, using Monocle 3. The resulting dataset derived from each genotype was corrected for batch effects and comprised 6,998 nuclei grouped in 13 clusters for Jemalong A17 (Figure S1A) and 9,926 nuclei in 15 clusters for *sunn-4* (Figure S1B). The presence of similar cell types in both genotypes was confirmed by the expression of marker genes involved in RNS (Figure S1C). Next, we integrated the eight datasets (four time points for each of two genotypes) and identified 24 clusters (Figure 1A) containing 16,211 nuclei and 35,131 expressed genes.

To attribute cell identities to each cluster, we analyzed the expression profile of *M. truncatula* homologs of cell-type-specific markers previously identified in *Arabidopsis* roots^{27–29} and markers validated in *M. truncatula* or other species capable of RNS (Table S2). We also evaluated the expression of genes specific to each cluster (Moran's I test, q-value < 0.05) in a laser capture microdissection (LCM) dataset³⁰ that investigated the response of *M. truncatula* to rhizobial infection (Figures S2–S4). As a result, we identified nuclei derived from six types of root cells: the epidermal non-hair cells (atrichoblast; cluster 2), epidermal root hair (trichoblast; cluster 11), cortex (clusters 1 and 3), endodermis (cluster 8) and suberized endodermis (cluster 10), pericycle (cluster 7), and stele (excluding pericycle; clusters 5, 13, 14, 17, 18, and 19) (Figures 1B and 1E). Cluster 4, while composed of a majority of epidermal non-hair cells, also contains a subset of cells that appear to be from the cortex.

While the cluster annotation relied on a combination of strategies, some of the most informative marker genes for each cell type are described below. Epidermal non-hair cells were detected based on the expression of the phosphate transporter *MtPT1*³¹ and a homolog of the *Arabidopsis* SMALL AUXIN UPREGULATED RNA (SAUR1). Epidermal root hair cells were identified by the expression of RESPIRATORY BURST OXIDASE HOMOLOG F (*MtRbohF*)³². In cluster 8, the expression of SCARECROW³³ (*MtSCR*) indicated it is composed of endodermal cells, while the expression of MYELOBLASTOSIS (MYB) transcription factor 106 (*MtMYB106*; ortholog of AT1G34670) and GLYCEROL-3-PHOSPHATE 2-O-ACYLTRANSFERASE 1 (*MtGPAT-1*; ortholog of AT2G38110) shows that cluster 10 is composed of the suberized endodermal cells.²² The detection of cortical cells was supported by the expression of homologs of the *Glycine max* gene *Glyma.18G300800*, shown to be specifically expressed in this cell type.³⁴ The stele was characterized by the expression of SHORT ROOT 1 (*MtSHR1*) and SHORT ROOT 2 (*MtSHR2*).^{33,35} Markers specific to the xylem included the homologs of *Arabidopsis* GALACTURONOSYLTRANSFERASE 12 (*AtGAUT12*)³⁶ and (*AtCOBL4*)³⁷ and, for phloem, the homologs of the *Arabidopsis* genes AT5G54660 (MtrunA17Chr2g0319211), AT2G16740 (MtrunA17Chr4g0037011), and AT1G10380 (MtrunA17Chr5g0415341) previously described as markers for

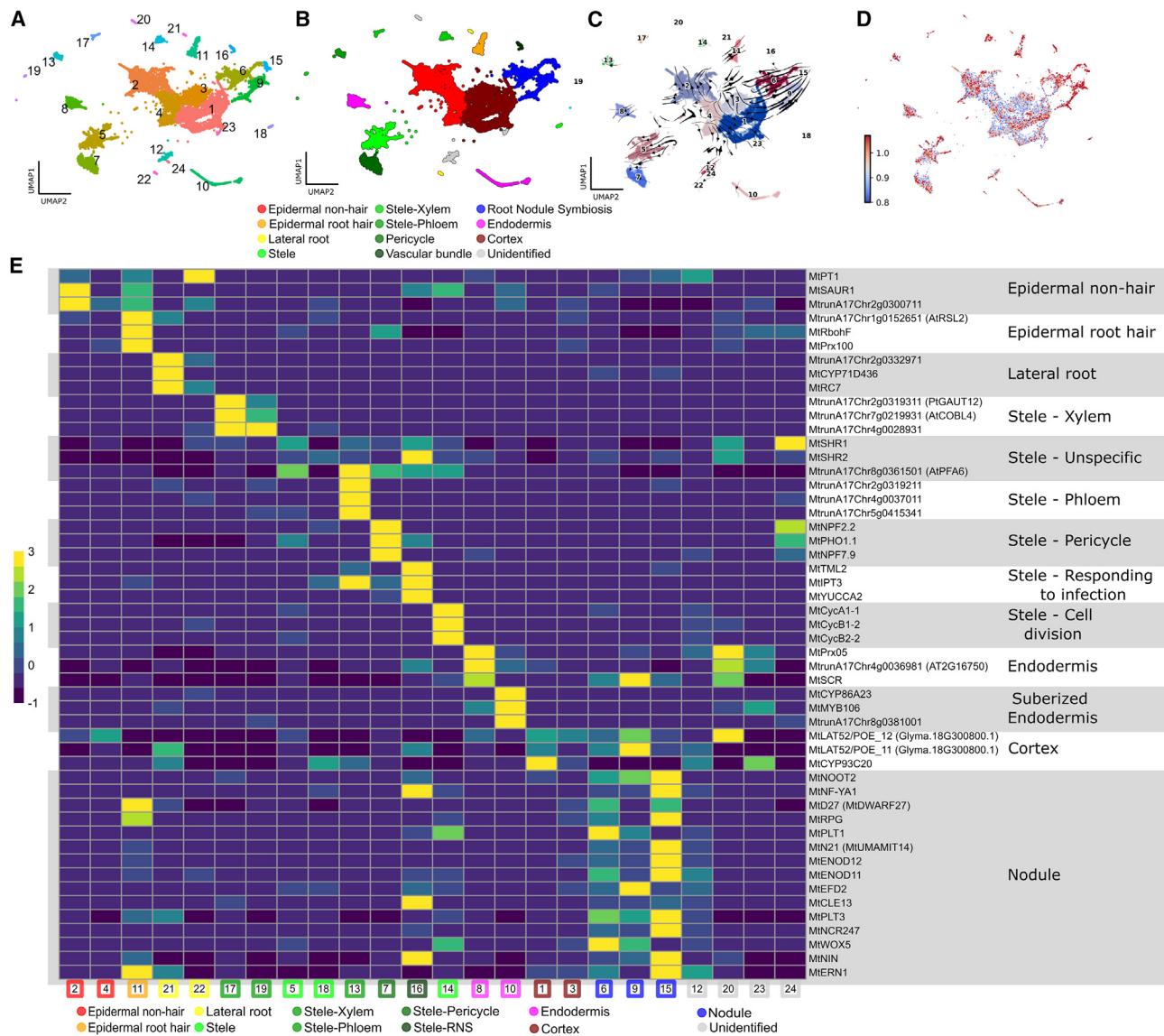


Figure 1. Single-nuclei transcriptomes from *M. truncatula* roots cluster into distinct cell types

(A) UMAP projection plot showing the distribution of the 16,211 cells from the two genotypes and four time points. Cells were grouped in 24 clusters according to their gene expression profiles.

(B) Six cell types were identified, in addition to clusters containing cells from the lateral root and nodule development.

(C) RNA velocity stream plot showing the dynamics of gene expression based on splicing kinetics.

(D) The UMAP plot describes the rate of differentiation represented by the length of the velocity vectors generated in the RNA velocity analysis, where longer vectors (red) indicate higher rates.

(E) The expression of known cell-type marker genes across cells reveals the identity of clusters. In (E), color shows the average expression across cells in each cluster. See also Table S2, which contains the complete list of cell-type markers used in the present study, and Figures S2–S4.

that cell type.²² Markers for early cell proliferation of the vascular procambial/cambial cells in the stele (e.g., homologs of *Arabidopsis CYCLIN A1;1*)³⁸ were detected in a cluster related to cell division (cluster 14). Cluster 7 is characterized by the expression of pericycle markers, including homologs of *Arabidopsis NITRATE TRANSPORTER 1.5* genes.³⁹ Clusters 21 and 22 represent cells from lateral roots based on the detection of genes expressed specifically during their development and unrelated to RNS⁶ (Figure 1E). The cell identity of clusters 12 (n = 285 cells),

20 (n = 75), 23 (n = 34), and 24 (n = 24) remained unknown due to the absence of molecular evidence (cluster 12) or because of the mixed expression of marker genes from multiple cell types (clusters 20, 23, and 24).

We also identified four clusters of cells involved in nodule development (clusters 6, 9, 15, and 16), as shown by the expression of genes involved in RNS and genes detected in the LCM data (Figures 1E and S2–S4). Cluster 15 is composed of infected cells that eventually become the nitrogen fixation zone of the

nodule, as shown by the expression of genes such as *NUCLEAR FACTOR Y SUBUNIT A 1* (*MtNF-YA1*) and *NODULE-SPECIFIC CYSTEINE-RICH PEPTIDE 247* (*MtNCR247*) (Figure 1 E). Both clusters 6 and 9 contain meristematic cells, expressing genes of the *PLETHORA* family. Cells in cluster 9 also express marker genes compatible with the presence of nodule primordia (e.g., *BASIC-HELIX-LOOP-HELIX* transcription factor 1, *MtbHLH1*), which are absent in cluster 6. Thus, it may represent cells differentiating from the cortex to become a nodule. RNA velocity analysis supports this inference, with cell dynamics pointing to a transition from cortex (C1) to nodule (C15), passing through C9 (Figures 1C and 1D). We note that a large fraction of cells in cluster 6 is already present in the control (0 hpi), suggesting the presence of meristematic cells that are not part of the nodule primordia (Video S1).

The annotation of the clusters was also supported by the expression profile of their 100 most specific genes (q-value < 0.05, sorted by the specificity value for Moran's I test) in the LCM data (Figures S2–S4). Table S2 describes the list of genes significantly enriched for expression in each cluster (Moran's I test, q-value < 0.05). In addition, we identified genes differentially expressed (Wilcox test, q-value < 0.05) in the comparison among time points for each cluster (Table S3).

Contrasting response of epidermal sub-cell types to infection

RNS begins with the perception of bacterial LCOs by epidermal root hair cells, followed by infection and the formation of an infection thread (IT). As expected, at 24 hpi, the transcriptome profile of a subset of epidermal root hair cells (cluster 11) included the expression of marker genes for the IT and infectosome, such as *RHIZOBIUM DIRECTED POLAR GROWTH* (*MtRPG*⁴⁰), *CYSTATHIONINE-B-SYNTHASE-LIKE DOMAIN-CONTAINING* (*MtCBS1*⁴¹), *NODULE PECTATE LYASE* (*MtNPL*⁴²), *FLOTILLIN 4* (*MtFLOT4*⁴³), and *VAPYRIN* (*MtVPY*⁴⁴). Cluster 11 also showed the highest expression of genes related to epidermal infection.⁶ A high expression of genes related to epidermal infection was also observed in cluster 15, composed of cells from the developing nodule, pointing to the contribution of these genes to the infection of the cortex-derived nodule cells.

Because a subset of the cells in the clusters containing epidermis (atrichoblasts, Figure 1, clusters 2 and 4, and trichoblasts, Figure 1, cluster 11) was identified as responding to the rhizobia, we reclustered them at a higher resolution to compare gene expression in infected cells relative to those uninfected. Reanalysis of epidermal root hair and epidermal non-hair cells identified 10 clusters (Figures 2A–2C), hereinafter referred to as RH1 to RH10, of which only RH6 was absent at 0 hpi but appeared at 24 hpi in the genotypes A17 and sunn-4 (Video S2). Annotation based on marker genes for epidermal root hair containing IT, including those described above, confirmed the identity of this emerging cluster. By investigating the expression of marker genes (e.g., *MtRbohF*³²), and the list of the most specifically expressed genes in each cluster (Table S2), we identified RH8 as representing the uninfected epidermal root hair. The expression profile of genes previously identified as markers for epidermal infection⁶ showed that RH6 is involved in this process (Figure 2D).

Next, we identified the genes most specifically expressed in the infected cells of RH6 (Moran's I test, q-value < 0.05) (Figure 2C; Table S2). In legumes, RNS begins with the recognition of LCOs secreted by the rhizobia by a heteromeric Lys motif (LysM)-type transmembrane receptor complex. In *Medicago*, this association involves the LYK-I and LYR-type receptors *LYSM DOMAIN CONTAINING RECEPTOR KINASE 3* (*MtLYK3*) and *NOD FACTOR PERCEPTION* (*MtNFP*), respectively.⁴⁵ While both are transcribed in cells of cluster RH6, the role of LysM domain-containing receptor kinases in RNS likely extends to other gene family members, including *MtLYK4* and *MtLYK10*, which were specifically expressed in that cluster (Table S2). The induction of these receptor kinases was accompanied by the expression of several genes in the strigolactones (SL) biosynthesis pathway—a class of phytohormones hypothesized to promote IT formation.⁴⁶ Genes encoding for proteins involved in the synthesis of SL precursors (carotenoids), including *MtZDS* (*ZETA-CAROTENE DESATURASE*) and *MtZ-ISO* (*15-CIS-ZETA-CAROTENE ISOMERASE*), were preferentially expressed in infected epidermal root hair cells (RH6). The expressions of SL pathway genes *MORE AXILLARY BRANCHES 1* (*MtMAX1a*) and *DWARF27* (*MtD27*)⁴⁷ were among the most specific to this cluster (Table S3) and were significantly upregulated (Wilcox test, q-value < 0.05; Table S3) compared to the uninfected epidermal root hair cells (cluster RH8). Markers for infected cells (e.g., *MtRPG*), as well as RNS uncharacterized genes related to plant defense (*PADRE DOMAIN-CONTAINING PROTEIN*, MtrunA17Chr8g0364591 and MtrunA17Chr4g0054191) and cell wall remodeling (*PECTINESTERASE*, MtrunA17Chr4g0069841) were significantly upregulated in infected (RH6) compared to uninfected epidermal root hair cells (RH8) (Table S3). Cell wall remodeling is required for the rhizobia to penetrate the epidermal root hair cell wall. Finally, several receptor-like protein kinases (RLK) were upregulated in RH6 (Table S3). This family of proteins plays critical roles in signaling pathways related to developmental and defense responses in plants. Among the differentially expressed RLKs, a gene of the *PROTEIN KINASE RLK-PELLE-LRR-XII-1 FAMILY* (MtrunA17Chr5g0408421) was only detected after rhizobial infection, with high specificity to the infected epidermal root hairs (RH6) and to the infected cells of the nodule (discussed below), pointing to a role during RNS establishment. MtrunA17Chr5g0408421 is downregulated in the *Medicago nfp*, but not in *lyk3*, mutant background⁴⁸; our data show a strong spatiotemporal co-expression between MtrunA17Chr5g0408421 and *MtNFP*.

In contrast to the epidermal root hair cells containing IT, those in RH8 were enriched for the expression of NADPH oxidases, or *RESPIRATORY BURST OXIDASE HOMOLOGS* (RBOHs). Reactive oxygen species (ROS) are abundant during nodulation and mainly generated by RBOHs. The *Medicago* RBOH gene family contains ten members, some of which catalyze the generation of ROS in response to infection when a protein complex is formed with *CALCIUM-DEPENDENT PROTEIN KINASES* (CDPKs)⁴⁹. When comparing the expression of RBOHs and CDPKs, we observed a significant, specific expression of *MtRbohB*, *MtRbohD*, *MtRbohF*, *MtRbohH*, *MtRbohI*, and *MtCDPK9* in epidermal root hair cells lacking IT (RH8; Moran's test, q-value < 0.05). In addition, *MtCDPK9*, *MtRbohD*,

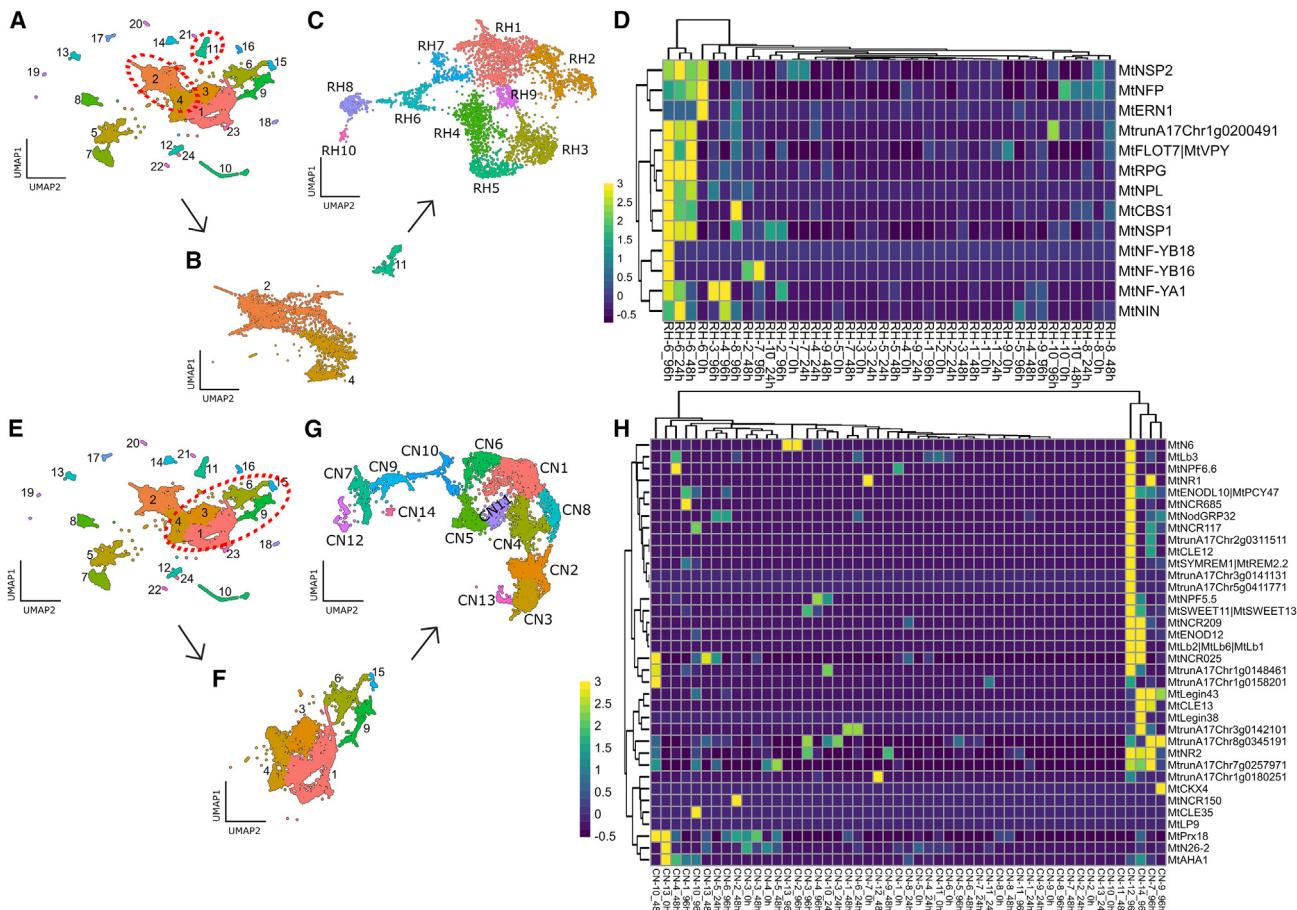


Figure 2. Reclustering cells of different types reveals sub-cell types responding to the infection

(A–D) Analysis of epidermal non-hair and epidermal root hair cells during RNS. From the whole dataset (A), cells belonging to clusters 2, 4, and 11 were isolated (B) and reclustered at a higher resolution (see [STAR Methods](#)), resulting in 10 subclusters (C). Using marker genes of epidermal infection identified previously in Schiessl et al.,⁶ RH8 was identified as epidermal root hair and RH6 as epidermal root hair containing IT (D).
 (E–H) Analysis of cortical and nodule cells during RNS. Cells from clusters 1, 3, 4, 6, 9, and 15 were isolated (F) and reclustered in 14 subclusters (G). Based on nodule-specific genes, described by Schiessl et al.,⁶ CN7, CN9, CN12, and CN14 were shown to comprise cells from the nodule (H). The heatmaps in (D) and (H) show the average expression of all cells in a cluster at each time point. Note that some clusters are only present at specific time points. See also [Table S2](#) for the list of genes shown in the heatmaps.

MtRbohF, and *MtRbohI* were downregulated in the infected cells compared to the uninfected in at least one of the time points (Wilcoxon test, q-value < 0.05), pointing to a coordinated activation of innate immunity in those cells, in contrast with those containing IT where their expression is suppressed. Investigating the RBOHs on the whole dataset shows that while suppressed in the infected epidermal root hair, some of those genes are activated in the cortex (C1; *RbohA*) and nodule-related clusters (C9; *RbohA*, *RbohC*, and *RbohE*) as the infection progresses.

Autoregulation of nodulation genes is activated in a subset of stele cells following infection

The first cell divisions triggered by the rhizobia occur in the pericycle as early as 24 hpi, despite several cell layers of separation from the epidermal root hair where the plant first perceives the LCOs. In conjunction with the endodermis, some of these cells later differentiate to form the nodule vascular bundle.³ Based

on marker genes, C7 was annotated as pericycle (Figures 1B and 1E). Because several reported pericycle markers were also expressed in other clusters, we generate a *promoter::beta-glucuronidase (GUS)* fusion for MtrunA17Chr3g0136151, a gene specifically expressed in C7 (Figure 3A; Table S2). Transgenic roots generated via *Agrobacterium rhizogenes*-mediated transformation of *M. truncatula* showed GUS activity under the MtrunA17Chr3g0136151 promoter primarily in the pericycle cells (Figure 3B). Cluster analysis of these cells at a higher resolution did not result in their separation into sub-types. Furthermore, analysis of the differential expression of the cells in this cluster, across time points, revealed only 47 differentially expressed genes (DEGs) (Wilcox test, q -value < 0.05), all unrelated to infection and RNS (Table S3). Thus, C7 appears to represent pericycle cells in a steady state, not responding to the rhizobia.

In contrast, we observed that C16, a cluster of cells that emerged mostly at 96 hpi (Video S1, Figure S5), is enriched

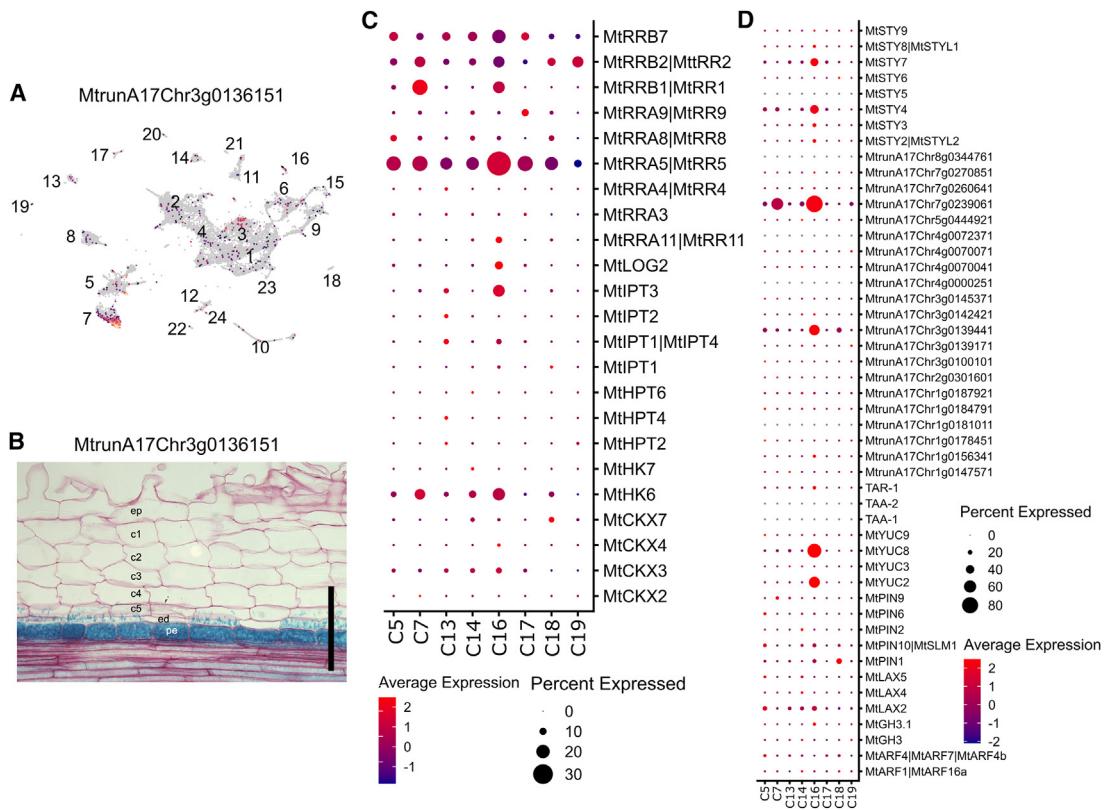


Figure 3. Cytokinin and auxin genes are differently expressed in stele clusters, including pericycle and differentiating cells

(A) Expression profile of the gene MtrunA17Chr3g0136151, a putative *TRIPEPTIDYL-PEPTIDASE II*, selected for localization experiments due to its specific expression in C7.

(B) GUS activity under the MtrunA17Chr3g0136151 promoter was primarily detected in the pericycle in Jemalong A17. Cluster 16 is composed of stele cells responding to the rhizobial infection. Expression profile of genes related to CK (C) and auxin (D) signaling, described by Schiessl et al.,⁶ show an increased expression of RNS genes, such as *MtIPT3*, *MtLOG2*, *MtYUC2*, and *MtYUC8*, in cluster 16 in comparison to all other clusters containing cells from the stele. Scale bar represents 30 µm. See also Table S3.

for the expression of RNS genes, such as *MADS BOX TRANSCRIPTION FACTOR* (*MtMADS1*⁵⁰), *MtNIN*⁵¹ and *MtNF-YA1*,⁵² *MtSHR2*, and *SHORT INTERNODES/STYLISH* 4 and 7 (*MtSTY4* and *MtSTY7*). Cluster C16 transcriptional profile is also consistent with the LCM data from nodule vasculature, a tissue that emerges from pericycle cells³ (Figure S3). The cluster-specific expression of the gene *ISOPENTENYL TRANSFERASE 3* (*MtIPT3*) further supports that C16 represents cells from the stele responding to the rhizobial infection, at the base of the dividing cells in the nodule primordia.⁵³ The highest correlation between the transcriptome of C16 and C5 ($r = 0.81$), and C7 ($r = 0.80$), strengthens the hypothesis that this cluster contains cells from the stele responding to the infection by rhizobia, in particular the pericycle cells that have undergone cell division. Cells from this cluster were mostly detected in *sunn-4* at 96 hpi. During nodulation, the pericycle undergoes fewer cell divisions compared to the cortex to form the nodule. Thus, sampling from a hypernodulating mutant such as *sunn-4* may be necessary to obtain a sufficient representation of this relatively rare cell sub-type.

To further investigate the transcriptional response triggered by RNS in C16, we compared gene expression in this cluster relative to all other clusters annotated as representing cell types pre-

sent in the stele (C5, C7, C13, C14, C17, C18, and C19). A total of 194 DEGs were identified (Table S3), including RNS genes involved in nodule-specific sucrose transport (*MtSWEET15c* [*MtN3*], *MtSWEET3c*, and *MtSWEET15d*), auxin biosynthesis regulators (e.g., *YUCCA* 2 and 8; *MtYUC2* and *MtYUC8*), and RNS transcription factors. The expression of genes involved in the AON was also concentrated in this cluster. Both *CLAVATA3/EMBRYO-SURROUNDING REGION 12* and *13* (*MtCLE12* and *MtCLE13*) are part of the root-to-shoot signaling in AON.⁵⁴ The signal is modulated by *MtSUNN* in the shoot, triggering the activating of the genes *TOO MUCH LOVE 1* and *2* (*MtTML1* and *MtTML2*) in the roots, inhibiting nodule formation. Except for *MtTML1*, these AON genes are highly and specifically expressed in C16. In addition, *MtCEP1* (*C-TERMINALLY ENCODED PEPTIDE*), which acts as a nitrogen “hunger” signal in plants and is transported from root to shoot,^{4,55} is also specifically expressed in C16. Altogether, the evidence suggests that C16 represents a group of cells that coordinate nodule development in *M. truncatula* through the AON. It should be noted that the detection of expression of *MtTML* genes in *sunn-4* conflicts with the expectation that their regulation is dependent on *MtSUNN*, pointing to alternative pathways involved in controlling

nodule number and distribution by the AON. The expression of *MtTML* genes in *sunn-4* mutants has also been reported by others.⁵⁶

Clusters are identified for cells in the nodule meristem and infection zone

Two clusters appeared after 0 hpi in A17 and *sunn-4*, reflecting the emergence of new cell types induced by the rhizobial infection (Figure S5). Cells in cluster 9 were first noted at 48 hpi, followed by an expansion at 96 hpi, when cluster 15 was also detected (Video S1). The number of cells in clusters 9 and 15 was higher in the hypernodulating mutant *sunn-4* genotype.

Cluster 9 comprised cortex cells responding to the infection, transitioning toward nodule formation (nodule primordia). Markers for the nodule primordia, such as *NON SPECIFIC LIPID TRANSFER PROTEIN* (*MtN5*)⁵⁷, *MtbHLH1*,⁵⁸ and the *AUXIN INFLUX TRANSPORTER LAX2* (*MtLAX2*)⁵⁹, were strongly expressed in this cluster. Moreover, RNA velocity shows a progression from cortex (cluster 1) to nodule (cluster 15) passing through this cluster (Figures 1C and 1D), reinforcing the hypothesis that these cells are in a developmentally transitioning state. The transcription profile of this cluster's most specific markers strongly overlaps with the expression detected in the LCM data obtained from developing nodules (Figure S3).

Cluster 15 was composed of cells that are part of the developing nodule, as shown by the localized expression of multiple nodule-specific genes⁶ and by the profile of the most specifically expressed genes of this cluster, in the LCM data (Figure S3). Based on the expression of known markers, cells from two of the nodule zones could be detected at 96 hpi. These include cells from the nodule meristem, expressing *PLETHORA* (*MtPLT*) 1, 3, and 4,⁶⁰ *WUSCHEL-RELATED HOMEOBOX 5* (*MtWOX5*)⁶⁰, *MtCLE13*,⁵⁴ and *NODULE ROOT2* (*MtNOOT2*)⁶¹. Also detected were genes expressed in the infection zone (ZII), such as *NODULE-SPECIFIC PLAT-DOMAIN 1* (*MtNPD1*)⁶², *MtNF-YA1*,⁶³ *MtNCR247*,⁶⁴ *USUALLY MULTIPLE ACIDS MOVE IN AND OUT TRANSPORTERS 14* (*MtUMAMIT14/MtN21*)⁶⁵, and *SWEET11*.⁶⁶ We searched for cells expressing NCR peptides to identify those representing the nitrogen-fixing zone (ZIII). Investigating the expression profile of all 675 NCR peptides in the *M. truncatula* genome revealed that genes encoding for 221 were expressed in our dataset (Table S2). However, NCR peptides specific to ZIII, such as *MtNCR035* and *MtNCR001*,⁶² were not detected. The lack of transcripts for both peptides suggests an absence of cells from the ZIII, compatible with the developing stage of the nodules at 96 hpi in *M. truncatula*.³ Alternatively, the number of cells from ZIII may be too limited to result in a discernible cluster.

In addition to clusters 9 and 15, cluster 6 also showed an expansion in the number of cells between 48 and 96 hpi. Cells in cluster 6 show a transcriptional profile that is characteristic of meristematic cells, including the expression of genes of the *PLETHORA* family. However, a subset of those cells was already present at 0 hpi (Video S1), suggesting that they represent cells from the lateral root meristem. These cells were removed from the dataset in the follow-up analysis.

Single-cell trajectory inference uncovers the cell lineage leading to nodule formation

The mitotic division of distinct cortical cell layers originates cell lineages that form each nodule compartment.³ In *M. truncatula*, the nodule meristem is derived from dividing cells of the 3rd cortical layer. Meanwhile, the 4th and 5th cell layers of the inner root cortex give rise to approximately eight nodule cell layers. These cells undergo an enlargement phase before being penetrated by the IT.³ At approximately 80 hpi, rhizobia are released from the IT and infect those cells. Concurrently, the nodule meristem begins to divide and add cells to the nodule.³ Thus, our dataset represents the lineage derived from the cortex to form nodule compartments, as well as the contribution of the nodule meristematic cells. To investigate these processes, we re-analyzed root cortex cells (Figures 2E and 2F; clusters 1, 3, and 4) detected at all time points and clusters that emerged from them after 0 hpi. This included cells from clusters 9 and 15, and cluster 6 excluding cells already present at 0 hpi (Figures 2E and 2F). Re-clustering resulted in the detection of fourteen groups of cells, hereinafter referred to as CN1 to CN14 (Figures 2G and 2H). Among these, clusters CN7, CN9, CN10, CN12, and CN14 are associated with nodule development (Video S3).

To reconstruct the cell lineage trajectories that result in nodule formation, we applied a trajectory inference analysis using Slingshot.⁶⁷ We defined the cortex cluster with most cells at 0 hpi (CN1) as the starting trajectory position but did not assign an end cluster to avoid biasing the direction of the trajectory. Among the lineages that emerged (Figure S6), the lineage from CN1 → CN6 → CN10 → CN9 → CN7 represents cells transitioning from cortex (CN1) to nodule meristem (CN7). Based on the LCM data and markers gene, we inferred that CN1 and CN6 are composed of cortical cells; these clusters were also detected throughout the experiment (0–96 hpi). Cells in both clusters were significantly enriched for the expression of *CHASE-DOMAIN CONTAINING HISTIDINE KINASE* receptors (*MtCHK3* and *MtCHK4* in CN1, *MtCHK3* and *MtCHK1/MtCRE1* in CN6; Table S2) and other CK signaling genes (e.g., *MtRRB24* (*TYPE-B RESPONSE REGULATOR*), Figure 4A). Genes related to CK signaling were also specifically expressed in CN10 (Moran's I test, q-value < 0.05), such as *TYPE-A RESPONSE REGULATOR 5* and *6* (*MtRRA5*) and (*MtRRA9*). Markers for the nodule primordia, such as *MtbHLH1*⁵⁸ (Figure 4A) and *MtLAX2*⁵⁹ were strongly expressed in CN9 and part of CN10. CN9 was particularly enriched for auxin biosynthesis, transport, and signaling genes. Among the genes specifically expressed in this cluster are several components of the *STYLISH* family (*MtSTY2*, *MtSTY3*, *MtSTY4*, *MtSTY7*, and *MtSTY8*). LOB-domain proteins regulate these genes and, together with genes of the *YUCCA* family, participate in generating a local auxin maximum required for the initial nodule developmental program.⁶ The same cluster-specific genes included auxin influx and auxin efflux carriers (e.g., *MtPIN2* (*PIN-FORMED AUXIN EFLUX CARRIER COMPONENT2*), Figure 4A, and *MtLAX2*). Finally, the expression of members of the *LIGHT SENSITIVE SHORT HYPOCOTYL* transcription factor gene family (*MtrunA17Chr1g0183761-MtLSH1* and *MtrunA17Chr7g0261381-MtLSH2*), recently shown to be required for the specification of nodule primordia,⁶⁸ was also enriched in CN9 and CN7. Markers for nodule meristem, including

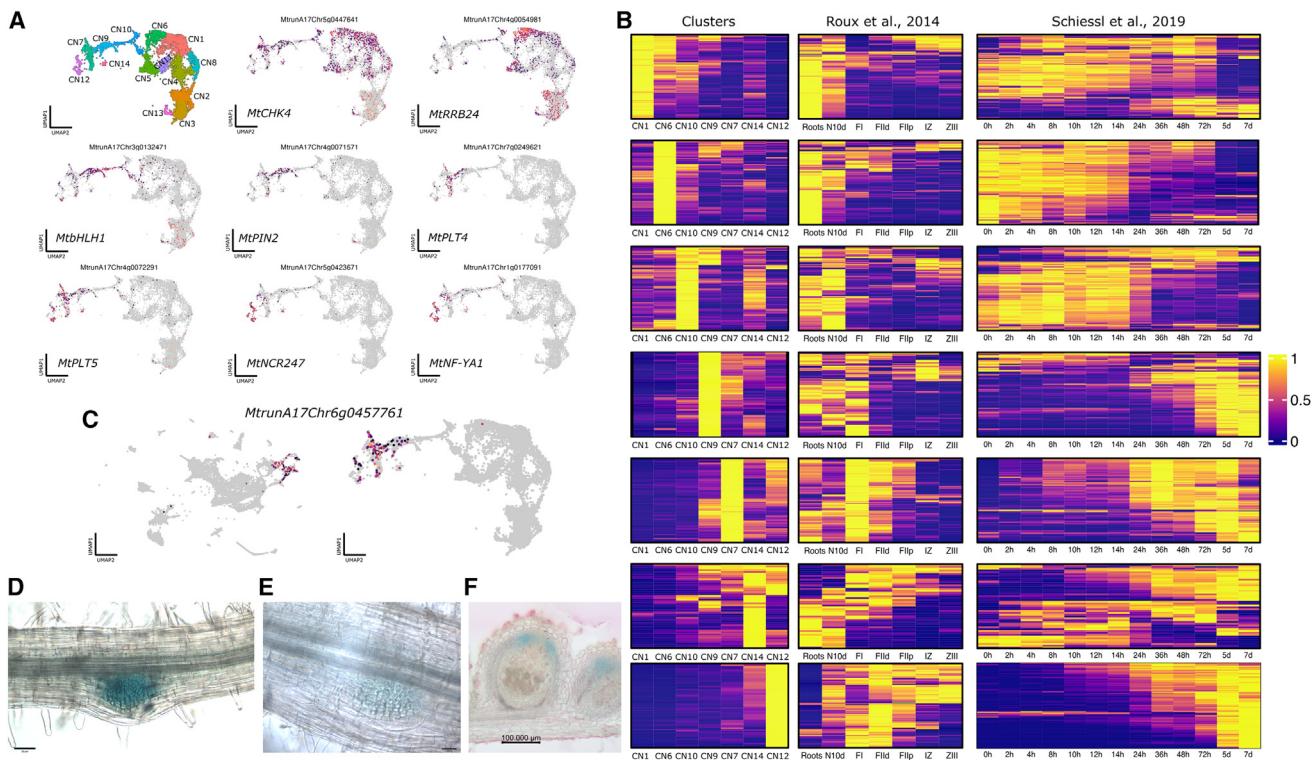


Figure 4. Reclustering cells at higher resolution reveals distinct nodule components

(A) From the complete dataset, cells belonging to clusters 1, 3, 4, 6, 9, and 15 were selected and reclustered (as described in Figure 2G). The remaining UMAP plots represent the expression profile of marker genes for cortex responding to the infection (*MtCHK4* and *MtRRB24*), nodule primordia (*MtbHLH1* and *MtPIN2*), nodule meristem (*MtPLT4* and *MtPLT5*), and for infected cells from the infection zone (*MtNCR247* and *MtNF-YA1*).

(B) Expression of the most specific genes ($n = 100$; q -value < 0.05 , sorted by the specificity value returned by the Moran's I test) detected in the CN clusters (left panel) and their transcriptional profile in two reference datasets, including nodule fractions obtained by laser capture microdissection³³ (middle panel) and bulk RNA sequencing⁶ (right panel). Each dataset was scaled and row-centered separately. The clusters are shown in the order they appear in the UMAP plots represented in (A).

(C) Expression localization of *MtrunA17Chr6g0457761* in the whole dataset (left panel) and in the cells from cortex and nodule after reclustering (right panel). (D–F) GUS activity under the *MtrunA17Chr6g0457761* promoter was detected in both lateral root primordia (D) and nodule primordia (E) of genotype Jemalong A17 at 6 dai. (F) At 21 dai, GUS activity is mainly detected in the nodule meristem. Scale bars represent 50 µm in (D), 25 µm in (E), and 100 µm in (F). Differentiating between lateral root primordia and nodule primordia was achieved by observing in what root layer cell division is more frequent: lateral root primordia if most divisions originated in the pericycle and nodule primordia if most divisions originated on the 3rd, 4th, and 5th inner layers of the cortex. See also Table S1.

*MtWOX5*⁶⁰, the *PLETHORAS* *MtPLT3*, *MtPLT4*, and *MtPLT5*⁶⁰ (Figure 4A), and *MtNOOT2*⁶¹ were expressed in CN7. Analysis of the expression of the most specific genes of these clusters in two reference datasets^{6,33} reinforces this annotation. Genes specific to CN7 were more highly expressed in infected than in control samples starting at 24 hpi in a bulk RNA sequencing dataset⁶ and were detected primarily in the fraction of the nodule that contains the meristem (Fl)³³ (Figure 4B). To further confirm the identity of this cluster, we selected the gene *MtrunA17Chr6g0457761* (Figure 4C), a hypothetical protein ranked among the most specifically expressed genes in CN7, for spatial and temporal localization using *promoter::GUS* fusion. GUS activity under the *MtrunA17Chr6g0457761* promoter was detected in both the nodule primordia and lateral root primordia 6 days after infection (dai; Figures 4D and 4E). At 21 dai, expression is specifically localized in the meristem of mature nodules (Figure 4E). We confirmed its expression in both lateral root and nodules in the dataset produced by Schiessl et al.⁶ The expression is substantially higher in

nodules at all time points measured at or after 72 hpi, further supporting that this cell lineage originates the nodule meristem in *M. truncatula*.

To investigate genes governing the developmental trajectory from CN1 (cortex) to the cells forming the nodule meristem (CN7), we applied TradeSeq⁶⁹ to identify DEGs in the inferred trajectory (Figures 5A–5C). A total of 6,779 DEGs were detected (Table S5); those known to be involved in RNS are shown in Figure 5D.

In addition to the cell lineage involved in the nodule meristem development, a cluster of infected cells (CN12) was also detected (Figure 4A). Trajectory inference analysis⁶⁷ identified CN12 as an isolated cluster, lacking connectivity to any other cluster or lineage. CN12 may represent existing root cortex cells that are infected at 96 hpi. Alternatively, CN12 may be part of a lineage, but the lack of cells at intermediate stages of the trajectory could limit our ability to make inferences about its relationship to other clusters. A similar limitation applies to

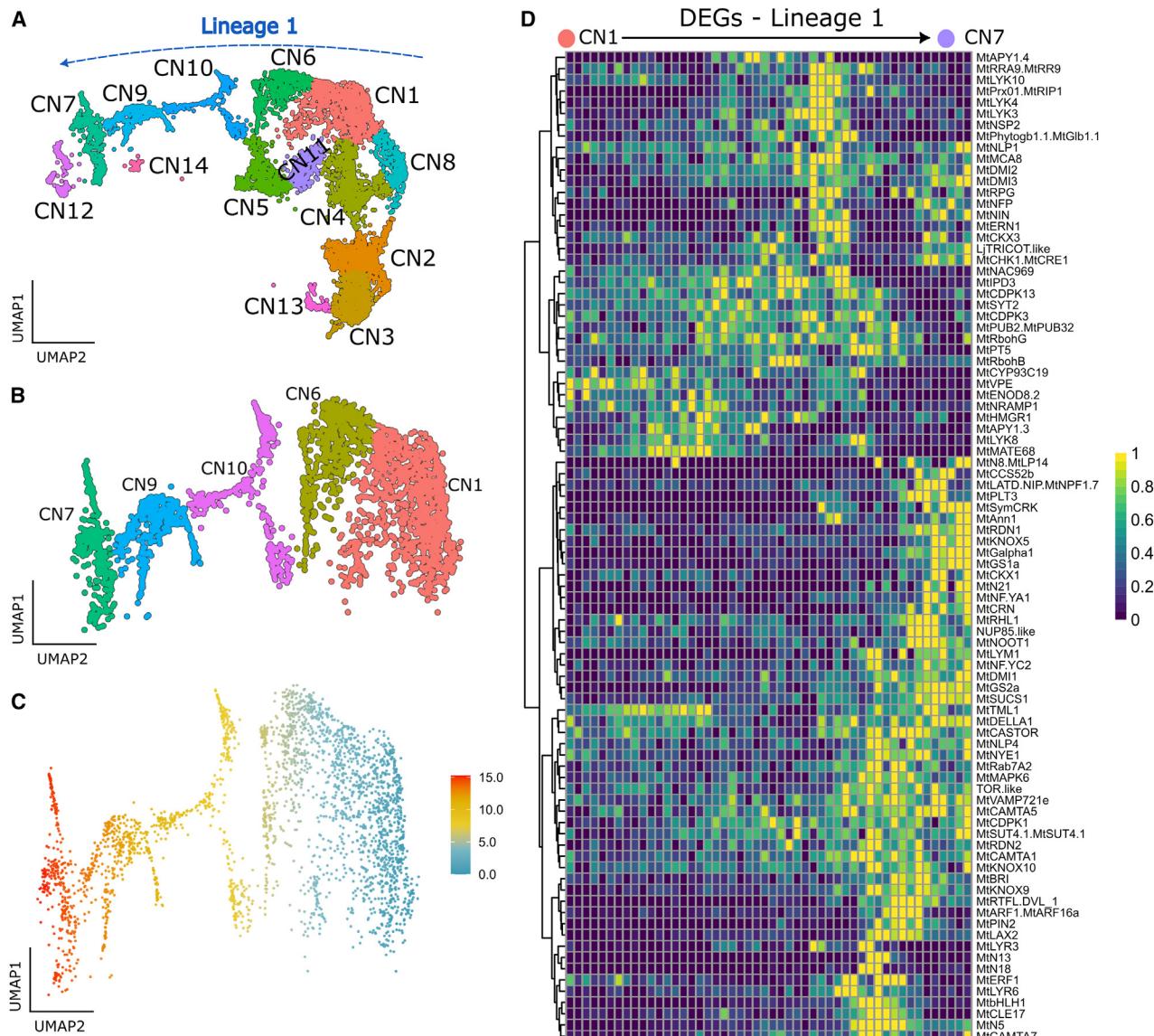


Figure 5. Cell lineage derived from the cortex forms the nodule meristem

(A) UMAP plot showing the fourteen groups of cells obtained after reclustering (see Figure 3).

(B) Cells from clusters CN1, CN6, CN10, CN9, and CN7 are part of the lineage originating the nodule meristem from cortical cells.

(C) The UMAP plot represents the cell distribution according to their pseudotimes.

(D) Expression profile of RNS genes detected as differentially expressed within the trajectory (TradeSeq, false discover rate < 0.001). Cells in the trajectory were binned into 50 continuous groups of equal size to facilitate data visualization. See also Table S5 and Figure S6.

CN14, a small cluster of 71 cells detected mostly at 96 hpi, expressing many of the genes detected in CN12. Among the genes specifically expressed in CN12 (Table S2), most encoded for proteins previously associated with RNS, such as nodulins and early nodulins (*MtN2*, *MtN6*, *MtN7*, *MtN15*, *MtN19*, *MtN20*, *MtN21*, *MtN26-4*, *MtENOD12*, and *MtENOD16*), NCRs (*MtNCR057*, *MtNCR247*, *MtNCR316*, *MtNCR514*, and *MtNCR657*), NFYs (*MtNFY-A1*, *MtNFY-A2*, *MtNF-YB16*, and *MtNF-YB18*), NPL, *NODULE-INDUCED RECEPTOR-LIKE KINASE 1* (*MtNRLK1*), *MtVPY*, *MtRPG*, *SYMBIOTIC PROTEIN KINASE 1* (*MtSPK*), *SYMBIOTIC CYS-RICH RECEPTOR-LIKE*

KINASE (*MtSymCRK*), and *MtSWEET13*, which are primarily expressed in the nodule infection zone (FIld fraction³³). Notably, a large number of genes expressed specifically in those infected cells remain uncharacterized (Table S2). For many of them, changes in the expression profile during RNS have also been captured in other published datasets (Figure 4B).

Weighted gene co-expression network analysis reveals regulators of RNS

The *M. truncatula* transcriptional response to rhizobia involves multiple cell types. Thus, an encompassing approach is

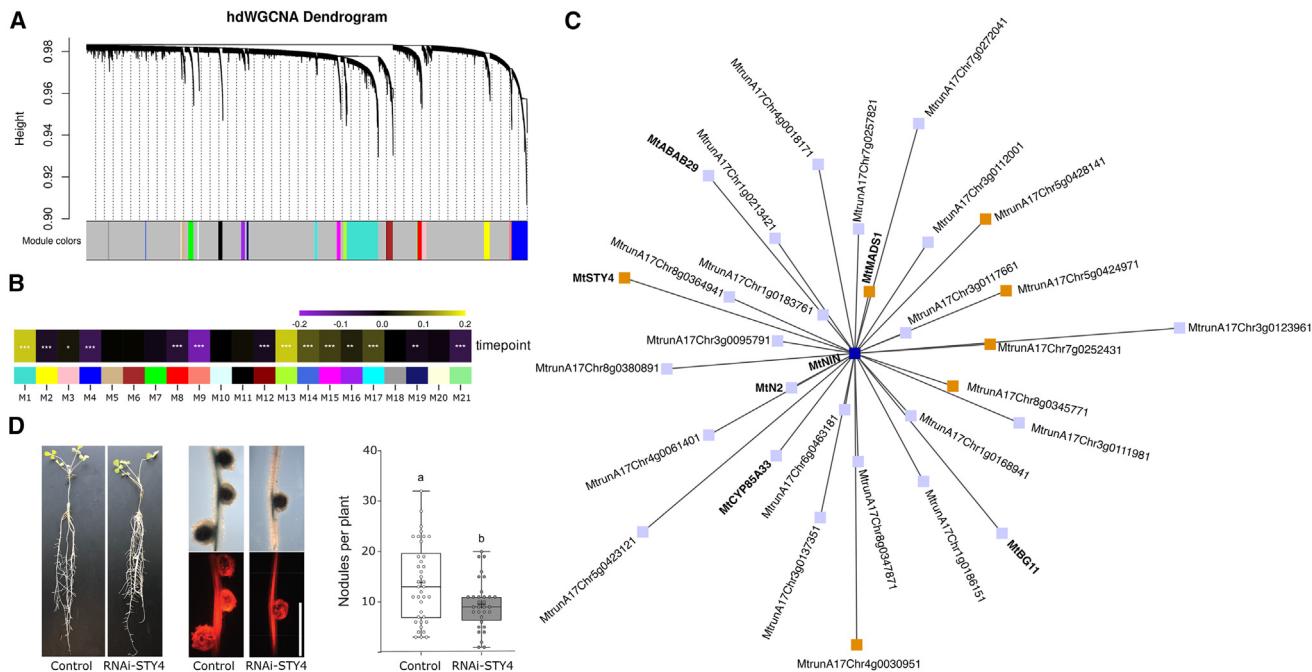


Figure 6. High-dimension WGCNA reveals new candidate genes in RNS

(A) Hierarchical clustering shows the modules of co-expressed genes, represented by different colors. Genes represented in gray were not assigned to any module.

(B) Relating the expression of the eigengene of each module with the time points reveals modules that are significantly correlated, either increasing (positively correlated, in yellow) or decreasing their expression (negatively correlated, in purple) as the infection progresses (* = $0.01 < p < 0.05$, ** = $0.001 < p < 0.01$, and *** = < 0.001).

(C) Weighted co-expression network representing the 30 genes most correlated to *MtNIN* (blue). Seven other transcription factors are present in this network and are represented in orange, including *MtSTY4*.

(D) Knockdown of *MtSTY4* causes a reduction in the number of nodules in *M. truncatula*. In the left and middle panels, composite plants inoculated with *S. meliloti* constitutively express *lacZ*. Detection of *lacZ* activity was used to monitor the successful nodule infection. Positive events of the transformation are identified as the ones emitting the red fluorescence of the dsRed protein. Scale bar: 2 mm. On the right is the nodule number of composite plants that carried the empty vector as a control (white box, $n = 38$ plants) or *MtSTY4* RNAi constructs (gray box, $n = 47$ plants) at 14 dai. For each box-and-whiskers plot, the center black line represents the median; “+” represents the mean; the box extends from the 25th to 75th percentiles; the whiskers are drawn down to the 10th percentile and up to the 90th. Different lowercase letters indicate significant differences, and p value of 0.005 was determined by the Student’s t test. See also Table S4.

necessary to detect gene networks and their regulators, instead of focusing on each cell type individually. We applied the weighted gene co-expression network analysis (WGCNA),⁷⁰ adapted to high-dimension datasets such as scRNA-seq (hdWGCNA⁷¹), to uncover transcriptional patterns related to *M. truncatula* response to infection. Using hdWGCNA, we detected 21 modules of co-expressed genes (Figure 6A; Table S4). The eigengenes (a statistical feature that describes the overall expression of genes within the module) of modules 1, 3, 13, 14, 15, 16, and 17 were positively correlated to the time points of the experiment, showing an overall expression increase as the infection progressed (Figure 6B). Seven modules were negatively correlated with the progression of the infection and may contain genes whose expression needs to be suppressed for the effective establishment of RNS (Figure 6B). All positively correlated modules have one or more known RNS genes (Table S4). As demonstrated in previous studies, using network metrics such as the degree of interconnectivity within a module can point to genes with important regulatory function⁷² or new candidates involved in the process.

hdWGCNA allows for interrogating what genes are more strongly correlated with any gene of interest in the network. Investigating the genes more strongly correlated with *MtNIN* revealed the network depicted in Figure 6C. For several genes in this network, their role in RNS has been demonstrated, such as *MADS box transcription factor* (*MtMADS1*)⁵⁰ and *NODULIN2* (*Mtn2*).⁷³ Another gene revealed by this analysis encodes a *LIGHT SENSITIVE SHORT HYPOCOTYL 1* (*MtLSH1*) hypothesized to be critical for nodule meristem formation in *M. truncatula*.⁶⁸ Six genes of this network are predicted to be transcription factors and are still uncharacterized, including *MtSTY4*, a member of the STYLISH family, homologs of the Arabidopsis gene *LATERAL ROOT PRIMORDIUM 1* (Figure 6C).

One or more members of the STYLISH family are hypothesized to be involved in the auxin regulation during lateral root⁷⁴ and nodule formation.⁶ To test if *MtSTY4* is indeed involved in RNS transgenic lines expressing *MtSTY4*, RNAi was generated. The knockdown of this gene caused a reduction in the number of nodules in *M. truncatula*, demonstrating the involvement of

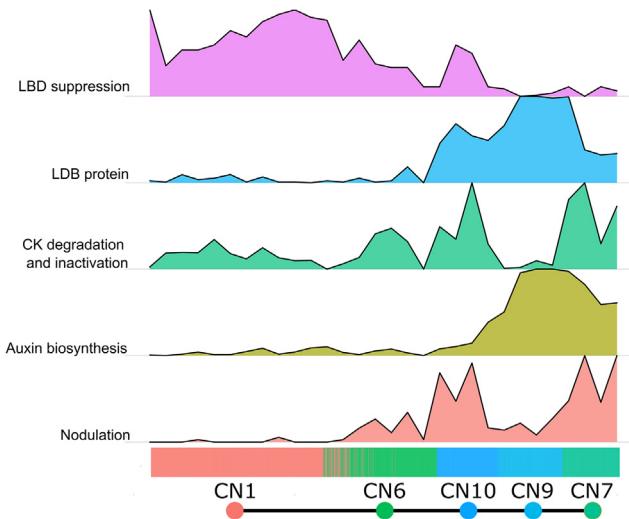


Figure 7. Relative expression of differentially expressed genes along the cell lineage trajectory from cortex to the nodule meristem
Cells in the trajectory (from CN1 to CN7) were sorted, and the expression of each DEG represented in the functional categories was binned as described previously (see Figure 6). The average expression of all genes in each category was calculated per bin and scaled for representation. The list of all DEGs used to generate the expression profile of each category is in Table S5.

MtSTY4 in RNS (Figure 6D). Whether *MtSTY4* is indeed involved in the auxin regulation or affects RNS in a different way remains to be experimentally investigated.

DISCUSSION

Here we present a single-cell transcriptome dataset encompassing various stages of RNS in *Medicago*. The plant cellular response to rhizobia affects most of the cell types in the root infection zone. However, only a limited fraction of cells within each of these types change their transcriptional state as only one or a few nodules typically develop in a root. Similarly, cell division and the establishment of cell lineages that form nodule compartments are constrained primarily to a few cell layers of the cortex and, to a smaller extent, the pericycle. To address the limitation imposed by the rarity of these cells within the root tissue, we integrated single-cell data from a wild-type genotype (A17 Jemalong) with that of the hypernodulating mutant *sunn-4*, significantly expanding the population of cells responding to the infection in this dataset. As a result, we were able to characterize the response to the rhizobia in subgroups of cells involved in RNS, including the infection of the epidermal root hair, the stelle activation of AON, and the differentiation of cortex cells into the nodule primordia and, later, its meristem. By isolating groups of cells responding to the infection within each of these cell types, we could identify the genes activated in their differentiation.

A significant advance enabled by single-cell genomics is the identification of cells that participate in lineages resulting in novel organs such as nodules and the transcriptome reprogramming that occurs as these cells differentiate. Our data describe for

the first time how gene expression is modulated along a cortex-derived cell lineage, resulting in the formation of a nodule meristem, and the interplay between auxin and CK-related genes where activation of one is often accompanied by mechanisms to suppress the other (Figure 7). For example, our data show that the gene *MtHB1* is highly expressed in the cortex cell clusters, coinciding with the downregulation of *LBD* genes. *MtHB1* suppresses the expression of *MtLBD1* and the establishment of an auxin maximum required for lateral root formation in *Medicago*.⁷⁵ The downregulation of *LBD* genes involved in nodulation suggests *MtHB1* plays a similar role in limiting the establishment of the auxin maxima required early in the inferred trajectory. The decrease in expression of *MtHB1* in the next steps of the cell lineage was followed by the broad activation of a suite of auxin biosynthesis genes (*MtSTY3*, *MtSTY4*, *MtSTY7*, *MtSTY8*, and *MtSTY9*) supporting the critical function of this phytohormone in the development of the nodule primordium. Finally, the specific activation of nodulation genes (e.g., *MtNIN*, *MtNF-YA1*, *MtNF-YB16*, and *MtNF-YB18*) in the developing nodule was accompanied by the expression of several genes involved in the inactivation (*MtAPT1* and *MtAPT2*⁷⁶) and degradation of CK (*MtCKX3* and *MtCKX7*⁷⁷). While the activation of several nodulation genes may require CK in the initial stages of nodule development,⁷⁸ the data point to suppression as the distinct compartments of the nodule become formed.

Limitations of the study

While this dataset uncovered the cells involved in the differentiation of cortex cells into a nodule meristem, the emergence of other nodule compartments remains to be studied. Furthermore, the time points when the data were collected did not capture cells from the ZIII. Finally, the lack of cohesiveness between clusters representing some of the cell types described in the UMAP plots suggests that our sampling procedure has not captured all lineages or developmental stages that lead to those differentiated cells, or that our conservative quality control may have excluded cells in some intermediate states.

Similarly, the lineage derived from the division of nodule meristematic cells that contribute to the development of the nodule remains to be uncovered. Despite these limitations, the discovery of cell types involved in nodule formation and associated lineages creates an opportunity for future single-cell comparative studies between lateral root and nodule development, as proposed recently.⁷⁹ Comparisons between the transcriptome of a developing lateral root and a nodule have been explored through bulk RNA sequencing.⁶ Expanding these inferences to the comparison of cell lineages⁷⁹ may identify where these lineages diverge to form distinct organs and point to the regulators required to modify the developmental program.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- **KEY RESOURCES TABLE**
- **RESOURCE AVAILABILITY**
 - Lead contact

- Materials availability
- Data and code availability
- **EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS**
 - Plant material and growth conditions
 - Rhizobia strain and growth conditions
- **METHOD DETAILS**
 - Inoculation assay and sample collection
 - Single nuclei RNA-seq and library preparation
 - Experimental validation of gene markers
 - Functional characterization of MtSTY4
- **QUANTIFICATION AND STATISTICAL ANALYSIS**
 - Quality control and cell clustering
 - Identification of novel *Medicago* cell-type-specific markers
 - Trajectory inference analysis
 - Weighted Correlation Network Analysis
 - Comparison of nodule number after MtSTY4 knock-down via RNAi

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at <https://doi.org/10.1016/j.cellrep.2024.113747>.

ACKNOWLEDGMENTS

We would like to thank Thomas B. Irving, currently at the University of Cambridge, for providing a list of genes involved in the root nodule symbiosis. We thank the United States Department of Energy, Office of Science, Biological and Environmental Research program for funding this study (DE-SC0018247 to M.K., S.R., and J.-M.A.). We also thank Clemson University's SUCCEEDS program for funding J.F. and F.A.F.

AUTHOR CONTRIBUTIONS

Conceptualization, C.D., D.C., F.A.F., J.F., J.-M.A., K.M.B., M.K., P.M.T., S.R., and W.J.P.; methodology, M.K. and W.J.P.; software, W.J.P.; validation, W.J.P.; formal analysis, J.B., K.M.B., and W.J.P.; investigation, C.N., C.D., D.C., H.W.S., J.B., P.M.T., and Y.G.; resources, J.F., J.-M.A., M.K., and S.R.; data curation, C.D., M.K., and W.J.P.; writing – original draft, C.D., J.B., M.K., and W.J.P.; writing – review & editing, C.N., C.D., D.C., F.A.F., H.W.S., J.B., J.F., K.M.B., M.T., M.K., N.C.P., P.M.T., S.C., S.A.K., S.R., T.B.I., Y.G., and W.J.P.; visualization, J.B., D.C., M.K., Y.G., and W.J.P.; supervision, C.D., J.F., J.-M.A., M.K., and S.R.; project administration, C.D., M.K., and W.J.P.; funding acquisition, J.F., F.A.F., J.-M.A., M.K., and S.R.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: June 21, 2023

Revised: October 31, 2023

Accepted: January 22, 2024

Published: February 7, 2024

REFERENCES

1. Breakspear, A., Liu, C., Roy, S., Stacey, N., Rogers, C., Trick, M., Morieri, G., Mysore, K.S., Wen, J., Oldroyd, G.E.D., et al. (2014). The root hair “infectorome” of *Medicago truncatula* uncovers changes in cell cycle genes and reveals a requirement for auxin signaling in rhizobial infection. *Plant Cell* 26, 4680–4701. <https://doi.org/10.1105/tpc.114.133496>.
2. Timmers, A.C., Auriac, M.C., and Truchet, G. (1999). Refined analysis of early symbiotic steps of the *Rhizobium-Medicago* interaction in relationship with microtubular cytoskeleton rearrangements. *Development* 126, 3617–3628. <https://doi.org/10.1242/dev.126.16.3617>.
3. Xiao, T.T., Schildeink, S., Moling, S., Deinum, E.E., Kondorosi, E., Franssen, H., Kulikova, O., Niebel, A., and Bisseling, T. (2014). Fate map of *Medicago truncatula* root nodules. *Development* 141, 3517–3528. <https://doi.org/10.1242/dev.110775>.
4. Roy, S., Liu, W., Nandety, R.S., Crook, A., Mysore, K.S., Pislaru, C.I., Frugoli, J., Dickstein, R., and Uvdardi, M.K. (2020). Celebrating 20 Years of Genetic Discoveries in Legume Nodulation and Symbiotic Nitrogen Fixation. *Plant Cell* 32, 15–41. <https://doi.org/10.1105/tpc.19.00279>.
5. Hirsch, A.M., Larue, T.A., and Doyle, J. (1997). Is the Legume Nodule a Modified Root or Stem or an Organ sui generis. *Crit. Rev. Plant Sci.* 16, 361–392. <https://doi.org/10.1080/07352689709701954>.
6. Schiessl, K., Lilley, J.L.S., Lee, T., Tamvakis, I., Kohlen, W., Bailey, P.C., Thomas, A., Luptak, J., Ramakrishnan, K., Carpenter, M.D., et al. (2019). *ODULE INCEPTION* Recruits the Lateral Root Developmental Program for Symbiotic Nodule Organogenesis in *Medicago truncatula*. *Curr. Biol.* 29, 3657–3668.e5, e5. <https://doi.org/10.1016/j.cub.2019.09.005>.
7. Lin, J., Frank, M., and Reid, D. (2020). No Home without Hormones: How Plant Hormones Control Legume Nodule Organogenesis. *Plant Commun.* 1, 100104. <https://doi.org/10.1016/J.XPLC.2020.100104>.
8. Jardinaud, M.F., Boivin, S., Rodde, N., Catrice, O., Kisiala, A., Lepage, A., Moreau, S., Roux, B., Cottret, L., Sallet, E., et al. (2016). A laser dissection-RNAseq analysis highlights the activation of cytokinin pathways by nod factors in the *Medicago truncatula* root epidermis. *Plant Physiol.* 171, 2256–2276. <https://doi.org/10.1104/pp.16.00711>.
9. Jarzyniak, K., Banasiak, J., Jamruszka, T., Pawela, A., Di Donato, M., Novák, O., Geisler, M., and Jasiński, M. (2021). Early stages of legume-rhizobia symbiosis are controlled by ABCG-mediated transport of active cytokinins. *Nat. Plants* 7, 428–436. <https://doi.org/10.1038/s41477-021-00873-6>.
10. Ferguson, B.J., Mens, C., Hastwell, A.H., Zhang, M., Su, H., Jones, C.H., Chu, X., and Gresshoff, P.M. (2019). Legume nodulation: The host controls the party. *Plant Cell Environ.* 42, 41–51. <https://doi.org/10.1111/pce.13348>.
11. Schnabel, E., Journet, E.-P., de Carvalho-Niebel, F., Duc, G., and Frugoli, J. (2005). The *Medicago truncatula* SUNN Gene Encodes a CLV1-like Leucine-rich Repeat Receptor Kinase that Regulates Nodule Number and Root Length. *Plant Mol. Biol.* 58, 809–822. <https://doi.org/10.1007/s11103-005-8102-y>.
12. Schnabel, E., Mukherjee, A., Smith, L., Kissaw, T., Long, S., and Frugoli, J. (2010). The lss supernodulation mutant of *Medicago truncatula* reduces expression of the SUNN gene. *Plant Physiol.* 154, 1390–1402. <https://doi.org/10.1104/pp.110.164889>.
13. Mergaert, P., Kereszt, A., and Kondorosi, E. (2020). Gene Expression in Nitrogen-Fixing Symbiotic Nodule Cells in *Medicago truncatula* and Other Nodulating Plants. *Plant Cell* 32, 42–68. <https://doi.org/10.1105/tpc.19.00494>.
14. Giacomello, S. (2021). A new era for plant science: spatial single-cell transcriptomics. *Curr. Opin. Plant Biol.* 60, 102041. <https://doi.org/10.1016/j.pbi.2021.102041>.
15. Saelens, W., Cannoodt, R., Todorov, H., and Saeys, Y. (2019). A comparison of single-cell trajectory inference methods. *Nat. Biotechnol.* 37, 547–554. <https://doi.org/10.1038/s41587-019-0071-9>.
16. Zhang, T.-Q., Chen, Y., and Wang, J.-W. (2021). A single-cell analysis of the *Arabidopsis* vegetative shoot apex. *Dev. Cell* 56, 1056–1074.e8. <https://doi.org/10.1016/j.devcel.2021.02.021>.
17. Conde, D., Triozzi, P.M., Pereira, W.J., Schmidt, H.W., Balmant, K.M., Knaack, S.A., Redondo-López, A., Roy, S., Dervinis, C., and Kirst, M. (2022). Single-nuclei transcriptome analysis of the shoot apex vascular

- system differentiation in *Populus*. *Development* 149, dev200632. <https://doi.org/10.1242/dev.200632>.
18. Liu, Z., Zhou, Y., Guo, J., Li, J., Tian, Z., Zhu, Z., Wang, J., Wu, R., Zhang, B., Hu, Y., et al. (2020). Global Dynamic Molecular Profiling of Stomatal Lineage Cell Development by Single-Cell RNA Sequencing. *Mol. Plant* 13, 1178–1193. <https://doi.org/10.1016/j.molp.2020.06.010>.
 19. Lopez-Anido, C.B., Vatén, A., Smoot, N.K., Sharma, N., Guo, V., Gong, Y., Anleu Gil, M.X., Weimer, A.K., and Bergmann, D.C. (2021). Single-cell resolution of lineage trajectories in the *Arabidopsis* stomatal lineage and developing leaf. *Dev. Cell* 56, 1043–1055.e4. <https://doi.org/10.1016/j.devcel.2021.03.014>.
 20. Chen, Y., Tong, S., Jiang, Y., Ai, F., Feng, Y., Zhang, J., Gong, J., Qin, J., Zhang, Y., Zhu, Y., et al. (2021). Transcriptional landscape of highly lignified poplar stems at single-cell resolution. *Genome Biol.* 22, 319. <https://doi.org/10.1186/s13059-021-02537-2>.
 21. Xie, J., Li, M., Zeng, J., Li, X., and Zhang, D. (2022). Single-cell RNA sequencing profiles of stem-differentiating xylem in poplar. *Plant Biotechnol. J.* 20, 417–419. <https://doi.org/10.1111/pbi.13763>.
 22. Cervantes-Pérez, S.A., Thibivilliers, S., Laffont, C., Farmer, A.D., Frugier, F., and Libault, M. (2022). Cell-specific pathways recruited for symbiotic nodulation in the *Medicago truncatula* legume. *Mol. Plant* 15, 1868–1888. <https://doi.org/10.1016/j.molp.2022.10.021>.
 23. Ye, Q., Zhu, F., Sun, F., Wang, T.-C., Wu, J., Liu, P., Shen, C., Dong, J., and Wang, T. (2022). Differentiation trajectories and biofunctions of symbiotic and un-symbiotic fate cells in root nodules of *Medicago truncatula*. *Mol. Plant* 15, 1852–1867. <https://doi.org/10.1016/j.molp.2022.10.019>.
 24. Liu, Z., Yang, J., Long, Y., Zhang, C., Wang, D., Zhang, X., Dong, W., Zhao, L., Liu, C., Zhai, J., and Wang, E. (2023). Single-nucleus transcriptomes reveal spatiotemporal symbiotic perception and early response in *Medicago*. *Nat. Plants* 9, 1734–1748. <https://doi.org/10.1038/s41477-023-01524-8>.
 25. Penmetsa, R.V., Frugoli, J.A., Smith, L.S., Long, S.R., and Cook, D.R. (2003). Dual Genetic Pathways Controlling Nodule Number in *Medicago truncatula*. *Plant Physiol.* 131, 998–1008. <https://doi.org/10.1104/pp.015677>.
 26. scRNA-seq of *Medicago truncatula* roots during Root Nodule Symbiosis. Shiny app - https://kirstlab.shinyapps.io/scrnaseq_medicago_truncatula/.
 27. Denyer, T., Ma, X., Klesen, S., Scacchi, E., Nieselt, K., and Timmermans, M.C.P. (2019). Spatiotemporal Developmental Trajectories in the *Arabidopsis* Root Revealed Using High-Throughput Single-Cell RNA Sequencing. *Dev. Cell* 48, 840–852.e5. <https://doi.org/10.1016/j.devcel.2019.02.022>.
 28. Shulse, C.N., Cole, B.J., Ciobanu, D., Lin, J., Yoshinaga, Y., Gouran, M., Turco, G.M., Zhu, Y., O’Malley, R.C., Brady, S.M., and Dickel, D.E. (2019). High-Throughput Single-Cell Transcriptome Profiling of Plant Cell Types. *Cell Rep.* 27, 2241–2247.e4. <https://doi.org/10.1016/j.celrep.2019.04.054>.
 29. Zhang, T.-Q., Xu, Z.-G., Shang, G.-D., and Wang, J.-W. (2019). A Single-Cell RNA Sequencing Profiles the Developmental Landscape of *Arabidopsis* Root. *Mol. Plant* 12, 648–660. <https://doi.org/10.1016/j.molp.2019.04.004>.
 30. Schnabel, E., Thomas, J., El-Hawaz, R., Gao, Y., Poehlman, W.L., Chavan, S., Pasha, A., Esteban, E., Provart, N., Feltus, F.A., and Frugoli, J. (2023). Laser capture microdissection transcriptome reveals spatiotemporal tissue gene expression patterns of *M. truncatula* roots responding to rhizobia. *MPMI (Mol. Plant-Microbe Interact.)* 36, 805–820. <https://doi.org/10.1094/MPMI-03-23-0029-R>.
 31. Damiani, I., Drain, A., Guichard, M., Balzergue, S., Boscarini, A., Boyer, J.-C., Brunaud, V., Cottaz, S., Rancurel, C., Da Rocha, M., et al. (2016). Nod Factor Effects on Root Hair-Specific Transcriptome of *Medicago truncatula*: Focus on Plasma Membrane Transport Systems and Reactive Oxygen Species Networks. *Front. Plant Sci.* 7, 794. <https://doi.org/10.3389/fpls.2016.00794>.
 32. Marino, D., Andrio, E., Danchin, E.G.J., Oger, E., Gucciardo, S., Lambert, A., Puppo, A., and Pauly, N. (2011). A *Medicago truncatula* NADPH oxidase is involved in symbiotic nodule functioning. *New Phytol.* 189, 580–592. <https://doi.org/10.1111/j.1469-8137.2010.03509.x>.
 33. Roux, B., Rodde, N., Jardinaud, M.-F., Timmers, T., Sauviac, L., Cottret, L., Carrère, S., Sallet, E., Courcelle, E., Moreau, S., et al. (2014). 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. <https://doi.org/10.1111/tpj.12442>.
 34. Song, J.H., Montes-Luz, B., Tadra-Sfeir, M.Z., Cui, Y., Su, L., Xu, D., and Stacey, G. (2022). High-Resolution Translatome Analysis Reveals Cortical Cell Programs During Early Soybean Nodulation. *Front. Plant Sci.* 13, 820348. <https://doi.org/10.3389/fpls.2022.820348>.
 35. Dong, W., Zhu, Y., Chang, H., Wang, C., Yang, J., Shi, J., Gao, J., Yang, W., Lan, L., Wang, Y., et al. (2021). An SHR-SCR module specifies legume cortical cell fate to enable nodulation. *Nature* 589, 586–590. <https://doi.org/10.1038/s41586-020-3016-z>.
 36. Persson, S., Caffall, K.H., Freshour, G., Hilly, M.T., Bauer, S., Poindexter, P., Hahn, M.G., Mohnen, D., and Somerville, C. (2007). The *Arabidopsis* irregular xylem8 Mutant Is Deficient in Glucuronoxylan and Homogalacturonan, Which Are Essential for Secondary Cell Wall Integrity. *Plant Cell* 19, 237–255. <https://doi.org/10.1105/tpc.106.047720>.
 37. Lasserre, E., Jobet, E., Llauro, C., and Delseny, M. (2008). AtERF38 (At2g35700), an AP2/ERF family transcription factor gene from *Arabidopsis thaliana*, is expressed in specific cell types of roots, stems and seeds that undergo suberization. *Plant Physiol. Biochem.* 46, 1051–1061. <https://doi.org/10.1016/j.plaphy.2008.07.003>.
 38. Zhang, J., Eswaran, G., Alonso-Serra, J., Kucukoglu, M., Xiang, J., Yang, W., Elo, A., Nieminen, K., Damén, T., Joung, J.-G., et al. (2019). Transcriptional regulatory framework for vascular cambium development in *Arabidopsis* roots. *Nat. Plants* 5, 1033–1042. <https://doi.org/10.1038/s41477-019-0522-9>.
 39. Lin, S.H., Kuo, H.F., Canivenc, G., Lin, C.S., Lepetit, M., Hsu, P.K., Tillard, P., Lin, H.L., Wang, Y.Y., Tsai, C.B., et al. (2008). Mutation of the *Arabidopsis* NRT1.5 nitrate transporter causes defective root-to-shoot nitrate transport. *Plant Cell* 20, 2514–2528. <https://doi.org/10.1105/tpc.108.060244>.
 40. Arrighi, J.-F., Godfroy, O., de Billy, F., Saurat, O., Jauneau, A., and Gough, C. (2008). The RPG gene of *Medicago truncatula* controls Rhizobium-directed polar growth during infection. *Proc. Natl. Acad. Sci. USA* 105, 9817–9822. <https://doi.org/10.1073/pnas.0710273105>.
 41. Sinharoy, S., Liu, C., Breakspear, A., Guan, D., Shailes, S., Nakashima, J., Zhang, S., Wen, J., Torres-Jerez, I., Oldroyd, G., et al. (2016). A *Medicago truncatula* Cystathione-β-Synthase-like Domain-Containing Protein Is Required for Rhizobial Infection and Symbiotic Nitrogen Fixation. *Plant Physiol.* 170, 2204–2217. <https://doi.org/10.1104/pp.15.01853>.
 42. Xie, F., Murray, J.D., Kim, J., Heckmann, A.B., Edwards, A., Oldroyd, G.E.D., and Downie, J.A. (2012). Legume pectate lyase required for root infection by rhizobia. *Proc. Natl. Acad. Sci. USA* 109, 633–638. <https://doi.org/10.1073/pnas.1113992109>.
 43. Liang, P., Stratil, T.F., Popp, C., Marín, M., Folgmann, J., Mysore, K.S., Wen, J., and Ott, T. (2018). Symbiotic root infections in *Medicago truncatula* require remorin-mediated receptor stabilization in membrane nanodomains. *Proc. Natl. Acad. Sci. USA* 115, 5289–5294. <https://doi.org/10.1073/pnas.1721868115>.
 44. Murray, J.D., Muni, R.R.D., Torres-Jerez, I., Tang, Y., Allen, S., Andrianakaja, M., Li, G., Laxmi, A., Cheng, X., Wen, J., et al. (2011). Vapyrin, a gene essential for intracellular progression of arbuscular mycorrhizal symbiosis, is also essential for infection by rhizobia in the nodule symbiosis of *Medicago truncatula*. *Plant J.* 65, 244–252. <https://doi.org/10.1111/j.1365-313X.2010.04415.x>.
 45. Limpens, E., Franken, C., Smit, P., Willemse, J., Bisseling, T., and Geurts, R. (2003). LysM Domain Receptor Kinases Regulating Rhizobial Nod Factor-Induced Infection. *Science* 302, 630–633. <https://doi.org/10.1126/science.1090074>.

46. McAdam, E.L., Hugill, C., Fort, S., Samain, E., Cottaz, S., Davies, N.W., Reid, J.B., and Foo, E. (2017). Determining the Site of Action of Strigolactones during Nodulation. *Plant Physiol.* 175, 529–542. <https://doi.org/10.1104/pp.17.00741>.
47. Mashiguchi, K., Seto, Y., and Yamaguchi, S. (2021). Strigolactone biosynthesis, transport and perception. *Plant J.* 105, 335–350. <https://doi.org/10.1111/tpj.15059>.
48. Meng, J., Yang, J., Peng, M., Liu, X., and He, H. (2020). Genome-Wide Characterization, Evolution, and Expression Analysis of the Leucine-Rich Repeat Receptor-Like Protein Kinase (LRR-RLK) Gene Family in *Medicago truncatula*. *Life* 10, 176. <https://doi.org/10.3390/life10090176>.
49. Yu, H., Xiao, A., Dong, R., Fan, Y., Zhang, X., Liu, C., Wang, C., Zhu, H., Duanmu, D., Cao, Y., and Zhang, Z. (2018). Suppression of innate immunity mediated by the CDPK-Rboh complex is required for rhizobial colonization in *Medicago truncatula* nodules. *New Phytol.* 220, 425–434. <https://doi.org/10.1111/nph.15410>.
50. Moreau, S., Verdenaud, M., Ott, T., Letort, S., de Billy, F., Niebel, A., Gouzy, J., de Carvalho-Niebel, F., and Gamas, P. (2011). Transcription Re-programming during Root Nodule Development in *Medicago truncatula*. *PLoS One* 6, e16463. <https://doi.org/10.1371/journal.pone.0016463>.
51. Barnett, M.J., Toman, C.J., Fisher, R.F., and Long, S.R. (2004). A dual-genome Symbiosis Chip for coordinate study of signal exchange and development in a prokaryote-host interaction. *Proc. Natl. Acad. Sci. USA* 101, 16636–16641. <https://doi.org/10.1073/pnas.0407269101>.
52. Combier, J.-P., Frugier, F., de Billy, F., Boualem, A., El-Yahyaoui, F., Moreau, S., Vernié, T., Ott, T., Gamas, P., Crespi, M., and Niebel, A. (2006). *MtHAP2-1* is a key transcriptional regulator of symbiotic nodule development regulated by microRNA169 in *Medicago truncatula*. *Genes Dev.* 20, 3084–3088. <https://doi.org/10.1101/gad.402806>.
53. Triozzi, P.M., Irving, T.B., Schmidt, H.W., Keyser, Z.P., Chakraborty, S., Balmant, K., Pereira, W.J., Dervinis, C., Mysore, K.S., Wen, J., et al. (2022). Spatiotemporal cytokinin response imaging and *I/SOPENTENYL-TRANSFERASE 3* function in *Medicago* nodule development. *Plant Physiol.* 188, 560–575. <https://doi.org/10.1093/plphys/kiab447>.
54. Mortier, V., Den Herder, G., Whitford, R., Van de Velde, W., Rombauts, S., D'haeseleer, K., Holsters, M., and Goormachtig, S. (2010). CLE Peptides Control *Medicago truncatula* Nodulation Locally and Systemically. *Plant Physiol.* 153, 222–237. <https://doi.org/10.1104/pp.110.153718>.
55. Okamoto, S., Tabata, R., and Matsubayashi, Y. (2016). Long-distance peptide signaling essential for nutrient homeostasis in plants. *Curr. Opin. Plant Biol.* 34, 35–40. <https://doi.org/10.1016/j.pbi.2016.07.009>.
56. Schnabel, E.L., Chavan, S.A., Gao, Y., Poehlman, W.L., Feltus, F.A., and Frugoli, J.A. (2023). A *Medicago truncatula* Autoregulation of Nodulation Mutant Transcriptome Analysis Reveals Disruption of the SUNN Pathway Causes Constitutive Expression Changes in Some Genes, but Overall Response to Rhizobia Resembles Wild-Type, Including Induction of TML1 and TML2. *Curr. Issues Mol. Biol.* 45, 4612–4631. <https://doi.org/10.3390/cimb45060293>.
57. Denny, T.P., Ganova-Raeva, L.M., Huang, J., and Schell, M.A. (1996). Cloning and Characterization of *tek*, the Gene Encoding the Major Extracellular Protein of *Pseudomonas solanacearum*. *MPMI* (Mol. Plant-Microbe Interact.) 9, 272–281. <https://doi.org/10.1094/MPMI-9-0272>.
58. Godiard, L., Lepage, A., Moreau, S., Laporte, D., Verdenaud, M., Timmers, T., and Gamas, P. (2011). *MtbHLH1*, a bHLH transcription factor involved in *Medicago truncatula* nodule vascular patterning and nodule to plant metabolic exchanges. *New Phytol.* 191, 391–404. <https://doi.org/10.1111/j.1469-8137.2011.03718.x>.
59. Roy, S., Robson, F., Lilley, J., Liu, C.-W., Cheng, X., Wen, J., Walker, S., Sun, J., Cousins, D., Bone, C., et al. (2017). *MtLAX2*, a Functional Homologue of the Arabidopsis Auxin Influx Transporter AUX1, Is Required for Nodule Organogenesis. *Plant Physiol.* 174, 326–338. <https://doi.org/10.1104/pp.16.01473>.
60. Franssen, H.J., Xiao, T.T., Kulikova, O., Wan, X., Bisseling, T., Scheres, B., and Heidstra, R. (2015). Root developmental programs shape the *Medicago truncatula* nodule meristem. *Development* 142, 2941–2950. <https://doi.org/10.1242/dev.120774>.
61. Magne, K., Couzigou, J.-M., Schiessl, K., Liu, S., George, J., Zhukov, V., Sahl, L., Boyer, F., Iantcheva, A., Mysore, K.S., et al. (2018). *MtNODULE ROOT1* and *MtNODULE ROOT2* Are Essential for Indeterminate Nodule Identity. *Plant Physiol.* 178, 295–316. <https://doi.org/10.1104/pp.18.00610>.
62. Van de Velde, W., Zehirov, G., Szatmari, A., Debreczeny, M., Ishihara, H., Kevei, Z., Farkas, A., Mikulass, K., Nagy, A., Tiricz, H., et al. (2010). Plant Peptides Govern Terminal Differentiation of Bacteria in Symbiosis. *Science* 327, 1122–1126. <https://doi.org/10.1126/science.1184057>.
63. Laporte, P., Lepage, A., Fournier, J., Catrice, O., Moreau, S., Jardinaud, M.-F., Mun, J.-H., Larrainzar, E., Cook, D.R., Gamas, P., and Niebel, A. (2014). The CCAAT box-binding transcription factor NF-YA1 controls rhizobial infection. *J. Exp. Bot.* 65, 481–494. <https://doi.org/10.1093/jxb/ert392>.
64. Farkas, A., Maróti, G., Durgó, H., Györgypál, Z., Lima, R.M., Medzihradszky, K.F., Kereszt, A., Mergaert, P., and Kondorosi, É. (2014). *Medicago truncatula* symbiotic peptide NCR247 contributes to bacteroid differentiation through multiple mechanisms. *Proc. Natl. Acad. Sci. USA* 111, 5183–5188. <https://doi.org/10.1073/pnas.1404169111>.
65. Garcia, K., Cloghessy, K., Cooney, D.R., Shelley, B., Chakraborty, S., Kafle, A., Busidan, A., Sonawala, U., Collier, R., Jayaraman, D., et al. (2023). The putative transporter *MtUMAMIT14* participates in nodule formation in *Medicago truncatula*. *Sci. Rep.* 13, 804. <https://doi.org/10.1038/s41598-023-28160-8>.
66. Kryvoruchko, I.S., Sinharoy, S., Torres-Jerez, I., Sosso, D., Pislaru, C.I., Guan, D., Murray, J., Benedito, V.A., Frommer, W.B., and Udvardi, M.K. (2016). *MtSWEET11*, a Nodule-Specific Sucrose Transporter of *Medicago truncatula*. *Plant Physiol.* 171, 554–565. <https://doi.org/10.1104/pp.15.01910>.
67. Street, K., Risso, D., Fletcher, R.B., Das, D., Ngai, J., Yosef, N., Purdom, E., and Dudoit, S. (2018). Slingshot: cell lineage and pseudotime inference for single-cell transcriptomics. *BMC Genom.* 19, 477. <https://doi.org/10.1186/s12864-018-4772-0>.
68. Schiessl, K., Lee, T., Orvosova, M., Bueno-Batista, M., Stuer, N., Bailey, P.C., Mysore, K.S., Wen, J., and Oldroyd, G.E.D. (2023). Light sensitive short hypocotyl (LSH) confer symbiotic nodule identity in the legume *Medicago truncatula*. Preprint at bioRxiv. <https://doi.org/10.1101/2023.02.12.528179>.
69. Van den Berge, K., Roux de Bézieux, H., Street, K., Saelens, W., Cannoodt, R., Saeys, Y., Dudoit, S., and Clement, L. (2020). Trajectory-based differential expression analysis for single-cell sequencing data. *Nat. Commun.* 11, 1201. <https://doi.org/10.1038/s41467-020-14766-3>.
70. Langfelder, P., and Horvath, S. (2008). WGCNA: an R package for weighted correlation network analysis. *BMC Bioinf.* 9, 559. <https://doi.org/10.1186/1471-2105-9-559>.
71. Morabito, S., Reese, F., Rahimzadeh, N., Miyoshi, E., and Swarup, V. (2023). hdWGCNA identifies co-expression networks in high-dimensional transcriptomics data. *Cell Rep. Methods* 3, 100498. <https://doi.org/10.1016/j.crmeth.2023.100498>.
72. Balmant, K.M., Noble, J.D., C Alves, F., Dervinis, C., Conde, D., Schmidt, H.W., Vazquez, A.I., Barbazuk, W.B., Campos, G.d.L., Resende, M.F.R., and Kirst, M. (2020). Xylem systems genetics analysis reveals a key regulator of lignin biosynthesis in *Populus deltoides*. *Genome Res.* 30, 1131–1143. <https://doi.org/10.1101/gr.261438.120>.
73. Gamas, P., Niebel, F. de C., Lescure, N., and Cullimore, J. (1996). Use of a subtractive hybridization approach to identify new *Medicago truncatula* genes induced during root nodule development. *Mol. Plant Microbe Interact.* 9, 233–242. <https://doi.org/10.1094/mpmi-9-0233>.
74. Singh, S., Yadav, S., Singh, A., Mahima, M., Singh, A., Gautam, V., and Sarkar, A.K. (2020). Auxin signaling modulates *LATERAL ROOT PRIMORDIUM1* (LRP1) expression during lateral root development in *Arabidopsis*. *Plant J.* 101, 87–100. <https://doi.org/10.1111/tpj.14520>.

75. Ariel, F., Diet, A., Verdenaud, M., Gruber, V., Frugier, F., Chan, R., and Crespi, M. (2010). Environmental Regulation of Lateral Root Emergence in *Medicago truncatula* Requires the HD-Zip I Transcription Factor HB1. *Plant Cell* 22, 2171–2183. <https://doi.org/10.1105/tpc.110.074823>.
76. Zhang, X., Chen, Y., Lin, X., Hong, X., Zhu, Y., Li, W., He, W., An, F., and Guo, H. (2013). Adenine Phosphoribosyl Transferase 1 is a Key Enzyme Catalyzing Cytokinin Conversion from Nucleobases to Nucleotides in Arabidopsis. *Mol. Plant* 6, 1661–1672. <https://doi.org/10.1093/mp/sst071>.
77. Schmülling, T., Werner, T., Riefler, M., Krupková, E., and Bartrina y Manns, I. (2003). Structure and function of cytokinin oxidase/dehydrogenase genes of maize, rice, Arabidopsis and other species. *J. Plant Res.* 116, 241–252. <https://doi.org/10.1007/s10265-003-0096-4>.
78. van Zeijl, A., Op den Camp, R.H.M., Deinum, E.E., Charnikhova, T., Franssen, H., Op den Camp, H.J.M., Bouwmeester, H., Kohlen, W., Bisseling, T., and Geurts, R. (2015). Rhizobium Lipo-chito-oligosaccharide Signaling Triggers Accumulation of Cytokinins in *Medicago truncatula* Roots. *Mol. Plant* 8, 1213–1226. <https://doi.org/10.1016/j.molp.2015.03.010>.
79. Conde, D., and Kirst, M. (2022). Decoding exceptional plant traits by comparative single-cell genomics. *Trends Plant Sci.* 27, 1095–1098. <https://doi.org/10.1016/j.tplants.2022.08.006>.
80. GEO Accession viewer - <https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE224539>.
81. GitHub - https://github.com/KirstLab/scrnaseq_medicago_A17_sun4.
82. Pereira, W.J. (2023). KirstLab/scrnaseq_medicago_A17_sun4, Zenodo. <https://doi.org/10.5281/zenodo.7712986>. <https://zenodo.org/doi/10.5281/zenodo.7712985>.
83. FigShare - Complementary Files of the Single-Cell Transcriptome Program of Nodule Development Cellular Lineages in *Medicago truncatula*. <https://doi.org/10.6084/m9.figshare.24461014.v2>. 10.6084/m9.figshare.24461014.v2.
84. CYG germination pouches (2014). CYG Seed Germination Pouch. <https://mega-international.com/tech-info/>.
85. Chakraborty, S., Driscoll, H.E., Abrahante, J.E., Zhang, F., Fisher, R.F., and Harris, J.M. (2021). Salt Stress Enhances Early Symbiotic Gene Expression in *Medicago truncatula* and Induces a Stress-Specific Set of Rhizobium-Responsive Genes. *MPMI (Mol. Plant-Microbe Interact.)* 34, 904–921. <https://doi.org/10.1094/MPMI-01-21-0019-R>.
86. Boisson-Dernier, A., Chabaud, M., Garcia, F., Bécard, G., Rosenberg, C., and Barker, D.G. (2001). Agrobacterium rhizogenes-Transformed Roots of *Medicago truncatula* for the Study of Nitrogen-Fixing and Endomycorrhizal Symbiotic Associations. *MPMI (Mol. Plant-Microbe Interact.)* 14, 695–700. <https://doi.org/10.1094/MPMI.2001.14.6.695>.
87. Conde, D., Triozzi, P.M., Balmant, K.M., Doty, A.L., Miranda, M., Boullosa, A., Schmidt, H.W., Pereira, W.J., Dervinis, C., and Kirst, M. (2021). A robust method of nuclei isolation for single-cell RNA sequencing of solid tissues from the plant genus Populus. *PLoS One* 16, e0251149. <https://doi.org/10.1371/journal.pone.0251149>.
88. Triozzi, P.M., Schmidt, H.W., Dervinis, C., Kirst, M., and Conde, D. (2021). Simple, efficient and open-source CRISPR/Cas9 strategy for multi-site genome editing in *Populus tremula × alba*. *Tree Physiol.* 41, 2216–2227. <https://doi.org/10.1093/treephys/tpab066>.
89. Pecrix, Y., Staton, S.E., Sallet, E., Lelandais-Brière, C., Moreau, S., Carrère, S., Blein, T., Jardinaud, M.-F., Latrasse, D., Zouine, M., et al. (2018). Whole-genome landscape of *Medicago truncatula* symbiotic genes. *Nat. Plants* 4, 1017–1025. <https://doi.org/10.1038/s41477-018-0286-7>.
90. Conde, D., González-Melendi, P., and Allona, I. (2013). Poplar stems show opposite epigenetic patterns during winter dormancy and vegetative growth. *Trees (Berl.)* 27, 311–320. <https://doi.org/10.1007/s00468-012-0800-x>.
91. Leong, S.A., Williams, P.H., and Ditta, G.S. (1985). Analysis of the 5' regulatory region of the gene for δ-aminolevulinic acid synthetase of *Rhizobium meliloti*. *Nucleic Acids Res.* 13, 5965–5976. <https://doi.org/10.1093/nar/13.16.5965>.
92. Lun, A.T.L., Riesenfeld, S., Andrews, T., Dao, T.P., Gomes, T., and participants in the 1st Human Cell Atlas Jamboree; and Marioni, J.C. (2019). EmptyDrops: distinguishing cells from empty droplets in droplet-based single-cell RNA sequencing data. *Genome Biol.* 20, 63. <https://doi.org/10.1186/s13059-019-1662-y>.
93. Monocle (2023). <https://github.com/cole-trapnell-lab/monocle3>, 3.
94. Germain, P.-L., Lun, A., Garcia Meixide, C., Macnair, W., and Robinson, M.D. (2022). Doublet identification in single-cell sequencing data using scDblFinder. *F1000Res.* 10, 979. <https://doi.org/10.12688/f1000research.73600.2>.
95. Traag, V.A., Waltman, L., and van Eck, N.J. (2019). From Louvain to Leiden: guaranteeing well-connected communities. *Sci. Rep.* 9, 5233. <https://doi.org/10.1038/s41598-019-41695-z>.
96. Haghverdi, L., Lun, A.T.L., Morgan, M.D., and Marioni, J.C. (2018). Batch effects in single-cell RNA-sequencing data are corrected by matching mutual nearest neighbors. *Nat. Biotechnol.* 36, 421–427. <https://doi.org/10.1038/nbt.4091>.
97. Cao, J., Spielmann, M., Qiu, X., Huang, X., Ibrahim, D.M., Hill, A.J., Zhang, F., Mundlos, S., Christiansen, L., Steemers, F.J., et al. (2019). The single-cell transcriptional landscape of mammalian organogenesis. *Nature* 566, 496–502. <https://doi.org/10.1038/s41586-019-0969-x>.
98. Carrere, S., Verdier, J., and Gamas, P. (2021). MtExpress, a Comprehensive and Curated RNAseq-based Gene Expression Atlas for the Model Legume *Medicago truncatula*. *Plant Cell Physiol.* 62, 1494–1500. <https://doi.org/10.1093/pcp/pcab110>.
99. Hao, Y., Hao, S., Andersen-Nissen, E., Mauck, W.M., Zheng, S., Butler, A., Lee, M.J., Wilk, A.J., Darby, C., Zager, M., et al. (2021). Integrated analysis of multimodal single-cell data. *Cell* 184, 3573–3587.e29. <https://doi.org/10.1016/j.cell.2021.04.048>.

STAR★METHODS

KEY RESOURCES TABLE

| REAGENT or RESOURCE | SOURCE | IDENTIFIER |
|------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------|
| Bacterial and virus strains | | |
| <i>Sinorhizobium meliloti</i> 1021 | Dr. Jean Michel-Ané, University of Wisconsin-Madison | N/A |
| <i>Agrobacterium rhizogenes</i> MSU440 | Dr. Jean Michel-Ané, University of Wisconsin-Madison | N/A |
| Biological samples | | |
| <i>Medicago truncatula</i> Jemalog A17 | Dr. Jean Michel-Ané, University of Wisconsin-Madison | N/A |
| <i>Medicago truncatula</i> sunn-4 mutant | Dr. Julia Frugoli, Clemson University | N/A |
| Chemicals, peptides, and recombinant proteins | | |
| Protector RNase Inhibitor | Sigma Aldrich | SKU #3335399001 |
| DAPI (4',6-diamidino-2-phenylindole) | Sigma Aldrich | PubChem Substance ID 24894305, SKU #D9542-5MG |
| BD FACS Flow Sheath Buffer | BD Biosciences | catalog number: 342003 |
| RNase inhibitor RibolockTM | Thermo Fisher Scientific | catalog number: EO0382 |
| Critical commercial assays | | |
| CYG germination pouches | https://mega-international.com/tech-info/ | N/A |
| Polypap stands | https://mega-international.com/tech-info/ | N/A |
| Humidity dome | Alberta LTD | Amazon Standard Identification Number: B089Q24D3N |
| 40 µm cell strainer | Greiner Bio-One | Item No.: 542040 |
| Miracloth | Calbiochem | Cat# 475855 |
| Chromium Next GEM Chip G Single Cell Kit, 16 rxn | 10x Genomics | 1000127 |
| MoClo Toolkit | Addgene | Kit #1000000044 |
| Kulzer Technovit 7100 resin | Electron Microscopy Sciences | 14653 |
| Phusion Taq | NEB | M0530S |
| Monarch DNA Plasmid Prep Kit | NEB | T1010S |
| pENTR Directional TOPO vector | Thermo Fisher | K240020 |
| Gatewa LR Clonas II Enzyme mix | Thermo Fisher | 11791100 |
| 10X Genomics Library Construction Kit | 10X Genomics | 1000190 |
| 10X Genomics Chromium Next GEM Single Cell 3' Kit v3.1, 4 rxns | 10X Genomics | 1000269 |
| MoClo Plant Parts Kit | Addgene | 1000000047 |
| Monarch DNA Gel Extraction Kit | NEB | T1020S |
| Deposited data | | |
| Single-cell datasets generated in this study. | NCBI's Gene Expression Omnibus | GSE224539 |
| Normalized expression data from Schiessl et al., 2019 and Roux et al., 2014. | Medicago truncatula RNA-seq Gene Expression Atlas Project | https://lpm-browsers.toulouse.inra.fr/pub/expressionAtlas/app/v3/ |
| LCM dataset | NCBI - Sequence Read Archive | PRJNA704996 |
| Clustered datasets | Figshare | https://doi.org/10.6084/m9.figshare.24461014 |
| <i>Medicago truncatula</i> reference genome | INRAE | https://medicago.toulouse.inra.fr/MtrunA17r5.0-ANR/ |
| Oligonucleotides | | |
| FWD 5'-CACCCATGAACCAGCAAAGAATGGACG-3' | this manuscript | N/A |
| REV 5'-AGTGCAAGAAATTAAACAAAG-3' | this manuscript | N/A |

(Continued on next page)

Continued

| REAGENT or RESOURCE | SOURCE | IDENTIFIER |
|-------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------|
| Recombinant DNA | | |
| pK7GW1GW2(II)-RedRoot | THE VIB-UGENT PSB PLASMID REPOSITORY | Vector ID: 3_27 |
| Software and algorithms | | |
| R v.4.1.2 | https://cran.r-project.org/ | N/A |
| Cell Ranger v. 7.0.1 | https://www.10xgenomics.com/ | N/A |
| Velocyto 0.17 | https://velocyto.org/ | N/A |
| Source code | https://github.com/KirstLab/scrnaseq_medicago_A17_sunn4 https://doi.org/10.5281/zenodo.7712986 | N/A |

RESOURCE AVAILABILITY

Lead contact

All additional information and requests for resources, reagents, and methods should be directed to the lead contact, Dr. Matias Kirst (mkirst@ufl.edu).

Materials availability

All unique reagents generated in this study will be made available from the [lead contact](#) (M. K.) and may require a completed materials transfer agreement.

Data and code availability

- The datasets supporting the conclusions of this article are available in the NCBI's Gene Expression Omnibus and are accessible through accession number GSE224539⁸⁰. Those include the raw data and the count matrices produced after processing using Cell Ranger.
- The source code to reproduce the analysis is available in GitHub⁸¹ and Zenodo.⁸² The clustered datasets were deposited in the R Data Serialization (RDS) format, which allows loading of the data into R and recovering any necessary information and is available on FigShare.⁸³
- Any additional information required to reanalyze the data reported in this paper is available from the [lead contact](#) upon request.

EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS

Plant material and growth conditions

In this study, two genotypes of *M. truncatula* were used, Jemalong A17 and a mutant for the gene SUNN, namely *sunn-4*.¹¹ The mutant *sunn-4* presents a hypernodulation phenotype.

M. truncatula seeds were scarified and stratified to initiate germination. Scarification was done by soaking seeds in sulfuric acid for 8 min before washing them with water 6 times. Seeds were sterilized in bleach (12% sodium hypochlorite) for 4 min, washed with sterile water 4 times, and incubated for 1 h in water. Seeds were sown on 1% agar plates with 1 µM of GA₃. The plates were kept at 4°C for three days to induce stratification. After incubating plates at 24°C for 12 h in the dark, four seedlings were transferred to CYG germination pouches.⁸⁴ Each pouch contained four seedlings and 10 mL of Modified Nodulation Medium.⁸⁵ Pouches were placed vertically in polypap stands,⁸⁴ and the stands were placed in a tray and 7.5" (L) x 11" (W) x 21.25" (H) humidity dome (Alberta LTD, Amazon Standard Identification Number: B089Q24D3N). The inside of the dome was sprayed daily with water to maintain humidity. After 3 days, pouches were replenished with 10 mL of Modified Nodulation Medium. The growth chamber was kept at 24°C with long-day light conditions (16 h light/8 h dark; 150 µmol m⁻² s⁻¹ light intensity).

Rhizobia strain and growth conditions

In this study, the *Sinorhizobium meliloti* 1021 strain of rhizobia was used for inoculation assays. To prepare the rhizobia for inoculation, a 5 mL liquid culture was grown overnight from a fresh plate of *S. meliloti* 1021. After 100–200 µL of the liquid culture was plated on Tryptone Yeast plates, they were grown for 48 h at 28°C. The entire film of rhizobia was collected in a 50 mL falcon tube and resuspended with autoclaved water. The rhizobia suspension was then diluted to a final OD600 = 0.1 using Fahræus medium without nitrogen.⁸⁶

METHOD DETAILS

Inoculation assay and sample collection

Plants were grown for one week in pouches before inoculation. Plants were inoculated by pipetting 1 mL of the *S. meliloti* 1021 suspension into each pouch. At the time of inoculation, the susceptibility zone of each plant was marked outside the pouch in the proximity of the root tips.

Ten growth pouches were used at each time point for a total of 40 plants each. At each time point, 4 cm sections around the marked susceptibility zone were harvested from each plant. The root segments from the 0 h time point (i.e., before inoculation) served as a control compared with expression data from the later time points. Root segments were harvested at 24, 48, and 96 hpi for each respective time point. The root segments collected at each time point were pooled for nuclei isolation.

Single nuclei RNA-seq and library preparation

The protocol used for nuclei isolation was previously described by Conde et al.⁸⁷ with minor adjustments. Nuclei isolation was carried out at 4°C, and all materials, tools, and solutions were pre-cooled to that temperature. Root segments were moved to a glass plate with 200 µL of modified Nuclei Isolation Buffer (NIB; 0 mM MES-KOH pH 5.7, 10 mM NaCl, 10 mM KCl, 3 mM MgCl₂, 2.5 mM EDTA, 250 mM sucrose, 0.1 mM spermine, 0.5 mM spermidine, 1 mM DTT, 0.2 U/µl Protector RNase Inhibitor⁸⁷) containing 0.5 U/mL of Protector RNase Inhibitor (Sigma Aldrich). After samples were fragmented with a sterile razor blade in two cycles of 2 min each time, the homogenate was washed into a 50 mL tube with modified NIB and incubated on a shaker for 5 min. All samples underwent at least two filtration stages, first with Miracloth (Calbiochem) and then with a 40 µm cell strainer (Greiner Bio-One), both pre-wet with modified NIB. The samples were centrifuged for 5 min at 600 g before the supernatant was discarded without disturbing the pellet. The pellet was resuspended with 4 mL of NIB WASH (10 mM MES-KOH pH 5.7, 10 mM NaCl, 10 mM KCl, 3 mM MgCl₂, 250 mM sucrose, 0.1% BSA and 0.2 U/µl Protector RNase Inhibitor)⁸⁷. Centrifugation and resuspension were repeated once more. After the final centrifugation, 300 µL of NIB WASH was used to resuspend the pellet. Samples that contained debris underwent an additional filtration step with a 40 µm cell strainer with an attached 5 mL tube (Falcon Corning). Next, the suspension was transferred to a sterile tube compatible with the BD FACS Aria IIU/III upgraded cell sorter. Nuclei were stained with 5 µg/mL DAPI for 5 min at 28°C before sorting. The sample was used for Florescence Activated Nuclei Sorting (FANS) at the Interdisciplinary Center for Biotechnology Research (ICBR) at the University of Florida to reduce the debris in the sample and minimize the likelihood of clogging the 10× Genomics microfluid chip. Nuclei were sorted in a BD FACS Aria IIU/III upgraded cell sorter with the following laser configuration (50mw 488 blue; 100mw 350 UV; 50mw 635 red; 100mw 405 violet and 100mw 531 yellow/green with 15 detectors and FSC/SSC) used to sort 20,000 nuclei at a final concentration of approximately 400 nuclei/µL. The sheath tank was filled with autoclaved BD FACS Flow Sheath Buffer (BD Biosciences, catalog number: 342003). RNase inhibitor RibolockTM (Thermo Fisher Scientific, catalog number: EO0382) was added to each liter of sheath buffer to total of 10 L of sterile sheath padded to the ARIA per run.

To generate snRNA-seq libraries for each genotype and time point, 10× Genomics microfluid chips were loaded with 16,500 nuclei. The Single Cell v3.1 Dual Index Gene Expression protocol was followed with the exception that NIB WASH was used to reach the final desired volume of nuclei suspension and a total of 15 PCR cycles were used to amplify the cDNA (Conde et al.⁸⁷). The amplified cDNA was used to construct libraries according to the manufacturer instructions, and the cDNA was sequenced at the ICBR using the NovaSeq 6000 System, 2 full S1 flow cells, and 2x100 bp sequencing kits but with the cycling of 28 bp for read 1, 10 bp for index 1, 10 bp for index 2, and 90 bp for read 2.

Experimental validation of gene markers

To confirm the cell type annotation of clusters of interest, Golen Gate MoClo⁸⁸ cloning was used to create promoter-beta-glucuronidase (GUS) expression cassettes for each marker gene. Briefly, the 2kb sequence upstream of the starting ATG (promoter) of each marker gene was identified using the *Medicago truncatula* genome.⁸⁹ The sequences were domesticated to remove any internal Bsal or BpI sites as needed. Bsal sites were added to either end to facilitate cloning and the sequences were synthesized by Synbio Technologies (Monmouth Junction, NJ). The promoters were cloned in frame with the beta-glucuronidase (GUS) CDS (pICH75111; MoClo Plant Parts Kit) and the 35S terminator into the cloning vector pICH47811 (MoClo Toolkit) as a level 1 reaction. To identify the positive events of the transformation, we used a fluorescent selection marker consisting of the CaMV 35S constitutive promoter driving the expression of the tdTomato protein cloned into the pICH47802 vector (MoClo Toolkit).⁸⁸ The constructs containing the marker gene's promoter and fluorescent selection marker were cloned together into the final expression vector pAGM4673 (MoClo Toolkit) using the level 2 reaction.

Composite *M. truncatula* A17 plants were generated as previously described.⁸⁶ One-day-old seedlings were inoculated with *Agrobacterium rhizogenes* (MSU440) harboring the tdTomato selection marker described above and the cluster-specific marker gene's promoter driving GUS expression. After 3 weeks of growing in plates containing Fahraeus medium supplemented with 15 mM of NH4NO₃, transgenic roots were selected by the presence of tdTomato fluorescence under an Olympus MVX10 fluorescence stereo microscope and transferred to the GUS staining solution (5 nM potassium ferrocyanide, 5 nM potassium ferricyanide, 0.1 M sodium phosphate buffer, 1 mM sodium EDTA, 1% Triton X- and 0.3% X-Gluc (previously dissolved in N,N-Dimethylformamide). To explore the expression of the marker gene in the nodule, 4-week-old composite *M. truncatula* plants were transferred to BNM. After 5 days, transgenic roots were treated with *S. meliloti* as described above. Roots or nodules were harvested (the timing after infection

depends on the cluster being evaluated and is described in the results), and stained with the GUS staining solution. Roots were incubated for 4 h and nodules overnight, at 37°C, in the staining solution. Then, transgenic roots or nodules were fixed in 4% formaldehyde as previously described.⁹⁰ Stained and fixed roots and nodules were then embedded using the Kulzer Technovit 7100 resin (Emgrid Australia), following the manufacturer's instructions. 8–10 µm thick sections were obtained using a Leica RM2045 microtome and stained with 0.1% Ruthenium Red (Sigma-Aldrich) in PBS for 10 min. Images were obtained at Zeiss Axioplan 2 microscope attached to a QImaging Retiga EXi Fast 1394 camera.

Functional characterization of MtSTY4

To characterize the function of the gene *MtSTY4*, we generated composite plants where the expression of the gene was knockdown via RNAi. A 153 bp fragment of the *MtSTY4* coding sequence was amplified by PCR (FWD 5'-CACCATGAACCAGCAAAGATG GACG-3', REV 5'-AGTCAAGAAATTAAACAAAG-3') from *M. truncatula* A17 root cDNA using Phusion Taq (NEB, Ipswich MA) according to the manufacturer's protocol and an annealing temperature of 55°C. Fragments were run through a 2% TBE agarose gel and cleaned with the Monarch DNA Gel Extraction Kit (NEB, Ipswitch, MA) and cloned into pENTR Directional TOPO vector following the manufacturer's instructions (Invitrogen, Carlsbad, California) to generate the Gateway entry clone vector. To generate the expression vector, the LR reaction was performed with the entry clone vector and the pK7GW1GW2(II)-RedRoot destination vector using the LR clonase kit (Invitrogen, Carlsbad, California).

Composite *M. truncatula* A17 plants were generated using constructs described above and methods previously reported.⁸⁶ Three weeks after transformation with *A. rhizogenes* MSU440, the roots were screened for the red fluorescence of dsRed, and the composite plants with red roots were transferred to growth pouches (Mega International, Newport, Minnesota) containing modified nodulation medium (MNM).⁸⁵ The plants were acclimated for one week, and then each pouch was inoculated with 1 mL of *S. meliloti* 1021 (OD 600 nm = 0.1), harboring the pXLGD4 and expressing lacZ under the hemA promoter.⁹¹ The MNM was replenished three days after transferring the plants and every week after that. Two weeks after inoculation, the total number of nodules per plant was counted as the number of nodules present in all the fluorescent roots in each plant under an Olympus MVX10 fluorescence stereo microscope. Then, live seedlings were stained for lacZ activity (5 mM potassium ferrocyanide, 5 mM potassium ferricyanide, and 0.08% X-gal in 0.1 M PIPES, pH 7) overnight at 37°C. Roots were rinsed in distilled water, and the bacteria inside the nodules of the transgenic roots were confirmed by observing the blue product of the reaction.

QUANTIFICATION AND STATISTICAL ANALYSIS

Quality control and cell clustering

The sequencing output was demultiplexed and processed using the software Cell Ranger (v7.0, 10× Genomics). Cell Ranger was also applied to generate counts for each gene in the *M. truncatula* genome (v.5.0, release 1.9⁸⁹). Cell Ranger was executed using default parameters. Next, using the raw counts matrix produced by Cell Ranger as input, the software EmptyDrops,⁹² DropletUtils R package, was used to identify and remove instances where a barcode represented an empty droplet. The filtered dataset was loaded on Monocle3 (commit 87f6e88⁹³), and only cells containing 400 or more detected unique molecular identifiers (UMIs) were selected and used for the analysis. Finally, the R package scDblFinder⁹⁴ was used to identify doublets, which were excluded from the dataset prior to further analysis.

Combining the samples and clustering was performed using Monocle3 and the Leiden community detection method.⁹⁵ Before clustering, the dataset was corrected for batch effect using Monocles' implementation of the matching mutual nearest neighbors method.⁹⁶ The default resolution of 0.0001 was applied during the clustering of the whole dataset, and a resolution of 0.001 was applied when reclustering specific cell types.

Identification of novel *Medicago* cell-type-specific markers

To identify the cell types that constitute each cluster of cells in our sample, we evaluated the expression of cell-type-specific marker genes previously described in the literature. In addition, we identified genes specifically expressed in each cluster and investigated their function as a proxy to define the cell type contained in the cluster. For each cluster, only genes expressed in more than 20% of the cells of the cluster and with an adjusted p value (q-value) ≤ 0.05, generated by Moran's I test of spatial autocorrelation, were selected. Moran's I test measures the dependence of a gene expression on the spatial location defined by the cluster and is effective in finding genes that vary in single-cell RNA-seq datasets.⁹⁷ Therefore, it can be interpreted as a differential expression test, identifying genes that vary between groups of cells in the UMAP space. When more than 100 marker genes were detected for a cluster, we selected the 100 genes with higher values for the specificity parameter as markers for that cluster for further exploration.

We evaluated the expression of those marker genes in a dataset originating from a laser capture microdissection (LCM) experiment investigating the initial hours (0 hpi, 24 hpi, 48 hpi, and 72 hpi) of *M. truncatula* response after infection with *S. medicae* ABS7³⁰ (BioProject PRJNA704996). Specific expression of a marker gene in a cell type captured by LCM indicates that the cluster represents cells of that type. A similar strategy was deployed to investigate the expression of the top 100 marker genes of each cluster in other relevant datasets publicly available as part of the *M. truncatula* RNA-seq Gene Expression Atlas Project v.2.⁹⁸

To identify differentially expressed genes in a comparison across time points, we used the function FindMarkers from the R package Seurat⁹⁹ to apply the Wilcox test. In each comparison, a cluster was only considered if it had more than 20 cells, and a gene was

only considered as differentially expressed with the FDR was smaller and 0.05 and the absolute value of the log₂ of the Fold Change was larger than one.

Trajectory inference analysis

During the response to rhizobia infection, multiple cell types undergo transcriptional regulation promoting developmental changes, for example, in the transition from the different layers of cortical cells to the different zones of the nodule. Trajectory inference analysis can reconstruct the different developmental stages by positioning the cells within an inferred developmental trajectory. Once the trajectory is inferred, it is possible to search for genes whose expression changes along the trajectory, pointing to the potential regulators and other critical genes governing the transition. In this study, we used the Bioconductor packages Slingshot⁶⁷ to model the trajectories and TradeSeq⁶⁸ to identify genes whose expression changes along the trajectory. A gene was considered differentially expressed within the trajectory if the Wald test, performed by the function associationTest() in TradeSeq, returned a p value corrected by the False Discovery Rate method smaller than 0.001.

Weighted Correlation Network Analysis

Weighted Correlation Network Analysis (WGCNA) is used to characterize patterns of co-expression in large gene expression datasets. We use the package hdWGCNA,⁷¹ an adaptation of the original WGCNA⁷⁰ method to single-cell datasets, to construct a co-expression network of genes expressed in the whole dataset.

Comparison of nodule number after *MtSTY4* knockdown via RNAi

Comparison between the nodule number in plants where the gene *MtSTY4* was knockdown via RNAi and control plants was performed by the student's t-test. The difference was considered significant when the p value <0.05.