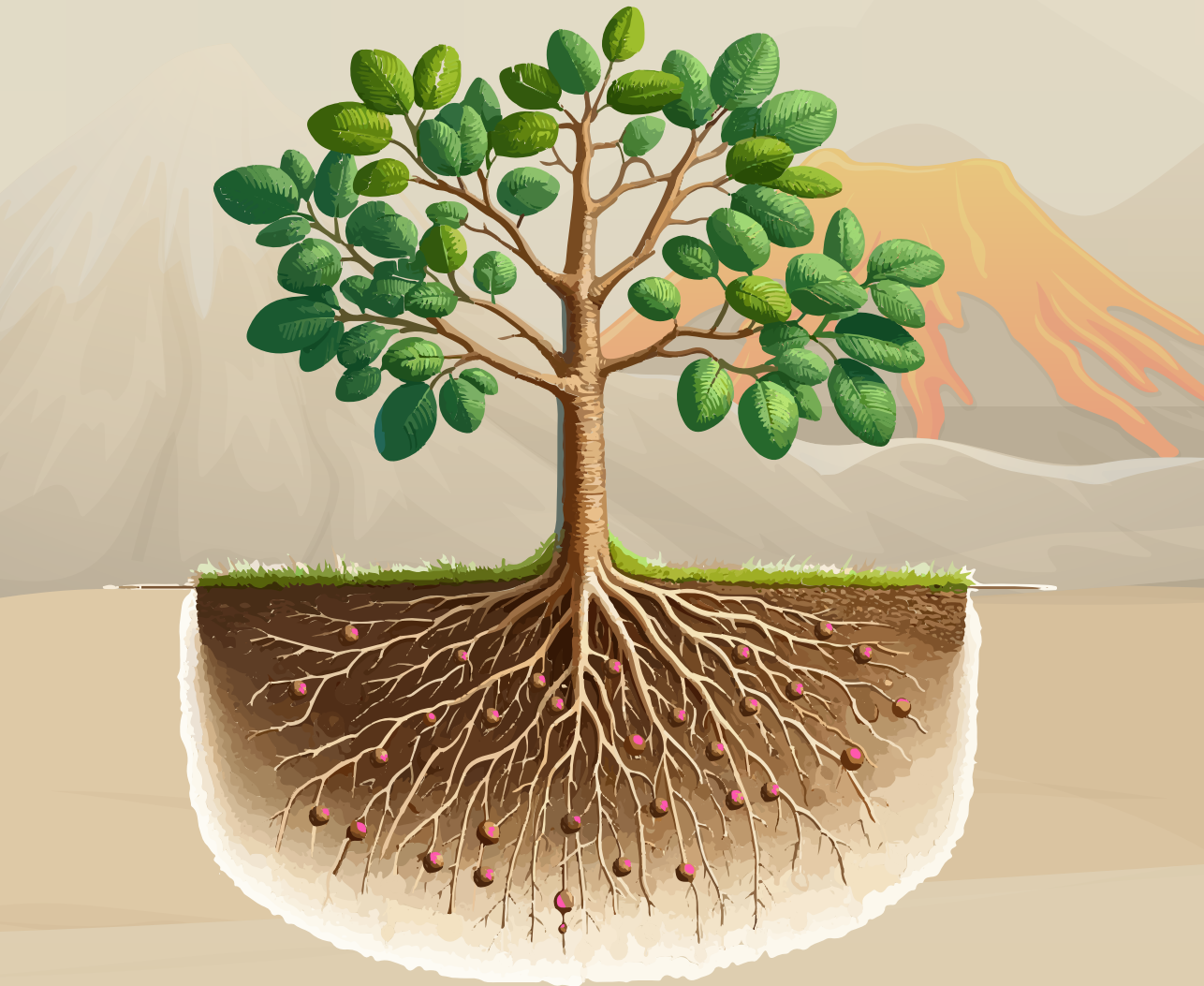
# Dissecting Symbiosis Permissiveness in *Parasponia andersonii*

Sultan Alhusayni



## Propositions

1. Ignoring rare non-canonical splice sites in ab initio gene prediction leads to misinterpretation of gene functionality.
   (this thesis)
2. The presence of the AON pathway in *Parasponia* provides a novel strategy for engineering nodulation in crops.
   (this thesis)
3. The trade-offs in plant development due to *NSP2* overexpression outweigh the benefits of increased mycorrhization (Li et al. 2022).
4. Vertical farming products demonstrate sustainability, but their high costs may worsen food deserts (Van Gerrewey et al. 2021).
5. COVID-19 lockdowns had a considerable positive impact on both society and the environment.
6. *"TOO MUCH LOVE* will kill you" DOES NOT hold for plants. (Brian May, Queen)

Propositions belonging to the thesis, entitled:

Dissecting Symbiosis Permissiveness in *Parasponia andersonii*

Sultan Alhusayni
Wageningen, 17 December 2024

# Dissecting Symbiosis Permissiveness in *Parasponia andersonii*

**Sultan Awwadh Sultan Alhusayni**

**Thesis committee**

**Promotors**
Prof. Dr T. Bisseling
Professor of Molecular biology
Wageningen University & Research

**Co-promotor**
Dr R. Geurts
Associate professor, Laboratory of Molecular Biology
Wageningen University & Research

**Other members**
Prof. Dr G. Smant, Wageningen University & Research
Dr R.V.Velzen, Wageningen University & Research
Dr K. Schiessl, University of Cambridge, UK
Dr A.V. Zeijl, Rijk Zwaan, Dinteloord

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

# Dissecting Symbiosis Permissiveness in *Parasponia andersonii*

**Sultan Awwadh Sultan Alhusayni**

**Thesis**

submitted in fulfillment 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 Tuesday 17 December 2024
at 1:00 p.m. in the Omnia Auditorium.

.

Sultan Awwadh Sultan Alhusayni
Dissecting Symbiosis Permissiveness in *Parasponia andersonii*,
250 Pages.

PhD thesis, Wageningen University, Wageningen, the Netherlands (2024)
with references, with summary in English.

ISBN: 978-94-6510-244-3
DOI: https://doi.org/10.18174/674020

*This thesis is dedicated to my family*

إهداء

إلى عائلتي الغالية

# Table of Contents

CHAPTER 1

# General Introduction

Sultan Alhusayni[1,2]

[1]Laboratory of Molecular Biology, Department of Plant Sciences, Wageningen University, Droevendaalsesteeg 1, 6708 PB Wageningen, The Netherlands

[2]Biological Sciences Department, College of Science, King Faisal University, Al-Ahsa 31982, Saudi Arabia

## Introduction

Nitrogen is the most abundant element in the atmosphere, yet it is often a limiting factor for plant growth and development. Farmers commonly use chemical nitrogen fertilizers to supplement the soil and support plant growth to address nitrogen deficiencies. However, excessive use of these fertilizers can have detrimental environmental effects. For instance, high nitrogen levels in the soil enhance microbial denitrification, leading to increased emissions of greenhouse gases such as nitrous oxide. Additionally, elevated nitrogen concentrations in aquatic ecosystems can cause eutrophication, characterized by the overgrowth of algae and aquatic plants. This nutrient enrichment disrupts aquatic ecosystems, reducing biodiversity (Martínez-Dalmau et al. 2021). A more sustainable approach to managing nitrogen deficiencies in soil involves plant-bacteria symbiosis, where nitrogen fixation is regulated by the plants' specific nitrogen needs (Peoples et al. 1995).



In general, plants interact with a wide range of beneficial microbes. For example, over 80% of land plants can form symbiotic relationships with arbuscular mycorrhizal (AM) fungi from the phylum Glomeromycota. According to fossil records, these fungi began colonizing land plants over 400 million years ago (Mya) (Remy et al. 1994; Smith and Read 2008). Another example is the symbiotic relationship between plants and diazotrophic bacteria, such as rhizobium (Gram-negative) and *Frankia* (Gram-positive), which can fix atmospheric dinitrogen ($N_2$) into ammonia and provide this to the plant (Benson and Silvester 1993; Spaink 1995; Frey-Klett et al. 2007; Oldroyd 2013). Rhizobium and *Frankia* bacteria form symbiotic relationships with only a limited number of flowering plant species that are taxonomically related. This so-called nitrogen-fixing clade (NFC) consists of four orders: Fabales, Fagales, Cucurbitales, and Rosales, comprising 28 families. Of these, only 10 families engage in symbiosis with nitrogen-fixing *Frankia* or rhizobium (Soltis et al. 1995; Geurts et al. 2012; Delaux et al. 2015; van Velzen et al. 2019). The rhizobium nitrogen-fixing symbiosis is restricted to

two groups of plant lineages. One group is represented by the legumes family (Fabaceae, order Fabales), which includes 750 genera containing about 20,000 species. The other group is defined by the genus *Parasponia* (Cannabaceae family, order Rosales), comprising five tropical tree species (Trinick 1973; Trinick and Galbraith 1976; Trinick 1979; Trinick and Galbraith 1980; Trinick and Hadobas 1988; Geurts et al. 2012; Delaux et al. 2015; van Velzen et al. 2019). In contrast, *Frankia* nitrogen-fixing symbiosis is found in eight families collectively known as actinorhizal plants. These families are dispersed among the orders Fagales (three families), Cucurbitales (two families), and Rosales (three families) (Geurts et al. 2012; Van Nguyen and Pawlowski 2017).

Both AM fungi and nitrogen-fixing bacteria can colonize the roots of plants and become involved in endosymbiosis, which is the most intimate plant-microbe interaction, hosting symbiont microbes inside the plant cells (Martin and Schwab 2012; Harris et al. 2020). In the AM symbiosis, the fungal hyphae colonize and penetrate the root epidermal layer, navigating towards the inner cortical cells. In the root cortex, the hyphae branch and colonize the apoplastic space. Subsequently, hyphae infect the cortical cells intracellularly and form feeding structures, so-called arbuscules, where fungi exchange phosphate, nitrogen and water for carbohydrates from the plants (Parniske 2008).

*Frankia* and rhizobium bacteria cannot penetrate the root in a fashion as AM fungi do. Instead, these bacteria trigger the formation of a de novo organ called a nodule, which originates from dividing cortical root cells. Physiologically, nodules provide the optimal environment to allow intracellular bacterial infection and to support the nitrogen fixation process by these bacteria. For example, in the case of legumes, *Parasponia* and some actinorhizal plants, nodules express a high level of leghemoglobin, which controls the oxygen homeostasis in the nodule, such that the bacterial nitrogenase enzyme complex is protected from oxidation (Ott et al. 2005; Geurts et al. 2016). To reach the dividing cortical cells that forms a nodule

primordium, rhizobium, and *Frankia* explore a range of infection strategies. These strategies range from crack entry in which apoplastic openings in the root are explored to sophisticated root hair-based infection (Pawlowski and Bisseling 1996). Which infection mechanism is employed is largely depending on the plant species.

Researchers have been trying to fully understand the mechanisms underpinning AM fungi and root nodule endosymbiosis for decades. Forward genetics has been employed to generate mutant populations in *Glycine max* (soybean) and *Pisum sativum* (pea) crops, as well as in two model species, *Medicago truncatula* and *Lotus japonicus* (Carroll et al. 1985; Duc and Messager 1989; Sagan et al. 1995; Szczyglowski et al. 1998; Penmetsa and Cook 2000; Madsen et al. 2005; Parniske 2008; Tadege et al. 2008; Fukai et al. 2012). Screening for impaired symbiotic phenotypes in these mutant populations revealed a list of genes involved in both symbiotic interactions, indicating that root nodule endosymbiosis recruited some of the genes utilized by the AM endosymbiosis. Furthermore, recent genomic and transcriptomic comparative studies reported the conservation of several intracellular infection signature genes in AM fungi symbiosis and across all known land plant lineages that have evolved an intracellular mutualistic relationship, such as arbuscular, ericoid and orchid mycorrhizal symbiosis, as well as root nodule symbiosis (Parniske 2008; Radhakrishnan et al. 2020). Taken together, these findings strongly indicate that genetic networks controlling the AM symbiosis were co-opted during the evolution of the nodule symbiosis.



**Evolutionary Route and Novelty of Root Nodule Symbiosis**

Early phylogenetic studies on the nodulating plant species using polymorphisms in the chloroplast gene *rbcL*, which encodes Ribulose 1,5-bisphosphate carboxylase/oxygenase (Rubisco), led to the coinage of the term the nitrogen-fixing clade (NFC), highlighting that nodulating plants species share a common ancestor (Soltis et al. 1995). The NFC clade consists

of non-nodulating lineages interspersed with ten nodulating lineages. For example, out of 10 genera in the Cannabaceae family, only species of the *Parasponia* genus can engage in root nodule symbiosis (Geurts et al. 2012). This interspersion of nodulating lineages among non-nodulating lineages led to the formulation of two hypotheses. The first hypothesis proposes that the common ancestor of the NFC gained a genetic predisposition, serving as a prerequisite for the evolution of the nodulation trait. Subsequently, ten nodulation-evolving events occurred independently. This multiple gain hypothesis predicts a gain of nodulation event in the legume family, in the *Parasponia* lineage, and up to 8 times in actinorhizal plant lineages. Since the symbiont and nodule ontogeny of legume, *Parasponia*, and actinorhizal types are different, it is seen as support for the occurrence of independent evolutionary origins of nodulation (Geurts et al. 2012; Geurts et al. 2016; Van Nguyen and Pawlowski 2017). Consequently, this predisposition-multiple gain of nodulation hypothesis is widely accepted and finds support even to date (Bruneau et al. 2008; Koenen et al. 2021).

On the contrary, the second hypothesis suggests that the common ancestor of the NFC evolved nodulation once, which was followed by an independent massive loss of nodulation in many of its subsequent lineages. This hypothesis, though initially less acknowledged, has gained increasing consideration in recent years as new findings emerge (Soltis et al. 1995).
Phylogenomic studies provide strong support for the “single gain massive loss” hypothesis by revealing parallel losses of putative orthologs of well-characterized genes, which are known to be essential for nodulation, in many non-nodulating plant species. For example, three key genes known to be essential for nodulation have been either lost or pseudogenized in plants that no longer have the ability to nodulate. These genes include a Lysin Motif receptor-like kinase that functions in rhizobial lipo-chitooligosaccharides (LCOs) perception, a transcription factor named NIN that acts as a master regulator of nodulation by controlling genes responsible for rhizobium infection and nodule organogenesis, and finally a regulator of root hair curling and infection thread progression toward nodule primordia

(Griesmann et al. 2018; van Velzen et al. 2018). Further insights into the evolution of root nodule symbiosis were gained through a genomic comparative study on nodulating and non-nodulating legume plant species, which revealed the conservation of a set of genes across all legume plant species that retained the root nodule symbiosis (Griesmann et al. 2018). Additionally, phylogenetic analysis of close homologs of well-characterized symbiotic genes, along with the nodule transcriptome of *M. truncatula* and *Parasponia andersonii* (*P. andersonii*) revealed a common utilization of 290 putative genes in root nodules. Among these, 26 genes were identified as putative orthologs of rhizobium-legume symbiosis genes, demonstrating the conservation of many genes governing rhizobium root nodule symbiosis in species within and outside the leguminous family (van Velzen et al. 2018). These findings collectively suggest that the common ancestor of the nitrogen-fixing clade was a nodulator, having co-opted components of existing gene networks to evolve this trait.



**What do we know about the last *Parasponia*-Legume common ancestor?**

In this thesis, I took the single gain massive loss hypothesis as the most likely scenario to explain the pattern of nodulating and non-nodulating plant lineages in the NFC. Accumulating evidence suggests that the first nodule in the last common ancestor of the nitrogen-fixing clade was the actinorhizal root nodule type induced by *Frankia* bacteria. A phylogenomic analysis of 3,500 species of flowering plants indicates that the earliest root nodule symbiosis evolved around 100 million years ago (Mya) (Werner et al. 2014; Persson et al. 2015). Additionally, fossil records revealed that the actinorhizal root nodule is the oldest nodule type. This is supported by the discovery of the oldest known nodule, which dates to the late Santonian (approximately 83.6 to 86.3 Mya) (Herendeen et al. 1999).

The filamentous actinobacteria genus Frankia comprises four taxonomic lineages (clusters I, II, III, and IV) capable of colonizing root nodules of actinorhizal plant species. However, regarding its nitrogen-fixing symbiosis

capacity, only clusters I, II, and III can fix nitrogen, not cluster IV. Recent phylogenetic analysis of 50 housekeeping genes across multiple genomes from *Frankia* revealed that cluster II is the basal *Frankia* group, with clusters I and III being its derived sister groups (Persson et al. 2015). Moreover, phylogenetic analysis of close homologs of the canonical rhizobial *nodABC* genes within the *Frankia* genus showed that homologs of these genes are conserved only in *Frankia* cluster II (strain Dg1) but not in clusters I and III. *nodA*, *nodB*, and *nodC* represent the core enzymes of the bio-synthetic machinery of LCOs in symbiotic rhizobia. LCOs are symbiotic signal molecules that activate nodule formation in most legumes and *Parasponia*. LCO signaling molecules haven't been detected in any *Frankia* strain; however, the *nodA*, *nodB* and *nodC* gene homologs of *Frankia* sp. Dg1 can functionally complement rhizobium *nodABC* mutants, suggesting that this *Frankia* cluster II strain can produce LCOs (Sen et al. 2014; Persson et al. 2015). It indicates that the common ancestor of the NFC evolved root nodule endosymbiosis with *Frankia* strains capable of producing LCOs, and this capacity was subsequently lost in descendant *Frankia* lineages in favor of a different mechanism to induce nodule formation in plants. In contrast, most rhizobia explore LCOs to communicate with their potential host plants and only a few bacterial strains switch to additional mechanisms. Studies in *Bradyrhizobium* species revealed that LCO-independent nodulation is based on symbiotic effector proteins that can activate the symbiotic signaling pathway in legumes (Fabre et al. 2015).

**Symbiosis signaling through LysM receptors**

Genetic screening of *M. truncatula* and *L. japonicus* mutant populations led to the discovery of a pair of Lysin Motif (LysM) receptors, which are involved in the perception of the rhizobial-secreted LCO molecules (also called Nodulation Factors or Nod Factors). The perception of rhizobia-produced LCOs through these receptors leads to the activation of the root nodule symbiotic signaling program (Dénarié et al. 1996; Oldroyd 2013). Moreover, the application of a low concentration of either purified or chemically synthesized LCOs corresponding to compatible rhizobia triggers root nodule

organogenesis on host plants, indicating that even in the absence of rhizobia, the perception of LCO signaling molecules by plant receptors is sufficient to initiate root nodule symbiosis (Dénarié et al. 1996; Demont-Caulet et al. 1999; Niwa et al. 2001).

LysM receptors are categorized into two clades/types based on the possession of a functional kinase domain (Clade I) or a non-functioning kinase domain (Clade II). *LysM Receptor Kinase 3* (*MtLYK3*) in *M. truncatula* and *Nod Factor Receptor 1* (*LjNFR1*, ortholog of *MtLYK3*) in *L. japonicus* belong to the LysM Clade I, whereas *Nod Factor Perception* (*MtNFP*) in *M. truncatula* and *Nod Factor Receptor 5* (*LjNFR5*, ortholog of *MtNFP*) in *L. japonicus* are members of the LysM Clade II. Moreover, mutant studies on legumes *MtLYK3/LjNFR1* and *MtNFP/LjNFR5* showed that both receptors are required for root nodule symbiosis, but not AM fungi symbiosis (Madsen et al. 2003; Radutoiu et al. 2003; Arrighi et al. 2006; Smit et al. 2007; Zhang et al. 2007). Similarly, in *Parasponia*, mutants of *PanLYK3*, ortholog of *MtLYK3/LjNFR1*, and *PanNFP2*, ortholog of *MtNFP/LjNFR5*, are required for root nodule symbiosis. However, unlike *MtLYK3/LjNFR1*, which is not required for AM symbiosis, *PanLYK3* is involved in AM fungi symbiosis and chitin immune response in *Parasponia* (Rutten et al. 2020). Furthermore, studies on LjNFR1 and LjNFR5 proteins showed that they are able to bind directly to a nanomolar range of LCOs and interact with each other forming a complex. In vitro, studies showed that *Lj*NFR1 functional kinase domain can trans-phosphorylate the pseudokinase domain of *Lj*NFR5, suggesting the formation of a heterodimeric complex that might be important for activating downstream signaling upon perception of LCOs (Madsen et al. 2011; Broghammer et al. 2012; Oldroyd 2013).

It is speculated that in legumes, LysM receptors function in the perception of LCOs produced by AM fungi, also called Mycorrhization Factors (Myc Factors). For example, in *M. truncatula* the application of a low concentration of the non-sulphated LCOs, which can be produced by the AM fungus *Rhizophagus irregularis,* induces lateral root formation, whereas this response is lost in the *Mtnfp* mutant background (Maillet et al. 2011).

However, analysis of the mycorrhization phenotype of *Mtnfp* and *Mtlyk3* mutants showed that the colonization level of both mutants is comparable to that of wild-type plants, indicating that their role in Myc Factors perception is redundant (Maillet et al. 2011). Moreover, in non-legume plants, it was shown that LysM-type receptors homologous *MtLYK3*/*LjNFR1* are functioning in AM symbiosis. For example, the Lysin Motif receptor-like kinase OsCERK1 in *Oryza sativa* (rice) is required for AM colonization, as it functions in the perception of short-chain chitin oligomers (COs) and LCOs signaling molecules. This is consistent with observations for *Parasponia* PanLYK3 (Miyata et al. 2014; Zhang et al. 2015; Carotenuto et al. 2017).

Symbiotic LysM receptors subsequently activate the so-called Common Symbiosis Signaling Pathway (CSSP) **(Figure 1)**. The CSSP component genes are positioned downstream of the LysM receptors in the symbiotic signaling pathway. The CSSP includes four signaling elements: an LRR-type transmembrane receptor (named LjSYMRK or MtDMI2), cation channels located in the nuclear envelope (named LjCASTOR and LjPOLLUX, and MtDMI1), a nuclear-localized calcium and calmodulin regulated kinase (LjCCAMK and MtDMI3), and a CCAMK interacting transcription factor (LjCYCLOPS and MtIPD3) (Catoira et al. 2000; Endre et al. 2002; Stracke et al. 2002; Ane et al. 2004; Kistner et al. 2005; Edwards et al. 2007; Charpentier et al. 2008; Mitra et al. 2004; Tirichine et al. 2006; Messinese et al. 2007; Yano et al. 2008). Further downstream of the CSSP, the signaling pathway diverges to initiate either AM symbiosis or root nodule symbiosis by inducing genes essential for each symbiosis (Oldroyd 2013).

One of the essential receptors involved in the perception of LCO signaling molecules is SYMRK, which is localized to the plasma membrane of epidermal cells (Stracke et al. 2002; Oldroyd 2013; Riely et al. 2013). In the absence of symbiotic signals, SYMRK protein undergoes cleavage of its extracellular malectin domain, making it an unstable protein that is rapidly degraded (Antolín-Llovera et al. 2014; Pan et al. 2018). This unstable form of SYMRK is strongly inclined to form extracellularly a heterodimeric

complex with the *MtNFP*/*LjNFR5* LysM-type receptor. This heterodimeric complex is highly stable and is involved in the perception of LCO signaling molecules (Antolín-Llovera et al. 2014). In addition to SYMRK's extracellular interaction with the *MtNFP*/*LjNFR5* type receptor, the cytoplasmic domain of SYMRK can interact with proteins, which are called SYMRK Interacting Protein (SIP1) and (SIP2) and SYMRK Interacting E3 Ubiquitin Ligase (SIE3). These components are implicated in early signaling (Zhu et al. 2008; Chen et al. 2012; Den Herder et al. 2012). Moreover, the SYMRK cytoplasmic domain can interact with and inhibit the well-known positive regulator of plant immunity BRASSINOSTEROID INSENSITIVE 1-Associated receptor Kinase 1 (LjBAK1), leading to the suppression of many immunity-related genes (Feng et al. 2021). This highlights the importance of SYMRK in LCO perception and the suppression of plant immunity.



*SYMRK* is essential for both AM fungi and root nodule symbioses. *symrk* mutants are unable to support AM fungi arbuscule formation or hyphae intracellular penetration. AM fungal hyphae on a *symrk* mutant plant exhibit hyphal abnormal swelling and arrest at the epidermal cell layer. In the context of nodulation, *symrk* knockout mutants fail to form infection pockets, to initiate infection threads and cortical cell divisions, or trigger calcium spiking in response to rhizobium inoculation (Catoira et al. 2000; Wais et al. 2000; Endre et al. 2002; Stracke et al. 2002). Interestingly, the R38 mutant in *M. truncatula* was found to have a single amino acid substitution (glycine to glutamic acid) in the *SYMRK* kinase domain, resulting in a truncated *SYMRK* protein missing the kinase domain. This mutant exhibited a surprising phenotype by obtaining normal AM fungi symbiosis but impaired root nodule symbiosis, indicating that the kinase domain of *SYMRK* could be dispensable for AM symbiosis (Endre et al. 2002).

*SYMRK* plays a role in the transcriptional activation of early nodulin genes that can be used as marker genes for early stages of infection. For example, the early nodulation marker leghemoglobin is activated within a few hours after LCO application; however, this activation is lost in the background of

the *symrk* mutant, affirming the involvement of *SYMRK* in the early responses induced by LCO signaling (Stracke et al. 2002). Subsequent

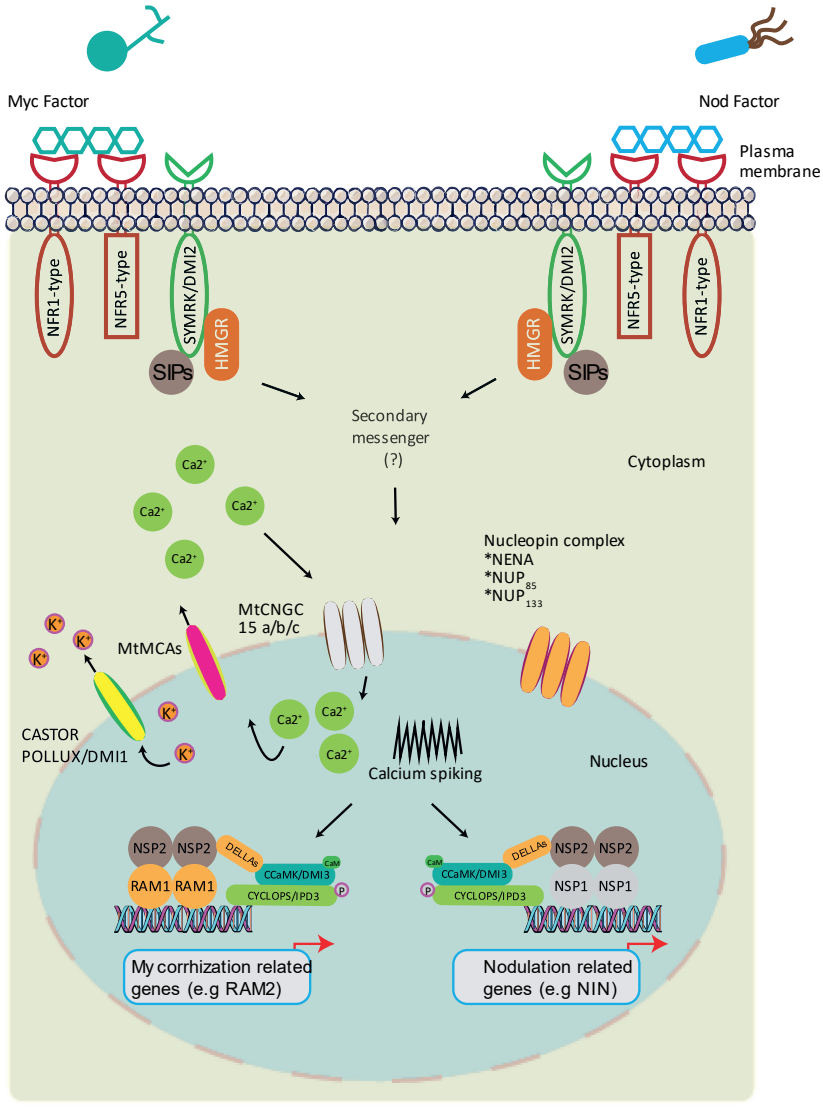



**Figure 1. Schematic representation of the common symbiotic signaling pathway (CSSP) in *Lotus japonicus* and *Medicago truncatula*.** LysM-type receptors on the plasma membrane perceive bacterial Nod factors or mycorrhizal Myc factors, triggering calcium spiking in the nucleus. This signal is decoded by CCaMK/DMI3, leading to the phosphorylation of CYCLOPS/IPD3. The CCaMK/DMI3-CYCLOPS/IPD3 complex interacts with DELLA and specific transcription factors to activate downstream genes. In nodulation, this complex interacts

with NSP1 and NSP2 to activate nodulation-related genes, while in mycorrhization, it interacts with NSP2 and RAM1 to activate mycorrhization-related genes (Genre and Russo 2016; Roy et al. 2020; Ren et al. 2022).

studies also showed that *SYMRK* also plays a role in the later stages of rhizobium infection. For instance, RNAi knockdown of the SYMRK ortholog in the nodule of *M. truncatula*, which forms root hair-based infection, and *Sesbania rostrata*, which forms crack-based infection, revealed the involvement of *SYMRK* in later stages of infection. This later-stage infection function is exhibited by the observed inhibition of bacterial release from the infection threads into nodule cells when *SYMRK* is knocked down (Capoen et al. 2005; Limpens et al. 2005).



Research on *SYMRK* gene expression has revealed that it is constitutively expressed in roots. For instance, Northern blotting studies showed that *SYMRK* mRNA expression was detected in *M. truncatula* roots and nodules, and this expression remained unaffected by rhizobium and nitrogen application (Bersoult et al. 2005). Furthermore, the spatial expression of the *SYMRK* promoter showed that *SYMRK* is constitutively expressed in the root under inoculated and non-inoculated conditions. For example, *M. truncatula* transgenic plants transformed with a construct carrying the β-glucuronidase (GUS) reporter gene driven by the promoter of *SYMRK* (a 3 kilobase (kb) region upstream of the transcription site) showed that under non-inoculation conditions, *SYMRK* was expressed along the whole root, except at the root tip. Histological sections of these roots showed that *SYMRK* expression occurred in the root hair, the epidermis, and the cortical cell layers (Bersoult et al. 2005). Moreover, in response to rhizobium inoculation, the 3 kb *SYMRK* promoter showed increased and more localized expression in the susceptible zone just above the root tip of emerging lateral roots. For instance, two days post-inoculation (2 dpi) showed that the *SYMRK* promoter expression was confined to discontinuous batches in the cortical cells of the susceptible zone, whereas at 3 and 5 dpi the expression significantly increased in the nodule primordium (Bersoult et al. 2005). A closer look at *SYMRK* spatial expression in the nodule showed that the gene is mainly expressed in the pre-infection zone between the meristematic and

infection zones. This pre-infection zone is marked by the expression localization of early nodulin genes such as *Early Nodulin 11* (*ENOD11*) and *ENOD12* (Pichon et al. 1992; Journet et al. 2001; Bersoult et al. 2005). Therefore, *SYMRK* is constitutively expressed in the root system at low levels under non-inoculation conditions, but upon rhizobium inoculation, this expression is markedly enhanced and restricted to the susceptible zone of the root and subsequently to the developing nodule.

Trans-complementation studies revealed the conservation of *SYMRK* function in nodulating and non-nodulating plant species. For example, the full-length *SYMRK* gene of plant species in the NFC such as *M. truncatula* and *Datisca glomerata*, as well as the full-length *SYMRK* gene of *Tropaeolum majus* in the sister clade of NFC, were able to fully complement nodulation and mycorrhization of the *L. japonicus Ljsymrk* mutant when driven by the native *LjSYMRK* promoter (Markmann et al. 2008). However, trans-complementation studies of *L. japonicus symrk* mutants using SYMRK homologous genes from species that are more distantly related successfully complement AM symbiosis but not root nodule symbiosis. For example, complementation studies on the *Ljsymrk* mutant using the full-length *SYMRK* gene of tomato and rice species were able to fully complement mycorrhization; however, in relation to nodulation, these genes were only able to complement nodule organogenesis but not infection. This failure to fully complement nodulation might be attributed to the shorter gene structure of *SYMRK* in tomato and rice which contain 14 and 12 exons, respectively (Markmann et al. 2008).

Deregulation of *SYMRK* expression in legumes rewires the symbiotic signaling pathway and can induce spontaneous nodule formation in the absence of rhizobia. For example, overexpression of *LjSYMRK* in the background of *Ljsymrk* mutant induces spontaneous nodules on non-inoculated plant roots (Ried et al. 2014). Another example is the overexpression of the *M. truncatula* and *Arachis hypogaea* intracellular domain of SYMRK, which was able to induce spontaneous nodulation on *M.*

*truncatula* TR25 (*dmi2* mutant) (Saha et al. 2014). However, overexpressing the intracellular domain of *SYMRK* led to a reduced number of properly infected nodules under rhizobial inoculation, indicating that the ectodomain of SYMRK is important for rhizobium recognition and proper infection (Saha et al. 2014). In **chapter 2,** I study *SYMRK* in *P. andersonii* and show that its function is conserved in both AM fungi and root nodule symbioses. Additionally, I show that ectopic expression of *P. andersonii SYMRK* induces spontaneous nodule formation in the absence of rhizobia.



In legumes, the perception of LCOs results in the induction of calcium oscillation in the nucleus of epidermal root cells. When *SYMRK* was mutated in both legumes and non-legumes, this calcium oscillation signal was completely abolished (Walker et al. 2000; Wais et al. 2000; Oldroyd et al. 2001; Kistner and Parniske 2002; Li et al. 2022). The molecular link between the perception of LCOs and the induction of calcium oscillation is not yet known. Research has shown that SYMRK can interact with 3-Hydroxy-3-Methylglutaryl Coenzyme A Reductase1 (HMGR1). This enzyme is involved in the catalysis of 3-hydroxy3-methyl-glutaryl-CoA into mevalonic acid, the rate-limiting step of the mevalonate pathway (Friesen and Rodwell 2004; Kevei et al. 2007). RNA silencing of *HMGR1* in *M. truncatula* results in abolishing of the calcium oscillation response, a severe reduction of nodule number, and low expression of early nodulin genes, such as *ENOD11* (Kevei et al. 2007; Venkateshwaran et al. 2015). Moreover, the application of mevalonic acid was able to induce calcium oscillation in various legume and non-legume plant species and rescued the reduced nodulation and calcium spiking phenotypes of *M. truncatula* lines carrying *HMGR1* silencing constructs (Kevei et al. 2007; Venkateshwaran et al. 2015). Therefore, it is concluded that HMGR1 is an integral part of the symbiotic signaling pathway.

Research investigating the symbiotic signaling pathway components directly involved in calcium oscillation led to the discovery of nuclear membrane-localized proteins. *LjPOLLUX* and *LjCASTOR* in *L. japonicus* and *Do Not Make*

*Infection 1* (*MtDMI1*, ortholog of *LjPOLLUX*) in *M. truncatula* are cation channels that play an essential role in calcium oscillation. These cation channels can form homodimers and are permeable for potassium (Charpentier et al. 2008; Ane et al. 2004). Lines carrying mutations in genes encoding these cation channels exhibit a defect in AM and root nodule symbioses, most probably due to the loss of induced calcium oscillation in response to LCO perception (Catoira et al. 2000; Kistner et al. 2005; Kosuta et al. 2008). Furthermore, a gain-of-function mutation in MtDMI1 (MtDMI1$^{S760N}$) in the absence of rhizobia and AM fungi results in constitutive activation of calcium oscillation, activation of nodulation- and mycorrhization-related genes, and the formation of spontaneous nodules. This observed effect of MtDMI1$^{S760N}$ requires functional three cyclic nucleotide-gated channels (CNGC15s) genes, which are localized to the nuclear membrane (Liu et al. 2022; Oldroyd 2013; Charpentier et al. 2016). Finally, deregulation of the *LjPOLLUX* and *LjCASTOR* genes in *L. japonicus* revealed their functional redundancy. For example, overexpression of *LjPOLLUX* in the background of *Ljcastor* revealed a trans-complementation ability of *LjPOLLUX* to rescue the nodulation phenotype of *Ljcastor* mutant (Charpentier et al. 2008).

The most downstream components of the CSSP are CCaMK and CYCLOPS, which are localized in the nucleus and positioned downstream of calcium oscillation. Legumes carrying mutations in these genes have retained the calcium oscillation response upon LCO perception, but lost the capacity to form AM and root nodule symbioses. Furthermore, *ccamk* and *cyclops* mutants have lost the early symbiotic response to rhizobium, such as root hair deformation related to infection and inner cortical cell division related to nodule initiation (Catoira et al. 2000; Messinese et al. 2007; Yano et al. 2008; Kistner et al. 2005). CCaMK primarily deciphers the induced calcium oscillation signal in the nucleus and forms a heterodimer with CYCLOPS protein (CCaMK-CYCLOPS) (Messinese et al. 2007; Yano et al. 2008). This heterodimer can activate downstream genes specific to nodulation. For example, an autoactive form of CCaMK (auto-CCaMK) can rewire the

symbiotic signaling pathway through the phosphorylation of CYCLOPS, which leads to the activation of nodulation-related genes such as *ENOD11*. In the absence of rhizobia, this activation results in the formation of spontaneous nodules in wild type and *symrk* mutant background (Gleason et al. 2006; Tirichine et al. 2006; Takeda et al. 2012). Similarly, the autoactive form of CYCLOPS (CYCLOPS-DD) can rewire the symbiotic signaling pathway and induce spontaneous nodules in wild type plants and in the background of *symrk* and *ccamk* mutants, indicating that CYCLOPS acts downstream of *SYMRK* and *CCaMK* and is the last component of the CSSP (Yano et al. 2008; Singh et al. 2014; Limpens and Bisseling 2014).



**Interplay of symbiotic signaling pathway components downstream of the CCSP**

Nodulation-related phenotype screening in *P. sativum*, *M. truncatula*, and *L. japonicus* mutant populations led to the discovery of two mutants that are defective in engaging in root nodule endosymbiosis but had only a mild phenotype in AM fungi endosymbiosis. Since these mutants show a strong impairment in nodulation, they were named *NODULATION SIGNALING PATHWAY 1* (*NSP1*) and *NSP2* (Catoira et al. 2000; Oldroyd and Long 2003; Gleason et al. 2006). Subsequently, reverse genetics knockout mutations in the putative orthologs of *NSP1* and *NSP2* in *P. andersonii* revealed that its symbiotic functioning is conserved (van Zeijl et al. 2018). *NSP1* and *NSP2* encode transcription factors containing a GRAS domain, named after Gibberellic Acid Insensitive (GAI), Repressor of GA1-3 (RGA), and Scarecrow (SCR) (Pysh et al. 1999; Kaló et al. 2005). GRAS proteins have a characteristic signature consisting of leucine heptad repeat I and II (LHRI and LHRII), VHIID, PFYRE, and SAW domains (Kaló et al. 2005; Hirsch et al. 2009).

Phylogenetic studies showed that NSP1 and NSP2 are conserved across diverse plant species. For example, it has been found that in the monocot *Oryza sativa* (which only engages in AM symbiosis), in dicot *M. truncatula* and *L. japonicus* (which engage in both AM and root nodule symbioses), and

in *Arabidopsis thaliana* (which lacks the ability to engage in both AM and root nodule symbiosis) at least a single copy of both NSP1 and NSP2 is present. This suggests that NSP1 and NSP2 functions are highly conserved in higher plants. For instance, trans-complementation studies have shown that *Oryza sativa NSP1* and *NSP2* possess the necessary molecular features to participate in nodulation-related processes, allowing them to restore the nodulation trait of *L. japonicus nsp1* and *nsp2* mutants, respectively (Heckmann et al. 2006; Liu et al. 2011; Yokota et al. 2010). Furthermore, *NSP1* and *NSP2* are involved in regulating genes associated with the carotenoid biosynthetic pathway and the downstream genetic pathway converting carotenoids into strigolactones and abscisic acid. For example, unlike in wild type plants, *PHYTOENE SYNTHASE* (*PSY*), the key rate-limiting enzyme of the carotenoid biosynthesis pathway, and *DWARF27* (*D27*), the first component gene of the strigolactone biosynthesis pathway, exhibit severely reduced gene expression in *nsp1* and *nsp2* mutants in *M. truncatula* and barley (*Hordeum vulgare*). Moreover, rice *nsp2* mutants showed significantly reduced expression levels of *D27* (Liu et al. 2011; Li et al. 2022; Kun Yuan et al. 2023). Taken together, *NSP1* and *NSP2* are highly conserved and involved in multiple processes in plants.

Studying the early signaling events in *nsp1* and *nsp2* mutants showed that both genes are involved in LCO signaling. For example, both *nsp1* and *nsp2* mutants respond to LCO application by inducing calcium oscillation in the nucleus and root hair deformation. However, no initiation of inner cortical cell divisions is observed in response to the LCOs, indicating that NSP1 and NSP2 might play similar roles in nodulation (Catoira et al. 2000; Tsyganov et al. 2002; Oldroyd and Long 2003; Kaló et al. 2005; Gleason et al. 2006; Heckmann et al. 2006). Furthermore, gain-of-function auto-CCaMK and CYCLOPS-DD experiments demonstrated that NSP1 and NSP2 are required for nodule organogenesis. For instance, in legumes, auto-CCaMK and CYCLOPS-DD can rewire the symbiotic signaling pathway, activating nodule organogenesis-related genes in the absence of rhizobia. This results in the formation of spontaneous nodules and the activation of *ENOD11* on the

roots of wild-type plants but not in *nsp1* and *nsp2* mutants (Gleason et al. 2006; Tirichine et al. 2006; Takeda et al. 2012; Limpens and Bisseling 2014; Singh et al. 2014). Surprisingly, CYCLOPS-DD can activate expression of the *NIN* gene independent of *NSP1* and *NSP2* in *L. japonicus*, suggesting that *NSP1* and *NSP2* are positioned downstream of or parallel to *NIN*. This discrepancy in the roles of NSP1 and NSP2 concerning nodule organogenesis and early symbiotic responses suggests they function as facilitators of symbiosis (Limpens and Bisseling 2014; Singh et al. 2014). This assumption is supported by studies on *MtNSP2* overexpression in *M. truncatula* and barley. For example, plants with constructs overexpressing *MtNSP2* in *M. truncatula* and barley showed upregulation of *SYMRK* and/or *CYCLOPS*, indicating a more complex role for NSP2. Therefore, this accumulated evidence suggests that *NSP2* plays a multifunctional role in symbiotic signaling, positioning it as a facilitator in the symbiotic signaling pathway (Li et al. 2022).



Research on *NSP1* and *NSP2* has provided insights into their tissue-specific expression and the subcellular localization of their proteins (Hirsch et al. 2009; Kang, H. *et al.* 2014). Both genes are mainly expressed in root tissue, and the encoded proteins co-localize in the nucleus. These data suggest that these transcription factors might physically interact. The co-expression of *MtNSP1* and *MtNSP2* proteins in *Nicotiana benthamiana* leaves and *Arabidopsis thaliana* protoplasts revealed that both proteins indeed can form a heterodimer (NSP1-NSP2) (Hirsch et al. 2009). Additionally, dissecting the essential domains of MtNSP2 for interaction with MtNSP1 led to the discovery that the LHRI domain of MtNSP2 is necessary to interact with MtNSP1, and the LHRI domain alone is sufficient for this interaction (Hirsch et al. 2009). Moreover, complementation of *Mtnsp2* mutants with a construct carrying an NSP2 version lacking the LHRI domain (NSP2 Δ-LHRI) failed to restore the nodulation phenotype, indicating that heterodimerization of NSP1 and NSP2 proteins might be required for nodule formation. Furthermore, a single amino acid substitution from alanine to valine in the LHRI domain of MtNSP2 led to a three-fold reduction in the

MtNSP1 and MtNSP2 interaction efficiency, which also resulted in a reduced number of nodules. Therefore, it is concluded that NSP1 and NSP2 heterodimerization is essential for nodulation (Hirsch et al. 2009).

In response to LCO perception and the subsequent induction of the symbiotic signaling cascade, NSP1 and NSP2, positioned downstream of the CSSP, form higher-order heterodimer complexes to activate genes specific to nodulation or mycorrhization. For instance, in response to rhizobium LCOs, NSP1-NSP2 can form a higher-order complex with CCaMK-IPD2, bridged by DELLA proteins. This complex can bind to and activate the promoter of the nodulation-related gene *ERF REQUIRED FOR NODULATION 1* (*ERN1*) (Hirsch et al. 2009; Fonouni-Farde et al. 2016; Jin et al. 2016). Furthermore, the NSP1-NSP2 heterodimer can form another higher-order complex with the INTERACTOR OF NSP2 PROTEIN (INP2) in response to rhizobium LCOs. This complex is capable of activating *NIN*, the master transcriptional regulator of nodulation, through direct binding of INP2 to the INP2-responsive element (INP2-RE) and NSP1 to the nodulation-responsive element (NRE) motifs in the *NIN* promoter (Hirsch et al. 2009; Xiao et al. 2020). Activation of these nodulation-specific transcription factors leads to transcriptional reprogramming of genes involved in rhizobium root nodule symbiosis. In contrast, the perception of AM fungal LCOs triggers the symbiotic signaling cascade, which promotes NSP2 (and likely NSP1) to form a heterodimer complex with the GRAS transcription factor specific to mycorrhization, REDUCED ARBUSCULAR MYCORRHIZA1 (RAM1). The formed NSP2-RAM1 heterodimer activates another mycorrhization-specific gene, *RAM2*, which facilitates AM colonization (Gobbato et al. 2012; Xue et al. 2015).

Deregulation of NSP2, but not NSP1, has revealed a role in promoting plant symbiotic permissiveness. For example, a construct using the *MtNSP2* native promoter to drive a modified version of the *MtNSP2* coding sequence (miRR-*MtNSP2*), where the recognition site for miRNA171h -a miRNA known to degrade *NSP2* messenger RNA (mRNA)- is silenced, exhibited enhanced AM

colonization under low phosphate conditions (Lauressergues et al. 2012). Furthermore, unlike control lines, transgenic lines transformed with the miRR-*MtNSP2* construct under a constitutive promoter showed enhanced mycorrhization even under high phosphate conditions. This indicates that overproduction and stabilization of *MtNSP2* mRNA from degradation by miRNA171h are sufficient to enhance plant AM fungi permissiveness (Li et al. 2022). In **Chapter 3**, I investigated the symbiotic and developmental phenotypes resulting from the overexpression of a miRNA171h-resistant version of *Parasponia NSP2* (*mNSP2*), demonstrating that *mNSP2* overexpression enhances mycorrhizal colonization under both low and high exogenous phosphate conditions. Additionally, I show that *mNSP2* overexpression negatively regulates root branching and root nodule formation in *Parasponia*, whereas its overexpression enhances shoot branching. Thus, the dual role of *Parasponia NSP2* highlights its critical involvement in balancing symbiotic interactions and plant development.



**Thesis outline**

The research presented in this thesis investigates how the manipulation of key genetic pathways involved in plant-microbe symbiosis can be used to engineer nodulation in non-nodulating plant species. Utilizing the nodulating *P. andersonii* and the non-nodulating *Trema orientalis* RG33 accession as experimental systems, the study focuses on three main aspects: (i) validating the symbiotic functionality of the *SYMRK* in the *Parasponia-Trema* lineages, (ii) evaluating the impact of overexpressing the GRAS transcription factor *NSP2* in *P. andersonii* to enhance plant symbiotic interactions, and (iii) investigating whether the Autoregulation of Nodulation (AON) mechanism, which regulates nodule number in legumes, is conserved in the non-legume *P. andersonii*.

**Chapter 2** involves comparative studies on mycorrhizal colonization dynamics in the roots of *P. andersonii* and *T. orientalis* RG33. While the overall mycorrhizal colonization level was lower in *T. orientalis RG33*

compared to *P. andersonii*, the number of arbuscule formed was comparable between the two species when examining 50 randomly selected 1 cm root segments. The conservation of *SYMRK* symbiotic functioning within the *Parasponia-Trema* lineage was investigated through two experiments. Initially, three independent CRISPR-Cas9-induced *Pansymrk* mutants in *P. andersonii* were generated, showing loss of both nodulation and mycorrhization capacity. Subsequently, overexpression of *PanSYMRK* under a constitutive promoter, in the absence of rhizobia, induced spontaneous root-like structures. A 'GA' mutation at the donor splice site of intron 12 in *SYMRK* of *T. orientalis* RG33 was identified, raising questions about its functionality. Testing through trans-complementation of a *Pansymrk* mutant with the 'GA' mutation restored nodulation and mycorrhization in the complemented roots, demonstrating the functionality of the *T. orientalis SYMR*K gene. The efficiency of the 'GA' splice site in *T. orientalis* RG33 was found to be approximately 95% compared to the more common 'GC' splice site of intron 12 in *P. andersonii*.

In **Chapter 3**, I explore the symbiotic benefits and trade-offs of overexpressing *mNSP2* in *P. andersonii*. Six transgenic lines with ectopic *mNSP2* expression were generated, showing varied effects on root and shoot developments. Elevated *mNSP2* expression positively correlated with increased mycorrhizal colonization under varying phosphate conditions. The role of NSP2 as a key regulator in the biosynthesis of strigolactones, important for mycorrhization, was further explored through RNA-seq analysis on roots of *mNSP2* overexpressing lines, revealing upregulation in the MEP and strigolactone biosynthesis pathways. An inverse relationship was observed between *mNSP2* expression levels and nodule formation. Abnormal cell divisions in the roots of lines with the highest *mNSP2* expression resembled enlarged pre-nodules, potentially serving as infection pockets for rhizobial entry.

**Chapter 4** reviews the autoregulation of nodulation (AON) mechanism, discussing its key genetic components and signaling networks that balance the advantages of nitrogen fixation with the plant's overall energy demands.

In **Chapter 5** I investigate the conservation of the AON mechanism in the non-legume *P. andersonii*. Using a CRISPR-Cas9-induced mutation system, two independent mutants of *P. andersonii CLAVATA1* (*PanCLV1*), analogous to *M. truncatula SUNN* and *L. japonicus HAR1*, were generated. These two mutants exhibited a hypernodulation phenotype and increased mycorrhizal colonization *P. andersonii*. Changes in root development were also noted, with an increase in primary root length and lateral root numbers.



**Chapter 6** summarizes and discusses the results obtained throughout the thesis, integrating these findings with existing knowledge in the field of nitrogen fixation. This chapter provides a forward-looking perspective on the potential for engineering nodulation in non-nodulating plant species.

**References**

**Ane, J.M. *et al.*** (2004) 'Medicago truncatula DMI1 required for bacterial and fungal symbioses in legumes', *Science*, **303**(5662), pp. 1364–1367.

**Antolín-Llovera, M., Ried, M.K. and Parniske, M.** (2014) 'Cleavage of the SYMBIOSIS RECEPTOR-LIKE KINASE Ectodomain Promotes Complex Formation with Nod Factor Receptor 5', *Current biology: CB*, **24**(4), pp. 422–427.

**Arrighi, J.-F. *et al.*** (2006) 'The Medicago truncatula Lysine Motif-Receptor-Like Kinase Gene Family Includes NFP and New Nodule-Expressed Genes', *Plant physiology*, **142**(1), pp. 265–279.

**Benson, D.R. and Silvester, W.B.** (1993) 'Biology of Frankia strains, actinomycete symbionts of actinorhizal plants', *Microbiological reviews*, **57**(2), pp. 293–319.

**Bersoult, A. *et al.*** (2005) 'Expression of the Medicago truncatula DM12 gene suggests roles of the symbiotic nodulation receptor kinase in nodules and during early nodule development', *Molecular plant-microbe interactions: MPMI*, **18**(8), pp. 869–876.

**Broghammer, A. *et al.*** (2012) 'Legume receptors perceive the rhizobial lipochitin oligosaccharide signal molecules by direct binding', *Proceedings of the National Academy of Sciences*, **109**(34), pp. 13859–13864.

**Bruneau, A. *et al.*** (2008) 'Phylogenetic patterns and diversification in the caesalpinioid legumesThis paper is one of a selection of papers published in the Special Issue on Systematics Research', *Botany*, **86**(7), pp. 697–718.

**Capoen, W. *et al.*** (2005) 'SrSymRK, a plant receptor essential for symbiosome formation', *Proceedings of the National Academy of Sciences of the United States of America*, **102**(29), pp. 10369–10374.

**Carotenuto, G. *et al.*** (2017) 'The rice LysM receptor-like kinase OsCERK1 is required for the perception of short-chain chitin oligomers in arbuscular mycorrhizal signaling', *The New phytologist*, **214**(4), pp. 1440–1446.

**Carroll, B.J., McNeil, D.L. and Gresshoff, P.M.** (1985) 'Isolation and properties of soybean [Glycine max (L.) Merr.] mutants that nodulate in the presence of high nitrate concentrations',

*Proceedings of the National Academy of Sciences of the United States of America*, **82**(12), pp. 4162–4166.

**Catoira, R. *et al.*** (2000) 'Four genes of Medicago truncatula controlling components of a nod factor transduction pathway', *The Plant cell*, **12**(9), pp. 1647–1666.

**Charpentier, M. *et al.*** (2008) 'Lotus japonicus CASTOR and POLLUX are ion channels essential for perinuclear calcium spiking in legume root endosymbiosis', *The Plant cell*, **20**(12), pp. 3467–3479.

**Charpentier, M. *et al.*** (2016) 'Nuclear-localized cyclic nucleotide-gated channels mediate symbiotic calcium oscillations', *Science*, **352**(6289), pp. 1102–1105.

**Chen, T. *et al.*** (2012) 'A MAP kinase kinase interacts with SymRK and regulates nodule organogenesis in Lotus japonicus', *The Plant cell*, **24**(2), pp. 823–838.

**Delaux, P.-M., Radhakrishnan, G. and Oldroyd, G**. (2015) 'Tracing the evolutionary path to nitrogen-fixing crops', *Current opinion in plant biology*, 26, pp. 95–99.

**Demont-Caulet, N. *et al.*** (1999) 'Nodule-inducing activity of synthetic Sinorhizobium meliloti nodulation factors and related lipo-chitooligosaccharides on alfalfa. Importance of the acyl chain structure', *Plant physiology*, **120**(1), pp. 83–92.

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

**Den Herder, G. *et al.*** (2012) 'Lotus japonicus E3 ligase SEVEN IN ABSENTIA4 destabilizes the symbiosis receptor-like kinase SYMRK and negatively regulates rhizobial infection', *The Plant cell*, **24**(4), pp. 1691–1707.

**Duc, G. and Messager, A.** (1989) 'Mutagenesis of pea (Pisum sativum L.) and the isolation of mutants for nodulation and nitrogen fixation', *Plant science: an international journal of experimental plant biology*, **60**(2), pp. 207–213.

**Edwards, A. *et al.*** (2007) 'Structural implications of mutations in the pea SYM8 symbiosis gene, the DMI1 ortholog, encoding a predicted ion channel', *Molecular plant-microbe interactions: MPMI*, **20**(10), pp. 1183–1191.

**Endre, G. *et al.*** (2002) 'A receptor kinase gene regulating symbiotic nodule development', *Nature*, **417**(6892), pp. 962–966.

**Fabre, S. *et al.*** (2015) 'Nod factor-independent nodulation in Aeschynomene evenia required the common plant-microbe symbiotic toolkit', *Plant physiology*, **169**(4), pp. 2654–2664.

**Feng, Y. *et al.*** (2021) 'Suppression of LjBAK1-mediated immunity by SymRK promotes rhizobial infection in Lotus japonicus', *Molecular plant*, 14(11), pp. 1935–1950.

**Fonouni-Farde, C. *et al.*** (2016) 'DELLA-mediated gibberellin signalling regulates Nod factor signalling and rhizobial infection', *Nature communications*, 7(1), pp. 1–13.

**Frey-Klett, P., Garbaye, J. and Tarkka, M.** (2007) 'The mycorrhiza helper bacteria revisited', *The New phytologist*, **176**(1), pp. 22–36.

**Friesen, J.A. and Rodwell, V.W.** (2004) 'The 3-hydroxy-3-methylglutaryl coenzyme-A (HMG-CoA) reductases', *Genome biology*, **5**(11), p. 248.

**Fukai, E. *et al.*** (2012) 'Establishment of a Lotus japonicus gene tagging population using the exon-targeting endogenous retrotransposon LORE1', *The Plant journal: for cell and molecular biology*, **69**(4), pp. 720–730.

**Geurts, R., Lillo, A. and Bisseling, T.** (2012) 'Exploiting an ancient signalling machinery to enjoy a nitrogen fixing symbiosis', *Current opinion in plant biology*, **15**(4), pp. 438–443.

**Geurts, R., Xiao, T.T. and Reinhold-Hurek, B.** (2016) 'What Does It Take to Evolve A Nitrogen-Fixing Endosymbiosis?', *Trends in plant science*, 21(3), pp. 199–208.

**Gleason, C. *et al.*** (2006) 'Nodulation independent of rhizobia induced by a calcium-activated kinase lacking autoinhibition', *Nature*, **441**(7097), pp. 1149–1152.

**Gobbato, E. *et al.*** (2012) 'A GRAS-type transcription factor with a specific function in mycorrhizal signaling', *Current biology: CB*, **22**(23), pp. 2236–2241.

**Griesmann, M. *et al.*** (2018) 'Phylogenomics reveals multiple losses of nitrogen-fixing root nodule symbiosis', *Science*, **361**(6398), p. eaat1743.

**Harris, J.M., Pawlowski, K. and Mathesius, U.** (2020) 'Editorial: Evolution of Signaling in Plant Symbioses', *Frontiers in plant science*, **11**, p. 456.

**Heckmann, A.B. *et al.*** (2006) 'Lotus japonicus nodulation requires two GRAS domain regulators, one of which is functionally conserved in a non-legume', *Plant physiology*, **142**(4), pp. 1739–1750.

**Herendeen, P.S. *et al.*** (1999) 'A Preliminary Conspectus of the Allon Flora from the Late Cretaceous (Late Santonian) of Central Georgia, U.S.A', *Annals of the Missouri Botanical Garden. Missouri Botanical Garden*, **86**(2), pp. 407–471.

**Hirsch, S. *et al.*** (2009) 'GRAS proteins form a DNA binding complex to induce gene expression during nodulation signaling in Medicago truncatula', *The Plant cell*, **21**(2), pp. 545–557.

**Jin, Y. *et al.*** (2016) 'DELLA proteins are common components of symbiotic rhizobial and mycorrhizal signalling pathways', *Nature communications*, 7(1), pp. 1–14.

**Journet, E.P. *et al.*** (2001) 'Medicago truncatula ENOD11: a novel RPRP-encoding early nodulin gene expressed during mycorrhization in arbuscule-containing cells', *Molecular plant-microbe interactions: MPMI*, **14**(6), pp. 737–748.

**Kaló, P. *et al.*** (2005) 'Nodulation signaling in legumes requires NSP2, a member of the GRAS family of transcriptional regulators', *Science*, **308**(5729), pp. 1786–1789.

**Kang, H. *et al.*** (2014) 'A MYB coiled-coil transcription factor interacts with NSP2 and is involved in nodulation in Lotus japonicus', *The New phytologist*, **201**(3), pp. 837–849.

**Kevei, Z. *et al.*** (2007) '3-hydroxy-3-methylglutaryl coenzyme a reductase 1 interacts with NORK and is crucial for nodulation in Medicago truncatula', *The Plant cell*, **19**(12), pp. 3974–3989.

**Kistner, C. *et al.*** (2005) 'Seven Lotus japonicus genes required for transcriptional reprogramming of the root during fungal and bacterial symbiosis', *The Plant cell*, **17**(8), pp. 2217–2229.

**Kistner, C. and Parniske, M.** (2002) 'Evolution of signal transduction in intracellular symbiosis', *Trends in plant science*, **7**(11), pp. 511–518.

**Koenen, E.J.M. *et al.*** (2021) 'The origin of the legumes is a complex paleopolyploid phylogenomic tangle closely associated with the Cretaceous-Paleogene (K-Pg) mass extinction event', *Systematic biology*, 70(3), pp. 508–526.

**Kosuta, S. *et al.*** (2008) 'Differential and chaotic calcium signatures in the symbiosis signaling pathway of legumes', *Proceedings of the*

*National Academy of Sciences of the United States of America*, **105**(28), pp. 9823–9828.

**Kun Yuan *et al.*** (2023) 'Low phosphorus promotes NSP1-NSP2 heterodimerization to enhance strigolactone biosynthesis and regulate shoot and root architecture in rice', *Molecular plant*, **16**(11), pp. 1811–1831.

**Lauressergues, D. *et al.*** (2012) 'The microRNA miR171h modulates arbuscular mycorrhizal colonization of Medicago truncatula by targeting NSP2', *The Plant journal: for cell and molecular biology*, **72**(3), pp. 512–522.

**Limpens, E. *et al.*** (2005) 'Formation of organelle-like N2-fixing symbiosomes in legume root nodules is controlled by DMI2', *Proceedings of the National Academy of Sciences of the United States of America*, **102**(29), pp. 10375–10380.

**Limpens, E. and Bisseling, T.** (2014) 'CYCLOPS: a new vision on rhizobium-induced nodule organogenesis', *Cell host & microbe*, **15**(2), pp. 127–129.

**Liu, H. *et al.*** (2022) 'Constitutive activation of a nuclear-localized calcium channel complex in Medicago truncatula', *Proceedings of the National Academy of Sciences of the United States of America*, **119**(34), p. e2205920119.

**Liu, W. *et al.*** (2011) 'Strigolactone Biosynthesis in Medicago truncatula and Rice Requires the Symbiotic GRAS-Type Transcription Factors NSP1 and NSP2', *The Plant cell*, 23(10), pp. 3853–3865.

**Li, X.-R. *et al.*** (2022) 'Nutrient regulation of lipochitooligosaccharide recognition in plants via NSP1 and NSP2', *Nature communications*, **13**(1), p. 6421.

**Madsen, E.B. *et al.*** (2003) 'A receptor kinase gene of the LysM type is involved in legume perception of rhizobial signals', *Nature*, 425(6958), pp. 637–640.

**Madsen, E.B. *et al.*** (2011) 'Autophosphorylation is essential for the in vivo function of the Lotus japonicus Nod factor receptor 1 and receptor-mediated signalling in cooperation with Nod factor receptor 5', *The Plant journal: for cell and molecular biology*, **65**(3), pp. 404–417.

**Madsen, L.H. *et al.*** (2005) 'LORE1, an active low-copy-number TY3-gypsy retrotransposon family in the model legume Lotus japonicus', *The Plant journal: for cell and molecular biology*, **44**(3), pp. 372–381.

**Maillet, F. *et al.*** (2011) 'Fungal lipochitooligosaccharide symbiotic signals in arbuscular mycorrhiza', *Nature*, 469(7328), pp. 58–63.

**Markmann, K., Giczey, G. and Parniske, M.** (2008) 'Functional Adaptation of a Plant Receptor- Kinase Paved the Way for the Evolution of Intracellular Root Symbioses with Bacteria', *PLoS biology*, **6**(3), p. e68.

**Martin, B. and Schwab, E.** (2012) 'Current Usage of Symbiosis and Associated Terminology', *International journal of biology*, **5**(1), p. 32.

**Martínez-Dalmau, J., Berbel, J. and Ordóñez-Fernández, R.** (2021) 'Nitrogen Fertilization. A Review of the Risks Associated with the Inefficiency of Its Use and Policy Responses', *Sustainability: Science Practice and Policy*, **13**(10), p. 5625.

**Messinese, E. *et al.*** (2007) 'A novel nuclear protein interacts with the symbiotic DMI3 calcium- and calmodulin-dependent protein kinase of Medicago truncatula', *Molecular plant-microbe interactions: MPMI*, **20**(8), pp. 912–921.

**Mitra, R.M., Shaw, S.L. and Long, S.R.** (2004) 'Six nonnodulating plant mutants defective for Nod factor-induced transcriptional changes associated with the legume-rhizobia symbiosis', *Proceedings of the National Academy of Sciences of the United States of America*, **101**(27), pp. 10217–10222.

**Miyata, K. *et al.*** (2014) 'The bifunctional plant receptor, OsCERK1, regulates both chitin-triggered immunity and arbuscular mycorrhizal symbiosis in rice', *Plant & cell physiology*, 55(11), pp. 1864–1872.

**Niwa, S. *et al.*** (2001) 'Responses of a model legume Lotus japonicus to lipochitin oligosaccharide nodulation factors purified from Mesorhizobium loti JRL501', *Molecular plant-microbe interactions: MPMI*, **14**(7), pp. 848–856.

**Oldroyd, G.E. *et al.*** (2001) 'Evidence for structurally specific negative feedback in the Nod factor signal transduction pathway', *The Plant journal: for cell and molecular biology*, **28**(2), pp. 191–199.

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

**Oldroyd, G.E.D. and Long, S.R**. (2003) 'Identification and characterization of nodulation-signaling pathway 2, a gene of Medicago truncatula

involved in Nod actor signaling', *Plant physiology*, **131**(3), pp. 1027–1032.

**Ott, T. *et al.*** (2005) 'Symbiotic leghemoglobins are crucial for nitrogen fixation in legume root nodules but not for general plant growth and development', *Current biology: CB*, **15**(6), pp. 531–535.

**Pan, H., Stonoha-Arther, C. and Wang, D.** (2018) 'Medicago Plants Control Nodulation by Regulating Proteolysis of the Receptor-Like Kinase DMI2', *Plant physiology*, 177(2), pp. 792–802.

**Parniske, M.** (2008) 'Arbuscular mycorrhiza: the mother of plant root endosymbioses', *Nature reviews. Microbiology*, 6(10), pp. 763–775.

**Pawlowski, K. and Bisseling, T.** (1996) 'Rhizobial and Actinorhizal Symbioses: What Are the Shared Features?', *The Plant cell*, **8**(10), pp. 1899–1913.

**Penmetsa, R.V. and Cook, D.R.** (2000) 'Production and characterization of diverse developmental mutants of Medicago truncatula', *Plant physiology*, **123**(4), pp. 1387–1398.

**Peoples, M.B., Herridge, D.F. and Ladha, J.K.** (1995) 'Biological nitrogen fixation: An efficient source of nitrogen for sustainable agricultural production?', in J.K. Ladha and M.B. Peoples (eds) *Management of Biological Nitrogen Fixation for the Development of More Productive and Sustainable Agricultural Systems: Extended versions of papers presented at the Symposium on Biological Nitrogen Fixation for Sustainable Agriculture at the 15th Congress of Soil Science, Acapulco, Mexico, 1994*. Dordrecht: Springer Netherlands, pp. 3–28.

**Persson, T. *et al.*** (2015) 'Candidatus Frankia Datiscae Dg1, the Actinobacterial Microsymbiont of Datisca glomerata, Expresses the Canonical nod Genes nodABC in Symbiosis with Its Host Plant', *PloS one*, **10**(5), p. e0127630.

**Pichon, M. *et al.*** (1992) 'Rhizobium meliloti elicits transient expression of the early nodulin gene ENOD12 in the differentiating root epidermis of transgenic alfalfa', *The Plant cell*, **4**(10), pp. 1199–1211.

**Pysh, L.D. *et al.*** (1999) 'The GRAS gene family in Arabidopsis: sequence characterization and basic expression analysis of the SCARECROW-LIKE genes', *The Plant journal: for cell and molecular biology*, **18**(1), pp. 111–119.

**Radhakrishnan, G.V. *et al.*** (2020) 'An ancestral signalling pathway is conserved in intracellular symbioses-forming plant lineages', *Nature plants*, 6(3), pp. 280–289.

**Radutoiu, S. *et al.*** (2003) 'Plant recognition of symbiotic bacteria requires two LysM receptor-like kinases', *Nature*, 425(6958), pp. 585–592.

**Remy, W. *et al.*** (1994) 'Four hundred-million-year-old vesicular arbuscular mycorrhizae', *Proceedings of the National Academy of Sciences of the United States of America*, **91**(25), pp. 11841–11843.

**Ried, M.K., Antolín-Llovera, M. and Parniske, M.** (2014) 'Spontaneous symbiotic reprogramming of plant roots triggered by receptor-like kinases', *eLife*, **3**. Available at: https://doi.org/10.7554/eLife.03891.

**Riely, B.K. *et al.*** (2013) 'Development of tools for the biochemical characterization of the symbiotic receptor-like kinase DMI2', *Molecular plant-microbe interactions: MPMI*, 26(2), pp. 216–226.

**Rutten, L. *et al.*** (2020) 'Duplication of Symbiotic Lysin Motif Receptors Predates the Evolution of Nitrogen-Fixing Nodule Symbiosis', *Plant physiology*, **184**(2), pp. 1004–1023.

**Sagan, M. *et al.*** (1995) 'Selection of nodulation and mycorrhizal mutants in the model plant Medicago truncatula (Gaertn.) after γ-ray mutagenesis', *Plant science: an international journal of experimental plant biology*, **111**(1), pp. 63–71.

**Saha, S. *et al.*** (2014) 'Intracellular catalytic domain of symbiosis receptor kinase hyperactivates spontaneous nodulation in absence of rhizobia', *Plant physiology*, **166**(4), pp. 1699–1708.

**Sen, A. *et al.*** (2014) 'Phylogeny of the class Actinobacteria revisited in the light of complete genomes. The orders "Frankiales" and Micrococcales should be split into coherent entities: proposal of Frankiales ord. nov., Geodermatophilales ord. nov., Acidothermales ord. nov. and Nakamurellales ord. nov', *International journal of systematic and evolutionary microbiology*, **64**(Pt 11), pp. 3821–3832.

**Singh, S. *et al.*** (2014) 'CYCLOPS, a DNA-binding transcriptional activator, orchestrates symbiotic root nodule development', *Cell host & microbe*, **15**(2), pp. 139–152.

**Smith, S.E. and Read, D.J.** (2008) 'Mycorrhizal Symbiosis, 3rd editio Edn'. Cambridge, MA: Academic Press. doi.

**Smit, P. *et al.*** (2007) 'Medicago LYK3, an Entry Receptor in Rhizobial Nodulation Factor Signaling', *Plant physiology*, **145**(1), pp. 183–191.

**Soltis, D.E. *et al.*** (1995) 'Chloroplast gene sequence data suggest a single origin of the predisposition for symbiotic nitrogen fixation in angiosperms', *Proceedings of the National Academy of Sciences of the United States of America*, **92**(7), pp. 2647–2651.

**Spaink, H.P.** (1995) 'The molecular basis of infection and nodulation by rhizobia: the ins and outs of sympathogenesis', *Annual review of phytopathology*, **33**, pp. 345–368.

**Stracke, S. *et al.*** (2002) 'A plant receptor-like kinase required for both bacterial and fungal symbiosis', *Nature*, **417**(6892), pp. 959–962.

**Szczyglowski, K. *et al.*** (1998) 'Nodule Organogenesis and Symbiotic Mutants of the Model Legume Lotus japonicus', *Molecular plant-microbe interactions: MPMI*, **11**(7), pp. 684–697.

**Tadege, M. *et al.*** (2008) 'Large-scale insertional mutagenesis using the Tnt1 retrotransposon in the model legume Medicago truncatula', *The Plant journal: for cell and molecular biology*, **54**(2), pp. 335–347.

**Takeda, N., Maekawa, T. and Hayashi, M.** (2012) 'Nuclear-localized and deregulated calcium- and calmodulin-dependent protein kinase activates rhizobial and mycorrhizal responses in Lotus japonicus', *The Plant cell*, **24**(2), pp. 810–822.

**Tirichine, L. *et al.*** (2006) 'Deregulation of a Ca2+/calmodulin-dependent kinase leads to spontaneous nodule development', *Nature*, **441**(7097), pp. 1153–1156.

**Trinick, M.J.** (1973) 'Symbiosis between Rhizobium and the Non-legume, Trema aspera', *Nature*, **244**(5416), pp. 459–460.

**Trinick, M.J.** (1979) 'Structure of nitrogen-fixing nodules formed by Rhizobium on roots of Parasponia andersonii Planch', *Canadian journal of microbiology*, 25(5), pp. 565–578.

**Trinick, M.J. and Galbraith, J.** (1976) 'Structure of root nodules formed by Rhizobium on the non-legume Trema cannabina var. scabra', *Archives of microbiology*, **108**(2), pp. 159–166.

**Trinick, M.J. and Galbraith, J.** (1980) 'THE *RHIZOBIUM* REQUIREMENTS OF THE NON-LEGUME *PARASPONIA* IN RELATIONSHIP TO THE CROSS-INOCULATION GROUP CONCEPT OF LEGUMES', *The New phytologist*, **86**(1), pp. 17–26.

**Trinick, M.J. and Hadobas, P.A.** (1988) 'Biology of theParasponia-Bradyrhizobium symbiosis', *Plant and soil*, **110**(2), pp. 177–185.

**Tsyganov, V.E. *et al.*** (2002) 'Genetic Dissection of the Initiation of the Infection Process and Nodule Tissue Development in the Rhizobium–Pea (Pisum sativum L.) Symbiosis', *Annals of botany*, **89**(4), pp. 357–366.

**Van Nguyen, T. and Pawlowski, K.** (2017) 'Frankia and Actinorhizal Plants: Symbiotic Nitrogen Fixation', in S. Mehnaz (ed.) *Rhizotrophs: Plant Growth Promotion to Bioremediation*. Singapore: Springer Singapore, pp. 237–261.

**van Velzen, R. *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 of the United States of America*, **115**(20), pp. 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 in plant science*, **24**(1), pp. 49–57.

**Venkateshwaran, M. *et al.*** (2015) 'A role for the mevalonate pathway in early plant symbiotic signaling', *Proceedings of the National Academy of Sciences*, **112**(31), pp. 9781–9786.

**Wais, R.J. *et al.*** (2000) 'Genetic analysis of calcium spiking responses in nodulation mutants of Medicago truncatula', *Proceedings of the National Academy of Sciences of the United States of America*, **97**(24), pp. 13407–13412.

**Walker, S.A., Viprey, V. and Downie, J.A.** (2000) 'Dissection of nodulation signaling using pea mutants defective for calcium spiking induced by nod factors and chitin oligomers', *Proceedings of the National Academy of Sciences of the United States of America*, **97**(24), pp. 13413–13418.

**Werner, G.D.A. *et al.*** (2014) 'A single evolutionary innovation drives the deep evolution of symbiotic N2-fixation in angiosperms', *Nature communications*, **5**(1), pp. 1–9.

**Xiao, A. *et al.*** (2020) 'Transcriptional regulation of NIN expression by IPN2 is required for root nodule symbiosis in Lotus japonicus', *The New phytologist*, **227**(2), pp. 513–528.

**Xue, L. *et al.*** (2015) 'Network of GRAS Transcription Factors Involved in the Control of Arbuscule Development in Lotus japonicus', *Plant physiology*, **167**(3), pp. 854–871.

**Yano, K. *et al.*** (2008) 'CYCLOPS, a mediator of symbiotic intracellular accommodation', *Proceedings of the National Academy of Sciences*, **105**(51), pp. 20540–20545.

**Yokota, K. *et al.*** (2010) 'Function of GRAS proteins in root nodule symbiosis is retained in homologs of a non-legume, rice', *Plant & cell physiology*, **51**(9), pp. 1436–1442.

**van Zeijl, A. *et al.*** (2018) 'CRISPR/Cas9-Mediated Mutagenesis of Four Putative Symbiosis Genes of the Tropical Tree Parasponia andersonii Reveals Novel Phenotypes', *Frontiers in plant science*, **9**, p. 284.

**Zhang, X. *et al.*** (2015) 'The receptor kinase CERK1 has dual functions in symbiosis and immunity signalling', *The Plant journal: for cell and molecular biology*, **81**(2), pp. 258–267.

**Zhang, X.-C. *et al.*** (2007) 'Molecular evolution of lysin motif-type receptor-like kinases in plants', *Plant physiology*, **144**(2), pp. 623–636.

**Zhu, H. *et al.*** (2008) 'A novel ARID DNA-binding protein interacts with SymRK and is expressed during early nodule development in Lotus japonicus', *Plant physiology*, **148**(1), pp. 337–347.

CHAPTER 2

# 2

# A rare non-canonical splice site in *Trema orientalis SYMRK* does not affect its dual symbiotic functioning in endomycorrhiza and rhizobium nodulation

Sultan Alhusayni[1,2a] | Yuda Purwana Roswanjaya[1,3a] | Luuk Rutten[1] | Rik Huisman[1] | Simon Bertram[1] | Trupti Sharma[1] | Michael Schon[1] | Wouter Kohlen[1] | Joël Klein[1] | Rene Geurts[1]

[1]Laboratory of Molecular Biology, Cluster of Plant Development, Plant Science Group, Wageningen University, Droevendaalsesteeg 1, 6708 PB Wageningen, The Netherlands

[2]Biological Sciences Department, College of Science, King Faisal University, Al-Ahsa 31982, Saudi Arabia

[3]Research Centre for Applied Microbiology, National Research and Innovation Agency (BRIN), Cibinong 16911, Indonesia.

[a]These authors contributed equally to this work.

*Published in BMC plant biology, 23(1), p. 587*
*doi: 10.1186/s12870-023-04594-0 (2023)*

## ABSTRACT

**Background:** Nitrogen-fixing nodules occur in ten related taxonomic lineages interspersed with lineages of non-nodulating plant species. Nodules result from an endosymbiosis between plants and diazotrophic bacteria; rhizobia in the case of legumes and *Parasponia* and *Frankia* in the case of actinorhizal species. Nodulating plants share a conserved set of symbiosis genes, whereas related non-nodulating sister species show pseudogenization of several key nodulation-specific genes. Signaling and cellular mechanisms critical for nodulation have been co-opted from the more ancient plant-fungal arbuscular endomycorrhizal symbiosis. Studies in legumes and actinorhizal plants uncovered a key component in symbiotic signaling, the LRR-type SYMBIOSIS RECEPTOR KINASE (*SYMRK*). *SYMRK* is essential for nodulation and arbuscular endomycorrhizal symbiosis. To our surprise, however, despite its arbuscular endomycorrhizal symbiosis capacities, we observed a seemingly critical mutation in a donor splice site in the *SYMRK* gene of *Trema orientalis,* the non-nodulating sister species of *Parasponia*. This led us to investigate the symbiotic functioning of *SYMRK* in the *Trema-Parasponia* lineage and to address the question of to what extent a single nucleotide polymorphism in a donor splice site affects the symbiotic functioning of *SYMRK*.

**Results**: We show that *SYMRK* is essential for nodulation and endomycorrhization in *Parasponia andersonii*. Subsequently, it is revealed that the 5'-intron donor splice site of *SYMRK* intron 12 is variable and, in most dicotyledon species, doesn't contain the canonical dinucleotide 'GT' signature but the much less common motif 'GC'. Strikingly, in *T. orientalis,* this motif is converted into a rare non-canonical 5'-intron donor splice site 'GA'. This *SYMRK* allele, however, is fully functional and spreads in the *T. orientalis* population of Malaysian Borneo. A further investigation into the occurrence of the non-canonical GA-AG splice sites confirmed that these are extremely rare.

**Conclusion**: *SYMRK* functioning is highly conserved in legumes, actinorhizal plants, and *Parasponia*. The gene possesses a non-common 5'-intron GC donor splice site in intron 12, which is converted into a GA in *T. orientalis* accessions of Malaysian Borneo. The discovery of this functional GA-AG splice site in *SYMRK* highlights a gap in our understanding of splice donor sites.

## Keywords

Non-canonical splice site, SYMRK, LRR-type transmembrane receptor kinase, Mutualistic endosymbiosis, Nitrogen-fixing nodulation symbiosis, Arbuscular mycorrhizal symbiosis, *Parasponia andersonii*, *Trema orientalis*, Common symbiosis signaling pathway, plant evolution.

## Introduction

Plants have evolved a range of mutualistic endosymbiotic partnerships with microbes to enhance nutrient uptake. The most ancient mutualistic endosymbiosis is the interaction between plant roots and Glomeromycota fungi, also known as arbuscular mycorrhizal (AM) fungi, which evolved over 400 million years ago [1]. Even today, AM endosymbiosis still occurs in ~72% of all higher plants [2]. Besides AM symbiosis, several plant lineages evolved additional or even alternative mutualistic endosymbiotic interactions, like orchid mycorrhiza, ericoid mycorrhiza, and diazotrophic rhizobia or *Frankia* bacteria hosted in root nodules. Interestingly, the evolution of these mutualistic endosymbiotic partnerships co-opted a signaling pathway critical for AM symbiosis. This pathway, known as the common symbiosis signaling pathway, is highly conserved and can be found in angiosperms, gymnosperms, monilophytes, and bryophytes [3].

The common symbiosis signaling pathway was first discovered in pea (*Pisum sativum*), showing to be critical for AM symbiosis and rhizobium-induced nodulation [4]. The subsequent molecular genetic characterization in the

legume models *Lotus japonicus* and *Medicago truncatula* revealed the pathway consists of four conserved components stretching from an LRR-type transmembrane receptor kinase down to the transcription factor *LjCYCLOPS/MtIPD3* [5, 6]. The LRR-type receptor kinase is generally named *SYMRK* (SYMBIOSIS SIGNALLING RECEPTOR KINASE), except for pea, *M. truncatula*, and *Medicago sativa,* where it is named PsSYM19, MtDMI2, and MsNORK, respectively [7]. The *SYMRK* extracellular structure varies between species, but in case of eudicots possesses a malectin domain, a conserved GDPC motif, and 2-3 LRR domains linked to a canonical intracellular serine-threonine kinase domain [7–10]. The malectin domain is cleaved in the absence of symbiotic signaling [11, 12]. Studies in *L. japonicus* showed that the remaining part of the SYMRK protein interacts with the LysM-type transmembrane receptor LjNFR5 [11, 12]. LjNFR5 is part of the receptor complex essential for recognizing rhizobium-secreted lipo-chitooligosaccharide (LCO) signal molecules [13, 14]. Legume *symrk* knockout mutants are blocked in rhizobium LCO-induced signaling through the common symbiosis signaling pathway. Subsequently, nodule formation is not initiated, nor is *Rhizobium* infection initiated in *symrk* mutants [7, 8, 15, 16]. *Lj*SYMRK also interacts with the innate immune receptor LjBAK1 (BRASSINOSTEROID INSENSITIVE 1-ASSOCIATED RECEPTOR KINASE 1), which may allow repression of immune responses upon symbiotic infection [17]. Such a role is supported by *symrk* mutant analysis, revealing fortification of the plant cell wall upon infection with *Glomus mosseae* AM fungus or rhizobium in mutant or RNA interference (RNAi) lines [18, 19].

Studies on *SYMRK* in non-legumes are limited. RNAi Knockdown studies in the actinorhizal plants *Datisca glomerata* and *Casuarina glauca* showed that, like in legumes, *SYMRK* is essential for nodulation [9, 20]. These findings demonstrate that the common symbiosis signaling pathway defines a conserved genetic basis for nodulation with rhizobia or *Frankia*. More recent phylogenomic studies support the hypothesis that the nodulation trait has a single evolutionary origin in the last common ancestor of the orders Fabales, Fagales, Cucurbitales and Rosales, representing all ten

nodulating plant lineages [20–22]. The occurrence of non-nodulating lineages in these four taxonomic orders allowed the identification of nodulation-specific genes, as such genes are prone to pseudogenization from the moment a plant lineage loses the nodulation trait. We identified seven of such nodulation-specific genes by comparing nodulating *Parasponia* species to their non-nodulating sister species of the genus *Trema* [23]. Among these is an *NFR5* orthologous LysM-type receptor named *NFP2,* essential for nodulation in *Parasponia* [24]. To our surprise, however, we also identified a seemingly critical mutation in *SYMRK* of *Trema orientalis* (accession RG33; *TorSYMRK*$^{RG33}$), originating from the Sabah Provence in Malaysian Borneo [25]. It suggests that the *TorSYMRK*$^{RG33}$ allele experiences pseudogenization, despite the fact *T. orientalis* accession RG33 can still establish an AM symbiosis [23].

*TorSYMRK*$^{RG33}$ has a conserved gene structure, though has a mutation in the conserved dinucleotide motif in the 5'-donor splice site of intron 12, converting this generally highly conserved dinucleotide motif into 'GA'. This led us to investigate the symbiotic functioning of *SYMRK* in the *Trema-Parasponia* lineage and investigate the impact of a seemingly critical SNP in an intron donor splice site in this gene.

## Results

### *Trema orientalis* and *Parasponia andersonii* differ in *Rhizophagus irregularis* colonization

Since *SYMRK* is known to be important for arbuscular mycorrhization in a range of species [7, 8, 20, 26], we first questioned whether *T. orientalis* accession RG33 can be effectively mycorrhized. To investigate this, we compared the mycorrhization dynamics of *T. orientalis* RG33 to *P. andersonii* (accession WU1). Both species are close relatives that diverged less than 20 million years ago [22], though have a somewhat different root architecture.

*T. orientalis* plantlets have a shorter main root, whereas its lateral roots are longer when compared to *P. andersonii* **(Figure S1)**.

To compare the mycorrhization efficiency, seedlings of both species were inoculated with 125 spores of *Rhizophagus irregularis* DOAM197198. Mycorrhization was quantified for 6 weeks, focusing on the frequency of mycorrhizal presence in the root system(F%), the intensity of mycorrhization in the root system (M%), the arbuscule abundance in the root system (A%), and the averaged arbuscule abundance in randomly selected infected root segments (a%) [27]. This revealed a clear difference in mycorrhization colonization dynamics between both species. The root system of *P. andersonii* is broadly colonized, showing an abundant presence of hyphae 4 weeks post-inoculation (F% >80%, M% >50%, **Figure 1A, C**). In contrast, *T. orientalis* RG33 showed a reduced mycorrhizal infection and a low abundance of mycorrhizal hyphae in the root (F% <20%, M% <10%, 4 weeks post-inoculation) **(Figure 1A, D)**. These reduced mycorrhizal infection rates of *T. orientalis* RG33 were also reflected in a reduced number of arbuscules found in the root system (A%). However, when evaluating the infected root segments, the arbuscule abundance (a%) was comparable to *P. andersonii* **(Figure 1B)**. This indicates that *T. orientalis* RG33 is infected less frequently by *R. irregularis* DOAM197198 when compared to *P. andersonii*. But once infected, the number of arbuscules formed in the infected root segment is similar between both species.



### *Parasponia andersonii SYMRK* is essential for arbuscular mycorrhization and nodulation

As *T. orientalis* RG33 can establish an arbuscular mycorrhizal symbiosis, we questioned whether *SYMRK* represents a single copy gene in the *Trema-Parasponi*a taxonomic lineage. We analyzed genome sequences of 20 species representing monocots and major clades of dicots, including Fabales, Fagales, Cucurbiales, and Rosales species. The closest SYMRK paralogs of *P. andersonii* and *T. orientalis* were included as an outgroup. This

revealed that *SYMRK* is a single-copy gene in the *Parasponia* - *Trema* lineage **(Figure S2)**.

Knock-down experiments in legumes and the actinorhizal species *C. glauca* and *D. glomerata* showed that *SYMRK* commits a dual role in establishing arbuscular mycorrhizal symbiosis and nodulation [7–9, 15, 16, 20, 28]. Furthermore, studies in *L. japonicus* revealed that ectopic expression of *LjSYMRK* results in the spontaneous onset of nodule organogenesis in absence of rhizobia [29]. To determine whether *SYMRK* in the *Parasponia-Trema* lineage fulfils a similar symbiotic role, two experiments were conducted. We generated CRISPR/Cas9 *symrk* knockout mutants in *P. andersonii* and conducted *PanSYMRK* ectopic expression studies in roots.

In total, three *Pansymrk* knockout mutant lines (homozygous line *Pansymrk-4* and the bi-allelic mutant lines *Pansymrk-5* and *Pansymrk-6*) were obtained by targeting the fourth and fifth coding exon using two single guide RNAs (sgRNAs) **(Figure S3A)**. All mutant alleles represent large deletions, only encoding a fragment of the extracellular domain **(Figure S3B)**. To determine whether *SYMRK* commits a key symbiotic function in *P. andersonii,* we first studied the nodulation phenotype of the *Pansymrk* mutants. *Pansymrk-4*, Pan*symrk-5* and Pan*symrk-6* plantlets were inoculated with *Mesohizobium plurifarium* BOR2, and the nodulation phenotypes were examined six weeks post-inoculation. The transgenic empty vector control plants (EV) were effectively nodulated, having nodule numbers ranging from 25 to 61 per plant. In contrast, the three *Pansymrk* mutant lines were unable to nodulate **(Figure 2A)**.

Next, we investigated the role of *PanSYMRK* in arbuscular mycorrhizal symbiosis. *Pansymrk-4, Pansymrk-5, Pansymrk-6*, and EV control plantlets were inoculated with an *R. irregularis* DAOM197198 spore suspension. Mycorrhization phenotypes were examined six weeks post-inoculation by quantifying four parameters; F%, M%, a%, and A%, as described above. The EV control plants interacted normally with the applied symbiont, with F%,

M%, a%, and A% of 65,4%, 36,8%, 77,1%, and 26,1%, respectively **(Figure 2B, C)**. Although some intraradical hyphae were observed in a minority of the

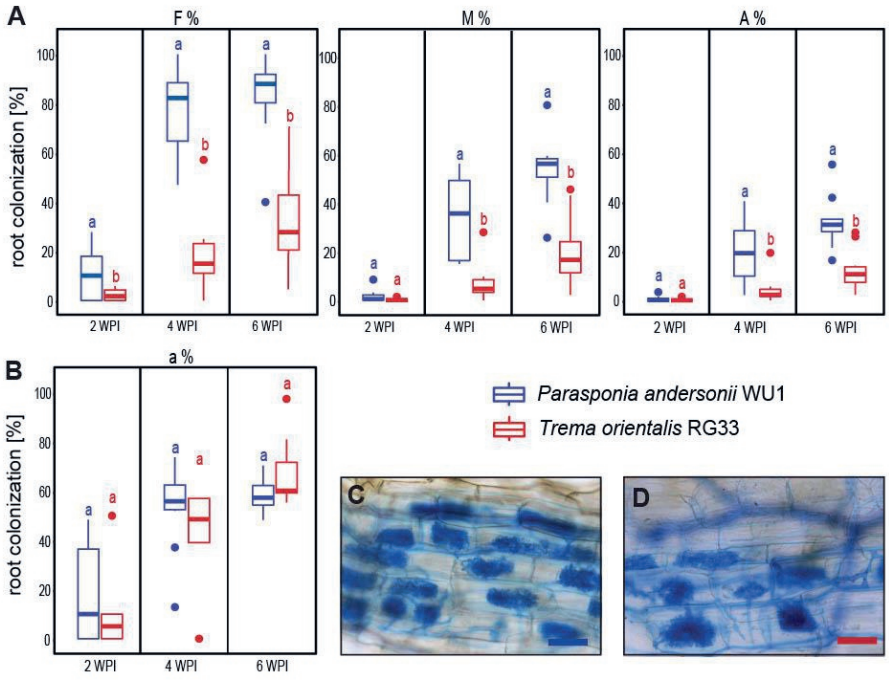



**Figure 1. *Trema orientalis* accession RG33 and *Parasponia andersonii* accession WU1 differ in mycorrhizal colonization.** (**A**) Comparison of mycorrhization efficiency in the root system of *P. andersonii* WU1 (blue) and *T. orientalis* RG33 (red) at 2, 4 and 6 weeks post-inoculation with *Rhizophagus irregularis* DOAM197198. F%: The frequency of mycorrhiza in the root system. M%: the intensity of mycorrhizal colonization in the root system. A%: Arbuscule abundance in the root system. (**B**) a%: Averaged arbuscule abundance detected in 50 randomly selected 1 cm infected segments of a root system. Error bars represent the SE of 10 biological replicates for each 50x 1cm root segment that has been analyzed. Analysis was done according to Trouvelot *et al.* (1986) [26]. (**C**) Toluidine blue-stained *P. andersonii* and (**D**) *T. orientalis* root segment visualizing *R. irregulates* arbuscules 6 weeks post-inoculation. Size bar = 10 µm. Difference letters above the box indicate statistical difference ($p<0.01$) as determined by Kruskal–Wallis test in combination with Fisher's post-hoc test.

*Pansymrk* root segments (7 out of 417, 6 out of 760, and 9 out of 1085 segments) **(Figure 2B, D-F)**, generally, no arbuscules were observed in any

of the tested *Pansymrk* mutant plantlets. This demonstrates that *SYMRK* is essential for nodulation and arbuscular mycorrhization of *P. andersonii* roots.

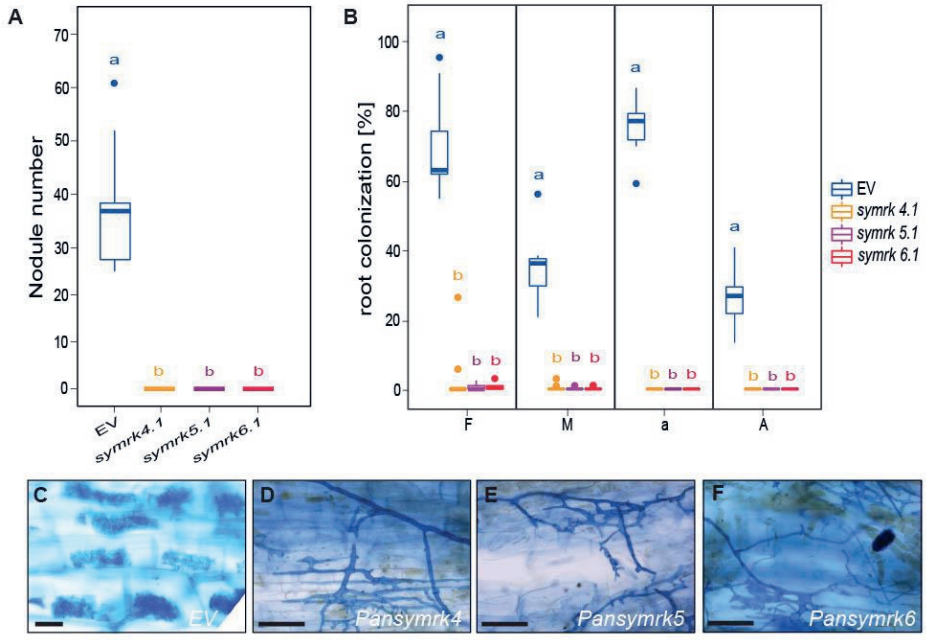



**Figure 2: *Parasponia andersonii SYMRK* is essential for mycorrhization and nodulation.** (**A**) Nodule numbers formed in *P. andersonii* empty vector control line (EV) and three *Pansymrk* mutant lines, 6 weeks post-inoculation with *Mesorhizobium plurifarium* BOR2. (**B**) mycorrhization efficiency in the root system of *P. andersonii* EV-control and three independent *Pansymrk* mutant lines 6 weeks post-inoculation with *Rhizophagus irregularis* DOAM197198. F%: The frequency of mycorrhiza in the infected root system. M%: the intensity of mycorrhizal colonization in the infected root system. A%: Arbuscule abundance in the infected root system. a%: Averaged arbuscule abundance detected in 50 randomly selected 1 cm segments of a root system. Error bars represent the SE of 10 biological replicates, for each 50x 1cm root segment that has been analyzed. Analysis was done according to Trouvelot *et al.* (1986) [26] (**C-F**): Toluidine blue-stained *P. andersonii* EV-control (**C**), *Pansymrk-4* (**D**), *Pansymrk-5* (**E**), and *Pansymrk-6* (**F**) root segment visualizing *R. irregulates* infections 6 weeks post-inoculation. Size bar = 10 µm. Difference letters above the box indicate statistical difference ($p<0.01$) as determined by Kruskal–Wallis test in combination with Fisher's post-hoc test.

Next, we questioned whether the ectopic expression of *PanSYMRK* is sufficient for spontaneous formation of nodule-like structures. We employed *Agrobacterium rhizogenes*-mediated (*A. rhizogenes*) root transformation to introduce *PanSYMRK* driven by the *L. japonicus UBIQUITIN 1* (*LjUBI1*) promoter. This revealed spontaneous formation of nodule-like structures on roots ectopically expressing *PanSYMRK* (n= 5/25) **(Figure 3A-C)**. Longitudinal sections revealed that these nodule-like structures originate from dividing cortical and pericycle cells, similar to genuine *Parasponia* nodules **(Figure 3D)**. This led us to conclude that *SYMRK* is an essential key regulatory LRR-type receptor kinase for the onset of the nodule developmental program in the non-legume *P. andersonii.*



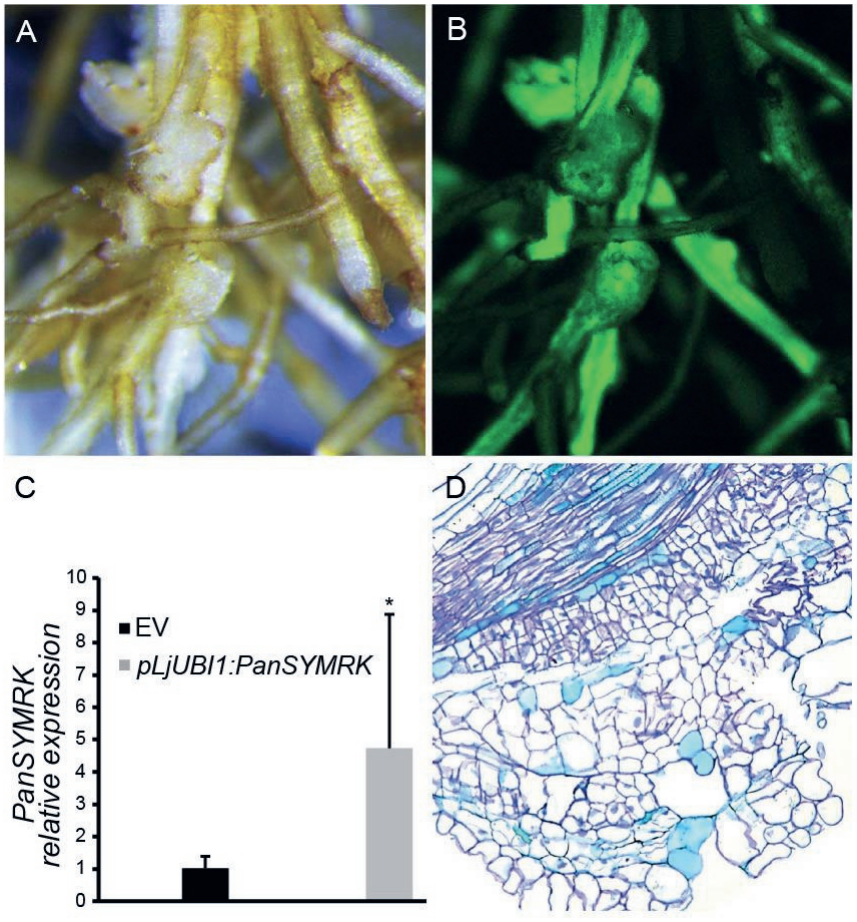


**Figure 3: *PanSYMRK* ectopic expression induces spontaneous nodulation in *Parasponia andersonii*.** (**A, B**) Bright-field (A) and green fluorescent image (**B**) of *P. andersonii A. rhizogenes*-transformed roots expressing *GFP* and *PanSYMRK* under control of the *pLjUBI1* promoter showing spontaneously formed nodule-like structures (6 weeks post planting). (**C**) Relative gene expression of *PanSYMRK* in *P. andersonii A. rhizogenes*-transformed roots containing an empty vector (EV) or *pLjUBI1:PanSYMRK* (n=3). (**D**) Longitudinal section of a spontaneously formed nodule-like structure visualizing cortical and pericycle cell divisions. Scale bar in (**A** and **B**) is 0.02mm and in (**D**) 50µM. Significant difference ($p<0.05$) was determined by student t-test.

## The GA mutation of the 5'-donor splice site of intron 12 doesn't affect *SYMRK* functionally

As *T. orientalis* RG33 -possessing a single *SYMRK* gene copy- can be mycorrhized effectively, it suggests that the *TorSYMRK*$^{RG33}$ allele encodes a functional protein to support this plant-fungus symbiosis. Earlier studies in *M. truncatula* revealed that the *SYMRK* requirements differ between mycorrhizal colonization and rhizobium nodulation [7]. The *M. truncatula* R38 *dmi2* mutant possesses a missense mutation converting a glycine to glutamic acid mutation at position 794 of the protein [7]. This mutation affects the kinase phosphorylation activity and the capacity of the protein to interact with it's downstream target 3-HYDROXY-3-METHYLGLUTARYL COENZYME A REDUCTASE1 (MtHMGR1) [30, 31]. *M. truncatula* R38 *dmi2* is affected in nodulation but not in mycorrhization, suggesting a SYMRK functional kinase domain is less critical for the latter interaction [7]. As the *T. orientalis SYMRK*$^{RG33}$ may encode -at least in part- a truncated SYMRK protein lacking essential domains of the kinase motif **(Figure 4A)**, we question to what extent this allele could function in nodulation.

To investigate this, first the native promoter region of *P. andersonii SYMRK* was identified. We used *A. rhizogenes* root transformation to show that a ~3 kb upstream region including the 5'-UTR driving the *PanSYMRK* gene functionally complemented the *Pansymrk*-5 mutant (4.9 nodules/plant at 8 wpi) **(Figure 4B; Figure S4B)**. Next, we used this promoter to drive a *PanSYMRK* gene mutant harbouring a GA at the donor site of intron 12,

mimicking the *TorSYMRK*$^{RG33}$ allele to determine its functionality in the *P. andersonii Pansymrk*-5 mutant background. Using *A. rhizogenes* root transformation, we found full complementation of the *Pansymrk* mutant phenotype **(Figure 4C-E; Figure S4C)**. On average, 13 nodules per plant were formed at 8 wpi. Sections of these nodules revealed a wild type cytoarchitecture, including a large zone of cells possessing fixation threads. This shows that the GA point mutation at the donor site of intron 12 is not affecting *SYMRK* gene functionality.

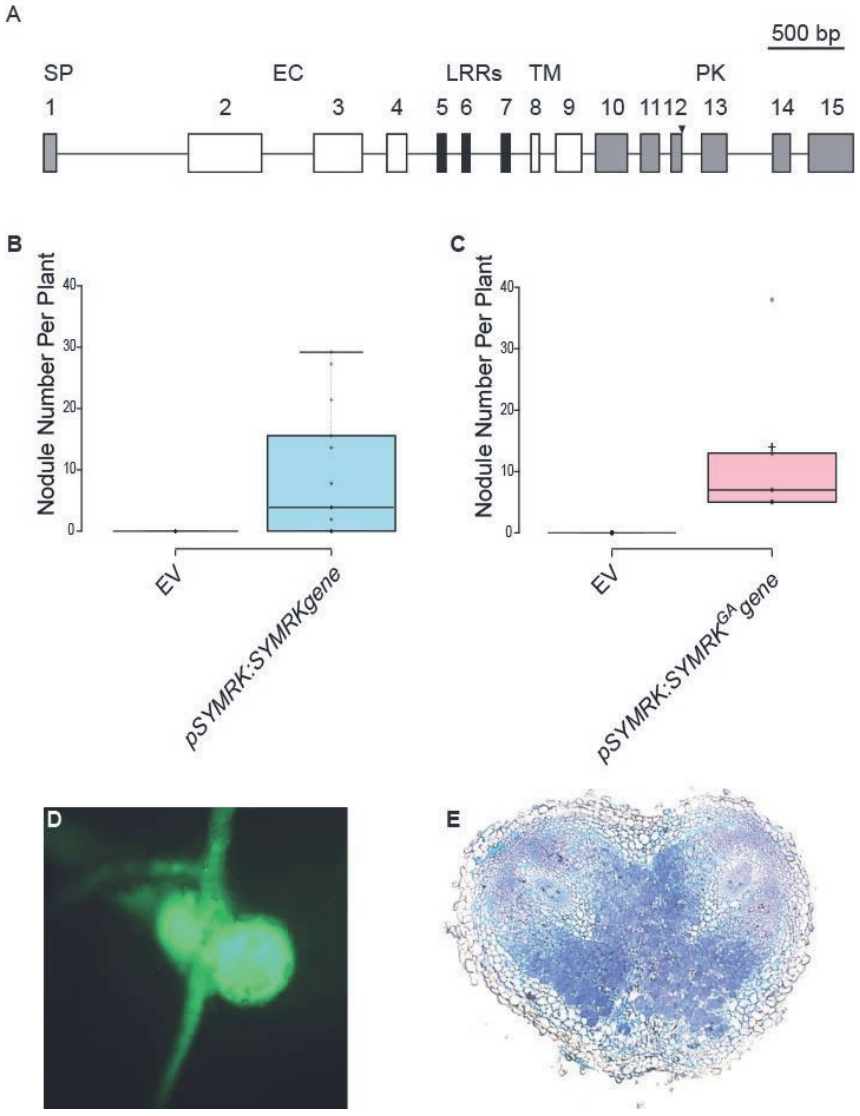


**Figure 4: *Parasponia symrk-5* mutant trans-complementation of root nodule symbiosis.** (**A**) Schematic representation of *P. andersonii SYMRK* gene structure. Arrowhead points to the location of the introduced GA mutation in *PanSYMRK* at the 5'-donor splice site of intron 12. (**B**) Nodule number per plant formed on *Pansymrk-5 A. rhizogenes* transformed root with *pPanSYMRK:PanSYMRK* gene (n=13). (**C-E**) Nodule number per plant (n=5) (**C**), representative image of green fluorescent protein (GFP) nodule (**D**), and a section through a mature nodule (**E**) of *Pansymrk-5 A. rhizogenes* transformed root with *pPanSYMRK:PanSYMRK*$^{GA}$ carrying a GA mutation at the 5'-donor splice site of intron 12. Nodules were harvested and analyzed at 8 weeks post inoculation with *Mesorhizobium plurifarium* BOR2 (OD600 = 0.025). Scale bar in (**D**) is 0.2mm and in (**E**) is 100µm.

## A GA 5'-donor splice site is very rare, though effectively spliced in *TorSYMRK*$^{RG33}$

We question how effective an intron that possesses a GA as the first two nucleotides of a donor splice site is spliced. To determine this, we aimed to compare the coverage of RNAseq reads of the 15 exons and 14 introns of the *SYMRK* gene of *T. orientalis* and *P. andersonii*. *SYMRK* is highly similar in both species, though introns show some variation in length **(Table 1)**. *SYMRK* is known to be expressed in the root [8]. We grew *T. orientalis* and *P. andersonii* seedlings *in vitro* on a low nitrate medium and subsequently isolated 1 cm regions of roots just above the root meristemic zone. RNA extracted from these samples was sequenced (in triplicates), mapped, and analyzed **(Figure S5A)**. When focusing on intron 12, we found a per base mean coverage of 4.6±1.0 for *TorSYMRK*$^{RG33}$, whereas in *P. andersonii,* the coverage of this intron is only 0.2±0.3 mean per base coverage **(Table 1).** Comparing the splice site efficiency of intron 12, we observe that GA splice site in *T. orientalis* splices efficiently at approximately 95%, while the GC splice site in *P. andersonnii* shows an efficiency of 99.9%. This difference in intron retention between *PanSYMRK* and *TorSYMRK*$^{RG33}$ was also observed by qRT-PCR on root mRNA **(Figure S5B)**. These data suggest that *SYMRK* intron 12 is spliced less efficiently in *T. orientalis* when compared to *P. andersonii.* However, a similar variance is observed for other introns, which possess canonical donor and acceptor splice sites; e.g. *PanSYMRK* intron 11 **(Table 1)**, suggesting some intron retention is not hampering gene function.

Therefore, we conclude that SYMRK$^{RG33}$ is fully functional, despite a non-canonical GA dinucleotide motif in the donor splice site.

**Table 1: Splicing efficiency of *Trema orientalis* and *Parasponia andersonii SYMRK* in the root susceptible zone for symbiotic engagement.**

| *SYMRK* | *T. orientalis* | root susceptible zone | | *P. andersonii* | root susceptible zone | |
|---|---|---|---|---|---|---|
| | length (bp) | coverage | s.d. | length (bp) | coverage | s.d. |
| exon1 | 99 | 28.1 | 8.9 | 99 | 94.6 | 39.6 |
| intron1 | 765 | 0.0 | 0.1 | 912 | 0.1 | 0.1 |
| exon2 | 525 | 35.2 | 13.2 | 525 | 96.5 | 17.6 |
| intron2 | 339 | 0.0 | 0.0 | 341 | 0.1 | 0.2 |
| exon3 | 475 | 51.2 | 9.1 | 475 | 114.6 | 17.1 |
| intron3 | 351 | 0.0 | 0.0 | 364 | 0.0 | 0.0 |
| exon4 | 150 | 59.3 | 23.4 | 150 | 127.0 | 24.9 |
| intron4 | 188 | 0.1 | 0.1 | 175 | 0.1 | 0.2 |
| exon5 | 71 | 89.2 | 27.6 | 71 | 166.1 | 29.4 |
| intron5 | 103 | 0.3 | 0.2 | 102 | 1.0 | 0.7 |
| exon6 | 68 | 84.4 | 20.7 | 68 | 133.1 | 22.2 |
| intron6 | 202 | 2.7 | 0.5 | 319 | 0.2 | 0.2 |
| exon7 | 71 | 70.7 | 18.3 | 71 | 125.0 | 13.2 |
| intron7 | 133 | 0.0 | 0.0 | 134 | 0.3 | 0.4 |
| exon8 | 71 | 62.2 | 16.9 | 71 | 138.8 | 15.5 |
| intron8 | 97 | 0.9 | 1.1 | 86 | 0.1 | 0.1 |
| exon9 | 193 | 76.1 | 16.1 | 193 | 142.9 | 22.8 |
| intron9 | 98 | 1.9 | 0.6 | 98 | 0.9 | 0.9 |
| exon10 | 235 | 79.6 | 13.5 | 235 | 155.3 | 18.6 |
| intron10 | 83 | 0.1 | 0.1 | 90 | 0.3 | 0.6 |
| exon11 | 126 | 100.8 | 13.3 | 126 | 158.0 | 35.0 |
| intron11 | 65 | 1.3 | 1.9 | 83 | 10.5 | 4.3 |
| exon12 | 86 | 91.8 | 19.3 | 86 | 158.9 | 34.5 |
| **intron12** | **118** | **4.6** | **1.0** | **118** | **0.2** | **0.3** |
| exon13 | 189 | 109.4 | 13.8 | 189 | 159.3 | 16.9 |
| intron13 | 276 | 1.6 | 1.0 | 273 | 0.7 | 0.6 |
| exon14 | 132 | 155.1 | 7.4 | 132 | 255.8 | 22.9 |
| intron14 | 112 | 4.2 | 4.3 | 112 | 0.4 | 0.7 |
| exon15 | 326 | 135.4 | 9.4 | 326 | 207.6 | 13.0 |

RNA-seq quantification for each intron and exon in the *SYMRK* gene of *Trema orientalis* and *Parasponia andersonii* is determined by the mean per base coverage of three biological replicates.

Next, we questioned how unique a GA donor splice site is in plants. For this, we analyzed all annotated introns in *T. orientalis*, *P. andersonii,* and the model plant species *L. japonicus, M. truncatula*, and *Arabidopsis thaliana* [23, 32–34]. This showed that a GA donor splice site is extremely rare, varying from none in the annotated gene models of *M. truncatula* to 14 in *A. thaliana* **(Table 2, Table S1)**.

**Table 2: Frequency of predicted canonical and non-canonical donor splice sites.** Splice site occurrences are based on existing gene models predictions for *Trema orientalis*, *Parasponia andersonii*, *Lotus japonicus*, *Medicago truncatula* and *Arabidopsis thaliana*.

| | *T. orientalis* | | *P. andersonii* | | *M.truncatula* | | *L. japonicus* | | *A. thaliana* | |
|---|---|---|---|---|---|---|---|---|---|---|
| **Splice motif** | **Total** | **%** | **Total** | **%** | **Total** | **%** | **Total** | **%** | **Total** | **%** |
| GT-AG | 102107 | 98.09% | 100852 | 98.36% | 213612 | 95.94% | 148570 | 97.99% | 159839 | 98.80% |
| GC-AG | 537 | 0.52% | 567 | 0.55% | 8962 | 4.03% | 2193 | 1.45% | 1742 | 1.08% |
| AT-AC | 13 | 0.01% | 14 | 0.01% | 79 | 0.04% | 656 | 0.43% | 90 | 0.06% |
| GA-AG | 11 | 0.01% | 8 | 0.01% | 0 | 0.00% | 6 | 0.00% | 14 | 0.01% |
| Others | 1426 | 1.37% | 1096 | 1.07% | 0 | 0.00% | 190 | 0.13% | 89 | 0.06% |
| Canonical | 102107 | 98.09% | 100852 | 98.36% | 213612 | 95.94% | 148570 | 97.99% | 159839 | 98.80% |
| Non-canonical | 1987 | 1.91% | 1685 | 1.64% | 9041 | 4.06% | 3045 | 2.01% | 1935 | 1.20% |
| Grand Total | 104094 | 100% | 102537 | 100 % | 222653 | 100 % | 151615 | 100 % | 161774 | 100 % |

## *Trema orientalis SYMRK$^{RG33}$* GA donor splice site is geographically limited

As *T. orientalis* RG33 possesses an extremely rare GA motif at the donor splice site of intron 12, we question to what extent such polymorphism is unique in *SYMRK*. First, we analyzed SYMRK orthologs in a broad phylogenetic context. This showed that a non-canonical GC donor splice site is common in *SYMRK* intron 12 of dicotyledon species **(Figure 5)**.

However, none of the analyzed *SYMRK* genes possesses a GA motif at this position. Subsequently, we analyzed *SYMRK* of the *Parasponia-Trema* species complex. Among others, *T. orientalis* accession RG33 was collected during an expedition in Sabah Provence, Malaysian Borneo, in 2012 [23, 25]. We analyzed 27 additional *T. orientalis* individuals collected from five distinct locations in Malaysian Borneo **(Figure 6A)**. All possess the rare GA intron 12 donor splice, whereas this mutation is absent in *Trema* and

*Parasponia* accessions sampled outside Borneo **(Figure 6; Table S2)**. This demonstrates that the *SYMRK*$^{RG3}$ allele is not unique, though it associates with the Borneo *T. orientalis* population.

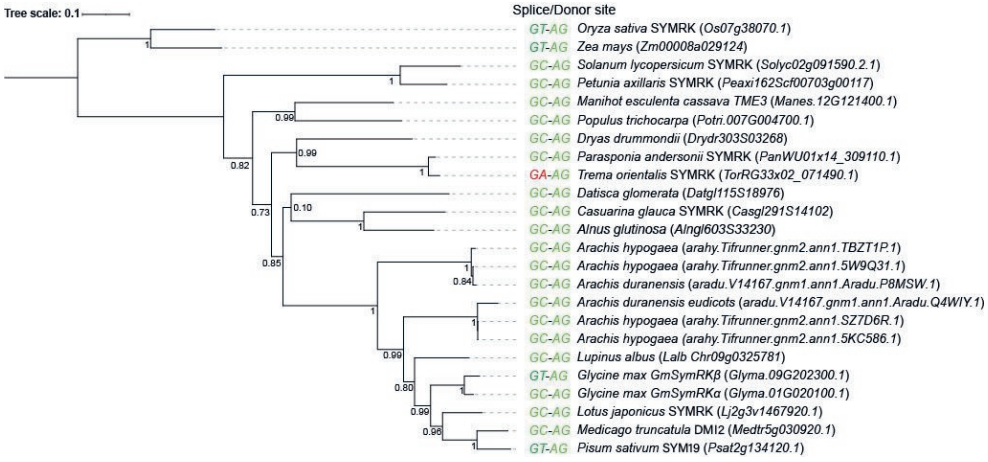



**Figure 5: Phylogeny of SYMRK including the splice site dinucleotide motifs for intron 12.** Phylogeny was reconstructed based on an alignment of SYMRK orthologous proteins from 19 species. Leaves are labelled by their respective species, gene name if available) and gene identifier. The non-canonical GC donor splice site is common in *SYMRK* intron 12 of dicotyledon species, except in *Glycine max SYMRKβ* and *Pisum sativum SYM19*, where GC is substituted by GT. In contrast, only *Trema orientalis* RG33 possesses a GA motif in this position (highlighted in red).

## Discussion

The LRR-type receptor kinase *SYMRK* is a critical component in the common symbiosis signaling pathway controlling endosymbioses. In legumes, *SYMRK* is essential for rhizobium LCO-induced signaling. We identified a seemingly critical mutation of the conserved dinucleotide motif in the 5'-donor splice site in *T. orientalis SYMRK* accession RG33*. T. orientalis* is a non-nodulating relative of nitrogen-fixing *Parasponia* species and has experienced pseudogenization of several key nodulation genes [23]. Here we show that despite a mutation in a splice site motif, *TorSYMRK*$^{RG33}$ remains a functional

allele that can be effectively spliced. The dominant occurrence of the *TorSYMRK*$^{RG33}$ allele in the Malaysian Borneo *T. orientalis* population underlines the splice site mutation is not affecting the fitness of the tree species.

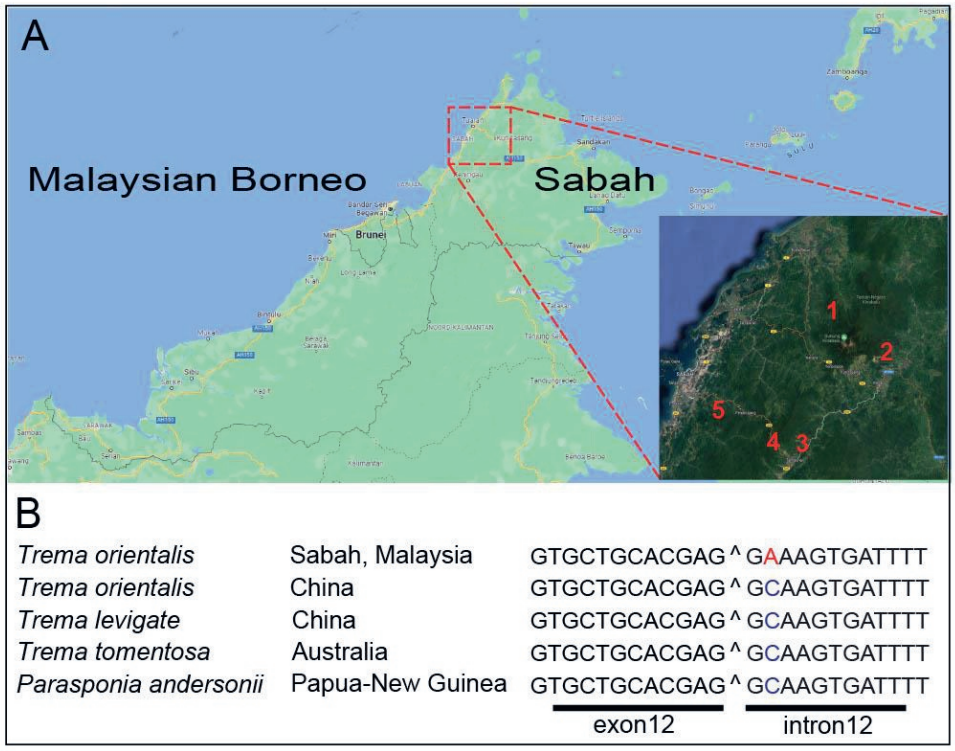



**Figure 6: *SYMRK* intron 12 unique non-canonical donor splice site occurs in a *Trema orientalis* population endogenous to Sabah, Malaysia. (A)** Locations of 28 *Trema orientalis* specimens collected in Malaysian Borneo, province of Sabah. 1: Sayap, 2: Poring, 3: Mahua, 4: Gunung Alab, and 5: Inobong. Plants were collected in 2012 as described in Merckx *et al*. (2015) [31] (see also **Table S2**). Map data © 2023 Google. (**B**) The 'GA' donor splice site of intron 12 is unique to *Trema orientalis* of Malaysia, Sabah, whereas related accessions and species possess a non-canonical 'GC' at this position in *SYMRK*.

Splicing is a highly conserved process in eukaryotes, requiring a spliceosome complex consisting of five small nuclear RNAs and several proteins. The vast majority of introns are spliced by the so-called U2-type spliceosome, recognizing two highly conserved di-nucleotide motifs at the start and end of the intron sequence, namely GT-AG. Bioinformatic studies in plant,

animal, and fungal species indicate that alternative dinucleotide motifs are used in less than 2% of cases, among which GC-AG is the most abundant non-canonical splice motif representing 1.5% of all introns annotated in plant gene models [35, 36]. The GA-AG splicing motif, as found in *TorSYMRK*$^{RG33}$ intron 12, is reported to occur in >0.03% of the cases [36].

The mechanism driving the evolution of rare non-canonical splice sites remains elusive. The GA-AG dinucleotide splicing motif was found in higher frequency in two non-related animal species; the copepod *Eurytemora affinis* and the tunicate *Oikopleura dioica* [36–38]. However, it remains unknown whether both species have gained these by convergent evolution or, alternatively, it is an ancestral trait preserved in only a few species [36]. In the case of *SYMRK*, we noted that in related species, *SYMRK* intron 12 possesses the more common non-canonical GC-AG dinucleotide splice motif. This may lead to the hypothesis that such a GC-AG motif is the ancestral state allowing the evolution of the even more rare GA-AG motif. We inserted the GC to GA mutation in the *P. andersonii SYMRK* gene and showed that this variant is fully functional when expressed under its native promoter. This suggests that a simple single nucleotide polymorphism is sufficient to allow the evolution of the GA-AG dinucleotide splicing motif in *TorSYMRK*$^{RG33}$. We analyzed genomes of five plant species for gene models possessing a GA dinucleotide motif in the donor splice site. We found that the GA motif is indeed present in the annotated gene models, albeit at very low frequency in the analyzed species.

Using CRISPR-Cas9 technology in *P. andersonii*, we demonstrated for the first time by mutant analysis that *SYMRK* commits a dual symbiotic role in essential nodulation and AM symbiosis in a non-legume. Earlier studies using RNAi in *C. glauca* and *D. glomerata* provided evidence that *SYMRK* is required for *Frankia*-induced nodulation and mycorrhization [9, 20]. *Parasponia*, *Casuarina*, and *Datisca,* together with legumes, represent all four taxonomic orders that contain nodulating species and for which *SYMRK* is an essential symbiotic gene. It supports the hypothesis that *SYMRK* -and

other components of the common symbiosis signaling pathway- have been recruited to function in nodulation in a common ancestor that lived before the divergence of the Fabales, Fagales, Cucurbitales, and Rosales orders.

## Conclusions

This study of the LRR-type receptor kinase SYMRK in the non-nodulating relative of nitrogen-fixing *Parasponia* species, *T. orientalis*, led to the identification of a functional splice site mutation in the gene. The discovery of this rare non-canonical GA-AG splice site motif in *SYMRK* raises questions about the evolution of such motifs and the mechanisms driving their occurrence. Furthermore, this study demonstrates the conservation of SYMRK functioning in nodulation and AM symbiosis in both legumes and non-legumes. The *Parasponia*-*Trema* comparative system is established to obtain insight into the evolutionary trajectory of the nodulation trait. It uncovered several genes critical for rhizobium-induced nodulation in a non-legume [23, 24, 39]. Eventually, *Trema* species can serve as an experimental test system to uncover essential genes to rebuild the nodulation trait. Additionally, we demonstrated that the *Parasponia*-*Trema* comparative system is equally valuable to uncovering the functionality of rare non-canonical splicing motifs. Overall, this study contributes to our understanding of both the common symbiosis signaling pathway and the mechanisms of gene splicing in plants.

## Material and Methods

### Plant materials and growth conditions

*Trema orientalis* plants used in this study were collected between September 10th and 25th, 2012, during the Crocker Range/Kinabalu Scientific Expedition. This expedition was conceived, organized, funded, and conducted jointly by Sabah Parks (Malaysia) and the Naturalis Biodiversity Center (The Netherlands). Detailed information about the expedition is available in Merckx *et al.,* 2015 [25]. Taxonomic analysis of the *T. orientalis*

samples has been previously published in van Velzen *et al.,* 2018. [23]. *P. andersonii* WU1 and *T. orientalis* RG33 were grown and maintained as described previously [40, 41]. Plantlets for nodulation and mycorrhization assay were vegetatively propagated *in vitro* and rooted [40, 41].

**Mycorrhization assays and trypan blue staining**

Mycorrhization assays were performed using a commercial spore of *Rhizopagus irregularis* (Agronutrion-DAOM197198, Carbonne, France). Spores inoculum, inoculation, and trypan blue staining were prepared and performed as described previously [41].



To quantify mycorrhization, a minimum of ~50 cm roots for each sample were cut into 1 cm fragments. 25-30 root fragments were placed on a single microscope slide, and 30% glycerol was added. Roots were covered with a cover glass and pressed until root fragments became flat. The frequency of mycorrhiza (%F), the intensity of mycorrhizal colonization (%M), and arbuscules abundance (%A) in the root system was scored and calculated according to Trouvelot et al. [27].

**Nodulation assay**

*P. andersonii* plantlets for nodulation were inoculated with *Mesorhizobium plurifarium* BOR2 (OD600 = 0.05) [23, 40, 41]. Plants were removed from the pots six weeks post-inoculation, roots were washed with running water to remove perlite, and nodules were counted. In (*trans*) complementation studies, plant roots were examined under fluorescent stereo microscopy, and nodule number was quantified for each transgenic root (eight weeks post-inoculation with *Mesorhizobium plurifarium* BOR2 (OD600= 0.025).

**Root growth assay**

Five seedlings of *P. andersonii* and *T. orientalis* RG33 were grown on ½

strength modified Hoagland medium in 12 cm square plates. Plants were grown vertically at a 60-degree angle for 21 days at 28$^{o}$C, 16/8h day-night regime. The primary root was determined as the main root that emerged from cotyledon, whereas lateral roots were determined as roots that emerged from the primary root. Per plants, primary root length, the average number of lateral roots, and lateral root density (per cm main root) were determined 21 days post germination. Primary root growth was measured by following its development every day for 21 days post germination. The average lateral root length was determined by measuring its size in five selected lateral roots 21 days post germination.

**Vectors and Constructs**

Single-guide RNAs (sgRNAs) were designed using the 'Find CRISPR Targets' function implemented in Geneious software version 9.1.5 (Biomatters, New Zealand) and subsequently checked against the *P. andersonii* genome for high identity off-targets. For CRISPR/Cas9-mediated mutagenesis and complementation studies, binary transformation constructs were created using Golden Gate assembly as described previously [40, 41], and a list of constructs generated from both studies is listed in **(Table S3)**. For CRISPR/Cas9-mediated mutagenesis, two sgRNAs were used to target the fourth and the fifth coding exons of *PanSYMRK* **(Figure S3)**. Selected sgRNAs were amplified using sequence-specific forward primers and a universal reverse primer **(Table S4)**, using Addgene plasmid no. 46966 as template [42]. To allow for Golden Gate cloning, BpiI and BsaI restriction sites in the putative promoter sequence of *PanSYMRK* were mutated by introducing single nucleotide substitution [43]. For the complementation study, the sequence of *P. andersonii SYMRK* promoter, 5' untranslated region (5' UTR), genomic DNA, 3' untranslated region (3' UTR), and terminator were synthesized. Also, a modified version of *P. andersonii SYMRK* genomic DNA was synthesized harboring a point mutation at the donor splice site of the 12th intron, mimicking *T. orientalis SYMRK*$^{RG33}$. (Invitrogen, Thermo Fisher Scientific, United States).

**Plant transformation**

*Agrobacterium tumefaciens*-mediated transformation and genotyping were done based on previously published protocols [40, 41]. Primers used for genotyping are listed in **(Table S4)**. Hairy root transformations were performed according to Cao *et al.* [44], where *A. rhizogenes* MSU440 or AR1193 harboring plasmid DNA of interest were used to infect micro-propagated plants wounded on their base. Infected plants were grown on agar plates of Schenk and Hildbrandt medium (SH medium) [45] and incubated at 21°C for one week on a 16/8h light/dark regime. Transformed plants were transferred to agar plates of SH medium supplemented with 10 g sucrose/L, cefotaxime 100 µg/mL, and kanamycin 50 µg/mL and subsequently incubated at 21°C for one week followed by 28°C for two weeks. Plants were checked for transgenic roots using a fluorescence stereo microscope.



**RNA Sequencing**

For RNA isolation, tissue was harvested from a ~1 cm region just above the meristematic zone of young growing roots and snap-frozen in liquid nitrogen. Material from ~5 plants was combined to form a single biological replicate. RNA was isolated in triplicate as previously described [23]. Library preparation and RNA sequencing was conducted by BGI (Schenzhen, China). Mapped RNA-sequencing reads covering the *SYMRK* gene in *P. andersonii* and *T. orientalis* were visualized using Integrative Genomics Viewer (IGV) [46]. Based on the different splice sites, two *SYMRK* splice variants were manually constructed. Functional protein domains for these variants were annotated using InterProScan 5 [47].

**Phylogenetic reconstruction**

Orthologs of SYMRK were identified among 49 publicly available proteomes by applying a Reciprocal Best Hits (RBH) approach, using *L. japonicus* SYMRK

(Lj2g3v1467920.1) as the query sequence. Identified orthologous proteins were aligned using Clustal Omega 1.2.3. [48]. A phylogenetic SYMRK tree was constructed using PhyML 3.0 [49] with LG substitution model 1,000 bootstrap replicates and rooted on the two Poales outgroup species. The tree was visualized using the Interactive Tree Of Life (iTOL) tree viewer [50]. A sub-selection of 20 species was extracted from the SYMRK orthogroup, and a tree was constructed using the same methods described above. Based on the SYMRK gene models for these 20 species, the splice site at intron 12 for each SYMRK ortholog was added.

**Statistical analysis**

Graphs and statistical analysis for mycorrhization quantification were performed using RStudio version 1.1.456. The Ramf R package was used to analyze and display quantitative AM fungal root colonization data [51]. Statistical tests on three classes of mycorrhization efficiency were done using Kruskal-Wallis test in combination with the post-hoc test using Fisher's least significant difference criterion. Statistical significance was defined as a $p<0.01$. A statistical test on root growth assays and for nodules number quantification on complementation study was done using a student t-test. Statistical significance for these parameters was defined as a $p<0.05$.

**Availability of data and materials**

The datasets analyzed during the current study are available in the NCBI SRA repository under BioProject numbers PRJNA272473 and PRJNA272482. Plant material and seeds used in this study can be obtained upon request from the corresponding author.

**Funding**

The project was funded by the Ministry of Research, Technology, and Higher Education of the Republic of Indonesia (grant no. 8245–ID to YPR), the

Ministry of Education, King Faisal University, Saudi Arabia (grant no. 10598 to SA) and the ENSA project funded by the Bill & Melinda Gates Foundation to the University of Cambridge (OPP1172165 to RG).

**Author information**

Sultan Alhusayni and Yuda Purwana Roswanjaya contributed equally to this work.

**Competing interests**

The authors declare no conflict of interest.



**Authors' Contributions**

Data presented in Fig. 1: YPR; Fig. 2: YPR; Fig. 3: SA, SB; Fig. 4: SA, SB, RH, LR; Fig. 5: JK, SA; Fig6: RG, YPR, WK; Fig S1: TS, WK; Fig. S2: JK; Fig. S3: YPR, LR, RH; Fig. S4: SA, Fig. S5: JK, SA; Table 1: JK; Table 2: JK; Table S1: JK, MS; Table S2: YPR, TS, WK, RG; Table S3: YPR, SA. Manuscript preparation: RG, JK, SA. All authors reviewed the manuscript. The authors read and approved the final manuscript.

**Corresponding authors**

Correspondence to Joël Klein or Rene Geurts.

**Ethics approval and consent to participate**

This study was conducted in accordance with local legislation and the Convention on the Trade in Endangered Species of Wild Fauna and Flora (CITES). No specific permits were required. The plant material explored here was not collected in this study, it was described in 2018 in van Velzen et al., 2018, Proc. Nat. Acad. Sci USA: doi.org/10.1073/pnas.1721395115, which is

referenced in the Materials and Methods section on line 424. In that publication, an in-depth phylogenetic analysis of the collected specimens is included (see **Figure S6** van Velzen *et al.,* 2018). The plant material was sampled between September 10th and 25th, 2012, by the participants of the Crocker Range/Kinabalu Scientific Expedition, which was conducted jointly by Sabah Parks (Malaysia) and Naturalis Biodiversity Center (The Netherlands) and described in detail Merckx *et al.* 2015 Nature (doi.org/10.1038/nature14949) (as indicated in the Materials and Methods section of the current manuscript: line 422-424) We have specified the permissions and licenses acquired for the plant material collection.

## Supplemental information

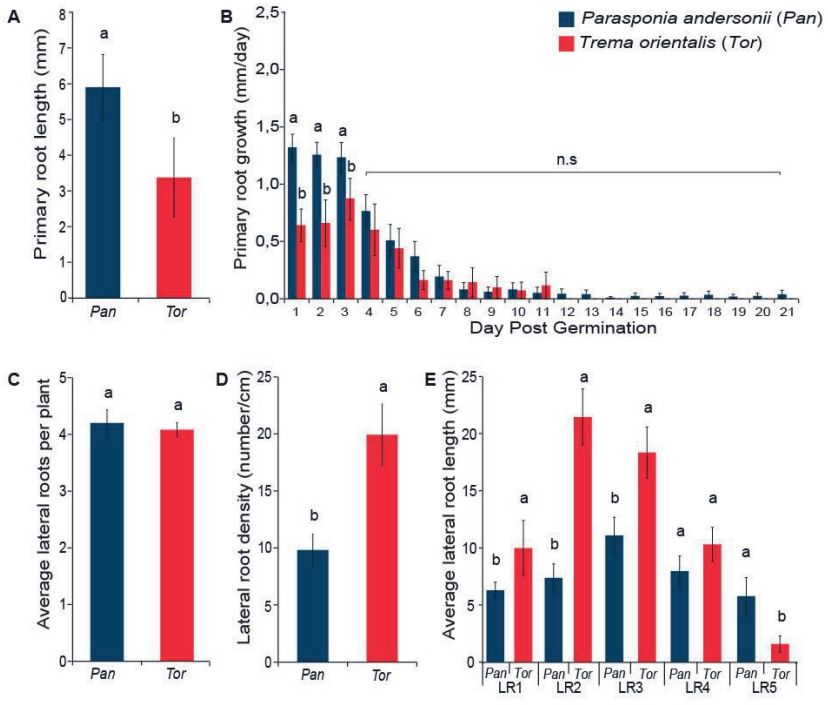



**Figure S1: Root architecture of *Trema orientalis* accession RG33 and *Parasponia andersonii* accession WU1 differs.** (**A**) Primary root length 21 days post-germination. (**B**) Growth of primary root 1-21 days post-germination (mm/day) (**C**) The average number of lateral roots per plant, and (**D**) lateral root density (cm-1 main root) 21 days post-germination (**E**) Average lateral root length in five selected root 21 days post-germination (mm). n=5 +/- s.e. Different letters above the bars indicate statistical significance ($p < 0.05$) as determined by student t-test. Tor, red: *T. orientalis* RG33, Pan, blue: *P. andersonii* WU1. Plants were grown in vitro on ½ strength modified Hoagland medium in 12 cm square plates.

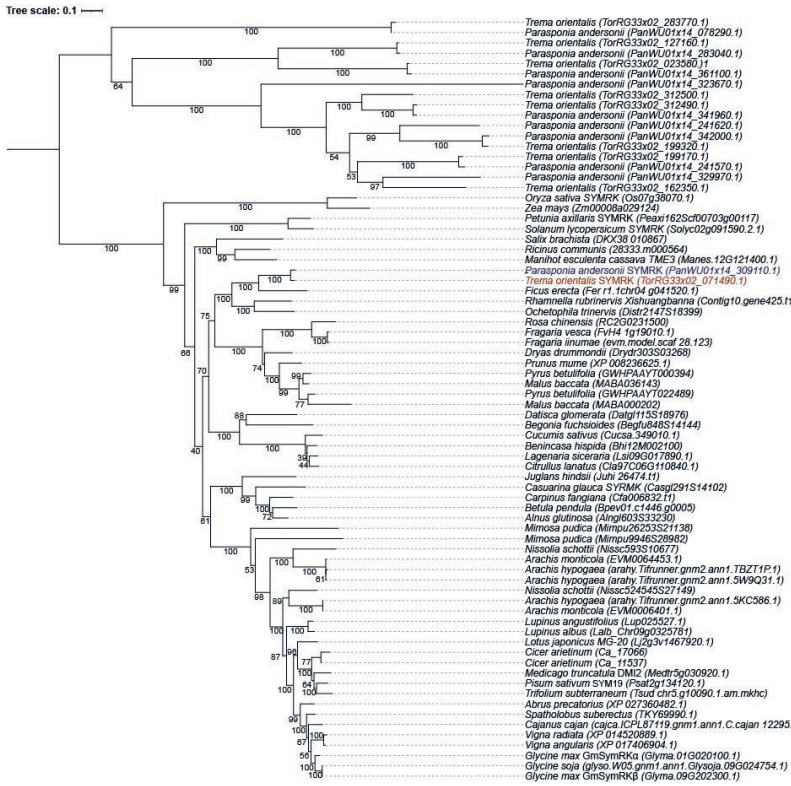



**Figure S2: Phylogenetic reconstruction of SYMRK orthologs.** Phylogeny was reconstructed on an alignment of SYMRK proteins from 51 plant species belonging to the Nitrogen fixation clade and two species belonging to the Solanales and two species of the Poales. In addition, *Trema* and *Parasponia* SYMRK homologous proteins were added to show that these groups are outside the SYMRK clade. Branch support is indicated by posterior probabilities. Lineages are labelled by species name and gene identifier.

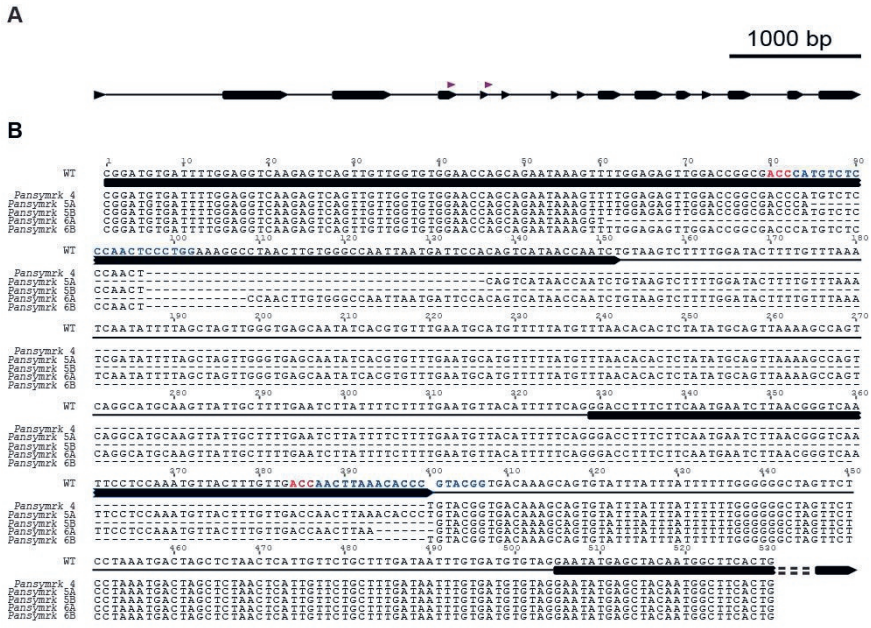



**Figure S3: *Parasponia andersonii symrk* CRISPR-Cas9 mutant alleles.** (**A**) Structure of *Pansymrk* gene spanning 7,280 bp and possessing 15 exons and 14 introns. Indicated are the positions of two sgRNAs (purple arrowheads) in exons 4 and 5. (**B**) Sequence alignment of the fourth and fifth exons of *PanSYMRK* in wild type (WT) and the three mutants *Pansymrk-4, Pansymrk-5,* and *Pansymrk-6*. Note: *Pansymrk-4* is a homozygote mutant possessing a 303 bp deletion whereas line 5 and 6 are bi-allelic. In the bi-allelic mutant lines, both alleles (**A** and **B**) are shown. Highlighted in blue and red are the sgRNA target sites and PAM sequences, respectively.

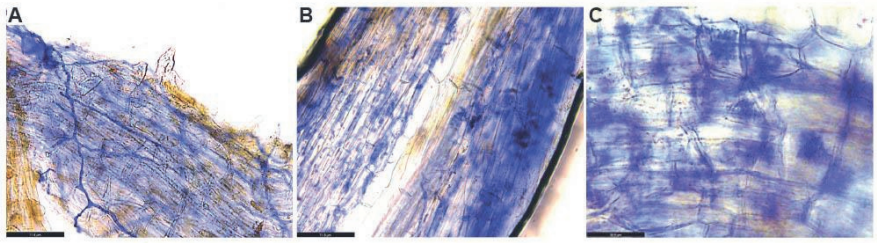



**Figure S4: *Parasponia symrk-5* mutant trans-complementation assay of mycorrhization.** (**A**) Representative image of *Pansymrk-5 A. rhizogenes* transformed root with empty vector (EV). (**B**) complementation with *pPanSYMRK:PanSYMRK* gene and (**C**) trans-complementation with *pPanSYMRK:PanSYMRK*$^{GA}$ gene. Visualization of *Rhizophagus irregularis* infection 6 weeks post-inoculation. Scale bar in (**A** abd **B**) is 72 µM, and in (**C**) is 36 µM.

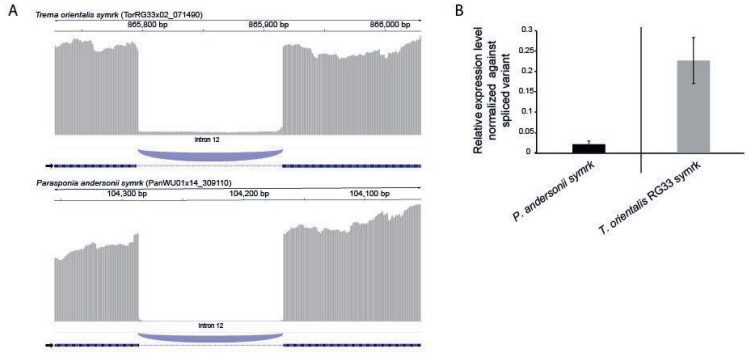



**Figure S5: Analysis of *SYMRK* Intron 12 splice variant expression in *T. orientalis* and *P. andersonii* roots.** (**A**) Mapping of root RNA sequence reads to *SYMRK* gene models of *T. orientalis* and *P. andersonii* showing a ~ 300 bp region around intron 12. (**B**) Difference in intron retention of *SYMRK* intron 12 and detected by qRT-PCR in *P. andersonii* and *T. orientalis* RG33. The bar plot represents the means of three biological replicates ± SD.

**Table S1: Frequency of GA-AG intron splice sites in four plant species.**

| Species + geneiD | # introns | GA-AG containing intron | Gene name | Description |
|---|---|---|---|---|
| ***T. orientalis*** | | | | |
| TorRG33x02_013800.1 | 6 | 5 | - | ABC transporter ABCE, partial |
| *TorRG33x02_020930.1* | 10 | 1 | - | Hypothetical protein |
| TorRG33x02_024300.1 | 3 | 3 | - | Hypothetical protein |
| *TorRG33x02_071490.1* | 14 | 12 | - | Receptor-like kinase |
| TorRG33x02_098400.1 | 2 | 2 | - | Hypothetical protein |
| *TorRG33x02_119330.1* | 1 | 1 | - | Ribonuclease |
| TorRG33x02_144520.1 | 2 | 1 | - | Hypothetical protein |
| *TorRG33x02_177120.1* | 1 | 1 | - | Hypothetical protein |
| TorRG33x02_254520.1 | 9 | 4 | - | Alpha/Beta hydrolase fold containing protein DNA polymerase III |
| TorRG33x02_314940.1 | 11 | 5 | - | DNA polymerase III, subunit gamma/ tau |
| TorRG33x02_315370.1 | 1 | 1 | - | Hypothetical protein |
| ***P. andersonii*** | | | | |
| PanWU01x14_022890.1 | 1 | 1 | - | Hypothetical protein |
| PanWU01x14_040010.1 | 2 | 1 | - | Hypothetical protein |
| PanWU01x14_044300.1 | 1 | 1 | - | Hypothetical protein |
| PanWU01x14_108930.1 | 10 | 3 | - | 43kDa postsynaptic protein |

| | | | | |
|---|---|---|---|---|
| PanWU01x14_144790.1 | 3 | 2 | - | Tyrosine-protein kinase |
| PanWU01x14_232680.1 | 7 | 1 | - | Zinc finger, C2H2 domain containing protein, partial |
| PanWU01x14_279230.1 | 20 | 8 | - | Leucine-tRNA ligase, archaeal |
| PanWU01x14_359300.1 | 8 | 7 | - | Carotenoid oxygenase |

***M. truncatula***

None

***L. japonicus***

| | | | | |
|---|---|---|---|---|
| Lj1g3v4833890.1 | 11 | 10 | - | Long chain acyl-CoA synthetase 1-like |
| Lj2g3v0632220.1 | 9 | 9 | - | Probable cyclic nucleotide-gated ion channel 20, chloroplastic-like |
| Lj3g3v2661170.1 | 1 | 1 | - | Lipoxygenase |
| Lj4g3v0353400.1 | 5 | 2 | - | Uncharacterized protein isoform 1 |
| Lj5g3v0837720.1 | 1 | 1 | - | F-box/kelch-repeat protein |
| Lj5g3v2045560.1 | 6 | 2 | - | Protein tesmin/TSO1-like CXC 2-like |

***A. thaliana***

| | | | | |
|---|---|---|---|---|
| AT1G21750.2 | 9 | 9 | PDI5 | PROTEIN DISULFIDE ISOMERASE 5 |
| AT1G62200.1 | 4 | 4 | PTR6 | PEPTIDE TRANSPORTER 6 |
| AT1G69020.1 | 10 | 5 | - | prolyl oligopeptidase family |
| AT1G73300.1 | 12 | 8 | SCPL2 | SERINE CARBOXYPEPTIDASE-LIKE 2 |

| | | | | |
|---|---|---|---|---|
| AT2G16960.2 | 17 | 4 | - | ARM repeat family |
| AT2G17770.2 | 2 | 1 | FDP | FD PARALOG |
| AT3G08900.1 | 3 | 2 | RGP3 | REVERSIBLY GLYCOSYLATED POLYPEPTIDE 3 |
| AT3G10350.1 | 10 | 5 | GET3B | GUIDED ENTRY OF TAIL-ANCHORED PROTEINS 3B |
| AT3G12520.2 | 17 | 1 | SULTR4;2 | SULFATE TRANSPORTER 4;2 |
| AT3G48190.1 | 77 | 73 | ATM | TAXIA-TELANGIECTASIA MUTATED 1 |
| AT3G62040.1 | 5 | 2 | - | haloacid dehalogenase-like hydrolase family |
| AT4G01800.1 | 19 | 4 | SECA1 | CHLOROPLAST SecA |
| AT4G27500.1 | 5 | 5 | PPI1 | PROTON PUMP INTERACTOR 1 |
| AT4G35900.1 | 2 | 1 | FD | FD |

Frequency of GA-AG intron splice sites in annotated gene models of *Parasponia andersonii*, *Trema orientalis*, *Lotus japonicus*, and *Medicago trancatula* genomes. Total number of introns with GA-AG splice sites are based on available annotations.

**Table S2: *T. orientalis* individuals collected in Malaysia, Sabah possess a GA donor splice site at intron 12.**

| species | accession | intron 12 donor splice site | origin | GPS MGRS | | Latitude & Longitude | altitude (meter) |
|---|---|---|---|---|---|---|---|
| *T. orientalis* | RG1 | CACGAG^GAAAGT | Malaysia, Sabah, Poring | 50NMM 58651 | 56071 | 5.958664, 116.603094 | 1018 |
| *T. orientalis* | RG2 | CACGAG^GAAAGT | Malaysia, Sabah, Poring | 50NMM 68352 | 67379 | 6.046485, 116.705216 | 540 |
| *T. orientalis* | RG4 | CACGAG^GAAAGT | Malaysia, Sabah, Poring | 50NMM 68352 | 67379 | 6.046485, 116.705216 | 540 |
| *T. orientalis* | RG5 | CACGAG^GAAAGT | Malaysia, Sabah, Poring | 50NMM 68312 | 67494 | 6.046124, 116.706255 | 496 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| *T. orientalis* | RG8 | CACGAG^GAAAGT | Malaysia, Sabah, Poring | 50NMM 68377 | 67137 | 6.046710, 116.703029 | 499 |
| *T. orientalis* | RG9 | CACGAG^GAAAGT | Malaysia, Sabah, Poring | 50NMM 68377 | 67137 | 6.046710, 116.703029 | 499 |
| *T. orientalis* | RG16 | CACGAG^GAAAGT | Malaysia, Sabah, Sayap | 50NMM 83196 | 51863 | 6.180670, 116.564897 | 852 |
| *T. orientalis* | RG17 | CACGAG^GAAAGT | Malaysia, Sabah, Sayap | 50NMM 83438 | 51942 | 6.182860, 116.565609 | 716 |
| *T. orientalis* | RG18 | CACGAG^GAAAGT | Malaysia, Sabah, Sayap | 50NMM 83446 | 51951 | 6.182932, 116.565690 | 802 |
| *T. orientalis* | RG19 | CACGAG^GAAAGT | Malaysia, Sabah, Sayap | 50NMM 82920 | 51893 | 6.178174, 116.565170 | 787 |
| *T. orientalis* | RG20 | CACGAG^GAAAGT | Malaysia, Sabah, Sayap | 50NMM 82479 | 51936 | 6.174185, 116.565562 | 835 |
| *T. orientalis* | RG23 | CACGAG^GAAAGT | Malaysia, Sabah, Poring | 50NMM 68279 | 67802 | 6.045827, 116.709038 | 467 |
| *T. orientalis* | RG27 | CACGAG^GAAAGT | Malaysia, Sabah, Poring | 50NMM 68282 | 68156 | 6.045856, 116.712237 | 445 |
| *T. orientalis* | RG28 | CACGAG^GAAAGT | Malaysia, Sabah, Poring | 50NMM 68251 | 67752 | 6.045573, 116.708587 | 461 |
| *T. orientalis* | RG29 | CACGAG^GAAAGT | Malaysia, Sabah, Mahua | 50NMM 40705 | 34597 | 5.796153, 116.409250 | 1065 |
| *T. orientalis* | RG31 | CACGAG^GAAAGT | Malaysia, Sabah, Mahua | 50NMM 40499 | 34905 | 5.794292, 116.412034 | 1046 |
| ***T. orientalis*** | **RG33** | **CACGAG^GAAAGT** | **Malaysia, Sabah, Mahua** | **50NMM 40387** | **35073** | **5.793281, 116.413552** | **1050** |
| *T. orientalis* | RG34 | CACGAG^GAAAGT | Malaysia, Sabah, Mahua | 50NMM 38236 | 36940 | 5.773840, 116.430435 | 895 |
| *T. orientalis* | RG36 | CACGAG^GAAAGT | Malaysia, Sabah, Mahua | 50NMM 38255 | 36887 | 5.774012, 116.429956 | 887 |
| *T. orientalis* | RG37 | CACGAG^GAAAGT | Malaysia, Sabah, Mahua | 50NMM 38442 | 36703 | 5.775701, 116.428292 | 895 |
| *T. orientalis* | RG38 | CACGAG^GAAAGT | Malaysia, Sabah, Gunug Alab | 50NMM 43115 | 27210 | 5.817880, 116.342505 | 1827 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| *T. orientalis* | RG39 | CACGAG^GAAAGT | Malaysia, Sabah, Gunug Alab | 50NMM 43132 | 27221 | 5.818034, 116.342604 | 1830 |
| *T. orientalis* | RG40 | CACGAG^GAAAGT | Malaysia, Sabah, Gunug Alab | 50NMM 40871 | 26424 | 5.797573, 116.335429 | 1560 |
| *T. orientalis* | RG41 | CACGAG^GAAAGT | Malaysia, Sabah, Gunug Alab | 50NMM 38135 | 27548 | 5.772835, 116.345609 | 1340 |
| *T. orientalis* | RG45 | CACGAG^GAAAGT | Malaysia, Sabah, Gunug Alab | 50NMM 35934 | 29782 | 5.752948, 116.365808 | 1004 |
| *T. orientalis* | RG50 | CACGAG^GAAAGT | Malaysia, Sabah, Crocker Rang, Inobong | 50NMM 49526 | 03620 | 5.875581, 116.129347 | 261 |
| *T. orientalis* | RG52 | CACGAG^GAAAGT | Malaysia, Sabah, Crocker Rang, Inobong | 50NMM 50645 | 03744 | 5.885704, 116.130451 | 205 |
| *T. orientalis* | RG53 | CACGAG^GAAAGT | Malaysia, Sabah, Crocker Rang, Inobong | 50NMM 50648 | 03738 | 5.885736, 116.130401 | 205 |
| *T. orientalis* | WU41 | CACGAG^GCAAGT | China | not determined | | not determined | not determined |
| *T. orientalis* | WU42 | CACGAG^GCAAGT | China | not determined | | not determined | not determined |
| *T. orientalis* | WU43 | CACGAG^GCAAGT | China | not determined | | not determined | not determined |
| *T. orientalis* | WU44 | CACGAG^GCAAGT | China | not determined | | not determined | not determined |
| *T. orientalis* | WU45 | CACGAG^GCAAGT | China | not determined | | not determined | not determined |
| *T. levigata* | WU50 | CACGAG^GCAAGT | China | not determined | | not determined | not determined |
| *T. tomentosa* | WU10 | CACGAG^GCAAGT | Australia | not determined | | not determined | not determined |
| *P. andersonii* | WU1 | CACGAG^GCAAGT | Papua-New Guinea | not determined | | not determined | not determined |

Twenty-eight *Trema orientalis* individuals collected at 5 distinct locations in Malaysia, Sabah all possess a non-canonical GA donor splice site at intron 12. *T. orientalis* plants collected from locations outside Malaysia were found to possess a GC donor splice site at intron 12.

**Table S3: List of Golden Gate constructs used in this study.**

| Construct | ID | Description | Level | Backbone | contains |
|---|---|---|---|---|---|
| 1 | EC75056 | *nptII* resistence cassette | 1 | pICH47802 | pICSL70004:*nptII* |
| 2 | EC74638 | $35S_{pro}$:ΩNLS-Cas9:$35S_{ter}$ | 1 | pICH47742 | pICH41388:$35S_{pro}$, pAGM5331:ΩNLS, pICH41308::aCas9, pICH41414:$35S_{ter}$ |
| 3 | EC74578 | PanSYMRKsgRNA1 | 1 | pICH47761 | pICSL01009:AtU6p, corresponding PCR amplicon |
| 4 | EC74579 | PanSYMRKsgRNA2 | 1 | pICH47751 | pICSL01009:AtU6p, corresponding PCR amplicon |
| 5 | EC74836 | CRISPR_ctrl | 2 | pICSL4723 | 1R: construct 1, 2F: construct 2, end-link pICH41744 |
| 6 | EC74796 | CRISPR_PanSYMRK | 2 | pICSL4723 | 1R: construct 1, 2F: construct 2, 3F: construct 3, 4F: construct 4; end-link pICH41766 |
| 7 | EC75091 | PanSYMRK Part1 $_{Pro}$:$SYMR_{gene}$: $SYMRK_{ter}$ | 1 | pICH47742 | EC75120:$SYMRK_{Pro,}$ EC75123: $SYMRK_{gene}$ part1, EC75124: $SYMRK_{gene}$ part2, EC75122: $SYMRK_{ter}$ |
| 8 | EC75092 | PanSYMRK Part1 $_{Pro}$:$SYMRK_{gene-GA}$ :$SYMRK_{ter}$ | 1 | pICH47742 | EC75120:$SYMRK_{Pro}$, EC75123: $SYMRK_{gene}$ part1, EC75125: $SYMRK_{gene-GA}$ part2, EC75122: $SYMRK_{ter}$ |
| 9 | EC75093 | PanSYMRK $Part2_{Pro}$ | 1 | pICH47732 | EC75121:$SYMRK_{Pro}$ part2 |
| 10 | EC75022 | $35S_{Pro}$:$erGFP_{gene}$:$t35S_{ter}$ | 1 | pICH47831 | pICH51277:$35S_{Pro}$, EC74047:$erGFP_{gene}$, pICH41414:$35S_{ter}$ |

| | | | | | |
|---|---|---|---|---|---|
| 11 | EC75220 | $35S_{Pro}$:erGFP$_{gene}$:t$35S_{ter}$ | 2 | MOB215_pICS L4723_modified to Spec R | Dummy1,Dummy2,Dummy3,4R: Construct 10, end-link pICH41780 |
| 12 | EC75223 | PanSYMRK$_{Pro}$ Part2, SYMRK$_{Pro}$ Part1, SYMRK$_{gene}$:SYMRK$_{ter}$, $35S_{Pro}$:erGFP$_{gene}$:$35S_{ter}$ | 2 | MOB215_pICS L4723_modified to Spec R | 1F: Construct 9, 2F: Construct 7, Dummy 3, 4R: Construct 10, end-link pICH41780 |
| 13 | EC75224 | PanSYMRK$_{Pro}$ Part2, SYMRK$_{Pro}$ Part1, SYMRK$_{gene-GA}$:SYMRK$_{ter}$, $35S_{Pro}$:erGFP$_{gene}$:$35S_{ter}$ | 2 | MOB215_pICS L4723_modified to Spec R | 1F: Construct 9, 2F: Construct 8, Dummy 3, 4R: Construct 10, end-link pICH41780 |
| 14 | EC75926 | LjUBQ1$_{Pro}$: PanSYMRK$_{gene}$ part1: tNOS$_{ter}$ | 1 | pICH47811 | EC74013:LjUBQ1$_{Pro,}$ EC75123: SYMRK$_{gene}$ part1, EC75124: SYMRK$_{gene}$ Part2, pICH41421:tNOS |
| 15 | EC75668 | LjpUBQ1$_{pro}$, PanSYMRK$_{gene}$, $35S_{Pro}$:erGFP$_{gene}$:35Ster | 2 | MOB215_pICS L4723_modified to Spec R | 1F: Construct 1, 2R:construct 14, Dummy 3, 4R: contstruct 10, end-link pICH41780 |

Each construct is identified by a unique number, an ID, and a brief description. The level of Golden Gate assembly for each vector is indicated, as well as the plasmid backbone into which the constructs were cloned.

## Table S4: Primers used in this study.

| Name | Purpose | Sequence |
|---|---|---|
| PanSYMRK_For | qRT-PCR | GTCCTCGGGTTCCAGTTTG |
| PanSYMRK_Rev | qRT-PCR | ATCACATCCGCATCATTTGG |
| PanEf-1a_For | qRT-PCR | AGACAAGGTTAAGCGTGCAG |
| PanEf-1a_Rev | qRT-PCR | TGCAACTGGGCAACAAACTC |
| PanACT_For | qRT-PCR | CCTCATTGGAATGGAAGCAC |
| PanACT_Rev | qRT-PCR | TTCCAGGAAACATGGTGGAC |
| PanSYMRK_unspliced_For | qRT-PCR | ACAGGAGAAGCATCAGCAAG |
| PanSYMRK_unspliced_Rev | qRT-PCR | ATGAACAAGCTAACATCCAAGG |
| PanSYMRK_spliced_For | qRT-PCR | ACAGGAGAAGCATCAGCAAG |
| PanSYMRK_spliced_Rev | qRT-PCR | ATGTGTCAATCCTCGTGCAG |
| TorSYMRK_unspliced_For | qRT-PCR | CCAACCAGACTTTCGATTGC |
| TorSYMRK_unspliced_Rev | qRT-PCR | CAAGCTAACATCAAAGGCACTG |
| TorSYMRK_spliced_For | qRT-PCR | ACAGGAGAAGCATCAGCAAG |
| TorSYMRK_spliced_Rev | qRT-PCR | ATGTGTCAATCCTCGTGCAG |
| sgRNA-Rv | CRISPR assembly | TGTGGTCTCCAAGCGTAATGCCAACTTTGTAC |

| | | |
|---|---|---|
| PanSYMRK_sgRNA1 | CRISPR assembly | TGTGGTCTCAATTGACCCATGTCTCCCAACTCCCGTTTTAGAGCTAGAAATAGCAAG |
| PanSYMRK_sgRNA2 | CRISPR assembly | TGTGGTCTCAATTGACCAACTTAAACACCCTGTAGTTTTAGAGCTAGAAATAGCAAG |
| geno_PanSYMRK-KO-Fw | Genotyping CRISPR mutants | TTCCAGTTTGGGGCCCATTT |
| geno_PanSYMRK-KO-Rv | Genotyping CRISPR mutants | GGAGGAAGAAGGAAAGTCCGG |

List of primers, their purpose, and their sequence that were used in various applications in this study.

**References**

1. **Remy W, Taylor TN, Hass H, Kerp H.** Four hundred-million-year-old vesicular arbuscular mycorrhizae. Proc Natl Acad Sci U S A. 1994;**91**:11841–3.
2. **Brundrett MC, Tedersoo L.** Evolutionary history of mycorrhizal symbioses and global host plant diversity. New Phytol. 2018;**220**:1108–15.
3. **Radhakrishnan GV, Keller J, Rich MK, Vernié T, Mbadinga Mbadinga DL, Vigneron N, et al.** An ancestral signalling pathway is conserved in intracellular symbioses-forming plant lineages. Nat Plants. 2020;**6:**280–9.
4. **Duc G, Trouvelot A, Gianinazzi-Pearson V, Gianinazzi S.** First report of non-mycorrhizal plant mutants (Myc−) obtained in pea (Pisum sativum L.) and fababean (Vicia faba L.). Plant Sci. 1989;**60**:215–22.
5. **Parniske M.** Arbuscular mycorrhiza: the mother of plant root endosymbioses. Nat Rev Microbiol. 2008;**6**:763–75.
6. **Oldroyd GED.** Speak, friend, and enter: signalling systems that promote beneficial symbiotic associations in plants. Nat Rev Microbiol. 2013;**11**:252–63.
**7. Endre G, Kereszt A, Kevei Z, Mihacea S, Kaló P, Kiss GB.** A receptor kinase gene regulating symbiotic nodule development. Nature. 2002;**417**:962–6.
8. **Stracke S, Kistner C, Yoshida S, Mulder L, Sato S, Kaneko T, et al.** A plant receptor-like kinase required for both bacterial and fungal symbiosis. Nature. 2002;**417**:959–62.
9. **Markmann K, Giczey G, Parniske M.** Functional adaptation of a plant receptor-kinase paved the way for the evolution of intracellular root symbioses with bacteria. PLoS Biol. 2008;**6**:0497–506.
10. **Kosuta S, Held M, Hossain MS, Morieri G, Macgillivary A, Johansen C, et al.** Lotus japonicus symRK-14 uncouples the cortical and epidermal symbiotic program. Plant J. 2011;**67**:929–40.
11. **Antolín-Llovera M, Ried MK, Parniske M.** Cleavage of the SYMBIOSIS RECEPTOR-LIKE KINASE ectodomain promotes complex formation with Nod factor receptor 5. Curr Biol. 2014;**24**:422–7.
12. **Pan H, Stonoha-Arther C, Wang D.** Medicago plants control nodulation by regulating proteolysis of the receptor-like kinase DMI2. Plant Physiol. 2018.

13. **Madsen EB, Madsen LH, Radutoiu S, Olbryt M, Rakwalska M, Szczyglowski K, et al.** A receptor kinase gene of the LysM type is involved in legume perception of rhizobial signals. Nature. 2003;**425**:637–40.
14. **Gysel K, Laursen M, Thygesen MB, Lironi D, Bozsóki Z, Hjuler CT, et al.** Kinetic proofreading of lipochitooligosaccharides determines signal activation of symbiotic plant receptors. Proc Natl Acad Sci U S A. 2021;118.
15. **Catoira R, Galera C, de Billy F, Penmetsa RV, Journet EP, Maillet F, et al.** Four genes of Medicago truncatula controlling components of a nod factor transduction pathway. Plant Cell. 2000;**12**:1647–66.
16. **Indrasumunar A, Wilde J, Hayashi S, Li D, Gresshoff PM.** Functional analysis of duplicated Symbiosis Receptor Kinase (SymRK) genes during nodulation and mycorrhizal infection in soybean (Glycine max). J Plant Physiol. 2015;**176**:157–68.
17. **Feng Y, Wu P, Liu C, Peng L, Wang T, Wang C, et al.** Suppression of LjBAK1-mediated immunity by SymRK promotes rhizobial infection in Lotus japonicus. Mol Plant. 2021. https://doi.org/10.1016/j.molp.2021.07.016.
18. **Calantzis C, Morandil D, Arnouldl C, Gianinazzi-Pearsonl V.** Cellular Interactions between G. mosseae and a Myc" dmi2 mutant in Medicago truncatula. 2001. https://dalspace.library.dal.ca/bitstream/handle/10222/77857/VOLUME%2030-NUMBERS%202-3-2001-PAGE%2097.pdf?sequence=1. Accessed 25 Jul 2022.
19. **Limpens E, Mirabella R, Fedorova E, Franken C, Franssen H, Bisseling T, et al.** Formation of organelle-like N2-fixing symbiosomes in legume root nodules is controlled by DMI2. Proc Natl Acad Sci U S A. 2005;**102**:10375–80.
20. **Gherbi H, Markmann K, Svistoonoff S, Estevan J, Autran D, Giczey G, et al.** SymRK defines a common genetic basis for plant root endosymbioses with arbuscular mycorrhiza fungi, rhizobia, and Frankiabacteria. Proc Natl Acad Sci U S A. 2008;**105**:4928–32.
21. **Griesmann M, Chang Y, Liu X, Song Y, Haberer G, Crook MB, et al.** Phylogenomics reveals multiple losses of nitrogen-fixing root nodule symbiosis. Science. 2018;**361**:eaat1743.
22. **van Velzen R, Doyle JJ, Geurts R.** A Resurrected Scenario: Single Gain and Massive Loss of Nitrogen-Fixing Nodulation. Trends Plant Sci.

2019;**24**:49–57.

23. **van Velzen R, Holmer R, Bu F, Rutten L, van Zeijl A, Liu W, et al.** Comparative genomics of the nonlegume Parasponia reveals insights into evolution of nitrogen-fixing rhizobium symbioses. Proceedings of the National Academy of Sciences. 2018;**115**:E4700–9.
24. **Rutten L, Miyata K, Roswanjaya YP, Huisman R, Bu F, Hartog M, et al.** Duplication of Symbiotic Lysin Motif Receptors Predates the Evolution of Nitrogen-Fixing Nodule Symbiosis. Plant Physiol. 2020;**184**:1004–23.
25. **Merckx VSFT, Hendriks KP, Beentjes KK, Mennes CB, Becking LE, Peijnenburg KTCA, et al.** Evolution of endemism on a young tropical mountain. Nature. 2015;**524**:347–50.
26. **Capoen W, Goormachtig S, De Rycke R, Schroeyers K, Holsters M.** SrSymRK, a plant receptor essential for symbiosome formation. Proc Natl Acad Sci U S A. 2005;**102**:10369–74.
27. **Trouvelot A., Kough J.L., Gianinazzi-Pearson V**. Mesure du taux de mycorhization VA d'un systeme radiculaire. Recherche de methodes d'estimation ayant une signification fonctionnelle. In: Gianinazzi-Pearson GSV, editor. Physiological and Genetic Aspects of Mycorrhizae. INRA press; 1986. p. 217–21.
28. **Demchenko K, Winzer T, Stougaard J, Parniske M, Pawlowski K.** Distinct roles of Lotus japonicus SYMRK and SYM15 in root colonization and arbuscule formation. New Phytol. 2004;**163**:381–92.
29. **Ried MK, Antolín-Llovera M, Parniske M.** Spontaneous symbiotic reprogramming of plant roots triggered by receptor-like kinases. Elife. 2014;**3**:1–17.
30. **Kevei Z, Lougnon G, Mergaert P, Horváth GV, Kereszt A, Jayaraman D, et al.** 3-Hydroxy-3-Methylglutaryl Coenzyme A Reductase1 Interacts with NORK and Is Crucial for Nodulation in Medicago truncatula. Plant Cell. 2007;**19**:3974–89.
31. **Jayaraman D, Richards AL, Westphall MS, Coon JJ, Ané J-M.** Identification of the phosphorylation targets of symbiotic receptor-like kinases using a high-throughput multiplexed assay for kinase specificity. Plant J. 2017;**90**:1196–207.
32. **Sato S, Nakamura Y, Kaneko T, Asamizu E, Kato T, Nakao M, et al.** Genome structure of the legume, Lotus japonicus. DNA Res.

2008;**15**:227–39.

33. **Pecrix Y, Staton SE, Sallet E, Lelandais-Brière C.** Whole-genome landscape of Medicago truncatula symbiotic genes. Nature Plants. 2018.
34. **Lamesch P, Berardini TZ, Li D, Swarbreck D, Wilks C, Sasidharan R, et al.** The Arabidopsis Information Resource (TAIR): improved gene annotation and new tools. Nucleic Acids Res. 2012;**40** Database issue:D1202–10.
35. **Parada GE, Munita R, Cerda CA, Gysling K.** A comprehensive survey of non-canonical splice sites in the human transcriptome. Nucleic Acids Res. 2014;**42**:10564–78.
36. **Frey K, Pucker B.** Animal, Fungi, and Plant Genome Sequences Harbor Different Non-Canonical Splice Sites. Cells. 2020;9.
37. **Lee CE.** Evolutionary mechanisms of habitat invasions, using the copepod Eurytemora affinis as a model system. Evol Appl. 2016;**9**:248–70.
38. **Robertson HM.** Noncanonical GA and GG 5’ Intron Donor Splice Sites Are Common in the Copepod Eurytemora affinis. G3 Genes|Genomes|Genetics. 2017;**7**:3967–9.
39. **Bu F, Rutten L, Roswanjaya YP, Kulikova O, Rodriguez-Franco M, Ott T, et al.** Mutant analysis in the non-legume Parasponia andersonii identifies NIN and NF-YA1 transcription factors as a core genetic network in nitrogen-fixing nodule symbioses. New Phytol. 2020.
40. **van Zeijl A, Wardhani TAK, Seifi Kalhor M, Rutten L, Bu F, Hartog M, et al.** CRISPR/Cas9-Mediated Mutagenesis of Four Putative Symbiosis Genes of the Tropical Tree Parasponia andersonii Reveals Novel Phenotypes. Front Plant Sci. 2018;**9**:284.
41. **Wardhani TAK, Roswanjaya YP, Dupin S, Li H, Linders S, Hartog M, et al.** Transforming, Genome Editing and Phenotyping the Nitrogen-fixing Tropical Cannabaceae Tree Parasponia andersonii. J Vis Exp. 2019. https://doi.org/10.3791/59971.
42. **Nekrasov V, Staskawicz B, Weigel D, Jones JDG, Kamoun S.** Targeted mutagenesis in the model plant Nicotiana benthamiana using Cas9 RNA-guided endonuclease. Nat Biotechnol. 2013;**31**:691–3.
43. **Engler C, Youles M, Gruetzner R, Ehnert T-M, Werner S, Jones JDG, et al.** A golden gate modular cloning toolbox for plants. ACS Synth Biol. 2014;**3**:839–43.

44. **Cao Q, den Camp RO, Kalhor MS, Bisseling T, Geurts R.** Efficiency of Agrobacterium rhizogenes–mediated root transformation of Parasponia and Trema is temperature dependent. Plant Growth Regul. 2012;**68**:459–65.
45. **Schenk RU, Hildebrandt AC.** Medium and techniques for induction and growth of monocotyledonous and dicotyledonous plant cell cultures. Can J Bot. 1972;**50**:199–204.
46. **Robinson JT, Thorvaldsdóttir H, Winckler W, Guttman M, Lander ES, Getz G, et al.** Integrative genomics viewer. Nat Biotechnol. 2011;29:24–6.
47. **Jones P, Binns D, Chang H-Y, Fraser M, Li W, McAnulla C, et al.** InterProScan 5: genome-scale protein function classification. Bioinformatics. 2014;**30**:1236–40.
48. **Sievers F, Wilm A, Dineen D, Gibson TJ, Karplus K, Li W, et al.** Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. Mol Syst Biol. 2011;**7**:539.
49. **Guindon S, Dufayard J-F, Lefort V, Anisimova M, Hordijk W, Gascuel O.** New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. Syst Biol. 2010;**59**:307–21.
50**. Letunic I, Bork P.** Interactive Tree Of Life (iTOL): an online tool for phylogenetic tree display and annotation. Bioinformatics. 2007;**23**:127–8.
51. **Chiapello M, Das D, Gutjahr C.** Ramf: An Open-Source R Package for Statistical Analysis and Display of Quantitative Root Colonization by Arbuscular Mycorrhiza Fungi. Front Plant Sci. 2019;**10**:118

CHAPTER 3

# 3

# Ectopic expression of the GRAS-type transcriptional regulator *NSP2* in *Parasponia* triggers contrasting effects on symbioses

Sultan Alhusayni[1,2] | Nick Kersten[1] | Rik Huisman[1] | Rene Geurts[1*] | Joël Klein[1*]

[1]Laboratory of Molecular Biology, Cluster of Plant Development, Plant Science Group, Wageningen University, Droevendaalsesteeg 1, 6708 PB Wageningen, The Netherlands

[2]Biological Sciences Department, College of Science, King Faisal University, Al-Ahsa 31982, Saudi Arabia

*Correspondence: Corresponding Authors: rene.geurts@wur.nl, joel.klein@wur.nl

**Keywords:** Arbuscular mycorrhiza, nodulation, Parasponia, *NSP2*, *CYCLOPS*, *PHYTOENE SYNTHASE* , carotenoids, strigolactones

*Published in Front. Plant Sci. 15*
*doi: 10.3389/fpls.2024.1468812 (2024)*

**Abstract**

**Introduction:** Plants strictly control root endosymbioses with nutrient-scavenging arbuscular endomycorrhizal fungi or nodule inducing diazotrophic bacteria. The GRAS-type transcriptional regulator NODULATION SIGNALING PATHWAY 2 (NSP2) is a conserved hub in this process. The NSP2-regulated transcriptional network is instrumental in balancing nutrient homeostasis with symbiotic interactions. NSP2 activity is modulated post-transcriptionally by a specific microRNA. Overriding this control mechanism by ectopic expression of a miRNA-resistant *NSP2* transgene enhances the symbiotic permissiveness to arbuscular endomycorrhizal fungi. Such engineered plants may possess enhanced capacities for nutrient uptake. However, the trade-off of this strategy on plant development or other symbiotic interactions, like nodulation, is yet to be fully understood.

**Method:** We used the nodulating *Cannabaceae* species *Parasponia andersonii* as an experimental system to study the effect of ectopic *NSP2* expression. Parasponia and legumes (Fabaceae) diverged 100 million years ago, providing a unique comparative system to dissect the nodulation trait.

**Results:** Six independent transgenic *Parasponia* lines were generated that differed in the level of *NSP2* expression in the root from 6 to 95-fold higher when compared to the empty vector control plants. Analysis of these plants revealed a positive correlation between mycorrhization and the NSP2 expression level, as well as with the expression of the symbiosis transcription factor *CYCLOPS* and the rate-limiting enzyme in the carotenoid biosynthetic pathway *PHYTOENE SYNTHASE1* (*PSY1*). Yet ectopic expression of *NSP2* affected plant architecture and root nodule organogenesis.

**Discussion:** This indicates a significant trade-off when leveraging *NSP2* over-expression to enhance endomycorrhization.

Chapter 3

## Introduction

Plants explore mutualistic relationships with soil microorganisms to enhance access to essential nutrients. The endosymbiotic interactions with arbuscular mycorrhizal fungi and nodulating nitrogen-fixing bacteria like rhizobia are the most advanced examples of such ecosystem services. As these microbes are engulfed in plant cells, they largely depend on the plant's supply of carbon sources. These intimate endosymbiotic associations necessitate stringent regulation to prevent the microbes from exploiting the plant, especially when excess exogenous nutrients are in the soil.

The GRAS-type transcriptional regulator NODULATION SIGNALING PATHWAY2 (NSP2) acts as a central hub integrating the plant's nutrient homeostasis and symbiotic engagement to orchestrate adaptive responses (Quilbé and Arrighi, 2021). *NSP2* is first identified in legumes, where it is essential for rhizobium-induced nodule formation (Oldroyd and Long, 2003; Kaló *et al.*, 2005; Heckmann *et al.*, 2006; Ali *et al.*, 2014; Shtark *et al.*, 2016; Peng *et al.*, 2021). Subsequent genome studies revealed the gene is present in many plant species and predates the evolution of angiosperms (Lauressergues *et al.*, 2012; Cenci and Rouard, 2017). *NSP2* of different species is, to a large extent, functionally conserved, as was shown in *trans*-complementation studies of a *Lotus japonicus nsp2* nodulation mutant using *OsNSP2* of rice (*Oryza sativa*) (Heckmann *et al.*, 2006). In legumes and non-legumes, NSP2 also controls mycorrhizal infection levels (Lauressergues *et al.*, 2012; Shtark *et al.*, 2016; Li *et al.*, 2022). Studies in barley (*Hordeum vulgare*) suggest that NSP2 may do so by regulating the expression of several common symbiosis signaling genes, including the transcription factor *HvCYCLOPS* (Li *et al*., 2022).

NSP2 also acts as a transcriptional regulator of the strigolactone biosynthesis pathway in a nutrient status-dependent manner (Liu *et al.*, 2011; Li *et al.*, 2022). Strigolactones act *in planta* and *ex planta. In planta,* strigolactones function as developmental hormones and control axillary bud and lateral root outgrowth (Ruyter-Spira *et al.*, 2011; Brewer *et al.*, 2015;

Marzec and Melzer, 2018). Upon phosphate starvation, plants exude strigolactones into their rhizosphere, affecting the microbial community composition, including the infectiveness of arbuscular mycorrhizal fungi (Schlemper *et al.*, 2017; Kobae *et al.*, 2018; Carvalhais *et al.*, 2019). Together, these findings highlight the complex regulatory role of NSP2 as a signaling hub.

NSP2 interacts with a suite of transcription factors to control the expression of target genes. Studies in heterologous systems showed it interacts with other GRAS-type transcriptional regulators like NSP1, DELLA, REQUIRED FOR ARBUSCULAR MYCORRHIZATION1 (RAM1), and REQUIRED FOR ARBUSCULAR DEVELOPMENT1 (RAD1), and the MYB-type transcription factor INTERACTING PROTEIN OF NSP2 (IPN2) to control nodulation and/or mycorrhization (Hirsch *et al.*, 2009; Gobbato *et al.*, 2012; Park *et al.*, 2015; Xue *et al.*, 2015; Fonouni-Farde *et al.*, 2016; Heck *et al.*, 2016; Jin *et al.*, 2016; Pimprikar *et al.*, 2016; Xiao *et al.*, 2020). These transcriptional modules are regulated in a complex manner. Studies in rice showed that *NSP2* expression is controlled by the MYB transcription factor PHOSPHATE STARVATION RESPONSE 2 (OsPHR2), in response to phosphate starvation (Das *et al.*, 2022; Yuan *et al.*, 2023). Additionally, *NSP2* translation is controlled post-transcriptionally by microRNA *miR171h* (Lauressergues *et al.,* 2012; Hofferek *et al.,* 2014). Transcript levels of *miR171h* correlate with available phosphate concentration, suggesting that *NSP2* is controlled transcriptionally and post-transcriptionally in a nutrient-dependent manner (Hofferek *et al.*, 2014).

To enhance sustainable agricultural productivity, reducing reliance on inorganic fertilizers is essential. This can be achieved by better utilizing the arbuscular mycorrhizal symbiosis. The ectopic expression of *NSP2* improves mycorrhization under exogenous phosphate concentrations that generally inhibit the symbiosis, suggesting it represents a potential biotechnological strategy to enhance nutrient uptake by the plant (Li *et al.*, 2022; Isidra-Arellano, Singh and Valdés-López, 2024; Yuan *et al.*, 2023). However, at this stage, it remains elusive what the trade-off of such a strategy will be on plant development or other symbiotic interactions, like nodulation.

Studies in the legume models showed that NSP2 overexpression had no effect on nodulation (*Medicago truncatula*) or resulted in clustering of nodules on the root though not affecting the total number of nodules (*L. japonicus*) (Lauressergues *et al*., 2012; Murakami *et al*., 2013). We used the nodulating species *Parasponia andersonii* (*Parasponia*) as a comparative experimental system complementary to legumes to study the effect of ectopic *NSP2* expression. The nodulation trait in *Parasponia* (Cannabaceae) and legumes (Fabaceae) share an evolutionary origin but diverged soon after (~100 million years ago) (van Velzen, Doyle and Geurts, 2019). *Parasponia* nodules have a different ontogeny when compared to legumes (Behm *et al*., 2014). Nevertheless, the same core set of symbiotic genes is used to establish nitrogen-fixing root nodules, including *NSP2* (van Velzen *et al*., 2018; van Zeijl *et al*., 2018; Bu *et al*., 2020; Rutten *et al*., 2020; Alhusayni *et al*., 2023). *Parasponia* is amenable for efficient *Agrobacterium tumefaciens*-mediated transformation, giving T0 transgenic plantlets within ~3 months that can subsequently be propagated vegetatively (Wardhani *et al.*, 2019). We generated independent transgenic *Parasponia* lines that differ in the level of *NSP2* expression from 6 up to 95-fold compared to the empty vector control plants. Analysis of these plants revealed a correlation between mycorrhization and the level of expression of *NSP2,* the transcription factor *CYCLOPS*, and all genes in the biosynthetic pathways required for the formation of strigolactones. Yet ectopic expression of *NSP2* affected plant shoot development and root nodule organogenesis.

## Results

### *Parasponia NSP2*ox lines have a root and shoot branching phenotype

*Parasponia PanNSP2* was studied previously, revealing it represents a single-copy gene essential for rhizobium-induced nodulation (van Zeijl *et al.*, 2018). To enable ectopic expression of a stable *PanNSP2* transcript, we first identified the putative *miR171h* target site (Lauressergues *et al.*, 2012). Such a putative site was identified in the coding region and subsequently

removed by introducing synonymous DNA mutations **(Supplementary Figure S1A)**. An additional alteration was included in the *PanNSP2* coding region, which may enhance expression stability; the insertion of an intron **(Supplementary Figure S1B)** (Feike *et al.*, 2019). Ultimately, two *miR171h* resistant *PanNSP2* versions (mNSP2) were used for *Parasponia* transformation, a construct with and one without an engineered intron, both driven by the constitutive *L. japonicus UBIQUITIN1* promoter (*pLjUBQ1*) **(Supplementary Figure S1B)** (Maekawa *et al.*, 2008).

Six transgenic *Parasponia* lines were selected, displaying a gradient level of *PanNSP2* ectopic expression ranging from a 6 to 95-fold increase in root tissue **(Figure 1A)**. In the shoots of these transgenic lines, m*NSP2* was also expressed; in tissues where *PanNSP2* transcripts are normally not detected **(Supplementary Figure S2A)**. Noteworthy, lines possessing the m*NSP2*intron construct generally have a lower transgene expression when compared to constructs possessing m*NSP2* with only an adapted *miR171h* target site.



To analyze whether morphogenic phenotypes are associated with m*NSP2* ectopic expression in *Parasponia*, we selected lines 1, 3, and 6 that possess a 6-, 51-, and 95-fold enhanced *NSP2* expression in root tissue compared to empty vector control plants. First, the root systems of 34-day-old plantlets grown *in vitro* on EKM medium containing low levels of ammonium nitrate (0.375 mM) and high levels of phosphate (3 mM) were analyzed. Control plants transformed with an empty vector showed a highly branched root system with up to 80 secondary lateral roots. This contrasts with what was observed in m*NSP2ox* lines 3 and 6, in which the root system was less branched, with only ~30 secondary lateral roots, whereas the primary root length is unaffected **(Figures 1B, C)**. Such root phenotype was not observed on m*NSP2ox* line 1. As this line has only moderately levels of transgene expression compared to lines 3 and 6, it suggests that repression of lateral root growth has a threshold of *PanNSP2* expression. Next, we analyzed the shoot architecture of 60-day-old plantlets grown on potting soil. In the case of shoot branching, control plants transformed with an empty vector exhibited an average of 11 branches. In contrast, *mNSP2ox* lines showed

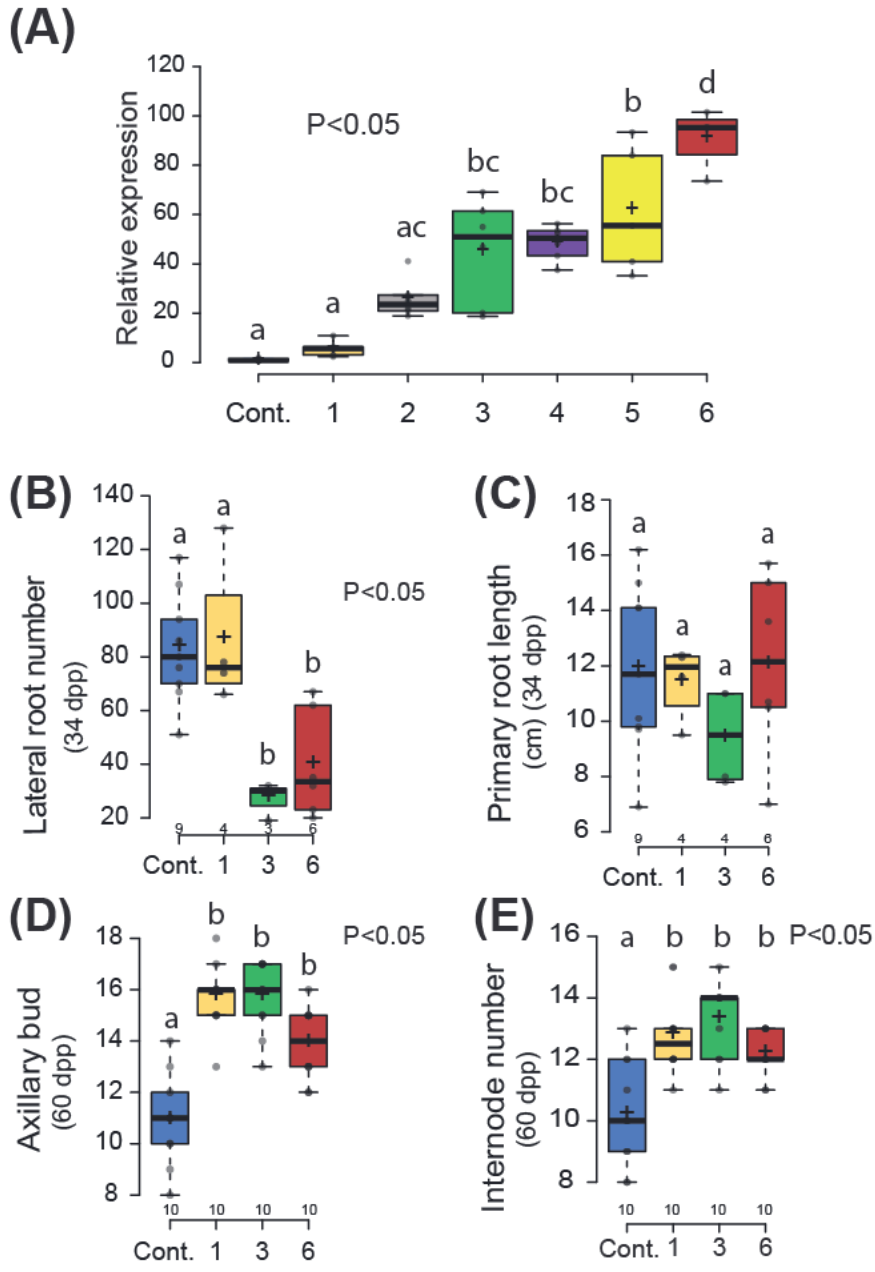



**Figure 1. Non-symbiotic phenotype of *Parasponia* m*NSP2ox* lines.** (**A**) Relative expression of *NSP2* in root tissue of 6 independent lines compared to an empty vector control line (Cont.). m*NSP2ox* lines 1 to 3 contain the transgene construct *pLjUBQ1:mPanNSP2intron*,

whereas lines 4 to 6 carry *pLjUBQ1:mNSP2*. Expression was measured using qPCR (n=3). (**B, C**) Quantification of the number of lateral roots (**B**) and primary root length (**C**) of control and transgenic lines 1,3, and 6 *in vitro* grown 34 days post planting (dpp). (**D, E**) Quantification of number of lateral shoot branching (**D**) and internode (**E**) of m*NSP2ox* lines 1, 3, and 6 at 60 days post planting (dpp). Different letters indicate significant differences ($p<0.05$) between these lines as determined by One-way ANOVA in combination with Tukey's post hoc test. All data are displayed in box plots, showing data points, the median, and the interquartile range (IQR).

increased shoot branching, with an average of 14 to 16 branches, respectively **(Figure 1D)**. This increased number of lateral branches is associated with an increased number of internodes **(Figure 1E)**. This phenotype is also observed in m*NSP2ox* line 1, indicating that NSP2-controlled developmental effects on plant architecture in root and shoot might have a different threshold. Quantification of internode diameter suggests that ectopic *PanNSP2* expression may also increase stem thickness, even though the trees were all similar in size **(Supplementary Figure S3)**. Taken together, the data show that ectopic m*NSP2* expression affects shoot and root development in *Parasponia.*



### *Parasponia NSP2*ox lines possess enhanced mycorrhizal infection

Over-expression of an *NSP2 miR171h* resistant allele in *M. truncatula* roots enhances arbuscular mycorrhizal colonization under phosphate-limited conditions (Lauressergues *et al.*, 2012). Furthermore, a more recent study showed that in *M. truncatula* and barley *miR171h* resistant *NSP2* over-expression promotes arbuscular mycorrhizal root colonization even under exogenous phosphate conditions that are less amenable for the mycorrhizal symbiosis. This response is associated with the up-regulation of the strigolactone biosynthesis pathway (Li *et al.*, 2022). We questioned whether *Parasponia* m*NSP2ox* lines also possess enhanced mycorrhization colonization levels. We tested two phosphate conditions; 20 µM and 3 mM exogenous Pi, and inoculated all six m*NSP2ox* lines with *Rhizophagus irregularis* DOAM197198 spores. Mycorrhization levels were scored 6 weeks post inoculation. Empty vector control plants can be colonized under both conditions, though the high exogenous phosphate concentration affects

arbuscular mycorrhizal infection levels, most notably the number of vesicles formed. In the m*NSP2ox* lines, the number of hyphae, arbuscules, and

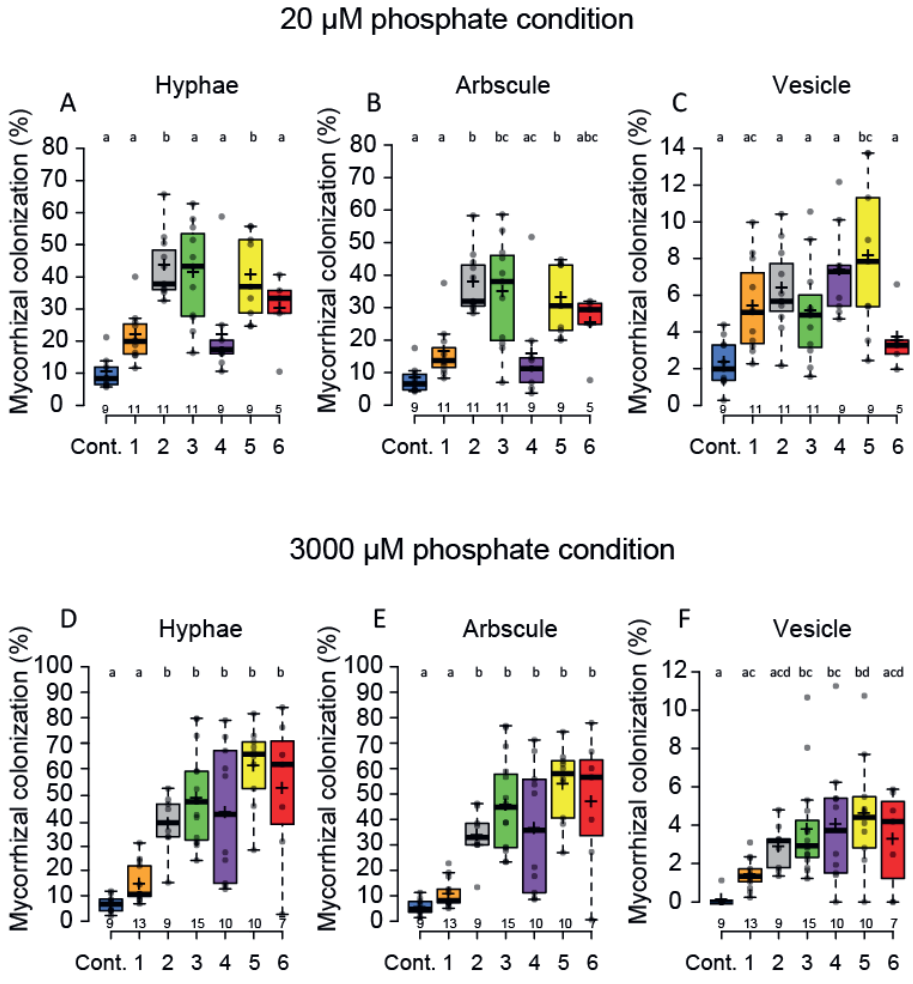



**Figure 2. Ectopic expression of *PanNSP2* in *Parasponia* enhances mycorrhizal colonization.** Mycorrhizal colonization grown at low exogenous phosphate condition (20 µM $PO_4^{3-}$) (**A-C**) and high exogenous phosphate condition (3000 µM $PO_4^{3-}$) (**D-F**). *Parasponia* lines 1 to 3 are transformed with *pLjUBQ1:mPanNSP2intron* construct, whereas lines 4 to 6 contain *pLjUBQ1:mPanNSP2*. Different letters indicate significant differences ($p<0.05$) between these lines as determined by One-way ANOVA followed by a Tukey's post hoc test. Plants were harvested 6 weeks post-inoculation (6 wpi) with 300 spores/plant. All data are displayed in box plots, showing data points, the median, and the interquartile range (IQR).

vesicles was significantly increased under both phosphate conditions when compared to the empty vector control line **(Figure 2)**. At high exogenous phosphate conditions, the level of mycorrhizal colonization is associated with the level of *PanNSP2* expression. Thus, we conclude that ectopic expression of *PanNSP2* in *Parasponia* promotes mycorrhization, which is consistent with reported data in *M. truncatula* and barley.

### ***NSP2*** **enhances the expression of genes in the carotenoid biosynthetic pathway under nutrient starved and nodulation permissive growth conditions**

To identify genes that are potentially directly or indirectly regulated by NSP2 in *Parasponia*, we investigated the transcriptional effect of increased m*NSP2* expression under two distinct nutrient conditions. First plants were grown for 6 weeks on nutrient-rich medium containing 24.72 mM $NO_3^-$, 2.6 mM $NH_4^+$, and 2.6 mM $PO_4^{3-}$ and subsequently transferred for 3 weeks to nutrient starved medium (no nitrogen (N) or phosphate (P) source). This treatment will enhance the mycorrhizal responsiveness of the plants. In the second treatment, the plants were transferred from the nutrient-rich medium to nodulation medium, which contains high phosphate levels (0.375 mM $NH_4NO_3$ and 3 mM $PO_4^{3-}$) (see Materials & Methods). The Pearson correlation test revealed that under nodulation permissive conditions, 2.73% (873 genes) exhibited a positive correlation with *NSP2* expression, while only 0.95% (305 genes) showed a negative correlation (p-value, BH < 0.05). Conversely, under nutrient starved conditions, 1.70% (525 genes) were positively correlated with *NSP2*, and 1.74% (536 genes) were negatively correlated **(Supplementary Table S1)**. Most genes displayed a non-significant correlation with *NSP2* under either condition **(Figure 3A)**.

We performed a hypergeometric test on the overlap between the positive and negative correlated gene sets under both nutrient conditions. The results indicated a p-value of 0 for the overlap in positively correlated genes and a p-value of 1.250571e-96 for the negatively correlated gene set. These findings led us to conclude that *NSP2* consistently regulates a core set of

genes, regardless of nutrient conditions, with a subset of genes showing differential regulation attributable to the nutrient environment **(Figure 3B)**.

Studies in *M. truncatula*, rice, and barley suggest that ectopic expression of *NSP2* enhances the colonization of plant roots by arbuscular mycorrhizal fungi, potentially through the modulation of the strigolactone biosynthetic pathway (Liu *et al*., 2011; Li *et al*., 2022). We investigated the activation of the methylerythritol phosphate (MEP), carotenoid, strigolactone, zaxinone, and abscisic acid (ABA) biosynthetic pathways in roots of *Parasponia* m*NSP2ox* lines **(Figure 4A; Supplementary Table S1)**. This revealed a positive correlation between the expression of *NSP2* and all genes encoding enzymes in the carotenoid, strigolactone, and zaxinone pathways **(Figure 4B, Supplementary Table S1)**. The encoded enzymes of these genes convert geranylgeranyl pyrophosphate at the start of the carotenoid pathway into strigolactones (Okada *et al*., 2000; Ablazov *et al*., 2023). In addition, the expression levels of genes required for the biosynthesis of the apocarotenoid zaxinone correlate with the *NSP2* expression. Zaxinone may act in a positive feedforward loop towards strigolactone biosynthesis, as was shown in rice (Votta *et al.*, 2022). Taken together, it is most probable that *Parasponia NSP2* ectopic expression leads to an increased biosynthesis of strigolactones in *Parasponia* root tissue.

Next, we questioned whether the transcriptional activation of the strigolactone biosynthetic pathway also occurs in the shoot of m*NSP2ox* lines. To test this, we used real-time polymerase chain reaction (qRT-PCR) to quantify the expression of three *Parasponia* genes, *DWARF27 (PanD27), CAROTENOID CLEAVAGE DIOXYGENASE7 (PanCCD7)*, and *PanCCD8*, in shoots of lines 1, 3, and 6 under both nutrient conditions **(Supplementary Figure S2 and Supplementary Table S4)**. This revealed that *NSP2* can also activate the expression of these genes in shoot tissue.

Together, these observations lead us to conclude that enhanced *NSP2* expression in *Parasponia* markedly enhances the expression of genes

essential for strigolactone biosynthesis and the apocarotenoid zaxinone across varying nutrient conditions.

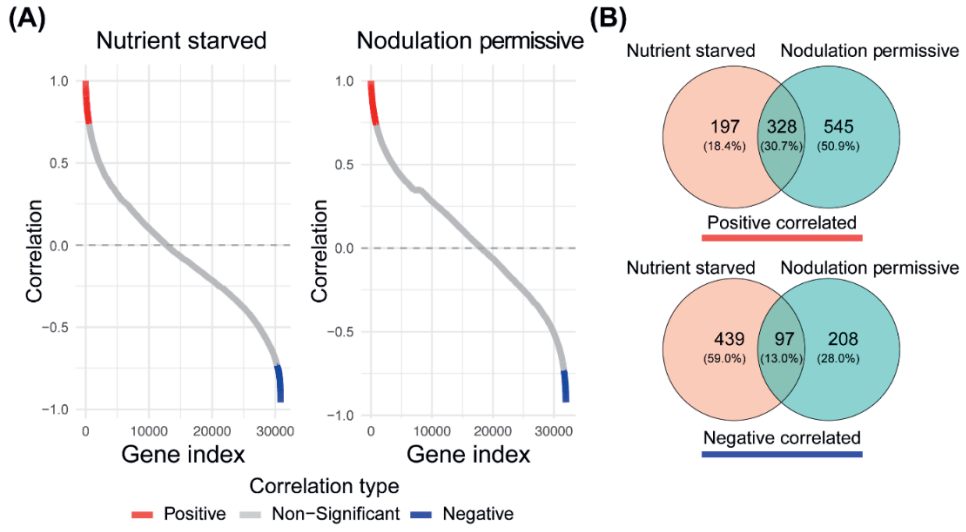



**Figure 3. Pearson correlation of all *Parasponia* genes versus *NSP2* profiles under arbuscular mycorrhizal and nodulation permissive nutrient conditions.** (**A**) Pearson correlation coefficients are sorted in ascending order. Genes that are significantly positively correlated are colored in red, while negatively correlated genes are colored in blue. (Benjamini-Hochberg adjusted p-value, BH < 0.05) (**B**) Venn diagrams illustrate the overlap between genes that positively and negatively correlate with *NSP2* expression for the two nutrient conditions.

## *NSP2* enhances the expression of a limited number of symbiosis genes among which is CYCLOPS

We also analyzed the expression dataset for symbiosis genes of which the expression correlates with m*NSP2* expression. This revealed that the expression of only a limited subset of the symbiosis-related genes positively correlated with m*NSP2* expression **(Figure 5; Supplementary Table S1)**. Among the positively correlated genes, the transcription factor *PanCYCLOPS*, the two transporters *STUNTED ARBUSCULE2* (*PanSTR2*) and *YELLOW STRIPE-LIKE 1* (*PanYSL1*), and ANNEXIN 1 (*PanANN1*) were shared between both growth conditions. Intriguingly, the genes encoding a nuclear envelope localized cation channel CASTOR and the VAMP-associated protein

VAPYRIN (*PanVPY*) show only positive correlation with *NSP2* expression under nodulation permissive conditions, whereas studies in legumes indicated they are also essential for mycorrhizal symbiosis (Imaizumi-Anraku *et al.*, 2005; Pumplin *et al.*, 2010; Murray *et al.*, 2011). Conversely, under nutrient starved conditions, a negative correlation with *NSP2* expression was observed with several transcription factor genes, including *NIN-LIKE PROTEIN1* (*PanNLP1*) and several *NUCLEAR FACTOR Y* genes: *PanNFYA5, PanNFYA6, PanNFYA7, PanNFYB6, PanNFYB8*, and *PanNFYC4* **(Supplementary Table S3)**. This indicates a complex regulatory network where enhanced *NSP2* expression differentially influences the expression of a few symbiosis-related genes depending on the nutrient status of the plant. This observation suggests a nuanced role of *NSP2* in the regulatory network governing symbiotic permissiveness.

To investigate the gene expression patterns during mycorrhization and nodulation in *Parasponia* and their interplay with *NSP2* expression, we plotted the $Log_2$fold changes of genes in mycorrhizal roots and young developing nodules **(Supplementary Table S3)** against their Pearson correlation coefficients with *PanNSP2* expression under both conditions **(Figure 6)**. This revealed that *PanSTR2* not only exhibits upregulation in mycorrhizal roots, but also is transcriptionally induced young developing *Parasponia* nodules. When focusing on the carotenoid and strigolactone biosynthetic pathways, only a single gene stood out; namely *PHYTOENE SYNTHASE1* (*PanPSY1*). *Parasponia* possesses three gene copies encoding phytoene synthases **(Supplementary Figure S4)**, of which only one correlates with *NSP2* expression. *PanPSY1* is transcriptionally enhanced in mycorrhizal roots (0.58 $Log_2$fold (1.5 fold) upregulated) and young nodules (6.47 $Log_2$fold (89.7 fold) upregulated) **(Supplementary Table S3)**. Phytoene synthase catalyzed the first step in carotenoid biosynthesis and is a major rate-limiting enzyme (Zhou *et al.*, 2022). This suggests that NSP2-controlled of *PanPSY1* expression represent a critical step in the biosynthesis of downstream products, including strigolactones in *Parasponia* roots that associate with arbuscular mycorrhizal fungi or nitrogen-fixing rhizobia.

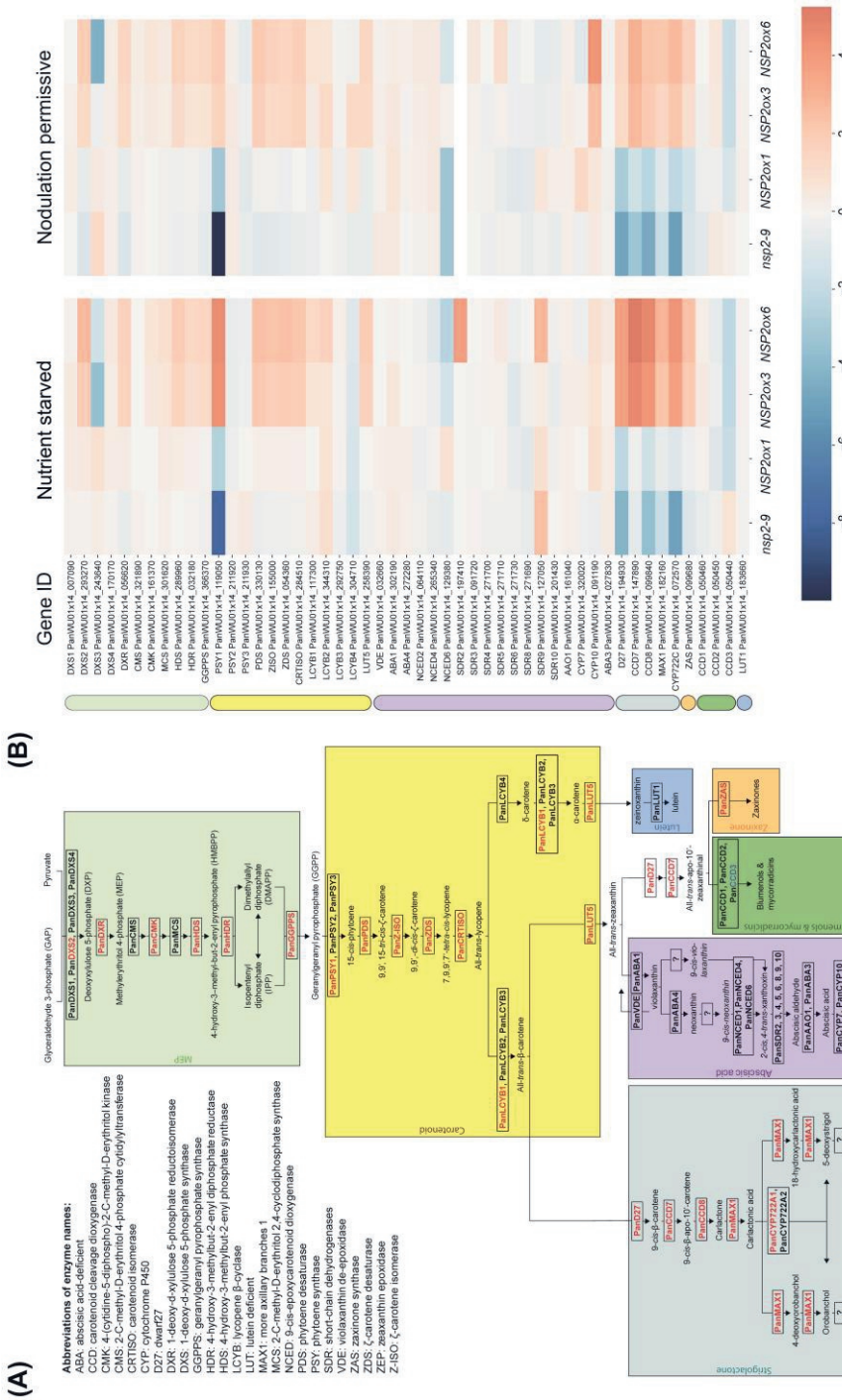

(A)
(B)
Gene ID
Nutrient starved
Nodulation permissive
nsp2-9
NSP2ox1
NSP2ox3
NSP2ox6
Log2FoldChange
Abbreviations of enzyme names:
ABA: abscisic acid-deficient
CCD: carotenoid cleavage dioxygenase
CMK: 4-(cytidine-5-diphospho)-2-C-methyl-D-erythritol kinase
CMS: 2-C-methyl-D-erythritol 4-phosphate cytidylyltransferase
CRTISO: carotenoid isomerase
CYP: cytochrome P450
D27: dwarf27
DXR: 1-deoxy-d-xylulose 5-phosphate reductoisomerase
DXS: 1-deoxy-d-xylulose 5-phosphate synthase
GGPPS: geranylgeranyl pyrophosphate synthase
HDR: 4-hydroxy-3-methylbut-2-enyl diphosphate reductase
HDS: 4-hydroxy-3-methylbut-2-enyl phosphate synthase
LCYB: lycopene β-cyclase
LUT: lutein deficient
MAX1: more axillary branches 1
MCS: 2-C-methyl-D-erythritol 2,4-cyclodiphosphate synthase
NCED: 9-cis-epoxycarotenoid dioxygenase
PDS: phytoene desaturase
PSY: phytoene synthase
SDR: short-chain dehydrogenases
VDE: violaxanthin de-epoxidase
ZAS: zaxinone synthase
ZDS: ζ-carotene desaturase
ZEP: zeaxanthin epoxidase
Z-ISO: ζ-carotene isomerase
Glyceraldehyde 3-phosphate (GAP)
Pyruvate
MEP
Carotenoid
Abscisic acid
Strigolactone
Lutein
Zaxinone
Blumenols & mycorradicins

**Figure 4. Schematic representation of the carotenoid biosynthetic pathway in *Parasponia* and heatmap of differential gene expression in root tissue in *Parasponia* under nutrient starved mycorrhiza and nodulation permissive conditions.** (**A**) *Parasponia* genes found to be positively correlated with *NSP2* over-expression level are colored in red, while non correlating genes are colored in black. The MEP biosynthetic pathway is shaded with light green, the carotenoid biosynthetic pathway is shaded with light yellow, the strigolactones biosynthetic pathway is shaded light cyan, the ABA biosynthesis pathway is shaded with light magenta, the zaxinone biosynthesis is shaded light orange, the blumenol and mycorradicins biosynthesis is shaded green, and the Lutein biosynthesis is shaded light cornflower blue. Enzyme names are based on annotations in *Arabidopsis thaliana* and *Parasponia andersonii*. Biosynthetic pathway is based on Wakabayashi *et al*., 2019; Wakabayashi *et al*., 2020; and Li *et al*., 2022. (**B**) Heatmap of gene expression in the carotenoid pathway across different lines: *Pannsp2*-9 knockout mutant, m*NSP2*ox1, m*NSP2*ox3, and m*NSP2*ox6 ectopic expression lines. Expression levels are indicated by color intensity, with the scale representing $\log_2$fold changes. Each row represents a gene from the carotenoid pathway.

## *Parasponia NSP2* ectopic expression induces epidermal cell proliferation upon rhizobium inoculation

Given that *NSP2* is critical for both mycorrhization and nodulation, we investigated whether overexpression of m*NSP2* could enhance symbiotic interactions with rhizobia. To test that, we grew all six m*NSP2*ox lines on nodulation medium (0.375 mM $NH_4NO^3$, 3 mM $PO_4^{3-}$) and scored the number of nodules at 8 weeks post inoculation with *Mesorhizobium plurifarium* BOR2. We observed fewer nodules on lines with the highest m*NSP2* expression **(Figure 7A)**. We examined the cytoarchitecture of these nodules, revealing a wild-type phenotype where infection threads and fixation threads were formed **(Figures 7B-D)**. A surprising observation was the occurrence of abnormal cell divisions on roots of plants highly expressing m*NSP2*. To determine if this cell proliferation was spontaneous due to the level of *PanNSP2* expression or a response to rhizobium application, we grew plantlets of three *PanNSP2* over-expressor lines with fold change in gene expression of 6, 51, and 95 on nodulation medium either non-inoculated or inoculated with *M. plurifarium* BOR2. Under the non-inoculated condition, all m*NSP2ox* lines showed no such epidermal cell divisions, whereas under inoculated conditions, m*NSP2ox* lines with 51 and

95-fold increase in *mNSP2* expression, but not 6-fold, exhibited foci of massive cell divisions **(Figures 7E-G)**. This shows that high levels of *NSP2*

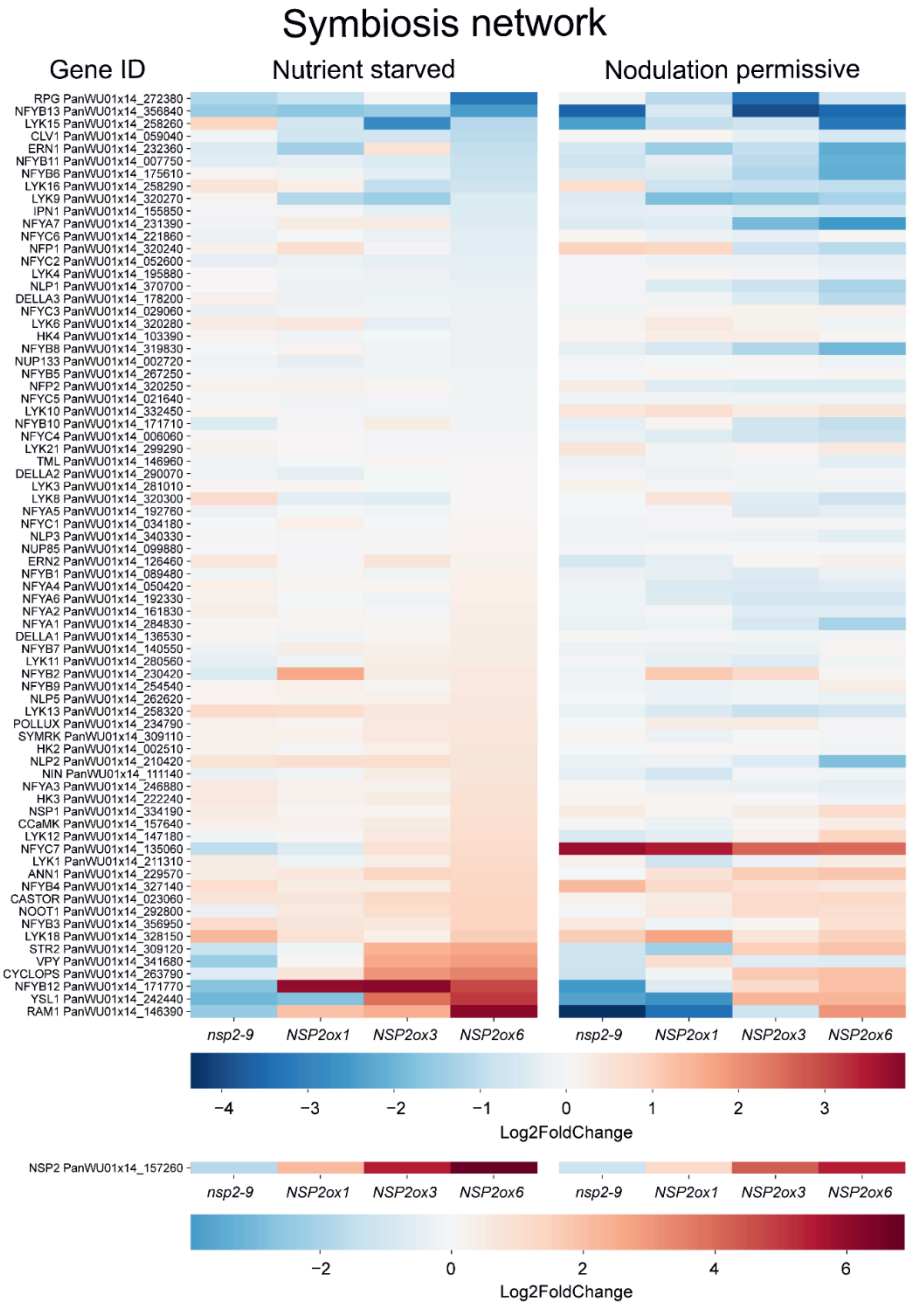


**Figure 5. Heatmap of genes involved in the symbiosis network across different lines**: *Pannsp2*-9 knockout mutant, m*NSP2*ox1, m*NSP2*ox3, and m*NSP2*ox6 ectopic expression lines. Expression levels are indicated by color intensity, with the scale representing $Log_2$fold changes. Each row represents a gene from the symbiosis pathway.

promote cell proliferation in response to rhizobium inoculation. Sections of these aberrant structures revealed mitotic activity in epidermal, cortical and endodermal cell layers, **(Figures 7H and I)**. We concluded that increased *NSP2* expression in *Parasponia* negatively regulates nodule formation and causes abnormal proliferation of root tissue in response to rhizobium inoculation.

## Discussion

The GRAS-type transcriptional regulator NSP2 serves as a conserved hub that regulates the expression of symbiotic genes and genes essential for strigolactone biosynthesis. As such, NSP2 performs a dual function. The gene is essential for rhizobium-induced root nodule formation in legumes and the Cannabaceae species *Parasponia* (Kaló *et al*., 2005; Heckmann *et al*., 2006; van Zeijl *et al*., 2018; Peng *et al*., 2021). Additionally, NSP2-controlled strigolactones are exuded into the rhizosphere, serving as stimulants for arbuscular mycorrhizal fungi (Liu *et al.*, 2011). Remarkably, ectopic expression of *NSP2* in barley and rice has been shown to increase arbuscular mycorrhizal colonization rates under various nutrient conditions, while the reported trade-offs in plant developmental traits are relatively minor (Li *et al*., 2022; Yuan *et al*., 2023). This opens up opportunities to investigate the effects of enhanced *NSP2* expression on plant development and endosymbioses in different species (Isidra-Arellano, Singh and Valdés-López, 2024). We focused on *Parasponia*, the only non-legume known to form nitrogen-fixing root nodules with rhizobium and revealed novel phenotypes associated with increased *NSP2* expression.

We generated six independent *Parasponia NSP2 ectopic* expression lines that differ in *NSP2* over-expression from 6 to 95-fold compared to the root tissue of control plants. These transgenic lines allowed us to correlate

phenotypic responses to *NSP2* expression levels. This revealed that the mycorrhizal infection levels positively correlate with *NSP2* expression.

We also explored the ectopic expression lines to identify genes of which the expression correlates positively with *NSP2* expression. Thereby we focused on two gene sets; the pathways that lead to strigolactone biosynthesis and a set of (putative) symbiosis genes. In the case of symbiosis genes, only a few genes showed a correlation with *NSP2* expression **(Supplementary Table S1)**. Of these *PanCYCLOPS* is most relevant. CYCLOPS is a transcription factor that functions in the common symbiosis signaling pathway as the last shared hub of arbuscular mycorrhizal- and rhizobium-induced signaling (Capoen and Oldroyd, 2008). Studies in a range of plant species showed that CYCLOPS is critical for both symbioses (Chen, Ané and Zhu, 2008; Yano *et al.*, 2008; Horváth *et al.*, 2011; Das *et al.*, 2019; Jin *et al.*, 2018; Prihatna *et al.*, 2018). CYCLOPS and NSP2 can form a complex having DELLA as an intermediate (Jin *et al.*, 2016). Studies in *M. truncatula* showed that the CYCLOPS orthologs INTERACTING PROTEIN OF DMI3 (*MtIPD3*) and IPD3-LIKE (*MtIPD3L*) are critical for mycorrhizal-induced signaling when plants are grown at higher exogenous phosphate levels (Lindsay *et al.*, 2019). Furthermore, constitutive expression of auto active variants of *MtIPD3* and *MtIPD3L* induces the expression of several arbuscular mycorrhizal symbiosis genes, including *MtVPY* (Lindsay *et al*., 2019; Lindsay *et al*., 2022). We found *PanVPY* to be associated with m*NSP2ox* expression in *Parasponia* when grown at nodulation permissive conditions, which contains a relatively high phosphate concentration, but not under nutrient starved (mycorrhizal permissive) conditions. This led us to conclude that *NSP2*-controlled *PanVPY* expression is phosphate-dependent in *Parasponia*, a response that may require CYCLOPS-DELLA-NSP2 complex formation.

Analyzing the strigolactone biosynthetic pathways revealed that the expression of all genes encoding essential enzymes for the biosynthesis of carotenoids, strigolactones, and zaxinone positively correlates with *NSP2* expression. In contrast, side branches leading to abscisic acid and lutein do

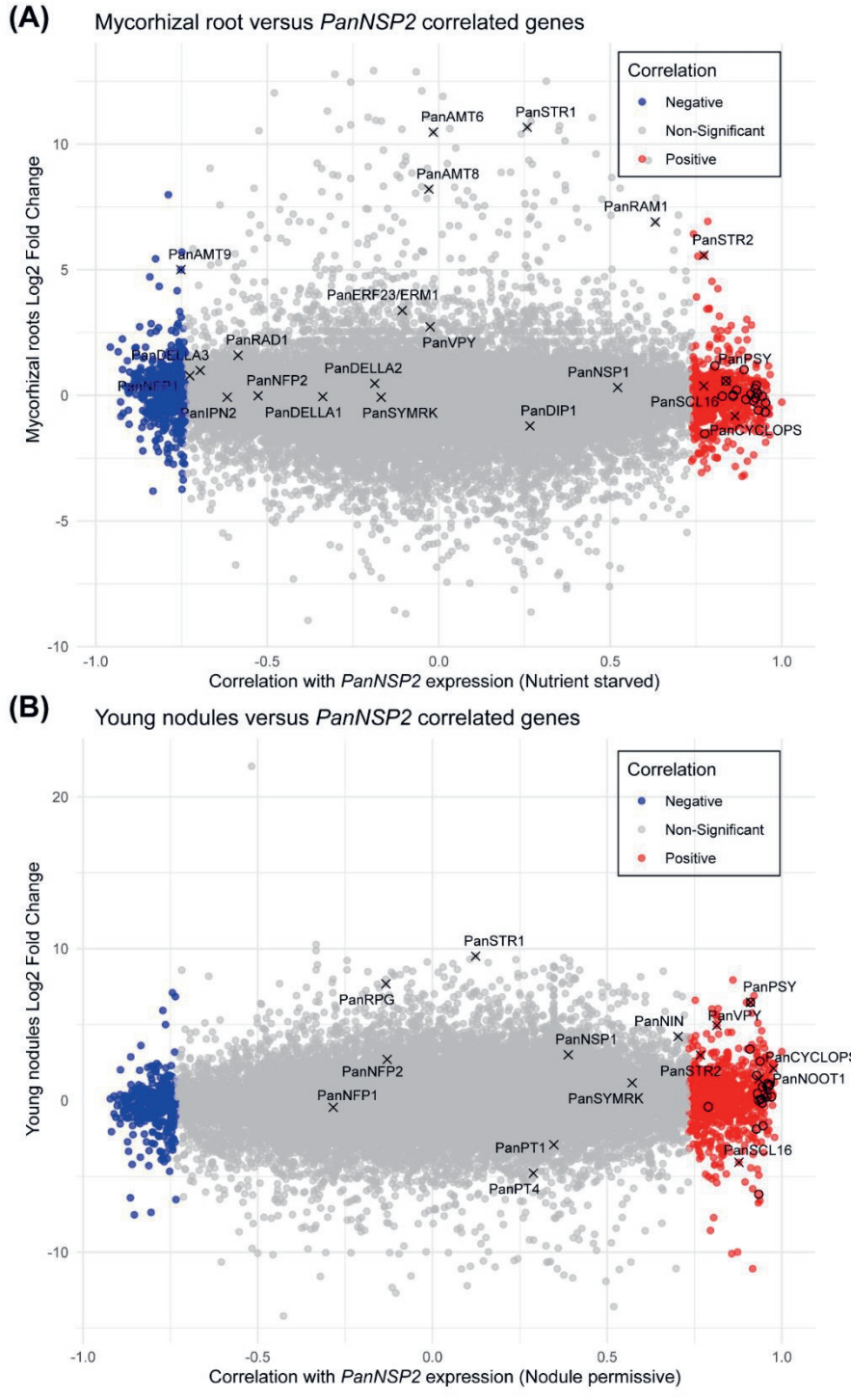

(A)
Mycorhizal root versus PanNSP2 correlated genes
Correlation
Negative
Non-Significant
Positive
PanAMT6
PanSTR1
PanAMT8
PanRAM1
PanSTR2
PanAMT9
PanERF23/ERM1
PanVPY
PanRAD1
PanDELLA3
PanDELLA2
PanNFP2
PanNSP1
PanPSY
PanIPN2
PanDELLA1
PanSYMRK
PanDIP1
PanSCL16
PanCYCLOPS
Mycorhizal roots Log2 Fold Change
Correlation with PanNSP2 expression (Nutrient starved)
(B)
Young nodules versus PanNSP2 correlated genes
Correlation
Negative
Non-Significant
Positive
PanSTR1
PanRPG
PanPSY
PanNIN
PanVPY
PanNSP1
PanCYCLOPS
PanNFP2
PanSTR2
PanNOOT1
PanNFP1
PanSYMRK
PanPT1
PanSCL16
PanPT4
Young nodules Log2 Fold Change
Correlation with PanNSP2 expression (Nodule permissive)

**Figure 6. Scatterplot displaying the expression correlation between *PanNSP2* and mycorrhizal-induced and nodulation-induced genes.** (**A**) Scatterplot representing the $Log_2$fold change in gene expression of wild type mycorrhizal *Parasponia* roots against the corresponding Pearson correlation **(Supplementary Table S1)** with *mNSP2* overexpression under nutrient starved conditions. (**B**) Scatterplot representing the $Log_2$fold change of gene expression in young wild type *Parasponia* nodules against the corresponding Pearson correlation **(Supplementary Table S1)** with *mNSP2* expression in nodulation permissive conditions. Each point indicates a *Parasponia* gene; blue points denote genes negatively correlated with *mNSP2* expression, red points represent genes with a positive correlation, and gray points signify non-significant correlations. Several marker genes of mycorrhization and/or nodulation are indicated. Genes representing the carotenoid pathway are indicated by a circle.

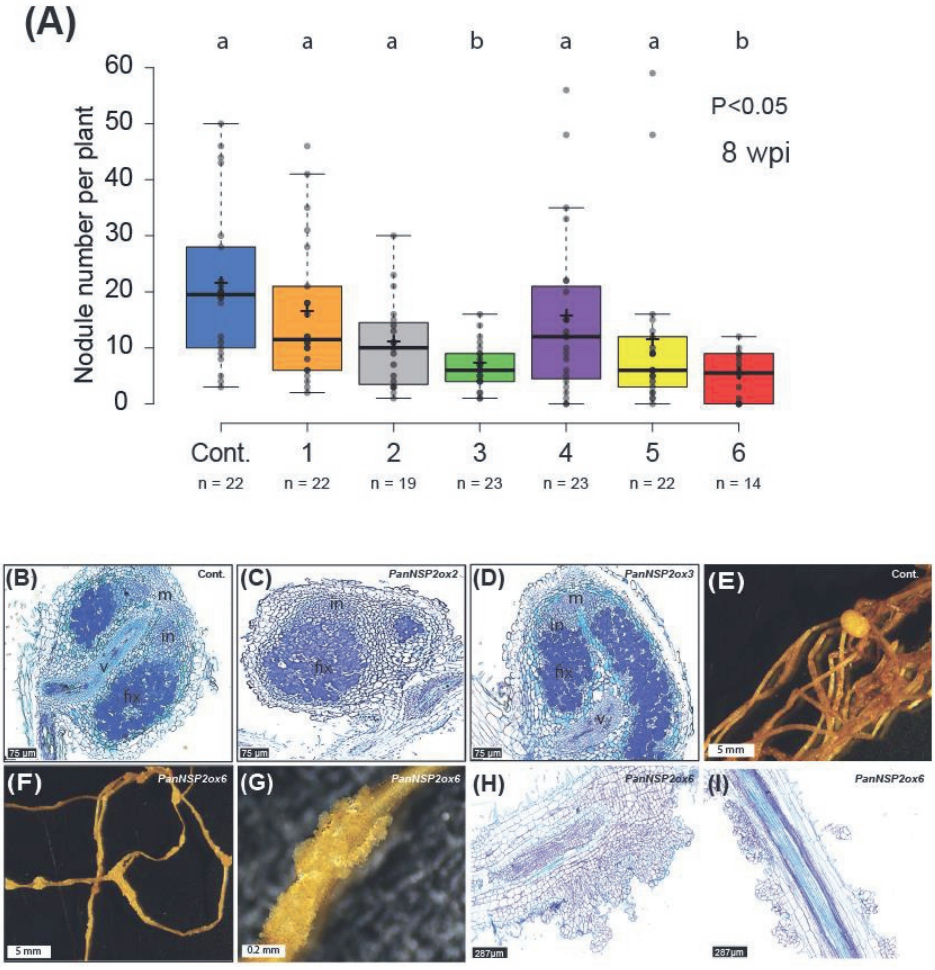


**Figure 7: Rhizobium-induced nodulation on *Parasponia NSP2*ox lines.** (**A**) Nodule number per plant of empty vector control (Cont.) and *mNSP2ox* lines 1 to 6 8 weeks post-inoculation (wpi) with *M. plurifarium* BOR2. Different letters indicate significant differences ($p<0.05$) between these lines as determined by One-way ANOVA followed by a Tukey's post hoc test. Data is displayed in a box plot, showing data points, the median, and the interquartile range (IQR) (**B-D**) Representative image of a section through a mature nodule formed on an empty vector control line (cont.) (**B**), *mNSP2ox* lines 2 (**C**), and m*NSP2*ox lines 3 ( **D**) (8 wpi). Scale bar is 75 µm. m, nodule meristem; in, infection zone; fix, fixation zone; v, nodule vasculature. (**E-G**) Bright-field image of nodulated root of the empty vector control line (cont.) (**E**), and the root of *mNSP2ox* line 6 showing rhizobium-induced irregular cell divisions (**F, G**) (10 wpi). Scale bar in (**E**) and (**F**) is 5 mm and in (**G**) is 0.2 mm. (**H, I**) Longitudinal sections of rhizobium-induced irregular cell divisions on the root of m*NSP2*ox line 6 (10 wpi). Scale bar is 287µm.

not show such a response. A critical gene in the carotenoid pathway, namely *PanPSY1*, is also enhanced upon rhizobium and mycorrhizal inoculation. PSY commits the first step in the carotenoid pathway converting geranylgeranyl pyrophosphate into 15-cis-phytoene and is generally considered the rate-limiting step in carotenoid biosynthesis (Zhou *et al.*, 2022). PSY functioning is controlled by several mechanisms, in which transcriptional regulation plays an important role (Stanley and Yuan, 2019). Enhancing *PSY* expression has been utilized extensively in plant biotech approaches to enhance carotenoid biosynthesis (Zhou *et al.*, 2022). We hypothesize that in *Parasponia* the symbiotic regulation of *PanPSY1* in response to arbuscular mycorrhizal infection or rhizobium-induced nodulation is a critical factor ultimately in the biosynthesis of strigolactones.

High expression levels of *NSP2* also cause pleiotropic phenotypes. Some of these, like reduced axillary bud outgrowth, increased secondary thickening of the stem, and reduced lateral root formation are known strigolactone - controlled responses (Waters *et al*., 2017). In addition, we observed irregular cell divisions upon rhizobium-inoculation in roots of some *Parasponia* m*NSP2ox* lines. The structures formed resemble enlarged pre-nodules, which can serve as infection pockets to facilitate rhizobial crack entry (Lancelle and Torrey, 1984). A function of NSP2 in the formation of rhizobium infection pockets is also known in the legume *Aeschynomene*

*avenia*. Like in *Parasponia*, rhizobium penetrates *A. avenia* by crack entry. Under nodulation permissive conditions, but in absence of rhizobium, *A. avenia* forms clusters of multicellular auxiliary hair-like cells at the base of newly formed lateral roots. These structures can be explored by rhizobium as an infection pocket that guides bacterial entry (Bonaldi *et al.*, 2011). *A. avenia nsp2* mutants do not form such infection pockets (Quilbé *et al.*, 2022). Together, these experiments suggest the importance of a balanced *NSP2* expression in the formation of pockets that ultimately can be explored by rhizobium as a starting point of intracellular infection.

Ectopic expression of *NSP2* also affects plant architecture. m*NSP2ox Parasponia* trees have a reduced number of lateral roots but more lateral shoot branches. Studies in rice revealed similar developmental phenotypes, where *OsNSP2ox* results in increased tiller formation and a reduced number of lateral roots (Liu *et al.*, 2011; Yuan *et al.*, 2023). These phenotypes question whether exploring *NSP2* ectopic expression in a biotech approach aiming to improve crop yield in low phosphate environments is a realistic strategy (Isidra-Arellano, Singh and Valdés-López, 2024). Our findings advocate using tissue-specific promoters to obtain enhanced mycorrhization while minimizing pleiotropic phenotypes in plant architecture. Studies in rice indicate that using native promoter elements to enhance expression of GRAS-type transcriptional regulators might be the way forward to improve agronomic traits under low and medium-phosphorus conditions (Yuan *et al.*, 2023).



## Materials and methods

### Plant materials and growth conditions

The tissue culture propagation and maintenance of *Parasponia* plants were done on agar plates supplemented with Parasponia propagation medium. This medium contains 3.2 g/L SH-basal salt, 1 g/L SH-vitamin mixture, 10 g/L sucrose, 1 mL/L BAP (1 mg/mL), 100 µL/L IBA (1 mg/mL), 3 mL/L MES (1 M, pH 5.8), and 8 g/L Daishin agar (van Zeijl *et al.*, 2018; Wardhani *et al.*, 2019).

Performance of nodulation and mycorrhization assays was carried out on platelets vegetatively grown and rooted following previously published procedures (van Zeijl *et al.*, 2018; Wardhani *et al.*, 2019).

**Nodulation assay and analysis**

Rooted plantlets were grown in 1 liter crystal-clear polypropylene pots (OS14BOX, Duchefa Biochemie, Netherlands) half-full, with a mixture (1:2 ratio) of river sand and agri perlite (Massmond-Westland, Netherlands). The mixture was watered with modified EKM medium [15 µM Fe-Citrate, 3 mM MES (pH 6.6), 6.6 µM $MnSO_4$, 4.1 µM $Na_2MoO_4$, 2.08 mM $MgSO_4$, 0.70 mM $Na_2SO_4$, 1.4 mM $CaCl_2$, 0.375 mM $NH_4NO_3$, 1.5 µM $ZnSO_4$, 0.88 mM $KH_2PO_4$, 1.6 µM $CuSO_4$, 2.07 mM $K_2HPO_4$, and 4 µM $H_3BO_3$] (Becking, 1983) and inoculated with *Mesorhizobium plurifarium* BOR2 (OD600 = 0.025). Inoculated plants were then incubated in a controlled climate room at 85% humidity, under a 16 /8 h day/night growth condition for eight weeks. Plants were removed from pots and washed under running tap water followed by quantification of total nodule number and harvest of some nodules for sectioning. Subsequently, harvested nodules were fixed in 0.1 M phosphate buffer (pH 7.2) containing 0.5% glutaraldehyde followed by vacuum application for 60 minutes. Next, nodules were embedded in infiltration plastic, Technovit 7100 (Heraeus-Kulzer, Germany), following the protocol provided by the manufacturer. A Leica RJ2035 microtome was used to make 4.5 µm plastic sections which were stained with 0.5% toluidine blue. A Leica DFC425c camera was used to make high resolution images of the cytoarchitecture of the fixed nodules.

**Mycorrhization assays and ink staining**

Rooted plantlets were placed in pots half-filled with a mixture (5:1 ratio) of river sand and commercial potting soil and watered with half-strength Hoagland's medium containing 20 µM phosphate (low P condition) or 3 mM (high P condition). Plants then were incubated in a controlled climate room

at 85% humidity, under a 16 /8 h day/night regime for two weeks followed by inoculation with 300 spores of *Rhizophagus irregularis* fungi (Agronutrion-DAOM197198, Carbonne, France) and subsequent incubation for 6 weeks. Afterwards, plants' roots were harvested and treated with 10% (w/v) KOH and boiled for 30 minutes at 90 °C and then stained in a staining solution contains 5% acetic acid and 3% ink (Waterman Vulpen, Zwolle, the Netherland) for 15 minutes at 90 ℃. Next, fungal colonization and formation of arbuscules and vesicles were quantified, using the gridline intersect method (Giovannetti and Mosse, 1980), and high-resolution images were taken, using a Leica CTR66000 microscope.

**Vector constructs**

All vectors used in this study were propagated in *Escherichia coli* strain DH5α and assembled, using Golden Gate Cloning (GGC) as was previously described (Engler *et al.*, 2009). Level one and two acceptors and binary vectors used for the assembly were obtained from the GGC toolbox (Engler *et al.*, 2014). *Parasponia* clone of *NSP2* coding sequence (CDS) with and without insertion of intron 10 of *Arabidopsis thaliana UBQ10* gene (intron) in the first putative splicing site in *NSP2* CDS, including silent mutation in microR171h putative site and golden gate BsaI or BpiI restriction sites was synthesized as level zero vectors. Subsequently, level zero vectors of *NSP2*, *Lotus japonicus UBQ1* (*LjUBQ1*) promoter, C-terminal Myc tag, and *Agrobacterium tumefaciens* nopaline synthase terminator (T-NOS) were recombined into level one acceptor. Finally, level one vectors of *NSP2* with or without (intron), including Kanamycin resistant level one vectors were recombined into Level 2 binary vectors. All binary GGC vectors were validated by restriction enzymes and sequencing before transformation.

**Plant transformation and Genotyping**

Level 2 binary vectors were transformed into *Agrobacterium tumefaciens* AGL1 strain, using Eppendorf Eporator (Eppendorf SE, Hamburg Germany).

Transformed AGL1 were then grown on lysogeny broth (LB) agar plate supplemented with appropriate antibiotics and incubated at 28 ℃ for 2 days. *Parasponia* plant transformation was carried out as previously described (van Zeijl *et al.*, 2018; Wardhani *et al.*, 2019), so young stems and petioles were collected from *Parasponia* trees grown in controlled greenhouse conditions at 85% humidity and under a 16 /8 h day/night regime. The surface of the stems and petioles was sterilized with 2.5% bleach with a few drops of Tween 20, and subsequently stems and petioles washed six times with sterile water followed by a co-cultivation with AGL1 strains carrying constructs of interest for 2 days. Afterwards, transformed stems and petioles were transferred to callus induction medium supplemented with cefotaxime (300µg/ml) and kanamycin (50µg/ml) antibiotics for one week followed by transferring to plant propagation medium supplemented with cefotaxime (300µg/ml) and kanamycin (50µg/ml) antibiotics, refreshed every 1-2 weeks until transgenic shoots were formed. Genotyping was performed using a pair of primers designed to anneal to the *Lotus japonicus UBIQUITIN 1 PROMOTER* (*pLjUBQ1*) and the C-terminal Myc tag, with amplification carried out using the Phire Plant Direct PCR kit (Thermo Fisher, F130WH), following the manufacturer's protocol.

**Plant growth conditions for RNA-sequencing**

*Parasponia* plant lines, consisting of three NSP2 overexpression mutants (*mNSP2ox1*, *mNSP2ox3*, and *mNSP2ox6*), a *nsp2* knockout (*nsp2-9*), and control line 46 with an empty vector were selected. The *mNSP2* overexpression mutants displayed *mNSP2* expression levels of 6, 51, and 95-fold higher than empty vector control. Selected plants were cultivated for 6 weeks on nutrient rich medium containing 24.72 mM $NO_3^-$, 2.6 mM $NH_4^+$ and 2.6 mM $PO_4^{3-}$. Next, cultivated plants were transferred for 3 weeks to either nutrient starved medium (no N and P resource), or the nodulation permissive medium (0.375 mM $NH_4NO_3$ and 3 mM $PO_4^{3-}$).

Subsequently, healthy plantlets were transferred to pots containing a 2:1 ratio of perlite to sand. Each pot was supplemented with 150 mL of EKM medium, with or without nitrogen and phosphorus, referred to as nodulation permissive (low nitrate, high phosphate) and nutrient starved (low nitrate and low phosphate) treatments, respectively. Pots were populated with three plantlets of a single genotype and grown under a 16-hour light/8-hour dark photoperiod for 21 days. Following this period, roots from each pot were pooled for RNA extraction.

**RNA extraction and quality control**

For each treatment, five root samples weighing 50 mg each were collected. Tissue lysis was performed using a QIAGEN TissueLyser LT, followed by sequential phenol:chloroform and chloroform extractions. Nucleic acid precipitation was performed using sodium acetate, followed by sequential washing steps with isopropanol and ethanol to purify the samples. Subsequently, the samples were treated with DNase to eliminate genomic DNA contamination. A second round of phenol:chloroform and chloroform extractions was conducted. The RNA was then re-precipitated in ethanol and re-suspended in nuclease-free water. The concentration of the isolated RNA was quantified using both a Biochrom SimpliNano™ Spectrophotometer and a Qubit 2.0 Fluorometer, utilizing Broad Range Assay Kits. RNA integrity was confirmed through gel electrophoresis. To validate NSP2 expression levels, quantitative polymerase chain reaction (qPCR) was employed. *ACTIN* and *Elongation Factor1 Alpha* (*EF1α*) served as the housekeeping genes for normalization.



**RNA-seq library preparation and high-throughput sequencing**

For each harvested sample, mRNA was isolated from total RNA utilizing poly-T oligo-attached magnetic beads. The resulting mRNA served as the template for first-strand cDNA synthesis, which was initiated with random hexamer primers, and subsequently followed by second-strand cDNA synthesis. Sample quality and nucleic acid concentration were quantitatively

assessed using Qubit and qPCR assays, using *EF1α* and *ACTIN* for normalization genes. Size distribution of the samples was further confirmed via bioanalyzer analysis. Three biological replicates for each sample exhibiting the best characteristics (quantity and quality) for sequencing were selected for library preparation and sequencing, conducted at Novogene UK (Cambridge, United Kingdom). The resulting libraries were pooled and subjected to paired-end sequencing on an Illumina NovaSeq 6000 platform. Each biological replicate generated approximately 25 million reads, with each read consisting of 150 base pairs.

The quality of the sequenced reads was assessed using FastQC and subsequently visually inspected through MultiQC. Remaining adapters and low-quality bases were trimmed using Trimmomatic with settings TRAILING:3, SLIDINGWINDOW:4:15, MINLEN:50, HEADCROP:12. Transcript abundances were quantified with Kallisto, and the resulting files were loaded into R Studio version 2022.07.2 using R version 4.1.1. Differential gene expression analysis was conducted using DESeq2 for each condition (Nodulation permissive and nutrient starved) separately **(Supplementary Table S2)**.

## Correlation Analysis between NSP2 Overexpression and Gene Expression

Pearson correlation analysis was performed for each nutrient condition (nodulation permissive and nutrient starved) to assess the expression correlation with *NSP2*. The analysis was conducted in R, calculating the correlation, p-value, and False Discovery Rate (FDR, or adjusted p-value). An FDR cutoff of 0.05 was applied and used to categorize the correlation as positive, non-significant, or negative. The Pearson correlation graphs were plotted using the ggplot2 package. Overlapping genes between nutrient conditions were visualized using the VennDiagram package in R.

Prominent pathways featuring upregulated genes under m*NSP2* overexpression conditions, as well as genes of interest related to nodulation and mycorrhization in *Parasponia*, were selected for detailed investigation.

The first set, named the "Symbiosis Network," consists of a gene set in *Parasponia* previously annotated **(Supplementary Table S1)**. The second set includes genes involved in the carotenoid biosynthesis pathway in *Parasponia*. Orthologous genes in *Parasponia* involved in carotenoid biosynthesis were identified using gene IDs from *Arabidopsis thaliana*, as reported in the study by (Li *et al*., 2022) **(Supplementary Table S1)**. Corresponding gene IDs for *Parasponia* were identified through a reciprocal best BLAST hit approach, using the *Arabidopsis* gene IDs as references. BLASTP was performed with the following parameters: E-value cutoff of 1e-5, BLOSUM62 scoring matrix, gap penalties of 11 for existence and 1 for extension, and a word size of 3. The top BLAST hits in *Parasponia* were then used for a reverse search against the TAIR Araport 11 protein set to verify orthologous relationships.

### Data analysis on mycorrhization and Young nodules data



*Parasponia* wild type young nodule data were previously published under PRJNA272473. Additionally, an experiment was conducted on *Parasponia* plants inoculated with *Rhizophagus irregularis* (MYC+) or left uninoculated (MYC-) (van Velzen *et al.*, 2018; van Zeijl *et al.*, 2018; Bu *et al.*, 2020; Rutten *et al.*, 2020; Alhusayni *et al.*, 2023). *Parasponia* plants were grown under controlled greenhouse conditions with a 16-hour light/8-hour dark photoperiod at 28°C and 85% humidity. The growth medium consisted of a mixture (5:1 ratio) of river sand and commercial potting soil and watered with half-strength Hoagland solution containing 20 µM $PO_4$ every other week. After placing the plants in pots, they were allowed to recover for 2 weeks before being inoculated with 300 *R. irregularis* spores. Root samples were harvested 42 days after inoculation.

Total RNA was extracted from collected samples using the RNeasy Plant Mini Kit (Qiagen). mRNA was isolated from total RNA using poly-T oligo-attached magnetic beads. The isolated mRNA served as the template for first-strand cDNA synthesis, initiated with random hexamer primers, followed by second-strand cDNA synthesis. Sample quality and nucleic acid

concentration were assessed using Qubit and qPCR assays. Size distribution of the samples was confirmed via bioanalyzer analysis. Three biological replicates per sample exhibiting the best characteristics (quantity and quality) were selected for library preparation and sequencing.

Sequencing was conducted at Novogene UK (Cambridge, United Kingdom) using an Illumina NovaSeq 6000 platform. Each biological replicate generated approximately 25 million paired end reads of 150 base pairs. The quality of the sequenced reads was assessed using FastQC and subsequently visually inspected through MultiQC.

To analyze the RNA-seq dataset from nodule stages 1, 2, and 3 versus roots, as well as mycorrhization data of *Parasponia* wild type MYC- and MYC+ data, the following procedures were performed using the R programming language. Transcript abundances were quantified using Kallisto. The resulting files were imported into R Studio version 2022.07.2 with R version 4.1.1 for subsequent analysis using DESeq2 as previously described **(Supplementary Table S2)**. The differential expression results were integrated with existing annotated datasets for nodulation permissive and nutrient starved conditions. Ggplot2 was used to plot $Log_2$fold change versus the NSP2 correlation **(Supplementary Table S3)**.

**Statistical analysis**

Box plot graphs were generated using BoxPlotR, an online tool developed by (Spitzer *et al*., 2014). Bar charts were created in Microsoft Excel for Mac (Version 16.89). The Kruskal–Wallis test, followed by Tukey's post-hoc test, was employed for statistical analysis. Statistical significance was established at a threshold of $p < 0.05$. Additionally, the quantification of internode diameter was assessed using a Student's t-test, with statistical significance also defined as $p < 0.05$.

**Data availability**

The RNA-seq datasets analyzed in this study are available in the NCBI SRA repository under BioProject number PRJNA1029573. Plant material and seeds used in this study can be obtained upon request from the authors. Scripts used in the differential gene expression and visualization of the heatmaps can be found on GitHub: https://github.com/kleinjoel/NSP2_overexpression_analysis. Young nodules data were previously published under BioProject number PRJNA272473.

**Author Contributions**

Data presented in Fig. 1: SA; Fig. 2: SA; Fig. 3: JK Fig. 4: SA, Jk, NK; Fig. 5: JK; Fig. 6: JK; Fig 7 SA; Fig S1: SA; Fig. S2: SA; Fig. S3: SA; Fig. S4: RG, JK; Supplementary Table S1: JK; Supplementary Table S2: JK; Supplementary Table S3: JK; Supplementary Table S4: SA; Manuscript preparation: RG, JK, SA. All authors reviewed the manuscript. The authors read and approved the final manuscript.

**Conflicts of interests**

The authors have carefully reviewed the content of this manuscript and declare that they have no financial, personal, or professional interests that could be construed as a conflict of interest regarding its publication.

**Funding**

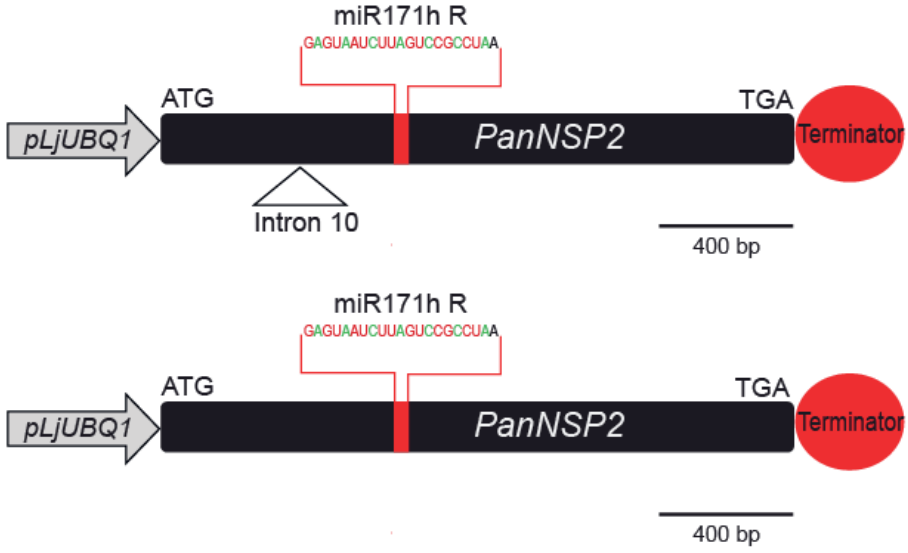
The project was funded by the ENSA project funded by the Bill & Melinda Gates Foundation to the University of Cambridge (to RG) [OPP1172165], and the Ministry of Education, King Faisal University, Saudi Arabia (grant no. 10598 to SA), RH is funded by Dutch Science Organization (Nederlandse Organisatie voor Wetenschappelijk Onderzoek VI.Veni.212.132), and JK by UK Research and Innovation (UKRI) under the UK government's MCSA Postdoctoral Fellowship funding guarantee [grant number EP/X023672/1].

## Supplementary information

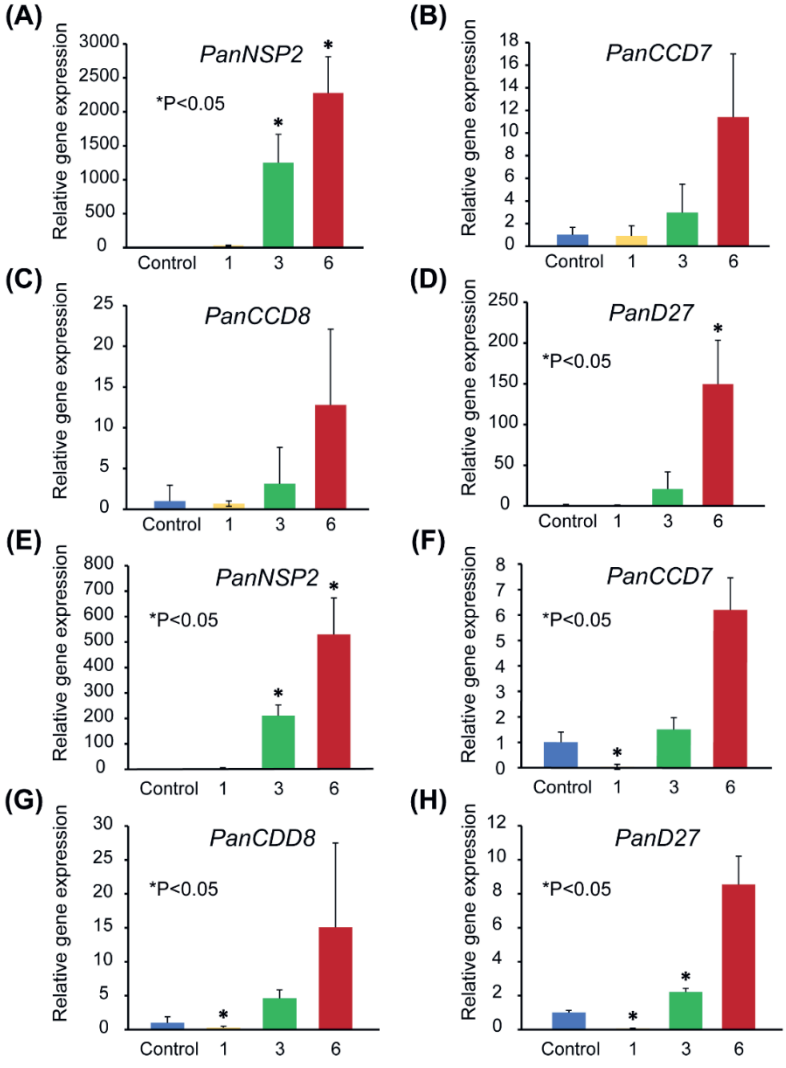
**(A)**

| | |
|---|---|
| miR171h binding site | GAGUGAUAUUGAUUCGGCUCG |
| *MtNSP2* - putative miR171h binding site | GAGUGAUAUUGAUCCGGCUCA |
| *PanNSP2* - putative miR171h binding site | GGGUGAUAUUGGUUCGGCUCA |
| *mNSP2* miR171h (silenced) | GAGUAAUCUUAGUCCGCCUAA |

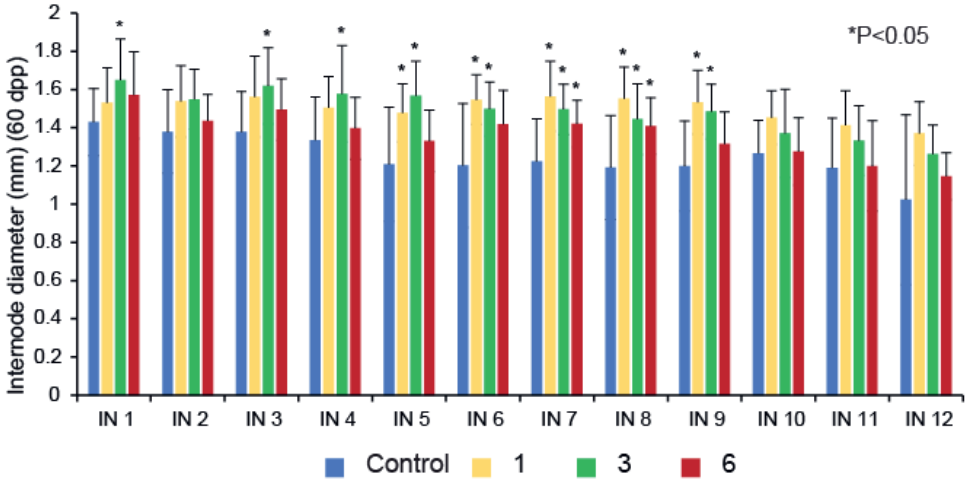
**(B)**



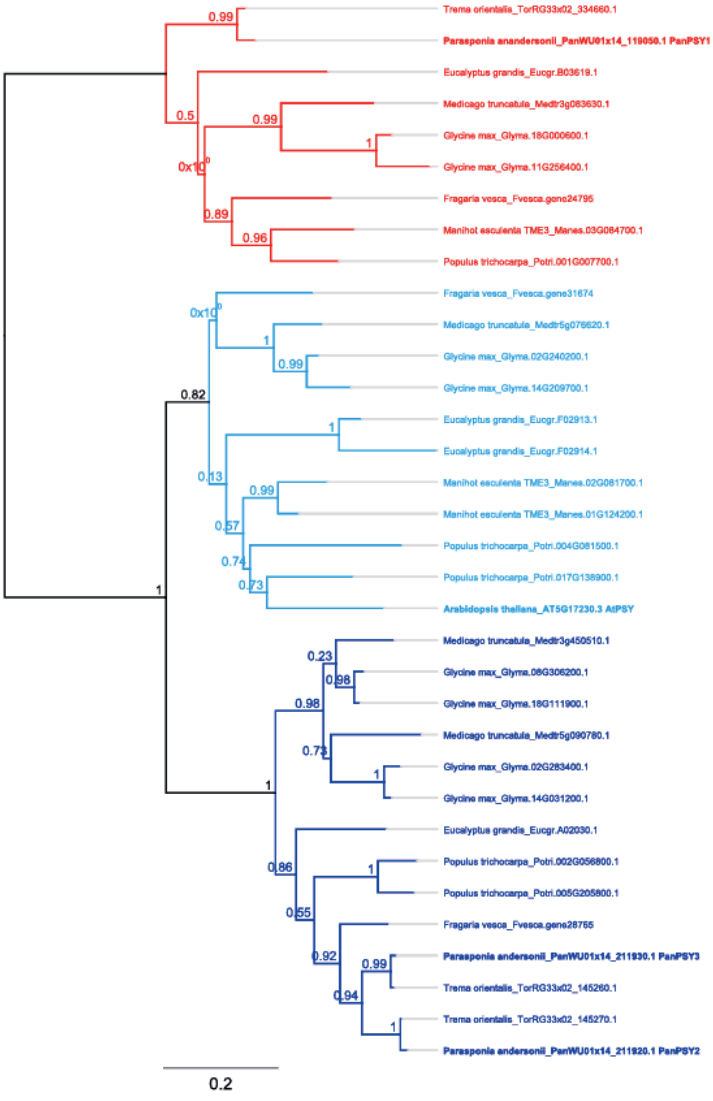

**Supplementary Figure S1: Construct design to ectopically express *Parasponia NSP2***. (**A**) Alignment of putative miR171h target sites in *Medicago truncatula* (Lauressergues *et al.*, 2012) and *Parasponia NSP2* as well as m*NSP2* a modified version of *Parasponia NSP2* in which the target site of miR171hR is silenced. Matching nucleotides at the miR171h binding site are shown in red, mismatching nucleotides in black, and silencing nucleotides in green. (**B**) A schematic representation of the transformation constructs to achieve ectopic expression of *PanNSP2* in *Parasponia*.




**Supplementary Figure S2: Ectopic *PanNSP2* expression enhances the expression of strigolactone biosynthesis genes in shoot tissue.** (**A-D**) qRT-PCR-based expression of *PanNSP2, PanCCD7, PanCDD8*, and *PanD27* in shoot of plants grown under nodulation permissive condition (n=3), and (**E-H**) nutrient starved condition (n=3). Asterisk indicates significant difference relative to cont. as determined by Student's t-test (*p < 0.05)




**Supplementary Figure S3: Shoot phenotype of *Parasponia NSP2*ox lines.** Quantification of internode diameter of the empty vector control line (cont.) and *mNSP2*ox lines 1,3 and 6. Asterisk indicates significant difference relative to cont. as determined by Student's t-test (*$p < 0.05$). Stem diameters were measured at the middle of internodes (IN) for the first 12 internodes, from the bottom to the top of 60 days old plants (n=10).




**Supplementary Figure S4**: **Phylogenetic relation of *Parasponia PSY* proteins.** Phylogenetic relation is reconstructed based on an alignment of protein sequences from the following species: *Parasponia andersonii, Trema orientalis, Arabidopsis thaliana, Medicago truncatula, Glycine max, Fragaria vesca, Eucalyptus grandis,* and *Manihot esculenta. Parasponia* and *Arabidopsis* proteins are in bold letter-type. Three orthogroups are recognized and marked in distinct colors. Node labels indicate posterior probability.

## References

**Alhusayni, S. *et al.*** (2023) 'A rare non-canonical splice site in Trema orientalis SYMRK does not affect its dual symbiotic functioning in endomycorrhiza and rhizobium nodulation'.

**Ali, L. *et al.*** (2014) 'Mapping and identification of a Cicer arietinum NSP2 gene involved in nodulation pathway', *TAG. Theoretical and applied genetics. Theoretische und angewandte Genetik*, **127**(2), pp. 481–488.

**Becking, J.H.** (1983) 'TheParasponia parviflora—Rhizobium symbiosis. Host specificity, growth and nitrogen fixation under various conditions', *Plant and soil*, **75**(3), pp. 309–342.

**Bonaldi, K. *et al.*** (2011) 'Nodulation of Aeschynomene afraspera and A. indica by photosynthetic Bradyrhizobium Sp. strain ORS285: the nod-dependent versus the nod-independent symbiotic interaction', *Molecular plant-microbe interactions: MPMI*, **24**(11), pp. 1359–1371.

**Brewer, P.B. *et al.*** (2015) 'Strigolactone Inhibition of Branching Independent of Polar Auxin Transport', *Plant physiology*, **168**(4), pp. 1820–1829.

**Bu, F. *et al.*** (2020) 'Mutant analysis in the nonlegume Parasponia andersonii identifies NIN and NF-YA1 transcription factors as a core genetic network in nitrogen-fixing nodule symbioses', *The New phytologist*, **226**(2), pp. 541–554.

**Capoen, W. and Oldroyd, G.** (2008) 'How CYCLOPS keeps an eye on plant symbiosis', *Proceedings of the National Academy of Sciences of the United States of America*, **105**(51), pp. 20053–20054.

**Carvalhais, L.C. *et al.*** (2019) 'The ability of plants to produce strigolactones affects rhizosphere community composition of fungi but not bacteria', *Rhizosphere*, **9**, pp. 18–26.

**Cenci, A. and Rouard, M.** (2017) 'Evolutionary Analyses of GRAS Transcription Factors in Angiosperms', *Frontiers in plant science*, **8**, p. 273.

**Chen, C., Ané, J.-M. and Zhu, H.** (2008) 'OsIPD3, an ortholog of the Medicago truncatula DMI3 interacting protein IPD3, is required for mycorrhizal symbiosis in rice', *The New phytologist*, **180**(2), pp. 311–315.

**Das, D. *et al.*** (2022) 'PHOSPHATE STARVATION RESPONSE transcription

factors enable arbuscular mycorrhiza symbiosis', *Nature communications*, **13**(1), p. 477.

**Das, D.R. et al.** (2019) 'Functional conservation of CYCLOPS in crack entry legume Arachis hypogaea', Plant science: an international journal of experimental plant biology, **281**, pp. 232–241.

**Engler, C. *et al.*** (2009) 'Golden gate shuffling: a one-pot DNA shuffling method based on type IIs restriction enzymes', *PloS one*, **4**(5), p. e5553.

**Engler, C. *et al.*** (2014) 'A golden gate modular cloning toolbox for plants', *ACS synthetic biology*, **3**(11), pp. 839–843.

**Feike, D. et al.** (2019) 'Characterizing standard genetic parts and establishing common principles for engineering legume and cereal roots', Plant biotechnology journal, **17**(12), pp. 2234–2245.

**Fonouni-Farde, C. *et al.*** (2016) 'DELLA-mediated gibberellin signalling regulates Nod factor signalling and rhizobial infection', *Nature communications*, **7**, p. 12636.

**Giovannetti, M. and Mosse, B.** (1980) 'An evaluation of techniques for measuring vesicular arbuscular mycorrhizal infection in roots', *The New phytologist*, **84**(3), pp. 489–500.

**Gobbato, E. *et al.*** (2012) 'A GRAS-Type Transcription Factor with a Specific Function in Mycorrhizal Signaling', *Current biology: CB*, **22**(23), pp. 2236–2241.

**Heck, C. *et al.*** (2016) 'Symbiotic Fungi Control Plant Root Cortex Development through the Novel GRAS Transcription Factor MIG1', *Current biology: CB*, **26**(20), pp. 2770–2778.

**Heckmann, A.B. *et al.*** (2006) 'Lotus japonicus nodulation requires two GRAS domain regulators, one of which is functionally conserved in a non-legume', *Plant physiology*, **142**(4), pp. 1739–1750.

**Hirsch, S. *et al.*** (2009) 'GRAS proteins form a DNA binding complex to induce gene expression during nodulation signaling in Medicago truncatula', *The Plant cell*, **21**(2), pp. 545–557.

**Hofferek, V. *et al.*** (2014) 'MiR171h restricts root symbioses and shows like its target NSP2 a complex transcriptional regulation in Medicago truncatula', *BMC plant biology*, **14**, p. 199.

**Horváth, B. *et al.*** (2011) 'Medicago truncatula IPD3 is a member of the common symbiotic signaling pathway required for rhizobial and mycorrhizal symbioses', *Molecular plant-microbe interactions: MPMI*, **24**(11), pp. 1345–1358.

**Imaizumi-Anraku, H. *et al.*** (2005) 'Plastid proteins crucial for symbiotic fungal and bacterial entry into plant roots', *Nature*, 433(7025), pp. 527–531.

**Isidra-Arellano, M.C., Singh, J. and Valdés-López, O.** (2024) 'Unraveling the potential of the strigolactones-NSP1/NSP2 friendship in crop improvement', Trends in plant science. Elsevier BV, **29**(5), pp. 501–503.

**Jin, Y. *et al.*** (2016) 'DELLA proteins are common components of symbiotic rhizobial and mycorrhizal signalling pathways', *Nature communications*, **7**, p. 12433.

**Jin, Y. *et al.*** (2018) 'IPD3 and IPD3L Function Redundantly in Rhizobial and Mycorrhizal Symbioses', *Frontiers in plant science*, **9**, p. 267.

**Kaló, P. *et al.*** (2005) 'Nodulation signaling in legumes requires NSP2, a member of the GRAS family of transcriptional regulators', *Science*, **308**(5729), pp. 1786–1789.

**Kobae, Y. *et al.*** (2018) 'Strigolactone Biosynthesis Genes of Rice are Required for the Punctual Entry of Arbuscular Mycorrhizal Fungi into the Roots', *Plant & cell physiology*, **59**(3), pp. 544–553.

**Lancelle, S.A. and Torrey, J.G.** (1984) 'Early development of Rhizobium-induced root nodules of Parasponia rigida. I. Infection and early nodule initiation', *Protoplasma*, 123, pp. 26–37.

**Lauressergues, D. *et al.*** (2012) 'The microRNA miR171h modulates arbuscular mycorrhizal colonization of Medicago truncatula by targeting NSP2', *The Plant journal: for cell and molecular biology*, **72**, pp. 512–522.

**Lindsay, P.L. *et al.*** (2019) 'A Phosphate-Dependent Requirement for Transcription Factors IPD3 and IPD3L During Arbuscular Mycorrhizal Symbiosis in Medicago truncatula', *Molecular plant-microbe interactions: MPMI*, **32**(10), pp. 1277–1290.

**Liu, W. *et al.*** (2011) 'Strigolactone biosynthesis in Medicago truncatula and rice requires the symbiotic GRAS-type transcription factors NSP1 and NSP2', *The Plant cell*, **23**(10), pp. 3853–3865.

**Li, X.-R. *et al.*** (2022) 'Nutrient regulation of lipochitooligosaccharide recognition in plants via NSP1 and NSP2', *Nature communications*, **13**(1), p. 6421.

**Maekawa, T. *et al.*** (2008) 'Polyubiquitin promoter-based binary vectors for overexpression and gene silencing in Lotus japonicus', *Molecular plant-microbe interactions: MPMI*, **21**(4), pp. 375–382.

**Marzec, M. and Melzer, M.** (2018) 'Regulation of Root Development and Architecture by Strigolactones under Optimal and Nutrient Deficiency Conditions', *International journal of molecular sciences*, **19**(7). Available at: https://doi.org/10.3390/ijms19071887.

**Murray, J.D. *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', *The Plant journal: for cell and molecular biology*, **65**(2), pp. 244–252.

**Oldroyd, G.E. and Long, S.R.** (2003) 'Identification and characterization of nodulation-signaling pathway 2, a gene of Medicago truncatula involved in Nod factor signaling', *Plant physiology*, **131**(3), pp. 1027–1032.

**Park, H.-J. *et al.*** (2015) 'Hyphal Branching during Arbuscule Development Requires Reduced Arbuscular Mycorrhiza1', *Plant physiology*, **169**(4), pp. 2774–2788.

**Peng, Z. et al.** (2021) 'Natural polymorphisms in a pair of NSP2 homoeologs can cause loss of nodulation in peanut', Journal of experimental botany. Oxford University Press (OUP), **72**(4), pp. 1104–1118.

**Pimprikar, P. *et al.*** (2016) 'A CCaMK-CYCLOPS-DELLA Complex Activates Transcription of RAM1 to Regulate Arbuscule Branching', *Current biology: CB*, **26**(8), pp. 987–998.

**Prihatna, C. *et al.*** (2018) 'Tomato CYCLOPS/IPD3 is required for mycorrhizal symbiosis but not tolerance to Fusarium wilt in mycorrhiza-deficient tomato mutant rmc', *Mycorrhiza*, **28**(5-6), pp. 495–507.

**Pumplin, N. *et al.*** (2010) 'Medicago truncatula Vapyrin is a novel protein required for arbuscular mycorrhizal symbiosis', *The Plant journal: for cell and molecular biology*, **61**(3), pp. 482–494.

**Quilbé, J.,Nouwen, N., Pervent, M., Guyonnet, R., Cullimore, J., Gressent, F. et al.** (2022) 'A mutant-based analysis of the establishment of Nod-independent symbiosis in the legume Aeschynomene evenia', Plant physiology. Oxford University Press (OUP), **190**(2), pp. 1400–1417.

**Quilbé, J. and Arrighi, J.-F.** (2021) 'NSP2, a key symbiotic regulator in the spotlight', *Journal of experimental botany*, **72**(4), pp. 959–963.

**Rutten, L. *et al.*** (2020) 'Duplication of Symbiotic Lysin Motif Receptors Predates the Evolution of Nitrogen-Fixing Nodule Symbiosis', *Plant*

*physiology*, **184**(2), pp. 1004–1023.

**Ruyter-Spira, C. *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**(2), pp. 721–734.

**Schlemper, T.R. *et al.*** (2017) 'Rhizobacterial community structure differences among sorghum cultivars in different growth stages and soils', *FEMS microbiology ecology*, **93**(8). Available at: https://doi.org/10.1093/femsec/fix096.

**Shtark, O.Y. *et al.*** (2016) 'Arbuscular mycorrhiza development in pea (Pisum sativum L.) mutants impaired in five early nodulation genes including putative orthologs of NSP1 and NSP2', *Symbiosis* , **68**(1-3), pp. 129–144.

**Stanley, L. and Yuan, Y.-W.** (2019) 'Transcriptional regulation of carotenoid biosynthesis in plants: So many regulators, so little consensus', *Frontiers in plant science*, **10**, p. 1017.

**van Velzen, R. *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 of the United States of America*, **115**(20), pp. 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 in plant science*, **24**(1), pp. 49–57.

**Votta, C. *et al.*** (2022) 'Zaxinone synthase controls arbuscular mycorrhizal colonization level in rice', *The Plant journal: for cell and molecular biology*, **111**(6), pp. 1688–1700.

**Wardhani, T.A.K. *et al.*** (2019) 'Transforming, genome editing and phenotyping the nitrogen-fixing tropical Cannabaceae tree Parasponia andersonii', Journal of visualized experiments: JoVE. MyJove Corporation, (150). doi: 10.3791/59971.

**Xiao, A. *et al.*** (2020) 'Transcriptional regulation of NIN expression by IPN2 is required for root nodule symbiosis in Lotus japonicus', *The New phytologist*, **227**(2), pp. 513–528.

**Xue, L. *et al.*** (2015) 'Network of GRAS transcription factors involved in the control of arbuscule development in Lotus japonicus', *Plant physiology*, **167**(3), pp. 854–871.

**Yano, K. *et al.*** (2008) 'CYCLOPS, a mediator of symbiotic intracellular

accommodation', *Proceedings of the National Academy of Sciences of the United States of America*, **105**(51), pp. 20540–20545.

**Yuan, K. *et al.*** (2023) 'Low phosphorus promotes NSP1-NSP2 heterodimerization to enhance strigolactone biosynthesis and regulate shoot and root architecture in rice', Molecular plant, 16(11), pp. 1811–1831.

**van Zeijl, A. *et al.*** (2018) 'CRISPR/Cas9-Mediated Mutagenesis of Four Putative Symbiosis Genes of the Tropical Tree Parasponia andersonii Reveals Novel Phenotypes', *Frontiers in plant science*, **9**, p. 284.

**Zhou, X. *et al.*** (2022) 'Phytoene synthase: The key rate-limiting enzyme of carotenoid biosynthesis in plants', Frontiers in plant science, 13, p. 884720.

CHAPTER 4

# A Comprehensive Review of Autoregulation of Nodulation (AON): Genetic Pathways and Evolutionary Insights into Plant Symbiosis

Sultan Alhusayni[1,2]

[1]Laboratory of Molecular Biology, Department of Plant Sciences, Wageningen University, Droevendaalsesteeg 1, 6708 PB Wageningen, The Netherlands

[2]Biological Sciences Department, College of Science, King Faisal University, Al-Ahsa 31982, Saudi Arabia

## Introduction

Nutrient availability in soil is one of the key limiting factors for plant growth and development. To overcome nutrient shortages, plants have evolved the capacity to engage in endosymbiosis with nutrient-scavenging microbes. For example, most land plants participate in symbiosis with arbuscular mycorrhizal (AM) fungi, whereas leguminous and actinorhizal plants can also form a root nodule endosymbiosis with nitrogen-fixing rhizobium or *Frankia* bacteria, respectively (Smith and Read, 2008; Doyle, 2011; Geurts, Lillo and Bisseling, 2012). The rhizobium-legume endosymbiosis has been extensively studied. In nitrogen-deficient soil, legumes engage with nitrogen-fixing rhizobium to satisfy their nitrogen needs. This interaction results in the formation of many root nodules, which host millions of bacteria inside the nodule cells. This intimate interaction comes at a cost. Maintaining bacterial fitness and driving the reduction of dinitrogen into ammonia require an investment of up to 25% of the plant's photosynthates (Minchin *et al.*, 1981; Schuize, Adgo and Merbach, 1999; Oono and Denison, 2010). Consequently, legume plants have evolved an autoregulation of nodulation (AON) mechanism to tightly control and balance the engagement with nitrogen-fixing rhizobium to prevent the overproduction of ammonia, which can negatively impact plant growth, development, and yield (Britto and Kronzucker, 2002; Liu *et al.*, 2013; Pan *et al.*, 2016; Liu and von Wirén, 2017).



This chapter highlights the progress made in comprehending the mechanism of AON in legumes and the genetic dissection of the key players involved in the AON pathway.

## A mechanism controlling infection and nodule organogenesis at early stages

Research in soybean (*Glycine max*) revealed that rhizobium inoculation triggers many infection events, of which only a few are successful, resulting in the formation of a limited number of functional nodules around the root

crown. Anatomical studies revealed that unsuccessful infection events are aborted in the epidermis and are accompanied by sub-epidermal cell division (Calvert *et al.*, 1984; Mathews, Carroll and Gresshoff, 1989; Reid *et al.*, 2011). In *G. max*, nodule formation is anatomically divided into 8 developmental stages. The first four stages are marked by hypodermal and cortical anticlinal cell divisions that form a young nodule meristem beneath the infected curled root hair. Research on these developmental stages has revealed that most infection events arrest at stages 1 to 4 **(Figure 1)** (Calvert *et al.*, 1984; Mathews, Carroll and Gresshoff, 1989; Reid *et al.*, 2011). This suggests that there may be an active mechanism controlling infection and nodule organogenesis at these early stages of the interaction.

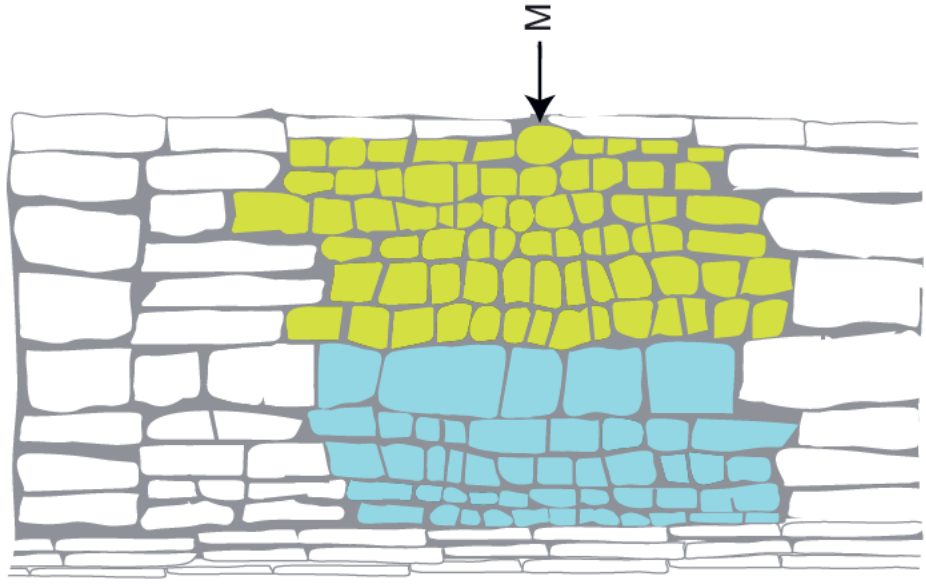



**Figure 1: Illustration of the stage of nodule development in soybean in which AON-triggered abortion of infection occurs**. The cortical cells that form the nodule primordium are colored. The outermost cortical cells are colored in yellow-green, and the inner cortical cells are colored in light blue. M: marks the place of infection. The illustration is based on Calvert et al. (1984).

Split root experiments in legumes were among the early studies investigating the AON mechanism. These simple but powerful experiments

revealed that AON is activated systemically through long-distance signals that travel from root to shoot and back to the root to inhibit nodule formation (Delves *et al.*, 1986; Olsson *et al.*, 1989). In a split root experiment, the plant root is separated into two sides, with one side inoculated some days earlier than the second (delayed inoculation). Split root experiments on *G. max, Pisum sativum, Medicago truncatula,* and *Lotus japonicus* have demonstrated that AON is activated within 1 to 4 days after infection. For example, a difference of two days in inoculation between both sides of the root system already resulted in a remarkably reduced number of nodules formed on the "delayed" side of the root, whereas, a seven-day inoculation window between both sides of the root system completely inhibited nodule formation on the 'delayed' side (Kosslak and Bohlool, 1984; Olsson *et al.*, 1989; Suzuki *et al.*, 2008; Li, Kinkema and Gresshoff, 2009; Kassaw and Frugoli, 2012).

In conclusion, AON regulates nodule number systemically through interaction between the root and the shoot. It is activated within the early days of infection, before plants fulfill their nitrogen needs by forming functional nodules. Finally, AON is a generic mechanism to control nodulation found in various legume species.



**Genetic dissection of AON**

Genetic screenings for genes involved in controlling nodule number have been saturating, leading to the discovery of many orthologous genes involved in the AON pathway in different legume species (Carroll, McNeil and Gresshoff, 1985; Sagan *et al.*, 1995; Krzysztof Szczyglowski *et al.*, 1998). Legume mutants in AON are characterized by a hypernodulation / supernodulation phenotype. Interestingly, nodulation in these mutants is also tolerant to exogenous nitrate, pointing to an overlap between the AON genetic network and the nitrate regulation pathway (Day *et al.*, 1986; Jeudy *et al.*, 2010; Okamoto and Kawaguchi, 2015). Components of the AON pathway **(Table 1)** are categorized into four groups: systemic root-derived

small peptide signals, shoot receptor signal receivers, systemic shoot-derived non-coding microRNA (miRNA) and cytokinin signals, and a root Kelch F-box protein targeted by the systemic miRNA signal. These four groups will be discussed below.

**The flow of the AON pathway**

Upon rhizobium and nitrate application *NODULE INCEPTION* (*NIN*) and NIN-like proteins (NLP) are activated in the root, respectively. **(Figure 2)** *NIN* and *NLP* genes encode for transcription factor proteins involved in the regulation of a group of root-specific genes, which encode *CLAVATA 3* (*CLV3*) and *EMBRYO SURROUNDING REGION* (*ESR*)-*Related* proteins collectively referred to as CLE peptides. For example, in response to rhizobium, NIN induces in *G. max RHIZOBIA-INDUCED CLE PEPTIDE 1* (*GmRIC1*) *and GmRIC2,* in *L. japonicus CLE-ROOT SIGNAL 1 (LjCLE-RS1)*, *LjCLE-RS2*, and *LjCLE-RS3*, and in *M. truncatula MtCLE12* and *MtCLE13*. Conversely, in response to nitrate, NLP proteins induce in *G. max NITRATE-INDUCED CLE 1 (GmNIC1),* and *GmNCI2,* in *L. japonicus LjCLE-RS2*, and in *M. truncatula MtCLE35* (Okamoto et al., 2009; Mortier et al., 2010; Lim, Lee and Hwang, 2011; Sasaki et al., 2014; Hastwell et al., 2017; Kassaw et al., 2017; Moreau, Gautrat and Frugier, 2021; Bashyal, Gautam and Müller, 2023). The induced CLE peptides in the root undergo post-translational modification and translocate to the shoot through the xylem. In the shoot, they are perceived by a leucine-rich repeats receptor-like kinase (LRR-LK). This receptor is called *NODULE AUTOREGULATION RECEPTOR KINASE* (*GmNARK*) in *G. max* (Delves *et al.*, 1986; Sheng and Harper, 1997; Searle *et al.*, 2003)*, SYM29 in P. sativum* (Sagan and Duc, 1996)*, HYPERNODULATED ABERRANT ROOT FORMATION 1* (*LjHAR1*) in *L. japonicus* (K. Szczyglowski *et al.*, 1998; Wopereis *et al.*, 2000; Krusell *et al.*, 2002; Nishimura *et al.*, 2002; Okamoto *et al.*, 2009), and *SUPER NUMERIC NODULE* (*MtSUNN*) in *M. truncatula* (Sagan *et al.*, 1995; Penmetsa *et al.*, 2003; Schnabel *et al.*, 2005). Activation of this receptor leads to the suppression of a shoot-derived signal, microRNA2111 (miRNA2111) which, in the absence of rhizobium and nitrate, is produced in the shoot and

transferred to the root through the phloem. The suppression miRNA2111 production in the shoot correlates with the reduction of its abundance level in the root. Low abundance of miRNA2111 in the root favors the accumulation of its main target *TOO MUCH LOVE* (*TML*), a root specific gene encodes for a Kelch Repeat-Containing F-box. The accumulation of TML in the root suppresses nodule formation and development.

**Table 1. Key players of AON pathway discovered by forward and reverse genetics**

| **Gene product** | ***G. max*** | ***P. sativum*** | ***M. truncatula*** | ***L. japonicus*** | **Place of production** | **Place of function** |
|---|---|---|---|---|---|---|
| CLV1-like LRR-RK | *NARK* | *Sym29* | *SUNN* | *HAR1* | Shoot/root | Shoot/root |
| CLV2-like LRR-RK | - | *SYM28* | *CVL2* | *CLV2* | Shoot/root | Shoot/root? |
| LRR-RK | - | - | - | *KLAVIER* | Shoot/root | Shoot/root? |
| LRR-RK | - | - | *CORYNE* | - | Shoot/root | Shoot/root? |
| NCT | - | - | *DAR* | | Root | Root |
| CLE peptide | *RIC1; RIC2; NIC1; NIC2* | - | *CLE12; CLE13; CLE35* | *CLE_RS1; CLE_RS2; CLE_RS3* | Root | Root |
| HPAT | - | *NOD3* | *RDN1* | *PLENTY* | Root | Root |
| KR-F-box | *TML1a; TML1b; TML2* | - | *TML1; TML2* | *TML* | Root | Root |
| pre-miRNA2111 | miRNA2111 | - | miRNA2111 | miRNA2111 | Shoot | Root |
| cytokinin | - | - | - | *IPT1; IPT3* | Shoot | Root |

HPAT= Hydroxyproline O-arabinosyltransferases, KR-F-box= Kelch Repeat-Containing F-box Protein, LRR-RK=leucine-rich repeat receptor kinase, NCT= non-cytoplasmic transmembrane spanning domains, and DAR= Defective in Autoregulation

Chapter 4

**CLE signal induction in root**

Deregulation of the *CLE* genes in legumes systematically modulates nodule number in the root. For example, overexpression of *GmRIC1*, and *GmRIC2, LjCLE-RS1*, *LjCLE-RS2*, *LjCLE-RS3, MtCLE12,* and *MtCLE13* systemically suppresses nodulation in *GmNARK/LjHAR1*/*MtSUNN-TML* dependent manner, pointing to their involvement in the AON pathway (Okamoto *et al.*, 2009; Mortier *et al.*, 2010; Lim, Lee and Hwang, 2011; Takahara *et al.*, 2013; Sasaki *et al.*, 2014; Kassaw *et al.*, 2017; Bashyal, Gautam and Müller, 2023). Additionally, overexpression of *LjCLE-RS1* and *LjCLE-RS2* in *L. Japonicus* and *MtCLE13 in M. truncatula* revealed their dependence on *LjHAR1*/*MtSUNN* to regulate lateral root and petiole number, respectively (Mortier *et al.*, 2010; Sasaki *et al.*, 2014). Finally, the induction of *GmNIC1, GmNIC2, LjCLE-RS2, and MtCLE35* through the modulation of nitrate concentration requires a fully functional GmNARK/LjHAR1/MtSUNN-TML pathway to suppress nodulation in root locally and systemically (Okamoto *et al.*, 2009; Reid, Ferguson and Gresshoff, 2011; Moreau, Gautrat and Frugier, 2021).

CLE Proteins in legumes- either those induced by rhizobium or nitrate- can undergo arabinosylation through a root-located LjPLENTY/MtRDN1/PsNOD3, which their corresponding mutants exhibit hypernodulation phenotype (Ishikawa *et al.*, 2008; Yoshida, Funayama-Noguchi and Kawaguchi, 2010; Kassaw *et al.*, 2017; Yoro *et al.*, 2019). Subsequently, nano-liquid chromatography-mass spectrometry (nano-LC-MS) on samples overexpressing *LjCLE-RS2* revealed that the encoded CLE protein undergoes cleavage into a small 12 or 13 amino acid peptide and the addition of three residues of arabinose at their 7th hydroxyproline residue (Okamoto *et al.*, 2013). Enzymatic and trans-complementation assays demonstrated that LjPLENTY/MtRDN1 has arabinosylation activity. Subcellular localization of these proteins in the Golgi apparatus provided additional support for their role in protein modification (Kassaw *et al.*, 2017; Yoro *et al.*, 2019). Furthermore, LjCLE-RS3 and MtCLE12-induced suppression of nodulation requires a functional *LjPLENTY/MtRDN1* gene.

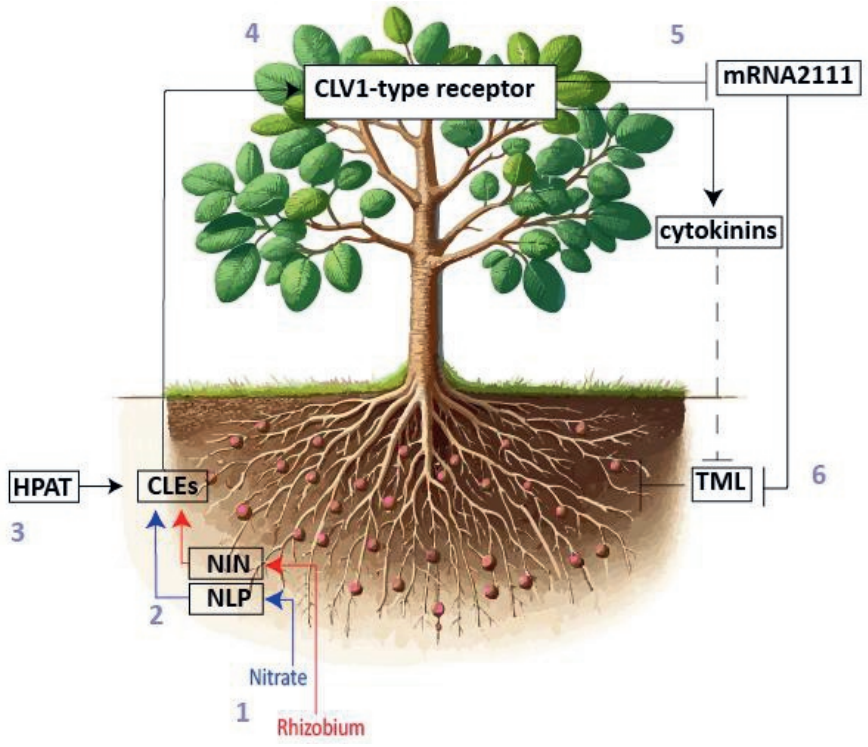



**Figure 2: Schematic representation of the autoregulation of nodulation (AON) pathway**. In the presence of rhizobium or nitrate, the transcription factors NODULE INCEPTION (NIN) or NIN-LIKE PROTEIN (NLP) encoding genes are induced in the root. The induction of NIN and NLP leads to the transcriptional activation of a group of root-specific genes encoding CLE peptides. These CLE peptides in the root undergo post-translational modification, such as cleavage into 12 or 13-amino-acid peptides and arabinosylation by Hydroxyproline O-arabinosyltransferases enzyme (HAPT). These now mature small amino-acid CLE peptides are then translocated to the shoot through the xylem, where they are perceived by CLAVATA1 (CLV1)-type receptor. Activation of the CLV1-type receptor leads to the suppression of shoot-derived signals such as microRNA2111 (miRNA2111) and cytokinins, which are transferred to the root through the phloem and target *TOO MUCH LOVE* (*TML*) messenger RNA, reducing its gene expression. Suppression of the shoot-derived signals by the CLV1-type receptor decreases their production and their abundance in the root, leading to the induction of *TML* gene expression in the root, which inhibits nodule formation and development.



In contrast, LjCLE-RS1, LjCLE-RS2, and MtCLE13 can suppress nodulation in the background of *Ljplenty/Mtrdn1* mutation, indicating that they do not undergo arabinosylation by LjPLENTY/MtRDN1, but possibly by other yet

unknown proteins. This was further supported through shoot feeding of none arabinosylated and arabioyslated synthetic peptides of LjCLE-RS1 (Ara3-LjCLE-RS1), Ara3-LjCLE-RS2, and Ara3-AtCLE3 to WT plants, revealing that only Ara3-LjCLE-RS1 and Ara3-LjCLE-RS2 were able to suppress nodulation. However, shoot feeding of these peptides to *Ljhar1* mutant plants did not result in suppression of nodulation, confirming that Ara3-LjCLE-RS1 and Ara3-LjCLE-RS2 suppression of nodulation is LjHAR1-dependent (Okamoto *et al.*, 2013; Kassaw *et al.*, 2017; Yoro *et al.*, 2019). Since the *LjCLE-RS2* gene is expressed in the root and LjHAR1 is located in the shoot, transportation of arabioyslated LjCLE-RS2 was further investigated in xylem sap of *G. max* plants with *Agrobacterium rhizogenes* transformed root expressing *LjCLE-RS2* gene under its native promoter. Xylem sap collection from the shoot and subsequent analysis using nano-LC-MS revealed the presence and transportation of arabinosylated LjCLE-RS2, produced and modified in the root, to the shoot (Okamoto *et al.*, 2013).

**Perception of CLE root signals in the shoot**

Earlier grafting and shoot apex decapitation experiments revealed that the function of *GmNARK, LjHAR1-1*, and *MtSUNN* is required in the shoot and suggested that these genes are active in the leaf (Delves, Higgins and Gresshoff, 1992; Wopereis *et al.*, 2000; Krusell *et al.*, 2002; Searle *et al.*, 2003; Schnabel *et al.*, 2005; Reid *et al.*, 2011). Subsequent research found that the encoded receptor proteins are located in leaf phloem cells, where they perceive CLE peptides to activate the AON pathway (Nontachaiyapoom *et al.*, 2007; Reid *et al.*, 2011; Sasaki *et al.*, 2014; Notaguchi and Okamoto, 2015). Mutants of *Gmnark, Ljhar1-1*, and *Mtsunn* fail to perceive CLE peptide root signals to control nodule number, resulting in high energy investment in forming a hyper number of nodules, negatively impacting overall plant growth and development. Unlike the scattered nodules on the wild-type root system, on *Gmnark, Ljhar1-1*, and *Mtsunn* mutant roots are massively covered with nodules. Sectioning of these nodules revealed a wild-type mode of infection (Krzysztof Szczyglowski *et al.*, 1998; K. Szczyglowski *et al.*,

1998; Wopereis *et al.*, 2000). Moreover, nodulation permissiveness of *Gmnark, Ljhar1-1*, and *Mtsunn* mutants show tolerance to high exogenous nitrogen ($NH_4NO_3$) concentrations. This contrasts with wild-type plants, which only form bump-like nodule structures under similar exogenous nitrogen conditions (Delves *et al.*, 1986; Sagan *et al.*, 1995; Wopereis *et al.*, 2000; Schnabel *et al.*, 2005). Finally, mutating *LjHAR1* in the background of the autoactive cytokinin receptor *LHK1*, named *spontaneous nodule 2* (*snf2*) mutant, results in an enormous number of spontaneous nodules, indicating that nodule organogenesis is sufficient to activate the AON feedback mechanism (Tirichine *et al.*, 2007).

Another interesting function of the AON pathway in legumes is its role in controlling root development. Unlike wild-type plants, non-inoculated roots of *Gmnark and Ljhar1-1,* but not *Mtsunn,* showed an increased density of lateral roots, most likely due to increased cell division in the pericycle and cortical cell layers (Wopereis *et al.*, 2000; Searle *et al.*, 2003; Schnabel *et al.*, 2005). Furthermore, non-inoculated roots of *Ljhar1-1* and *Mtsunn* mutant plants exhibit shorter primary roots, and the severity of this phenotype is increased upon rhizobium inoculation. Additionally, a closer examination of *Ljhar1-1* roots revealed a decrease in root diameter compared to WT, most likely due to impaired root cell expansion (Wopereis *et al.*, 2000; Schnabel *et al.*, 2005; Chaulagain *et al.*, 2023).



Expression analyses showed that *GmNARK, LjHAR1* and *MtSUNN* are expressed in both shoots and roots of *G. max, L. Japonicus* and *M. truncatula,* respectively. Their expression is not altered upon rhizobium and nitrate treatment, indicating that they are transcriptionally regulated by another factor (Krusell *et al.*, 2002; Searle *et al.*, 2003; Schnabel *et al.*, 2010). This led to the hypothesis that root-produced CLE peptides are transferred to the shoot to activate the GmNARK, LjHAR1 and MtSUNN receptors. This was confirmed through photoaffinity labeling in *L. Japonicus*, demonstrating that LjHAR1 Halo Tag (HAR1-HT) protein is bound by Ara3-LjCLE-RS1 and Ara3-LjCLE-RS2, but not Ara3-AtCLV3. In contrast, Ljhar1$^{L246F}$ mutated

protein, which carries a mutation in Leu$^{246}$, mimicking the mutation of *Ljhar1-4* mutant, was unable to bind to Ara3-LjCLE-RS1, Ara3-LjCLE-RS2 and Ara3-AtCLV3. This confirmed that LjCLE-RS1 and LjCLE-RS2 are the ligands of the LjHAR1 receptor in the shoot (Okamoto *et al.*, 2013).

In *M. truncatula*, a receptor with an inactive kinase domain named CORYNE (MtCRN) was identified as the closest ortholog of AtCRN. A *Tnt1* (*transposable element of Nicotiana tabacum*) insertion mutation in the first exon of *MtCRN* results in a hyperrnodulation phenotype, revealing its involvement in the AON pathway in legumes (Crook, Schnabel and Frugoli, 2016; Nowak, Schnabel and Frugoli, 2019). Subcellular localization studies showed that MtCRN is localized in the plasma membrane and the endoplasmic reticulum, while MtSUNN is localized only in the plasma membrane. Co-expression of MtCRN and MtSUNN in *N. benthamiana* leaves demonstrated that they can form a heterodimer complex in the plasma membrane (Crook, Schnabel and Frugoli, 2016). In addition to the interaction of MtSUNN on the protein level, gene expression studies revealed that *MtSUNN* has a genetic interaction with *Like Sunn Supernodulator* (*MtLSS*) locus (Schnabel *et al.*, 2010). *Mtlss* mutant exhibited ethylene and nitrate insensitivity, as well as a hypernodualtion phenotype, which is very similar to the phenotype of the *Mtsunn* mutant (Schnabel *et al.*, 2010). Although *Mtlss* mutant does not show any mutations in the *MtSUNN* gene or its surrounding region, a mutant cross between *Mtlss* and *Mtsunn* could not rescue the hypernodualtion phenotype. This ambiguity was explained by the drastic reduction in *MtSUNN* gene expression in the *Mtlss* mutant, indicating that *MtSUNN* is regulated by *MtLSS* locus (Schnabel *et al.*, 2010).

An ortholog of *A. thaliana* AtCLV2 was also identified in *L. japonicus, M. truncatula*, and *P. sativum*, called *CLAVATA2* (*LjCLV2, MtCLV2,* and *PsCLV2*, respectively) (Bleckmann et al., 2010; Zhu et al., 2010; Krusell et al., 2011; Crook, Schnabel and Frugoli, 2016). The *CLV2* gene encodes a receptor lacking the intracellular kinase domain, and a mutation in *LjCLV2* and

*PsCLV2*, but not *MtCLV2*, results in a hyperproduction phenotype (Krusell *et al.*, 2011; Crook, Schnabel and Frugoli, 2016). The *Ljclv2* mutant has a *Tnt1* insertion in the coding sequence that disrupts the gene, and *Psclv2* mutant has a single nucleotide substitution, leading to an early premature stop codon. However, the *Mtclv2* mutant has a *Tnt1* insertion in the non-coding region upstream of the start transcript site, indicating that *MtCLV2* gene expression is minimally disrupted (Krusell *et al.*, 2011; Crook, Schnabel and Frugoli, 2016). Transient co-expression of MtCLV2 and MtSUNN in *N. benthamiana* leaves showed that they can form heterodimers; however, co-expression of LjCLV2 *and* LjHAR1 failed to form such an interaction. This inconsistency is explained by their protein subcellular localization: MtCLV2 and MtSUNN co-localized in the plasma membrane, whereas LjCLV2 and LjHAR1 are localized differently, with LjCLV2 in the endoplasmic reticulum and LjHAR1 in the plasma membrane (Krusell *et al.*, 2011; Crook, Schnabel and Frugoli, 2016). Finally, transient co-expression of MtCLV2, MtSUNN, and MtCRN in *N. benthamiana* leaves revealed that they could form a complex similar to what has been reported for their orthologs in *A. thaliana* (Bleckmann et al., 2010; Zhu et al., 2010; Krusell et al., 2011; Crook, Schnabel and Frugoli, 2016).



Another interactor of LjHAR1 in *L. japonicus* is a leucine-rich repeat receptor-like kinase protein encoded by *LjKLAVIER (LjKLV).* An out-of-frame deletion in the coding sequence of *LjKLV* gene resulted in a hypernodulation phenotype, with nodules covering about 80 % of the root system (Oka-Kira *et al.*, 2005). In contrast to wild-type plants, the *Ljklv* mutant grown under $KNO_3$ (5 mM) was able to form nodules, and even under a 20 mM $KNO_3$ concentration, it continued to form nodules, showing tolerance to $KNO_3$. However, it is worth noting that the observed tolerance to $KNO_3$ is less pronounced compared to the *Ljhar1* mutant, as the *Ljklv* hypernodulation level was slightly reduced under high $KNO_3$ (Oka-Kira *et al.*, 2005). A grafting experiment between wild-type and *Ljklv* shoot and root parts showed that the graft composed of wild-type shoot but not root could rescue the hypbernodulation phenotype of *Ljklv* mutant, indicating that also *LjKLV*

functioning is mainly in the shoot. Like LjHAR1, LjKLV protein is required for LjCLE-RS1 and LjCLE-RS2-triggered suppression of nodulation (Oka-Kira *et al.*, 2005; Miyazawa *et al.*, 2010). Moreover, like *LjHAR1*, *LjKLV* gene is required for the activation of the AON pathway induced by the gain of function *snf2* mutant, as the *Ljklv Ljsnf2* double mutant shows massive formation of spontaneous nodules. Furthermore, the *Ljklv* mutant also exhibits non-symbiotic phenotypes such as aberrant leaf veins, increased flower number, and delayed flowering time, revealing that it has a non-symbiotic function in plants (Oka-Kira *et al.*, 2005; Miyazawa *et al.*, 2010). The symbiotic and non-symbiotic phenotypes of the *Ljhar1 Ljklv* double mutant are similar to those observed in the *Ljhar1* and *Ljklv* single mutants, indicating that they function in the same pathway, and none of them is epistatic to the other. Finally, tissue-specific expressions of *LjHAR1* and *LjKLV* reveal that they are co-expressed in vasculature cells, suggesting a physical interaction is possible. This was confirmed through their transient expression in *N. benthamiana* leaves, which demonstrated that LjHAR1 and LjKLV proteins can form homodimers and heterodimers (Miyazawa *et al.*, 2010).

**Shoot derive signals**

Recent research in legumes has revealed that MicroRNA2111 (miRNA2111) is the most prominent shoot-to-root signal, functioning downstream of *GmNARK, LjHAR1* and *MtSUNN* receptors and upstream of its direct target *TML*. It also highlighted the involvement of miRNA2111 in the regulation of plant susceptibility to rhizobium infection and nodulation. For instance, non-inoculated roots of *G. max*, *L. japonicus,* and *M. truncatula* exhibited high levels of miRNA2111, while *TML* expression level is low. This contrasts with rhizobium inoculated roots, which show significantly reduced levels of miRNA2111 and an alleviated transcript level of *TML*. Moreover, the observed abundance of miRNA2111 in the shoot and root requires a functional LjHAR1 and GmNARK proteins in the shoot (Tsikou *et al.*, 2018; Okuma *et al.*, 2020; Zhang *et al.*, 2020). In agreement with this, *M.*

*truncatula* plants carrying *Agrobacterium rhizogenes* transgenic roots overexpressing miRNA2111, but not miRNA2111 short tandem target mimic (miRNA2111STTM), show increases in infection threads and nodule number, as well as reduction of *TML* expression (Tsikou *et al.*, 2018; Zhang *et al.*, 2020). Additionally, overexpression of miRNA2111 in the background of *Gmnark* mutant reduces its hypernodulation phenotype, indicating that miRNA2111 is suppressed by *GmNARK* in the shoot (Zhang *et al.*, 2020). Furthermore, grafting experiments in *L. japonicus* showed that grafts composed of a shoot from a miRNA2111 overexpressing plant and a root from an *Ljtml* mutant plant did not manifest an additive effect on the hypernodulation phenotype of *Ljtml* mutant, indicating that miRNA2111 and *LjTML* are functioning in the same pathway (Okuma et al., 2020). Similarly, in *L. japonicus,* grafting experiments showed that grafts composed of a shoot from a miRNA2111 overexpressing plant and a root from wild-type plant revealed that miRNA2111 accumulation in the shoot is sufficient to increase nodules number on the root (Okuma *et al.*, 2020). Likewise, in *G. max* double-stranded synthetic miRNA2111 feeding through petioles enhanced nodule number on the wild-type plant root, suggesting that miRNA2111 is transported through the phloem. This was further supported by the expression of miRNA2111 in the phloem of mature leaves, and the detection of miRNA2111 in the phloem sap of *G. max* and *L. japonicus*. Taken together, these data confirm that miRNA2111 plays a key role in the AON pathway and systematically regulates root nodule number through modulation of *TML* gene expression in the root (Tsikou *et al.*, 2018; Zhang *et al.*, 2020).

In *L. japonicus,* cytokinins, another important but less prominent shoot-to-root signal, were found to function downstream of *LjHAR1.* For example, rhizobium inoculation or overexpression of *LjCLE-RS1* and *LjCLE-RS2* induce expression of the cytokinin biosynthesis *ISOPENTENYL TRANSFERASE 1* (*LjIPT1*) and *LjIPT3* in the shoot in LjHAR1 dependent manner, with a more pronounced induction of *LjIPT3* gene expression than *LjIPT1*. Moreover, isopentenyl riboside monophosphate (iPRPs), *LjIPT* products, are highly

accumulated in the shoot of rhizobium-inoculated wild-type plants, and in the shoot of *LjCLE-RS1* and *LjCLE-RS2* overexpressing plants, with full dependence on *LjHAR1* gene. Furthermore, the spatial expression of the *LjIPT3* promoter driving The beta-glucuronidase (GUS) reporter gene showed that *LjIPT3* is expressed in the leaf phloem, similar to *LjHAR1* (Nontachaiyapoom et al., 2007; Sasaki et al., 2014). Interestingly, the mutation of *LjIPT3* in the shoot, but not the root, resulted in an increased nodule number. In contrast, overexpression of *LjIPT3* both in the shoot and the root reduced nodule number, indicating that *LjIPT3* negatively regulates nodulation systematically from the shoot. Furthermore, shoot feeding of Benzoylaminopurine (BOP), a synthetic form of cytokinin, reduces the number of nodules and lateral roots in a *TML*-dependent manner. Therefore, it is concluded that shoot-produced cytokinins are transferred to the root to negatively regulate nodulation (Sasaki *et al.*, 2014).

**Root targets of shoot-derived signals**

Genetic screening in legumes discovered a root-specific gene that plays a role in the AON pathway downstream of the LRR-LK shoot receptor. This gene is called *TOO MUCH LOVE* (*TML*), and in legumes there is at least one copy of this gene. For example, *L. japonicus* has a single copy *TML* gene, whereas in *M. truncatula* and *G. max TML* has undergone duplications, resulting in *MtTML1 and MtTML2* and *GmTML1a, GmTML1b,* and *GmTML2*, respectively (Magori et al., 2009; Gautrat et al., 2019; Zhang et al., 2020). The *TML* gene is expressed in roots and nodules and encodes a Kelch Repeat F-box protein containing two nuclear localization signals. The kelch repeat domain suggests that it can interact with other proteins, whereas the F-box domain implies that TML is involved in protein degradation. Subcellular localization studies revealed that TML localizes to the nucleus (Takahara *et al.*, 2013). Moreover, *tml* mutants in legumes are characterized by hypernodulation, tolerance to exogenous nitrate concerning nodulation permissiveness, and a short primary root (Chaulagain *et al.*, 2023). Interestingly, a mutant carrying mutations in both, the gain of function *LHK1* (*snf2*) and knockout of *tml* exhibits a phenotype of massive spontaneous

nodulation, indicating that AON functions downstream of the cytokinin receptor (Takahara *et al.*, 2013). Finally, in the absence of rhizobia and nitrate, *TML* gene expression is barely detectable, whereas in response to rhizobium and nitrate application, its gene expression is remarkably increased in a GmNARK, LjHAR1 and MtSUNN-dependent manner (Magori *et al.*, 2009; Takahara *et al.*, 2013; Gautrat *et al.*, 2019; Gautrat, Laffont and Frugier, 2020; Zhang *et al.*, 2020; Chaulagain *et al.*, 2023; Schnabel *et al.*, 2023). Taken together, TML plays a key role in the suppression of nodulation and is the most downstream component of the AON pathway.

**AON regulation of mycorrhization**

Split-root experiments have shown that systemic suppression of arbuscular mycorrhizal (AM) fungi colonization occurs in the "delayed" side of the root in barley, soybean, and alfalfa. Pre-inoculation with AM fungi, Nod factors, or rhizobium on one side of the root has been found to systematically suppress mycorrhization on the other side, suggesting an overlap between AON suppression of nodulation and mycorrhization (Vierheilig *et al.*, 2000; Catford *et al.*, 2003; Vierheilig, 2004; Meixner *et al.*, 2005, 2007; Sakamoto, Ogiwara and Kaji, 2013). In legumes, mutations in various genes involved in the AON pathway, both in the shoot and the root, lead to increased mycorrhizal colonization (Morandi *et al.*, 2000; Shrihari *et al.*, 2000; Zakaria Solaiman *et al.*, 2000; Meixner *et al.*, 2005; Chaulagain et al. 2024). Interestingly, similar effects are observed in non-legumes with mutations in orthologous AON pathway genes, where enhanced mycorrhization occurs. For example, a mutation in the ortholog of the legume *CLV2* gene in tomato increases both mycorrhizal colonization and arbuscule formation (Wang, Reid and Foo, 2018). In conclusion, this body of evidence suggests that multiple components of the AON pathway regulate plant associations with both rhizobia and AM fungi, highlighting a shared regulatory mechanism in controlling interactions with these beneficial microbes.

Chapter 4

**Regulation of shoot meristem and lateral root formation in *Arabidopsis thaliana***

In *Arabidopsis thaliana* (*A. thaliana*), *CLAVATA3/ESR* (*CLV3*) and *CLAVATA1* (*CLV1*) are components of a signaling pathway in the shoot devoted to regulate the maintenance of shoot apical meristem. The AtCLV3 peptides undergo modification such as arabinosylation and cleavage into functionally active small peptides, 12-13 amino acids. These small peptides function as ligands that directly bind to the ectodomain of AtCLV1 receptor promoting it to form a heterocomplex with two other interacting receptors AtCRN and AtCLV2 to regulate shoot apical meristem (Clark, Running and Meyerowitz, 1993; Kondo *et al.*, 2006; Ni and Clark, 2006; Ogawa *et al.*, 2008; Bleckmann *et al.*, 2010; Zhu *et al.*, 2010; Ni *et al.*, 2011). The *A. thaliana* CLV1 controlled pathway also regulates root meristem differentiation in response to external factors such as nitrate deficiency. For example, under nitrogen-starved conditions, the production of AtCLE3 root-mobile signaling peptide is induced, reducing lateral root formation as an adaptive strategy response. This observed reduction in lateral root through AtCLE3 requires a functional AtCLV1 receptor (Araya *et al.*, 2014; Bashyal, Gautam and Müller, 2023).

In contrast, under nitrogen starvation conditions, the dicot species *L. japonicus* and *A. thaliana* have evolved mechanisms to promote lateral number initiation through a root-specific gene called *TOO MUCH LOVE* (*LjTML*) and its homolog *HOMOLOGUE OF LEGUME TML* (*AtHOLT*), respectively. The *LjTML and AtHOLT* genes are orthologous. Taking *AtHOLT* as an example, under nitrogen starvation, *AtHOLT* is highly induced in the root, and its induction positively correlates with an increase in lateral root initiation. However, under nitrogen-sufficient conditions, the expression level of *AtHOLT* is drastically reduced, resulting in a corresponding decrease in lateral root initiation. This reduction in *AtHOLT* is systemically regulated through nitrate-responsive microRNAs (miRNA2111a and MiRNA2111b) produced in the shoot. For example, overexpression of miRNA2111 in the shoot is sufficient to systematically reduce the expression level of *AtHOLT* in

root and decrease the number of lateral roots initiated. This observation is consistent with the reduced lateral root initiation phenotype of *Atholt* mutant, indicating that *AtHOLT* is a positive regulator of lateral root initiation and is systematically regulated through miRNA2111 (Hsieh *et al.*, 2009; Pant *et al.*, 2009; Erdmann and Barciszewski, 2011; Sexauer *et al.*, 2023). In conclusion, this suggests that the systemic regulation of lateral root development is an ancient mechanism conserved in dicot species, which evolved in legumes to regulate nodule number.

**Other pathways regulating nodule number**

**COMPACT ROOT ARCHITECTURE 2 (CRA2)**

In *M. truncatula*, forward genetics screening discovered a gene called *COMPACT ROOT ARCHITECTURE 2* (*MtCRA2*), which encodes a Leucine-Rich Repeat Receptor-like Kinase that regulates nodulation systemically and root architecture locally. Mutation of *MtCRA2* leads to reduced nodule number and shorter primary root length, while increasing lateral root number (Huault *et al.*, 2014). *MtCRA2*-dependent regulation of nodulation and root architecture requires upstream functional root-induced C-TERMINALLY ENCODED PEPTIDEs (CEPs). *CEP* genes are transcriptionally upregulated in the root in response to nitrogen starvation and the absence of rhizobia, highlighting their role in enhancing plant symbiotic permissiveness (Imin *et al.*, 2013; Zhu *et al.*, 2021). For example, overexpression of *MtCEP1* or treatment with a synthetic form of MtCEP1 peptide enhances nodule formation and infection but reduces lateral root emergence (Imin *et al.*, 2013; Mohd-Radzman *et al.*, 2016). Moreover, the application of synthetic MtCEP1 peptides significantly increases hypernodulation of *Mtsunn* and *Mtskl* mutants (Mohd-Radzman *et al.*, 2016). The *MtCEP1-MtCRA2* pathway systematically modulates nodule number in the root through the regulation of miRNA2111 expression in the shoot, which in turn reduces the expression levels of *MtTML1* and *MtTML2* in the root, resulting in enhanced nodule formation (Gautrat, Laffont and Frugier, 2020; Laffont and Frugier, 2024).

Therefore, this indicates an overlap between the MtCEPs-MtCRA2 and MtCLEs-MtSUNN pathways, independently but consecutively regulating nodule number.

**References**

**Araya, T. *et al.*** (2014) 'CLE-CLAVATA1 peptide-receptor signaling module regulates the expansion of plant root systems in a nitrogen-dependent manner', *Proceedings of the National Academy of Sciences of the United States of America*, **111**(5), pp. 2029–2034.

**Bashyal, S., Gautam, C.K. and Müller, L.M.** (2023) 'CLAVATA signaling in plant–environment interactions', *Plant physiology*, p. kiad591.

**Bleckmann, A. *et al.*** (2010) 'Stem cell signaling in Arabidopsis requires CRN to localize CLV2 to the plasma membrane', *Plant physiology*, **152**(1), pp. 166–176.

**Britto, D.T. and Kronzucker, H.J.** (2002) 'NH4+ toxicity in higher plants: a critical review', *Journal of plant physiology*, **159**(6), pp. 567–584.

**Calvert, H.E. *et al.*** (1984) 'Anatomical analysis of the development and distribution of Rhizobium infections in soybean roots', *Canadian journal of botany. Journal canadien de botanique*, **62**(11), pp. 2375–2384.

**Carroll, B.J., McNeil, D.L. and Gresshoff, P.M.** (1985) 'Isolation and properties of soybean [Glycine max (L.) Merr.] mutants that nodulate in the presence of high nitrate concentrations', *Proceedings of the National Academy of Sciences of the United States of America*, **82**(12), pp. 4162–4166.

**Catford, J.-G. *et al.*** (2003) 'Suppression of arbuscular mycorrhizal colonization and nodulation in split-root systems of alfalfa after pre-inoculation and treatment with Nod factors', *Journal of experimental botany*, **54**(386), pp. 1481–1487.

**Chaulagain, D. *et al.*** *(2023) 'TML1 AND TML2 SYNERGISTICALLY REGULATE NODULATION BUT NOT ARBUSCULAR MYCORRHIZA IN MEDICAGO TRUNCATULA', bioRxiv : the preprint server for biology [Preprint]. Available at: https://doi.org/10.1101/2023.12.07.570674.*

**Chaulagain, D. *et al.*** (2024) 'Tml1 and tml2 synergistically regulate nodulation and affect arbuscular mycorrhiza in Medicago truncatula', bioRxiv.org: the preprint server for biology [Preprint]. Available at: https://doi.org/10.1101/2023.12.07.570674.

**Clark, S.E., Running, M.P. and Meyerowitz, E.M.** (1993) 'CLAVATA1, a regulator of meristem and flower development in Arabidopsis', *Development* , **119**(2), pp. 397–418.

**Crook, A.D., Schnabel, E.L. and Frugoli, J.A.** (2016) 'The systemic nodule

number regulation kinase SUNN in Medicago truncatula interacts with MtCLV2 and MtCRN', *The Plant journal: for cell and molecular biology*, **88**(1), pp. 108–119.

**Day, D.A. *et al.*** (1986) 'Growth comparisons of a supernodulating soybean (Glycine max) mutant and its wild-type parent', *Physiologia plantarum*, 68(3), pp. 375–382.

**Delves, A.C. *et al.*** (1986) 'Regulation of the soybean-Rhizobium nodule symbiosis by shoot and root factors', *Plant physiology*, **82**(2), pp. 588–590.

**Delves, A.C., Higgins, A. and Gresshoff, P.** (1992) 'Shoot apex removal does not alter autoregulation of Modulation in soybean', *Plant, cell & environment*, **15**(2), pp. 249–254.

**Doyle, J.J.** (2011) 'Phylogenetic perspectives on the origins of nodulation', *Molecular plant-microbe interactions: MPMI*, **24**(11), pp. 1289–1295.

**Erdmann, V.A. and Barciszewski, J.** (2011) *Non Coding RNAs in Plants*. Springer Science & Business Media.

**Gautrat, P. *et al.*** (2019) 'Unraveling new molecular players involved in the autoregulation of nodulation in Medicago truncatula', *Journal of experimental botany*, **70**(4), pp. 1407–1417.

**Gautrat, P., Laffont, C. and Frugier, F.** (2020) 'Compact Root Architecture 2 Promotes Root Competence for Nodulation through the miR2111 Systemic Effector', *Current biology: CB*, **30**(7), pp. 1339–1345.e3.

**Geurts, R., Lillo, A. and Bisseling, T**. (2012) 'Exploiting an ancient signaling machinery to enjoy a nitrogen fixing symbiosis', *Current opinion in plant biology*, **15**(4), pp. 438–443.

**Hastwell, A.H. *et al.*** (2017) 'CLE peptide-encoding gene families in Medicago truncatula and Lotus japonicus, compared with those of soybean, common bean and Arabidopsis', *Scientific reports*, **7**(1), p. 9384.

**Hsieh, T.-F. *et al.*** (2009) 'Genome-wide demethylation of Arabidopsis endosperm', *Science*, **324**(5933), pp. 1451–1454.

**Huault, E. *et al.*** (2014) 'Local and systemic regulation of plant root system architecture and symbiotic nodulation by a receptor-like kinase', *PLoS genetics*, **10**(12), p. e1004891.

**Imin, N. *et al.*** (2013) 'The peptide-encoding CEP1 gene modulates lateral root and nodule numbers in Medicago truncatula', *Journal of experimental botany*, **64**(17), pp. 5395–5409.

**Ishikawa, K. *et al.*** (2008) 'Isolation of a novel root-determined hypernodulation mutant rdh1 of Lotus japonicus', *Soil science and plant nutrition*, **54**(2), pp. 259–263.

**Jeudy, C. *et al.*** (2010) 'Adaptation of Medicago truncatula to nitrogen limitation is modulated via local and systemic nodule developmental responses', *The New phytologist*, **185**(3), pp. 817–828.

**Kassaw, T. *et al.*** (2017) 'ROOT DETERMINED NODULATION1 Is Required for M. truncatula CLE12, But Not CLE13, Peptide Signaling through the SUNN Receptor Kinase', *Plant physiology*, **174**(4), pp. 2445–2456.

**Kassaw, T.K. and Frugoli, J.A.** (2012) 'Simple and efficient methods to generate split roots and grafted plants useful for long-distance signaling studies in Medicago truncatula and other small plants', *Plant methods*, **8**(1), p. 38.

**Kondo, T. *et al.*** (2006) 'A plant peptide encoded by CLV3 identified by in situ MALDI-TOF MS analysis', *Science*, **313**(5788), pp. 845–848.

**Kosslak, R.M. and Bohlool, B.B.** (1984) 'Suppression of nodule development of one side of a split-root system of soybeans caused by prior inoculation of the other side', *Plant physiology*, **75**(1), pp. 125–130.

**Krusell, L. *et al.*** (2002) 'Shoot control of root development and nodulation is mediated by a receptor-like kinase', *Nature*, **420**(6914), pp. 422–426.

**Krusell, L. *et al.*** (2011) 'The Clavata2 genes of pea and Lotus japonicus affect autoregulation of nodulation', *The Plant journal: for cell and molecular biology*, **65**(6), pp. 861–871.

**Laffont, C. and Frugier, F.** (2024) 'Rhizobium symbiotic efficiency meets CEP signaling peptides', *The New phytologist*, **241**(1), pp. 24–27.

**Li, D., Kinkema, M. and Gresshoff, P.M.** (2009) 'Autoregulation of nodulation (AON) in Pisum sativum (pea) involves signaling events associated with both nodule primordia development and nitrogen fixation', *Journal of plant physiology*, **166**(9), pp. 955–967.

**Lim, C.W., Lee, Y.W. and Hwang, C.H.** (2011) 'Soybean nodule-enhanced CLE peptides in roots act as signals in GmNARK-mediated nodulation suppression', *Plant & cell physiology*, **52**(9), pp. 1613–1627.

**Liu, Y. *et al.*** (2013) 'Ammonium inhibits primary root growth by reducing the length of meristem and elongation zone and decreasing

elemental expansion rate in the root apex in Arabidopsis thaliana', *PloS one*, **8**(4), p. e61031.

**Liu, Y. and von Wirén, N.** (2017) 'Ammonium as a signal for physiological and morphological responses in plants', *Journal of experimental botany*, **68**(10), pp. 2581–2592.

**Magori, S. *et al.*** (2009) 'TOO MUCH LOVE, a Root Regulator Associated with the Long-Distance Control of Nodulation in Lotus japonicus', *Molecular plant-microbe interactions: MPMI*, **22**(3), pp. 259–268.

**Mathews, A., Carroll, B.J. and Gresshoff, P.M.** (1989) 'Development ofBradyrhizobium infections in supernodulating and non-nodulating mutants of soybean (Glycine max [L.] Merrill)', *Protoplasma*, **150**(1), pp. 40–47.

**Meixner, C. *et al.*** (2005) 'Lack of mycorrhizal autoregulation and phytohormonal changes in the supernodulating soybean mutant nts1007', *Planta*, **222**(4), pp. 709–715.

**Meixner, C. *et al.*** (2007) 'Two defined alleles of the LRR receptor kinase *GmNARK* in supernodulating soybean govern differing autoregulation of mycorrhization', *Physiologia plantarum*, **130**(2), pp. 261–270.

**Minchin, F.R. *et al.*** (1981) 'Carbon and nitrogen nutrition of nodulated roots of grain legumes', *Plant, cell & environment*, **4**(1), pp. 5–26.

**Miyazawa, H. *et al.*** (2010) 'The receptor-like kinase KLAVIER mediates systemic regulation of nodulation and non-symbiotic shoot development in Lotus japonicus', *Development* , **137**(24), pp. 4317–4325.

**Mohd-Radzman, N.A. *et al.*** (2016) 'Different Pathways Act Downstream of the CEP Peptide Receptor CRA2 to Regulate Lateral Root and Nodule Development', *Plant physiology*, **171**(4), pp. 2536–2548.

**Morandi, D. *et al.*** (2000) 'Influence of genes determining supernodulation on root colonization by the mycorrhizal fungus Glomus mosseae in Pisum sativum and Medicago truncatula mutants', *Mycorrhiza*, **10**(1), pp. 37–42.

**Moreau, C., Gautrat, P. and Frugier, F.** (2021) 'Nitrate-induced CLE35 signaling peptides inhibit nodulation through the SUNN receptor and miR2111 repression', *Plant physiology*, **185**(3), pp. 1216–1228.

**Mortier, V. *et al.*** (2010) 'CLE Peptides Control Medicago truncatula Nodulation Locally and Systemically', *Plant physiology*, **153**(1), pp. 222–237.

**Ni, J. *et al.*** (2011) 'Characterization of a CLE processing activity', *Plant molecular biology*, **75**(1-2), pp. 67–75.

**Ni, J. and Clark, S.E.** (2006) 'Evidence for functional conservation, sufficiency, and proteolytic processing of the CLAVATA3 CLE domain', *Plant physiology*, **140**(2), pp. 726–733.

**Nishimura, R. *et al.*** (2002) 'HAR1 mediates systemic regulation of symbiotic organ development', *Nature*, **420**(6914), pp. 426–429.

**Nontachaiyapoom, S. *et al.*** (2007) 'Promoters of orthologous Glycine max and Lotus japonicus nodulation autoregulation genes interchangeably drive phloem-specific expression in transgenic plants', *Molecular plant-microbe interactions: MPMI*, **20**(7), pp. 769–780.

**Notaguchi, M. and Okamoto, S.** (2015) 'Dynamics of long-distance signaling via plant vascular tissues', *Frontiers in plant science*, **6**, p. 161.

**Nowak, S., Schnabel, E. and Frugoli, J.** (2019) 'The Medicago truncatula CLAVATA3-LIKE CLE12/13 signaling peptides regulate nodule number depending on the CORYNE but not the COMPACT ROOT ARCHITECTURE2 receptor', *Plant signaling & behavior*, **14**(6), p. 1598730.

**Ogawa, M. *et al.*** (2008) 'Arabidopsis CLV3 peptide directly binds CLV1 ectodomain', *Science*, **319**(5861), p. 294.

**Oka-Kira, E. *et al.*** (2005) 'klavier (klv), a novel hypernodulation mutant of Lotus japonicus affected in vascular tissue organization and floral induction', *The Plant journal: for cell and molecular biology*, **44**(3), pp. 505–515.

**Okamoto, S. *et al.*** (2009) 'Nod factor/nitrate-induced CLE genes that drive HAR1-mediated systemic regulation of nodulation', *Plant & cell physiology*, **50**(1), pp. 67–77.

**Okamoto, S. *et al.*** (2013) 'Root-derived CLE glycopeptides control nodulation by direct binding to HAR1 receptor kinase', *Nature communications*, **4**, p. 2191.

**Okamoto, S. and Kawaguchi, M.** (2015) 'Shoot HAR1 mediates nitrate inhibition of nodulation in Lotus japonicus', *Plant signaling & behavior*, **10**(5), p. e1000138.

**Okuma, N. *et al.*** (2020) 'MIR2111-5 locus and shoot-accumulated mature miR2111 systemically enhance nodulation depending on HAR1 in Lotus japonicus', *Nature communications*, **11**(1), p. 5192.

**Olsson, J.E. *et al.*** (1989) 'Lack of Systemic Suppression of Nodulation in Split Root Systems of Supernodulating Soybean (Glycine max [L.] Merr.) Mutants', *Plant physiology*, **90**(4), pp. 1347–1352.

**Oono, R. and Denison, R.F.** (2010) 'Comparing symbiotic efficiency between swollen versus nonswollen rhizobial bacteroids', *Plant physiology*, **154**(3), pp. 1541–1548.

**Pant, B.D. *et al.*** (2009) 'Identification of nutrient-responsive Arabidopsis and rapeseed microRNAs by comprehensive real-time polymerase chain reaction profiling and small RNA sequencing', Plant physiology, **150**(3), pp. 1541–1555.

**Pan, W.L. *et al.*** (2016) 'Ammonia/ammonium toxicity root symptoms induced by inorganic and organic fertilizers and placement', *Agronomy journal*, **108**(6), pp. 2485–2492.

**Penmetsa, R.V. *et al.*** (2003) 'Dual genetic pathways controlling nodule number in Medicago truncatula', *Plant physiology*, **131**(3), pp. 998–1008.

**Reid, D.E. *et al.*** (2011) 'Molecular mechanisms controlling legume autoregulation of nodulation', *Annals of botany*, **108**(5), pp. 789–795.

**Reid, D.E., Ferguson, B.J. and Gresshoff, P.M.** (2011) 'Inoculation- and nitrate-induced CLE peptides of soybean control NARK-dependent nodule formation', *Molecular plant-microbe interactions: MPMI*, **24**(5), pp. 606–618.

**Sagan, M. *et al.*** (1995) 'Selection of nodulation and mycorrhizal mutants in the model plant Medicago truncatula (Gaertn.) after γ-ray mutagenesis', *Plant science: an international journal of experimental plant biology*, **111**(1), pp. 63–71.

**Sagan, M. and Duc, G.** (1996) 'Sym28 and sym29, two new genes involved in regulation of nodulation in pea (Pisum sativum L.)', *Symbiosis* , **20**, pp. 229–245.

**Sakamoto, K., Ogiwara, N. and Kaji, T.** (2013) 'Involvement of autoregulation in the interaction between rhizobial nodulation and AM fungal colonization in soybean roots', *Biology and fertility of soils*, **49**(8), pp. 1141–1152.

**Sasaki, T. *et al.*** (2014) 'Shoot-derived cytokinins systemically regulate root nodulation', *Nature communications*, **5**, p. 4983.

**Schnabel, E. *et al.*** (2005) 'The Medicago truncatula SUNN gene encodes a CLV1-like leucine-rich repeat receptor kinase that regulates nodule

number and root length', *Plant molecular biology*, **58**(6), pp. 809–822.

**Schnabel, E. *et al.*** (2010) 'The lss supernodulation mutant of Medicago truncatula reduces expression of the SUNN gene', *Plant physiology*, **154**(3), pp. 1390–1402.

**Schnabel, E.L. *et al.*** (2023) 'A Medicago truncatula Autoregulation of Nodulation Mutant Transcriptome Analysis Reveals Disruption of the SUNN Pathway Causes Constitutive Expression Changes in Some Genes, but Overall Response to Rhizobia Resembles Wild-Type, Including Induction of TML1 and TML2', *Current issues in molecular biology*, **45**(6), pp. 4612–4631.

**Schuize, J., Adgo, E. and Merbach, W.** (1999) 'Carbon Costs Associated with $N_2$ Fixation in *Vicia faba* L and *Pisum sativum* 1. over a 14-Day Period', *Plant biology* , **1**(6), pp. 625–631.

**Searle, I.R. *et al.*** (2003) 'Long-distance signaling in nodulation directed by a CLAVATA1-like receptor kinase', *Science*, 299(5603), pp. 109–112.

**Sexauer, M. *et al.*** (2023) 'A micro RNA mediates shoot control of root branching', *Nature communications*, **14**(1), p. 8083.

**Sheng, C. and Harper, J.E.** (1997) 'Shoot versus Root Signal Involvement in Nodulation and Vegetative Growth in Wild-Type and Hypernodulating Soybean Genotypes', *Plant physiology*, **113**(3), pp. 825–831.

**Shrihari, P.C. *et al.*** (2000) 'Interaction between supernodulating or non-nodulating mutants of soybean and two arbuscular mycorrhizal fungi', *Mycorrhiza*, **10**(3), pp. 101–106.

**Smith, S.E. and Read, D.J.** (2008) 'Mycorrhizal Symbiosis, 3rd editio Edn'. Cambridge, MA: Academic Press. doi.

**Suzuki, A. *et al.*** (2008) 'Split-root study of autoregulation of nodulation in the model legume Lotus japonicus', *Journal of plant research*, **121**(2), pp. 245–249.

**Szczyglowski, K. *et al.*** (1998) 'Har1, an Incomplete Dominant Mutation in Lotus japonicus, Confers Aberrant Root Development and an Unusual Hypernodulation Response (HNR)', in C. Elmerich, A. Kondorosi, and W.E. Newton (eds) *Biological Nitrogen Fixation for the 21st Century: Proceedings of the 11th International Congress on Nitrogen Fixation, Institut Pasteur, Paris, France, July 20–25, 1997*. Dordrecht: Springer Netherlands, pp. 305–306.

**Szczyglowski, K. *et al.*** (1998) 'Nodule Organogenesis and Symbiotic

Mutants of the Model Legume Lotus japonicus', *Molecular plant-microbe interactions: MPMI*, **11**(7), pp. 684–697.

**Takahara, M. *et al.*** (2013) 'TOO MUCH LOVE, a Novel Kelch Repeat-Containing F-box Protein, Functions in the Long-Distance Regulation of the Legume–Rhizobium Symbiosis', *Plant & cell physiology*, **54**(4), pp. 433–447.

**Tirichine, L. *et al.*** (2007) 'A gain-of-function mutation in a cytokinin receptor triggers spontaneous root nodule organogenesis', *Science*, **315**(5808), pp. 104–107.

**Tsikou, D. *et al.*** (2018) 'Systemic control of legume susceptibility to rhizobial infection by a mobile microRNA', *Science*, **362**(6411), pp. 233–236.

**Vierheilig, H. *et al.*** (2000) 'Systemic suppression of mycorrhizal colonization of barley roots already colonized by AM fungi', *Soil biology & biochemistry*, **32**(5), pp. 589–595.

**Vierheilig, H.** (2004) 'Further root colonization by arbuscular mycorrhizal fungi in already mycorrhizal plants is suppressed after a critical level of root colonization', *Journal of plant physiology*, **161**(3), pp. 339–341.

**Wang, C., Reid, J.B. and Foo, E.** (2018) 'The Art of Self-Control - Autoregulation of Plant-Microbe Symbioses', *Frontiers in plant science*, **9**, p. 988.

**Wopereis, J. *et al.*** (2000) 'Short root mutant of Lotus japonicus with a dramatically altered symbiotic phenotype', *The Plant journal: for cell and molecular biology*, **23**(1), pp. 97–114.

**Yoro, E. *et al.*** (2019) 'PLENTY, a hydroxyproline O-arabinosyltransferase, negatively regulates root nodule symbiosis in Lotus japonicus', *Journal of experimental botany*, **70**(2), pp. 507–517.

**Yoshida, C., Funayama-Noguchi, S. and Kawaguchi, M.** (2010) 'plenty, a novel hypernodulation mutant in Lotus japonicus', *Plant & cell physiology*, **51**(9), pp. 1425–1435.

**Zakaria Solaiman, M. *et al.*** (2000) 'Characterization of Mycorrhizas Formed by Glomus sp. on Roots of Hypernodulating Mutants of Lotus japonicus', *Journal of plant research*, **113**(4), pp. 443–448.

**Zhang, M. *et al.*** (2020) 'Shoot-derived miR2111 controls legume root and nodule development', *Cold Spring Harbor Laboratory*. Available at: https://doi.org/10.1101/2020.09.02.278739.

**Zhu, F. *et al.*** (2021) 'Multigene editing reveals that MtCEP1/2/12

redundantly control lateral root and nodule number in Medicago truncatula', *Journal of experimental botany*, **72**(10), pp. 3661–3676.

**Zhu, Y. *et al.*** (2010) 'Analysis of interactions among the CLAVATA3 receptors reveals a direct interaction between CLAVATA2 and CORYNE in Arabidopsis', *The Plant journal: for cell and molecular biology*, **61**(2), pp. 223–233.

CHAPTER 5

# 5

# *Parasponia* Mutant Analysis Revealed That *CLAVATA1* Regulates Symbiotic Interactions Beyond Legumes

Sultan Alhusayni[1,2] | Luuk Rutten[1] | Titis A. K. Wardhani[1] | Rene Geurts[1]

[1]Laboratory of Molecular Biology, Department of Plant Sciences, Wageningen University, Droevendaalsesteeg 1, 6708 PB Wageningen, The Netherlands

[2]Biological Sciences Department, College of Science, King Faisal University, Al-Ahsa 31982, Saudi Arabia

*Manuscript in preparation*

**ABSTRACT**

Plant engagement in root nodule and endomycorrhizal symbioses is tightly controlled by the plant in response to nutrient availability in the soil. This control is exerted through the autoregulation of nodulation (AON) pathway, a mechanism involving both the shoot and root parts of the plant. Mutations in orthologs of this gene in *Arabidopsis*, tomato, and cereals lead to enlarged shoot apical meristems, resulting in fasciation phenotypes. In legumes, the CLV1-type receptor is a key player in the AON pathway, transducing long-distance signals to control plant symbiotic interactions in the root. Legumes carrying mutations in this gene exhibit enhanced nodulation and mycorrhization phenotypes. However, it remains elusive whether the functioning of the CLV1-type receptor in the AON pathway is a legume-specific adaptation or a generic mechanism conserved in other plant species. To investigate the conservation of the AON pathway outside of legumes, we studied *Parasponia andersonii* (Cannabaceae), which can engage in nodulation with rhizobia and endomycorrhization with arbuscular mycorrhizal fungi. We generated two independent mutants of *P. andersonii CLAVATA1* (*PanCLV1*). *Panclv1* mutants exhibited hypernodulation and increased mycorrhization phenotypes, revealing that AON is conserved in non-legume plant species and functions in controlling plant symbiotic interactions. Additionally, *Panclv1* mutants manifest enhanced primary root length and an increased lateral root number, as well as reduced internode diameter, indicating that AON in *P. andersonii* plays a role in root and shoot development. These findings extend our understanding of the symbiotic role of CLV1 receptor beyond legume species, highlighting its potential as a target for biotechnological applications aimed at enhancing symbiotic interactions in crops.

Chapter 5

# Introduction

Plant growth is highly dependent on nutrient availability, specifically phosphate and nitrogen. Under phosphate and nitrogen limitations, plants form mutualistic interaction with micro-symbionts. For example, under phosphate limitation, more than 75% of land plant species engage in endomycorrhizal symbiosis with arbuscular mycorrhizal (AM) fungi of the Glomeromycota phylum (Remy *et al.*, 1994; Smith and Read, 2008). In contrast, under nitrogen limitation, a limited number of species of four phylogenetically related orders, Fabales, Fagales, Cucurbitales, and Rosales engage in endosymbiosis with nitrogen-fixing bacteria of the genus *Frankia* or rhizobia (Soltis *et al.*, 1995; Geurts, Lillo and Bisseling, 2012; Delaux, Radhakrishnan and Oldroyd, 2015; van Velzen, Doyle and Geurts, 2019). In both endosymbioses, the plant invests a significant portion of their photosynthates to maintain the symbiotic relationship (Minchin *et al.*, 1981; Schuize, Adgo and Merbach, 1999; Smith and Read, 2008; Oono and Denison, 2010).

Under nutrient sufficiency, plants prohibit an interaction with both AM fungi and nodulating bacteria. Legume plants evolved a long-distance signaling mechanism to tightly control nodule number. This mechanism is known as the Autoregulation of Nodulation Pathway (AON) (Carroll, McNeil and Gresshoff, 1985) (discussed in detail in **chapter 4**). Among the key players in this pathway is a gene that encodes a Leucine Rich Repeat Receptor Kinase, called *NODULE AUTOREGULATION RECEPTOR KINASE* (*GmNARK*) in soybean (*G. max*) (Delves *et al.*, 1986; Sheng and Harper, 1997; Searle *et al.*, 2003)*, PsSYM29* in pea (*Pisum sativum*) (Sagan and Duc, 1996)*, HYPERNODULATED ABERRANT ROOT FORMATION 1* (*LjHAR1*) in *Lotus japonicus* (Szczyglowski *et al.*, 1998; Wopereis *et al.*, 2000; Krusell *et al.*, 2002; Nishimura *et al.*, 2002; Okamoto *et al.*, 2009), and *SUPER NUMERIC NODULE* (*MtSUNN*) in *Medicago truncatula* (Sagan *et al.*, 1995; Penmetsa *et al.*, 2003; Schnabel *et al.*, 2005). Mutants carrying mutations in this gene in legume plant species exhibited a hypernodulation / supernodulation phenotype.

*GmNARK/PsSYM29/LjHAR1/MtSUNN* are putative orthologous to *CLAVATA1* of *Arabidopsis thaliana* (Sagan and Duc, 1996; Krusell et al., 2002; Nishimura et al., 2002; Searle et al., 2003; Schnabel et al., 2005). Mutations in this gene, as well as in its orthologs in soybean (*Glycine max*) tomato (*Solanum Lycopersicum*), cucumber (*Cucumis sativus*) rice (*Oryza stavia*), and maize (*Zea mays*) show enlarged shoot and floral meristems, resulting inflorescence and flower fasciation (Suzaki *et al.*, 2004; Bommert *et al.*, 2005; Xu *et al.*, 2015; Mirzaei *et al.*, 2017; Rodriguez-Leal *et al.*, 2019; Cheng *et al.*, 2022). Furthermore, *clv1* knockout mutants in *A. thaliana* and tomato show enhanced lateral root formation, especially at nitrogen-deprivation (Stahl *et al.*, 2013; Araya *et al.*, 2014; Xu *et al.*, 2015; Whitewoods *et al.*, 2018; Nakagami *et al.*, 2023) A similar subtle effect on root growth is observed the legume *Gmnark, Mtsunn* and *Ljhar1* mutants, indicating that the CLV1-type receptor commits a conserved non-symbiotic function in meristem formation and maintenance (Day *et al.*, 1986; Wopereis *et al.*, 2000; Schnabel *et al.*, 2012; Goh, Nicotra and Mathesius, 2019; Lagunas *et al.*, 2019).

In legumes, the GmNARK/PsSYM29/LjHAR1/MtSUNN receptor controls nodule number from the shoot. Upon rhizobium-induced signaling a set of genes encoding for *CLAVATA3* (*CLV3*) and *EMBRYO SURROUNDING REGION* (*ESR*)-*Related* (CLE) is activated (Okamoto *et al.*, 2009; Mortier *et al.*, 2010; Lim, Lee and Hwang, 2011; Sasaki *et al.*, 2014; Hastwell *et al.*, 2017; Moreau, Gautrat and Frugier, 2021; Bashyal, Gautam and Müller, 2023). These *CLE*s are called in *G. max Rhizobia-Induced CLE peptide 1* (*GmRIC1*) *and GmRIC2,* in *L. japonicus CLE-Root Signal 1 (LjCLE-RS1)*, *LjCLE-RS2*, and *LjCLE-RS3*, and in *M. truncatula MtCLE12,* and *MtCLE13*. Encoded proteins of the CLE genes undergo post-translational modifications and cleavage to 12-13 amino acids small peptides which are subsequently transported to the shoot through the xylem (Ishikawa *et al.*, 2008; Yoshida, Funayama-Noguchi and Kawaguchi, 2010; Kassaw *et al.*, 2017; Yoro *et al.*, 2019). In the shoot, transported CLE peptides are then perceived by the leaf phloem-localized CLV1-type receptor leading to an increase in its activity to suppress the leaf phloem-

produced microRNA2111s (miRNA2111s) (Okamoto *et al.*, 2013; Tsikou *et al.*, 2018; Okuma *et al.*, 2020; Zhang *et al.*, 2020). Prior to the suppression of miRNA2111s production, miRNA2111s are produced and translocated to the root through phloem to target and degrade the messenger RNA (mRNA) of the *TOO MUCH LOVE* (*TML*) gene. *TML* encodes a Kelch Repeat F-box protein containing two nuclear localization signals. Although its functioning is not yet uncovered, TML might trigger in the degradation of an unknown positive regulator of nodulation (Magori *et al.*, 2009; Takahara *et al.*, 2013; Gautrat *et al.*, 2019; Gautrat, Laffont and Frugier, 2020; Zhang *et al.*, 2020; Chaulagain *et al.*, 2024; Schnabel *et al.*, 2023).

Besides nodulation, CLV1-type receptors in legumes also regulate arbuscular mycorrhizal (AM) symbiosis through a process known as autoregulation of mycorrhization (AOM) (Wang, Reid, and Foo, 2018). For instance, in *Medicago truncatula*, the CLE53 gene is upregulated in roots in response to AM fungi, initiating a long-distance signaling pathway that modulates the level of AM fungal colonization in the roots (Müller et al., 2019; Karlo et al., 2020). This regulatory mechanism depends on a functional MtSUNN receptor in the shoot, as overexpression of *MtCLE53* in an *Mtsunn* mutant background did not lead to a reduction in mycorrhizal colonization levels, unlike in wild-type plants (Müller et al., 2019). Additionally, proper control over mycorrhization requires the *MtTML1* and *MtTML2* genes, as *tml1 tml2* double mutants in *M. truncatula* exhibited increased levels of mycorrhizal colonization compared to the control. (Chaulagain et al., 2024).

As mycorrhization is considered ancient, predating the evolution of nodulation, symbiotic CLE-CLV1 signaling might be a conserved function of this signaling module. However, till today, this has not been tested, so it remains elusive whether CLE-CLV1 signaling was co-opted in legumes to function in a symbiotic context, or, alternatively, is more ancient. To address this question, we made use of *Parasponia andersonii* (*Parasponia*), the only non-legume able to engage in root nodule symbiosis with rhizobium bacteria

and AM endomycorrhizal symbiosis. The legume family (Fabales order) and the *Parasponia* genus (Rosales order) diverged about 110 million years ago (van Velzen, Doyle and Geurts, 2019). By exploring *Agrobacterium tumefaciens*-mediated transformation in *Parasponia*, we aimed to generate a CRISPR-Cas9 mutants of *Parasponia TOO MUCH LOVE (PanTML)* and *CLAVATA1* (*PanCLV1*), of which only the latter was successful. Phenotypic analysis revealed that the AON pathway is conserved in *Parasponia* and plays a role in regulating plant symbiotic interactions and plant development.

## Results

### Mutating *Parasponia TML* is lethal for *Parasponia* regeneration.

To investigate the symbiotic functioning of the systemic CLE-CLV1 signaling in *Parasponia*, we initially aimed to generate CRISPR-Cas9 knock mutations in two genes; *CLV1* and *TML*. Phylogenetic analysis indicates both genes represents a single gene in *Parasponia,* which we named PanCLV1 (gene ID: PanWU01x14_059040) and PanTML (gene ID: PanWU01x14_146960) **(Supplemental Figure S1)**. Analysis of available RNAseq data on a range of tissues, showed that both genes are transcriptionally active, with *PanCLV1* highest expressed in stem internodes and *PanTML* in nodule primordia and mature nodules **(Figure 1)**.

Knockout mutants of the *TML* single gene in *L. japonicus* and the duplicated gene copies *TML1* and TML*2* in *M. truncatula* resulted in hypernodulation phenotype and increased mycorrhization (Magori *et al.*, 2009; Chaulagain *et al.*, 2024). To determine if *PanTML* serves a similar function in *Parasponia,* we generated two CRISPR-Cas9 constructs aiming to create a knockout mutant of *PanTML*. Each CRISPR-Cas9 construct contains 2 guide RNAs (gRNA) targeting the 5’ end of *PanTML* single exon. Next, *P. andersonii* stems and petioles were co-cultivated with *Agrobacterium tumefaciens* containing the *PanTML* sgRNAs, CRISPR-Cas9, and NPTII selectable markers (van Zeijl *et al.*, 2018). More than six independent transformation experiments were

performed; however, all were unsuccessful. We were unable to generate transgenic calli from neither *Parasponia* stems nor petioles **(Supplemental Figure S2)**. As an alternative strategy, we employed RNA interference (RNAi) to knock-down *PanTML* mRNA in *Parasponia*. For that, we generated a hairpin RNA construct containing a 250 pb homologous sequence to *PanTML*, driven by the constitutive *Cauliflower Mosaic Virus 35S promoter* (*CaMV35S*) and NPTII and DsRed selection markers. Subsequently, the *PanTML* silencing construct was then transformed into *Parasponia* stems and petioles. Initially, no transgenic calli were formed on the *Parasponia* optimized transgenic shoot induction medium (Wardhani *et al.*, 2019), but increasing the cytokinin concentration from 1 mg/L to 3 mg/L and 6 mg/L, led to the formation of a few transgenic calli which produced a total of six transgenic *TML* RNAi lines **(Supplemental Figure S3)**.

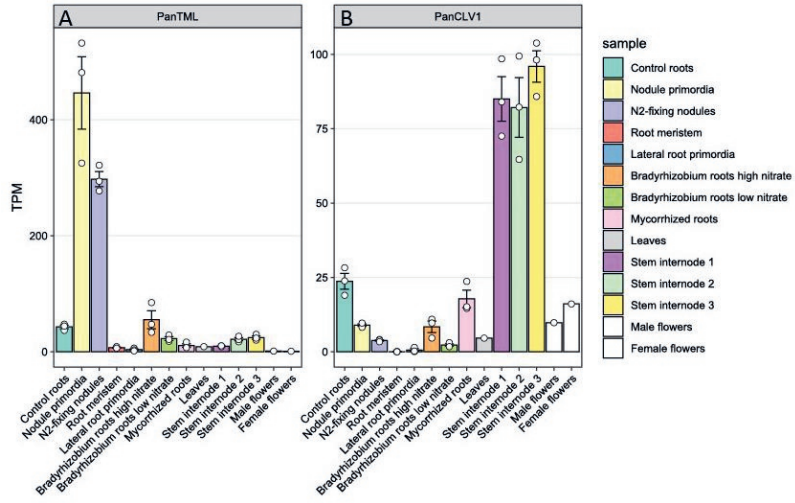



**Figure 1: Gene Expression Patterns of *Parasponia CLV1* and *TML***. Expression of *PanTML* (**A**) and *PanCLV1* (**B**) across various *Parasponia* tissues under symbiotic and non-symbiotic conditions, showing tissue-specific regulation in response to symbiosis.

To determine if the obtained six *PanTML* RNAi lines exhibit a root nodule symbiosis phenotype, and correlate the observed phenotype with the silenced level of *PanTML* gene expression, we conducted two experiments: nodulation assay and Real-Time Polymerase Chain Reaction (RT-PCR) on

root tissue of all six lines. *PanTML* RNAi lines were inoculated with *Mesorhizobium plurifarium* BOR2, and scored for nodulation 5 weeks post inoculation (5 wpi). We noticed a significant increase in the nodule number of lines 4 and 6, but not line 1, 2, 3 and 5 **(Supplemental Figure S4A)**. To determine if the observed increased nodule number is due to silenced *PanTML* gene expression, we extracted total RNA of the 6 lines, followed by complementary DNA (cDNA) synthesis, and performed RT-PCR. Our analysis revealed that neither lines 4 and 6 nor the other four lines had a statistically significant reduction of *PanTML* transcripts compared to control **(Supplemental Figure S4B)**, suggesting that they are escape lines. In conclusion, the failed knock-out and knockdown experiments of *PanTML* lead us to conclude that in *Parasponia PanTML* may represent a critical gene functioning in plant regeneration. Subsequently, knockdown or knockout mutations, might be lethal.

### *Parasponia CLAVATA1* functions as a negative regulator of nodulation

Previous studies on *LjHAR1* and *MtSUNN* revealed that these genes control the number of nodules a plant will form (Nishimura *et al.*, 2002; Searle *et al.*, 2003). This led us to question whether *PanCLV1* fulfills a similar symbiotic function in *Parasponia*. To investigate that, we generated two *Panclv1* homozygous mutants (*Panclv1-16 and Panclv1-21*) using 3 sgRNAs targeting the 5' end of the first coding exon **(Figure 2A)**. Subsequently, sequencing alignment of the target sites in *Panclv1-16* and *Panclv1-21* mutants revealed out-of-frame deletion 247 and 248 bp, respectively **(Figure 2B)**. These mutated alleles encode truncated PanCLV1 proteins of 19 and 57 amino acids, of which only the first 8 or 9 amino acids correspond to PanCLV1, covering a fragment of the signaling peptide. To determine if *PanCLV1* functions in regulating nodulation in *Parasponia*, we evaluated the nodulation phenotype of an empty vector control and *Panclv1-16 and Panclv1-21* mutants 4 weeks post-inoculation with *M. plurifarium* BOR2. The transgenic control (EV) plants had 16 nodules on average (n=16), whereas *Panclv1-16 and Panclv1-21* mutant plants exhibited a hypernodulation phenotype with 64 and 89 nodules on average (n=13), respectively **(Figure**

**3 A-C)**. We sectioned some of the nodules formed on roots of the *Panclv1* mutant plants and observed wild type phenotype **(Figure 3 D and E)**. In line with findings, we conclude that *PanCLV1* functions as a regulator of nodulation in *Parasponia*, similar as observed in legumes.

***Parasponia Panclv1* mutants show enhanced mycorrhizal colonization**

Previous studies in legumes reported the involvement of LjHAR1 and MtSUNN in AOM pathway, as mutants in these genes causes enhanced mycorrhizal colonization (Morandi *et al.*, 2000; Shrihari *et al.*, 2000; Zakaria Solaiman *et al.*, 2000; Meixner *et al.*, 2005). To investigate if *PanCLV1* plays a similar role in regulating mycorrhizal colonization, we inoculated *Panclv1-16 and Panclv1-21* mutant plantlets with 300 spores of *Rhizophagus irregularis* fungi, and analyzed their mycorrhization phenotype compared to the control six weeks post-inoculation (6wpi). We observed a colonization level of (3.3%, 2.9%, and 0.1%) for hyphae, arbuscule, and vesicle, respectively, on the transgenic EV control (n=15), whereas the colonization rate of hyphae, arbuscule, and vesicle on the *Panclv1-16* mutant (n=14) was about 45%, 41%, and 3%, on *Panclv1-21* mutant (n=15) and 68%, 63%, and 4%, respectively **(Figure 4).** In line with this finding, we conclude that *Parasponia PanCLV1* controls mycorrhization levels, most probably in an AOM pathway.

**Knockout mutant of *PanCLV1* alters shoot and root branching**

Studies on the CLV1 receptor mutants in *A. thaliana* and tomato have revealed developmental phenotypes in shoot and floral organs as well as root branching (Clark, Running and Meyerowitz, 1993; Xu *et al.*, 2015). To determine if *PanCLV1* is involved in shoot development, we grew empty vector control (n=10), *Panclv1-16* (n=10) and *Panclv1-21* (n=10) mutants in nutrient-rich potting soil (without rhizobium inoculation) for 60 days and subsequently analyzed the shoot branching phenotype as well as the diameter of internodes. First, we observed that empty vector control, *Panclv1-16 and Panclv1-21* mutants formed a similar number of axillary

buds and internodes with on average 16-17 axillary buds and 14-16 internodes **(Figure 5 A and B)**. Next, we measured the diameter of the first 12 internodes (with the 1st at the bottom and the 12th at the top of the stem). This revealed a decreasing trend in internode diameter in the *Parasponia Panclv1-16 and Panclv1-21* mutants compared to the empty vector control **(Supplemental Figure S5)**. This observation suggests that *PanCLV1* is not essential for shoot branch formation, but might play a positive role in controlling internode thickness.

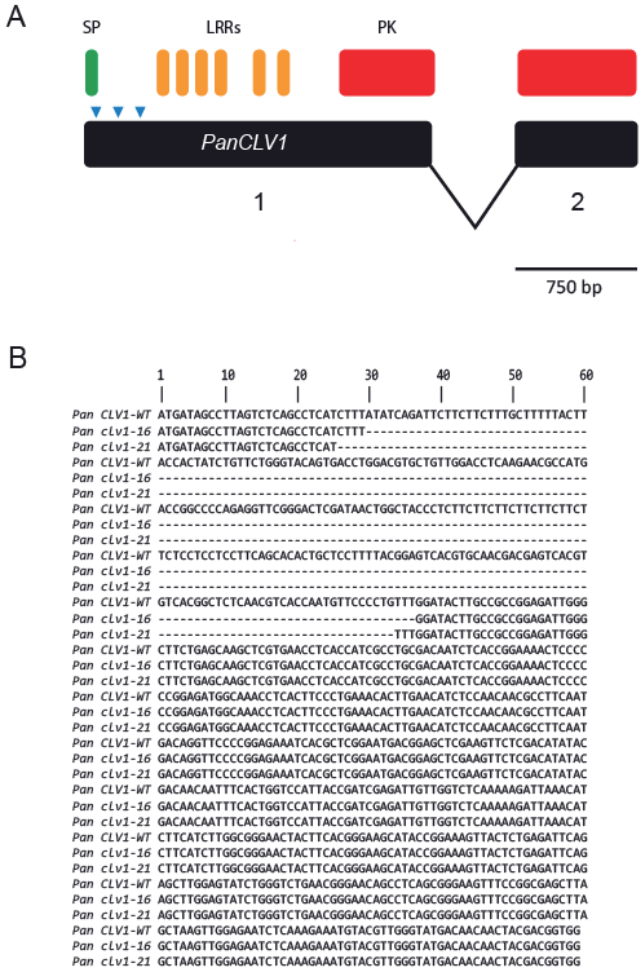



**Figure 2: CRISPR-Cas9 Induced Mutations in *PanCLV1* Gene of *Parasponia andersonii*.** (**A-B**) Structural and sequence analysis of the *PanCLV1* gene and its CRISPR-Cas9 induced mutations. (**A**) The schematic illustrates the 3553 bp *PanCLV1* gene structure, including two

exons and one intron, with the CRISPR-Cas9 guide RNA target site located at the 5' end of the first exon. (**B**) Sequence alignment of the first 719 nucleotides of the *PanCLV1* gene in wild type and two mutant lines (*Panclv1-16* and *Panclv1-21*) reveals out-of-frame deletions resulting in premature stop codons in the mutants.

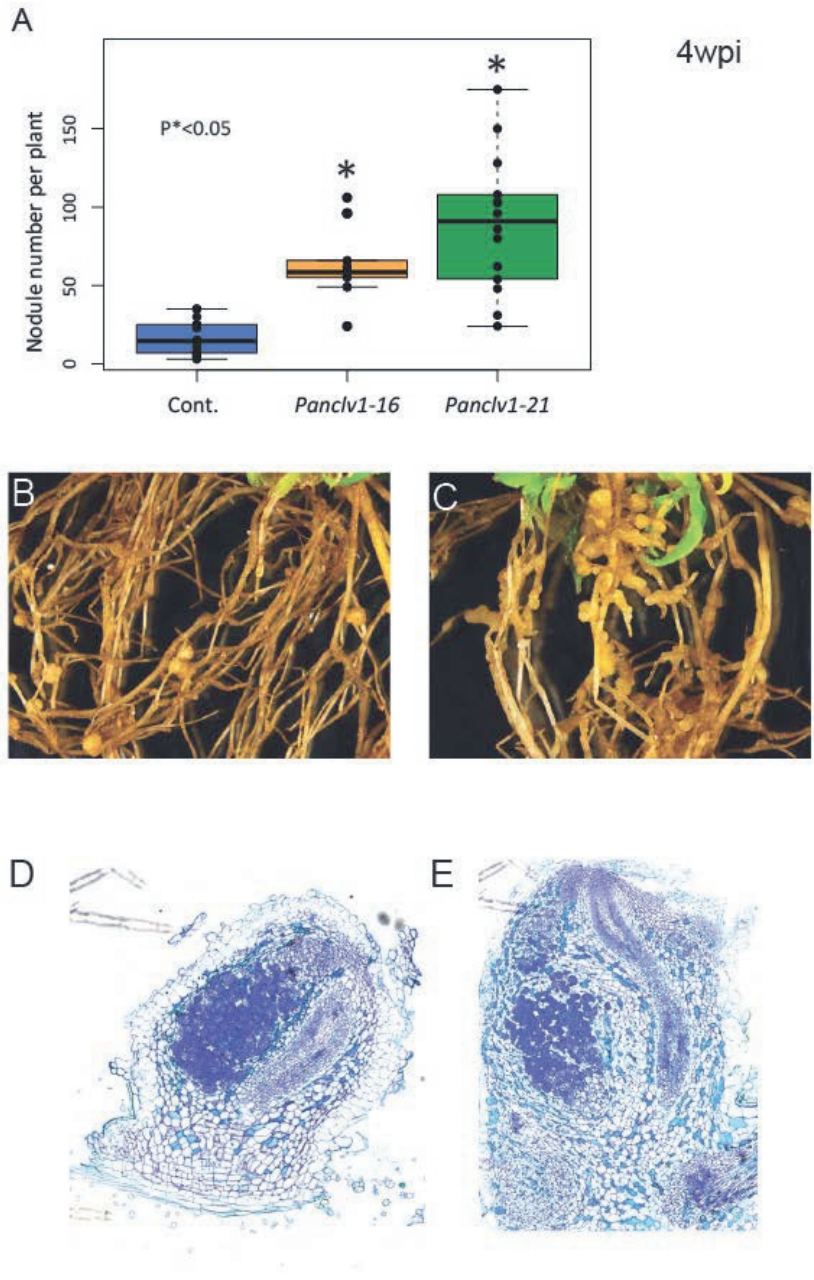


**Figure 3: Hypernodulation in *Parasponia andersonii clv1* Mutants**(**A-E**) Quantification and visual representation of nodulation in control plants (cont.) versus *Panclv1-16* and *Panclv1-21* mutants. (**A**) The graph shows a significant increase in nodule number per plant in *Panclv1-16* and *Panclv1-21* mutants at 4 weeks post-inoculation (wpi), with mutants displaying a higher nodule count compared to controls (*P < 0.05, Student's t-test). (**B-C**) Representative images of nodules formed on the roots, with (**B**) showing a typical nodule from a control plant and (**C**) displaying the hypernodulation observed in *Panclv1-16* and *Panclv1-21* mutants. (**D-E**) Cross-sections of mature nodules, comparing the structure of nodules from control plants (**D**) to those from *Panclv1* mutants (**E**). Scale bar in (**B** and **C**) is 5mM and in (**D** and **E**) is 75µM.

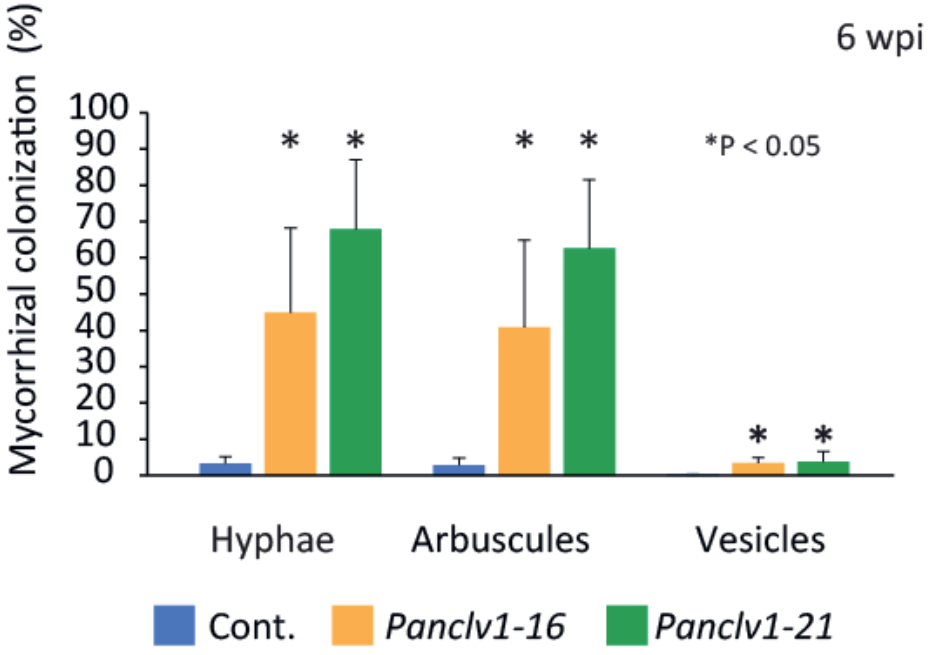



**Figure 4: Knock-out of *PanCLV1* Enhances Mycorrhizal Colonization.** Quantification of mycorrhizal colonization levels was performed in *Panclv1-16* (n=14), *Panclv1-21* (n=15) mutants, and empty vector control plants (cont.) (n=15) at 6 weeks post-inoculation (6 wpi) with 500 spores per plant. The analysis includes measurements of hyphae, arbuscules, and vesicles colonization percentages. Error bars represent standard deviation (SD), and asterisks indicate significant differences relative to control (*P < 0.05, Student's t-test).

Next, the root phenotype was characterized. *Parasponia* empty vector control plants and the *Panclv1-16 and Panclv1-21* mutant plants were grown in vitro on agar plates for 20 days without rhizobium inoculation, and subsequently, quantified their lateral root number and primary root length.

Compared to the empty vector control (n=17), *Panclv1*-16 (n=14) *and Panclv1-21* (n=12) showed an increase in lateral root number with an average of 19, 43 and 31 lateral roots, respectively **(Figure 5 C)**. In relation to primary root length *Panclv1-16* showed an increased primary root length with an average of 6 cm compared to 5 cm in both *Panclv1-21* mutant and the empty vector control **(Figure 5 D)**. In conclusion, our data suggest that *PanCLV1* negatively controls lateral root development and to some extent inhibits primary root length.

**Discussion**

The CLV1-type LRR Receptor Kinase in legumes plays a crucial role in regulating the AON/AOM pathways in the shoot, which are responsible for controlling symbiotic interactions with rhizobial bacteria and AM fungi in the roots. Legumes carrying mutations in *GmNARK/PsSYM29/LjHAR1/MtSUNN* exhibit enhanced nodulation and mycorrhization phenotypes under high or moderate levels of exogenous nitrogen and phosphate conditions, respectively (Sagan and Duc, 1996; Morandi *et al.*, 2000; Shrihari *et al.*, 2000; Nishimura *et al.*, 2002; Searle *et al.*, 2003; Schnabel *et al.*, 2005; Müller *et al.*, 2019). We investigated the conservation of this symbiotic function of the CLV1 type receptor outside the legume family using *Parasponia*, the only non-legume plant species capable of forming root nodule symbiosis with rhizobium. Through CRISPR-Cas9 mutagenesis, two independent *Parasponia Panclv1* mutants were generated. Studying these mutants revealed, for the first time, that the regulation of nodulation in a non-legume is—like in legumes—dependent on this CLV1 receptor. Therefore, we conclude that the AON pathway is not a legume-specific adaptation but rather a mechanism already present in the last shared common ancestor of legumes and *Parasponia*, which lived over 100 million years ago.

Studies on the non-symbiotic functions of the CLV1-type receptor in plant development have highlighted its significant role in regulating both shoot and root architecture. For example, mutating *CLV1* in *A. thaliana* resulted in shoot and floral meristem fasciation, as well as the expansion of lateral root

primordia (Clark, Running and Meyerowitz, 1993; Araya *et al.*, 2014). Similarly, mutations in the CLV1-type receptor in cucumber (*Cucumis sativus*) led to fasciation in the shoot, floral meristem, and internodes (Cheng *et al.*, 2022). Building on these findings, we explored the non-symbiotic role of the CLV1-type receptor in *Parasponia*. Our findings show that *PanCLV1* plays a role in *Parasponia* plant development, as it promotes internode diameter in the shoot while simultaneously suppressing primary root length and lateral root formation.

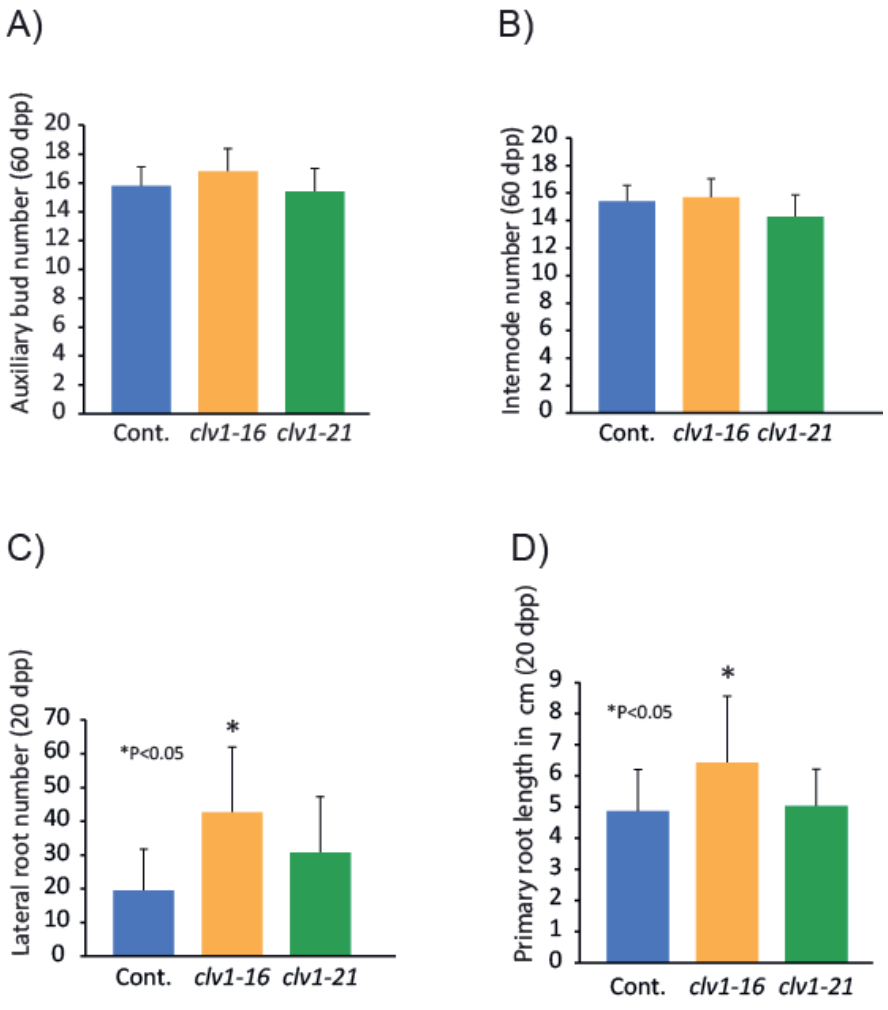


**Figure 5: Non-Symbiotic Root and Shoot Growth Traits in *Parasponia Panclv1* Mutant Lines**. Comparison of root and shoot growth traits between empty vector control (cont.) and *Panclv1* mutant lines. (**A-B**) Number of axillary buds (**A**) and internodes (**B**) as shoot growth traits in empty vector control (n=10), *Panclv1-16* (n=10), and *Panclv1-21* (n=10) mutant lines at 60 days post-planting (dpp). (**C-D**) Number of lateral roots **(C)** and primary root length (**D**) as root growth traits in empty vector control (n=17), *Panclv1-16* (n=14), and *Panclv1-21* (n=12) mutant lines in vitro grown for 20 days post-planting (dpp). Error bars represent standard deviation (SD), and asterisks indicate significant differences relative to control (*P < 0.05, Student's t-test).

We also aimed to study the function of *TML* in *Parasponia* using a mutant analysis approach, as this gene functions downstream of the CLE-CLV1 systemic signaling pathways. Genetic screening of mutagenized seed populations in *L. japonicus* led to the identification of the *tml* mutant, while reverse genetics approaches in *M. truncatula* produced RNAi-silenced and CRISPR-Cas9-induced *tml* mutants through tissue culture (Nishimura *et al.*, 2002; Gautrat *et al.*, 2019; Chaulagain *et al.*, 2024). However, our studies in *Parasponia* revealed that this approach is not universally applicable: no CRISPR-Cas9 mutants could be generated, and RNAi knockdown levels were relatively low in the transgenic lines that were regenerated. This raises the question of why *tml* mutants could not be generated in *Parasponia* via tissue culture, whereas such an approach was possible in *M. truncatula*. It suggests that *PanTML* may play a role in regulating embryogenesis in *Parasponia* and raises the question on whether the *TML* genes in legumes and *Parasponia* have functionally diverged.

The putative role of *TML* in regulating plant regeneration is supported by the expression of the *Manihot esculenta TML* gene (Manes.17G103600), which is highly expressed in friable embryogenic callus, suggesting a function in plant embryogenesis **(Supplemental Figure S6)** (Wilson *et al.*, 2017). *Parasponia* has a single *TML* gene copy that may also be expressed in embryogenic callus. *PanTML* is predicted to encode a Kelch Repeat F-box protein containing two nuclear localization signals, similar to the TML protein of *L. japonicus*, which has been shown to localize to the nucleus (Takahara *et al.*, 2013). The Kelch Repeat and F-box domains are known for

mediating protein-protein interactions and protein degradation, respectively. This suggests that in the nuclei of embryogenic cells in the callus, TML might interact with other proteins to form a complex that either enhances transcriptional activators or degrades transcriptional suppressors involved in cell differentiation during embryogenesis. Further research is needed to support this hypothesis, such as determining *PanTML* expression in *Parasponia* embryonic calli to clarify whether TML plays such a role.

Studies in *L. japonicus* revealed that cytokinins function in AON and control nodulation in a *TML* dependent manner (Takahara et al., 2013; Sasaki et al., 2014). For example, overexpression of the cytokinin biosynthesis gene isopentenyl transferase 3 (*LjIPT* 3) in the shoot reduces nodule number. This reduction is proposed to be systemic, as shoot feeding of Benzoylaminopurine (BAP), a synthetic form of cytokinins, reduces the number of nodules and lateral roots in a *TML*-dependent manner. If TML plays a broader function in cytokinin signaling, then a functional *TML* in tissue culture plant regeneration is highly relevant since artificial cell differentiation in *Parasponia* tissue culture is induced by hormones like cytokinins and auxin. To mitigate probable effect on cytokinin signaling, during our round of transformations, we tried to optimize the concentration of cytokinin in our transformation and growth medium to facilitate the formation of transgenic calli on stems and petioles transformed with CRISPR Cas9 construct targeting *TML*, but all our efforts were not successful. However, optimizing cytokinin concentration on stems and petioles transformed with an RNAi construct targeting *TML* mRNA partly worked as a few transgenic calli were formed and gave rise to six lines with a minor reduction in *TML* levels. Therefore, we concluded that knockdown of *Parasponia TML* allows some degree of regeneration, but complete knockout is detrimental to *Parasponia* plant regeneration.

## Material and Method

### Plant materials and growth condition

In this study, *Parasponia andersonii* WU1 plant or plantlet were used as previously described (Op den Camp *et al.*, 2011; Van Velzen *et al.*, 2018). Plant growth and maintenance were done as described (Van Velzen *et al.*, 2018; Wardhani *et al.*, 2019). Plant micropropagation for root and shoot architecture, nodulation, and mycorrhization assays was done following the published protocol (Van Velzen *et al.*, 2018; Wardhani *et al.*, 2019).

### Vectors and constructs

To generate *PanCLV1* mutants, a CRISPR/Cas9 mediated mutagenesis system was used. Three single-guide RNAs (sgRNA) targeting the 5' end of the *PanCLV1* first exon were designed, using the 'Find CRISPR Targets' feature implemented in Geneious software version 11.0.15 (Biomatters, New Zealand). The designed sgRNAs were amplified using *PanCLV1* specific forward primer and a generic reverse primer, using Addgene plasmid no. 46966 as a template (Nekrasov *et al.*, 2013). Golden Gate Cloning was used to generate the binary transformation vector as described (Van Velzen *et al.*, 2018; Wardhani *et al.*, 2019). To generate *PanTML* silenced mutants, an RNA interference system was used following (Limpens *et al.*, 2004). Gateway cloning was used to generate a binary transformation vector (Hartley, Temple and Brasch, 2000). *PanTML* specific forward and reverse primers were used to amplify 250 bp of the *PanTML* coding sequence. Next, the amplified sequence was cloned into pENTY/D-TOPO vector (Addgene), and subsequently, cloned into the 35S-driven expression vector pK7GWIWG2(II) (VIB-UGent Center for Plant Systems Biology, Belgium).

### Nodulation assay and histochemical analysis

*P. andersonii* control and *clv1* mutant plantlets were inoculated with *Mesorhizobium plurifarium* BOR2 (OD600=0.025). Four weeks post

inoculation, plants were taken out of pots 1 L crystal-clear polypropelene pots (OS14BOX, Duchefa Biochemie, Netherlands) and subsequently roots were washed under running water to remove adhered sand and perlite. Next, the nodule number was quantified and stored in a fixative solution (5% glutaraldehyde in 0.1 M phosphate buffer, pH 7.2). Fixed nodules were vacuumed for 30–60 minutes, and subsequently went through a dehydration process, followed by embedding in plastic polymerizing solution Technovit 7100 (Heraeus-Kulzer, Germany). Thereafter, an RJ2035 microtome was used to generate (6 µM) sections which were subsequently stained with 0.05 % Toludine Blue, and observed and imaged using a DM5500B microscope supplied with a DFC425c camera (Leica microsystems).

**Mycorrhization assays and ink staining**

*P. andersonii* plantlets were grown on river sand mixed with potting soil in polypropylene pots and watered with ½ Hoagland low phosphate (20 µM) medium. Next, plants were inoculated with 300 spores of *Rhizophagus irregularis* (Agonutrion- DAOM197198). Six weeks post inoculation, plants were removed from pots and roots were washed under running water to remove residual sand. Next, roots were treated with 10 % KOH and boiled at 95 ℃ for 30 minutes, followed by rinsing in water 6 times. Afterwards, roots were stained with a solution of (5% acetic acid and 3 % ink) at 95 ℃ for 15 minutes. Mycorrhizal colonization was assessed using the intersection method (Giovannetti and Mosse, 1980). Images of mycorrhizal roots were taken, using a DM5500B microscope equipped with a DFC425c camera (Leica microsystems).



**Shoot and root architecture phenotyping assay**

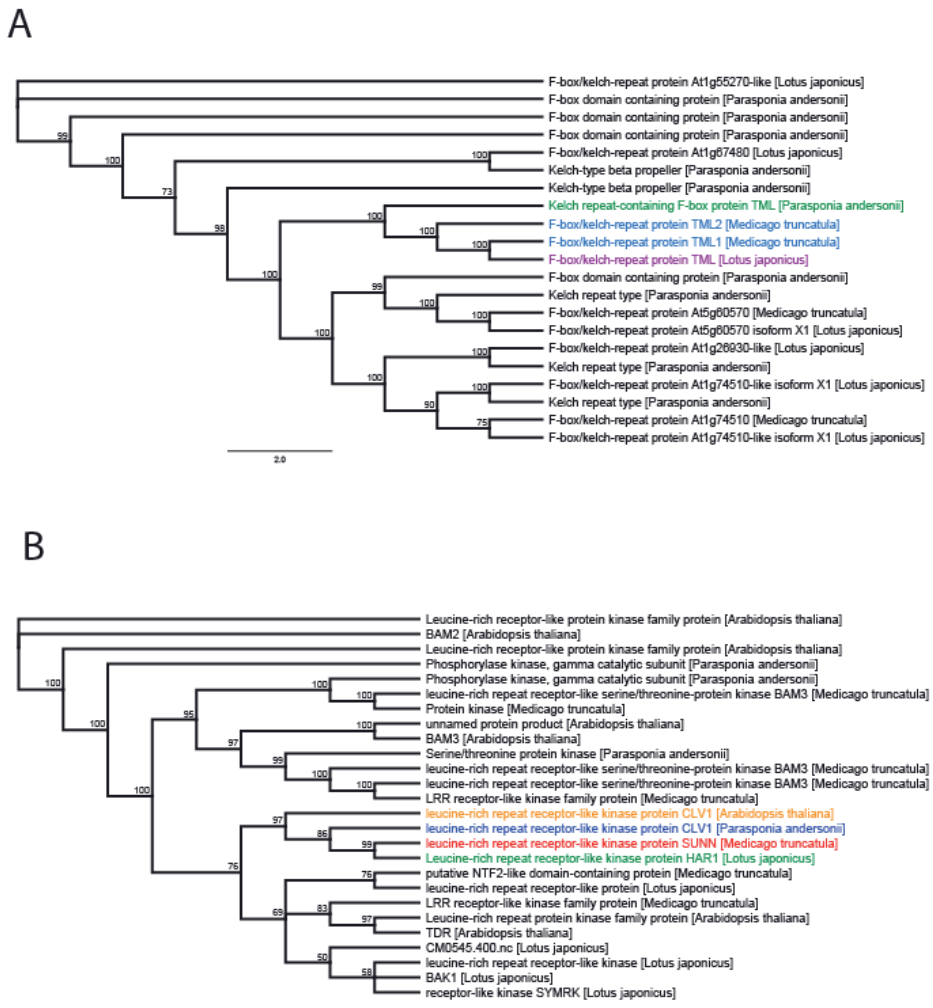
For shoot architecture, rooted plantlets of *P. andersonii* control and *clv1* mutants were grown on watered potting soil in polypropylene pots for 60 days without rhizobium inoculation. Next, plants were taken out of the pots and the formed axillary buds, and internodes were quantified. Internode diameter was measured at the center of the internode using a

digital vernier caliper. For root architecture, rooted plantlets of *P. andersonii* control and *clv1* mutants were grown between two cellophane membranes placed on squared petri dish plates supplement with 2 % agar in EKM medium [15 µM Fe-Citrate, 3 mM MES (PH 6.6), 6.6 µM $MnSO_4$, 4.1 µM $Na_2MoO_4$, 2.08 mM $MgSO_4$, 0.70 mM $Na_2SO_4$, 1.4 mM $CaCl_2$, 0.375 mM $NH_4NO_3$, 1.5 µM $ZnSO_4$, 0.88 mM $KH_2PO_4$, 1.6 µM $CuSO_4$, 2.07 mM $K_2HPO_4$, and 4 µM $H_3BO_3$] (Becking, 1983). Subsequently, plates were incubated vertically for 20 days at 85 % humidity and under a 16 /8 h day/night regime in a controlled growth chamber. Finally, primary root length and lateral root number were assessed.

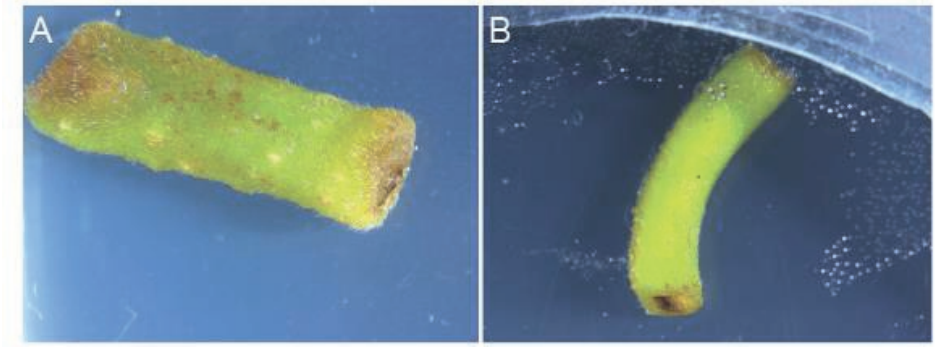
**Phylogenetic reconstruction**

Protein sequences of *L. japonicus* HAR1 (NP_001414787.1) (Nishimura *et al.*, 2002) and TML (XP_057437509.1), *M. truncatula* SUNN (AAW71475.1) (Schnabel *et al.*, 2005), TML1 (XP_013448012.1) and TML2 (XP_013451409.1 1) were obtained from NCBI. *Parasponia* CLV1 (PON73400.1) and TML (PON61319.1) protein sequences were acquired from www.parasponia.org (Van Velzen *et al.*, 2018). MUSCLE Alignment and FastTree tools implemented in Geneious software version 11.0.15 (Biomatters, New Zealand) were used to align protein sequences, and subsequently generate phylogenetic trees.

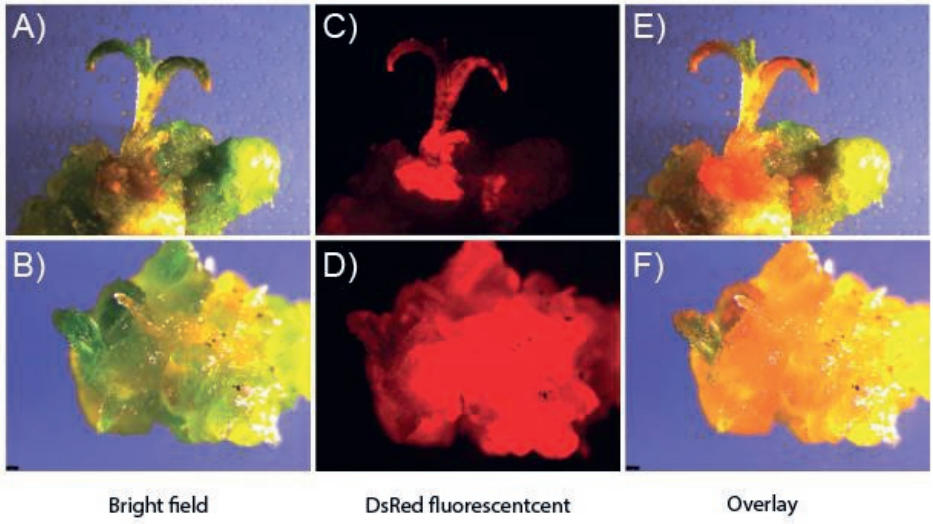
## Supplemental information



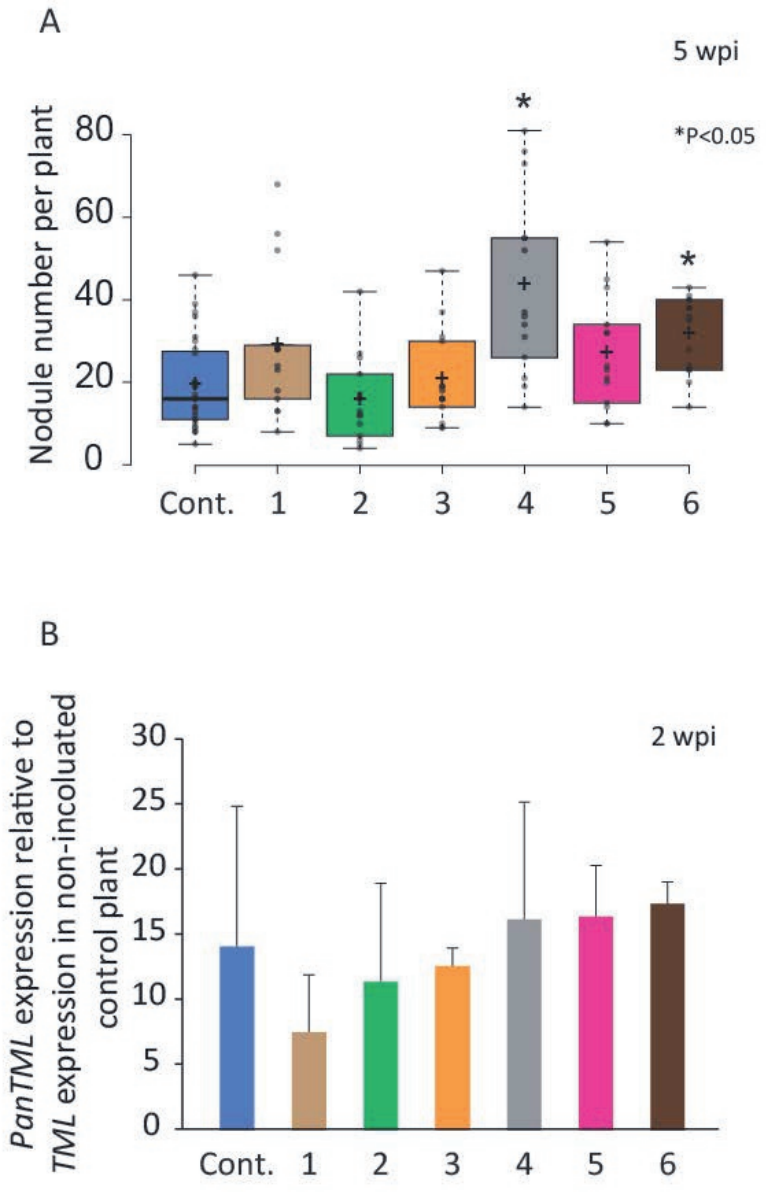

**Supplemental Figure S1: Phylogenetic Analysis of TML and CLV1 Proteins in *Parasponia andersonii* and Related Species**. Phylogenetic relationships of TML and CLV1 proteins from *Parasponia andersonii*, *Medicago truncatula*, *Lotus japonicus*, and *Arabidopsis thaliana* based on protein sequence alignments. (**A**) TML proteins: *P. andersonii* (green), *M. truncatula* (blue), and *L. japonicus* (pink). (**B**) CLV1-type proteins: *P. andersonii* (blue), *M. truncatula* (red), *L. japonicus* (green), and *A. thaliana* (orange). Bootstrap values are indicated at the nodes.




**Supplemental Figure S2: *Parasponia andersonii* Stem and Petiole Response to *PanTML* Knock-out**. (**A-B**) Images of *Parasponia andersonii* stems and petioles 1.5 months post co-cultivation with *Agrobacterium tumefaciens* carrying a plasmid with two CRISPR-Cas9 guide RNAs targeting the *PanTML* coding region. The transformation resulted in the failure to form transgenic calli.

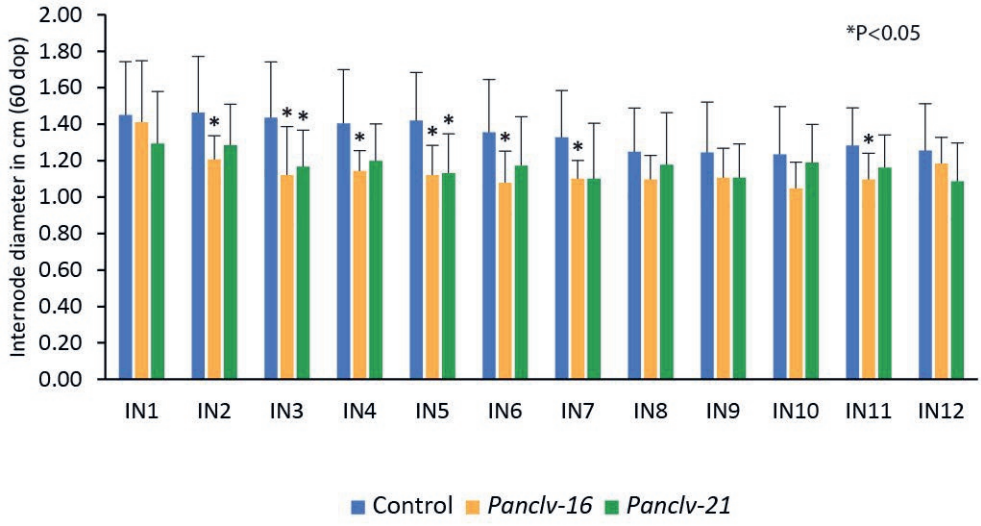



**Supplemental Figure S3: Expression of RNAi Hairpin Targeting *PanTML* mRNA in *Parasponia andersonii* Transgenic Shoots**. (**A-F**) Images of *Parasponia andersonii* transgenic shoots expressing an RNAi hairpin targeting *PanTML* mRNA along with a DsRed fluorescent marker. (**A, B**) Bright field (**A**) and DsRed fluorescence (**B**) of transgenic shoots with an empty vector (EV) RNAi construct. (**C, D**) Bright field (**C)** and DsRed fluorescence (**D**) of transgenic shoots with the *PanTML* RNAi construct. (**E, F**) Overlay images showing the EV RNAi construct (**E**) and the *PanTML* RNAi construct (**F**).

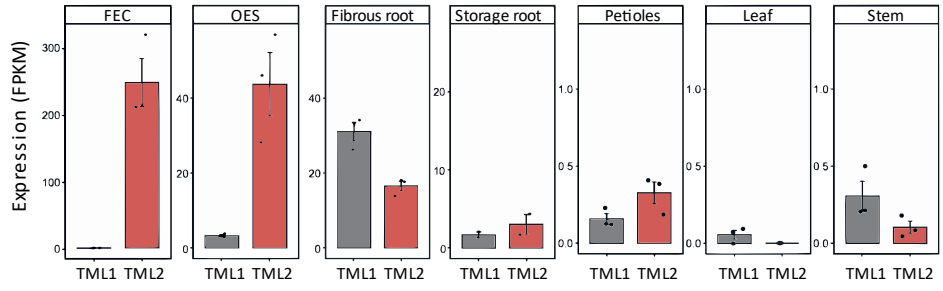



**Supplemental Figure S4: Quantification of Nodule Number and *PanTML* Gene Expression in RNAi Lines**. (**A-B**) Analysis of six transgenic RNAi lines of *PanTML* compared to the empty vector control. (**A**) Nodule number per plant at 5 weeks post-inoculation (wpi) with *Mesorhizobium plurifarium* BOR2 (OD600 = 0.025). (**B**) Relative expression of the *PanTML* gene in control and RNAi lines at 2 weeks post-inoculation, with expression levels

normalized to *PanTML* expression in non-inoculated control plants (*P < 0.05, Student's t-test). (n=3)




**Supplemental Figure S5: Internode Diameter in Mutant Lines of *Parasponia* CLV1**. Comparison of internode diameter in empty vector control and *Panclv1-16* and *Panclv1-21* mutant lines at 60 days post-inoculation (n=10). The data illustrate the impact of CLV1 mutation on shoot development in *Parasponia* (*P < 0.05, Student's t-test).




**Supplemental Figure S6: Gene Expression Patterns of Cassava (Manihot esculenta) *TML1* and *TML2***. Expression levels of *TML1* and *TML2* in various tissues of Cassava (accession 204), including friable embryogenic callus (FEC), organized embryogenic structure (OES), fibrous root, storage root, petioles, leaf, and stem. Expression is quantified as Fragments Per Kilobase of transcript per Million mapped reads (FPKM) based on RNA-seq data published by Wilson et al. (2017).

**References**

**Araya, T. *et al.*** (2014) 'CLE-CLAVATA1 peptide-receptor signaling module regulates the expansion of plant root systems in a nitrogen-dependent manner', *Proceedings of the National Academy of Sciences*, **111**(5), pp. 2029–2034.

**Bashyal, S., Gautam, C.K. and Müller, L.M.** (2023) 'CLAVATA signaling in plant–environment interactions', *Plant physiology*, p. kiad591.

**Becking, J.H.** (1983) 'TheParasponia parviflora—Rhizobium symbiosis. Host specificity, growth and nitrogen fixation under various conditions', *Plant and soil*, **75**(3), pp. 309–342.

**Bommert, P. *et al.*** (2005) 'thick tassel dwarf1 encodes a putative maize ortholog of the Arabidopsis CLAVATA1 leucine-rich repeat receptor-like kinase', *Development (Cambridge, England)*, **132**(6), pp. 1235–1245.

**Carroll, B.J., McNeil, D.L. and Gresshoff, P.M.** (1985) 'Isolation and properties of soybean [Glycine max (L.) Merr.] mutants that nodulate in the presence of high nitrate concentrations', *Proceedings of the National Academy of Sciences of the United States of America*, **82**(12), pp. 4162–4166.

**Chaulagain, D. *et al.*** (2024) 'Tml1 and tml2 synergistically regulate nodulation and affect arbuscular mycorrhiza in Medicago truncatula', bioRxiv.org: the preprint server for biology [Preprint]. Available at: https://doi.org/10.1101/2023.12.07.570674.

**Cheng, F. *et al.*** (2022) 'A SNP mutation in the CsCLAVATA1 leads to pleiotropic variation in plant architecture and fruit morphogenesis in cucumber (Cucumis sativus L.)', *Plant science: an international journal of experimental plant biology*, **323**(111397), p. 111397.

**Clark, S.E., Running, M.P. and Meyerowitz, E.M.** (1993) 'CLAVATA1, a regulator of meristem and flower development in Arabidopsis', *Development* , **119**(2), pp. 397–418.

**Day, D.A. *et al.*** (1986) 'Growth comparisons of a supernodulating soybean (Glycine max) mutant and its wild-type parent', *Physiologia plantarum*, **68**(3), pp. 375–382.

**Delaux, P.-M., Radhakrishnan, G. and Oldroyd, G.** (2015) 'Tracing the evolutionary path to nitrogen-fixing crops', *Current opinion in plant biology*, **26**, pp. 95–99.

**Delves, A.C. *et al.*** (1986) 'Regulation of the soybean-Rhizobium nodule

symbiosis by shoot and root factors', *Plant physiology*, 82(2), pp. 588–590.

**Gautrat, P. *et al.*** (2019) 'Unraveling new molecular players involved in the autoregulation of nodulation in Medicago truncatula', *Journal of experimental botany*, **70**(4), pp. 1407–1417.

**Gautrat, P., Laffont, C. and Frugier, F.** (2020) 'Compact Root Architecture 2 Promotes Root Competence for Nodulation through the miR2111 Systemic Effector', *Current biology: CB*, **30**(7), pp. 1339–1345.e3.

**Geurts, R., Lillo, A. and Bisseling, T.** (2012) 'Exploiting an ancient signalling machinery to enjoy a nitrogen fixing symbiosis', *Current opinion in plant biology*, **15**(4), pp. 438–443.

**Giovannetti, M. and Mosse, B.** (1980) 'An evaluation of techniques for measuring vesicular arbuscular mycorrhizal infection in roots', *The New phytologist*, **84**(3), pp. 489–500.

**Goh, C.-H., Nicotra, A.B. and Mathesius, U.** (2019) 'Genes controlling legume nodule numbers affect phenotypic plasticity responses to nitrogen in the presence and absence of rhizobia: Autoregulation genes affect root architecture', *Plant, cell & environment*, **42**(5), pp. 1747–1757.

**Hartley, J.L., Temple, G.F. and Brasch, M.A.** (2000) 'DNA cloning using in vitro site-specific recombination', *Genome research*, **10**(11), pp. 1788–1795.

**Hastwell, A.H. *et al.*** (2017) 'CLE peptide-encoding gene families in Medicago truncatula and Lotus japonicus, compared with those of soybean, common bean and Arabidopsis', *Scientific reports*, **7**(1), p. 9384.

**Ishikawa, K. *et al.*** (2008) 'Isolation of a novel root-determined hypernodulation mutant rdh1 of Lotus japonicus', *Soil science and plant nutrition*, **54**(2), pp. 259–263.

**Karlo, M. *et al.*** (2020) 'The CLE53–SUNN genetic pathway negatively regulates arbuscular mycorrhiza root colonization in Medicago truncatula', *Journal of experimental botany*, **71**(16), pp. 4972–4984.

**Kassaw, T. *et al.*** (2017) 'ROOT DETERMINED NODULATION1 Is Required for M. truncatula CLE12, But Not CLE13, Peptide Signaling through the SUNN Receptor Kinase', *Plant physiology*, 174(4), pp. 2445–2456.

**Krusell, L. *et al.*** (2002) 'Shoot control of root development and nodulation is mediated by a receptor-like kinase', *Nature*, **420**(6914), pp. 422–426.

**Lagunas, B. *et al.*** (2019) 'Regulation of resource partitioning coordinates nitrogen and rhizobia responses and autoregulation of nodulation in Medicago truncatula', *Molecular plant*, **12**(6), pp. 833–846.

**Lim, C.W., Lee, Y.W. and Hwang, C.H.** (2011) 'Soybean nodule-enhanced CLE peptides in roots act as signals in GmNARK-mediated nodulation suppression', *Plant & cell physiology*, **52**(9), pp. 1613–1627.

**Limpens, E. *et al.*** (2004) 'RNA interference in Agrobacterium rhizogenes-transformed roots of Arabidopsis and Medicago truncatula', *Journal of experimental botany*, **55**(399), pp. 983–992.

**Magori, S. *et al.*** (2009) 'TOO MUCH LOVE, a Root Regulator Associated with the Long-Distance Control of Nodulation in Lotus japonicus', *Molecular plant-microbe interactions: MPMI*, **22**(3), pp. 259–268.

**Meixner, C. *et al.*** (2005) 'Lack of mycorrhizal autoregulation and phytohormonal changes in the supernodulating soybean mutant nts1007', *Planta*, **222**(4), pp. 709–715.

**Minchin, F.R. *et al.*** (1981) 'Carbon and nitrogen nutrition of nodulated roots of grain legumes', *Plant, cell & environment*, **4**(1), pp. 5–26.

**Mirzaei, S. *et al.*** (2017) 'Neodiversification of homeologous CLAVATA1-like receptor kinase genes in soybean leads to distinct developmental outcomes', *Scientific reports*, **7**(1), p. 8878.

**Morandi, D. *et al.*** (2000) 'Influence of genes determining supernodulation on root colonization by the mycorrhizal fungus Glomus mosseae in Pisum sativum and Medicago truncatula mutants', *Mycorrhiza*, **10**(1), pp. 37–42.

**Moreau, C., Gautrat, P. and Frugier, F.** (2021) 'Nitrate-induced CLE35 signaling peptides inhibit nodulation through the SUNN receptor and miR2111 repression', *Plant physiology*, **185**(3), pp. 1216–1228.

**Mortier, V. *et al.*** (2010) 'CLE Peptides Control Medicago truncatula Nodulation Locally and Systemically', *Plant physiology*, **153**(1), pp. 222–237.

**Müller, L.M. *et al.*** (2019) 'A CLE–SUNN module regulates strigolactone content and fungal colonization in arbuscular mycorrhiza', *Nature Plants*, **5**(9), pp. 933–939.

**Nakagami, S. *et al.*** (2023) 'CLE3 and its homologs share overlapping functions in the modulation of lateral root formation through CLV1 and BAM1 in Arabidopsis thaliana', *The Plant journal: for cell and molecular biology*, **113**(6), pp. 1176–1191.

**Nekrasov, V. *et al.*** (2013) 'Targeted mutagenesis in the model plant Nicotiana benthamiana using Cas9 RNA-guided endonuclease', *Nature biotechnology*, **31**(8), pp. 691–693.

**Nishimura, R. *et al.*** (2002) 'HAR1 mediates systemic regulation of symbiotic organ development', *Nature*, **420**(6914), pp. 426–429.

**Okamoto, S. *et al.*** (2009) 'Nod factor/nitrate-induced CLE genes that drive HAR1-mediated systemic regulation of nodulation', *Plant & cell physiology*, **50**(1), pp. 67–77.

**Okamoto, S. *et al.*** (2013) 'Root-derived CLE glycopeptides control nodulation by direct binding to HAR1 receptor kinase', *Nature communications*, **4**, p. 2191.

**Okuma, N. *et al.*** (2020) 'MIR2111-5 locus and shoot-accumulated mature miR2111 systemically enhance nodulation depending on HAR1 in Lotus japonicus', *Nature communications*, **11**(1), p. 5192.

**Oono, R. and Denison, R.F.** (2010) 'Comparing symbiotic efficiency between swollen versus nonswollen rhizobial bacteroids', *Plant physiology*, **154**(3), pp. 1541–1548.

**Op den Camp, R. *et al.*** (2011) 'LysM-type mycorrhizal receptor recruited for rhizobium symbiosis in nonlegume Parasponia', *Science*, **331**(6019), pp. 909–912.

**Penmetsa, R.V. *et al.*** (2003) 'Dual genetic pathways controlling nodule number in Medicago truncatula', *Plant physiology*, **131**(3), pp. 998–1008.

**Remy, W. *et al.*** (1994) 'Four hundred-million-year-old vesicular arbuscular mycorrhizae', *Proceedings of the National Academy of Sciences of the United States of America*, **91**(25), pp. 11841–11843.

**Rodriguez-Leal, D. *et al.*** (2019) 'Evolution of buffering in a genetic circuit controlling plant stem cell proliferation', *Nature genetics*, **51**(5), pp. 786–792.

**Sagan, M. *et al.*** (1995) 'Selection of nodulation and mycorrhizal mutants in the model plant Medicago truncatula (Gaertn.) after γ-ray mutagenesis', *Plant science: an international journal of experimental plant biology*, **111**(1), pp. 63–71.

**Sagan, M. and Duc, G.** (1996) 'Sym28 and sym29, two new genes involved in regulation of nodulation in pea (Pisum sativum L.)', *Symbiosis* , **20**, pp. 229–245.

**Sasaki, T. *et al.*** (2014) 'Shoot-derived cytokinins systemically regulate root nodulation', *Nature communications*, **5**, p. 4983.

**Schnabel, E. *et al.*** (2005) 'The Medicago truncatula SUNN gene encodes a CLV1-like leucine-rich repeat receptor kinase that regulates nodule number and root length', *Plant molecular biology*, 58(6), pp. 809–822.

**Schnabel, E. *et al.*** (2012) 'The M. truncatula SUNN gene is expressed in vascular tissue, similarly to RDN1, consistent with the role of these nodulation regulation genes in long distance signaling', *Plant signaling & behavior*, 7(1), pp. 4–6.

**Schnabel, E.L. *et al.*** (2023) 'A Medicago truncatula Autoregulation of Nodulation Mutant Transcriptome Analysis Reveals Disruption of the SUNN Pathway Causes Constitutive Expression Changes in Some Genes, but Overall Response to Rhizobia Resembles Wild-Type, Including Induction of TML1 and TML2', *Current issues in molecular biology*, **45**(6), pp. 4612–4631.

**Schuize, J., Adgo, E. and Merbach, W.** (1999) 'Carbon Costs Associated with $N_2$ Fixation in *Vicia faba* L and *Pisum sativum* 1. over a 14-Day Period', *Plant biology* , **1**(6), pp. 625–631.

**Searle, I.R. *et al.*** (2003) 'Long-distance signaling in nodulation directed by a CLAVATA1-like receptor kinase', *Science*, **299**(5603), pp. 109–112.

**Sheng, C. and Harper, J.E.** (1997) 'Shoot versus Root Signal Involvement in Nodulation and Vegetative Growth in Wild-Type and Hypernodulating Soybean Genotypes', *Plant physiology*, **113**(3), pp. 825–831.

**Shrihari, P.C. *et al.*** (2000) 'Interaction between supernodulating or non-nodulating mutants of soybean and two arbuscular mycorrhizal fungi', *Mycorrhiza*, **10**(3), pp. 101–106.

**Smith, S.E. and Read, D.J.** (2008) 'Mycorrhizal Symbiosis, 3rd editio Edn'. Cambridge, MA: Academic Press. doi.

**Soltis, D.E. *et al.*** (1995) 'Chloroplast gene sequence data suggest a single origin of the predisposition for symbiotic nitrogen fixation in angiosperms', *Proceedings of the National Academy of Sciences of the United States of America*, **92**(7), pp. 2647–2651.

**Stahl, Y. *et al.*** (2013) 'Moderation of Arabidopsis root stemness by CLAVATA1 and ARABIDOPSIS CRINKLY4 receptor kinase complexes', *Current biology: CB*, **23**(5), pp. 362–371.

**Suzaki, T. *et al.*** (2004) 'The gene FLORAL ORGAN NUMBER1 regulates floral meristem size in rice and encodes a leucine-rich repeat receptor kinase orthologous to Arabidopsis CLAVATA1',

*Development (Cambridge, England)*, 131(22), pp. 5649–5657.

**Szczyglowski, K. *et al.*** (1998) 'Har1, an Incomplete Dominant Mutation in Lotus japonicus, Confers Aberrant Root Development and an Unusual Hypernodulation Response (HNR)', in C. Elmerich, A. Kondorosi, and W.E. Newton (eds) *Biological Nitrogen Fixation for the 21st Century: Proceedings of the 11th International Congress on Nitrogen Fixation, Institut Pasteur, Paris, France, July 20–25, 1997*. Dordrecht: Springer Netherlands, pp. 305–306.

**Takahara, M. *et al.*** (2013) 'TOO MUCH LOVE, a Novel Kelch Repeat-Containing F-box Protein, Functions in the Long-Distance Regulation of the Legume–Rhizobium Symbiosis', *Plant & cell physiology*, **54**(4), pp. 433–447.

**Tsikou, D. *et al.*** (2018) 'Systemic control of legume susceptibility to rhizobial infection by a mobile microRNA', *Science*, **362**(6411), pp. 233–236.

**Van Velzen, R. *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 of the United States of America, **115**(20), pp. 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 in plant science*, **24**(1), pp. 49–57.

**Wang, C., Reid, J.B. and Foo, E.** (2018) 'The Art of Self-Control - Autoregulation of Plant-Microbe Symbioses', *Frontiers in plant science*, **9**, p. 988.

**Wardhani, T.A.K. *et al.*** (2019) 'Transforming, Genome Editing and Phenotyping the Nitrogen-fixing Tropical Cannabaceae Tree Parasponia andersonii', Journal of visualized experiments: JoVE, (150). doi: 10.3791/59971.

**Whitewoods, C.D. *et al.*** (2018) 'CLAVATA was a genetic novelty for the morphological innovation of 3D growth in land plants', *Current biology: CB*, **28**(15), pp. 2365–2376.e5.

**Wilson, M.C. *et al.*** (2017) 'Gene expression atlas for the food security crop cassava', *The new phytologist*, **213**(4), pp. 1632–1641.

**Wopereis, J. *et al.*** (2000) 'Short root mutant of Lotus japonicus with a dramatically altered symbiotic phenotype', *The Plant journal: for cell and molecular biology*, 23(1), pp. 97–114.

**Xu, C. *et al.*** (2015) 'A cascade of arabinosyltransferases controls shoot meristem size in tomato', *Nature genetics*, **47**(7), pp. 784–792.

**Yoro, E. *et al.*** (2019) 'PLENTY, a hydroxyproline O-arabinosyltransferase, negatively regulates root nodule symbiosis in Lotus japonicus', *Journal of experimental botany*, **70**(2), pp. 507–517.

**Yoshida, C., Funayama-Noguchi, S. and Kawaguchi, M.** (2010) 'plenty, a novel hypernodulation mutant in Lotus japonicus', *Plant & cell physiology*, 51(9), pp. 1425–1435.

**Zakaria Solaiman, M. *et al.*** (2000) 'Characterization of Mycorrhizas Formed by Glomus sp. on Roots of Hypernodulating Mutants of Lotus japonicus', *Journal of plant research*, **113**(4), pp. 443–448.

**Van Zeijl, A. *et al.*** (2018) 'CRISPR/Cas9-Mediated Mutagenesis of Four Putative Symbiosis Genes of the Tropical Tree Parasponia andersonii Reveals Novel Phenotypes', *Frontiers in plant science*, **9**, p. 284.

**Zhang, M. *et al.*** (2020) 'Shoot-derived miR2111 controls legume root and nodule development', *Cold Spring Harbor Laboratory*. Available at: https://doi.org/10.1101/2020.09.02.278739.

CHAPTER 6

# 6

# General discussion

Sultan Alhusayni[1,2]

[1]Laboratory of Molecular Biology, Department of Plant Sciences, Wageningen University, Droevendaalsesteeg 1, 6708 PB Wageningen, The Netherlands

[2]Biological Sciences Department, College of Science, King Faisal University, Al-Ahsa 31982, Saudi Arabia

**Introduction**

Plants are stationary organisms, and being fixed in the soil exposes them to many environmental challenges, such as drought, high salinity, extreme temperatures, and, most importantly, nutrient starvation, namely phosphate and nitrogen (Smith and Read, 2008). To overcome phosphate and nitrogen limitations in soil, plants engage with nutrient-scavenging microbes (Parniske, 2000; Geurts, Lillo and Bisseling, 2012).

An outstanding example of mutualistic interaction with nutrient-providing microbes is the nitrogen-fixing root nodulation trait. Root nodules can be formed by plant species within ten related taxonomic lineages (known collectively as the Nitrogen-Fixation Clade), with the legume family (Fabaceae) being the most notable. Legumes comprise nearly 20,000 species, most of which can establish a root nodule endosymbiosis with a diverse range of diazotrophic rhizobium bacteria. These bacteria reside within nodule cells and are surrounded by plant-derived membranes. When inside the nodule cells, the bacteria appear organelle-like, resembling nitroplasts (Nowack and Weber, 2018; Coale *et al.*, 2024). Nodules provide a low-oxygen environment, which maintains the stability and functionality of the oxygen-sensitive nitrogenase enzyme in rhizobium, which is essential for converting atmospheric nitrogen ($N_2$) into ammonia (Tjepkema and Yocum, 1974; Tjepkema, 1983). This highly energy-demanding process is powered by photosynthates and represents a significant investment from the plant (Minchin *et al.*, 1981; Schuize, Adgo and Merbach, 1999; Oono and Denison, 2010). In exchange, the plant benefits by receiving ammonia from the bacteria, enabling it to thrive in soils deficient in exogenous fixed nitrogen sources.



Engineering the nodulation trait into non-legume crop species has been a long-standing goal (Burrill and Hansen, 1917; Beringer and Hirsch, 1984; de Bruijn, Jing and Dazzo, 1995; Saikia and Jain, 2007; Charpentier and Oldroyd, 2010; Beatty and Good, 2011; Untergasser *et al.*, 2012; Mus *et al.*, 2016;

Huisman and Geurts, 2020). This research thesis also aligns with this objective. In this discussion chapter, I will reflect on how the data obtained on *SYMRK* in *Trema orientalis* (**Chapter 2**), *NSP2* overexpression, and *CLAVATA1* knockout studies in *Parasponia andersonii* (**Chapters 3** and **5**) might contribute to an engineering strategy.

Studies in legume models such as *Lotus japonicus* and *Medicago truncatula*, along with several legume crop species, have uncovered hundreds of genes involved in nodule formation and function (Roy *et al.*, 2020). The research led to the discovery that the nitrogen-fixing nodulation trait has co-opted several genes also utilized by the arbuscular mycorrhizal (AM) endosymbiosis with Glomeromycota fungi. These nutrient-scavenging AM fungi penetrate plant roots and form feeding structures in the root cortical cells, known as arbuscules (Parniske, 2008). Like rhizobia in nodule cells, arbuscules are encased in a plant-derived membrane, serving as an interface for nutrient exchange. The fungi deliver minerals, especially phosphates, which they gather using their extensive extraradical hyphal network in the soil. In return, they receive carbohydrates, mainly lipids, from the plant host (Parniske, 2008). Unlike the nodulation trait, AM symbiosis is more common among plants, with about 75% of all plant species capable of interacting with AM fungi. This interaction is considered ancient, dating back over 400 million years (Remy *et al.*, 1994).

A critical genetic module shared between nodulation and AM endosymbioses is the common symbiotic signaling pathway (CSSP). The CSSP comprises four components: the SYMBIOTIC SIGNALING RECEPTOR KINASE (SYMRK), the nuclear calcium-releasing cation channels CASTOR and POLLUX, the calcium-decoding kinase protein CALCIUM CALMODULIN-DEPENDENT PROTEIN KINASE (CCaMK), and the transcription factor CYCLOPS (Catoira *et al.*, 2000; Endre *et al.*, 2002; Stracke *et al.*, 2002; Kistner *et al.*, 2005; Edwards *et al.*, 2007; Messinese *et al.*, 2007; Charpentier *et al.*, 2008; Yano *et al.*, 2008). Mutations in any of these CSSP genes eliminate the symbiotic ability to interact with AM fungi and rhizobium bacteria (Catoira

*et al.*, 2000; Mitra *et al.*, 2004; Godfroy *et al.*, 2006; Tirichine *et al.*, 2006; Messinese *et al.*, 2007; Banba *et al.*, 2008; Markmann, Giczey and Parniske, 2008; Yano *et al.*, 2008; Chen *et al.*, 2009; Li *et al.*, 2022). Bioinformatic comparisons of the gene structures of the CSSP components between nodulating and non-nodulating plant species have shown that the encoded proteins are structurally very similar (Banba *et al.*, 2008; Markmann, Giczey and Parniske, 2008). Furthermore, the trans-complementation of legume mutants using CSSP genes from non-nodulating plant species has revealed that the CSSP genes are functionally highly conserved (Banba *et al.*, 2008; Markmann, Giczey and Parniske, 2008). These findings indicate that CSSP genes in non-nodulating plant species possess the molecular capacity to initiate the nodulation program.

To identify nodulation genes that represent critical engineering targets, phylogenomic approaches were used to compare nodulating legumes and non-legume species with their closely related non-nodulating sister species (Griesmann *et al.*, 2018; van Velzen *et al.*, 2018). These studies yielded two major findings. First, non-nodulating species within the Nitrogen-Fixing Clade have lost a limited number of nodulation genes, most notably the LysM-type receptor NFP/NFR5 and the transcription factor NIN (Geurts & Huisman, 2023). The NFP/NFR5 receptor specifically recognizes microbial-secreted lipo-chitooligosaccharides (LCOs). Together with the heterologous LysM-type receptor LYK3/NFR1 and the malectin-like LRR-type receptor SYMRK, it activates the CSSP and subsequently induces the expression of *NIN* (Radutoiu et al., 2003; Arrighi et al., 2006; Antolín-Llovera, Ried and Parniske, 2014; van Zeijl et al., 2018). NIN acts as a key regulator of nodulation and controls the transcriptional networks essential for nodule organogenesis, bacterial infection, and the conditions necessary for nitrogen fixation. In short, *NFP/NFR5* and *NIN* are considered essential targets for strategies for engineering the nodulation trait in other plant species (Schauser *et al.*, 1999; Griesmann *et al.*, 2018; van Velzen *et al.*, 2018; Geurts and Huisman, 2024). Second, the conserved pattern of gene loss in non-nodulating sister species within the Nitrogen Fixing Clade, while

homologs of NFP/NFR5 remain highly conserved in species outside this clade, suggests that these genes specifically functioned in nodulation and were lost when the nodulation trait was lost (Griesmann *et al.*, 2018; van Velzen *et al.*, 2018). Following this line of reasoning, it is hypothesized that nodulation originated from a single evolutionary event approximately 100 million years ago in the last common ancestor of the Nitrogen-Fixation Clade (Herendeen *et al.*, 1999; Werner *et al.*, 2014; Persson *et al.*, 2015).

**Relevance of a functional *SYMRK* allele in *Trema orientalis***

SYMRK is a transmembrane receptor featuring a malectin domain, a conserved glycine- aspartate-proline-cysteine (GDPC) motif, and 2-3 Leucine Rich Repats (LRRs) in its extracellular domain (Endre *et al.*, 2002; Stracke *et al.*, 2002; Markmann, Giczey and Parniske, 2008; Kosuta *et al.*, 2011). The malectin domain can be cleaved, allowing the processed transmembrane-localized receptor to interact with the NFP/NFR5 receptor (Antolín-Llovera, Ried and Parniske, 2014; Pan, Stonoha-Arther and Wang, 2018). Upon phosphorylation, SYMRK activates downstream signaling, initiating the root nodule developmental program (Abel et al. 2024). This functionality is not unique to SYMRK of nodulating plants; for example, SYMRK from Indian cress (*Tropaeolum majus*) and rice (*Oryza sativa*) can functionally complement nodule formation in an *L. japonicus symrk* knockout mutant (Markmann et al. 2008), suggesting that SYMRK functioning is highly conserved in plants capable of establishing AM and/or nodule symbiosis.

Two strategies regarding target species can be employed when aiming to engineer the nodulation trait. One approach targets relevant crops such as maize (*Zea mays*), barley (*Hordeum vulgare*), or cassava (*Manihot esculenta*), all of which are outside the Nitrogen Fixation Clade. Alternatively, the nodulation trait could be reverse-engineered in a non-nodulating species within the Nitrogen-Fixation Clade. *Trema* sp. is considered an attractive option in this case as it is closely related to the nodulating non-legume *Parasponia* (Cannabaceae) (Huisman & Geurts, 2020). Unlike other non-legume nodulators, *Parasponia* can form root

nodules with rhizobia, similar to legumes, though *Parasponia* nodules are less advanced than most legumes (Behm, Geurts and Kiers, 2014). Phylogenetic studies suggest nodulating *Parasponia* is nested within the *Trema* clade (Yang *et al.*, 2013). The loss of *NFP/NFR5* and *NIN* genes in *Trema* is estimated to have occurred only 17 million years ago, marking it a relatively recent event (Lie et al. 2015; Van Velzen et al. 2019).

A high-quality genome sequence and annotation have been generated for *Trema orientalis* accession RG33, collected in Malaysian Borneo (van Velzen *et al.*, 2018). Comparative analysis with *P. andersonii* revealed only limited gene loss events in *T. orientalis*. Surprisingly, *SYMRK* was annotated as a pseudogene in *T. orientalis* RG33, even though this species can still effectively engage in mycorrhization (van Velzen *et al.*, 2018). I investigated *SYMRK* more thoroughly and concluded that in several *Parasponia* and *Trema* species, *SYMRK* is a single-copy gene with a conserved intron-exon structure. However, only *T. orientalis* RG33 *SYMRK* $^{RG33}$ possesses a critical mutation at the donor splice site of intron 12, converting a GC -a highly efficient non-canonical splice site- into a GA, a motif unknown to function in a donor splice site. Reverse genetic studies in *P. andersonii* confirmed that *SYMRK* is essential for nodulation and mycorrhization (Alhusayni *et al.*, 2023). This suggests that also *SYMRK*$^{RG33}$ is a functional genes. I then complemented the *P. andersonii symrk* CRISPR knockout mutant with a construct carrying the GA mutation (*SYMRK*$^{RG33}$ mutation mimic) at the donor splice site of intron 12 in the *P. andersonii SYMRK* gene. This mutation did not impair the functionality of the *P. andersonii SYMRK* gene, as the construct successfully restored both nodulation and arbuscule formation in transformed roots of the *P. andersonii symrk* mutant. In summary, *T. orientalis* RG33 possesses a functionally conserved symbiotic *SYMRK* gene, despite an extremely rare GA donor splice site at intron 12. I conclude that *T. orientalis* accession RG33 can be used as an experimental system for reverse engineering the nitrogen-fixing nodulation trait.

Chapter 6

**NSP2-induced upregulation of the carotenoid biosynthetic pathway doesn't enhance the permissiveness for nodulation**

*NSP2* encodes a plant-specific GRAS-type transcriptional regulator protein. The initial discovery of *NSP2* was in legumes, revealing that plants with mutations in this gene lost the capacity to form nodules (Oldroyd and Long, 2003). Subsequent studies have shown that *NSP2* also facilitates AM fungi colonization, with *nsp2* mutants in legumes exhibiting reduced levels of AM colonization (Maillet *et al.*, 2011). These findings are consistent with recent data from our lab, which demonstrated that *P. andersonii nsp2* mutants cannot form nodules and show reduced levels of AM root colonization compared to controls (van Zeijl *et al.*, 2018). Furthermore, overexpression studies of *NSP2* in barley, rice, and *M. truncatula* have indicated NSP2's role in regulating the symbiotic permissiveness of plants to AM fungi. For instance, under low exogenous phosphate conditions, barely and *M. truncatula* plants exhibited high permissiveness to AM colonization, but upon an increase in phosphate content, this permissiveness is suppressed. Interestingly, overexpression of *NSP2* in barely and *M. truncatula* override phosphate suppression of AM colonization. This effect is observed in the roots of plants grown on medium and high phosphate levels (0.5mM and 1mM, respectively), which show enhanced levels of AM colonization and strigolactone content compared to wild-type plants. The observed enhancement of AM colonization on *NSP2* overexpressor lines under phosphate suppressive conditions is potentially through the modulation of gene expression of components crucial for the biosynthesis of strigolactones. Strigolactones are plant signaling molecules secreted by roots during phosphate starvation to attract AM fungi. Strigolactones enhance AM fungi spore germination, promote hyphal growth and branching, and modulate the internal activity of AM fungi (Akiyama et al. 2005; Besserer et al. 2006; Parniske 2008). Ultimately, strigolactones can enhance AM symbiosis, even though studies in rice indicated they are not essential (Servanté *et al.*, 2024).

Studies in *L. japonicus* have shown that overexpression of *NSP2* also affects nodulation. *NSP2* overexpression lines exhibit clustering of nodules, although the total number of nodules is not immediately affected (Murakami *et al.*, 2013). This suggests that *NSP2* expression may also regulate permissiveness to rhizobium. Our aim was to explore the trade-offs of *NSP2* overexpression using *P. andersonii* as an experimental system, which allows us to study nodulation in a non-legume context. We generated six lines ectopically expressing miRNA171h-resistant version of *Parasponia NSP2* (*mNSP2*), with overexpression levels ranging from 6 to 95-fold. These lines demonstrated enhanced mycorrhization under both low and high exogenous phosphate conditions, associated with the upregulation of the carotenoid and strigolactone biosynthetic pathways, consistent with studies in barley and *M. truncatula* (Li *et al.*, 2022). Surprisingly, we observed that *mNSP2* overexpressor lines exhibited a reduction in nodule number, with the severity of this reduction intensifying as m*NSP2* expression levels increased. This suggests that overexpression of m*NSP2* is not an appropriate strategy for enhancing nodulation permissiveness. Additionally, we observed rhizobium-induced aberrant cell divisions in the root epidermal, cortical, and endodermal cell layers in *P. andersonii* lines with high levels of m*NSP2* expression. These clusters of dividing cells may resemble enlarged pre-nodule structures, serving as infection pockets for the crack entry mode of infection in *Parasponia* (Lancelle and Torrey, 1984). This highlights the sensitivity of nodulation to *NSP2* expression levels and emphasizes the importance of maintaining balanced *NSP2* levels for successful nodule formation.



### *Parasponia CLAVATA1* functions in autoregulation of nodulation and mycorrhization

Early research on the regulation of nodulation revealed that legumes strictly control the formation of nodules through the autoregulation of nodulation (AON) pathway, which gets activated in response to exogenous nitrate and rhizobium inoculation as early as 1 to 4 days post-inoculation. The AON

pathway in legumes underlines stringent regulation of nodulation. Legumes suppress nodulation even before fully functional nodules are formed to satisfy their needs for nitrogen (Kosslak and Bohlool, 1984; Olsson *et al.*, 1989; Suzuki *et al.*, 2008; Li, Kinkema and Gresshoff, 2009; Kassaw and Frugoli, 2012). The AON pathway acts systemically. Legume plants carrying mutations in key genes essential for the long-distance signaling in the AON pathway exhibited enhanced symbiotic interactions. For example, *GmNARK, PsSYM29, LjHAR1*, and *MtSUNN* genes all encode a structurally similar Leucine-Rich Repeat Receptor Kinase protein that acts in the shoot of the plant. Knockout mutations in this gene enhance nodulation and mycorrhization (Morandi *et al.*, 2000; Shrihari *et al.*, 2000; Meixner *et al.*, 2005). In addition, legume species with a mutated *TML* gene that encodes a Kelch Repeat-Containing F-box protein in the root showed enhanced nodulation and mycorrhization (Magori, S. *et al.,* 2009; Chaulagain, D. *et al.,*2024). This suggests impairing AON signaling in legume species enhances their susceptibility to symbiotic interactions. In **Chapter 4**, we aimed to investigate (i) if the AON pathway is conserved beyond legumes, like in the non-legume *Parasonia* species, and (ii) if deregulation of the AON pathway could be a strategy to engineer nodulation in non-nodulating plant species. To determine that, I targeted the *P. andersonii* putative orthologs of two key genes in legumes AON pathway: *P. andersonii CLAVATA1 (PanCLV1,* ortholog of *LjHAR1* and *MtSUNN*) and *TOO MUCH LOVE* (*PanTML,* ortholog of *LjTML* and *MtTML1* and *MtTML2*).

Generating CRISPR Cas9-induced mutants of *PanTML* in *P. andersonii* was problematic as seven independent transformation experiments were performed to knock out *PanTML,* but all attempts were unsuccessful. No transgenic calli were formed to give rise to transgenic platelets. Additionally, trials to generate *TML* RNAi silencing mutant lines were fruitless, as six transgenic lines were acquired from transgenic calli, but analysis of their *TML* gene expression showed minimal reduction. So, we focused our efforts on generating mutants of *PanCLV1* and succeeded in generating two independent CRISPR Cas9-induced mutants, *Panclv1-16* and *Panclv1-21.*

First, analyzing the nodulation phenotype of *Panclv1-16* and *Panclv1-21* mutants revealed that they have hypernodulation phenotype. This phenotype is consistent with the phenotype observed on legume mutants that have impaired orthologous genes *GmNARK, PsSYM29, LjHAR1*, or *MtSUNN* (Sagan and Duc, 1996; Krusell et al., 2002; Searle et al., 2003; Schnabel et al., 2005). Therefore, these findings indicate that the AON pathway is a conserved mechanism utilized by both legume and non-legume species to regulate nodulation. *Panclv1-16* and *Panclv1-21* mutants also showed enhanced mycorrhizal colonization on the roots compared to empty vector control. This enhanced mycorrhization phenotype on *Panclv1* mutants is consistent with report data on enhanced mycorrhization on roots of *Mtsunn*, *Gmnark*, *Ljhar1*, and *Pssym19* legumes mutants (Morandi *et al.*, 2000; Shrihari *et al.*, 2000; Zakaria Solaiman *et al.*, 2000; Meixner *et al.*, 2005; Wang, Reid and Foo, 2018). Thus, our data further support the hypothesis of overlap between AON and autoregulation of mycorrhization pathway (AOM) to regulate mycorrhizal colonization in legume and non-legume plant species. Next, evaluating the shoot and root architecture of *Panclv1-16* and *Panclv1-21* revealed that AON is involved in root branching as these mutants exhibited increased primary root length and number of lateral roots compared to empty vector control. Whereas in the shoot, AON seems to play a role in controlling internode thickness. Our findings partially align with the published data related to the root phenotype of soybean *Gmnark* and *L. japonicus Ljhar1-1* mutants, which showed enhanced lateral root formation (Wopereis *et al.*, 2000; Searle *et al.*, 2003). This suggests a common role of AON in regulating root growth in legume and non-legume plant species. In conclusion, these studies in *Parasponia* emphasize *CLV1* commits a conserved function in the AON pathway, controlling nodule number, permissiveness to AM and root branching.



**Concluding remark**

In this thesis, I explored potential strategies that could contribute to achieving the long-standing objective of engineering nodulation in non-

nodulating plant species. First, I investigated whether the seemingly pseudogenized *TorSYMRK*$^{RG33}$ allele plays a role in the loss of nodulation in *Trema orientalis* RG33 accession. The key finding from this part of the thesis is that the *TorSYMRK*$^{RG33}$ allele is fully functional in both nodulation and mycorrhization, providing valuable insight for the Wageningen team's efforts to engineer nodulation in *Trema orientalis* RG33. Second, I assessed the feasibility of overexpressing *Parasponia mNSP2* to enhance symbiotic interactions in *P. andersonii*. The main conclusion here is that the trade-offs associated with *mNSP2* overexpression outweigh the benefits of increased symbiotic interactions, indicating that overexpressing m*NSP2* is not an appropriate strategy for this purpose. Finally, I explored the viability of utilizing the autoregulation of nodulation (AON) pathway to enhance symbiotic interactions in plants. By generating *Panclv1* mutants in *P. andersonii*, and analyzing their symbiotic phenotypes, I observed a significant enhancement in their symbiotic interaction. This finding suggests that components of the AON pathway, such as *CLV1*, could serve as promising targets for engineering nodulation in non-nodulating plant species.

**References**

**Alhusayni, S. *et al.*** (2023) 'A rare non-canonical splice site in Trema orientalis SYMRK does not affect its dual symbiotic functioning in endomycorrhiza and rhizobium nodulation', *BMC plant biology*, **23**(1), p. 587.

**Antolín-Llovera, M., Ried, M.K. and Parniske, M.** (2014) 'Cleavage of the SYMBIOSIS RECEPTOR-LIKE KINASE ectodomain promotes complex formation with Nod factor receptor 5', *Current biology: CB*, **24**(4), pp. 422–427.

**Arrighi, J.-F. *et al.*** (2006) 'The Medicago truncatula Lysine Motif-Receptor-Like Kinase Gene Family Includes NFP and New Nodule-Expressed Genes', *Plant physiology*, **142**(1), pp. 265–279.

**Banba, M. *et al.*** (2008) 'Divergence of evolutionary ways among common sym genes: CASTOR and CCaMK show functional conservation between two symbiosis systems and constitute the root of a common signaling pathway', *Plant & cell physiology*, **49**(11), pp. 1659–1671.

**Beatty, P.H. and Good, A.G.** (2011) 'Plant science. Future prospects for cereals that fix nitrogen', *Science (New York, N.Y.)*, **333**(6041), pp. 416–417.

**Behm, J.E., Geurts, R. and Kiers, E.T.** (2014) 'Parasponia: a novel system for studying mutualism stability', *Trends in plant science*, **19**(12), pp. 757–763.

**Beringer, J.E. and Hirsch, P.R**. (1984) 'Genetic engineering and nitrogen fixation', *Biotechnology & genetic engineering reviews*, **1**(1), pp. 65–88.

**de Bruijn, F.J., Jing, Y. and Dazzo, F.B.** (1995) 'Potential and pitfalls of trying to extend symbiotic interactions of nitrogen-fixing organisms to presently non-nodulated plants, such as rice', in *Management of Biological Nitrogen Fixation for the Development of More Productive and Sustainable Agricultural Systems*. Dordrecht: Springer Netherlands, pp. 225–240.

**Burrill, T.J. and Hansen, R.** (1917) 'Is symbiosis possible between legume bacteria and non-legume plants?', *Bulletin* [Preprint].

**Catoira, R. *et al.*** (2000) 'Four genes of Medicago truncatula controlling components of a nod factor transduction pathway', *The Plant cell*, 12(9), pp. 1647–1666.

**Charpentier, M. *et al.*** (2008) 'Lotus japonicus CASTOR and POLLUX are ion channels essential for perinuclear calcium spiking in legume root endosymbiosis', *The Plant cell*, **20**(12), pp. 3467–3479.

**Chaulagain, D. *et al.*** (2024) 'Tml1 and tml2 synergistically regulate nodulation and affect arbuscular mycorrhiza in Medicago truncatula', bioRxiv.org: the preprint server for biology [Preprint]. Available at: https://doi.org/10.1101/2023.12.07.570674.

**Charpentier, M. and Oldroyd, G.** (2010) 'How close are we to nitrogen-fixing cereals?', *Current opinion in plant biology*, **13**(5), pp. 556–564.

**Chen, C. *et al.*** (2009) 'Antiquity and function of CASTOR and POLLUX, the twin ion channel-encoding genes key to the evolution of root symbioses in plants', *Plant physiology*, **149**(1), pp. 306–317.

**Coale, T.H. *et al.*** (2024) 'Nitrogen-fixing organelle in a marine alga', *Science*, **384**(6692), pp. 217–222.

**Edwards, A. *et al.*** (2007) 'Structural implications of mutations in the pea SYM8 symbiosis gene, the DMI1 ortholog, encoding a predicted ion channel', *Molecular plant-microbe interactions: MPMI*, **20**(10), pp. 1183–1191.

**Endre, G. *et al.*** (2002) 'A receptor kinase gene regulating symbiotic nodule development', *Nature*, **417**(6892), pp. 962–966.

**Geurts, R. and Huisman, R**. (2024) 'Innovations in two genes kickstarted the evolution of nitrogen-fixing nodules', *Current opinion in plant biology*, 77, p. 102446.

**Geurts, R., Lillo, A. and Bisseling, T.** (2012) 'Exploiting an ancient signalling machinery to enjoy a nitrogen fixing symbiosis', *Current opinion in plant biology*, **15**(4), pp. 438–443.

**Godfroy, O. *et al.*** (2006) 'A rice calcium- and calmodulin-dependent protein kinase restores nodulation to a legume mutant', *Molecular plant-microbe interactions: MPMI*, **19**(5), pp. 495–501.

**Griesmann, M. *et al.*** (2018) 'Phylogenomics reveals multiple losses of nitrogen-fixing root nodule symbiosis', *Science*, **361**(6398), p. eaat1743.

**Herendeen, P.S. *et al.*** (1999) 'A Preliminary Conspectus of the Allon Flora from the Late Cretaceous (Late Santonian) of Central Georgia, U.S.A', *Annals of the Missouri Botanical Garden. Missouri Botanical Garden*, **86**(2), pp. 407–471.

**Huisman, R. and Geurts, R.** (2020) 'A roadmap towards engineered

nitrogen-fixing nodule symbiosis', *Plant Communications*, p. 100019.

**Kassaw, T.K. and Frugoli, J.A.** (2012) 'Simple and efficient methods to generate split roots and grafted plants useful for long-distance signaling studies in Medicago truncatula and other small plants', *Plant methods*, **8**(1), p. 38.

**Kistner, C. *et al.*** (2005) 'Seven Lotus japonicus genes required for transcriptional reprogramming of the root during fungal and bacterial symbiosis', *The Plant cell*, 17(8), pp. 2217–2229.

**Kosslak, R.M. and Bohlool, B.B.** (1984) 'Suppression of nodule development of one side of a split-root system of soybeans caused by prior inoculation of the other side', *Plant physiology*, 75(1), pp. 125–130.

**Kosuta, S. *et al.*** (2011) 'Lotus japonicus symRK-14 uncouples the cortical and epidermal symbiotic program: GDPC sequence and the evolution of SYMRK function', *The Plant journal: for cell and molecular biology*, **67**(5), pp. 929–940.

**Krusell, L. *et al.*** (2002) 'Shoot control of root development and nodulation is mediated by a receptor-like kinase', *Nature*, **420**(6914), pp. 422–426.

**Lancelle, S.A. and Torrey, J.G.** (1984) 'Early development ofRhizobium-induced root nodules ofParasponia rigida. I. Infection and early nodule initiation', *Protoplasma*, **123**(1), pp. 26–37.

**Li, D., Kinkema, M. and Gresshoff, P.M.** (2009) 'Autoregulation of nodulation (AON) in Pisum sativum (pea) involves signalling events associated with both nodule primordia development and nitrogen fixation', *Journal of plant physiology*, **166**(9), pp. 955–967.

**Li, X.-R. *et al.*** (2022) 'Nutrient regulation of lipochitooligosaccharide recognition in plants via NSP1 and NSP2', *Nature communications*, **13**(1), p. 6421.

**Magori, S. *et al.*** (2009) 'TOO MUCH LOVE, a Root Regulator Associated with the Long-Distance Control of Nodulation in Lotus japonicus', *Molecular plant-microbe interactions: MPMI*, **22**(3), pp. 259–268.

**Maillet, F. *et al.*** (2011) 'Fungal lipochitooligosaccharide symbiotic signals in arbuscular mycorrhiza', *Nature*, **469**(7328), pp. 58–63.

**Markmann, K., Giczey, G. and Parniske, M.** (2008) 'Functional Adaptation of a Plant Receptor- Kinase Paved the Way for the Evolution of Intracellular Root Symbioses with Bacteria', *PLoS biology*, **6**(3), p.

e68.

**Meixner, C. *et al.*** (2005) 'Lack of mycorrhizal autoregulation and phytohormonal changes in the supernodulating soybean mutant nts1007', *Planta*, **222**(4), pp. 709–715.

**Messinese, E. *et al.*** (2007) 'A novel nuclear protein interacts with the symbiotic DMI3 calcium- and calmodulin-dependent protein kinase of Medicago truncatula', *Molecular plant-microbe interactions: MPMI*, **20**(8), pp. 912–921.

**Minchin, F.R. *et al.*** (1981) 'Carbon and nitrogen nutrition of nodulated roots of grain legumes', *Plant, cell & environment*, **4**(1), pp. 5–26.

**Mitra, R.M. *et al.*** (2004) 'A Ca2+/calmodulin-dependent protein kinase required for symbiotic nodule development: Gene identification by transcript-based cloning', *Proceedings of the National Academy of Sciences of the United States of America*, **101**(13), pp. 4701–4705.

**Morandi, D. *et al.*** (2000) 'Influence of genes determining supernodulation on root colonization by the mycorrhizal fungus Glomus mosseae in Pisum sativum and Medicago truncatula mutants', *Mycorrhiza*, **10**(1), pp. 37–42.

**Murakami, Y. *et al.*** (2013) 'Down-Regulation of NSP2 Expression in Developmentally Young Regions of Lotus japonicus Roots in Response to Rhizobial Inoculation', *Plant & cell physiology*, **54**(4), pp. 518–527.

**Mus, F. *et al.*** (2016) 'Symbiotic Nitrogen Fixation and the Challenges to Its Extension to Nonlegumes', *Applied and environmental microbiology*, **82**(13), pp. 3698–3710.

**Nowack, E.C.M. and Weber, A.P.M.** (2018) 'Genomics-Informed Insights into Endosymbiotic Organelle Evolution in Photosynthetic Eukaryotes', *Annual review of plant biology*, **69**, pp. 51–84.

**Oldroyd, G.E.D. and Long, S.R.** (2003) 'Identification and characterization of nodulation-signaling pathway 2, a gene of Medicago truncatula involved in Nod actor signaling', *Plant physiology*, **131**(3), pp. 1027–1032.

**Olsson, J.E. *et al.*** (1989) 'Lack of Systemic Suppression of Nodulation in Split Root Systems of Supernodulating Soybean (Glycine max [L.] Merr.) Mutants', *Plant physiology*, **90**(4), pp. 1347–1352.

**Oono, R. and Denison, R.F.** (2010) 'Comparing symbiotic efficiency between swollen versus nonswollen rhizobial bacteroids', *Plant physiology*, **154**(3), pp. 1541–1548.

**Pan, H., Stonoha-Arther, C. and Wang, D.** (2018) 'Medicago Plants Control Nodulation by Regulating Proteolysis of the Receptor-Like Kinase DMI2', *Plant physiology*, **177**(2), pp. 792–802.

**Parniske, M.** (2000) 'Intracellular accommodation of microbes by plants: a common developmental program for symbiosis and disease?', *Current opinion in plant biology*, **3**(4), pp. 320–328.

**Parniske, M.** (2008) 'Arbuscular mycorrhiza: the mother of plant root endosymbioses', *Nature reviews. Microbiology*, **6**(10), pp. 763–775.

**Persson, T. *et al.*** (2015) 'Candidatus Frankia Datiscae Dg1, the Actinobacterial Microsymbiont of Datisca glomerata, Expresses the Canonical nod Genes nodABC in Symbiosis with Its Host Plant', *PloS one*, **10**(5), p. e0127630.

**Radutoiu, S. *et al.*** (2003) 'Plant recognition of symbiotic bacteria requires two LysM receptor-like kinases', *Nature*, **425**(6958), pp. 585–592.

**Remy, W. *et al.*** (1994) 'Four hundred-million-year-old vesicular arbuscular mycorrhizae', *Proceedings of the National Academy of Sciences of the United States of America*, **91**(25), pp. 11841–11843.

**Roy, S. *et al.*** (2020) 'Celebrating 20 Years of Genetic Discoveries in Legume Nodulation and Symbiotic Nitrogen Fixation', *The Plant cell*, **32**(1), pp. 15–41.

**Sagan, M. and Duc, G.** (1996) 'Sym28 and sym29, two new genes involved in regulation of nodulation in pea (Pisum sativum L.)', *Symbiosis* , **20**, pp. 229–245.

**Saikia, S.P. and Jain, V.** (2007) 'Biological nitrogen fixation with non-legumes : An achievable target or a dogma?', *Current Science*, **92**(3), pp. 317–322.

**Schauser, L. *et al.*** (1999) 'A plant regulator controlling development of symbiotic root nodules', *Nature*, **402**(6758), pp. 191–195.

**Schnabel, E. *et al.*** (2005) 'The Medicago truncatula SUNN gene encodes a CLV1-like leucine-rich repeat receptor kinase that regulates nodule number and root length', *Plant molecular biology*, **58**(6), pp. 809–822.

**Schuize, J., Adgo, E. and Merbach, W.** (1999) 'Carbon Costs Associated with $N_2$ Fixation in *Vicia faba* L and *Pisum sativum* 1. over a 14-Day Period', *Plant biology* , **1**(6), pp. 625–631.

**Searle, I.R. *et al.*** (2003) 'Long-distance signaling in nodulation directed by a CLAVATA1-like receptor kinase', *Science*, **299**(5603), pp. 109–112.

**Servanté, E.K. *et al.*** (2024) 'Independent regulation of strigolactones and

blumenols during arbuscular mycorrhizal symbiosis in rice', *The Plant journal: for cell and molecular biology*, **119**(3), pp. 1289–1298.

**Shrihari, P.C. *et al.*** (2000) 'Interaction between supernodulating or non-nodulating mutants of soybean and two arbuscular mycorrhizal fungi', *Mycorrhiza*, **10**(3), pp. 101–106.

**Smith, S.E. and Read, D.J.** (2008) 'Mycorrhizal Symbiosis, 3rd editio Edn'. Cambridge, MA: Academic Press. doi.

**Stracke, S. *et al.*** (2002) 'A plant receptor-like kinase required for both bacterial and fungal symbiosis', *Nature*, **417**(6892), pp. 959–962.

**Suzuki, A. *et al.*** (2008) 'Split-root study of autoregulation of nodulation in the model legume Lotus japonicus', *Journal of plant research*, **121**(2), pp. 245–249.

**Tirichine, L. *et al.*** (2006) 'Deregulation of a Ca2+/calmodulin-dependent kinase leads to spontaneous nodule development', *Nature*, **441**(7097), pp. 1153–1156.

**Tjepkema, J.D.** (1983) 'Oxygen Concentration Within the Nitrogen-Fixing Root Nodules of Myrica gale L', *American journal of botany*, **70**(1), pp. 59–63.

**Tjepkema, J.D. and Yocum, C.S.** (1974) 'Measurement of oxygen partial pressure within soybean nodules by oxygen microelectrodes', *Planta*, **119**(4), pp. 351–360.

**Untergasser, A. *et al.*** (2012) 'One-step Agrobacterium mediated transformation of eight genes essential for rhizobium symbiotic signaling using the novel binary vector system pHUGE', *PloS one*, **7**(10), p. e47885.

**Van Velzen, R. *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 of the United States of America*, **115**(20), pp. E4700–e4709.

**Wang, C., Reid, J.B. and Foo, E.** (2018) 'The Art of Self-Control - Autoregulation of Plant-Microbe Symbioses', *Frontiers in plant science*, **9**, p. 988.

**Werner, G.D.A. *et al.*** (2014) 'A single evolutionary innovation drives the deep evolution of symbiotic N2-fixation in angiosperms', *Nature communications*, **5**(1), pp. 1–9.

**Wopereis, J. *et al.*** (2000) 'Short root mutant of Lotus japonicus with a dramatically altered symbiotic phenotype', *The Plant journal: for*

*cell and molecular biology*, **23**(1), pp. 97–114.

**Yang, M.-Q. *et al.*** (2013) 'Molecular phylogenetics and character evolution of Cannabaceae', *Taxon*, 62(3), pp. 473–485.

**Yano, K. *et al.*** (2008) 'CYCLOPS, a mediator of symbiotic intracellular accommodation', *Proceedings of the National Academy of Sciences*, **105**(51), pp. 20540–20545.

**Zakaria Solaiman, M. *et al.*** (2000) 'Characterization of Mycorrhizas Formed by Glomus sp. on Roots of Hypernodulating Mutants of Lotus japonicus', *Journal of plant research*, **113**(4), pp. 443–448.

**Van Zeijl, A. *et al.*** (2018) 'CRISPR/Cas9-Mediated Mutagenesis of Four Putative Symbiosis Genes of the Tropical Tree Parasponia andersonii Reveals Novel Phenotypes', *Frontiers in plant science*, **9**, p. 284.

# Summary

Sultan Alhusayni[1,2]

[1]Laboratory of Molecular Biology, Department of Plant Sciences, Wageningen University, Droevendaalsesteeg 1, 6708 PB Wageningen, The Netherlands

[2]Biological Sciences Department, College of Science, King Faisal University, Al-Ahsa 31982, Saudi Arabia

Nutrient limitation is a crucial factor restricting plant growth and development. To address this issue, farmers often supply their plants with chemical fertilizers. However, the overuse of these chemicals is harmful to the environment, leading to high concentrations in soil and water, which results in soil degradation and water pollution, respectively. To minimize reliance on chemical fertilizers, it is beneficial to explore biological nitrogen fixation, a sustainable biochemical process in which nitrogen-fixing bacteria inside plant cells fix the atmospheric nitrogen. For decades, the scientific community has focused on understanding this process, particularly the nitrogen-fixing endosymbiosis, where rhizobium bacteria are hosted in root lateral organs called nodules. These nodules provide a suitable environment for nitrogen fixation and exchange between the bacteria and the host plant. Forward and reverse genetics in two legume models, *Medicago truncatula* and *Lotus japonicus,* and soybean and pea crops have led to the discovery of over 200 genes involved in this process. The accumulated knowledge about the genetic network regulating nodule symbiosis has inspired recent initiatives to move from exploration to application, with aims to (i) enhance this trait in plants that naturally possess nitrogen-fixing endosymbiosis, like legumes, (ii) re-engineer this trait in plants that have lost it, such as non-nodulating plants within the nitrogen-fixing clade, and (iii) engineer this trait in plants that have never obtained it, such as rice, maize, and cassava. My research thesis aligns with aims (ii) and (iii), as they highlight promising strategies for the engineering of nodulation in non-nodulating plants. This work is based on research on the non-legume *Parasponia andersonii* that can form root nodule symbiosis with rhizobium bacteria.

In **Chapter 2,** I show that *Trema orientalis* RG33 is a valuable experimental system for reverse engineering the nitrogen-fixing nodulation trait. *Trema orientalis* is a non-nodulating plant closely related to the nodulating *Parasponia*. Previous comparative genomics between *Parasponia* and *Trema* species showed that several genes specific to nodulation experienced either loss or pseudogenization in *Trema* species. Notably, Among marked genes as pseudogenized in *Trema* species is only the *SYMRK* of *Trema*

*orientalis* RG33 (*TorSYMRK*$^{RG33}$), which possesses a critical mutation at the donor splice site of intron 12, converting a GC—a highly efficient non-canonical splice site (99.99%)—into a GA a rare motif not known to function in a donor splice site. The *SYMRK* gene encodes a Leucine-Rich Repeat Receptor-Like Kinase protein essential for nodulation and mycorrhization. As *Trema orientalis* RG33 can be mycorrhized, it remains unclear if the rare GA donor splice site in *TorSYMRK*$^{RG33}$ allele can also support nodulation. To investigate this, we generated *P. andersonii symrk* CRISPR knockout mutants, confirming that *SYMRK* is essential for nodulation and mycorrhization. We then complemented the *P. andersonii symrk* mutant with a construct mimicking the *SYMRK*$^{RG33}$ mutation at the donor splice site of intron 12 in the *P. andersonii SYMRK* gene. The introduced GA mutation in *P. andersonii SYMRK* gene did not affect its functionality, as the construct successfully restored nodulation and arbuscular formation in the *P. andersonii symrk* mutant. In summary, *T. orientalis* possesses a functionally conserved symbiotic *SYMRK* gene, despite the presence of an extremely rare GA donor splice site at intron 12. This work highlights the potential of *T. orientalis* as an experimental system for reverse engineering the nitrogen-fixing nodulation trait.

In **Chapter 3**, I show that overexpression of a miRNA171h-resistant version of *Parasponia NSP2* (*mNSP2*) is a beneficial strategy to enhance mycorrhizal colonization, but harmful for nodulation and plant root development. I generated six transgenic lines over-expressing *mNSP2* with an expression level ranging from a 6 to 95-fold increase compared to wild-type plants. Analysis of the mycorrhization phenotype of the *PanNSP2* over-expressing lines under both high and low exogenous phosphate conditions showed an increase in hyphae, arbuscule, and vesicle colonization. This increase in colonization level positively correlates with the increased expression level of *mNSP2. NSP2* is known to be a key regulator of the biosynthetic pathway of strigolactones, a plant-secreted hormone important for AM fungi root colonization and spore germination. Transcriptome analysis of *mNSP2* over-expressing lines showed upregulation of all genes in the strigolactone

signaling pathway and most component genes of upstream pathways, including the MEP and carotenoid pathways. Next, we assessed the nodulation phenotype of *mNSP2* over-expressing lines, and we observed that the nodule number was reduced, and this reduction negatively correlates with the increase in m*NSP2* expression level. Moreover, evaluation of the root system of *mNSP2* over-expressing lines under non-symbiotic conditions reveals a decrease in lateral root number only on m*NSP2* over-expressing lines with high expression levels. In conclusion, *NSP2* over-expression has the potential to be used as a biotechnological tool to promote mycorrhization. However, *NSP2* over-expression is not the right tool for engineering nodulation in non-nodulating plants.

In **Chapter 4**, I gave an overview of the Autoregulation of Nodulation (AON) mechanisms by which plants, particularly legumes, control their root nodule and AM symbiotic interactions.

In **Chapter 5**, I showed that the AON mechanism, which controls nodule number, is conserved in the non-legume *Parasponia*. I generated CRISPR Cas9-induced mutations in *Parasponia CLAVATA1* (*PanCLV1*), the putative ortholog of *Medicago truncatula SUNN* and *Lotus japonicus HAR1* genes that are known to play a key role in the AON pathway in legumes. I obtained two independent mutants of *PanCLV1* and named them *Panclv1-16* and *Panclv1-21*. Assessment of the nodulation phenotypes of *Panclv1-16* and *Panclv1-21* mutants revealed a hypernodulation phenotype compared to control plants. This finding is consistent with the nodulation phenotype observed in *Mtsunn* and *Ljhar1* mutants, suggesting that the AON pathway is an ancient mechanism to control nodule number in nodulating legume and non-legume plant species. Next, we evaluated the mycorrhization phenotype of the *Pancvl1* mutants and observed a significant increase in mycorrhizal colonization in their roots compared to control plants. This result indicates an overlap between AON and the more ancient autoregulation of mycorrhization (AOM) pathways. Finally, we assessed the impact of mutating *PanCLV1* on root development and observed an increase in the

primary root length and lateral roots number of *Panclv1* mutants compared to control plants. This finding aligns with the reported compact lateral roots phenotype of *Mtsunn* and *Ljhar1* mutants. In summary, the AON mechanism for controlling nodule number is not a legume-specific adaptation but a generic mechanism in both legume and non-legume plants that probably evolved alongside the nitrogen fixation trait in the last common ancestor of the nitrogen-fixing clade.

In **Chapter 6**, I discuss the implications of my research in the context of engineering nodulation in non-nodulating plants. The key conclusion is that *Trema orientalis* RG33 is a viable experimental system for re-engineering the nitrogen-fixing trait. Furthermore, I highlight that while overexpressing *mNSP2* is not an effective strategy for engineering nodulation in non-nodulating plants, the potential of the autoregulation of nodulation (AON) pathway presents a more promising strategy for future research in this area.

في النباتات غير البقولية، لكن ليس أداة مناسبة لتعزيز العقد الجذرية في النباتات غير البقولية أو لهندسة العقد الجذرية في النباتات غير العقدية

في الفصل الرابع، قدمت نظرة عامة على آلية التنظيم الذاتي التي طورتها النباتات البقولية للتحكم في عدد العقد الجذرية. قادنا ذلك إلى طرح تساؤلات حول ما إذا كانت هذه الآلية تم تطويرها بشكل مستقل أم أنها آلية قديمة طورها الجد المشترك وورثها لسلالاته.

في الفصل الخامس، استكشفت آليات التنظيم الذاتي للتعايش العقدي في النباتات غير البقولية مستخدمًا نبات *P. andersonii*، حيث قمت بتوليد طفرات باستخدام تقنية كريسبر كاس 9 في جين *CLAVATA1*، الذي يُعتبر نظيرًا لجين *SUNN* وجين *HAR1* في النباتات البقولية، والمعروفين بدورهما الأساسي في آليات التنظيم الذاتي للعقد الجذرية. وقد أظهرت النتائج أن الطفرة في جين *CLAVATA1* أدت إلى زيادة هائلة في عدد العقد الجذرية مقارنة بالنباتات البرية، مما يشير إلى أن الطفرة المستحدثة في هذا الجين عطلت قدرة النبات على التحكم في عدد العقد الجذرية، مما يقترح أن آليات التنظيم الذاتي للتعايش العقدي قديمة وتعود إلى الجد المشترك لطبقة تثبيت النيتروجين الجوي. في نهاية دراسة هذا الفصل، توصلنا إلى أن تعطيل آليات التنظيم الذاتي للتعايش العقدي تعتبر أداة حيوية واعدة لجعل النباتات غير العقدية أكثر قابلية لتكوين العقد الجذرية على جذورها.

في الفصل السادس، ناقشت النتائج التي توصلت إليها في رسالة الدكتوراه هذه، وربطتها بمستجدات الأبحاث في مجال تثبيت النيتروجين حيويًا.

تتماشى رسالة الدكتوراه هذه مع الأهداف الثلاثة، إذ تُبرز استراتيجيات واعدة لهندسة العقد الجذرية في النباتات غير العقدية. وبما أن معظم الدراسات السابقة لفهم عملية تثبيت النيتروجين حيويًا تمت على أنواع نباتية من العائلة البقولية، فقد ركزت في هذه الرسالة على دراسة العقد الجذرية في النباتات غير البقولية مثل نبات *Parasponia andersonii* (*P. andersonii*) ، الذي يمكنه تكوين تعايش جذري مع بكتيريا الرايزوبيوم. تُعطينا دراسة العقد الجذرية على جذور نبات *P. andersonii* صورة أوضح عن الجينات المشتركة في عملية تثبيت النيتروجين بين نباتات العائلة البقولية وغير البقولية.

في الفصل الأول من الرسالة، تناولت بإسهاب آخر الأبحاث حول عملية تثبيت النيتروجين على المستوى التطوري والجيني.

في الفصل الثاني، أجريت دراسة جينية مقارنة بين النبات غير العقدي *Trema orientalis RG33* (*T. orientalis*$^{RG33}$) والنبات العقدي *P. andersonii*، حيث تم اختيار هذين النباتين للمقارنة نظرًا لقربهما الجيني من بعضهما البعض واشتراكهما في جد مشترك، مما يقترح وجود اختلافات جينية محدودة بينهما. ركزت في هذا الفصل على جين *SYMRK*$^{RG33}$ في نبات*T. orientalis*$^{RG33}$، حيث تُظهر الدراسات السابقة في النباتات البقولية أن هذا الجين يلعب دورًا رئيسيًا في استقبال الإشارات المرسلة من بكتيريا الرايزوبيوم. وقد تبيّن أن هذا الجين في *T. orientalis*$^{RG33}$ تعرض لعملية تحوير، ما أثار تساؤلًا حول ما إذا كان التحور في*SYMRK*$^{RG33}$ هو أحد الأسباب الرئيسية لفقدان هذا النبات القدرة على تثبيت النيتروجين. وتوصلنا في نهاية هذا الفصل إلى أن هذا التحور لم يكن سبب في فقدان القدرة على تثبيت النيتروجين، إذ إننا حين أحدثنا تحورًا مشابهًا في جين *SYMRK* لنبات *P. andersonii*، بقي النبات قادرًا على التواصل مع بكتيريا الرايزوبيوم واستضافتها داخل عقد جذرية مشابهة لتلك الموجودة على جذور النبات غير المعدل.

في الفصل الثالث، قمت بدراسة جين *NSP2* في نبات *P. andersonii*، حيث أظهرت الدراسات السابقة أن هذا الجين يلعب دورًا رئيسيًا في تنظيم تواصل النبات البقولي وغير البقولي مع بكتيريا الرايزوبيوم وفي نفس الوقت ينظم تواصله مع فطريات المايكورايزا الشجرية الداخلية. وقد أظهرت دراسة حديثة أن الإفراط في التعبير الجيني لجين *NSP2* يُعدّ استراتيجية مفيدة لتعزيز استعمار فطريات المايكورايزا لجذور النباتات البقولية. ومن خلال إنتاج ستة خطوط معدلة وراثيًا تعبّر جين *NSP2* بمستويات تتراوح بين 6 إلى 95 ضعفًا مقارنة بالنباتات البرية، توصلت إلى أن الإفراط في التعبير عن *NSP2* يعزز استعمار فطريات المايكورايزا الشجرية لجذور نبات *P. andersonii*، لكنه يرتبط سلبًا بعدد العقد الجذرية. في نهاية دراسة هذا الفصل، توصلنا إلى أن الإفراط في التعبير الجيني لجين *NSP2* يعتبر أداة حيوية واعدة لتعزيز استعمار فطريات المايكورايزا الشجرية الداخلية

مقدمة

يُعدّ شح عنصر النيتروجين العضوي في التربة عاملًا حاسمًا يقيّد نمو وتطور النباتات. لمواجهة هذه المشكلة، يلجأ المزارعون غالبًا إلى تزويد النبات بالأسمدة الكيميائية، لكن الإفراط في استخدام هذه المواد يؤدي إلى تراكمها بكميات كبيرة في التربة والمياه، مما يسبب تدهور التربة وتلوث مصادر المياه. لتقليل الاعتماد على الأسمدة الكيميائية وتزويد النبات بالكميات الضرورية من النيتروجين، يُعدّ الاستفادة من العملية الحيوية لتثبيت النيتروجين خيارًا صديقًا للبيئة ومستدامًا. في هذه العملية، يتم تثبيت النيتروجين الجوي من الحالة الغازية إلى الحالة الثابتة داخل خلايا النبات بواسطة بكتيريا مثبتة للنيتروجين. على مدار عدة عقود، ركز المجتمع العلمي على فهم هذه العملية، ووجد أن بعض النباتات تستطيع تثبيت النيتروجين حيويًا من خلال استضافة أنواع معينة من بكتيريا الرايزوبيوم في أعضاء طرفية على الجذر تُسمى العقد الجذرية، حيث توفر هذه العقد بيئة مناسبة لتثبيت النيتروجين حيويًا وتبادل العناصر الغذائية بين البكتيريا والنبات المضيف. تتواجد هذه القدرة على تثبيت النيتروجين في نباتات تنتمي لأربع رتب نباتية هي Fabales و Fagales و Rosales و Cucurbitales، ويُطلق عليها مجتمعةً طبقة تثبيت النيتروجين.

أظهرت الدراسات الجينية أن عملية تثبيت النيتروجين حيويًا تطورت في الجد المشترك لهذه الطبقة قبل حوالي 110 مليون سنة، من خلال الاستفادة من جينات طورتها أسلافها قبل حوالي 450 مليون سنة للتعايش المفيد مع فطريات المايكورايزا الشجرية
(Arbuscular Mycorrhizal Fungi - AM Fungi)، التي تساعد النبات على توسيع نطاق الحصول على العناصر الغذائية من التربة. دراسة العديد من سلالات هذا الجد المشترك أظهرت أن عددًا كبيرًا من هذه السلالات فقد القدرة على تثبيت النيتروجين الجوي وسُمِّيت بالنباتات (غير العقدية)، بينما حافظت سلالات محدودة على هذه القدرة وسُمِّيت بالنباتات (العقدية). وقد أوضح تصنيف النباتات العقدية أنها تنتمي فقط إلى عشر عوائل نباتية، من أهمها العائلة البقولية.

أظهرت الدراسات الجينية التي استمرت لاكثر من عقدين على بعض النباتات البقولية، مثل *Medicago trancatula* و *Lotus japonicus* وفول الصويا والبازلاء، أن أكثر من 200 جين يشارك في عملية تثبيت النيتروجين. ألهمت المعرفة المتراكمة حول عمل هذه الشبكة الجينية العديد من المبادرات الحديثة للانتقال من طور البحث والاستكشاف إلى طور التطبيق، بهدف (أ) تعزيز عملية تثبيت النيتروجين حيويًا في النباتات القادرة على ذلك، مثل البقوليات، (ب) إعادة هندسة عملية تثبيت النيتروجين حيويًا في النباتات التي فقدت هذه القدرة، مثل نباتات طبقة تثبيت النيتروجين غير العقدية، (ج) هندسة عملية تثبيت النيتروجين حيويًا في النباتات التي لم تكتسب هذه العملية مسبقًا، مثل الأرز والذرة.

# ملخص الرسالة

# Acknowledgments

Acknowledgements

First and foremost, I am grateful to Allah; without His mercy and support I would have not reached this stage of my PhD. This journey was not a short road but took five years of my life, filled with both joyful and challenging moments. Thankfully, I was not alone; I was surrounded by many enthusiastic people who continually lifted my spirits and motivated me to keep moving forward. I want to express my deep appreciation to all those who supported me, both professionally and personally, throughout this journey.

**René,** In early 2019, a mutual friend introduced me to your work, and I reached out via email, hoping to master CRISPR-Cas9 gene editing technology in your lab. In your reply, you shared that this technology plays a central role in generating mutants in *Parasponia* to study the evolution of root nodule symbiosis. Not only did you introduce me more to your work, but you also proposed a Skype meeting to discuss it further and explore any overlapping interests. I am grateful that meeting took place, as it led me to realize that studying root nodule symbiosis is not only fascinating but also highly relevant to my country, which suffers from some of the poorest soils in the world and relies heavily on chemical fertilizers in agriculture. That meeting convinced me that this was the path I wanted to pursue, so I gratefully accepted your kind offer to do a PhD in your lab. Throughout my PhD journey, you guided me with patience and kindness, teaching me how to critically analyze my data, communicate my findings effectively in our discussions, and present them confidently at international conferences. On a personal level, you also emphasized the importance of building meaningful connections, which inspired me to actively expand my professional network.

**Ton**, I want to express my gratitude for being my promotor. I have great admiration for your seniority and expertise in science. Your unpredictable questions during the MOLBI Friday afternoon meetings were often tough to answer, but I am deeply thankful for this stimulating environment, which has pushed me to think more critically. I am also grateful for introducing me

to your work on *Indigofera argenteai* and for offering me the opportunity to collaborate and continue this research after completing my PhD.

**Rachid Ibrahim**, I would like to sincerely thank you for accepting to be my external advisor. During our discussions, we reflected on my progress throughout my PhD, and your insights were invaluable. I truly appreciate your help and support during this journey.

**Joel**, we joined René's group just a few months apart. When we first met, you asked if I was from Palestine, since you had googled my name and found someone from Palestine with a similar family name. When I told you I was actually from Saudi Arabia, you jokingly started calling me "Sheikh," even though I am definitely not one—LoL. Not long after that, we collaborated on the *SYMRK* manuscript, which we successfully published in 2023. A year later, we collaborated together again on the *NSP2* manuscript, and we managed to get it published just two months before my PhD defense, even though you were busy with your Marie Curie fellowship and planning your wedding second celebration in China. I am deeply grateful for all your support throughout my PhD, and I wish you a great success in both your professional and personal life.

**Rik**, I still remember the first day we met. We shook hands, you introduced yourself, and you said we will have fun working together. Just a few days later, we sat down to design some constructs for my experiments in chapters 2 and 3. In the lab, our benches were side by side, and you always took the time to demonstrate new techniques to me, showing me how to do things I had not tried before. Whenever I came to you to quickly check my results, you always found the time to help, even when you were busy with your own work. I'll never forget when I asked if you could inoculate my plants while I was away in Saudi Arabia, and you easily agreed, even though it was a lot of work. Thank you for everything you did for me, and I can confidently say that I really did have fun working with you.

**Titis**, I am happy to have you around during my PhD. You were always of a help. Whenever I face some problems in my experiments you always were welling to discuss then with me. I still remember when I struggled to generate *tml* mutant in *Parasponia*; you gave me some suggestions to optimize my transformation protocol. I am also grateful for taking care of my newly generated transgenic plants while I was away in Saudi Arabia. I know it took a lot of effort, especially having to refresh them every three days to eliminate the remaining Agrobacteria. I also appreciate your contribution to chapter 5. Thank you so much!

**Jieyu**, I want to thank you for capturing a realistic bright-field image of my *Parasponia* spontaneous nodule, which I was delighted to use in Figure 3 of my *SYMRK* paper. I am also grateful for the many conversations we had during writing my thesis. A particularly memorable one was our one-hour-long talk outside, where you gave me helpful advice on how to approach my discussion chapter. It was really useful, so thank you for that.

**Luuk**, when I first met you, I had the impression that you were someone difficult to start a conversation with. But over time, that impression completely changed. For instance, on our way back to Wageningen from the ENSA meeting in Freiburg, you and I talked continuously for 4 hours while other people were sleeping in the back seat. That day, I realized that you may not speak often, but when you do, you are incredibly knowledgeable and very funny. I also appreciate our multiple discussion about my results that were always insightful, as well as your contribution to Chapter 2 and 5. Thank you!

**Asma**, I am happy to have had you around during my PhD, especially since we were the only Arabic speakers in MOLBI. I truly value our friendship and food we had together at your place with **Maha**, and **Wasim** who I am happy to meet during my PhD journey. I also greatly appreciate you taking on the responsibility of repeating the mycorrhization experiment for the *clacata1* mutant in *Parasponia*. Thank you so much for that!

My paranymph, **Jundi**, we started our PhD journeys almost at the same time. Over the years, we often sat down to discuss our progress, sharing both the happy and stressful moments of our projects. I have truly valued our business-oriented conversations and look forward to the opportunity to collaborate in the future. I am grateful for the chance to taste the special food from your hometown at the Uyghur restaurant in The Hague, as well as for the gift you brought me from China. Most of all, thank you for agreeing to stand by my side and share the stage with me on my defense day. I wish you all the best in the final stretch of your PhD, and I appreciate your valuable friendship.

My paranymph, **Thomas**, I first got to know you during your master's thesis in our group, and then, two years later, as a fellow PhD candidate. Your character is truly unique; you have an incredible ability to make friends with everyone and a gentle influence on those around you. Like some of our labmates, they have embraced your style and begun wearing earrings as well! Thank you for agreeing to be my second paranymph and for sharing the stage with me on my defense day. I wish you all the best as you continue your PhD journey.

**Siqi and Yueyang,** we began our PhDs around the same time, and I am delighted that we are also finishing at the same time. It has been a pleasure to have you as colleagues and friends. We have learned so much from each other and shared valuable information to benefit our work. I wish you both great success in your professional and personal lives.

The former and current Parasponia/SYM team members; **Simon**, **Yuda**, **Lucas**, **Harm**, **Min**, **Pugang**, **Wenying, Jing, Erik, Xiaofan**, **Zihang**. **Simon**, thank you for trusting me to take care of your plants while you were on your summer holiday. I wish you success in your PhD. **Yuda**, I greatly appreciate your contribution to Chapter 2, which we published together in *BMC Plant Biology*. I wish you great success in your future career. **Lucas**, thank you for trusting me to drive your car in Belgium, even though I didn't have my driver's license with me! I truly enjoyed our friendship as colleagues, and I wish you the best in your PhD. **Harm**, you've been part of our team since your bachelor thesis, and sitting next to or across from each other has always led to great conversations and laughter. I wish you lots of success on

your PhD journey. **Min**, I truly appreciate your friendship and the enjoyable chats we had throughout my PhD. Wishing you all the best and great success in your future career. **Pugang**, even though you joined our group more recently, I've enjoyed our time working together in the office on weekends. **Wenying**, since joining our group, you've been a kind officemate, even when my things sometimes (unintentionally!) spill over onto your desk. Thank you for your patience, and I wish you success in your PhD. **Jing**, our conversations often take an interesting dive into Middle Eastern politics; have you ever considered a career in that field? I see real talent in you! **Erik**, I appreciate your insightful questions during our work discussion. They were always challenging but open up new horizons for me, pushing my thinking to the next level. **Xiaofan**, before you move you office place, we used to set next to each other, and I have fond memories of chatting with you and your husband who would sometimes accompany you while you worked in the lab on weekends. **Zihang**, I wish you great success in your PhD journey. Thank you for the meaningful conversations we had from time to time; I enjoyed them.

The current and former MOLBI lab technician, **Marijke**, **Caroline**, **Federico, and Jacqueline. Marijke**, I am so grateful for your support at the beginning of my PhD. You were incredibly kind and patient, teaching me how to handle *Parasponia* in tissue culture and guiding me through the transformation process. I also appreciate the time you invited me, my wife, and my son to meet your wonderful mother; we truly enjoyed that visit. **Caroline**, thank you for your assistance in the lab, especially when you calibrated my pipettes twice! I also really appreciate your help in finding *Indigofera argenteai* seeds. I'm looking forward to germinating them together soon. **Federico**, even though you've just joined our group to work on cassava, I already wish you great success in your work. **Jacqueline**, even though you worked in our group for only a short time, it was enough for me to see how kind you are. Thank you for the enjoyable conversations we shared!

**Maria** and **Marie-José**. **Maria**, thank you so much for your steady support throughout my PhD journey. You were always there when I needed administrative assistance, and I appreciate your dedication and kindness. **Marie-José**, I am deeply grateful for all the help you provided with my PhD-related financial matters. You ensured course fees were paid, even when I was running behind, and you were always there to assist with arranging

residency paperwork for my family and me. Without your support, managing these things on my own would have been challenging. Your assistance over these years at MOLBI has meant so much to me; thank you.

I am grateful for my BSc/MSc students which I supervised during my PhD journey: **Simon** and **Nick**. I would like to tell you that I enjoyed working with you, and I learned a lot from you. I wish you a bright future.

All staffs, former and current members of MOLBI/PDB/Cell Biology; **Joan**, **Renze**, **Viola**, **Tijs**, **Ronald**, **Wilma**, **Michael N**, **Andres**, **Gabriela**, **Richard**, **Michael S**, **Henk**, **Olga**, **Defeng**, **Huchen**, **Peng**, **Jelle**, **Peter van G**, **Michiel L**, **Jiancun**, **Pierangela**, **Bowei**, **Sanne**, **Siddhant**, **Alysha**, **Kyra**, **Ton**, **Xu**, **Martinus**, **Amina**, **Jiangyong**, **Renan**, **Jeroen**, **Menno**, **Anneke**, **Vera, Peter G**, **Wouter**, **Francesca**, **Sophia**, **Joyce**, **Kavya**, **Amber**, **Honglei**, **Merijn**, **Norbert**, **Alejandra, Kiki**, **Hannie**, **Andrea**, **Otto**, **Chris**, **Kelvin**, **Yvet**, **Isha**, **Eline**, **Jordi**, **Max**, **Harald**, **Michiel K**, **Roman**, **Dichuan**, **Fengjiao**, **Tristan**, **Jana**, **Zhuang**, **Jiawei**, **Zifan**, **Carolin G**, **Liu, Alessio, Tijs S, Haolin** and **Ben.** I truly enjoyed having all of you around during my PhD journey. Thank you for creating such a comforting and supportive environment; it made a real difference. **Renze, Otto, and Haolin,** I appreciate the interesting discussion we had during the research proposal brainstorming meeting. We came up with many interesting ideas, I hope we can collaborate on one of them in the future. **Michael S**, it's been nice knowing you. We shared the experience of living in one of the top party cities in the United States (Iowa City) and also shared co-authorship on the SYMRK paper. Thanks for the good memories and collaboration! **Amina,** thank you for warmly introducing me to the lab and for showing me how to use the pipette tip refill system for the first time. **Wouter,** thank you for your thoughtful questions during our work discussions and for your valuable contributions to the SYMRK paper. I really appreciate your support, especially when helping to address the reviewers' comments during the rebuttal process. It was great to have you as a co-author on that paper! **Norbert,** thank you for inviting me to your home for dinner with you and your wife while my family was a way. I really enjoyed your company and hospitality. **Henk,** you are not only incredibly funny but also genuinely kind. I remember when I brought my 6-month-old son to the lab; despite usually being restless around unfamiliar people, he stayed calm and comfortable with you, even in my absence. I feel fortunate to have gotten to know you, and I wish you a wonderful life.

To my friends in the Netherlands, **Rashad,** I am grateful to our mutual friend **Abdelfattah** for introducing us. We first met in person in 2018 in Al-Hofuf, Saudi Arabia, while you were working at Wageningen University and collaborating with KFU. You encouraged me to apply for a PhD position at Wageningen University, and you kindly connected me with **Eric Schranz**, who then introduced me to **René Geurts**. René graciously accepted to be my co-promoter, giving me the opportunity to join his group and pursue my PhD. I truly appreciate your help in securing this position and your invaluable advice to start searching for an apartment before moving to Wageningen, as finding one can be quite challenging. I thank Allah for guiding me to meet you; without your support, I would not be where I am today. All I can say is: "شكراً برشا."**Jouke,** you are one of the nicest people I have ever met in my life. I am grateful for the DEMI water adventure you and **Ilias** were on when we crossed paths. What a coincidence it was to meet you that day! You were so surprised to see a Saudi in Wageningen that you ran back to your office, and a few minutes later, you brought another colleague just to show him the Saudi you had unexpectedly found (hahaha). That moment was truly hilarious. **Ilias**, I remember how you asked me where I was from the day we met. Without that question, we might not have developed such a great friendship. You are such a kind and likable person. Both of you—**Jouke** and **Ilias** — have invited me and my family to your homes for dinner, where we got to meet your lovely families. I was equally happy to have you both accept my invitation to try Saudi food in my apartment. Those wonderful moments we shared together will always stay with me. Thank you for that!

Aan mijn buren in Belmonte, **Els Koek**, ik wil je bedanken voor het geweldige buur zijn en voor de vriendschap die je hebt geboden aan mijn vrouw, **Fatimah**, en als een lieve "oma" voor mijn zoon, **Abdulrahman**. In de beginperiode van onze reis in Nederland begeleidde je **Fatimah** op wandelingen in het park en tijdens winkeltrips in het centrum, terwijl je haar hielp met het oefenen van haar Engels. Je steun strekte zich ook uit naar mij, toen je als mijn advocaat optrad om me te helpen de Nederlandse zorgverzekering te begrijpen. Omdat ik als inwoner van Nederland werd beschouwd, moest ik overgaan van een internationale naar een lokale zorgverzekering, en met jouw ervaring hebben we de zaak gewonnen. Ik bewonder hoe je het persoonlijk maakte, jezelf uitdaagde om te bewijzen dat je juridische vaardigheden nog steeds scherp waren, zelfs na zoveel jaren van pensioen. Ik wil je ook bedanken, samen met je man, voor de uitnodiging

voor het diner in jullie mooie appartement terwijl **Fatimah** op vakantie was in Saudi-Arabië. Jullie vriendelijkheid en steun betekenen veel voor mij. Ik zou je graag uitnodigen om ons ooit in Saudi-Arabië te bezoeken, zodat we je de schoonheid van ons land kunnen laten zien. **C.E. de Boer** en **C.B. de Boer-Mijnhart**, ik wil u graag mijn oprechte dank betuigen voor uw vriendelijkheid tegenover mij, mijn vrouw **Fatimah** en onze zoon **Abdulrahman**. Ondanks de taalbarrière, omdat wij geen Nederlands spreken, hebt u ons altijd met warmte ontvangen. U nodigde ons uit in uw huis voor koffie en diner en zorgde zelfs dat er iemand aanwezig was om te vertalen, zodat we elkaar beter konden begrijpen. We zijn vooral dankbaar voor de attente cadeaus die u ons gaf, vooral die voor onze kleine jongen. Ze gaven ons het gevoel echt welkom te zijn in Belmonte, zelfs wanneer sommigen misschien minder blij waren met het idee van een kind in de buurt. De avonden waarop u zwaaide om mij welkom te heten na een lange dag op het werk, betekenden veel voor me. Ik voel me gelukkig dat ik u als buren heb gehad, en ik wens u beiden gezondheid en een lang leven. **B.W. van Druten**, ik ben dankbaar dat ik jou als buur heb, ook al ben je pas recent in je appartement komen wonen. Je bent altijd vriendelijk geweest tegen mijn vrouw **Fatimah** en mijn zoon Abdulrahman. Bedankt voor het cadeau dat je aan mijn zoon hebt gegeven, vooral de schildpad die veel voor je betekende. Ik ben dankbaar voor de avonden waarop je naar ons zwaaide wanneer ik na mijn werk met mijn zoon ging wandelen; dat betekende echt veel voor me. Ik voel me gezegend dat ik jou als buurman heb gehad, en ik wens je een gezond en lang leven toe. **M. Huijskes**, Ik ben blij dat je mijn buurman bent. Jij en mijn vrouw, **Fatimah**, zijn goede vrienden. Bedankt voor het uitnodigen van haar voor een kopje koffie in je huis. We zijn ook blij je bij ons thuis te ontvangen voor een kopje koffie. Ik ben dankbaar voor de cadeaus die je voor mijn zoon, **Abdulrahman**, hebt gegeven. **Silvia Jacobs**, je bent een geweldige secretaresse met een unieke en vriendelijke houding. Heel erg bedankt dat je voor ons huis hebt gezorgd toen we weg waren, vooral tijdens onze lange vakantie in Saudi-Arabië. Ik ben ook dankbaar voor je hulp bij het reserveren van de grote zaal in ons gebouw voor mijn feest na de verdediging. Het was heel attent van je om contact op te nemen met de verantwoordelijke persoon en de gebruikelijke wekelijkse activiteit voor die avond te annuleren, zodat ik de ruimte de hele avond kon gebruiken. Jouw steun betekent echt veel voor mij.

To my Saudi friends nearby Wageningen**, Mohammed Almansouri,** I am grateful to **Mostafa** for introducing us, which led to a wonderful friendship during the four years of my PhD journey. You were the first Saudi person I got to know in the Netherlands, and for more than a year, you remained my only Saudi friend here. This past years are full of moments that I consider some of the happiest I have had in the Netherlands. We went on numerous family trips together to several countries—Germany, Belgium, Poland, Austria, Slovakia, Hungary, France, Switzerland, to name a few. These travels gave us countless funny and memorable moments, like our spontaneous drive to Kleve, Germany, late in the evening just to buy halloumi burgers (hahaha), and the rush to make it home before curfew started during the COVID-19 lockdown. I am also deeply grateful for all the times you helped me, like when I was buying my car—you drove me to The Hague three times in a row and were willing to go a fourth if necessary. I truly appreciate your kindness and generosity. I would also like to express my sincere gratitude to your wife, **Eman**, and your children, **Romysa** and **Yousef**. Thank you, **Eman**, for your friendship with my wife, **Fatimah**, and for staying over at our apartment to keep her company when I was away in England. That meant so much to me. **Romysa**, thank you for being patient as I struggled to pronounce your name correctly at first (calling you "Romysah"☺ repeatedly). Even though you often teased me by saying you liked "**Mohamed Alharbi**" more than me, I always found it adorable. **Yousef**, when I first met you, you were only six months old, and I had the pleasure of watching you grow into a curious and active little boy. You even bravely was about to jumped over a bridge in Poland! Playing with you **Romysa** and **Yousef** in the playground at the Roermond Outlet gave me some of the best moments, and it prepared me well to play with my son, **Abdulrahman**, now.

**Mohamed Alharbi**, what a privilege it has been to know you and your family. I am grateful for that day when your wife, **Manal**, and my wife, **Fatimah**, both felt the urge to have a cup of coffee from the same coffee house, which led to them meeting each other. From that moment, they became friends and eventually introduced us. You became the second Saudi I got to know in the Netherlands and have always been a special friend to me. I have learned from you how to be kind and patient. You never get bothered by anything and always look at the positive side of any situation. I even once asked you if you have ever felt anger, and your expected answer was a confident "No." I am also grateful for the books you bought for me from Saudi Arabia, even

though people asked you why you were buying books by authors considered liberal in our conservative society. Finally, I would like to thank **Manal** for her friendship with **Fatimah** which has created some unforgettable moments in our journey in the Netherlands.

**Ayman Iskandrani**, I am so glad to have gotten to know you. You are a genuinely funny person, you have got the most colorful vocabulary I have ever heard; truly a poetic masterpiece! (hahahah). Thanks to your wife ,**Khadija**, and son ,**Qays**, for the joyful moments they shared with my family and for the thoughtful gifts they gave to my son.

**Yasser Alhejji**, I appreciate your friendship and the wonderful photo you took of me during our outing. I wish you great success in your PhD, and I hope we stay in touch in the near future.

To my other Saudi friends all around the Netherlands, **Abdulrahman Alasiri**, I really appreciate the time we spent together in the Netherlands. You are a very likable person, funny, and an excellent cook. I learned a lot from you, especially about how to respect the minds of children. I saw this in the way you interact with your son, **Amar**, and your lovely daughter, **Dareen**, during our regular outings. I am also grateful for your wife, Amal, and her friendship with my wife, **Fatimah**, and for staying over at our apartment to keep her company when I was away in England. **Mohammed Alasiri**, I am so happy to have gotten to know you. You are a funny and sarcastic person, always managing to deliver your witty remarks without ever being impolite. I still remember that one time at the Afghan restaurant when you couldn't read the menu because it was so dirty. Instead of complaining, you sarcastically thanked the waiter for how clean the menu was; I laughed about that for the entire day (hahaha). Congratulations on defending your PhD a few months ago, and I wish you all the success in your future career. **Abdulrahman Alshalani**, I want to thank you for the great times we shared. You are a very likable person, and I truly enjoyed your company, especially during our trip to Germany. I wish you all the best in your future endeavors. **Ali Alsamil**, although we didn't have many chances to sit and talk, the few times we did meet left a lasting impression on me. You are one of the kindest people I've encountered during my PhD journey. I wish you every success in life. **Khalid Alameer**, Thank you for organizing the Saudis in the Netherlands outings. We had so much fun during those gatherings, even though we often

found ourselves arguing a lot. I hope we catch up again in the near future. I would like to thank some more of my Saudi friend **Ahmed Zelfawei**, **Naif Farhan**, **Abdulrahman Qari**, **Elias Hawsawi.**

To my dear Saudi friends in the outing group recently renamed "Arykah group☺". **Fouad Asiri**, I first got to know you through your interview on the Thmanyah broadcast, and I was fortunate to meet you in person. You are incredibly knowledgeable and have a youthful spirit. I truly appreciate your friendship throughout my PhD, even though you always tease me; especially when I was entrusted by Hisham Abu-Tawil to be his secret keeper (hahaha). I wish you nothing but success in your future endeavors. **Hisham Abu-Tawil**, although I've only known you for about a year, it honestly feels like we have been friends for many years. You are humble and incredibly likable. The past two months spent with you have been more memorable than the entire five years of my PhD. You trusted me with your great secret (don't worry, no blackmail here, haha), and the ongoing, humorous debates with our mutual friends as they tried to persuade me to spill the beans were some of the funniest moments I have had. I laughed more in those two months than I did throughout my entire PhD journey. Thank you for your friendship, and I wish you great success in completing your PhD. Oh, and I can't forget to mention—you are a fantastic cook! The "makshan samak" you made last time was absolutely perfect☺. **Mohammed Alharthi,** I truly appreciate your friendship. You are a very quiet person and a great listener, but when you do speak, you are wise and knowledgeable. I wish you all the best in your PhD journey. **Ahmed Alhammad** you are such a funny and likable person. I still remember the first time we met when you jokingly suggested sharing your school calendar with everyone, so we could plan our meetings around your important dates. I found that absolutely hilarious. I also enjoyed making jokes about who might be divorced, and I remember how you always wrongly picked on **Fouad Asiri**, pretending he was divorced (hahaha). Thank you for your friendship and I wish you success in your future. **Mohammed Hamzi**, I really appreciate your friendship during the final stretch of my PhD. You are well-informed, especially when it comes to poetry and history. Like **Fouad Asiri**, you always found joy in teasing me—especially to a whole new level after I became the keeper of **Hisham Abu-Tawil**'s secret. I wish you all the best in your PhD journey. **Abdullah Aljifri**, I would like to thank you for your friendship. You organized our first outing in Rotterdam, which allowed me to meet many wonderful Saudi friends. Congratulations on recently

defending your thesis and securing a postdoc position in your group. I wish you great success and a bright future ahead. **Ammar Banafea** it is often said that Saudi students affiliated with Aramco tend to form small, close-knit groups. You, however, have been a remarkable exception; you always enjoyed connecting with other Saudi students, and you even invited some of us over for dinner at your place. I am so glad to have gotten to know you, and I wish you all the success in your future. **Mohammed Alqhtani**, I got to know you during the final stretch of my PhD. You are a very kind person, and I am happy to add you to the list of wonderful Qahtani friends I've had the pleasure of knowing. I wish you all the best in your PhD journey.

**رسالة شكر وعرفان لكل من ساهم في هذا النجاح**

**إلى أبي وأمي،**
شكرًا لكما على الدعم الذي لم ينضب منذ نعومة أظافري وحتى هذه المرحلة من حياتي. ربيتماني وأحسنتما تربيتي، واليوم تقطفان بعض ثمار جهودكما بهذا النجاح. كنتما دائماً خير عون لي للوصول إلى ما أنا عليه اليوم. لا يمكنني أن أنسى حنانكما ووجودكما حولي، ودعواتكما التي تمطراني بها في كل وقت، والتي تزداد غزارتها في أوقاتي العصيبة. لولا رحمة الله ثم دعواتكما، لما كان لهذا النجاح أن يرى النور. فشكراً لكما من أعماق قلبي، وأطال الله في عمركما، وأمدكما بالصحة والعافية.

**إلى أفراد عائلتي،**
شكراً لكم جميعاً على ما قدمتموه لي من محبة ودعم في جميع مراحل حياتي، كان له الأثر الكبير في وصولي إلى ما أنا عليه اليوم. أسأل الله أن يمنح من بقي منكم حياً التوفيق والنجاح، وصلاح النفس والذرية، ولمن انتقلوا إلى دار الآخرة الرحمة والمغفرة.

**إلى أخي عبدالرحمن وعائلته،**
شكراً لك ولعائلتك على حسن وكرم الضيافة طوال أولى مراحل دراستي الجامعية. بعد توفيق الله، ثم دعمكم اللامحدود في تلك الفترة المفصلية من حياتي وما تلاها، وصلتُ إلى هذا النجاح. فشكراً لكم على ما قدمتموه لي.

**إلى كل من ساهم في هذا النجاح بدعوة، أو بنصيحة، أو تجشم عناء حضور حفل التخرج،**
شكراً جزيلاً لكم، وأسأل الله أن يثيبكم على ما قدمتم، وأن يجزيكم خير الجزاء في الدنيا والآخرة.

To conclude, but most importantly, I want to express my deepest gratitude to my beloved family—my wife, **Fatimah**, and our son, **Abdulrahman**. Our journey as a family began when **Fatimah** and I got married in July 2018 in Taif, Saudi Arabia. Marriage was a new milestone for us, and shortly after, we moved to Al-Hofuf, where I spent a year working at King Faisal University. In early 2019, our family made the important decision to move to the Netherlands after I was accepted to pursue a PhD in **René Geurts**' group at Wageningen University. Preparing for the move took months of searching for an appropriate apartment, and ultimately, we secured the only one available to us in Wageningen—a small place in Belmonte. Our journey to the Netherlands began on October 20, 2019, and we had no idea where this adventure would lead us. On our arrival, we were captivated by the beauty of the Netherlands in the autumn, with thousands of leaves scattered across the ground and sidewalks. It gave us a sense that we would enjoy our time in Wageningen, and indeed, the first three months were wonderful. However, in February 2020, the Netherlands was struck by the COVID-19 pandemic. Multiple lockdowns and curfews forced us to isolate ourselves in our small apartment. Despite the challenges, we bravely kept ourselves safe. A year and a half later, COVID-19 restrictions were completely lifted, and we resumed our normal lives, exploring nearby cities and even venturing to neighboring countries.
In late 2022, our family was blessed with a beautiful new addition, our son **Abdulrahman**. He is the most incredible gift, and everyone remarked that he is a "copy-paste" of me. Watching him grow and develop into a courageous, funny, and adorable little boy has been one of my greatest joys. **Fatimah** and **Abdulrahman**, I am incredibly thankful to Allah for blessing me with both of you in my life.

**Fatimah**, I used to think that studying abroad as a single person would be easier since I wouldn't have responsibilities beyond my own. But with you, I realized how wrong I was. True, being with you meant new responsibilities, but I was only seeing the half-empty side of the cup. I never anticipated how beautifully you would take over so many of those responsibilities. You never complained about managing the household—handling groceries, keeping everything running smoothly, and making our home feel like my own personal restaurant, especially when you'd announce, "On the menu today, we have this and that." Those moments truly meant a lot to me. Even during

times when I struggled with a recurring upset stomach, you were incredibly kind, always finding creative ways to help me feel better.

You always wanted to continue your education and pursue a master's degree in the Netherlands. To do this, you needed to improve your English, so you traveled to Amsterdam three times a week for intensive English courses, as there was no nearby institute offering them. I watched your English skills improve so quickly, and you soon reached a level that qualified you for acceptance at a Dutch university. But as I was nearing the end of my PhD, and we welcomed **Abdulrahman** into our lives, you chose to postpone your plans until we return to Saudi Arabia. Thank you for putting our family's needs first; your selflessness is deeply appreciated.

Throughout this journey, you took on so much—caring for **Abdulrahman** all day. Even when I came home exhausted or unwell, you would still take charge of keeping him occupied so I could get a moment to rest. After the grace of Allah, I owe the completion of this journey to your unwavering support, love, and care. Words are not enough to express my gratitude, but what I can say is: THANK YOU!

**Abdulrahman**—or "**MaMan**," as you like to call yourself; you should know that I am so proud of you. Having you in my life has been, and will always be, the most beautiful gift. Your arrival in the final stretch of this journey brought new energy and made me rethink my balance between work and life, so I could savor moments with you. Whether it's playing with you as you ride on your bike—your "catati," as you like to call it—or reading your favorite books *London Bus Cloth Book* and *Gruffalo, Where Are You?*, these moments are priceless. Words can't fully capture my love and care, but what I can say is: *Is MaMan Habibi? And today, there's no room for a "no" from you!*(☺)

# About the author

Sultan Alhusayni was born in April 1990 in a small village near Taif, Saudi Arabia, where he completed his early education. In 2008, Sultan enrolled in the Biology Department at Taif University and earned his Bachelor's degree in 2011. In recognition of his academic potential, Sultan was awarded the prestigious King Abdullah Scholarship (may Allah rest his soul in peace) in 2012, which enabled him to pursue his education in the United States.

Sultan completed his Master's degree in Biology at the University of Central Missouri, Missouri State, in 2015, graduating with excellent academic standing. Building on this, Sultan transitioned from science to the hospitality industry in early 2016 and worked at the Movenpick Hotel in the holy city of Makkah until late 2017. During this period, he developed important skills in management and communication.

In late 2017, Sultan was selected as a lecturer in the Department of Biological Sciences at King Faisal University in Al-Ahsa, Saudi Arabia, where he contributed to the teaching and mentorship of undergraduate students for 14 months. His passion for research and education led him to pursue a PhD at Wageningen University & Research in the Netherlands, where he joined the group of René Geurts. His doctoral research focused on plant-microbe beneficial interactions, specifically nitrogen fixation in *Parasponia andersonii*. Sultan's work has led to the publication of two chapters of his thesis in peer-reviewed journals.

As an expert in molecular biology, Sultan has presented his research at international conferences, contributing to the global scientific community. He was also a participant in the ENSA (Enabling Nutrient Symbioses in Agriculture) consortium, where he collaborated with leading scientists to

advance knowledge in sustainable agriculture and plant-microbe interactions.

Sultan's academic and professional journey reflects his dedication to research and education. His work in molecular biology and plant science contributes to a better understanding of agricultural practices, with an emphasis on sustainable solutions through scientific research.

# List of publications

**Alhusayni, S. *et al.*** (2023) 'A rare non-canonical splice site in Trema orientalis SYMRK does not affect its dual symbiotic functioning in endomycorrhiza and rhizobium nodulation', *BMC plant biology*, 23(1), p. 587.

**Alhusayni, S. *et al.*** (2024) 'Ectopic expression of the GRAS-type transcriptional regulator NSP2 in Parasponia triggers contrasting effects on symbioses', *Frontiers in plant science*, 15. Available at: https://doi.org/10.3389/fpls.2024.1468812.

**Alhusayni, S. *et al.*** (to be submitted) '*Parasponia* Mutant Analysis Revealed That *CLAVATA1* Regulates Symbiotic Interactions Beyond Legumes'

**Alhusayni, S.** (to be submitted) 'A Comprehensive Review of Autoregulation of Nodulation (AON): Genetic Pathways and Evolutionary Insights into Plant Symbiosis'

**Colophon**

The research described in this thesis was financially supported by ENSA Project funded by Bill & Melinda Gates Foundation to the University of Cambridge to the Wageningen University to Rene Geurts (OPP1172165), and the Ministry of Education, King Faisal University, Saudi Arabia to Sultan Alhusayni (10598).

Financial support from Wageningen University for printing this thesis is gratefully acknowledged.

**Cover design**: Sultan Alhusayni and Stefanie van den Herik | HerikMedia
**Layout**: Sultan Alhusayni
**Printed by**: Digiforce – Proefschriftmaken, Ede, NL