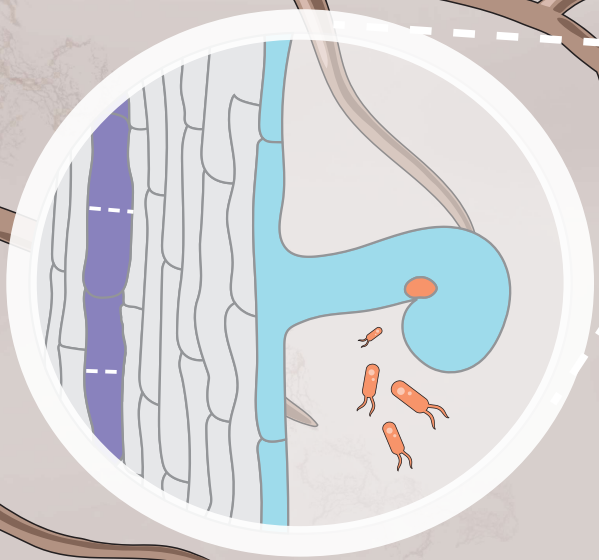
Decoding Nodules

The spatiotemporal hormonal regulation of tissue-specific nodule initiation in *Medicago truncatula*



Sophia Müller

Propositions

1. Localized changes in expression patterns are masked in bulk samples.
   (this thesis)
2. Single-cell transcriptomics does not provide answers but enables us to ask better questions.
   (this thesis)
3. Science is limited by what we can prove.
4. Without explaining the nature of Genetically Modified Organisms, society will never accept nitrogen-fixing cereals.
5. Micronutrient deficiencies as the cause of obesity should be a focus of agricultural research.
6. Plan B is the dead of plan A.
7. By making higher education the ultimate achievement, we destabilize the pillars of society.

Propositions belonging to the thesis, entitled:

Decoding Nodules: The spatiotemporal hormonal regulation of tissue-specific nodule initiation in *Medicago truncatula*

Sophia Müller

Wageningen, 5th June 2026

Decoding Nodules

The spatiotemporal hormonal regulation of tissue-specific nodule initiation in *Medicago truncatula*

Sophia Müller

Thesis Committee

Promotor

Prof. Dr V. Willemsen
Professor of Cell and Developmental Biology
Wageningen University & Research

Co-promotor

Dr W. Kohlen
Assistant professor in Cell and Developmental Biology
Wageningen University & Research

Other members

Prof. Dr *ir.* L. M. Trindade, Wageningen University & Research
Dr K. Kajala, Utrecht University
Dr C. Drapek, Phytoform, Harpenden, UK
Prof. Dr H. M. Schneider, The Leibniz Institute of Plant Genetics and Crop Plant Research (IPK), Gatersleben, Germany

This research was conducted under the auspices of the Graduate School of Experimental Plant Sciences.

Decoding Nodules

The spatiotemporal hormonal regulation of tissue-specific nodule initiation in *Medicago truncatula*

Sophia Müller

Thesis

submitted in fulfilment of the requirements for the degree of doctor
at Wageningen University
by the authority of the Rector Magnificus,
Prof. Dr C. Kroeze,
in the presence of the
Thesis Committee appointed by the Academic Board
to be defended in public
on Friday 5 June 2026
at 3.30 p.m. in the Omnia Auditorium.

Sophia Müller
Decoding Nodules - The spatiotemporal hormonal regulation of tissue-specific nodule initiation in *Medicago truncatula*,
233 pages

PhD thesis, Wageningen University, Wageningen, the Netherlands (2026)
With references, with summaries in English, Dutch and German
https://doi.org/10.18174/681603

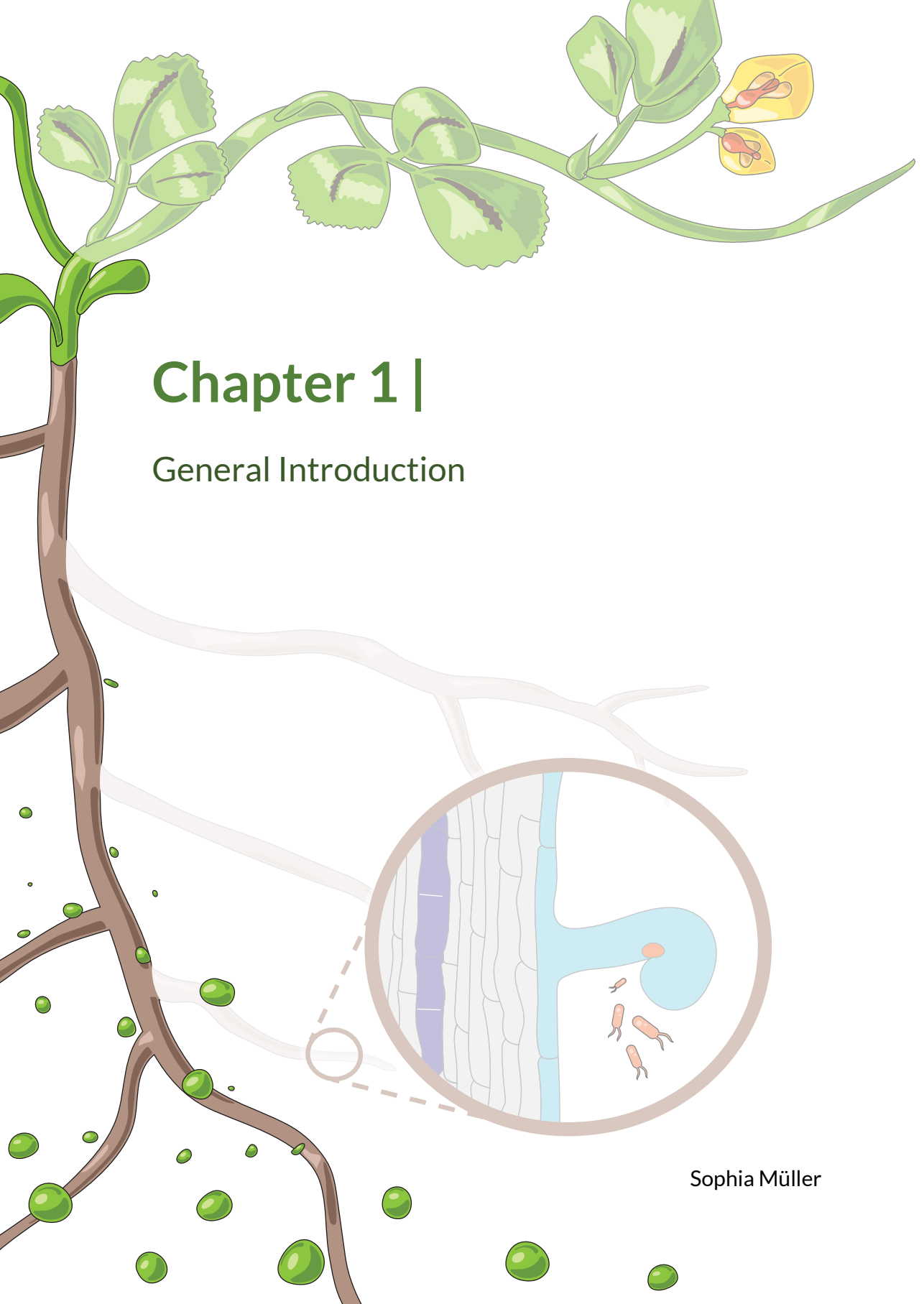
Table of Contents

Chapter 1 |	General Introduction	**7**
Chapter 2 |	Nodule organogenesis in *Medicago truncatula* requires local stage-specific auxin biosynthesis and transport	**29**
Chapter 3 |	Spatiotemporal dynamics of ethylene biosynthesis shape infection and nodule initiation in *Medicago truncatula*	**73**
Chapter 4 |	Establishing robust workflows for single-cell and single-nuclei RNA sequencing in the model legume *Medicago truncatula*	**117**
Chapter 5 |	Single-Cell Transcriptomics-Guided Identification of Transcription Factors with Pericycle-Enriched Expression in *Medicago truncatula*	**161**
Chapter 6 |	General discussion	**199**
Appendix |		**217**
	Summary	**218**
	Samenvatting	**220**
	Zusammenfassung	**222**
	Authorship Statement	**224**
	Acknowledgements	**226**
	About the Author	**231**
	List of Publications	**232**



Chapter 1 |

General Introduction



Sophia Müller

GENERAL INTRODUCTION

Legumes are widely recognized for their high protein content and constitute an essential component of global food and feed systems (Choręziak et al., 2025). In addition to serving as a major protein source, legumes are also employed as green manure to enhance soil fertility and increase crop yields (Sandhya Rani et al., 2022). Their elevated protein levels arise from an evolutionarily derived symbiosis with nitrogen-fixing soil bacteria collectively known as rhizobia (Peoples et al., 1995). Because amino acids, the fundamental building blocks of proteins, require substantial amounts of nitrogen, plants face the limitation that they cannot access atmospheric dinitrogen (N_2), which constitutes ~78% of the Earth's atmosphere (Schlesinger and Bernhardt, 2020).

Not only legumes, but all crop species require nitrogen for optimal growth, and agricultural productivity has long depended on the application of synthetic nitrogen fertilizers to sustain and increase yields (Ladha et al., 2005). However, only 30–50% of applied nitrogen is typically taken up by crops, with the remainder lost through leaching, volatilization, or surface runoff. These losses contribute substantially to environmental degradation, including eutrophication of surface waters, disruption of terrestrial and aquatic ecosystems, and emissions of potent greenhouse gases such as nitrous oxide (Ladha et al., 2005; Zhang et al., 2015). *Vice versa*, soils subjected to intensive agricultural use without nitrogen supplementation rapidly experience nutrient depletion, resulting in significantly reduced yields (Bhatt et al., 2025). As global food demand continues to rise with population growth, combined with a reduction in arable land, improving the efficiency and sustainability of nitrogen use will be essential for future food security (Bhatt et al., 2025; Zhang et al., 2015).

In contrast to plants, certain soil bacteria possess the ability to convert atmospheric N_2 into bioavailable nitrogen forms, such as am monium (NH_4^+), through the process of biological nitrogen fixation (Koushal et al., 2025; Peoples et al., 1995). Among these, rhizobia establish a symbiosis with legumes, in which fixed nitrogen is exchanged for plant-derived photosynthates, primarily carbon-rich sugars produced through photosynthesis (Koushal et al., 2025; Peoples et al., 1995; Roy et al., 2020). The plant additionally provides intracellular accommodation for the bacteria in the form of specialized root organs called nodules, where rhizobia differentiate into endosymbiotic forms capable of sustained nitrogen fixation (Peoples et al., 1995; Roy et al., 2020). Despite being commonly described as mutualistic, the extensive bacterial differentiation and loss of independent viability within nodules raises the question of whether this interaction reflects a controlled exploitation of rhizobial nitrogen-fixing capacity, with the plant maintaining bacterial survival through carbon supply. Nonetheless, the requirement for rhizobial production of Nod factors (NFs), a key signal initiating nodule organogenesis, indicates that rhizobia have also evolved adaptations promoting active engagement with the host. Rhizobia are known to adjust their symbiotic performance depending on the legume genotype, enhancing compatibility; however, clear evidence that such adjustments improve rhizobial evolutionary fitness remains lacking (Quides et al., 2021). Notably, the rhizobial perspective in symbiosis

research remains relatively underexplored, highlighting the need for further studies investigating the benefits, constraints, and evolutionary drivers of nodulation from the microbial viewpoint.

Building on the capacity of symbiotic nitrogen fixation to enhance nitrogen-use efficiency, the transfer of the nodulation trait to non-legume crops has been proposed as a long-term solution and is often described as the “holy grail” of nitrogen symbiosis (Jhu and Oldroyd, 2023). Achieving this would integrate nitrogen acquisition directly into crop physiology, theoretically enabling near-complete utilization of fixed nitrogen while minimizing environmental losses. Although conceptually appealing, engineering nodulation in non-nodulating species has proven far more complex than initially envisioned when the idea emerged nearly three decades ago.

As a result, substantial research has focused on elucidating the molecular, developmental, and evolutionary foundations of symbiotic nitrogen fixation. Whether the evolutionary loss of nodulation in many plant lineages (Van Velzen et al., 2019) reflects constraints or trade-offs that might impede its reintroduction remains an open question, one that paradoxically might only be fully understood once successful engineering becomes feasible.

NODULATION IN A NUTSHELL

To engineer, we need to understand how nodulation functions as a developmental and symbiotic trait. Therefore, it is essential to outline the underlying molecular framework. Nearly 200 genes have been implicated in various stages of nodulation, highlighting the remarkable complexity of this process. Comprehensive reviews by Roy et al. (2020) and Shen and Feng (2023) summarize two decades of progress, including recent insights into the pivotal transcription factor NODULE INCEPTION (NIN). In the following sections, I provide an overview of the key mechanisms most relevant to this thesis, focusing on *Medicago truncatula* (Medicago) and well-characterized orthologs in *Lotus japonicus* (Lotus). These mechanisms can be broadly categorized into: (i) rhizobial signaling, (ii) rhizobial infection, and (iii) nodule organogenesis (Roy et al., 2020; Shen and Feng, 2023; Velandia et al., 2022).

Rhizobial Signaling

Because nodule formation is energetically costly, the initiation of nodulation is tightly regulated by the plant's nitrogen status. Under low nitrate conditions, roots release a mixture of compounds, including specific flavonoids, that recruit compatible rhizobia from the surrounding soil (**Figure 1A|1**) (Peters et al., 1986; Redmond et al., 1986). In response, rhizobia perceive these flavonoids and synthesize lipochitooligosaccharides (LCOs), commonly referred to as the before mentioned Nod factors (NFs) (**Figure 1A|2**) (Dénarié and Cullimore, 1993).

On the plant side, NF perception occurs at the plasma membrane of epidermal root hair cells via heterodimeric LysM receptor kinases (**Figure 1A|3**). In Medicago, key receptors include LysM RECEPTOR KINASE 3 (LYK3), NOD FACTOR PERCEPTION (NFP), and DOES NOT MAKE INFECTIONS 2 (DMI2) (Arrighi et al., 2006; Endre et al., 2002; Limpens et al., 2003; Smit et al., 2007; Stracke et al., 2002). Activated NF receptor complexes initiate a conserved symbiotic signaling cascade by inducing characteristic Ca^{2+} oscillations ("Ca^{2+} spiking") within the nucleus (Charpentier and Oldroyd, 2013). This is mediated by the cation channel DMI1 and the calcium channels CYCLIC NUCLEOTIDE-GATED CHANNELS a-c (CNGC15a–c) (Ané et al., 2004; Charpentier and Oldroyd, 2013; Charpentier et al., 2016).

Nuclear Ca^{2+} spiking activates the Ca^{2+}/calmodulin-dependent protein kinase (CCaMK) DMI3, which phosphorylates INTERACTING PROTEIN OF DMI3 (IPD3) (Gleason et al., 2006; Singh et al., 2014; Tirichine et al., 2006; Yano et al., 2008). The phosphorylated form, IPD3-P, directly promotes transcription of NIN, the master regulator of nodulation (Singh et al., 2014). This activation is reinforced by the GRAS transcription factors NODULATION SIGNALING PATHWAY 1 and 2 (NSP1/NSP2), which cooperate with DMI3 and IPD3-P in a higher-order transcriptional complex that includes DELLA proteins (Hirsch et al., 2009; Jin et al., 2016; Kaló et al., 2005; Singh et al., 2014; Smit et al., 2005). DELLA factors, best known for their role in gibberellic acid signaling, serve as molecular scaffolds and act as positive regulators of nodulation (Fonouni-Farde et al., 2019; Jin et al., 2016).

NIN was the first transcription factor identified as essential for nodulation and remains one of the most studied and central regulators in the nodulation process (Akamatsu et al., 2021; Liu et al., 2019a; Schauser et al., 1999; Schiessl et al., 2019; Shen and Feng, 2023; Soyano et al., 2013). At this stage, the shared signaling pathway between rhizobial nodulation and arbuscular mycorrhizal (AM) symbiosis diverges: NIN drives the rhizobial branch, while the homolog REDUCED ARBUSCULAR MYCORRHIZAL 1 (RAM1) activates the AM fungal pathway (Geurts et al., 2016; Gobbato et al., 2012; Oldroyd, 2013; Oldroyd et al., 2011; Pimprikar et al., 2016; Singh et al., 2014). Once induced, NIN triggers extensive downstream transcriptional programs that coordinate two major processes: rhizobial infection in the epidermis and nodule organogenesis in the inner cortex, endodermis, and pericycle (Liu et al., 2019a; Roy et al., 2020; Shen and Feng, 2023; Soyano et al., 2013; Soyano et al., 2014; Xiao et al., 2014; Yano et al., 2008).

Rhizobial Infection

In the epidermis, approximately 54% of the genes required for infection are dependent on NIN, underscoring its central role in coordinating the transcriptional programs that regulate cell cycle progression, cell wall remodeling, hormonal signaling, and nutrient uptake necessary to accommodate invading rhizobia (Liu et al., 2019a; Shen and Feng, 2023). Within minutes of NF perception, the microtubule cytoskeleton reorganizes to trap rhizobia at the root hair tip, thereby forming the infection pocket, a microenvironment in which rhizobia proliferate prior to the initiation of the infection thread (IT) (**Figure 1A|4**) (Oldroyd et al., 2011; Sprent and James, 2007; Timmers, 2008).

A key transcription factor in this early infection process is ETHYLENE RESPONSE FACTOR REQUIRED FOR NODULATION 1 (ERN1) (Kawaharada et al., 2017). In Lotus, ERN1 is directly activated by the CCaMK/CYCLOPS complex (the functional equivalent of DMI3/IPD3 in Medicago) and additionally depends on NIN for full activation (Cerri et al., 2017; Liu et al., 2023). This could be an indirect regulation as direct binding of NIN to the ERN1 promoter has not been demonstrated (**Figure 1B|5**). ERN1 subsequently induces the expression of EARLY NODULIN 11 and 12 (ENOD11/ENOD12), both of which are essential for proper IT initiation and progression (**Figure 1C|6**) (Andriankaja et al., 2007).

Formation of the infection thread requires localized degradation and remodeling of the cell wall at the root hair tip, enabling the plasma membrane to invaginate around the dividing rhizobia (Van Spronsen et al., 1994). This membrane-lined tubular structure elongates through the epidermis and traverses the underlying cortical cell layers, guided by coordinated cytoskeletal and vesicle trafficking processes (Liu et al., 2019b; Murray et al., 2011; Oldroyd et al., 2011; Pumplin et al., 2010). Meanwhile, nodule primordium formation proceeds within the root interior. In Medicago, rhizobia are released from the IT into cells derived from cortical layers 4 and 5 at stage IV of nodule development (**Figure 1D|7**) (Gage, 2004; Xiao et al., 2014).

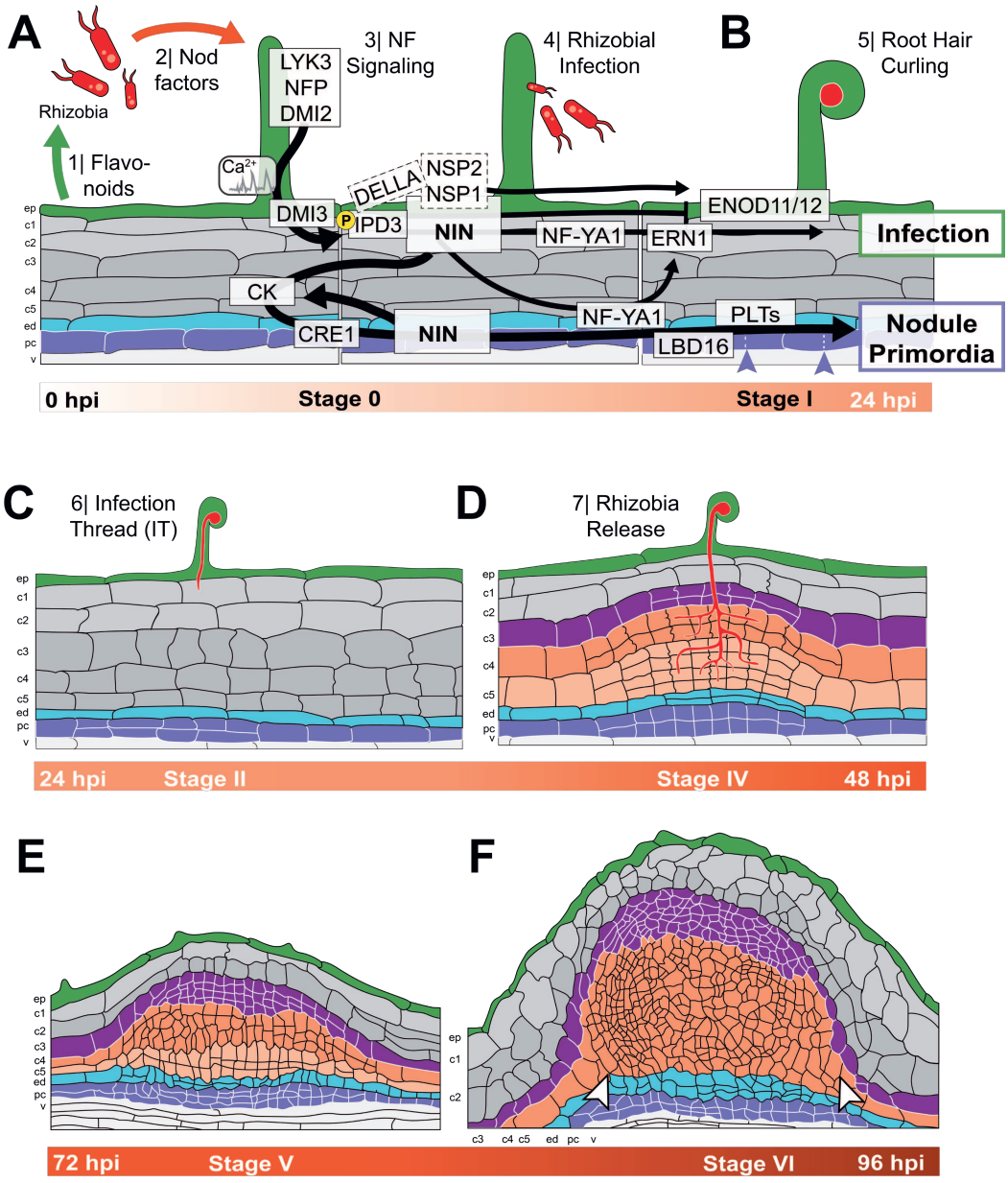


Figure 1 | Molecular signaling and developmental progression of nodule initiation in *Medicago truncatula*. (A-F) A schematic overview of the key signaling events and tissue-level changes underlying rhizobial infection and nodule organogenesis from 0 to 96 hours post inoculation (hpi) (Xiao et al., 2014). **(A-B)** Early signaling and activation of epidermal and inner root tissues (0–24 hpi). Following Nod factor (NF) perception by the LysM receptor kinases LYK3, NFP, and DMI2 (Stage 0), calcium oscillations activate the CCaMK DMI3 and its phosphorylation target IPD3. Together with DELLA, NSP1, and NSP2, this complex induces the expression of the master regulator NIN. In the epidermis, NIN activates early infection genes including ERN1, NF-YA1, and ENOD11/12, enabling infection pocket formation and infection thread (IT) initiation. Concurrently, NIN-dependent cytokinin (CK) signaling is triggered in the epidermis, and CK is perceived in the pericycle by its receptor CRE1, initiating NIN expression in inner tissues. *(continued on next page)*

Figure 1 (continued) | NIN subsequently activates NF-YA1, LBD16, and PLETHORA (PLT) transcription factors to prime the pericycle, endodermis, and inner cortex for nodule organogenesis (Stage I, 24 hpi). **(C-D)** Early primordium development and infection thread progression (24–48 hpi). Stage II (24 hpi): Anticlinal cell divisions begin in the pericycle (purple), while endodermis and inner cortical layers (C3–C5) exhibit the first periclinal and anticlinal divisions. The infection thread continues to grow through the epidermis toward the cortex. Stage IV (48 hpi): The nodule primordium becomes morphologically distinct as inner cortical layers (C4/C5) proliferate (orange/red). The infection thread penetrates the cortex, and rhizobia are released into newly divided host cells derived from C4/C5. **(E-F)** Nodule primordium patterning and establishment of the nodule structure (72–96 hpi). Stage V (72 hpi): Pericycle, endodermis, and cortical layers undergo further proliferation. C4/C5-derived cells begin endoreduplication and enlarge, preparing for symbiosome formation. Stage VI (96 hpi): Distinct nodule zones are established. C3-derived cells form the nascent nodule meristem (pink), C4/C5-derived cells develop into the future infection and fixation zones (orange), and the pericycle and endodermis form the nodule base (purple/turquoise). Rhizobia-containing symbiosomes differentiate within the central tissue, marking the onset of nitrogen fixation. Gene names are defined in the Glossary.

Inside these infected cells, both the plant nucleus and the internalized rhizobia undergo endoreduplication, resulting in the formation of enlarged, metabolically active structures, required for symbiotic exchange (Cebolla et al., 1999; Mergaert et al., 2006; Vinardell et al., 2003). Rhizobia differentiate into bacteroids and become enclosed within plant-derived membrane compartments known as symbiosomes, the functional units of nitrogen fixation (Roth and Stacey, 1989). Within symbiosomes, rhizobia reduce atmospheric N_2 to ammonium, while receiving plant-derived photosynthates as carbon sources, thus completing the mutualistic exchange central to the symbiosis (Peoples et al., 1995; Roy et al., 2020).

Nodule Organogenesis

In parallel to the infection process at the epidermis, the plant initiates the formation of the nodule primordium, which will ultimately serve as the host tissue for nitrogen-fixing rhizobia. Nodule organogenesis in Medicago has been systematically divided into discrete developmental stages based on morphological criteria (Xiao et al., 2014). For the purposes of this thesis, the stages corresponding to 24, 48, 72, and 96 hours post inoculation (hpi) are most relevant.

During stage I (24 hpi), the earliest visible response is the onset of anticlinal cell divisions in the pericycle (**Figure 1B**, purple arrowheads), accompanied by the formation of the infection pocket in the root hair. In stage II (27–44 hpi), periclinal divisions emerge in the pericycle, and anticlinal divisions appear in the endodermis and inner cortical layers (C5, C4, C3). Concurrently, the infection thread begins to advance toward the cortex (**Figure 1C**).

Stage III also involves continued proliferation of the endodermis, pericycle, and cortical layers C3–C5, accompanied by further progression of the infection thread. By stage IV (48 hpi), a distinct nodule primordium has formed from the dividing cortical cells, and rhizobia are released from the infection thread into cells derived from cortical layers C4 and C5 (**Figure 1D|7**). These inner cortical layers (C4/C5; orange/red) give rise to the early fixation zone. Cortex layer C3 (violet) becomes the future meristem, which will provide a steady supply of infectable cells to the infection zone and thus extend the fixation zone, while the endodermis (turquoise) and pericycle (purple) constitute the nodule base, tissues that remain uninfected and serve as structural support.

During stage V (72 hpi), all layers continue to proliferate until the C4/C5-derived region reaches approximately eight cell layers and the endodermis/pericycle reach about six layers (**Figure 1E**). Although cortical cells begin to enter endoreduplication as early as 24 hpi, this trait becomes more prominent by this stage. In stage VI (80–96 hpi), cell division within the C4/C5, endodermal, and pericycle layers is largely complete. C4/C5-derived cells begin differentiating into symbiosome-containing cells, while C3-derived cells develop into the nodule meristem, which remains active and provides new cells for ongoing nodule growth. Simultaneously, the nodule vasculature initiates formation at the periphery (**Figure 1F**, white arrowheads).

Following stage VI, the developing nodule transitions into its mature organization, comprising: (i) a persistent nodule meristem (C3-derived), (ii) an infection zone, (iii) a nitrogen-fixation zone, and (iv) eventually a senescence zone (not depicted). In this mature structure, newly formed cells are successively infected by rhizobia in the infection zone and differentiate into nitrogen-fixing symbiosomes in the fixation zone. Senescence then proceeds in a controlled manner.

In addition to stages I–VI, stage 0 (**Figure 1A**) encompasses root hair curling and early transcriptional responses triggered by NF perception and is also relevant to this thesis. Because these early epidermal processes can be induced solely through NF application, they provide an informative system for dissecting the initial signaling mechanisms independent of rhizobial colonization.

The molecular mechanisms of nodule organogenesis

Nodule organogenesis is orchestrated by the transcription factor NODULE INCEPTION (NIN), which functions as a central regulator downstream of NF perception (Shen and Feng, 2023). In the epidermis, NIN induces genes controlling cytokinin (CK) biosynthesis and homeostasis, thereby establishing a positive feedback loop that reinforces NIN activation and facilitates communication with inner root tissues (Hayashi et al., 2010; Heckmann et al., 2011; Jin et al., 2016; Liu et al., 2019a). In Medicago, a key CK-activating enzyme is LOG1 (LONELY GUY 1), which converts inactive CK precursors into bioactive forms (Mortier et al., 2014). In Lotus, NIN directly induces the expression of *ISOPENTENYLTRANSFERASE 3* (*IPT3*) and activates *CYTOKININ OXIDASE/DEHYDROGENASE 3* (*CKX3*) to maintain CK homeostasis (Chen et al., 2014; Reid et al., 2016). Although direct orthologs have not been fully established in Medicago, multiple candidate genes have been proposed.

A significant insight into the separation of infection and organogenesis came from the Lotus *daphne* mutant, which carries a weak *nin* allele (Yoro et al., 2014). While capable of forming infection threads, this mutant fails to initiate cell divisions in the inner cortex, preventing primordium formation (Yoro et al., 2014). A similar phenotype in Medicago re-sparked the hypothesis that CK could act as a mobile signal transmitted from the epidermis to the pericycle during nodule initiation (Liu et al., 2019a). In the pericycle, CK is perceived by the receptor CYTOKININ RESPONSE 1 (CRE1), which activates NIN expression in the pericycle, endodermis, and inner cortex (Liu et al., 2019a).

Downstream of NIN, several transcription factors play essential roles in primordium initiation and organization. NUCLEAR TRANSCRIPTION FACTOR Y, SUBUNIT A1 (NF-YA1) is a direct NIN target required for establishing and maintaining the nodule meristem (Laporte et al., 2014; Soyano et al., 2013). NIN also induces the root meristem regulators *PLETHORA1-4 (PLT1–4)* as early as stage II, supporting primordium proliferation and differentiation (Franssen et al., 2015). Another NIN-dependent regulator, *LATERAL ORGAN BOUNDARIES DOMAIN 16 (LBD16*), is notable because it is also essential for lateral root initiation, marking a developmental convergence between lateral roots and nodules (Schiessl et al., 2019). In both contexts, LBD16 promotes the establishment of an auxin maximum, which is thought to be a prerequisite for cell cycle re-entry. However, whereas lateral root development proceeds entirely from the pericycle, nodule development involves a transient pericycle contribution limited to the nodule base, while the bulk of the primordium originates from cortical cells (Schiessl et al., 2019; Xiao et al., 2014).

Understanding how NIN-mediated activation of the pericycle contributes to the formation of a largely cortex-derived organ is a central focus of this thesis. Because phytohormones play critical roles in coordinating organogenesis, the following sections examine their spatiotemporal regulation during early nodulation.

HORMONAL REGULATION OF NODULATION

Plant hormones are fundamental regulators of plant growth, development, and environmental responses, and the complex process of nodulation is no exception. Considerable research has focused on elucidating the hormonal networks underlying rhizobial infection and nodule development, as recently reviewed by Velandia et al. (2022), Lin et al. (2020), and Roy et al. (2020). In the context of this thesis, emphasis is placed on three phytohormones, auxin (AUX), cytokinin (CK), and ethylene (ET), which together constitute the central regulatory trio, here referred to as "ACE." These hormones have particularly prominent roles during infection, nodule initiation, and subsequent organ development.

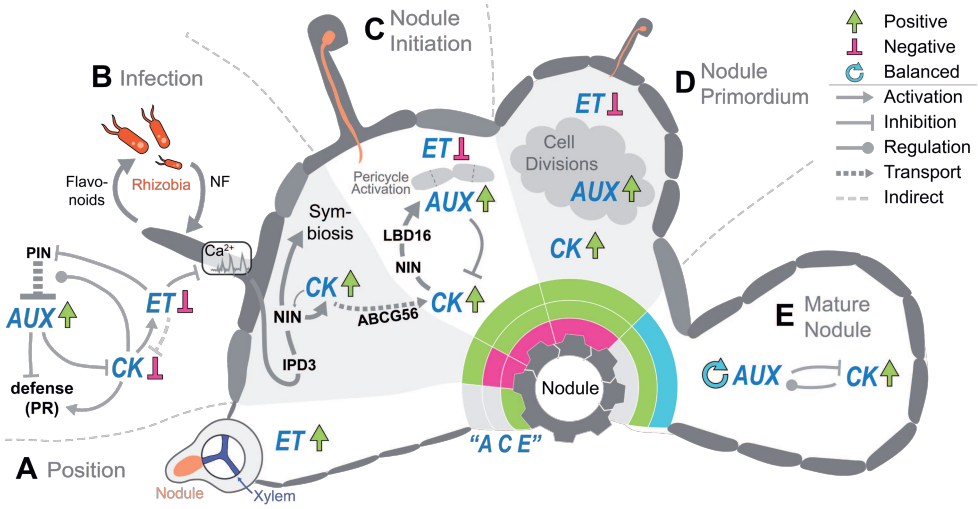


Figure 2 | Stage-specific hormonal regulation of nodulation in *Medicago truncatula*. A schematic representation of the coordinated roles of auxin (AUX), cytokinin (CK), and ethylene (ET) ("ACE") during successive stages of rhizobial infection, nodule initiation, and nodule development (modified from (Müller and Kohlen, 2024)). Panels A–E illustrate how finely tuned interactions between AUX, CK, and ET orchestrate infection, organ initiation, and maturation during legume nodulation. **(A) Nodule positioning**, ethylene (ET) restricts nodule formation to the xylem pole, acting as a positive positional cue. **(B) Rhizobial infection**, upon Nod factor (NF) perception, rhizobia trigger Ca^{2+} spiking and early symbiotic signaling in the epidermis. Ethylene inhibits Ca^{2+} oscillations and NF signaling, while CK further enhances ET responses and independently suppresses infection. Auxin positively regulates infection by promoting an AUX maximum through inhibition of PIN-mediated auxin transport, while also repressing pathogenesis response genes to enable symbiosis. **(B-C)** Signal transduction and **nodule initiation**, in the epidermis, NF signaling activates IPD3 and subsequently NIN, which induces CK biosynthesis. CK is transported inward via ABCG56, where it activates NIN, LBD16, and AUX biosynthesis in the pericycle and inner cortex. Together, AUX and CK promote the earliest pericycle cell divisions that initiate the nodule primordium, while ET negatively regulates this process. **(D) Nodule primordium formation**, AUX and CK maintain a positive regulatory environment for continued cell divisions in the developing primordium. ET suppresses primordium development and must remain low for proper organogenesis. Distinct hormonal domains emerge in the pericycle, cortex, and incipient meristem, establishing the early spatial organization of the nodule. **(E) Mature nodule**, AUX and CK continue to support meristem function, vascular development, and tissue patterning. ET does not have a known regulatory role at this stage. Hormonal cross-talk ensures the coordinated development of the meristem, infection zone, and fixation zone. Gene names are defined in the Glossary.

Auxin and cytokinin are among the most extensively studied hormones in plant developmental biology. Their effects are not determined by their absolute concentrations, but by their relative ratio, which dictates developmental outcomes (Schaller et al., 2015). A classic example is plant regeneration: a high auxin-to-cytokinin ratio induces callus formation by reactivating cellular pluripotency; maintaining high auxin promotes root formation, whereas shifting to high cytokinin induces shoot development (Motte et al., 2014). Ethylene, in contrast, is historically associated with fruit ripening in species such as tomato (Liu et al., 2015) and with inhibition of elongation growth, as exemplified by the ethylene-dependent triple response in etiolated seedlings (Bleecker et al., 1988).

Overview of hormonal crosstalk during nodulation

During nodulation, AUX and CK generally act as positive regulators, while ET functions predominantly as a negative regulator (**Figure 2**) (Berrabah et al., 2018; Breakspear et al., 2014; Larrainzar et al., 2015; Lin et al., 2020; Miri et al., 2019; Murray et al., 2007; Nadzieja et al., 2018; Ng et al., 2015; Oldroyd et al., 2001; Penmetsa and Cook, 1997; Plet et al., 2011; Reid et al., 2016; Suzaki et al., 2012; Tirichine et al., 2006; Tirichine et al., 2007; Velandia et al., 2022). Notably, their roles are context-dependent, and in certain stages their functions are reversed. For example, during nodule positioning, ET has a positive regulatory role by restricting nodule formation to the xylem pole (**Figure 2A**), likely through localized inhibition of surrounding tissues (Heidstra et al., 1997; Penmetsa et al., 2003). Conversely, CK negatively impacts rhizobial infection (**Figure 2B**) (Held et al., 2014; Miri et al., 2019; Murray et al., 2007; Reid et al., 2016). A schematic overview of this hormonal regulation is provided in **Figure 2.**

During epidermal infection, ET negatively regulates the NF signaling pathway by suppressing nuclear Ca^{2+} spiking (**Figure 2B**) (Oldroyd et al., 2001). CK enhances ET signaling in a CRE1-dependent manner and independently inhibits infection, whereas ET suppresses CK accumulation and signaling (Jardinaud et al., 2016; Penmetsa et al., 2008; Plet et al., 2011; Reid et al., 2018; van Zeijl et al., 2015). Together, CK and ET act to prevent hyperinfection. In contrast, AUX is required to promote infection (Kohlen et al., 2018; Ng and Mathesius, 2018). A key mechanism is the CK- and ET-enhanced inhibition of PIN-FORMED (PIN) auxin efflux carriers, which supports the establishment of a localized auxin maximum at the infection site, an essential prerequisite for rhizobial entry in Medicago (Kohlen et al., 2018; Ng and Mathesius, 2018; Ng et al., 2015; Plet et al., 2011; Prayitno et al., 2006; Rightmyer and Long, 2011). In addition, AUX suppresses pathogenesis-related (PR) genes, which must be downregulated to prevent activation of defense responses during (Kazan and Manners, 2009). CK acts antagonistically by inducing PR genes, reinforcing its inhibitory role during infection (Naseem and Dandekar, 2012). AUX also negatively regulates CK biosynthesis, helping maintain hormonal balance (Nadzieja et al., 2018; Ng et al., 2015).

As mentioned before, once early signaling has been successfully mediated through the NF cascade, Ca^{2+} spiking activates IPD3, which induces the key transcription factor NIN. NIN

activates a suite of symbiotic genes, including those involved in CK biosynthesis, which in turn amplifies NIN expression. Importantly, CK produced in the epidermis is hypothesized to be transported via ABCG56 to inner root tissues, where it is postulated to activate *NIN*, *LBD16*, and subsequently AUX biosynthesis to trigger pericycle cell divisions (**Figure 2B–C**) (Jarzyniak et al., 2021; Liu et al., 2019a; Schiessl et al., 2019).

Hormonal control of nodule initiation and development

Both AUX and CK promote the cell divisions required for nodule primordium formation. Exogenous CK or CK overexpression alone is sufficient to induce nodule primordia (Cooper and Long, 1994; Gauthier-Coles et al., 2019; Reid et al., 2017; Tirichine et al., 2007), likely by activating the NIN-CK positive feedback loop and stimulating auxin biosynthesis. This CK–AUX interaction mirrors the hormonal logic of lateral root formation, where CK establishes a local field that directs auxin accumulation in the pericycle (Schaller et al., 2015). Throughout nodulation, AUX inhibition of CK biosynthesis and CK regulation of PIN localization are recurring themes in hormonal cross-regulation (**Figure 2B–E**).

ET consistently acts as a negative regulator of nodule initiation and development (**Figure 2B–D**), highlighting the necessity of its tight spatiotemporal control. In mature nodules (**Figure 2E**), AUX and CK continue to play essential roles in sustaining meristem activity and tissue organization, whereas ET appears to have no major regulatory function. AUX remains enriched in the nodule meristem and vasculature, and CK contributes to the stabilization of auxin patterns (Breakspear et al., 2014; Chen et al., 2014; Mathesius et al., 1998; Reid et al., 2016; Reid et al., 2017; Suzaki et al., 2012; Triozzi et al., 2022; van Zeijl et al., 2015).

Nodulation represents a striking example of sophisticated hormonal crosstalk controlling a specialized developmental program. Although AUX, CK, and ET form the core regulatory triad, additional hormones, including jasmonic acid, gibberellins, brassinosteroids, and strigolactones, provide further fine-tuning. Notably, the CK-dependent activation of NIN reveals a specialized regulatory module directing auxin biosynthesis specifically during nodulation. The context-dependent, sometimes opposing roles of “ACE” hormones underscore the importance of precise temporal and spatial regulation in determining whether cells undergo infection, division, or organogenic differentiation.

OUTLINE OF THIS THESIS

The nitrogen-fixing symbiosis between rhizobia and legumes has long attracted interest due to its agricultural relevance and potential for sustainable crop production. Efforts to transfer this trait to non-nodulating species require a comprehensive understanding of the molecular and developmental events underlying nodule formation. Of particular importance is the controlled initiation of nodulation, which shares regulatory features with lateral root development and arbuscular mycorrhizal (AM) symbiosis, yet ultimately diverges to produce a distinct organ. During nodule initiation, pericycle cells are the first to re-enter the cell cycle, but unlike in lateral root development, these cells do not contribute directly to organ formation; instead, the bulk of the nodule originates from the inner cortical cell layers. Understanding how pericycle activation is coordinated and redirected represents a central knowledge gap in the field.

To shed light on this specialization within the nodulation program, this thesis first investigates the hormonal regulation of auxin and ethylene, two phytohormones central to the earliest stages of nodule initiation. In **Chapter 2**, we focus on auxin, which is essential for triggering the first cell divisions during primordium formation. Using spatiotemporal analyses, we characterize auxin biosynthesis and transport throughout early nodulation stages. **Chapter 3** examines ethylene, a gaseous hormone with a predominantly inhibitory role in nodulation. Remarkably, we find that different isoforms of the ethylene biosynthesis genes *MtACS3* and *MtACS10* are oppositely regulated upon rhizobial inoculation, generating a spatial shift of ethylene production from the inner to the outer root tissues without altering overall ACC concentrations. These findings demonstrate that high spatial resolution is critical, as the regulatory dynamics observed in either chapter would not have been detectable in bulk tissue analyses.

This emphasis on spatial resolution motivated the work described in **Chapter 4**, where we optimized and compared nuclei and protoplast isolation methods for single-cell and single-nucleus RNA sequencing. With a critical eye on data quality, we generated (i) a root protoplast reference atlas and (ii) a high-resolution transcriptional map of nodule development at 24, 48, 72, and 96 hours post inoculation. This dataset enabled new biological insights by resolving cell-type-specific transcriptional profiles across the early nodulation time course.

In **Chapter 5**, we applied this dataset to validate predicted cell-type assignments and explore the regulatory landscape of pericycle activation. We identified three transcription factors enriched in the pericycle and vasculature, confirmed their spatial expression patterns using *promoter*::GUS fusions, and perturbed their function through overexpression and RNA interference. While only mild phenotypes were observed, primarily in overexpression lines, these results highlight both the reliability of cell-type identification in single-cell datasets and the need for more stringent candidate selection criteria when dissecting gene function in nodulation.

Finally, **Chapter 6** integrates the findings from all chapters with current literature and outlines future perspectives for understanding pericycle activation, cell-type specification, and the broader goal of engineering nitrogen fixation in non-legume crops.

ACKNOWLEDGMENTS

I would like to thank Viola Willemsen and Wouter Kohlen for their critical feedback on an earlier version of this work.

GLOSSARY

Abbreviation	Full name
AUX	Auxin
CCaMK	CALCIUM/CALMODULIN-DEPENDENT PROTEIN KINASE
CK	Cytokinin
CRE1	CYTOKININ RESPONSE 1
DELLA	DELLA proteins (GA signaling repressors)
DMI2	DOES NOT MAKE INFECTIONS 2 (receptor-like kinase)
DMI3	DOES NOT MAKE INFECTIONS 3 (CCaMK in Lotus)
ENOD	EARLY NODULIN
ERN1	ETHYLENE RESPONSE FACTOR REQUIRED FOR NODULATION 1
ET	Ethylene
GRAS	GIBBERELLIC-ACID INSENSITIVE (GAI), REPRESSOR of GAI (RGA) and SCARECROW (SCR) Transcription factor family
IPD3	INTERACTING PROTEIN OF DMI3 (also known as CYCLOPS in Lotus)
LBD16	LATERAL ORGAN BOUNDARIES DOMAIN 16
LYK3	LYSM-DOMAIN RECEPTOR-LIKE KINASE 3
NFP	NOD FACTOR PERCEPTION
NIN	NODULE INCEPTION
NF-YA1	NUCLEAR FACTOR Y, SUBUNIT A1
NSP1/2	NODULATION SIGNALING PATHWAY PROTEIN (NSP1/NSP2)
PIN	PIN-FORMED auxin efflux carrier
PLT	PLETHORA transcription factor

REFERENCES

Akamatsu, A., Nagae, M., Nishimura, Y., Romero Montero, D., Ninomiya, S., Kojima, M., Takebayashi, Y., Sakakibara, H., Kawaguchi, M. and Takeda, N. (2021). Endogenous gibberellins affect root nodule symbiosis via transcriptional regulation of NODULE INCEPTION in Lotus japonicus. *Plant Journal* 105, 1507–1520.

Andriankaja, A., Boisson-Dernier, A., Frances, L., Sauviac, L., Jauneau, A., Barker, D. G. and De Carvalho-Niebel, F. (2007). AP2-ERF Transcription Factors Mediate Nod Factor–Dependent Mt ENOD11 Activation in Root Hairs via a Novel cis-Regulatory Motif. *Plant Cell* 19, 2866–2885.

Ané, J. M., Kiss, G. B., Riely, B. K., Penmetsa, R. V., Oldroyd, G. E. D., Ayax, C., Lévy, J., Debellé, F., Baek, J. M., Kalo, P., et al. (2004). Medicago truncatula DMI1 required for bacterial and fungal symbioses in legumes. *Science* 303, 1364–1367.

Arrighi, J. F., Barre, A., Ben Amor, B., Bersoult, A., Soriano, L. C., Mirabella, R., De Carvalho-Niebel, F., Journet, E. P., Ghérardi, M., Huguet, T., et al. (2006). The Medicago truncatula lysin [corrected] motif-receptor-like kinase gene family includes NFP and new nodule-expressed genes. *Plant Physiol* 142, 265–279.

Berrabah, F., Balliau, T., Aït-Salem, E. H., George, J., Zivy, M., Ratet, P. and Gourion, B. (2018). Control of the ethylene signaling pathway prevents plant defenses during intracellular accommodation of the rhizobia. *New Phytologist* 219, 310–323.

Bhatt, R., Kunal, Moulick, D., Bárek, V., Brestic, M., Gaber, A., Skalicky, M. and Hossain, A. (2025). Sustainable strategies to limit nitrogen loss in agriculture through improving its use efficiency—aiming to reduce environmental pollution. *J Agric Food Res* 22, 101957.

Bleecker, A. B., Estelle, M. A., Somerville, C. and Kende, H. (1988). *Insensitivity to Ethylene Conferred by a Dominant Mutation in Arabidopsis thaliana*.

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 "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.

Cebolla, A., Vinardell, J. M., Kiss, E., Oláh, B., Roudier, F., Kondorosi, A. and Kondorosi, E. (1999). The mitotic inhibitor ccs52 is required for endoreduplication and ploidy-dependent cell enlargement in plants. *EMBO J* 18, 4476–4484.

Cerri, M. R., Wang, Q., Stolz, P., Folgmann, J., Frances, L., Katzer, K., Li, X., Heckmann, A. B., Wang, T. L., Downie, J. A., et al. (2017). The ERN1 transcription factor gene is a target of the CCaMK/CYCLOPS complex and controls rhizobial infection in Lotus japonicus. *New Phytol* 215, 323–337.

Charpentier, M. and Oldroyd, G. E. D. (2013). Nuclear calcium signaling in plants. *Plant Physiol* 163, 496–503.

Charpentier, M., Sun, J., Martins, T. V., Radhakrishnan, G. V., Findlay, K., Soumpourou, E., Thouin, J., Véry, A. A., Sanders, D., Morris, R. J., et al. (2016). Nuclear-localized cyclic nucleotide-gated channels mediate symbiotic calcium oscillations. *Science (1979)* 352, 1102–1105.

Chen, Y., Chen, W., Li, X., Jiang, H., Wu, P., Xia, K., Yang, Y. and Wu, G. (2014). Knockdown of LjIPT3 influences nodule development in Lotus japonicus. *Plant Cell Physiol* 55, 183–193.

Choręziak, A., Rosiejka, D., Michałowska, J. and Bogdański, P. (2025). Nutritional Quality, Safety and Environmental Benefits of Alternative Protein Sources—An Overview. *Nutrients* 17, 1148.

Cooper, J. B. and Long, S. R. (1994). Morphogenetic Rescue of Rhizobium meliloti Nodulation Mutants by trans-Zeatin Secretion. *Plant Cell* 6, 215–225.

Dénarié, J. and Cullimore, J. (1993). Lipo-oligosaccharide nodulation factors: A new class of signaling molecules mediating recognition and morphogenesis. *Cell* 74, 951–954.

Endre, G., Kereszt, A., Kevei, Z., Mihacea, S., Kaló, P. and Kiss, G. B. (2002). A receptor kinase gene regulating symbiotic nodule development. *Nature 2002 417:6892* 417, 962–966.

Fonouni-Farde, C., Miassod, A., Laffont, C., Morin, H., Bendahmane, A., Diet, A. and Frugier, F. (2019). Gibberellins negatively regulate the development of Medicago truncatula root system. *Sci Rep* 9,.

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 (Cambridge)* 142, 2941–2950.

Gage, D. J. (2004). Infection and Invasion of Roots by Symbiotic, Nitrogen-Fixing Rhizobia during Nodulation of Temperate Legumes. *Microbiology and Molecular Biology Reviews* 68, 280–300.

Gauthier-Coles, C., White, R. G. and Mathesius, U. (2019). Nodulating legumes are distinguished by a sensitivity to cytokinin in the root cortex leading to pseudonodule development. *Front Plant Sci* 9, 430352.

Geurts, R., Xiao, T. T. and Reinhold-Hurek, B. (2016). What Does It Take to Evolve A Nitrogen-Fixing Endosymbiosis? *Trends Plant Sci* 21, 199–208.

Gleason, C., Chaudhuri, S., Yang, T., Muñoz, A., Poovaiah, B. W. and Oldroyd, G. E. D. (2006). Nodulation independent of rhizobia induced by a calcium-activated kinase lacking autoinhibition. *Nature 2006 441:7097* 441, 1149–1152.

Gobbato, E., Marsh, J. F., Vernié, T., Wang, E., Maillet, F., Kim, J., Miller, J. B., Sun, J., Bano, S. A., Ratet, P., et al. (2012). A GRAS-Type Transcription Factor with a Specific Function in Mycorrhizal Signaling. *Current Biology* 22, 2236–2241.

Hayashi, T., Banba, M., Shimoda, Y., Kouchi, H., Hayashi, M. and Imaizumi-Anraku, H. (2010). A dominant function of CCaMK in intracellular accommodation of bacterial and fungal endosymbionts. *Plant Journal* 63, 141–154.

Heckmann, A. B., Sandal, N., Bek, A. S., Madsen, L. H., Jurkiewicz, A., Nielsen, M. W., Tirichine, L. and Stougaard, J. (2011). Cytokinin Induction of Root Nodule Primordia in Lotus japonicus Is Regulated by a Mechanism Operating in the Root Cortex. *https://doi.org/10.1094/MPMI-05-11-0142* 24, 1385–1395.

Heidstra, R., Yang, W. C., Yalcin, Y., Peck, S., Emons, A. M., Van Kammen, A. and Bisseling, T. (1997). Ethylene provides positional information on cortical cell division but is not involved in Nod factor-induced root hair tip growth in Rhizobium-legume interaction. *Development* 124, 1781–1787.

Held, M., Hou, H., Miri, M., Huynh, C., Ross, L., Hossain, M. S., Sato, S., Tabata, S., Perry, J., Wang, T. L., et al. (2014). Lotus japonicus Cytokinin Receptors Work Partially Redundantly to Mediate Nodule Formation. *Plant Cell* 26, 678–694.

Hirsch, S., Kim, J., Muñoz, A., Heckmann, A. B., Downie, J. A. and Oldroyd, G. E. D. (2009). GRAS Proteins Form a DNA Binding Complex to Induce Gene Expression during Nodulation Signaling in Medicago truncatula. *Plant Cell* 21, 545–557.

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.

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.

Jhu, M. Y. and Oldroyd, G. E. D. (2023). Dancing to a different tune, can we switch from chemical to biological nitrogen fixation for sustainable food security? *PLoS Biol* 21, 1–24.

Jin, Y., Liu, H., Luo, D., Yu, N., Dong, W., Wang, C., Zhang, X., Dai, H., Yang, J. and Wang, E. (2016). DELLA proteins are common components of symbiotic rhizobial and mycorrhizal signalling pathways. *Nature Communications 2016 7:1* 7, 1–14.

Kaló, P., Gleason, C., Edwards, A., Marsh, J., Mitra, R. M., Hirsch, S., Jakab, J., Sims, S., Long, S. R., Rogers, J., et al. (2005). Nodulation Signaling in Legumes Requires NSP2, a Member of the GRAS Family of Transcriptional Regulators. *Science (1979)* 308, 1786–1789.

Kawaharada, Y., James, E. K., Kelly, S., Sandal, N. and Stougaard, J. (2017). The Ethylene Responsive Factor Required for Nodulation 1 (ERN1) Transcription Factor Is Required for Infection-Thread Formation in Lotus japonicus. *Mol Plant Microbe Interact* 30, 194–204.

Kazan, K. and Manners, J. M. (2009). Linking development to defense: auxin in plant-pathogen interactions. *Trends Plant Sci* 14, 373–382.

Kohlen, W., Ng, J. L. P., Deinum, E. E. and Mathesius, U. (2018). Auxin transport, metabolism, and signalling during nodule initiation: Indeterminate and determinate nodules. *J Exp Bot* 69, 229–244.

Koushal, S., HM, A. U. R., Yashaswini, M., Vishnoi, M., Yadav, P. and Yasmin, C. (2025). Innovative approaches to biological nitrogen fixation: Enhancing ecological and agricultural sustainability. *International Journal of Research in Agronomy* 8, 776–783.

Ladha, J. K., Pathak, H., Krupnik, T. J., Six, J. and van Kessel, C. (2005). Efficiency of Fertilizer Nitrogen in Cereal Production: Retrospects and Prospects. *Advances in Agronomy* 87, 85–156.

Laporte, P., Lepage, A., Fournier, J., Catrice, O., Moreau, S., Jardinaud, M. F., Mun, J. H., Larrainzar, E., Cook, D. R., Gamas, P., et al. (2014). The CCAAT box-binding transcription factor NF-YA1 controls rhizobial infection. *J Exp Bot* 65, 481–494.

Larrainzar, E., Riely, B. K., Kim, S. C., Carrasquilla-Garcia, N., Yu, H. J., Hwang, H. J., Oh, M., Kim, G. B., Surendrarao, A. K., Chasman, D., et al. (2015). Deep sequencing of the Medicago truncatula root transcriptome reveals a massive and early interaction between nodulation factor and ethylene signals. *Plant Physiol* 169, 233–265.

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 (1979)* 302, 630–633.

Lin, J., Frank, M. and Reid, D. (2020). No Home without Hormones: How Plant Hormones Control Legume Nodule Organogenesis. *Plant Commun* 1,.

Liu, M., Pirrello, J., Chervin, C., Roustan, J. P. and Bouzayen, M. (2015). Ethylene Control of Fruit Ripening: Revisiting the Complex Network of Transcriptional Regulation. *Plant Physiol* 169, 2380.

Liu, J., Rutten, L., Limpens, E., Van Der Molen, T., Van Velzen, R., Chen, R., Chen, Y., Geurts, R., Kohlen, W., Kulikova, O., et al. (2019a). A remote cis-regulatory region is required for nin expression in the pericycle to initiate nodule primordium formation in medicago truncatula. *Plant Cell* 31, 68–83.

Liu, C. W., Breakspear, A., Stacey, N., Findlay, K., Nakashima, J., Ramakrishnan, K., Liu, M., Xie, F., Endre, G., de Carvalho-Niebel, F., et al. (2019b). A protein complex required for polar growth of rhizobial infection threads. *Nature Communications 2019 10:1* 10, 2848-.

Liu, M., Kameoka, H., Oda, A., Maeda, T., Goto, T., Yano, K., Soyano, T. and Kawaguchi, M. (2023). The effects of ERN1 on gene expression during early rhizobial infection in Lotus japonicus. *Front Plant Sci* 13, 995589.

Mathesius, U., Schlaman, H. R. M., Spaink, H. P., Sautter, C., Rolfe, B. G. and Djordjevic, M. A. (1998). Auxin transport inhibition precedes root nodule formation in white clover roots and is regulated by flavonoids and derivatives of chitin oligosaccharides. *Plant Journal* 14, 23–34.

Mergaert, P., Uchiumi, T., Alunni, B., Evanno, G., Cheron, A., Catrice, O., Mausset, A. E., Barloy-Hubler, F., Galibert, F., Kondorosi, A., et al. (2006). Eukaryotic control on bacterial cell cycle and differentiation in the Rhizobium–legume symbiosis. *Proc Natl Acad Sci U S A* 103, 5230.

Miri, M., Janakirama, P., Huebert, T., Ross, L., McDowell, T., Orosz, K., Markmann, K. and Szczyglowski, K. (2019). Inside out: root cortex-localized LHK1 cytokinin receptor limits epidermal infection of Lotus japonicus roots by Mesorhizobium loti. *New Phytologist* 222, 1523–1537.

Mortier, V., Wasson, A., Jaworek, P., De Keyser, A., Decroos, M., Holsters, M., Tarkowski, P., Mathesius, U. and Goormachtig, S. (2014). Role of LONELY GUY genes in indeterminate nodulation on Medicago truncatula. *New Phytologist* 202, 582–593.

Motte, H., Vereecke, D., Geelen, D. and Werbrouck, S. (2014). The molecular path to in vitro shoot regeneration. *Biotechnol Adv* 32, 107–121.

Müller, S. and Kohlen, W. (2024). Jazzin' up nodules: The groovy role of jasmonic acid during nodulation. *Mol Plant* 17, 1639–1641.

Murray, J. D., Karas, B. J., Sato, S., Tabata, S., Amyot, L. and Szczyglowski, K. (2007). A cytokinin perception mutant colonized by Rhizobium in the absence of nodule organogenesis. *Science (1979)* 315, 101–104.

Murray, J. D., Muni, R. R. D., Torres-Jerez, I., Tang, Y., Allen, S., Andriankaja, 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.

Nadzieja, M., Kelly, S., Stougaard, J. and Reid, D. (2018). Epidermal auxin biosynthesis facilitates rhizobial infection in Lotus japonicus. *Plant Journal* 95, 101–111.

Naseem, M. and Dandekar, T. (2012). The Role of Auxin-Cytokinin Antagonism in Plant-Pathogen Interactions. *PLoS Pathog* 8, e1003026.

Ng, J. L. P. and Mathesius, U. (2018). Acropetal auxin transport inhibition is involved in indeterminate but not determinate nodule formation. *Front Plant Sci* 9, 1–13.

Ng, J. L. P., Perrine-Walker, F., Wasson, A. P. and Mathesius, U. (2015). The control of auxin transport in parasitic and symbiotic root–microbe interactions. *Plants* 4, 606–643.

Oldroyd, G. E. D. (2013). Speak, friend, and enter: signalling systems that promote beneficial symbiotic associations in plants. *Nature Reviews Microbiology 2013 11:4* 11, 252–263.

Oldroyd, G. E. D., Engstrom, E. M. and Long, S. R. (2001). Ethylene Inhibits the Nod Factor Signal Transduction Pathway of Medicago truncatula. *Plant Cell* 13, 1835–1850.

Oldroyd, G. E. D., Murray, J. D., Poole, P. S. and Downie, J. A. (2011). The rules of engagement in the legume-rhizobial symbiosis. *Annu Rev Genet* 45, 119–144.

Penmetsa, R. V. and Cook, D. R. (1997). A legume ethylene-insensitive mutant hyperinfected by its rhizobial symbiont. *Science (1979)* 275, 527–530.

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.

Penmetsa, R. V., Uribe, P., Anderson, J., Lichtenzveig, J., Gish, J. C., Nam, Y. W., Engstrom, E., Xu, K., Sckisel, G., Pereira, M., et al. (2008). The Medicago truncatula ortholog of Arabidopsis EIN2, sickle, is a negative regulator of symbiotic and pathogenic microbial associations. *Plant Journal* 55, 580–595.

Peoples, M. B., Herridge, D. F. and Ladha, J. K. (1995). Biological nitrogen fixation: An efficient source of nitrogen for sustainable agricultural production? *Plant Soil* 174, 3–28.

Peters, N. K., Frost, J. W. and Long, S. R. (1986). A plant flavone, luteolin, induces expression of Rhizobium meliloti nodulation genes. *Science* 233, 977–980.

Pimprikar, P., Carbonnel, S., Paries, M., Katzer, K., Klingl, V., Bohmer, M. J., Karl, L., Floss, D. S., Harrison, M. J., Parniske, M., et al. (2016). A CCaMK-CYCLOPS-DELLA Complex Activates Transcription of RAM1 to Regulate Arbuscule Branching. *Current Biology* 26, 987–998.

Plet, J., Wasson, A., Ariel, F., Le Signor, C., Baker, D., Mathesius, U., Crespi, M. and Frugier, F. (2011). MtCRE1-dependent cytokinin signaling integrates bacterial and plant cues to coordinate symbiotic nodule organogenesis in Medicago truncatula. *Plant Journal* 65, 622–633.

Prayitno, J., Rolfe, B. G. and Mathesius, U. (2006). The ethylene-insensitive sickle mutant of Medicago truncatula shows altered auxin transport regulation during nodulation. *Plant Physiol* 142, 168–180.

Pumplin, N., Mondo, S. J., Topp, S., Starker, C. G., Gantt, J. S. and Harrison, M. J. (2010). Medicago truncatula Vapyrin is a novel protein required for arbuscular mycorrhizal symbiosis. *Plant J* 61, 482–494.

Quides, K. W., Weisberg, A. J., Trinh, J., Salaheldine, F., Cardenas, P., Lee, H.-H., Jariwala, R., Chang, J. H. and Sachs, J. L. (2021). Experimental evolution can enhance benefits of rhizobia to novel legume hosts. 288, 1–10.

Redmond, J. W., Batley, M., Djordjevic, M. A., Innes, R. W., Kuempel, P. L. and Rolfe, B. G. (1986). Flavones induce expression of nodulation genes in Rhizobium. *Nature 1986 323:6089* 323, 632–635.

Reid, D., Heckmann, A. B., Novák, O., Kelly, S. and Stougaard, J. (2016). CYTOKININ OXIDASE/DEHYDROGENASE3 Maintains Cytokinin Homeostasis during Root and Nodule Development in Lotus japonicus. *Plant Physiol* 170, 1060–1074.

Reid, D., Nadzieja, M., Novák, O., Heckmann, A. B., Sandal, N. and Stougaard, J. (2017). Cytokinin Biosynthesis Promotes Cortical Cell Responses during Nodule Development. *Plant Physiol* 175, 361–375.

Reid, D., Liu, H., Kelly, S., Kawaharada, Y., Mun, T., Andersen, S. U., Desbrosses, G. and Stougaard, J. (2018). Dynamics of ethylene production in response to compatible nod factor. *Plant Physiol* 176, 1764–1772.

Rightmyer, A. P. and Long, S. R. (2011). Pseudonodule formation by wild-type and symbiotic mutant Medicago truncatula in response to auxin transport inhibitors. *Molecular Plant-Microbe Interactions* 24, 1372–1384.

Roth, L. E. and Stacey, G. (1989). Bacterium release into host cells of nitrogen-fixing soybean nodules: the symbiosome membrane comes from three sources. *Eur J Cell Biol* 49, 13–23.

Roy, S., Liu, W., Nandety, R. S., Crook, A., Mysore, K. S., Pislariu, C. I., Frugoli, J., Dickstein, R. and Udvardi, M. K. (2020). Celebrating 20 Years of Genetic Discoveries in Legume Nodulation and Symbiotic Nitrogen Fixation. *Plant Cell* 32, 15–41.

Sandhya Rani, Y., Jamuna, P., Triveni, U., Patro, T. S. S. K. and Anuradha, N. (2022). Effect of in situ incorporation of legume green manure crops on nutrient bioavailability, productivity and uptake of maize. *J Plant Nutr* 45, 1004–1016.

Schaller, G. E., Bishopp, A. and Kieber, J. J. (2015). The Yin-Yang of Hormones: Cytokinin and Auxin Interactions in Plant Development. *Plant Cell* 27, 44–63.

Schauser, L., Roussis, A., Stiller, J. and Stougaard, J. (1999). A plant regulator controlling development of symbiotic root nodules. *Nature* 402, 191–195.

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). NODULE INCEPTION Recruits the Lateral Root Developmental Program for Symbiotic Nodule Organogenesis in Medicago truncatula. *Current Biology* 29, 3657-3668.e5.

Schlesinger, W. H. and Bernhardt, E. S. (2020). The Atmosphere. *Biogeochemistry* 51.

Shen, L. and Feng, J. (2023). NIN—at the heart of Nitrogen-fixing Nodule symbiosis. *Front Plant Sci* 14, 1284720.

Singh, S., Katzer, K., Lambert, J., Cerri, M. and Parniske, M. (2014). CYCLOPS, A DNA-binding transcriptional activator, orchestrates symbiotic root nodule development. *Cell Host Microbe* 15, 139–152.

Smit, P., Raedts, J., Portyanko, V., Debellé, F., Gough, C., Bisseling, T. and Geurts, R. (2005). NSP1 of the GRAS Protein Family Is Essential for Rhizobial Nod Factor-Induced Transcription. *Science (1979)* 308, 1789–1791.

Smit, P., Limpens, E., Geurts, R., Fedorova, E., Dolgikh, E., Gough, C. and Bisseling, T. (2007). Medicago LYK3, an entry receptor in rhizobial nodulation factor signaling. *Plant Physiol* 145, 183–191.

Soyano, T., Kouchi, H., Hirota, A. and Hayashi, M. (2013). NODULE INCEPTION Directly Targets NF-Y Subunit Genes to Regulate Essential Processes of Root Nodule Development in Lotus japonicus. *PLoS Genet* 9, e1003352.

Soyano, T., Hirakawa, H., Sato, S., Hayashi, M. and Kawaguchi, M. (2014). NODULE INCEPTION creates a long-distance negative feedback loop involved in homeostatic regulation of nodule organ production. *Proc Natl Acad Sci U S A* 111, 14607–14612.

Sprent, J. I. and James, E. K. (2007). Legume Evolution: Where Do Nodules and Mycorrhizas Fit In? *Plant Physiol* 144, 575.

Stracke, S., Kistner, C., Yoshida, S., Mulder, L., Sato, S., Kaneko, T., Tabata, S., Sandal, N., Stougaard, J., Szczyglowski, K., et al. (2002). A plant receptor-like kinase required for both bacterial and fungal symbiosis. *Nature 2002 417:6892* 417, 959–962.

Suzaki, T., Yano, K., Ito, M., Umehara, Y., Suganuma, N. and Kawaguchi, M. (2012). Positive and negative regulation of cortical cell division during root nodule development in Lotus japonicus is accompanied by auxin response. *Development* 139, 3997–4006.

Timmers, A. C. J. (2008). The role of the plant cytoskeleton in the interaction between legumes and rhizobia. *J Microsc* 231, 247–256.

Tirichine, L., Imaizumi-Anraku, H., Yoshida, S., Murakami, Y., Madsen, L. H., Miwa, H., Nakagawa, T., Sandal, N., Albrektsen, A. S., Kawaguchi, M., et al. (2006). Deregulation of a Ca2+/calmodulin-dependent kinase leads to spontaneous nodule development. *Nature 2006 441:7097* 441, 1153–1156.

Tirichine, L., James, E. K., Sandal, N. and Stougaard, J. (2007). Spontaneous Root-Nodule Formation in the Model Legume Lotus japonicus: A Novel Class of Mutants Nodulates in the Absence of Rhizobia. *https://doi.org/10.1094/MPMI-19-0373* 19, 373–382.

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 ISOPENTENYLTRANSFERASE 3 function in Medicago nodule development. *Plant Physiol* 188, 560–575.

Van Spronsen, P. C., Bakhuizen, R., Van Brussel, A. A. N. and Kijne, J. W. (1994). Cell wall degradation during infection thread formation by the root nodule bacterium Rhizobium leguminosarum is a two-step process. *Eur J Cell Biol* 64, 88–94.

Van Velzen, R., Doyle, J. J. and Geurts, R. (2019). A Resurrected Scenario: Single Gain and Massive Loss of Nitrogen-Fixing Nodulation. *Trends Plant Sci* 24,.

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-chitooligosaccharide Signaling Triggers Accumulation of Cytokinins in Medicago truncatula Roots. *Mol Plant* 8, 1213–1226.

Velandia, K., Reid, J. B. and Foo, E. (2022). Right time, right place: The dynamic role of hormones in rhizobial infection and nodulation of legumes. *Plant Commun* 3,.

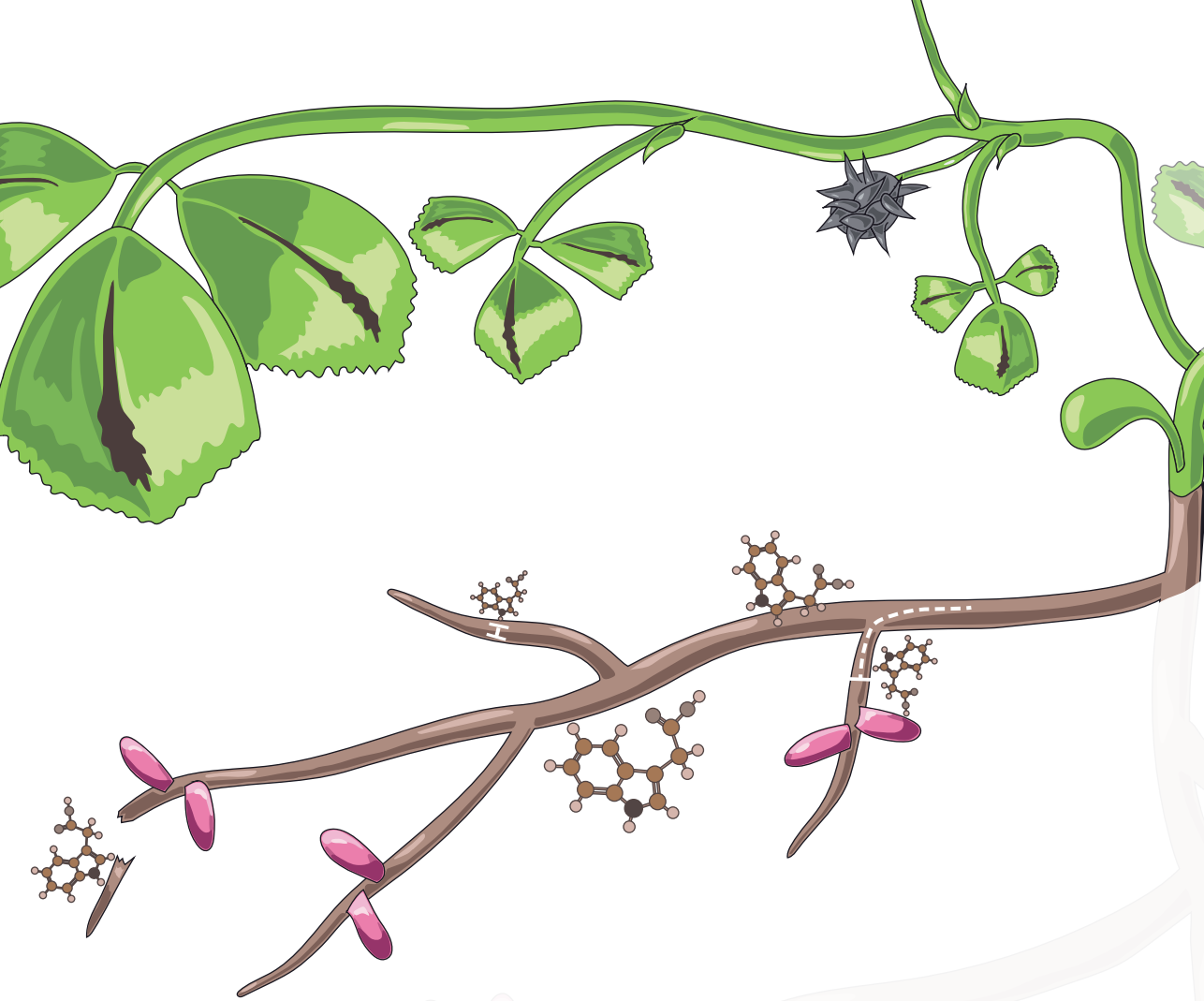
Vinardell, J. M., Fedorova, E., Cebolla, A., Kevei, Z., Horvath, G., Kelemen, Z., Tarayre, S., Roudier, F., Mergaert, P., Kondorosi, A., et al. (2003). Endoreduplication mediated by the anaphase-promoting complex activator CCS52A is required for symbiotic cell differentiation in Medicago truncatula nodules. *Plant Cell* 15, 2093–2105.

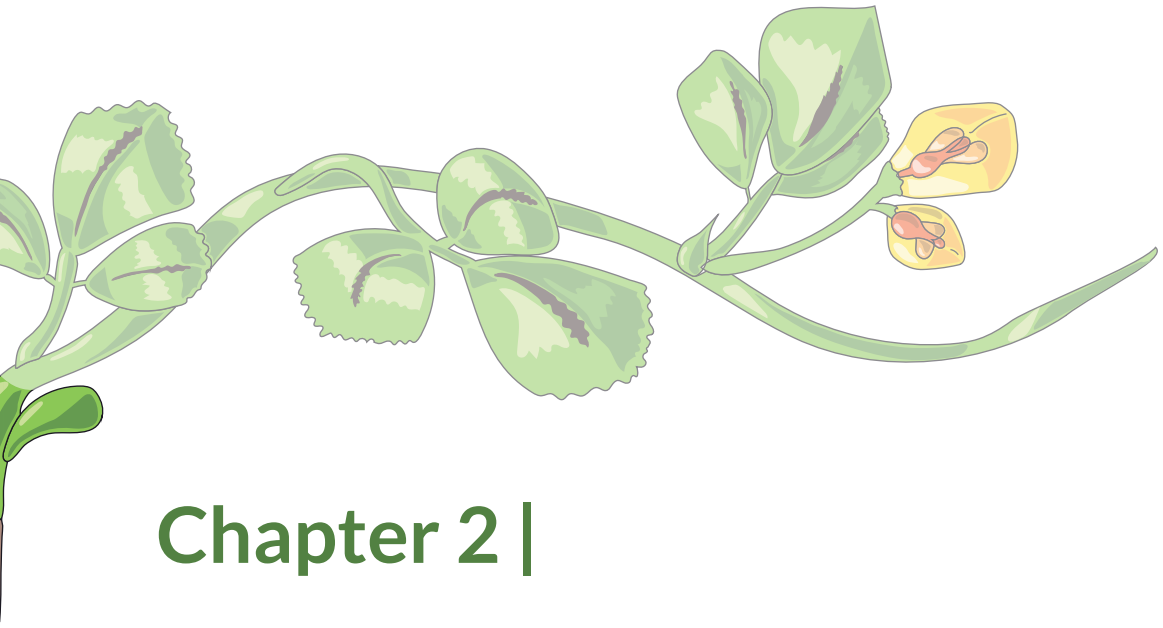
Xiao, T. T., Schilderink, S., Moling, S., Deinum, E. E., Kondorosi, E., Franssen, H., Kulikova, O., Niebel, A. and Bisseling, T. (2014). Fate map of Medicago truncatula root nodules. *Development (Cambridge)* 141, 3517–3528.

Yano, K., Yoshida, S., Müller, J., Singh, S., Banba, M., Vickers, K., Markmann, K., White, C., Schuller, B., Sato, S., et al. (2008). CYCLOPS, a mediator of symbiotic intracellular accommodation. *Proc Natl Acad Sci U S A* 105, 20540.

Yoro, E., Suzaki, T., Toyokura, K., Miyazawa, H., Fukaki, H. and Kawaguchi, M. (2014). A Positive regulator of nodule organogenesis, NODULE INCEPTION, acts as a negative regulator of rhizobial infection in Lotus japonicus. *Plant Physiol* 165, 747–758.

Zhang, X., Davidson, E. A., Mauzerall, D. L., Searchinger, T. D., Dumas, P. and Shen, Y. (2015). Managing nitrogen for sustainable development. *Nature 2015 528:7580* 528, 51–59.





Chapter 2 |

Nodule organogenesis in *Medicago truncatula* requires local stage-specific auxin biosynthesis and transport

Ting Ting Xiao†, Sophia Müller†, Shen Defeng†, Jieyu Liu, Kelvin Adema, Amber van Seters, Henk Franssen, Ton Bisseling, Olga Kulikova*, Wouter Kohlen*

† These authors contributed equally to this work

* corresponding authors

Published in Plant Physiology
https://doi.org/10.1093/plphys/kiaf133

ABSTRACT

The importance of auxin in plant organ development including root nodule formation is well established. Using auxin reporter constructs the spatiotemporal auxin distribution pattern during nodule development has previously been illustrated. However, our understanding of how this pattern is built-up and maintained still remains elusive.

To this end, we studied how the auxin gradient as inferred from DR5 expression patterns at different stages of nodule development in *Medicago truncatula* (Medicago), is correlated with the spatiotemporal expression patterns of known auxin biosynthesis and auxin transport genes. In addition, we record the *Mt*PIN10-GFP expression pattern and polar positioning on the cell plasma membranes during nodule primordium development to investigate the auxin flux. RNA interference and the application of auxin biosynthesis blockers were used to demonstrate the significance of biosynthesis and transport at the initial stages of the nodulation process.

Our results show that upon rhizobium inoculation, preceding the first mitotic activity, a specific set of *MtYUCs* and *MtPINs* as well as *MtLAX2* are expressed in the pericycle contributing to the creation of an auxin maximum. Overall, we demonstrate that dynamic spatiotemporal expression of both, *MtYUCs* and *MtPINs*, results in specific auxin outputs in subsequent stages of nodule primordia and nodule meristem formation.

INTRODUCTION

In legumes, rhizobium can establish a symbiotic relationship leading to the formation of nitrogen-fixing root nodules, involving two different developmental programs: infection thread initiation in the epidermis and nodule primordium formation in the inner root cell layers. Nodule formation is initiated by lipo-chitooligosaccharides, called Nod factors, secreted by rhizobia (reviewed by Dénarié *et al.*, 1996; and Geurts and Bisseling, 2002). In *Medicago truncatula* (Medicago), rhizobia enter the root via root hairs, where tubular structures, called infection threads, are formed (reviewed by Brewin, 2004; and Gage, 2004). Concomitantly, differentiated root cells, including pericycle and cortical cells are mitotically re-activated in an organized manner by which the Medicago nodule primordium is formed (reviewed by Oldroyd and Downie, 2006; Xiao *et al.*, 2014). The infection thread that contains rhizobia, grows towards the primordium and bacteria are released in cells derived from the inner cortex. A meristem which facilitates further growth is formed at the apex of the nodule primordium.

In general, two types of nodules can be distinguished based on the lifespan of the meristem. In determinate nodules (e.g. *Lotus japonicus*, Lotus) the meristem is active for a short time, while in indeterminate nodules (e.g. Medicago) a persistent meristem is formed (reviewed by Hirsch, 1992; and Sprent and James, 2007). These plants also differ in the location of the first cell divisions initiated upon rhizobial inoculation. Indeterminate nodules are initiated in the pericycle and the inner cortical cells and determinate nodules start in the outer and middle cortical cells (reviewed by Hirsch, 1992). In each nodule type the initiation of cell divisions is correlated with auxin accumulation (Mathesius *et al.*, 1998*b*; Pacios-Bras *et al.*, 2003; Wasson *et al.*, 2006; Mathesius, 2008; Zhang *et al.*, 2009; Ng *et al.*, 2015; Herrbach *et al.*, 2017; Ng and Mathesius, 2018).

Irrespective of the nodule type formed, auxin is involved in cell cycle control, vascular tissue differentiation and rhizobial infection during nodule development (reviewed by Kohlen *et al.*, 2018; and Lin *et al.*, 2020). Auxin is mainly synthetized in the aerial part of the plant, and can be transported towards the root apex both passively along the phloem and actively by auxin transporters (Swarup *et al.*, 2001; Vanneste and Friml, 2009; Swarup and Péret, 2012; Swarup and Bennett, 2014). Several studies have shown that inhibitors that block this acropetal auxin transport induce the formation of pseudo-nodules by creating a local auxin maximum (Allen *et al.*, 1953; Hirsch *et al.*, 1989; Scheres *et al.*, 1992; Schnabel and Frugoli, 2004; Subramanian *et al.*, 2006; Wasson *et al.*, 2006; Zhang *et al.*, 2009; Rightmyer and Long, 2011; Shen *et al.*, 2015; Sańko-Sawczenko *et al.*, 2016). An inhibition of auxin transport towards the root apex at the very start of nodule initiation was observed during the formation of indeterminate, but not determinate, nodules. On the contrary, at the initiation of determinate nodules an increase in acropetal auxin transport was observed (Allen *et al.*, 1953; Hirsch *et al.*, 1989; Rubery and Jacobs, 1990; Mathesius *et al.*, 1998*b*; Boot *et al.*, 1999; Pacios-Bras *et al.*, 2003; Prayitno *et al.*, 2006; Subramanian *et al.*, 2006). Polar auxin transport (PAT) is mediated by efflux carriers, among them the best studied are PIN-formed (PINs), and auxin influx carriers AUXIN

RESISTANT1 (AUX1) and LIKE-AUX1 (LAX) (Petrásek *et al*., 2006; Swarup and Bhosale, 2019). The formation of a PIN-dependent local auxin gradient is a central part of several developmental processes in plants (Benková *et al*., 2003; Wisniewska *et al*., 2006). In Medicago, 12 *PIN* genes were previously identified (Schnabel and Frugoli, 2004; Shen *et al*., 2015; Sańko-Sawczenko *et al*., 2016; Kohlen *et al*., 2018). In experiments where Nod-factors or *Sinorhizobium meliloti* was applied to Medicago (A17) it was shown that the earliest induced *PINs* in Medicago are *MtPIN2*, *MtPIN4* and *MtPIN10* (Plet *et al*., 2011; van Zeijl *et al*., 2015; Schiessl *et al*., 2019). The expression of these *MtPINs* during nodule initiation and development, and reported phenotypes of *MtPIN-RNAi* plants in Medicago indicate their importance in the creation of the auxin landscape during nodule development (Huo *et al*., 2006; Shen *et al*., 2015; Sańko-Sawczenko *et al*., 2016; Schiessl *et al*., 2019).

In addition to PINs, AUX1/LAX proteins function as auxin influx carriers. There are five members of the AUX1/LAX family in Medicago (Schnabel and Frugoli, 2004). The expression of *MtLAX2*, a homologue of *AtAUX1*, is highly induced during the initiation of nodule primordia (Roy *et al*., 2017; Schiessl *et al*., 2019). Moreover, two identified Medicago *Tnt1* insertion mutants effected in this gene, *Mtlax2-1 and Mtlax2-2*, developed fewer nodules and lateral roots, suggesting the same requirements for auxin influx activity for development of both lateral organs (de Billy *et al*., 2001; Schnabel and Frugoli, 2004; Swarup and Péret, 2012; Roy *et al*., 2017). In line with the putative function of the *Mt*LAX2 protein, both mutants displayed decreased *DR5::GUS* activity associated with rhizobial infection (Roy *et al*., 2017).

Auxin maxima are not only established as the result of polar auxin transport (PAT) but also as a consequence of local auxin biosynthesis (reviewed by Zhao, 2018; Cao *et al*., 2019). Although there are several pathways that lead to the production of auxin, the TAA/YUC pathway is reported to be the primary biosynthesis route for endogenous auxin (reviewed by Zhao, 2010, 2018). The TAA/YUC pathway is a two-step pathway in which TAA (TRYPTOPHAN AMINOTRANSFERASE) first converts tryptophane into indole-3-pyruvic acid (IPyA). Next, YUCCA, a flavin-containing monooxygenase, catalyses the rate-limiting step in biosynthesis of indole-3-acetic acid (IAA), the main natural auxin, from indole-3-pyruvic acid (IPyA) (Mashiguchi *et al*., 2011; Stepanova *et al*., 2011; Won *et al*., 2011). In Medicago, four *YUCCA* (*YUC*) genes were previously reported to be upregulated during nodule primordium formation: *MtYUC1*, *MtYUC2*, *MtYUC8* and *MtYUC9* and their expression is (in)directly regulated by the root nodule symbiosis specific transcription factor NIN (Liu *et al*., 2019; Schiessl *et al*., 2019).

Overall, during nodule initiation, it is likely that defined auxin concentrations are needed to induce and guide nodule organogenesis. However, how the auxin related gene network contributes to the establishment of such concentrations remains unknown. To unravel the complexities of these highly interconnected molecular physiological networks, we have employed a combination of techniques, including *DR5::GUS* based visualization of auxin output and the spatiotemporal expression profiling of genes associated with auxin biosynthesis and transport, known to be expressed during the formation of a nodule primordium. We observed

that the accumulation of auxin precedes cell divisions and, in subsequent stages of Medicago primordium development, remains closely associated with mitotic re-activation. Our investigations have revealed both, distinct and overlapping spatiotemporal expression patterns for *MtYUCs* and *MtPINs* genes. Additionally, we have pinpointed the localization of *Mt*PIN10 at the inception of nodule initiation and during subsequent developmental stages. Furthermore, through the utilization of *RNAi* and pharmacological interventions, we have validated that in addition to auxin transport, local auxin biosynthesis is an indispensable contributor to the initiation and development of root nodules in Medicago.

RESULTS

Defining the auxin response pattern during nodule initiation

To create a frame of reference for spatiotemporal analyses of genes involved in auxin patterning, we first determined the dynamics of the auxin gradient during nodule primordium initiation and development as previously described (Xiao *et al.*, 2014). For this, we created a stable transgenic Medicago R108 line containing the synthetic auxin-response promoter DR5 fused to β-glucuronidase (*DR5::GUS*).

Plants were grown on Fähraeus plates containing 1 µM Aminoethoxyvinylglycine (AVG) as previously described (Fähraeus, 1957; Xiao *et al.*, 2014). Roots of these plants were collected at different time points after spot-inoculation with *Sinorhizobium meliloti* 2011 (*Sm2011*), and incubated in GUS substrate buffer. Segments of stained roots were cut, fixed, embedded in plastic, and subsequently sectioned (**Figure 1**). The earliest *DR5* activity was observed in pericycle cells prior to periclinal divisions. As at this stage (12-24 hrs post inoculation, hpi) pericycle cell divisions were not yet visible we refer to this stage as stage 0 of nodule initiation. Stage 0 directly precedes the previously described stage I with detectable pericycle cell divisions (Xiao *et al.*, 2014). Stage 0 can be distinguished from lateral root initiation in the number of pericycle cells activated. Where lateral root initiation sites in Medicago are usually between 2-4 pericycle cells wide, at nodule stage 0 they can vary between 8-12 cells (**Figure 1, Supplemental Figure S1**).

Slightly later in stage I (~24 hpi), *DR5* activity was observed in dividing pericycle and endodermal cells, and not yet dividing cells of the inner root cortex (C4/5). At stage II (~30 hpi), *DR5* activity was maintained in these inner layers, and additionally had progressed to the outer cortex layers (C1-3) and epidermis. At stage III (~48 hpi) and IV (~72 hpi), *DR5* was highly active in C4/5 derived cells compared to pericycle and endodermis cells. The latter correlates with a decrease or completion of cell divisions in these cell layers at those stages.

During stage V (between 72 and 96 hpi), DR5 activity in the C4/5 derived cells in the central region of the nodule primordium markedly decreased and was no longer detectable in these cells during stage VI ($\geq$ 96 hpi). In contrast, *DR5* was highly active in C3 derived cells, an activity possibly linked to the establishment of the future nodule meristem from these cells.

In addition, in young nodules the infected cells directly adjacent to the nodule meristem, as well as cells of the vascular bundles, displayed relatively high *DR5* activity (**Supplemental Figure S2**). This pattern is reminiscent of the pattern of *DR5* activity in young Medicago A17 nodules (Demina *et al.*, 2019). Since the majority of experiments that will be described here, were conducted in Medicago Jemalong A17 (A17), we compared *DR5* expression patterns between R108 and A17 during early stages of nodule development. To this end, we introduced *DR5::GUS* to A17 by *Rhizobium rhizogenes* (formerly *Agrobacterium rhizogenes;* (Young *et al.*, 2001))-mediated hairy root transformation. The composite plants were grown in perlite for 7 days and

then inoculated with *Sm2011*. At 7 days post-inoculation (7 dpi) roots were collected and incubated in GUS substrate buffer. Segments of roots with stained nodule primordia at different stages were cut, fixed, embedded in plastic and sectioned. A comparison of R108 sections with those representing stages of A17 nodule formation (**Supplemental Figure S3**) showed that, despite R108 being stably transformed and A17 transiently transformed, the *DR5* patterns in both ecotypes were very similar.

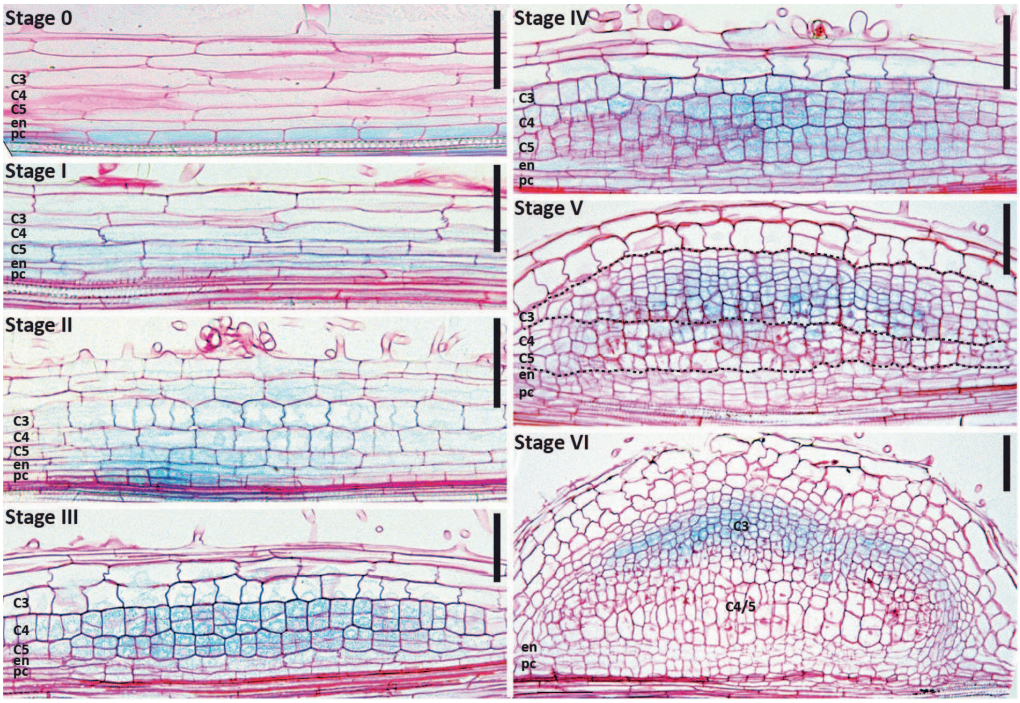


Figure 1 | DR5::GUS expression dynamics during the different stages of nodule primordium development in Medicago (cv R108). Longitudinal plastic sections (10mm) of spot inoculated root segments counter stained with ruthenium red. **Stage 0** (between 12-24 hpi), DR5 is activated in pericycle cells preceding their division. **Stage I** (~24 hpi), first anticlinal pericycle divisions have occurred, the DR5::GUS signal extends to the cortical cell layers prior to their activation. **Stage II** (~30 hpi), anticlinal divisions extend to cortex layers 3, 4 and 5, broad DR5 activity is detected in the entire nodule primordia. **Stage III** (~48 hpi), periclinal cell divisions in cortex layers 4 and 5, DR5 is highly expressed in actively dividing cortical cells of these layers. **Stage IV** (~72 hpi), periclinal cell divisions in cortex layer 3, DR5 signal extends to the third cortical cell layer. **Stage V** (between 72-96 hpi), cells derived from the C3 layer actively divide to form the future nodule meristem; DR5 activity is highest in these cells. Pericycle, endodermis and cortex layers 4 and 5 have stopped dividing and DR5 activity is no longer detected in the center of these cell layers. **Stage VI** (>96 hpi), vasculature bundles are formed, nodule meristem becomes active, DR5 activity is restricted to the nodule vasculature, nodule meristem and infected cells directly adjacent to it. For each time point 10- 15 spot inoculated root segments were sectioned. Most representative images for each stage were selected to be shown. C3-C5, cortical cell layers; ed, endodermis; pc, pericycle; scale bars 75 µm.

Expression patterns of *MtYUCs* during nodule initiation and development

DR5 activity results from auxin gradients established by a dynamic interplay between auxin biosynthesis and transport (reviewed by Zhao, 2018; Cao *et al.*, 2019). We first focused on the auxin biosynthesis genes previously reported to be expressed in the root susceptible zone and during nodule primordium formation in Medicago, *MtYUC1*, *2*, *8*, and *9* (van Zeijl *et al.*, 2015; Schiessl *et al.*, 2019). To dissect the role of these *MtYUC* genes in establishing the auxin patterning during nodule development as reflected by *DR5* activity, we first analysed their spatiotemporal expression patterns in the root tip and susceptible zone.

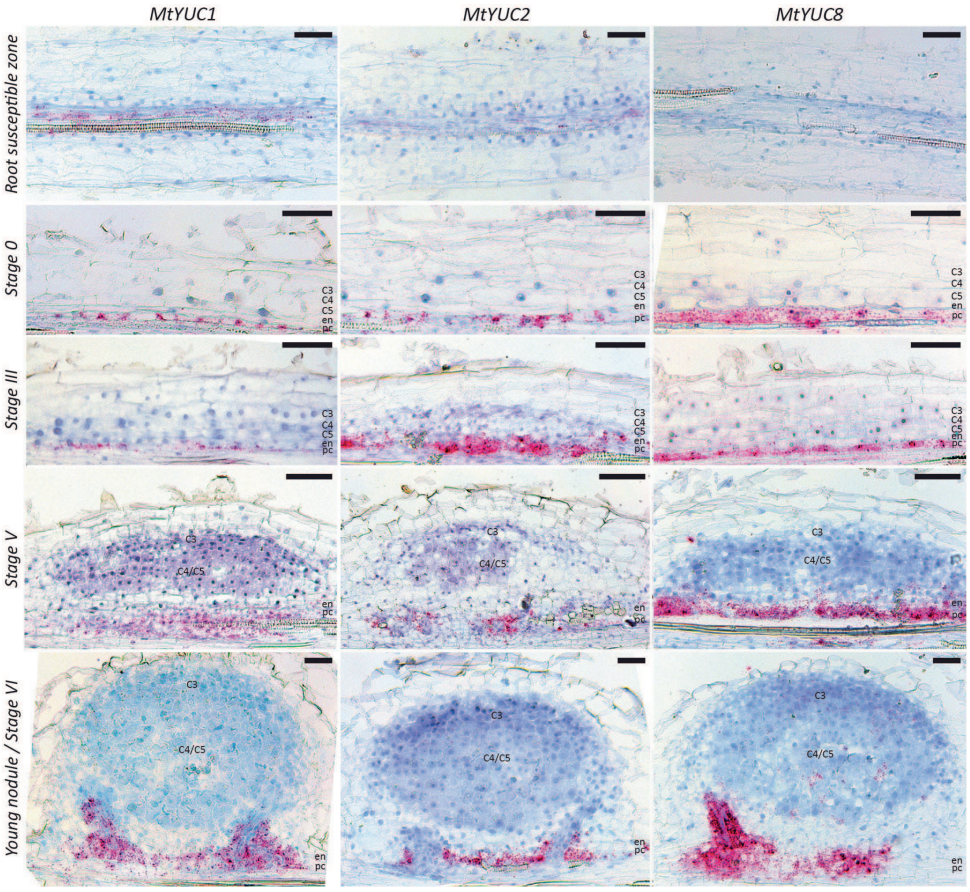


Figure 2 | The spatiotemporal expression dynamics of *YUCCAs* in root susceptible zone and during nodule primordium formation. Representative images of RNA *in situ* hybridizations with *MtYUC1*, *MtYUC2* or *MtYUC8* probe sets on longitudinal sections of root segments and nodule primordia at different stages of development. Red dots are hybridization signals. C3-C5, cortical cell layers; ed, endodermis; pc, pericycle; scale bars 75 µm.

RNA *in situ* hybridization demonstrated that only *MtYUC1* transcripts were clearly detectable in the vasculature of the root susceptible zone (**Figure 2**). All four *MtYUC* genes were expressed in the root elongation zone, and the expression of *MtYUC2* and *8* was extended slightly further to

the root differentiation zone (**Supplemental Figure S4**). In addition, *MtYUC8* and *9* transcripts were observed in the lateral root cap and columella cells of the root tip (**Supplemental Figure S4**).

To map the spatiotemporal expression dynamics of these *MtYUC*s onto the *DR5* pattern during nodule primordium organogenesis, we spot-inoculated Medicago roots and performed RNA *in situ* hybridization on longitudinal root sections accommodating nodule primordium developmental stages 0 - VI (**Figure 2**). For three out of four *MtYUCs*, *MtYUC1*, *2* and *8*, a strong hybridization signal was observed in the pericycle at stage 0. At the stages III to VI, *MtYUC1*, *2* and *8* transcripts remained restricted to the pericycle and the forming vasculature. In agreement with previously published data, we did not observe *MtYUC9* transcripts at the early stages of nodule primordia development (**Supplemental Figure S5**, Schiessl *et al.*, 2019). From stage IV onwards, *MtYUC9* transcripts were detectable at the periphery of the nodule primordium. Probably, these expression domains correlate to cells of the developing nodule vasculature. Across all analysed stages *MtYUC9* expression levels were relatively low (**Supplemental Figure S5**).

Functional analysis of *MtYUCs* during nodule initiation

Based on the spatiotemporal dynamics of *MtYUC* expressions during nodule initiation, we hypothesized that local auxin biosynthesis in the pericycle is required for nodule initiation. This raises the question whether inhibition or down-regulation of *MtYUCs* leads to a reduction in nodulation. To test this, we adopted pharmacological and *RNA* interference approaches (**Figure 3**).

It has previously been reported that 10 µM of 4-phenoxyphenylboronic acid (PPBo) is very efficient in blocking YUC enzymatic activity in Arabidopsis (Kakei *et al.*, 2015). To validate that 10 µM PPBo is able to block IAA biosynthesis in our experiments, we first set out to test its efficiency on Medicago roots. We sampled the lower 1.5 cm of uninoculated roots into 3x5 mm segments (**Figure 3A**) and determined the IAA concentration in these segments (**Figure 3B**). This revealed that IAA levels were approximately 2 times higher in the tip of the root compared to the other two segments (16.34, 7.18, 7.05 pmol•g^{-1} FW respectively). The application of 10 µM PPBo for 3 hrs reduced the level of IAA in the root tip by ~40%, while not effecting the level of IAA in the two other segments (9.57, 7.78, 6.68 pmol•g^{-1} FW respectively, **Figure 3B**). The fact that under these conditions *MtYUC* expression is mostly restricted to the root tip (**Figure 2, Supplemental Figure S4**) suggests that 10 µM PPBo was able to block local IAA biosynthesis in the root tip while not affecting the IAA derived from the acropetal auxin transport. Next, we treated roots for 24 hrs with 10 µM PPBo or mock, and subsequently spot-inoculated these roots with rhizobia in the root susceptible zone (5-10 mm). These regions and the 5 mm root segments above (10-15 mm) were harvested for IAA quantification at 24 hpi. In line with previous reports (Ng *et al.*, 2015; Schiessl *et al.*, 2019), *Sm2011* spot-application to the susceptible zone led to a ~60% increase in the level of IAA (**Figure 3C-D**). This increase was not

observed in the 5 mm segment above. Additionally, the spot-inoculated susceptible zone of roots pre-treated with PPBo for 24 hrs also did not show an increase in IAA levels (**Figure 3D**), suggesting that the observed increase in IAA in the susceptible zone is likely due to local YUC activity.

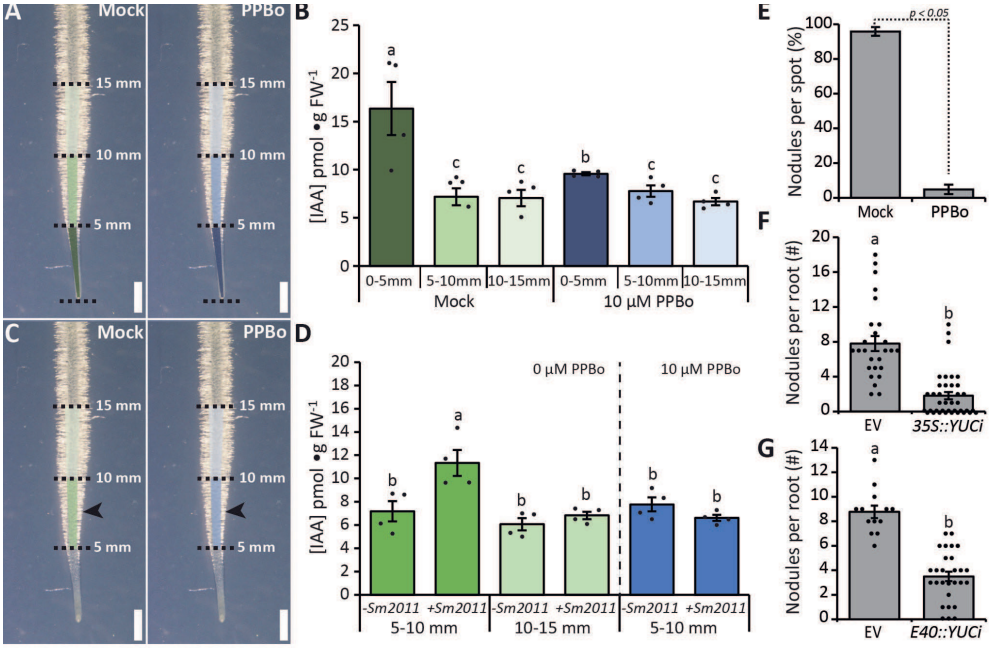


Figure 3 | The role of auxin biosynthesis during nodulation. (A, B) The measurement of auxin (IAA) in 3 zones of roots (root tip, susceptible zone, differentiation zone) after 3 hrs Mock or 10 µM PPBo application, bars represent averages, + standard error, dots show individual data points per sample, different letters indicate significant differences with $p < 0.05$, n=4. **(C, D)** The effect of 24 hrs 10 µM PPBo application on rhizobia induced specific auxin (IAA) accumulation in the root susceptible zone, bars represent averages, + standard error, dots show individual data points per sample, different letters indicate significant differences with $p < 0.05$, n=4. **(E)** The effect of pharmacological YUCCA inhibition by 10 µM PPBo on the nodulation efficiency of rhizobia spot-applications, bars represent combined averages of 4 independent experiments each containing >10 roots, $p < 0.05$, total n>40. **(F)** Number of nodules formed on *35S*::YUC1/2/8i (*35S*::YUCi) transgenic roots compared to Empty Vector (EV) control roots, bars represent averages, + standard error, dots show individual datapoints per transgenic root, different letters indicate significant differences with $p < 0.05$ according to ANOVA and Tukey post-hoc test (n>25). **(G)** Number of nodules formed on *ENOD40*::YUC1/2/8i (E40::YUCi) transgenic roots compared to EV control roots, bars represent averages, + standard error, dots show individual datapoints per transgenic root, different letters indicate significant differences with $p < 0.05$ according to ANOVA and Tukey post-hoc test, n>13.

To test whether addition of 10 µM PPBo prior to inoculation affects nodulation, we transferred 3-day-old Medicago seedlings to plates containing either 10 µM PPBo or mock. After 24 hrs, the root susceptible zones of these seedlings were spot-inoculated with *Sm2011*, and at 7 dpi, nodule numbers were scored. This revealed that, on average, nodules formed per spotted root segment were reduced by ~90% in plants growing on media containing 10 µM PPBo (**Figure 3E**).

Apart from blocking rhizobia-induced IAA accumulation in the susceptible zone, 10 µM PPBo also reduced auxin levels in the Medicago root tip (**Figure 3B**). Therefore, we questioned, is IAA biosynthesis at the root tip contributing to the auxin levels required for nodulation in the susceptible zone? For this reason, we re-sectioned the Medicago root tip (removal of the lower 3-5 mm) a few minutes prior to *Sm2011* spot application. Although a relatively harsh treatment, surprisingly it did not affect nodulation compared to the intact roots (**Supplemental Figure S6**).

To further investigate the relation between local auxin biosynthesis and nodulation, we created composite Medicago plants bearing transgenic roots in which *MtYUC*s expression was targeted by *RNAi*. It has been shown that in Arabidopsis, the 11 YUC genes have strong functional redundancies, and only higher-order *yucca* mutants display developmental defects (Cheng *et al.*, 2006, 2007). For this reason, we targeted all three early induced *MtYUCs* (i.e. *MtYUC1*, *2* and *8*) simultaneously (*YUCi*). We first used the *35S* promoter to drive this construct. The expression of *MtYUC1*, *MtYUC2* and *MtYUC8* was reduced by 50-60% in *35S::YUCi* transgenic roots compared to an empty vector (*EV*) control (**Supplemental Figure S7**). At 3 weeks post inoculation (wpi), nodulation in these transgenic roots was reduced by roughly 80% when compared to the *EV* control roots (**Figure 3F**). Next, we used the *ENOD40* promoter to drive the *YUCi* construct. The *ENOD40* promoter is specifically active in the pericycle prior to nodule initiation (Compaan *et al.*, 2001) and is further activated in the pericycle within 3 hrs of Nod factor application (Xiao *et al.*, 2014; van Zeijl *et al.*, 2015). At 3 wpi nodulation on *ENOD40:YUCi* transgenic roots was reduced by roughly 65% compared to the *EV* control roots (**Figure 3G**).

So far, we have shown that *MtYUC1*, *2* and *8* are all expressed in the pericycle of the root susceptible zone in response to rhizobia application. Additionally, we demonstrated that application of PPBo blocks the local auxin biosynthesis. As the first step in this process is the induction of pericycle divisions followed by cortical cell divisions, we questioned whether local auxin biosynthesis is critical in either of these cell division processes. Although *MtYUC1*, *2* and *8* are not expressed in the cortex, we first wondered if interfering with their function in auxin biosynthesis had any effect on the mitotic reactivation of these cortical cells. For this, we pre-treated Medicago roots for 24 hrs with either 0 (Mock) or 10 µM PPBo (PPBo) prior to spot inoculation with *Sm2011*. Root segments were collected at 72 and 96 hpi, embedded in plastic and sectioned. This revealed that spot application induced cortical cell divisions in almost all roots with the Mock pre-treatment, but rarely in plants pre-treated with 10 µM PPBo (**Figure 4A-D**). Next, if the root cortex is not mitotically reactivated when pre-treated with PPBo; what about the pericycle? As the effect of rhizobia spot application on the pericycle cells already occurs roughly at 24 hpi we repeated the spot inoculation experiment as described above and collected inoculated root segments at 24 hpi. To facilitate the identification of early initiation events we used a stable A17 *ENOD11::GUS* line (Journet *et al.*, 2001). GUS signal was readily found in the susceptible zone of Mock pre-treated plants spotted with *Sm2011*, but not in any of the other treatments (**Figure 4E-I**). To visualize the effect of PPBo application on the mitotic reactivation of the pericycle we measured the length of the pericycle cells in the root susceptible zone at the application site. This revealed the pericycle cells pre-treated with Mock and spotted

with *Sm2011* were roughly 50% smaller than those not inoculated (**Figure 4E-F, G**). This suggests that these cells had undergone a round of cell divisions prior to harvesting. This was not the case in plants pre-treated with 10 µM PPBo. In this treatment, *Sm2011* application did not affect pericycle cell length, suggesting that these cells were not dividing (**Figure 4E, H-I**).

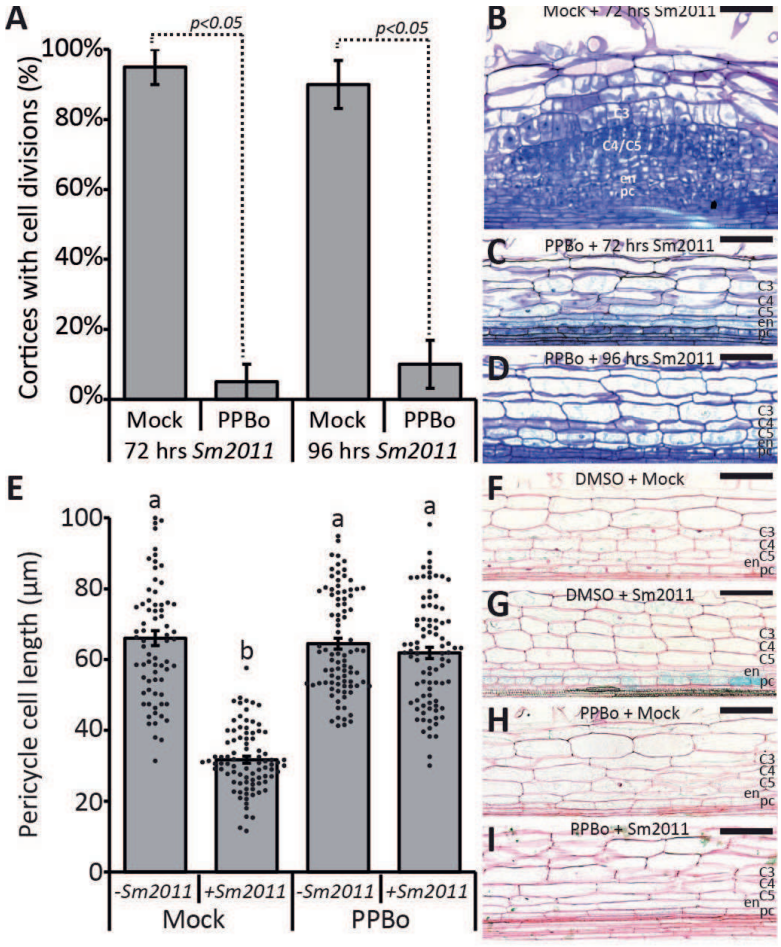


Figure 4 | The effect of pharmacological inhibition of *Mt*YUCCA on nodule initiation. **(A)** Percentage of spotted roots showing mitotic reactivation of cortical cells in the 0 µM (Mock) or 10 µM PPBo at 72 or 96 hpi. Bars represent averages, ± standard error, n=20. **(B-D)** Representative images of the Medicago roots after spot inoculation **(B)**, a Mock treated + *Sm2011* for 72 hrs **(C)** a 10 µM PPBo treated + *Sm2011* for 72 hrs, and **(D)** a 10 µM PPBo treated *Sm2011* for 96 hrs. **(E)** Average length of pericycle cells in the root susceptible zone at the application site without (*-Sm2011*) or with (*+Sm2011*) *Sm2011* spot application in the absence (Mock) or presence (PPBo) of 10 µM PPBo. Bars represent averages, ± standard error, dots show the length of individual pericycle cells, different letters indicate significant differences with $p < 0.05$, n>100. **(F-I)** Representative images of the Medicago root susceptible zone at 24 hpi; the blue staining marks the nodule primordium initiation in *pENOD11::GUS* stable transgenic A17 line. **(F)** a 24 hrs Mock treated - *Sm2011*, **(G)** a 24 hrs Mock treated *+Sm2011*, **(H)** a 24 hrs 10 µM PPBo treated *-Sm2011* , **(I)** a 24 hrs 10 µM PPBo treated + *Sm2011*. C3-C5, cortical cell layers; ed, endodermis; pc, pericycle; scale bars 100 µm.

Spatiotemporal expression of auxin transporters during nodule primordium development

The observation that the auxin biosynthesis genes *MtYUC1*, *2*, *8* and *9* are not expressed in cortical cells during primordium formation raises the question; "how do these cells acquire the

auxin as inferred from *DR5::GUS*?". As no other *MtYUCs* were detected during nodule primordium development, an obvious answer to this question is polar auxin transport. Polar auxin transport requires the presence of auxin export carrier *Mt*PIN proteins, positioned on the plasma membrane outwards to direct auxin flow from the pericycle, via the endodermis towards the cortex. To investigate this, we first determined the spatial expression pattern of the early induced *MtPIN2*, *MtPIN4* and *MtPIN10* (Plet *et al*., 2011; van Zeijl *et al*., 2015; Kohlen *et al*., 2018; Schiessl *et al*., 2019).

In the root susceptible zone, hybridization signals were only observed for *MtPIN4* and *MtPIN10*, in the stele (**Figure 5**), which agrees with the reported expression levels of these genes (van Zeijl *et al*., 2015). *MtPIN2* transcripts were not detected in the root susceptible zone. To exclude inefficient *MtPIN2* probe set affinity, we used for RNA *in situ* hybridization longitudinal sections of root tips as a control. In the root tip *MtPIN2*, *4* and *10* displayed specific spatial expression patterns (**Supplemental Figure S8**), to some extent matching the expression domains of their Arabidopsis orthologs. This demonstrated that *MtPIN2* probe set is capable of detecting *MtPIN2* transcripts, and that this gene is likely not expressed in the root susceptible zone.

Next, we studied the spatiotemporal expression pattern of *MtPIN2*, *MtPIN4* and *MtPIN10* at different developmental stages of nodule primordium development by RNA *in situ* hybridization (**Figure 5**). At stages 0-I, *MtPIN2* was not visible in the pericycle. *MtPIN4* was detected in the vasculature and pericycle, and this expression pattern was not different from that in the susceptible zone of un-inoculated plants (**Figure 5**). *MtPIN10* transcripts were observed in root vasculature and with highest intensity in the pericycle, suggesting that *MtPIN10* expression was induced in the pericycle at the start of nodule primordium initiation.

At stages III-IV, *MtPIN2* transcripts were seen in both dividing and non-dividing cortical cells, as well as in the developing nodule vasculature. At stage V-VI, *MtPIN2* is expressed mainly in the nodule vasculature and the future nodule meristem. Different from *MtPIN2*, *MtPIN4* remained restricted to the pericycle at stage I-III and to the forming vasculature at stages IV-VI. *MtPIN10* transcripts were observed also in cortex derived cells at stage III and at stage IV at the periphery of the nodule primordium. At later stages V and VI, *MtPIN4* and *MtPIN10* hybridization signals were visualized mainly in the developing vasculature. Like *MtPIN2*, *MtPIN10* transcripts were also detectable in the future nodule meristem at stage VI.

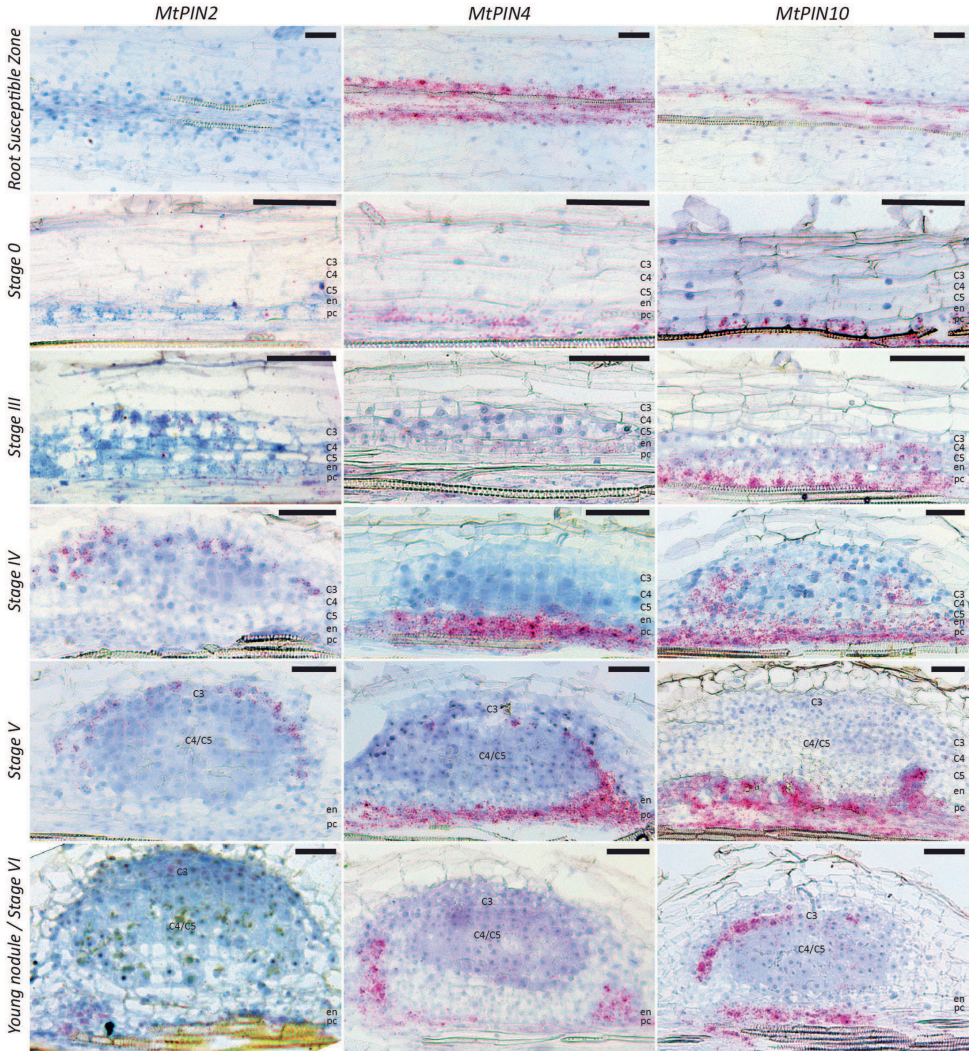


Figure 5 | The spatiotemporal expression dynamics of MtPINs in the root susceptible zone and during nodule primordium formation. Representative images of RNA in situ hybridizations with *MtPIN2*, *MtPIN4* or *MtPIN10* probe sets on longitudinal sections of root segments and nodule primordia at different stages of development. Red dots are hybridization signals. In addition, for the *MtPIN2* transcript detection at stage 0 and III 2-plex RNA *in situ* was used that included a *MtNF-YA1* probe set Type 6 as a marker for nodule primordium initiation (blue hybridization signals). C3-C5, cortical cell layers; ed, endodermis; pc, pericycle; scale bars 75 µm.

MtPIN6 is orthologous to the Arabidopsis *PIN6* and has a unique position within the family of PIN auxin transporters as it displays a dual localization, at the plasma membrane and endoplasmic reticulum (Simon *et al*., 2016). As a result, *At*PIN6 is believed to mediate both, polar auxin transport and intracellular auxin homeostasis. Additionally, *At*PIN6 has been demonstrated to be involved in organogenesis (Cazzonelli *et al*., 2013; Simon *et al*., 2016). By using RNA *in situ* hybridization, we could not detect *MtPIN6* in the root tip, root susceptible zone or at early nodule primordium stages 0-I. However, *MtPIN6* transcripts were visualized in root transition zone and at later time points of nodule development, when nodule vasculature was initiated (**Supplemental Figure S9**).

Apart from auxin efflux, auxin influx is a contributing factor to establish directional auxin transport (Bennett *et al*., 1996). The auxin import carrier *MtLAX2* is highly induced during the early stages of nodule development (de Billy *et al*., 2001; Roy *et al*., 2017; Schiessl *et al*., 2019). We performed *in situ* RNA hybridisation on the Medicago root, including the root tip and susceptible zone. Our results confirmed previously published expression pattern of *MtLAX2* that was detected in the lateral root cap, in some of columella cells and developing vasculature (de Billy *et al*., 2001) (**Supplemental Figure S10**). In addition, strong *MtLAX2* hybridization signals were detected in the vasculature of the susceptible zone. To resolve *MtLAX2* expression dynamics during nodule primordium formation, we performed RNA *in situ* hybridization on longitudinal sections of spot-inoculated root segments. This revealed that, at stage 0, *MtLAX2* transcripts were located in the root vasculature and pericycle (**Supplemental Figure S10**). At stage I, II, and III, *MtLAX2* expression was detected mainly in the pericycle and forming nodule vasculature. At later stages, expression was observed in the future nodule meristem, albeit at very low levels (**Supplemental Figure S10**).

Taken together, our spatiotemporal studies on auxin transporter gene expression revealed that *MtLAX2*, *MtPIN4* and *MtPIN10* could be involved in local auxin accumulation at stage 0 and I in the pericycle, while *MtPIN2*, *MtPIN6* and *MtPIN10* could participate in creating the auxin pattern in cortical cells.

MtPIN10 dynamics during nodule formation

If auxin biosynthesis at stage I and II of nodule primordium initiation does not occur locally in the cortex and both *MtPIN4* and *MtPIN10* are expressed in the pericycle, it implies that auxin might be produced in the pericycle and actively transported towards the cortex. To determine if cellular localization of the expressed MtPIN proteins could facilitate such transport, we created a *MtPIN10*::MtPIN10-eGFP (MtPIN10-GFP) protein fusion. We attempted to create a similar construct for *Mt*PIN4, but unfortunately without success.

First, we tested whether this construct would mimic the expected expression domains. Compound plants baring *Mt*PIN10-GFP transgenic roots showed a similar expression pattern in the root tip (**Supplemental Figure S11**) as observed by RNA *in situ* hybridisation for *MtPIN10*

(**Supplemental Figure S8**). However, no GFP signal was detected in the susceptible zone for MtPIN10 where this gene is expressed (**Supplemental Figure S11**).

Next, we spot-inoculated roots transiently expressing MtPIN10-GFP with Sm2011. Root segments were harvested at different time points after inoculation, longitudinal hand-made sections through the vasculature were made, and immediately observed under confocal microscope. To determine subcellular MtPIN10 localization, we analysed ~5-7 primordia for each stage and representative images are shown (**Figure 6**).

At stage 0, MtPIN10 was clearly visible in the pericycle at the site of nodule primordium induction, and, although not exclusively, in part localized towards the endodermis and root cortex (arrowheads). At stage I, MtPIN10 is extended to the endodermis (arrowheads). At stage II, MtPIN10 was visible in the endodermis and inner cortex. At stage III, MtPIN10 was localized in the dividing cortical cells, and later, at stage IV, the primordium periphery positioned towards the centre of the activated cortical cells. At stages V and VI, MtPIN10 localized in the periphery of the nodule primordia, likely the nodule vasculature. Here, subcellular localization of MtPIN10 was orientated towards the nodule meristem. At these stages, MtPIN10 is hardly detectable in the central part of the primordia.

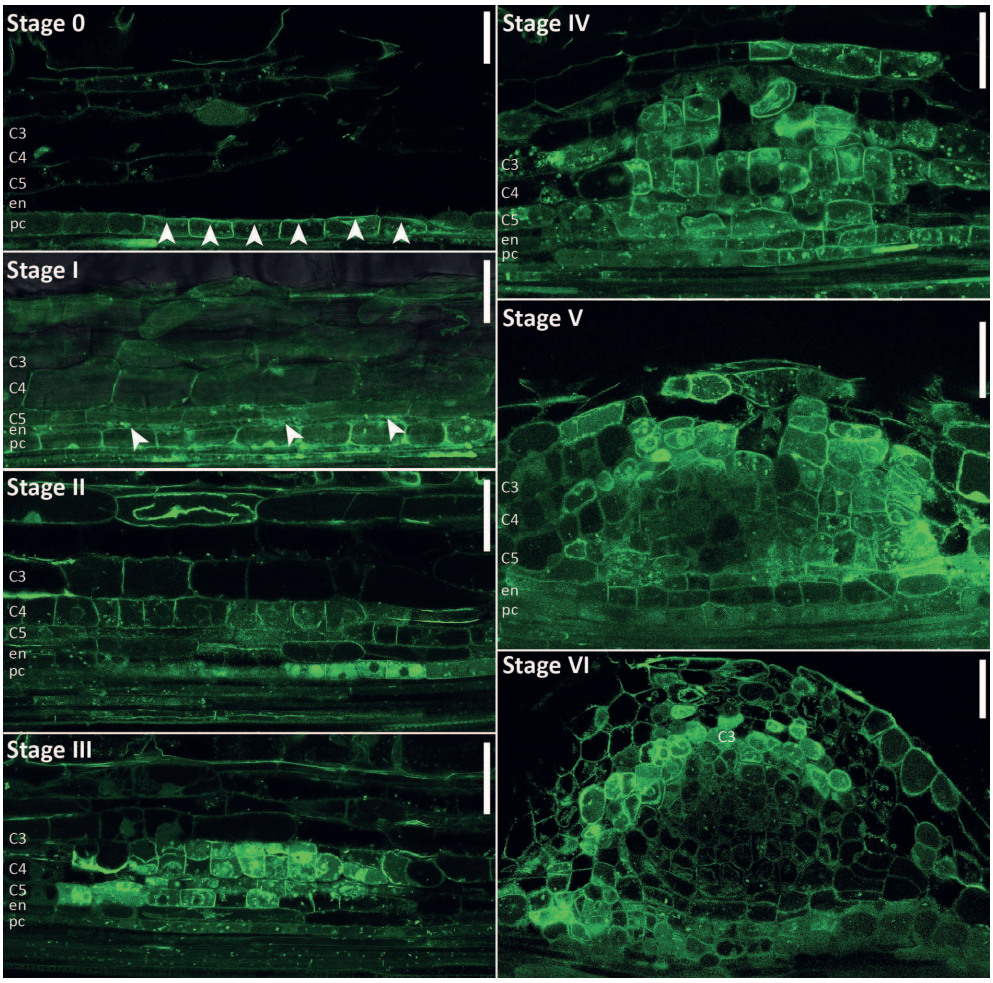


Figure 6 | MtPIN10 localization during Medicago nodule primordium formation. Stage 0, MtPIN10-GFP is detected in pericycle cells, partially orientated towards the cortex (arrowheads). **Stage I**, MtPIN10-GFP is extended to the endodermis (arrowheads). **Stage II**, MtPIN10 levels are increased in cortical cells with no apparent polarity. **Stage III**, MtPIN10 is present in pericycle only at low levels. High MtPIN10 levels are detected in dividing cortical cells with no clear polarity. MtPIN10 is also detected in the root vasculature positioned on the plasma membranes towards the root tip. **Stage IV**, MtPIN10 is positioned towards the center of the primordia in the cells located at the primordium periphery and no clear polarity is visible in the cells located at the center of the primordium. **Stage V/ VI**, MtPIN10 is hardly detectable in the central part of the primordia, at the periphery, in developing nodule vasculature cells, it is positioned towards the (future) nodule meristem. C3-C5, cortical cell layers; ed, endodermis; pc, pericycle; scale bars, 75 mm.

Functional analysis of *MtPINs* during nodulation

A reported Medicago *pin2* mutant did not show a nodulation-related phenotype (Ng *et al.*, 2020). These observations imply that there is a certain functional redundancy of *MtPINs*. Such redundancy is known in Arabidopsis roots where PIN proteins exhibit synergistic interactions when the loss of a specific PIN protein is compensated by auxin dependent ectopic expression of its homologs (Vieten et al., 2005). Additionally, two mutant lines have been isolated and described for Medicago *PIN10*, loss-of-function *pin10/slm1* (*SMOOTH LEAF MARGIN1*) (Peng and Chen, 2011; Zhou *et al.*, 2011*a*,*b*). These mutant lines display severe pleiotropic phenotypes in different organs including leaf and flower development. In addition, both lines are sterile. For this reason, these mutants must be maintained as heterozygous and for neither mutant a nodule phenotype has been described.

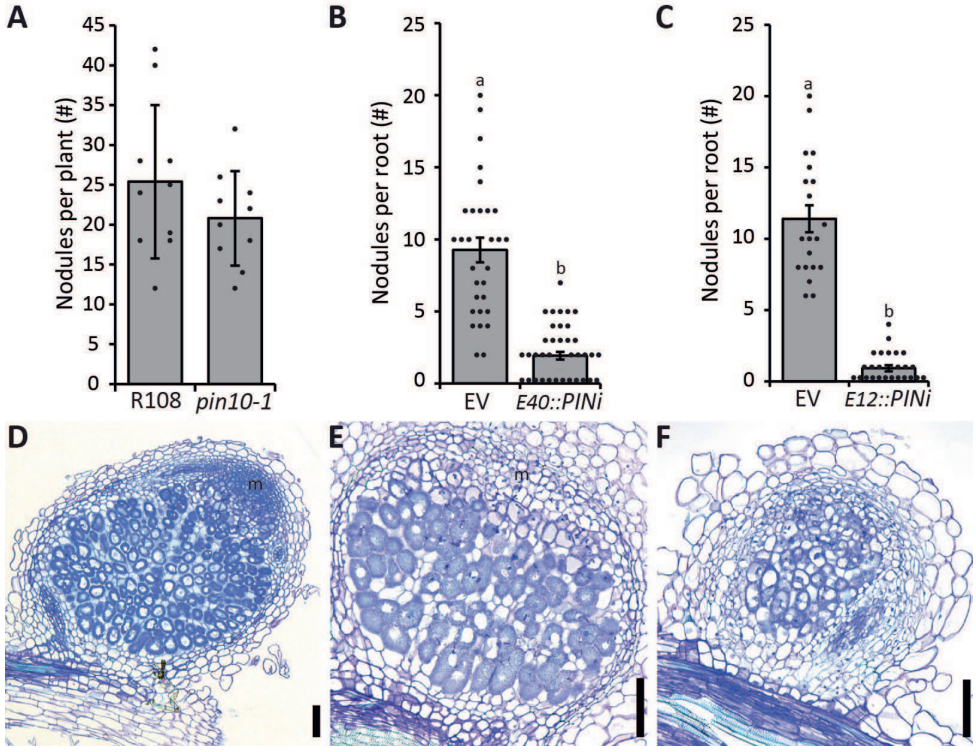


Figure 7 | The role of auxin transport during nodulation in Medicago. (A) The number of nodules formed on roots of *pin10-1* compared to wild type (R108). Bars represent averages, ± standard error, dots show nodules per individual plant (n>13). **(B)** The number of nodules formed on *ENOD40::PIN2/4/10i (E40::PINi)* transgenic roots compared to *EV* control roots, bars represent averages, ± standard error, dots show individual data points per transgenic root, different letters indicate significant differences with $p < 0.05$ according to ANOVA and Tukey post-hoc test, n>30. **(C)** The number of nodules formed on *ENOD12::PIN2/4/10i (E12::PINi)* transgenic roots compared to *EV* control roots, bars represent averages, ± standard error, dots show individual data points per transgenic root, different letters indicate significant differences with $p < 0.05$ according to ANOVA and Tukey post-hoc test, n>20. **(D-F)** Representative images of **(D)** an *EV* control nodule containing a fully developed nodule meristem, **(E)** *an E12:PINi* nodule with a short and underdeveloped meristem, and **(F)** an *E12:PINi* nodule without a meristem. m, meristem; scale bars, 100 µm.

We inoculated Medicago *pin10-1* mutant plants with rhizobia and analysed nodules at 3 wpi. The *pin10-1* mutation had no significant effect on the number of nodules formed per plant compared to wild-type R108 (**Figure 7A**). This observation suggests that *MtPIN10* either plays no role during nodulation, or its involvement is masked by a functional redundancy. In fact, *MtPIN4* has a similar expression pattern as *MtPIN10* (**Figure 5**), and therefore might function redundantly to *MtPIN10*. To test this, we generated *RNAi* construct to knock-down *MtPIN2*, *MtPIN4* and *MtPIN10* simultaneously.

We first used the *35S* promoter to drive the *PINi* construct, reducing expression of all three targeted *PINs* by 70-90% in transgenic roots (**Supplemental Figure S12A**). This led to a severe root phenotype where roughly 70% of all roots formed were less than 3 cm long (**Supplemental Table S1**). Closer inspection revealed that several of these roots failed to maintain a functional apical root meristem (**Supplemental Figure S12B-D**). Overall, nodulation on roots transformed with the *35S::PINi* construct was reduced by roughly 75% compared to an *EV* control (**Supplemental Figure S12E**). Additionally, roots expressing the *35S::PINi* construct formed several small nodules with a short and underdeveloped meristem (**Supplemental Figure S12F**) and nodule like bumps (**Supplemental Figure S12G**).

Table 1: E12::PINi nodule phenotype

***RNAi* composite plants**	**Nodules sectioned**	**Normal**	**Small or no meristem**
EV	25	22	3
E12:PINi	22	7	15

Next, we used the *ENOD40* promoter to silence these *MtPIN* genes at the very beginning of nodule initiation (Compaan *et al.*, 2001). The introduction of this construct *in planta* by hairy root transformation reduced nodule number by ~70% compared to the *EV* control (**Figure 7B**). This demonstrates that expression of *MtPIN4* and *MtPIN10* in the pericycle is required for nodule initiation. However, as the *ENOD40* promoter is also active in non-inoculated roots (Crespi *et al.*, 1994) and *MtPIN4*, *MtPIN10*, *and MtPIN2* are expressed in the pericycle of the Medicago root tip independent of rhizobia application (**Figure 5**), we cannot exclude that down-regulation of the expression of these *MtPINs* at this position in the root has an effect on nodulation. Therefore, we silenced *MtPIN4*, *MtPIN10* and *MtPIN2* under the symbiosis specific promoter *ENOD12* (Li *et al.*, 2021). Introduction of this construct led to a similar reduction in nodule numbers compared to the *ENOD40* promoter (**Figure 7B-C**). Moreover, the few nodules formed were small compared to wild type, and plastic sections revealed that they rarely (7 out of 22 vs 22 out of 25 for the EV control) formed a meristem (**Figure 7D-F, Table 1**). Combined, this shows that downregulation of PINs under control of the *ENOD40* and *ENOD12* promoters have a similar effect on nodule numbers, but no obvious effect on root development. This suggests that downregulation of these *PINs* under the *ENOD40* prior to rhizobium application might not have a strong additional

effect on nodule numbers and that local auxin transport, similar to local auxin biosynthesis, could be essential for nodulation.

No effect of blocking shoot derived auxin on nodulation

It is a well-studied effect that auxin produced in the shoot contributes to the organisation of root morphology (McDAVID *et al.*, 1972; Bhalerao *et al.*, 2002; Guo *et al.*, 2005; Overvoorde *et al.*, 2010). Additionally, it was previously demonstrated that polar auxin transport plays a role during nodulation (Mathesius *et al.*, 1998*b*; Wasson *et al.*, 2006; Deinum *et al.*, 2016). Still, whether auxin produced in the shoot contributes to the formation of root nodules remains an open question. To test if shoot derived auxin is needed during nodulation we blocked polar auxin transport at the hypocotyl-root junction using small agar droplets containing either 0 (Mock) or 50 µM N-1-naphthylphthalamic acid (NPA, **Figure 8A**). To validate that NPA application at the hypocotyl-root junction was affecting the plant roots, we scored additional auxin related root phenotypes previously linked to shoot derived auxin such as primary root growth and lateral root initiation (Reed *et al.*, 1998; Bhalerao *et al.*, 2002; Overvoorde *et al.*, 2010; Roychoudhry and Kepinski, 2022). This revealed that primary root growth was mildly, but significantly, reduced (~20%, **Figure 8B**) and lateral root development was blocked almost completely (**Figure 8C**) when NPA was applied at the hypocotyl-root junction compared to the mock application. Nevertheless, NPA application at the hypocotyl-root junction had no significant effect on nodulation (**Figure 8D**).

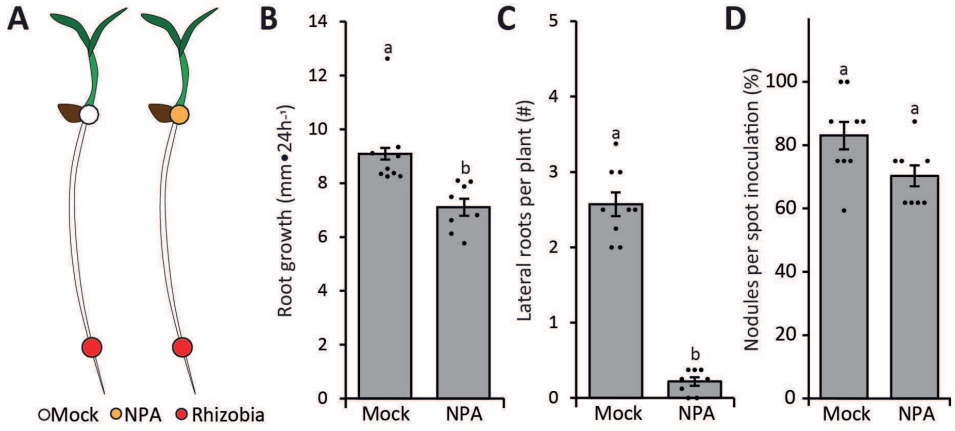


Figure 8 | The effect of N-1-naphthylphthal-amic acid (NPA) application at the hypocotyl-root junction on root development and nodulation. (A) Schematic represen-tation of the experimental design including location of application of either 0 µM NPA (Mock) or 50 µM NPA 24 hrs prior to Sm2011 spot application. **(B)** Averaged root growth (mm•24h-1) at 8 dpi. **(C)** The average number of later roots per plant. **(D)** Percentage of spot-inoculated roots that formed a nodule on the inoculation site. Bars represent averages, + standard error, dots show average datapoints per plate, different letters indicate significant differences with $p < 0.05$ according to ANOVA and Tukey post-hoc test (n>20).

DISCUSSION

Here, we presented dynamics of genes and proteins involved in auxin biosynthesis and transport and propose their role in establishing the *DR5* expression patterning during sequential stages of nodule formation. From this, the following picture emerges with respect to how the auxin landscape links to nodulation in Medicago; *DR5* activity is concentrated in pericycle cells preceding their first division, and in a wave like motion, only a few cells wide, the *DR5* activity moves through the endodermal and cortical cells that give rise to the nodule primordium, vasculature and ultimately the nodule meristem. Based on this *DR5* patterning, four major events during nodule development in which auxin likely plays an important role can be defined as: pericycle activation, cortex activation, vascularisation, and meristem establishment.

Local auxin accumulation in the pericycle is a prerequisite for nodule initiation.

The models presented by Deinum and colleagues (2012, 2016) gave priority to a local reduction of auxin efflux (PINs) to explain the auxin maximum preceding nodule formation (Xiao, 2015). It has been suggested that this reduction in transport capacity leads to a local increase of auxin levels in the pericycle and inner cortex at the site of nodule primordium initiation (Mathesius *et al.*, 1998*b*,*a*; Ng and Mathesius, 2018). However, a direct comparison between auxin transport inhibition and rhizobia application is difficult as the auxin response patterns and timing of the initiation between these two treatments are different (Ng and Mathesius, 2018).

We demonstrate that a local auxin transport network of *MtPIN4/10* and *MtLAX2* are expressed in the stele of the root susceptible zone, and that *MtLAX2* is highly expressed in the pericycle at the start of nodule initiation. It is possible that together these auxin transporters create a conduit for an auxin flow from the stele into the pericycle. However, our data show that during the first stages of nodule primordium development (stage 0-I) the expression of the auxin biosynthesis genes *MtYUC1*, *2* and *8* correlates with the mitotic re-activation of the few cells in the pericycle that initiate a nodule primordium. Pericycle specific *RNAi* against these *MtYUCs (ENOD40::YUCi)* and pharmacological applications with the YUCCA inhibitor PPBo demonstrates that local activity of *MtYUC1*, *2 and 8*, and therefore likely local auxin biosynthesis, in the pericycle is crucial for nodule initiation and further progression. It was previously reported that in Soybean (*Glycine max*) the expression of *GmYUC2a* correlates with determinate nodule development (Wang *et al.*, 2019). Interestingly, the fact that ectopic expression of *GmYUC2* under control of the *35S* promoter severely reduced nodule numbers, similar to the *RNAi* and pharmacological approach used in our study, suggests that during nodulation the auxin levels likely need to be precisely balanced to form root nodules.

Our findings that the addition of PPBo, while blocking local auxin biosynthesis and nodule initiation, has no effect on the auxin pool originating from the shoot suggests that acropetal auxin transport might be less important in delivering the auxin required for the establishment

of the nodule-initiating auxin maximum. Indeed, when blocking acropetal auxin transport at the hypocotyl-root junction this did not affect nodulation whereas it reduced root growth and severely hampered lateral root development. This indicates that the NPA treatments were successful, and that shoot derived auxin is less important during nodulation.

It was previously reported that also basipetal auxin transport is not necessary for nodulation (Ng *et al.*, 2020). Similarly, removal of the auxin biosynthesis at the root meristem by re-sectioning the Medicago root tip also had no effect on nodulation. In Arabidopsis it was demonstrated that root wounding induced local auxin biosynthesis near the wound site to regulate root regeneration (Matosevich *et al.*, 2020). We cannot exclude a similar process occurs in Medicago, however, in our experiments the Medicago root regenerated roughly 50% of the time whereas root nodules were formed at the spot-inoculation side always, even when the roots were re-sectioned. Nevertheless, our results with those of Ng et al., (2020) indicates that under normal conditions, auxin produced in the root tip is not required for nodulation.

It has been demonstrated that in Arabidopsis, local auxin biosynthesis and long-distance auxin transport are collaborating in regulation of the auxin homeostasis required for main and lateral root growth (Marchant *et al.*, 2002; Laskowski *et al.*, 2008; Chen *et al.*, 2014; Guo *et al.*, 2014; Tang *et al.*, 2017). Our data suggest that a similar mechanism could underly Medicago primary and secondary root growth and development, but not nodule initiation. On the other hand, it is likely that local auxin transport attenuation is still needed to sustain an auxin maximum created by local auxin biosynthesis in the pericycle upon nodule initiation.

Pericycle derived auxin triggers mitotic reactivation of cortical cells.

As the expression of *MtYUC1*, *2*, and *8* during the early stages of nodule initiation is restricted to the pericycle, their spatial expression alone cannot explain the *DR5* pattern beyond stage I. The only other *MtYUC* reported to be expressed during nodule primordia formation is *MtYUC9*. However, this gene is only expressed at later stages of nodulation (Schiessl *et al.*, 2019). Our *in situ* hybridization experiments demonstrated that *MtYUC9* expression is only detectable from stage IV onward, and only at the periphery of the nodule primordium. Therefore, it is unlikely that the expression of *MtYUC9* is responsible for the auxin required to initiate the cortical *DR5* activity during the preceding stages II and III.

Nevertheless, *MtYUCs* are not expressed in the root cortex of the susceptible zone with or without rhizobia application, indicating that the inhibitory effect of PPBo on cortical cell divisions can only be indirect. The most likely place where PPBo can exert its effect on *Mt*YUCs is the pericycle. This suggests that auxin locally biosynthesized in the pericycle has to be translocated into the cortex to mitotically reactivate these cells. This could also explain the observed *DR5* activity in the cortex prior to cell divisions in the absence of *MtYUC* expression. Our results further suggest that the activated pericycle cells could serve as an auxin source for this activation.

Translocation towards the cortex is likely mediated by auxin efflux- and influx carriers. *MtPIN4*, *MtPIN10* and *MtLAX2* are expressed in the pericycle during the early stages of nodule initiation, and at least for *Mt*PIN10 protein, we have shown a localization in the pericycle and endodermis and we hypothesize that this could create an auxin transport channel from the pericycle through the endodermis into the cortex. *MtPIN2* is activated later, in cortical cell layer 5, possibly taking a role in the transport of auxin between different cortical cell layers. *MtPIN6* is activated even later, and is likely associated with the formation of the nodule vasculature. Combined, our results with those of Roy et al., (2017), suggest that *Mt*PIN4 and 10, in concert with *Mt*LAX2, are possible candidates to facilitate this putative auxin flux from the pericycle into the cortex. These findings are in line with previous reports on the effect of PIN1 in Soybean (*Glycine max*) where a triple CRIPPR-CAS9 knockout of *GmPIN1a*, *1b* and *1c*, reduced nodule numbers by approximately 30% (Gao *et al*., 2021). Different from indeterminate nodules, the initiation of determinate nodules on Soybean roots follows a different pattern. There, outer cortical cells are mitotically re-activated first after which cell divisions progress inwards (Turgeon and Bauer, 1982). Although the initiation of these nodule types is different, the fact that also in Soybean stacked *pin1* mutants show a reduction in nodule numbers suggests a more conserved role for polar auxin transport during nodulation (reviewed by Kohlen *et al*., 2018; Gao *et al*., 2021). Nevertheless, the fact that NPA application can trigger the formation of pseudonodules in the absence of rhizobia several weeks post application (Allen *et al*., 1953; Hirsch *et al*., 1989; Scheres *et al*., 1992; Schnabel and Frugoli, 2004; Subramanian *et al*., 2006; Wasson *et al*., 2006; Zhang *et al*., 2009; Rightmyer and Long, 2011; Shen *et al*., 2015; Sańko-Sawczenko *et al*., 2016) while NPA physically blocks all acropetal, basipetal and radial auxin transport (Ung *et al*., 2022) leaves open the possibility that other means of auxin translocation, i.e. through plasmodesmata or diffusion, could play a role during nodulation.

Indeed, it was previously demonstrated that the symplastic fluxes facilitated by plasmodesmata are linked to nodule development (Gaudioso-Pedraza *et al*., 2018). Additionally, plasmodesmata have been implied to be involved in the transport of several plant hormones, including auxin (reviewed by Tee and Faulkner, 2024). It is therefore interesting to investigate the role of plasmodesmata in auxin translocation from the pericycle to the cortex. However, Gaudioso-Pedraza and colleagues (2018) also demonstrated that symplastic connectivity is required for the transcriptional activation of the key nodulation transcription factor *NIN* and some of its downstream targets. This will complicate the distinction between cause and effect, as almost all of the auxin biosynthesis and transport genes reported in our study are (in)directly dependent on NIN (Schiessl *et al*., 2019).

Is the nodule meristem an auxin sink or source?

At stages V and VI, the expression of *MtYUC1*, *2* and *8* is confined to the base of the nodule and nodule vasculature, while *DR5* activity is highly concentrated in the C3 derived cortical cells where the nodule meristem is about to be formed. As such, also at this stage *DR5* activity and local auxin biosynthesis do not fully match.

In Arabidopsis, locally produced auxin by *At*YUCs is a principal factor for vascular strand formation (Cheng *et al.*, 2006; Cao *et al.*, 2019). The expression of *MtYUC1*, *2* and *8* in the developing vasculature indicates that in Medicago *Mt*YUCs might fulfil a similar role in vascular strand formation. However, the strong *MtYUC1*, *2* and *8* expression at the base of the nodule, in a cell layer likely derived from the pericycle, does not lead to *DR5* activity in these cells. This could mean two things. Either, as the synthetic auxin reporter *DR5* is based on a single auxin response factor binding motif (i.e. motif bound by AtARF1) (Ulmasov *et al.*, 1997), the observed *DR5* activity might not reflect the total auxin dynamics. On the other hand, as *DR5* activity requires auxin to be nuclear localized (reviewed by Yu *et al.*, 2022), the absence of a strong *DR5* signal here could therefore imply that most of the auxin produced in the cells at the base of the nodule is immediately transported away. A putative auxin flow towards the forming nodule meristem would explain the *DR5* activity observed in the C3 derived cells. It is possible that auxin, biosynthesized in the base of the nodule, the nodule vascular bundles, or even from the acropetal auxin transport stream, is translocated through the nodule vasculature towards the future nodule meristem.

MtPIN2, *4* and *10* genes are expressed in the nodule vasculature and therefore likely candidates to function in such auxin transport. Medicago mutants that fail to develop proper nodule vascular bundles as *Mtnoot1/2* (Magne *et al.*, 2018; Shen *et al.*, 2020) or *Mtlin4* (Guan *et al.*, 2013), while being impaired in nodule meristem development or maintenance are supporting the hypothesis that nodule vascular bundles are involved in nodule meristem development or maintenance.

If so, this would suggest that the auxin content of the nodule meristem is regulated through transport of auxin produced elsewhere. We demonstrated that, when the level of *MtPIN2*, *4* and *10* is reduced during nodule initiation, possibly in the forming nodule vasculature, the formation of the nodule meristem is hampered. Combined, this suggests that auxin transport is needed for the nodule meristem to be formed, and as a consequence the nodule meristem is likely an auxin sink. This auxin can either be transported directly through the C5-C4 layers, or through the developing nodule vasculature.

Based on the canalisation hypothesis the latter is the most likely option. This hypothesis states that a high source of auxin finds its way to an auxin sink, and in doing so paves the way for the formation of a vascular strand (reviewed by Hajný *et al.*, 2022). During leaf vascular patterning, the tip of the leaf primordia is an auxin source, and as a consequence vascular strand development progresses from the tip of the leaf towards the base (Scarpella *et al.*, 2006;

reviewed by Biedroń and Banasiak, 2018). If the nodule meristem is indeed an auxin sink, this canalisation hypothesis implies that vascularisation would progress from the base of the nodule towards the nodule meristem, which would be in contrast to leaf vascularization. Such a nodule meristem could be considered a non-autonomous acting meristem.

In conclusion, our data suggest that a precise spatiotemporal regulation of auxin outputs are guiding multiple stages of nodule primordium initiation and development. This spatiotemporal regulation is likely the result of a balanced interplay between local auxin biosynthesis and transport. We propose that during the early stages of nodule initiation, auxin is biosynthesized in the pericycle, from where it is translocated into the cortex to mitotically re-activate the cells in the C4 and C5 layers. At later stages, when the auxin flow from the pericycle into cells derived from these C4 and C5 layers diminishes, these cells stop dividing and start to differentiate. At the same time, auxin biosynthesized elsewhere (possibly in the pericycle, developed nodule vasculature and/or acropetal auxin transport stream in the main root) is needed to initiate meristem formation and later ensure its persistence. At this moment, the nodule meristem is a sink, draining the surrounding cells of their auxin. Proximal to the nodule meristem, in the cells derived from the C4 layer and later from the nodule meristem, this auxin reduction triggers cell differentiation leading to the infection zone.

MATERIALS AND METHODS

Plant material and bacterial strains

Medicago R108 seedlings were used to make the stable *DR5::GUS* transgenic line by using *Rhizobium radiobacter* (formerly *Agrobacterium tumefaciens*) (strain AGL1) according to the protocol described by (Chabaud *et al.*, 2003). Medicago Jemalong A17 plants were used to generate *A. rhizogenes* (strain MSU440) mediated transgenic roots as previously described by Limpens et al. (2004) for RNAi and protein GFP fusion constructs. *Mtpin10-1/slm1* (NF3969) were previously described by (Zhou *et al.*, 2011*a*). Surface-sterilization and germination of Medicago seeds were performed as previously described by Limpens et al. (2004). *Sinorhizobium meliloti 2011* (*Sm2011*) was used to induce root nodule formation on seedlings growing in perlite or on plates with low nitrate Fähraeus (Fä) medium (Fähraeus, 1957). 1 µM Aminoethoxyvinylglycine (AVG) was added to plates growing R108 plants only.

Constructs

DNA fragments were amplified from Medicago genomic DNA using primer combinations listed in supplemental Table S2 and Phusion™ High-Fidelity DNA Polymerase (Finnzymes). To create the constructs, pENTR™/D-TOPO® Cloning Kits (Invitrogen) and Gateway® technology (Invitrogen) were used to generate the entry clones for RNAi constructs, promoter-GUS and PIN10-GFP constructs. First, 14 synthetic DR5 DNA fragments (Ulmasov *et al.*, 1997) were introduced in the entry clone. Then, the entry vector was recombined into Gateway®-compatible binary vector pKGW-RR, that contains the GUS reporter gene and *AtUBQ10::DsRED1* as a selection marker (Limpens *et al.*, 2004), by using Gateway® LR Clonase® II enzyme mix (Invitrogen). PCR fragments of about 400- 600 bp of single *MtYUCs* and *MtPINs* genes for RNAi constructs were generated on cDNA made from Medicago nodule or root RNA. These fragments were combined by subsequential PCR steps using primers with a complementary 15 bp overhang to generate one amplicon of two or three *DNA* fragments. The final DNA fragments were cloned into pENTR-D-TOPO and recombined into the Gateway-compatible binary vector pK7GWIWG2(II)-UBQ10::DsRED driven by two different nodule-specific promoters, *ENOD12* or *ENOD40*, or *CaMV 35S* (*35S*) promoter (Compaan *et al.*, 2001; Limpens *et al.*, 2004) to create the final RNAi constructs. The control for *RNAi* constructs was *ENOD12::Empty Vector (EV)* containing no coding DNA sequencies. For *Mt*PIN10-GFP fusion constructs, the promoter and first part of the DNA fragment was introduced into Gateway® donor vector pENTR2-1, eGFP DNA fragment without start codon were introduced into pENTR1-2 and the rest of the gene including the terminator was introduced into pENTR2-3, using Gateway® BP Clonase® II enzyme mix. These entry vectors were recombined into Gateway®-compatible binary vector pKGW-RR-MGW containing *AtUBQ10::DsRED1* as a selection marker using Gateway® LR Clonase® II Plus enzyme mix (Invitrogen). Insertion sites for the eGFP reporter in *PIN* genes were selected based on functional AtPINs::GFP protein

fusion constructs in Arabidopsis (Xu and Scheres, 2005), which are at position 1237 and 1565 for *MtPIN2* and *MtPIN10*, respectively. All used primers are listed in supplemental Table S2.

Tissue embedding, sectioning and staining

Root segments were fixed in 4% paraformaldehyde (w/v), 5% glutaraldehyde (v/v), 0.05 M sodium phosphate buffer (pH7.2) at 4 °C overnight. The fixed material was dehydrated in an ethanol series and subsequently embedded in Technovit 7100 (Heraeus Kulzer) according to the manufacturer's protocol. Sections (7 µm) were made by using a RJ2035 microtome (Leica Microsystems, Rijswijk, The Netherland), stained 5 min in 0.1% Ruthenium Red (Sigma, Germany). Sections were analysed by using a DM5500B microscope equipped with a DFC425C camera (Leica Microsystems, Wetzlar, Germany).

Suppression of YUCCAs by 10 µM PPBo

The germinated Medicago seeds were placed on Fä plates for 4 days and then seedlings were transferred to new plates containing Fä medium with either 10 µM PPBo or mock (Fähraeus, 1957; Kakei *et al.*, 2015). After 24 hrs, roots of these seedlings were spot inoculated with *Sm2011* (OD_{600}=0.02). Plants were grown further for 1, 2, 3, 4 or 7 more days depending on the experiment.

Suppression of Polar auxin transport 50 µM NPA

The germinated seeds were placed on Fä plates for 4 days and then treated with Naphthylphtalamic acid (NPA). NPA [50 mM] was dissolved in DMSO and diluted to an end concentration of 50 µM in Fä medium with 0.8 % Daishin agar. 200 µL agar droplets containing either NPA or Mock (0.1% DMSO) were solidified and locally applied on the hypocotyl–root junction. After 24 hrs, roots were spot-inoculated with *Sm2011* (OD_{600}=0.02). Plants were grown further for 8 days to form nodules. Successful nodulation, root growth, and the number of lateral roots were assessed.

Auxin quantification

Auxin (Indole-3-acidic acid (IAA)) was quantified through Liquid Chromatography-tandem Mass Spectrometry (LC-MS/MS). Medicago roots were collected and flash frozen in liquid nitrogen. Collected tissue was ground to a fine powder at -80°C using 3 mm stainless steel beads at 50 Hz for 2*30 seconds in a TissueLyser LT (Qiagen, Germantown, USA), and between 10-15 mg of ground tissue per sample was used for IAA extraction. Samples were extracted with 1 mL of cold methanol containing [phenyl $^{13}C_6$]-IAA (0.1 nmol/mL) as an internal standard in a 2-mL Eppendorf tube and purified as previously described (Ruyter-Spira *et al.*, 2011). Samples were filtered through a 0.45 µm Minisart SRP4 filter (Sartorius, Goettingen, Germany) and measured on the same day. Auxin was analyzed on a Waters Xevo TQs tandem quadruple mass spectrometer as previously described (Ruyter-Spira *et al.*, 2011; Gühl *et al.*, 2021).

RNA isolation and qRT-PCR

RNA was isolated from transgenic *35S::YUC1/2/8i* (*35S::YUCi*), *35S::PIN2/4/10i* (*35S::PINi*) and *EV* root tips (5-7 mm) using RNA Easy kit (Qiagen). 500 ng RNA was used for cDNA synthesis using the iScript cDNA synthesis kit (Bio-Rad). qRT-PCR was performed using SYBR Green Supermix and CFX real-time system (Bio-Rad). Gene expression was normalized using *ACTIN2* as a reference gene. Primers used for qPCR are listed in supplemental Table S2.

RNA *in situ* hybridization

RNA *in situ* hybridization was conducted using Invitrogen ViewRNA ISH Tissue 1-Plex and 2-Plex Assay kits (ThermoFisher Scientific) using the manufacture's protocol optimized for Medicago root and nodules sections (Kulikova *et al.*, 2018). RNA ISH probe sets were designed and synthesized at ThermoFisher Scientific. Assay ID of used probe sets are presented in Table S3. For each stage of nodule primordium development and with each probe set at least 3 *in situ* hybridizations were performed, each hybridization procedure included 2-3 different root segments on a slide. As a negative control for each hybridization run one slide was used without any probe set. *MtNF-YA1* probe set Type 6 was used as a marker for nodule primordium initiation (blue hybridization signals). The images were taken with an AU5500B microscope equipped with a DFC425c camera (Leica).

Detection of MtPIN10-GFP

Transgenic roots or root nodules were manually sectioned in a longitudinal direction, mounted on microscope slides and immediately observed under a Leica SP8 confocal microscope. GFP is visualized by excitation at 488 nm and by detection at 505–530 nm.

Statistics

When appropriate, data were subjected to the Student's t test (Microsoft Excel). All other data were subjected to one-way ANOVA. Individual differences were then identified using a post hoc Tukey test ($P < 0.05$). All analyses were performed using SAS_9.20 (http://www.sas.com/).

AUTHOR CONTRIBUTIONS

Conceptualization, HF, TB, OK, and WK; Methodology OK, and WK; Investigation, TTX, SM, DS, JL, KA, AvS, HF, OK, and WK; Writing – Original Draft, TTX, HF, TB, OK, and WK; Writing – Review & Editing, all authors; Writing – final version, HF, OK, and WK; Funding Acquisition, TB, and WK; Supervision, OK, and WK.

ACKNOWLEDGMENTS

This research is funded by European Research Council (ERC-2011-AdG-294790) to TB, NWO-VENI (863.15.010) and NWO-VIDI (VI.Vidi.193.119) to WK, and Graduate School 'Experimental Plant Science'.

REFERENCES

Allen EK, Allen ON, Newman AS. 1953. Pseudonodulation of Leguminous Plants Induced by 2-Bromo-3,5-Dichlorobenzoic Acid. American journal of botany 40, 429–435.

Benková E, Michniewicz M, Sauer M, Teichmann T, Seifertová D, Jürgens G, Friml J. 2003. Local, efflux-dependent auxin gradients as a common module for plant organ formation. Cell 115, 591–602.

Bennett MJ, Marchant A, Green HG, May ST, Ward SP, Millner PA, Walker AR, Schulz B, Feldmann KA. 1996. Arabidopsis AUX1 gene: a permease-like regulator of root gravitropism. Science 273, 948–950.

Bhalerao RP, Eklöf J, Ljung K, Marchant A, Bennett M, Sandberg G. 2002. Shoot-derived auxin is essential for early lateral root emergence in Arabidopsis seedlings. The Plant journal: for cell and molecular biology 29, 325–332.

Biedroń M, Banasiak A. 2018. Auxin-mediated regulation of vascular patterning in Arabidopsis thaliana leaves. Plant cell reports 37, 1215–1229.

de Billy F, Grosjean C, May S, Bennett M, Cullimore JV. 2001. Expression studies on AUX1-like genes in Medicago truncatula suggest that auxin is required at two steps in early nodule development. Molecular plant-microbe interactions: MPMI 14, 267–277.

Boot KJM, van Brussel AAN, Tak T, Spaink HP, Kijne JW. 1999. Lipochitin Oligosaccharides from Rhizobium leguminosarum bv. viciae Reduce Auxin Transport Capacity in Vicia sativa subsp. nigra Roots. Molecular plant-microbe interactions: MPMI 12, 839–844.

Brewin NJ. 2004. Plant Cell Wall Remodelling in the Rhizobium–Legume Symbiosis. Critical reviews in plant sciences 23, 293–316.

Cao X, Yang H, Shang C, Ma S, Liu L, Cheng J. 2019. The Roles of Auxin Biosynthesis YUCCA Gene Family in Plants. International journal of molecular sciences 20.

Cazzonelli CI, Vanstraelen M, Simon S, *et al.* 2013. Role of the Arabidopsis PIN6 auxin transporter in auxin homeostasis and auxin-mediated development. PloS one 8, e70069.

Chabaud M, de Carvalho-Niebel F, Barker DG. 2003. Efficient transformation of Medicago truncatula cv. Jemalong using the hypervirulent Agrobacterium tumefaciens strain AGL1. Plant cell reports 22, 46–51.

Chen Q, Dai X, De-Paoli H, Cheng Y, Takebayashi Y, Kasahara H, Kamiya Y, Zhao Y. 2014. Auxin overproduction in shoots cannot rescue auxin deficiencies in Arabidopsis roots. Plant & cell physiology 55, 1072–1079.

Cheng Y, Dai X, Zhao Y. 2006. Auxin biosynthesis by the YUCCA flavin monooxygenases controls the formation of floral organs and vascular tissues in Arabidopsis. Genes & development 20, 1790–1799.

Cheng Y, Dai X, Zhao Y. 2007. Auxin synthesized by the YUCCA flavin monooxygenases is essential for embryogenesis and leaf formation in Arabidopsis. The Plant cell 19, 2430–2439.

Compaan B, Yang W-C, Bisseling T, Franssen H. 2001. ENOD40 expression in the pericycle precedes cortical cell division in Rhizobium-legume interaction and the highly conserved internal region of the gene does not encode a peptide. Plant and soil 230, 1–8.

Crespi MD, Jurkevitch E, Poiret M, d'Aubenton-Carafa Y, Petrovics G, Kondorosi E, Kondorosi A. 1994. enod40, a gene expressed during nodule organogenesis, codes for a non-translatable RNA involved in plant growth. The EMBO journal 13, 5099–5112.

Deinum EE, Geurts R, Bisseling T, Mulder BM. 2012. Modeling a cortical auxin maximum for nodulation: different signatures of potential strategies. Frontiers in plant science 3, 96.

Deinum EE, Kohlen W, Geurts R. 2016. Quantitative modelling of legume root nodule primordium induction by a diffusive signal of epidermal origin that inhibits auxin efflux. BMC plant biology 16, 254.

Demina IV, Maity PJ, Nagchowdhury A, Ng JLP, van der Graaff E, Demchenko KN, Roitsch T, Mathesius U, Pawlowski K. 2019. Accumulation of and Response to Auxins in Roots and Nodules of the Actinorhizal Plant Datisca glomerata Compared to the Model Legume Medicago truncatula. Frontiers in plant science 10, 1085.

Dénarié J, Debellé F, Promé JC. 1996. Rhizobium lipo-chitooligosaccharide nodulation factors: signaling molecules mediating recognition and morphogenesis. Annual review of biochemistry 65, 503–535.

Fähraeus G. 1957. The infection of clover root hairs by nodule bacteria studied by a simple glass slide technique. Journal of general microbiology 16, 374–381.

Gage DJ. 2004. Infection and invasion of roots by symbiotic, nitrogen-fixing rhizobia during nodulation of temperate legumes. Microbiology and molecular biology reviews: MMBR 68, 280–300.

Gao Z, Chen Z, Cui Y, *et al.* 2021. GmPIN-dependent polar auxin transport is involved in soybean nodule development. The plant cell 33, 2981–3003.

Gaudioso-Pedraza R, Beck M, Frances L, Kirk P, Ripodas C, Niebel A, Oldroyd GED, Benitez-Alfonso Y, de Carvalho-Niebel F. 2018. Callose-regulated symplastic communication coordinates symbiotic root nodule development. Current biology: CB 28, 3562-3577.e6.

Geurts R, Bisseling T. 2002. Rhizobium nod factor perception and signalling. The Plant cell 14 Suppl, S239-49.

Guan D, Stacey N, Liu C, *et al.* 2013. Rhizobial infection is associated with the development of peripheral vasculature in nodules of Medicago truncatula. Plant physiology 162, 107–115.

Gühl K, Holmer R, Xiao TT, Shen D, Wardhani TAK, Geurts R, van Zeijl A, Kohlen W. 2021. The Effect of Exogenous Nitrate on LCO Signalling, Cytokinin Accumulation, and Nodule Initiation in Medicago truncatula. Genes 12.

Guo Y, Chen F, Zhang F, Mi G. 2005. Auxin transport from shoot to root is involved in the response of lateral root growth to localized supply of nitrate in maize. Plant science: an international journal of experimental plant biology 169, 894–900.

Guo J, Wei J, Xu J, Sun M-X. 2014. Inducible knock-down of GNOM during root formation reveals tissue-specific response to auxin transport and its modulation of local auxin biosynthesis. Journal of experimental botany 65, 1165–1179.

Hajný J, Tan S, Friml J. 2022. Auxin canalization: From speculative models toward molecular players. Current opinion in plant biology 65, 102174.

Herrbach V, Chirinos X, Rengel D, *et al.* 2017. Nod factors potentiate auxin signaling for transcriptional regulation and lateral root formation in Medicago truncatula. Journal of experimental botany 68, 569–583.

Hirsch AM. 1992. Developmental biology of legume nodulation. The New phytologist 122, 211–237.

Hirsch AM, Bhuvaneswari TV, Torrey JG, Bisseling T. 1989. Early nodulin genes are induced in alfalfa root outgrowths elicited by auxin transport inhibitors. Proceedings of the National Academy of Sciences of the United States of America 86, 1244–1248.

Huo X, Schnabel E, Hughes K, Frugoli J. 2006. RNAi Phenotypes and the Localization of a Protein::GUS Fusion Imply a Role for Medicago truncatula PIN Genes in Nodulation. Journal of plant growth regulation 25, 156–165.

Journet EP, El-Gachtouli N, Vernoud V, de Billy F, Pichon M, Dedieu A, Arnould C, Morandi D, Barker DG, Gianinazzi-Pearson V. 2001. Medicago truncatula ENOD11: a novel RPRP-encoding early nodulin gene expressed during mycorrhization in arbuscule-containing cells. Molecular plant-microbe interactions: MPMI 14, 737–748.

Kakei Y, Yamazaki C, Suzuki M, *et al.* 2015. Small-molecule auxin inhibitors that target YUCCA are powerful tools for studying auxin function. The Plant journal: for cell and molecular biology 84, 827–837.

Kohlen W, Ng JLP, Deinum EE, Mathesius U. 2018. Auxin transport, metabolism, and signalling during nodule initiation: indeterminate and determinate nodules. Journal of experimental botany 69, 229–244.

Kulikova O, Franken C, Bisseling T. 2018. In Situ Hybridization Method for Localization of mRNA Molecules in Medicago Tissue Sections. Methods in molecular biology 1822, 145–159.

Laskowski M, Grieneisen VA, Hofhuis H, Hove CAT, Hogeweg P, Marée AFM, Scheres B. 2008. Root system architecture from coupling cell shape to auxin transport. PLoS biology 6, e307.

Li H, Schilderink S, Cao Q, Kulikova O, Bisseling T. 2021. Plant-specific histone deacetylases are essential for early and late stages of Medicago nodule development. Plant physiology 186, 1591–1605.

Limpens E, Ramos J, Franken C, Raz V, Compaan B, Franssen H, Bisseling T, Geurts R. 2004. RNA interference in Agrobacterium rhizogenes-transformed roots of Arabidopsis and Medicago truncatula. Journal of experimental botany 55, 983–992.

Lin J, Frank M, Reid D. 2020. No Home without Hormones: How Plant Hormones Control Legume Nodule Organogenesis. Plant communications 1, 100104.

Liu J, Rutten L, Limpens E, *et al.* 2019. A Remote cis-Regulatory Region Is Required for NIN Expression in the Pericycle to Initiate Nodule Primordium Formation in Medicago truncatula. The Plant cell 31, 68–83.

Magne K, Couzigou J-M, Schiessl K, *et al.* 2018. MtNODULE ROOT1 and MtNODULE ROOT2 Are Essential for Indeterminate Nodule Identity. Plant physiology 178, 295–316.

Marchant A, Bhalerao R, Casimiro I, Eklöf J, Casero PJ, Bennett M, Sandberg G. 2002. AUX1 promotes lateral root formation by facilitating indole-3-acetic acid distribution between sink and source tissues in the Arabidopsis seedling. The Plant cell 14, 589–597.

Mashiguchi K, Tanaka K, Sakai T, *et al.* 2011. The main auxin biosynthesis pathway in Arabidopsis. Proceedings of the National Academy of Sciences of the United States of America 108, 18512–18517.

Mathesius U. 2008. Goldacre paper: Auxin: at the root of nodule development? Functional plant biology: FPB 35, 651–668.

Mathesius U, Bayliss C, Weinman JJ, Schlaman HRM, Spaink HP, Rolfe BG, McCully ME, Djordjevic MA. 1998*a*. Flavonoids Synthesized in Cortical Cells During Nodule Initiation Are Early Developmental Markers in White Clover. Molecular plant-microbe interactions: MPMI 11, 1223–1232.

Mathesius U, Schlaman HR, Spaink HP, Of Sautter C, Rolfe BG, Djordjevic MA. 1998*b*. Auxin transport inhibition precedes root nodule formation in white clover roots and is regulated by flavonoids and derivatives of chitin oligosaccharides. The Plant journal: for cell and molecular biology 14, 23–34.

Matosevich R, Cohen I, Gil-Yarom N, Modrego A, Friedlander-Shani L, Verna C, Scarpella E, Efroni I. 2020. Local auxin biosynthesis is required for root regeneration after wounding. Nature plants 6, 1020–1030.

McDAVID CR, Sagar GR, Marshall C. 1972. THE EFFECT OF AUXIN FROM THE SHOOT ON ROOT DEVELOPMENT IN *PISUM SATIVUM* L. The new phytologist 71, 1027–1032.

Ng JLP, Hassan S, Truong TT, Hocart CH, Laffont C, Frugier F, Mathesius U. 2015. Flavonoids and Auxin Transport Inhibitors Rescue Symbiotic Nodulation in the Medicago truncatula Cytokinin Perception Mutant cre1. The Plant cell 27, 2210–2226.

Ng JLP, Mathesius U. 2018. Acropetal Auxin Transport Inhibition Is Involved in Indeterminate But Not Determinate Nodule Formation. Frontiers in plant science 9, 169.

Ng JLP, Welvaert A, Wen J, Chen R, Mathesius U. 2020. The Medicago truncatula PIN2 auxin transporter mediates basipetal auxin transport but is not necessary for nodulation. Journal of experimental botany 71, 1562–1573.

Oldroyd GED, Downie JA. 2006. Nuclear calcium changes at the core of symbiosis signalling. Current opinion in plant biology 9, 351–357.

Overvoorde P, Fukaki H, Beeckman T. 2010. Auxin control of root development. Cold Spring Harbor perspectives in biology 2, a001537.

Pacios-Bras C, Schlaman HRM, Boot K, Admiraal P, Langerak JM, Stougaard J, Spaink HP. 2003. Auxin distribution in Lotus japonicus during root nodule development. Plant molecular biology 52, 1169–1180.

Peng J, Chen R. 2011. Auxin efflux transporter MtPIN10 regulates compound leaf and flower development in Medicago truncatula. Plant signaling & behavior 6, 1537–1544.

Petrásek J, Mravec J, Bouchard R, *et al.* 2006. PIN proteins perform a rate-limiting function in cellular auxin efflux. Science 312, 914–918.

Plet J, Wasson A, Ariel F, Le Signor C, Baker D, Mathesius U, Crespi M, Frugier F. 2011. MtCRE1-dependent cytokinin signaling integrates bacterial and plant cues to coordinate symbiotic nodule organogenesis in Medicago truncatula. The Plant journal: for cell and molecular biology 65, 622–633.

Prayitno J, Rolfe BG, Mathesius U. 2006. The Ethylene-insensitive sickle mutant of Medicago truncatula shows altered auxin transport regulation during nodulation. Plant physiology 142, 168–180.

Reed RC, Brady SR, Muday GK. 1998. Inhibition of auxin movement from the shoot into the root inhibits lateral root development in Arabidopsis. Plant physiology 118, 1369–1378.

Rightmyer AP, Long SR. 2011. Pseudonodule formation by wild-type and symbiotic mutant Medicago truncatula in response to auxin transport inhibitors. Molecular plant-microbe interactions: MPMI 24, 1372–1384.

Roy S, Robson F, Lilley J, *et al.* 2017. MtLAX2, a Functional Homologue of the Arabidopsis Auxin Influx Transporter AUX1, Is Required for Nodule Organogenesis. Plant physiology 174, 326–338.

Roychoudhry S, Kepinski S. 2022. Auxin in root development. Cold Spring Harbor perspectives in biology 14, a039933.

Rubery PH, Jacobs M. 1990. Auxin transport and its regulation by flavonoids. Plant Growth Substances 1988. Berlin, Heidelberg: Springer Berlin Heidelberg, 428–440.

Ruyter-Spira C, Kohlen W, Charnikhova T, *et al.* 2011. Physiological effects of the synthetic strigolactone analog GR24 on root system architecture in Arabidopsis: another belowground role for strigolactones? Plant physiology 155, 721–734.

Sańko-Sawczenko I, Łotocka B, Czarnocka W. 2016. Expression Analysis of PIN Genes in Root Tips and Nodules of Medicago truncatula. International journal of molecular sciences 17.

Scarpella E, Marcos D, Friml J, Berleth T. 2006. Control of leaf vascular patterning by polar auxin transport. Genes & development 20, 1015–1027.

Scheres B, McKhann HI, Zalensky A, Löbler M, Bisseling T, Hirsch AM. 1992. The PsENOD12 Gene Is Expressed at Two Different Sites in Afghanistan Pea Pseudonodules Induced by Auxin Transport Inhibitors. Plant physiology 100, 1649–1655.

Schiessl K, Lilley JLS, Lee T, *et al.* 2019. NODULE INCEPTION Recruits the Lateral Root Developmental Program for Symbiotic Nodule Organogenesis in Medicago truncatula. Current biology: CB 29, 3657-3668.e5.

Schnabel EL, Frugoli J. 2004. The PIN and LAX families of auxin transport genes in Medicago truncatula. Molecular genetics and genomics: MGG 272, 420–432.

Shen D, Xiao TT, van Velzen R, Kulikova O, Gong X, Geurts R, Pawlowski K, Bisseling T. 2020. A Homeotic Mutation Changes Legume Nodule Ontogeny into Actinorhizal-Type Ontogeny. The Plant cell 32, 1868–1885.

Shen C, Yue R, Bai Y, Feng R, Sun T, Wang X, Yang Y, Tie S, Wang H. 2015. Identification and Analysis of Medicago truncatula Auxin Transporter Gene Families Uncover their Roles in Responses to Sinorhizobium meliloti Infection. Plant & cell physiology 56, 1930–1943.

Simon S, Skůpa P, Viaene T, *et al.* 2016. PIN6 auxin transporter at endoplasmic reticulum and plasma membrane mediates auxin homeostasis and organogenesis in Arabidopsis. The New phytologist 211, 65–74.

Sprent JI, James EK. 2007. Legume evolution: where do nodules and mycorrhizas fit in? Plant physiology 144, 575–581.

Stepanova AN, Yun J, Robles LM, Novak O, He W, Guo H, Ljung K, Alonso JM. 2011. The Arabidopsis YUCCA1 flavin monooxygenase functions in the indole-3-pyruvic acid branch of auxin biosynthesis. The Plant cell 23, 3961–3973.

Subramanian S, Stacey G, Yu O. 2006. Endogenous isoflavones are essential for the establishment of symbiosis between soybean and Bradyrhizobium japonicum. The Plant journal: for cell and molecular biology 48, 261–273.

Swarup R, Bennett M. 2014. Auxin transport: Providing plants with a new sense of direction. The biochemist 36, 12–15.

Swarup R, Bhosale R. 2019. Developmental Roles of AUX1/LAX Auxin Influx Carriers in Plants. Frontiers in plant science 10, 1306.

Swarup R, Friml J, Marchant A, Ljung K, Sandberg G, Palme K, Bennett M. 2001. Localization of the auxin permease AUX1 suggests two functionally distinct hormone transport pathways operate in the Arabidopsis root apex. Genes & development 15, 2648–2653.

Swarup R, Péret B. 2012. AUX/LAX family of auxin influx carriers—an overview. Frontiers in plant science 3.

Tang LP, Zhou C, Wang SS, Yuan J, Zhang XS, Su YH. 2017. FUSCA3 interacting with LEAFY COTYLEDON2 controls lateral root formation through regulating YUCCA4 gene expression in Arabidopsis thaliana. The New phytologist 213, 1740–1754.

Tee EE, Faulkner C. 2024. Plasmodesmata and intercellular molecular traffic control. The new phytologist 243, 32–47.

Turgeon BG, Bauer WD. 1982. Early events in the infection of soybean by Rhizobium japonicum. Time course and cytology of the initial infection process. Canadian journal of botany. Journal canadien de botanique 60, 152–161.

Ulmasov T, Murfett J, Hagen G, Guilfoyle TJ. 1997. Aux/IAA proteins repress expression of reporter genes containing natural and highly active synthetic auxin response elements. The Plant cell 9, 1963–1971.

Ung KL, Winkler M, Schulz L, Kolb M, Janacek DP, Dedic E, Stokes DL, Hammes UZ, Pedersen BP. 2022. Structures and mechanism of the plant PIN-FORMED auxin transporter. Nature 609, 605–610.

Vanneste S, Friml J. 2009. Auxin: a trigger for change in plant development. Cell 136, 1005–1016.

Wang Y, Yang W, Zuo Y, Zhu L, Hastwell AH, Chen L, Tian Y, Su C, Ferguson BJ, Li X. 2019. GmYUC2a mediates auxin biosynthesis during root development and nodulation in soybean. Journal of experimental botany 70, 3165–3176.

Wasson AP, Pellerone FI, Mathesius U. 2006. Silencing the flavonoid pathway in Medicago truncatula inhibits root nodule formation and prevents auxin transport regulation by rhizobia. The Plant cell 18, 1617–1629.

Wisniewska J, Xu J, Seifertová D, Brewer PB, Ruzicka K, Blilou I, Rouquié D, Benková E, Scheres B, Friml J. 2006. Polar PIN localization directs auxin flow in plants. Science 312, 883.

Won C, Shen X, Mashiguchi K, Zheng Z, Dai X, Cheng Y, Kasahara H, Kamiya Y, Chory J, Zhao Y. 2011. Conversion of tryptophan to indole-3-acetic acid by TRYPTOPHAN AMINOTRANSFERASES OF ARABIDOPSIS and YUCCAs in Arabidopsis. Proceedings of the National Academy of Sciences of the United States of America 108, 18518–18523.

Xiao TT. 2015. Root and Nodule: Lateral Organ Development in N2-Fixing Plants. *in press*.

Xiao TT, Schilderink S, Moling S, Deinum EE, Kondorosi E, Franssen H, Kulikova O, Niebel A, Bisseling T. 2014. Fate map of Medicago truncatula root nodules. Development 141, 3517–3528.

Xu J, Scheres B. 2005. Dissection of Arabidopsis ADP-RIBOSYLATION FACTOR 1 function in epidermal cell polarity. The Plant cell 17, 525–536.

Young JM, Kuykendall LD, Martínez-Romero E, Kerr A, Sawada H. 2001. A revision of Rhizobium Frank 1889, with an emended description of the genus, and the inclusion of all species of Agrobacterium Conn 1942 and Allorhizobium undicola de Lajudie et al. 1998 as new combinations: Rhizobium radiobacter, R. rhizogenes, R. rubi, R. undicola and R. vitis. International journal of systematic and evolutionary microbiology 51, 89–103.

Yu Z, Zhang F, Friml J, Ding Z. 2022. Auxin signaling: Research advances over the past 30 years. Journal of integrative plant biology 64, 371–392.

van Zeijl A, Op den Camp RHM, Deinum EE, Charnikhova T, Franssen H, Op den Camp HJM, Bouwmeester H, Kohlen W, Bisseling T, Geurts R. 2015. Rhizobium Lipo-chitooligosaccharide Signaling Triggers Accumulation of Cytokinins in Medicago truncatula Roots. Molecular plant 8, 1213–1226.

Zhang J, Subramanian S, Stacey G, Yu O. 2009. Flavones and flavonols play distinct critical roles during nodulation of Medicago truncatula by Sinorhizobium meliloti. The Plant journal: for cell and molecular biology 57, 171–183.

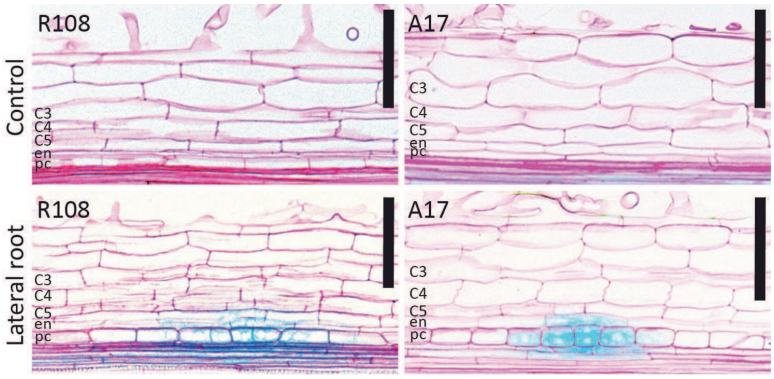
Zhao Y. 2010. Auxin biosynthesis and its role in plant development. Annual review of plant biology 61, 49–64.

Zhao Y. 2018. Essential Roles of Local Auxin Biosynthesis in Plant Development and in Adaptation to Environmental Changes. Annual review of plant biology 69, 417–435.

Zhou C, Han L, Hou C, Metelli A, Qi L, Tadege M, Mysore KS, Wang Z-Y. 2011*a*. Developmental analysis of a Medicago truncatula smooth leaf margin1 mutant reveals context-dependent effects on compound leaf development. The Plant cell 23, 2106–2124.

Zhou C, Han L, Wang Z-Y. 2011*b*. Potential but limited redundant roles of MtPIN4, MtPIN5 and MtPIN10/SLM1 in the development of Medicago truncatula. Plant signaling & behavior 6, 1834–1836.

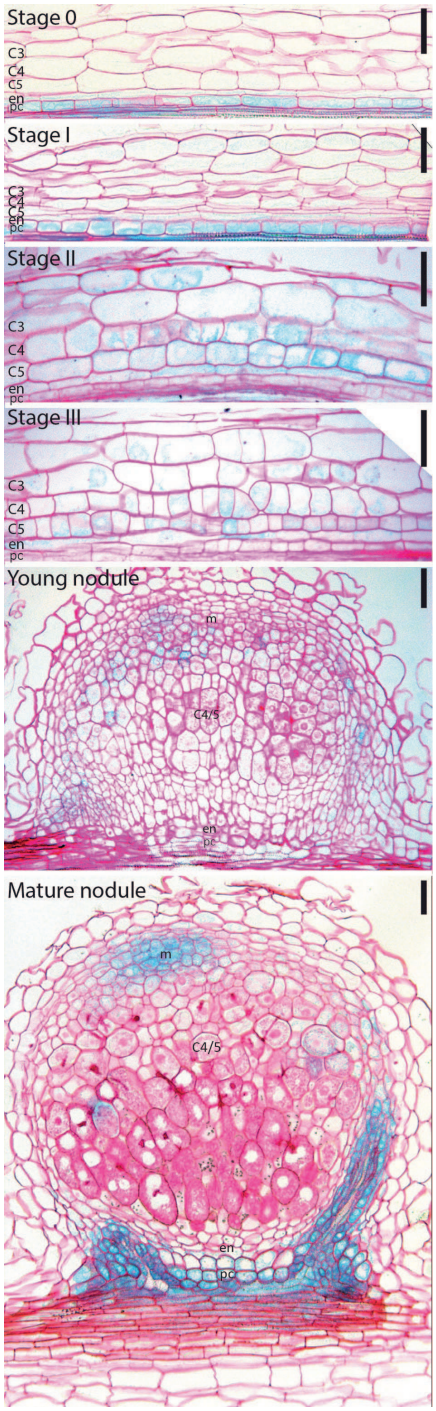
SUPPLEMENTAL FIGURES



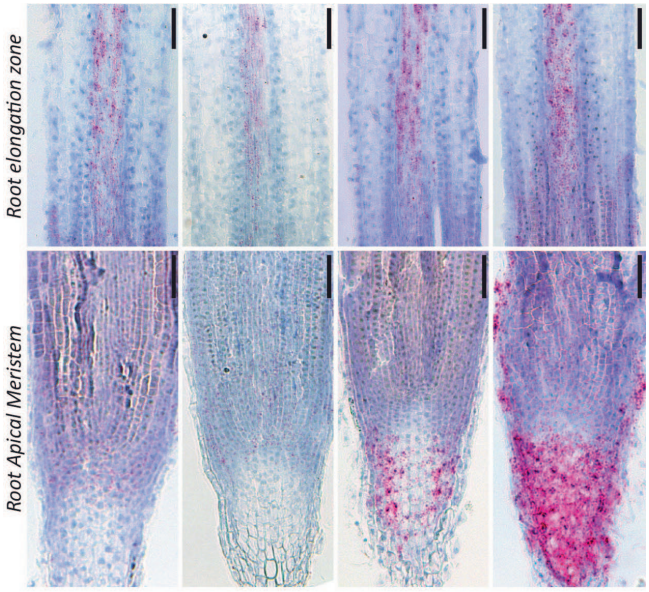
Supplemental Figure S1 | *DR5::GUS* expression pattern in Medicago roots of R108 and A17. Representative images of a control root and young lateral root primordia highlighting the narrow initiation site of lateral root development. Longitudinal plastic sections of root segments stained with ruthenium red. C3-C5, cortical cell layers; en, endodermis; pc, pericycle; scale bars 75 µm.



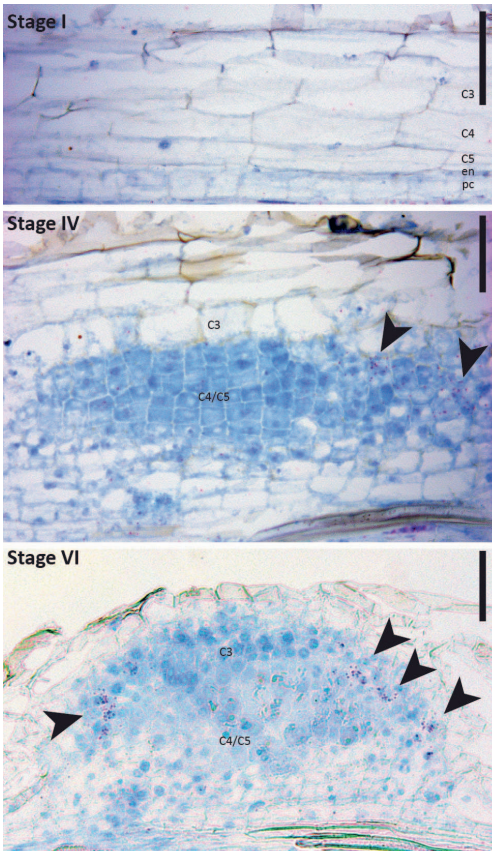
Supplemental Figure S2 | DR5::GUS expression pattern in Medicago R108 nodule. Longitudinal plastic sections of a mature nodule stained with ruthenium red (C4-C5, cortical cell layers; en, endodermis; pc, pericycle; m, meristem; scale bars 75 µm).



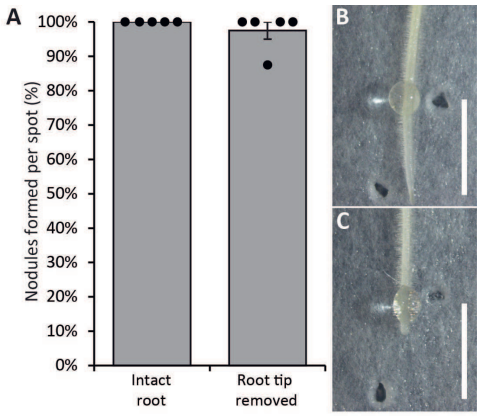
Supplemental Figure S3 | Dynamics of *DR5::GUS* expression patterns during nodule development in Medicago A17. The *DR5::GUS* construct was introduced in A17 by transient hairy root transformation. Longitudinal plastic sections of root segments with different stages of nodule development were stained with ruthenium red. C3-C5, cortical cell layers; en, endodermis; pc, pericycle; m, meristem; scale bars 75 µm.



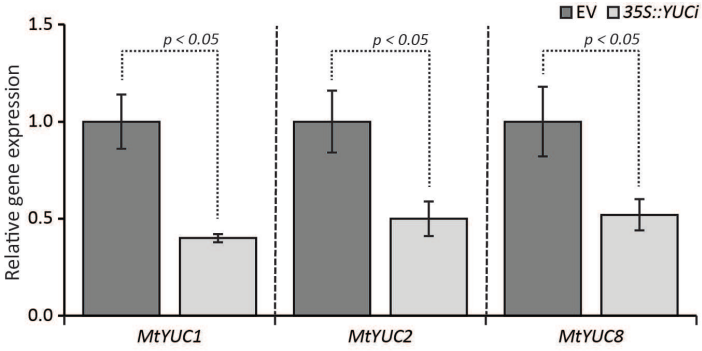
Supplemental Figure S4 | The spatiotemporal expression patterns of *YUCCAs* in root tip of Medicago A17. Representative images of RNA *in situ* hybridizations with *MtYUC1*, *MtYUC2*, *MtYUC8* or *MtYUC9* probe sets on longitudinal sections of the root tip including elongation zone and root meristem (red dots are hybridization signals; scale bars 75µm).



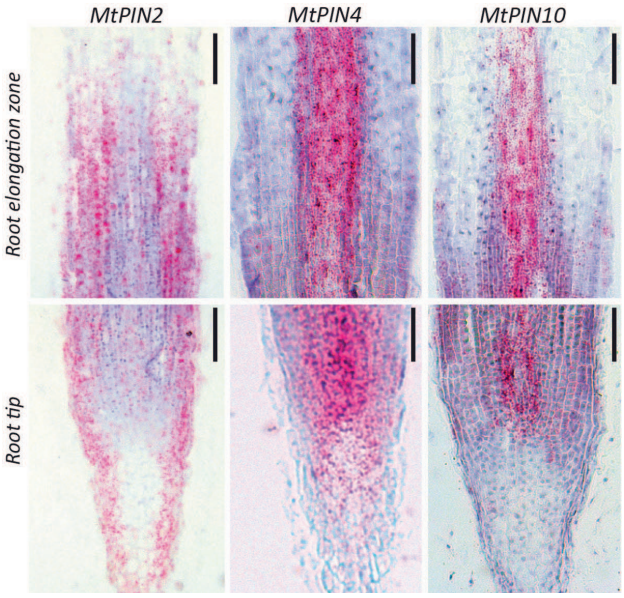
Supplemental Figure S5 | The spatiotemporal expression patterns of *MtYUC9* during nodule primordium formation in Medicago A17. Representative images of RNA *in situ* hybridization with *MtYUC9* probe set on longitudinal sections of nodule primordia stages I, IV and VI (red dots are hybridization signals; arrows indicate regions with hybridization signals). C3-C5, cortical cell layers; ed, endodermis; pc, pericycle; scale bars 75 µm.



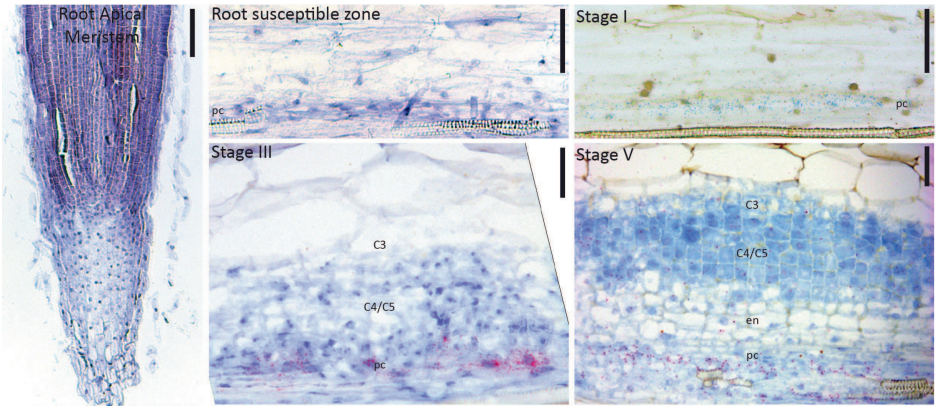
Supplemental Figure S6 | Effect of re-sectioning of the primary root on nodulation. (A) Average percentage of nodules formed per plate after rhizobia spot-applications on intact or re-sectioned Medicago roots, bars represent averages, ± standard error, dots show percentage of spotted roots to form a nodule per plate, n=5 (5 plates with each ~6-8 Medicago plants). **(B)** Application spot of the intact root, **(C)** Application spot and cut site of the re-sectioned root. Scale bars 5 mm.



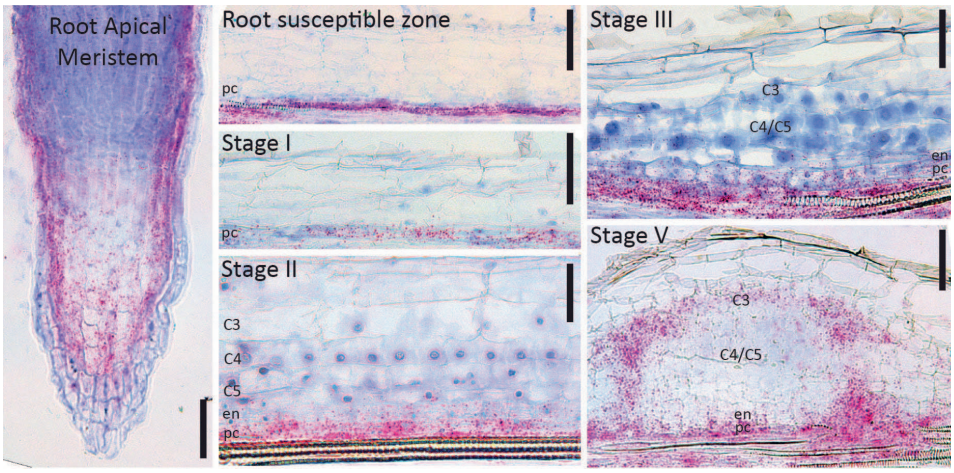
Supplemental Figure S7 | Validation of RNA interference on *MtYUC1*, *MtYUC2* and *MtYUC8* targeted by *YUCi* in composite Medicago A17 plants. Relative expression of *MtYUC1*, *2* and *8* in transgenic *35S::YUC1/2/8i* (*35S::YUCi*) roots compared to *EV* control roots, bars represent averages, ± standard error, n=3.



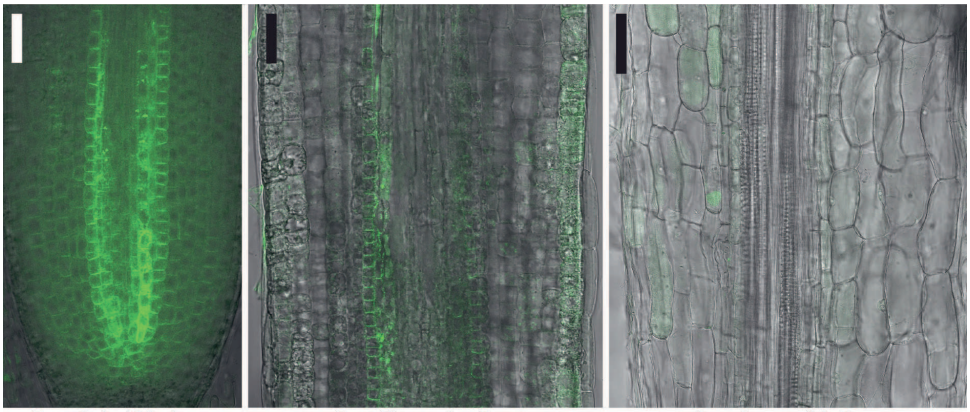
Supplemental Figure S8 | The spatiotemporal expression patterns of *PINs* in the root tip of Medicago A17. Representative images of RNA *in situ* hybridizations with *MtPIN2*, *MtPIN4*, or *MtPIN10* probe set on longitudinal sections of the root tip including elongation zone and root meristem (red dots are hybridization signals; scale bars 75μm).



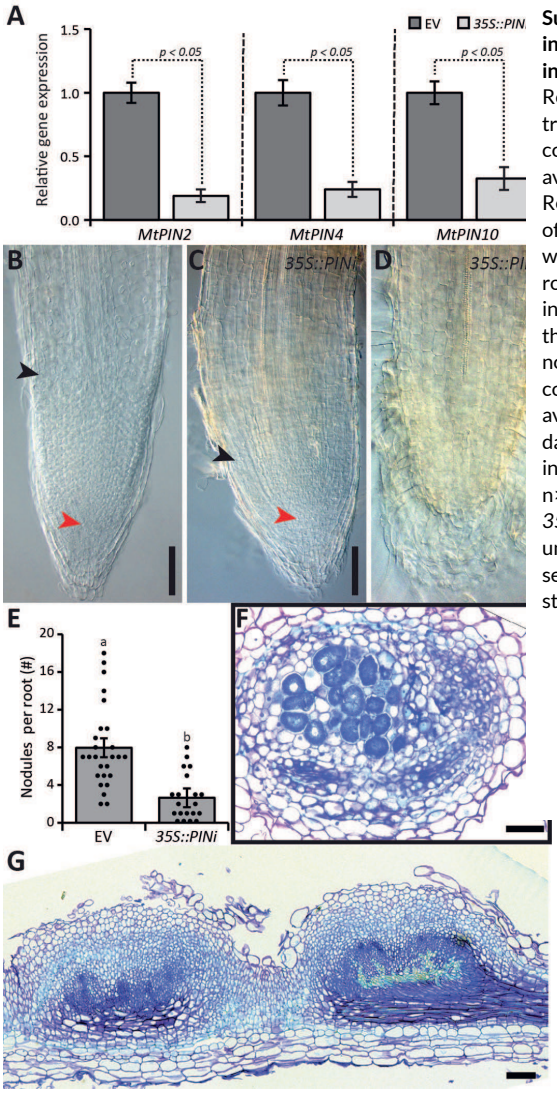
Supplemental Figure S9 | The spatiotemporal expression patterns of *MtPIN6* during nodule primordium formation. Representative images of RNA *in situ* hybridization with *MtPIN6* probe sets on longitudinal sections of the root tip, the susceptible zone and nodule primordia at stage I, III, and V. For stage I, *NF-YA1* was used as a marker for nodule primordium initiation. (Red dots are *MtPIN6* hybridization signals, blue dots are *MtNF-YA1* hybridization signals). C3-C5, cortical cell layers; ed, endodermis; pc, pericycle; scale bars 75 μm.



Supplemental Figure S10 | The expression patterns of *MtLAX2* in the root and nodule primordia at different stages of development in Medicago A17. Representative images of RNA *in situ* hybridization with *MtLAX2* probe set on longitudinal sections of the root tip, the root susceptible zone, and nodule primordia at stages I, II, III, and V (red dots are hybridization signals). C3-C5, cortical cell layers; en, endodermis; pc, pericycle; scale bars 75 µm.



Supplemental Figure S11 | *Mt*PIN10 pattern in the Medicago root. *Mt*PIN10-GFP is accumulated in vasculature and detected in cortical cells of root meristem, it is basal localized there. *Mt*PIN10-GFP level decreased in elongation zone and became undetectable in root susceptible zone (scale bars 75 µm).



Supplemental Figure S12 | Validation of RNA interference on *MtPIN2*, *MtPIN4* and *MtPIN10* in roots of composite Medicago A17 plants. (A) Relative expression of *MtPIN2*, *4* and *10* in transgenic *35S::PIN2/4/10i* (*35S::PINi*) roots compared to *EV* control roots, bars represent averages, ± standard error, n=3. **(B-D)** Representative Nomarski microscopy images of **(B)** an *EV* control root, **(C)** a *35S::PINi* root with a short root meristem and **(D)** a *35S::PINi* root without a root meristem. Arrowhead indicates quiescent center (red) and the end of the root meristem (black). **(E)** Number of nodules formed on *35S::PINi* transgenic roots compared to *EV* control roots, bars represent averages, ± standard error, dots show individual datapoints per transgenic root, different letters indicate significant differences with $p < 0.05$, n>20. **(F-G)** Longitudinal plastic sections of a *35S::PINi* **(F)** nodule with a short and underdeveloped meristem and **(G)** root segment containing nodule-like bumps (both stained with toluidine blue; scale bars 50 µm).

SUPPLEMENTAL TABLES

Supplemental Table S1 | Phenotype of *35S::PIN2/4/10 RNAi* transgenic roots

RNAi composite plants	Total roots analysed	Short roots (<3cm)	Longer roots (>3cm)
EV	87	9	78
35S::PINi	89	62	27

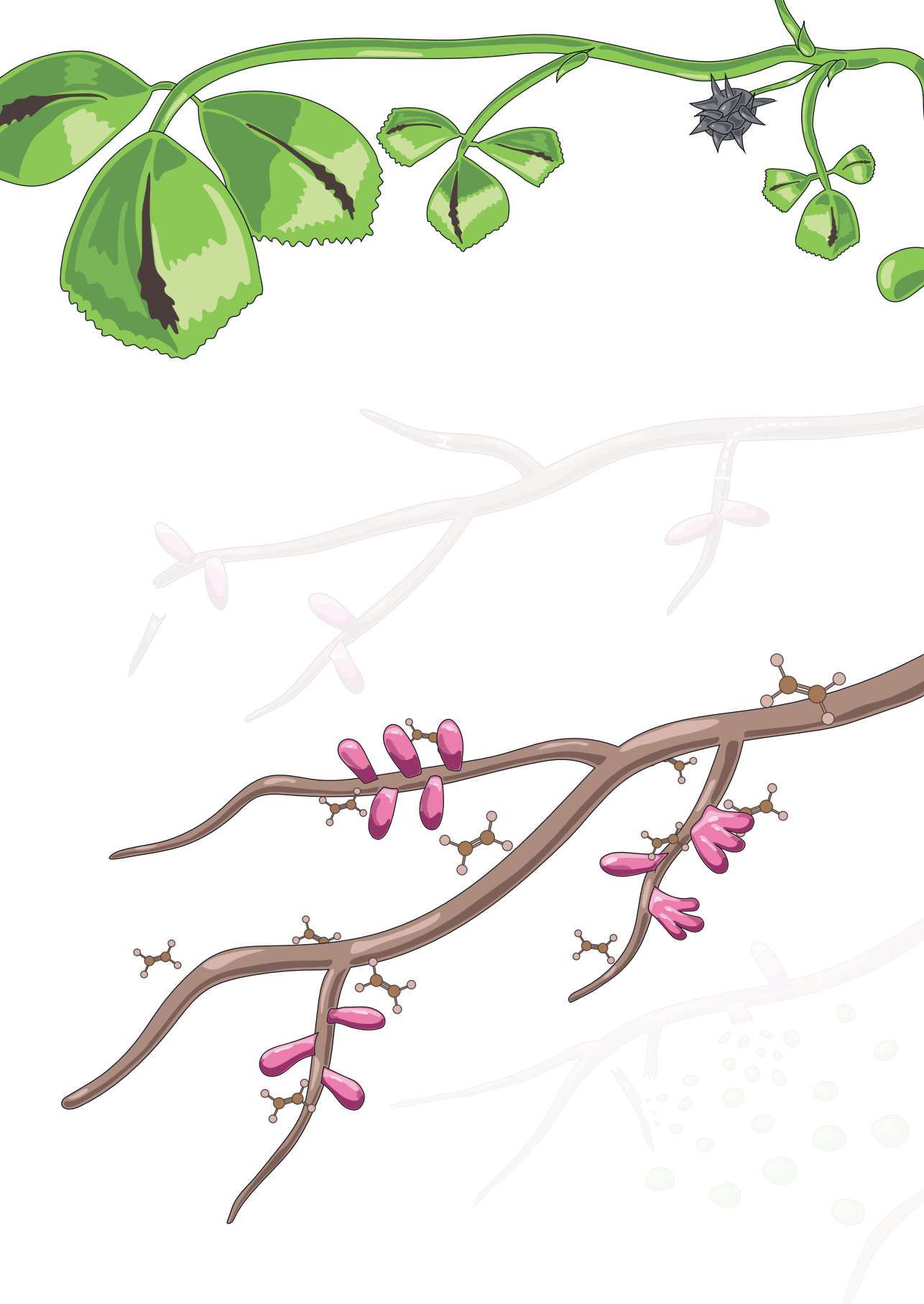
Supplemental Table S2 | List of primers

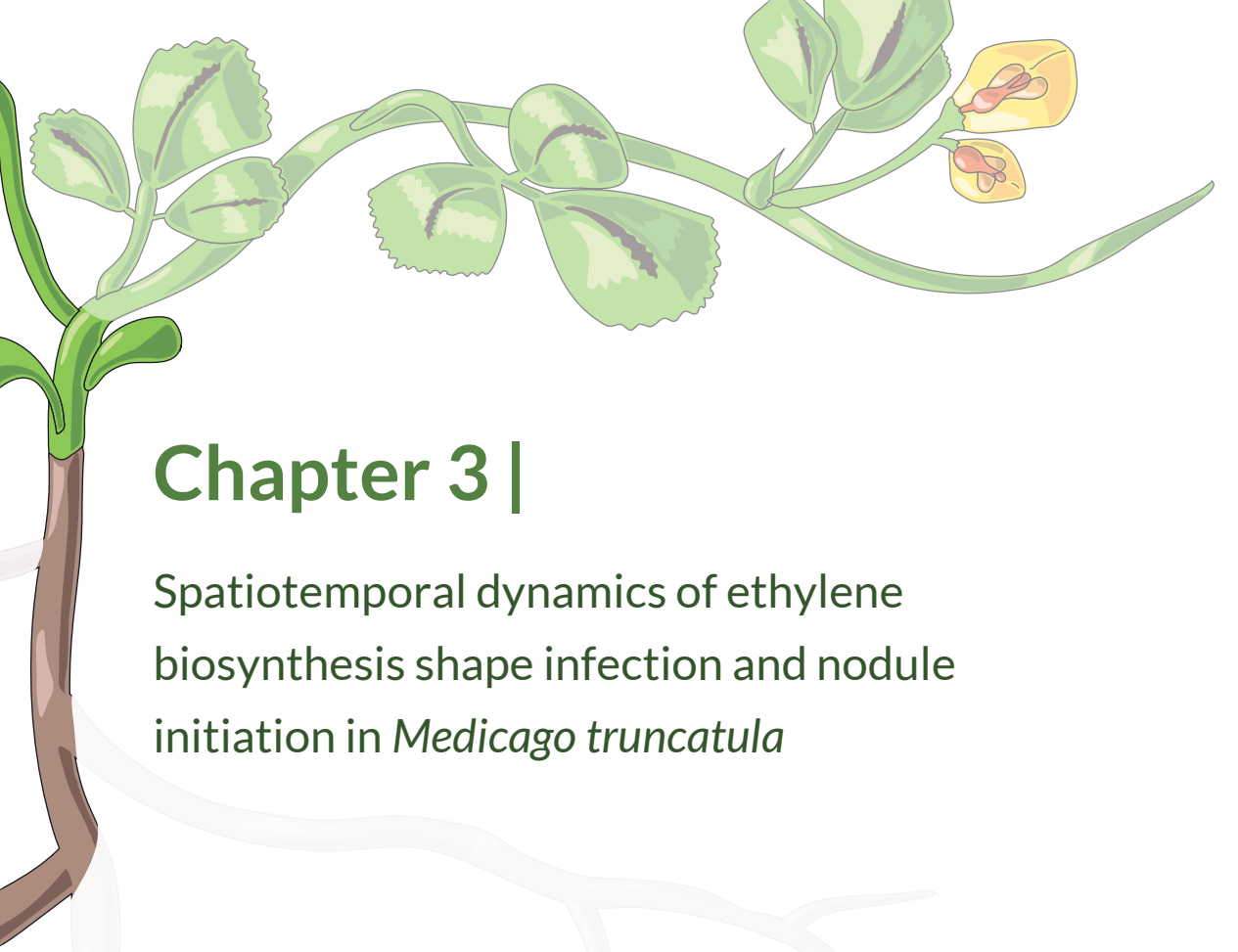
MtPINs protein GFP fusion constructs		
Name	**Sequence**	
MtPIN10-F1	GGGGACAACTTTGTATAGAAAAGTTGCGTGCTCGAATATTAGCCCAAC	
MtPIN10-R1	GGGGACTGCTTTTTTGTACAAACTTGCCATAGCTTCACCTTGAGATCCAG	
MtPIN10-F2	GGGGACAGCTTTCTTGTACAAAGTGGAGCCAACAAACATGCCAC	
MtPIN10-R2	GGGGACAACTTTGTATAATAAAGTTGGTGGCGGAATTCTTC	
eGFP-F	GGGGACAAGTTTGTACAAAAAAGCAGGCTCCCGGGGGTACCTGGTGAGCAAGGGC GAGGA	
eGFP-R	GGGGACCACTTTGTACAAGAAAGCTGGGTACTTGTACAGCTCGTCCATGC	
Primers for MtPINs promoter GFP fusion constructs		
Name	**Sequence**	
pMtPIN4-F2	TCGTGGTCAAGAACGTTCTC	
pMtPIN4-R	GTTTTTGAGTTGAGGTTTGAAGAAG	
pMtPIN10-F	TGCTCGAATATTAGCCCAAC	
pMtPIN10-R	TTTGTTTGGCTTATTGAAGTTTTG	
Primers for qRT-PCR		
Name	**Sequence**	**Gene ID Mt4 / r5.0**
YUC1qF	CTCAACTTGGGTTGGATCGT	Medtr3g109520 / MtrunA17_Chr3g0139441
YUC1qR	CCTTGAGCCAATAGGGAACA	
YUC2qF	GGGTGTGGAAATTCAGGTATGGAG	Medtr6g086870 / MtrunA17_Chr6g0485621
YUC2qR	AGGATGGATGAGCATTATGGTTGC	
YUC8qF	AGACTTTCTCTCACGCCGTTGC	Medtr7g099330 / MtrunA17_Chr7g0262591
YUC8qR	ACCAGATGGACCTGCACCTATG	
MtPIN2_01F	AAGAAAAGGGGAGGGAGGAG	Medtr4g127100 / MtrunA17_Chr4g0071571
MtPIN2_01R	CTTCTTTGTGCCAGCACTTG	
MtPIN4_01F	CTGTCAGATGCAGGACTTGG	Medtr6g069510 / MtrunA17_Chr6g0478431
MtPIN4_01R	CACACCCATTGAAAAAGCTG	
MtPIN10_02F	AAACACCGTTGCTTCTTTCG	Medtr7g089360 / MtrunA17_Chr7g0255941
MtPIN10_02R	CAATGTGCAACAGAACTCCC	

MtACTIN2_01F	TGGCATCACTCAGTACCTTTCAACAG	Medtr2g008050 / MtrunA17_Chr2g0278591
MtACTIN2_01R	ACCCAAAGCATCAAATAATAAGTCAACC	
RNAi constructs for Single MtPINi and MtYUCi		
Name	**Sequence**	
MtPIN2-F	CACCATGATTACCGGTAAGGATATATAC	
MtPIN2-R	CATTTCACCTATCTCTGCATCAG	
MtPIN4-F	CACCCTACCATGTCATGACAGCAATG	
MtPIN4i-R	TCCAATGGTGTTCTACCATCCAATG	
MtPIN10i-F	CACCATACCATGTCCTCACAGCAG	
MtPIN10i-R	ACCATCTTCACCAACTTCAGC	
MtYUC1-F	GATCAGCTAGACGAAACATGCTACCACTC	
MtYUC1-R	GCACTGTCTCATTAAACCGTG	
MtYUC2-F	CACCATGGCACATGATCATCAAAAG	
MtYUC2-R	AACATGCTACCACTCTTATAC	
MtYUC8-F	CACCTCGTCTAGCTGATCACCAAGAC	
MtYUC8-R	TCAAGTGAGAGTTCCATTCCAG	
Double and triple MtPINi and MtYUCi		
Name	**Sequence**	
MtPin4/10-R	AAGTCATGTCCACCAACCATCTTCACCAAC	
MtPin10/4-F	GTTGGTGAAGATGGTTGGTGGACATGACTT	
MtPin2/10-R	GTGAGGACATGGTATCATCTTCACCTATCT	
MtPin10/2- F	AGATAGGTGAAGATGATACCATGTCCTCAC	
MtYUC8/2-F	GAGTGGTAGCATGTTTCGTCTAGCTGATC	
MtYUC2/8-R	GATCAGCTAGACGAAACATGCTACCACTC	
MtYUC1/2-R	ATGATCATGTGCCATTCAAGTGAGAGTTCC	
MtYUC2/1-F	TTAATGAGACAGTGCATGGCACATGATCAT	

Supplemental Table S3 | Gene probe sets used for ViewRNA *in situ* hybridization

Gene name	Gene ID Mt4 / r5.0	Assay ID according to ThermoFisher Scientific
MtNF-YA1	Medtr1g056530 / MtrunA17_Chr1g0177091	VPRWEK2 Type 6
MtYUC1	Medtr3g109520 / MtrunA17_Chr3g0139441	VPKA3EK Type 1
MtYUC2	Medtr6g086870 / MtrunA17_Chr6g0485621	VF1-6000670 Type 1
MtYUC8	Medtr7g099330 / MtrunA17_Chr7g0262591	VF1-6000768 Type 1
MtYUC9	Medtr1g069275 / MtrunA17_Chr1g0182991	VPCE3WM Type 1
MtPIN2	Medtr4g127100 / MtrunA17_Chr4g0071571	VF1-17677 Type 1
MtPIN4	Medtr6g069510 / MtrunA17_Chr6g0478431	VF1-20313 Type 1
MtPIN6	Medtr1g029190 / MtrunA17_Chr1g0159341	VPDJXGJ Type 1
MtPIN10	Medtr7g089360 / MtrunA17_Chr7g0255941	VF1-17681 Type 1
MtLAX2	Medtr7g067450 / MtrunA17_Chr7g0241841	VPZTD33 Type 1





Chapter 3 |

Spatiotemporal dynamics of ethylene biosynthesis shape infection and nodule initiation in *Medicago truncatula*

Sophia Müller, Thijs Stegmann, Kelvin Adema, Rens Holmer, Amber van Seters, Robin van Velzen, Olga Kulikova, Tristan Wijsman, Joel Klein, Josefina-Patricia Fernandez-Moreno, Anna N. Stepanova, Jose M. Alonso, Henk Franssen, Estibaliz Larrainzar, Arjan van Zeijl, Wouter Kohlen

Accepted in The Plant Cell

ABSTRACT

Ethylene is a well-established negative regulator of nodulation, yet how ethylene biosynthesis and perception are spatially coordinated during early symbiotic signaling remains unresolved. Here, we investigate the dynamics of ethylene responses in *Medicago truncatula* using transcriptomics, promoter–reporter analyzes, loss-of-function approaches and a synthetic reporter. We show that the activity of the ethylene-responsive *EBSn* reporter shifts from inner root tissues under non-symbiotic conditions to the outer cortex and epidermis following rhizobial inoculation, revealing a spatial reprogramming of ethylene signaling. Among the eight Medicago *1-AMINOCYCLOPROPANE-1-CARBOXYLIC ACID SYNTHASE* (*ACS*) genes, upon rhizobia application *MtACS3* is induced in outer root cell layers, while *MtACS10* is repressed in the inner cortex and pericycle, mirroring the shift in ethylene perception. Functional analysis demonstrates that *MtACS10* restricts nodule initiation, whereas *MtACS3* modulates infection thread number, prevents nodule clustering, and contributes to radial positioning of nodule primordia. Rhizobial induced ectopic *ACS* expression in the root interior counteracts *MtACS10* repression and blocks nodulation, highlighting the requirement for spatially confined downregulation of ethylene biosynthesis. Together, these findings establish a framework in which localized shift in ethylene biosynthesis, mediated by distinct Medicago *ACS* genes, balances infection and organogenesis while co-defining the spatial limits of the root susceptible zone.

INTRODUCTION

Plants constantly face environmental challenges such as fluctuating water, nutrients, light, temperature, humidity, and biotic stresses. To cope with these dynamic conditions, plants finely tune growth and development through phytohormones, which transmit signals within tissues and activate transcriptional programs (Evert and Eichhorn 2013). Among these adaptations, legumes have evolved a remarkable developmental innovation: the ability to establish nitrogen-fixing symbioses with rhizobia under nitrogen-limiting conditions, forming specialized root organs called nodules (Roy et al. 2020). Rhizobia activate nodule initiation only in a specific region of the early root differentiation zone. This region is therefore referred to as the root susceptible zone (Bhuvaneswari et al. 1980). These nodules provide a controlled environment for bacteria to convert atmospheric nitrogen into ammonium, supplying the plant with essential nutrients at the cost of high metabolic energy (Ferguson et al. 2019; Roy et al. 2020). Because of this energetic expense, nodule formation is tightly regulated by hormonal signals that integrate internal and external cues (Ferguson et al. 2019; Lin et al. 2020).

Nodule organogenesis is initiated upon perception of rhizobium-derived lipochitooligosaccharide (LCO) signals, called Nod factors, which activate key plant transcriptional regulators, including the master symbiotic regulator *NODULE INCEPTION* (*NIN*), and trigger localized accumulation of cytokinins such as isopentenyl adenine (iP) and trans-zeatin (tZ) (Breakspear et al. 2014; Larrainzar et al. 2015; Limpens et al. 2015; van Zeijl et al. 2015b; Reid et al. 2017; Gühl et al. 2021). These signals collectively induce cell divisions in the root pericycle and cortex, generating a nodule primordium (Timmers et al. 1999; Xiao et al. 2014). Concurrently, LCO perception initiates a rhizobial infection process that begins in responsive root hairs. Root hair curling traps rhizobia in infection pockets, from which cell wall-bound infection threads guide bacteria to the developing nodule primordium. Inside the nodule, rhizobia are released and develop into symbiosomes: host membrane-bound compartments that facilitate nutrient exchange between plant and microbe (Roy et al. 2020; Adema and Kohlen 2024).

In the model legumes *Medicago truncatula* (Medicago) and *Lotus japonicus* (Lotus), ethylene-insensitive mutants (*sickle*/*Mtein2*) display hyper-infection and excessive numbers of, often clustered, nodules (Penmetsa and Cook 1997; Oldroyd et al. 2001; Penmetsa et al. 2008; Miyata et al. 2013), showing that ethylene is a key regulator in these processes. In contrast to auxins and cytokinins, ethylene acts as a negative regulator of nodule initiation (Penmetsa and Cook 1997; Gonzalez-Rizzo et al. 2006; Penmetsa et al. 2008; Guinel 2015; van Zeijl et al. 2015b; Xiao et al. 2025). Additionally, in Medicago, several *1-AMINOCYCLOPROPANE-1-CARBOXYLIC ACID SYNTHASE (ACS)* genes are transcriptionally induced following LCO perception (Larrainzar et al. 2015; van Zeijl et al. 2015b; Herrbach et al. 2017), and elevated ethylene production has been reported in alfalfa (*Medicago sativa*) and Lotus within hours of inoculation (Ligero et al. 1986, 1987; Reid et al. 2018).

Beyond nodulation, ethylene is widely recognized for roles in plant development and stress responses. These roles include promoting fruit ripening, root growth inhibition, shoot thickening, leaf senescence, and defence against pathogens (Verma et al. 2016; Dubois et al. 2018; Salvi et al. 2021). In *Arabidopsis thaliana* seedlings, ethylene induces the classical "triple response"; a phenotype characterized by inhibition of hypocotyl and root elongation, radial organ swelling, and exaggerated apical hook formation (Knight et al. 1910; Merchante and Stepanova 2017). The study of this response has been instrumental in identifying key components of the ethylene biosynthesis and signaling pathway (Binder 2020).

Ethylene biosynthesis is tightly regulated, possibly due to its rapid diffusion. The pathway begins with methionine, which is converted to S-adenosylmethionine (SAM) by SAM SYNTHASE. SAM is then converted into 1-aminocyclopropane-1-carboxylic acid (ACC) in the cytosol by ACC SYNTHASE (ACS), often considered the rate-limiting step of ethylene biosynthesis (Kende 1993). ACC is subsequently oxidized to ethylene by ACC OXIDASE (ACO), but it has been suggested that this step can also occur non-enzymatically via free oxygen (Depaepe and Van Der Straeten 2020; Pattyn et al. 2021). Additionally, ethylene production can be modulated by the catabolism of ACC to α-ketobutyrate via ACC DEAMINASE (ACD) activity (Plett et al. 2009).

Ethylene perception occurs at the endoplasmic reticulum (ER) membrane through a family of ethylene-specific receptors. In Arabidopsis these include ETHYLENE RECEPTOR1 (ETR1), ETR2, ETHYLENE RESPONSE SENSOR1 (ERS1), ERS2, and ETHYLENE INSENSITIVE4 (EIN4), which can form both homodimers and heterodimers (Gao et al. 2008). The ethylene-binding domain faces the ER lumen, while the cytosolic side interacts with protein complexes containing the negative regulator CONSTITUTIVE TRIPLE RESPONSE 1 (CTR1) (Gao et al. 2008; Shakeel et al. 2013). In the absence of ethylene, CTR1 phosphorylates the C-terminal end of the highly conserved ETHYLENE INSENSITIVE2 (EIN2), maintaining the pathway in an inactive state (Ju et al. 2012; Bowman et al. 2019). Ethylene binding triggers conformational changes that release CTR1, allowing EIN2 dephosphorylation, cleavage, and nuclear translocation of its C-terminal fragment. In the nucleus, EIN2-C activates transcription factors ETHYLENE INSENSITIVE3 (EIN3) and EIN3-LIKE1 (EIL1), which in turn regulate second-tier transcription factors, ETHYLENE RESPONSE FACTORs (ERFs), propagating transcriptional responses that mediate diverse ethylene-dependent phenotypes (Qiao et al. 2012; Wen et al. 2012; Ma and Dong 2021). Loss-of-function mutations in *EIN2* cause near-complete ethylene insensitivity, highlighting the central role of the EIN2 protein in the pathway (Guzmán and Ecker 1990).

The small and gaseous nature of ethylene poses major analytical challenges, complicating its detection and hindering detailed functional characterization. Unlike other key phytohormones such as auxin, abscisic acid, and gibberellic acid (Rizza et al. 2017; Herud-Sikimić et al. 2021; Rowe et al. 2023), ethylene currently lacks a dynamic *in vivo* biosensor. Moreover, ACC-based sensors do not directly measure ethylene activity and therefore cannot distinguish canonical ethylene signaling from ethylene-independent roles of ACC, which has been shown to act as a growth regulator in its own right (Van de Poel and Van Der Straeten 2014). Nonetheless,

ethylene responses have been effectively monitored in Arabidopsis using *35S::EIN3-GFP* and *35S::EIL1-GFP* reporter lines to track ethylene-dependent stabilization of EIN3/EIL1 proteins (Yanagisawa et al. 2003; An et al. 2010; Ju et al. 2015), as well as the *EBS::GUS* transcriptional reporter to visualize EIN3-mediated transcriptional activation in specific tissues (Stepanova et al. 2007). The *EBS* reporter has since been refined and optimized in subsequent studies (Fernandez-Moreno et al. 2024, 2025). Attempts to apply previous versions of the *EBS* reporter in Medicago have been unsuccessful, and without functional ethylene sensors in this species, the spatiotemporal basis of ethylene-mediated repression of nodulation remains unresolved. The recently developed *EBSn* (*EIN3-Binding Site New*) reporter, which exhibits strong ethylene responsiveness in Arabidopsis and *Solanum lycopersicum* (Fernandez-Moreno et al. 2025), holds considerable promise for extending such analyzes to Medicago and other legumes.

In Medicago, early transcriptional activation of *ACC synthases* (*ACS*) following rhizobial signal perception suggests that ethylene biosynthesis is dynamically reprogrammed during nodule organogenesis (Larrainzar et al. 2015; van Zeijl et al. 2015b). However, while ethylene is recognized as a negative regulator of nodulation, the mechanisms governing its spatiotemporal regulation in roots remain largely unknown. Here we reveal that early symbiotic signaling in Medicago triggers a dynamic spatial reorganization of ethylene biosynthesis, marked by a shift in *EBSn* promoter activity from inner to outer root layers following rhizobial inoculation. We show that among the ACC synthases, *MtACS3* and *MtACS10* exhibit antagonistic transcriptional regulation and distinct spatial expression domains, with *MtACS3* upregulated in the outer cell layers of the root and *MtACS10* downregulated specifically in cells destined to form nodule primordia. Functional analyzes show that loss of *MtACS10* increases nodule number, whereas loss of *MtACS3* leads to fused nodules, affects infection thread formation, and disrupts radial nodule positioning. Moreover, we demonstrate that ectopically maintaining *MtACS10* expression in the root interior suppresses nodulation. Additionally, we report that ethylene signaling defines the size and competence of the root susceptible zone, as seen in the ethylene-insensitive *sickle/Mtein2* mutant, which displays expanded symbiotic susceptibility along the root. Together, these findings indicate that localized ethylene biosynthesis and its transcriptional control by *ACS* genes finely coordinate where and how nodulation occurs in Medicago, revealing a highly dynamic spatiotemporal regulation of ethylene potential.

RESULTS

The *EBSn* promoter is functional in the Medicago root and shows a spatial shift in activity during early symbiotic signaling

To investigate the spatial dynamics of ethylene signaling during early rhizobial interactions, we generated composite Medicago plants with transgenic roots expressing the synthetic ethylene-responsive reporter *EBSn::GUS* (Fernandez-Moreno et al. 2025). Composite plants have normal shoots and partially transgenic roots grown from a callus—used to study root-specific gene functions and phenotypes (Limpens et al. 2004). Under non-symbiotic (mock) conditions, strong *EBSn* promoter activity was observed in the inner tissues of the root susceptible zone (**Figure 1A**). Upon inoculation with *Sinorhizobium meliloti* 2011 (*Sm2011*), *EBSn* activity shifted within 24 hours to the outer root layers (**Figure 1B**). At 48 hours post inoculation (HPI), promoter activity was mainly restricted to the epidermis and developing nodule primordia (**Figure 1C**).

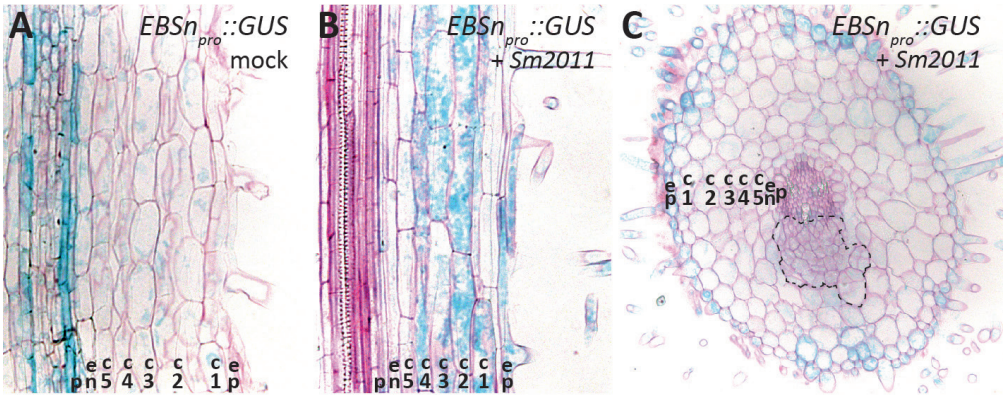


Figure 1 | Ethylene responses in the root susceptible zone of Medicago during rhizobia-induced signaling. (A–C) Activity of the synthetic ethylene-responsive reporter *EBSn::GUS* **(A)** after 24 h mock treatment, **(B)** at 24 h post-inoculation (24 hpi) with *Sm2011* (*Sinorhizobium meliloti* 2011), and **(C)** in a root cross-section at 48 hpi. Dotted line indicates developing nodule primordia, ep, epidermis;c1-c5, cortical cell layers; en, endodermis; pc, pericycle.

Given that ethylene is a gaseous phytohormone assumed to be capable of diffusing across tissues, we hypothesized that this distinct spatiotemporal pattern of *EBSn* activity reflects either a direct diffusion pattern or a shift in localized ethylene biosynthesis. Ethylene is synthesized via the sequential action of ACS and ACO (Kende 1993) (**Figure 2A**). Previous phylogenetic analysis of the Medicago genome revealed five *ACO*, one *ACD* and eight *ACS* genes (**Figure 2B-D, Supplemental File S1, Supplemental Table S1**, Gómez-Fernández et al. 2025). To explore candidate biosynthesis genes possibly involved in the observed ethylene dynamics, we performed RNA-seq analysis on the root susceptible zone (a 5 mm region starting 5 mm from the root tip) treated with either mock or LCOs. In pea (*Lathyrus oleraceus*, formerly *Pisum sativum*), *PsACO* expression has been reported to correlate with nodule initiation (Heidstra et al. 1997). For this reason, we first focused on the Medicago *ACO* genes. In our system (FA medium,

0.25 mM NO_3^-, no addition of aminoethoxyvinylglycine (AVG)), several *MtACO* genes exhibited robust expression, including *MtACO1* (~20 TPM), *MtACO2* (~160 TPM), *MtACO3* (~390 TPM), and *MtACO4* (~45 TPM), while *MtACO5* expression was below 1 TPM (**Figure 2E**). In addition, *MtACD1* was expressed at approximately 50 TPM (**Figure 2F**). Apart from a reduction in *MtACO4* expression, none of these genes were significantly affected by early LCO-induced signaling (**Figure 2E-F**).

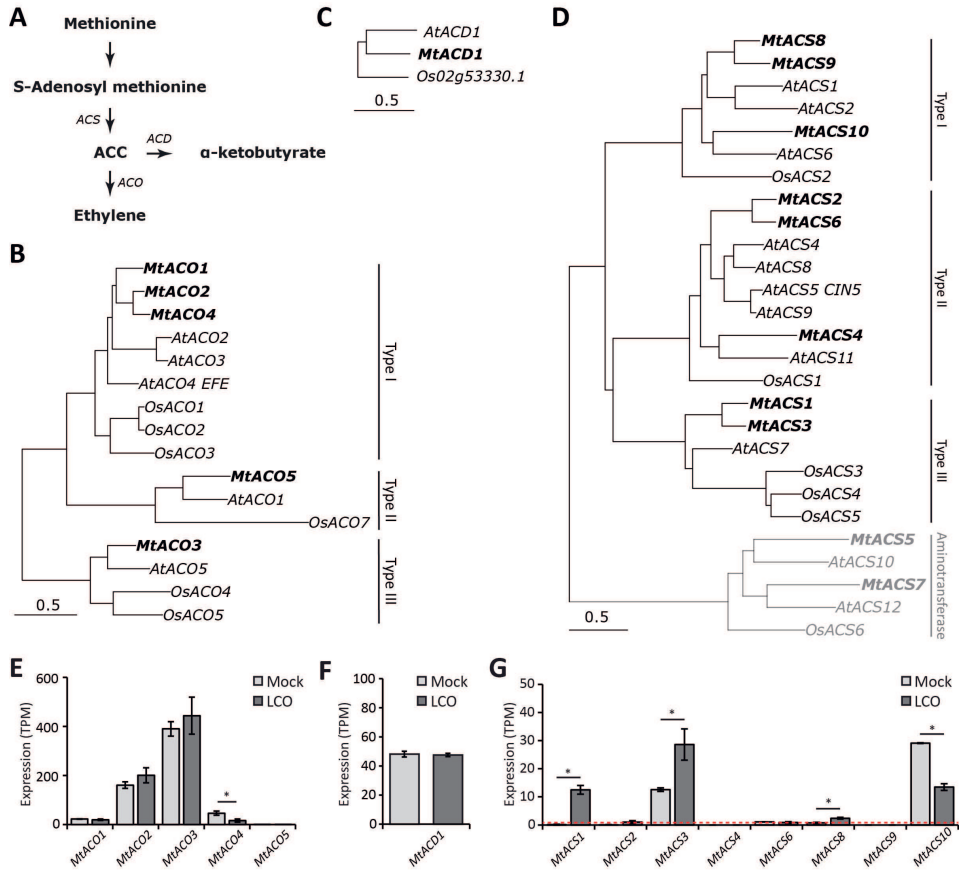


Figure 2 | Identification and analysis of Ethylene biosynthesis genes in Medicago. (A) schematic representation of the ethylene biosynthesis pathway in plants, **(B-D)** Summary of a phylogenetic analysis of **(B)** ACC Oxidase **(C)** ACC deaminase and **(D)** ACC synthase amino acid sequences (*Mt* = *Medicago truncatula*, *At* = *Arabidopsis thaliana*, *Os* = *Oryza sativa*, for full phylogenetic trees including gene IDs see Supplemental file S2, in grey putative aminotransferases previously misidentified as ACC synthases (for review see Gómez-Fernández et al. 2025); scale bars indicate substitutions per site. **(E-G)** Expression of ethylene biosynthesis genes under control and symbiotic (LCO-treated) conditions at 3 h post-LCO application. **(E)** *MtACO1*, *2*, *3*, *4*, and *5*; **(F)** *MtACD*. **(G)** *MtACS1*, *2*, *3*, *4*, *6*, *8*, *9*, and *10*; Bars represent means ± SD (n = 2). Asterisks indicate significant differences between treatments (P < 0.05). Dotted line in **(E)** represents the used expression confident threshold (1 TPM), data available at the European Nucleotide Archive (ENA) under accession number PRJEB38785.

Next, we analyzed the expression of the *ACS* genes. This revealed that under mock conditions, only *MtACS3* and *MtACS10* were expressed well above our expression confidence threshold of 1 TPM (~12.5 and ~29 TPM, respectively;Fig**ure 2G**). Except for *MtACS6*, which was expressed at 1.07 TPM, all remaining *ACS* genes were expressed at levels well below 1 TPM (Figure 2G). In contrast to the *MtACO* and *MtACD* genes, several of the *MtACSs* were differentially regulated by LCO signaling. Specifically, *MtACS1* (~0 TPM to ~12 TPM, mock vs LCO treatment respectively), *MtACS3* (~12.5 to ~28 TPM), and *MtACS8* (~0 TPM to ~3 TPM) were upregulated, whereas *MtACS10* was downregulated (~29 to 13.4 TPM) (**Figure 2G**). These results are in line with changes in the expression previously observed for these genes (**Supplemental Figure S1**, Larrainzar et al. 2015; van Zeijl et al. 2015b; Schiessl et al. 2019).

MtACS3 and *MtACS10* exhibit differential transcriptional regulation and spatial expression patterns in response to LCO signaling

Assuming that genes involved in the regulation of susceptibility need to be expressed prior to inoculation, *MtACS3* and *MtACS10* represent interesting candidates, as they were the only *ACS* genes expressed well above our confidence threshold in the *Medicago* root susceptible zone before nodulation and both responded transcriptionally to LCO signaling. Therefore, we focused our subsequent analyzes on these two genes. An independent qRT-PCR experiment confirmed the antagonistic regulation of the *MtACS3* and *MtACS10* expression (**Figure 3A, B**), a pattern that we also identified after re-analysis of a published dataset **(Supplemental Figure S1A**, Schiessl et al. 2019). This convergence across independent studies strengthens the robustness of the observation and indicates that antagonistic regulation of these genes is a consistent feature of the early nodulation response.

Moreover, we found that the cumulative expression of *MtACS3* and *MtACS10* was similar between LCO- and mock-treated samples (**Supplemental Figure S2A**). This correlated with the observation that no significant differences in ACC levels could be detected in bulk samples from the root susceptible zone 3 hours after LCO application (**Supplemental Figure S2B**). These findings suggest that differential regulation of *MtACS3* and *MtACS10* can only influence nodule initiation if these two genes are expressed in distinct temporospatial domains.

To investigate this, we first determined the spatial expression patterns of *MtACS3* and *MtACS10*. To this end, we created composite plants bearing transgenic roots in which expression of the *GUS* reporter was driven by the putative promoter of either *MtACS3* or *MtACS10*. Analysis of the root susceptible zone revealed that, under mock conditions, the *MtACS3*$_{pro}$*::GUS* reporter construct is mostly inactive in the root susceptible zone (**Figure 3C**). As the first cell divisions have been reported to occur at 24 HPI (Xiao et al. 2014), we decided to analyze the effect of rhizobia spot inoculation just prior to this event. At approximately 20 HPI, the *MtACS3*$_{pro}$*::GUS* promoter construct is activated in the vasculature, outer cortical cells, and the epidermis (**Figure 3D**).

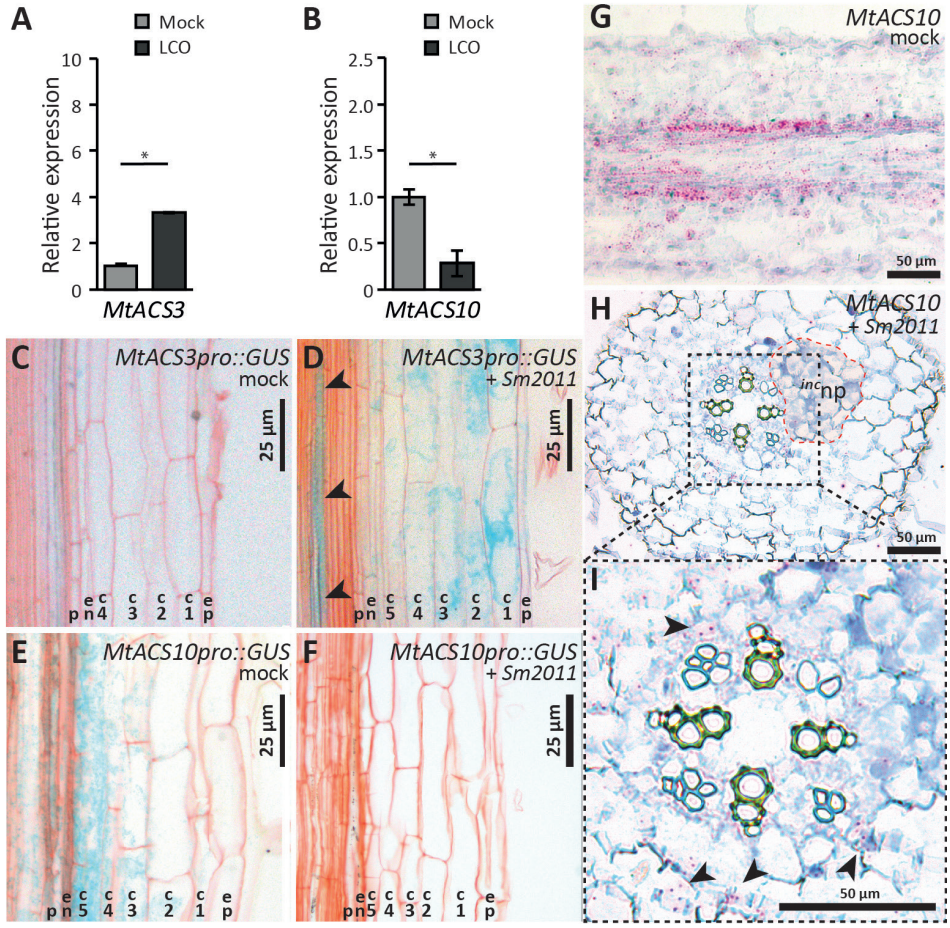


Figure 3 | Expression Analysis of *MtACS3* and *MtACS10* in Medicago roots under symbiotic and control conditions. **(A, B)** Expression levels of **(A)** *MtACS3* and **(B)** *MtACS10* following mock and LCO application, measured at 3 hours post-treatment. Bars represent mean ± SE (n=3). Asterisks (*) indicate significant differences between treatments ($P < 0.05$). (C-F) *MtACS3* and *MtACS10* promoter activity and expression domains under mock and rhizobia inoculation. (**C, D**) *MtACS3pro::GUS* reporter activity at 20 hours post **(C)** mock or **(D)** *Sm2011* (*Sinorhizobium meliloti 2011*) spot-application, Arrows indicating GUS expression in the vasculature after **(D)** *Sm2011* treatment. **(E, F)** *MtACS10pro::GUS* activity at 20 hours post **(E)** mock or **(F)** *Sm2011* spot-application. **(G, I)** *MtACS10* expression domains in the pericycle and inner cortex as the result of RNA *in situ* hybridization. **(G)** Longitudinal section of an untreated Medicago wild-type (A17) root susceptible zone. **(H-I)** Cross-section of a root susceptible zone 36 hours post *Sm2011* application. **(H)** highlighting the position of an incipient nodule primordium (^{inc}np) highlighted by the dotted red line, **(I)** Zoom in of vasculature and inner root cell layers. Hybridization signals are absent in reactivated pericycle and cortical cells of the incipient primordium. Hybridization signals appear as red dots (arrowheads in I), ep, epidermis;c1-c5, cortical cell layers; en, endodermis; pc, pericycle in C-F.

In addition, the *MtACS10*$_{pro}$*::GUS* reporter construct is active in the pericycle, endodermis, and inner cortical cell layers under non-inoculated conditions (**Figure 3E**). When treated with *Sm2011*, the activity of the *MtACS10*$_{pro}$*::GUS* reporter was abolished in all interior root cells close to the site of application of the root susceptible zone (**Figure 3F**). Taken together, these reporter line observations suggest a spatial shift in the domain of *ACS* expression, and possibly ethylene biosynthesis.

To further investigate the cellular localization of *MtACS3* and *MtACS10* transcripts, we performed RNA *in situ* hybridization. RNA *in situ* hybridization against *MtACS3* revealed no signal in the root susceptible zone, although a *MtACS3* hybridization signal was detected in the epidermis and outer cortex at the elongation zone and start of the root differentiation zone (**Supplemental Figure S3**). This suggests that the observed *MtACS3* gene expression prior to LCO application might be due to a co-sampling of the expression domain at the start of the root differentiation zone, just below the susceptible zone. Transcripts of *MtACS10* were predominantly detected in the pericycle and inner cortex of the elongation and untreated root susceptible zone (**Figure 3G,**Supplemental Figure **S3**). Transverse sections of a 36 HPI *Sm2011*-treated root susceptible zone confirmed expression in a specific domain of inner cortical cells, surrounding the stele (**Figure 3H, I**, arrowheads in I). However, *MtACS10* mRNA was excluded from the central vascular and incipient nodule primordia (**Figure 3H, I**). Combined with the $MtACS10_{pro}$*::GUS* reporter data, this suggests that upon inoculation with rhizobia, *MtACS10* is removed from the cells that will initiate division to form a nodule primordia. Although we were unable to detect *MtACS3* expression in the epidermis, results published by Breakspear at al., (2014) show *MtACS3* to be upregulated by both LCO and bacteria application in their root hair samples (**Supplemental Figure 1B**). Taken together, the observed *EBSn* pattern could be explained by *MtACS10* basal activity in the inner root tissues, followed by *MtACS3* induction in the outer cortex and epidermis after inoculation, suggesting that localized activity of these two *ACS* genes sequentially shapes the spatial dynamics of ethylene signaling during nodule initiation in Medicago roots.

Different roles of MtACS3 and MtACS10 in the regulation of nodule development and infection thread formation in Medicago

Next, to investigate the functional role of MtACS3 and MtACS10 during nodulation, we generated composite plants with transgenic roots expressing RNA interference constructs targeting either *MtACS3* ($ACS3^i$) or *MtACS10* ($ACS10^i$). RNAi-mediated silencing specifically reduced *MtACS3* and *MtACS10* transcript levels by approximately 50% and 75%, respectively, without off-target effects on any of the other Mt*ACS* genes, nor the closely related putative aminotransferases (**Supplemental Figure S4**).

To assess whether knockdown of *MtACS3* or *MtACS10* was sufficient to alter nodulation, we inoculated these plants with *Sm2011*. At four weeks post inoculation (WPI), $ACS10^i$ transgenic roots developed approximately twice as many nodules as either the empty vector control (EV) or the non-transgenic roots harvested from the same composite plants (**Figure 4A, Supplemental Figure S5**). In contrast, $ACS3^i$ did not exhibit a change in nodule number; however, transgenic roots in $ACS3^i$ composite plants frequently formed fused nodules, a phenotype not often observed in EV nor $ACS10^i$ roots (**Figure 4B**).

We next screened the *Medicago* Tnt1 retrotransposon insertion collection for mutants with disruptions in *MtACS3* or *MtACS10* (Tadege et al. 2008). We identified an insertion line in exon 2 of *MtACS3* and exon 1 of *MtACS10*, and designated them *Mtacs3-1* (NF1952) and *Mtacs10-1* (NF15254) (**Supplemental Figure S6**).

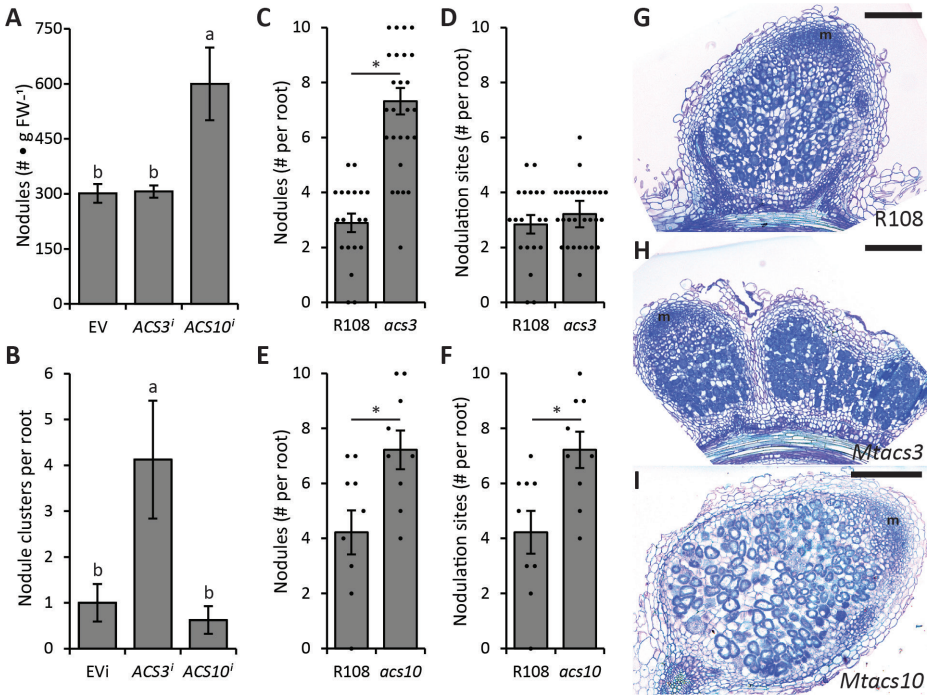


Figure 4 | Nodulation phenotypes of *MtACS3* and *MtACS10* knockdown or knockout mutants. (A-B) Effect of *ACS3*^{i} or *ACS10*^{i} knockdown on nodulation (n>10). **(A)** Average number of root nodules per gram FW. **(B)** Average number of nodule clusters per root system. Transformed and untransformed roots were separated based on *DsRed* expression (for untransformed roots, see **Supplemental Figure S5**). Bars represent mean ± SE; different letters or an asterisk (*) indicate significant differences ($P < 0.05$). **(C-F)** Effect of *ACS3* and *ACS10* knockout mutations on nodulation. **(C)** Average number of root nodules formed on R108 and *Mtacs3-1* (*acs3*) mutants (n>19). **(D)** Number of nodule initiation sites per root (n>19). **(E)** Average number of root nodules formed on R108 and *Mtacs10-1* (*acs10*) mutants (n>8). **(F)** Number of nodule initiation sites per root (n>8). **(G-I)** Representative images of nodules formed on **(G)** wild-type R108, **(H)** *Mtacs3-1* mutant, or **(I)** *Mtacs10-1* mutant. Scale bars: 300 µm; m, nodule meristem.

When grown on plates, *Mtacs3-1* roots were significantly longer than those of wild-type R108, a phenotype reminiscent of the *sickle* root, and not observed in *Mtacs10-1* (**Supplemental Figure S7A, B**) (Penmetsa and Cook 1997).

To evaluate nodulation in these mutants, *Mtacs3-1* and *Mtacs10-1* plants were inoculated with *Sm2011* and analyzed at 2 WPI. The *Mtacs3-1* mutant formed a substantial number of nodule clusters (**Figure 4C**). Although total nodule number increased ~2.5-fold in *Mtacs3-1*, the number of distinct nodulation sites per root was comparable to wild-type R108 (**Figure 4D**). This is in agreement with *ACS3i* data on the increase of nodule clusters and suggests a reduced ability to control nodule numbers during nodule initiation. In agreement with the *ACS10*^{i} phenotype, *Mtacs10-1* mutant plants exhibited a ~90% increase in nodule number compared to wild-type

R108 (**Figure 4E**), and, proportional to R108, only occasionally produced fused nodules when grown on plates. Both the number of nodules and the number of nodulation sites were similarly elevated in *Mtacs10-1* relative to wild type (**Figure 4E, F**). Aside from nodule clusters in *Mtacs3-1*, nodules in both mutants appeared morphologically similar to wild-type R108 nodules (**Figure 4G–I**).

Previous studies reported that the ethylene-insensitive *sickle* mutant forms a greater number of infection threads compared to wild-type plants (Penmetsa and Cook 1997). The role of ethylene during root hair infection was further supported by findings that treatment with the ethylene biosynthesis inhibitor AVG increased the number of infection threads per plant (Oldroyd et al. 2001). As we observed relatively high GUS activity in the epidermis at 48 HPI with *Sm2011* during our *EBSn* reporter assays (**Figure 1**), we hypothesized that the previously reported effects of ethylene on limiting surplus infections may be attributed to an increase in local ethylene biosynthesis in the outer cell layers of the root susceptible zone. To investigate whether *MtACS3* and/or *MtACS10* contribute to this response, we inoculated wild-type R108, *Mtacs3-1*, and *Mtacs10-1* lines with a *Sm2011* strain constitutively expressing GFP, and quantified infection thread (IT) formation at two days post inoculation (DPI). On average, we detected two ITs per root in both R108 and *Mtacs10-1*. In contrast, *Mtacs3-1* roots exhibited approximately a fourfold increase in the number of ITs compared to wild-type R108 (**Figure 5**).

In summary, *MtACS10* is expressed in the inner cortex pre-inoculation, and its mutant forms more nodules but a similar number of infection threads compared to wild-type R108. By contrast, *MtACS3* is induced in the outer root layers post-inoculation and likely functions in the outer cortex and epidermis, and its mutant shows both increased nodule numbers and elevated infection thread formation. Combined, the differential expression of *MtACS3* and *MtACS10* suggests an interesting, less straightforward, mechanistic link between infection threads and nodule primordium initiation that warrants further investigation.

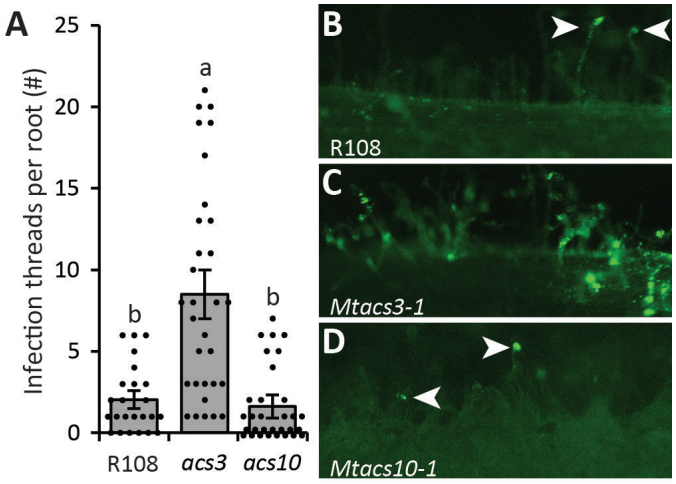


Figure 5 | Effect of *Mtacs3-1* and *Mtacs10-1* loss-of-function mutations on infection thread formation. (A) Quantification of infection threads per spot-inoculation site following inoculation with GFP-labeled *Sm2011* (Sm2011-GFP) on wild-type R108, *Mtacs3-1* (*asc3*), *Mtacs10-1* (*acs10*). Data represent mean ± SE; different letters indicate statistical differences ($P < 0.05$) (B–D) Representative confocal images of Medicago (R108) roots at 2 days post-inoculation with *Sm2011*-GFP. (B) Wild-type R108, (C) *Mtacs3-1* mutant, and (D) *Mtacs10-1* mutant. Arrowheads indicate individual infection threads in (B) and (D).

Loss off *MtACS3*, but not *MtACS10*, affects radial positioning of Medicago nodule primordia

Previous studies in pea reported that nodules form almost exclusively opposite the xylem poles, and that this spatial positioning is at least partially regulated by ethylene (Heidstra et al. 1997). A similar effect has been described for Medicago, though with slightly more variability: approximately 80% of nodules form opposite the xylem poles, while around 20% are located opposite the phloem poles (Penmetsa et al. 2003). Furthermore, the regulatory influence of ethylene appears to be more pronounced in Medicago, as the *sickle* mutant displays an approximately equal distribution of nodule primordia between xylem and phloem poles (Penmetsa et al. 2003).

To determine whether local ethylene biosynthesis by MtACS3 or MtACS10 contributes to radial positioning of nodules along the root axis, we grew wild-type R108, *Mtacs3-1*, and *Mtacs10-1* lines in perlite inoculated with a *Sm2011* strain constitutively expressing GFP. After 7 days, plants were harvested and the GFP signal was used to identify and isolate early rhizobia infection sites. This resulted in the collection of roughly 20–25 root segments per line containing early nodulation events. Root segments were fixed, embedded in plastic, and sectioned for analysis.

Although the central vasculature of Medicago roots can occasionally form a fourth xylem pole, all roots in this experiment developed the typical triarch configuration. In this arrangement, the three xylem poles are spaced approximately 120° apart, with a phloem pole located midway between each pair, at ~60° intervals (**Figure 6**). We sectioned all root segments and successfully traced 18–20 nodule primordia per genotype. Using ImageJ, we drew a line from the central vascular bundle through a xylem pole and a second line from the center to the midpoint of the dividing cortical cells forming the primordium. The angle between these two lines was then calculated and categorized into six bins: 0–10°, 10–20°, 20–30°, 30–40°, 40–50°, and 50–60° (**Table 1**). Primordia with angles <30° were classified as associated with the xylem pole, while those >30° were considered phloem associated.

Table 1 | Angle distribution of the nodule primordia towards the xylem poles.

Line	BIN1 (0-10˚)	BIN2 (10-20˚)	BIN3 (20-30˚)	BIN4 (30-40˚)	BIN5 (40-50˚)	BIN6 (50-60˚)
R108	13	5	2	0	0	0
Mtacs3-1	5	2	4	3	3	1
Mtacs10-1	12	3	5	0	0	0

In wild-type R108, all nodule primordia were positioned opposite xylem poles, with the majority falling within the 0–10° range, averaging the angle between the xylem and nodule primordia at 8.7° (**Figure 6A-C, Table 1**). In contrast, *Mtacs3-1* roots displayed a more randomized spatial distribution of nodule primordia, leading to an average angle of roughly 25° (**Figure 6A, B, D, E**). In some cases, the primordium appeared to originate at a xylem pole but expanded across a broader radial domain (**Figure 6D**), while in others, a clearly defined infection thread was observed initiating opposite the phloem pole (**Figure 6E**, Arrowhead). *Mtacs10-1* roots, similar to R108, exhibited a strong preference for xylem pole-associated nodules, averaging the angle between the xylem and nodule primordia at 10.1° (**Figure 6A, B, F**), demonstrating that the ethylene restriction on the radial positioning of nodules is co-depending on MtACS3, not MtACS10.

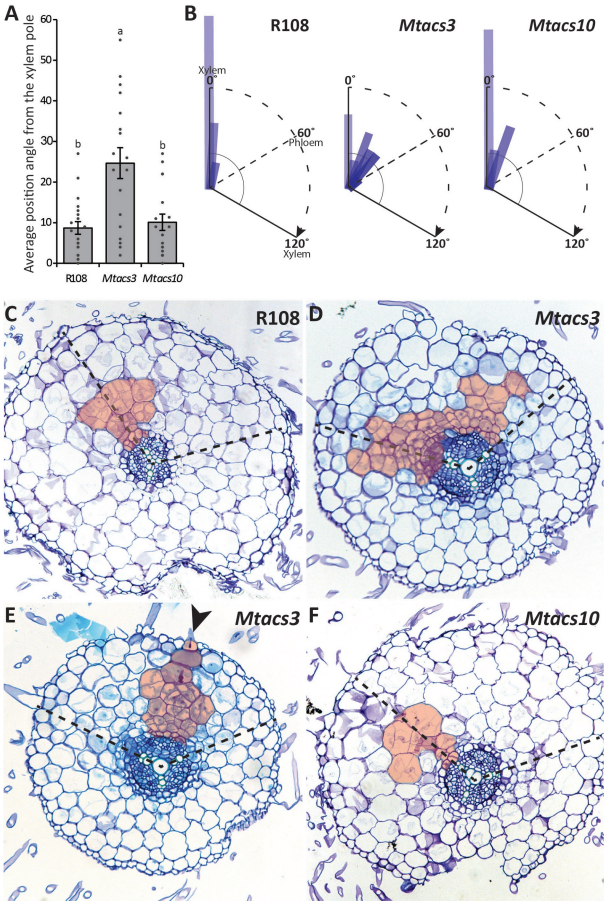


Figure 6 | Nodule position in relation to xylem and phloem poles. (A) Average position angle of the nodule primordia from the xylem pole **(B-D)** schematic representation of one-third of the Medicago vasculature triarch, with a 120° angle between two xylem poles and the phloem in between at 60° showing **(B)** the relative positioning of nodule primordia at 3 days post-inoculation in wild-type R108, *Mtacs3-1*, and *Mtacs10-1* mutants (n>18). **(C-F)** Representative cross-sections of young nodule primordia (highlighted in red) 3 days after inoculation with a GFP-labeled *Sm2011* strain (*Sm2011-GFP*), formed on **(C)** wild-type *R108*, **(D-E)** *Mtacs3-1* and **(F)** *Mtacs10-1* mutants. Dotted lines indicate the two closest xylem poles. Arrowhead in E points at infection thread.

Ectopic expression of *MtACS3* or *MtACS10* under control of the *MtCRE1* promoter leads to reduced nodulation

Next, we investigated whether the reduction in *MtACS10* expression could be linked to the root's susceptibility to nodulation. To address this, we conducted a complementation experiment in which we inhibited the LCO-induced transcriptional repression of *MtACS10*.

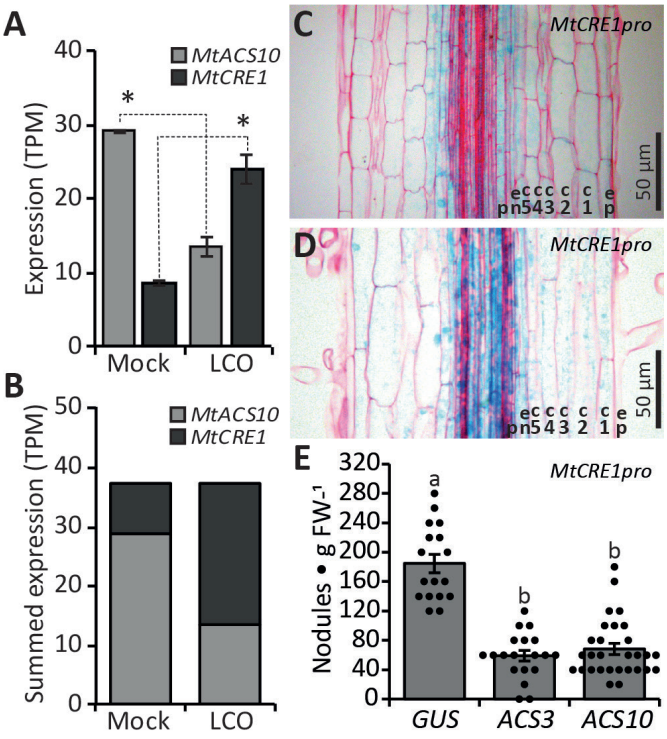


Figure 7 | Counteracting *MtACS10* reduction and its effect on nodulation. **(A-B)** Antagonistic expression responses of *MtACS10* and *MtCRE1* under mock and LCO treatment, shown as **(A)** individual and **(B)** summed expression bar graphs (n=3). **(C-D)** *MtCRE1pro::GUS* reporter activity at 24 hours post-inoculation with **(C)** mock or **(D)** *Sm2011*. **(E)** Effect of local ectopic induction of *ACS* gene expression on nodulation. Nodule numbers in control (*MtCRE1pro::GUS*) and two independent ectopic induced *ACS* expression lines (*MtCRE1pro::ACS3* and *MtCRE1pro::ACS10*) (n>15). Transgenic and non-transgenic roots (for untransformed roots seeSupplemental Figure **S8**) were distinguished by *DsRed* expression. Bars represent mean ± SE; different letters or an asterisk (*) indicate significant differences ($P < 0.05$). FW, fresh weight, ep, epidermis;c1-c5, cortical cell layers; en, endodermis; pc, pericycle in C, D.

For this purpose, we leveraged the observation that *MtACS10* and *MtCRE1* are expressed in overlapping spatial domains (**Figure 3E**, Gonzalez-Rizzo et al. 2006; Liu et al. 2019), and respond antagonistically to LCO application (**Figure 7A, B**). While *MtACS10* transcript levels decrease by approximately 70% upon LCO treatment, *MtCRE1* expression increases by roughly twofold (**Figure 7A**), and their cumulative expression remains comparable between treatments (**Figure 7B**).

To test the functional relevance of the *MtACS10* repression, we generated composite plants bearing transgenic roots that ectopically express either *GUS* or *MtACS10* under control of the *MtCRE1* promoter. To rule out a gene-specific effect of *MtACS10*, we also created a construct in which *MtACS3* is driven by the *MtCRE1* promoter. The *MtCRE1pro::GUS* construct served both as a negative control and to validate that the promoter is active in the appropriate root cell layers (i.e. the pericycle, endodermis, and inner cortex, **Figure 7C**) and is induced following LCO

perception (**Figure 7D**). As expected, expression of *MtCRE1pro::GUS* had no detectable effect on nodulation (**Figure 7E**). In contrast, ectopic expression of either *MtACS3* or *MtACS10* under the *MtCRE1* promoter led to a 50–60% reduction in nodulation (**Figure 7E**). This effect was absent in the non-transgenic roots harvested from the same composite plants (**Supplemental Figure S8**). These findings demonstrate that temporal ectopic expression of either *ACS* gene in root interior represses nodulation.

Ethylene co-defines the size of the Medicago root susceptible zone in apical-basal direction

As nodulation in Medicago seems to rely on the removal of a local ethylene potential by a transcriptional repression of *MtACS10*, we next asked whether ethylene might also define the spatial limits of the root susceptible zone. To test this, we flooded the root of Medicago wild-type A17 and *sickle* root with *Sm2011* LCOs for 3 hrs and sampled the lower 2.5 cm of the root in five 5 mm segments. These segments were named from the root tip upwards as Z1 to Z5 (**Figure 8A**). From these segments, we isolated RNA for cDNA synthesis and subsequent qRT-PCR against one of the key genes in nodulation, *NODULE INCEPTION* (*NIN*). This revealed that in wild type, *NIN* induction is mostly restricted to Z2, the susceptible zone (Figure 8B). We observed *NIN* induction in Z1 and Z3; however, these increases were marginal compared to the approximately 90-fold induction detected in Z2. This limited activation in the flanking zones may reflect natural variation in the position or extent of the susceptible zone, which can occasionally extend beyond the 5 mm segment defined as Z2. Similarly, the induction observed in Z4 suggests that the susceptible zone may sometimes expand further along the root or may not be strictly confined to a fixed location. In *sickle*, *NIN* was induced in all segments tested, although to a lesser extent in Z1 (**Figure 8B**). This suggests that the response to LCO application is broader in this mutant. Next, we tested if nodulation could occur outside the susceptible zone. For this, we spot-inoculated wild-type A17 and *sickle* at different positions on the root (i.e. either Z1, Z2, Z3, Z4 or Z5) and counted the number of nodules at 7 DPI. This revealed that in wild type, nodules were consistently formed on Z2, the susceptible zone, and Z1 (**Figure 8C**). The latter is probably because during the remainder of the experiment the root keeps growing and this Z1 at the start of the experiment progressed through a stage in which it became susceptible. This is not the case for the other zones. Apart from Z3, that only formed nodules occasionally, Z4 and Z5 never formed any nodules. When we performed this experiment in *sickle*, the results were different. In this mutant, nodules were formed on all zones, although marginally less efficient in Z5 (**Figure 8C**).

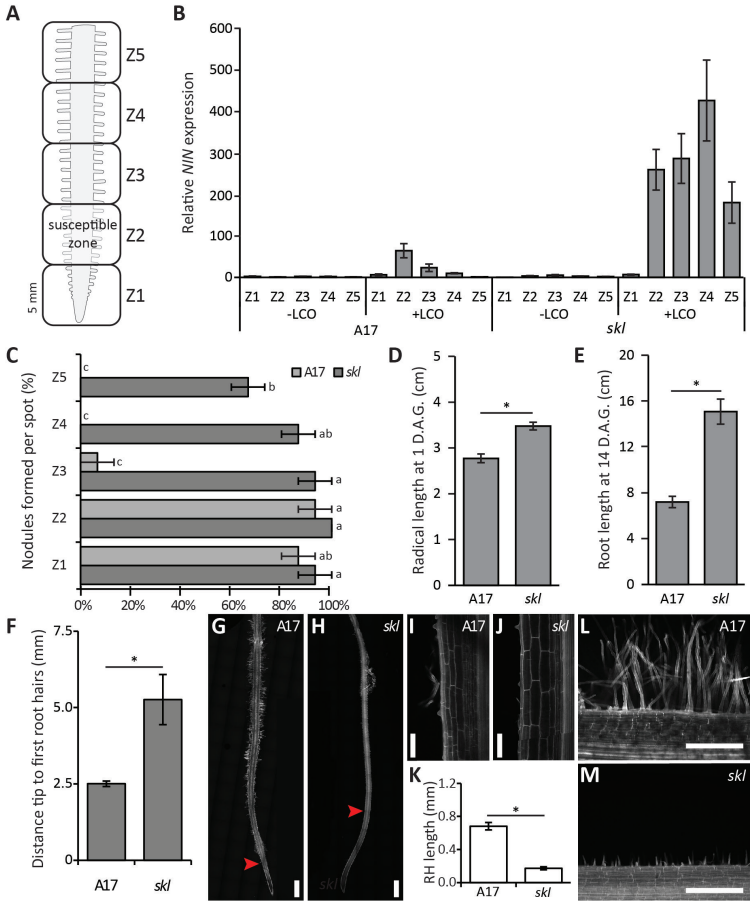


Figure 8 | The role of ethylene in defining the root susceptible zone in Medicago. (A) Schematic representation of the Medicago root separation in 5 mm zones numbered bottom up, Zone (Z)1-5. **(B)** Relative *NIN* expression (n = 3) in Z1 to Z5 in response to *Sm2011* Nod factors (LCO) in wild-type A17 and the *sickle(skl)/ethylene insensitive2* mutant of Medicago. **(C)** Percentage of nodule formed per spotted location at different distances from the root tip in wild-type A17 and *skl*. Root length of wild-type A17 and *skl* at **(D)** 1 and **(E)** 14 days after germination (D.A.G.). **(F)** The average distance from the Medicago root tip where the first root hairs are formed in wild-type A17 and *skl*. **(G-H)** Representative confocal tile image of the Medicago **(G)** wild-type A17 and (H) *skl* root tips. Arrowhead points at first root hair. **(I-J)** Zoom in of first root hairs in **(I)** wild-type A17 and **(J)** *skl*. **(K)** Average length of the root hairs in the susceptible zone measured at roughly 5 mm distance from the root tip in wild-type A17 and 10 mm in *skl* **(L-M)** Representative confocal image of root hairs in the root susceptible zone of **(L)** wild-type A17 and **(M)** *skl*. Bars represent mean ± SE; different letters or an asterisk (*) indicate significant differences ($P < 0.05$), scale bars 1 mm in G, 100 µm in I and J, and 500 µm in L and M.

A key feature of the *sickle* mutant root is its length (Penmetsa and Cook 1997; Penmetsa et al. 2003). By 1 day after germination (DAG), the radical of the *sickle* mutant was approximately 25% longer than those of wild-type A17 (**Figure 8D**). This difference in primary root length persisted and increased over time, reaching roughly a two-fold difference by 14 DAG (**Figure 8E**). Additionally, we observed that the distance from the epidermal cells producing the first root hairs to the root tip was increased by roughly two-fold compared to wild-type A17 (~2.5 vs 5.0 mm,Fig**ure 8F-J**). Interestingly, *sickle* root hairs remained small and few over the entire length of the root susceptible zone, whereas in wild-type A17, root hairs became substantially longer within a short distance from their initiation (**Figure 8K-M**). This phenotype was reminiscent of that observed in *Mtacs3-1* (**Supplemental Figure S9A**). Similar to *sickle*, root hairs in this mutant emerged further from the root tip and remained smaller than in wild-type R108, a characteristic not observed in the *Mtacs10-1* mutant (**Supplemental Figure S9B-J**).

DISCUSSION

This study provides insights into the spatial and functional dynamics of ethylene biosynthesis and signaling during the early stages of legume-rhizobium symbiosis in *Medicago truncatula* (Medicago). By employing promoter-reporter constructs, mutant analysis, and transgenic approaches, we reveal a complex yet coordinated regulatory mechanism by which ethylene biosynthesis, and specifically MtACS3 and MtACS10, shapes the spatial domain and developmental outcome of nodule formation (**Figure 9A-D**). Our findings expand our understanding of how local ethylene biosynthesis and response contribute to nodule number, infection thread formation, and radial positioning (**Figure 9E**).

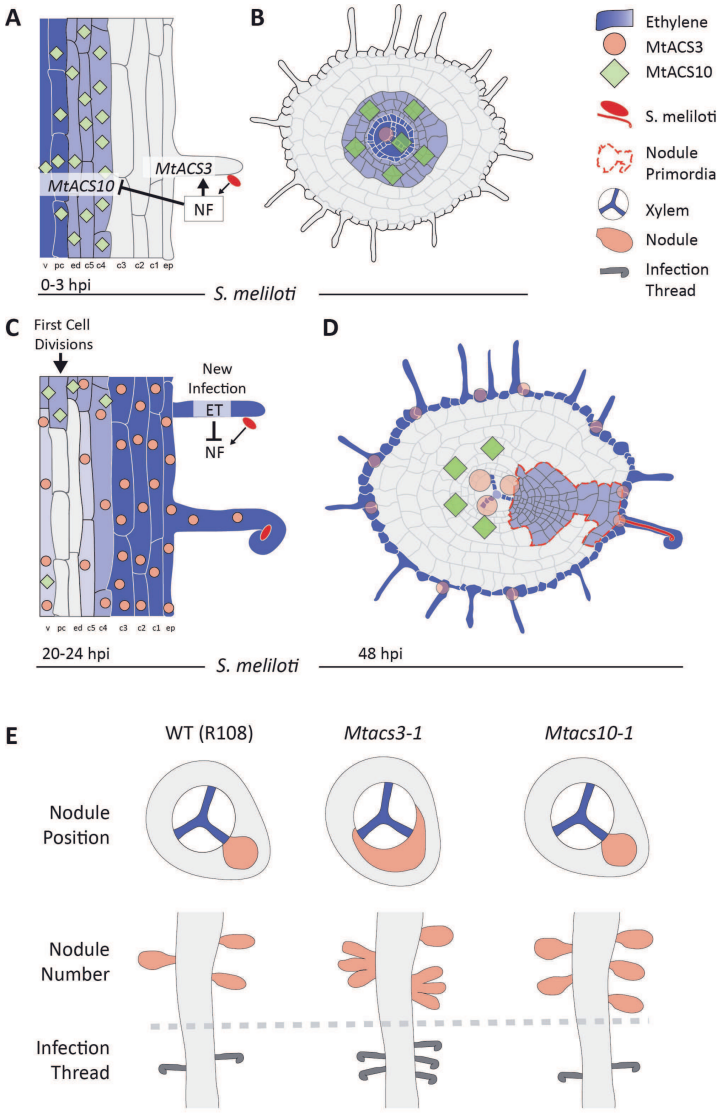


Figure 9 | A model of ethylene redistribution during nodule initiation in *Medicago truncatula*. (A, B) In the uninduced susceptible zone, ethylene is produced internally via MtACS10. Upon *Sinorhizobium meliloti* inoculation or Nod factor (NF) perception, *MtACS10* is suppressed while *MtACS3* is activated in the outer root layers, including the epidermis. (C, D) By 20–24 hpi, ethylene biosynthesis is shifted to the outer root tissues via *MtACS3*, which inhibits additional root hair infection. At 48 hours post inoculation (hpi), the *EBSnew* ethylene reporter marks developing nodule primordia, with *MtACS3* proposed as a candidate source. (E) Phenotypes of *Mtacs3-1* and *Mtacs10-1*: wild type (WT) and *Mtacs10-1* nodules remain xylem pole positioned, whereas this positional specificity is lost in *Mtacs3-1*. *Mtacs10* mutants form more nodules without changing the number of Infection Threads, while *Mtacs3* shows more infection threads, fused nodules, but a similar total number of nodulation events.

We first analyzed the spatiotemporal activity of the synthetic ethylene-responsive promoter *EBSn* during early rhizobial signaling. The initial localization of promoter activity in the inner tissues of the root susceptible zone under mock conditions likely reflects baseline ethylene biosynthesis, possibly through *MtACS10*. Upon rhizobial inoculation, however, *EBSn* activity shifts to the outer root layers, and eventually to the epidermis, suggesting either a redistribution of ethylene or, given our MtACS results, more likely a spatially restricted repression and induction of local ethylene biosynthesis. Given that ethylene is a gaseous molecule, one might expect diffusion to obscure such sharp expression boundaries. It is often assumed that ethylene can diffuse freely between cells and tissues (Park et al. 2017). Nevertheless, our observations show that localization of biosynthesis is crucial for nodulation. Ethylene biosynthesis occurs inside the cell and is perceived by receptors that are located on the endoplasmic reticulum (Chen et al. 2002; Wang et al. 2002; Ma et al. 2006; Ju and Chang 2015). This suggests that for ethylene to be functional, it has to be dissolved in the luminal fluid of the endoplasmic reticulum. The diffusion rate of dissolved ethylene in water (~1–2 × 10^{-5} cm^2 s^{-1}) is three to four orders of magnitude lower than that of ethylene in air (~1.4 × 10^{-1} cm^2 s^{-1}) (Huq and Wood 1968; Pritchard and Currie 1982). As such, it is likely that spatial separation of ethylene biosynthesis in different tissues can create concentration gradients, similar as is observed for other hormones (Friml et al. 2003; Grieneisen et al. 2007; Antoniadi et al. 2015; Rizza and Jones 2019). The localized changes in *MtACS* and *EBSn* promoter activity strongly argue for cell-type-specific biosynthesis, regulated transcriptionally. Shifting ethylene biosynthesis from inner to outer root tissues could also explain the increased ethylene emissions after rhizobia application previously reported (Ligero et al. 1986, 1987; Reid et al. 2018), even without changes in net ACC content, as ethylene is likely to leak more freely into the rhizosphere from the root epidermis.

Through RNA sequencing, qRT-PCR, and promoter-reporter analyzes, we identified *MtACS3* and *MtACS10* as the primary *ACS* genes expressed in the root susceptible zone under non-symbiotic conditions. RNA *in situ* hybridization partially confirmed the expression patterns of these two genes, with *MtACS10* confined to inner cortical domains. *MtACS3* was undetectable at the mRNA level in the susceptible zone but active in reporter assays upon inoculation. This apparent discrepancy may reflect rapid transcriptional activation of *MtACS3* or limitations in probe sensitivity and highlights the importance of complementary methods to visualize dynamic gene expression. Nevertheless, the dynamic and spatially distinct regulation of these two *ACS* genes represents a mechanism to reconfigure ethylene biosynthesis to balance the competing requirements of infection control and nodule initiation.

The functional relevance of these transcriptional shifts is highlighted by RNAi and mutant analyzes. Knockdown or loss-of-function of *MtACS10* results in a significant increase in nodule number, suggesting that this gene acts as a local suppressor of nodulation. Its downregulation after LCO treatment appears to be a critical step in rendering specific pericycle and/or cortical cells competent for division and nodule primordium formation. This interpretation is supported by the finding that ectopic expression of either *MtACS3 or MtACS10* under the *MtCRE1* promoter represses nodulation. This supports the hypothesis that downregulation of *MtACS10*, and likely

ethylene biosynthesis, in specific cell layers is a prerequisite for successful nodule initiation in Medicago. By preventing this repression, the developmental program is hindered, reinforcing the functional importance of spatial gene regulation. Conversely, MtACS3 is likely affecting nodule number indirectly by influencing infection thread frequency and when mutated particularly promoting nodule clustering. As we expected this response to be regulated from the inner root tissues, the *MtACS3* expression domain indicates the observed nodule clusters might be the result of secondary infections. This would suggest that MtACS3 fine-tunes the infection process and possibly spatial coordination among neighboring infection events, potentially by modulating ethylene levels in the epidermis. The induction of *MtACS3* expression in the root hairs observed by Breakspear at al. (2014) supports this latter statement. We hypothesize that an epidermal increase in ethylene levels in response to LCO perception is part of a negative feedback loop that restricts further signaling after a successful symbiotic interaction has been initiated. This hypothesis is consistent with the hyper-infection phenotype observed in the *sickle* mutant (Penmetsa and Cook, 1997).

Importantly, the radial positioning of nodule primordia, a process previously shown to be ethylene-sensitive in pea and Medicago (Heidstra et al. 1997; Penmetsa et al. 2003, 2008), is altered in the *Mtacs3-1* mutant. Wild-type Medicago roots preferentially form nodules opposite xylem poles, a pattern disrupted in *Mtacs3-1* where nodule initiation is more randomly distributed. These results suggest that local ethylene biosynthesis via MtACS3 contributes to defining or reinforcing polarity cues for nodule positioning. We cannot exclude that the observed vasculature promoter activity upon rhizobial inoculation is sufficient to create a local restriction on nodulation near the phloem poles. Nevertheless, even if a vasculature-associated expression domain of *MtACS3* would be sufficient to direct nodule initiation toward the xylem poles, the fact that *MtACS3* promoter activity is only observed after rhizobial application raises questions on what determines the position of this local ethylene biosynthesis. Our results suggest that ethylene produced via MtACS3 may not establish positional information itself but instead acts downstream of an unknown cue that determines the radial bias. However, in the *Mtacs3* mutant, some nodules seem to spread wider after initiation, whereas at other nodule primordia, infection threads that originate in the epidermis opposite the phloem poles are clearly observed. This suggests that epidermal ethylene biosynthesis could indirectly co-regulate nodule numbers and taxis and indicates that secondary infections in the *Mtacs3-1* mutant lead, at least in part, to the observed disruption in radial nodule positioning. Interestingly, although *MtACS10* is expressed in the pericycle and inner cortex, *Mtacs10-1* mutants retain wild-type spatial bias, reinforcing the notion that MtACS10 primarily functions in restricting nodule number rather than spatial placement.

The ethylene-insensitive *sickle* mutant demonstrates that ethylene not only regulates local infection and nodule formation but also co-defines the spatial extent of the root susceptible zone. In wild-type roots, the LCO-induced expression of *MtNIN* is largely confined to a narrow region roughly 5 mm distal from the root tip, while in *sickle*, this response is, with the exception of the root tip, broadened to the entire 20 mm root segment analyzed. This expansion is

consistent with the broader nodulation capacity of *sickle* roots, which form nodules outside the conventional susceptible zone. Additionally, the altered root growth dynamics in *sickle*, including extended elongation zone and altered root hair development, underscore the broader developmental roles of ethylene that indirectly influence symbiotic competence.

Ethylene serves as a central developmental "timer" in plant tissues, promoting their maturation and eventual senescence (Khan et al. 2014). Additionally in Arabidopsis, ethylene also enhances root hair elongation by modulating key regulatory pathways (Pitts et al. 1998). In line with this, our findings suggest that MtACS3 may act as a key regulator of root hair maturation and infection competence in Medicago. RNA *in situ* hybridisation shows that *MtACS3* is expressed in the epidermis at the start of the differentiation zone, precisely where root hairs are emerging. This is consistent with the observation that in the *Mtacs3-1* mutant, root hairs remain small and poorly developed, a phenotype strongly reminiscent of the ethylene-insensitive *sickle/Mtein2* mutant. Although small and poorly developed, the root hairs of *Mtacs3-1* still support rhizobial infection, indicating that ethylene is not strictly required to establish rhizobia susceptibility. Instead, we suggest that ethylene regulates the duration of the susceptibility window. When root hairs emerge at the start of the differentiation zone, ethylene may initiate a physiological timer that defines both the temporal and spatial extent of the susceptible zone. In this framework, wild-type root hairs in the susceptible zone remain competent for rhizobial entry, whereas those positioned beyond it have matured past susceptibility. A rhizobia-induced burst of ACS3-derived ethylene in the epidermis upon inoculation could then accelerate maturation of neighboring, non-infected hairs, effectively pre-closing the infection window. In this model, a strong upregulation of *MtACS3* would function as a switch—allowing infection of a limited number of hairs while shielding the root from undesired supernumerary infections. In the absence of *MtACS3* activity, or if ethylene signaling is compromised, as in *Mtacs3-1* or *sickle/Mtein2*, respectively, both maturation controls (i.e., physiological timer and switch) are lost, leading to prolonged susceptibility, excess infection threads, and clustered nodules. Thus, ACS3-mediated ethylene production may function as a que that synchronizes the transition from susceptibility to resistance at the root surface.

Taken together, these findings establish a new framework in which ethylene signaling, mediated by distinct *ACS* genes, functions in a spatially and temporally resolved manner to balance infection thread formation, restrict excessive nodulation, and co-defines the physical limits of the root susceptible zone. The selective repression of *MtACS10* emerges as a key step in nodule initiation in Medicago, while *MtACS3* modulates infection-related parameters. This dual regulation of ethylene biosynthesis, both activation and repression depending on tissue context and developmental timing, illustrates once more the sophistication of hormonal control in root nodule symbiosis. Moreover, this dynamic regulation of ethylene biosynthesis provides insight into how plants use spatial and functional partitioning of gene activity to integrate multiple signaling layers to fine-tune developmental responses like nodulation with precision. For legume-rhizobia interactions, this spatial specificity allows the plant to limit the cost of nodulation by tightly controlling where and how nodules form. Future work should test whether

similar regulatory mechanisms operate in other legumes, particularly crops, and how autoregulation of nodulation or environmental factors such as nitrate and abiotic stress modulate this ethylene-dependent spatial control.

To date, engineering nodulation in non-legumes has largely emphasized activating positive regulators, while inhibitory pathways remain underexplored (Untergasser et al. 2012; Mus et al. 2016; Huisman and Geurts 2020). Our findings suggest that ethylene homeostasis in Medicago restricts pericycle and cortical divisions, mirroring its conserved role in lateral root initiation in *Arabidopsis* (Nodzon et al. 2004; Prasad et al. 2010). Given the proposed shared origin of nodules and lateral roots (Nutman 1948; Hirsch et al. 1997; Mathesius et al. 2000; Franssen et al. 2015; Schiessl et al. 2019; Soyano et al. 2019), ethylene or similar repressors may likewise limit nodule initiation in other species. Accounting for such inhibitory mechanisms will be essential for engineering nodulation in non-legume crops.

MATERIALS AND METHODS

Plant material, growth conditions and treatments

Medicago seeds (wild-type Jemalong A17 and R108, *Mtacs3-1* (NF1952) and *Mtacs10-1* (NF15254)) were treated with concentrated H_2SO_4 for 7 minutes, rinsed five times with MilliQ water, and sterilized for 10 minutes using normal household bleach. Seeds were again washed with sterile MilliQ water five times and placed on round Petri dishes containing Fåhraeus medium (0.25 mM NO_3^-, 1% Daishin agar, 0.01% (w/v) carbendazim) at 4˚C in darkness for stratification. After 48 hours, seeds were transferred to room temperature to germinate for an additional 24 hours. Germinated seedlings were transferred to the appropriate growth system. For all experiments, plants were grown in an environmentally controlled growth chamber at 20ºC/18ºC with a 16h-light/8h-dark cycle and 70% relative humidity. A modified Fåhraeus medium (0.12 g/L $MgSO_4 \bullet 7H_2O$, 0.10 g/L KH_2PO_4, 0.15 g/L $Na_2HPO_4 \bullet 2H_2O$, 1 ml/L 15 mM Fe-Citrate, 2.50 ml/L Spore-elements β- ($CuSO_4 \bullet 5H_2O$ 0.0354g/L, $MnSO_4 \bullet H_2O$ 0.462g/L, $ZnSO_4 \bullet 7H_2O$ 0.974g/L, H_3BO_3 1.269 g/L, $Na_2MoO_4 \bullet 2H_2O$ 0. 398 g/L), pH 6.7) (Fähraeus 1957) with 0.125 mM $Ca(NO_3)_2$ was used in all experiments. For square (12cmx12cm) plate grown plants, 1% Daishin agar was added. Eight seedlings were grown per plate, and four plates per were used per replicate. All plates were partially covered with tin foil to avoid light grown roots. From the growing roots, the root susceptible zones (± 0.5 cm) of ca. 20-30 roots were collected and snap frozen in liquid nitrogen. All samples were stored at -80 ºC until further use.

LCO or Rhizobia application

Sinorhizobium meliloti strain 2011 (*Sm2011*) LCOs were purified and applied as previously described (Spaink et al. 1991; van Zeijl et al. 2015b) with minor modifications. *Sm2011* LCO stocks were stored in 100% DMSO and diluted 100-fold (~10-9 M) in Fåhraeus (0.125 mM $Ca(NO_3)_2$) medium prior to application. The same Fåhraeus medium with 1% DMSO was used as mock treatment. LCO or mock treatments were pipetted on the root susceptible zone. Roots were exposed for 3 hours and subsequently the root susceptible zone (0.5 cm root segments) were cut just 5 mm from the root tip and snap-frozen in liquid nitrogen (n=3) for gene expression analysis. *Sm2011* was grown on liquid YEM medium, spined down and the pallet resuspended in Fåhraeus medium to an OD_{600} of 0.02 or 0.1 for direct application on either plates or inoculation in plants grown on perlite, respectively.

Phylogenetic analyzes

ACS, ACO, and ACD protein sequences were retrieved from various plant genome databases, including that of Medicago (see supplemental table S1), using the BLASTP algorithm (Altschul et al. 1990). Homologous sequences were aligned using MAFFT version 7.450 with automatic selection of the appropriate algorithm and using the BLOSUM62 scoring matrix (Henikoff and Henikoff 1992; Katoh et al. 2002). Phylogenetic trees were estimated based on maximum likelihood using the IQTree webserver version 1.6.11 (Trifinopoulos et al. 2016; Minh et al. 2020) with automatic selection of the best-fitting model of protein evolution based on BIC using

ModelFinder (ACS and ACO: JTT+F+I+G4; ACD: LG+G4 (Kalyaanamoorthy et al. 2017)) and 1000 ultrafast bootstrap replicates to assess clade support (Hoang et al. 2017).

RNA isolation, cDNA synthesis and Quantitative RT-PCR

RNA was isolated from snap-frozen roots samples using the plant RNA kit (E.Z.N.A, Omega Biotek, Norcross, USA) according to the manufacturer's protocol. This RNA was either send for sequencing (Beijing Genomics Institute, Hong Kong, China) or 1 µg total RNA was used to synthesize cDNA using the i-script cDNA synthesis kit (Bio-Rad, Hercules, USA) as described in the manufacturer's protocol. Real time qRT-PCR was set up in 10 µl reactions with 2× iQ SYBR Green Super-mix (Bio-Rad, Hercules, USA). Experiments have been conducted on a CFX Connect optical cycler, according to the manufacturers protocol (Bio-Rad, Hercules, USA). All primers including the genes used for normalization (*MtUBQ10* and *MtPTB*) are given in Supplemental table S2. Data analysis was performed using CFX Manager 3.0 software (Bio-Rad, Hercules, USA). Cq values of 32 and higher were excluded from the analysis, though still checked for transcriptional induction. Statistical significance was determined based on student's t-test ($p<0.01$).

RNA Sequencing

RNA was sequenced at BGI Tech Solutions (Beijing Genomics Institute, Hong Kong, China) using the Illumina Truseq (Transcriptome) protocol utilizing an Illumina Hiseq 2000 instrument generating 150 bp paired-end reads. In total 24–33 million clean reads were generated for each sample. Sequencing data were analyzed by pseudoaligning RNA-seq reads against the *Medicago truncatula* genome annotation (Mt4.0v1) using kallisto v0.46.2 (Bray et al. 2016) with default settings. All reads are available on the European Nucleotide Archive (ENA) under accession number PRJEB38785.

RNA in situ hybridisation

RNA in situ hybridisation was performed as previously described (Kulikova et al. 2018). RNA ISH probe set for *MtACS3* and *MtACS10* was designed and synthesized by request at Thermo Fisher Scientific (catalogue numbers VF1-6000770 and VF1-6000771 for *MtACS3* and *MtACS10*, respectively). As a negative control any probe set was omitted for hybridization. Probe IDs are given in Supplemental table S3.

Vectors and constructs

For RNAi-mediated knockdown of *MtACS3* or *MtACS10*, two fragments (332 bp and 310 bp, respectively (**Supplemental Table S4**) were amplified from Medicago Jemalong A17 root cDNA, using specific primer pairs (**Supplemental Table S2**), and cloned into pENTR-D-TOPO (Invitrogen, Carlsbad, USA). Both RNAi fragments were recombined into the DsRed-modified gateway vector pK7GWIWG2(II)-RR driven by the *CaMV35S* promoter (Limpens et al. 2004) to obtain the binary constructs pK7GWIWG2(II)-RR-*p35S-MtACS3*-RNAi and pK7GWIWG2(II)-RR-*p35S-MtACS10*-RNAi. For the empty vector control, the binary plasmid pK7GWIWG2(II)-RR-*p35S*-RNAi-control as previously described (van Zeijl et al. 2015a) was used. For ectopic

expression and promoter GUS studies, the coding sequences, putative promoters (3.5 kb region upstream of start codon) and putative 3'UTRs (1 kb region downstream of stop codon) of *MtACS3* and *MtACS10* were synthesized at Thermo Fisher Scientific. These, as well as a GUS-encoding sequence (pICH75111), *MtCRE1* promoter (Gonzalez-Rizzo et al. 2006) and *35S* terminator (pICH41414) were assembled together with the into vectors pICH47811 to create EC74831; MtCRE1p-MtACS3-t35S; EC74832; MtCRE1p-MtACS10-t35s; EC74833; MtCRE1p-GUS-t35s; EC74963; MtACS3p-GUS-ACS3-3`UTR; EC74964; MtACS10p-GUS-ACS10-3`UTR (for sequences see **Supplemental Table S5**) using Golden Gate cloning (Engler et al. 2014). Binary transformation constructs were created by assembling the resulting clones together with clone pICH47732-*AtUBQ10::DsRED::tNOS* into vector pICSL4723 (Engler et al. 2014). All constructs are available from our laboratory upon request.

Plant transformation and histology

Agrobacterium rhizogenes-mediated root transformation was used to transform Medicago (Jemalong A17) as previously described (Limpens et al. 2004). Transgenic roots were selected based on *DsRED1* expression. Plants with transgenic roots were, depending on their use, transferred to either plates or perlite containing Fåhraeus medium (0.125 mM $Ca_2(NO_3)_2$), and inoculated with the *S. meliloti* strain *2011* (*Sm2011*, OD_{600} = 0.1) between 20-48 hours for most GUS work and two-three weeks in the case of nodulation assays after transfer.

ACC extraction, detection and quantification by liquid chromatography-tandem mass spectrometry

ACC analysis from the Medicago root were performed as previously described (Bours et al. 2013) with adaptations as described by Gühl and colleagues (Gühl et al. 2021).

Statistical Analysis

When appropriate, data were subjected to the Student's t test (Microsoft Excel). All other data were subjected to one-way ANOVA. Individual differences were then identified using a post hoc Tukey test ($P < 0.05$). For these, all analyzes were performed using SAS_9.20 (http://www.sas.com/).

ACKNOWLEDGMENTS

We would like to acknowledge Marijke Hartog and Sidney Linders for technical support of this project. Colleen Drapek for suggestions and critically reading of the manuscript.

AUTHOR CONTRIBUTIONS

Conceptualization, SM, AvZ, WK; Methodology, RvV, OK, RH, WK; Investigation, SM, TS, KA, RH, AvS, RvV, OK, TW, JK, HF, EL, AvZ, WK; Formal Analysis, SM, TS, AvZ and WK; Contribution of a new analytic tool; JFM, AS, JA; Visualization, SM, WK; Writing – Original Draft, SM, TS, AvZ and WK; Writing – Review & Editing, SM, TS, AvS, RvV, OK, RH, JFM, AS, JA, HF, EL, AvZ, WK; Writing – Final version, SM, TS, AvZ and WK; Funding Acquisition, WK; Supervision, WK.

FUNDING

This work was supported by NWO-Veni (863.15.010) and NWO-Vidi (VI.Vidi.193.119) to WK, and by grant PID2021-122740OB-I00 funded by MCIN/AEI/10.13039/501100011033 to EL.

REFERENCES

Adema K and Kohlen W. The symbiosome-a transient organelle in evolution. J Exp Bot. 2024:75(11):3209–3213.

Altschul SF, Gish W, Miller W, Myers EW, and Lipman DJ. Basic local alignment search tool. J Mol Biol. 1990:215(3):403–410.

An F, Zhao Q, Ji Y, Li W, Jiang Z, Yu X, Zhang C, Han Y, He W, Liu Y, et al. Ethylene-induced stabilization of ETHYLENE INSENSITIVE3 and EIN3-LIKE1 is mediated by proteasomal degradation of EIN3 binding F-box 1 and 2 that requires EIN2 in Arabidopsis. Plant Cell. 2010:22(7):2384–2401.

Antoniadi I, Plačková L, Simonovik B, Doležal K, Turnbull C, Ljung K, and Novák O. Cell-type-specific cytokinin distribution within the Arabidopsis primary root apex. Plant Cell. 2015:27(7):1955–1967.

Bhuvaneswari TV, Turgeon BG, and Bauer WD. Early Events in the Infection of Soybean (Glycine max L. Merr) by Rhizobium japonicum: I. LOCALIZATION OF INFECTIBLE ROOT CELLS. Plant Physiol. 1980:66(6):1027–1031.

Binder BM. Ethylene signaling in plants. J Biol Chem. 2020:295(22):7710–7725.

Bours R, van Zanten M, Pierik R, Bouwmeester H, and van der Krol A. Antiphase light and temperature cycles affect PHYTOCHROME B-controlled ethylene sensitivity and biosynthesis, limiting leaf movement and growth of Arabidopsis. Plant Physiol. 2013:163(2):882–895.

Bowman JL, Briginshaw LN, Fisher TJ, and Flores-Sandoval E. Something ancient and something neofunctionalized-evolution of land plant hormone signaling pathways. Curr Opin Plant Biol. 2019:47:64–72.

Bray NL, Pimentel H, Melsted P, and Pachter L. Near-optimal probabilistic RNA-seq quantification. Nat Biotechnol. 2016:34(5):525–527.

Breakspear A, Liu C, Roy S, Stacey N, Rogers C, Trick M, Morieri G, Mysore KS, Wen J, Oldroyd GED, 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. 2014:26(12):4680–4701.

Chen Y-F, Randlett MD, Findell JL, and Schaller GE. Localization of the ethylene receptor ETR1 to the endoplasmic reticulum of Arabidopsis. J Biol Chem. 2002:277(22):19861–19866.

Depaepe T and Van Der Straeten D. Tools of the ethylene trade: A chemical kit to influence ethylene responses in plants and its use in agriculture. Small Methods. 2020:4(8):1900267.

Dubois M, Van den Broeck L, and Inzé D. The pivotal role of ethylene in plant growth. Trends Plant Sci. 2018:23(4):311–323.

Engler C, Youles M, Gruetzner R, Ehnert T-M, Werner S, Jones JDG, Patron NJ, and Marillonnet S. A golden gate modular cloning toolbox for plants. ACS Synth Biol. 2014:3(11):839–843.

Evert RF and Eichhorn SE. Regulating growth and development: The plant hormones. . In. Raven Biology of Plants. (Macmillan Learning: New York), pp. 638–659.

Fähraeus G. The infection of clover root hairs by nodule bacteria studied by a simple glass slide technique. J Gen Microbiol. 1957:16(2):374–381.

Ferguson BJ, Mens C, Hastwell AH, Zhang M, Su H, Jones CH, Chu X, and Gresshoff PM. Legume nodulation: The host controls the party. Plant Cell Environ. 2019:42(1):41–51.

Fernandez-Moreno J-P, Fenech M, Yaschenko AE, Zhao C, Neubauer M, Davis HN, Marchi AJ, Concannon R, Keren-Keiserman A, Reuveni M, et al. EBSn, a robust synthetic reporter for monitoring ethylene responses in plants. Plant Biotechnol J. 2025. https://doi.org/10.1111/pbi.70302

Fernandez-Moreno J-P, Yaschenko AE, Neubauer M, Marchi AJ, Zhao C, Ascencio-Ibanez JT, Alonso JM, and Stepanova AN. A rapid and scalable approach to build synthetic repetitive hormone-responsive promoters. Plant Biotechnol J. 2024:22(7):1942–1956.

Franssen HJ, Xiao TT, Kulikova O, Wan X, Bisseling T, Scheres B, and Heidstra R. Root developmental programs shape the Medicago truncatula nodule meristem. Development. 2015:142(17):2941–2950.

Friml J, Vieten A, Sauer M, Weijers D, Schwarz H, Hamann T, Offringa R, and Jürgens G. Efflux-dependent auxin gradients establish the apical-basal axis of Arabidopsis. Nature. 2003:426(6963):147–153.

Gao Z, Wen C-K, Binder BM, Chen Y-F, Chang J, Chiang Y-H, Kerris RJ 3rd, Chang C, and Schaller GE. Heteromeric interactions among ethylene receptors mediate signaling in Arabidopsis. J Biol Chem. 2008:283(35):23801–23810.

Gómez-Fernández GO, van Velzen R, Mun J, Cook DR, Kohlen W, and Larrainzar E. Ethylene biosynthesis in legumes: gene identification and expression during early symbiotic stages. J Exp Bot. 2025:eraf069.

Gonzalez-Rizzo S, Crespi M, and Frugier F. The Medicago truncatula CRE1 cytokinin receptor regulates lateral root development and early symbiotic interaction with Sinorhizobium meliloti. Plant Cell. 2006:18(10):2680–2693.

Grieneisen VA, Xu J, Marée AFM, Hogeweg P, and Scheres B. Auxin transport is sufficient to generate a maximum and gradient guiding root growth. Nature. 2007:449(7165):1008–1013.

Gühl K, Holmer R, Xiao TT, Shen D, Wardhani TAK, Geurts R, van Zeijl A, and Kohlen W. The Effect of Exogenous Nitrate on LCO Signaling, Cytokinin Accumulation, and Nodule Initiation in Medicago truncatula. Genes . 2021:12(7). https://doi.org/10.3390/genes12070988

Guinel FC. Ethylene, a hormone at the center-stage of nodulation. Front Plant Sci. 2015:6:1121.

Guzmán P and Ecker JR. Exploiting the triple response of Arabidopsis to identify ethylene-related mutants. Plant Cell. 1990:2(6):513–523.

Heidstra R, Yang WC, Yalcin Y, Peck S, Emons AM, van Kammen A, and Bisseling T. Ethylene provides positional information on cortical cell division but is not involved in Nod factor-induced root hair tip growth in Rhizobium-legume interaction. Development. 1997:124(9):1781–1787.

Henikoff S and Henikoff JG. Amino acid substitution matrices from protein blocks. Proc Natl Acad Sci U S A. 1992:89(22):10915–10919.

Herrbach V, Chirinos X, Rengel D, Agbevenou K, Vincent R, Pateyron S, Huguet S, Balzergue S, Pasha A, Provart N, et al. Nod factors potentiate auxin signaling for transcriptional regulation and lateral root formation in Medicago truncatula. J Exp Bot. 2017:68(3):569–583.

Herud-Sikimić O, Stiel AC, Kolb M, Shanmugaratnam S, Berendzen KW, Feldhaus C, Höcker B, and Jürgens G. A biosensor for the direct visualization of auxin. Nature. 2021:592(7856):768–772.

Hirsch AM, Larue TA, and Doyle J. Is the Legume Nodule a Modified Root or Stem or an Organ sui generis? CRC Crit Rev Plant Sci. 1997:16(4):361–392.

Hoang DT, Chernomor O, Haeseler A von, Minh BQ, and Vinh LS. UFBoot2: Improving the ultrafast bootstrap approximation. bioRxiv. 2017. https://doi.org/10.1101/153916

Huisman R and Geurts R. A Roadmap toward Engineered Nitrogen-Fixing Nodule Symbiosis. Plant Commun. 2020:1(1):100019.

Huq A and Wood T. Diffusion coefficient of ethylene gas in water. J Chem Eng Data. 1968:13(2):256–259.

Ju C and Chang C. Mechanistic insights in ethylene perception and signal transduction. Plant Physiol. 2015:169(1):85–95.

Ju C, Van de Poel B, Cooper ED, Thierer JH, Gibbons TR, Delwiche CF, and Chang C. Conservation of ethylene as a plant hormone over 450 million years of evolution. Nat Plants. 2015:1(1):14004.

Ju C, Yoon GM, Shemansky JM, Lin DY, Ying ZI, Chang J, Garrett WM, Kessenbrock M, Groth G, Tucker ML, et al. CTR1 phosphorylates the central regulator EIN2 to control ethylene hormone signaling from the ER membrane to the nucleus in Arabidopsis. Proc Natl Acad Sci U S A. 2012:109(47):19486–19491.

Kalyaanamoorthy S, Minh BQ, Wong TKF, von Haeseler A, and Jermiin LS. ModelFinder: fast model selection for accurate phylogenetic estimates. Nat Methods. 2017:14(6):587–589.

Katoh K, Misawa K, Kuma K-I, and Miyata T. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. Nucleic Acids Res. 2002:30(14):3059–3066.

Kende H. Ethylene biosynthesis. Annu Rev Plant Physiol Plant Mol Biol. 1993:44(1):283–307.

Khan M, Rozhon W, and Poppenberger B. The role of hormones in the aging of plants - a mini-review. Gerontology. 2014:60(1):49–55.

Knight LI, Rose RC, and Crocker W. Effects of various gases and vapors upon etiolated seedlings of the sweet pea. Science. 1910:311(1):635–636.

Kulikova O, Franken C, and Bisseling T. In Situ Hybridization Method for Localization of mRNA Molecules in Medicago Tissue Sections. Methods Mol Biol. 2018:1822:145–159.

Larrainzar E, Riely BK, Kim SC, Carrasquilla-Garcia N, Yu H-J, Hwang H-J, Oh M, Kim GB, Surendrarao AK, Chasman D, et al. Deep sequencing of the Medicago truncatula root transcriptome reveals a massive and early interaction between Nodulation factor and ethylene signals. Plant Physiol. 2015:169(1):233–265.

Ligero F, Lluch C, and Olivares J. Evolution of Ethylene from Roots of Medicago sativa Plants Inoculated with Rhizobium meliloti. J Plant Physiol. 1986:125(3–4):361–365.

Ligero F, Lluch C, and Olivares J. Evolution of ethylene from roots and nodulation rate of alfalfa (Medicago sativa L.) plants inoculated with Rhizobium meliloti as affected by the presence of nitrate. J Plant Physiol. 1987:129(5):461–467.

Limpens E, Ramos J, Franken C, Raz V, Compaan B, Franssen H, Bisseling T, and Geurts R. RNA interference in Agrobacterium rhizogenes-transformed roots of Arabidopsis and Medicago truncatula. J Exp Bot. 2004:55(399):983–992.

Limpens E, van Zeijl A, and Geurts R. Lipochitooligosaccharides modulate plant host immunity to enable endosymbioses. Annu Rev Phytopathol. 2015:53(1):311–334.

Lin J, Frank M, and Reid D. No Home without Hormones: How Plant Hormones Control Legume Nodule Organogenesis. Plant Commun. 2020:1(5):100104.

Liu J, Rutten L, Limpens E, van der Molen T, van Velzen R, Chen R, Chen Y, Geurts R, Kohlen W, Kulikova O, et al. A Remote cis-Regulatory Region Is Required for NIN Expression in the Pericycle to Initiate Nodule Primordium Formation in Medicago truncatula. Plant Cell. 2019:31(1):68–83.

Ma B, Cui M-L, Sun H-J, Takada K, Mori H, Kamada H, and Ezura H. Subcellular localization and membrane topology of the melon ethylene receptor CmERS1. Plant Physiol. 2006:141(2):587–597.

Ma Q and Dong C-H. Regulatory functions and molecular mechanisms of ethylene receptors and receptor-associated proteins in higher plants. Plant Growth Regul. 2021:93(1):39–52.

Mathesius U, Weinman JJ, Rolfe BG, and Djordjevic MA. Rhizobia can induce nodules in white clover by "hijacking" mature cortical cells activated during lateral root development. Mol Plant Microbe Interact. 2000:13(2):170–182.

Merchante C and Stepanova AN. The triple response assay and its use to characterize ethylene mutants in Arabidopsis. Methods Mol Biol. 2017:1573:163–209.

Minh BQ, Schmidt HA, Chernomor O, Schrempf D, Woodhams MD, von Haeseler A, and Lanfear R. IQ-TREE 2: New models and efficient methods for phylogenetic inference in the genomic era. Mol Biol Evol. 2020:37(5):1530–1534.

Miyata K, Kawaguchi M, and Nakagawa T. Two distinct EIN2 genes cooperatively regulate ethylene signaling in Lotus japonicus. Plant Cell Physiol. 2013:54(9):1469–1477.

Mus F, Crook MB, Garcia K, Garcia Costas A, Geddes BA, Kouri ED, Paramasivan P, Ryu M-H, Oldroyd GED, Poole PS, et al. Symbiotic nitrogen fixation and the challenges to its extension to nonlegumes. Appl Environ Microbiol. 2016:82(13):3698–3710.

Nodzon LA, Xu W-H, Wang Y, Pi L-Y, Chakrabarty PK, and Song W-Y. The ubiquitin ligase XBAT32 regulates lateral root development in Arabidopsis. Plant J. 2004:40(6):996–1006.

Nutman PS. Physiological Studies on Nodule Formation. Ann Bot. 1948:12(2):81–96.

Oldroyd GE, Engstrom EM, and Long SR. Ethylene inhibits the Nod factor signal transduction pathway of Medicago truncatula. Plant Cell. 2001:13(8):1835–1849.

Park J, Lee Y, Martinoia E, and Geisler M. Plant hormone transporters: what we know and what we would like to know. BMC Biol. 2017:15(1):93.

Pattyn J, Vaughan-Hirsch J, and Van de Poel B. The regulation of ethylene biosynthesis: a complex multilevel control circuitry. New Phytol. 2021:229(2):770–782.

Penmetsa RV and Cook DR. A legume ethylene-insensitive mutant hyperinfected by its rhizobial symbiont. Science. 1997:275(5299):527–530.

Penmetsa RV, Frugoli JA, Smith LS, Long SR, and Cook DR. Dual genetic pathways controlling nodule number in Medicago truncatula. Plant Physiol. 2003:131(3):998–1008.

Penmetsa RV, Uribe P, Anderson J, Lichtenzveig J, Gish J-C, Nam YW, Engstrom E, Xu K, Sckisel G, Pereira M, et al. The Medicago truncatula ortholog of Arabidopsis EIN2, sickle, is a negative regulator of symbiotic and pathogenic microbial associations. Plant J. 2008:55(4):580–595.

Pitts RJ, Cernac A, and Estelle M. Auxin and ethylene promote root hair elongation in Arabidopsis. Plant J. 1998:16(5):553–560.

Plett JM, McDonnell L, and Regan S. Plant encoded 1-aminocyclopropane-1-carboxylic acid deaminase activity implicated in different aspects of plant development. Plant Signal Behav. 2009:4(12):1186–1189.

Prasad ME, Schofield A, Lyzenga W, Liu H, and Stone SL. Arabidopsis RING E3 ligase XBAT32 regulates lateral root production through its role in ethylene biosynthesis. Plant Physiol. 2010:153(4):1587–1596.

Pritchard DT and Currie JA. Diffusion of coefficients of carbon dioxide, nitrous oxide, ethylene and ethane in air and their measurement. J Soil Sci. 1982:33(2):175–184.

Qiao H, Shen Z, Huang S-SC, Schmitz RJ, Urich MA, Briggs SP, and Ecker JR. Processing and subcellular trafficking of ER-tethered EIN2 control response to ethylene gas. Science. 2012:338(6105):390–393.

Reid D, Liu H, Kelly S, Kawaharada Y, Mun T, Andersen SU, Desbrosses G, and Stougaard J. Dynamics of ethylene production in response to compatible Nod factor. Plant Physiol. 2018:176(2):1764–1772.

Reid D, Nadzieja M, Novák O, Heckmann AB, Sandal N, and Stougaard J. Cytokinin Biosynthesis Promotes Cortical Cell Responses during Nodule Development. Plant Physiol. 2017:175(1):361–375.

Rizza A and Jones AM. The makings of a gradient: spatiotemporal distribution of gibberellins in plant development. Curr Opin Plant Biol. 2019:47:9–15.

Rizza A, Walia A, Lanquar V, Frommer WB, and Jones AM. In vivo gibberellin gradients visualized in rapidly elongating tissues. Nat Plants. 2017:3(10):803–813.

Rowe J, Grangé-Guermente M, Exposito-Rodriguez M, Wimalasekera R, Lenz MO, Shetty KN, Cutler SR, and Jones AM. Next-generation ABACUS biosensors reveal cellular ABA dynamics driving root growth at low aerial humidity. Nat Plants. 2023:9(7):1103–1115.

Roy S, Liu W, Nandety RS, Crook A, Mysore KS, Pislariu CI, Frugoli J, Dickstein R, and Udvardi MK. Celebrating 20 Years of Genetic Discoveries in Legume Nodulation and Symbiotic Nitrogen Fixation. Plant Cell. 2020:32(1):15–41.

Salvi P, Manna M, Kaur H, Thakur T, Gandass N, Bhatt D, and Muthamilarasan M. Phytohormone signaling and crosstalk in regulating drought stress response in plants. Plant Cell Rep. 2021:40(8):1305–1329.

Schiessl K, Lilley JLS, Lee T, Tamvakis I, Kohlen W, Bailey PC, Thomas A, Luptak J, Ramakrishnan K, Carpenter MD, et al. NODULE INCEPTION Recruits the Lateral Root Developmental Program for Symbiotic Nodule Organogenesis in Medicago truncatula. Curr Biol. 2019:29(21):3657-3668.e5.

Shakeel SN, Wang X, Binder BM, and Schaller GE. Mechanisms of signal transduction by ethylene: overlapping and non-overlapping signaling roles in a receptor family. AoB Plants. 2013:5(0):lt010.

Soyano T, Shimoda Y, Kawaguchi M, and Hayashi M. A shared gene drives lateral root development and root nodule symbiosis pathways in Lotus. Science. 2019:366(6468):1021–1023.

Spaink HP, Sheeley DM, van Brussel AA, Glushka J, York WS, Tak T, Geiger O, Kennedy EP, Reinhold VN, and Lugtenberg BJ. A novel highly unsaturated fatty acid moiety of lipo-oligosaccharide signals determines host specificity of Rhizobium. Nature. 1991:354(6349):125–130.

Stepanova AN, Yun J, Likhacheva AV, and Alonso JM. Multilevel interactions between ethylene and auxin in Arabidopsis roots. Plant Cell. 2007:19(7):2169–2185.

Tadege M, Wen J, He J, Tu H, Kwak Y, Eschstruth A, Cayrel A, Endre G, Zhao PX, Chabaud M, et al. Large scale insertional mutagenesis using Tnt1 retrotransposon in the model legume Medicago truncatula. The Plant Journal. 2008:54(2):335–347.

Timmers AC, Auriac MC, and Truchet G. Refined analysis of early symbiotic steps of the Rhizobium-Medicago interaction in relationship with microtubular cytoskeleton rearrangements. Development. 1999:126(16):3617–3628.

Trifinopoulos J, Nguyen L-T, von Haeseler A, and Minh BQ. W-IQ-TREE: a fast online phylogenetic tool for maximum likelihood analysis. Nucleic Acids Res. 2016:44(W1):W232-5.

Untergasser A, Bijl GJM, Liu W, Bisseling T, Schaart JG, and Geurts R. One-step Agrobacterium mediated transformation of eight genes essential for rhizobium symbiotic signaling using the novel binary vector system pHUGE. PLoS One. 2012:7(10):e47885.

Van de Poel B and Van Der Straeten D. 1-aminocyclopropane-1-carboxylic acid (ACC) in plants: more than just the precursor of ethylene! Front Plant Sci. 2014:5:640.

Verma V, Ravindran P, and Kumar PP. Plant hormone-mediated regulation of stress responses. BMC Plant Biol. 2016:16(1):86.

Wang KL-C, Li H, and Ecker JR. Ethylene biosynthesis and signaling networks. Plant Cell. 2002:14 Suppl:S131-51.

Wen X, Zhang C, Ji Y, Zhao Q, He W, An F, Jiang L, and Guo H. Activation of ethylene signaling is mediated by nuclear translocation of the cleaved EIN2 carboxyl terminus. Cell Res. 2012:22(11):1613–1616.

Xiao TT, Müller S, Shen D, Liu J, Adema K, van Seters A, Franssen H, Bisseling T, Kulikova O, and Kohlen W. Nodule organogenesis in Medicago truncatula requires local stage-specific auxin biosynthesis and transport. Plant Physiol. 2025:197(4). https://doi.org/10.1093/plphys/kiaf133

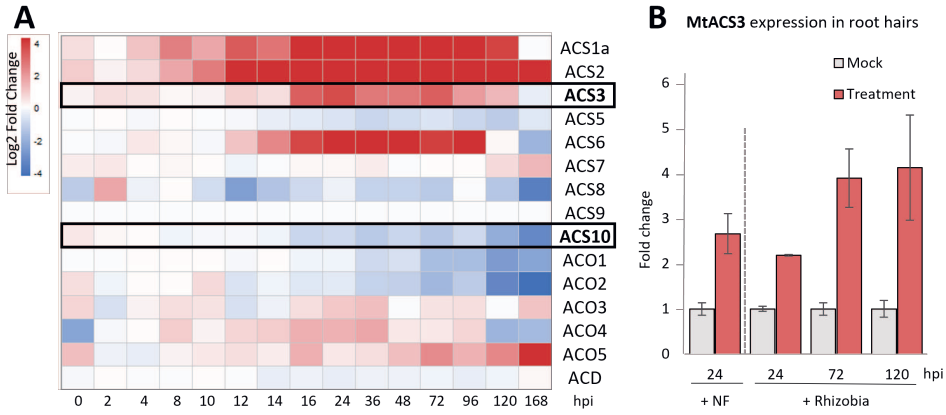
Xiao TT, Schilderink S, Moling S, Deinum EE, Kondorosi E, Franssen H, Kulikova O, Niebel A, and Bisseling T. Fate map of Medicago truncatula root nodules. Development. 2014:141(18):3517–3528.

Yanagisawa S, Yoo S-D, and Sheen J. Differential regulation of EIN3 stability by glucose and ethylene signaling in plants. Nature. 2003:425(6957):521–525.

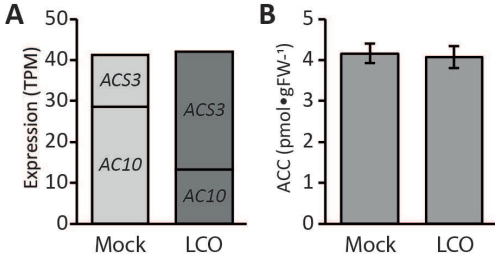
van Zeijl A, Liu W, Xiao TT, Kohlen W, Yang W-C, Bisseling T, and Geurts R. The strigolactone biosynthesis gene DWARF27 is co-opted in rhizobium symbiosis. BMC Plant Biol. 2015a:15:260.

van Zeijl A, Op den Camp RHM, Deinum EE, Charnikhova T, Franssen H, Op den Camp HJM, Bouwmeester H, Kohlen W, Bisseling T, and Geurts R. Rhizobium Lipo-chitooligosaccharide Signaling Triggers Accumulation of Cytokinins in Medicago truncatula Roots. Mol Plant. 2015b:8(8):1213–1226.

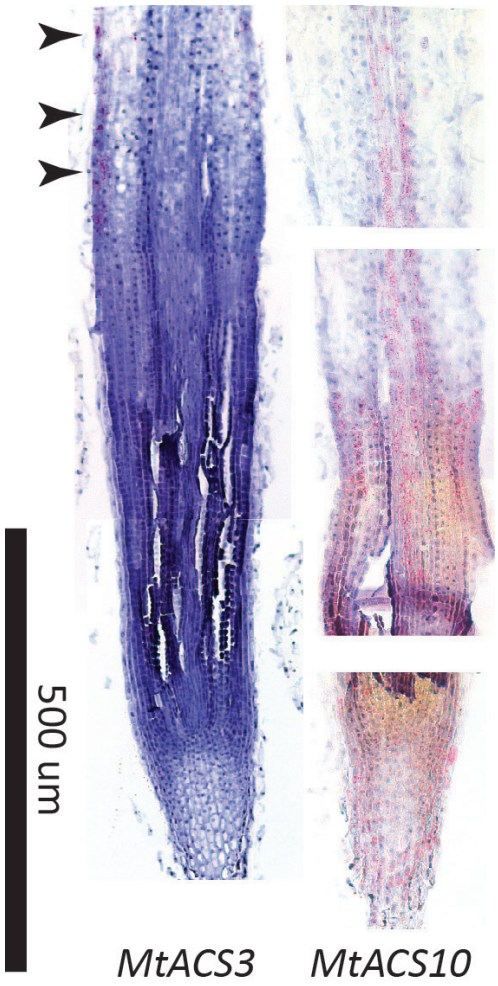
SUPPLEMENTAL FIGURES



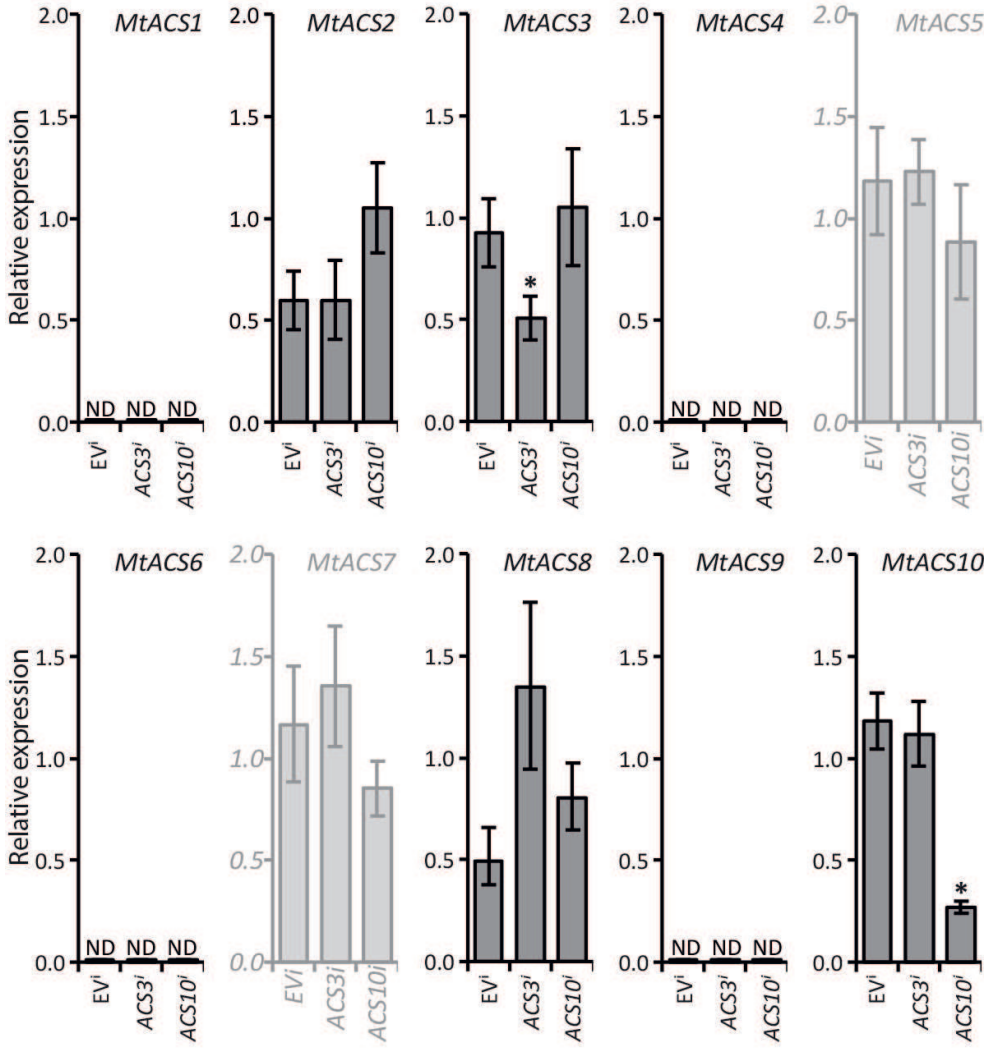
Supplemental Figure S1 | Ethylene biosynthesis gene expression from published datasets. For reference, the gene names were unified according to Supplemental Table S1. **(A)** From Schiessl et al., (2019). **(B)** From Breakspear et al., (2014).



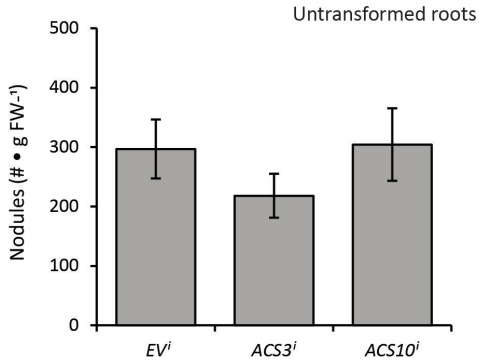
Supplemental Figure S2 | The effect of 3 hours LCO application on **(A)** cumulative *MtACS3* and *MtACS10* expression and **(B)** ACC levels. All bars represent means ± SE, (n=6).



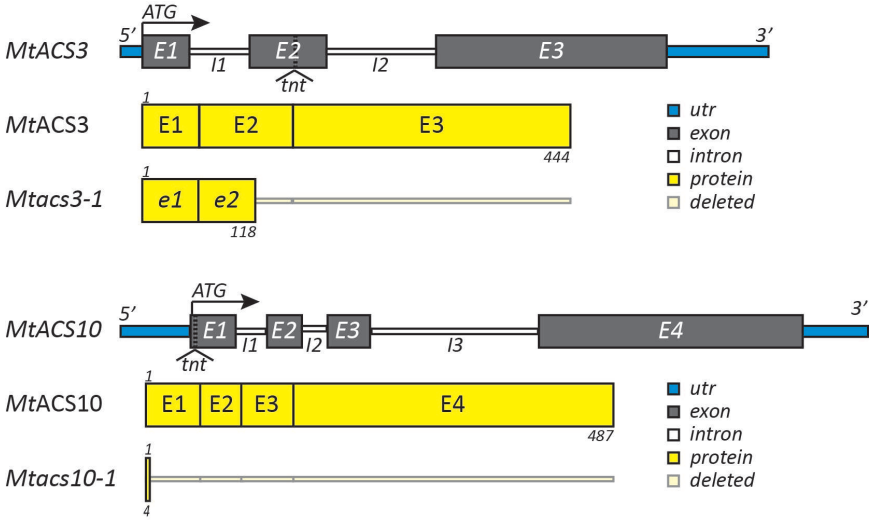
Supplemental Figure S3 | Longitudinal sections of the untreated Medicago A17 wild-type root tip with *MtACS3* or *MtACS10* expression pattern visualized via RNA *in situ* hybridization. Hybridization signals appear as red dots (arrowheads highlight distinct expression domain of *MtACS3* in the epidermis and c1 layer of the start of the differentiation zone).



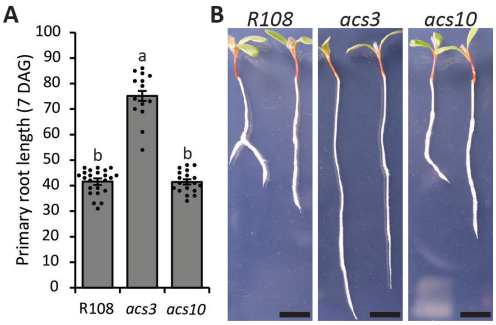
Supplemental Figure S4 | Silencing and off-target effects of *ACS3*^{i} and *ACS10*^{i} on *MtACS1*, *MtACS2*, *MtACS3*, *MtACS4*, *MtACS5*, *MtACS6*, *MtACS7*, *MtACS8*, *MtACS9*, and *MtACS10* expression under mock conditions. Bars represent means ± SE, n =3, bars labelled with an asterisk (*) differ significantly between the Empty Vector (*EV*^{i}) and *ACS3*^{i} or *ACS10*^{i} line ($P < 0.05$). Putative aminotransferases previously misidentified as ACC synthases in gray.



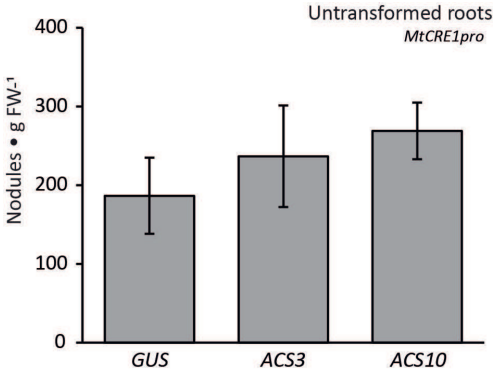
Supplemental Figure S5 | Numbers of nodules formed on non-transgenic roots belonging to the composite plants of the Empty Vector (*EV*^{i}) control and the two RNA interference (*ACS3*^{i} and *ACS10*^{i}) lines (n>10). Bars represent means ± SE. Non-transgenic roots were harvested based on absence of *DsRed* expression. No significant difference found ($P > 0.05$).



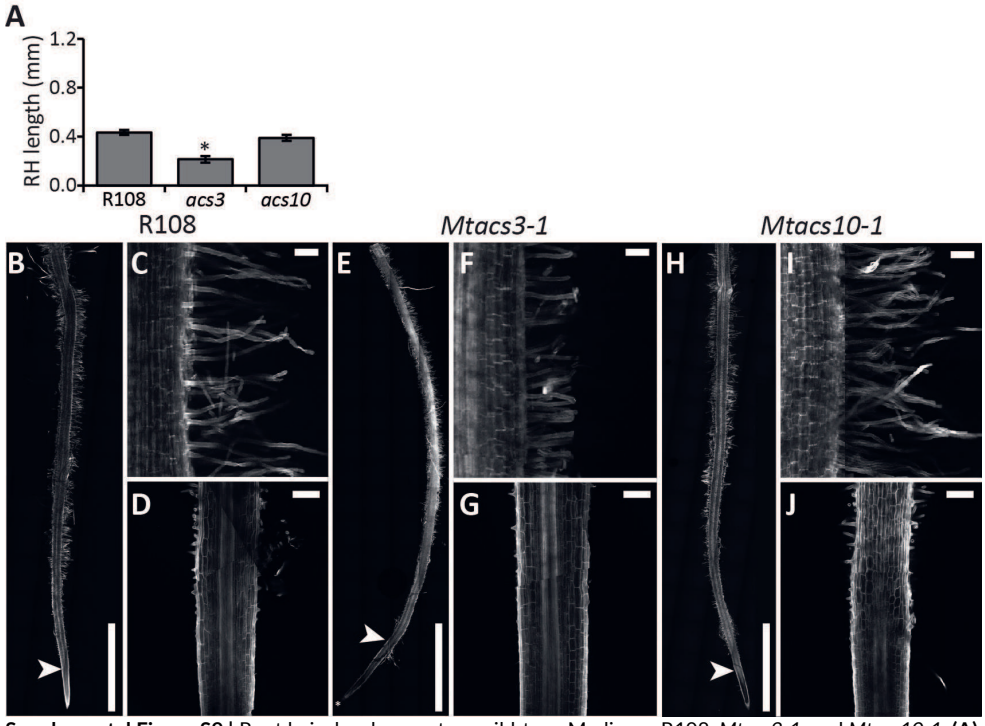
Supplemental Figure S6 | *Tnt1* insertion mutations in *MtACS3* and *MtACS10*. Schematic representation of the gene structures of *MtACS3* (top) and *MtACS10* (bottom), showing untranslated regions (blue), exons (grey), introns (white), and predicted protein-coding regions (yellow). Positions of *Tnt1* retrotransposon insertions are indicated by triangles. The *Mtacs3-1* allele carries an insertion in exon 2, resulting in a truncated predicted protein of 118 amino acids compared to the full-length 444 aa MtACS3. The *Mtacs10-1* allele carries an insertion in exon 1, leading to a severely truncated predicted protein of only 4 amino acids compared to the 487 aa full-length MtACS10. Deleted regions are indicated in light grey.



Supplemental Figure S7 | Loss of *MtACS3*, but not *MtACS10*, increases primary root length in Medicago. **(A)** Quantification of primary root length at 7 days after germination (DAG) in wild-type R108, *Mtacs3*, and *Mtacs10* mutant seedlings (n>14). Each dot represents an individual root; bars show mean ± SE Different letters indicate statistical significance, one-way ANOVA followed by Tukey's HSD test ($P < 0.001$ for *Mtacs3* vs. R108 and *Mtacs3* vs. *Mtacs10*; n.s., not significant for *Mtacs10* vs. R108). **(B)** Representative seedlings of R108, *Mtacs3*, and *Mtacs10* at 7 DAG. Scale bars, 1 cm.



Supplemental Figure S8 | Numbers of nodules formed on non-transgenic roots belonging to the composite plants of the control *GUS* (*MtCRE1pro::GUS*) and two independent ectopically induced *ACS* expression lines, *ACS3* (*MtCRE1pro::ACS3)* and *ACS10* (*MtCRE1pro::ACS10*) (n>15). Bars represent means ± SE. Non-transgenic roots were harvested based on absence of *DsRed* expression. No significant difference found ($P > 0.05$)



Supplemental Figure S9 | Root hair development on wild-type Medicago R108, *Mtacs3-1*, and *Mtacs10-1*. **(A)** Average length of the root hairs in the susceptible zone measured at roughly 4 mm distance from the root tip in wild-type R108, *Mtacs3-1*, and *Mtacs10-1*. Bars represent mean ± SE; asterisk (*) indicate significant differences ($P < 0.05$, $n>10$). **(B-J)** Representative confocal images of the Medicago **(B-D)** wild-type R108, **(E-G)** *Mtacs3-1*, and **(H-J)** *Mtacs10-1* root tips and root hair zones. **(B, E, H)** Confocal tile image of the Medicago **(B)** wild-type R108, **(E)** *Mtacs3-1*, and **(H)** *Mtacs10-1* root (Arrowhead points at first root hair, asterisk (*) in E marks root tip, scale bars 5 mm). **(C, F, I)** Confocal image of root hairs in the root susceptible zone of Medicago **(C)** wild-type R108, **(F)** *Mtacs3-1*, and **(I)** *Mtacs10-1* (scale bars 100 µm). (D, G, J) Confocal zoom in image of first root hairs of Medicago (D) wild-type R108, (G) *Mtacs3-1*, and (J) *Mtacs10-1* (scale bars 200 µm).

SUPPLEMENTAL TABLES

Supplemental Table S1 | Overview of ethylene biosynthesis genes with version ID and common name used in different publications. Based on Gómez-Fernández et al., (2025). "-" gene absent from the dataset, "nn" gene was not named with a number.

Common Gene Name	A17 (MtrunA17r5.0-ANR)	A17 (Mt4.0v1)	Breakspear et al. 2014	Larrainzar et al. 2015	Van Zeijl et al. 2015	Schiessl et al. 2019
ACS1a	MtrunA17_Chr8g0389311	Medtr8g101750	-	-	**ACS1**	*nn*
ACS1b	MtrunA17_Chr8g0389321	Medtr8g101820	-	**ACS7**	-	-
ACS2	MtrunA17_Chr4g0054371	Medtr4g097540	-	**ACS4** contig_12436_1	**ACS2**	*nn*
ACS3	MtrunA17_Chr6g0488041	Medtr6g091760	Mtr.20234.1.S1_at	*nn*	**ACS3**	*nn*
ACS4	MtrunA17_Chr3g0135251	Medtr3g103550	-	-	-	-
ACS5	MtrunA17_Chr5g0398431	Medtr5g011400	-	-	-	*nn*
ACS6	MtrunA17_Chr5g0400831	Medtr5g015020	-	**ACS8**	-	*nn*
ACS7	MtrunA17_Chr6g0475681	Medtr6g463260	-	-	-	*nn*
ACS8	MtrunA17_Chr7g0248681	Medtr7g079080	-	*nn*	-	*nn*
ACS9	MtrunA17_Chr8g0348351	Medtr8g028600	-	-	-	*nn*
ACS10	MtrunA17_Chr8g0387481	Medtr8g098930	-	contig_49261_1	-	*nn*
ACO1	MtrunA17_Chr2g0289341	Medtr2g025120	-	*nn*	-	*nn*
ACO2	MtrunA17_Chr3g0121141	Medtr3g083370	-	*nn*	-	*nn*
ACO3	MtrunA17_Chr3g0125201	Medtr3g088565	-	-	-	*nn*
ACO4	MtrunA17_Chr5g0439011	Medtr5g085330	-	*nn*	-	*nn*
ACO5	MtrunA17_Chr6g0488511	Medtr6g092620	-	-	-	*nn*
ACD	MtrunA17_Chr8g0393311	Medtr8g107670	-	-	-	*nn*

Supplemental Table S2 | Primers and Gene IDs (Medicago genome v4.01) used in this study.

Primer name	Sequence	Gene ID Mt4
MtACO1_Fw	ATTGGAGAAACTGGCAGAGG	Medtr2g025120
MtACO1_Rv	CCAGCATCTGTGTGTGCTCT	
MtACO2_Fw	CAAAGGGACCAACTTTTGGA	Medtr3g083370
MtACO2_Rv	GGGAGGGACATCTACCCAGT	
MtACO3_Fw	CCGGAACTTGTGAATGGTCT	Medtr3g088565
MtACO3_Rv	ATCTGCCATTGCTCAGGACT	
MtACO4_Fw	CATCGTTGTCAACCTTGCTG	Medtr5g085330
MtACO4_Rv	GGAGCAGGGTAAATGACAGC	
MtACO5_Fw	GATGCTGGTGGAATCATCCT	Medtr6g092620
MtACO5_Rv	TCCATTCTTGTCAGGCATCA	
MtACS1_Fw	CGGAGATGCTTTGCTTGTTCC	Medtr8g101750
MtACS1_Rv	CAGCACTCCACTCACTTTCATG	
MtACS2_Fw	TGGTTTCGTGTGTGTTTCGC	Medtr4g097540
MtACS2_Rv	TTGCCAGTGGACTCAACAAC	
MtACS3_Fw	TGGGCTTGGCAGAAAATCAAG	Medtr6g091760
MtACS3_Rv	ACCACCTTTTCCCCATGTTG	
MtACS4_Fw	TGCAACTTTTGAGGCAGAGA	Medtr3g103550
MtACS4_Rv	AATCCTTCGCATTGCAACTT	
MtACS5_Fw	CTGCTGGGCTGACATGAGTA	Medtr5g011400
MtACS5_Rv	CCCGATCCGTTCCACTACTA	
MtACS6_Fw	TCATGTTTTGCCTTGCTGAA	Medtr5g015020
MtACS6_Rv	AAGCCGATTCTGTGATTTGG	
MtACS7_Fw	CAACGGTTTTCTCCTGGGTA	Medtr6g463260
MtACS7_Rv	CGAAATAGAGCCACGAGAGG	
MtACS8_Fw	GGGTCTTCCGGGTTTTAGAG	Medtr7g079080
MtACS8_Rv	AAAATTCACGTCGCTCTGGT	
MtACS9_Fw	ATCCAGGATTCGTTCGTGAC	Medtr8g028600
MtACS9_Rv	CTAATGGGTTGGAGGGGTTT	
MtACS10_Fw	CGAACTCGTCATCAGACAGC	Medtr8g098930
MtACS10_Rv	TGACCGTAACCTCGTTCACA	
MtACD_Fw	AAACAAAGTGCGGAAATTGG	Medtr8g107670
MtACD_Rv	GGGATCTTGGTCAACGAGAA	
MtNIN_Fw	GGGAGAAAGTCCGGGGACAA	Medtr5g099060
MtNIN_Rv	GACACACACCGATGCTCTTTGC	
MtACT_Fw	GCAAAGGCAGAATATGATGAAT	Medtr2g008050
MtACT_Rv	CCACTATGACTGCCAGAACACTTA	
MtPTB_Fw	TGAACCAGTGCCTGGAATCCT	Medtr3g090960
MtPTB_Rv	CGCCTTGTCAGCATTGATGTC	
MtUBQ_Fw	CACCTCCAATGTAATGGTCTTTCC	Medtr4g091580
MtUBQ_Rv	CCCTTCATCTTGTCCTTCGTCTG	
MtACS3_RNAi_Fw	CACCCCCTACTTTGCTGGATGGAA	Medtr3g088565
MtACS3_RNAi_Rv	CAACTGCGGCTAATGAGCTT	
ACS10_RNAi_Fw	CACCAACCACCCCATACTTCATTCC	Medtr8g098930
ACS10_RNAi_Rv	CCAACCAAAAATCCTCAAGG	

Supplemental Table S3 | Gene probe sets used for ViewRNA *in situ* hybridization

Gene name	Gene ID Mt4 / r5.0	Assay ID according to ThermoFisher Scientific
MtACS3	Medtr6g091760 / MtrunA17_Chr6g0488041	catalogue number VF1-6000770
MtACS10	Medtr8g098930 / MtrunA17_Chr8g0387481	catalogue number VF1-6000771

Supplemental Table S4 | Sequences use for RNAi constructs.

MtACS3^{i} _fragment
CACCCCCTACTTTGCTGGATGGAAAGCATATGATGAAAACCCTTATCATGAATTAACTAACTCTTCTGGTG
TTATACAAATGGGATTGGCAGAAAATCAAGTTTCATTTGATTTGGTAGAAAAGTATTTGGAAGTGCACC
CGGAAGATTACAATGGTTTCAGAGAAAATGCATTATTTCAAGACTATCATGGACTTAAATCATTCAGAAC
TGCAATGGCAAGTTTCATGGAACAAATAAGAGGTGGTAAAGCTACATTTGATTCGGAAAGAATAGTCAT
CACTGCCGGAGCAACTGCGGCTAATGAGCTT
Note: CACC added for TOPO cloning

MtACS10^{i}_fragment
CACCAACCACCCCATACTTCATTCCTACACATTGCAATCTCTTCTAACTTATTCTTACACCTTCAAACATAA
GAAGAATTCCCACACAAATTCGCTTCTTCCTCGTACTTTTTTGTTTTTGTTTTTGAATCTCATTGATTTCA
TCACATTTTCTATTAAATTAATTAATATTTTAGACTATAGTAATTAATAATGGGACTTGTGAGCATGGACCA
ACCTCAATTGTTGTCCAAGATAGCCACTGGTGATGGACATGGTGAAACATCATCTTACTTTGATGGATG
GAAAGCTTATGATAAAAACCCTTTTCATCCAACCAAAAATCCTCAAGG
Note: CACC added for TOPO cloning

Supplemental Table S5. Sequences used for ectopic expression of *MtACS3* and *MtACS10*.

EC74831; *MtCRE1p-MtACS3-t35S*
EC74832; *MtCRE1p-MtACS10-t35S*
EC74833; *MtCRE1p-GUS-t35S*
EC74963; *MtACS3p-GUS-ACS3-3`UTR*
EC74964; *MtACS10p-GUS-ACS10-3`UTR*

MtACS3
ATGGGTCTTGAGATTGAACAAGAACACCCTTGTGTTGAACTTTCAAATATTGCAACTTCTGAAACTCATGGAGAAAATTCTCCATACTTTGCTGGATGGAAAGCCTATGATGAAAA
CCCTTATCATGAAATAACTAACCCTTCTGGAGTTATACAAATGGGCTTGGCAGAAAATCAAGTATCATTTGATTTACTTGAAAAATACTTGGAAGAACACTCAGAGGCTTCAACAT
GGGGAAAAGGTGGTTCAAGTTTTAGAGATAATGCATTATTTCAAGACTATCATGGACTTAAATCATTCAGAAAAGCAATGGCAAGTTTCATGGAAAAAATTAGAGGAAATAAAGC
AAAATTTGATTATGAAAGAATCGTCCTCACTGCTGGTGCTACTGCTGCCAATGAGCTCTTGACTTTCATTCTTGCAAATCCAGGAGATGCTTTACTTGTTCCAACACCATACTATCC
TGGATTTGATAGAGATTTGAGATGGAGAACTGGTGTAAACATAATTCCAATCCATTGTGATAGCTCAAACAATTTTCAAATCACACTTGAAGCATTAGAAACTGCATACAAAAATG
CAGAATCCATGAACATGAAAGTAAAAGCAGTACTTATAACCAACCCATCAAATCCATTAGGCATATCGATTCAACGTTCAGTTCTCGAGGACATTCTGAACTTCGTGACTCGCAA
GAACATACACCTTGTCTCAGACGAAATCTACTCGGGCTCAGTTTTCTCTTCACATGAATTCATAAGCATAGCCGAGATTCTTGAATCTCGTCAATACAAGAGCGCGGAAAGATGTC
ACATTGTTTATAGTCTTTCTAAAGATCTCGGTCTACCAGGTTTCAGAGTCGGAACAATTTATTCCTACAACGATAAAGTTGTTACAACAGCACGAAGAATGTCGAGTTTTACCTTA
ATATCTTCACAAACACAACATCTTTTAGCATCAATGTTGTCAGATGAAAGTTTCACTGATAATTACATCAAGGTCAATAGAGAAAGATTAAGGAAAAGATATGAAATGATCATTGAA
GGTTTGAAAAGTGCTGGAATTGAATGCTTGAAAGGTAATGCAGGGTTGTTTTGTTGGATGAATATGAGTCCAATGTTGGAAAGTAATACAAGAGAAGGTGAATTGAAGCTTTGGA
ATGAGATTTTGAATGAAGTTAAGCTTAATATTTCACCAGGGTGTTCTTGTCATTGTTCCGAACCCGGTTGGTTTAGGGTTTGTTTTGCAAATATGAGTGAAGAAACACTTGAACTT
GCACTCAAAAGAATACGTGATTTCATGAATAACAAGGACAGAAAGGATATAGGAATATAA

MtACS10
ATGGGACTTGTGAGCATGGACCAACCTCAATTGTTGTCCAAGATAGCCACTGGTGATGGACATGGTGAAACATCATCTTACTTTGATGGATGGAAAGCTTATGATAAAAACCCTT
TTCATCCAACCAAAAATCCTCAAGGTGTTATCCAAATGGGTCTTGCAGAGAATCAGCTTACTGCTGATTTGGTTCAAAATTGGATAATGAGTAACCCAGAAGCCTCAATTTGTACT
CTAGAAGGAGTACACAATTTCAAAGAAATGGCTAATTTTCAGGATTATCATGGTCTACCAGAGTTCAGAAATGCTGTGGCTAAATTCATGTCAAGAACAAGAGGAAATAGAGTGA
CATTTGATCCTGATCGTATTGTCATGAGTGGTGGAGCAACTGGAGCACATGAGGCCACTGCCTTTTGTTTGGCAGATCCTGGTGATGCTTTTTTGGTGCCTACACCTTACTATCCA
GGATTTGATCGAGATTTGAGGTGGAGAACAGGGGTTAAACTTGTTCCAGTTATCTGCGAAAGTGCAAACAATTTCAAATTAACAAAACAAGCCTTAGAAGAAGCATATGAAAAA
GCCACAGAAGATAACATCAGAATTAAAGGTTTACTCATAACAAATCCCTCAAATCCATTAGGCACAGTTATGGACAGAAACACATTAAGAACCGTTGTAAATTTCATCAACGAAA

AGCGTATTCACTTAATAAGCGATGAAATTTACGCTGCAACGGTTTTTAGCCACCCAAGTTTCATAAGCATAGCTGAAATATTAGAACATGACACAGACATTGAATGTGACCGTAAC
CTCGTTCACATAGTTTACAGTCTTTCAAAAGACATGGGATTCCCTGGTTTTAGAGTTGGTAATATACTCTTATAATGATACCGTTGTAAATTGTGCACGAAAAATGTCAAGTTTTG
GATTAGTTTCAACACAGACACAATACTTGATGGCGAAAATGCTGTCTGATGACGAGTTCGTTAAAAAGTTTCTTACTGAAAGTGCAAAGAGGTTAGCACAAAGGTACAGAATTTT
CACCAGTGGATTAACCAAAGTTGGAATTAATTGTTTACAAAGTAACGGTGGACTTTTTGTGTGGATGGATTTGAGAGGACTTCTTAAGGAAGCTACATTTGAATCAGAATTGGAA
CTATGGAGAGTGATTATTCACGAAGTTAAGATTAATGTTTCACCTGGAGTTCTTTTCATTGTTCTGAGCCAGGGTGGTTTAGAGTGTGTTATGCTAACATGGATGATAGAGATGTG
CAAATTGCTTTACAAAGGATTAGGTCATTTGTGGTTCAGAATAATAAGGAGGTTATGGTGTCTGAGAAGAACACTAAACCTTGTTGGCATAGTAATTTGAGGTTAAGCCTTAAAAC
AAGAAGGTTTGATGATATTGTAATGTCACCTCATTCTCCATTTCCTCAGTCACCTCTTGTTAAAGCCACTACTTGA

GUS (pICH75111)

ATGGGTCAGTCCCTTATGTTACGTCCTGTAGAAACCCCAACCCGTGAAATCAAAAAACTCGACGGCCTGTGGGCATTCAGTCTGGATCGCGAAAACTGTGGAATTGATCAGCGT
TGGTGGGAAAGCGCGTTACAAGAAAGCCGGGCAATTGCTGTGCCAGGCAGTTTTAACGATCAGTTCGCCGATGCAGATATTCGTAATTATGCGGGCAACGTCTGGTATCAGCG
CGAAGTCTTTATACCGAAAGGTAAGTCTTACTCTCTCTTTTTTGGTCTGTATTTTTAATTTTTTGAAGTATACTATTTGTACTGACGCTAATAATCTTTTTTCAGGTTGGGCAGGCCA
GCGTATCGTGCTGCGTTTCGATGCGGTCACTCATTACGGCAAAGTGTGGGTCAATAATCAGGAAGTGATGGAGCATCAGGGCGGCTATACGCCATTTGAAGCCGATGTCACGCC
GTATGTTATTGCCGGGAAAAGTGTACGTATCACCGTTTGTGTGAACAACGAACTGAACTGGCAGACTATCCCGCCGGGAATGGTGATTACCGACGAAAACGGCAAGAAAAAGC
AGTCTTACTTCCATGATTTCTTTAACTATGCCGGAATCCATCGCAGCGTAATGCTCTACACCACGCCGAACACCTGGGTGGACGATATCACCGTGGTGACGCATGTCGCGCAAGA
CTGTAACCACGCGTCTGTTGACTGGCAGGTACTTCATGCTTCAACGTGTAACTTAAGAGATACTGTGTGAAATTTTATATTTCCATACATTTGCTTGACCTTTGCTTTTTGTCAATTT
TTTTCCCCTTACAGGTGGTGGCCAATGGTGATGTCAGCGTTGAACTGCGTGATGCGGATCAACAGGTGGTTGCAACTGGACAAGGCACTAGCGGGACTTTGCAAGTGGTGAAT
CCGCACCTCTGGCAACCGGGTGAAGGTTATCTCTATGAACTGTGCGTCACAGCCAAAAGCCAGACAGAGTGTGATATCTACCCGCTTCGCGTCGGCATCCGGTCAGTGGCAGT
GAAGGGCGAACAGTTCCTGATTAACCACAAACCGTTCTACTTTACTGGCTTTGGTCGTCATGAAGATGCGGACTTGCGTGGCAAAGGATTCGATAACGTGCTGATGGTGCACGA
CCACGCATTAATGGACTGGATTGGGGCCAACTCCTACCGTACCTCGCATTACCCTTACGCTGAAGAGATGCTCGACTGGGCAGATGAACATGGCATCGTGGTGATTGATGAAAC
TGCTGCTGTCGGCTTTAACCTCTCTTTAGGCATTGGTTTCGAAGCGGGCAACAAGCCGAAAGAACTGTACAGCGAAGAGGCAGTCAACGGGGAAACTCAGCAAGCGCACTTA
CAGGCGATTAAAGAGCTGATAGCGCGTGACAAAAACCACCCAAGCGTGGTGATGTGGAGTATTGCCAACGAACCGGATACCCGTCCGCAAGGTGCACGGGAATATTTCGCGC
CACTGGCGGAAGCAACGCGTAAACTCGACCCGACGCGTCCGATCACCTGCGTCAATGTAATGTTCTGCGACGCTCACACCGATACCATCAGCGATCTCTTTGATGTGCTGTGCC
TGAACCGTTATTACGGATGGTATGTCCAAAGCGGCGATTTGGAAACGGCAGAGAAGGTACTGGAAAAAGAACTTCTGGCCTGGCAGGAGAAACTGCATCAGCCGATTATCATC
ACCGAATACGGCGTGGATACGTTAGCCGGGCTGCACTCAATGTACACCGACATGTGGAGTGAAGAGTATCAGTGTGCATGGCTGGATATGTATCACCGCGTCTTTGATCGCGTC
AGCGCCGTCGTCGGTGAACAGGTATGGAATTTCGCCGATTTTGCGACCTCGCAAGGCATATTGCGCGTTGGCGGTAACAAGAAAGGGATCTTCACTCGCGACCGCAAACCGA
AGTCGGCGGCTTTTCTGCTGCAAAAACGCTGGACTGGCATGAACTTCGGTGAAAAACCGCAGCAGGGAGGCAAACAATGA

MtCRE1$_{pro}$

GGAGCCTAGAACCAATATAAAGACTATTTTTATTGTCAAAATCTAGACTTATTAGAACTTTTTGTTTTCTAATTTCCAAATGAGTTATATTATATGAACAATTTTTTCTTGTGACAAC
TTGATTGACAATCAAATTATACAAAGAAATTTAGATAAATATACAAAAATCAAAGCAATAGAGAGAAAGTAGATAAATAATGTGAGTATGAGAGATAAAATTGTCACAAAAGTT
GTCAAAAATGATTGTTCAATTATCATTTTTCTTTTCAAATATTGAACTTTTTCTTACAAGATACAAATGAGGCATTCTTCGATAGTCACGAACACAAATGCATAGTGTTTGGACACAC
ACATATAATTTATATGTCATTTCTTGAAAATACTAAATTATTTTATTATCAATTTCTTTTTCATTATTAAAAAAAATTGTAATGTGTACGCGCCTAAACACTATCCATTTGTATTTGTGATT
TCTCGATGCCAAGACTAAACTTGATGGCAATTATTTTATTGAAAAATTTGATAGCAATTAGTACCACTTAAGAATTAAATTGGTGGTATATTCATCTATTCTTTCTTTAGAGATGTAAT
CATGCCTTTATTTGAATAAAGTGCCACAACGAGATAAGGTCACCTTAATAATTTGGAAGAGGACATTCTATCTTTCTTTTTTTAACATTACATTTATTATTTTTAATGAAGTCGCAAA
GATGTAAAAAAATTGTCCAAACAACAAAGTAGGGAACCAATCAAACCGTTTTTTCTTTTTCGTTGCAAAGTTAAGCTAAATGAAATGAGTAGCAATATTTGAATAACCATGTTGGA
TAACTTATGTGATAATTTTTTTTCTCTCTTTTTATTGGTCAAAATCAATGGAAAGAGAAAAATAAGAGAGAGAAAAAAAGTATAATATGAGTATGAGAGAGAAATTTGTAAAAAAA
AAAATCATAAAATGGTTATACAAATGTAATTTATCAAAAGAAATAACTAAATTAACTGTTTACCAACAATGAATGTATGAGTACTACTATACATAGATTATTAAGGATACATTAGTCAC
AACATTAATTAATACAAAACAACAAAGTTAATTAATGTTGATTATCCATTGAGGGTAGTTATGGAACTCTAAATAATAAAAAGGGTGGGTGTTTAGGGTAATTCACATGGGAATAAG
ATCCTAGCATTGTCGCAAGTCACTATCTTTCAGATTTGTGATTGTGATTTTTTCTATCTTTCTCTCTCCTCTGTTCCACTACACAACATTGTTTAACGTTGGAACACATAGAAATAGT
GAGAAAGACCCATTTGAAGATTCAACTCTAGAATGTGAAAAAGTTTCAATTCTTACATTCATTTTCCAAAGTTAGTAATAATAGCTTAACTGGGTCAGTTCATTCTCCATTGAAGCT
TCTTTTCAAAGTGCTAGTTGAAAAAAGATGCTAACTTGAGGAGAAAATATTGTTGGGTTGAAAGAGAGAGTTTTGACTTGACTAGAGTGTGAAGTAGAAGGGAGACAACAACAG
CGTATCAGAGAAAATCGAAAAAAAATTCGATTTAGTTTTAGTAATGTCCTGTTTTGTGCTTAAAAAATAAAGAAAGTAAAGTAAGAAAGCATATATAAAAAAGGTTGTATAAAAAAACA
AAAACTAGTACAGAAAGAGAAAGGTGCAGTGCAGTGCAGATCAAGACGAACCCAACCAACCAATCTGCTCTCAACTGTGTTCATTGTATTGTTTCTCTTTCCTATTCCTA
TACAGGTAAACATTGTTTCTCTAACTTACTCATTGGGTTCTGTGTGTGCTTCGCTGTTTTTTTTTTTTTTTTTTGTTTAGTATGGTTTGTGTGTTTTATTTTTGAGGTCAATGTTTGGTT
TTTTTGATTAAACAAATAAAAGTGTGAACGTGCTTTTGTTTGTCCTTGAGCAGAGACTGGTTTCACGCCTAGAGAGCATGATTAGAAGTCTCTTGAATTTTTAATTCATAATATTGG
CTGTTGAGAGAAAGGAGAAGGTGCTTCTTTGGAATTTGAGCTGTTTTTCCATCTTTTGAGAGCCATGCTTGTCTCTATTTCCCTTCTTTTTAATTCTTTCTCTCTTTTCTCTCACTTC
TTTTTTTTATTCTCTAATTTTCTGTGTAATTTTTCAAATGAATTTTCCTCAATCATTCATCATGCTTGGTTTTGCTTTTTTATTTGATTCTTGTTTCCTAAGAGGCTTTGTTTTGGTTTG
GTTTGAATCACGCAAAACAATGCTGTAATGATGCTGTACTTTGGTTTTTTGTTCTGTGAACTTTTTCTTTCAACCATCAGATAAAAGTGTCTGAGATTTTAGGACTACTTTACACTTTA
AGCAAAGTAAAAGCTACTTCCTATATATAAAAAAGTTTGTGCTCTGTTAATTGAGTATCTTTTGGGACTTCCATAGATATGAAGAAGAGTTCAGAGAAATAGGCAGAAAGAAAGAG
ATTTGGTGTTAGTTGGTGTTGTAAATG

t35S (pICH41414)

CTTCTCTAGCTAGAGTCGATCGACAAGCTCGAGTTTCTCCATAATAATGTGTGAGTAGTTCCCAGATAAGGGAATTAGGGTTCCTATAGGGTTTCGCTCATGTGTTGAGCATATAA
GAAACCCTTAGTATGTATTTGTATTTGTAAAATACTTCTATCAATAAAATTTCTAATTCCTAAAACCAAAATCCAGTACTAAAATCCAGAT

MtACS3p

AACCCCCTAATTTCCTAACAATAACAAATTGTTTGGTTCAGTTTTAACGTTAAAAATATGAATTGGACACTCAACTCAAACCAAAGACAATTGATCAGATTCAGTAATGAGTTTGG
TCAAATTAGCCCAACCCAACCTCCAAACACTCCTAATCAGAGGTGGATAAAGTGGTGGTCGTCAATGGTGGGTGGTGGTTAGGGTTTACTGGGGGTAACATGAGTGGTGGTAGA
AGGTGGTCAAAATAATGGTCAGATAGGTGATAAGAGTTGGCCAACAGTGGTCAACAACCGACCACTGGACGGTGGTACGCGTTGTGGACTCTATAATACTTCAATATATTTAATA
CTATAAAAAAAAAAAAAAAATTTAATATATTTAAGAGATTTTTTTGGTCAAGTAGACTAGTGGTTAGAATTCCCCTTTTTCAAAGCGAATAAGTGGGGTGTCCGGGTTCAAATCCG
GACCCCTGCATATAATAATGCATATCTTTTCCAACTGAGCTATGCTCACCGGAATTATATTTAAGAGATTTATTGTCGTAGTATGTTTATTTTTTTCATTTTTGAAAGGAGGCATTGTT
ATAGTTTAAAAAATCATAAAAAAGAAAAAAAATTGATTACGTTAGAAACTTTACAAAAGACTGATGTAATCCAAATTAGTTTTGCCAAAATACTAAAAGTGTACATTTTCTTCTTCT
AAAAGTTTGTCTTGTCACAAAAGAAAAAAGTTTGACATAGAAAATGAGAATTCATTTTTTTTAGTATCATTTCTAATAATTAAAATTCAACTGAAACAATTTCCAAGCCTATATTCGTT
TTGAAATTCTTCGTCTTTCGTCAGATTTATTACTTTTTCTTTTGTCGTTGCTCATTAGTTTATTTATTTATTTTTTCATGTTGACTTTTGCATGGAATACAACTTCTTCAATTTCTCA
AATTGAAAGTTTGTGTGAACCTGTGAATTATGTGCAAAGATGATTAAGCATTTCACTCATATTAGTATACTTATTAATTAATTTATACTTAATGAATAAGTATACGTAAATTCTCAATAA
GGTTGTGGTTCTCTTAATTTTTTAATATTAAAAGAAAAATAGCTTTCATTTTCATAATGATTTTTCTTAAAGGATTCATTTTCTTAATGATGTTATTTTGCTTTTAATATAAGCTCTCAT
TGAGCCACGGTGTATATATATATATTCCTAATAAGTAAACGTCCCATAAATGATATTTGTACAACCATTTTGTAATAACTTTTGTGACAACTTTCTCTCTCATACTCACATATGTTTTCA
TTTTATCTTTCTATTGCTTTGATATTTGTGCTAAAACTTATTTTTTCTTTATAAATTTATGGTTGTCAAATAAATTGTCTATCAAAAGGTTATTCAAATAACACACCTCTTTTATCAATAT
ACACCTCATCCTTAAAAAATAAAATCAATATACACCTCATAAAAAAGAACTACAATATGCCAAAAATAGCAAAAGTTAAACCATCCCTTTAAAGAATTTTACTCTCCAAGTCTTAA
CACTAGTATTAATCAAGCAGACAGCTATAATTACATATGTTTAGCTGTTTGTACCAAAAATATTAATATTGTTGGTTTTGAATTCCGGATAAATGTTCTGAGTCATAGGTATGCCCTTA
ACCCTAATGAATCCCGGAGCCCAACCCATATAAACTTTATCTTATAAAAAAAATAACTCTAGAAGTATATGAAAATTTTGTGATATTTTTCATTTAACAAGACAATATTATTTAAGGCC
AATGATAAAACATATAATTTTGAAATGATGAATTATGTGATATTTTCCATAGATAACTAATCTATTATTGAAATAATTGTAAATTATGATTATTGTTTGTGAGATGCCTATCATTTAGTTT
TATTGATCGCCTATAATTTTTTCCGATGATAGTTAAGTCTAGTAGAATGAGTTAGTCTCCTACAATTTCGCAAACCACAATTTGATTACTCATTATATAGTGGCACCCATATTTAGT
GTCCTATTTCTACTCAAATAAGATTGCCTCTAGATTGCCTTTCATTTCATCACCAAGGAAAAAATGGTGCTCCAAATATTTAGGAAGTGCACAAGTAAATGACTAATTAGAACAAG
GTGCACCACAATTAGATTAGATACAACTACAAACTTTAAAGGTTGTATATATTGAAGAGATTTGATGCATAAACCCTAAAAATATATTTGATTATACCCAAAAAAAAAATCATATAGA
AATTGATTGTATTGGATTGGATTGGAGTACTCATTGTGCTCCTTTATCCATAATCTTAATTTGTTCTTCAAATTAAGGGAGATGTTTTATTTAAATCTCACAAAACTCATTGATCAAT
TTTGTTGTTTGGTATTGATGCAAAAGTGGGGTCACAAGATATCCCAAGATGTTTCTCTTTTGTCATAGTTCTTCTTGCAAAATCAGAATCTTAATCATCTTTTGGGTTACATGGTATC
TTTGTCTGGTCATCAAAACTTAACTACTCCTTGCTAGAATTTTCTCCATTTTTTTCTTATATGCTAAACACCCAATTCAGATCTTAAATCATGATTTGGTTTCCACAATGGAATTTTTG
GTGTTTATTTGGCATTGGATTTGATTAGTCGTAGGTTGAGAATAGTAGAATTGTTTTTTCAAGAACTCAACCACCTGTAAATTTAATAAACAAGTATGAAACCATAAATTAATTAAG
CCACAAACAAAGTTATGAACAATCTGCTCCATTAAAACAAAGGTTATTTTATTTTTTTATAAAATGTTGAAGTAAGTTGAACCCGAAGAGTTGCGAGGATCAAGAAAAATTTCAAG
ATTCAGAGACGACAAATACAAATACTTACATATTATCTATTAACGATGTTCTCAAACTAACAACGAACATCAAACTCAGGTCTTTGAAGAATCTGAGCGGAATCCTTATCAACGAA
GATCAAGACAAATTTCAAGATTCATAGACCACAGATACAAGTACTTACATATCATCTAATAACGATGTTCTTAAACTAACAATGAACATCAAACCCAGATCTTTGAAGAATCTGAGC
GGAATCCTTATCAACTTCACCAAACAAAGTTGGTAATTTGTAAGCAAAATTAGGGAACTATAATAGCAACTTAATCAAAACATCCATCTCCAAACAACCACCATATAATTTCCGTG
TGGATCTGTCACAGTTGTCACAAGTAGCAACAGATAAACTATGTGAAATTAAATTAAATTAAAGCTAAAAGGTAAAGCTTTGATTTGCTTACACTACAAAATCAATAATTATATTA
TATACATTTTATGCATTATTTATGCTGAATAAGCCAGCCAAGTTGAAAAGGGACCTTTTTTCCTTGTCTTTAAGCCTCAAAGCCTATATATAGAGAAATTTCACACAAACTTTTCATT
CAACTCTCTCTCACCCATTTTCTCTCTCTACCTATCTAACTACCTTCTCTCTCTTTCCCCTCCTTCTAAAAATAGTTAGT

MtACS3-3`UTR

TTTTTACCATCACTCTACCATGATTAAATAGTATTTTTTTAGGGTCAAATCAATCAGAAGAGTCAGAGTTTGAATCATGACAAAAATAATCATTTGTCAAGTTTTACTTATCTCCCAA
ACGAATTCCGATTACCAGTTTCTTTCGCCTTGAAATGGGAGGGATAAGACCAAAACATTAGAGTCTTTGTTTATTTATTTTTATTTTTTTTCAAATGGTTGAGGCCATTTTGGTATAT
TCTTTTAAAAGCACATAATGTAAAGGTGAATATTCAATTGAAGTGTATCAATCCCTTGTTTGCTTGCTTAATTTACTTAAGCAAGTTTTTATTACAAAATTAATCTTTTGTATTATTTTG
TTTTATAATGAATGAAATTTGTTACTATTAAATTATGTGTCTTAATTACTACATTATTTGTATGATGAGTTAGTATAGTGTCTATATCATCTTGAACATCTTGATTGGTTGTGTAGTAAAA
AATTAATCCTTATTGAATGAGTATTGGTCTTGAATTTATCAACATCATTTGACTTAGTAATTACTGATTTAATTTCAATTCATATCATAGTTGACTTAGAGTTTGGAGTTTAAAGTACGT
AACAGCTAGCAAAACACACCAACCAGTTATAGCTTTCAGCTTCATGAAAGGGTTGGAAACTTGGTATGAATAAAATATTAACATGCCTCAATCCCTACAGTACATGATGATGTGCA
TATTTTTAATTCAAGTCTAATTGCCATGTTTATTGGATAATTAGTAGTATATAAGTATGATTTGATTATAACATTGTAATTAATTATAAGTATAGTATAAGTATATACTTTGTATTAATTCC
TATAGTAGTGTCGGGAGATATTCTAGAACAAATTAAAAATGGTCCACCCCACGTGATGTACACTGCTTGAATTTATATTTGTATATATAACTGCATCTCTTTGTTGTTGTTGTTGTGTT
GTGTGTGTGTGAGGAAAATGATGTTGATATTGATTGATATTAAAAT

MtACS10p

GACCTATCGTGATTATGGTAAACACTTGAGATCCAACTCACTAAACGATAACAATATGATTTAAATTCATGTAATTGCCATCAAATTTATATGAGATTAACATGATTTAGTGAAACTG
ATGTGAATTAGGTTCAGCAGTGTATTGTGTGTGTGGAAAGTGGAAACCATAGTATTCAGTTCCACAATTCTGTCACCGACCAACAAGGGTAAGCAGTGCATATGGCGCCAACAAC
ACGGCAATAATTCTATTCCAATTGTGAGTATCTTTGTTAGGTTAGTACTTAGTGGTTTGGTTACACTCTCTTTTCATTTCATTTCATGGGATGGGACATAATGCCACTTGTGTTTTCAT
TTAATGCTCACAGCACAAAGTCACACTCTTCTTCTTCCAACAATTCCACTCTCAATTAATACTACGTGTCAAAAATTGAAACTGTGACCCACCTTTGTAGACAGAAATAAATGAAT
ACGTCCCATGAAAATCTATGGGTTGGTCTGGTGTAGGGTTGTACAAATCCCTCTCATGAAGGGCATCACCCTAATATAGAAAAGATATGAGGACAAAAGGTTGTCCAAGAACATT
TTTTTTACGATCGAGTAAATACCAATATGGACTCTATCTTTATAAGTTATTATCAAATTAGTTTCTCATTTTATTAATTTTTTAAAATAGTCCCTCAAAAGTTTCTCAATTATATCATTGC
CTCTATGTTTTTGTCATATGACGAGGAATTTAATTGACTTAATTGAAAAACTTTTGACAAAGACATGAACTATTTTTAAATATTCTCAAAAGCTTCCAATTAGATCCTTTTCTATTTGT
TTTTGTCACAAGACGATAGACTTAATTGAAAAATTTTTGACAAAAATAAAAACTAATTTAAAATATTAATAGAGTCAGAGACTGATTTGAAAATTACTTACAAAGATAGAGAGCAA
ATTAACAGATTACTCATTTTTTTACTTATACATTTGATAATATATATTATGAAAAAAAGTCTCAAATAATTTTATCTATATTTAATTGGAGTTTTTGATATATTTAATTTTCACCAAATTAA
ATATAGATATTGAATCACTATCTTAACTTGCTTAAGCACTCAATCAAATTTTTCATCAGCCTTTTTGTCATGGTAGGTGGCATAATGCTTTCTTCATTCTTTCAATTTATTTTTTTCCAA
TACATATATATTTATGATTTAATTTGTGTAGGCCGATAATTGTAAAGCTTTTACACATTTAAATTACTCTACAATGTCCTCTTTTTAGGTTGTTCCTAAAACATCTCTACACATATCTAT
TTAACATTAATTTGATTAGATGCATATAAAAATGTTTTAGACTTTTTGTGCATAGAAATGAGTCTCTATATTTACATATTGATTTAAGTATTTGTCAATAAAAACTATTATATATTATTTT
AGTGCGTGTTGTGTAGATGGATTACTAACTACGAAGATATCTGAATCTACATTCATATCCCCCTCATTTTTTCAAATAATTATTGGTGCTCATGCCAATATCTATAATGGTTTTTTACT
CGACCAATTAATATAATTTGGTTTGCAGGTATCAAATAAATAAATAAGGCAACTTCTTCTGCGGTATACCTCACAAATTGGGGTGTACCGGTACTCCTTGCTTCATAAATTGATAAA
TTAACGATCATTTTAGAAAGCAAAGAGCTATTTGTCAAGGTTATTCTTTAGAAAACATCATTTCATAAGAAAAAAAAAACATAATTTCATTGTTTATAAGAATCATATTAAATTTTTAA
CATTTTCTCATAAAAAAATAATCAATAATCTTCATTATGAGTAATAAAAATTCAAAAGAATTTAATTTAATTTCATCGACATTTTGGTAAATAAGATTTAATTTTTATAATTGTCAATGAT
AGATTTCTTCAAAAATTCATCTAATTTAATTAATAATTTAACAATTTGGTGTATCGGTACACTCAAATATTTGGGTGTACCATATAATCCACTAGTCTCTGCCCTATAAGACACTCTCA
AATAAACTATAAGATTATGAGTATTTTAAAACAATCTCTATTCTCTATATTTAACCACTTACTATTAATTTTACATTAATTAAGTCATTAAATTTTAACAATTTACTATCGATTTCCCGCA
AAAAAAAAATTAATATCAATTTAATTCATAAATTTATCATTCACCATCAAATTAATCATACCATATTTTTTAAAGAACTAATATAAGTAAGATTTGATAAAGTCAAAAATATTTTTAAAA
TATTCATAATGCTAATTACCTATTTGAGAACGTCCTATAAAGTCTGGAATTAATTTGTTGATTTAGAGACTAATCTTCATCTCTGCACCCCTGTCCAATCCATAACTCTTTCCCATCTT
TATGACCCTGGCCACATCAACTTCCACCAACAAATAGATAACAGATCAGAGCGTACGCAAAAATACAAATAAGTAGGTGAAACTATATGTCAATTGTATAATAAATAACAATTATA
CTATTTTTTTAAATAACAATTATACTAATATTATTATACTTTTTTTTGGGTGTACATAATAATATTATTATACAAATAAAAAGTAATATTCCGTCTTATGATTATTAAAAAAAAACAGTA
TTATGCATAGGAACGTATACAACTGTAGTACAACACAAAACCTGAATACAAAGGTCATCACAAACTTGAAGTCTAATATTTTTCTCCCACTTTGTTGACACAAAAACAAAATACAA
GTGGCATTTTGCTATTTGAACGACAATTATTACTAGTTATTGTCTGCCTACCAGCATCTGTCTAATTTTCAACTGAAGCTGCTGGCGTGGCATCATAGCTGAAAATGCTGCCAGAGT
CCAAAAATAATTGCTCATTCACTATTATTAGGAGATATCGTTTAATCGATTGATTAGACTCAAATGATTAAAGAGTTTTTCTTAAAAGGAAGTGTTAGGTTTTGTTGGAGAAACTCTT
TAAAGATTTAAATTTTTTTATTTTTTATAAAAGGATAAATATTTTGATTTTCAATGCATTAATTACATATATTTTTCATAAAAAAATTATCATTTAAAAGTATTAAAGAGTGTCCGGAGAAA
CTCGTCCATATTTTCCAAATAATTTTTGGATTTTGGTTTCCTATGACCCATTTTTCCTAAAAATTAAAGGCTTAACAAAAAAATATATTTGCTAGCTAGACTCTTACGCTCTCACCAT
TACCTTCATGTGAACCTTCCTTTTTTATATTTGGACTAACTAAAAAAAATATTCTTATGTCTGAAAATCAAAATATATTTGTGCATATGCGCATGCTATAAATAACCACCCCATACTTCA
TTCCTACACATTGCAATCTCTTCTAACTTATTCTTACACCTTCAAACATAAGAAGAATTCCCACACAAATTCGCTTCTTCCTCGTACTTTTTTGTTTTTGTTTTTGAATCTCATTGATT
TCATCACATTTTCTATTAAATTAATTAATATTTTAGACTATAGTAATTAATA

MtACS10-3`UTR

ATTATTTAGCATATAGAAGAAGTAATAACTGATAATTTTTTTTTTTTTTTTTTGGTTTGTTGATTTATCAGTTTGGGAAACCACATGTTGTGTTCCTTTAATGTTGGGGGCATTTTTTT
CTTTTTTCACTTACTGAATCACATTGGTTTGTGGTGAAAGAAGTTCATAGAATTTTTAAGGTCAATGATGAATGAATTGTTAATAAACTTCTTCAATTATTCTACATGTTTTATTGTTG
TTTTATAAACTGTTAAAGATAAATCAACTTTAGATTTGTACTAGTATGATAATGATACATATCAAATAGTATAAAGGAATTGAAGTGGATGTATGAGTAATTCACTAAAGATAACAAG
ACTACTGAATAAACATTCAATTTTTTATTTTTGGAATAAAGAAATTTACAAACAAGAAATTGCAGGGAGATTTACTCAAAAATATTTGTATTGAAAGTGATTCAAGTACTACCGTCC
AACTTGCAAACCATCACCATGTCCACATGATTCACTTTCACGGTTTCTTTGCTATAAGATGATGCAAAGATTCAGTCTGGTCAAAAAATTGTAAAATACTATTAAAACGCAGTCTCA
ATTTGGATCCGTTGAATGAAAATTTCACAATTCATATTCCAAATTCATAGTTTGAATATGACTTAAGGTTGCACAAATTGAATATGAGCAGTTTCATTTGCTAAAAAAAATAAGGAATC
GGAATTCGGACATGATGACAACTCTTACGGTACCCTATATGATTTACACATTGGTCACGAGATTGGCCAATATAGGGGCAACGAACTGGATAATACATAGGACAAAAGACTAAAC
AAAAACATAAATTCTTGTCCTTAATTTTTTTTTTTATTTGGAATACAGTCTACTTGGTTGGAAGGTTATGTTTGACTTCAACTCAAGATTTTTTTTCTTTTTCTTGAAGAGGCTCAAG
AAATATTTCTTGAGTTGAGACACGTGAAATAGGTCGCGGAGGCAATGTTAACTGATACC



Chapter 4 |

Establishing robust workflows for single-nuclei and single-cell RNA sequencing in the model legume *Medicago truncatula*

Sophia Müller, Kelvin Adema, Amber van Seters, Jana Witmer, Carolin Grones[#], Wouter Kohlen[#]

# contributed equally to this work

Manuscript in preparation

ABSTRACT

Legume nodulation relies on highly localized cell state transitions in the root, but in plants these events remain difficult to resolve at single-cell resolution because of the cell wall and the fragility of isolated cells. Here we systematically benchmark single-nuclei and single-cell RNA sequencing workflows for *Medicago truncatula* roots and nodules, comparing multiple nuclei (N1–N3) and protoplast (P1–P4) isolation protocols across two high-throughput platforms (10x Genomics Chromium and BD Rhapsody™). We define critical quality-control checkpoints from tissue dissociation to data filtering, and show that small changes in buffer composition, digestion time, and size selection have large effects on yield, viability, and transcriptome quality. Nuclei isolations consistently suffered from RNA leakage, high ambient RNA and low mapping rates, such that stringent filtering retained only a small fraction of reads with limited cell-type resolution. In contrast, optimized protoplast workflows produced higher mapping rates, more genes and Unique Molecular Identifier (UMIs) per cell, and robust recovery of major root and nodule cell types, particularly when combined with BD Rhapsody. Nonetheless, nuclei can be integrated to contribute complementary information, such as RNA velocity or underrepresented populations. Our work provides a practical decision framework for choosing between single-cell and single-nuclei strategies in legumes and offers broadly applicable guidelines for generating high-quality plant single-cell transcriptomes.

INTRODUCTION

The symbiotic process of nodulation in legumes has long intrigued plant biologists, both for its agronomic importance, enabling atmospheric nitrogen fixation and reducing fertilizer demand (Jhu and Oldroyd, 2023; Roy et al., 2020), and for its remarkable developmental plasticity. Nodulation is initiated through a highly specialized interaction between legumes and soil-dwelling rhizobia (Roy et al., 2020). Under nitrogen-limiting conditions, plant-derived flavonoids trigger the production of Nod factors in compatible rhizobia, which activate early nodulation regulators such as NODULE INCEPTION (NIN). These early signaling events induce root hair curling and divisions in internal root layers, enabling rhizobial infection and ultimately nodule organogenesis. Within mature nodules, rhizobia reduce atmospheric nitrogen (N_2) into ammonium (NH_4^+), which is exchanged for plant-derived carbon.

The model legume *Medicago truncatula* (Medicago) forms indeterminate nodules sustained by a persistent meristem (Kohlen et al., 2018; Timmers et al., 1999). Notably, the nodule meristem originates from differentiated root tissues that must undergo extensive developmental reprogramming. Cell divisions first initiate in the pericycle, then expand into the endodermis and inner cortex, with the future meristem emerging from the third inner cortex layer (Xiao et al., 2014). Despite decades of research, the molecular basis for why these specific tissues retain competence for dedifferentiation, and how they differ from surrounding cell layers, remains poorly understood. Bulk RNA-seq cannot resolve the small populations of responsive cells involved in early nodulation, making single-cell RNA sequencing (scRNA-seq) a powerful approach to dissect these cell-type-specific events with higher resolution.

Modern single-cell platforms achieve this resolution by combining high-throughput compartmentalization, via microfluidic droplets (10x Genomics Chromium) or microwells (BD Rhapsody™), with barcoding strategies that assign individual transcripts to their cell of origin (Adema et al., 2024; Seyfferth et al., 2021). Unique molecular identifiers (UMIs) further mitigate PCR-induced biases. These technologies routinely profile tens of thousands of cells per experiment and capture thousands of transcripts per cell. Although originally optimized for animal systems, they are increasingly being applied to plant tissues (Grones et al., 2024; Lv et al., 2024; Sun et al., 2024).

A major bottleneck for plant scRNA-seq, however, is the efficient isolation of intact single cells. The plant cell wall demands enzymatic digestion to release protoplasts, but digestion efficiency varies across tissues and species and often requires extensive optimization (Ren et al., 2021). Moreover, cell wall removal can induce substantial transcriptional artifacts, affecting hundreds to thousands of genes in *Arabidopsis thaliana* (Arabidopsis) roots (Birnbaum et al., 2003; Denyer et al., 2019). Many specialized cell types remain inaccessible to enzymatic release, and even successfully isolated protoplasts are fragile, prone to bursting and RNA loss during downstream handling. Their large size (up to 100 µm in diameter) also exceeds the recommended limits for microfluidic platforms such as 10x Genomics (~30 µm), increasing the risk of clogging and sample loss. Microwell-based platforms like BD Rhapsody can mitigate size constraints to some

extent, yet proper bead-to-well coverage remains challenging when handling buoyant or irregularly sized cells.

Single-nuclei RNA sequencing (snRNA-seq) provides an alternative that avoids cell wall digestion. Nuclei can be isolated rapidly, more uniformly across tissues, and even from frozen samples (Conde et al., 2021; Thibivilliers et al., 2020). Transcriptomes derived from nuclei correlate well with those from protoplasts and bulk tissue (Farmer et al., 2021), although nuclei typically yield fewer detected genes per cell, up to ~4,000 fewer (Bakken et al., 2018) and are more susceptible to RNA leakage and ambient RNA contamination (Grones et al., 2024). Despite these limitations, both protoplast- and nuclei-based approaches have been successfully applied to plant systems, including Medicago (Lv et al., 2024; Sun et al., 2024).

To date, studies dissecting early nodulation in Medicago have all relied on nuclei-based scRNA-seq using 10x Genomics workflows, typically focusing on early signaling responses up to 96 hpi (Cervantes-Pérez et al., 2022; Liu et al., 2023; Pereira et al., 2024), and most published workflows rely on fluorescence-activated nuclei sorting (FANS) and centrifugation steps adapted from a commercial nuclei isolation buffer (Thibivilliers et al., 2020). The only protoplast-based dataset available is from fully developed nodules at 14 dpi (Ye et al., 2022).

Here, we present an optimized protocol for isolating both nuclei and protoplasts from Medicago roots and nodules, combined with high-throughput single-cell transcriptome profiling after rhizobial infection or Nod factor exposure. We systematically benchmark nuclei- versus protoplast-derived datasets across two major platforms (i.e.10x Genomics® and BD Rhapsody™) and provide detailed troubleshooting guidance for both approaches. Contrary to earlier reports, we found that nuclei preparations yielded very few high-quality single nuclei and lacked clear cell-type resolution. In contrast, protoplast-derived scRNA-seq produced more cells with deeper transcriptome coverage and robust cell-type assignments. Together, our results offer a practical framework for selecting between single-nuclei and single-cell approaches and optimizing platform choice, enabling more efficient and reliable generation of high-resolution transcriptomic datasets in Medicago and other plant systems.

RESULTS

Overview of workflow development

To establish reliable single-nuclei and single-cell RNA sequencing (scRNA-seq and snRNA-seq) in Medicago roots and nodules, we compared multiple strategies for isolating either nuclei or protoplasts, combined with two single-cell platforms: the microfluidics-based 10x Genomics® Chromium system and the microwell-based BD Rhapsody™ platform (**Figure 1**).

We assessed each step of the workflow and provide guidance on key considerations, including (1) nuclei or protoplast isolation, (2) sample quality control, (3) sample loading onto the chosen platform, (4) cDNA synthesis and library construction, (5) sequencing and downstream data processing, and (6) stringent data quality evaluation.

During optimization of the single-cell workflow, protocols were adjusted in response to issues encountered at various downstream steps (**Figure 1A**,**B**). The position of blocking, and thereby the relative progression of each isolation method, is represented by arrow length (**Figure 1B**). The key modification in the nuclei isolation methods (N1–N3), the protoplast isolation methods (P1–P4), and the single-cell platform, is highlighted (**Figure 1B**).

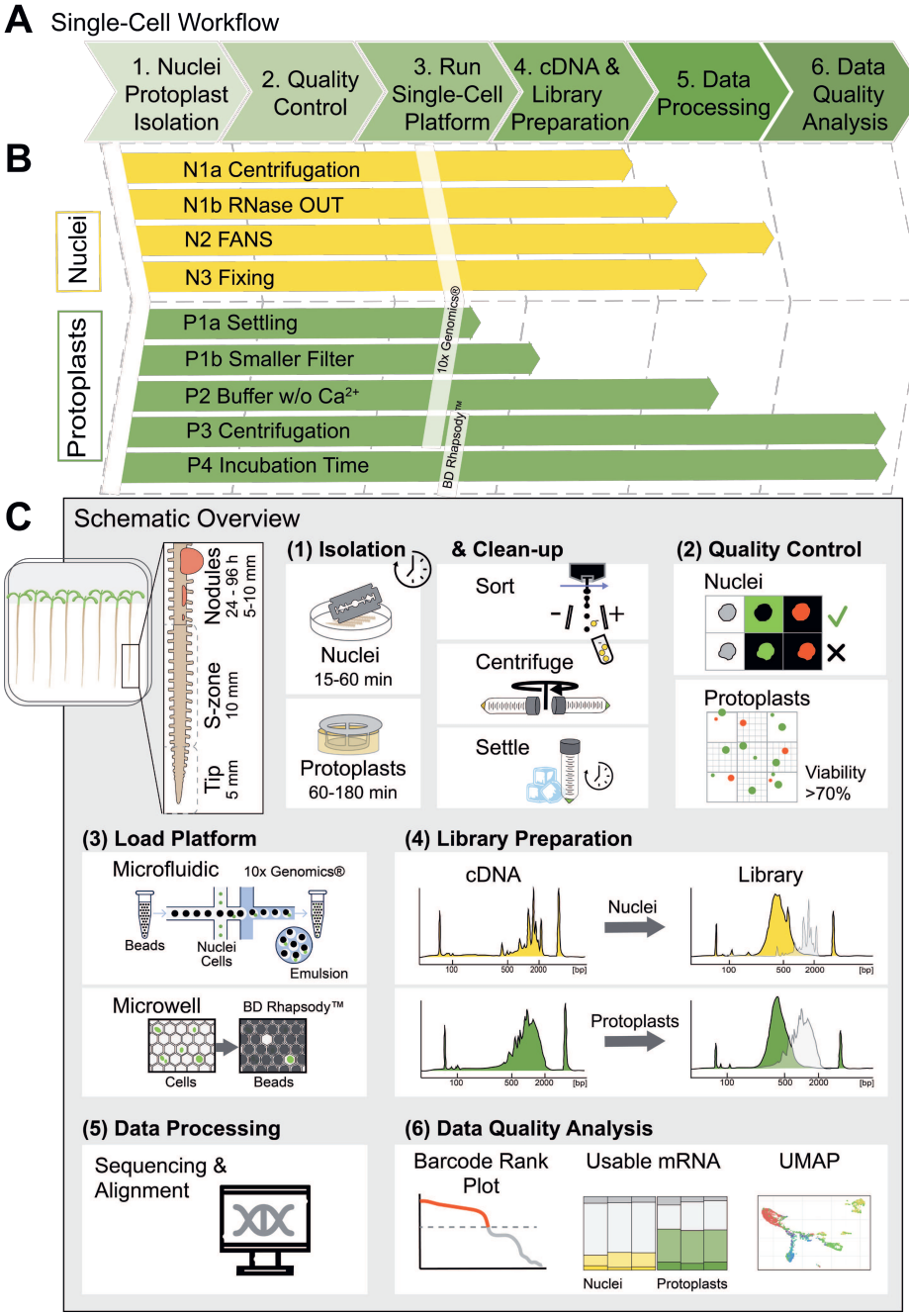


Figure 1 | Schematic representation of (A-B) optimizations tested for each step during sample preparation and processing and (C) a detailed workflow. The progression of each isolation method along this workflow is indicated by the length of the arrow (nuclei or protoplast, N1-3 yellow, or P1-4 green, respectively). The prominent changes in the protocol compared to the previous method is noted as well as the switch of the single-cell platform. An exemplary depiction of the sample preparation and each step is shown.

STEP 1-2: ISOLATION AND QUALITY CONTROL

Nuclei isolation optimization and comparison of buffers

Fast isolation and high purity are crucial for successful single-nuclei RNA-seq because they help preserve the integrity of biological data and minimize technical artifacts that can compromise downstream analyses. Rapid isolation limits the time nuclei are exposed to unfavorable conditions, reducing the risk of RNA degradation and stress-induced changes in gene expression. Delays can cause transcriptional artifacts or the loss of low-abundance transcripts, which bias results and mask true biological states (Kersey et al., 2025). Furthermore, high purity ensures that the nuclei preparation is free of debris, ambient RNA, and damaged nuclei, all of which can contaminate sequencing libraries. Impurities such as ribosomal and mitochondrial RNAs can overwhelm datasets, obscure cell type identification, and necessitate costly resequencing. Clean, intact nuclei yield more reliable, reproducible data and allow for accurate integration with other omics information (Nadelmann et al., 2021). To identify optimal conditions to obtain high purity nuclei, we compared nuclei quantity, purity and RNA yield between two nuclei isolation buffers (Nuclei Isolation Buffer (NIB; methods N1 and N2) (Thibivilliers et al., 2020) and Galbraith Buffer (GB; method N3) (Kao et al., 2021)) with and without further purification using fluorescence-activated nuclei sorting (FANS) (**Figure 2**). FANS analysis distinguished subpopulations cluster 1 (C1, red) and cluster 3 (C3, green) based on forward scatter (FSC-A) and propidium iodide (PI) fluorescence. Nuclei (C3) were more clearly separated from background events when isolated with NIB, consistent with reduced debris (**Figure 2A-B**). RNA yield per milligram of fresh weight was similar across buffers, independent of FANS (**Figure 2C–D**). Addition of RNaseOUT during isolation with NIB buffer further increased RNA yield (**Figure 2E**).

All three methods (N1–N3) yielded intact nuclei, confirmed by PI staining and fluorescence microscopy (**Figure 2F**). Isolation efficiency was broadly comparable, although method N3 may result in some nuclei loss due to the post-sorting centrifugation step (**Figure 2G**). GB is further limited for use with the 10x Genomics® workflow, as its high $MgCl_2$ concentration (45 mM) may interfere with downstream enzymatic reactions (per manufacturers recommendations).

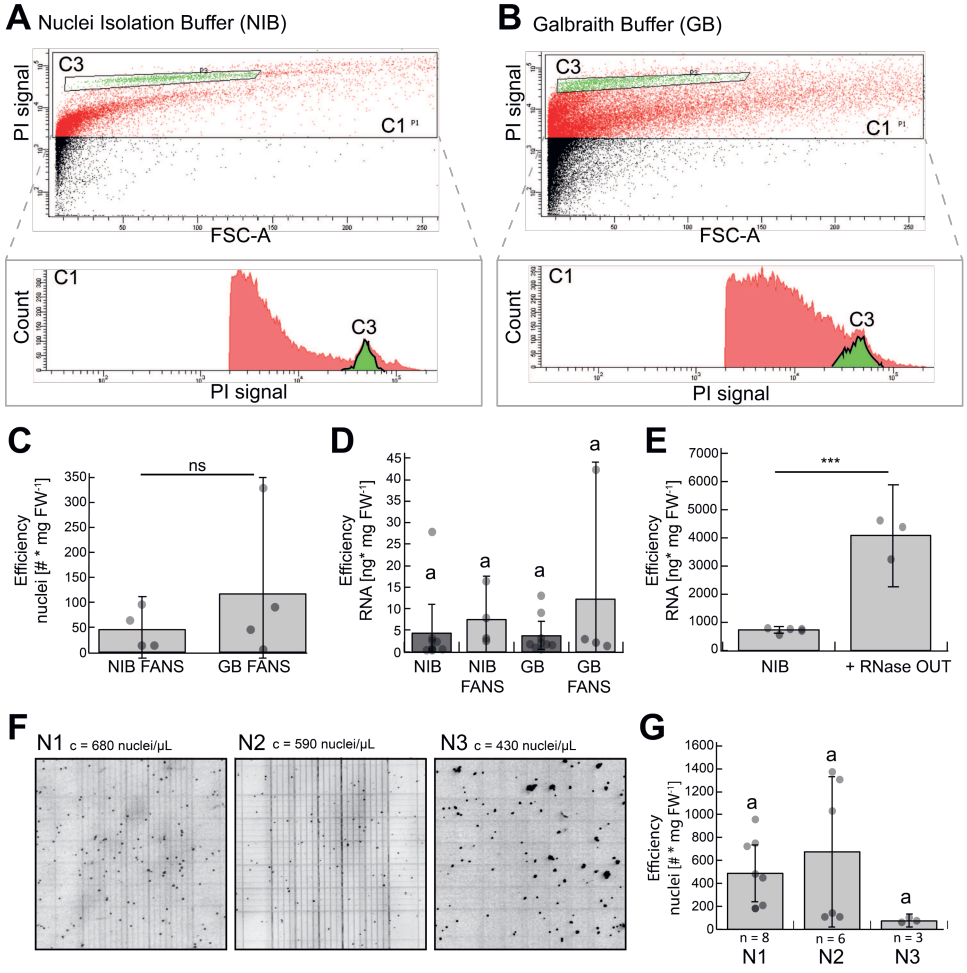


Figure 2 | Comparison of nuclei isolation methods and buffer performance. (A-B) FANS profiles showing forward scatter (FSC-A) versus propidium iodide (PI) fluorescence. Nuclei (Cluster C3, green) are gated from the broader population 1 (cluster C1, red). Lower panels display event counts within C1, with nuclei (cluster C3) highlighted in green. Isolation using **(A)** the commercially available Nuclei Isolation Buffer (NIB; CELLTYPN1, Sigma-Aldrich) or **(B)** the homemade Galbraith Buffer (GB). **(C)** Nuclei isolation efficiency using NIB or GB. **(D)** RNA yield from nuclei isolated with NIB or GB, with or without FANS purification. **(E)** Effect of RNase OUT on RNA recovery from NIB-isolated nuclei. **(F)** Fluorescence microscopy images of nuclei isolated using methods N1–N3, visualized on a hemacytometer with corresponding concentrations. **(G)** Comparison of overall nuclei isolation efficiency across methods N1–N3. Bars represent mean with 95% CI (n ≥ 2). Different letters indicate statistically significant differences (ANOVA, Tukey post hoc test, $p \leq 0.05$). Asterisks indicate significance between treatments (*** $p \leq 0.001$).

Assessment of nuclei quality prior to snRNA-seq

In addition to nuclei yield, we assessed nuclei quality prior to single-cell library preparation. FANS profiles revealed three subpopulations corresponding to 2n, 4n, and 8n ploidy levels (**Figure 3A–C**). Differences in size were visually confirmed, and the frequency of sorted events per fraction corresponds with the counted concentrations of 2n to 8n (**Figure 3A–C**).

Nuclei integrity was further evaluated using fluorescent-antibody staining. Intact nuclei showed no green fluorescence, while compromised nuclei accumulated dye (**Figure 3D**). This method confirmed intactness but was too time-consuming for representative quantification.

As an additional analysis of the quality of the nuclei samples, we tried to measure ambient RNA. For that, we isolated RNA from supernatant and pellet of nuclei isolated with methods N1 and N2 after a final centrifugation step (**Supplemental Figure S1**). High RNA content in the supernatant would indicate high levels of ambient RNA. Unfortunately, low RNA yields prevented direct comparisons by light absorption measurements (NanoDrop, 260 nm), or agarose gel analysis (**Supplemental Figure S1A,B**). To increase sensitivity, RNA levels were assessed by qRT-PCR. As all samples were treated, we analyzed the induction of the central nodulation gene *MtNODULE INCEPTION* (*MtNIN*). The induction of *MtNIN* is triggered by both benzylaminopurine (BAP) and Nod factors (NF). However, low RNA yields and high Cq values for both *MtACTIN1* (*MtACT1*, housekeeping) and induced *MtNIN* in supernatant and pellet precluded reliable conclusions (**Supplemental Figure S1**).

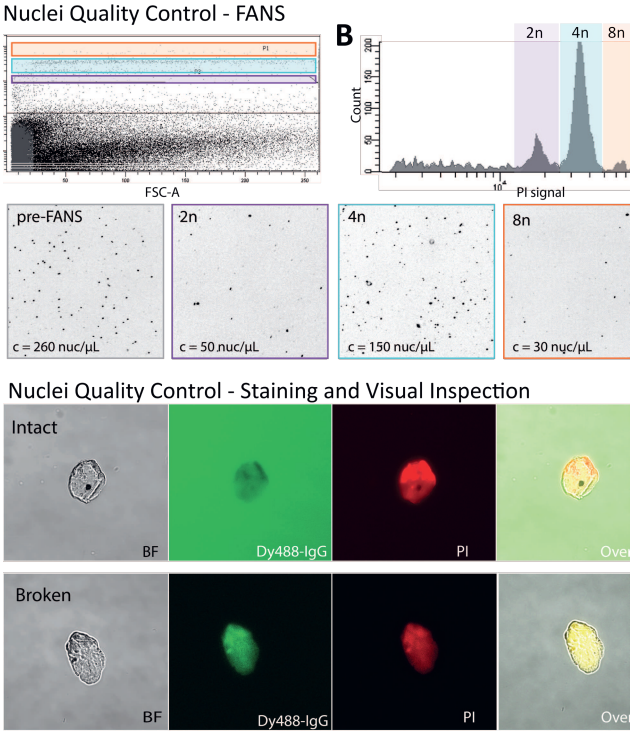


Figure 3 | Nuclei quality control and assessment of ploidy and integrity. (A) FANS profiles displaying forward scatter (FSC-A) versus propidium iodide (PI) fluorescence, used to assess nuclei populations. **(B)** Quantification of nuclei events from **(A),** classified by ploidy: 2n (purple, smallest), 4n (blue, majority), and 8n (orange, largest). **(C)** Fluorescence microscopy images of unsorted nuclei (pre-FANS) and sorted 2n, 4n, and 8n populations visualized on a hemacytometer. **(D)** Confocal microscopy of individual nuclei showing brightfield (BF), green-fluorescent GFP-tagged negative control antibody (Dy488-IgG), red-fluorescent DNA stain (PI), and merged images. Top panel shows an intact nucleus with no GFP signal; bottom panel shows a disrupted nucleus with GFP accumulation, indicating compromised nuclear integrity.

Optimization and evaluation of protoplast isolation methods for distinct Medicago root sections

To overcome the limitations observed during our nuclei-based attempts, we set out to evaluate protoplast-based single-cell RNA-seq as an alternative approach. For this, protoplasts were isolated from Medicago root tissues using four methods (P1–P4) adapted from Jia et al. (2018), and their viability was assessed by live/dead staining (**Figure 4**). To account for developmental heterogeneity, root tissue was divided into two regions: the 2–5 mm root tip, where cells are actively dividing and differentiating, and the 10 mm susceptible zone (S-zone) above, where cells are fully differentiated and root hairs are developing or mature (**Figure 4A**). The S-zone is of particular interest as the primary site of *Rhizobium* infection and nodulation initiation. However, its composition of highly differentiated and specialized cell types that have undergone cell wall modifications poses technical challenges for successful protoplast isolation.

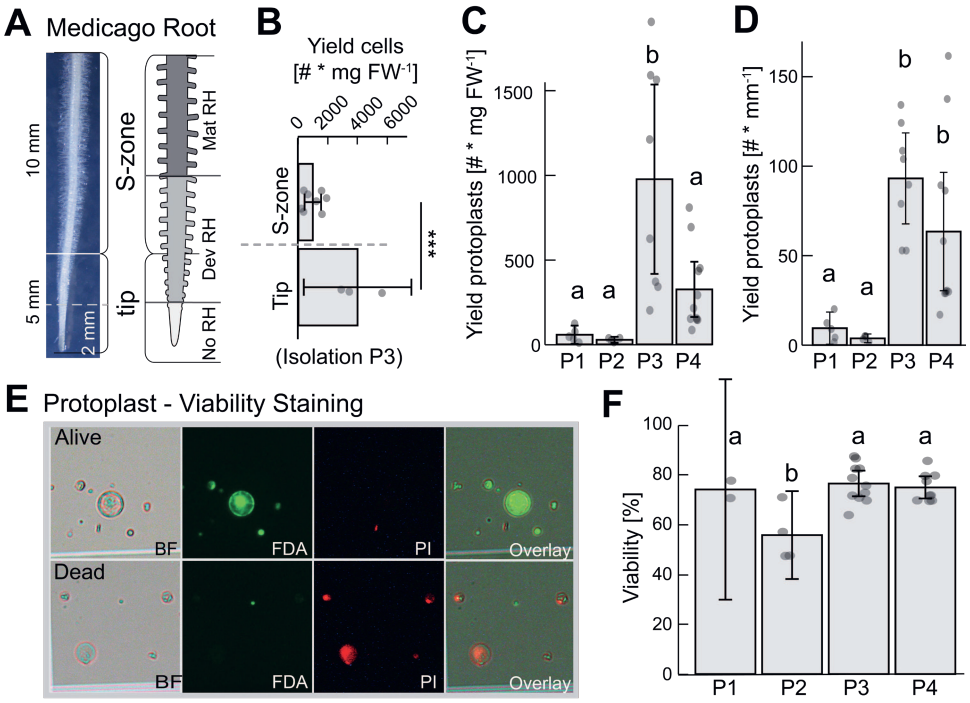


Figure 4 | Protoplast isolation from Medicago roots. (A) Image and schematic of a Medicago root indicating sampled regions: the 2–5 mm root tip without root hairs (No RH) and the 10 mm susceptible zone (S-zone) above the tip, where root hairs are developing (Dev RH) or mature (Mat RH). **(B)** Comparison of protoplast yield from root tip versus S-zone using method P3. **(C-D)** Comparison of protoplast isolation methods P1–P4. **(C)** Yield expressed as protoplasts per mg fresh weight (FW), excluding root tip samples. **(D)** Yield expressed as protoplasts per mm total length, excluding root tip samples. **(E)** Fluorescence microscopy of isolated protoplasts showing, from left to right: brightfield (BF), green-fluorescent live stain (fluorescein diacetate, FDA), red-fluorescent dead stain (propidium iodide, PI), and the merged image. The top panel displays an intact, viable protoplast with FDA signal; the bottom panel shows a lysed cell lacking FDA but positive for PI. **(F)** Viability of isolated protoplasts across methods. Bars represent mean with 95% CI ($n \geq 2$). Different letters indicate statistically significant differences (ANOVA, Tukey post hoc test, $p \leq 0.05$). Asterisks indicate significance between treatments (* $p \leq 0.001$).

This difference in isolation efficiency between zones is reflected in our results (**Figure 4B**). Using method P3, protoplast yield was approximately fourfold higher in root tips than in the S-zone (3980 vs. 970 protoplasts per mg fresh weight; **Figure 4B**). Methods P1 and P2 combined tissues from both regions to increase yield, while P3 and P4 processed them separately to preserve spatial resolution. For comparisons of efficiency and viability (**Figure 4C–D,F**), only S-zone samples were used for P3 and P4. Among all methods, P3 yielded the highest efficiency, showing a 3-, 20-, and 50-fold improvement compared to P4, P2, and P1, respectively (**Figure 4C–D**). When normalized to root length, P3 and P4 performed similarly (**Figure 4D**). Improvements in P3 and P4 included use of digestion buffer instead of PBS+ for washing and centrifugation-based recovery rather than passive settling.

Cell viability was measured using fluorescein diacetate (FDA) and propidium iodide (PI) staining (**Figure 4E**). Viable protoplasts retained FDA fluorescence (green), whereas non-viable protoplasts accumulated PI (red). Method P2, which used PBS+, reduced viability to 55%, whereas P1, P3, and P4 maintained viability between 75–80%, all above the 70% threshold recommended for both single-cell platforms (**Figure 4F**).

Nod factor treatment and optimization of protoplast viability

Next, we checked if our optimized protocol for protoplasts isolation is also compatible with root cell isolation after external Nod factor application. To this end, 1% dimethyl sulfoxide (DMSO) was initially added to enhance uptake and trigger nodulation responses, including induction of the transcription factor *MtNIN*. However, at this concentration, protoplast viability decreased substantially and did not reach the 70% threshold recommended for single-cell workflows (data not shown). Reducing DMSO to 0.1% maintained *MtNIN* induction while preserving intact protoplasts (**Supplemental Figure S2E–F**).

The effect of Nod factor treatment was neither extended nor lost during protoplast isolation (**Supplemental Figure S2G**). Therefore, roots were exposed to Nod factors for 3 hrs prior to digestion to ensure induction of *MtNIN* expression, while the subsequent 2 hrs protoplast isolation did not alter this response (**Supplemental Figure S2G**).

STEP 3-5: SINGLE-CELL PLATFORMS AND SAMPLE PROCESSING

Adaptations for protoplast loading on single-cell platforms

One of the main causes of failure when working with droplet-based single cell platform is the risk of cell accumulation ('clogging') in the microfluidics channels. The risk of clogging is substantially higher in samples with high cell debris and/or high number of larger cells (cell size restriction is set to 40 µm in diameter). To prevent clogging with the microfluidic-based 10x Genomics® platform, nuclei were loaded after filtration through a 20 µm strainer while strainers with a maximum pore size of 40 µm were essential to filter larger cells and cell debris for each protoplast sample. Use of larger strainers (50 or 70 µm) led to system clogging and complete sample loss, as encountered during method P1a. Switching to 40 µm strainers resolved this problem without loss of specific cell populations, as qPCR analysis of endodermis and epidermis markers showed no depletion (**Supplemental Figure S2A**,**B**).

For the microwell-based BD Rhapsody™ platform, 70 µm strainers were used, because sample clogging does not apply here, as protoplasts settle into the wells and larger debris can be washed away. Bovine serum albumin (BSA) was excluded from the buffer, since it appeared to interfere with cell settling. While the omission of BSA did not increase protoplast aggregation, attempts to quantify differences in settling efficiency under varying BSA concentrations were inconclusive due to sample variability and limitations in lane recycling.

For viability staining, fluorescein diacetate (FDA) was more effective than Calcein, yielding stronger fluorescent labeling and more accurate cell counts by the scanner. No transcriptomic differences or significant shifts in cell-type composition were observed between staining methods (**Supplemental Figure S3, Figure 9B**).

cDNA quality control during library preparation

To ensure that downstream comparisons between single nuclei and single cell RNA-seq data were not confounded by technical biases, we first evaluated the cDNA quality and fragment size differences between both sample types. Differences in fragment size between nuclei and protoplast-derived cDNA synthesis is a well-documented phenomenon (10x Genomics: CG000124, CG000315). Naturally, nuclei-derived cDNA is synthesized from longer, less processed transcripts, whereas protoplast-derived cDNA is synthesized from a high abundant and uniform pool of mature cytoplasmic mRNAs. During 10x Genomics® single-cell sample processing, cDNA quality was assessed at two stages: total cDNA synthesis and cDNA library preparation (**Figure 5**). In step 1, analysis of total cDNA revealed clear differences in fragment sizes between nuclei and protoplast inputs. Nuclei libraries displayed distinct peaks in the bioanalyzer profile ranging from 400–3000 bp and generally yielded lower concentrations, even with two additional amplification cycles (**Figure 5A–C**, step 1). In contrast, protoplast-derived cDNA showed a Gaussian-like distribution with minor peaks between 400–2000 bp, consistent with expected profiles (**Figure 5C**, step 1).

In step 2, fragmentation to 300–600 bp produced uniform distributions across both nuclei- and protoplast-derived libraries (**Figure 5A–C**, step 2). Addition of extra dithiothreitol (DTT) during nuclei isolation with method N3 did not interfere with cDNA synthesis or library preparation (**Figure 5B**).

Single-cell processing with the BD Rhapsody™ whole-transcriptome workflow follows principles comparable to those of the 10x Genomics® platform. The first quality control, performed after random primer extension (RPE) PCR, yielded fragments ranging from 150–2000 bp with two prominent peaks at ~350 and ~800 bp (**Figure 5d**, step 1). Following library construction by whole-transcriptome analysis (WTA) index PCR, fragment distributions became more uniform, spanning 250–1000 bp (**Figure 5d**, step 2).

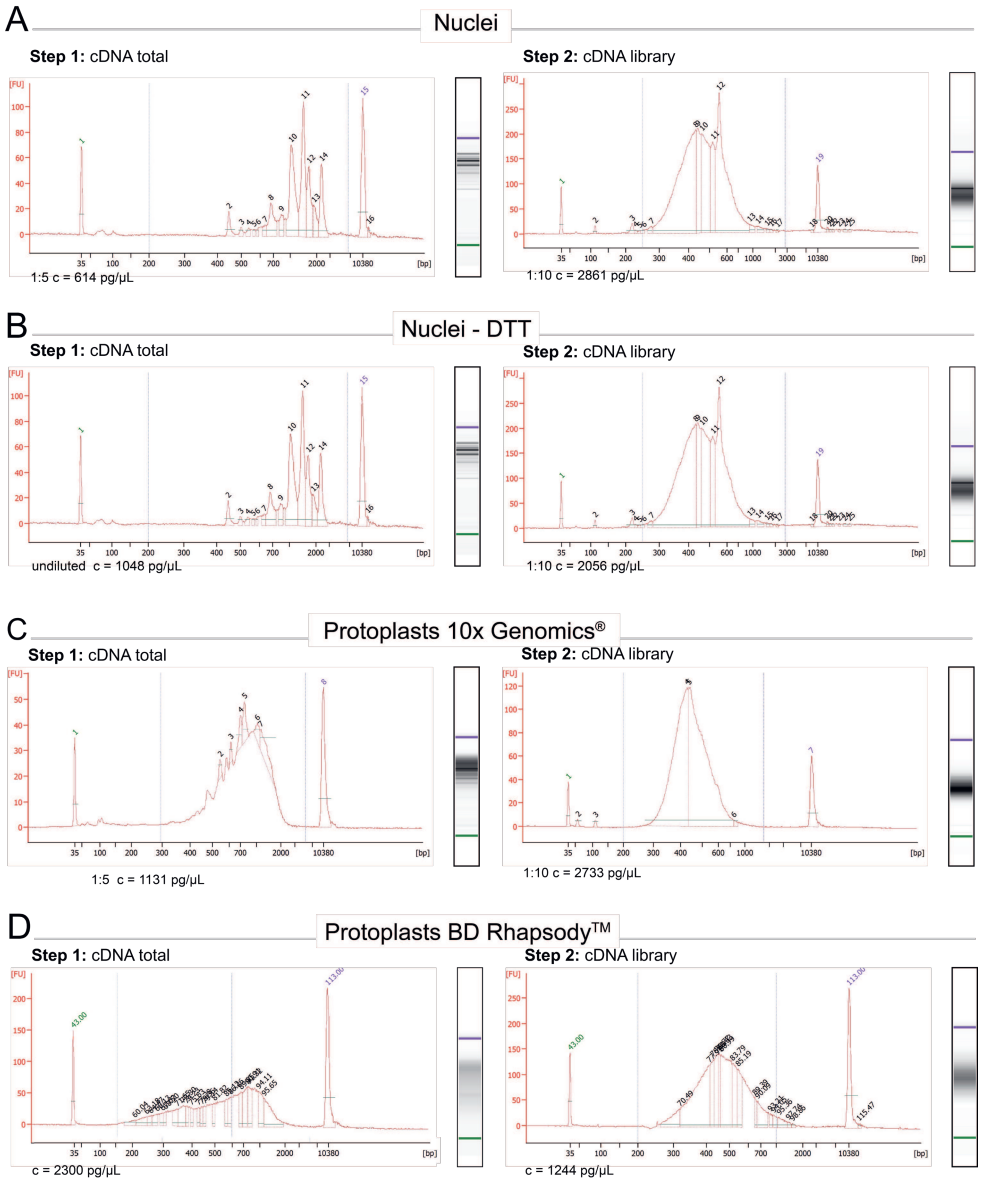


Figure 5 | Bioanalyzer profiles of nuclei and protoplasts during cDNA synthesis and library preparation. Fragment size distributions were analyzed at key stages of 10x Genomics **(A-C)** and BD Rhapsody **(D)** sample processing: cDNA synthesis (Step 1) and final library preparation (Step 2), corresponding to protocol steps 2.4 and 3.7 for 10x Genomics, and RPE PCR and WTA index PCR for BD Rhapsody, respectively. **(A-B)** Nuclei-derived cDNA displays distinct fragment peaks during Step 1, indicating more variable fragment sizes. **(C)** Protoplast-derived cDNA exhibits a smooth, Gaussian-like distribution in Step 1 In both sample types, the cDNA fragment distribution becomes more uniform after library preparation in Step 2, as expected from size selection and amplification steps. **(D)** Protoplast cDNA after Random Primer Extension (RPE) PCR has shorter fragments with two distinct peaks in step 1, that become equally uniform after Whole Transcriptome Analysis (WTA) index PCR in step 2.

STEP 6: DATA QUALITY CONTROL

Data quality assessment of nuclei and protoplast isolations

We next assessed the effect of each protocol optimization approach based on sample quality parameters after sequencing, including e.g. mapping rates, ambient RNA levels, number of UMIs and genes/cell (**Figure 6**). Although nuclei and protoplasts were obtained with all isolation methods (N1–N3, P1–P4), method P1 yielded no cDNA and was excluded from further analysis.

Reads were first mapped to the Medicago genome (version 5, annotation version 1.9, **Figure 6A**). Nuclei-derived data consistently showed lower mapping efficiency than protoplast-derived data. For nuclei, ≤25% of reads mapped successfully (N2) and only ~5% of reads were retained as single-cell specific after ambient mRNA filtering (**Figure 6A,B**). Method N3 performed poorest, with ~5% mapping and ~1% usable reads (**Figure 6A,B**). In contrast, protoplasts showed higher mapping rates: ~40% (P3) with ~15% usable reads, ~50% (P4) with ~13% usable reads, and ~60% (P2) but only ~10% usable reads, reflecting higher ambient mRNA (**Figure 6A,B**).

As expected, nuclei libraries contained a larger proportion of unspliced transcripts, whereas protoplasts contained 65–80% spliced transcripts after removing the ambiguous fraction (**Figure 6C,D**). N2 and N3 yielded the lowest fraction of spliced reads (~25%), while N1 retained ~45%, likely due to incomplete removal of cytosolic mRNA (**Figure 6D**). Method N1 also showed elevated chloroplast and mitochondrial transcript fractions, indicating contamination (**Figure 6E,F**).

Ambient mRNA levels were highest in nuclei samples and in P2, resulting in reduced usable reads (**Figure 6B,G**). An additional parameter to assess sample quality is the barcode-rank plot. Here, the cell barcodes are ranked by the number of unique molecular identifiers (UMIs), revealing the separation between genuine cells and background noise. A steep curve is associated with low sample quality, and possibly high ambient RNA. In our experiments, the protoplast barcode-rank plot showed a steadier curve compared to nuclei (**Supplemental Figure S4A**).

Cell recovery efficiency was similarly low for N1–N3 and P2. P3 showed modest improvement, while P4 (BD Rhapsody™) enabled higher cell loading and more accurate recovery estimates, though overall recovery remained ~40% of the expected input (**Figure 6H,I**).

Single-cell quality was further assessed by UMIs, number of genes per cell, sparsity, and complexity (**Figure 6J–M**). N1–N3 and P2 showed high sparsity, with ~90% of overall genes shared across cells, while P3 and P4 showed lower sparsity and higher complexity, indicating greater transcriptome diversity (**Figure 6J,M**). UMIs and detected genes per cell were also significantly higher in P3 and P4 (**Figure 6K,L**). Among nuclei methods, N2 (FANS-based, without centrifugation) showed the most favorable trend (**Figure 6J–M**).

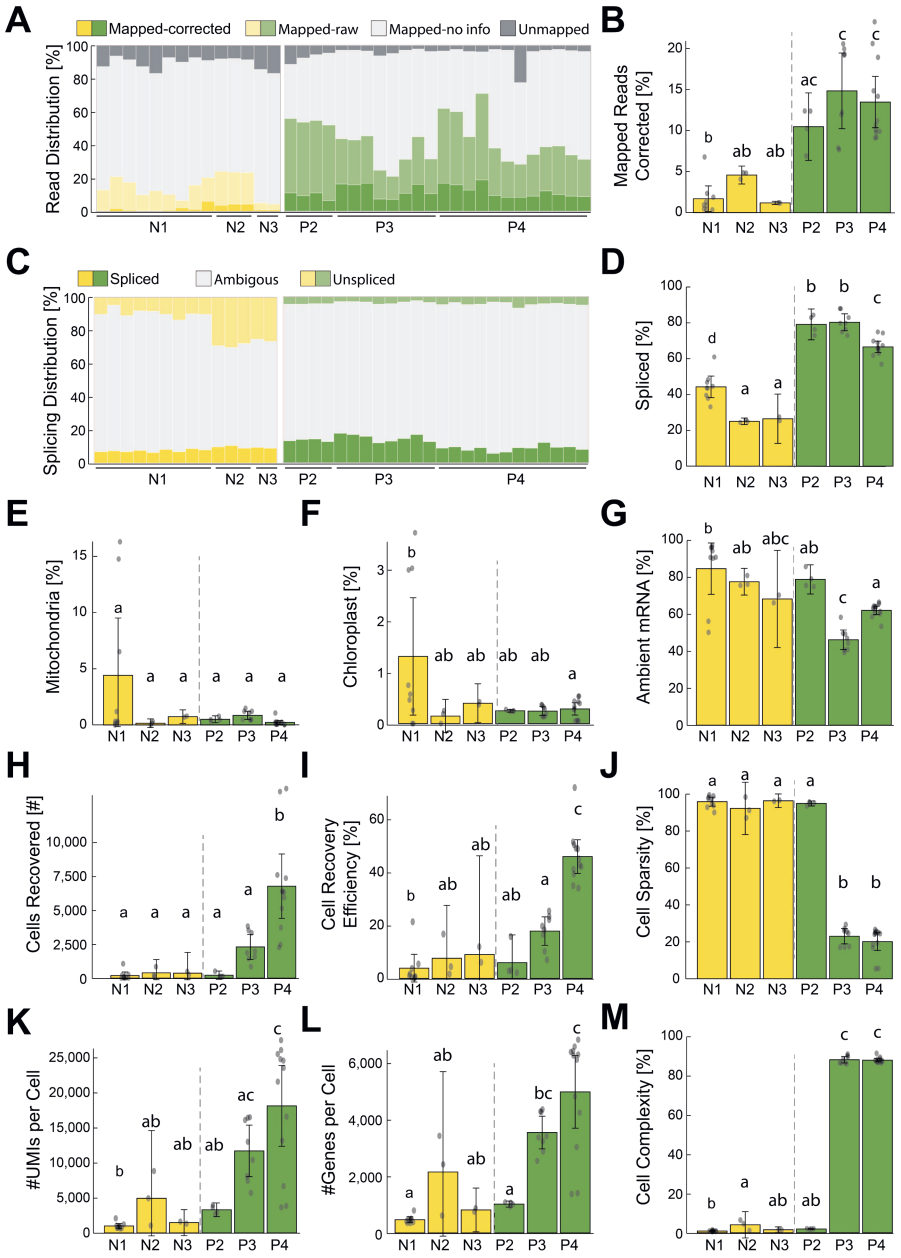


Figure 6 | Quality Control parameters of snRNA-seq and scRNA-seq. Comparison of the different nuclei and protoplast isolation methods (nuclei N1-3 in yellow and protoplasts P2-4 in green). **(A)** Total reads from sequencing distribution of Unmapped (dark grey), Mapped-no information (light grey), Mapped-raw reads (light yellow/green), and Mapped-corrected reads (dark yellow/green). **(B)** Percentage of mapped-corrected reads, that are used for further analysis. **(C)** Fractions of spliced (light yellow/green), ambiguous (grey), and unspliced (dark yellow/green) genes. **(D)** Percentage of spliced genes without ambiguous fractions. **(E)** Percentage of mitochondria. **(F)** Percentage of chloroplast. **(G)** Percentage of ambient RNA, that is removed during filtering. This corresponds to the fraction of removed raw reads. **(H)** Total number of nuclei or protoplasts confidently identified after data analysis. (*continued on next page*)

Figure 6 (continued) | (I) Recovery efficiency of nuclei or protoplasts; known input vs. recovered amount. **(J)** Percentage of cell sparsity. A high sparsity indicates that each cell has a high percentage of all combined genes, and therefore similar information. Low sparsity shows that the cells are more unique compared to each other. **(K)** Amount of Unique Molecular Identifiers (UMIs) per cell, corresponding to transcripts per cell. **(L)** Amount of identified genes per cell. **(M)** Percentage of cell complexity. Bars represent mean with 95% CI ($n \geq 2$). Different letters indicate statistically significant differences (ANOVA, Tukey post hoc test, $p \leq 0.05$).

Comparative analysis of protoplast datasets and cell-type composition

For downstream biological analyses, protoplast datasets were prioritized due to higher overall quality compared with nuclei datasets (**Figure 6**, **Supplemental Figure S4A**). We compared two single-cell platforms (10x Genomics® and BD Rhapsody™), root sampling zones (tip and susceptible zone, S-zone), and treatments with Nod factors and *Rhizobium*. Three datasets were assembled: the first with a nuclei and protoplast sample (NP), the second with predominantly 10x samples and two BD samples (A1), and the third with predominantly BD samples and three reference 10x samples (A2) (**Table 1**).

Table 1 | Composition of combined samples for data analysis.

Dataset	Isolation method	Single-cell platform	Sample composition
NP (**Figure S4**)	N2	10x Genomics	S-zone 10 mm (mock)
	P3		S-zone 10 mm (mock)
A1 (**Figure 7**)	P3	10x Genomics	Tip 2 mm (3x)
			Rhizobia application (24, 48, 72 and 96 h)
			S-zone 10 mm (mock)
	P4	BD Rhapsody	Tip 2 mm
			Tip 10 mm
A2 (**Figure 8**)	P3	10x Genomics	Tip 2 mm (3x)
	P3	BD Rhapsody	S-zone 10 mm (mock)
	P4		Tip 2 mm
			Tip 10 mm
			S-zone 10 mm (Nod factors), 1+2h digestion
			S-zone Nod factors (with Calcein or FDA), 1h digestion
			Rhizobia application (12, 24, 48, 72 and 96 h), 1+2h digestion

After filtering and curating, we were able to obtain Uniform Manifold Approximation and Projections (UMAPs) as a way of depicting the similarity of cell-type-specific transcriptomes for all three datasets (**Figure 7-8, Supplemental Figures S4-S6**). To enable the integration of the nuclei and protoplast dataset, comparable sequencing depth, and further filtering is required (NP; **Supplemental Figure S4B–E**).

To identify the different cell types that are associated with the clusters gained from UMAP analysis, we used already known markers of Medicago or Arabidopsis orthologs (**Figure 7-8, Supplemental Figures S4-S6, Supplemental Table S2**).

In the integrated nuclei dataset (NP), we were able to identify all expected cell types, with the limitation of the highly underrepresented endodermis (**Supplemental Figure S4B-E**). Nevertheless, the addition of nuclei recovered more epidermal and trichoblast cells, and the

corresponding level of nuclei-specific parameters such as RNA velocity, could be further implemented (**Supplemental Figure S4B**).

In the protoplast datasets (A1, A2), the cluster association differs between analysis A1 and A2, especially the nodule cluster (red) during A1 is split into two clusters where one is more associated with the pericycle and the other with the cortex (**Figure 7A**). As these clusters are also present in the uninduced susceptible zone, where no nodule cluster is expected, we assume these to be meristematic clusters, most likely of lateral root primordia (**Figure 7A, 9A**). Furthermore, there is a split in the pericycle cluster (purple) between tip and S-zone, where the more mature S-zone, and origin of the nodule, positions closer to the future nodule (**Figure 7A**). In analysis A2, the epidermis (green) is the only cell type, that separates into two distinct clusters depending on their origin from the tip or S-zone, indicating a strong transcriptomic change during differentiation (**Figure 8A**). In A1, we cannot replicate this distinction, as we are generally lacking epidermis cells of the more mature S-zone (**Figure 7-9**).

Cell-type composition varied with tissue differentiation state and, to some extent, with protoplast digestion duration (**Figure 9, Supplemental Figure S2C,D**). Analysis methods A1 and A2 also differed in cluster assignments (**Figure 9D**). As expected, root cap and meristematic cells were absent from the S-zone due to tissue maturation (**Figure 9C,D**). However, epidermis and endodermis fractions were also reduced in the S-zone compared to tip samples (**Figure 9C,D**). This trend was already visible when comparing 2 mm versus 10 mm tip samples and became pronounced in most S-zone datasets (**Figure 9A–C**).

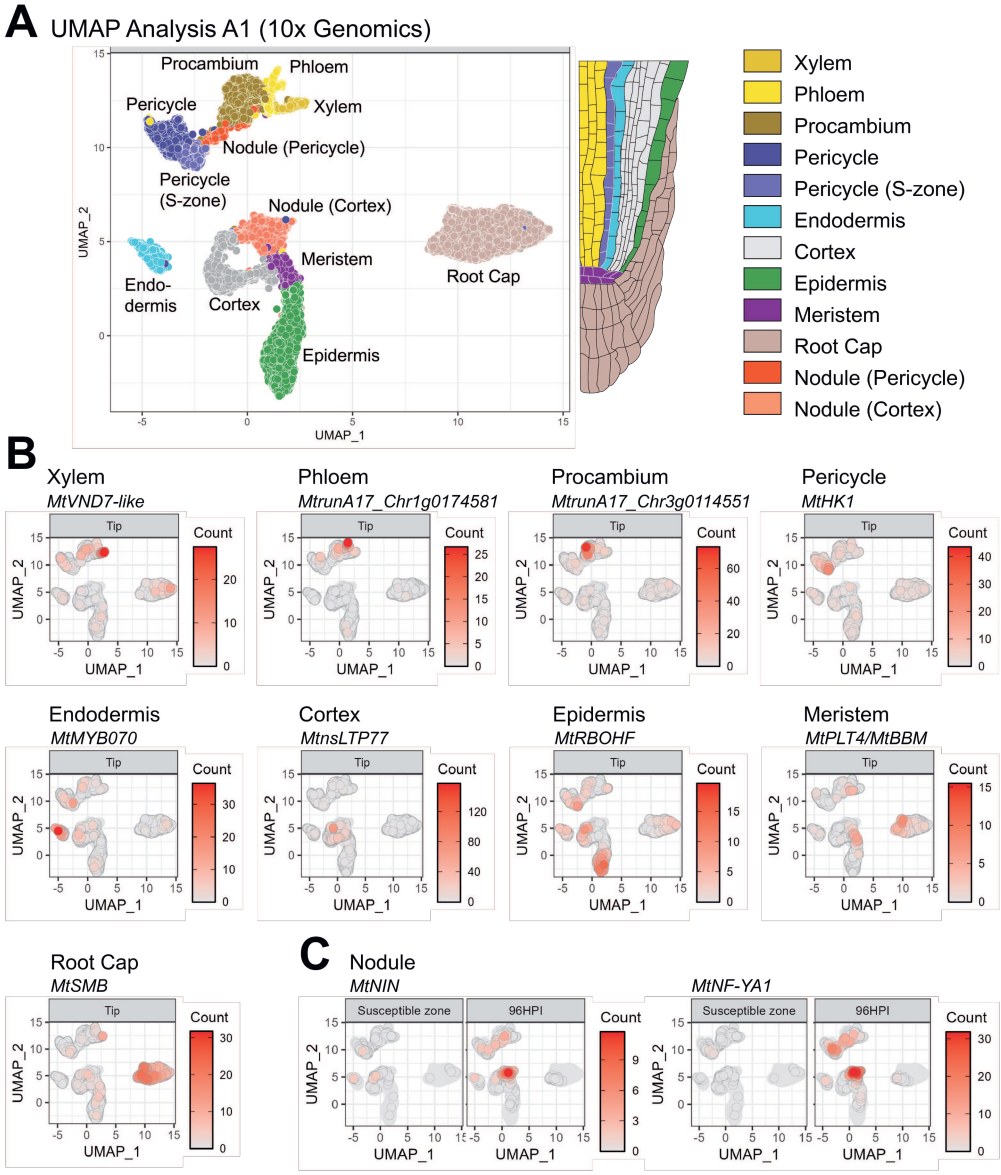


Figure 7 | UMAP Cluster Analysis A1 – 10x Genomics. Assigned cell types **(A)** according to the expression of cell-type specific markers **(B),** and nodule induced markers **(C). (A)** UMAP of all data combined with cell-type specific resolution, marked in the corresponding color and indicated in the schematic of the root tip. **(B)** Cell-type specific marker expression in the Tip sample. **(C)** Comparison of susceptible zone uninduced (left) and 96 hpi with *Sinorhizobium meliloti* (right). Expression in counts per cell. Additional markers in **Supplemental Figure S5**.

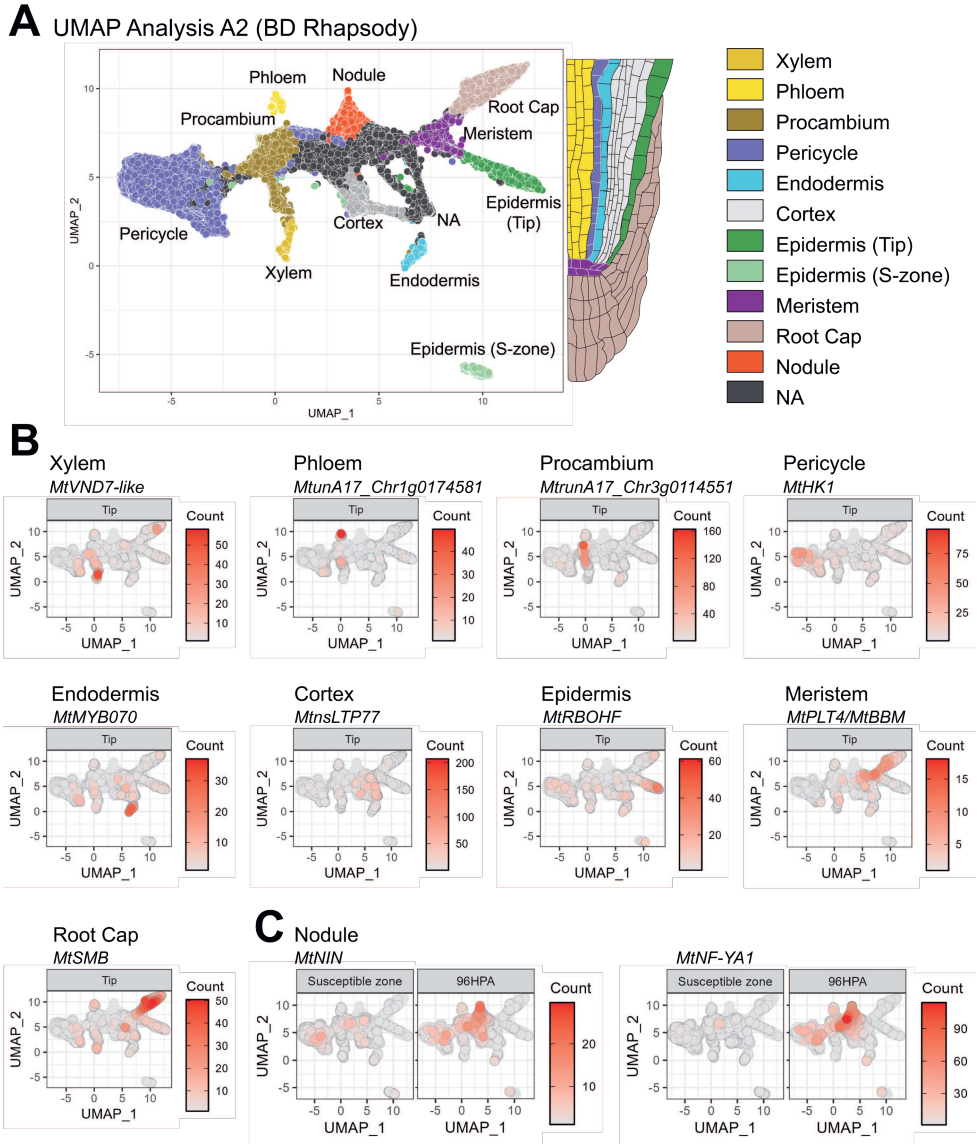


Figure 8 | UMAP Cluster Analysis A2 – BD Rhapsody. Assigned cell types **(A)** according to the expression of cell-type specific markers **(B)**, and nodule induced markers **(C). (A)** UMAP of all data combined with cell-type specific resolution, marked in the corresponding color and indicated in the schematic of the root tip. **(B)** Cell-type specific marker expression in the Tip sample. **(C)** Comparison of susceptible zone uninduced (left) and 96 hpi with *Sinorhizobium meliloti* (right). Expression in counts per cell. Additional markers in **Supplemental Figure S6**.

Potential causes of epidermis and endodermis loss were evaluated. First, cell strainer size was tested by comparing 40 µm and 100 µm filters. No depletion of marker expression was observed; epidermis markers were slightly enriched in the 40 µm fraction, ruling out size-based filtering as a cause (**Supplemental Figure S2A,B**). Second, incubation time during protoplast isolation was

tested. Epidermis marker expression decreased significantly after 3 hrs digestion compared to 1 h (**Supplemental Figure S2C**). This was consistent with FDA staining, which showed a larger epidermis fraction in S-zone samples digested for 1 h compared to 2 hrs or combined 1+2 hrs incubations (**Figure 9B**). Endodermis markers were less consistent: *MtCASPARIAN STRIP MEMBRANE PROTEIN 1 and 2* (*MtCASP1/MtCASP2)* decreased after 2 hrs and recovered after 3 hrs; *MYELOBLASTOSIS TRANSCRIPTION FACTOR 70* (*MtMYB070)* was unchanged; *MtSCARECROW* (*MtSCR)* increased after 2 hrs and decreased after 3 hrs (**Supplemental Figure S2D**). Based on these results, method P4 was adapted to include separate 1 and 2 hrs digestions, pooled at a 1:1 ratio, to improve recovery of epidermis cells while maintaining yield and minimizing transcriptomic alterations. In pooled Nod factor–treated S-zone samples, both epidermis and endodermis cells were detected (**Figure 9B**).

Nodule-associated clusters increased during nodulation and correlated with developmental progression (**Figure 9A**). In dataset A1 (predominantly 10x), assignment to the nodule cluster was less distinct than in dataset A2 (predominantly BD) (**Figure 9A, Table 1**). At 0 h, cells were already assigned to a meristematic nodule or lateral root (LR) cluster (light red), reflecting limited resolution between these states. Despite attempts to exclude lateral roots during harvest, increased meristematic and root cap fractions suggested their inclusion (**Figure 9A**).

Differences in cluster composition were also observed between analysis methods. In A1, pericycle cells were absent, and meristematic cells reduced compared to A2 (**Figure 9D**). These discrepancies reflect differences in dataset composition (**Table 1**), highlighting the influence of sample context on clustering outcomes.

The apparent reduced diversity in S-zone samples from the BD Rhapsody™ dataset was not representative (**Figure 9C**). This was the first BD run, performed with Calcein staining, which led to overestimation of input cell numbers, reduced loading efficiency, and ultimately fewer recovered cells.

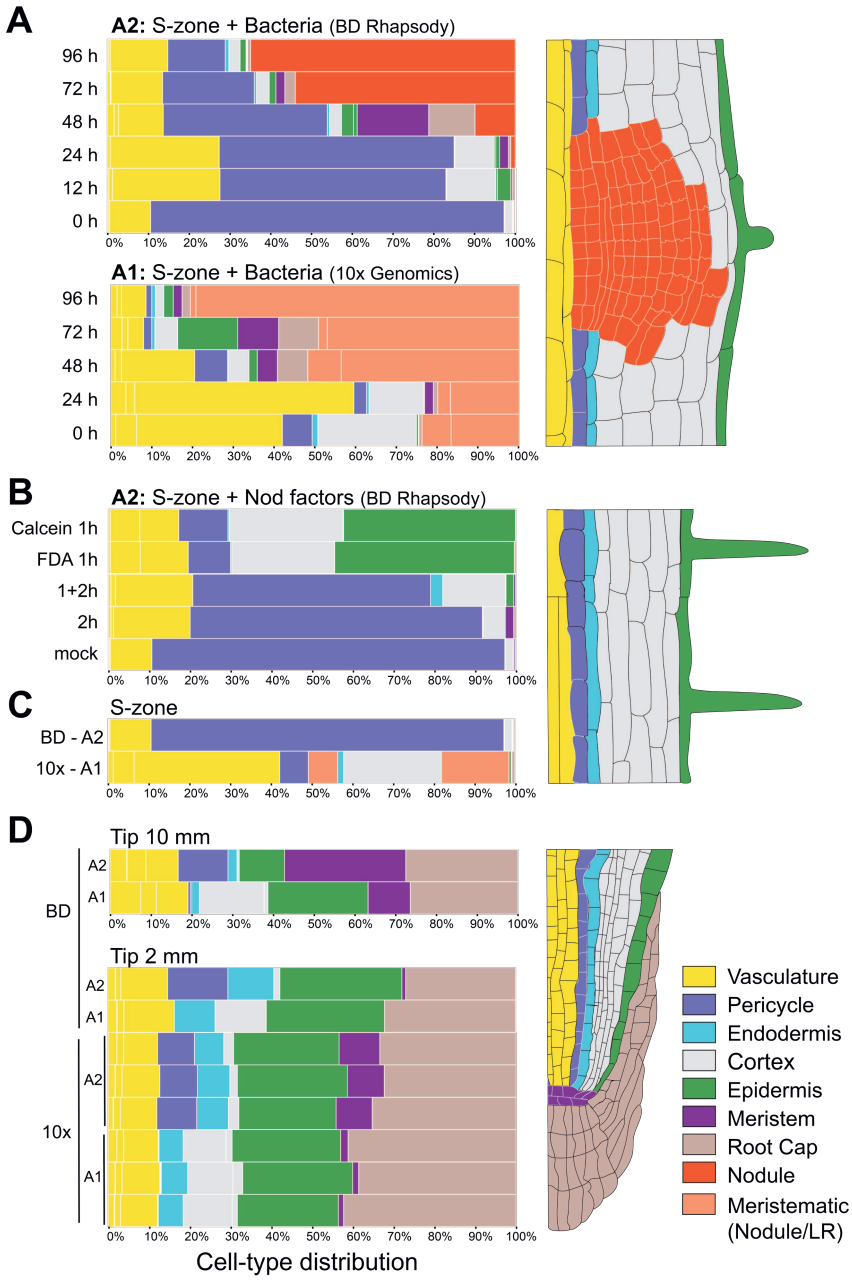


Figure 9 | scRNA-seq cell-type distribution of different zones and treatments from isolated protoplasts. **(A)** Early nodulation with *Sinorhizobium meliloti 2011* inoculation for 0-96 hrs. Upper and lower panel show the data from analysis A2 based on BD Rhapsody (BD), and analysis A1 based on the 10x Genomics (10x) platform, respectively. **(B)** Mock and Nod factor treated S-zone, compared with different digestion incubation times (1, 2 or 1+2 h), and the different cell stains Calcein or fluorescein diacetate (FDA). **(C)** Untreated S-zone comparison of BD and 10x. **(D)** Untreated tip samples of either 2 or 10 mm total length. Comparison between the platforms BD and 10x, as well as the two different analysis methods A1 and A2 (Table 1). LR = Lateral Root.

DISCUSSION

In this study, we identified critical steps in the nuclei and cell isolation and processing workflow to achieve high-quality single cell transcriptomes from Medicago roots. Furthermore, our improved protocols allowed us to compare data quality and content when nuclei and protoplasts were used as starting material for single-cell RNA sequencing in Medicago, using multiple isolation protocols and two single-cell platforms. Our results demonstrate that protoplasts provide more reliable input for single-cell transcriptomics, producing higher mapping rates, greater transcript complexity, and improved cell recovery compared with nuclei (Figure 6). Nevertheless, nuclei can be integrated with protoplasts to access extra information, such as RNA velocity. In addition, nuclei could potentially increase the number of otherwise underrepresented or inaccessible cell populations

Nuclei: opportunities and limitations

Single-nuclei RNA-seq is frequently described as more versatile than protoplast-based methods, with reported advantages including the ability to use frozen tissue, simplified isolation steps, reduced cell-type bias, and the recovery of rare populations (Bakken et al., 2018; Farmer et al., 2021; Slyper et al., 2020; Thibivilliers et al., 2020; Wang et al., 2023; Wu et al., 2019). In our hands, however, nuclei proved challenging due to RNA leakage, ambient RNA contamination, and the need for extensive data filtering. The best-performing nuclei protocol (N2, FANS-based) avoided centrifugation after sorting, which likely minimized leakage, but still required strong bioinformatic quality filtering that reduced usable data to as little as 1–5% of total reads. By contrast, methods involving fixation or post-sorting centrifugation (N3) yielded the lowest quality, likely due to buffer incompatibilities or increased leakage.

A major limitation of nuclei-based approaches is the lack of reliable quality control prior to sequencing. Although nuclear morphology can be inspected microscopically, RNA integrity and leakage cannot be directly assessed. Attempts to monitor nuclei quality using fluorescent negative staining were impractical, as imaging was too slow and inconsistent, and RNA leakage continued during handling, further reducing reproducibility (Kimble et al., 2022). Improvements to nuclei sorting, such as stringent cleaning procedures and the inclusion of RNase inhibitors (Conde et al., 2021), may help mitigate these issues and should be evaluated in future legume studies.

RNA leakage during nuclei processing produces high background signals, making extensive filtering necessary. A common proxy for nuclei intactness is the signal-to-noise ratio; however, the highest ratios are often observed in multimers—droplets containing more than one nucleus or cell. In our datasets, ~95% of nuclei had to be filtered out. The remaining ~5% closely matched the expected multimer rate of ~4% for 5,000 nuclei on the 10x platform, suggesting that most retained data likely originated from random multimers. If this is the case, only ~1% of reads would derive from true single nuclei (i.e. equivalent to ~50 individual nuclei per sample) yet distinguishing these from multimers is challenging and would benefit from a reference data set.

Consequently, our single-nuclei datasets resemble partially bulk sequencing, lacking the intended cellular resolution.

For comparison, Pereira et al. (2024) loaded 20,000 nuclei per sample, targeting recovery of ~12,000, and after filtering obtained ~4,600 nuclei per sample. Their improved yield could reflect a more suitable experimental setup, or alternatively, less stringent filtering. In either case, our findings indicate that, despite the common perception that nuclei are easier to handle than protoplasts, they may not offer the same level of reliability. Future work should directly compare our N2 protocol with an additional post-FANS centrifugation step, as applied in other Medicago snRNA-seq studies (Cervantes-Pérez et al., 2022; Liu et al., 2023; Pereira et al., 2024), to determine whether centrifugation reduces ambient RNA without inducing additional leakage. Nevertheless, nuclei can provide valuable cell-type information, particularly for populations underrepresented in protoplast datasets, but only when very large numbers are loaded and sequenced at depth to offset substantial losses.

Protoplasts: robustness with caveats

In contrast, protoplasts are relatively stable once isolated (Sheen, 2001), can be readily assessed for viability, and produce consistently higher-quality transcriptomes. Protoplasts also offered straightforward quality assessment via FDA/PI staining and could be washed by gentle centrifugation without specialized equipment. Although isolation protocols required tailoring to Medicago tissue zones, once optimized, protoplasts provided a robust input for both microfluidic and microwell single-cell platforms.

However, a major limitation of protoplasts is incomplete representation of specific cell types, particularly epidermis and endodermis in more differentiated regions of the root. This can be problematic as the lack of specific cell types will hinder the proper analysis of complex biological systems. For example, in our studies, the susceptible zone, where rhizobia interact with root hairs and initiate organogenesis from the inner root tissues (reviewed by Roy et al., 2020) lacking cell types, especially epidermal, will be extremely problematic as the primary interaction between host and symbiont will occur at the epidermis.

Isolating protoplasts from endodermal cells is particularly difficult because Casparian strip formation and subsequent suberization reinforce the cell wall with lignin and suberin, reducing susceptibility to enzymatic digestion (Geldner, 2013; Su et al., 2023). In Medicago, the suberized endodermis cells start 7.5 mm from the tip, which correlates with our decrease of endodermis cells in more differentiated tissues (Yan et al., 2025). These modifications limit the effectiveness of standard cellulases and pectinases and could therefore be responsible for the low recovery of these cells. Improved recovery may therefore require additional treatments, for example with enzymes such as esterases, lipases, or cutinases, which have the potential to target lignin- or suberin-based barriers (Jabloune et al., 2020; Zhang et al., 2022). Although this has not widely been implemented in legumes, in combination with carefully optimized chemical pretreatments, such approaches might improve endodermal cell wall degradation and enhance the yield of viable protoplasts.

Additionally, epidermis cell numbers were reduced after prolonged digestion. This could reflect either transcriptomic changes induced by protoplasting (Birnbaum et al., 2003; Tucker et al., 2018), physical constraints of elongated root hairs (Miller et al., 1997; Shaw et al., 2000), a developmental trajectory leading to an identity shift, or a combination. By combining protoplasts from short (1 h) and longer (2 hrs) digestions, we increased recovery of these difficult cell types while maintaining overall yield. However, epidermal cells were less abundant in older root tissues (i.e. tissues that were harvested further away from the root tip) suggesting that at least in part an identity shift does seem to occur.

Platform choice: BD Rhapsody versus 10x Genomics

We directly compared two single-cell platforms: the droplet-based 10x Genomics® Chromium and the microwell-based BD Rhapsody™. For plant cells, BD Rhapsody proved more suitable, offering the ability to load larger cells, tolerate debris, and incorporate additional control steps such as direct cell counting in wells and repeat loading attempts. This flexibility allowed us to adapt conditions, including removal of BSA to improve settling and substitution of FDA for Calcein to achieve more accurate viability measurements. Although recovery was lower than expected, BD Rhapsody consistently outperformed 10x Genomics in terms of practical usability with protoplasts. Importantly, both platforms suffered ~90 % cell loss from loading to filtering, but BD's ability to load more protoplasts partially compensated for this limitation.

The flexibility of the microwell system will become increasingly valuable if multiplexing strategies, currently developed for animal systems (e.g. 10x Genomics Flex Gene Expression), can be adapted to plants. Such developments could enable cost-effective analysis of multiple conditions and larger cell numbers, even if size restrictions or debris continue to limit droplet-based approaches.

Biological insights and combined datasets

Despite technical challenges, our optimized protoplast datasets captured key biological processes, including nodule-associated clusters that correlated with developmental progression. Nuclei datasets, although of lower quality, contained additional cell types missing from protoplasts. Integration of both data sources increased overall resolution and improved biological representation, providing a more complete view of nodulation-associated transcriptomes.

Conclusion and outlook

Our results emphasize that single-cell RNA-seq in plants remains technically demanding, with outcomes highly dependent on the choice of input material, isolation method, and platform. Protoplasts provided the most reliable data in Medicago, especially when processed with the BD Rhapsody platform, but nuclei remain valuable for accessing specific cell types and frozen material. Each approach carries trade-offs, and careful optimization is essential to avoid overinterpreting noisy data.

In the future, combining single-cell with spatial transcriptomics will be crucial to resolve both cell-type and positional information (Adema et al., 2024), as demonstrated recently in arbuscular mycorrhizal symbiosis (Serrano et al., 2024). For legumes, capturing early nodulation events in pericycle and cortical cells would provide unprecedented insight into dedifferentiation and infection responses. Although building a comprehensive single-cell root atlas for Medicago may be more challenging than for Arabidopsis, the integration of protoplast and nuclei datasets, together with spatial methods, will be a powerful resource for the nodulation community.

MATERIALS AND METHODS

Plant material and growth conditions

Medicago truncatula Jemalong A17 wild-type seeds were surface-sterilized (7 min in concentrated H_2SO_4, six rinses with sterile water; 7 min in commercial bleach, six rinses with sterile water), imbibed for 1–3 hrs, and stratified overnight at 4°C on Fåhraeus medium (0.125 mM $Ca(NO_3)_2$, 0.9% Daishin agar, 1 g/L carbendazim). After 20–24 hrs germination in the dark, seed coats were removed and seedlings transferred to square plates containing Fåhraeus medium (0.125 mM $Ca(NO_3)_2$, no carbendazim). Plants were grown vertically at 22 °C under long day conditions (16 hrs light/8 hrs dark). Roots were protected from excess light by wrapping plates in foil.

Rhizobial culture and application

Sinorhizobium meliloti 2011 (*Sm2011*, GFP-reporter strain) was grown in YEM medium supplemented with tetracycline (Limpens et al., 2003). Log-phase cultures (OD_{600} 0.6–0.9) were diluted to OD_{600} 0.02 in low-nitrate Fåhraeus medium (0.25 mM NO_3^-) and applied as three 2 µL droplets in the root susceptible zone (5–15 mm above the tip).

Nod factor and cytokinin treatments

Purified *Sm2011* Nod factors were diluted to ~10^{-9} M in Fåhraeus medium containing 0.1% DMSO (Heidstra et al., 1994). Whole roots were incubated for 3 hrs prior to harvesting. For cytokinin treatment, 1 µM 6-benzylaminopurine (BAP) in 0.1% DMSO was applied similarly.

Nuclei isolation for snRNA-seq

Nuclei were isolated according to either Thibivilliers et al. (2020) or Kao et al. (2021), followed by purification using centrifugation and/or fluorescence-activated nucleus sorting (FANS). For all methods, roots were processed on ice, gently chopped with razor blades in isolation buffer, filtered through 20–40 µm strainers, and stained with 2 µg/mL propidium iodide (PI). Isolated nuclei were loaded on the 10x Genomics Chromium Controller targeting 5,000–10,000 nuclei (expected multiplet rate ~4–8%).

Method N1 – Centrifugation (Thibivilliers protocol): Roots were incubated in NIB (Sigma CELLYTPN1), filtered, and centrifuged (700 ×g, 10 min, 4 °C). Pellets were washed in PBS++ (PBS with 0.04% BSA, 0.2 U/µL RNase inhibitor). RNase inhibitor was essential, as omission caused severe RNA degradation. Final suspensions were counted by hemocytometer prior to loading.

Method N2 – FANS (Thibivilliers protocol): NIB-isolated nuclei were directly stained with PI and sorted using a BD FACS Aria II (488 nm laser, 70 µm nozzle, 70 psi). Approximately 25,000 events (nuclei) of all ploidy levels were sorted and of this between 8,000–10,000 were nuclei loaded per sample.

Method N3 – Crosslink, centrifuge, FANS (Kao protocol): To reduce RNA leakage, tissues were fixed in Galbraith buffer containing 500 µM DSP, quenched in Tris-GB, and washed with hexylene glycol-containing buffer (HG-GB). Nuclei were released, purified by sequential centrifugation and FANS, and finally de-crosslinked during first-strand cDNA synthesis by supplementing the 10x master mix with dithiothreitol (DTT, 12.5–50 mM).

Protoplast isolation for scRNA-seq

Protoplasts were prepared from root segments according to Jia et al. (2018) with modifications to improve yield and transcriptome integrity. Roots were cut into 1–3 mm pieces, vacuum-infiltrated in digestion buffer (0.4 M D-sorbitol, 10 mM $CaCl_2$, 10 mM MES, pH 5.7) supplemented with cellulase, macerozyme, viscozyme, and BSA, and incubated at 23–25 °C for 1–3 hrs. Protoplasts were filtered, washed, and resuspended with minimal mechanical stress.

Method P1 – Original protocol (Jia et al. 2018): Digestion with 2% macerozyme for 2 h; protoplasts washed with W5 buffer. Issues: clogging of 10x device and cDNA synthesis failure likely due to high calcium.

Method P2 – Reduced enzyme and PBS+ wash: Macerozyme reduced to 0.5% (w/v). Protoplasts washed in PBS+ (PBS + 0.04% BSA) to reduce calcium interference. Yield remained low, but RNA integrity improved.

Method P3 – Mechanical aid and DB wash: Tissues gently squeezed during digestion to increase cell release. Protoplasts washed with digestion buffer (DB) instead of PBS+ to improve viability. Yield and transcriptome quality improved; centrifugation (100 ×g, 3 min, 4 °C) used instead of settling.

Method P4 – Adaptation for BD Rhapsody: To recover more epidermal and endodermal cells, tissues were split into short (1 h) and long (2 hrs) digestion batches. Larger strainers (70 µm) could be re-used in the microwell system, reducing loss. BSA was removed from the wash buffer to prevent buoyancy. Despite improved recovery, 55,000–80,000 protoplasts were loaded to recover 5,000–14,000 viable protoplasts.

Quality control of nuclei and protoplasts

Integrity was assessed microscopically. Viable protoplasts were stained with fluorescein diacetate (FDA, 8 µg/mL), dead protoplasts and nuclei with PI (2 µg/mL) (Huang et al., 1986). Viability >70% and concentration >500 cells/µL were required for downstream use. Nuclei integrity was further validated by confocal microscopy using Dylight488-coupled antibodies, which only enter leaky or ruptured nuclei.

Protoplast isolation for FDA and Calcein treatment

Arabidopsis thaliana Col-0 seedlings were grown on half-strength Murashige and Skoog (½ MS) medium for 6 days. Approximately 30 mL of digestion solution (0.4 M mannitol, 20 mM KCl, 10 mM $CaCl_2$, 20 mM MES; 1.5% cellulase R10, 0.01 % pectolyase, 0.3% macerozyme, pH 5.7) was used to digest protoplasts from whole seedlings. Seedlings were placed into digestion solution

for 1.5 h, after which the protoplast suspension was filtered through a 70 µm strainer, washed with 5 ml wash buffer (0.4M mannitol, 20 mM KCl, 10 mM $CaCl_2$, 20 mM MES; pH 5.7), and divided into 6 equal samples of 5 mL each. Protoplasts were pelleted at 150 xg for 6 min and resuspended in 1 mL wash buffer.

FDA treatment: 1 mL of protoplast suspension was mixed with 9 µL fluorescein diacetate (FDA; 5 mg/mL diluted in aceton). For the Mock_FDA treatment, 1 mL of protoplast suspension was mixed with 9 µL acetone diluted in wash buffer. Following treatments with FDA and mock for 15 min in darkness at room temperature, cells were pelleted (150 xg, 6 min), resuspended in 200 µL wash buffer and frozen in liquid nitrogen. RNA was extracted from protoplasts using the RNeasy Micro Kit (Qiagen) according to the manufacturer's instructions.

Calcein treatment: 1 ml protoplast suspension was mixed with 9 µL 2 mM Calcein AM stock diluted in DMSO (dimethyl sulfoxide). For the Mock_Calcein treatment, 1 mL protoplast suspension was treated with 9 µL DMSO. Each treatment was done for 15 min in darkness at room temperature and further processed as described for FDA treatment.

Single-cell platforms

Single-cell libraries were generated using two platforms. The microfluidic-based 10x Genomics Chromium iX was employed for both nuclei and protoplast samples, with libraries prepared using the Chromium Next GEM Single Cell 3' v3.1 kit (CG000315); for nuclei-derived libraries, the number of PCR amplification cycles during cDNA synthesis was increased by two to improve yield. In parallel, the microwell-based BD Rhapsody HT system was applied exclusively to protoplasts, as nuclei preparations showed excessive mRNA leakage. Protoplast viability on this platform was assessed using fluorescein diacetate (FDA) and propidium iodide (PI) staining in place of Calcein AM and DRAQ7, which proved less effective in plant material.

RNA isolation, cDNA synthesis, and qRT-PCR

Total RNA was extracted from frozen tissues, nuclei, or protoplasts using the E.Z.N.A. Plant RNA Kit (Omega Bio-Tek). RNA was eluted in 50 µL nuclease-free water at 60 °C. cDNA was synthesized with iScript™ (Bio-Rad) using up to 500 ng RNA input, diluted 1:5, and analyzed by qRT-PCR using SYBR® Green (Applied Biosystems). Primers as described in Supplemental Table 1 were used.

Gene expression quantification and differential expression analysis

For transcript quantification, a kallisto (v0.46.1) index was pre-built from the *Arabidopsis thaliana* reference transcriptome ARAPORT11. Differential expression analysis was carried out in R (v4.4.1) using the DESeq2 package (Love et al., 2014). The workflow included the following steps:

Estimated counts from kallisto were loaded, sample metadata was created to define the experimental conditions (Acetone, FDA, DMSO, Calcein), each with three biological replicates. A DESeqDataSet was constructed using the count matrix and metadata, with experimental

condition specified as the design formula (~ condition). Lowly expressed genes (total counts < 10 across all samples) were removed to increase statistical power. Gene-wise dispersions were estimated (estimateDispersionsGeneEst), and the corresponding estimates were directly assigned to the DESeq2 object. Wald tests were applied and pairwise contrasts were performed for: FDA vs Acetone, Calcein vs DMSO, DMSO vs Acetone. For each contrast, log2 fold changes, standard errors, and adjusted p-values (Benjamini–Hochberg FDR correction) were computed. Variance-stabilizing transformation (VST) was applied to the count data to reduce heteroscedasticity.

Statistical analysis and data visualization

Statistical analyses were performed using Student's *t*-tests in RStudio (v4.4.1) or Microsoft Excel. Significance thresholds were $P \leq 0.05$ *(*), $P \leq 0.01$ (**), and $P \leq 0.001$ (***)*. For multiple factor analysis, ANOVA with Tukey HSD correction in RStudio (v4.4.1) were applied, with different letters indicating significant differences ($p \leq 0.05$). Graphical visualizations were generated using ggplot2 (RStudio) for UMAPs and QC parameters, and Microsoft Excel for distribution analysis (QC parameters, cell types) and qRT-PCR results.

Sequencing and data analysis

Libraries were sequenced by Novogene or BGI (30–75 Gb/sample, paired-end PE150 bp) or BGI (PE100). Data analysis was performed following provider pipelines with additional custom filtering steps.

scRNA-seq data analysis and QC parameters

For detailed script and calculations, refer to: https://github.com/scPlantCoder/scRNA-seq

scRNA-seq mapping and preprocessing: Raw FASTQ files generated by Novogene or BGI were mapped to the *Medicago truncatula* reference genome (V1.9: https://medicago.toulouse.inra.fr/MtrunA17r5.0-ANR/) using alevin-fry implemented via simpleaf V0.15.1. We constructed a spliced + unspliced reference index using parameters --ref-type spliceu and the default intron flanking read length of 90 basepairs (--rlen 90). The corresponding V1.9 GTF annotation file was provided with minor modifications; the transcript_id tag was modified to be specified as MtrunA17_XXtXXXXXX instead of MtrunA17_XXgXXXXXX. Barcode structures were adjusted according to the manufacturer's recommendations in line with the provided format by 10x Genomics® or BD Rhapsody™.

Ambiguous Reads: Ambiguous reads across datasets were assigned to either unspliced or spliced in the same ratio as confidently assigned reads, or if not possible, assigned depending on the sample origin to either unspliced in nuclei samples, or to spliced in protoplast samples.

Ambient RNA removal: The combined resolved matrix (resolved spliced + resolved unspliced) was subject to ambient RNA removal using CellBender V0.3.0 in –remove-background mode using a learning rate of $5*10^{-5}$. This ambient RNA corrected count matrix was used for all downstream processing unless otherwise stated.

Quality control metric computation: Raw and corrected count matrices were processed in python (V3.7) and stored into AnnData objects, dedicating a layer to each count type (spliced, unspliced, ambiguous, resolved spliced, resolved unspliced, total raw counts and corrected counts). This layered AnnData object was subsequently used to compute a series of quality control (QC) metrics from the most relevant data source. QC metrics were computed aiming to evaluate cellular complexity and sparsity, and transcript composition.

Cellular complexity and sparsity: To assess the information content of each cell, we computed complexity and sparsity values. These metrics capture per-cell transcript diversity and cellular transcriptome completeness. Low complexity values indicate domination of the cell's transcriptome by a few genes, potentially reflecting poor capture efficiency. In addition to complexity, sparsity was used to assess the 'completeness' of a cell's transcriptome. High sparsity values reflect transcriptionally relatively depleted cells, indicative of leaking, damaged or otherwise low-quality cells.

Transcript composition: Metrics describing transcript compositions were calculated to identify signatures of cellular stress, organelle disruption, leakage, and/or ambient RNA contamination. Mitochondrial fractions, chloroplast fractions, and ribosomal RNA fractions were defined as the proportion of UMIs mapping to a given gene set of interest and computed for each cell. Comparatively high levels of mitochondrial and chloroplast fractions were interpreted as indicators of cell damage and leakage, due to their protection by the organelle-membrane. Elevated ribosomal RNA fractions were primarily taken as indications of technical biases, such as inefficient rRNA depletion during library preparation.

Cell filtering

Following computation of QC-metrics, we performed cell filtering using adaptive threshold based on multiple QC-metrics. Initial putative droplets were filtered based on a knee-point in the barcode-rank plot, setting the minimum number of UMIs to 500. Cells were next filtered based on within-sample distributions of total UMI counts, total gene count, complexity, mitochondrial fraction, chloroplast fraction, and ribosomal RNA fraction. Any cell outside the three median absolute deviation (MAD) boundary were flagged as outliers. Additionally, we modelled the relationship between the total number of UMIs and gene counts. Cells with residuals exceeding three MAD were flagged.

For multi-metric QC, we fitted a gaussian mixture model on all continuous metrics that were either approximately normal or could be log/logit-transformed to be approximately normal (two or more of: spliced fraction, mitochondrial fraction, chloroplast fraction, ribosomal RNA fraction, total UMIs, total gene count and ambient fraction). Cells were flagged as outliers when their Mahalanobis distance to the cluster centroid exceeded a chi-squared threshold.

In parallel, we utilized all computed QC-metrics to fit an isolation forest model. Any cell below the contamination-threshold was flagged as an outlier. Finally, we used a distance-of-distance transformation to remove noise points (Liu and Vinck, 2022).

Nuclei Integration and UMAP visualization

To integrate samples, we fitted a scVI model using scVI-tools (v1.3.1), with model hyperparameters optimized using the autotune module. For visualization, UMAP embeddings were computed using the Scanpy implementation, using the learned scVI latent representation as input. Cosine distance was used as the neighborhood metric, and the number of nearest neighbors was set to the square root of the number of observations, providing adaptive neighborhood scaling across datasets.

ACKNOWLEDGEMENTS

We would like to thank Viola Willemsen for constructive feedback on this work.

REFERENCES

Adema, K., Schon, M. A., Nodine, M. D. and Kohlen, W. (2024). Lost in space: what single-cell RNA sequencing cannot tell you. *Trends Plant Sci* 29, 1018–1028.

Bakken, T. E., Hodge, R. D., Miller, J. A., Yao, Z., Nguyen, T. N., Aevermann, B., Barkan, E., Bertagnolli, D., Casper, T., Dee, N., et al. (2018). Single-nucleus and single-cell transcriptomes compared in matched cortical cell types. *PLoS One* 13,.

Birnbaum, K., Shasha, D. E., Wang, J. Y., Jung, J. W., Lambert, G. M., Galbraith, D. W. and Benfey, P. N. (2003). A Gene Expression Map of the Arabidopsis Root. *Science (1979)* 302, 1956–1960.

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.

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.

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.

Farmer, A., Thibivilliers, S., Ryu, K. H., Schiefelbein, J. and Libault, M. (2021). Single-nucleus RNA and ATAC sequencing reveals the impact of chromatin accessibility on gene expression in Arabidopsis roots at the single-cell level. *Mol Plant* 14, 372–383.

Geldner, N. (2013). The endodermis. *Annu Rev Plant Biol* 64, 531–558.

Grones, C., Eekhout, T., Shi, D., Neumann, M., Berg, L. S., Ke, Y., Shahan, R., Cox, K. L., Gomez-Cano, F., Nelissen, H., et al. (2024). Best practices for the execution, analysis, and data storage of plant single-cell/nucleus transcriptomics. *Plant Cell* 36, 812–828.

Heidstra, R., Geurts, R., Franssen, H., Spaink, H. P., Van Kammen, A. and Bisseling, T. (1994). Root Hair Deformation Activity of Nodulation Factors and Their Fate on Vicia sativa. *Plant Physiol* 105, 787–797.

Huang, C. N., Cornejo, M. J., Bush, D. S. and Jones, R. L. (1986). Estimating viability of plant protoplasts using double and single staining. *Protoplasma* 135, 80–87.

Jabloune, R., Khalil, M., Ben Moussa, I. E., Simao-Beaunoir, A. M., Lerat, S., Brzezinski, R. and Beaulieu, C. (2020). Enzymatic Degradation of p-Nitrophenyl Esters, Polyethylene Terephthalate, Cutin, and Suberin by Sub1, a Suberinase Encoded by the Plant Pathogen Streptomyces scabies. *Microbes Environ* 35,.

Jhu, M. Y. and Oldroyd, G. E. D. (2023). Dancing to a different tune, can we switch from chemical to biological nitrogen fixation for sustainable food security? *PLoS Biol* 21, 1–24.

Jia, N., Zhu, Y. and Xie, F. (2018). An efficient protocol for model legume root protoplast isolation and transformation. *Front Plant Sci* 9,.

Kao, P., Schon, M. A., Mosiolek, M., Enugutti, B. and Nodine, M. D. (2021). Gene expression variation in Arabidopsis embryos at single-nucleus resolution. *Development (Cambridge)* 148, 1–14.

Kersey, H. N., Acri, D. J., Dabin, L. C., Hartigan, K., Mustaklem, R., Park, J. H. and Kim, J. (2025). Comparative analysis of nuclei isolation methods for brain single-nucleus RNA sequencing. *bioRxiv* 2025.03.25.645306.

Kimble, A. L., Silva, J., Omar, O. M., Murphy, M., Hensel, J. A., Nicholas, S. A. E., Jellison, E. R., Reese, B. and Murphy, P. A. (2022). A method for rapid flow-cytometric isolation of endothelial nuclei and RNA from archived frozen brain tissue. *Laboratory Investigation* 102, 204–211.

Kohlen, W., Ng, J. L. P., Deinum, E. E. and Mathesius, U. (2018). Auxin transport, metabolism, and signalling during nodule initiation: Indeterminate and determinate nodules. *J Exp Bot* 69, 229–244.

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 (1979)* 302, 630–633.

Liu, J. and Vinck, M. (2022). Improved visualization of high-dimensional data using the distance-of-distance transformation. *PLoS Comput Biol* 18, e1010764.

Liu, Z., Yang, J., Long, Y., Zhang, C., Wang, D., Zhang, X., Dong, W., Zhao, L., Liu, C., Zhai, J., et al. (2023). Single-nucleus transcriptomes reveal spatiotemporal symbiotic perception and early response in Medicago. *Nat Plants*.

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

Lv, Z., Jiang, S., Kong, S., Zhang, X., Yue, J., Zhao, W., Li, L. and Lin, S. (2024). Advances in Single-Cell Transcriptome Sequencing and Spatial Transcriptome Sequencing in Plants. *Plants (Basel)* 13,.

Miller, D. D., De Ruijter, N. C. A., Mie, A. and Emons, C. (1997). *From signal to form: aspects of the cytoskeleton-plasma membrane-cell wall continuum in root hair tips*.

Nadelmann, E. R., Gorham, J. M., Reichart, D., Delaughter, D. M., Wakimoto, H., Lindberg, E. L., Litviňukova, M., Maatz, H., Curran, J. J., Ischiu Gutierrez, D., et al. (2021). Isolation of Nuclei from Mammalian Cells and Tissues for Single-Nucleus Molecular Profiling. *Curr Protoc* 1,.

Pereira, W. J., Boyd, J., Conde, D., Triozzi, P. M., Balmant, K. M., Dervinis, C., Schmidt, H. W., Boaventura-Novaes, C., Chakraborty, S., Knaack, S. A., et al. (2024). The single-cell transcriptome program of nodule development cellular lineages in Medicago truncatula. *Cell Rep* 43,.

Ren, R., Gao, J., Yin, D., Li, K., Lu, C., Ahmad, S., Wei, Y., Jin, J., Zhu, G. and Yang, F. (2021). Highly Efficient Leaf Base Protoplast Isolation and Transient Expression Systems for Orchids and Other Important Monocot Crops. *Front Plant Sci* 12, 626015.

Roy, S., Liu, W., Nandety, R. S., Crook, A., Mysore, K. S., Pislariu, C. I., Frugoli, J., Dickstein, R. and Udvardi, M. K. (2020). Celebrating 20 Years of Genetic Discoveries in Legume Nodulation and Symbiotic Nitrogen Fixation. *Plant Cell* 32, 15–41.

Serrano, K., Bezrutczyk, M., Goudeau, D., Dao, T., O'Malley, R., Malmstrom, R. R., Visel, A., Scheller, H. V. and Cole, B. (2024). Spatial co-transcriptomics reveals discrete stages of the arbuscular mycorrhizal symbiosis. *Nature Plants 2024 10:4* 10, 673–688.

Seyfferth, C., Renema, J., Wendrich, J. R., Eekhout, T., Seurinck, R., Vandamme, N., Blob, B., Saeys, Y., Helariutta, Y., Birnbaum, K. D., et al. (2021). Advances and Opportunities in Single-Cell Transcriptomics for Plant Research. *Annu Rev Plant Biol* 72, 847–866.

Shaw, S. L., Dumais, J. and Long, S. R. (2000). *Cell Surface Expansion in Polarly Growing Root Hairs of Medicago truncatula*.

Sheen, J. (2001). Signal transduction in maize and Arabidopsis mesophyll protoplasts. *Plant Physiol* 127, 1466–1475.

Slyper, M., Porter, C. B. M., Ashenberg, O., Waldman, J., Drokhlyansky, E., Wakiro, I., Smillie, C., Smith-Rosario, G., Wu, J., Dionne, D., et al. (2020). A single-cell and single-nucleus RNA-Seq toolbox for fresh and frozen human tumors. *Nat Med* 26, 792–802.

Su, Y., Feng, T., Liu, C. Bin, Huang, H., Wang, Y. L., Fu, X., Han, M. L., Zhang, X., Huang, X., Wu, J. C., et al. (2023). The evolutionary innovation of root suberin lamellae contributed to the rise of seed plants. *Nature Plants 2023 9:12* 9, 1968–1977.

Sun, Y., Sun, J., Lin, C., Zhang, J., Yan, H., Guan, Z. and Zhang, C. (2024). Single-Cell Transcriptomics Applied in Plants. *Cells* 13,.

Thibivilliers, S., Anderson, D. and Libault, M. (2020). Isolation of Plant Root Nuclei for Single Cell RNA Sequencing. *Curr Protoc Plant Biol* 5, 1–12.

Timmers, A. C. J., 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.

Tucker, M. R., Lou, H., Aubert, M. K., Wilkinson, L. G., Little, A., Houston, K., Pinto, S. C. and Shirley, N. J. (2018). Exploring the Role of Cell Wall-Related Genes and Polysaccharides during Plant Development. *Plants 2018, Vol. 7, Page 42* 7, 42.

Wang, K., Zhao, C., Xiang, S., Duan, K., Chen, X., Guo, X. and Sahu, S. K. (2023). An optimized FACS-free single-nucleus RNA sequencing (snRNA-seq) method for plant science research. *Plant Science* 326, 111535.

Wu, H., Kirita, Y., Donnelly, E. L. and Humphreys, B. D. (2019). Advantages of single-nucleus over single-cell RNA sequencing of adult kidney: Rare cell types and novel cell states revealed in fibrosis. *Journal of the American Society of Nephrology* 30, 23–32.

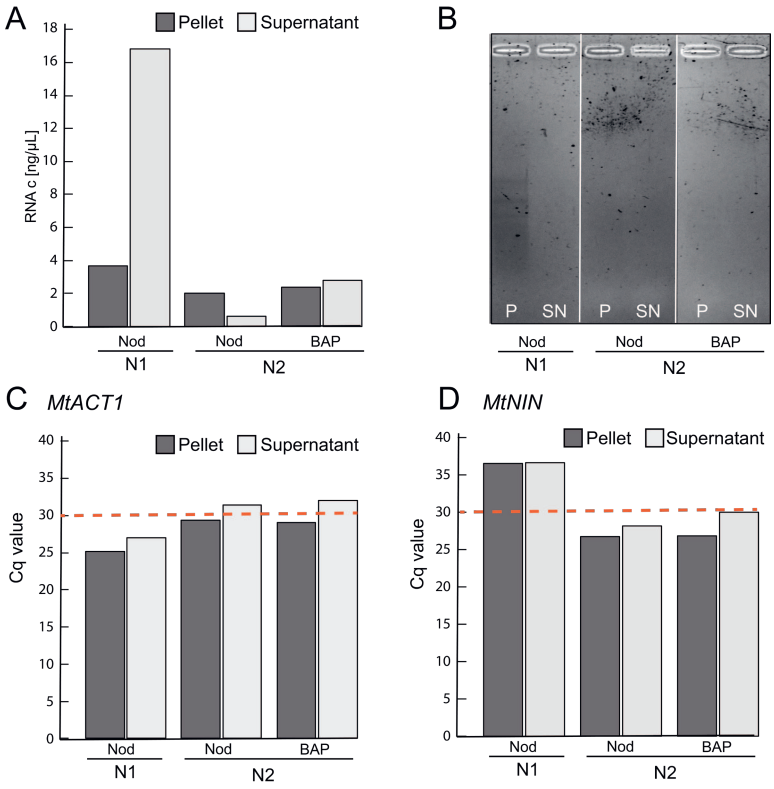
Xiao, T. T., Schilderink, S., Moling, S., Deinum, E. E., Kondorosi, E., Franssen, H., Kulikova, O., Niebel, A. and Bisseling, T. (2014). Fate map of Medicago truncatula root nodules. *Development (Cambridge)* 141, 3517–3528.

Yan, J., Bisseling, T., Cao, Q., Li, H., Yan, J., Bisseling, T., Cao, Q. and Li, H. (2025). The adaptive root: endodermal plasticity and microbial composition.

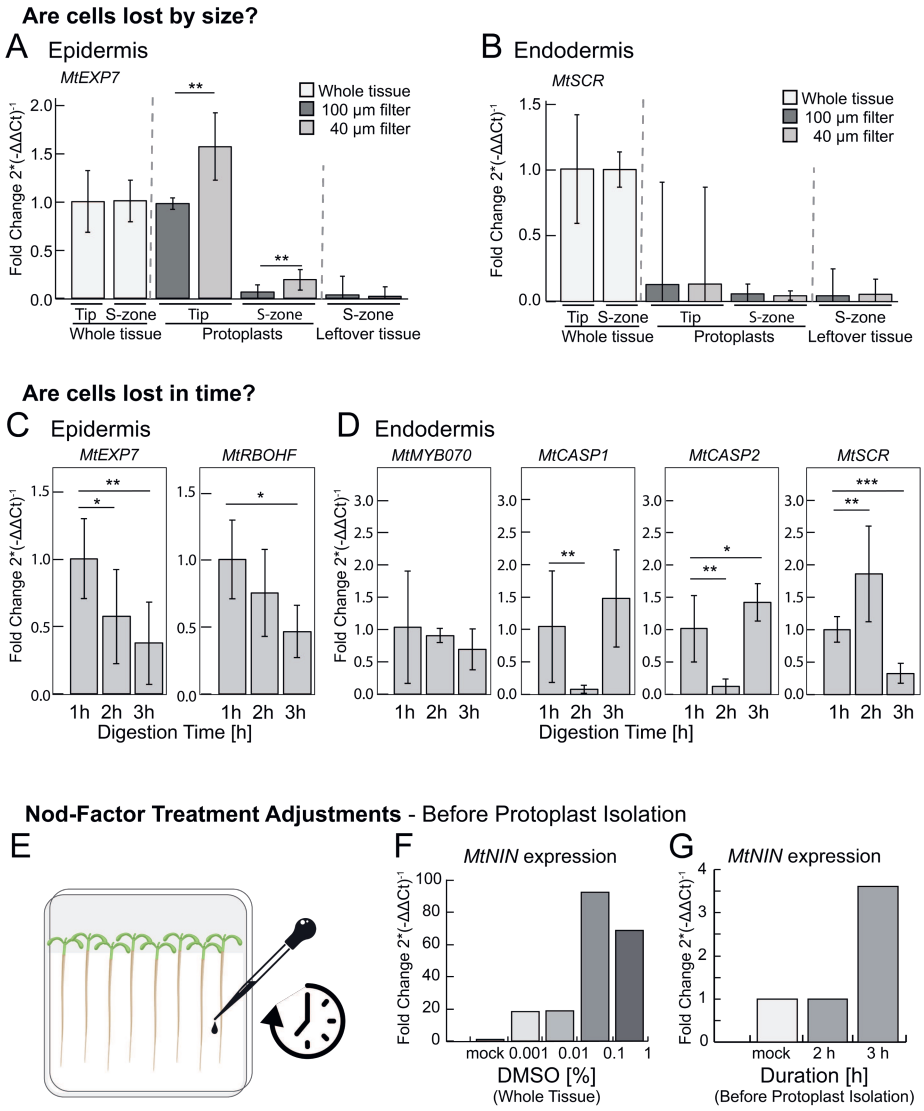
Ye, Q., Zhu, F., Sun, F., Wang, T. C., Wu, J., Liu, P., Shen, C., Dong, J. and Wang, T. (2022). Differentiation trajectories and biofunctions of symbiotic and un-symbiotic fate cells in root nodules of Medicago truncatula. *Mol Plant* 15, 1852–1867.

Zhang, S., Dong, Z., Shi, J., Yang, C., Fang, Y., Chen, G., Chen, H. and Tian, C. (2022). Enzymatic hydrolysis of corn stover lignin by laccase, lignin peroxidase, and manganese peroxidase. *Bioresour Technol* 361,.

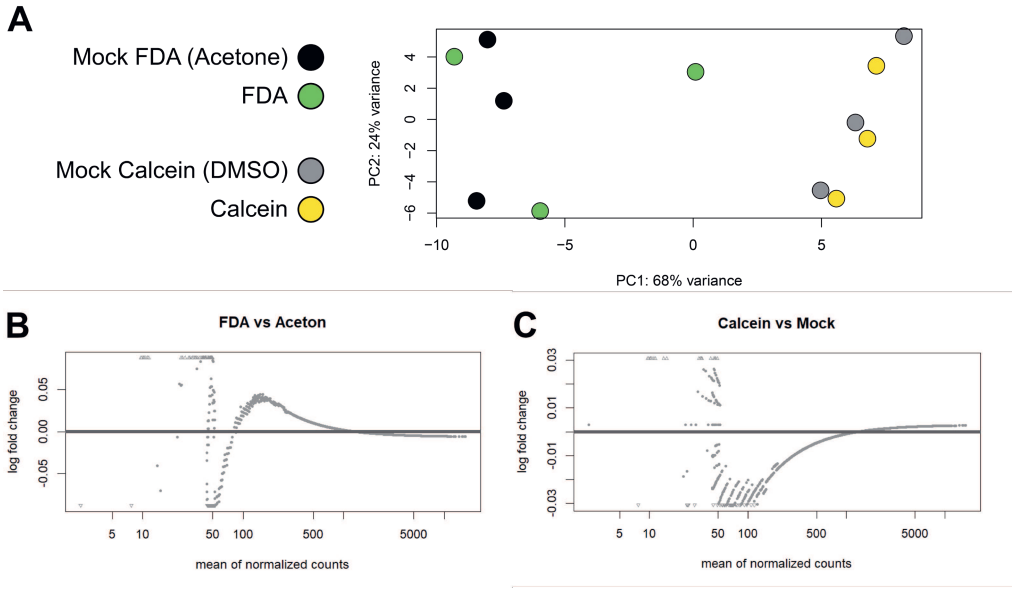
SUPPLEMENTAL FIGURES



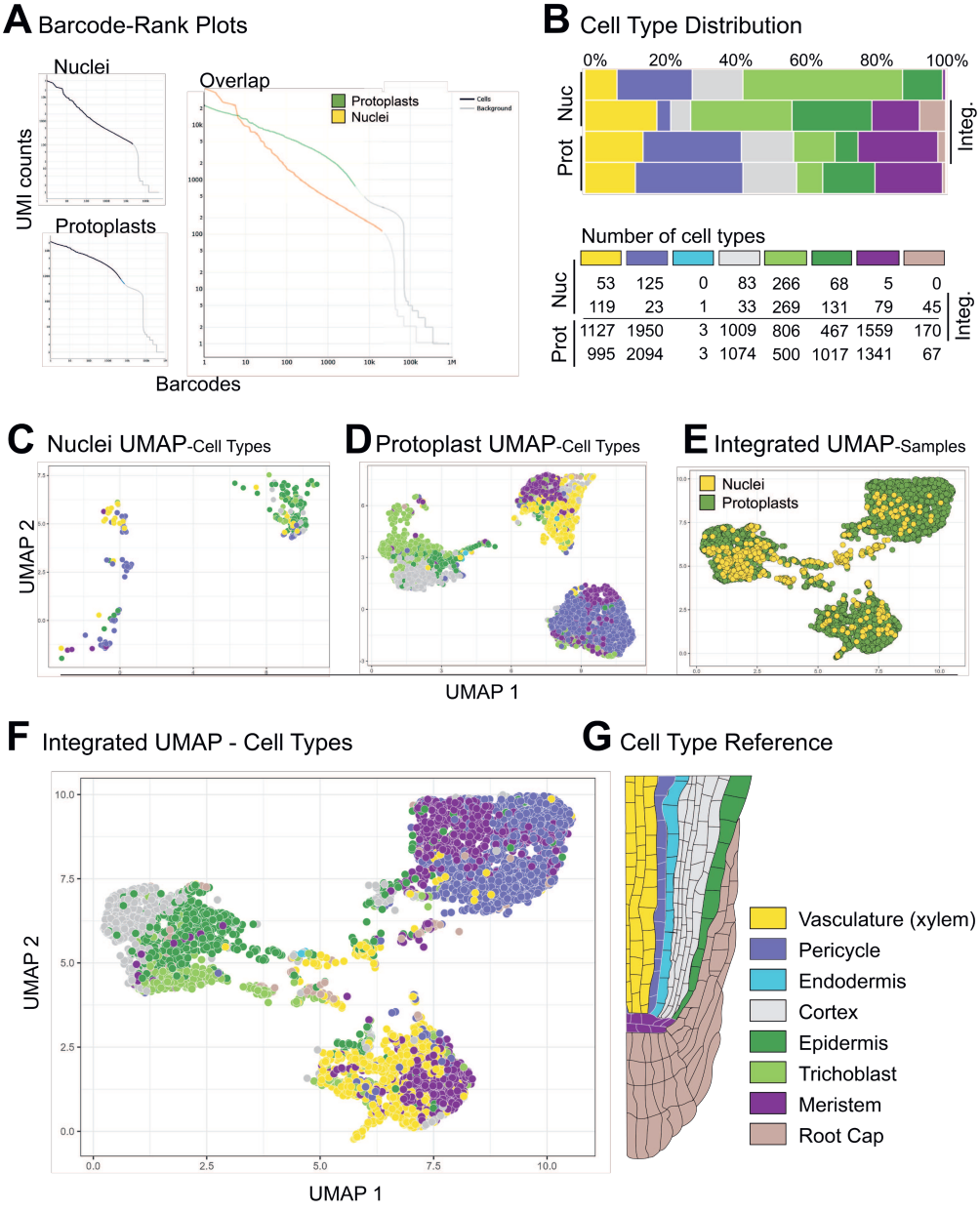
Supplemental Figure S1 | mRNA localization in pellet versus supernatant is insufficient to evaluate nuclei isolation methods. Nuclei isolation methods N1 (centrifugation-based) and N2 (FACS-based) were compared under two treatments known to strongly induce *MtNIN* expression: the cytokinin benzylaminopurine (BAP) and Nod factors. To assess potential ambient RNA contamination, RNA was extracted separately from the pellet (P) and supernatant (SN) fractions. **(A)** RNA concentration in each sample. **(B)** 20 ng of RNA resolved on a 1% agarose gel stained with ethidium bromide. **(C)** Cq values from qPCR targeting the housekeeping gene *MtACT*, showing differential abundance between pellet and supernatant fractions. **(D)** Cq values for *MtNIN*, a BAP- and Nod-responsive gene. The orange dotted line marks the Cq threshold of 30, above which expression is considered very low or not detected.



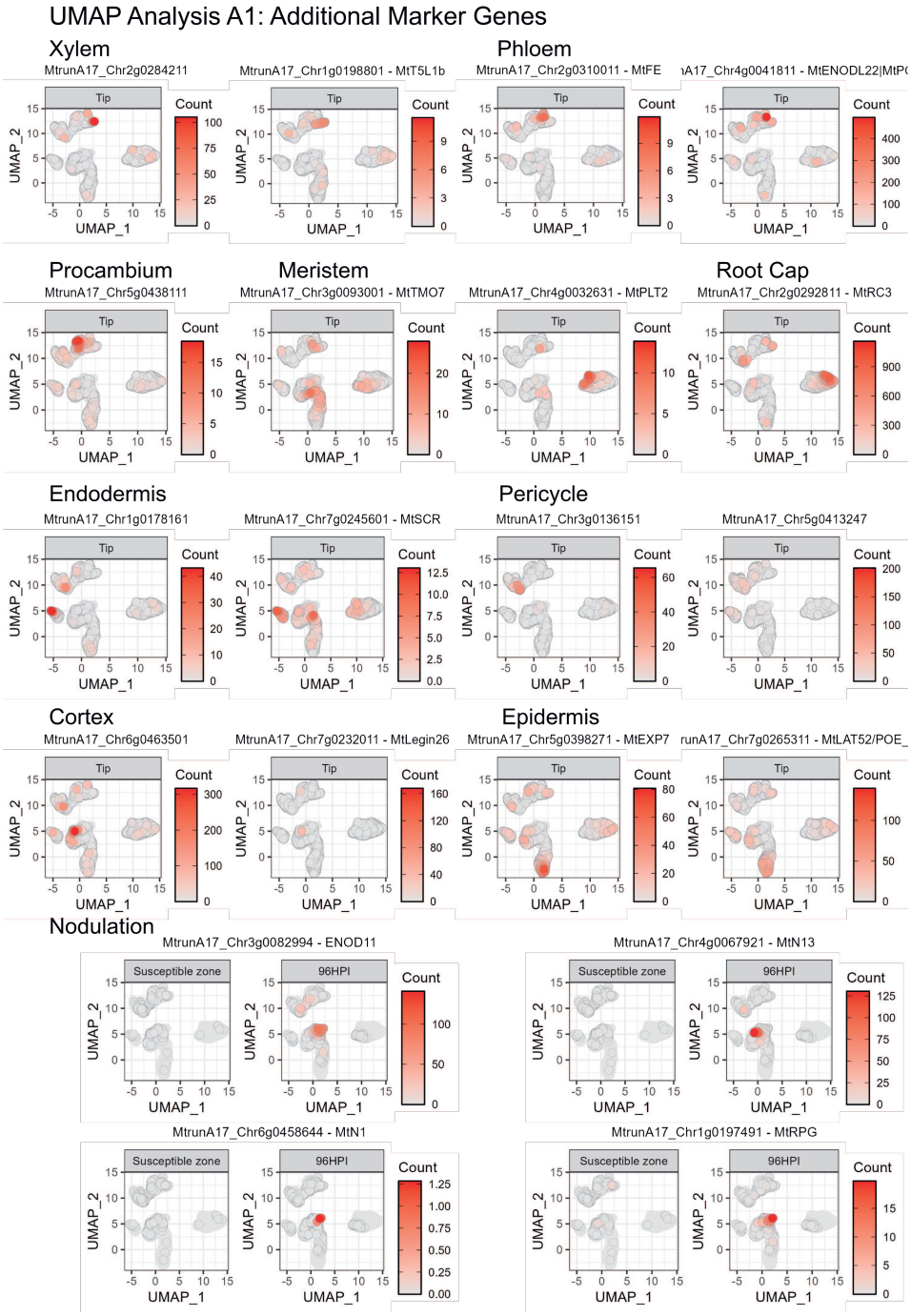
Supplemental Figure S2 | Protocol adaptations before and during protoplast isolation. (A–B) RNA was isolated from protoplasts filtered through either 100 µm or 40 µm strainers, from root tip or susceptible zone (S-zone) tissues, and compared to RNA from whole tissue and residual material post-filtration. **(A)** qPCR for the epidermis marker *MtRBOHF*. **(B)** qPCR for the endodermis marker *MtSCR*. **(C–D)** Impact of digestion duration (1, 2, or 3 hours) on cell-type–specific transcript recovery from protoplasts. **(C)** qPCR for epidermal markers *MtEXP7* and *MtRBOHF*. **(D)** qPCR for endodermal markers *MtMYB70*, *MtCASP1*, *MtCASP2*, and *MtSCR*. **(E–F)** Optimization of treatment conditions prior to protoplast isolation. (E) Schematic of in planta Nod factor application. **(F)** Effect of DMSO concentration on *MtNIN* induction following Nod factor treatment, assessed via qPCR on RNA from whole root tissue. **(G)** Comparison of Nod factor treatment duration (2 vs. 3 hrs with 0.1% DMSO) prior to 2 hrs protoplast isolation. RNA was extracted from isolated protoplasts, and *MtNIN* expression was assessed by qPCR. Fold Change normalized to the housekeeping gene *MtUBQ*. **(A-D)** Bars represent mean with 95% CI (n ≥ 2). Asterisks indicate significance between treatments (* p ≤0.05, ** p ≤ 0.01, *** p ≤ 0.001). **(F-G)** Bars represent single replicate.



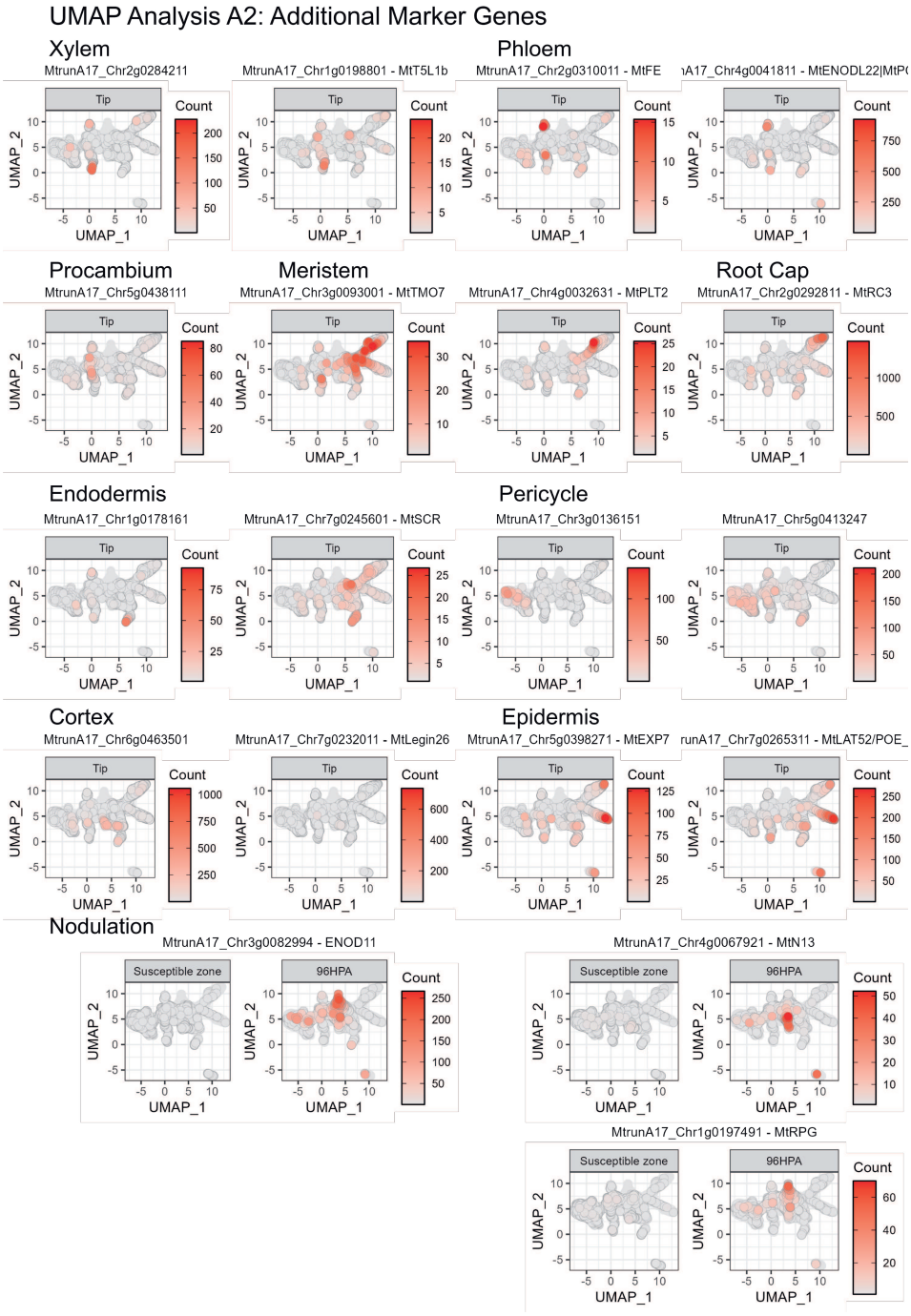
Supplemental Figure S3 | Transcriptomic analysis of *Arabidopsis thaliana* protoplasts, stained with the cell dyes fluorescein diacetate (FDA) or Calcein. (A) Principal Component Analysis (PCA) identifies acetone or DMSO (dimethyl sulfoxide) as main difference between samples. Either Calcein (yellow) or FDA (green) do not influence the transcriptome. **(B,C)** MA plots to visualize transcriptomic differences between **(B)** FDA and **(C)** Calcein against their respective control. X-axis: average expression strength of each gene across all samples (log scale). Y-axis: log2 fold change of expression between tested conditions. Grey points: no statistically significant difference between treatments (adjusted p-value >0.05). Lower expressed genes (left) are more variable in fold change as a consequence of low overall counts. Highly expressed genes (right) cluster around 0, indicating a stable expression between treatments.



Supplemental Figure S4 | Data comparison and integration of nuclei (n=600) and protoplasts (n=7091) from untreated susceptible zone during scRNA-seq analysis. (A) Barcode-Rank plots of nuclei and protoplasts separately and overlayed. **(B)** Cell type distribution of nuclei and protoplast datasets before (top, bottom) and after (inside) integration, with actual numbers of cell type distribution according to color scheme of (F). **(C-F)** UMAP representation of **(C)** nuclei, **(D)** protoplasts, **(E)** integrated nuclei with protoplasts, colored by sample, and **(F)** integrated data, colored by cell types. **(G)** Schematic representation of cell types and corresponding coloring scheme. **Abbreviations**: Nuc, nuclei; Prot, protoplasts; Integ., integrated; UMAP, Uniform Manifold Approximation and Projection.



Supplemental Figure S5 | UMAP Analysis A1-10x Genomics. Additional markers for cell-type identification. Expression in counts per cell.



Supplemental Figure S6 | UMAP Analysis A2-BD Rhapsody. Additional markers for cell-type identification. Expression in counts per cell.

SUPPLEMENTAL TABLE

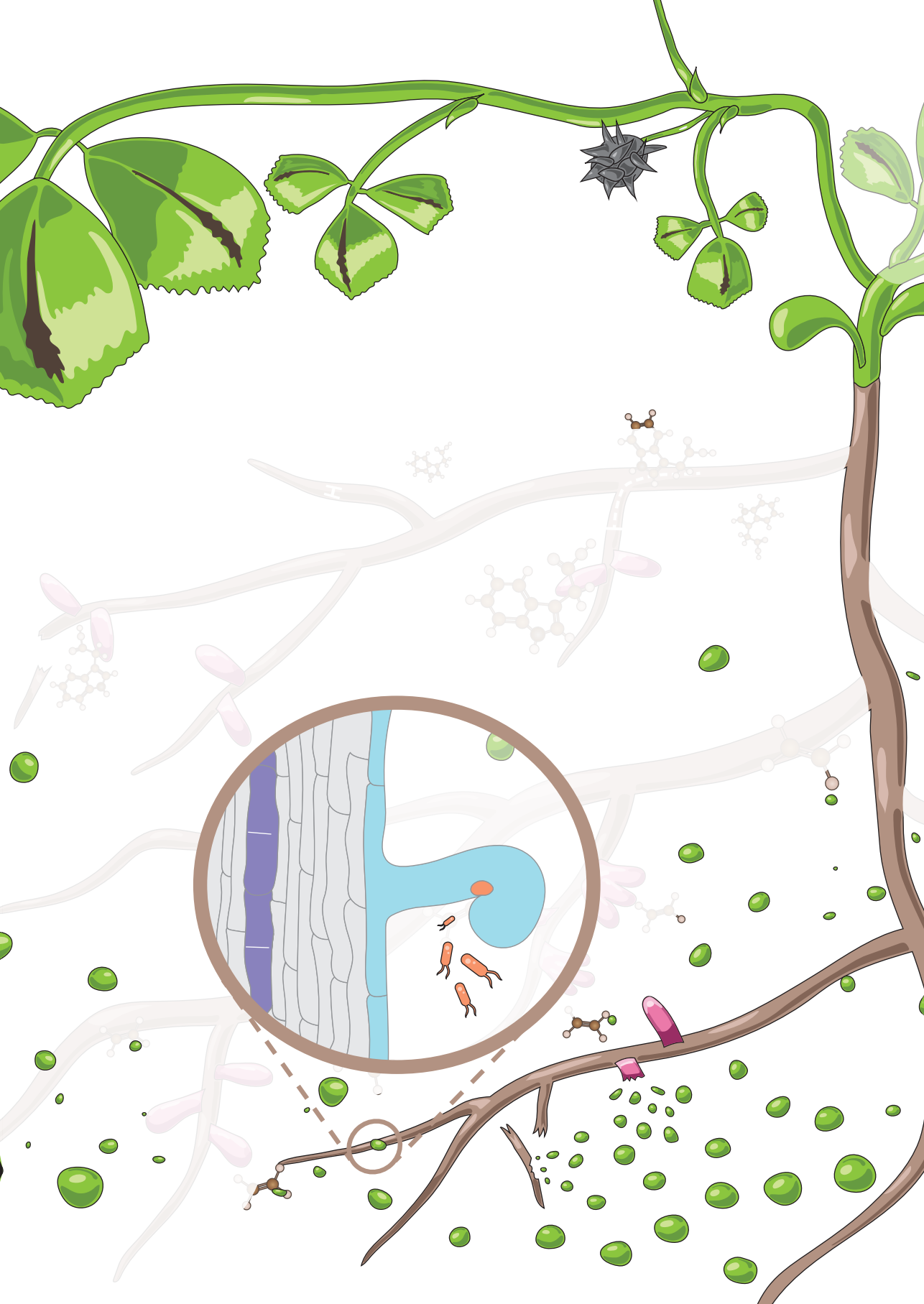
Supplemental Table S1 | Primers used for qRT-PCR, with their respective target, their gene ID based on MtrunA17r5.0, their name, and their sequence

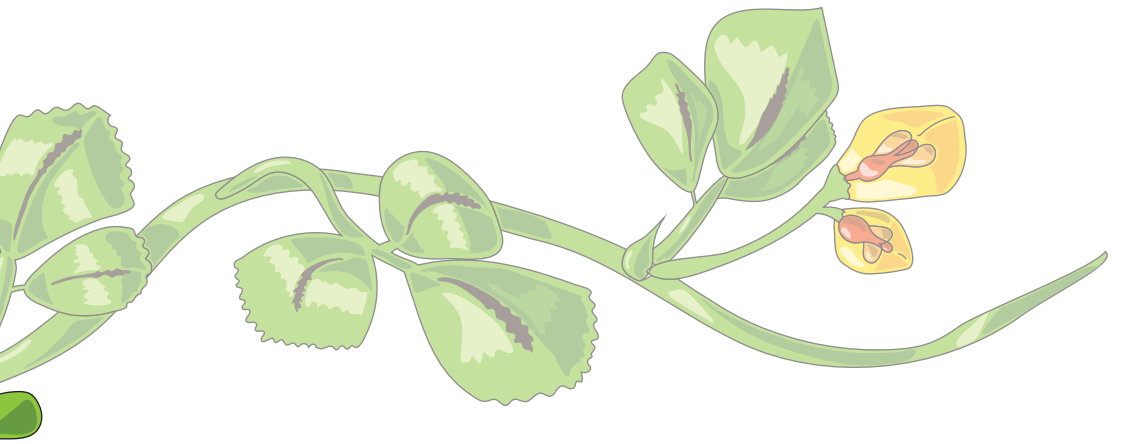
Target	Gene ID (MtrunA17r5.0)	Primer Name	Sequence
Housekeeping	MtUBQ MtrunA17_Chr4g0048591	MtUBQ_Fw	CACCTCCAATGTAATGGTCTTTCC
		MtUBQ_Rv	CCCTTCATCTTGTCCTTCGTCTG
Endodermis	MtMYB70 or MtMYB36 MtrunA17_Chr4g0072751	qEnMYB36_F	AAGCATTGGAAGCAGGTGGT
		qEnMYB36_R	TCGCTTTCCGAGTAATTTCTTCT
	MtCASP1 MtrunA17_Chr1g0178161	qEnCASP1_F	CCCTTTCTTCACTCAGTTCCT
		qEnCASP1_R	AGGGATAAGACCAGGTAGCCA
	MtCASP2 MtrunA17_Chr1g0178181	qEnCASP2_F	TCACTCAGTTCTTTCAGTTTGAAGC
		qEnCASP2_R	AGACAAGACCAAATAGCCAGCA
	MtCASP3 MtrunA17_Chr1g0178191	qEnCASP3_F	AGCCATCATTCGCCCTCATG
		qEnCASP3_R	CTTGATTTCCATTATGTGCCAAGT
	MtSCR MtrunA17_Chr7g0245601	qEnSCR_F	TGTTGCTGTTCATTGGTTGCA
		qEnSCR_R	ACTGTTACCACTTTAGGTGCCAA
Epidermis	MtEXP7 MtrunA17_Chr1g0178191	qEpEXP7_F	CACCTCCGACGTTCATCACA
		qEpEXP7_R	AGCTGGTATCGTCCCTGTCA
	MtRBOHF MtrunA17_Chr7g0238371	qEpRBOHF_F	TGACAAGAATGGAGATGGTAGGT
		qEpRBOHF_R	TCAGCATGTTGTTTAAGATTTCCCA
Nodule induction	MtNIN MtrunA17_Chr1g0182631	qMtNIN_F	GCCTTCTTCTGCTTCTGAACC
		qMtNIN_R	CCAGTTGCAGGTTGTGACTT
	MtNF-YA1 MtrunA17_Chr1g0177091	qMtNF-YA1_F	TGCAATGCGTAGACCAAGAG
		qMtNF-YA1_R	CTCCGTTTGTCTCCTTCGAC

Supplemental Table S2 | Cell-type specific marker genes used for cluster identification of scRNA-seq UMAPs during data analysis.

Tissue	Gene ID (MtrunA17r5.0)	Common Name	Function	Arabidopsis Ortholog
Epidermis	MtrunA17_Chr7g0238371	MtRbohF	RESPIRATORY BURST OXIDASE HOMOLOG F	AT5G60010.1, AT3G45810.1
	MtrunA17_Chr5g0398271	MtEXP7	EXPANSIN7	AT1G62980.1, AT1G12560.1
	MtrunA17_Chr7g0265311	MtLAT52/ POE_16	LATE ANTHER TOMATO52/POLLEN OLE E 1 ALLERGEN AND EXTENSIN	AT3G09925
Endodermis	MtrunA17_Chr4g0072751	MtMYB070	MYELOBLASTOSIS (MYB) TRANSCRIPTION FACTOR 70	AT5G57620.1
	MtrunA17_Chr1g0178161		PUTATIVE CASPARIAN STRIP MEMBRANE PROTEIN	AT2G36100.1, AT3G11550.1, AT5G06200.1
	MtrunA17_Chr7g0245601	MtSCR	SCARECROW (in situ hybridization in Roux et al., 2014)	AT3G54220.1
Xylem	MtrunA17_Chr1g0200561	MtVND7-like	VASCULAR-RELATED NAC-DOMAIN7 (VND7)-like	AT1G71930.1
	MtrunA17_Chr1g0198801	MtT5L1b	TMO5-LIKE1b (bHLH transcription factor)	AT3G25710.1, AT1G68810.1
	MtrunA17_Chr2g0284211		Putative cellulose synthase (UDP-forming)	AT4G18780.1

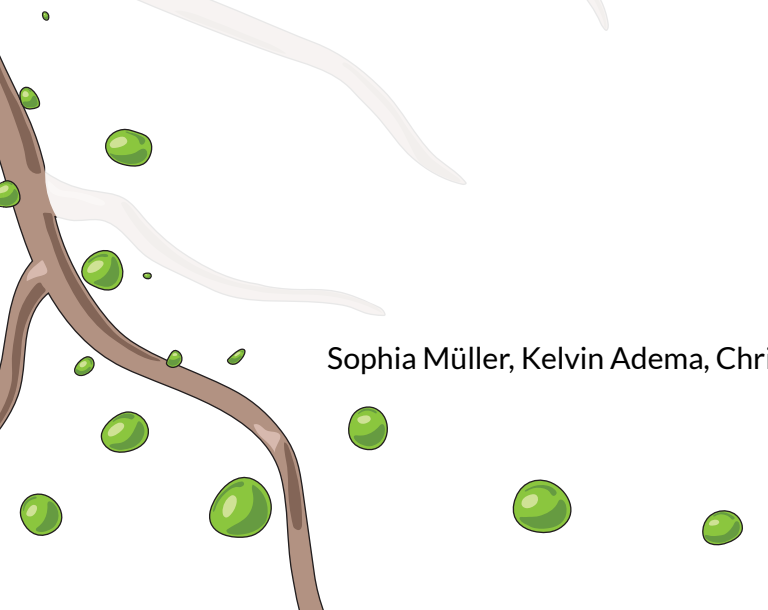
Phloem	MtrunA17_Chr2g0310011	MtFE	FLOWERING LOCUS E	AT1G79430.2
	MtrunA17_Chr4g0041811	MtENODL22/ MtPCY54	ENOD-like 22|Plantacyanin/Chemocyanin 54	AT3G20570.1
	MtrunA17_Chr1g0174581		Putative C2 domain, phosphoribosyltransferase, C2 domain superfamily	AT5G06850.1
Cortex	MtrunA17_Chr3g0120411	MtnsLTP77	non-specific Lipid Transfer Protein	-
	MtrunA17_Chr7g0232011	MtLegin26	LEGINSULIN	-
	MtrunA17_Chr6g0463501		putative protein	-
Pericycle	MtrunA17_Chr5g0413247			-
	MtrunA17_Chr5g0405701	MtHK1	HISTIDINE KINASE 1 (no CHASE domain)	AT2G17820.1
	MtrunA17_Chr3g0136151		Putative tripeptidyl-peptidase II	AT1G20150.1, AT1G20160.1
Procambium	MtrunA17_Chr3g0114551		Putative WEB family protein	AT3G56270.1, AT2G40480.1
	MtrunA17_Chr5g0438111		Putative WEB family protein	AT3G56270.1, AT2G40480.1
Meristem	MtrunA17_Chr3g0093001	MtTMO7	TARGET OF MONOPTEROS 7	AT1G74500.1
	MtrunA17_Chr7g0249621	MtPLT4/ MtBBM	PLETHORA ERF|BABYBOOM	-
	MtrunA17_Chr4g0032631	MtPLT2	PLETHORA 2 (ERF transcription factor)	AT3G20840.1, AT1G51190.1
Root Cap	MtrunA17_Chr2g0309331	MtSMB	SOMBRERO (auxin-responsive, Putative transcription factor NAM family)	AT1G79580.1
	MtrunA17_Chr2g0292811	MtRC3	ROOT CAP 3	AT3G19430.1
Nodule	MtrunA17_Chr5g0448621	MtNIN	NODULE INCEPTION	-
	MtrunA17_Chr3g0082994	MtENOD11	EARLY NODULIN-11	-
	MtrunA17_Chr1g0177091	MtNF-YA1	NUCLEAR FACTOR Y SUBUNIT A1	AT5G06510.1, AT3G05690.1
	MtrunA17_Chr1g0197491	MtRPG	RHIZOBIUM-DIRECTED POLAR GROWTH	-
	MtrunA17_Chr6g0458644	MtN1	MTNODULIN 1 (Y10456 GENBANK)	-
	MtrunA17_Chr4g0067921	MtN13	MTODULIN 13 (Y10455.1 GENBANK)	-





Chapter 5|

Single-Cell Transcriptomics-guided identification of transcription factors with pericycle-enriched expression in *Medicago truncatula*



Sophia Müller, Kelvin Adema, Chris Roelofsen, Viola Willemsen, Wouter Kohlen

ABSTRACT

Legume nodulation relies on the remarkable ability of differentiated root cells to re-entering the cell cycle and forming a new symbiotic organ. Yet, the transcriptional regulators that confer pericycle and cortical competence to achieve this remain poorly defined. Here, we combine bulk and single-cell RNA sequencing (scRNA-seq) in *Medicago truncatula* to identify transcription factors (TFs) that mark and potentially regulate pericycle activation during early nodule initiation. By intersecting whole-root time-course RNA-seq with our root scRNA-seq atlas, we shortlisted ~600 pericycle-enriched, nodulation-responsive genes and focused on three TFs from distinct families: *MtbZIP43* (bZIP), *MtDREB3* (AP2/ERF), and *MtSCL14* (GRAS). scRNA-seq and *promoter::GUS* analyses reveal that all three genes are expressed in the pericycle and vasculature with gene-specific expression patterns and dynamic regulation upon rhizobial inoculation, including a striking pericycle-specific induction of *MtSCL14* in the susceptible zone. RNA interference did not measurably alter nodulation, likely reflecting either redundancy or incomplete knockdown, whereas constitutive overexpression of each TF reduced nodule number, with *MtDREB3* overexpression additionally increasing root biomass. Our work establishes *MtbZIP43*, *MtDREB3* and *MtSCL14* as robust pericycle and vascular markers, illustrates both the power and caveats of scRNA-seq–driven candidate discovery, and provides a framework for refining transcriptional regulators that control early stages of symbiotic organogenesis.

INTRODUCTION

Legumes are renowned for their high protein content, a trait made possible through their symbiotic association with rhizobial bacteria capable of fixing atmospheric nitrogen (Peoples et al., 1995). This natural nitrogen-fixing ability has long attracted scientific interest, particularly for its potential to reduce reliance on synthetic fertilizers by transferring this trait to non-nodulating plant species (Jhu and Oldroyd, 2023). The evolution of a specialized organ, the nodule, provides a unique system for investigating fundamental questions in plant developmental biology. *Medicago truncatula* (Medicago) has been established as a model legume for studying nodulation, a process initiated by the recognition of rhizobial Nod factors (Roy et al., 2020). Successful nodule organogenesis requires remarkable developmental plasticity, as differentiated root cells must be reprogrammed to re-enter the cell cycle and proliferate to form the nodule primordium. Early during this process, cell divisions are initiated in the pericycle and subsequently expand toward the endodermis and inner cortex, ultimately giving rise to the developing nodule primordium (Schiessl et al., 2019; Xiao et al., 2014; Xiao et al., 2025). Key transcriptional regulators such as NODULE INCEPTION (NIN), LATERAL ORGAN BOUNDARIES DOMAIN PROTEIN16 (LBD16), and NUCLEAR TRANSCRIPTION FACTOR Y SUBUNIT A1 (NF-YA1) have been identified as central players in activating these cell divisions and establishing nodule identity (Schauser et al., 1999; Schiessl et al., 2019; Soyano et al., 2013). However, the molecular mechanisms that confer pericycle competence during symbiotic reprogramming remain poorly understood. In particular, how these pathways function at the level of individual cells, and which developmental regulators are specifically activated within the pericycle, remain open questions. Recent advances in single-cell RNA sequencing (scRNA-seq, see Chapter 4) now offer unprecedented resolution to dissect cell-type-specific transcriptional trajectories during the earliest stages of nodule initiation.

Besides the nodule-specific transcription factors (TFs) NIN, and NF-YA1, other TFs traditionally associated with root development, such as LBD16 and PLT1-4 of the PLETHORA family, also contribute to nodule formation (Franssen et al., 2015; Schiessl et al., 2019). Owing to their central regulatory roles in developmental processes, TFs represent compelling candidates for investigating the transcriptional control underlying pericycle activation and nodule initiation. Structurally, TFs typically contain a DNA-binding domain, an oligomerization site, a transcriptional regulation region, and a nuclear localization signal, enabling them to modulate plant development and respond to a variety of biotic and abiotic cues (Liu et al., 1999; Riechmann and Ratcliffe, 2000).

By cross-referencing our single-cell data set (**Chapter 4**) with previously available bulk RNA-seq data (Schiessl et al., 2019), we identified potential candidate TFs belonging to three major families: bZIP (basic leucine zipper), AP2/ERF (APETALA2/ETHYLENE-RESPONSIVE FACTOR), and GRAS (GAI [GIBBERELLIN-INSENSITIVE], RGA [REPRESSOR OF ga1-3], and SCR [SCARECROW]). These TF families exhibit extensive functional diversity, highlighting their

potential roles in mediating transcriptional reprogramming during pericycle activation and early nodule initiation.

In this study, we leverage the single-cell RNA sequencing dataset described in Chapter 4 to identify new pericycle marker genes and investigate their transcriptional dynamics during nodulation. Specifically, we focus on three representative TFs from distinct families: MtbZIP43 (BASIC LEUCINE ZIPPER 43; bZIP family; *MtrunA17_Chr6g0456061*), MtDREB3 (DEHYDRATION-RESPONSIVE ELEMENT-BINDING PROTEIN 3; AP2/ERF family; *MtrunA17_Chr3g0140161*), and MtSCL14 (SCARECROW-LIKE PROTEIN 14; GRAS family; *MtrunA17_Chr4g0031721*). These TFs are consistently expressed in the pericycle and display dynamic expression changes upon rhizobial inoculation. Although their loss-of-function phenotypes are subtle, these findings underscore the importance of refining candidate selection criteria for future studies aiming to uncover nodule-specific regulatory mechanisms.

RESULTS

Candidate gene selection

To identify genes involved in the initial pericycle cell divisions after nodule induction, we used the existing whole tissue RNA sequencing dataset by Schiessl et al. (2019) as a reference. By combining bulk differentially expressed gene information with our tissue-specific single-cell RNA sequencing (scRNA-seq) data, we identified ~600 potential candidate genes expressed in the pericycle and differentially regulated during nodulation (**Supplemental Figure S1A**).

Given that the first cell divisions in the pericycle initiate around 1 day post-inoculation (dpi), we refined our candidate list by focusing on early differential gene regulation at 1 dpi and 2 dpi in the scRNA-seq dataset. To identify regulators of early pericycle activation, we concentrated on three transcription factor (TF) families and manually selected one representative from each that exhibited noteworthy expression dynamics: *MtbZIP43* (bZIP family), *MtDREB3* (AP2/ERF family), and *MtSCL14* (GRAS family). All three TFs, *MtbZIP43*, *MtDREB3*, and *MtSCL14*, were expressed in the pericycle of the root tip and/or the susceptible zone (S-zone) (**Figure 1**). Importantly, none of these genes showed altered expression in response to the protoplasting procedure used for scRNA-seq (**Supplemental Figure S1B**). Instead, all selected genes were specifically upregulated in the pericycle or pericycle-associated nodule clusters following rhizobial inoculation (**Figure 2**).

Throughout the 4 day nodulation time course, *MtbZIP43* and *MtDREB3* maintained highly pericycle-specific expression patterns (**Figure 2A, D**). In contrast, *MtSCL14* displayed a broader expression domain, especially in the 2 mm root tip sample (**Figure 1G, 2G**). Notably, expression became progressively more restricted as root tissues matured: in the 10 mm tip sample, *MtSCL14* expression was already more focused on vascular tissues and the pericycle, a pattern that became even clearer in the S-zone (**Figure 1G**). In the mature S-zone sample, *MtSCL14* expression was largely confined to the vasculature (including the procambium and xylem), the pericycle, and the cortex, with additional activity in developing nodule or lateral root primordia (**Figure 1G**). These observations highlight a maturation-dependent modification of *MtSCL14* expression, which contrasts with the consistently vasculature and pericycle-specific profiles of *MtbZIP43* and *MtDREB3*.

MtbZIP43, *MtDREB3* and *MtSCL14* are expressed in the pericycle and vasculature of the root susceptible zone

To validate the expression specificity observed in our scRNA-seq dataset, we cloned the putative promoter and terminator regions of the selected genes. We used the 4.3 kb, 3.5 kb, and 2.2 kb upstream sequence of *MtbZIP43*, *MtDREB3* and *MtSCL14*, respectively.

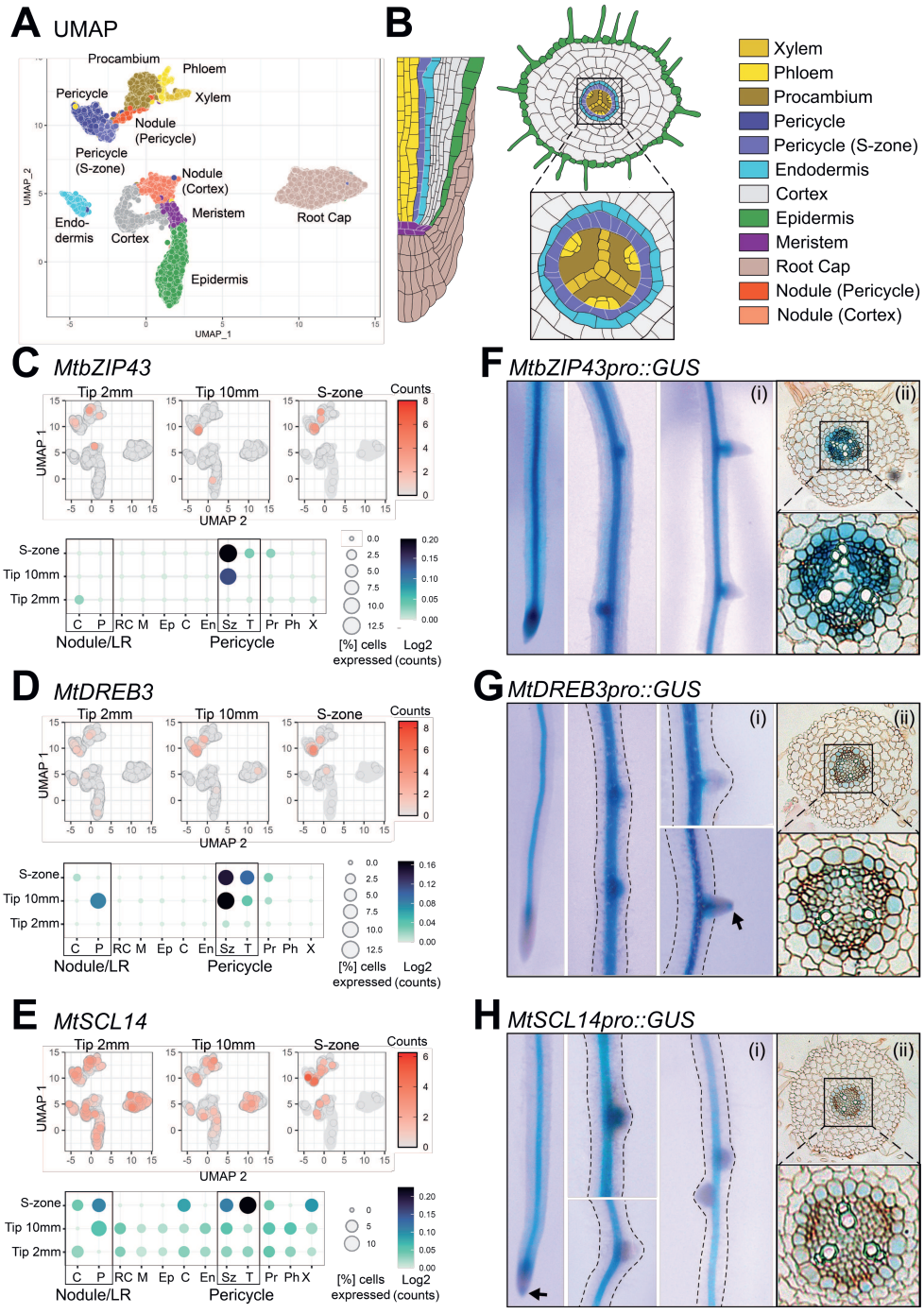


Figure 1 | Localization of *MtbZIP43*, *MtDREB3* and *MtSCL14*. **(A)** UMAP projection of the scRNA-seq dataset showing annotated cell-type clusters, color-coded according to **(B)** schematic representations of a Medicago root in longitudinal (left) and cross-sectional (right) views. (*continued on next page*)

Figure 1 (continued) | (C–E) scRNA-seq expression profiles of *MtbZIP43*, *MtDREB3*, and *MtSCL14*, respectively, shown for 2-mm and 10-mm root tip segments and the susceptible zone (S-zone; 10-mm region located 5 mm above the tip). Upper panels: UMAP expression patterns displayed as counts per transcript. Lower panels: dot-plot analyses across annotated cell types; color scale indicates log2-transformed counts, and dot size denotes the percentage of cells expressing the gene. **(F–H)** Promoter activity analyses of *MtbZIP43pro(4.3 kb)::GUS*, *MtDREB3pro(3.5 kb)::GUS*, and *MtSCL14pro(2.2 kb)::GUS*. (i) Whole-mount images of root tips, early lateral root primordia, and emerging lateral roots; outlines are traced for clarity and arrows indicate GUS staining in the root cap. (ii) Cross-sections of the S-zone, with close-up views of the vasculature. **Abbreviations:** C, cortex; P, pericycle; RC, root cap; M, meristem; Ep, epidermis; Pr, procambium; Ph, phloem; X, xylem; Sz, S-zone; T, tip; LR, lateral root.

To include the 3' UTR, 533, 500, and 500 bp downstream of the coding region we fused these fragments to the 3' end of the β-glucuronidase (GUS) reporter gene. For *in planta* validation, the constructs were introduced using the Medicago transient hairy root transformation system (Limpens et al., 2004).

From a macroscopic perspective, all three promoter–GUS reporter constructs exhibited activity in the inner root tissues, primarily within the vasculature and adjacent cell layers (**Figure 1F–H** (i)). GUS activity was also detected in early lateral root primordia and became prominent in the vasculature of emerging lateral roots. To more precisely define the activity domains of these promoters, we analyzed root cross-sections (**Figure 1F–H** (ii)). These confirmed that promoter activity was localized within the vasculature, including the pericycle, procambium, phloem, and xylem, with gene-specific nuances in distribution patterns. Consistent with our scRNA-seq data, all three *promoter*::GUS fusions displayed activity in the pericycle (**Figure 1C–H**).

Closer examination of the vascular tissues revealed distinct promoter activity profiles for each transcription factor. The putative *MtbZIP43* promoter showed strong activity in the procambium, with weaker but detectable activity in the phloem and xylem, consistent across both scRNA-seq and *promoter*::GUS analyses (**Figure 1C, F**). *MtDREB3* exhibited pronounced promoter activity in the xylem, which was not observed in our scRNA-seq dataset, as well as a weaker but consistent procambium-associated signal in more mature root regions (Tip 10 mm and S-zone) (**Figure 1D, G**). Occasional promoter activity was also observed in the columella root cap (**Figure 1G** (i), arrow). The *MtSCL14* promoter displayed overall weaker activity compared to the other TFs but showed a clear pericycle-associated signal consistent with the scRNA-seq data, along with additional activity in the root cap and a broader signal in the root tip region (**Figure 1E, H**). This broader tip-associated activity of *MtSCL14* corresponded with the observed expression pattern of this gene in our scRNA-seq dataset. Although a similar pattern was observed across all three *promoter*::GUS reporters, a broader tip-associated expression for *MtbZIP43* and *MtDREB3* was not reflected in our scRNA-seq data (**Figure 1C-H**). Taken together, our *promoter*–GUS analyses largely corroborate the cell-type specificity inferred from the scRNA-seq dataset, confirming pericycle-associated and vasculature expression for all three candidates.

Dynamic regulation of *MtbZIP43*, *MtSCL14* and *MtDREB3* during nodulation

Next, we examined the regulation of our candidate genes during nodulation. All three genes exhibited differential regulation in both whole-root RNA-seq data (Schiessl et al., 2019) and our scRNA-seq dataset (**Figure 2D–F; Supplemental Figures S3, S4**). In whole-tissue RNA-seq data, *MtbZIP43* and *MtSCL14* were downregulated, while *MtDREB3* showed consistent upregulation. Interestingly, our scRNA-seq analysis revealed both overlapping and tissue-specific regulatory patterns (**Figure 2D–F**).

In our scRNA-seq data, *MtDREB3* mirrored the expression observed in the whole-tissue data, displaying clear upregulation specifically within its expression domain (the pericycle and the developing, pericycle-derived nodule clusters (**Figure 2E**)). In the whole-tissue data, expression peaked at 2 days post-inoculation, showing a significant increase compared to the 0-day time point. Combined with our scRNA-seq results, this suggests a potential role in early pericycle activation. *MtbZIP43*, in contrast, was downregulated in the mature pericycle after 1 day, consistent with the trend observed in bulk RNA-seq (**Figure 2D**). Although slight activation was detected in other pericycle and nodule clusters, these changes remained low in magnitude.

Of these three genes, the most intriguing regulatory behavior was observed for *MtSCL14*. Despite an overall downregulation in both the whole-tissue dataset and the tip pericycle cluster of our scRNA-seq data, *MtSCL14* showed a distinct and highly specific upregulation in the pericycle of the susceptible zone at 1 and 2 days post-inoculation (**Figure 2F**), underscoring the importance of cell-type-resolved transcriptomic approaches for uncovering regulatory dynamics that are obscured in bulk tissue analyses.

We were unable to assess promoter activity in very early nodule primordia using the transient *promoter*::GUS reporter lines. However, we obtained nodules at later developmental stages, 4 and 7 dpi (**Figure 2G–I**). At 4 dpi, *MtbZIP43* and *MtDREB3* displayed strong GUS activity, reminiscent of their patterns in early lateral root primordia, whereas *MtSCL14* did not show detectable promoter activity (**Figure 2G–I**). By 7 dpi, all three constructs exhibited GUS activity in the nodule vasculature and meristematic regions, resembling their promoter activity domains in the root vasculature and tip (**Figure 1F–H, 2G–I**). Overall, these results indicate that while all three candidates respond transcriptionally to rhizobial inoculation, *MtDREB3* and *MtSCL14* show the clearest evidence for nodulation-associated, domain-specific regulation in the pericycle, supporting their prioritization for further functional analysis.

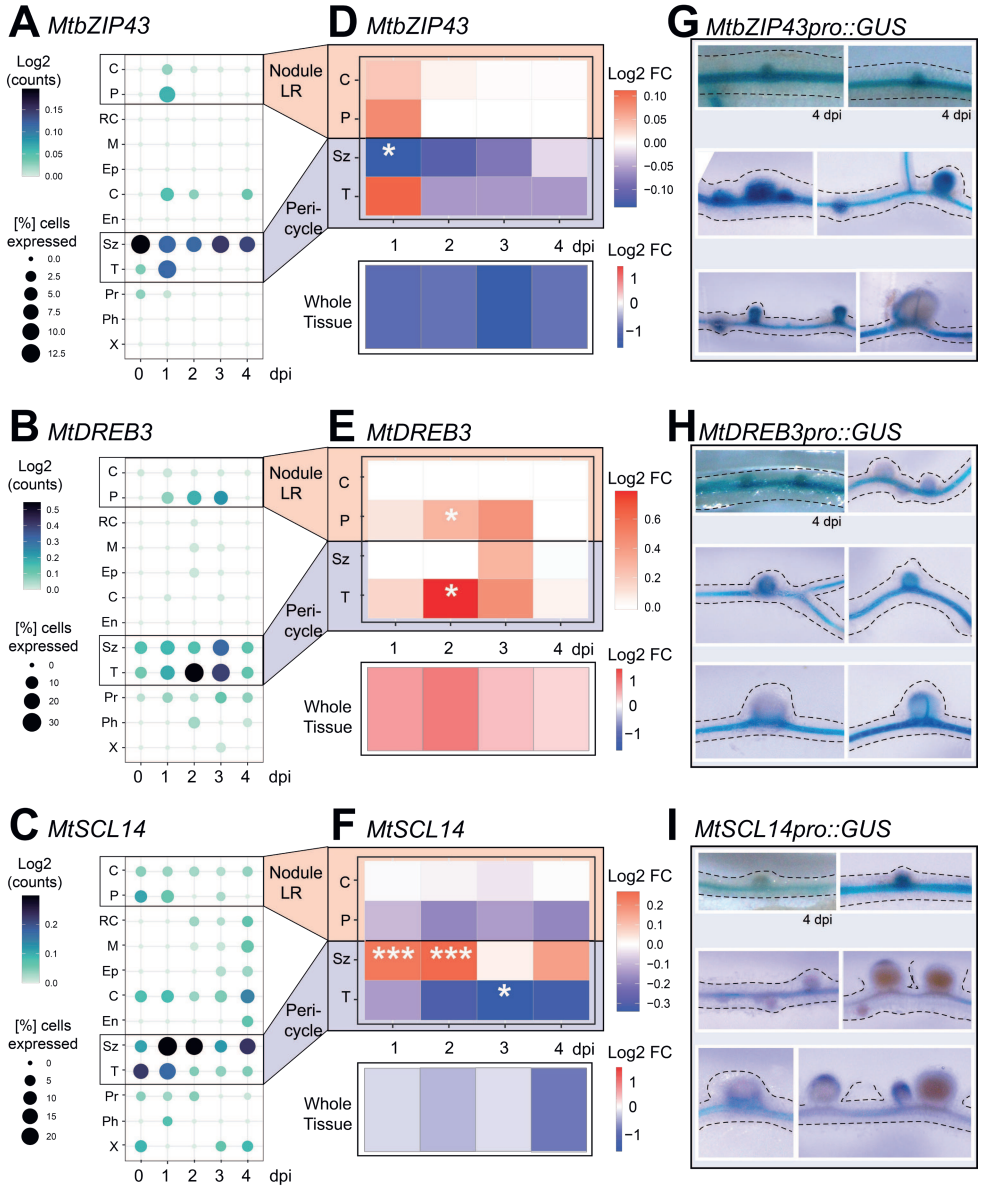


Figure 2 | Nodulation specific expression patterns of *MtbZIP43*, *MtDREB3*, and *MtSCL14*. Candidate gene expression following inoculation with *Sinorhizobium meliloti* 2011 for 0, 1, 2, 3, and 4 days post-inoculation (dpi) **(A–F)** or for 4 and 7 days post-inoculation (dpi) **(G–I)**. (A–C) Dot-plot analyses showing gene expression across cell types; color scale represents log2-transformed counts and dot size represents the percentage of expressing cells. **(D–F)** Heatmaps of pericycle (S-zone or tip) and nodule/lateral root (cortex or pericycle) cluster expression. Lower panels: whole-tissue RNA-seq data adapted from Schiessl et al. (2019), presented as log2 fold change relative to 0 hpi. **(G–I)** Promoter::GUS analyses of *MtbZIP43pro(4.3 kb)::GUS*, *MtDREB3pro(3.5 kb)::GUS*, and *MtSCL14pro(2.2 kb)::GUS*. Images show whole-mount tissues at representative nodule developmental stages (7 dpi unless indicated); root outlines are traced for clarity. **Abbreviations:** C, cortex; P, pericycle; RC, root cap; M, meristem; Ep, epidermis; Pr, procambium; Ph, phloem; X, xylem; Sz, S-zone; T, tip; LR, lateral root. Asterisks denote significant differences between treatments (*P ≤ 0.05; ***P ≤ 0.001).

Overexpression of *MtbZIP43*, *MtSCL14* or *MtDREB3* reduces nodule numbers

To investigate the functional roles of the three candidate genes during nodulation, we altered their expression using RNA interference (RNAi) and overexpression (ox) approaches. For RNAi, ~300 bp fragments targeting the endogenous mRNAs were used to reduce transcript abundance (**Supplemental Table S4**) (Limpens et al., 2004).

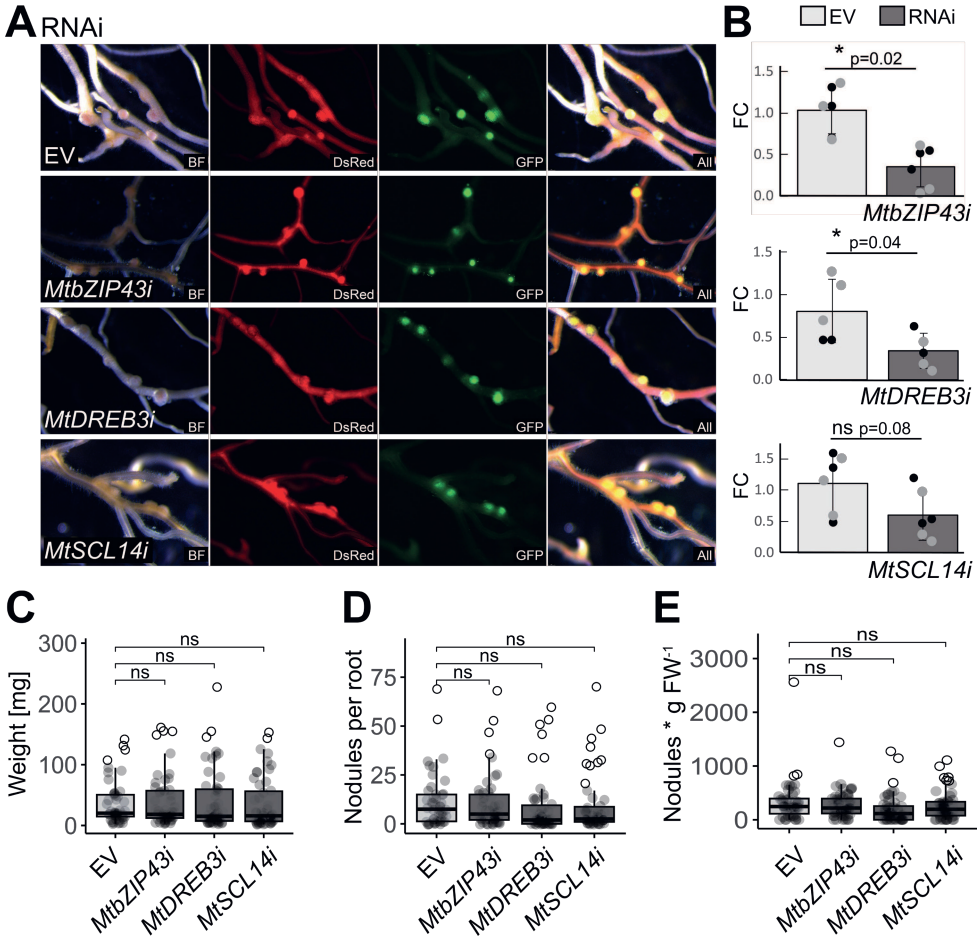


Figure 3 | RNA interference (RNAi) analysis of *MtbZIP43*, *MtDREB3*, and *MtSCL14*. Representative images of transformed roots shown from left to right: bright-field (BF), DsRed fluorescence marking transgenic roots, *Sinorhizobium meliloti* 2011::*GFP* signal (GFP), and merged channels (All). Rows show plants transformed with the empty vector (EV), *MtbZIP43*i, *MtDREB3*i, and *MtSCL14*i. **(B)** qRT-PCR validation of target gene knockdown in EV control (light gray) and corresponding RNAi lines (dark gray). Fold change (FC) relative to *MtUBQ*; bars represent mean ± SD (n ≥ 5), with colored points representing independent experiments. **(C–E)** Boxplots of single-root phenotypes (n ≥ 97): **(C)** root fresh weight (mg), **(D)** nodule number, and **(E)** nodules per gram fresh weight (FW). Asterisks denote significant differences between treatments (*P ≤ 0.05; **P ≤ 0.01).

For overexpression, the full-length coding sequences were placed under the control of a constitutive promoter and terminated with their native 3′ regions (**Supplemental Table S5**). Both constructs were introduced into Medicago via transient hairy root transformation (Limpens et al., 2004), and transgenic roots were selected based on red fluorescence. Nodulation was subsequently induced using *Sinorhizobium meliloti 2011* constitutively expressing GFP (*Sm2011-GFP*), allowing visualization of nodules in transformed roots (**Figures 3, 4; Supplemental Figure S6**).

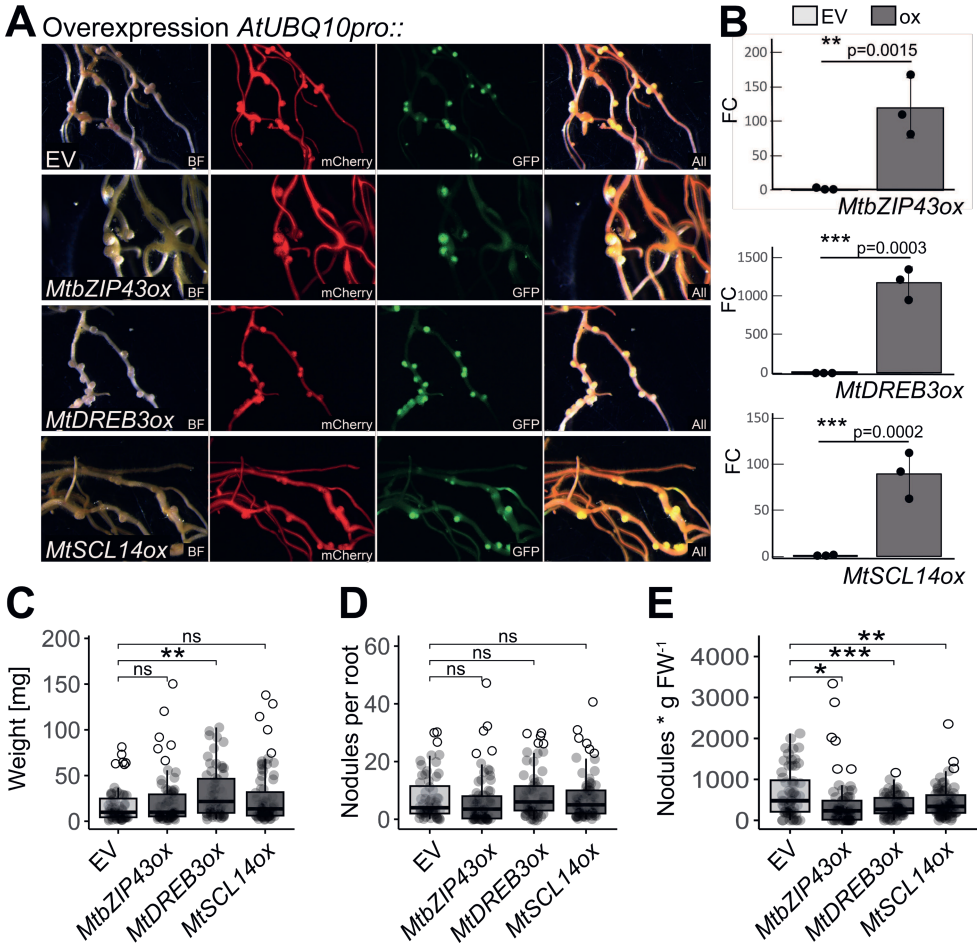


Figure 4 | Overexpression (ox) analysis of *MtbZIP43*, *MtDREB3*, and *MtSCL14*. (A) Representative images of transformed roots shown from left to right: bright-field (BF), DsRed fluorescence marking transgenic roots, *Sinorhizobium meliloti* 2011::GFP signal (GFP), and merged channels (All). Rows show plants transformed with the empty vector (EV), *MtbZIP43*ox, *MtDREB3*ox, and *MtSCL14*ox. **(B)** qRT-PCR validation of target gene overexpression in EV (light gray) and corresponding ox lines (dark gray). Fold change (FC) relative to *MtUBQ*; bars represent mean ± SD (n = 3). **(C–E)** Boxplots of single-root phenotypes (n ≥ 60): **(C)** root fresh weight (mg), **(D)** nodule number, and **(E)** nodules per gram fresh weight (FW). Asterisks denote significant differences between treatments (*P ≤ 0.05; **P ≤ 0.01; ***P ≤ 0.001).

In the RNAi experiments, *MtbZIP43* and *MtDREB3* transcript levels were reduced by approximately 70% and 60%, respectively. Additionally, *MtSCL14* RNAi achieved an average reduction of only 46%, which was not statistically significant. Nevertheless, no significant effects on nodulation were detected for any of these constructs (**Figure 3B–E**).

Overexpression led to robust transcriptional activation of all three genes, with *MtbZIP43ox* and *MtSCL14ox* lines showing ~80-fold increases, and *MtDREB3ox* reaching up to 800-fold induction (**Figure 4B**). Whereas RNAi-mediated silencing did not significantly alter nodulation phenotypes (**Figure 3**), overexpression of all three candidate genes resulted in a significant reduction in the number of nodules per gram fresh weight (g FW^{-1}) across all constructs (**Figure 4E**). Additionally, *MtDREB3ox* roots displayed a pronounced increase in total root biomass, emphasizing the necessity of a well thought out normalization strategy for scoring nodule density to accurately compare genotypic effects (**Figure 4C**).

Apart from the quantitative reduction in nodule number, no obvious morphological abnormalities were observed in nodules formed on either RNAi or overexpression roots (**Figure 3A, 4A**). The nodules appeared macroscopically similar to control nodules, although detailed histological analyses were not conducted (**Figure 3A, 4A; Supplemental Figure S6**). Non-transgenic roots did not display significant differences across measured parameters compared to the empty vector control, confirming that all plants were grown and treated under comparable conditions (**Supplemental Figure S5**).

It should be noted that the *MtSCL14* overexpression construct contained a single amino acid substitution (E49G), introduced to remove an internal BpiI restriction site during cloning (**Supplemental Figure S7; Table S2**). Although this residue is conserved (**Supplemental Figure S7D**), structural modeling indicated that the overall protein fold is maintained (**Supplemental Figure S7B–C**) and that no known functional domains are disrupted (**Supplemental Figure S7A**). Nonetheless, this substitution may potentially influence protein–protein interactions, and the *MtSCL14ox* phenotype should therefore be interpreted with appropriate caution.

Although none of these transcription factors were knocked down by more than 60%, collectively, these functional analyses suggest that none of the three candidates is individually essential for nodule formation, but that their misexpression is sufficient to perturb nodulation efficiency, consistent with roles as modulatory regulators whose proper dosage and spatial restriction are critical for balanced root growth and possibly nodule organogenesis.

DISCUSSION

In this study, we used the single-cell RNA sequencing (scRNA-seq) dataset generated in Chapter 4 to identify transcription factors expressed in the pericycle and assess their roles during nodule initiation in Medicago. By integrating scRNA-seq with existing whole-root RNA-seq data (Schiessl et al., 2019), we identified ~600 candidate genes that were both expressed in the pericycle and differentially regulated during the first 2 days after rhizobial inoculation. Because pericycle cell divisions initiate around 1 day post-inoculation (dpi), we focused specifically on genes up- or downregulated at 1 and 2 dpi. From three major transcription factor (TF) families known for developmental regulation (bZIP, AP2/ERF, and GRAS), we decided to focus one representative per family (i.e *MtbZIP43*, *MtDREB3*, and *MtSCL14*) showing the before mentioned expression pattern. All three genes showed pericycle-associated expression in either the root tip or the susceptible zone (S-zone), were unaffected by the protoplasting process, and displayed distinct transcriptional responses to rhizobial inoculation.

MtbZIP43, *MtSCL14* and *MtDREB3* are expressed in the pericycle and vasculature of the root susceptible zone

In line with our scRNA-seq dataset, we validated a *promoter*::GUS signal of all our three TFs in the pericycle and vasculature of the susceptible zone. Additionally, we can distinguish unique vascular expression patterns for each TF. *MtbZIP43* marks the whole vasculature, including a strong signal from the procambium and weaker signals from the xylem and phloem. *MtDREB3* only marks the procambium and the xylem. For *MtSCL14* it remains challenging to clarify the vascular expression, because of the weak GUS signal even with longer incubation times in the GUS reaction buffer. Therefore, we can only validate the expression of *MtSCL14* in the pericycle.

Comparisons of the expression domains found for the three selected genes with publicly available single-nucleus RNA-seq datasets (Cervantes-Pérez et al., 2022; Pereira et al., 2024) revealed only a partial overlap in pericycle and vascular expression. Discrepancies between datasets are possibly due to differences in tissue origin, cell types captured, sequencing depth, and growth or treatment conditions. For instance, the strong epidermal expression of MtDREB3 reported by Pereira et al. (2024) is not observed in our dataset, which may in part reflect the low representation or absence of epidermal cells in our mature S-zone samples. However, the lack of epidermal activity in our *promoter*::GUS analyses suggests that this previously detected epidermal signal may represent a dataset-specific artefact rather than a robust biological feature.

Additionally, analysis of the Arabidopsis root single-cell atlas (rootcellatlas.org) shows that the orthologs of our candidate genes exhibit partially similar cell-type-specific expression patterns, further supporting the biological relevance of our findings. *AtbZIP42*, the closest ortholog of *MtbZIP43*, shows a divergent expression pattern restricted to the meristematic zone, whereas *AtbZIP43*, which is less similar at the sequence level, is expressed in the procambium and more closely resembles the vascular expression of *MtbZIP43* in Medicago. *AtDREB3/TINY2*, the

closest ortholog of *MtDREB3*, is expressed in the vasculature, particularly in the procambium and xylem, in agreement with the Medicago scRNA-seq data. *AtSCL14* is broadly expressed throughout the root, similar to *MtSCL14*, with expression becoming more restricted to the pericycle and xylem in more mature tissues. Altogether, the partial overlap in cell-type-specific expression between Medicago genes and their Arabidopsis orthologs supports the idea that overall, our scRNA-seq dataset closely matches the expression patterns of the selected candidate genes. Additionally, it also suggests that these transcription factors may share evolutionary conserved functions in root and vascular development.

Overexpression of candidate TFs suggests their possible role during nodulation

Based on the expression patterns observed during nodule induction, we initially expected that RNAi and overexpression of the three transcription factors would produce contrasting phenotypes. MtbZIP43 is consistently downregulated after inoculation, suggesting that its repression, or repression of downstream targets, may be required for the onset of nodule organogenesis, similar to the pattern described for MtACS10 in Chapter 3. Under this model, RNAi of MtbZIP43 would be predicted to increase nodule numbers, whereas overexpression should reduce them. Although RNAi did not alter nodulation, the reduction in nodules in the overexpression lines is consistent with this hypothesis.

In contrast, both *MtDREB3* and *MtSCL14* are upregulated upon inoculation, with *MtSCL14* showing a particularly strong and specific induction in the pericycle. This would suggest a positive role during early nodule initiation, in which case RNAi would be expected to decrease, and overexpression increase, nodule numbers. However, RNAi again produced no measurable phenotype, and overexpression unexpectedly reduced nodule numbers for both genes.

Together, these findings indicate that RNAi silencing was too weak and/or too variable to reveal reliable phenotypes. Moreover, it should be noted that the transient hairy root system used in this study also presents limitations for detecting subtle phenotypes. RNAi rarely achieves complete transcript depletion, and even CRISPR-Cas9 typically yields mosaic or monoallelic edits in this system (H. Zhang et al., 2020). indicating that strong knockout phenotypes are difficult to obtain. Although stable transformation would circumvent these issues, it remains technically challenging, if not currently unattainable, in the used Medicago model genotype Jemalong A17.

Additionally, functional redundancy may further obscure knockdown effects. This is especially likely for bZIP transcription factors, where redundancy within subgroups is well documented (Dietrich et al., 2011; Pedrotti et al., 2018). Eliminating redundancy sometimes requires knocking out entire subfamilies, as shown by Pedrotti et al. (2018) for the bZIP S1 subgroup. More broadly, redundancy and stress-dependent functions are common features of many TF families (Riechmann and Meyerowitz, 1998). As such, paralogs of *MtbZIP43* or other S-group bZIPs could compensate for reduced *MtbZIP43* expression levels. Indeed, when analyzing the

expression profile of the *MtbZIP43* closest paralog *MtbZIP43_65* in our scRNA-seq data, we discovered that it displayed a similar expression pattern (**Supplemental Figure S8**), suggesting that these TFs could act redundantly. If so, *MtbZIP43_65* could mask any putative nodulation phenotype of the RNAi line targeting *MtbZIP43*.

Moreover, phenotypes can be conditionally dependent on additional stresses such as drought or temperature. This has previously been demonstrated to be the case for ectopic expression of *AtCBF1* (AP2/ERF family) and *OsGRAS29* (GRAS family). Both produced no obvious phenotype under control conditions, but conferred enhanced freezing or drought tolerance, respectively (Stockinger et al., 1997; Xu et al., 2015). We cannot exclude that under specific conditions nodulation might be effected in any of our RNA lines.

While overexpression produced clear phenotypic effects, the fact that all three genes reduced nodule number in a similar manner suggests that these phenotypes may reflect a global mis-regulation rather than gene-specific functions. The use of the constitutive *AtUBQ* promoter likely disrupts the spatial and temporal regulation required for coordinated nodule development. Alternatively, considering that in the root susceptible zone all selected TFs are specifically expressed in the pericycle and vasculature, they might be involved in pericycle and/or vasculature specific functions or identity. Ubiquitous expression of any of these TFs could in theory disrupt the true identities of the ectopic expressing transgenic tissues. As a result, such tissues might be hindered in proper nodule initiation. Employing a pericycle- or nodule-specific promoter, such as *MtENOD11pro*, *MtENOD40pro* (**Chapter 2**), or *MtCRE1pro* (**Chapter 3**), could provide a more targeted approach and help determine whether localized overexpression affects pericycle activation in a more physiologically meaningful way.

Phenotypes associated with *MtSCL14* require additional caution in interpretation. Although RNAi trended toward ~50% knockdown, the high variability between samples rendered this reduction statistically insignificant. Moreover, the *MtSCL14* overexpression construct carries an unavoidable E49G substitution introduced to enable Golden Gate cloning (Engler et al., 2014; Weber et al., 2011). Although structural modeling suggests that this mutation does not disrupt protein folding or known functional domains, subtle alterations in protein–protein interactions cannot be excluded. Nevertheless, because the phenotype of *MtSCL14ox* matches that of the other two overexpression lines, global mis-regulation caused by ubiquitous expression remains the most likely explanation. Still, impaired function of the MtSCL14^E49G protein cannot be entirely ruled out.

scRNA-seq: High resolution data with important caveats

Taken together, this work demonstrates the value of scRNA-seq for uncovering cell-type-specific transcriptional dynamics, such as the pericycle induction of *MtSCL14* in the susceptible zone, which is completely masked in whole-tissue datasets. However, scRNA-seq also presents challenges for candidate identification. The technique is characterized by transcriptional sparsity, often capturing <20% of the transcriptome per cell (Ryu et al., 2025). Even in high-quality protoplast datasets, where up to ~11,000 genes per cell may be detected (Bakken et al., 2018; Grones et al., 2024), low-abundance transcripts remain difficult to recover. Detection sensitivity is also affected by sequencing depth, filtering strategies, and the choice between nuclei and whole-cell preparations (Grones et al., 2024; Zhang et al., 2020b).

These limitations emphasize the need for stringent candidate selection criteria. In our case, refining the analysis to genes activated specifically in the pericycle in response to rhizobia significantly improved the relevancy of chosen markers and will likely be essential in future efforts to pinpoint nodule-specific regulators. At the same time, the ease with which scRNA-seq enabled identification of robust pericycle and vascular markers highlights the overall high quality and strong resolution of the dataset and underscores its potential for further discovery.

MATERIALS AND METHODS

Plant material and growth conditions

Medicago truncatula Jemalong A17 wild-type seeds were prepared one week prior to experimentation. Seeds were scarified in concentrated H_2SO_4 for 7 min, rinsed seven times with deionized water, and surface-sterilized in commercial bleach (>2.5% active chlorine) for 7 min, followed by seven washes with sterile Milli-Q water. After imbibition in sterile water for 1–3 h, seeds were transferred to modified Fåhraeus medium plates (Fahraeus, 1957), 0.125 mM $Ca(NO_3)_2$, 0.9% [w/v] Daishin agar, 1 g L^{-1} carbendazim) and stratified overnight at 4 °C. Plates were then inverted and incubated in the dark at room temperature for 20–24 h to allow germination. Seed coats were removed to ensure uniform growth, and groups of five seedlings were placed on filter paper overlays positioned on Fåhraeus medium (without carbendazim; 0.125 mM $Ca(NO_3)_2$) in round plates. Plates were sealed with Leukopor™ tape, leaving the top edge open for gas exchange. Seedlings were grown vertically for 5 days under a 16-h light/8-h dark photoperiod at 22 °C before transformation.

Rhizobia growth conditions and application

Sinorhizobium meliloti 2011 (Sm2011; also known as *Ensifer meliloti* GFP-reporter strain), was used to induce nodule formation (Limpens et al., 2003). Cultures were grown in YEM medium to mid–log phase (OD_{600} = 0.6–0.9), then diluted to OD_{600} = 0.02 for spot inoculation (~2 µL droplets applied to the susceptible zone on plates), or to OD_{600} = 0.1 for mass inoculation of plants grown in perlite (2 mL per plant), using Fåhraeus medium supplemented with 0.25 mM NO_3^-.

Vectors and constructs

Promoter::GUS and overexpression constructs were generated using the Golden Gate cloning system (Engler et al., 2014). Putative native promoter regions (2.2–4.3 kb upstream of the start codon), full-length coding sequences, and native 3′ terminators (500 bp downstream of the stop codon) were amplified from genomic DNA using Phusion™ High-Fidelity polymerase (Thermo Fisher Scientific). Internal BpiI sites were removed by site-directed mutagenesis using overlapping primers (Supplemental Tables S3, S6). Vector components were obtained from the Golden Gate cloning toolbox (Engler et al., 2014). or provided through the ENSA program (https://www.ensa.ac.uk/) (Supplemental Table S5). A Kanamycin resistance cassette combined with a red fluorescent marker served as the empty vector (EV) control; *promoter*::GUS units or *AtUBQ10pro*::coding sequence cassettes were inserted at position 4 to generate the constructs *MtbZIP43pro::GUS*, *MtDREB3pro::GUS*, *MtSCL14pro::GUS*, *MtbZIP43ox*, *MtDREB3ox*, and *MtSCL14ox*.

For RNA interference (RNAi), ~300 bp gene fragments (Supplemental Table S4) were amplified from genomic DNA using primers listed in Supplemental Table S3. As the target genes lack introns, cDNA templates were not required. Fragments were first cloned into pENTR-D-TOPO

(Invitrogen) through CACC overhangs and subsequently recombined via LR clonase into a DsRed-modified pK7GWIWG2(II)-RR-CaMV35Spro RNAi vector (Limpens et al., 2004), generating *MtbZIP43i*, *MtDREB3i*, and *MtSCL14i*. The EV control was pK7GWIWG2(II)-RR-p35S-RNAi-control (van Zeijl et al., 2015).

Genomic DNA extraction

Genomic DNA was isolated from Medicago A17 roots using SDS/Tris/NaCl/EDTA extraction (Edwards et al., 1991) with an added potassium acetate clarification step (Dellaporta et al., 1983). Approximately 30–100 mg fresh tissue was frozen in liquid nitrogen and ground using a tissue lyser. Extraction buffer (400 µL; 250 mM Tris-HCl pH 7.5, 250 mM NaCl, 25 mM EDTA, 0.5% SDS) was added, followed by 150 µL 3 M potassium acetate (pH 5). After centrifugation (12,000 rpm, 5 min), the supernatant was mixed with 550 µL ice-cold 100% isopropanol. DNA pellets were washed with 70% ethanol, air-dried, and resuspended in 50–100 µL TE buffer.

Plant transformation, histology and phenotyping

Transient hairy root transformation with *Agrobacterium rhizogenes* was performed as described (Limpens et al., 2004). Transgenic roots expressing GUS constructs were selected using mCherry fluorescence and transferred to square plates or perlite containing Fåhraeus medium (0.125 mM $Ca(NO_3)_2$). Plates were opened at the top to allow upright growth and gas exchange. Plants were grown for 14 days before spot (OD_{600} = 0.02) or mass inoculation (OD_{600} = 0.1) with Sm2011::GFP. Plate-grown plants were harvested at 1, 2, 4, and 7 dpi; perlite-grown plants at 7 dpi. Tissue was fixed in 90% acetone at −20 °C (<2 weeks), then incubated in GUS reaction buffer for 1 h (MtbZIP43pro::GUS, MtDREB3pro::GUS) or 1.5 h (MtSCL14pro::GUS) (Dedow et al., 2022). For sectioning, samples were fixed in 4% paraformaldehyde and embedded in Technovit® 7100.

RNAi and overexpression roots were selected by DsRed or mCherry fluorescence and transferred to perlite. After two weeks, roots were mass inoculated (OD_{600} = 0.1; 2 mL/plant). Individual transgenic and selected non-transgenic roots were harvested at 14–20 dpi, weighed, and nodules counted. For qRT-PCR, multiple transgenic roots were pooled (three biological replicates per construct).

RNA isolation, cDNA synthesis and quantitative RT-PCR

Transgenic roots were frozen in liquid nitrogen and stored at −80 °C. Total RNA was extracted using the E.Z.N.A. Plant RNA Kit (Omega Bio-tek) with on-column DNase I digestion. cDNA was synthesized from 0.5–1 µg RNA using the iScript cDNA Synthesis Kit (Bio-Rad). qRT-PCR reactions (10 µL) were prepared using iQ SYBR Green Supermix (Bio-Rad) and run on a CFX Connect cycler. Primer sequences, including the reference gene *MtUBQ*, are listed in Supplemental Table S3. Statistical significance of ΔΔCq values was assessed using a Student's *t*-test; fold changes were calculated as FC = 2^(−ΔΔCq).

Data analysis of scRNA-seq dataset

scRNA-seq data were processed and filtered as described in Chapter 4. Differentially expressed genes identified from whole-root RNA-seq (Schiessl et al., 2019) were screened in the scRNA-seq dataset to define pericycle-enriched genes (Supplemental Figure S1). From approximately 600 candidates, three transcription factor genes were selected based on cell-type specificity. Analyses were performed in RStudio (v4.4.1) using ggplot2 (v3.5.2), ggpubr (v0.6.0), dplyr (v1.1.4), rlang (v1.1.4), and patchwork (v1.3.1).

Protein data analyses

Paralog identification used BLASTp (NCBI) against the Medicago proteome (taxid:3880) with standard parameters (expect threshold 0.05, BLOSUM62). Candidate homologs (Supplemental Table S1) were selected based on ≥70% query coverage (≥50% for bZIP43). Alignments were generated with Clustal Omega (EMBL-EBI; (Madeira et al., 2024)). and visualized using Jalview (Waterhouse et al., 2009). Phylogenetic trees were produced in Jalview (Neighbour-Joining, BLOSUM62). Motif and domain annotations were obtained from UniProt. Protein structures (MtSCL14 and MtSCL14^E49G) were modeled with AlphaFold (Fleming et al., 2025; Jumper et al., 2021) and CoLab AlphaFold, and visualized in PyMol (v2.3.4). Cross-species homologs were identified with BLASTp and included in multiple sequence alignments and phylogenies using Clustal Omega (Madeira et al., 2024).

Statistical analysis and data visualization

Statistical analyses were performed using Student's *t*-tests in RStudio (v4.4.1) or Microsoft Excel. Significance thresholds were $P \leq 0.05$ (*), $P \leq 0.01$ (**), and $P \leq 0.001$ (***). Graphical visualizations were generated using ggplot2 for scRNA-seq and phenotyping data, and Microsoft Excel for qRT-PCR results.

REFERENCES

Bakken, T. E., Hodge, R. D., Miller, J. A., Yao, Z., Nguyen, T. N., Aevermann, B., Barkan, E., Bertagnolli, D., Casper, T., Dee, N., et al. (2018). Single-nucleus and single-cell transcriptomes compared in matched cortical cell types. *PLoS One* 13,.

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.

Dedow, L. K., Oren, E. and Braybrook, S. A. (2022). Fake news blues: A GUS staining protocol to reduce false-negative data. *Plant Direct* 6, e367.

Dellaporta, S. L., Wood, J. and Hicks, J. B. (1983). A plant DNA minipreparation: Version II. *Plant Mol Biol Report* 1, 19–21.

Dietrich, K., Weltmeier, F., Ehlert, A., Weiste, C., Stahl, M., Harter, K. and Dröge-Lasera, W. (2011). Heterodimers of the Arabidopsis Transcription Factors bZIP1 and bZIP53 Reprogram Amino Acid Metabolism during Low Energy Stress. *Plant Cell* 23, 381–395.

Edwards, K., Johnstone, C. and Thompson, C. (1991). A simple and rapid method for the preparation of plant genomic DNA for PCR analysis. *Nucleic Acids Res* 19, 1349.

Engler, C., Youles, M., Gruetzner, R., Ehnert, T. M., Werner, S., Jones, J. D. G., Patron, N. J. and Marillonnet, S. (2014). A Golden Gate modular cloning toolbox for plants. *ACS Synth Biol* 3, 839–843.

Fahraeus, G. (1957). The Infection of Clover Root Hairs by Nodule Bacteria Studied by a Simple Glass Slide Technique. *Microbiology (N Y)* 16, 374–381.

Fleming, J., Magana, P., Nair, S., Tsenkov, M., Bertoni, D., Pidruchna, I., Lima Afonso, M. Q., Midlik, A., Paramval, U., Žídek, A., et al. (2025). AlphaFold Protein Structure Database and 3D-Beacons: New Data and Capabilities. *J Mol Biol* 437, 168967.

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 (Cambridge)* 142, 2941–2950.

Grones, C., Eekhout, T., Shi, D., Neumann, M., Berg, L. S., Ke, Y., Shahan, R., Cox, K. L., Gomez-Cano, F., Nelissen, H., et al. (2024). Best practices for the execution, analysis, and data storage of plant single-cell/nucleus transcriptomics. *Plant Cell* 36, 812–828.

Jhu, M. Y. and Oldroyd, G. E. D. (2023). Dancing to a different tune, can we switch from chemical to biological nitrogen fixation for sustainable food security? *PLoS Biol* 21, 1–24.

Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Žídek, A., Potapenko, A., et al. (2021). Highly accurate protein structure prediction with AlphaFold. *Nature* 596, 583–589.

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 (1979)* 302, 630–633.

Limpens, E., Ramos, J., Franken, C., Raz, V., Compaan, B., Franssen, H., Bisseling, T. and Geurts, R. (2004). RNA interference in Agrobacterium rhizogenes-transformed roots of Arabidopsis and Medicago truncatula. *J Exp Bot* 55, 983–992.

Liu, L., White, M. J. and MacRae, T. H. (1999). Transcription factors and their genes in higher plants: Functional domains, evolution and regulation. *Eur J Biochem* 262, 247–257.

Madeira, F., Madhusoodanan, N., Lee, J., Eusebi, A., Niewielska, A., Tivey, A. R. N., Lopez, R. and Butcher, S. (2024). The EMBL-EBI Job Dispatcher sequence analysis tools framework in 2024. *Nucleic Acids Res* 52, W521–W525.

Pedrotti, L., Weiste, C., Nägele, T., Wolf, E., Lorenzin, F., Dietrich, K., Mair, A., Weckwerth, W., Teige, M., Baena-González, E., et al. (2018). Snf1-RELATED KINASE1-controlled C/S1-bZIP signaling activates alternative mitochondrial metabolic pathways to ensure plant survival in extended darkness. *Plant Cell* 30, 495–509.

Peoples, M. B., Herridge, D. F. and Ladha, J. K. (1995). Biological nitrogen fixation: An efficient source of nitrogen for sustainable agricultural production? *Plant Soil* 174, 3–28.

Pereira, W. J., Boyd, J., Conde, D., Triozzi, P. M., Balmant, K. M., Dervinis, C., Schmidt, H. W., Boaventura-Novaes, C., Chakraborty, S., Knaack, S. A., et al. (2024). The single-cell transcriptome program of nodule development cellular lineages in Medicago truncatula. *Cell Rep* 43,.

Riechmann, J. L. and Meyerowitz, E. M. (1998). *The AP2/EREBP Family of Plant Transcription Factors*.

Riechmann, J. L. and Ratcliffe, O. J. (2000). A genomic perspective on plant transcription factors. *Curr Opin Plant Biol* 3, 423–434.

Roy, S., Liu, W., Nandety, R. S., Crook, A., Mysore, K. S., Pislariu, C. I., Frugoli, J., Dickstein, R. and Udvardi, M. K. (2020). Celebrating 20 Years of Genetic Discoveries in Legume Nodulation and Symbiotic Nitrogen Fixation. *Plant Cell* 32, 15–41.

Ryu, K. H., Zhu, Y. and Schiefelbein, J. (2025). Plant Cell Identity in the Era of Single-Cell Transcriptomics. *Annual Reviews* 22, 17.

Schauser, L., Roussis, A., Stiller, J. and Stougaard, J. (1999). A plant regulator controlling development of symbiotic root nodules. *Nature* 402, 191–195.

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). NODULE INCEPTION Recruits the Lateral Root Developmental Program for Symbiotic Nodule Organogenesis in Medicago truncatula. *Current Biology* 29, 3657-3668.e5.

Soyano, T., Kouchi, H., Hirota, A. and Hayashi, M. (2013). NODULE INCEPTION Directly Targets NF-Y Subunit Genes to Regulate Essential Processes of Root Nodule Development in Lotus japonicus. *PLoS Genet* 9, e1003352.

Soyano, T., Shimoda, Y., Kawaguchi, M. and Hayashi, M. (2019). A shared gene drives lateral root development and root nodule symbiosis pathways in Lotus. *Science (1979)* 366, 1021–1023.

Stockinger, E. J., Gilmour, S. J. and Thomashow, M. F. (1997). Arabidopsis thaliana CBF1 encodes an AP2 domain-containing transcriptional activator that binds to the C-repeat/DRE, a cis-acting DNA regulatory element that stimulates transcription in response to low temperature and water deficit. *Proc Natl Acad Sci U S A* 94, 1035–1040.

van Zeijl, A., Liu, W., Xiao, T. T., Kohlen, W., Yang, W. C., Bisseling, T. and Geurts, R. (2015). The strigolactone biosynthesis gene DWARF27 is co-opted in rhizobium symbiosis. *BMC Plant Biol* 15, 1–15.

Waterhouse, A. M., Procter, J. B., Martin, D. M. A., Clamp, M. and Barton, G. J. (2009). Jalview Version 2—a multiple sequence alignment editor and analysis workbench. *Bioinformatics* 25, 1189–1191.

Weber, E., Engler, C., Gruetzner, R., Werner, S. and Marillonnet, S. (2011). A modular cloning system for standardized assembly of multigene constructs. *PLoS One* 6,.

Xiao, T. T., Schilderink, 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 (Cambridge)* 141, 3517–3528.

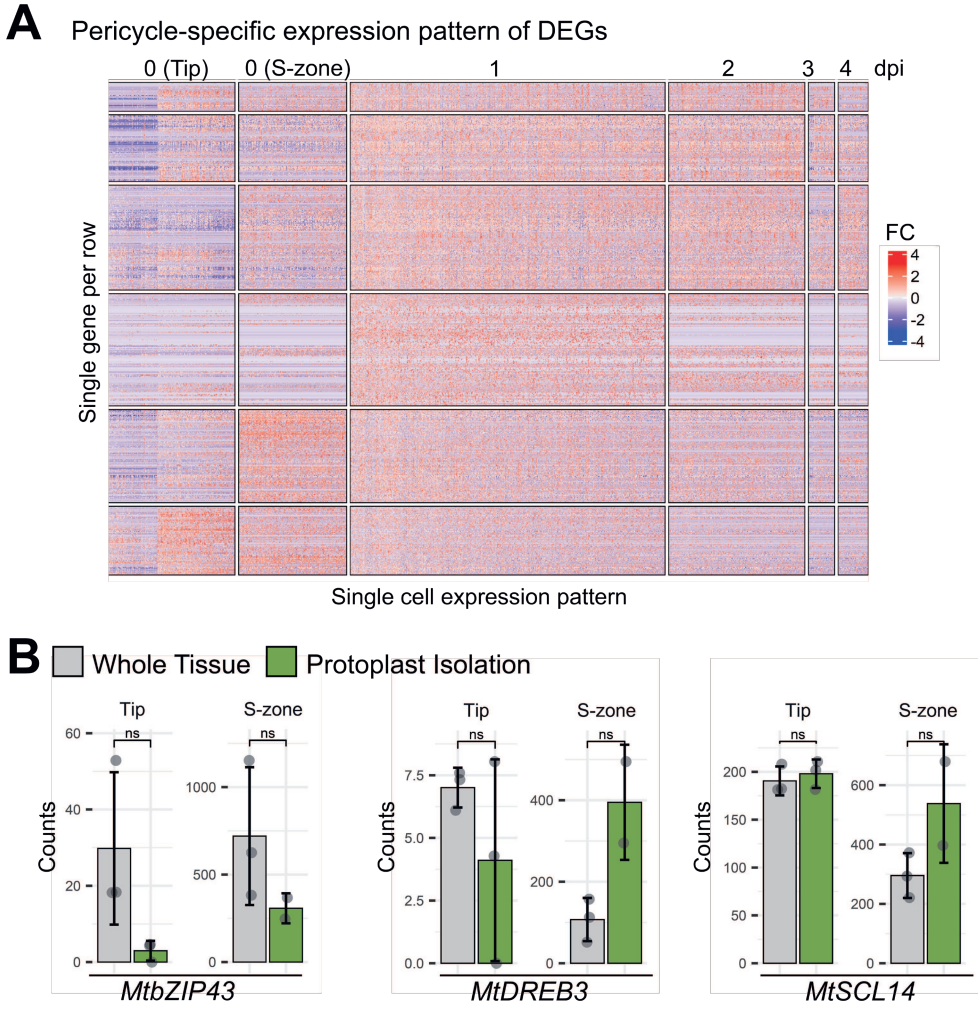
Xiao, T. T., Müller, S., Shen, D., Liu, J., Adema, K., van Seters, A., Franssen, H., Bisseling, T., Kulikova, O. and Kohlen, W. (2025). Nodule organogenesis in Medicago truncatula requires local stage-specific auxin biosynthesis and transport. *Plant Physiol* 197,.

Xu, K., Chen, S., Li, T., Ma, X., Liang, X., Ding, X., Liu, H. and Luo, L. (2015). OsGRAS23, a rice GRAS transcription factor gene, is involved in drought stress response through regulating expression of stress-responsive genes. *BMC Plant Biol* 15,.

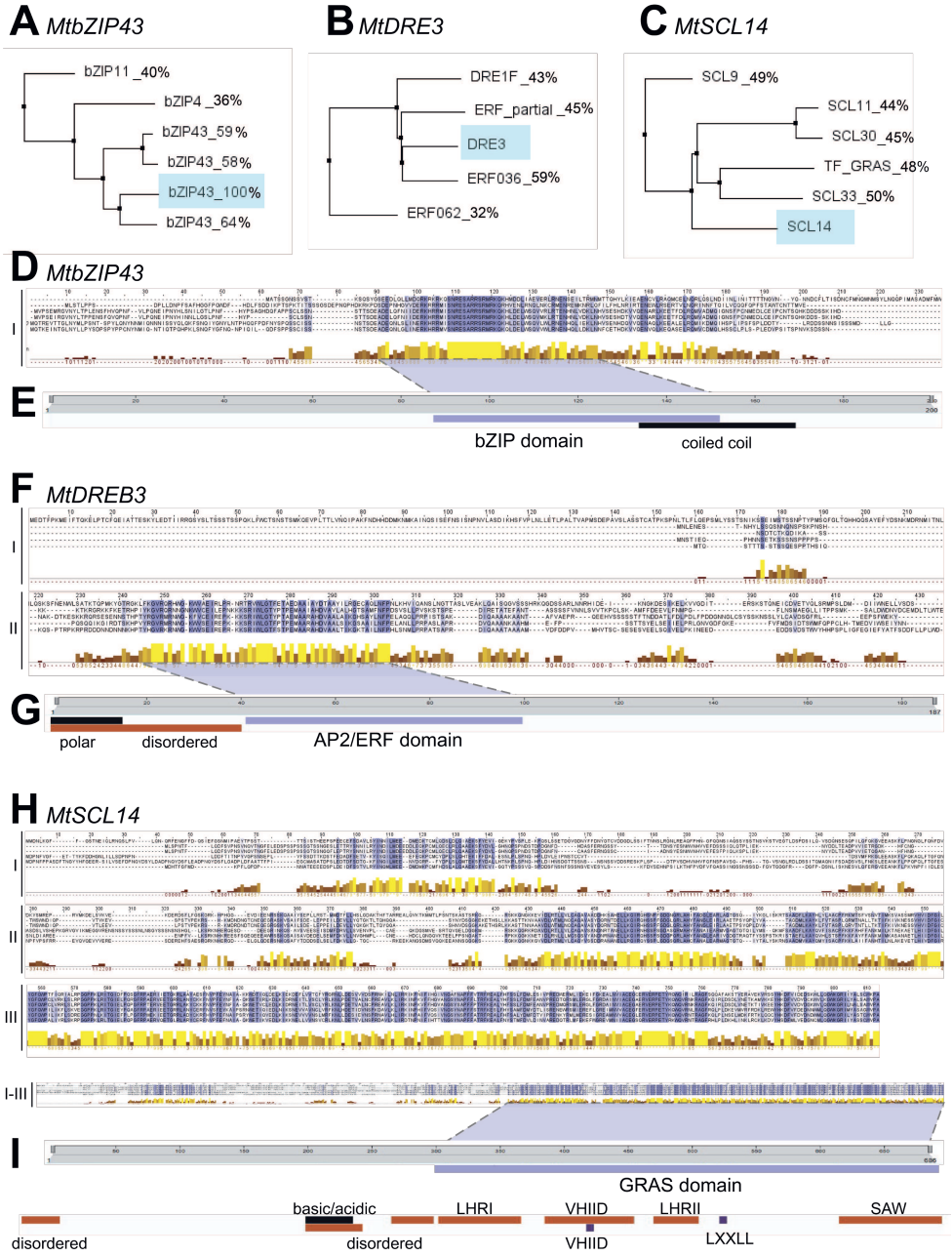
Zhang, H., Cao, Y., Zhang, H., Xu, Y., Zhou, C., Liu, W., Zhu, R., Shang, C., Li, J., Shen, Z., et al. (2020a). Efficient Generation of CRISPR/Cas9-Mediated Homozygous/Biallelic Medicago truncatula Mutants Using a Hairy Root System. *Front Plant Sci* 11, 474707.

Zhang, M. J., Ntranos, V. and Tse, D. (2020b). Determining sequencing depth in a single-cell RNA-seq experiment. *Nat Commun* 11, 1–11.

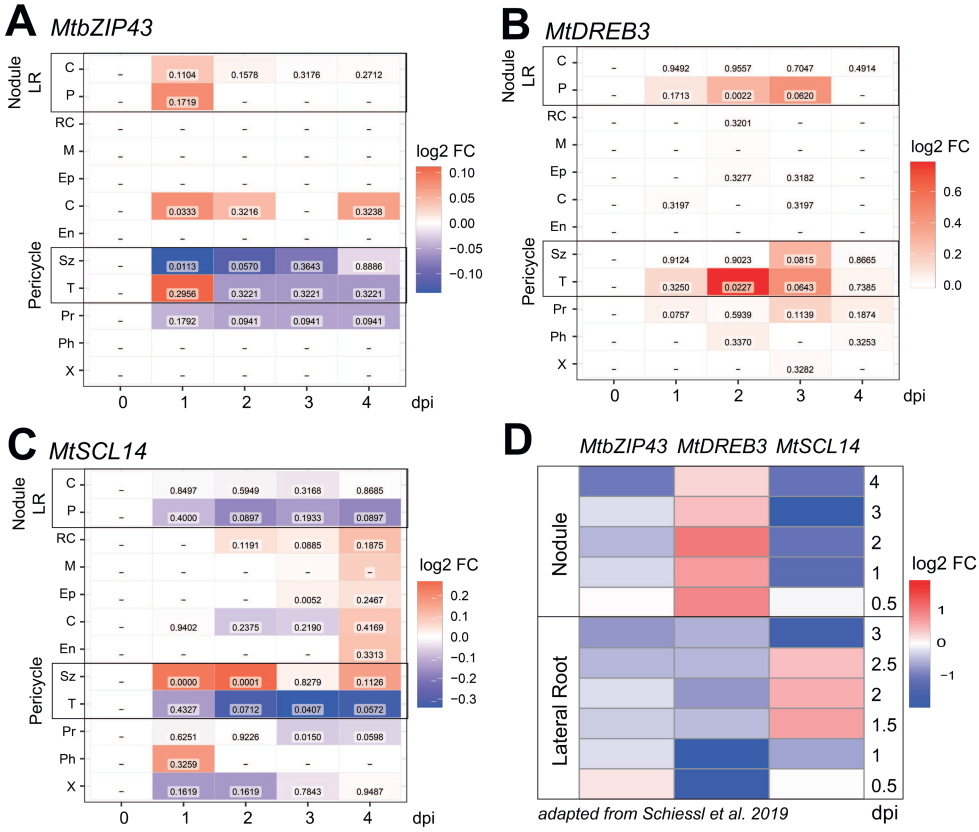
SUPPLEMENTAL FIGURES



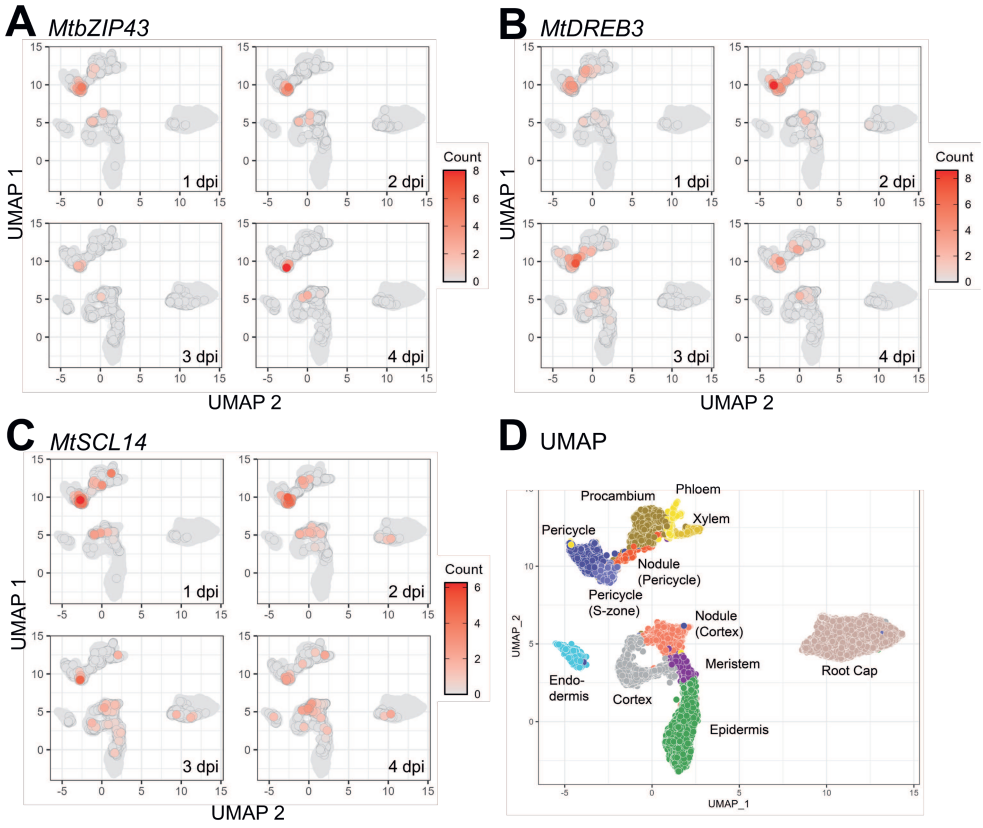
Supplemental Figure S1 | Pericycle specific gene expression patterns. (A) Differentially expressed genes (DEGs) from Schiessl et al. (2019) mapped onto the scRNA-seq dataset, focusing on pericycle-enriched genes. Each row represents a gene and each column a single cell across annotated timepoints (0, 1, 2, 3 and 4 days post-inoculation [dpi]). Values represent relative expression displayed as log2 fold change (FC) compared to 0 hpi. **(B)** Influence of protoplast isolation on gene expression. Comparison of counts per transcript in whole-tissue RNA versus protoplast-isolated RNA for the 5-mm root tip and 10-mm susceptible zone (S-zone) above the tip (whole tissue, gray; protoplasts, green).



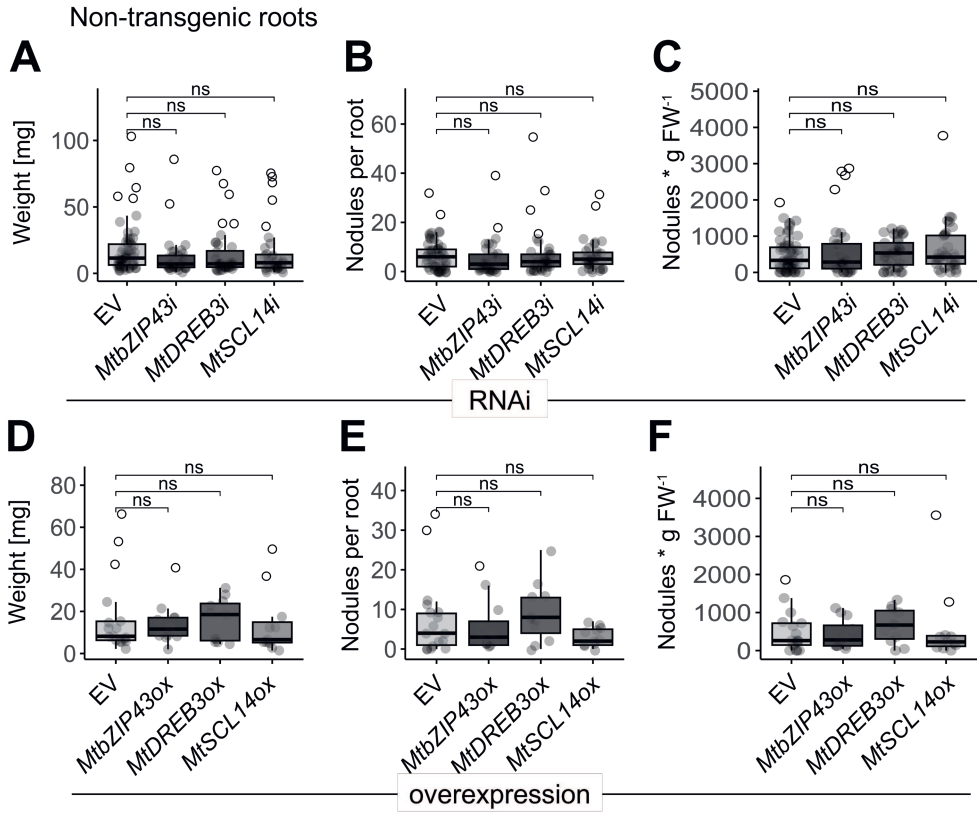
Supplemental Figure S2 | Paralog analysis of MtbZIP43, MtDREB3, and MtSCL14. (A–C) Phylogenetic trees of the most similar proteins (≥70% sequence coverage by BLASTp, NCBI). The focal candidate protein is highlighted in blue; percent sequence identity is indicated. **(D, F, H)** Sequence alignments of related proteins (purple, arranged according to the corresponding phylogenetic tree) with consensus amino acids shown in yellow. Increasing purple intensity indicates higher sequence conservation. **(E, G, I)** UniProt-predicted protein structures showing motifs (black), domains (red), and family domains (purple). Phylogenetic analyses and alignments were performed using ClustalOmega (EMBL-EBI), and visualized in Jalview.



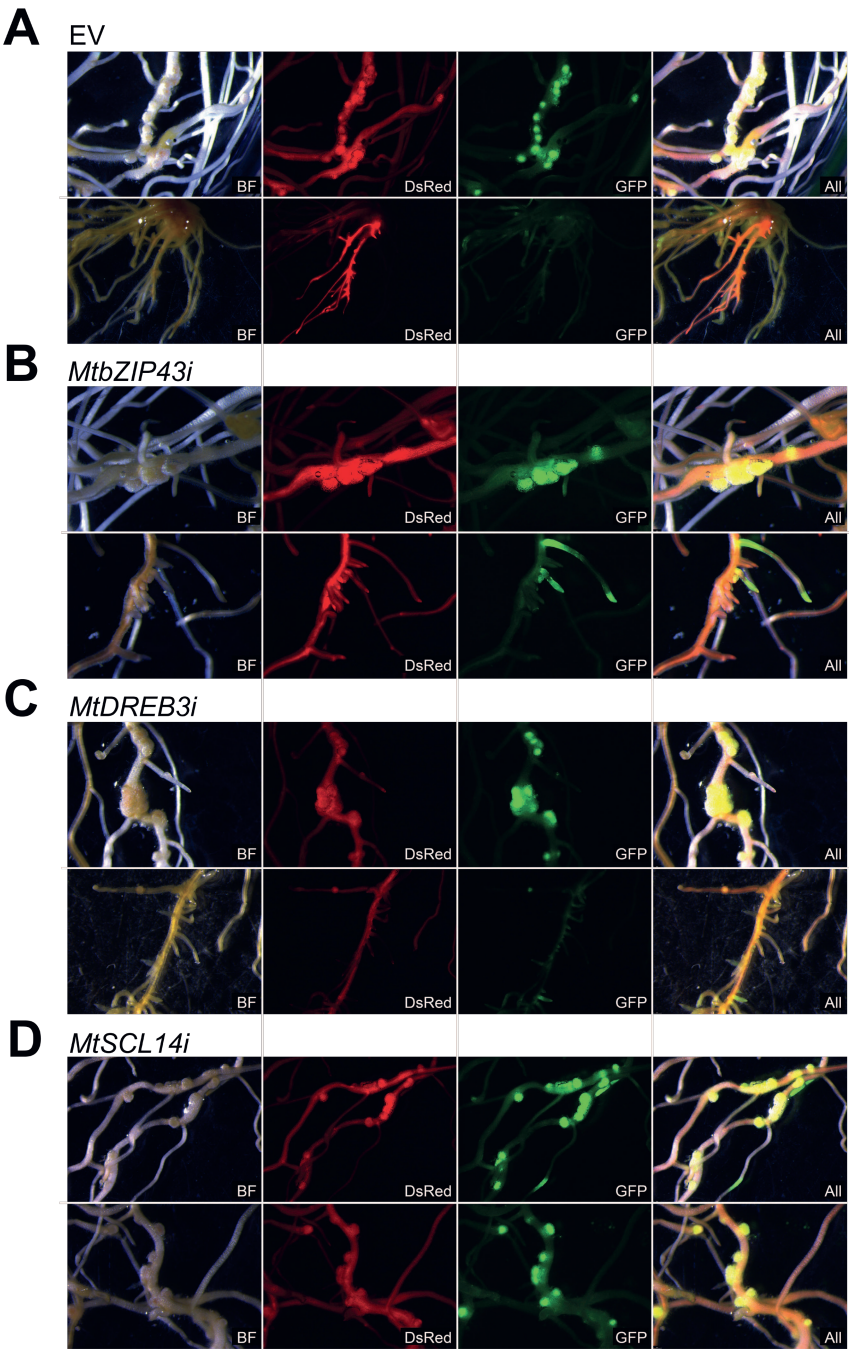
Supplemental Figure S3 | Additional heatmaps of MtbZIP43, MtDREB3, and MtSCL14. (A–C) Heatmaps of candidate gene expression across all cell types after *Sinorhizobium meliloti* 2011 treatment for 0, 1, 2, 3, or 4 days post-inoculation(dpi). Values represent log2 fold change (FC) relative to 0 dpi. Corresponding *P*-values for each comparison are indicated. **(D)** Heatmap of candidate gene expression during nodule and lateral root development, adapted from Schiessl et al. (2019). **Abbreviations:** C, cortex; P, pericycle; RC, root cap; M, meristem; Ep, epidermis; Pr, procambium; Ph, phloem; X, xylem; Sz, S-zone; T, tip; LR, lateral root.



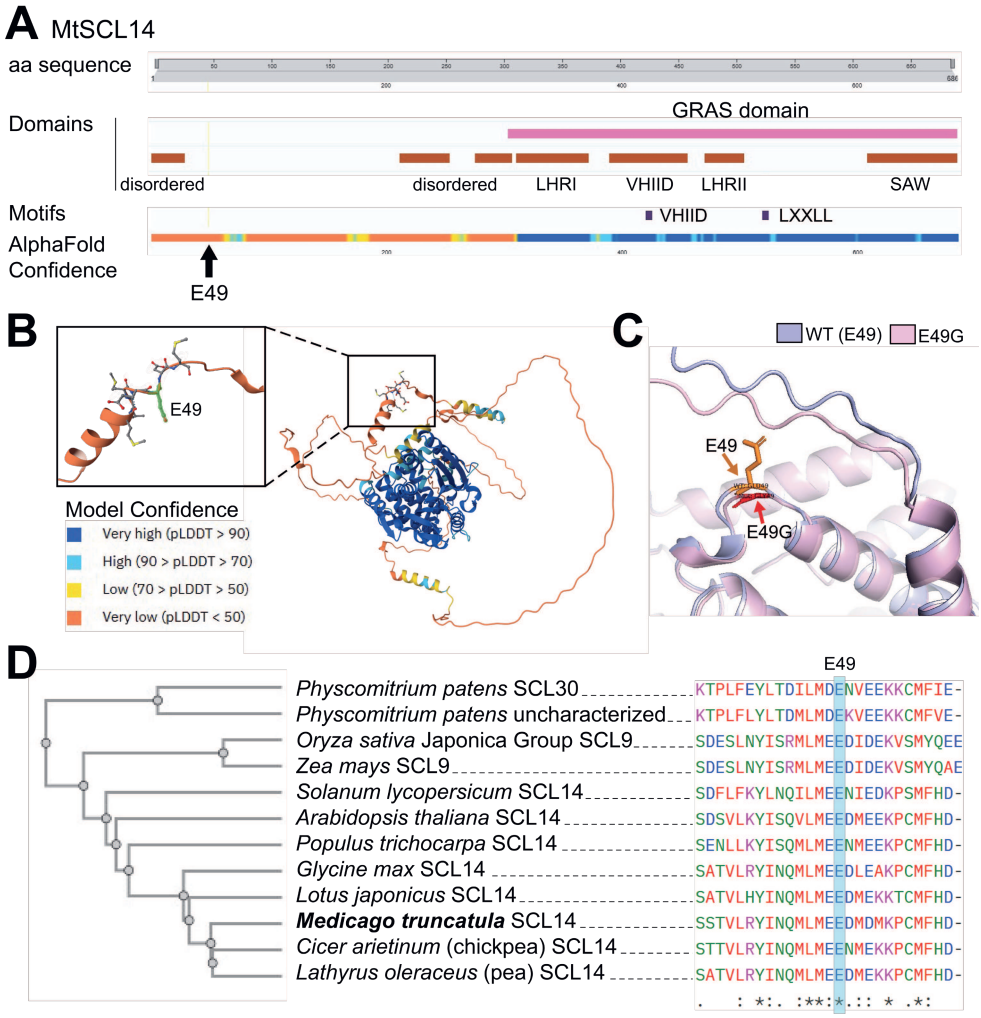
Supplemental Figure S4 | Expression domains of selected candidate genes in the scRNA-seq UMAPs after bacteria application over time. UMAP projections showing counts per transcript for candidate genes after *Sinorhizobium meliloti* 2011 inoculation at 1, 2, 3, and 4 days post inoculation (dpi). Panels show expression of **(A)** *MtbZIP43*, **(B)** *MtDREB3*, and **(C)** *MtSCL14*. **(D)** UMAP projection of the scRNA-seq dataset showing annotated cell-type clusters, color-coded according to Figure 1B.



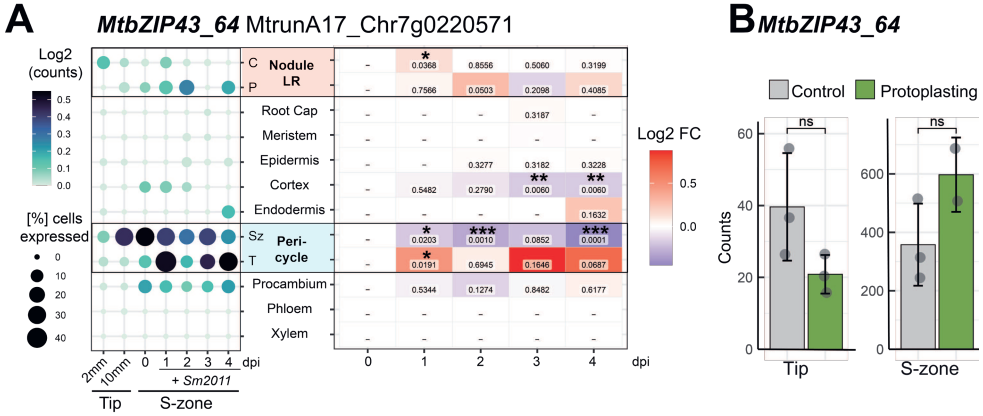
Supplemental Figure S5 | Non-transgenic roots of RNAi and overexpression lines show no phenotypic differences. Boxplots of single-root parameters for **(A–C)** RNAi lines and **(D–F)** overexpression (ox) lines (RNAi: n ≥ 13; ox: n ≥ 11): **(A, D)** root fresh weight (mg), **(B, E)** nodule number, and (C, F) nodules per gram fresh weight (FW).



Supplemental Figure S6 | Additional RNAi imaging. Representative images of roots transformed with **(A)** empty vector (EV), **(B)** *MtbZIP43*i, **(C)** *MtDREB3*i, and **(D)** *MtSCL14*i. For each construct, images are shown from left to right: bright-field (BF), DsRed fluorescence marking transgenic roots, *Sinorhizobium meliloti* 2011::GFP signal (GFP), and merged channels (All).



Supplemental Figure S7 | *In silico* analysis *of the MtSCL14* E49G mutation. (A) *MtSCL14* amino acid sequence showing predicted family domain (pink), domains (red), motifs (black), and AlphaFold structural confidence (UniProt ID: G7JUT7_MEDTR). The E49G substitution (glutamic acid to glycine at position 49) is highlighted. **(B)** AlphaFold-predicted *MtSCL14* structure with residue E49 marked and enlarged. **(C)** Structural comparison of wild-type (WT; E49, purple) and mutant (E49G, pink) proteins predicted using CoLab AlphaFold2.0 and visualized in PyMol. **(D)** Phylogenetic tree of proteins most similar to *MtSCL14* across diverse species (BLASTp, NCBI), with sequence alignment of the region surrounding residue 49. Alignment symbols: "*" conserved amino acid, ":" conserved property, "." similar property. Trees and alignments were generated with ClustalOmega (EMBL-EBI).



Supplemental Figure S8 | scRNA-seq analysis of *MtbZIP43_64*, sharing 64% identity with the candidate gene *MtbZIP43*.(A) Gene expression of the tip (2 and 10 mm), and the susceptible zone (S-zone) following inoculation with *Sinorhizobium meliloti* 2011 for 0, 1, 2, 3, and 4 days post-inoculation (dpi). **(A-left)** Dot-plot analyses showing gene expression across cell types; color scale represents log2-transformed counts and dot size represents the percentage of expressing cells. **(A-right)** Heatmap across cell types, presented as log2 fold change relative to 0 hpi. **(B)** Influence of protoplast isolation on gene expression. Comparison of counts per transcript in whole-tissue RNA versus protoplast-isolated RNA for the 5-mm root tip and 10-mm susceptible zone (S-zone) above the tip (whole tissue, gray; protoplasts, green). **Abbreviations:** C, cortex; P, pericycle; Sz, S-zone; T, tip; LR, lateral root. Asterisks denote significant differences between treatments (*P ≤ 0.05; ***P ≤ 0.001).

SUPPLEMENTAL TABLES

Supplemental Table S1 | Protein IDs used for paralog analyses

Protein name	Protein accession (NCBI)	Gene ID (v1.9)	Query coverage	E value	Percentage identity
bZIP43					
bZIP43_100	XP_013451170.1	MtrunA17_Chr6g0456061	100%	3.00E-147	100
bZIP43_65	XP_013447699.1	MtrunA17_Chr7g0220571	95%	2.00E-65	64.29
bZIP43_59	XP_013448026.1	MtrunA17_Chr7g0225661	80%	8.00E-54	59.01
bZIP43_58	XP_013444229.1	MtrunA17_Chr8g0341311	80%	2.00E-53	58.75
bZIP4	XP_003608557.1	MtrunA17_Chr4g0054291	60%	5.00E-16	35.92
bZIP11	XP_013466147.1	MtrunA17_Chr1g0155551	51%	3.00E-17	40.2
DREB3					
DREB3	XP_013462028.1	MtrunA17_Chr3g0140161	100%	7.00E-140	100
ERF_partial	UUW38918.1	n.a.	85%	2.00E-44	44.69
ERF062	XP_003591645.2	MtrunA17_Chr1g0196781	79%	2.00E-20	32.35
DRE1F	XP_039687047.1	MtrunA17_Chr2g0321071	75%	3.00E-32	42.57
ERF036	XP_003626994.1	MtrunA17_Chr8g0340311	72%	6.00E-45	58.62
SCL14					
SCL14	XP_003606670.1	MtrunA17_Chr4g0031721	100%	0	100
SCL9	XP_013448874.1	MtrunA17_Chr7g0238891	97%	0	49.35
GRAS_TF	RHN76102.1	MtrunA17_Chr2g0328391	97%	0	48.41
SCL33	XP_003606669.1	MtrunA17_Chr4g0031711	96%	0	49.78
SCL11	XP_013465417.1	MtrunA17_Chr2g0328411	96%	0	44.36
SCL30	XP_013465418.1	MtrunA17_Chr2g0328431	96%	0	44.81

Supplemental Table S2 | Protein IDs used for phylogenetic analysis of MtSCL14, and characterization of MtbZIP43 and MtDREB3

Organism	Protein name	Protein accession (NCBI)
Physcomitrium patens	SCL30	XP_024401617.1
Physcomitrium patens	Uncharacterized	XP_024400152.1
Oryza sativa Japonica Group	SCL9	XP_015630939.1
Zea mays	SCL9	NP_001167714.1
Solanum lycopersicum	SCL14	XP_004241220.1
Arabidopsis thaliana	SCL14	NP_172233.1
Populus trichocarpa	SCL14	XP_052311923.1
Glycine max	SCL14	XP_003539083.1
Lotus japonicus	SCL14	XP_057449184.1
Medicago truncatula	SCL14	XP_003606670.1
Cicer arietinum (chickpea)	SCL14	XP_012572738.1
Lathyrus oleraceus (pea)	SCL14	XP_050901067.1
Arabidopsis thaliana	bZIP42/bZIP43	NP_189674.1/NP_198696.1
Arabidopsis thaliana	DREB3/TINY2	NP_196720.1

Supplemental Table S3 | Primers and Gene IDs used in this study. GGC – Golden Gate Cloning

Purpose	Primer Name	Primer Sequence (5' – 3')
***MtbZIP43*, MtrunA17_Chr6g0456061**		
Promoter, GGC	pbZIP43_1_F	ttGAAGACttGGAGTTCGCAGTTGCGTGCGA
	pbZIP43_1_R	ttGAAGACttGAGACTGTTATGTTGAAAGC
	pbZIP43_2_R	ttGAAGACttCATTGGTTTTTTTTAAGGAATA
	pbZIP43_2_F	ttGAAGACttTCTCCCATTATAGTTATTATAC
CDS, GGC	bZIP43_SC_F	ttGAAGACttAATGATGCAAACAAGAGAAGTA
	bZIP43_SC_R	ttGAAGACttAAGCTCAACCAAGAAGATCCA
Terminator,	tbZIP43_F	tGAAGACttGCTTTTTCAAGACAAATTTTTA
GGC	tbZIP43_R	tGAAGACttAGCGGACAGAAACACAATCATG
RNAi	RNAi_bZIP43_F	CACCGTAATTACCTAAACACCCC
	RNAi_bZIP43_R	GCATTCTCTTGAACAACTTG
qPCR	qbZIP43_3_F	ACATGCTCCCATCAAACACATCAC
	qbZIP43_3_R	TGCTTATGCATGATGATTGTGGAC
Supplemental Table S3 (continued)|		
***MtDREB3*, MtrunA17_Chr3g0140161**		
Promoter, GGC	pDREB3_3_F	ttGAAGACttGGAGATGTCAATTACGGTAAAC

	pDREB3_2_R	ttGAAGACttCATTATTGTTTGGTTTTGGCAA
CDS, GGC	DREB3_SC_F	ttGAAGACttAATGATGAATAGTACTATTGAAC
	DREB3_SC_R	ttGAAGACttAAGCCTAGTTATTATAAATTTC
Terminator,	tDREB3_F	tGAAGACttGCTTTTTTCAGATGAGTTTGTT
GGC	tDREB3_R	tGAAGACttAGCGAGCCAACAGCAACAAATT
RNAi	RNAi_DREB3_F	CACCCATAAGAGACACAAGCAAG
	RNAi_DREB3_R	GAAGTTGTTGAAGATGGAAAC
qPCR	qDREB3_3_F	CCCGAGATGTTCAAGCAGCAG
	qDREB3_3_R	CCCAAACGAGGAAGCTCAATTATCTC
***MtSCL14*, MtrunA17_Chr4g0031721**		
Promoter, GGC	pSCL14_3_F	ttGAAGACttGGAGCTGTTATTATTACTAGTGG
	pSCL14_2_R	ttGAAGACttCATTCTATGAGTGATTTGGGG
CDS, GGC	SCL14_SC_1_F	ttGAAGACttAATGATGGATCACACCTTTGG
	SCL14_SC_1_R	ttGAAGACttTCTCCTTCCATAAGCATTTGG
	SCL14_SC_2_F	ttGAAGACttGAGACATGGATATGAAACCATG
	SCL14_SC_2_R	ttGAAGACttAAGCCTATGCAGGTACCCAGC
Terminator,	tSCL14_F	tGAAGACttGCTTGACTTGGCAATGATCTG
GGC	tSCL14_R	tGAAGACttAGCGCTGTTGTGAGTGTTGTG
RNAi	RNAi_SCL14_F	CACCCTATGAAGTCATTGGTCAG
	RNAi_SCL14_R	CTCAAATTGCAATACAGACTC
qPCR	qSCL14_F	CTTTGCGACACGGTTCAAGG
	qSCL14_R	TTCCGAACCTGCCATTTCCT
Housekeeping Gene *MtUBQ*, MtrunA17_Chr4g0048591		
qPCR	MtUBQ_Fw	CACCTCCAATGTAATGGTCTTTCC
	MtUBQ_Rv	CCCTTCATCTTGTCCTTCGTCTG

Supplemental Table S4 | Sequences used for RNA interference. CACC overhang is added to primers to enable TOPO cloning.

MtbZIP43i (308 bp)

CACCGTAATTACCTAAACACCCCTCATCAACAATTTCCTGACTTCAACTATAGTCCACAATCATCATGCATAAGCAGCAACTCAACTTCTGATGAAGCAGATGAACAAAACCT
TAGTCTCATCAATGAGAGGAAACACAGAAGGATGATATCAAACCGCGAATCAGCGCGTCGTTCACGCATGAGGAAGCAGAAACACCTTGATGAGCTTTGGTCACAAGTTCT
TTGGCTAAGGAATGAAAATCATCAACTTATAGAGAAACTTAACCATGTTTCTGAGAATCATGATCAAGTTGTTCAAGAGAATGC

MtDREB3i (307 bp)

CACCCATAAGAGACACAAGCAAGCATCCGGTATATCGTGGTGTCCGAATGCGAAATTGGGGAAAATGGGTGTCCGAAATTCGTGAGCCTAAGAAAAAATCCCGAATATGGC
TCGGCACATTTCCCACACCGGAAATGGCAGCTCGAGCACACGATGTAGCTGCTCTCAGTATAAAAGGAAGTGCCGCCATTCTCAACTTCCCTGAGTTAGCAAACTTGCTTC
CTCGACCGGCTTCACTTGCTCCCCGAGATGTTCAAGCAGCAGCTAATAAAGCAGCTCACATGGAGTTTCCATCTTCAACAACTTC

MtSCL14i (308 bp)

CACCCTATGAAGTCATTGGTCAGACTTACCCTTCTTCTTCTGTACAAAGCCCTGATGATAGTTTTAACAGTAATTTCAGTAGTAGTAATTCTGTTGAAGTTGAATCATATTTGAG
TAAATTTGATGTTTCTGAGCATAATCCTTCTATATTGAAAACCCATTTTCCTGTTGATTTTGTTTTTCAAGCTAGTTCAATCAATGGTAGTAGTAATAGCAGTAGCGATTTCGGTG
TAACCGGTGGTGGATTTAGGGATGGATTTCCTGATTCAAATTTGATTAGTAAGAATGAGTCTGTATTGCAAATTTGAG

Supplemental Table S5 | Golden Gate cloning constructs

Final constructs	
EV (empty vector)	pL2B-KAN-D2-p35S:mCherry
MtbZIP43pro::GUS	pL2B-KAN-D2-p35S:mCherry-*bZIP43pro*::GUS::tbZIP43
MtDREB3pro::GUS	pL2B-KAN-D2-p35S:mCherry-*DREB3pro*::GUS::tDREB3
MtSCL14pro::GUS	pL2B-KAN-D2-p35S:mCherry-*SCL14pro*::GUS::tSCL14
MtbZIP43ox	pL2B-KAN-D2-p35S:mCherry-*AtUBQ10pro*::bZIP43::tbZIP43
MtDREB3ox	pL2B-KAN-D2-p35S:mCherry-*AtUBQ10pro*:DREB3::tDREB3
MtSCL14ox	pL2B-KAN-D2-p35S:mCherry-*AtUBQ10pro*:SCL14::tSCL14
Used parts from Golden Gate Toolkit (Engler et al., 2014) and Engineering Nitrogen Symbiosis for Africa (ENSA) (https://www.ensa.ac.uk/) (ECxxxxx)	
EC50507	Level 2 backbone
EC15029	*pNOS-Kanamycin* resistance (position 1)
pICH54022	Dummy 2 (position 2)
EC75021	pL1M-R3-p35S::mCherry (position 3)
pICH47831	Level 1 backbone (position 4, reverse, used for gene specific constructs)
pICH41780	End-linker (position 4)
pICH75111	pL0M-SC-GUS (GUS protein)
pICH41295	*AtUBQ10pro*, PU module (promoter)
pICH41295	Level 0, PU acceptor module (for promoter sequence)
pICH41308	Level 0, SC acceptor module (for protein sequence)
pICH41276	Level 0, T acceptor module (for terminator sequence)

Supplemental Table S6 | Gene IDs and Sequences used in this study. Based on *Medicago truncatula* genome version 5, annotation 1.9. For cloning purposes, any BpiI sites were removed by single base mutations (blue highlighted) to enable Golden Gate cloning.

MtbZIP43 - MtrunA17_Chr6g0456061, bZIP family of TF

bZIP43pro (promoter) 4.3 kb

TTCGCAGTTGCGTGCGAAAAATAACTAATTTATAAAAAAAAAATAGTTAGCTTTTACTAAATTATATCCTCTATCCTCTATTAAATGATATAAGCATTTATATTTTTTTGGGAACTA
AGCATATTATAGTACTTTGTAATCATATTGAAACACTTTAATGTAAGTTTCCACATGTGTGGAAAAATAAATAAAGACAAAGTAAAGAACAATCAAAAGACAAGAATGAAAGAT
GGGTCAAAATCTGTCTTATGATATTGTTGACCATTCAATCCACTCAACTCAACTAGACACACACGGGTAAGTGCATAGCATTTGTCTAACAAGATTTCATGTTTCTCATCTTTCT
CAATGCATTTAGCTGGCTTAAGGGTTAGTGGAATATACAATTTAAAGTACAATTCTTGTTCAATAATTTCCACTACATTTGTTATATATGACTTTTTTCCATGAATGGTTAAGATA
TCCATTCAATCTCTTTTCTTTCCATTTTTCCTAAATGAGTAATTAAGCTTGTGTCAATTTGTCAAAAAAAAAAAAAAAAAAAAAGATTAATCAAACTTGTGTCAATATAGCAAATA
ATATAACAACACTCAATAAACTATATGGATATCAATTTGAGGAAGTAGTTCAGAGAGACAATGATCAATGTATTGGATAACAATGTTTGAAATAGTGTCCCTCTAGGGATCACT
TGTTCCCCATTAGAAATTCCCCAGCATTCATGTAATAATATGATTTTTGACACAATTTGAGATCTATGGTCCATATACTAAGGCTTTGTTTAGGAGTTTCAATGCAGGGGAAAG
GAGGGTTTTGAAGAGTTGGAATATAAGGGTAAATGAAAAGATCTTTTACATTTTTTTATAGACTTAATTGCAGTTTTGACCTCCTAAGTATCATAATTTTGTGATTTTGGCTCTC
TATAAAAAAAACAATAACAAATCAGCCCCTAAATTTACCCTTTTTTTACAACTTTAACCCTCGAAACCAATTTTGATTGATTCAATGCATACGTGACACCGAGATCAATCACCG
GCCAAACATTACTCTACTCTAAACATTAGTAGTACGTTGCAATTATCTTAGCTTACTTTGGAGGATAGTCTATACTCTACACTAAAATAAAAAGCATCAATCACTGGCTATGGA
AACTTTCTCAAATTAGCACCAACTAAAAGACAAAGATGTGGAAAAGAAATGTAACCAACTACTTTACAATTTAGAAAGTTTGCAACATAAGAAAAATGTATCCTTTACATTTA
CCTTCATAACATAAATGTTTGTTGTCTGTTTTTTATTAGCCAAAGCAATAATCTAACTTATTATTACATGAATGAATTGACTAACACTACTTGAAACTCACCTATATAACTGTGAG
ATCCTGAGTTTAAATCCAATAAAAGTGTTCAGTCTAACAATATTGATCCTACTAGTTGAGCTATACATTTTGGACGGTGCATAACCAAATATTTTAACAGTACACCTTATATTAA
TAAACACCTATTCATATCTTTGACTGTTCACCATCATGAAACCATGTTCAAGTTAATTAAGGGCAATTTGATGTAGAATAATCATGAGAATATGCCAGTAAATATGTAAATTATTT
AAAATCAATGGTAACTTTAAAGTTTACTATAGAATAGGCATGTGGAACTAACGCAATGAGCATAATGCATTAATTCAACATTTAAGCTGGCATACAATTGGCATGTGGTCTTGA
AACTTCTAGTTGTGTTCCCTTCATTGTGATTATACTGTTGAAGTCGTCATTACCATGGCTTAATAATATACTTTGATTATCAAGTACCTAACCCTACTAAATAAAACAAGTGATTG
CTGATGGGTTAGTGATTTGGCTTCTATTTAACCCACATATACACATTTTGTTTGAGGTCCCAATAGAAGTGGTCATTGAGTCATCAATATTCTTAATCTTTCTTTATTTGTCTCTC
ATGATTTGTTGCCAATTTAGAAAGTTGTGAATTTCACAATTACCTTTTTTATTTTTTTTATTTTTTTTTATTTTTATCTAAAGTCTCTTACACCATGGTATGTATTGAGTCTTGTGCA
AGTTCTCATGATTTTATCCTTTGTAAAAGACTTAAATAGTTTATAAAAAAATCACACTCCTCAATTCACCGAGACGTCTTTAACAAAAAGCAAAACTGTGATATGTGAATTGGAA
TTGGTGTAGCGGACTTGAAATTGGAATTTTTGTACAAGTAATATAGGACCTACGCCACACTCGATTACTTAAAATATTTTTAGGCACGAAATTTTATTCTTATGTCCGTAAGAC
ATACATCAACTGACAATGTTGATATTGTTAGGATGCATAACTTCACCACATTATAAACTTGAGACGTTATAGTTGTGTGTGTTTTAATAGTGTTTGTCATTTTGTCTATATGTATG
GGGAATTTTTTGTGTCCCTATCCAATAATGTATTTCATATTTCATGAATGGAAGCAAGCTTCTAATCTTAATATAAATAAAATAAACCGTAATATTACACTCGTTCTTTGAAGATA
TGGACCCTCTTGATTAAAATTTCTAAAGTTTCCCTCAACTCAAATATCAAATTGCCTGCTTTGGACATAGTGTAATTTGTTACTTAAGATTTTATTTTCCATTGAGTAGTGATTGC
AGTTATGATACAATCTTTCAAACAATAAAATATTACAACTAAGGATGATGCAACCTTAATTAATTACAATTGCAATGAATGATTGAAGAAACCTAAAAAAAACATTAACTTACA
ACCACAATTTTGATTATACTTTAAAATATTGATTATATTTAGTCCAATAAAATCTTAGCCAACCTAATAAGCCTTTCTCAACCTTGTTTTTCCACAATTCCACCAAAATCTTATGT
CACTTAAGTGAATACTATCCACATCAAAACAGATAATAAGCTTGGAAGTACAAAACACATGTATTATAATAAAAATGACTTTAAACTCCAAACATTGTATTGTGCTTGCTTTCA
ACATAACAGTCT(T->C)CCATTATAGTTATTATACTCCCCATGATGTGAAATCCCTCTGTCTCTATATTACGTAAGCATATTTTTTTTTTCTTCAAAATGATTATATCCACAAAACCA
AATCGCATTCACCCTTTGGGCCACTTGTGTGCATCATCCATATACATAAACATGTCACATGTATGTTGTTATGCACATAAATACACACAAATGAAATTAGATTTTTTTTTTCCCTT
GATGAAATATTCATACACATTAAACTCCTCAAGTGGTTAACACTACCAAAAAACCAAAATTAGTGAGGGAAAACAAGAAATTCCTCTCTAATTTTGTCATTAAATTTTGCGATC
AAAGAATTTGTGAGAGATTTTATTTTTTCCGTCATTAATTTTCCTCCCTAATTTTGCTTTTCCTAATAGTGTAAGGCAAGAATCAATTCTTATATATAAGAACCCACCAAAGAGT
CTACATTTAGAGATGTTCCAAGTACAACTAATAAAACTGTCATATATATTAATAAATTAATTTAACTCAAGTGGTTAACGAAGTTCAATAAAAGTCAAATCGATTAAGAGTATGA
AGTTTGAACCGTGTTTTTTATAATTATTAGTCACACTTTACTTATTTTTCGTTTGATCAACGAATTCATAGAGCCACTTTTTCTTGAAAATCGTTGGGTTAAAAAAATTCACCAC
ATACACATGATGTGCATCTCCAACATGATCGCGACATTATGAGCTTCAAATGATGCAATTACTCAAGTGAGAATCGATTCTTTTAAAGAAGTGTTAAGAGTCCCACATATCAGA
TATGTGATGGTCTGACAATGTGTTTATAAGGGGGCAATCCTCACCTGACAAGTTGGTTTTGTGGGGTTGTGTTAGGCCCAACCACAATTCTTAAGAATAAGAACCCGCACGC
AAACGAGTATGCATAAATAGAGATGTTCCAAGTGCAACTAATTAAATTCATCATATATTAAGTTTTTGGATTCAGATGATCAATAAACTTCAATGAATGTTGATTTAATCAGATAT
ACTTAGTTCGAATCTTGGTTGGAACAATTCTTTATCGGACTATATTTATTTCCCGGAAGAAATTTATATTAACAAACCCTTTCCCCTTAAAACCAAAAAGTTTTATTCAAAAATA
AGCCACATATTCATGATGTGCATGTCATTCCATCAAAATACAATAAGCTTAATTATAACAAAACATCAACATTGTGAACTCTACTATTATCCAAGTAGGGCCAATTCATCAATTT
CCCCTCCCTATATAATACCTACTCTGCTCACCTCACTCTTCAAACACAATTGAACCTTCCACCACTCTTATTACTCAACAAAACCCTTTCAATCCTTAAATTTTTCTCATATATTC
CTTAAAAAAAACC

bZIP43 (CDS)

ATGCAAACAAGAGAAGTAACAACTGGACTCAATTACATGCTCCCATCAAACACATCACCTTATCTACAAAACTATAACAACATGATTCAAAACAATAACATTTCCTCATATCAA
CTCCAAAAATTCTCTAACCAAATATATGGTAATTACCTAAACACCCCTCATCAACAATTTCCTGACTTCAACTATAGTCCACAATCATCATGCATAAGCAGCAACTCAACTTCT
GATGAAGCAGATGAACAAAACCTTAGTCTCATCAATGAGAGGAAACACAGAAGGATGATATCAAACCGCGAATCAGCGCGTCGTTCACGCATGAGGAAGCAGAAACACCT
TGATGAGCTTTGGTCACAAGTTCTTTGGCTAAGGAATGAAAATCATCAACTTATAGAGAAACTTAACCATGTTTCTGAGAATCATGATCAAGTTGTTCAAGAGAATGCTCAGC
TTAAAGAAGAAGCTTTGGAACTTCGACAAATGATAAAAGATATGCAGATTCATAGTCCTTTAATCCCTTCTTTTAGTCCTTTGGATGATACATATCTTAGAGATGATTCTTCAAA
CAATTCCATCTCAAGTTCAATGGATCTTCTTGGTTGA

tbZIP43 (terminator) 500 bp

TTTCAAGACAAATTTTTATCTATTTGGTTTAATTTCTTTACTTTTTTTAGTTTGTTTGTGTCTTGTGTTGATGTCAAAATCAGTAGTTTTCTCCTCCTTTTGAGATATGTATGAGGT
GAAAAGGATAAATGGCAATAACTTGAGTTGAGTTTCATTATTTGAGCCTTTTCACTTCATTAACTATGAAACTACCTCTCATTGTATATAAAAGAGTAGTATCTAAATTTGTATGTC
TTTAGCAATAAAATTTTATGAGCCTCAAGAGTTAGAGTAAACAAGTGTATTCACTATAGTCTCTTAATGCATTTAATATTAGGGTTACGCTAACGAGTATCCTTAGGTCTTAAGG
GCACTTGTTAAAGACTTAAAAAAGAAAATAATTGTTAAATTTCATGTAGAAAAAGTTTCTTTTTAATCGTTCAATAAGTTGAATACACAATTTCTAAGATAAAGTTGATACTTTA
TGTTTCTTATCAAGGGTCGTTAGGGCACTTGTTAGCATTTTTCTTAATATTATCATGATTGTGTTTCTGTC

Supplemental Table S6 (continued) |

MtDREB3 - MtrunA17_Chr3g0140161, AP2-EREBP family of TF

DREB3pro (promoter) 3.5 kb

ATGTCAATTACGGTAAACTAATTTATCTTCTTAAAAAAATATTAAATAATAGTTTATGAATGATGATATTAAAATGATGAAGTTAACAAAAATTTGTTTATAGTTTAGTCAATTTAC
TTTTTTTTACTAATATTTGTCACTGATTTTTTTTTGTCTGAACCCTTTTTTATGAATTTGTTTGAACCCTCTTGTTTCTAAGAAAAATAATATTGATAATTCAAAGTTTATTTATGAT
AATTATATATCCACGTTATGTGATATTACTATCAATATTACTTGCACATAAATAGGAGGCATAACATTAATGGGCTTAGGAGACATAACATTGTTATGCATAACCTTTTGAACTCT
ATGTAGAAGGTCAACAACCTCTGATGCTAGAAAGTCGAAAGTGTCAAAAATAAATGGTATAAAAGCATGTTGGTTGTCATAACATGCTTTTTCATGTTTGACGCAGTTTTTAG
GGTTGTTTGTCCTATCGTAAAAGTCTCGACCCCTAAACCCACAAGTGGTGAAACCTTAATCAAGTCCACACATGATCAAATATTGTTTCAAACAACATTTAAATTTGGGTTCTT
CCTCAAAAAGTCAACTTTGATGTGAAATGGACCAAACAAGCGAGAAAAAAAAATATAAGTAGTAATATTTCTATATTTAGTTACAATAATAATATATATTTAGATGCTCCAAATAA
ATATAAATATTATTACACAATTAAAAAATTTATTCTTCATAGTTTGGTACTTTCAGTCTCTATAGCAACTACAAATGAAAATTTGTTAGTAATGAAAAAAGTAAAGTCAGAATTAA
AAATAAAAATAAAAATAAAAACTATGCTGCTATAATTTGTATTGATTTTTTTAATCACATCCCTTAAAAAATTGCTTTACTCATCTTCCTCATTTTCTCTCACTAAATCTCTCAACT
TCCAATATATGAAGGGTAAAGACATTTTTGTTACAAGACAATGATGCCACATGTCAATAATCATCTAAAGACAAGTTCCATTTTTCATATAATGTAATGTAAAGTGAAAAGTTTC
ATGTATTGACAAGTTGCAAGATTCATAAGATGCAAAGTGACCTGGTGCACACTTCATGTATTGCTTGTTGTTTGTAAGAAAATTTGGTTTACCTCTTCTATAATTAGCAAGCTC
ATTTATAGACTTCATCATACCATCAAGTAGTAACACACCAACAAGCACTAAAACTTGTTTCAGAGTGAGAAGAAACAAGGTTAATACCTCAAAAAATCTTTGTAGTTTTGTGA
GGTTTGAAAGTAGGCAAATTCTATGGTACACCAAGTCGTAAAATTAATAATAAATGAGTTTTTTTTTTTTGAAGAAAAATAATTATTTTCTATCATTGACAACTAAGATAATTAA
ATTTTATATACTAAAAGTTTCGACGAAATAAAATTTAATTTTTAAGAATTTTTATTACTCATAACAATGAATATTGATTATTTTTTATGACAAAATGTAAATTTTGTAATATGATCCT
TATAAACAATGAAATGATGGTTTTTCTTATGAAATGATGTTTTCTAAAATATAAATATTGACAAATACTTTTTTATTTCATTAAAGTGATCTTTAATTTATATATTTTGTGAAACAAG
AGTACCGGTACACTCAAAATTGTGAGTGTACCATAGAAGTTCCCGGCTACTAGGTCTTTTTTAAATTGATTCAATTGTGAAATAAAATATCATGACATGTGTATTGTGTATCTAG
GGAGAATTCACTTTGAAATGACTAATCCCTAGATACATCTATTTCATTTTAGACCTATTCTGAATATATAGATTTGTGCGAATTTTTAGAAAAAATAAATGATGAAAAAAAATCC
AATGGATTTCAAATTTATGCAAAATGCTTTTCGATTTGTAAAATGTTTAATATATTTTTTACATCTTCTACACGAAAATGTTCAATTTTCATAAGGTAGCCCCCATAAAAAAAATA
AAAATAAAAATAAAAATAAAAAAAACAAATTTCACCTTTGTAATTCAGGTTCTTCATATTTTAACCTTCACCCTTAAAAAAATAGTTGAGCGAAGCCTAAGAATAACCCTTGTA
AAGAGTTAAAAATATGAAGAATCTAAATTACGAAATGAAATCTGTGTCTGTGTTTTACAAGGATTATTTTCATACCTCGTCCAAATTAAAGTGATATTTGATGTACTAACCTAAA
GAATAATTAGTTTCTCATGTATGGTTGTTCCTTATGATAAATCCTTCTACTCTTACCACGTAGTTTTTCGTTTCTGAGTTTTGAGCCAGTCGAAAATAATAAGTAATAGTAATATA
TTTAATTTGAGACTCGCACTCAGATAATTAATAACCTCAAATAAAACTCTTACAAGTGATACTCACATTAAGTCAACATAATCGTATAGCAGAATCCCTACCCCAATAAAGTTA
AGATCATGAAAACCACGTTAAATTGAAAGGCATAAGAAAGCCTAAAGAGCCAATTAAAATGAGAATATGTAAAGTATAAATGAAGGCCTAATGAGCTATTAGCGAGTCCTTG
CCACCACATCAACAACCTTTATAAATCCCACTAACGAACCAAACTTGTCGCTGCCATTATGTATGTACCTTCTTTTTTTTTAGGTAATGTATGTACCTTAAAATACTCACTCCTA
TACATATATGTAGCAAAACTAGGGGCAAATTGATGATAAACGGGAATATTTACATTGCATCAATTTGAACAAAACAAGAGAACCCATCCTTTTGGTTCTGCAGAAATTGCTGA
TAAGGGAATATACATGATCAGTATAATTATATGGGTGAAGAAACAGGAGGCAGTGGTTAAAAAAATAACATGAACCAAACAAACAACAGTACATAATATACAGCAAAGTCAC
AGAGAGGCCAGAGTTTGTGTGGCACAAGGGCGGTACGTGTACCACTTGTAATTACCTTGCCGGGTATTGGATTTGAGTTATGTGGGGTCGATCACATTTAGACCGTCCGATT
AAGATCGCACTTGACTGTATAGTTCCAATTTTAATGTAGATTTGATTTCTTATGCCTAAGATATCAAATTTGAAAGGATCAAGGACACTGGGTAAGGATACTAAAAAAGAAAGT
TATCCATCATATTTTATTATTGATCTACACCTTACTCTTTTGAAGGATTGATCTACATCTTACTTGTAGTGGAGCAGAATGTATGTTCTTCAATTGCCGGAATGCATGGTAATGTT
AGCACTAGAGGATTGTTGTTAGAAGTTTTACATTGATATGATTTTTAAAGTTACGAAATTCTCTTACCTAAAATCCTAGTGTTTTTTGAGATTCAACTTATTTTGAGTTAAAATTT
ATGACTAATTTTCTATCGTGGGTCCCAGGTCCGTCCTTTACACTATACTACTCTTTTCAATATCAACTCAACCTTTATAAGTGTGATTGGCAAGTGACGGAAAAAACGGTCACG
GCCCATTAATTTCCTTTTATTTCTCTCCTTCCCATTGCCAAAACCAAACAAT

DREB3 (CDS)

ATGAATAGTACTATTGAACAACCTCATAATAACTCAGAAACTAAAAGTAGCTCAAATTCACCACCACCACCATCACCACAATCCCAACAAATAAAAGGCATAAGAGACACAA
GCAAGCATCCGGTATATCGTGGTGTCCGAATGCGAAATTGGGGAAAATGGGTGTCCGAAATTCGTGAGCCTAAGAAAAAATCCCGAATATGGCTCGGCACATTTCCCACAC
CGGAAATGGCAGCTCGAGCACACGATGTAGCTGCTCTCAGTATAAAAGGAAGTGCCGCCATTCTCAACTTCCCTGAGTTAGCAAACTTGCTTCCTCGACCGGCTTCACTTG
CTCCCCGAGATGTTCAAGCAGCAGCTAATAAAGCAGCTCACATGGAGTTTCCATCTTCAACAACTTCATATGAATTGAGTGAGATAATTGAGCTTCCTCGTTTGGGAAACGTT
GGAGATTTTGGAAAGGAGTTTGTATTTATGGATTCCATTGATACTTCATGGATGTTTCAGCCTCCTTGCTTGCACACCATGGAAGATGTGATTTGGAGTGAAATTTATAATAAC
TAG

tDREB3 (terminator) 500 bp

TTTTCAGATGAGTTTGTTTAATGTTGCTATAAGTGGTCCTTCCATTTTTTTCAATTATTTTATGTGAATGTCTTAGTGTGTTTATCATGTTTTGTTTATTATTATCAAATATTATCATA
GAGACGGAGGCAAACACCTTATTGACAAACGAGGTGTCAATTCTGTATTCTTGTTTTGTATAAATCTCTCTCTATCTTTCTTTCATTTTGTTTTTTGTTTTTTAGTTTTTGAGTGT
CTCTGCTGCAGCCACTTGTGGTTTTTCTCTTATGAGTATGCTGCTCTTGTCCATTGTGTTGTTTTTTTCTTTTCTCTTCCTCTTAGACTCAGTTGAGTTGCAATGTAATTATGGTTTA
GAAATTATAATAATAAAATGAAAGTGGTAAATTATTTGATTTTCCTTTTTCTGTTTTGGCAACGAATGTTTTTATGATGACATGCAGGTGGTTGCCTCGTTCTTTGAGACTGTCA
AAATTATAAAGTCTTTAATTTGTTGCTGTTGGCT

Supplemental Table S6 (continued)|

MtSCL14 - MtrunA17_Chr4g0031721, GRAS family of TF

SCL14pro (putative promoter) 2.2 kb

CTGTTATTATTACTAGTGGTGACTGGTGTTCAAGAGATAAGTAAGTTTGTCCCGGCACCAATCATAAATGTTTACTAAGCCATTTTTTTCAATCTCAATCTTGACCCCAAACCT
GAGTAAAATTTATAATGTAATAGTGTAATTATGTTTGTGTTTATGAAGGATATAATCGAATTGATCTAAAGGACAACAAAGGAAAATTACAAAATAATCGTGCAGTGAAACCCG
ACCTAAGCAATGTATCTAATGAATCAGAATTGATCATAAACTTTGTTCCTGTATAAATCGACACTGATAATTCAGTCACAATGCAATGAGGTCATACCTATTTTCCTTCAACTTT
TTGAAAGCAATCTAGAATATTTTCTGCCAGTTTTGTTTCCCTTTTTTACTTCAGACAAACTTGATTCTGCAGTTCCTATCTTCTCAATGACTCAATCAAAACTGAATGAAGATGT
TTCTGGCAGTAATTGGTAATCTATTATGAAGCAATCACAATCTTAGTAATTGGTAATCTTTGATTCAGATGAGATGCTGAATTGGTTATCTTCACATCTCCATTGTAACTTCCATT
TCCTCAAACCAAATAATATGGAGTTTTGTAAAGAGCCCTCACTGTGTCGTTTTCCCTGTCCACCTTGCCATAAGATAGAAACCGCGAAGTGGAAAACCTTAGAGTGCTTCGA
GTTCAAGAAAGGTGTCTCAAATTTTGGACCAAATATTCTTCTTCTTTTGAGCTTTGTTTGCGAGTTTTGAGGGAAAGGAAAGGAAGGGAAGGGAGATAGGAGGAAAAATCT
TCTTTTGTTTGAGAGTTAAAAAACCAACAAAAAAAAGGAGATGGATGACTTATGTTATAATTTTACTATTTTATTCTTTTTCTCTTTAAATCAATACATTGTGTTATACTTTGTA
ATTTTTCTTGTGATTTATTCAAGATAACTTATCTCACGAAAAACAAATTATTACTATCTATTTTTGAAAAACACTTATTCAAAATAGTTTATTTATAAAAAAAAAACTTATTACGATC
TCTTTTTAAAAAATAACTTATTTAAAACAGCTTATTACGGTATCTTTCTAAAAAAACAACTTATCTAATGAAAAACAAATTATTACGATCTCTTTTTAAAAAACAGCTTATTACGA
TAAGTTGTTTTCAATATTTTTTTTTAAAAAAGAGATCGTAATAAAAAAAAAATTTGAAAAAAAAAAATCGTAATAAGTTGTTTTGTATAAATAAACTGTTTTTTTTTTTATAAAAAAAAA
AAGTGATACAATAATGATAAATTAATTTCACTATATATACAAAATTATATACAATAATTGTAATTATATTTCCTTTTATTTCATAATAATTATTTTTGGGTACAAATTTCATTAATAAT
TTAAAAAAATAAAAGAATTGAAATTTTAAATTATTGTAAGGGTAAAATTGTAAATTTGTATCTTTAACCCTCAAAATCCATCAATTTTGGAGAACTTTAAAGTGGACAAATAAAT
CCCTTTAAAACCGCCAAAACTTTTCCCTCTTTTTTTAAAACTCTAAAACAAAGGATTACCTCTCCATAAGCCTTTGATTCCCCCGCTTTCCCTCCAAAACCCAACTCACAAAG
CCTAAAAAACAACTATTTTTTTTTTTCATTTTCAAATCATTAATAATTTTTTTAGATATATGTTAACATTATCCATAAAAAAATCCAAATTACCGGTCAAAATTAATGCAATCAGGG
TGATATGTGAGGACAGACTTTGAAGATCATGGAGTCAAGATTGAACAAACAAGTTGAAATTGAAGGTTGACAATAGGTCCCAACAAGATTCTGTGATCTGTCACCGTTCTCTT
TCACTCTTGTTAATGCCATCTCCACCAACTACATTAGATAACGGAATCAAACCAATAAATTGAGATAAAATGAAATAATTAATGATCTTCATTAAAAATTAGTTCAATAAAATGT
TCTTTTATTAAACTTTGGATCAAACTTTAAATTACTAGACTAGATTTTCTTTTTCTGTGATAACCGAAACATCAGGTTATCTAAAAATAAAAATAATATCTCACATCTTCTCTGTC
AGAAACAACACTCACTCAATCATTTATTCATTTTTGTTAATTAAAACCTCCTTTCTTTCTTCTTCCACTTCATAATAATCACCGCTTCAAGATCTTAAATTATTAAACATTATCCC
CAAATCACTCATAG

SCL14 (CDS)

ATGGATCACACCTTTGGATTTATGGACAAACCTTTTCTTCAACCTGATCCAAACAATGCTACAGAAGAAGAAGAAGATTCACCTTTGGATGAAATTGATTTCTCTTCCACTGT
TCTTCGTTACATAAACCAAATGCTTATGGAAG(A->G)AGACATGGATATGAAACCATGTATGTTTCATGATTCATTGGCTTTACAAGCTGCTGAAAAATCTTTCTATGAAGTCAT
TGGTCAGACTTACCCTTCTTCTTCTGTACAAAGCCCTGATGATAGTTTTAACAGTAATTTCAGTAGTAGTAATTCTGTTGAAGTTGAATCATATTTGAGTAAATTTGATGTTTCTG
AGCATAATCCTTCTATATTGAAAACCCATTTTCCTGTTGATTTTGTTTTTCAAGCTAGTTCAATCAATGGTAGTAGTAATAGCAGTAGCGATTTCGGTGTAACCGGTGGTGGATT
TAGGGATGGATTTCCTGATTCAAATTTGATTAGTAAGAATGAGTCTGTATTGCAATTTGAGAGGGGTGTGGAGGAAGCTAATAAGTTTCTACCTAAAGTGAATCCATTTATTAT
TGATTTGGAGAGTAACCCTTTTGTTCCTTCTTTTAGAAGGGAGGTTCAGGTTCAGGAGGTAGTGGTTGAGAGGGAGAGCGATGAAAGGGAACATTTTTCAGCTGAGTCGAG
GGGGAGGAAGAATCATGAGCGGGGAGATGAATTGGGTTTACAAGATGAGAGAAGTAACAAGCAGTCAGCTTTTTATACTGATGATGACAGTGAGCTATCCGAGCTGTTTGA
TAAGGTGCTGCTTGGAACGGGGTGCAGAAAGGAAGAAAAGGCTAATGGGTCGGATATGAGTGTGCAGCAGAAAGAAGAAGCAAATAATTCTGGTGGTGGAAAAAGCCGT
GGTAAGAAACAAGGCAACAAGAAGGGTGTTGTGGATTTGAGGACAATGTTGGTTCTGTGTGCTCAATATGTTTCTTCTGATGATCGTGCAAATGCTAATGAGTTATTGAGGC
AGATAAGGCAGTATTCTTCTCCACTGGGTGATGGATCTCAGAGGCTGGCACATTGTTTTGCGAATGCTCTTGAAGCACGCATGGCTGGCACTGGAACTCAGATTTATACTGC
TTTATACTCCAAAAGAAACTCTGCTGCCGACATGGTGAAAGCGTACCAAATGTATATTTCAGCCTGTCCATTCAAGAAGCTTGCAATCATTTTCGCGAACCATACAATTCTGA
ATCTAGCCAAGGAAGTAGAAACTCTTCACATTGTAGATTTTGGCATCCGTTACGGCTTTCAATGGCCTGCCCTTATTTATCGTTTATCAAAGCGACCTGGGGGGCCTCCTAAG
CTGCGCTTGACAGGGATAGAACTCCCCCAACCTGGCTTCAGGCCTGCAGAGAGAGTCCAGGAGACAGGACTTCGCCTGGCAAGGTATTGTGAACGTTTTAATGTTCCGTT
TGAGTTCAATGCTATTGCACAGAAATGGGAAACCATCAAAGTTGAGGACTTGAAGATAAAGAAAAATGAACTTCTTGTGGTGAATTCTGTTTGTCGGCTTAAAAATTTACTTG
ATGAGACAGTTGTGTTGAATAGTCCCAGGGATGCTGTTTTGAAATTAATTAGGGACACAAATCCCAATATCTTATACATACTACTGTCAATGGGAGCTACAATGCCCCATTCT
TTGCGACACGGTTCAAGGAGGCTCTTTTCAATTACTCTACAATGTTTGATGTGCTTGACATCAATGTTGCTCGTGAAGATCAAACGAGGTTGATGTTTGAGAAAGAGTTCTGG
GGGAGGGAAGTAATGAATATTATAGCCTGTGAAGGTTCTCAGAGGGTTGAGAGGCCAGAGACATACAGGAAATGGCAGGTTCGGAATACAAGGGCTGGATTTAGGCATCT
GCCCTTGGATAAACATCTCATTAACAAGTTAAGATGCAAGTTGAAAGACGTGTATCATAGTGATTTCATGCTTGTTGAAGATGGAAACTGCATGCTGCAAGGATGGAAGGGC
CGTATAATTTACGCTTCCTCCTGCTGGGTACCTGCATAG Glu → Gly E49G

tSCL14 (terminator) 500 bp

GACTTGGCAATGATCTGCTCAAGGAATCTGCATAGGGCTATGGGTGACATATTCTAGAAGCTGACCCCTGCCAATGGAATGGTATAGAACAGTTCAGTTTCAGTAGCCATAAT
ATTTTTATTTGGTTAAGTTTGTGTATGATATAACTAATGCTATGTAAATGACGTATTTAGATGCATATTAGCTGAACTTTTCTTTTGTTGGGTAGTCGCTCATAAGGATGCAGCAG
AATCGTGAGTACTTCCCCGTCTCACAATAAGTGTTCTCTTTATTATTTTTTGTTTTGTCTTAAAATATTTGTCCCTTTTACAATACCAATACAACATTTATTATTCTTTTCTACTAA
CACACTCTTATTTATTGTATTTCAATCACCTAACTACTCCACTAACTATAATTAATAATGATGTTTTAGTAAATTAAACATATTTTACCATTAAAATCAACACTACTAATCATTTCC
TTAATAACTGCGTAAAACTCACAACACTCACAACAG

Additional Sequences used in this study

mCherry (CDS)

ATGGTGAGCAAGGGCGAGGAGGATAACATGGCCATCATCAAGGAGTTCATGCGCTTCAAGGTGCACATGGAGGGCTCCGTGAACGGCCACGAGTTCGAGATCGAGGGCG
AGGGCGAGGGCCGCCCCTACGAGGGCACCCAGACCGCCAAGCTGAAGGTGACCAAGGGTGGCCCCCTGCCCTTCGCCTGGGACATCCTGTCCCCTCAGTTCATGTACG
GCTCCAAGGCCTACGTGAAGCACCCCGCCGACATCCCCGACTACTTGAAGCTGTCCTTCCCCGAGGGCTTCAAGTGGGAGCGCGTGATGAACTTCGAGGACGGCGGCGT
GGTGACCGTGACCCAGGACTCCTCCCTGCAGGACGGCGAGTTCATCTACAAGGTGAAGCTGCGCGGCACCAACTTCCCCTCCGACGGCCCAGTAATGCAGAAGAAAACC
ATGGGCTGGGAGGCCTCCTCCGAGCGGATGTACCCCGAGGACGGCGCCCTGAAGGGCGAGATCAAGCAGAGGCTGAAGCTGAAGGACGGCGGCCACTACGACGCTGA
GGTCAAGACCACCTACAAGGCCAAGAAGCCCGTGCAGCTGCCCGGCGCCTACAACGTCAACATCAAGTTGGACATCACCTCCCACAACGAGGACTACACCATCGTGGAA
CAGTACGAACGCGCCGAGGGCCGCCACTCCACCGGCGGCATGGACGAGCTGTACAAGT

Supplemental Table S6 (continued)|

Additional Sequences used in this study

GUS (CDS)

ATGGGTCAGTCCCTTATGTTACGTCCTGTAGAAACCCCAACCCGTGAAATCAAAAAACTCGACGGCCTGTGGGCATTCAGTCTGGATCGCGAAAACTGTGGAATTGATCAG
CGTTGGTGGGAAAGCGCGTTACAAGAAAGCCGGGCAATTGCTGTGCCAGGCAGTTTTAACGATCAGTTCGCCGATGCAGATATTCGTAATTATGCGGGCAACGTCTGGTAT
CAGCGCGAAGTCTTTATACCGAAAGGTAAGTCTTACTCTCTCTTTTTTGGTCTGTATTTTTAATTTTTTGAAGTATACTATTTGTACTGACGCTAATAATCTTTTTTCAGGTTGGG
CAGGCCAGCGTATCGTGCTGCGTTTCGATGCGGTCACTCATTACGGCAAAGTGTGGGTCAATAATCAGGAAGTGATGGAGCATCAGGGCGGCTATACGCCATTTGAAGCCG
ATGTCACGCCGTATGTTATTGCCGGGAAAAGTGTACGTATCACCGTTTGTGTGAACAACGAACTGAACTGGCAGACTATCCCGCCGGGAATGGTGATTACCGACGAAAACG
GCAAGAAAAAGCAGTCTTACTTCCATGATTTCTTTAACTATGCCGGAATCCATCGCAGCGTAATGCTCTACACCACGCCGAACACCTGGGTGGACGATATCACCGTGGTGAC
GCATGTCGCGCAAGACTGTAACCACGCGTCTGTTGACTGGCAGGTACTTCATGCTTCAACGTGTAACTTAAGAGATACTGTGTGAAATTTTATATTTCCATACATTTGCTTGA
CCTTTGCTTTTTGTCAATTTTTTTCCCCTTACAGGTGGTGGCCAATGGTGATGTCAGCGTTGAACTGCGTGATGCGGATCAACAGGTGGTTGCAACTGGACAAGGCACTAGC
GGGACTTTGCAAGTGGTGAATCCGCACCTCTGGCAACCGGGTGAAGGTTATCTCTATGAACTGTGCGTCACAGCCAAAAGCCAGACAGAGTGTGATATCTACCCGCTTCGC
GTCGGCATCCGGTCAGTGGCAGTGAAGGGCGAACAGTTCCTGATTAACCACAAACCGTTCTACTTTACTGGCTTTGGTCGTCATGAAGATGCGGACTTGCGTGGCAAAGG
ATTCGATAACGTGCTGATGGTGCACGACCACGCATTAATGGACTGGATTGGGGCCAACTCCTACCGTACCTCGCATTACCCTTACGCTGAAGAGATGCTCGACTGGGCAGA
TGAACATGGCATCGTGGTGATTGATGAAACTGCTGCTGTCGGCTTTAACCTCTCTTTAGGCATTGGTTTCGAAGCGGGCAACAAGCCGAAAGAACTGTACAGCGAAGAGGC
AGTCAACGGGGAAACTCAGCAAGCGCACTTACAGGCGATTAAAGAGCTGATAGCGCGTGACAAAAACCACCCAAGCGTGGTGATGTGGAGTATTGCCAACGAACCGGAT
ACCCGTCCGCAAGGTGCACGGGAATATTTCGCGCCACTGGCGGAAGCAACGCGTAAACTCGACCCGACGCGTCCGATCACCTGCGTCAATGTAATGTTCTGCGACGCTCA
CACCGATACCATCAGCGATCTCTTTGATGTGCTGTGCCTGAACCGTTATTACGGATGGTATGTCCAAAGCGGCGATTTGGAAACGGCAGAGAAGGTACTGGAAAAAGAACT
TCTGGCCTGGCAGGAGAAACTGCATCAGCCGATTATCATCACCGAATACGGCGTGGATACGTTAGCCGGGCTGCACTCAATGTACACCGACATGTGGAGTGAAGAGTATCA
GTGTGCATGGCTGGATATGTATCACCGCGTCTTTGATCGCGTCAGCGCCGTCGTCGGTGAACAGGTATGGAATTTCGCCGATTTTGCGACCTCGCAAGGCATATTGCGCGTT
GGCGGTAACAAGAAAGGGATCTTCACTCGCGACCGCAAACCGAAGTCGGCGGCTTTTCTGCTGCAAAAACGCTGGACTGGCATGAACTTCGGTGAAAAACCGCAGCAGG
GAGGCAAACAATGA

35Spro (promoter short)

GGAGGTCAACATGGTGGAGCACGACACTCTGGTCTACTCCAAAAATGTCAAAGATACAGTCTCAGAAGATCAAAGGGCTATTGAGACTTTTCAACAAAGGATAATTTCGGG
AAACCTCCTCGGATTCCATTGCCCAGCTATCTGTCACTTCATCGAAAGGACAGTAGAAAAGGAAGGTGGCTCCTACAAATGCCATCATTGCGATAAAGGAAAGGCTATCATT
CAAGATCTCTCTGCCGACAGTGGTCCCAAAGATGGACCCCCACCCACGAGGAGCATCGTGGAAAAAGAAGAGGTTCCAACCACGTCTACAAAGCAAGTGGATTGATGTGA
CATCTCCACTGACGTAAGGGATGACGCACAATCCCACTATCCTTCGCAAGACCCTTCCTCTATATAAGGAAGTTCATTTCATTTGGAGAGGACACGCTCGAGTATAAGAGCT
CATTTTTACAACAATTACCAACAACAACAAACAACAAACAACATTACAATTACATTTACAATTATCGATACAGA

t35S (terminator)

GCTTCTCTAGCTAGAGTCGATCGACAAGCTCGAGTTTCTCCATAATAATGTGTGAGTAGTTCCCAGATAAGGGAATTAGGGTTCCTATAGGGTTTCGCTCATGTGTTGAGCAT
ATAAGAAACCCTTAGTATGTATTTGTATTTGTAAAATACTTCTATCAATAAAATTTCTAATTCCTAAAACCAAAATCCAGTACTAAAATCCAGAT

pNOS (promoter)

AGCGGAGAATTAAGGGAGTCACGTTATGACCCCCGCCGATGACGCGGGACAAGCCGTTTTACGTTTGGAACTGACAGAACCGCAACGTTGAAGGAGCCACTCAGCCGCG
GGTTTCTGGAGTTTAATGAGCTAAGCACATACGTCAGAAACCATTATTGCGCGTTCAAAAGTCGCCTAAGGTCACTATCAGCTAGCAAATATTTCTTGTCAAAAATGCTCCAC
TGACGTTCCATAAATTCCCCTCGGTATCCAATTAGAGTCTCATATTCACTCTCAAATG

NPTII (Neomycin Phosphotransferase II, Kanamycin resistance) (CDS)

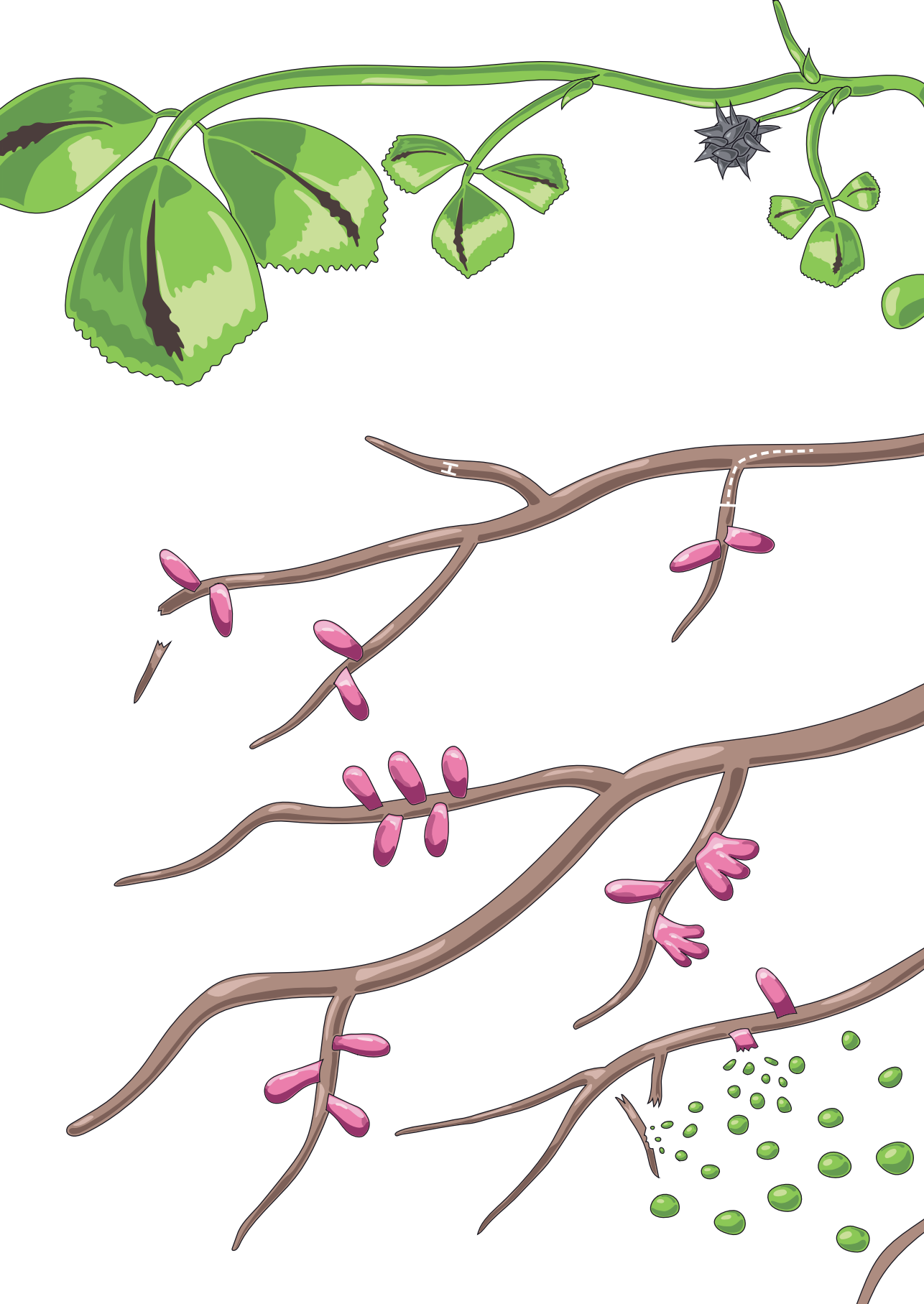
ATGATTGAACAAGATGGATTGCACGCAGGTTCTCCGGCCGCTTGGGTGGAGAGGCTATTCGGCTATGACTGGGCACAACAGACAATCGGCTGCTCTGATGCCGCCGTGTTC
CGGCTGTCAGCGCAGGGGCGCCCGGTTCTTTTTGTCAAGACCGACCTGTCCGGTGCCCTGAATGAACTCCAAGACGAGGCAGCGCGGCTATCGTGGCTGGCCACGACGG
GCGTTCCTTGCGCAGCTGTGCTCGACGTTGTCACTGAAGCGGGAAGGGACTGGCTGCTATTGGGCGAAGTGCCGGGGCAGGATCTCCTGTCATCTCACCTTGCTCCTGCC
GAGAAAGTATCCATCATGGCTGATGCAATGCGGCGGCTGCATACGCTTGATCCGGCTACCTGCCCATTCGACCACCAAGCGAAACATCGCATCGAGCGAGGACGTACTCG
GATGGAAGCCGGTCTTGTCGATCAGGATGATCTGGACGAAGAGCATCAGGGGCTCGCGCCAGCCGAACTGTTCGCCAGGCTCAAGGCGCGGATGCCCGACGGCGAGGA
TCTCGTCGTGACCCAGGGCGATGCCTGCTTGCCGAATATCATGGTGGAAAATGGCCGCTTTTCTGGATTCATCGACTGTGGCCGGCTGGGTGTGGCGGACCGCTATCAGGA
CATAGCGTTGGCTACCCGTGATATTGCTGAAGAGCTTGGCGGCGAATGGGCTGACCGCTTCCTCGTGCTTTACGGTATCGCCGCTCCCGATTCGCAGCGCATCGCCTTCTA
TCGCCTTCTTGACGAGTTCTTCTGA

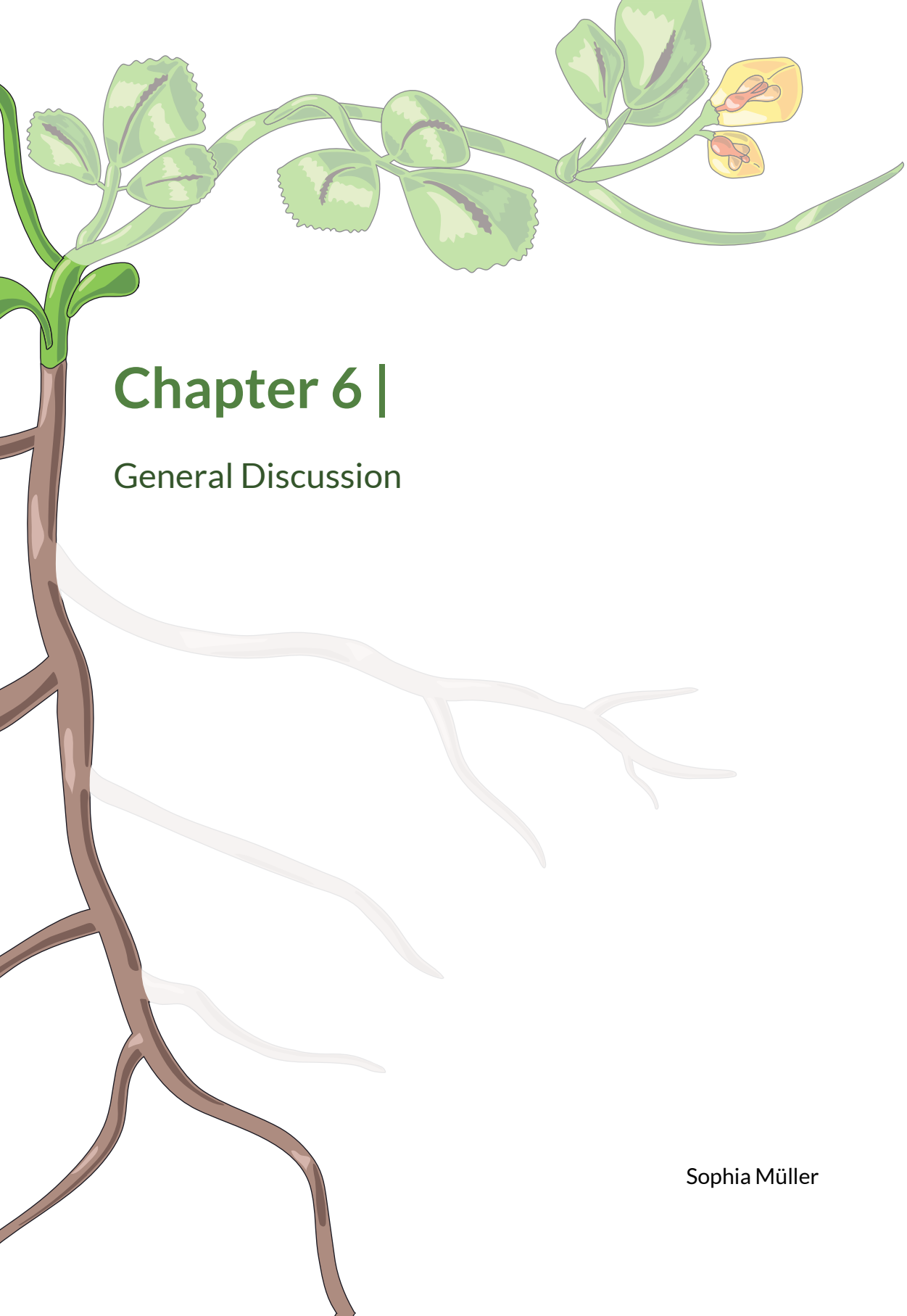
tNOS (terminator)

GCTTGTCAAGCAGATCGTTCAAACATTTGGCAATAAAGTTTCTTAAGATTGAATCCTGTTGCCGGTCTTGCGATGATTATCATATAATTTCTGTTGAATTACGTTAAGCATGTAA
TAATTAACATGTAATGCATGACGTTATTTATGAGATGGGTTTTTATGATTAGAGTCCCGCAATTATACATTTAATACGCGATAGAAAACAAAATATAGCGCGCAAACTAGGATA
AATTATCGCGCGCGGTGTCATCTATGTTACTAGATCGA

pAtUBQ10 (promoter)

GGAGCGACGAGTCAGTAATAAACGGCGTCAAAGTGGTTGCAGCCGGCACACACGAGTCGTGTTTATCAACTCAAAGCACAAATACTTTTCCTCAACCTAAAAATAAGGCAA
TTAGCCAAAAACAACTTTGCGTGTAAACAACGCTCAATACACGTGTCATTTTATTATTAGCTATTGCTTCACCGCCTTAGCTTTCTCGTGACCTAGTCGTCCTCGTCTTTTCTT
CTTCTTCTTCTATAAAACAATACCCAAAGAGCTCTTCTTCTTCACAATTCAGATTTCAATTTCTCAAAATCTTAAAAACTTTCTCTCAATTCTCTCTACCGTGATCAAGGTAAATT
TCTGTGTTCCTTATTCTCTCAAAATCTTCGATTTTGTTTTCGTTCGATCCCAATTTCGTATATGTTCTTTGGTTTAGATTCTGTTAATCTTAGATCGAAGTCGATTTTCTGGGTTTG
ATCGTTAGATATCATCTTAATTCTCGATTAGGGTTTCATAGATATCATCCGATTTGTTCAAATAATTTGAGTTTTGTCGAATAATTACTCTTCGATTTGTGATTTCTATCTAGATCT
GGTGTTAGTTTCTAGTTTGTGCGATCGAATTTGTAGATTAATCTGAGTTTTTCTGATTAACA





Chapter 6 |

General Discussion

Sophia Müller

Nodulation as a constrained developmental innovation

Symbiotic nodulation represents a highly specialized developmental innovation that enables plants to access atmospheric dinitrogen through association with rhizobia (Wang et al., 2022). While often portrayed as a broadly advantageous trait, its phylogenetic distribution reveals a far more constrained evolutionary trajectory (Griesmann et al., 2018; van Velzen et al., 2018). Nodulation likely originated 110 million years ago in the *Rosids* clade but was subsequently lost in most derived lineages, persisting mainly in Fabales and in a limited number of species belonging to the *Rosales*, *Cucurbitales*, and *Fagales* (Soltis et al., 1995). This repeated loss strongly suggests that nodulation imposes substantial developmental and ecological constraints that outweigh its benefits under many environmental conditions (Van Velzen et al., 2019). Understanding nodulation therefore requires more than identifying its molecular components; it demands insight into how these components are spatially and temporally restricted to ensure that symbiosis remains productive rather than deleterious. In this sense, nodulation emerges not as a permissive gain-of-function trait, but as a tightly constrained developmental program.

Nodulation and lateral root development: shared modules, distinct outcomes

In legumes such as *Medicago truncatula* (Medicago), nodulation depends on the formation of a new root organ which development is tightly controlled (Roy et al., 2020). Although there are similarities between lateral root and nodule development (Hirsch et al., 1997; Mathesius et al., 2000; Schiessl and Jhu, 2025; Schiessl et al., 2019; Soyano et al., 2019), there remain striking differences. Nodule initiation is confined to a narrow susceptible zone located several millimeters behind the root tip, where root hairs are elongated but not yet fully mature. This spatial restriction contrasts sharply with lateral root formation, which can occur along much of the root axis (Santos Teixeira and ten Tusscher, 2019). The existence of this susceptible zone suggests that nodulation is not simply an alternative outcome of a generic organogenic program but rather a carefully constrained developmental response that integrates hormonal signaling, tissue maturity, and microbial cues (Lin et al., 2020; Roy et al., 2020).

At the developmental level, nodules and lateral roots share early mechanistic features, including the activation of pericycle cells and the establishment of an auxin maximum preceding cell-cycle reentry (Xiao et al., 2014). However, the two organs diverge substantially in their tissue contributions, growth, and further development. In Arabidopsis, lateral roots arise almost exclusively from pericycle-derived divisions and establish an autonomous meristem (Santos Teixeira and ten Tusscher, 2019). In contrast, nodules require coordinated proliferation across multiple tissues. In Medicago, the pericycle contributes to early divisions and vascular connections, but the bulk of the organ derives from inner cortex layers, with the C3 layer giving rise to the future nodule meristem (Xiao et al., 2014). This multi-tissue origin increases the regulatory complexity and necessitates precise spatial coordination.

Hormonal rewiring enables nodule-specific cell activation

Auxin, cytokinin, and ethylene form the core hormonal network regulating nodule initiation. As introduced in **Chapter 1**, I refer to this triad as the "ACE" of nodulation. Interestingly, initiation of both lateral roots and nodules requires auxin accumulation in the pericycle, yet cytokinin plays opposing roles in the two processes (Roy et al., 2020; Santos Teixeira and ten Tusscher, 2019; Schiessl et al., 2019; Xiao et al., 2014). In lateral root development, cytokinin antagonizes auxin biosynthesis and inhibits founder-cell activation, whereas in nodulation cytokinin is a positive regulator and is sufficient to induce pseudonodules (Cooper and Long, 1994; Gauthier-Coles et al., 2019; Laplaze et al., 2008). This apparent contradiction is likely the result of the recruitment of the nodulation-specific transcription factor NODULE INCEPTION (NIN), which activates LATERAL ORGAN BOUNDARIES DOMAIN 16 (LBD16) expression in the pericycle in a cytokinin-dependent manner (Liu et al., 2019; Schiessl et al., 2019). Through this regulatory innovation, nodulation uses the auxin-driven developmental module, while circumventing the inhibitory effects of cytokinin. This example illustrates that nodulation is not achieved by inventing entirely new pathways but by rewiring existing ones to operate under different regulatory constraints (Lee et al., 2024; Schiessl and Jhu, 2025; Schiessl et al., 2019).

Although the importance of auxin accumulation during nodulation has long been recognized, the mechanisms underlying its spatial control were poorly understood. In **Chapter 2**, I disentangled the contributions of auxin transport and biosynthesis and demonstrated that local auxin biosynthesis is a critical driver of early nodule initiation (**Figure 1**). While long-distance auxin transport from the shoot contributes to overall auxin availability, neither acropetal nor basipetal auxin flow is strictly required to establish the auxin maximum that precedes pericycle divisions. Instead, induction of the auxin biosynthesis genes *YUCCAs MtYUC1*, *MtYUC2*, and *MtYUC8* in vascular and pericycle-associated tissues generates a localized auxin maximum that triggers cell-cycle reentry (**Figure 1B**). These findings position local auxin biosynthesis, rather than transport alone, as a critical regulatory step in nodule initiation.

The spatial restriction of auxin accumulation is particularly striking when considered alongside the negative regulator ethylene. Ethylene is known for its negative and growth inhibitory effects during plant development (Bleecker et al., 1988; Lee and LaRue, 1992; Merchante and Stepanova, 2017; Penmetsa and Cook, 1997). Additionally, ethylene is closely linked to developmental timing processes, acting as an internal signal that promotes tissue maturation in processes such as root hair differentiation, fruit ripening, and the onset of senescence (Khan et al., 2014; Liu et al., 2015; Pitts et al., 1998). Although ethylene has long been known to inhibit nodulation (Penmetsa and Cook, 1997; Reid et al., 2018), bulk transcriptomic studies consistently report induction of ethylene biosynthesis genes upon Nod factor or rhizobial treatment (Breakspear et al., 2014; Larrainzar et al., 2015; Reid et al., 2018; Schiessl et al., 2019; van Zeijl et al., 2015). This apparent paradox is explored in **Chapter 3**, where I demonstrated that ethylene regulation during nodulation is fundamentally spatiotemporally regulated. Prior to nodulation, ethylene biosynthesis and signaling are detectable in inner root tissues, including

the vasculature, pericycle, and endodermis, largely driven by *Medicago truncatula ACC SYNTHASE 10* (*MtACS10*) expression (**Figure 1A**). Upon Nod factor perception, ethylene biosynthesis is actively redistributed (**Figure 1B**). Expression of *MtACS10* is downregulated in inner tissues, while *MtACS3* is induced in the epidermis and outer cortex layers (**Figure 1B**). As a result, ethylene production is reduced precisely where cell divisions must occur and increased where infection should be restrained.

This spatial reorganization resolves the apparent contradiction between increased ethylene production and its inhibitory role (Penmetsa and Cook, 1997; Reid et al., 2018). Ethylene is not fully suppressed during nodulation; instead, it is repositioned to fulfill distinct functions in different tissues (**Chapter 3**). In the epidermis, ethylene interferes with Nod factor signaling and limits excessive or ectopic infection. In the inner tissues, the reduction of ethylene biosynthesis relieves inhibition of cell-cycle reentry, enabling nodule primordium formation. Nodulation thus depends on the precise spatial separation of positive and negative hormonal inputs rather than on uniform changes of hormone levels in whole tissues.

A dynamic model of auxin and ethylene coordination during early nodulation

Integrating the findings from **Chapters 2** and **3** leads to a spatiotemporal model in which auxin and ethylene form a coordinated regulatory circuit governing nodule initiation. Before rhizobial signaling, auxin is primarily transported through the vasculature (Grunewald et al., 2009), contributing to root growth and patterning, while ethylene signaling is detectable at low levels in inner tissues (**Figure 1A**). This baseline state likely reflects a balance between developmental competence and maturation, ensuring that only tissues within the susceptible zone remain responsive to symbiotic cues.

Following Nod factor perception, local auxin biosynthesis is induced in the pericycle, generating an auxin maximum that precedes the first cell divisions (**Figure 1B**). At the same time, ethylene biosynthesis is actively suppressed in these inner tissues through downregulation of *MtACS10* (**Figure 1B**). It remains to be determined if, and how, these changes in auxin and ethylene levels are linked to each other. Nevertheless, this dual change in the inner tissues likely creates a permissive environment for pericycle cells to be re-activated. Like ethylene, auxin accumulation is not stationary; once pericycle divisions are initiated, auxin spreads to adjacent tissues, including the endodermis and inner cortex, again preceding cell divisions (**Figure 1C**). Importantly, this spread is possibly mediated by diffusion or auxin transporters such as the PIN-FORMED MtPIN4 and MtPIN10 rather than by local biosynthesis within the dividing cortex cells themselves. This separation between biosynthesis and response highlights the importance of directional auxin transport in shaping the developing nodule primordium.

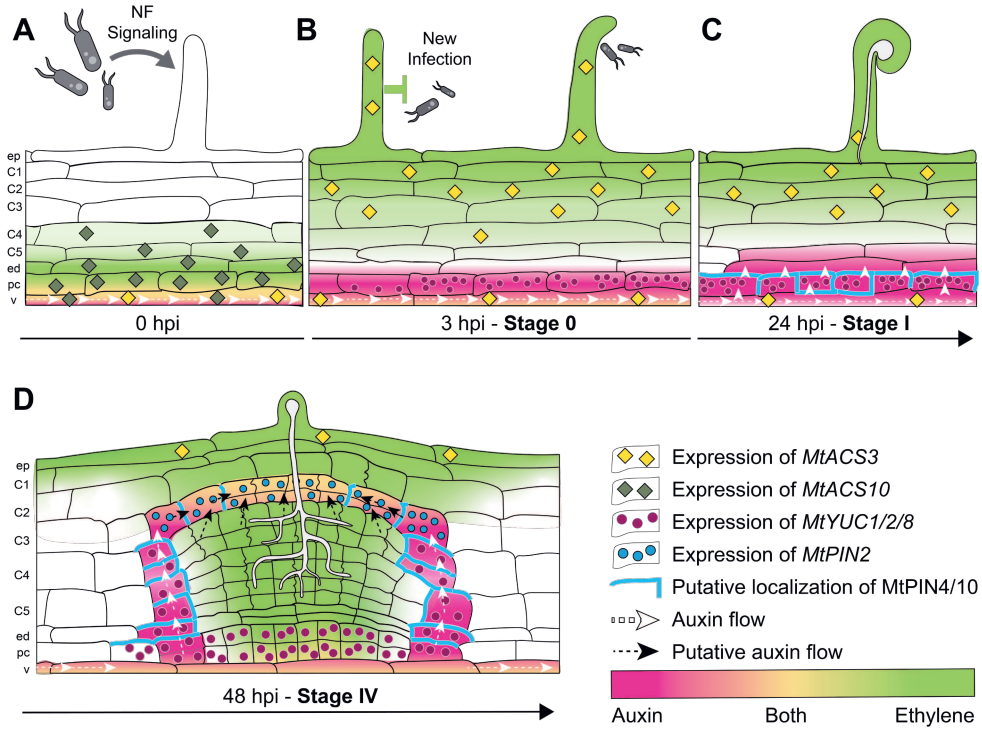


Figure 1| Spatiotemporal dynamics of auxin and ethylene during early nodulation. (A) Prior to rhizobial induction (0 hours post inoculation, hpi), auxin (pink) is predominantly transported through the root vasculature, while ethylene (green) is synthesized in inner root tissues, including the vasculature, pericycle, endodermis, and cortex layers C4 and C5, largely via the ethylene biosynthesis gene *MtACS10* (green squares). **(B)** Upon perception of rhizobial Nod factors (NF; grey), local auxin biosynthesis mediated by *MtYUC1*, *MtYUC2*, and *MtYUC8* (dark pink circles) establishes an auxin maximum in the pericycle that precedes the first cell divisions. Concurrently, ethylene biosynthesis is spatially redistributed from inner tissues to the outer root layers, including the epidermis and cortex layers C1–C4, through activation of *MtACS3* (yellow squares) and repression of *MtACS10*. Epidermal ethylene restricts excessive rhizobial infection. **(C)** At 24 hpi, pericycle cell divisions are initiated, and auxin is transported into the endodermis and cortex layers ahead of subsequent divisions, primarily via the auxin exporters MtPIN4 and MtPIN10 (turquoise). **(D)** By 48 hpi, a distinct nodule primordium is visible. Auxin biosynthesis persists in the pericycle and becomes enriched at the periphery of the developing nodule, where future vascular tissues will form. Auxin is transported toward the C3-derived nodule meristem (white arrows), potentially guided by additional expression of *MtPIN2* (blue circles), which may direct auxin into the meristematic zone (black arrow). Ethylene signaling remains active in the epidermis, likely through continued *MtACS3* expression, and is also detectable within the developing nodule primordium Abbreviations: ep, epidermis; C1–C5, cortex layers 1–5; ed, endodermis; pc, pericycle; v, vasculature.

As the nodule primordium becomes morphologically distinct, auxin dynamics change again. Auxin biosynthesis remains associated with the pericycle and developing vascular domains, while auxin is transported toward the periphery of the primordium, where future nodule vasculature will form (**Figure 1D**). This pattern closely resembles auxin canalization during leaf vascular development and suggests that auxin gradients guide the establishment of vascular connections between the root stele and the developing nodule (Biedroń and Banasiak, 2018; Cheng et al., 2006). Notably, auxin is located in the future nodule meristem, which derives from the C3 cortex layer, but auxin biosynthesis genes are largely absent (**Figure 1D**). This

observation raises the possibility that, unlike lateral root meristems, the nodule meristem functions as an auxin sink that remains dependent on imported auxin (Du and Scheres, 2018; Ljung et al., 2005). Such a dependence would represent a fundamental developmental distinction between nodules and lateral roots and may contribute to the unique growth dynamics of nodules.

Ethylene signaling, although reduced in inner tissues during the earliest stages of nodule initiation (**Figure 1B-C)**, reappears in the developing primordium at later stages (**Figure 1D**). At first glance, the presence of an inhibitory hormone within a proliferating organ appears contradictory. However, genetic evidence argues against a simple permissive role for ethylene in nodule meristem formation. Ethylene-insensitive mutants such as *sickle/Mtein2* display hyperinfection and altered nodule numbers but do not show a complete loss of meristematic activity (Penmetsa and Cook, 1997; Penmetsa et al., 2003; Penmetsa et al., 2008). Instead, these phenotypes suggest that ethylene acts as a modulatory signal that constrains proliferation and enforces developmental boundaries. Ethylene may therefore function to prevent excessive or disorganized growth within the primordium and to coordinate organogenesis with infection progression.

Ethylene has also been implicated in nodule positioning, with classical models proposing that localized ethylene production provides positional information that restricts nodules to sites opposite xylem poles (Heidstra et al., 1997; Penmetsa et al., 2003; Penmetsa et al., 2008). Although we detected only weak and transient expression of *MtACS10* near developing nodules, the observation of activity of the ethylene reporter within the primordium itself suggests that positioning may be achieved through more indirect mechanisms. One possibility is that ethylene acts as a maturation timer for epidermal cells and root hairs, thereby defining the spatial window of infection competence. I postulate that *MtACS3* expression at the transition zone where root hairs emerge could accelerate maturation in specific domains, limiting future infections and indirectly shaping nodule spacing. This hypothesis aligns with the broader role of ethylene in tissue maturation and senescence and offers an alternative explanation for nodule positioning that does not rely solely on local ethylene biosynthesis adjacent to the forming primordium.

Spatial regulation explains apparent contradictions in bulk transcriptomics

A recurring theme throughout this thesis is that bulk transcriptomic data can obscure, or even invert, biologically meaningful regulatory patterns. The simultaneous induction and repression of functionally similar genes, particularly within biosynthetic gene families such as ACC SYNTHASES (*ACS*) or *YUCCAs* (YUCs), often appears contradictory when averaged across tissues. In **Chapters 2** and **3** we demonstrated that such patterns frequently arise because different isoforms operate in different tissues and could be regulated in opposite directions. Without spatial resolution, these opposing trends cancel each other out or produce misleading averages.

The differential regulation of *MtACS3* and *MtACS10* provides a particularly clear example. Bulk RNA-seq datasets report upregulation of ethylene biosynthesis genes during early nodulation, a result that seems incompatible with the inhibitory role of ethylene. By resolving expression at the tissue level, it becomes clear that ethylene biosynthesis is indeed reduced where nodules must form and increased where infection must be restrained. This insight would have been inaccessible without distinguishing between isoforms and mapping their expression spatially.

Similar considerations apply to auxin metabolism and transport genes. Bulk datasets frequently show concurrent induction of auxin biosynthesis and auxin inactivation pathways, raising questions about how auxin levels can increase under such conditions (Schiessl et al., 2019; van Zeijl et al., 2015). The spatial analysis I presented in my thesis reveals that these pathways are often activated in different tissues or at different developmental stages, allowing the plant to fine-tune auxin distribution rather than simply increasing or decreasing overall levels. These observations underscore the necessity of integrating spatial information into transcriptomic analyses of highly organized and complex spatiotemporal developmental processes such as during nodulation.

Single-cell transcriptomics as a tool to resolve nodulation complexity

The need to move beyond bulk tissue analyses motivated the development of a single-cell RNA sequencing dataset in **Chapter 4**. When I initiated this work, no single-cell datasets for nodulation in Medicago were available. Since then, several single-nucleus RNA sequencing studies have been published, highlighting the need and potential value of cell-resolved approaches for studying symbiotic development (Cervantes-Pérez et al., 2022; Liu et al., 2023; Pereira et al., 2024). The work I presented here complements these studies by providing a direct comparison between protoplast-based scRNA-seq and nucleus-based snRNA-seq approaches.

Protoplast-derived scRNA-seq offers clear advantages in terms of transcriptome complexity and gene detection, reflecting the inclusion of cytoplasmic RNA. This increased sensitivity proved valuable for identifying cell-type-specific expression patterns and for detecting induction of nodulation-related genes. However, protoplast isolation also introduced significant challenges. In more mature tissues, particularly within the susceptible zone, certain cell types such as the endodermis and epidermis were underrepresented. The likely causes include the presence of the Casparian strip and suberin lamellae in the endodermis, as well as the fragility and complex morphology of epidermal root hair cells (Geldner, 2013; Shaw et al., 2000; Su et al., 2023; Yan et al., 2025). These limitations are particularly problematic for nodulation studies, where interactions between epidermal infection programs and inner tissue organogenesis are central. Accordingly, the inclusion of enzymes capable of degrading components of the casparian strip represents a plausible strategy to improve endodermal protoplast recovery release (Jabloune et al., 2020; Zhang et al., 2022). however, this assumption remains to be experimentally validated, and additional, currently unidentified factors may also limit endodermis isolation efficiency.

In contrast, single-nucleus RNA sequencing is less biased by tissue maturity and can recover cell types that are difficult to isolate as protoplasts. However, not only in our hands, nuclei-based datasets are strongly affected by ambient mRNA contamination, which reduced cell-type specificity and caused expression profiles to resemble bulk RNA-seq when not thoroughly filtered (Adema et al., 2024). While other groups have published snRNA-seq performed on nodulating tissues, their results likely suffer from similar restrictions. This highlights the sensitivity of nuclei-based approaches to experimental conditions and underscores the need for rigorous optimization and filtering.

Taken together, these observations suggest that no single approach is sufficient on its own. Protoplast-based scRNA-seq provides rich transcriptomic information but incomplete cell-type coverage, while snRNA-seq offers broader coverage at the cost of reduced specificity. Additionally, snRNA-seq would allow to investigate RNA velocity, not possible when using only protoplast (Bergen et al., 2021; Gorin et al., 2022; La Manno et al., 2018). Therefore, combining these approaches, either experimentally or computationally, may provide a more complete picture of nodulation at cellular resolution. In addition, spatial transcriptomic data would likely further improve the usefulness of these data sets (Adema et al., 2024).

Candidate gene discovery in sparse single-cell datasets

In **Chapter 5**, I leveraged the scRNA-seq dataset to identify pericycle-enriched transcription factors that were differentially regulated during nodulation. Although *promoter*–GUS analyses confirmed the predicted spatial expression patterns, functional assays revealed only weak or inconsistent nodulation phenotypes. These results highlight inherent challenges in candidate discovery from single-cell data. An inherent property of scRNA-seq datasets is that the expression data per cell is relatively sparse. Consequently, lowly expressed regulatory genes may be detected in only a fraction of cells. Apparent cell-type specificity can therefore reflect sampling effects rather than robust biological regulation.

Moreover, the pericycle itself is a heterogeneous tissue, and only a small subset of pericycle cells initiates nodulation (Schiessl et al., 2019; Xiao et al., 2014). Genes that are broadly expressed in the pericycle may not be specifically involved in the initiation of a nodule primordium. The weak phenotypes observed in RNA interference and ectopic expression lines suggest that many of the candidates identified may play modulatory rather than essential roles, contributing to robustness, timing, or spatial precision rather than serving as primary drivers of nodule formation. This is further supported by the fact that there are no reports that any of these genes were previously identified in mutant screens.

These findings indicate that stricter selection criteria are necessary when using scRNA-seq data for candidate discovery. In particular, genes that are expressed in the pericycle under control conditions but only modestly induced during nodulation may be less informative than genes that are absent or lowly expressed in uninduced pericycle cells and strongly activated in nodule-initiating subsets. Integrating scRNA-seq with bulk RNA-seq, laser capture microdissection,

spatial transcriptomics and published snRNA-seq datasets can further help identify candidates with consistent, reproducible expression patterns across experiments and platforms.

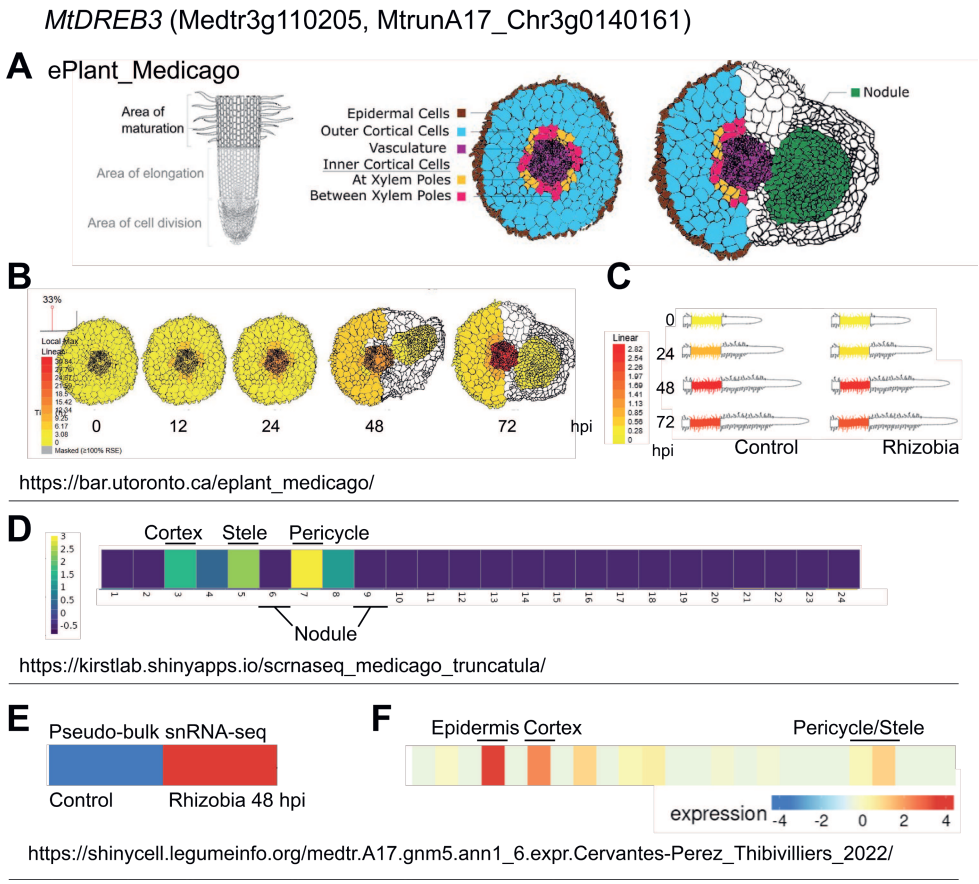


Figure 2| *MtDREB3* expression across independent transcriptomic datasets. (A–C) Visualization of *MtDREB3* expression using the ePlant_Medicago platform. **(A)** Schematic representation of root tissue identities. **(B)** Tissue-resolved expression derived from laser capture microdissection of Medicago roots following rhizobial inoculation, as reported by Schnabel et al. (2023). **(C)** Root zone–specific expression based on bulk RNA-seq analysis, illustrating *MtDREB3* transcript abundance across developmental zones and time points. **(D)** Cell-type–specific expression of *MtDREB3* inferred from single-nucleus RNA-seq (snRNA-seq) data published by Pereira et al. (2024). **(E–F)** snRNA-seq data from Cervantes-Pérez et al. (2022). **(E)** Pseudo-bulk RNA-seq generated by aggregating all cells from rhizobia-treated samples (48 hours post inoculation, hpi) and untreated controls, respectively. **(F)** Tissue- and cell-type–specific expression of *MtDREB3* within the same snRNA-seq dataset.

The importance of data integration and visualization

The analysis of *Medicago truncatula DEHYDRATION-RESPONSIVE ELEMENT BINDING PROTEIN* 3 (*MtDREB3)* across multiple datasets illustrates the value of combining diverse data sources (**Figure 2**). Depending on the dataset and sampling strategy, *MtDREB3* appeared either broadly

induced during root maturation or specifically upregulated in response to rhizobia (**Figure 2**, Cervantes-Pérez et al., 2022; Pereira et al., 2024; Waese et al., 2017). Only by comparing scRNA-seq, bulk RNA-seq, laser microdissection data, snRNA-seq and possible future spatial transcriptomics, could these patterns be resolved and attributed to differences in tissue maturity and sampling regions (**Figure 2**). This example underscores the need for online interfaces, where data is integrated and visualized to allow researchers to compare expression patterns across developmental stages, tissues, and experimental conditions.

Existing tools such as ePlant demonstrate how multi-level datasets can be made accessible for hypothesis generation (Waese et al., 2017). Extending such platforms to include single-cell and single-nucleus datasets, with explicit annotation of root zonation and nodulation stages, would greatly enhance the ability to select robust candidates and to interpret seemingly conflicting results. As the number of high-resolution nodulation datasets continues to grow, integration rather than further fragmentation will be essential.

Single-cell data as validation of mechanistic models

Beyond candidate discovery, scRNA-seq provides an independent means of validating mechanistic models derived from imaging and genetic analyses. The expression patterns of *MtYUC1*, *MtYUC8*, and *MtYUC9* in the scRNA-seq dataset were largely consistent with the spatial and temporal dynamics inferred from RNA *in situ* hybridization and reporter analyses in **Chapter 2** (**Figure 3**). *MtYUC1* showed strong enrichment in vascular and pericycle-associated populations and was induced upon nodulation, supporting its proposed role in establishing the pericycle auxin maximum (**Figure 3A**). *MtYUC8* displayed induction patterns consistent with nodulation, although interpretation was complicated by its induction to protoplast isolation (**Figure 3B**). *MtYUC9* remained lowly expressed, in line with its weak and late expression in tissue-level analyses (**Figure 3C**).

These results provide mutual validation. The scRNA-seq dataset captures biologically meaningful patterns consistent with independent experimental approaches, while the agreement with **Chapter 2** further supports the overall quality of the single-cell data generated in **Chapter 4**. At the same time, limitations remain evident. The absence of epidermal cells in mature zones restricts the ability to validate ethylene redistribution, and the lack of positional information means that rare, initiating pericycle cells cannot be unambiguously distinguished from their unresponsive neighbor cells.

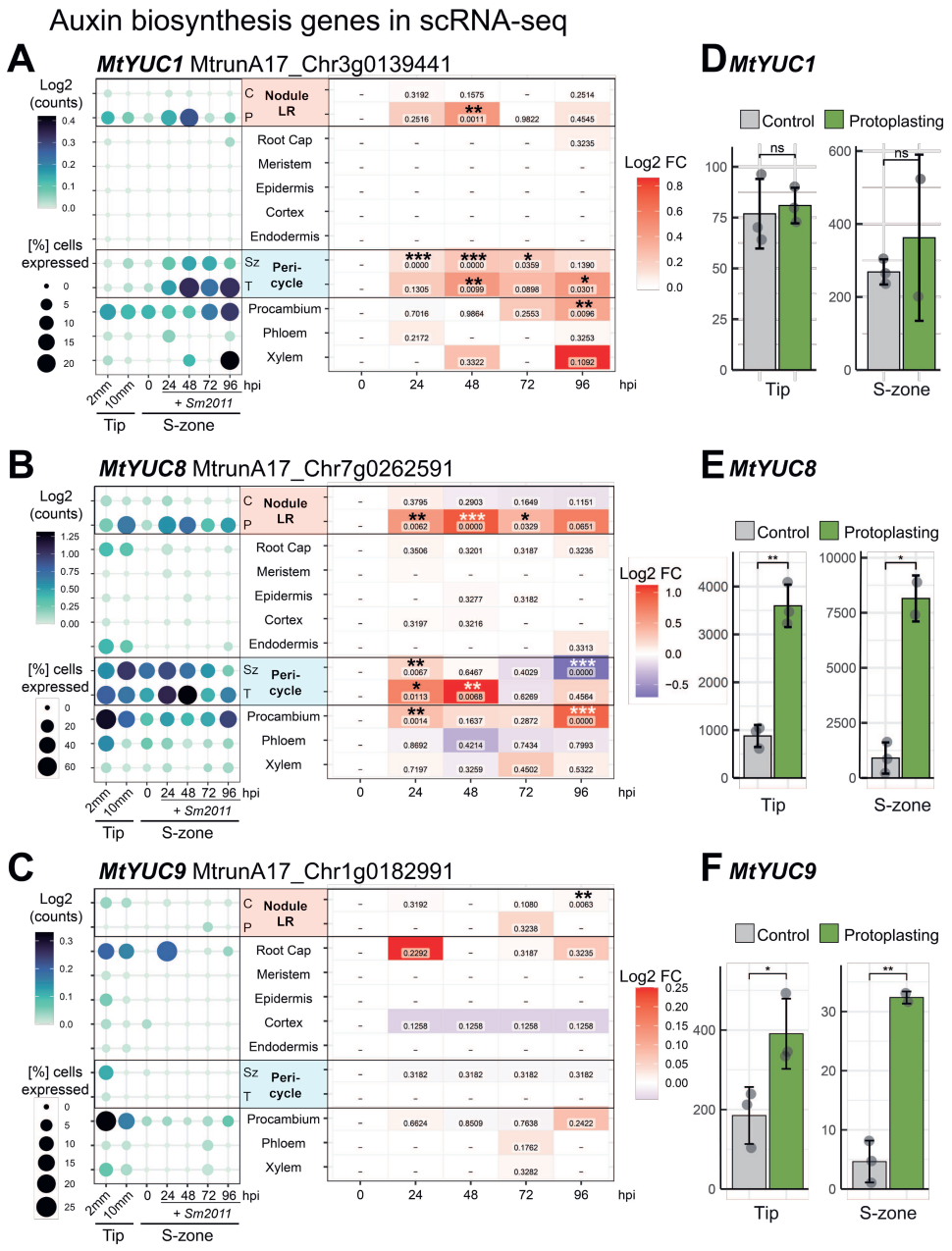


Figure 3 | Cell-type–resolved expression of *MtYUC1*, *MtYUC8*, and *MtYUC9* in scRNA-seq data. (A–C) Dot plots depicting expression patterns of *MtYUC1*, *MtYUC8*, and *MtYUC9* across tissues and cell types identified in the scRNA-seq dataset (left panels), alongside the corresponding $\log_2$ fold changes in response to *Sinorhizobium meliloti* 2011 inoculation at 24, 48, 72, and 96 hours post inoculation (hpi) relative to uninduced controls (right panels). **(D–F)** Changes in transcript abundance for *MtYUC1*, *MtYUC8*, and *MtYUC9* following protoplast isolation in the root tip and susceptible zone (S-zone), expressed as raw counts, illustrating procedure-associated effects on gene expression. Statistical significance between treatments was assessed using Student's *t*-test, with asterisks indicating significance levels ($p \leq 0.05$, $^{*}p \leq 0.01$, $^{**}p \leq 0.001$).

The need for spatial transcriptomics

Despite the advances enabled by scRNA-seq, the most critical events in nodule initiation occur in a small number of cells defined not only by their identity but also by their precise position within the root (Schiessl et al., 2019; Xiao et al., 2014). Single-cell approaches inherently lose this positional information (Adema et al., 2024). As a result, transcriptional changes specific to the earliest nodule-initiating pericycle cells are likely diluted within larger cell populations. Spatial transcriptomics therefore represents an essential next step for the field. By preserving tissue architecture while capturing transcriptomic information, spatial approaches could directly reveal how hormonal gradients, transcriptional programs, and cell fate transitions are organized *in situ* during the earliest stages of nodulation. Targeted approaches such as MERFISH with a specifically selected set of genes could be used to anchor scRNA-seq data sets to a spatial domain (Moffitt and Zhuang, 2016).

Implementing spatial transcriptomics in nodulation studies will be technically challenging, particularly because the relevant stages precede visible morphological changes. Nonetheless, recent advances in plant spatial transcriptomics and co-transcriptomics demonstrate that these challenges can be overcome. The integration of spatial transcriptomic data with existing single-cell and bulk datasets would provide a detailed, multi-tissue view of nodulation as a coordinated developmental process.

Implications for engineering nodulation and nitrogen use efficiency

One of the long-term motivations for understanding nodulation at a mechanistic level is the possibility of transferring this trait to non-nodulating crops (Geurts and Radutoiu, 2025; Huisman and Geurts, 2020; Kundu et al., 2025). While nodulation offers clear advantages in terms of nitrogen use efficiency, I would like to emphasize that the trait is inherently complex and tightly regulated. Simply introducing key positive regulators such as NIN is unlikely to be sufficient (Untergasser et al., 2012), because nodulation depends on the coordinated activation and suppression of multiple pathways in specific tissues and at specific times (Roy et al., 2020; Velandia et al., 2022).

Auxin-mediated pericycle activation represents a promising entry point for engineering, given its shared yet distinct roles in lateral root and nodule initiation. The findings I presented in **Chapter 2** suggest that inducing local auxin biosynthesis may be sufficient to trigger organogenic divisions without recreating the full complexity of indeterminate nodule meristems. Determinate, nodule-like structures may already provide a suitable niche for rhizobia, particularly if infection and nitrogen fixation can be established.

Equally important, however, is the role of negative regulators such as ethylene. Ethylene cannot simply be removed, as it is essential for restricting infection, enforcing developmental boundaries, and maintaining plant defenses (Berrabah et al., 2018; Bleecker et al., 1988; Khan et al., 2014; Lee and LaRue, 1992; Liu et al., 2015; Merchante and Stepanova, 2017; Pitts et al., 1998). The spatial redistribution of ethylene during nodulation provides a clear example of how

negative regulation can be locally relieved while being maintained in other tissues. Engineering approaches that mimic this spatial logic, potentially in coordination with microbial partners that locally suppress ethylene biosynthesis, may provide a more realistic path toward controlled synthetic symbiosis.

Given the complexity of these regulatory networks, it is likely that alternative strategies, including optimized fertilizer use, crop rotation, and microbial consortia, will remain critical for improving nitrogen use efficiency in the near future (Bhatt et al., 2025; Gan et al., 2015; Liu et al., 2020). Nonetheless, a detailed mechanistic understanding of nodulation, such as addressed in my thesis, will be a necessary foundation for any future attempts at trait transfer.

Conclusions and future perspectives

With this thesis, I demonstrate that controlled nodule initiation in Medicago depends on the precise spatiotemporal regulation of phytohormones, particularly auxin and ethylene. Local auxin biosynthesis in the pericycle initiates cell divisions, while spatial redistribution of ethylene removes inhibition from inner tissues and reinforces restriction of infection at the epidermis. These dynamics cannot be understood from bulk measurements alone and require resolution at the level of tissues, isoforms, and individual cells.

The generation of a protoplast-based scRNA-seq dataset provided new insights into cell-type-specific regulation during nodulation and highlighted both the power and the limitations of single-cell approaches in less well defined model species. Future progress will depend on integrating single-cell data with spatial transcriptomics and on combining multiple datasets to identify robust regulatory patterns. As cytokinin dynamics, cortex dedifferentiation, and meristem formation are further dissected, the conceptual framework I started to develop here will help place new findings into a coherent developmental model.

Ultimately, nodulation emerges not as a simple gain-of-function trait but as a finely nuanced developmental program that integrates positive and negative signals across tissues. Understanding how this balance is achieved brings us closer to explaining how nodulation evolved, why it was lost so often, and how its principles might one day be harnessed for more sustainable agricultural practices.

ACKNOWLEDGMENTS

I would like to thank my supervisors Viola Willemsen and Wouter Kohlen for their constructive feedback on this work.

REFERENCES

Adema, K., Schon, M. A., Nodine, M. D. and Kohlen, W. (2024). Lost in space: what single-cell RNA sequencing cannot tell you. *Trends Plant Sci* 29, 1018–1028.

Bergen, V., Soldatov, R. A., Kharchenko, P. V and Theis, F. J. (2021). RNA velocity—current challenges and future perspectives. *Molecular Systems Biology 2021 17:8* 17, MSB202110282-.

Berrabah, F., Balliau, T., Aït-Salem, E. H., George, J., Zivy, M., Ratet, P. and Gourion, B. (2018). Control of the ethylene signaling pathway prevents plant defenses during intracellular accommodation of the rhizobia. *New Phytologist* 219, 310–323.

Bhatt, R., Kunal, Moulick, D., Bárek, V., Brestic, M., Gaber, A., Skalicky, M. and Hossain, A. (2025). Sustainable strategies to limit nitrogen loss in agriculture through improving its use efficiency—aiming to reduce environmental pollution. *J Agric Food Res* 22, 101957.

Biedroń, M. and Banasiak, A. (2018). Auxin-mediated regulation of vascular patterning in Arabidopsis thaliana leaves. *Plant Cell Reports 2018 37:9* 37, 1215–1229.

Bleecker, A. B., Estelle, M. A., Somerville, C. and Kende, H. (1988). *Insensitivity to Ethylene Conferred by a Dominant Mutation in Arabidopsis thaliana*.

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 "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.

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.

Cheng, Y., Dai, X. and Zhao, Y. (2006). Auxin biosynthesis by the YUCCA flavin monooxygenases controls the formation of floral organs and vascular tissues in Arabidopsis. *Genes Dev* 20, 1790–1799.

Cooper, J. B. and Long, S. R. (1994). Morphogenetic Rescue of Rhizobium meliloti Nodulation Mutants by trans-Zeatin Secretion. *Plant Cell* 6, 215–225.

Du, Y. and Scheres, B. (2018). Lateral root formation and the multiple roles of auxin. *J Exp Bot* 69, 155–167.

Gan, Y., Hamel, C., O'Donovan, J. T., Cutforth, H., Zentner, R. P., Campbell, C. A., Niu, Y. and Poppy, L. (2015). Diversifying crop rotations with pulses enhances system productivity. *Scientific Reports 2015 5:1* 5, 14625-.

Gauthier-Coles, C., White, R. G. and Mathesius, U. (2019). Nodulating legumes are distinguished by a sensitivity to cytokinin in the root cortex leading to pseudonodule development. *Front Plant Sci* 9, 430352.

Geldner, N. (2013). The endodermis. *Annu Rev Plant Biol* 64, 531–558.

Geurts, R. and Radutoiu, S. (2025). Self-fertilizing crops: Plant Biotech's Hardest, Most Rewarding Task. *EMBO Rep* 26, 5644.

Gorin, G., Fang, M., Chari, T. and Pachter, L. (2022). RNA velocity unraveled. *PLoS Comput Biol* 18, e1010492.

Griesmann, M., Chang, Y., Liu, X., Song, Y., Haberer, G., Crook, M. B., Billault-Penneteau, B., Lauressergues, D., Keller, J., Imanishi, L., et al. (2018). Phylogenomics reveals multiple losses of nitrogen-fixing root nodule symbiosis. *Science (1979)* 361,.

Grunewald, W., Van Noorden, G., Van Lsterdael, G., Beeckman, T., Gheysen, G. and Mathesius, U. (2009). Manipulation of Auxin Transport in Plant Roots during Rhizobium Symbiosis and Nematode Parasitism. *Plant Cell* 21, 2553–2562.

Heidstra, R., Yang, W. C., Yalcin, Y., Peck, S., Emons, A. M., Van Kammen, A. and Bisseling, T. (1997). Ethylene provides positional information on cortical cell division but is not involved in Nod factor-induced root hair tip growth in Rhizobium-legume interaction. *Development* 124, 1781–1787.

Hirsch, A. M., Larue, T. A. and Doyle, J. (1997). Is the Legume Nodule a Modified Root or Stem or an Organ sui generis ? . *CRC Crit Rev Plant Sci* 16, 361–392.

Huisman, R. and Geurts, R. (2020). A Roadmap toward Engineered Nitrogen-Fixing Nodule Symbiosis. *Plant Commun* 1, 100019.

Jabloune, R., Khalil, M., Ben Moussa, I. E., Simao-Beaunoir, A. M., Lerat, S., Brzezinski, R. and Beaulieu, C. (2020). Enzymatic Degradation of p-Nitrophenyl Esters, Polyethylene Terephthalate, Cutin, and Suberin by Sub1, a Suberinase Encoded by the Plant Pathogen Streptomyces scabies. *Microbes Environ* 35,.

Khan, M., Rozhon, W. and Poppenberger, B. (2014). The role of hormones in the aging of plants - a mini-review. *Gerontology* 60, 49–55.

Kundu, A., Moraes, T. A., Jordan Price, R., Harrison, R. J. and Oldroyd, G. E. D. (2025). Getting to the Route: The Evolution of Nitrogen-Fixing Nodules. *Annu Rev Cell Dev Biol* 41, 457–478.

La Manno, G., Soldatov, R., Zeisel, A., Braun, E., Hochgerner, H., Petukhov, V., Lidschreiber, K., Kastriti, M. E., Lönnerberg, P., Furlan, A., et al. (2018). RNA velocity of single cells. *Nature 2018 560:7719* 560, 494–498.

Laplaze, L., Benkova, E., Casimiro, I., Maes, L., Vanneste, S., Swarup, R., Weijers, D., Calvo, V., Parizot, B., Herrera-Rodriguez, M. B., et al. (2008). Cytokinins Act Directly on Lateral Root Founder Cells to Inhibit Root Initiation. *Plant Cell* 19, 3889–3900.

Larrainzar, E., Riely, B. K., Kim, S. C., Carrasquilla-Garcia, N., Yu, H. J., Hwang, H. J., Oh, M., Kim, G. B., Surendrarao, A. K., Chasman, D., et al. (2015). Deep sequencing of the Medicago truncatula root transcriptome reveals a massive and early interaction between nodulation factor and ethylene signals. *Plant Physiol* 169, 233–265.

Lee, K. H. and LaRue, T. A. (1992). Exogenous Ethylene Inhibits Nodulation of Pisum sativum L. cv Sparkle. *Plant Physiol* 100, 1759–1763.

Lee, T., Orvosova, M., Batzenschlager, M., Bueno Batista, M., Bailey, P. C., Mohd-Radzman, N. A., Gurzadyan, A., Stuer, N., Mysore, K. S., Wen, J., et al. (2024). Light-sensitive short hypocotyl genes confer symbiotic nodule identity in the legume Medicago truncatula. *Current Biology* 34, 825-840.e7.

Lin, J., Frank, M. and Reid, D. (2020). No Home without Hormones: How Plant Hormones Control Legume Nodule Organogenesis. *Plant Commun* 1,.

Liu, M., Pirrello, J., Chervin, C., Roustan, J. P. and Bouzayen, M. (2015). Ethylene Control of Fruit Ripening: Revisiting the Complex Network of Transcriptional Regulation. *Plant Physiol* 169, 2380.

Liu, J., Rutten, L., Limpens, E., Van Der Molen, T., Van Velzen, R., Chen, R., Chen, Y., Geurts, R., Kohlen, W., Kulikova, O., et al. (2019). A remote cis-regulatory region is required for nin expression in the pericycle to initiate nodule primordium formation in medicago truncatula. *Plant Cell* 31, 68–83.

Liu, K., Bandara, M., Hamel, C., Knight, J. D. and Gan, Y. (2020). Intensifying crop rotations with pulse crops enhances system productivity and soil organic carbon in semi-arid environments. *Field Crops Res* 248, 107657.

Liu, Z., Yang, J., Long, Y., Zhang, C., Wang, D., Zhang, X., Dong, W., Zhao, L., Liu, C., Zhai, J., et al. (2023). Single-nucleus transcriptomes reveal spatiotemporal symbiotic perception and early response in Medicago. *Nat Plants*.

Ljung, K., Hull, A. K., Celenza, J., Yamada, M., Estelle, M., Normanly, J. and Sandberg, G. (2005). Sites and Regulation of Auxin Biosynthesis in Arabidopsis Roots. *Plant Cell* 17, 1090–1104.

Mathesius, U., Weinman, J. J., Rolfe, B. G. and Djordjevic, M. A. (2000). Rhizobia can induce nodules in white clover by "hijacking" mature cortical cells activated during lateral root development. *Mol Plant Microbe Interact* 13, 170–182.

Merchante, C. and Stepanova, A. N. (2017). The Triple Response Assay and Its Use to Characterize Ethylene Mutants in Arabidopsis. *Methods in Molecular Biology* 1573, 163–209.

Moffitt, J. R. and Zhuang, X. (2016). RNA Imaging with Multiplexed Error-Robust Fluorescence In Situ Hybridization (MERFISH). *Methods Enzymol* 572, 1–49.

Penmetsa, R. V. and Cook, D. R. (1997). A legume ethylene-insensitive mutant hyperinfected by its rhizobial symbiont. *Science (1979)* 275, 527–530.

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.

Penmetsa, R. V., Uribe, P., Anderson, J., Lichtenzveig, J., Gish, J. C., Nam, Y. W., Engstrom, E., Xu, K., Sckisel, G., Pereira, M., et al. (2008). The Medicago truncatula ortholog of Arabidopsis EIN2, sickle, is a negative regulator of symbiotic and pathogenic microbial associations. *Plant Journal* 55, 580–595.

Pereira, W. J., Boyd, J., Conde, D., Triozzi, P. M., Balmant, K. M., Dervinis, C., Schmidt, H. W., Boaventura-Novaes, C., Chakraborty, S., Knaack, S. A., et al. (2024). The single-cell transcriptome program of nodule development cellular lineages in Medicago truncatula. *Cell Rep* 43,.

Pitts, R. J., Cernac, A. and Estelle, M. (1998). Auxin and ethylene promote root hair elongation in Arabidopsis. *Plant J* 16, 553–560.

Reid, D., Liu, H., Kelly, S., Kawaharada, Y., Mun, T., Andersen, S. U., Desbrosses, G. and Stougaard, J. (2018). Dynamics of ethylene production in response to compatible nod factor. *Plant Physiol* 176, 1764–1772.

Roy, S., Liu, W., Nandety, R. S., Crook, A., Mysore, K. S., Pislariu, C. I., Frugoli, J., Dickstein, R. and Udvardi, M. K. (2020). Celebrating 20 Years of Genetic Discoveries in Legume Nodulation and Symbiotic Nitrogen Fixation. *Plant Cell* 32, 15–41.

Santos Teixeira, J. A. and ten Tusscher, K. H. (2019). The Systems Biology of Lateral Root Formation: Connecting the Dots. *Mol Plant* 12, 784–803.

Schiessl, K. and Jhu, M. Y. (2025). From roots to nodules: regulation of organogenesis in nitrogen-fixing symbiosis. *Curr Opin Plant Biol* 86, 102755.

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). NODULE INCEPTION Recruits the Lateral Root Developmental Program for Symbiotic Nodule Organogenesis in Medicago truncatula. *Current Biology* 29, 3657-3668.e5.

Shaw, S. L., Dumais, J. and Long, S. R. (2000). *Cell Surface Expansion in Polarly Growing Root Hairs of Medicago truncatula*.

Soltis, D. E., Soltis, P. S., Morgan, D. R., Swensen, S. M., Mullin, B. C., Dowd, J. M. and Martin, P. G. (1995). Chloroplast gene sequence data suggest a single origin of the predisposition for symbiotic nitrogen fixation in angiosperms. *Proc Natl Acad Sci U S A* 92, 2647–2651.

Soyano, T., Shimoda, Y., Kawaguchi, M. and Hayashi, M. (2019). A shared gene drives lateral root development and root nodule symbiosis pathways in Lotus. *Science (1979)* 366, 1021–1023.

Su, Y., Feng, T., Liu, C. Bin, Huang, H., Wang, Y. L., Fu, X., Han, M. L., Zhang, X., Huang, X., Wu, J. C., et al. (2023). The evolutionary innovation of root suberin lamellae contributed to the rise of seed plants. *Nature Plants 2023 9:12* 9, 1968–1977.

Untergasser, A., Bijl, G. J. M., Liu, W., Bisseling, T., Schaart, J. G. and Geurts, R. (2012). One-step Agrobacterium mediated transformation of eight genes essential for rhizobium symbiotic signaling using the novel binary vector system pHUGE. *PLoS One* 7,.

van Velzen, R., Holmer, R., Bu, F., Rutten, L., van Zeijl, A., Liu, W., Santuari, L., Cao, Q., Sharma, T., Shen, D., et al. (2018). Comparative genomics of the nonlegume Parasponia reveals insights into evolution of nitrogen-fixing rhizobium symbioses. *Proceedings of the National Academy of Sciences* 115, E4700–E4709.

Van Velzen, R., Doyle, J. J. and Geurts, R. (2019). A Resurrected Scenario: Single Gain and Massive Loss of Nitrogen-Fixing Nodulation. *Trends Plant Sci* 24,.

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-chitooligosaccharide Signaling Triggers Accumulation of Cytokinins in Medicago truncatula Roots. *Mol Plant* 8, 1213–1226.

Velandia, K., Reid, J. B. and Foo, E. (2022). Right time, right place: The dynamic role of hormones in rhizobial infection and nodulation of legumes. *Plant Commun* 3,.

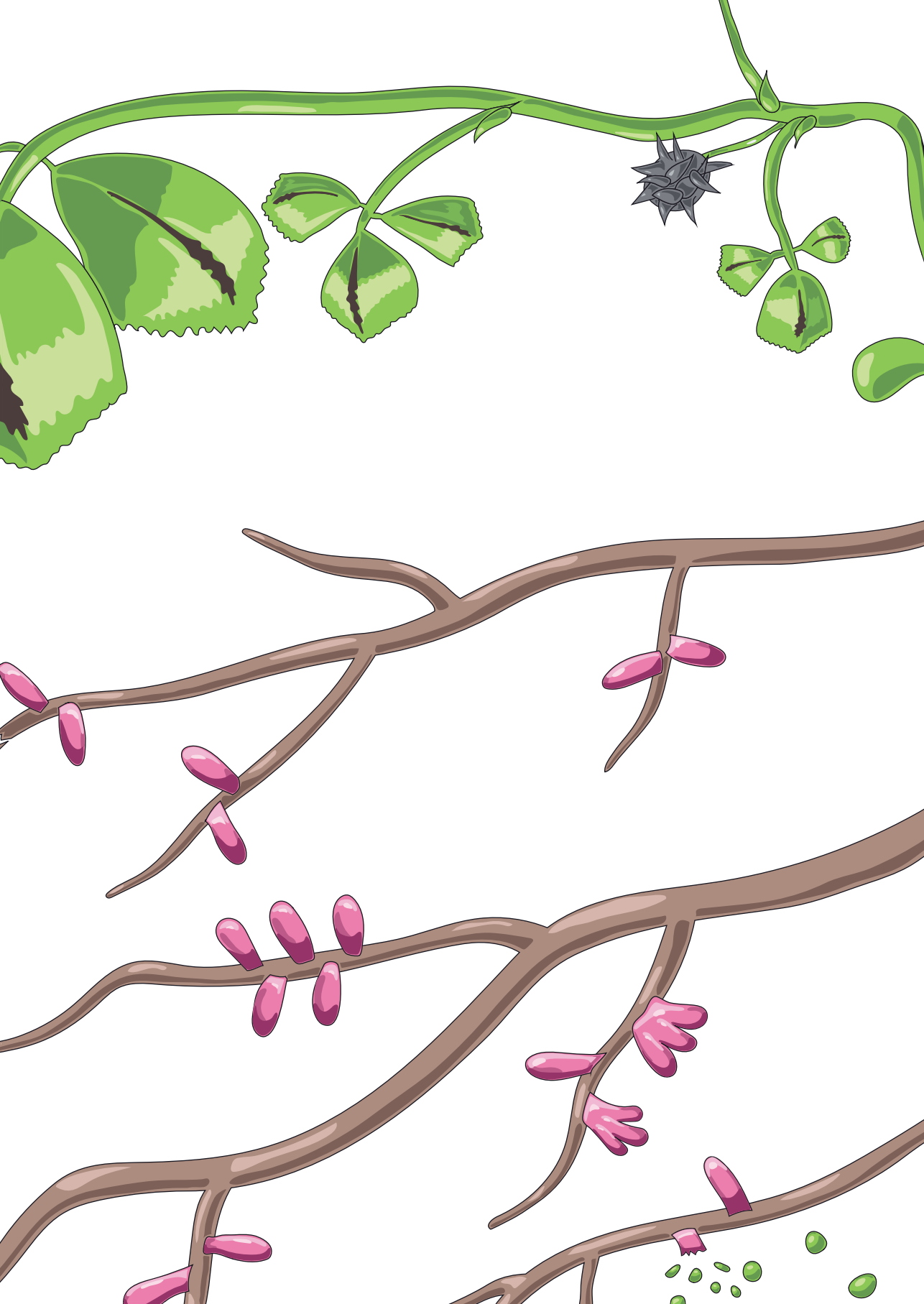
Waese, J., Fan, J., Pasha, A., Yu, H., Fucile, G., Shi, R., Cumming, M., Kelley, L. A., Sternberg, M. J., Krishnakumar, V., et al. (2017). ePlant: Visualizing and Exploring Multiple Levels of Data for Hypothesis Generation in Plant Biology. *Plant Cell* 29, 1806–1821.

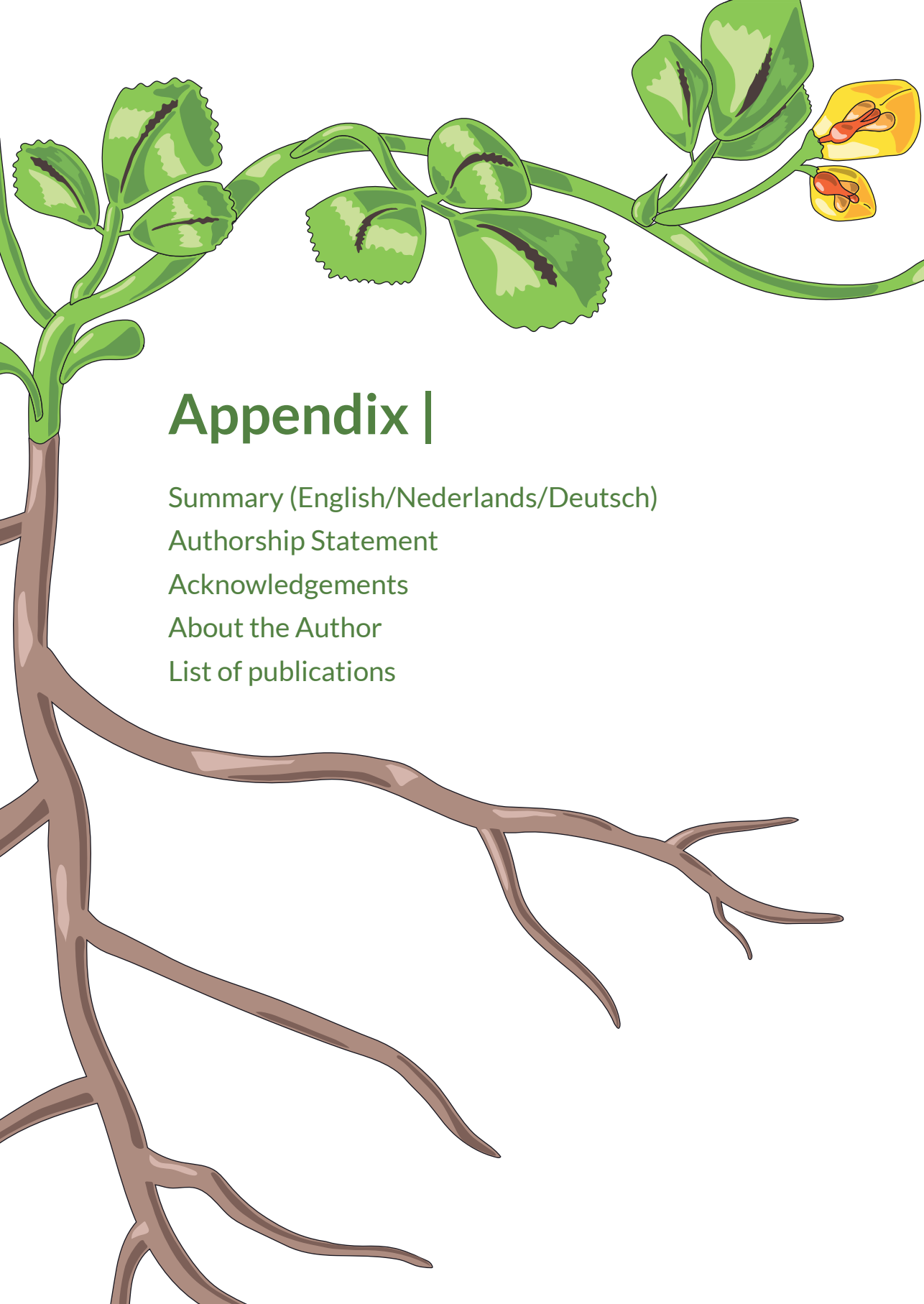
Wang, D., Dong, W., Murray, J. and Wang, E. (2022). Innovation and appropriation in mycorrhizal and rhizobial Symbioses. *Plant Cell* 34, 1573.

Xiao, T. T., Schilderink, 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 (Cambridge)* 141, 3517–3528.

Yan, J., Bisseling, T., Cao, Q., Li, H., Yan, J., Bisseling, T., Cao, Q. and Li, H. (2025). The adaptive root: endodermal plasticity and microbial composition.

Zhang, S., Dong, Z., Shi, J., Yang, C., Fang, Y., Chen, G., Chen, H. and Tian, C. (2022). Enzymatic hydrolysis of corn stover lignin by laccase, lignin peroxidase, and manganese peroxidase. *Bioresour Technol* 361,.





Appendix |

Summary (English/Nederlands/Deutsch)

Authorship Statement

Acknowledgements

About the Author

List of publications

Summary

Nodulation is a symbiotic interaction between soil-borne bacteria, rhizobia, and legumes such as the model organism *Medicago truncatula*. In this symbiosis, rhizobia fix atmospheric nitrogen (N_2) into a form usable by the plant, while the plant provides carbon sources, such as sugars, and a protected environment, the nodule, in exchange. This highly complex process is of interest both for its agricultural potential and for its unique developmental program, which is absent in most plants.

This thesis investigates the spatiotemporal regulation of early nodule development in *Medicago truncatula*, with a focus on hormonal regulation and the use of high-resolution transcriptomics. **Chapter 1** introduces legume nodulation at a general level, with an emphasis on hormone biology, and outlines why understanding when and where root cells change identity is essential for explaining nodule initiation. Current gaps in knowledge of hormonal regulation during early symbiosis are highlighted, emphasizing the need for spatiotemporally resolved approaches.

Building on this, **Chapter 2** addresses how auxin gradients are established during the earliest stages of nodule formation. While auxin maxima have long been associated with organ development, the mechanisms establishing these patterns during nodulation were previously unclear. Correlating DR5 auxin signaling with the expression of auxin biosynthesis and transport genes reveals that specific *MtYUCs*, *MtPINs*, and *MtLAX2* are activated in the pericycle immediately after rhizobial inoculation, prior to cell division. Furthermore, we demonstrate that coordinated auxin biosynthesis and polar transport are required for the initiation of the nodule primordium and meristem. **Chapter 3** extends the hormonal analysis to ethylene, a negative regulator of nodulation whose spatial dynamics had not been resolved. Transcriptomics, synthetic reporters, and mutant analyses show that ethylene signaling is reprogrammed following rhizobial inoculation. The responsiveness to ethylene, detected through the EBSn reporter, shifts from inner root tissues to the outer cortex and epidermis. This shift coincides with the induction of *MtACS3* and repression of *MtACS10*. MtACS10 restricts nodule initiation, whereas MtACS3 regulates infection thread progression and limits excessive or mispositioned primordia, demonstrating the importance of spatially confined ethylene regulation.

To resolve early nodulation processes at higher resolution, we developed and optimized single-cell and single-nucleus RNA-seq workflows for roots and nodules in **Chapter 4**. Comparison of isolation protocols and high-throughput platforms identifies critical factors affecting RNA quality, cell viability, and transcriptome coverage. Optimized protoplast-based approaches provide higher mapping rates and broader cell-type representation, while integration with nuclei-based methods offers access to complementary information and underrepresented populations. Leveraging these optimized approaches, **Chapter 5** identifies transcription factors that may confer pericycle competence during nodule initiation. Integration of bulk time-course RNA-seq with our root single-cell atlas highlights ~600 pericycle-enriched, nodulation-responsive genes, and lead us to focus on three transcription factors—*MtbZIP43*, *MtDREB3*, and

MtSCL14. These genes display distinct but overlapping expression patterns in the pericycle and vasculature and are dynamically induced after inoculation. RNA interference lines show no strong phenotypes, whereas overexpression reduces nodule numbers.

Finally, **Chapter 6** correlates these findings with each other and puts them in a broader context, showing how the combined hormonal, transcriptional, and methodological insights advance our understanding of the spatial regulation underlying symbiotic organogenesis.

Samenvatting

Nodulatie is een symbiotische interactie tussen bodembewonende bacteriën, de rhizobia, en vlinderbloemigen zoals het modelorganisme *Medicago truncatula*. In deze symbiose fixeren rhizobia atmosferische stikstof (N_2) in een voor de plant bruikbare vorm, terwijl de plant in ruil daarvoor koolstofbronnen, zoals suikers, en een beschermde omgeving, het wortelknolletje, verschaft. Dit hoogcomplexe proces is van groot belang, zowel vanwege het landbouwkundig potentieel als vanwege het unieke ontwikkelingsprogramma, dat bij de meeste planten ontbreekt.

Dit proefschrift onderzoekt de ruimtelijk-tijdelijke regulatie van de vroege ontwikkeling van wortelknolletjes in *Medicago truncatula*, met een focus op hormonale regulatie en het gebruik van transcriptomics met hoge resolutie. **Hoofdstuk 1** introduceert de nodulatie bij vlinderbloemigen op algemeen niveau, met nadruk op de hormoonbiologie, en beschrijft waarom inzicht in wanneer en waar wortelcellen van identiteit veranderen essentieel is voor het begrijpen van knolinitiatie. Daarnaast worden huidige kennislacunes in de hormonale regulatie tijdens de vroege symbiose besproken, wat de noodzaak van ruimtelijk-tijdelijk opgeloste benaderingen onderstreept.

In **Hoofdstuk 2** wordt onderzocht hoe auxinegradiënten tot stand komen tijdens de vroegste stadia van de wortelknolvorming. Hoewel auxinemaxima al lange tijd in verband worden gebracht met orgaanontwikkeling, waren de mechanismen die deze patronen tijdens nodulatie bepalen tot nu toe onduidelijk. Door DR5-auxinesignalering te correleren met de expressie van genen betrokken bij auxinebiosynthese en -transport, blijkt dat specifieke *MtYUCs*, *MtPINs* en *MtLAX2* direct na rhizobiale inoculatie in de pericykel worden geactiveerd, voorafgaand aan celdeling. Verder wordt aangetoond dat een nauw gereguleerde auxinebiosynthese en polair auxinetransport noodzakelijk zijn voor de initiatie van het wortelknolprimordium en meristeem. **Hoofdstuk 3** breidt deze analyse uit naar ethyleen, een negatieve regulator van nodulatie waarvan de ruimtelijke dynamiek tot dusver niet bekend was. Transcriptomics, synthetische reporters en mutantanalyse tonen aan dat ethyleensignalering na rhizobiale inoculatie sterk wordt hergeprogrammeerd. De ethyleenrespons, gedetecteerd met de EBSn-reporter, verschuift van de binnenste wortelweefsels naar de buitenste cortex en epidermis. Deze verschuiving gaat samen met inductie van *MtACS3* en repressie van *MtACS10*. Waar MtACS10 de knolinitiatie beperkt, reguleert MtACS3 de voortgang van infectiedraden en voorkomt het overmatige of verkeerd gepositioneerde primordia, wat het belang van ruimtelijk begrensde ethyleenregulatie aantoont.

Om vroege nodulatieprocessen met hogere resolutie te bestuderen, worden in **Hoofdstuk 4** single-cell- en single-nucleus-RNA-seq-workflows voor wortels en wortelknolletjes ontwikkeld en geoptimaliseerd. Vergelijking van verschillende isolatieprotocollen en high-throughputplatforms identificeert cruciale factoren die RNA-kwaliteit, celviabiliteit en transcriptomdekking beïnvloeden. Geoptimaliseerde protoplast-gebaseerde benaderingen

leveren hogere mappingpercentages en een bredere representatie van celtypen, terwijl integratie met nucleus-gebaseerde methoden toegang biedt tot aanvullende informatie en ondervertegenwoordigde populaties. Op basis van deze geoptimaliseerde benaderingen identificeert **Hoofdstuk 5** transcriptiefactoren die mogelijk bijdragen aan pericykelcompetentie tijdens wortelknolinitiatie. Integratie van bulk time-course RNA-seq met een single-cell wortelatlas brengt ~600 pericykel-verrijkte, nodulatie-responsieve genen aan het licht en leidt tot een focus op drie transcriptiefactoren—*MtbZIP43*, *MtDREB3* en *MtSCL14*. Deze genen vertonen onderscheidende maar overlappende expressiepatronen in de pericykel en het vaatweefsel en worden dynamisch geïnduceerd na inoculatie. RNA-interferentielijnen vertonen geen uitgesproken fenotypen, terwijl overexpressie resulteert in een verminderd aantal wortelknolletjes.

Ten slotte plaatst Hoofdstuk 6 deze bevindingen in een breder kader en laat zien hoe de gecombineerde hormonale, transcriptionele en methodologische inzichten bijdragen aan een beter begrip van de ruimtelijke regulatie van symbiotische organogenese.

Zusammenfassung

Wurzelknöllchenbildung (Nodulation) ist eine symbiotische Interaktion zwischen bodenbewohnenden Bakterien, den Rhizobien, und Leguminosen wie dem Modellorganismus *Medicago truncatula*. In dieser Symbiose fixieren Rhizobien atmosphärischen Stickstoff (N_2) in eine für die Pflanze nutzbare Form, während die Pflanze Kohlenstoffquellen wie Zucker und eine geschützte Umgebung, das Wurzelknöllchen, bereitstellt. Dieser hochkomplexe Prozess ist sowohl wegen seines landwirtschaftlichen Potenzials als auch wegen seines einzigartigen Entwicklungsprogramms, das bei den meisten Pflanzen fehlt, von großem Interesse.

Diese Dissertation untersucht die räumlich-zeitliche Regulation der frühen Wurzelknöllchenentwicklung in *Medicago truncatula*, mit Schwerpunkt auf hormoneller Regulation und der Anwendung hochauflösender Transkriptomik. **Kapitel 1** führt die Nodulation bei Leguminosen auf allgemeiner Ebene ein, mit Fokus auf der Hormonregulation, und erläutert, warum das Verständnis des Zeitpunkts und der Lokalisation von Veränderungen der Zellidentität in den Wurzeln entscheidend für die Initiierung der Wurzelknöllchen ist. Aktuelle Wissenslücken zur hormonellen Regulation während der frühen Symbiose werden aufgezeigt, wobei die Notwendigkeit räumlich-zeitlicher Analyseansätze betont wird.

Aufbauend darauf behandelt **Kapitel 2** die Entstehung von Auxin-Gradienten in den frühesten Stadien der Wurzelknöllchenbildung. Obwohl Auxin-Maxima schon lange mit Organentwicklung in Verbindung gebracht werden, waren die Mechanismen, die diese Muster während der Nodulation erzeugen, bisher unklar. Die Korrelation von DR5-Auxin-Signalisierung mit der Expression von Auxinsynthese- und Transportgenen zeigt, dass spezifische *MtYUCs*, *MtPINs* und *MtLAX2* unmittelbar nach der Rhizobieninokulation im Perizykel aktiviert werden, noch bevor die Zellteilung beginnt. Koordinierte Auxinsynthese und polare Transportmechanismen sind erforderlich, um die Initiierung des Wurzelknöllchenprimordiums und des Meristems zu gewährleisten. **Kapitel 3** erweitert die hormonelle Analyse auf Ethylen, einen negativen Regulator der Nodulation, dessen räumliche Dynamik bislang nicht aufgeklärt war. Transkriptomik, synthetische Reporter und Analysen mit Mutanten zeigen, dass die Ethylensignalkaskade nach Inokulation mit Rhizobien umprogrammiert wird. Die Responsivität zu Ethylene, nachgewiesen mit dem EBSn-Reporter, verlagert sich von den inneren Wurzelgeweben zum äußeren Cortex und der Epidermis. Diese Verschiebung korreliert mit der Induktion von *MtACS3* und der Repression von *MtACS10*. MtACS10 begrenzt die Initiierung von Wurzelknöllchen, während MtACS3 den Verlauf der Infektionsfäden reguliert und übermäßige oder fehlplatzierte Primordien verhindert, was die Bedeutung räumlich begrenzter Ethylenregulation verdeutlicht.

Um frühe Nodulationsprozesse in höherer Auflösung zu untersuchen, wurden in **Kapitel 4** Single-Cell- und Single-Nukleus-RNA-seq-Workflows für Wurzeln und Wurzelknöllchen entwickelt und optimiert. Der Vergleich von Isolationsprotokollen und Hochdurchsatzplattformen identifiziert kritische Faktoren, die die RNA-Qualität, Zellvitalität

und Transkriptomabdeckung beeinflussen. Optimierte Protoplast-basierte Ansätze ermöglichen höhere Mapping-Raten und eine breitere Zelltypenrepräsentation, während die Integration von Nukleus-basierten Methoden Zugang zu komplementären Informationen und unterrepräsentierten Zellpopulationen bietet. **Kapitel 5** identifiziert mithilfe dieser optimierten Ansätze Transkriptionsfaktoren, die die Kompetenz des Perizykels während der Wurzelknöllcheninitiierung vermitteln könnten. Die Integration von Bulk-Time-Course-RNA-seq mit unserem Single-Cell-Atlas identifiziert ~600 perizykel-angereicherte, nodulationsresponsive Gene, von denen wir uns auf drei Transkriptionsfaktoren fokussieren – *MtbZIP43*, *MtDREB3* und *MtSCL14*. Diese Gene zeigen klare, aber überlappende Expressionsmuster im Perizykel und im Gefäßgewebe und werden nach Inokulation dynamisch induziert. RNA-Interferenz-Linien zeigen keine starken Phänotypen, während Überexpression die Anzahl der Wurzelknöllchen reduziert.

Abschließend verknüpft **Kapitel 6** diese Ergebnisse miteinander und stellt sie in einen breiteren Kontext, wobei gezeigt wird, wie die Kombination hormoneller, transkriptioneller und methodischer Erkenntnisse das Verständnis der räumlichen Regulation der symbiotischen Organogenese voranbringt.

Authorship statement

Chapter 1 | General introduction
I was responsible for the full conceptualization and writing of this chapter. Minor comments and suggestions from my supervisors were incorporated following a single round of feedback.

Chapter 2 | Nodule organogenesis in *Medicago truncatula* requires local stage-specific auxin biosynthesis and transport
The conceptualization of this chapter was carried out by Henk Franssen, Olga Kulikova, and Wouter Kohlen, and the work builds on earlier unpublished research by Ting Ting Xiao, Defeng Shen, and Ton Bisseling. Together with Jieyu Liu, Kelvin Adema, Amber van Seters, Henk Franssen, Olga Kulikova, and Wouter Kohlen, I executed part of the experiments. DR5 localization studies were conducted by Ting Ting Xiao, Defeng Shen, and Olga Kulikova. Auxin measurements and pharmacological experiments were performed jointly by Wouter Kohlen and me. *RNAi* constructs targeting *MtYUCCA1*, *MtYUCCA2*, and *MtYUCCA8* were generated by Olga Kulikova and transgenic plants were created and analysed by Olga Kulikova and me, while PIN localization studies were performed by Olga Kulikova, with construct cloning completed by Olga Kulikova and Henk Franssen. *RNAi* constructs targeting *MtPIN2*, *MtPIN4*, and *MtPIN10* were generated by Olga Kulikova while knock down experiments using these constructs were carried out by Olga Kulikova and me. I conducted the physiological experiments related to auxin. I contributed to the preparation of the initial draft of the manuscript, which was primarily written by Henk Franssen, Olga Kulikova, and Wouter Kohlen, and together with Wouter Kohlen, I made the figures. All authors, including myself, reviewed and edited the final version.

Chapter 3 | Spatiotemporal dynamics of ethylene biosynthesis shape infection and nodule initiation in *Medicago truncatula*
The conceptualization of this chapter was carried out by me, and Wouter Kohlen and the work builds on earlier unpublished research by Arjan van Zeijl, Henk Franssen and Wouter Kohlen. Josefina-Patricia Fernandez-Moreno, Anna N. Stepanova, Jose M. Alonso contributed *EBSn*, a new analytical tool to visualize ethylene responses which we used in Medicago roots. These ethylene responses in the root susceptible zone during rhizobia-induced signalling were analysed by Thijs Stegman and me. Identification and characterization of ethylene biosynthesis genes in Medicago were performed by Robin van Velzen Joel Klein and Estibaliz Larrainzar.

Expression analysis of *MtACS3* and *MtACS10* under symbiotic and control conditions was conducted by Olga Kulikova, Arjan van Zeijl, and Wouter Kohlen, with *promoter*::GUS constructs generated by Arjan van Zeijl and *in situ* hybridizations performed by Olga Kulikova.

Nodulation phenotypes of *MtACS3* and *MtACS10* knockdown and knockout mutants were assessed by Amber van Seters, Arjan van Zeijl, and Wouter Kohlen, with mutant identification and *RNAi* construct generation completed by Arjan van Zeijl. Nodulation assays in perlite and on plates were carried out by me with assistance from Amber van Seters and Wouter Kohlen. The effects of the *mtacs3-1* and *mtacs10-1* loss-of-function mutations on infection thread formation were investigated by me. Nodule positioning relative to xylem and phloem poles was analysed in perlite by Amber van Seters and myself. The study examining whether counteracting

MtACS10 reduction alters nodulation was based on constructs generated by Arjan van Zeijl and experiments performed by Wouter Kohlen.

Physiological experiments to determine the role of ethylene in establishing the root susceptible zone were conducted by Wouter Kohlen and myself, with imaging performed by Tristan Wijsman. The model figure was created by me. Figure preparation was carried out by me together with Wouter Kohlen. The original draft was written by me and Wouter Kohlen. All authors contributed to review and editing of the manuscript, and the final version was prepared by me and Wouter Kohlen.

Chapter 4 | Establishing robust workflows for single-cell and single-nuclei RNA sequencing in the model legume *Medicago truncatula*

The conceptualization of this chapter was carried out by me, Carolin Grones and Wouter Kohlen, and forms one part of a diptych set of manuscripts developed in collaboration with Kelvin Adema. With support from Amber van Seters, I performed all experiments required to generate samples for single-cell and single-nucleus analyses. I conducted the nuclei extractions and, together with Jana Wittmer, designed and executed all protoplasting tests and optimization experiments. Bioinformatic analyses, including sample quality assessment and cell-type identification, were performed by me in collaboration with Kelvin Adema. I carried out all wet-lab validation experiments. In collaboration with Carolin Grones, I constructed all single-cell and single-nucleus libraries.

All artwork in this chapter was prepared by me. The original draft was written by Carolin Grones, Wouter Kohlen, and me. All authors contributed to reviewing and editing the chapter, and the final version was completed by Carolin Grones, Wouter Kohlen, and myself.

Chapter 5 | Single-Cell Transcriptomics-Guided Identification of Transcription Factors with Pericycle-Enriched Expression in *Medicago truncatula*

The conceptualization of this chapter was carried out by myself, Kelvin Adema, and my promotors. A list of candidate genes derived from the single-cell dataset was generated by Kelvin Adema. Candidate gene selection for further analysis was performed jointly by Kelvin Adema, Viola Willemsen, Wouter Kohlen, and me. Constructs for *promoter*::GUS, *RNAi*, and overexpression analyses were designed by me. All constructs were transiently introduced into *Medicago truncatula* roots by me and validated together with Chris Roelofsen. *Promoter*::GUS images were analysed in collaboration with my promotors. All artwork in this chapter was prepared by me. The original draft was written by me. All authors contributed to reviewing and editing the chapter, and the final version was completed by Wouter Kohlen and myself.

Chapter 6 | General discussion

I determined the outline and selected the topics to be included, and I wrote the chapter in full. I subsequently revised the chapter once, incorporating suggestions and comments provided by my promoters.

Acknowledgements

Everything comes to an end, and so does this thesis. I would like to express my gratitude to everyone who has accompanied me on this journey—through direct contributions to this work, by fostering a positive and supportive environment, or by caring for my well-being. If I have unintentionally failed to mention anyone by name, please know that your support and contributions are equally appreciated. Thank you all for making this possible.

First, I would like to thank my supervisor **Wouter**, who has made all of this possible. Without you, this project would never have come to life. From the very beginning, you made me feel welcome and created an atmosphere where I always felt comfortable approaching you—even during challenging personal times. You supported my decisions to be with my family for important events, such as the birth of my nieces, and enabled me to go to South Korea for a conference. Whenever I felt uncertain, our enthusiastic discussions left me feeling motivated and inspired. Thank you for giving me the opportunity to be your first PhD student and for supporting me all the way to the finish line.

Secondly, I would like to thank you, **Viola,** for being my promotor and supporting me especially in this final phase of the PhD, which proved to be particularly challenging. Your perspective greatly enriched our discussions and broadened my thinking. And, of course, thank you for all the wonderful dancing sessions.

I would also like to thank you, **Ben**, as my initial promotor, for your contributions to the early stages of this project and our insightful discussions. It was equally great to share some time on the dancefloor with you.

My work in the lab would not have been as efficient (maybe sometimes more efficient) or fun without all the people from our cluster of the Molecular Biology and the Cell and Developmental Biology subgroups. **Alessio, Amber, Andrea, Anneke, Asma, Alysha, Ben, Bowei, Caro, Carolien, Chris, Eline, Erik, Fengjiao, Francesca, Harm, Harald, Henk, Honglei, Isha, Jana, Jeroen, Jiawei, Jieyu, Job, Joël, Jordi, Joyce, Jundi, Kelvin, Kyra, Leonardo, Luuk, Marijke, Max, Merijn, Michael N., Michael S., Michiel, Norbert, Olga, Otto, Peter G., Peter v. G., Pierangela, René, Rens, Renze, Richard, Rik, Robert, Roman, Ronald, Sanne, Siqi, Sultan, Syafii, Thijs, Thomas, Tijs, Titis, Ton, Tristan, Vera, Viola, Wenying, Wilma, Wouter, Xiaofan, Yvet, Zhuang**, and **Zifan** – thank you for making the time in the lab and offices feel less like work.

The many conversations we shared—about science and life alike—are among the moments I cherish most from this time. I am truly grateful for them.

Thank you, Anneke, for creating the best PhD songs - now together with **Max -** and for keeping the literature discussion on track. **Alysha, Sanne,** and **Kyra** although we did not spend much time together, it was nice to see you join the lab, and have little chitchats. **Carolien**, thank you for ensuring everything ran smoothly and for always knowing where to find what we needed. **Chris,** your help toward the end of my thesis—especially with cloning and qPCR analysis—was invaluable, and I am grateful to have benefited from your expertise.

Eline, Pierangela, Ton, Joyce, Harald, **Zifan** and **Jeroen** – although you were in the office on "the other side" of the building, it was always fun to talk to you and see your different perspectives from barley and physcomitrium. Thank you, **Harald,** for making my self-designed magnetic rack a reality. **Tijs** and **Otto**, it was always inspiring to talk with you and see your optimistic and enthusiastic perspectives on the world.

Fengjiao, thank you for the tote bag—I truly love it and hope your research accelerates future discoveries. **Francesca**, I very much enjoyed our lunches, conversations, and the PhD videos we created together - I am happy to have gotten to know you. **Harm, Alessio, Max**, and **Thijs**, thank you for your musical contributions, taking over from the PDBand. It was always fun with you in the lab and in the office.

Honglei, Jiawei, and **Zhuang** thank you for the nice little conversations here and there that always lighten the day – especially since a problem shared, is a problem halved. **Joël** and **Titis**, I hope that the Chamaecrista work will eventually pay off – thank you for all the nice conversations we had, and all the best in your new places.

Jordi – it was always a blast having our century lunches. Luckily not only every century. I am looking forward to our next one. **Merijn** and **Michiel** —M&M—I still fondly remember the PhD Jamboree where you introduced me to Monty Python; since then, it has always been a joy spending time together. **Rens**, thank you for your bioinformatics support and the fun times at the Efteling. **Roman**, I truly enjoyed talking to you about fantasy and medieval worlds, as well as about board games.

Marijke, thank you for sharing your knowledge and support. **Michael N**. and **Michael S.**, I appreciate your help with nuclei work and bioinformatics. **Norbert**, thank you for your assistance in my unknowledgeable attempts with the microscope. **Peter v. G.,** thank you for making me feel so rich, when you shared your freshly arrived hemacytometer cover slips with me. **René**, thank you for including me in your ENSA group during the early years—an experience I greatly value. **Sultan**, thank you for sharing insights into your culture and making time at the sterile bench pass more quickly.

Thomas, it has been a pleasure having you in the lab, and I always appreciated your easy-going nature. **Tristan**, thank you for your legendary protocols and solutions, as well as your unique personality, always lifting ones spirit. **Vera**, although our time overlapped only briefly, I truly enjoyed getting to know you—and even more so each time we meet again. **Yvet**, I had a great time creating the PhD video for Vera with you; I wish you all the best for your PhD. **Wilma**, thank you for your highly competent Agrobacteria—even if they did not yield transgenic plants in my case.

Andrea, thank you for being the supportive and engaging person that you are. It was always a pleasure to work and talk with you. Good luck with everything.

Renze, I still remember the day you walked into the coffee room with that characteristic smirk to tell us the FACS was broken. I was not sure whether you were joking or not – of course you

were not. It perfectly captures how I will remember you. Thank you for your help, your teaching, and our conversations—always marked by your unique seemingly-sarcastic, yet optimistic attitude.

A special thank you goes to the Medicago group of **Rik, Luuk, Siqi, Jieyu**, and **Erik** – you were always helping me with my questions about Medicago, transformation, cloning, bacteria or whatever I needed from you. **Olga** and **Henk**, you were particularly important during the early phase of my PhD and had always good advice to make things more efficient or work at all. I hope you will make the best of your retirement, and I am happy that I still experienced you as an integral part of the lab.

Jana, Thijs and **Caro**, the term protoplasting-Leidensgenossen (fellow sufferer) is indeed a very fitting one. You cannot imagine how glad I was to discover that this seemingly simple procedure does not only suck the life out of me, but of everyone else, too. Thank you, **Jana,** for pioneering protoplasting with me in our lab. Thank you, **Thijs,** for suffering with me through the tomato shoot apical meristem dissection. And thank you, **Caro,** especially for validating that we are not crazy, and for taking over the single-cell procedure, so I was not only done after protoplasting, but actually finished. This really made a difference.

Maria and **Marie-José**, thank you for all the essential background work, your open ears, and your constant willingness to help wherever possible.

I would also like to thank the students I supervised. **Wessel, Jelle, Jorrick, Thijs d.V., Thijs S.**, and **Jordi d.G.** I hope you learned something in the time with me. I sure did learn from teaching you, and I wish you all the best in your future paths, wherever they may lead you.

Kilian, Jana, Rianne, and **Alessio** organizing the EPSR together with you was a great experience and I had a lot of fun with you, especially after all the hard work paid off.

I am especially grateful that as a reward for the hard work, we established our "Spa gang", that has led **Jana, Rianne,** and me to multiple relaxing sessions. Thank you for rewinding together at the end of the day. I hope we will keep doing this for many years to come.

Alessio, I have always admired your ability to clearly communicate your complex formulas with non-formula people. You are really good at what you are doing – as far as I can tell at least. Thank you for bringing the different viewpoint to our group. I am happy that we became friends, and I am looking forward to your defense and your future.

Amber, thank you for being there. Your social nature has made me feel very welcome right from the start, and I will always treasure our shared office time and cooking evenings. You showed me all the where's and what's in the lab, and your very specialized techniques that I gratefully applied. Wherever your life may lead you – I wish you all the best and a lot of happiness in it. I am excited to see where it takes you.

Caro, I'm glad that we got along so well right from the start after you became my new desk neighbour. Maybe we chatted a little too much from time to time, but it's always nice to have

someone who understands you. Even though the time didn't feel very long, I'm grateful to have had you by my side, especially during this final phase. I hope your single-cell platform will be realized soon, so you can truly benefit from it and spend more time doing science and less time on administration.

Colleen, thank you for providing me with the constructs for cloning in the beginning, and welcoming me in Cambridge.

Kaisa, thank you for being my external supervisor. I especially enjoyed our shared lunch and meeting you at all the conferences.

I also would like to especially thank my paranymphs that will be by my side to support me during my defense:

Jana – I'm glad that over time we didn't just share our struggles, but also many wonderful moments. I must admit your encouragement to join the goal trainers did put me through hardships at times, but that was to be expected. I'm especially grateful to you for introducing me to Jirka. Thank you for accompanying me on this final step toward my doctoral degree and for being there for me as a friend. I am excited to see to which delicious alternative protein solutions you will contribute.

Kelvin – I am happy that you joined our team and that you could take over all the data analysis, that I could not do. Without you, parts of this project would not have been possible. Thank you for being there and helping me. I am glad that we became friends and maybe one day we will finally manage our boardgame evening. I am looking forward to your defense and I wish you a smooth final stretch – it will be over one day.

Thank you, **Laura W.** and **Sarah,** to have been there for me during the last 26 years. I am very thankful that we, with our friend group from high school, had the chance to grow even closer, although I again moved further away. Thank you, **Laura W., Sarah, Anna, Astrid, Theresa**, and **Laura E.,** for the festivals, birthday parties, vacations, and the wedding! A great time to recharge for my PhD journey.

My friends from my Bachelor studies **Ronja, Judy**, and **Raphi**: I am thankful that since almost 13 years ago, until this day, we meet as much as possible and always have great times with and without science. **Ronja**, you always make me feel better about myself and less doubtful about my abilities, and enjoying nature with you always grounds me. **Judy**, your journey made me realize that there are a lot of alternative paths to success, and your strength to follow your wishes, always encouraged me and gave me a sense of security, that there is no wrong way. **Raphi**, you always pulled me to all the different corners in the world, and I am looking forward to visiting you in the yet closest destination in Switzerland. Thank you for being my friends, no matter what happens.

Of course, all this would not have been possible without my family. **Mama**, you raised me to be a strong, independent person who left the nest with joy. **Pupupapa**, you gave me the freedom to pursue my own path without worrying, including this PhD position in the Netherlands. My bonus Italian Mama **Lavinia**, since you joined our family 15 years ago, you have further strengthened my confidence and were always there with a delicious meal to fuel my batteries. All of you have always supported me—whether I was full of doubt and considering an entirely new direction, or highly motivated on track and working long hours. With your support behind me, I know I can achieve anything I want.

I am also thankful for my older siblings **Jonathan** and **Lea**, and my younger bonus sister **Laura**.

Lea and **Mark**, thank you for giving me two wonderful nieces during my PhD. Thanks to **Lina** and **Mali**, I saw you more often than I otherwise would have. **Lea**, you are an amazing mother and sister, and I'm grateful to have you. Being with you completely took my mind off work.

My Italian sister **Laura** (mein Amoore), and **Marko**, you have always supported my academic journey with outstanding celebrations—and especially you **Laura**, by providing a distraction with your seemingly endless master's thesis, which I was glad to support you with.

At this place, I would also like to acknowledge the newest furry member of our family, **Neo**. Thank you for successfully distracting me from writing this thesis in favor of extended walks and fetching and seeking "sein Ball". It is impossible to resist you, our little "Wunderhund" ("wonder dog").

Last, I want to thank my partner, **Jirka**. I am very grateful that we met in the last two years of my PhD. Although moving in with you in Wageningen might have seemed a bit quick, it always felt like the right thing to do – and I conveniently could reduce my commute time to university. With you, everything just fits effortlessly, may it be keeping the right orientation of the toilet paper role, or beating our enemies in *Dice Throne*. Thank you for motivating and supporting me to finish this important chapter of my life. I am curious and excited to see where our journey leads next. With you by my side, I know it will be amazing.

About the Author



Sophia in South Korea

Sophia Müller was born on 8 June 1994 in a small village in Germany. She completed her secondary education with the initial intention of pursuing veterinary medicine. However, an early internship led her to realize that her interests lay less in understanding disease and pathology and more in uncovering the fundamental principles by which biological systems function.

Following a six-month internship at a kindergarten in Gambia, where she gained exposure to diverse cultural and social contexts, she enrolled in the Molecular Life Sciences program in Salzburg. This interdisciplinary curriculum, combining biology, and chemistry, at the molecular level, strongly aligned with her scientific curiosity. After completing this joint degree between Salzburg and Linz, she travelled through parts of Southeast Asia and East Asia.

She continued her academic career with a Master's degree in Biology in Düsseldorf, during which she refined her focus on plant biology. Motivated by the conviction that sustainable food production represents one of the most direct ways in which biological research can contribute to society, she specialized in plant molecular biology. During her Master's research, she worked on the plant hormone ethylene, investigating components of its signaling pathway involved in fruit ripening.

After completing her Master's degree right before the pandemic, she temporarily stepped away from academic research, engaging in practical work such as house renovation and food cultivation. This period further strengthened her interest in applied biological systems. She subsequently began her doctoral research in Wageningen under the supervision of Wouter Kohlen, where her work integrates molecular biology with questions of agricultural relevance.



Neo, der Wunderhund

Her current interests increasingly extend toward applied research and knowledge translation, with the aim of contributing to solutions in sustainable agriculture and food systems. Alongside her scientific work, she remains committed to gaining experience across cultures and maintaining a close connection to nature.

List of Publications

Sophia Müller, Thijs Stegmann, Kelvin Adema, Rens Holmer, Amber van Seters, Robin van Velzen, Olga Kulikova, Tristan Wijsman, Joel Klein, Josefina-Patricia Fernandez-Moreno, Anna N. Stepanova, Jose M. Alonso, Henk Franssen, Estibaliz Larrainzar, Arjan van Zeijl, Wouter Kohlen (2025) Spatiotemporal dynamics of ethylene biosynthesis shape infection and nodule initiation in *Medicago truncatula*. *The Plant Cell*, *accepted*.

Ting Ting Xiao[*], **Sophia Müller**[*], Defeng Shen[*], Jieyu Liu, Kelvin Adema, Amber van Seters, Henk Franssen, Ton Bisseling, Olga Kulikova[+], Wouter Kohlen[+] (2025) Nodule organogenesis in Medicago truncatula requires local stage-specific auxin biosynthesis and transport. *Plant Physiology*, *Volume 197*, *Issue 4*, *kiaf133*, DOI: 10.1093/plphys/kiaf133. [*,+ contributed equally to this work]

Sophia Müller and Wouter Kohlen (2024) Jazzin' up nodules: The groovy role of jasmonic acid during nodulation, *Molecular Plant*, *17*, *Issue 11*, *p1639-1641*, *DOI: 10.1016/j.molp.2024.10.001*

Jana Wittmer, **Sophia Müller**, Rianne C. M. Kluck Kilian Duijts (2024) Bringing Europe's early career plant scientists back together, *New Phytologist*, *242*, *Issue 3*, *p878-880*, *DOI: 10.1111/nph.19665*

Alexander Hofmann, **Sophia Müller**, Thomas Drechsler, Mareike Berleth, Katharina Caesar, Leander Rohr, Klaus Harter, Georg Growth (2020) High-Level Expression, Purification and Initial Characterization of Recombinant Arabidopsis Histidine Kinase AHK1, *Plants*, *9*, *Issue 3*, *DOI: 10.3390/plants9030304*

The research described in this thesis was performed in the Laboratory of Cell and Developmental Biology and was supported by NWO-VIDI (VI.Vidi.193.119) granted to Wouter Kohlen. Financial support from Wageningen University for printing this thesis is gratefully acknowledged.

Cover design by Sophia Müller
Layout by Sophia Müller
Printed by ProefschriftMaken | proefschriftmaken.nl

