PNAS

Supplementary Information for

Intracellular infection by symbiotic bacteria requires the mitotic kinase AURORA1

Jin-Peng Gao (高锦鹏)[a,b], Suyu Jiang[a], Yangyang Su[c], Ping Xu[c], Junjie Wang[a,b], Wenjie Liang[a,b], Cheng-Wu Liu[d], and Jeremy D. Murray[a,e,*]

* corresponding author: Jeremy D. Murray
Email: jeremy.murray@jic.ac.uk

This PDF file includes:

Supplementary Information Text

SI Materials and Methods

Promoter GUS Analysis. For promoter-GUS analysis in *M. truncatula* hairy roots, promoters (MYB3R1$_{pro}$, 1990 bp; AUR1$_{pro}$, 1423 bp; CYCA3;1$_{pro}$, 1722 bp; CYCD1;1$_{pro}$, 1949 bp; CDKB2;2$_{pro}$, 1855 bp; CDKC2;1$_{pro}$, 2251 bp; TPXL2$_{pro}$, 1939 bp; TPXL3$_{pro}$, 1922 bp; MAP65-1$_{pro}$, 1969 bp; and MAP65-9$_{pro}$, 1695 bp) were cloned into pBGWFS7 vector by LR reaction (11791019, Invitrogen). These vectors were then transformed into *A. rhizogenes* strain Arqua1 (AC1060L, Weidi) for hairy root transformation. Transgenic roots of more than 20 composite A17 plants harboring the corresponding promoter-GUS reporter construct were stained with a solution comprised of 0.5 mM potassium ferricyanide, 0.5 mM potassium ferrocyanide, 0.5 mg/mL 5-bromo-4-chloro-3-indolyl-β-D-glucuronic acid (X-gluc, 10904ES03, Yeasen), 0.1% (v/v) Triton X-100, and 0.1 M sodium phosphate buffer (pH=7.0) that was vacuum infiltrated 5 min, and subsequently incubated at 37ºC in the dark for 2-6 h. GUS-staining patterns were observed and photographed under a microscope (Olympus SZX7 or BX53). For some *AUR1$_{pro}$:GUS* transgenic roots, after GUS staining the roots were fixed in 2.5% glutaraldehyde for 1 h, then stained in 100 mM sodium phosphate buffer (pH=7.4) containing 0.8 mg/mL magenta-gal (1758-0400-100MG, Inalco), 10 mM KCl and 1 mM $MgSO_4$ at 28°C in dark overnight to stain the lacZ-tagged rhizobia.

Yeast One Hybrid. The CDS of *MYB3R1* and *MYB2R1* were each cloned into pGADT7 by LR reactions (11791019, Invitrogen). The different *AUR1* promoter fragments (-400/-200 bp or -200/-1 bp) were amplified by PCR and a mutated form of -200/-1 bp was synthesized (Genscript) and were inserted into pHIS2 by recombination reactions (NR005-01A, Novoprotein). Different pairs of constructs were co-transformed into yeast strain Y187 (YC1020S, Weidi), and the cells were grown on SD-Trp/-Leu medium. Background (leaky) *HIS3* expression was controlled by adding 90 mM 3-amino-1,2,4-triazole (3AT) to the SD-Trp/-Leu/-His medium. Yeast growth was monitored for up to 3 days, and photos were taken using a camera (100D18-55, Canon).

Yeast Two Hybrid. The full CDS of *AUR1*, *WDL*, *TPXL2*, *TPXL3*, *MAP65-1*, *MAP65-3*, *MAP65-5*, *MAP65-8*, *MAP65-9* and partial CDS of *TPXL2* (*TPXL2-N*, amino acids 1-85; *TPXL2△N*, aa 86-372), *TPXL3* (*TPXL3-N*, aa 1-93; *TPXL3△N*, aa 94-475) were cloned into pGADT7 and pGBKT7 by LR reaction (11791019, Invitrogen). Different pairs of constructs were co-transformed into yeast strain AH109 (YC1010M, Weidi), and the yeast cells were grown on SD-Trp/-Leu medium. Then three single clones from SD-Trp/-Leu plates were transferred to SD-Trp/-Leu/-His and SD-Trp/-Leu/-His/-Ade plates. Yeast growth was monitored for up to 3 days, and photos were taken using a camera (100D18-55, Canon).

Co-Immunoprecipitation. Using recombination reactions, the CDS of *AUR1*, *TPXL2*, and *TPXL3* were cloned into the pUB-GFP-HA vector, and the CDS of *AUR1*, *MAP65-1* and *MAP65-9* were cloned into pUB-GFP-FLAG vector (NR005-01A, Novoprotein). The resulting plasmids were transformed into *A. tumefaciens* GV3101 (AC1001M, Weidi) and then co-infiltrated into *N. benthamiana* leaves with *A. tumefaciens* GV3101 carrying *35S$_{pro}$:P19* for transient expression. After 3 days growth, the leaves were ground in liquid nitrogen, and total proteins were extracted using buffer containing 50 mM HEPES (pH=7.5), 150 mM NaCl, 1 mM EDTA, 50 mM β-glycerophosphate, 5% glycerol, 0.5% Triton X-100, and EDTA-free protease inhibitor cocktail (K1011-10, APExBIO). The protein extract was then incubated with pre-washed anti-FLAG M2 magnetic beads (A2220-5ML, Sigma-Aldrich) for 2 h at 4°C on a roller shaker. The beads were washed 5 times with a buffer containing 50 mM HEPES (pH=7.5), 150 mM NaCl, 1 mM EDTA, 50 mM β-glycerophosphate, 5% glycerol, 0.5% Triton X-100, and used directly for western blot

analyses using anti-HA (H6908-100UL, Sigma-Aldrich) or anti-FLAG (SAB4301135-100UL, Sigma-Aldrich) antibody.

Split Luciferase Biomolecular Complementation. *AUR1* and *MYB3R1* was individually cloned into a modified *$35S_{pro}$:nLUC* vector and *TPXL2*, *TPXL2-N*, *TPXL2△N*, *TPXL3*, *TPXL3-N*, *TPXL3△N*, *MAP65-1*, *MYB3R2* was respectively cloned into a modified *$35S_{pro}$:cLUC* vector. nLUC and cLUC empty vectors were used as controls. The resulting plasmids were transformed into *A. tumefaciens* GV3101 (AC1001M, Weidi) which were then co-infiltrated into *N. benthamiana* leaves with GV3101 carrying the *$35S_{pro}$:P19* vector. After three days of growth, the *N. benthamiana* leaves were sprayed with 1 mM luciferin (E1500, Promega) and the luciferase activity was monitored by a Tanon 5200 detector.

RNA Extraction and Real-Time PCR. Total RNA was extracted from ~100 mg plant tissues using an Ultrapure RNA kit (DP419, Tiangen). RNA concentrations were determined using a NanoDrop and first strand cDNA was synthesized with oligo (dT)18 primer from 1-2 μg RNA using the One-Step gDNA Removal and cDNA Synthesis Super Mix (AT311-03, TransGen). Then the cDNA was diluted 10-20 times for subsequent experiments. The qRT-PCR reactions were performed using 2×RealStar Green Fast Mixture kit (A303-05, GenStar) and assayed with the real-time PCR detection system (CFX96, BioRad). The PCR conditions were as follows: 40 cycles of 95°C for 15 s, 60°C for 15 s, and 72°C for 15 s. *MtEF1* (*Elongation Factor 1*) was used as a reference gene (1). The sequences of the primers used are listed in *SI Appendix*, Table S1.

Protein Expression. *AUR1*, *TPXL2-N* (aa 1-85), *TPXL3-N* (aa 1-93), *MAP65-1*, and *MYB3R1-N* (aa 1-200) were cloned into vector pGEX-4T-1 upstream and in-frame with GST by recombination reactions (NR005-01A, Novoprotein). Plasmids were transformed into the *E. coli* strain Rosetta (DE3; EC1010, Weidi) for expression as GST fusion proteins. Expression of the fusion proteins was induced with 0.5 mM isopropylthio-β-galactoside (IPTG, 10902ES10, Yeasen) for 16-24 h at 16°C. The bacteria cultures were placed on ice slurry and sheared using a sonicator (JY92-IIN, Ningbo Scientz). The supernatant was then transferred to Poly-Prep Chromatography Columns (7311550, Bio-Rad) filled with pre-washed glutathione sepharose 4B (17075601, GE Healthcare). The columns were then incubated for 2-4 h at 4°C on a roller shaker. After three washes of phosphate buffer saline (pH=7.4), GST elution buffer (20 mM reduced glutathione S-transferase, 50 mM Tris-HCl and 0.5% Triton X-100, pH=8.0) was applied to the columns and collected. The purified recombinant proteins were analyzed using a 4-12% Precast SDS–PAGE gel (180-8011H, Tanon Biofuraw) and identified by Coomassie brilliant blue staining.

***In Vitro* Phosphorylation.** GST-AUR1, GST-TPXL2-N (aa 1-85), GST-TPXL3-N (aa 1-93), GST-MAP65-1 and GST purified proteins were incubated in kinase reaction buffer containing 50 mM HEPES (pH=7.5), 10 mM $MgCl_2$, 1 mM DTT, 1 mM ATP, and ^{32}P ATP at 25°C for 30 min. The samples were subsequently analyzed using a 10% Precast-GLgel Hepes SDS-PAGE gel (C621101, Sangon) followed by autoradiography. A similar gel was subjected to Coomassie brilliant blue staining. The gel was monitored using a Typhoon FLA 9000 or Tanon 5200 detector.

Electrophoretic Mobility Shift Assay. Electrophoretic mobility shift assays (EMSA) were performed using the N-terminus of MYB3R1 (aa 1-200), which includes the DNA binding domain, purified from *E. coli* (GST-MYB3R1) and probes labeled with 5' CY5 universal primer. For the assay, 0.1, 0.3, or 0.9 μg of purified GST-MYB3R1 protein and 0.04 pmol labeled DNA probe (-197/-90 bp of *AUR1* promoter) were mixed and incubated at 28°C for 30 min. Competition experiments were performed by adding 0.8, 2.0, or 4.0 pmol of unlabeled DNA probes. The protein/DNA samples were loaded on a 6% native gel and the gel was run at 100V for 90 min in a 4°C cold room. The gel was scanned with a Typhoon FLA 9000 (FUJIFILM FLA 9000 plus DAGE).

Transactivation Assay in Yeast. The CDS of full length MYB3R1, N-terminal of MYB3R1 (MYB3R1-N, aa 1-200), C-terminal of MYB3R1 (MYB3R1-C, aa 201-903), and full length NSP2 were cloned by LR reaction (11791019, Invitrogen) into pGBKT7 to fuse them to the GAL4 binding domain. These constructs were transformed into yeast strain AH109 and screened on selective medium (SD/-W or SD/-WHA). The empty vector was used as a negative control, and transcription factor NSP2 served as a positive control. Yeast growth was monitored for up to 3 days.

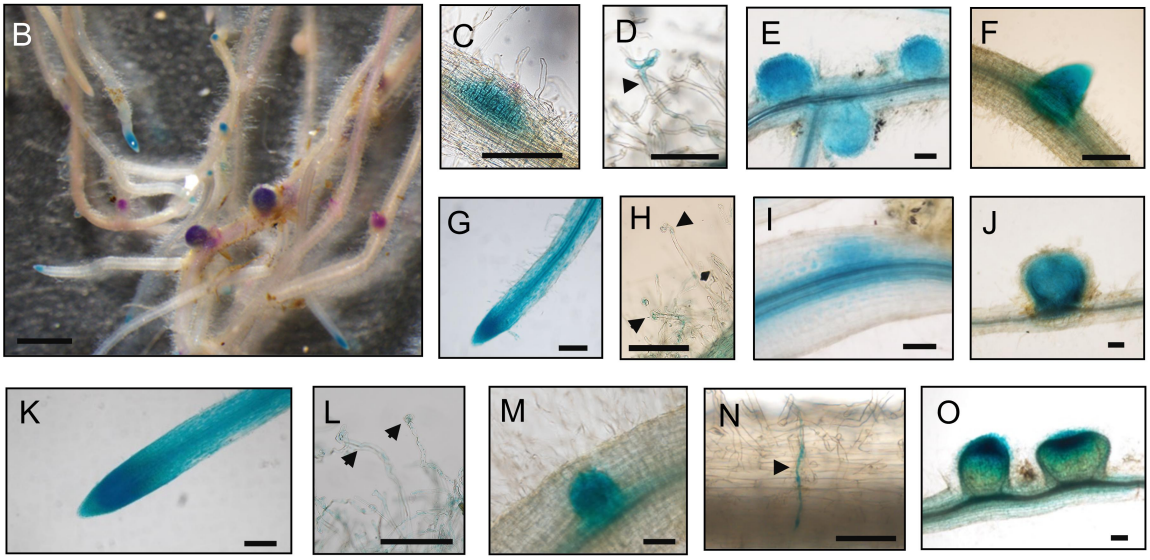
Transactivation Assay in *N. benthamiana* and GUS Quantification. Various constructs containing pUB-GFP-FLAG-MYB3R1 and pBGWFS7-AUR1$_{pro}$ were transformed into *A. tumefaciens* GV3101 and then co-infiltrated into *N. benthamiana* leaves with *35S$_{pro}$:P19*. After 36 hours, the leaves were vacuum infiltrated for 5 min in GUS histochemical staining buffer and incubated at 37°C for 2-6 h. To stop the reaction and clear the tissue, the leaves were washed with 75% ethanol five times. Then the leaves were observed and photographed using a digital camera (100D18-55, Canon). For GUS quantification, total protein from six leaves were extracted in liquid nitrogen and mixed with a protein extraction buffer (50 mM disodium hydrogen phosphate, 10 mM Na_2EDTA, β-mercaptoethanol, 0.1% Sarcosyl and 0.1% TritonX-100). The slurry was centrifuged and the supernatant was mixed with 4-methylumbelliferyl β–D-glucuronide (4-MUG; M9130-100MG, Sigma-Aldrich), and incubated at 37°C in the dark for 60 min. Then GUS activity was measured using a microplate reader at 355 nm excitation and 460 nm emission filter (Varioskan Flash, Thermo Scientific).

Chromatin Immunoprecipitation Analysis. The chromatin immunoprecipitation (ChIP) assay was performed using a Plant ChIP Kit (EPT-P-2014-24, EpiQuik). Briefly, 1 g of transgenic hairy roots (~100 plants) expressing *LjUBQ$_{pro}$:MYB3R1-FLAG* was harvested and cross-linked with 20 ml 1% formaldehyde for 5-10 min under a vacuum. After quenching the reaction with 2.5 ml 1 M glycine solution, the roots were then ground to a fine power in liquid nitrogen and nuclei were isolated through two layers of Miracloth (475855-1R, Millipore). The cross-linked chromatin was sheared using a sonicator (UCD-200, Bioruptor) for 15 s at 40% duty cycle, until the majority of the DNA fragments produced had a size between 200 and 600 bp. Immunoprecipitation was performed using anti-FLAG antibody (SAB4301135-100UL, Sigma-Aldrich). qPCR analysis then was performed, and the sequences of primers are listed in Table S1.

Statistical Analysis. For evaluation of phenotypes, at least 10 individual plants were analysed per experiment; sample sizes and statistical parameters are stated in the Figure legends. Means of infection threads, nodule numbers or relative gene expression were compared using two-tailed Student's *t* test (statistical significance, * indicates $P < 0.05$, ** indicates $P < 0.01$, *** indicates $P < 0.001$) using Microsoft Excel 2016 or GraphPad Prism 7 software.

A

gene symbol	gene model	1dpi	5dpi	skl	NF	ZIII	IZ	FIIP	FIID	F1	root	root meristem	leaf	vegetative bud
CYCA3;1	Medtr3g102530	1.6	6.0	5.3	3.4	658.0	485.0	791.0	1351.0	1165.0	3.2	10.1	2.6	4.6
CYCD1;1	Medtr8g063120	1.2	1.4	3.2	3.2	33.0	23.0	101.0	480.0	1720.0	3.1	4.2	2.6	6.1
CDKB2;2	Medtr1g075610	1.3	2.3	4.1	0.6	0.0	3.0	76.0	619.0	2366.0	3.2	6.1	1.2	19.6
CDKC2;1	Medtr1g098300	3.0	5.6	12.2	6.4	377.0	274.0	392.0	751.0	415.0	2.6	13.1	1.6	2.8
AUR1	Medtr3g110405	1.0	3.2	3.9	1.1	5.0	29.0	87.0	400.0	732.0	5.6	10.6	2.9	7.9
MYB3R1	Medtr3g110028	1.0	2.1	2.6	2.1	1.0	0.0	20.0	118.0	354.0	3.3	6.6	1.8	11.8



P

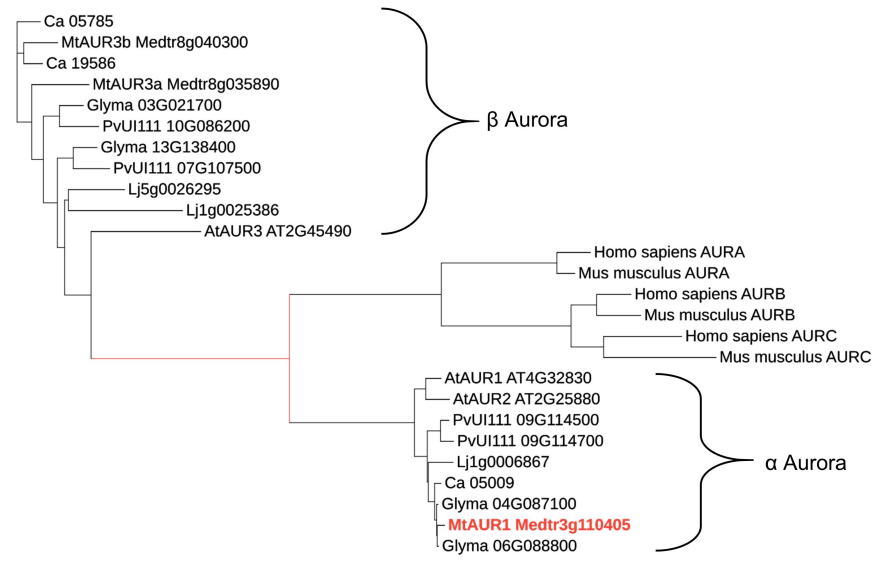


Fig. S1. Symbiotic expression of cell cycle related genes. (***A***) Data for symbiotic expression of cell cycle related genes was extracted from previously published data. The first chart is based on data from early report (2). 1 and 5 dpi indicate the expression change in root hairs of seedlings

at 1 and 5 days post inoculation with rhizobia Rm1021 in the wild type A17; *skl*, in the hyperinfected *sickle* mutant (impaired in the sensing of ethylene) at 5 dpi; NF, in the root hairs of wild type A17 24 hours after Nod factor treatment. The second chart represents the normalized RNAseq reads of cell cycle related genes in five nodule zones based on laser capture microdissection data from early report (3). FI indicates meristematic zone; FIID indicates infection zone-distal fraction; FIIP indicates infection zone-proximal fraction; IZ indicates interzone; and ZIII indicates nitrogen-fixing zone. The third and fourth charts represent the expression of the same set of genes in root, root meristem and in leaf, vegetative buds based on the data from the *Medicago truncatula* Gene Expression Atlas (https://medicago.toulouse.inrae.fr/MtExpress). (***B***) An overall image of *A. rhizogenes*-induced hairy roots expressing $AUR1_{pro}$*:GUS* after staining. The blue color indicates GUS staining, and the magenta color indicates LacZ staining of *S. meliloti* Rm2011. Scale bar, 1 mm. (***C***) A image showing $AUR1_{pro}$*:GUS* activity in a nodule primordium at 5 dpi with rhizobia. (***D-F***) Images showing $CYCA3;1_{pro}$*:GUS* activity in an infected root hair (D), in young nodules (E), and in lateral root primordium (F) at 5 dpi. (***G-J***) Images showing $CYCD1;1_{pro}$*:GUS* activity in a root tip (G), in infected root hairs (H), in a nodule primordium (I), and in a young nodule (J) at 5 dpi. (***K-M***) Images showing $CDKB2;2_{pro}$*:GUS* activity in a root tip (K), in infected root hairs (L), and in a nodule primordium (M) at 5 dpi. (***N*** and ***O***) Images showing $CDKC2;1_{pro}$*:GUS* activity in an infected root hair (N) and in elongated nodules (O) at 5 dpi. Scale bars, 200 μm (C-O). (***P***) Phylogenetic tree of Aurora kinases from different plant and animal species. At, *Arabidopsis thaliana*; Ca, *Cicer arietinum*; Glyma, *Glycine max*; Lj, *Lotus japonicus*; Pv, *Phaseolus vulgaris*; Mt, *Medicago truncatula*. The mammal AURs comprise A, B, C three groups; and the plant AURs comprise two groups, α-AURs and β-AURs. The red color indicates the single α-Aurora of *M. truncatula* (MtAUR1).

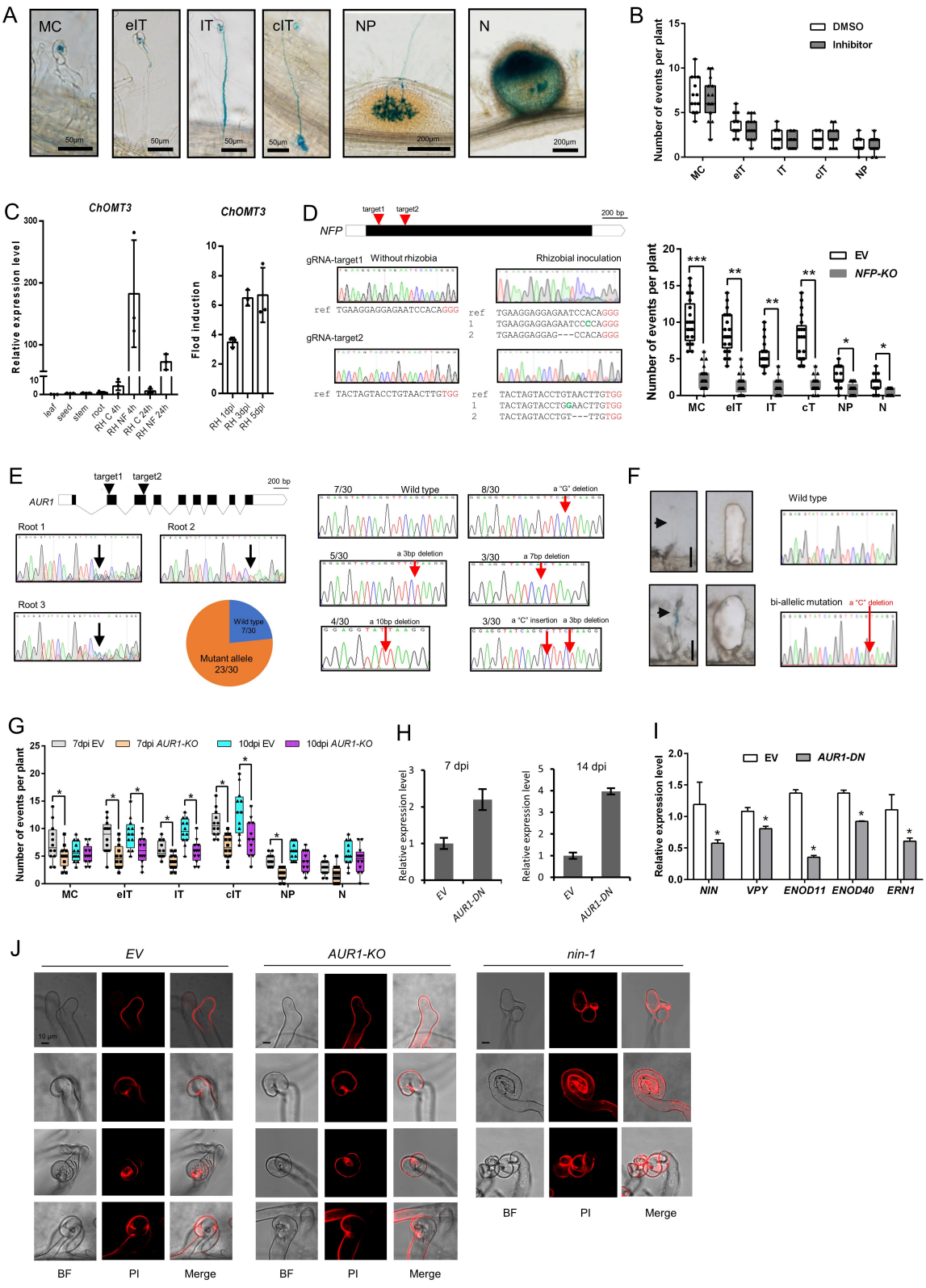


Fig. S2. Cell cycle genes are involved in nodulation. (***A***) The four categories of infection events which were used to characterize the infection phenotype in this study are shown. MC,

rhizobia microcolony; eIT, elongating infection thread in root hair; IT, infection thread that fully traversed the root hair; cIT, infection thread that traversed the root hair and ramified into the cortex; NP, nodule primordium; N, nodule. (***B***) Infection events of *M. truncatula* wild type seedlings with 0.6 nM Aurora kinase inhibitor treatment. Tozasertib, a small molecule inhibitor of mammalian α- and β-Aurora kinase activity (4, 5), was used to treat rhizobia inoculated *M. truncatula* A17 seedlings, and DMSO was used as a control. All samples (with and without the inhibitor) were germinated and grown on plates in the same incubator, and the infection events were counted at 7 dpi with rhizobia. (***C***) The expression level of *ChOMT3* in leaf, seed, stem, root and root hairs based on the data form the *Medicago truncatula* Gene Expression Atlas (left). Fold induction of *ChOMT3* in root hairs during infection (right), the data from early study (2). RH C, root hair control; RH NF, root hair with Nod factor treatment. RH 1/3/5 dpi indicates the expression change in root hairs at 1/3/5 days post inoculation with *S. meliloti* 1021 vs a matched control strain unable to produce Nod factors. (***D***) Tissue-specific CRISPR/Cas9 approach (*ChOMT3*$_{pro}$*:Cas9*) in legume using *NFP-KO* as an example. *Nod Factor Perception* (*NFP*) is required for infection thread and nodule formation (6, 7). The guide RNA target positions in *NFP* are indicated by red arrows (upper panel) and Sanger sequencing results of the PCR amplicons from *NFP-KO* roots are shown in the lower panel. Sanger sequencing results showing that *NFP* was not edited in non-inoculated *NFP-KO* roots (left), and that it was edited in roots inoculated with *S. meliloti* (right). The wild type sequence (ref) at the targeting sites and two primary edited forms are aligned, where the substituted bases are indicated in green, deleted bases by underscores, and the red letters indicate the PAM site. The right panel showed quantification of infection events and nodule primordia in *NFP-KO* and EV control plants at 5 dpi. The results showed that the *NFP-KO* transgenic roots had strongly restricted nodule formation and developed ~80% fewer infection threads than control roots. (***E***) CRISPR/Cas9 mediated gene disruption of *AUR1-KO* in hairy roots. The guide RNA target positions in *AUR1* exons are indicated by the black triangle. Sequences having multiple peaks around the guide RNA target1 are indicated by the black arrow, which indicates gene editing. Three independent hairy roots that were positively edited as shown by sequence chromatography (root 1, root 2 and root 3) were further analyzed for sequence editing effects. Their PCR products were pooled and cloned into a vector, and 30 clones were sequenced. Results from sequence analysis of these clones revealed 23 mutated and 7 wild type sequences as shown in the pie chart and chromatographs where the mutated sequences were labeled with red arrows. (***F***) *AUR1* PCR product sequence from DNA isolated from an individual root hair from *AUR1-KO* roots. Individual root hairs were collected using laser capture microdissection and their corresponding sequences at the target site are shown. The two panels show the magnified views of an un-infected (upper panel) and infected root hair (bottom panel) captured before (left) and after (right) microdissection. The root hairs were immediately used as templates for 2 rounds of PCR, with each round of PCR having 32 cycles. Representative sequence of individual root hair from *AUR1-KO* roots are shown. Black arrows point to root hairs and red arrow indicates the mutated sequence. Scale bars, 50 μm. (***G***) Quantification of infection events and nodule primordia in the EV and *AUR1-KO* at 7 dpi and 10 dpi. (***H***) Expression analysis of *AUR1* in *AUR1-DN* transgenic roots at 7 dpi and 14 dpi. Expression levels were normalized against the reference gene *MtEF-1*, and data are mean ±SD from three independent biological repeats. (***I***) Transcript levels of multiple nodulation genes (*NIN*, *VPY*, *ENOD11*, *ENOD40*, and *ERN1*) in EV or *AUR1-DN* transgenic roots at 14 dpi with *S. meliloti* Rm2011. Expression levels were normalized against *MtEF1*, and data are mean ±SD from three biological repeats. (***J***) Images showing root hair deformation and curling after propidium iodide (PI) staining at 3 dpi. The red color indicates cell wall by propidium iodide staining. Scale bars, 10 μm. Boxes show the first quartile, median and third quartile; whiskers show minimum and maximum values; dots show data points. Means were compared using Student's *t*-test, * indicates $P<0.05$, ** indicates $P<0.01$, *** indicates $P<0.001$. Experiments were carried out twice with similar results.

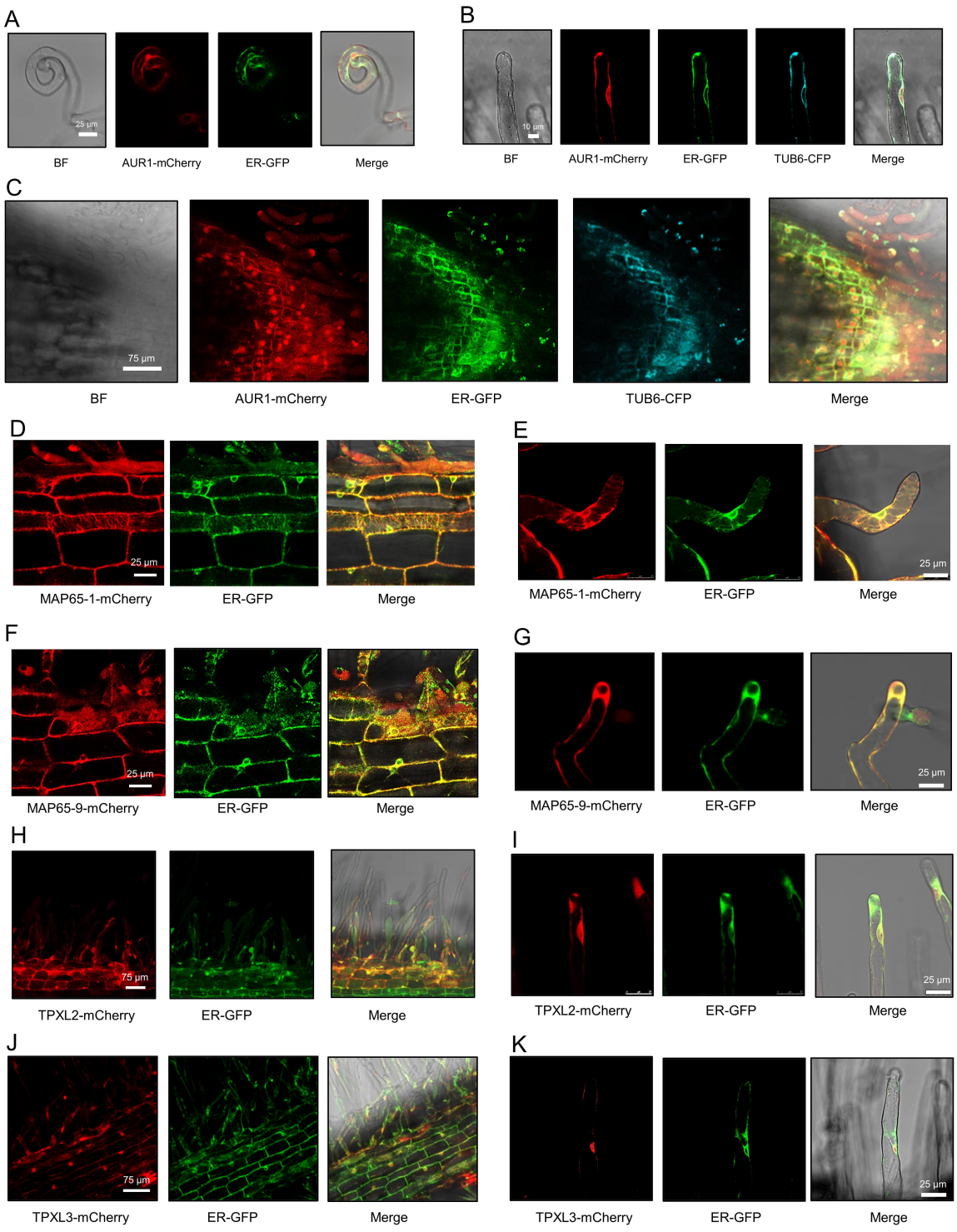


Fig. S3. Confocal images showing the subcellular localization of AUR1, TPXL and MAP65. (***A***) AUR1-mCherry driven by *LjUBQ* promoter in a root hair of *M. truncatula* A2 line (ER-GFP). (***B*** and ***C***) Co-localization of AUR1-mCherry driven by *LjUBQ* promoter and TUB6-CFP by *AtUBQ* promoter in a root hair (B) and other root cells (C) of A2 plants. (***D*** and ***E***) Confocal images of MAP65-1-mCherry driven by *LjUBQ* promoter in root (D) and root hair (E). MAP65-1 localizes to

the ER. (***F*** and ***G***) Confocal images of MAP65-9-mCherry driven by *LjUBQ* promoter in root (F) and root hair (G). MAP65-9 localizes to the ER. (***H*** and ***I***) Confocal images of TPXL2-mCherry driven by *LjUBQ* promoter in transgenic roots. TPXL2 is partially localizes to the ER. (***J*** and ***K***) Confocal images of TPXL3-mCherry driven by *LjUBQ* promoter in hairy roots. Images were captured in *M. truncatula* A2 plants which contain the ER marker (*35S$_{pro}$:GFP-HDEL*) at 5 dpi with *S. meliloti* Rm2011. ER, endoplasmic reticulum. Experiments were carried out three times with similar results.

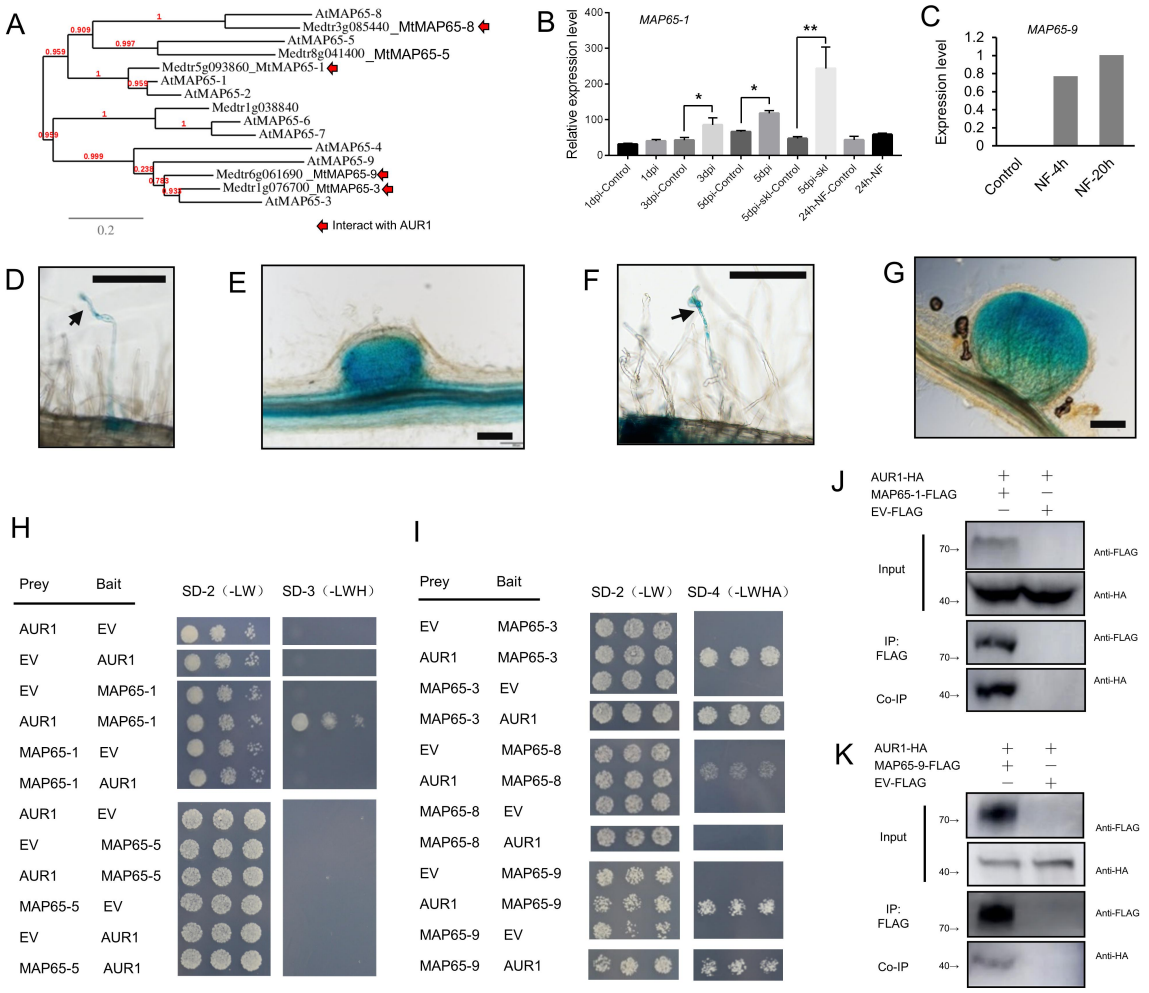


Fig. S4. Interaction between AUR1 and MAP65. (***A***) Phylogenetic tree of MAP65 homologs from *M. truncatula* and *A. thaliana*. The red arrows indicate the MAP65s that interact with AUR1 based on the Y2H assays in (H and I). (***B***) Relative expression of *MAP65-1* after rhizobial infection and Nod factor (NF) treatment from early report (2). (***C***) Relative expression of *MAP65-9* after Nod factor (NF) treatment based from previous report (8). (***D*** and ***E***) Images showing *MAP65-1*$_{pro}$*:GUS* activity in a rhizobia infected root hair (D) and in a nodule primordium (E) at 5 dpi. (***F*** and ***G***) Images showing *MAP65-9*$_{pro}$*:GUS* activity in an infected root hair (F) and in a young nodule (G) at 5 dpi. (***H*** and ***I***) Y2H assays between AUR1 and MAP65-1/-3/-5/-8/-9. The combinations of proteins expressed in either the AD vector (pGADT7) or the BD vector (pGBKT7) are indicated alongside the yeast colonies. Three yeast colonies were plated for each interaction combination on SD-2/-Leu-Trp or SD-3/-Leu-Trp-His or SD-4/-Leu-Trp-His-Ade medium. (***J*** and ***K***) Co-IP assays of AUR1-HA and MAP65-1-FLAG (J) or MAP65-9-FLAG (K) in *N. benthamiana* leaves. Proteins were immunoprecipitated (IP) with anti-FLAG-M2 beads, and analyzed by Western blot using HRP-conjugated anti-FLAG or anti-HA antibody. Experiments carried out three times with similar results. Scale bars, 200 μm (D-G).

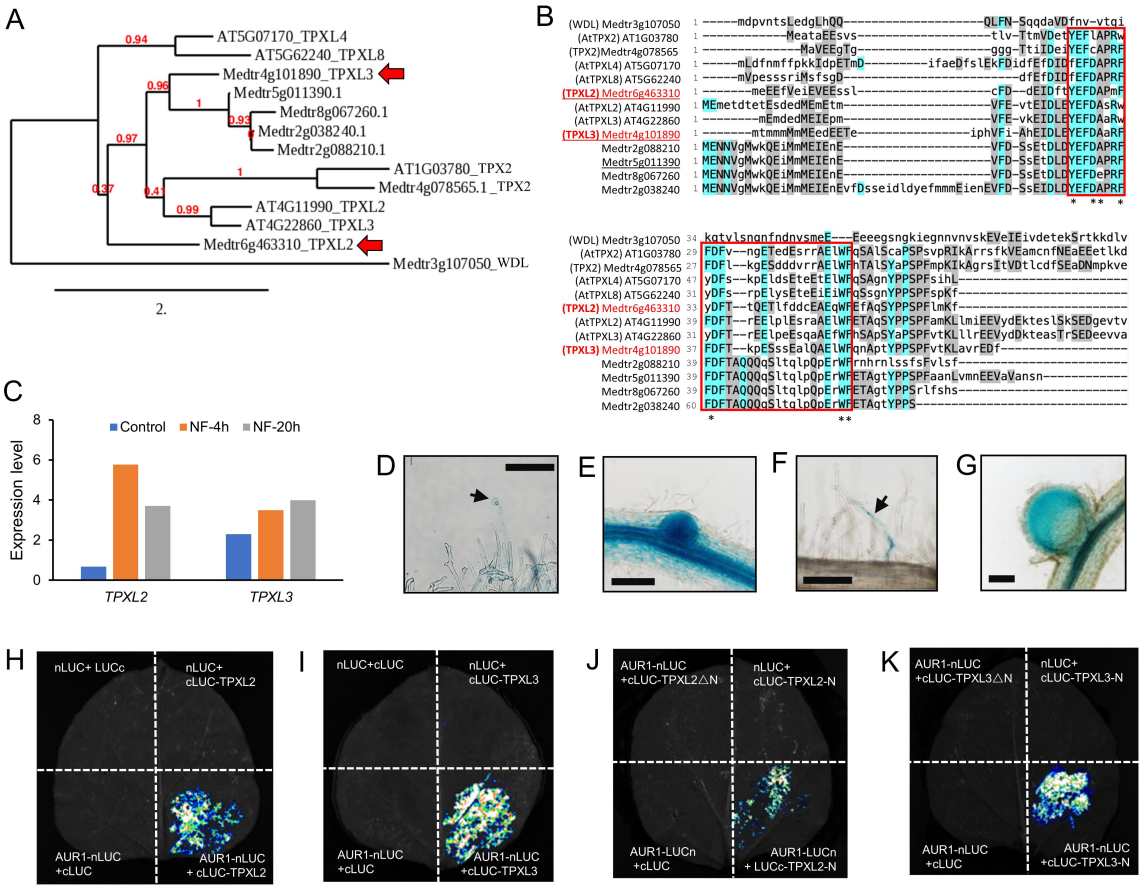


Fig. S5. Interaction between AUR1 and TPXL. (***A***) Phylogenetic tree of *A. thaliana* TPX2/TPXL and their homologs in *M. truncatula*. A Wave Dampened-Like (WDL) homolog is shown as a reference. *M. truncatula* TPXL2/3 are indicated with red arrows. Analysis was carried out using the Phylogeny.fr server (9) as follows. Protein sequences were aligned using MUSCLE (10) and the phylogenetic tree was constructed using PhyML (11). (***B***) Alignment of the N-terminal regions of sequences in (A). The predicted Aurora kinase binding site is boxed in red, with key conserved residues identified by early report (12) indicated by asterisks. The *M. truncatula* group A TPXL proteins that contain both Aurora binding and TPX importin domains are underlined. *M. truncatula* TPXL2/3 are indicated with red color. (***C***) Relative expression of *TPXL2* and *TPXL3* genes after Nod factor (NF) treatment based on the previous transcriptomic study (8). (***D*** and ***E***) Images showing *TPXL2*$_{pro}$*:GUS* activity in infected root hairs (D) and in a nodule primordium (E) at 5 dpi. (***F*** and ***G***) Images showing *TPXL3*$_{pro}$*:GUS* activity in an infected root hair (F), and in a young nodule (G) at 5 dpi. (***H*** and ***I***) Split luciferase complementation assays between AUR1 and TPXL2 (H) or TPXL3 (I). nLUC-tagged AUR1 was co-infiltrated into *N. benthamiana* leaves along with the cLUC-tagged TPXL2 or TPXL3. (***J*** and ***K***) Split luciferase complementation assays between AUR1 and TPXL2-N (J) or TPXL3-N (K). TPXL2-N indicates N-terminus of TPXL2 (1-85 aa), TPXL2ΔN indicates deletion of N-terminus (86-372 aa); TPXL3-N indicates N-terminus of TPXL3 (1-93 aa), TPXL3ΔN indicates deletion of N-terminus of TPXL3 (94-475 aa). Experiments were carried out three times with similar results. Scale bars, 200 μm (D-G).

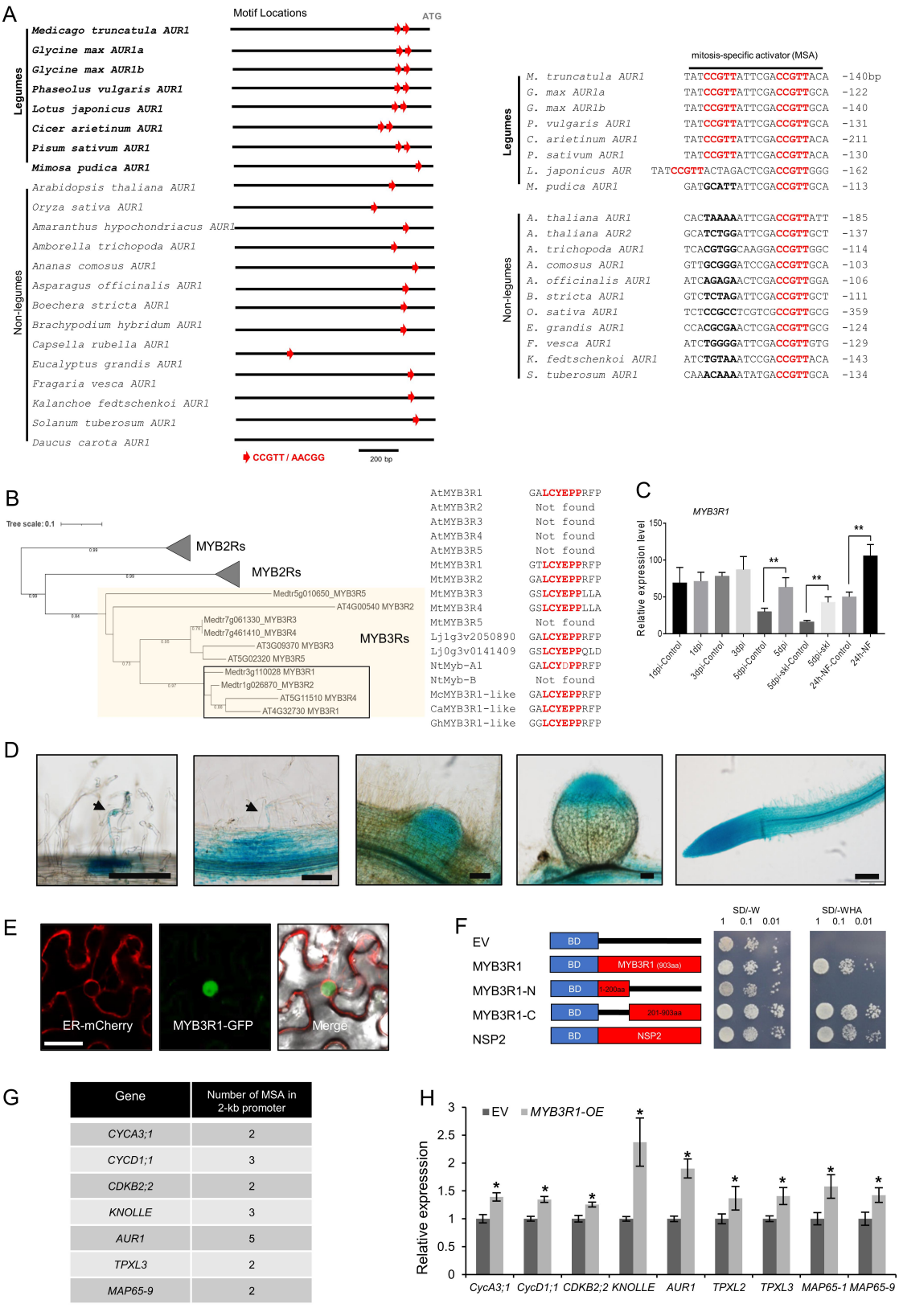


Fig. S6. MYB3R1 regulates expression of cell cycle genes. (***A***) Analysis of *AUR1* promoters from different species revealed that most have at least one core MSA element sequence

(AACGG) located close to the transcriptional initiation site (ATG) while legumes have two closely spaced MSAs (left). Some *AUR1* promoter sequences from legumes and non-legumes are aligned in the right panel. (***B***) A phylogenetic tree of *M. truncatula* and *A. thaliana* MYB3Rs and closest homologs. Sequences were aligned using T-COFFEE, and the phylogenetic tree was constructed using PhyML and presented using iTOL. Bootstrap support values >0.70 are shown. LCYEPP was found in many MYB3R1 proteins (right panel), At, *Arabidopsis thaliana*; Lj, *Lotus japonicus*; Mt, *Medicago truncatula*; Nt, *Nicotiana tabacum*; Mc, *Momordica charantia*; Ca, *Capsicum annuum*; Gh, *Gossypium hirsutum*. (***C***) Relative expression of *MYB3R1* genes in root hairs after rhizobial infection or Nod factor (NF) treatment from early report (2). (***D***) Images, from left to right, showing $MYB3R1_{pro}$*:GUS* activity in an infected root hair, in a nodule primordium, in a young nodule, in an elongated nodule and in a root tip at 7 dpi. (***E***) Subcellular localization of MYB3R1-GFP and ER-mCherry in *N. benthamiana* leaf epidermal cells. (***F***) Transactivation activity assay of MYB3R1 in yeast. The pGBKT7 vector expresses proteins fused to a GAL4 DNA-binding domain (BD); Empty vector (EV) served as a negative control, and transcriptional activator NSP2 served as a positive control. MYB3R1-N contains the N-terminal 200 aa fragment which contains the DNA binding domain. (***G***) The number of mitosis-specific activator (MSA) elements in the 2-kb promoters of some cell cycle related genes. (***H***) Relative gene expression of cell cycle related genes in EV and *MYB3R1-OE* hairy roots. Relative expression was normalized to *MtEF-1*. The presented data are from three biological repeats. Data are mean ± SD. Means were compared using Student's *t*-test, * indicates $P<0.05$, ** indicates $P<0.01$. Experiments were carried out twice (H) or three times (D-F) with similar results. Scale bars, 200 μm (D), 25 μm (E).

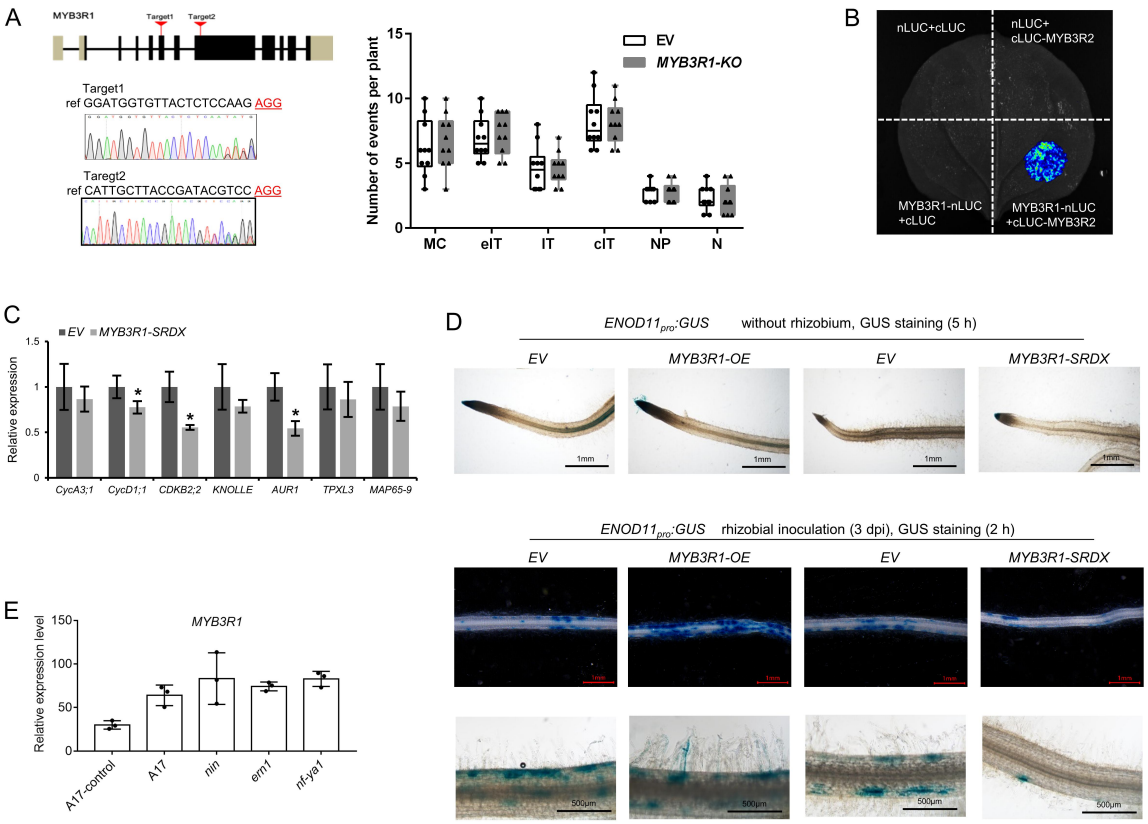


Fig. S7. MYB3R1 positively regulates nodulation. (***A***) No phenotype was observed after knockout of *MYB3R1* in *M. truncatula* hairy roots compared with the control. The guide RNA target positions in *MYB3R1* gene are indicated by red arrows (upper left panel) and Sanger sequencing results of the corresponding PCR amplicons from *MYB3R1-KO* roots show double peaks around PAM (bottom left panel). Red letters indicate the PAM sites, and 'ref' indicates reference sequences. The right panel shows the results from quantification of infections and nodules in the EV control and *MYB3R1-KO* transgenic roots of the composite plants at 7 dpi with *S. meliloti* Rm2011. MC, rhizobia microcolony; eIT, elongating infection thread in root hair; IT, infection thread that fully traversed the root hair; cIT, infection thread that traversed the root hair and ramified into the cortex; NP, nodule primordium; N, nodule. Boxes show the first quartile, median and third quartile; whiskers show minimum and maximum values; dots show data points. (***B***) Split luciferase complementation assay between MYB3R1 and MYB3R2. nLUC-tagged MYB3R1 was co-infiltrated into *N. benthamiana* leaves along with the cLUC-tagged MYB3R2. (***C***) Relative expression of cell cycle related genes in EV and *MYB3R1-SRDX* transgenic hairy roots. Their expression was normalized to *MtEF-1*. The presented data are from three independent biological repeats. Data are mean ± SD. Means were compared using Student's *t*-test, * indicates $P<0.05$. (***D***) GUS staining of *ENOD11*$_{pro}$*:GUS* stable transgenic plants expressing *MYB3R1-OE* or *MYB3R1-SRDX*. Scale bars, 1 mm in top and middle panels, 500 μm in the bottom panel. (***E***) Relative expression of *MYB3R1* gene in root hairs at 5 dpi. The data are from previous transcriptomic dataset (2, 13). Experiments were carried out twice with similar results.

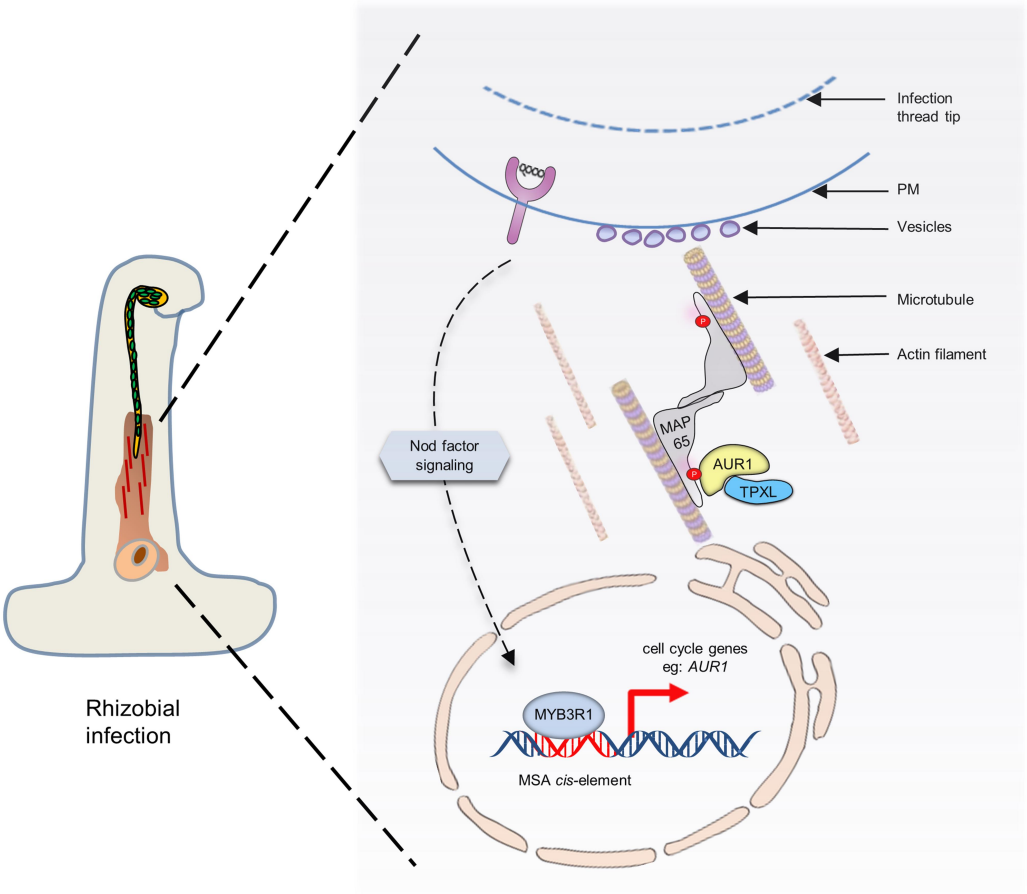


Fig. S8. A model showing MYB3R-TPXL-AUR1-MAP65 regulates rhizobial infection.

Table S1. Primers used in this study.

primer name	sequence (from 5'- to -3')	purpose
pAUR1-F	CGGTTCTTATGGAATACCATAATC	promoter-GUS
pAUR1-R	ATCGATCTCAGAGAGAGAGAGAG	promoter-GUS
pMYB3R1-F	CAACTTAATTCAATGGTTGCAACC	promoter-GUS
pMYB3R1-R	AGTGTACCGGAATTTCGACTTGTC	promoter-GUS
pCYCA3;1-F	AGAGACACTGAAACCTGCATG	promoter-GUS
pCYCA3;1-R	CTCAGCTCAGATTGAAGGAG	promoter-GUS
pCYCD1;1-F	ACTTCACGTCATGCCTGCTG	promoter-GUS
pCYCD1;1-R	TGCTTTGTCATGTTGAGTCACC	promoter-GUS
pCDKB2;2-F	GCCATGATGCATGGATACAAG	promoter-GUS
pCDKB2;2-R	GATTGAATCGATCGAGAAAATTAG	promoter-GUS
pCDKC2;1-F	GACTCAAATAGACTCAACCCAC	promoter-GUS
pCDKC2;1-R	ATTAATGGACCCTTGTAGATATTC	promoter-GUS
pTPXL2-F	GTTTTCATTCCAAGGAAGGTG	promoter-GUS
pTPXL2-R	GTTTTTTTTCTTCTTCTTTCTCATAC	promoter-GUS
pTPXL3-F	CACGAAGTGATGAACGTTTAAG	promoter-GUS
pTPXL3-R	CAATGATTCGTAACGAGCTAAC	promoter-GUS
pMAP65-1-F	GCAATTCCAACATGTTTGGTTC	promoter-GUS
pMAP65-1-R	GAAACAGAGTGTGTGTTGTGTT	promoter-GUS
pMAP65-9-F	ACGGAGGGAGTATATTCATAAC	promoter-GUS
pMAP65-9-R	TTTGGAGCACCTTTATCAACTTA	promoter-GUS
AUR1-gRNA1-F	TGGTCTCGATTGGAGGTATCAGGTTCAGCTAGTTTTAGAGCTAGAAATAGC	CRISPR
AUR1-gRNA2-R	TGGTCTCGAAACTGGCAATATATAAGGGCTCGCAATCACTACTTCGACTCTAG	CRISPR
MYB3R1-gRNA1-F	TGGTCTCGATTGGATGGTGTTACTCTCCAAGGTTTTAGAGCTAGAAATAGC	CRISPR
MYB3R1-gRNA2-R	TGGTCTCGAAACGGACGTATCGGTAAGCAATGCAATCACTACTTCGACTCTAG	CRISPR
NFP-gRNA1-F	TGGTCTCGATTGTGAAGGAGGAGAATCCACAGTTTTAGAGCTAGAAATAGC	CRISPR
NFP-gRNA2-R	TGGTCTCGAAACCAAGTTACAGGTACTAGTA CAATCACTACTTCGACTCTAG	CRISPR
pAUR1-MSA-T1-F	TGGTCTCGATTGAGAGGCAGGATTGGAACACGGTTTTAGAGCTAGAAATAGC	CRISPR
pAUR1-MSA-T2-R	TGGTCTCGAAACTCGTTCATTCATTCGTTAAT CAATCACTACTTCGACTCTAG	CRISPR
MtEF1-qF	CTTTGCTTGGTGCTGTTTAGATGG	qPCR
MtEF1-qR	ATTCCAAAGGCGGCTGCATA	qPCR
AUR1-qF	GATCGATATGGCCATCGCAAC	qPCR
AUR1-qR	GATCCTTAGCTGAACCTGATAC	qPCR
MYB3R1-qF	ATGGGAAGTGAGTCGACAATTCC	qPCR
MYB3R1-qR	TATCCTCCTCGGGTGTCCATTG	qPCR
CYCA3;1-qF	CTTAGGCCAAAGAAATCTGATGC	qPCR
CYCA3;1-qR	TCTGAACCTCCATTGTGCGAAG	qPCR
CYCD1;1-qF	GTCGTCATTATTTGCCGAGGAAG	qPCR
CYCD1;1-qR	CTCATGTACCTTGAGAATCCATG	qPCR
CDKB2;2-qF	GAAGAAGACTCGTCTTCATGAAG	qPCR
CDKB2;2-qR	TCCAACAACCTAACAACATGTGG	qPCR
TPXL2-qF	GACGATTGTGAAGCAGAACAATG	qPCR

TPXL2-qR	GTTGCACCACCATTTCCGTATC	qPCR
TPXL3-qF	AGCTCGTTACGAATCATTGATGAC	qPCR
TPXL3-qR	GTGCTGCGTCGAATTCGTATTC	qPCR
MAP65-1-qF	GAACTTCAATGGCCGTAACTGAAG	qPCR
MAP65-1-qR	CCTGCTCTAACTGAAGAAGCATC	qPCR
MAP65-9-qF	CACTAAACTTGTTGATAAAGGTGC	qPCR
MAP65-9-qR	CTGAAGTTCTTGAAGAAGTGATCC	qPCR
NIN-qF	GCAATGTGGGGATTTAGAGATT	qPCR
NIN-qR	GGAAGATTGAGAGGGGAAG	qPCR
Vapyrin-qF	TCATCCTCCACAACAACAAGGT	qPCR
Vapyrin-qR	TCAAGCACTTCTCTTATGTCATCCATTG	qPCR
KNOLLE-qF	CTCTCTGGACTTAAAGATGGTTCA	qPCR
KNOLLE-qR	GCCTAAGTCCCTGAAACTCCAT	qPCR
ENDO11-qF	TTCTTGTACTCGCTAGGGTTAGTGTT	qPCR
ENDO11-qR	GAGGCTTGTAAGTAGGAGGAGGC	qPCR
MSA-qF	AAGAATAGAGAGAGGCAGGATTG	qPCR
MSA-qR	GCGATGAAGATGAATGCAGCTC	qPCR
CT1-qF	CAACAACAACACACCGAACTAAC	qPCR
CT1-qR	CATTTATGAAAGTTGAAGTCTGAC	qPCR
CT2-qF	ATCATATCGTTGCGTTGAAAGTAC	qPCR
CT2-qR	AGGGTGTCGCAGATGACTCTG	qPCR
probe-F	AGCCAGTGGCGATAAGAGAATAGAGAGAGGCAGGATTG	EMSA
probe-R	AGCCAGTGGCGATAAGTGCTTCTTCCTTCGTTCATTCAT	EMSA
AUR1-771-F	GGACGAGCTCGGTACCCGGGATCCATGGCCATCGCAACAGAGACTC	LUC
AUR1-771-R	GGGACGCGTACGAGATCTGGTCGACGCTCCTGTATATACCGGAG	LUC
TPXL2-772-F	TCGTACGCGTCCCGGGGCGGTACCATGGAAGAAGAGTTCGTGG	LUC
TPXL2-772-R	CGAACGAAAGCTCTGCAGGTCGACTTAACATACCTCATGCTCCTG	LUC
TPXL2-N-772-F	TACGCGTCCCGGGGCGGTACCATGGAAGAAGAGTTCGTGG	LUC
TPXL2-N-772-R	AACGAAAGCTCTGCAGGTCGACTCAGTGAGGATCCCCACAAGTA	LUC
TPXL3-N-772-F	TACGCGTCCCGGGGCGGTACCATGACGATGATGATGATGATG	LUC
TPXL3-N-772-R	AACGAAAGCTCTGCAGGTCGACTCATTGTTTATCATTATCATCCACA	LUC
TPXL3-772-F	TCGTACGCGTCCCGGGGCGGTACCATGACGATGATGATGATGATG	LUC
TPXL3-772-R	CGAACGAAAGCTCTGCAGGTCGACTCACCGAATTCCCAAGCTCC	LUC
MAP65-1-772-F	TCGTACGCGTCCCGGGGCGGTACCATGGCCGTAACTGAAGCTC	LUC
MAP65-1-772-R	CGAACGAAAGCTCTGCAGGTCGACTTAGGGTGATGTTGGGATGG	LUC
MAP65-9-772-F	GTACGCGTCCCGGGGCGGTACCATGCATAAACCTCATAATGATC	LUC
MAP65-9-772-R	ACGAAAGCTCTGCAGGTCGACTCATGCTAACAAAAAACCAAGC	LUC
TPXL2-coip-F	TTGTTGATGTGATTACAGTCTAGAATGGAAGAAGAGTTCGTG	Co-IP
TPXL2-HA-R	TGGAACATCGTATGGGTAGGTACCACATACCTCATGCTCCTGG	Co-IP
AUR1-coip-F	TTGTTGATGTGATTACAGTCTAGAATGGCCATCGCAACAGAGACTC	Co-IP
AUR1-flag-R	GTCGTGGTCCTTATAGTCGGTACCGCTCCTGTATATACCGGAG	Co-IP
AUR1-HA-R	TGGAACATCGTATGGGTAGGTACCGCTCCTGTATATACCGGAG	Co-IP
TPXL3-coip-F	TTGTTGATGTGATTACAGTCTAGAATGACGATGATGATGATGATG	Co-IP

TPXL3-HA-R	TGGAACATCGTATGGGTAGGTACCCCGAATTCCCAAGCTCCTC	Co-IP
MAP65-1-coip-F	TTGTTGATGTGATTACAGTCTAGAATGGCCGTAACTGAAGCTC	Co-IP
MAP65-1-flag-R	GTCGTGGTCCTTATAGTCGGTACCGGGTGATGTTGGGATGGGTT	Co-IP
MAP65-9-coip-F	TTGTTGATGTGATTACAGTCTAGAATGCATAAACCTCATAATGATC	Co-IP
MAP65-9-flag-R	GTCGTGGTCCTTATAGTCGGTACCTGCTAACAAAAAACCAAGCC	Co-IP
GST-AUR1-F	CCGCGTGGATCCCCGGAATTCATGGCCATCGCAACAGAGACTC	fusion protein
GST-AUR1-R	GATGCGGCCGCTCGAGTCGACCTAGCTCCTGTATATACCGGAG	fusion protein
GST-TPXL3-N-F	CCGCGTGGATCCCCGGAATTCATGACGATGATGATGATGATG	fusion protein
GST-TPXL3-N-R	GATGCGGCCGCTCGAGTCGACTTGTTTATCATTATCATCCACATTTTC	fusion protein
GST-TPXL2-N-F	CCGCGTGGATCCCCGGAATTCATGGAAGAAGAGTTCGTGG	fusion protein
GST-TPXL2-N-R	GATGCGGCCGCTCGAGTCGACTCAGTGAGGATCCCCACAAGTA	fusion protein
GST-MAP65-1-F	CCGCGTGGATCCCCGGAATTCATGGCCGTAACTGAAGCTC	fusion protein
GST-MAP65-1-R	GATGCGGCCGCTCGAGTCGACTTAGGGTGATGTTGGGATGG	fusion protein
GST-MYB3R1-F	CCGCGTGGATCCCCGGAATTCATGGGAAGTGAGTCGACAATTC	fusion protein
GST-MYB3R1-R	GATGCGGCCGCTCGAGTCGACCTATCTACAGCCCTTCAACAAATAC	fusion protein
GST-MYB3R1-N-R	GATGCGGCCGCTCGAGTCGACctaATTTTGAAACTGAGCAAGCAAG	fusion protein

Table S2. Genes mentioned in this study.

gene symbol	gene model	description
AUR1	Medtr3g110405	Serine/threonine-protein kinase Aurora
AUR3a	Medtr8g035890	Serine/threonine-protein kinase Aurora
AUR3b	Medtr8g040300	Serine/threonine-protein kinase Aurora
ChOMT3	Medtr7g011900	Chalcone-O-Methyltransferase
CDKB2;2	Medtr1g075610	Cyclin-Dependent Kinase
CDKC2;1	Medtr1g098300	Cyclin-Dependent Kinase
CYCA3;1	Medtr3g102530	A-type Cyclin
CYCD1;1	Medtr8g063120	D-type Cyclin
ENOD11	Medtr3g415670	Early Nodulin 11
ENOD40	Medtr8g069785	Early Nodulin 40
ERN1	Medtr7g085810	Ethylene Response Factor required for Nodulation 1
KNOLLE	Medtr5g012010	SNARE protein Syntaxin 1 and related protein
MAP65-1	Medtr5g093860	65-KD Microtubule-associated protein
MAP65-3	Medtr1g076700	65-KD Microtubule-associated protein
MAP65-5	Medtr8g041400	65-KD Microtubule-associated protein
MAP65-8	Medtr3g085440	65-KD Microtubule-associated protein
MAP65-9	Medtr6g061690	65-KD Microtubule-associated protein
MYB2R1	Medtr1g100667	R2R3-Myb superfamily Transcription factor
MYB3R1	Medtr3g110028	R1R2R3-Myb superfamily Transcription factor
MYB3R2	Medtr1g026870	R1R2R3-Myb superfamily Transcription factor
NFP	Medtr5g019040	Nod Factor Perception
NIN	Medtr5g099060	Nodule Inception
NSP2	Medtr3g072710	Nodulation Signaling Pathway 2
TPXL2	Medtr6g463310	Targeting protein for Xklp2 (TPX2) like protein
TPXL3	Medtr4g101890	Targeting protein for Xklp2 (TPX2) like protein
TUB6	Medtr8g107250	Beta-tubulin
WDL	Medtr3g107050	Wave dampened like protein
VPY	Medtr6g027840	Vapyrin
AtAUR1	At4g32830	Serine/threonine-protein kinase Aurora
AtAUR2	At2g25880	Serine/threonine-protein kinase Aurora
AtAUR3	At2g45490	Serine/threonine-protein kinase Aurora
AtMAP65-1	At5g55230	65-KD Microtubule-associated protein
AtMAP65-2	At4g26760	65-KD Microtubule-associated protein
AtMAP65-3	At5g51600	65-KD Microtubule-associated protein
AtMAP65-4	At3g60840	65-KD Microtubule-associated protein
AtMAP65-5	At2g38720	65-KD Microtubule-associated protein
AtMAP65-6	At2g01910	65-KD Microtubule-associated protein
AtMAP65-7	At1g14690	65-KD Microtubule-associated protein
AtMAP65-8	At1g27920	65-KD Microtubule-associated protein

AtMAP65-9	At5g62250	65-KD Microtubule-associated protein
AtMYB3R1	At4g32730	R1R2R3-Myb superfamily Transcription factor
AtMYB3R2	At4g00540	R1R2R3-Myb superfamily Transcription factor
AtMYB3R3	At3g09370	R1R2R3-Myb superfamily Transcription factor
AtMYB3R4	At5g11510	R1R2R3-Myb superfamily Transcription factor
AtMYB3R5	At5g02320	R1R2R3-Myb superfamily Transcription factor
AtTPX2	At1g03780	Targeting protein for Xklp2
AtTPXL2	At4g11990	Targeting protein for Xklp2 (TPX2) like protein
AtTPXL3	At4g22860	Targeting protein for Xklp2 (TPX2) like protein

SI References

1. S. K. Gomez et al., *Medicago truncatula* and *Glomus intraradices* gene expression in cortical cells harboring arbuscules in the arbuscular mycorrhizal symbiosis. *BMC Plant Biol*. **9**, 10 (2009).
2. A. Breakspear et al., The root hair "infectome" of *Medicago truncatula* uncovers changes in cell cycle genes and reveals a requirement for auxin signaling in rhizobial infection. *Plant Cell* **26**, 4680-4701 (2014).
3. B. Roux et al., An integrated analysis of plant and bacterial gene expression in symbiotic root nodules using laser-capture microdissection coupled to RNA sequencing. *Plant J.* **77**, 817-837 (2014).
4. T. Li et al., Phosphorylation and chromatin tethering prevent cGAS activation during mitosis. *Science* **371**, 1221-+ (2021).
5. Y. Kim, A. J. Holland, W. Lan, D. W. Cleveland, Aurora kinases and protein phosphatase 1 mediate chromosome congression through regulation of CENP-E. *Cell* **142**, 444-455 (2010).
6. J. F. Arrighi et al., The *Medicago truncatula* lysine motif-receptor-like kinase gene family includes *NFP* and new nodule-expressed genes. *Plant Physiol.* **142**, 265-279 (2006).
7. B. B. Amor et al., The *NFP* locus of *Medicago truncatula* controls an early step of Nod factor signal transduction upstream of a rapid calcium flux and root hair deformation. *Plant J.* **34**, 495-506 (2003).
8. I. Damiani et al., 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 (2016).
9. A. Dereeper et al., Phylogeny.fr: robust phylogenetic analysis for the non-specialist. *Nucleic Acids Res.* **36**, W465-W469 (2008).
10. R. C. Edgar, MUSCLE: a multiple sequence alignment method with reduced time and space complexity. *BMC Bioinformatics* **5**, 1-19 (2004).
11. S. Guindon, O. Gascuel, A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Syst. Biol.* **52**, 696-704 (2003).
12. E. D. Tomastikova et al., Functional divergence of microtubule-associated TPX2 family members in *Arabidopsis thaliana*. *Int. J. Mol. Sci.* **21** (2020).
13. C.-W. Liu et al., NIN acts as a network hub controlling a growth module required for rhizobial infection. *Plant Physiol.* **179**, 1704-1722 (2019).