

*Divergence of Rhizobium-Induced  
Cytokinin Signalling in  
Nodulating Species*

**Titis Anggraeny Kusuma Wardhani**

# Propositions

1. *Parasponia* should be considered as an excellent model to study the nodulation trait alongside the legume species.  
(this thesis)
2. A conserved mechanism is not necessarily based on orthologous genes.  
(this thesis)
3. The ability to adapt to the environment does not make plants intelligent (Trewavas A.J., 2003, *Annals of Botany* 9: 1–20).
4. Measurements on facial expressions do not fully reflect the emotional state in mammals (Dolensek et al., 2020, *Science* 368, 89–94).
5. The current global pandemic of COVID-19 contributes to the acceleration of pseudo-science.
6. Stereotypes limit the ability to identify an individual's uniqueness.
7. Being able to solve global issues should be considered a privilege.

Propositions belonging to the thesis, entitled:  
“Divergence of Rhizobium-Induced Cytokinin Signalling in Nodulating Species”

Titis Anggraeny Kusuma Wardhani  
Wageningen, 16 December 2020

# Divergence of Rhizobium-Induced Cytokinin Signalling in Nodulating Species

**Titis Anggraeny Kusuma Wardhani**

## **Thesis committee**

### **Promotor**

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 F.P.M. Govers, Wageningen University & Research  
Prof. Dr T. Ott, Universität Freiburg, Germany  
Dr D. Reid, Aarhus University, Denmark  
Dr R. Karlova, Wageningen University & Research

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

# Divergence of Rhizobium-Induced Cytokinin Signalling in Nodulating Species

**Titus Anggraeny Kusuma Wardhani**

## **Thesis**

submitted in fulfilment of the requirements for the degree of doctor  
at Wageningen University

by the authority of the Rector Magnificus,

Prof. Dr A.P.J. Mol,

in the presence of the

Thesis Committee appointed by the Academic Board

to be defended in public

on Wednesday 16 December 2020

at 4.00 p.m. in the Aula.

Titis Anggraeny Kusuma Wardhani

Divergence of Rhizobium-Induced Cytokinin Signalling in Nodulating Species,  
256 pages.

PhD thesis, Wageningen University, Wageningen, the Netherlands (2020)

With references, with summaries in English

ISBN: 978-94-6395-576-8

DOI: <https://doi.org/10.18174/532477>

*This thesis is dedicated to my family*



# Table of Contents

<b>Chapter 1</b>	General Introduction	<b>8</b>
<b>Chapter 2</b>	Dissecting the Role of Cytokinin Signalling and NIN during nodulation in <i>Chamaecrista mimosoides</i>	<b>28</b>
<b>Chapter 3</b>	Transforming, Genome Editing and Phenotyping the Nitrogen-Fixing Tropical Cannabaceae Tree <i>Parasponia andersonii</i>	<b>78</b>
<b>Chapter 4</b>	CRISPR/Cas9-Mediated Mutagenesis of Four Putative Symbiosis Genes of the Tropical Tree <i>Parasponia andersonii</i> Reveals Novel Phenotypes	<b>110</b>
<b>Chapter 5</b>	The <i>Parasponia andersonii</i> Cytokinin Receptor HISTIDINE KINASE3 is Required for Intracellular Rhizobium Infection of Nodule Cells	<b>152</b>
<b>Chapter 6</b>	General Discussion: Synergy of Cytokinin Signalling and Transcriptional Signalling Networks in Symbiotic Rhizobia Association	<b>188</b>
<b>Reference</b>		<b>204</b>
<b>Summary</b>		<b>226</b>
<b>Appendices</b>		<b>242</b>



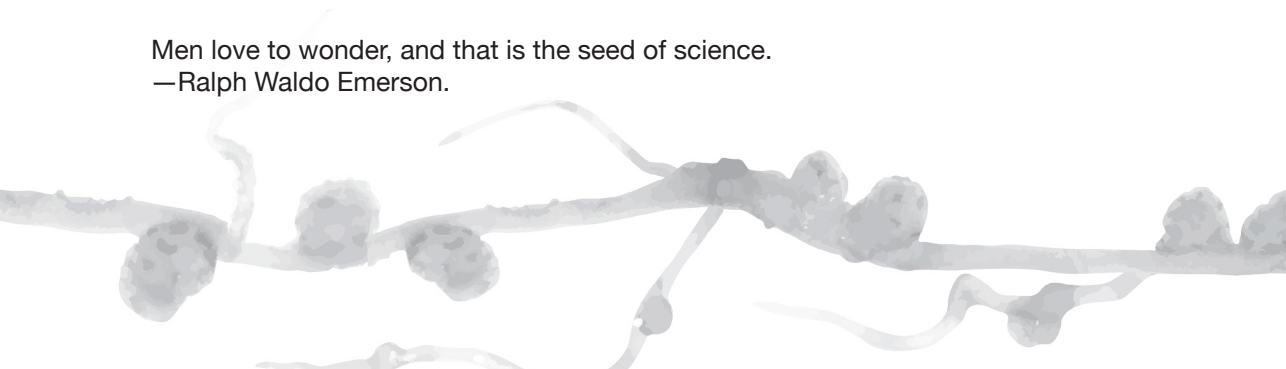
# Chapter 1

## General Introduction

**Titis A. K. Wardhani<sup>1</sup>**

<sup>1</sup>Laboratory of Molecular Biology, Department of Plant Sciences, Wageningen University, Droevendaalsesteeg 1, 6708 PB Wageningen, The Netherlands

Men love to wonder, and that is the seed of science.  
—Ralph Waldo Emerson.





## Introduction

Integration of sustainable agricultural practices is an important strategy to mitigate environmental risks. In modern agriculture, manufactured fertilisers and massive application of synthetic fertilisers to the soils are costly both economically and environmentally, and the forecast to cope with ever increasing food demands is not prosperous. In nature, mutualism between plants and microorganisms are widespread and provide vital contributions to ecosystems. Root association with arbuscular mycorrhizal (AM) fungi is the most pervasive symbiosis, which evolved ~450 million years ago (MYA) (reviewed by (Parniske, 2008). In this symbiosis the fungus supplies nutrients, such as phosphates, to the host plants. ‘A more recent evolution’ (~110 MYA) is an interaction between plants and nitrogen-fixing bacteria. This mutualistic interaction leads to the formation of specific organs on the root of the plant; so-called root nodules. In these nodules the bacteria find optimized physiological conditions to convert atmospheric dinitrogen ( $N_2$ ) into ammonia. To date, the nitrogen-fixing nodule symbiosis can occur in ten taxonomic lineages either with the gram-positive filamentous *Frankia* bacteria, or the gram-negative rhizobia (Soltis et al., 1995; Werner et al., 2014; Geurts et al., 2016). To this end, it allows the host plants to grow in the nitrogen-deprived soils without depending on the fertilisers.

Both, the AM and nitrogen-fixing symbioses, lead to intracellular hosting of the microorganism in the plant root cells. Such intracellular hosting will allow optimized nutrient and energy exchange between the microsymbiont and the host plants. However, to establish such an interaction coordinated signalling exchange between both organisms is required (Oldroyd, 2013; Delaux et al., 2015; Geurts et al., 2016). Upon an AM symbiosis, a specialised membrane in root cortical cells hosts arbuscules; the branching hyphae of AM fungi. Similarly, in the nitrogen-fixing symbiosis the bacteria are surrounded in a plant derived membrane, either in the form of a thread (so-called fixation thread), or as individual bacteria. The latter is only occurring with rhizobia that infect legumes. The bacteria inside nodule cells have an organelle-like appearance, and are named symbiosomes (Parniske, 2008; Oldroyd et al., 2011).

Although the nitrogen-fixing symbiosis is considered as one of the most efficient systems for a plant to obtain fixed nitrogen when it is not available in the soil, the nodulation trait is restricted to a subset of genera belonging to a paraphyletic clade in angiosperms; ‘the Nitrogen-Fixing Clade’. The Nitrogen-Fixing Clade comprises four related orders; Fabales, Fagales, Cucurbitales and Rosales, which are dispersed into ten families. Species belonging to eight families engage in the symbiotic interaction with *Frankia*, and thereby are classified as actinorhizal plants. Species of two additional families nodulate with rhizobium; namely legumes (Fabaceae, Fabales) and *Parasponia* sp (Cannabaceae,

Rosales) (Soltis et al., 1995; Werner et al., 2014). Legumes are studied in most detail in regards to the endosymbiotic interaction with rhizobia. *Medicago truncatula* (*Medicago*) and *Lotus japonicus* (*Lotus*) are two legume species of the Papilionoideae subfamily that are used as models. Genetic screens (either via chemical, or physical mutagens) have resulted in a large mutant collection that is affected in nodulation. These mutants are indispensable in the discovery of genes involved in the rhizobial-induced nitrogen-fixing symbiosis. In addition, legumes are also economically relevant, as this includes crop species such as soybean (*Glycine max*), peanut (*Arachis hypogaea*), and pea (*Pisum sativum*). Meanwhile, *Parasponia* (Cannabaceae, Rosales), which diverged ~100 MYA from legumes and comprised five species, represent the only non-legume genus that can establish the symbiosis with rhizobia (Becking, 1983; Lancelle and Torrey, 1984; Becking, 1992; Geurts et al., 2012).

Questions on how some plant lineages evolved the nitrogen-fixing nodule symbioses while others did not have been a major driving force for many genetic studies over the last 20 years and resulted in breakthrough discoveries. Forward and reverse genetics analyses in several legumes species have uncovered that some genes are not only essential for nodulation, but also the AM symbiosis (Stracke et al., 2002; Ané et al., 2004; Lévy et al., 2004a; Imaizumi-Anraku et al., 2005; Kanamori et al., 2006a; Saito et al., 2007a; Groth et al., 2010a; Horváth et al., 2011). These include key genes encoding components in the signalling machinery of the plant that are essential to recognize the symbiotic microbe. This suggests that the evolutionary younger nitrogen-fixing symbiosis recruited similar pathways as the much older AM fungi symbiosis (Parniske, 2008). This raises questions on how the existing AM symbiotic pathways were evolutionary rewired to accommodate the nitrogen-fixing symbioses.

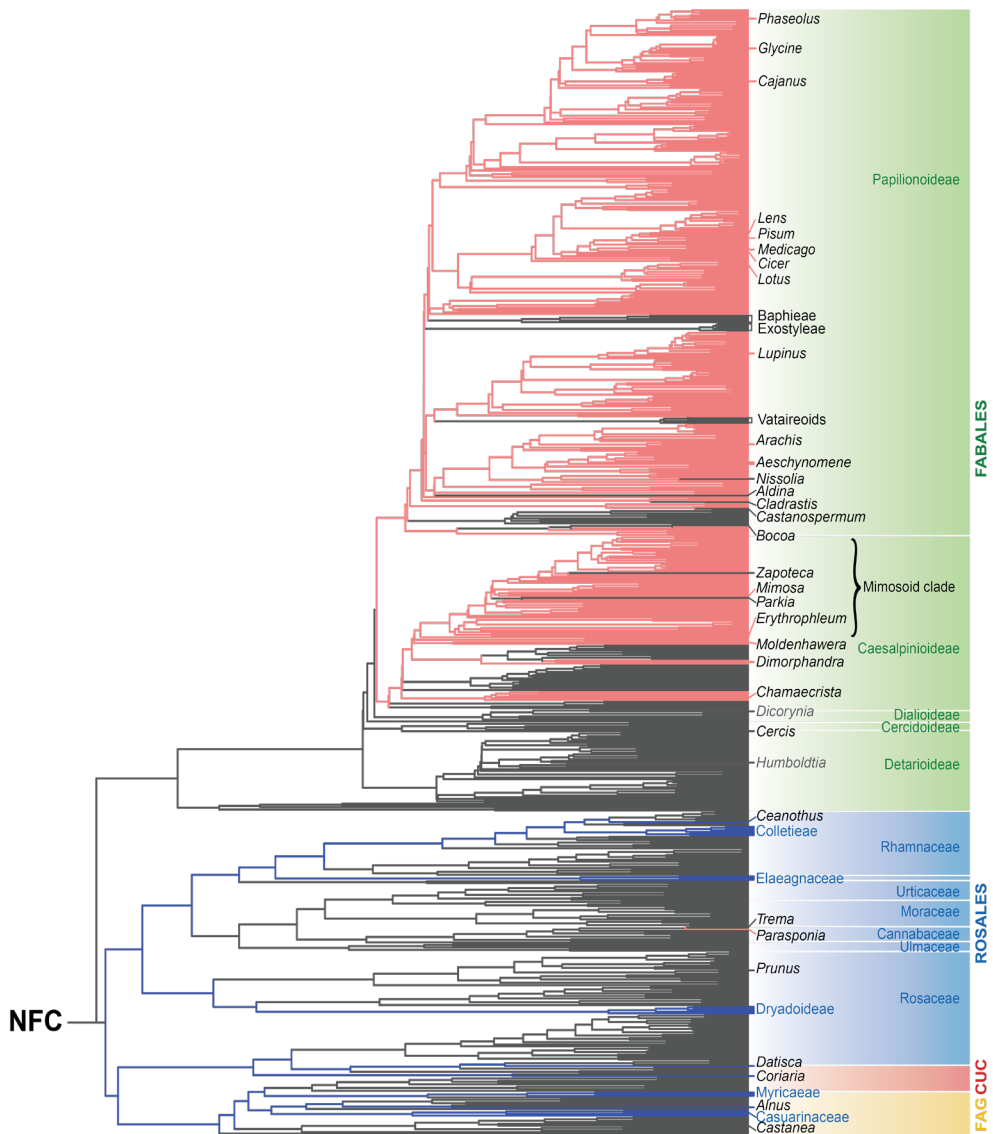
Considering the potential of nitrogen-fixing symbioses in providing the more environmentally fixed nitrogen to the soils, it leads to a question whether it is possible to extend the host range of nitrogen-fixing microbes into the non-nodulating crops (such as wheat, maize, or rice). Apart from the massive body of knowledge from studies in model legumes, this engineering vision remains a highly topical challenge to solve. To this end, it is very important to identify the core set of nodulation genes that is conserved in nodulating species of different taxonomic clades. One of the strategies to identify such a core set of nodulation genes is to expand the genetic studies to non-model species of the Rosales, Cucurbitales and/or Fagales orders. Such a strategy is also relevant in order to avoid legume-biased knowledge on nodulation. Additionally, functional studies in other nodulating clades may offer information on lineage-specific genetic adaptations. In this thesis research, I made use of two such species; *Chamaecrista mimosoides* and *Parasponia andersonii* (**Figure 1**). *C. mimosoides* belongs to the legume subfamily of

Caesalpinioideae, in which the occurrence of nodulating species are less common than in the subfamily of Papilionoideae. *Chamaecrista* also represents the basal node in the divergence of the legume family, and therefore the *Chamaecrista*-based functional studies may provide insights on basal adaptations in the legume-rhizobium interactions. Meanwhile, since *Parasponia* is distantly related with legumes and did not experience a recent genomic duplication, and therefore studying *Parasponia* in the context of rhizobial symbiosis may provide a more accessible platform to study the rhizobial-induced symbioses.

### Evolutionary trajectory on the Nitrogen-fixing endosymbioses

Quantitative and comparative phylogenetics and genomics studies have contributed to our understanding of the evolutionary trajectory on the associative nitrogen-fixing symbioses. Earlier studies postulated that the nitrogen-fixing symbioses had evolved -at least- ten times. Eight times for the symbiosis with actinorhizal bacteria *Frankia*, and two times for the symbiosis with rhizobia (Werner et al., 2015). This hypothesis found support by the fact that the nodule ontogeny of actinorhizal plants and legumes is different, which suggests specific recruitment of molecular mechanisms (Doyle, 1998; Swensen, 2010). As nodulating species are limited to only the Nitrogen-Fixing Clade, predisposition event was hypothesized, which serves as a prerequisite for evolving nodulation independently before the significant divergence in this clade (Soltis et al., 1995; Werner et al., 2014).

Recently, phylogenomic comparisons of nodulating and non-nodulating species resulted in a new hypothesis on the evolutionary trajectory of nodulation (**Figure 1**) (Griesmann et al., 2018; van Velzen et al., 2018; van Velzen et al., 2019). The main discovery is the phenomenon of multiple independent loss-of-function of essential genes for nodulation in non-nodulating species of the Nitrogen-Fixing Clade. These genes include putative orthologs of *NFP* (*NOD FACTOR PERCEPTION*) that encodes a receptor for bacterial secreted lipo-chitooligosaccharide (LCO) signal molecules, the transcription factor *NIN* (*NODULE INCEPTION*) that is essential for nodule organogenesis, and the coil-coiled protein *RPG* (*RHIZOBIUM POLAR GROWTH*) that is crucial for intracellular bacterial infection. Additionally, the studies also show that the nodulating species utilise a shared subset of genes that are transcriptionally induced in nodules. An example is from the transcriptomic analyses between *P. andersonii* and *M. truncatula*, which revealed ~290 orthologous genes with a nodule enhanced expression profile (van Velzen et al., 2018). Altogether, this discovery challenges the previous hypothesis on parallel evolution of nodulation predated by a predisposition event. More probable the nodulating lineages originated from a nodulating common ancestor.



**Figure1. Nitrogen Fixing Clade (NFC).** Phylogenetic relationship distribution of the nodulating and non-nodulating species in orders of Fabales, Fagales (FAG), Cucurbitales (CUC), and Rosales which collectively form the NFC. The phylogenetic relationship is based ‘single gain-massive loss’ hypothesis, in which the branch colours depict the status of symbiotic capacity and type of microsymbionts. Red colour indicates the interaction with rhizobium. Blue colour indicates the symbiosis with *Frankia*. Grey colour shows a secondary loss of the symbiotic capacity. The phylogenetic tree is adapted from (van Velzen et al., 2019).

If the current hypothesis on ‘a single gain-massive loss’ is correct, then we would expect that the nodulation trait in all plants, including *C. mimosoides* and *P. andersonii*, is founded on a conserved genetic network.

## Early interactions and common signalling symbiotic pathway

In the nitrogen-fixing symbioses, exchanging signals between the host plants and the microsymbionts mark the initiation of symbiotic interaction. Under N-limiting conditions, the host plants secrete root exudates, like flavonoids, to attract the compatible bacteria. The flavonoids perception leads to the expression of bacterial nodulation (*nod*) operon genes to produce LCOs, which are also known as Nod-Factors (reviewed by (Limpens et al., 2015). Nod-Factors consist of four to five  $\beta$ -1,4-linked N-acetyl glucosamine residues coupled to a fatty acid chain (Lerouge et al., 1990), and structurally similar to LCOs produced by AM fungi (so-called Myc-LCOs) (Maillet et al., 2011). Specific substitutions on the terminal or non-terminal residues of rhizobial LCOs define the LCOs affinity with the host plants' receptors (Spaink et al., 1991; Masson-Boivin et al., 2009; Masson-Boivin and Sachs, 2018). Earlier studies have demonstrated that the presence of compatible rhizobium is essential for activating early response genes, such as early nodulins (Hirsch et al., 1989; Franssen et al., 1990), and eventually induce the formation of root nodules. In some cases, the purified LCOs molecules can induce a complete nodule developmental program even without the presence of rhizobia, implicating that the LCOs perception is sufficient for nodule organogenesis (Truchet et al., 1991; Mergaert et al., 1993; Relic et al., 1994; Stokkermans and Peters, 1994).

In the symbiotic association with *Frankia*, studies on the molecular dialogue between plant and microbe are less extensive compared to the legume-rhizobium model systems. Nodulating *Frankia* are classified in three taxonomic lineages, named cluster 1, 2 and 4, of which cluster 2 represents the most basal lineage. Recent genome sequencing revealed the occurrence of Nod-Factors biosynthesis genes -encoded by the *nodABC* operon- in cluster 2 *Frankia* species (Persson et al., 2015), similar as found in rhizobia. No such genes are present *Frankia* in clusters 1 and 4 (Persson et al., 2015; Nguyen et al., 2016). This suggests that *Frankia* species of the latter two lineages nodulate actinorhizal plants in an LCO-independent manner (Nguyen and Pawlowski, 2017).

Mutant analyses in model legumes, such as *Medicago* (*Mt*) and *Lotus* (*Lj*), impaired in perceiving the LCOs and the subsequent nodule formation, have led to the discovery of Nod factor receptor genes. These genes include the orthologous pairs of *MtLYK3/LjNFR1* (LysM-containing RECEPTOR LIKE-KINASES, *LysM-RLK*) and *MtNFP/LjNFR5* (NOD FACTOR PERCEPTION, *NFP*), which encode the receptor kinases containing three lysin motif domains (Limpens et al., 2003; Madsen et al., 2003; Radutoiu et al., 2003; Arrighi et al., 2006). At the plasma membrane of epidermal cells, the LysM domains of these receptor kinases can form dimers, which are essential for the LCOs binding. In *Parasponia andersonii* (*Pan*), the orthologous of these receptor genes -*PanNFP2* and *PanLYK3*- are also essential for nodulation (Rutten et al., 2020). Interestingly, *PanLYK3*

functions also in AM symbiosis as well as in chitin-triggered immune signalling (Rutten et al 2020). In actinorhizal plants, a little is known about the *Frankia*-secreted Nod factors receptors. A recent phylogenomic study annotated the putative orthologues genes encoding LCO receptors in some actinorhizal plants, which showed that NFP/NFR5-type receptors are limited to species that are nodulated by *Frankia* cluster 2 strains (Rutten et al 2020). Although functional studies on this receptor have not been performed, it suggests that this receptor is important for LCO-based nodulation.

Upon LCOs perception the signalling network leading to nodule formation and bacterial infection become activated. Interestingly, a large part of the rhizobial-induced signalling network in legumes is also required for symbiotic signal transduction systems in the AM fungi symbiosis. This extensive genetic network is collectively referred to as ‘a common symbiotic signalling pathway’ (CSSP) (**Figure 2.**), which is activated either after the recognition of rhizobial-LCOs or the Myc-LCOs. Genetic studies revealed that the CSSP comprises an LRR-type transmembrane receptor kinase MtDMI2/LjSYMRK, a nuclear-localized potassium channel MtDMI1/LjCASTOR and LjPOLLUX, nuclear pore complex proteins such as LjNUP133, LjNUP85, and LjNENA, and a calcium calmodulin dependent kinase MtDMI3, and the CCAMK interacting transcription factor MtIPD3/LjCYCLOPS. This signalling cascade can activate two different developmental programs, leading either to nodule formation, or arbuscules formation. Detail functions on each gene involved in the CSSP is described below.

The LRR-transmembrane receptor kinase *Lotus* SYMBIOTIC RECEPTOR KINASE (LjSYMRK)/*Medicago* DOES NOT MAKE INFECTIONS2 (MtDMI2) can form a heterodimer complex with LjNFR5/MtNFP and exhibits essential functions for both rhizobia and AM fungi symbioses (Endre et al., 2002; Stracke et al., 2002). *P. andersonii* *Pansymrk* mutant lines revealed a conserved function as this line abolished nodule and arbuscules formation (Roswanjaya et al., unpublished data). It was shown that MtDMI2 also interacts with MtHMGR1, an enzyme involved in mevalonate biosynthesis that evokes nuclear calcium spiking (Venkateshwaran et al., 2015). This calcium spiking consists of regular oscillations of perinuclear calcium concentrations, a notable feature triggered by the rhizobial-LCOs and Myc-LCO perceptions. Symbiotic calcium spiking is dependent on nuclear based proteins such as a nuclear-localized calcium-dependent adenosine triphosphatase (MtMCA8), a nuclear-localized potassium channel (LjCASTOR/MtDMI1, LjPOLLUX) and nuclear-localized calcium channels (MtCNGC15a-c) (Stracke et al., 2002; Ané et al., 2004; Imaizumi-Anraku et al., 2005; Charpentier et al., 2016). Genetic evidence also identified the involvement of other nuclear pore complex proteins such as NUP133, NUP85, and NENA, which based on its subcellular localisation may regulate the movement of components required for calcium spiking (Kanamori et al., 2006b; Saito et al., 2007b; Groth et al., 2010b). The calcium- and calmodulin-dependent protein

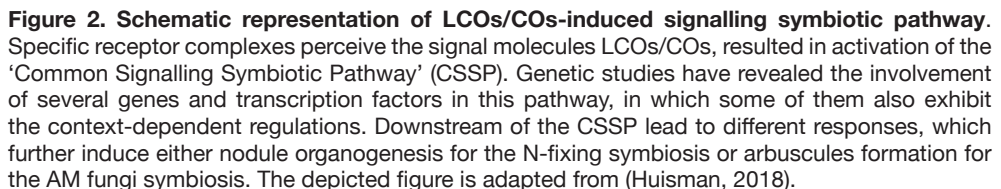
kinase CCaMK decodes the calcium spiking (Lévy et al., 2004b; Sieberer et al., 2009), which can interact with and phosphorylate the transcription factor LjCYCLOPS/MtIPD3 (Yano et al., 2008; Singh et al., 2014; Limpens and Bisseling, 2014). Phosphorylation of CYCLOPS by CCaMK marks the divergence of the rhizobial symbiosis and the AM fungi symbiosis. The transcription factors CYCLOPS can regulate different transcriptional reprogramming to either generate the nodule formation or develop the highly branched AM fungi, so-called arbuscules.

CYCLOPS activates other transcriptional factors such as *MtNSP1* and *MtNSP2* (NODULATION SIGNALING PATHWAY) of the GIBBERELLIC ACID INSENSITIVE (GAI), REPRESSOR of GAI, and SCARECROW (GRAS) family (Kaló et al., 2005; Smit et al., 2005). NSP1 and NSP2 interact to form either homo- or heterodimers that are important for binding to and activating Nod factor-responsive markers, including *EARLY NODULIN 11* (*ENOD11*), *NODULE INCEPTION* (*NIN*) AND *ETHYLENE RESPONSE FACTOR REQUIRED FOR NODULATION1* (*ERN1*) (Cerri et al., 2012; Kawaharada et al., 2017). Eventually, the output of this symbiotic pathway leads to rhizobial infection and synchronised nodule organogenesis, both of which are inextricably intertwined with generic plant hormonal networks. Meanwhile, in the AM fungi symbiosis, CYCLOPS activates the GRAS-type transcription factor RAM1, which induces the downstream signalling cascades for successful penetration of AM fungi into the root cells (Pimprakar et al., 2016).

### Into the journey of intracellular infection events

Rhizobial-LCO perception and activation of downstream signalling cascades are essential for accommodating rhizobial colonisation into the root cells. Legumes control the rhizobial colonisation via diverse infection processes, although this mostly occurs via ‘root hair curling’. Nod factor producing rhizobia attach to tips of growing root hairs. This induces tip-growth reorientation and entraps the bacteria between root hair cell walls of the curled root hair, thereby forming ‘a closed infection pocket’. Here, the bacteria can enter the root cells, divide and form an infection threads surrounded by a plant derived membrane. Eventually, the infection threads progress towards the cortical cells, which concomitantly undergo a reprogramming and become mitotically activated. This process results in the formation of nodule primordia of which cells are susceptible to rhizobial infection. In these cells, the bacteria are released as transient nitrogen-fixing organelle-like structures, and surrounded by the symbiosome membrane (Xiao et al., 2014; Roy et al., 2020).

In contrast to legumes, the early infection process in *Parasponia* involves cell division of epidermis and outer cortex, which provides apoplastic spaces for the bacteria to



penetrate in between the cells. This process is known as ‘a crack-entry’. Crack entry is not unique to *Parasponia*, but also occurs *Frankia* infected actinorhizal plants with the Rosales and Cucurbitales orders, as well as in a few legume species (Boogerd and Van Rossum, 1997; Ibáñez et al., 2016). During the formation of *Parasponia* nodule primordia, bacterial infection switches from intercellular to intracellular. Once inside a nodule primordium cell, the rhizobia remain associated within a thread, which is called a fixation thread, which usually have a relatively thin, lightly staining wall (Becking, 1983; Lancelle and Torrey, 1984; Lancelle and Torrey, 1985). In *Parasponia* nodules, the rhizobia are never released as symbiosomes in the fixation zones.

During early infection events, many components are involved for the reorientation of the root hair tip, remodelling of the cell wall and the cell membrane, and subsequently the formation of the infection thread. One prominent example is the production and delivery of the cell wall modifying enzyme NODULE PECTATE LYASE (NPL), to the infection pocket, which is required for successful development of the infection thread (Xie et al., 2012). Defects in *Lotus npl* resulted in a reduction in the growth of IT and an increase of uninfected nodules (Xie et al., 2012). Additionally, the presence of membrane trafficking proteins also contributes to the successful progression and elongation of the infection thread, and this includes FLOTILIN2 (MtFLOT2), MtFLOT4, and MtSYMREM1 (SYMBIOTIC REMORIN 1) in *M. truncatula* (Haney and Long, 2010; Lefebvre et al., 2010; Liang et al., 2018). Knockdown of *Mtflot2* significantly reduces the number of infection threads, while a mutant of *Mtflot2* resulted in lethality, suggesting that the *MtFLOT2* has a broader function than just regulating rhizobium infection (Liang et al., 2018). Additionally, it was found that MtFLOT4 and MtSYMREM1 can interact with the receptor protein MtLYK3, to mediate endocytosis of the MtLYK3 receptor. Mutants on *Mtflot4* and *Mtsymrem1* are affected in the elongation of infection thread growth (Haney and Long, 2010; Lefebvre et al., 2010; Liang et al., 2018).

Genetic analyses on other host components upon rhizobial inoculation have identified the involvement of a protein named MtVPY (MtVAPYRIN) and its two interacting partners; MtLIN (LUMPY INFECTION) and the exocyst complex subunit MtEXO70 H4 (EXOCYST Subunit H4) (Pumplin et al., 2010; Murray et al., 2011; Liu et al., 2019a). Any mutations in either *Mtvpv* or *Mtlin* resulted in a reduction in nodule numbers, of which the developed nodules are small and uninfected. Another protein involved in actin rearrangement and infection is MtRPG. Mutants in *Mtrpg* showed a delay in the formation of infection chambers in the root hair curls, although no substantial delay in the formation of infection threads. A distinctive feature of *Mtrpg* nodules are the unusually thick infection threads with bulbous protrusions, suggesting active bacterial multiplication (Arrighi et al., 2008). Amongst several transcription factors essential for rhizobial infection, NIN is one of the key players (Vernié et al., 2015). Transcriptomic data on root hairs of *nin* mutants revealed

that NIN has a large impact on the symbiotic transcriptional networks. About 10% of NIN targets are associated with cell wall remodelling, such as *NPL*, *RPG*, and *Cystathionine- $\beta$ -Synthase-like 1 (CBS1)* (Liu et al., 2019b). *Medicago Mtnin* mutants showed a defect in the development of the infection chamber due to the lack of exocytosis. A potential cause is due to lack VAMP721e accumulation in the surrounding membrane domain (Liu et al., 2019b).

Several plant hormones also regulate the progression of the infection process positively or negatively. Mutant analysis in *Medicago Mtskl (sickle)*, which is a loss-of-function in the ethylene signalling component EIN2 (ETHYLENE INSENSITIVE2), shows a defect in rhizobial infection (Penmetsa and Cook, 1997). *Medicago* and *Lotus skl* mutants exhibit massive proliferation of infection threads and a reduce development of mature nodules (Penmetsa and Cook, 1997; Miyata et al., 2013). In *Medicago*, the *Mtskl* mutant is not only hyper-sensitive to rhizobium, but also hyper-susceptible to infection by fungal pathogens, suggesting that ethylene signalling plays a crucial role in fine-tuning interactions with microbes (Penmetsa et al., 2008). Additionally, mutations in *MtARF16 (AUXIN RESPONSE FACTOR)* lead to a reduction in the number of infection events, which consequently reduce the number of functional nodules (Breakspear et al., 2014). A recent study by Kim et.al., (2019) describes the integration of gibberellic acid in the infection processes. Targeted mutagenesis on *MtGA2ox10 (GIBBERELLIC ACID OXIDASE10)* resulted in fewer infection events and reduced the number and size of root nodules (Kim et al., 2019).

In contrast to the positive roles of several hormones in rhizobial infection, cytokinin signalling may negatively regulate the epidermal infection. In legumes, cytokinin signalling in nodule organogenesis is mainly regulated by the functional cytokinin receptors MtCRE1/LjLHK1, which is orthologous to *Arabidopsis thaliana* AtAHK4/AtCRE1 (Tirichine et al., 2007; Heckmann et al., 2011; Plet et al., 2011). Murray, et al., (2007) showed that the local regulation of cytokinin signalling via a cytokinin receptor *Lotus Ljhhk1 (histidine kinase1)* exhibits a negative role on the rhizobial infection. A hyper-infection on the epidermis was observed in the *Lotus Ljhhk1* mutant background, which resulted in an infection switch from the root hair curling-based into the crack-entry mechanism that subsequently delayed and reduced the nodulation capacity in this mutant line (Murray et al., 2007; Held et al., 2014).

### Hormonal and regulatory networks for fine-tuning in nodulation capacity

Progression of rhizobial infection from the epidermis to the nodule primordium is a significant milestone in the endosymbiotic interaction between plants and rhizobium or

*Frankia*, leading to *de novo* nodule organogenesis. In addition to roles in infection, plant hormones also play a role in nodule development.

Genetic analyses in *Medicago* and *Lotus* of the cytokinin receptor *MtCRE1/LjLHK1* showed that this gene is essential for nodule formation. *Mtcre1/Ljlhk1* knockout mutants are severely delayed in nodule formation, resulting in a dramatic reduction in nodule numbers (Murray et al., 2007; Plet et al., 2011). Additionally, *Lotus* mutants of other cytokinin receptors, namely *LjLHK1A* and *LjLHK3*, together with *LjLHK1* completely abolished the nodule formation, which suggests that the *LjLHK1A* and *LjLHK3* are partially redundant with *LjLHK1* during nodule organogenesis (Held et al., 2014). Similar observations have been made also in *Medicago*, in which RNAi-mediated silencing of other cytokinin receptors in an *Mtcre1-1* mutant background adds to the severity of the nodulation phenotype (Boivin et al., 2016). Additionally, it was shown that exogenous application of cytokinins (Cooper and Long, 1994; Gauthier-Coles et al., 2019) or a gain-of-function mutant of *LjLHK1* (*Ljsnf2*) (Tirichine et al., 2006; Tirichine et al., 2007) can induce spontaneous nodule formation independent of rhizobium. Additional evidence that pinpoints the integration of cytokinin signalling in legume nodule organogenesis is the finding that rhizobial Nod factor induced gene expression is largely dependent on *MtCRE1* in *Medicago* (van Zeijl et al., 2015a; Jardinaud et al., 2016; Schiessl et al., 2019). These include the Nod factor responsive genes *NSP1*, *NSP2*, *ERN1* and *NIN*, which are all essential for nodulation. Other studies revealed that the cytokinin biosynthesis (ISOPENTENYL TRANSFERASE3 (*LjIPT3*) in *Lotus* and LONELY GUY1 (*MtLOG1*) in *Medicago*) and cytokinin degradation-related genes (*Lotus* CYTOKININ OXIDASE/DEHYDROGENASE3 (*LjCKX3*)) serve as an integral function in nodule organogenesis, and that cytokinin biosynthesis enhanced upon Nod factor perception (Mortier et al., 2014; Reid et al., 2015; van Zeijl et al., 2015a; Reid et al., 2017). Taken-together, these results underline the importance of cytokinin signalling in legume nodule formation.

Further, the integration of cytokinin signalling upon LCO perception and nodule organogenesis is also strengthened by a spatial analysis using the synthetic cytokinin sensor *TCS* (Müller and Sheen, 2008) and its improved version *TCSn* (Zürcher et al., 2013). In *Lotus*, the first induction of cytokinin signalling coincides with the dividing cortical cells and later extended to the epidermal activity, which requires a functional *LjLHK1* gene (Reid et al., 2017). Conflicting reports in the cytokinin spatial expression in *Medicago* resulted from *TCS* (van Zeijl et al., 2015a) and *TCSn* promoter (Jardinaud et al., 2016). The *TCS* reporter shows spatial expressions on the cortical cells upon the LCOs perception (van Zeijl et al., 2015a), while the *TCSn* reporter shows spatial expressions on the epidermis (Jardinaud et al., 2016). This epidermal activation of cytokinin signalling supported by laser-capture microdissection RNAseq data that revealed an induction of

cytokinin responsive-genes in the epidermis (Jardinaud et al., 2016). A difference in sensitivity between the *TCSn* and *TCS* can be the causal of this discrepancy. Despite the differences in the integration of cytokinin signalling in *Lotus* and *Medicago*, there is a consensus regarding the essential role of cytokinin signalling in nodule organogenesis. To this end, it raises the next questions on the mechanism by which the cytokinin signalling regulates this process, and whether the cytokinin signalling also commits a similar role in the distantly related nodulating species *Parasponia*.

Genetic screens in model legumes have shown that *NIN* is indispensable for nodulation and has context-dependent functions in regulating rhizobial infection or nodule organogenesis (Soyano et al., 2015; Vernié et al., 2015; Liu et al., 2019b). In legumes, *NIN* is both Nod factor and cytokinin responsive. Upon LCOs perception, transactivation of *NIN* expression via CYCLOPS in the root epidermis is essential for root hair-based infection (Singh et al., 2014). Additionally, expression of *MtNIN* under the control of an enhanced epidermal promoter is sufficient to induce nodule organogenesis via activating MtCRE1-dependent cytokinin signalling in the inner cortical cells (Vernié et al., 2015) (**Figure 2**). Rhizobial application on wild-type roots of *Medicago* revealed the induction of cytokinin responsive-gene *MtRR4* (*RESPONSE REGULATOR4*) in the inner root cortex, while such induction was not observed in the *Mtnin* mutant background (Vernié et al., 2015). This shows that *NIN* is essential to rhizobium-induced cytokinin signalling. The activation of *NIN* in the inner root cortex is essential to initiate nodule formation and to transcriptionally activate *NF-YA1*, a transcription factor involved in several steps of nodulation (Soyano et al., 2013). In legume *nin* mutants, the rhizobial infection is blocked and nodule formation is abolished. This mutant phenotype can only be rescued by introducing *NIN* promoter containing a CYCLOPS binding site and a conserved region that includes several cytokinin B-type RESPONSE FACTOR (RR) binding sites (Liu et al., 2019c). Altogether, it suggests that the cytokinin signalling and *NIN* are interconnected, of which the expression of *NIN* and the cytokinin response displays a positive feedback loop in the cortical cells that will initiate the nodule primordium.

Similar to cytokinin, auxin can trigger cortical cell divisions, leading to the formation of nodule-like structures in the absence of bacteria. For instance, an exogenous application of synthetic auxin (50 mM); 2,4-Dichlorophenoxyacetic acid (2,4-D), or of 1-N-naphthylphthalamic acid (NPA) or auxin transport inhibitors 2,3,5-triodobenzoic acid (TIBA) induces the nodule-like structures on legume or rice roots (Hiltbrand et al., 2016). In the *Medicago Mtcre1* mutant, the application of certain flavonoids or TIBA rescued the initiation of nodule primordium via restoring the local auxin accumulation in the cortical cells (Ng et al., 2015). Also, an overexpression of the indole-3-pyruvate

(IPyA) monooxygenase *YUCCA2* (*GmYUC2a*) gene in soybean, which is required for the last step of indole-3-acetic acid (IAA) auxin biosynthesis, altered root hair deformation and reduced nodule number (Wang et al., 2019). Further, the auxin influx transporter *MtLAX2* is required at the initiation of nodule primordia, (Roy et al., 2017). A similar observation was done by reducing the level of expression of *PIN-FORMED 2* (*MtPIN2*), *MtPIN3*, and *MtPIN4* via RNA interference, resulting in reduced nodule numbers (Huo et al., 2006). Recently, overlapping regulation between the lateral root and nodule formation was found. This involves the activation of cytokinin-inducible transcription factor *NIN* to induce *LOB-DOMAIN PROTEIN 16* (*LBD16*), which regulates the auxin biosynthesis genes *STYLISH* (*STY*) and *YUCCA* (*YUC*). Mutations in *LBD16* resulted in defective nodule and lateral root initiation, while overexpression of *LBD16* or *YUC* promoted lateral root formation (Schiessl et al., 2019). Altogether, these results indicate that *YUCCA*-dependent auxin biosynthesis, PIN-mediated, and LAX-dependent auxin transport are involved in the regulation of nodule initiation (**Figure 2**).

In contrast with the positive roles of cytokinin and auxin, the phytohormone ethylene is proposed to act as the negative regulator in different processes of nodulation, including regulation of total nodule numbers, infection thread formation, nodule morphology, and nodule positioning (see review (Ferguson and Mathesius, 2003)). A first observation regarding the involvement of ethylene in nodulation came from the application of its biosynthetic precursor ACC (1-amino-cyclopropane-carboxylic acid), which suppresses nodule formation. Conversely, that application of chemical inhibitors of ethylene perception (i.e. Ag<sup>+</sup>) or biosynthesis (i.e. aminoethoxyvinyl glycine, AVG) can increase nodule numbers (reviewed by Ma et al. 2002). Early in the nodulation process, application of ACC dampens the rhizobial-LCO induced calcium spiking and blocks the initiation of infection threads (Oldroyd et al., 2001). Inhibition of spontaneous nodulation in the *LjLHK1* gain-of-function mutants by ethylene precursors places ethylene inhibition of nodulation upstream of cytokinin signalling (Tirichine et al., 2006; Tirichine et al., 2007). During the nodule development, an ethylene gradient dictates the positioning of nodule cells to be radially opposite of the xylem poles (Heidstra et al., 1997). This positioning effect is affected in a number of ethylene insensitive mutants (Penmetsa and Cook, 1997; Lohar et al., 2009). Further, rhizobia can manipulate ethylene levels in root nodule symbiosis. Many rhizobial species contain an ACC deaminase (*acdS*), a gene coding for an enzyme to catalyse the degradation of ACC into ammonium and α-ketobutyrate. In this way rhizobia can reduce ethylene levels in root cells (Grichko and Glick 2000; Kaneko 2000). *AcdS* was significantly upregulated in bacteroids, while an *acdS* mutant showed a reduction in infectability of *Pisum sativum* (Ma et al., 2002), suggesting that ethylene regulate bacteroid functioning.

In addition to NIN, the genetic screens for transcription factors essential in nodulation revealed two the GRAS (GAI, RGA, SCR)-type proteins; *NSP1* and *NSP2* (Kaló et al., 2005; Smit et al., 2005). These two transcription factors demonstrate indispensable role for strigolactone biosynthesis in the legume *Medicago* and in rice (*Oryza sativa*) (Liu et al., 2011). The disturbed strigolactone biosynthesis in *nsp1 nsp2* mutant backgrounds correlates with reduced expression of DWARF27, a gene essential for strigolactone biosynthesis (Liu et al., 2011). Additionally, *Medicago NSP1* and *NSP2* are expressed in nodules and commit integrated roles in nodulation by mediating the rhizobial-induced responses, such as root hair deformation, infection thread formation, and expression of early nodulation genes such as *ERN1* to transcriptionally activate *ENOD11* (Cerri et al., 2012; Kawaharada et al., 2017). Interestingly, a study shows that NSP1 and NSP2 are required for CYCLOPS-induced nodule organogenesis, but not CYCLOPS-induced *NIN* expression (Singh et al., 2014). To this end, it remains unknown on how NSP1-NSP2 activity is associated with CCaMK-CYCLOP module, or alternatively, whether NSP1-NSP2 act in an independent and parallel pathway affecting Nod factor induced signalling (Singh et al., 2014).

Gibberellic acid signalling also shows to have functions in the root nodulation, although its functioning varies between plant species (Fonouni-Farde et al., 2016; Jin et al., 2016). Bioactive gibberellic acids are perceived by GID1 (Gibberellic acid insensitive dwarf 1) (Ueguchi-Tanaka et al., 2005) and induces the degradation of DELLA, a nuclear protein that restrains the cell proliferation and expansion that drive plant growth (Harberd et al., 2009). Upon the gibberellic acid binding, DELLA degradation is mediated by the proteasome through the SCF (SLY/GID2) E3 ubiquitin ligase complex (Itoh et al., 2008). In *Medicago*, gibberellic acid signalling mediated by DELLA is inhibiting rhizobial infection and controlling the expression of *ENOD11*. A constitutively active DELLA1 protein expressed in the epidermis is sufficient to induce *ENOD11* expression even in absence of Nod factors. In a heterologous system, it was shown that DELLA interacts with NSP1/NSP2 and NF-YA1. Additionally, DELLA can bind to the *ERN1* promoter, activating its expression, which subsequently regulates rhizobial infection (Fonouni-Farde et al., 2016).

### **Nodule ontogeny in legumes and non-legume *Parasponia***

Although the key players as well as the hormonal involvement in nodulation are conserved among legumes, there are clear differences in nodule ontogeny of legumes as well as of nodulating non-legumes. This raises questions whether signalling networks are conserved across the nodulating species. Alternatively, the observed differences in nodule ontogeny are the result of the read out of a conserved signalling network with minor adaptations in different species.

In *Medicago*, initiation of nodule primordia starts by cell divisions in the pericycle, and is followed by divisions in the cortex and endodermis (Xiao et al., 2014). The cortical-derived cells will be susceptible to infection and form the central tissue of the nodule. Thus, the mature nodule consists of the outer cortex, endodermis, peripheral vascular bundles, and the cortex-derived cells that can contain symbiosomes (Xiao et al., 2014). The location in the cortex of the cells ultimately forming the nodule primordium can vary from one legume to the other. Whereas in *Medicago* these cells are positioned in the inner cortex (Gage, 2004; Xiao et al., 2014), in soybean they are from the outer cortex, while in *Lotus* they originate from cells in the middle part of the cortex. Based on the nodule ontogeny in legumes, they are identified as indeterminate or determinate nodules. Indeterminate nodules show a persistent nodule meristem and a more elongated growth of nodule structure, like in *Medicago*. While, the determinate nodules do not possess the persistent nodule meristem and generate more spherical types of nodules, such as in *Lotus* and soybean.

In *Parasponia* sp, nodule primordia originate from clusters of dividing cells of the epidermis and outer cortical cells, which become susceptible for rhizobium infection. Unlike legumes, the nodule vascular bundles in *Parasponia* originate from the dividing cells in the inner cortical cells and pericycle, which further develops into a central vascular bundle. Such central vascular bundle in *Parasponia* nodules resemble the ontogeny of actinorhizal-type nodules. *Parasponia* nodules have a persistent meristem at the tip of the nodule vascular bundle, which can grow and add new cell layers that are susceptible to infection. In mature nodules of *Parasponia*, where the infected cells host the differentiated bacteria that start to fix the atmospheric dinitrogen. Phenolic compounds are observed in the surrounding of the infected lobes, which serve as a protective layer to avoid pathogenic infections (Lancelle and Torrey, 1984; Lancelle and Torrey, 1985; Bender et al., 1987; Becking, 1992). Interestingly, when such phenolic compounds are present in the legume-type nodules, this suggests that the nodules are going to senesce (Van de Velde et al., 2006; Domonkos et al., 2017).

Despite the differences on the structure and cell origins of the nodule vascular bundles between *Parasponia* and legumes, a detailed observation by Shen et al., (2020) showed that *Parasponia*-type nodules and actinorhizal-type nodules are actually more similar to legume-type nodules than it is generally assumed. As in both *Parasponia* and legume-type nodules, the infected cells are from mitotically activated cortical cells. Additionally, mutant analysis in *M. truncatula* on the *NODULE ROOT1* (*MtNOOT1*) revealed the conversion of legume-type nodule to the actinorhizal-type nodule, which suggests that both *Parasponia* and legume-type nodules may have shared an evolutionary origin (Shen et al., 2020).

## Research strategy

As described above, nodulation potentially originated from a shared common ancestor. This suggests the occurrence of conserved symbiotic signalling mechanisms leading to nodulation (**Figure 2**) and the occurrence of specific lineage adaptations. Exploiting new experimental systems alongside the current model legumes is highly relevant to pinpoint the core genetic networks allowing successful symbiosis among the nodulating lineages. Better understanding of nodule organogenesis and infection means a closer step of designing engineering strategies to transfer the nodulation trait into the non-nodulating crops.

Appreciating the knowledge gained from studies in model legumes, it showed that the cytokinin signalling is one of the essential players in rhizobial-induced nodule organogenesis and possibly in nodule infection. However, whether the cytokinin signalling plays a conserved role in other nodulating species is unknown. To this end, I formulated the primary research question in this thesis as follows.

***“What is the role of cytokinin signalling in the nodulating species outside the Papilionoideae subfamily?”***

To answer this question, I investigated the integration of cytokinin signalling in the rhizobial-induced root nodulation by utilising two novel experimental systems, (1) *C. mimosoides*, which belongs to the basal legume Caesalpinioideae subfamily, and (2) *P. andersonii*, the distantly related non-legume species belonging to the Cannabaceae, Rosales. I hypothesise that cytokinin signalling is a conserved element in the nodulation trait.

In **Chapter 2**, I performed studies in *C. mimosoides* to unravel the role of cytokinin signalling and its connection with NIN during nodulation. My focus of study is divided into three aspects as follows. Firstly, I questioned whether a short time (24 h and 48 h) rhizobial (*Bradyrhizobium* sp WUR 077) application induces the activation of cytokinin signalling by quantifying the expressions of the cytokinin-signalling responsive genes (*Type ARR*s, *A-RESPONSE REGULATORS*). I also observed the expression of early gene marker *NIN* during this rhizobia treatment, since it is known from the model legumes that *NIN* is rhizobial-LCOs responsive. Secondly, I developed the transient transformation protocol to subsequently performed the functional studies in *C. mimosoides*. Utilising the recently published genomic database of *C. fasciculata* and in combination with the transcriptomic database of *C. mimosoides*, we identified the orthologues of *NIN* and cytokinin signalling related-genes in *C. mimosoides*. We used this information to design an RNA interference-based reverse genetic approach targeting *NIN* and each copy of

the cytokinin receptors, and subsequently assessing their nodulation capacities. Thirdly, I investigated whether the cytokinin signalling and *NIN* demonstrated an interconnected role in nodulation.

In **Chapter 3**, I present ‘a mini handbook’ of *P. andersonii* for as a new experimental system. This chapter includes the elaborative protocols on performing reverse genetic analyses in *P. andersonii* based on the CRISPR/Cas9-gene editing and via the stable transformation mediated by *Agrobacterium tumefaciens*. We also provided two experimental platforms on utilising the generated mutant lines in *P. andersonii*, either to study the N-fixing symbiosis or the AM fungi symbiosis. We argued that a common bottleneck in reverse genetic analyses is the availability of genetic manipulation tools. Therefore, establishing a robust and efficient stable transformation protocol is necessary to dissect the biological and molecular aspects of the symbioses.

In **Chapter 4**, I present a proof of concept of the reverse genetic analyses in *P. andersonii*. Utilising the recently published genomic database of *P. andersonii*, we designed constructs to perform the CRISPR/Cas9-based editing on *PanHK4*, *PanEIN2*, *PanNSP1*, or *PanNSP2*. These genes are known to be involved in the cytokinin, ethylene and strigolactones signalling, which in model legumes, these hormones exhibit a function in nodulation. In this chapter, we presented phenotypic observations on these mutant lines both in symbiotic and non-symbiotic contexts.

In **Chapter 5**, I present an in-depth analysis on the role of cytokinin signalling in the *P. andersonii*-rhizobium interaction. I made use of the synthetic cytokinin reporter, *TCSn*, to study whether the cytokinin signalling is a part of early LCO-responsive cascades. We performed reverse genetic analyses on other copies of the cytokinin receptors (*PanHK2*, *PanHK3*) either via RNA-interference or the CRISPR/Cas9-based gene-editing to analyse the function of these cytokinin receptors in nodulation. This revealed a novel role of *PanHK3* in controlling rhizobium intracellular infection.

In **Chapter 6**, I summarise the results and significance obtained from this PhD project, integrated them with the published data, and subsequently provided future perspectives on engineering the nodulation trait into non-nodulating crops.

## Acknowledgments

The author would like to thank Yuda Roswanjaya for sharing the data prior publication.



# Chapter 2

## **Dissecting the Role of Cytokinin Signalling and *NIN* during nodulation in *Chamaecrista mimosoides***

**Titis A. K. Wardhani<sup>1</sup>, Luuk Rutten<sup>1</sup>, Arjan van Zeijl<sup>1</sup>, Joël Klein<sup>1</sup>, René Geurts<sup>1</sup> and Wouter Kohlen<sup>1</sup>**

<sup>1</sup>Department of Plant Sciences, Cluster of Plant Developmental Biology, Wageningen University, Droevendaalsesteeg 1, 6708 PB, Wageningen, The Netherlands.

***Manuscript in preparation***

Nothing in life is to be feared, it is only to be understood.  
Now is the time to understand more, so that we may fear less.  
—Marie Curie





## ABSTRACT

During evolution, several plant lineages of the Fabales, Fagales, Cucurbitales, and Rosales orders established an endosymbiotic relation with nitrogen-fixing bacteria. Two recent phylogenetic studies suggest that the nodulation trait is likely gained once, and experienced massive parallel loss in the descending lineages. This hypothesis strengthens the idea that the nodulation signalling pathway is based on a shared genetic network. Studies in the Fabales sub-family of Papilionoideae, which focused on two model species (i.e. *Medicago truncatula* and *Lotus japonicus*), demonstrated that the transcription factor Nodule inception (NIN) is a key regulator in nodulation, and that its function is highly integrated with cytokinin signalling. However, whether these mechanisms are conserved in nodulating species outside the Papilionoideae remains to be seen. To investigate the function of *NIN* and its relation to cytokinin signalling in a nodulating non-Papilionoideae legume, we utilized a species belonging to the Caesalpinioideae legume sub-clade; *Chamaecrista mimosoides*. In *C. mimosoides*, we identified two copies of *NIN* (*NIN1* and *NIN2*), and one copy of the three cytokinin histidine kinase receptors (*HK2*, *HK3*, *HK4*) described in other plant species each. We demonstrate that both copies of *NIN* are transcriptionally activated within 24 hours when treated with either *Bradyrhizobium* sp WUR077 or cytokinin. In addition, knockdown of either *NIN* gene affects nodulation drastically, whereas *NIN2* knockdown also blocked bacteria release. Functional studies of HK receptors revealed that in *C. mimosoides* a observed reduction in nodule numbers correlates with the knockdown of *HK4*, and that the transcriptional activation of *NIN2* might be depending on *HK4* mediated cytokinin signalling. This suggests that, in line with the single gain, massive loss hypothesis, the mechanisms controlling nodulation in *C. mimosoides*, and possibly Caesalpinioideae, are similar to those found in model Papilionoideae.

## Introduction

The symbiotic association between plants and nitrogen-fixing bacteria contributes greatly to the amount of fixed nitrogen in soils. During this interaction, the plant produces a niche root organ in which bacteria are accommodated intracellularly, the so-called root nodule. In exchange for photosynthates, the bacteria provides reduced nitrogen to the plant (Oldroyd, 2013). The ability to form root nodules with nitrogen-fixing bacteria only occurs in four plant orders (i.e. Fabales, Fagales, Cucurbitales, and Rosales), commonly referred to as the 'Nitrogen-Fixing Clade' (Soltis et al., 1995; Werner et al., 2014). Within this clade, eight plant families associate with the filamentous *Frankia* bacteria and are called actinorhizal plants (Pawlowski and Demchenko, 2012). In addition, several plant species belonging to legumes (Fabales) and five *Parasponia* species (Cannabaceae, Rosales) establish root nodules with nitrogen-fixing rhizobia (Geurts et al., 2016).

The holy grail of nitrogen-fixing root nodule symbiosis is to expand the nodulation capacity to non-nodulating crops. To achieve this, we need to unravel the conserved mechanisms required for the successful establishment of such symbioses. Over the past 20 years, genetic research on nitrogen-fixing symbiosis primarily focused on two model legumes; *Medicago truncatula* and *Lotus japonicus*. These two species both belong to the highly successful legume sub-family of Papilionoideae. Available mutant libraries and omics datasets have greatly improved our understanding of symbiotic signaling pathways (Roy et al., 2020). However, recent phylogenetic studies suggest that the nitrogen-fixing symbioses likely originated from a common ancestor once, and experienced massive parallel loss in the descending lineages (Griesmann et al., 2018; van Velzen et al., 2018). In line with this hypothesis, 'single gain-massive loss' implies that conserved signaling pathways have been recruited in the last common ancestor of the nitrogen-fixing clade. Therefore, exploring additional experimental models representing genetic diversity within the nitrogen-fixing clade is highly relevant to identify key genetic components of root nodule formation. In addition, it will also enable to separate key genetic components from genetic diversity that is the result of lineage-specific adaptation.

Legumes are the third largest family in the angiosperms and, until recently, divided into six sub-families (i.e. Papilionoideae, Mimosoideae, Caesalpinioideae, Dialioideae, Cercidoideae, and Detarioideae). However, a recent phylogenetic reconstruction placed the Mimosoideae sub-clade as part of the Caesalpinioideae, leaving only five legumes subfamilies (van Velzen et al., 2019). Based on ecological studies and phylogenetic characterization, it was demonstrated that the ability to nodulate with rhizobia is a trait widely found in the Papilionoideae subfamily while it is less common in the more

‘ancestral’ subfamilies (e.g. Caesalpinioideae) (Griesmann et al., 2018; van Velzen et al., 2018; van Velzen et al., 2019).

A successful symbiotic interaction between legumes and rhizobia requires a synchronization of bacterial infection and the formation of a nodule primordium. First, the plant secretes flavonoids which are perceived by the rhizobium and trigger the production and secretion of lipo-chitooligosaccharides (LCOs, or Nod-Factors). In turn, these LCOs are then perceived by the plant via a receptor complex including the LysM-RLK NFP, located on the cell membrane of root epidermal cells. LCO perception triggers the activation of a conserved signaling cascade, the so-called ‘Common Symbiotic Signalling Pathway’ (Oldroyd, 2013) and induces several downstream responses. These responses include the transcriptional activation of the transcription factor *NODULE INCEPTION (NIN)* (Breakspear et al., 2014; van Zeijl et al., 2015a; Jardinaud et al., 2016; Schiessl et al., 2019). *NIN* has been demonstrated to be a major regulator of both rhizobia infection and nodule organogenesis (Soyano et al., 2013; Vernié et al., 2015; Liu et al., 2019c).

In papilionoidea, epidermal *NIN* is required to activate cortical cytokinin responses in relation to rhizobia symbiosis. Cytokinins play a major role in many developmental and environmental responses. During these, cytokinin response gradients between individual tissues and cells can be established and maintained by regulating either biosynthesis, transport, degradation, and/or perception (Hirose et al., 2008; Frebort et al., 2011).

Studies in *Arabidopsis thaliana* showed that the first and rate-limiting step of cytokinin biosynthesis is catalyzed by adenosine phosphate isopentenyl transferase (IPT), leading to the formation of the isopentenyladenine riboside-n-phosphate (iPRP) type of cytokinins (Takei et al., 2001). iPRP can be converted to trans-zeatin riboside-n-phosphate (tZRP) through the activity of the cytochrome P450 monooxygenase (CYP735A family) (Takei et al., 2004). Finally, both ribosided forms of cytokinin are converted to their active form through the activity of LONELY GUY (LOGs) (Kieber and Schaller, 2010; Kieber and Schaller, 2018).

During early LCOs signalling, short time exposure leads to induction of the bioactive cytokinins isopentenyladenine (iP) and *trans*-zeatin (tZ) (van Zeijl et al., 2015a; Jardinaud et al., 2016). This is consistent with a transcriptional induction of genes related to cytokinin biosynthesis such as *IPT3*, *CYP735A* and *LOG1* (van Zeijl et al., 2015a; Jardinaud et al., 2016; Reid et al., 2017). In addition, the early cytokinin response upon the LCO perception is required for many transcriptional changes downstream of the LCOs-signalling cascades, as many of these responses are dependent on the cytokinin receptor MtCRE1/LjLHK1 (van Zeijl et al., 2015a; Jardinaud et al., 2016; Reid et al.,

2017). Moreover, a gain-of-function cytokinin receptor in *Lotus* (*snf2/lhk1*) induces the formation of nodule-like structures even without the presence of rhizobium, whereas the loss-of-function of the cytokinin receptor *Ljlhk1* or *Mtcre1* leads to a massive reduction in nodule numbers and affects nodule ontogeny (Gonzalez-Rizzo et al., 2006; Murray et al., 2007; Plet et al., 2011). In addition, exogenous cytokinin application is able to trigger cortical cell divisions that resemble those induced by rhizobia, leading to the formation of nodule-like structures (Cooper and Long, 1994; Gauthier-Coles et al., 2019) and is able to activate early symbiotic responsive genes; *ENOD11* (*EARLY NODULIN GENE11*) and *NIN* (Heckmann et al., 2011, Liu et al., 2019c). Combined, this demonstrates the important role of cytokinin biosynthesis and perception during nodulation in model Papilionoids.

The morphology of rhizobia-induced nodules is very well-described for Papilionoidea e.g. *Medicago* and *Lotus*, which represent indeterminate and determinate nodules, respectively. The indeterminate nodule is defined by the presence of a nodule meristem, which keeps dividing and contributes to the growth of the nodule. In contrast, determinate nodules do not have a persistent nodule meristem (Hirsch, 1992; Gage, 2004). Information on nodule organogenesis in Caesalpinioideae is scant. One of only a handful of nodulating genera in the Caesalpinioideae sub-family is *Chamaecrista*. This genus comprises ~330 species which are mainly distributed across the tropical regions of the Americas, especially Brazil (Santos et al., 2017). The *Chamaecrista* genus was recently separated from the *Cassia* and *Senna* genera (van Velzen et al., 2019) and several *Bradyrhizobium* spp are characterized to nodulate herbaceous *Chamaecrista* (Moench) species (Santos et al., 2017). Based on a study by Ren and colleagues (2018), it was suggested that the *Chamaecrista* nodule (Caesalpinioideae) belongs to the determinate nodule type, although a pool of dividing cells that contribute to the number of infected cells is initiated later when a mature nodule is already formed (Ren, 2018).

As Caesalpinioideae and Papilionoideae diverged roughly 60 MYA, this places the *Chamaecrista* genus at a phylogenetically interesting position in relation to legume evolution. Several studies focused on *Chamaecrista fasciculata* as a complementary research model alongside *Medicago* and *Lotus* (Singer et al., 2009; Society and Journal, 2009). However, *C. fasciculata* is difficult to propagate under laboratory conditions, as it depends on the presence of long-tongued bees (Lee and Bazzaz, 1982). This outcrossing makes *C. fasciculata* highly polymorphic. *Chamaecrista mimosoides* is not dependent on bees for pollination, is easily selfed and produces ample amounts of seeds. In addition to its being amenable to molecular tools this makes *C. mimosoides* a more accessible model to study nodulation in Caesalpinioideae. The role of cytokinin signalling and *NIN* have been extensively shown to be required during nodule organogenesis in model Papilionoideae. However, if, and how, these are important

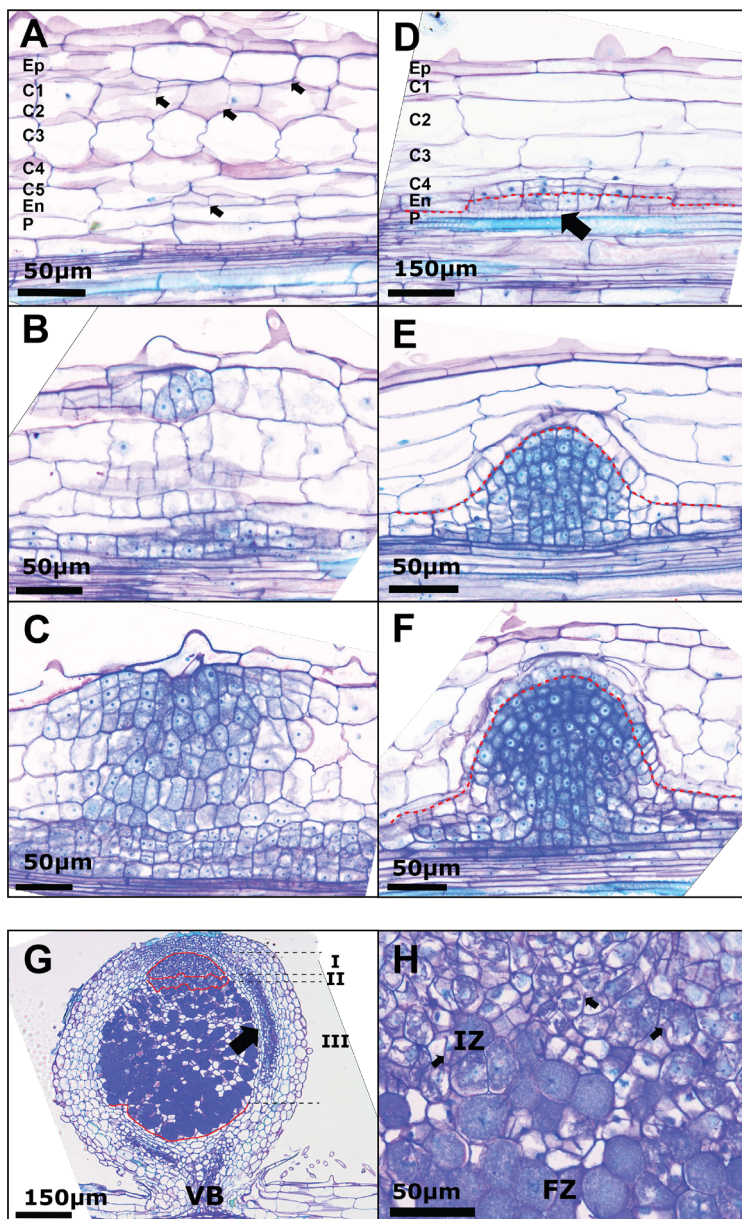
during nodulation in Caesalpinioideae remains to be seen. To answer this question, we utilized *C. mimosoides* as our research model. We set up two nodulation assays (plates and pots) and implemented a transient transformation protocol using *Agrobacterium rhizogenes*. We identify *HK2*, *HK3*, *HK4*, and two copies in *NIN* in *C. mimosoides* and demonstrate that *HK4*, possibly *HK2*, and both copies of *NIN* have a function during nodule initiation in *C. mimosoides*.

## Results

### Nodule developmental stages in *C. mimosoides*

To enable the functional analysis of *NIN* and the cytokinin receptors (i.e. *HK2*, *HK3* and *HK4*) during nodule organogenesis in *Chamaecrista mimosoides*, we needed to better understand the onset of nodule initiation in this species. In *M. truncatula* the first cell divisions leading to the formation of a nodule primordia occur in the pericycle after 24 hours and roughly 48 hours later the endodermis and the inner cortical layers are triggered to divide (Xiao et al., 2014). However, the precise timing of nodule initiation in *C. mimosoides* is unknown. Therefore, we inoculated 10-day old plate grown *C. mimosoides* seedlings with *Bradyrhizobium* WUR077 (previously named *Bradyrhizobium* sp RPG001 (Schneijderberg et al., 2018) ( $OD_{600} = 0.015$ ) and followed nodule development over time. We chose to analyse the initiation of cell divisions at 1- and 3-day post inoculation. However, as we were not sure these time points are appropriate for analysing the initiation of cell divisions in *C. mimosoides* we added a 6-day post inoculation time point. In addition, several inoculated plants were set aside to analyse the morphology of mature nodules after 3 weeks.

We observed that in *C. mimosoides* several cells in the more outer cortical layers (C2), endodermis (En) and pericycle (P) are mitotically reactivated with anticlinal divisions as early as 1-day post inoculation (**Figure 1. A**). This particular response is not described for either *Medicago* or *Lotus*. As we lack resolution, it is difficult to assess which cell divisions are triggered first, but it is clear that during early nodulation in *C. mimosoides* cells are activated in the outer and inner root cells. In the slightly further developed primordia (3 dpi), several anticlinal and periclinal cell divisions are observed in C1-C2, endodermis and pericycle, and further expanded to the C3-C5 inner layers of the root (**Figure 1. B-C**). Next, the cells from the more outer cortical cell layers are continued to divide to form a nodule primordium (**Figure 1. C, Supplemental Figure 1. A**), while the dividing cells derived from the inner layers and pericycle will develop a vasculature that connects the primordium and root. When comparing nodule (**Figure 1. A-C**) to lateral root initiation (**Figure 1. D-F**) we observed that during the latter, cell divisions are restricted to the pericycle and endodermis (**Figure 1. D-F**), with the pericycle-derived



**Figure 1. Representative images of determinate nodule and lateral root development in *C. mimosoides*.** (A-C) Micrographs of Toluidine Blue stained longitudinal sections of nodule primordia at (A) 1 day, (B) 3 days, and (C) 6 days post inoculation with *Bradyrhizobium* sp WUR077. (D-F) Micrographs of Toluidine Blue stained longitudinal sections during lateral root development. (G) Micrograph of a Toluidine Blue stained longitudinal section of mature nodule displaying a cluster of dividing cells (I), few layers of infection zone (II), fixation zone (III), and peripheral vasculature (Black arrowheads). (H) Zoom in micrograph of infection threads, IZ (Black arrowheads) and symbiosomes, FZ in zone III. Arrows highlight the location of early cell divisions in (A) nodule primordia and (D) emerging lateral root. Red line in (D-F) defines pericycle-derived cell divisions. Ep = epidermis. C1-C5 = cell layers1-5. En = endodermis. P = pericycle.

cells continue to divide to form lateral root primordia. This data suggests that the activation of the outer cortical cells is limited to the rhizobial-induced response leading to nodule initiation.

When analysing these early cell divisions, we noticed that during *C. mimosoides* nodule initiation rhizobia are maintained apoplastically between the cells of the nodule primordium (**Supplemental Figure 1. B**). In the young nodule (**Supplemental Figure 1. C**), these apoplastic bacteria are able to penetrate the cells to form new infection threads (**Supplemental Figure 1. D**), which are subsequently released as symbiosomes (**Figure 1. G, H**).

Three-week-old mature nodules consist of a cluster of dividing cells at the nodule apex that are potentially, few layers of infection zone and fixation zone (**Figure 1. G**). Interestingly, the persistent meristem is not formed during nodule primordia formation (**Supplemental Figure 1. A&C**), and in the mature nodule of *C. mimosoides* shows a peripheral vasculature, which may fuse with the nodule parenchyme (**Figure 1. G**). Based on this observation, we concluded that *C. mimosoides* develops the determinate nodule. This is in line with previous reports for *C. fasciculata* (Ren, 2018) and suggests that the infection process during nodulation of *Chamaecrista* is very similar to that of species belonging to the genistoids, *indigofera*, milletioids and dalbergioids clades of the Papilionoideae subfamily (Naisbitt et al., 1992; Ibáñez et al., 2017; Ren, 2018).

### Identification of the putative orthologous *NIN* and *HK* genes in *C. mimosoides*

Next, we set out to identify the putative *NIN*, *HK2*, *HK3* and *HK4* orthologues of *C. mimosoides*. To this end, we used the available *C. fasciculata* genomic data and had mRNA sequenced from *C. mimosoides* root, stem, shoot and nodule tissues to construct a *de novo* transcriptome database (**Supplemental Table 7**). Putative homologues were identified by BLAST (tblastx, Geneious software v8.1) using the *M. truncatula* *NIN*, *HK2*, *HK3* and *HK4/CRE1* as queries. All hits were manually curated to assure that correct protein coding transcripts were found. To further determine orthology, candidate protein sequences were aligned together with sequences retrieved from additional species and a phylogenetic tree was constructed using the FastTree method (**Supplemental Table 7, Supplemental Figures 2, 3**) (Price et al, 2009). Based on this analysis we determined that, in contrast, where most Papilionoideae only have one copy of *NIN*, we were able to identify two *NIN* copies in both *C. fasciculata* and *C. mimosoides*, which we randomly named *NIN1* and *NIN2*, respectively (**Supplemental Table 7, Supplemental Figure 2**). Similar to most other species, *C. mimosoides* has three genes encoding cytokinin receptors. These were named *HK2*, *HK3*, and *HK4* according to the orthogroup they belonged to (**Supplemental Table 7, Supplemental Figure 3**).

## Development of a nodulation assay and *A. rhizogenes*-mediated transformation protocols

We first set out to design a suitable growth system to conduct robust nodulation assays in *C. mimosoides*. To this end, we explored four different substrates, ranging from soil, perlite, and two mixtures of vermiculite-river sand (v:v 1:1 or 1:9) (**Supplemental Figure 4**). All substrates were supplemented with Fåhrens nutrient solution containing 0.25 mM of  $\text{Ca}(\text{NO}_3)_2$ . 14-day old *C. mimosoides* seedlings were transferred to these substrates and inoculated with WUR077. The number of nodules, fresh shoot and root weight were determined 4 weeks post inoculation. Plants grown on soil were the biggest, both in shoot and root system (**Supplemental Figure 4. A-B**). However, this substrate resulted in the lowest number of nodules (**Supplemental Figure 4. C**). Both perlite and vermiculite:sand (1:9) produced considerably higher number of nodules, although not statistically different from any of the other substrates. As nodule numbers in perlite were highly variable and the leaves of plants grown on this substrate consistently turned yellow (**Supplemental Figure 4. D**), we decided to use the vermiculite:sand (1:9) mixture for the remaining of our experiments. To rule out any negative effect of these growth systems on nodule architecture and development we sectioned ~10 nodules from each substrate. No abnormalities were observed, indicating that our selected growth medium is suitable to analyse nodule functionality in *Chamaecrista* (**Supplemental Figure 4. E, F**). Next, we turned our attention to setting up *Agrobacterium rhizogenes* transformation protocols in *C. mimosoides*. For this, we used the existing *M. truncatula* methods (Limpens et al., 2004) as a template (see **Materials and Methods**). This resulted in protocols where a single compound *C. mimosoides* plant bears up to 90% transgenic roots (**Supplemental Figure 5**).

### *NIN* plays a conserved role in *C. mimosoides* nodulation

In model Papilionoideae, *NIN* has been shown to be essential to successfully initiate a nitrogen-fixing symbiosis (see review by (Liu and Bisseling, 2020)). Moreover, evolutionary studies highlight that presence of a functional *NIN* is one of the key determinants whether the plant species maintains its nodulation capacity (Griesmann et al., 2018; van Velzen et al., 2018). Here, we hypothesise that *NIN* in *C. mimosoides* is also required for nodulation.

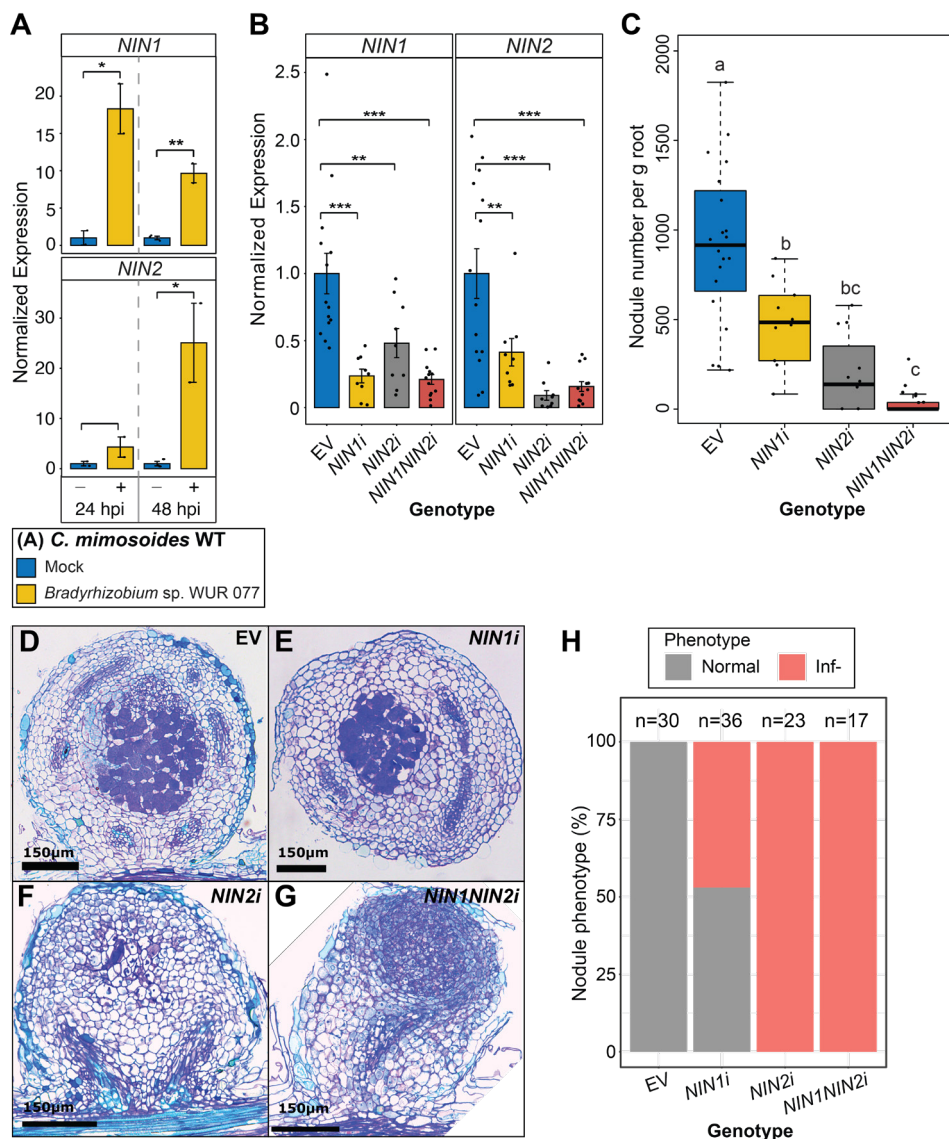
In addition, the expression of *NIN* is highly LCO-responsive in model Papilionoideae (van Zeijl et al., 2015a; Vernié et al., 2015; Jardinaud et al., 2016; Reid et al., 2017). To test whether these responses are also conserved in *C. mimosoides*, we grew *C. mimosoides* seedlings for 10 days on agar plates supplemented with Fåhrens nutrient containing 0.25 mM of  $\text{Ca}(\text{NO}_3)_2$  and treated these plants with WUR077 ( $\text{OD}_{600} = 0.015$ ).

We previously demonstrated that the use of WUR077 at this optical density is sufficient to successfully induce nodule formation on our plate system (**Figure 1**). As 24h and 48h post inoculation are time points where symbiotic responses are clearly initiated (**Figure 1. A-B**) we selected these time-points to investigate the effect of WUR077 on *C. mimosoides* *NIN1* and *NIN2* expression. For this, we isolated a 5 mm region of root at the transition of the elongation to the differentiation zone at roughly 5-10 mm distance from the root tip, commonly referred to as the root susceptible zone (**Supplemental Figure 6. A**). This revealed that *NIN1* and *NIN2* are both induced by rhizobia, suggesting that the harvested root segments are indeed responsive to rhizobia and that these responses include the transcriptional activation of both *NIN1* and *NIN2*. However, it might be worth mentioning that the expression of *NIN1* seems to be induced slightly earlier compared to *NIN2* (24 hours post inoculation vs 48 hours, **Figure 2. A**).

Next, we used RNA-interference to knock down *NIN1* and/or *NIN2* in *C. mimosoides*. Based on the *C. fasciculata* genome database we designed RNAi constructs that target either exon 1 or 2 for *NIN1* and *NIN2*, respectively. This resulted in compound plants in which we could identify three distinct patterns of *NIN1* and/or *NIN2* knockdown. Although off-target silencing between *NIN1* and *NIN2* does occur, these compound plants have either an emphasis on the reduction of *NIN1*, *NIN2* or an equal strong reduction of both *NIN1* and *NIN2* expression (**Figure 2. B**). For this reason, we named these lines *NIN1i*, *NIN2i* or *NIN1NIN2i*, respectively (**Figure 2. B**). Four weeks post inoculation with WUR077, we scored the nodule numbers and collected nodules for sectioning. Nodules were formed on all four lines (**figure 2. C**). We observed that all three RNAi constructs targeting *NINs* in *C. mimosoides* resulted in a reduction in nodule numbers compared to the empty vector control (EV), although the effect is strongest in double *NIN1NIN2i* lines (**Figure 2. C**). Here, several transgenic roots were unable to form any nodules at all (**Figure 2. C**). However, when comparing *NIN1i* to *NIN2i* we noticed that the latter showed a higher reduction in nodule numbers compared to the former (**Figure 2. C**). Next, we analysed sections from nodules collected from all four lines (**Figure 2. D-G**). All the EV nodules formed normal infection threads and symbiosomes (**Figure 2. D, H**), while we observed that only ~51% of the nodules collected from *NIN1i* did so (**Figure 2. E, H**). In contrast, nodules collected from *NIN2i* and *NIN1NIN2i* showed that, although most nodules contained cells filled with infection threads, symbiosomes were never detected (**Figure 2. J-H**). This suggests that *NIN2* might also be involved in the release of rhizobia into symbiosomes.

### Cytokinin perception is required during *C. mimosoides* nodule organogenesis

Studies in model Papilionoideae demonstrated that nodule organogenesis is dependent on cytokinin biosynthesis and signalling (Murray et al., 2007; Plet et al., 2011; van Zeijl et



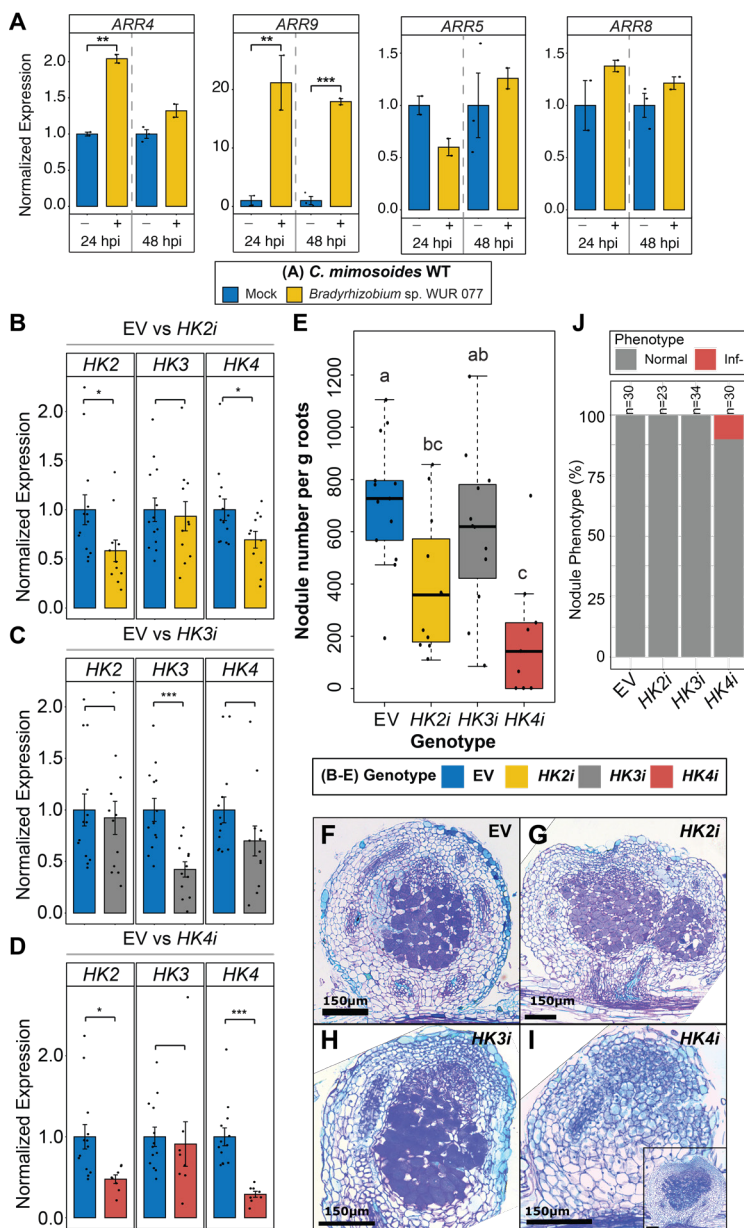
**Figure 2. The role of *NIN* during nodulation in *C. mimosoides*.** (A) Normalised gene expression of *NIN1* and *NIN2* after 24 and 48 hours with (+) or without (-, mock treatment) *Bradyrhizobium* sp WUR077 inoculation (n=3, student *t*-test). (B) Normalised gene expression of *NIN1* and *NIN2* in compound plants bearing transgenic roots expressing EV, *NIN1i* and/or *NIN2i* (n>10, student *t*-test). (C) Number of nodules on compound plants bearing transgenic roots expressing EV, *NIN1i* and/or *NIN2i* analysed at 4 weeks post inoculation with *Bradyrhizobium* sp. WUR077 (n>10, One-way ANOVA followed by post-hoc Tukey HSD for multiple comparisons). (A-C) Bars represent averages  $\pm$  SEM, (A-B) Significance:  $p < 0.001$  (\*\*\*\*),  $p < 0.01$  (\*\*\*),  $p < 0.05$  (\*\*), (C) bars with different letters differ significantly ( $p < 0.05$ ) (D-G) Representative micrographs of Toluidine Blue stained longitudinal sections of mature EV, *NIN1i*, *NIN2i* and *NIN1NIN2i* nodules. (H) Percentage nodules lacking symbiosomes in EV, *NIN1i*, *NIN2i* and *NIN1NIN2i* harvested 4 weeks post inoculation with *Bradyrhizobium* sp WUR077 (Normal, fully infected nodules. Inf-, nodules display only infection threads and apoplastic colonisation).

al., 2015a; Vernié et al., 2015; Jardinaud et al., 2016; Reid et al., 2017). In *M. truncatula* and *L. japonicus*, several cytokinin biosynthesis and signalling genes are induced early upon LCO application. To investigate if cytokinins are part of the early *C. mimosoides* response to rhizobia, we quantified the expression of the cytokinin biosynthesis (*IPT1*, *CYP735A*, *LOG1*) and responsive (*ARR4*, *ARR5*, *ARR8*, and *ARR9*) genes in response to WUR077.

Based on our manual curation against the *de novo* transcriptome assembly of *C. mimosoides*, we were able to identify the orthologs of *IPT1*, *CYP735a*, *LOG3* and a total of four *ARRs* (*ARR4*, *ARR5*, *ARR8* and *ARR9*, **Supplemental Table 1, Supplemental Figures 7-10**). Four *ARR* genes is a lower number than the seven *ARRs* genes annotated in *C. fasciculata* (**Supplemental File 7**). However, we suspect that additional copies of *ARR* genes are only very low expressed in the *C. mimosoides de novo* transcriptome assembly we created, and therefore not detected. We tested cytokinin responsiveness of these genes in *C. mimosoides* and found that, with the exception of *CYP735a*, all were induced within 48 hours of treatment with 1  $\mu$ M of the synthetic cytokinin 6-Benzyladenine (BAP) (**Supplemental Figure 6. B**).

After inoculation with WUR077, *ARR4* and *ARR9* were significantly induced after 24h or 24 and 48h respectively, whereas *ARR5*, *ARR8* were not responsive (**Figure 3. A**). The cytokinin biosynthesis genes *IPT1*, *CYP735a* and *LOG3* were not significantly induced by inoculation with WUR077 (**Supplemental Figure 6. C**), suggesting that different from *M. truncatula* and *L. japonicus*, cytokinin biosynthesis in *C. mimosoides* is not affected by rhizobia inoculation.

For several rhizobia strains it has been suggested that they are able to produce cytokinin-like molecules (Phillips and Torrey, 1972; Sturtevant and Taller, 1989; Kisiala et al., 2013). To exclude that the observed effect of WUR077 on gene expression in *C. mimosoides* is the result of such molecules, we tested the effect of WUR077 on *MtARR4* and *MtNIN* expression in *M. truncatula* jemalong A17 wild type and *Mtdmi3* mutant. The natural symbiotic partner of *M. truncatula*, *Sinorhizobium meliloti* 2011, was used as positive control. A slight (~1.7 fold), but not significant induction of *MtARR4* in the *M. truncatula* A17 wilt type inoculated with WUR077 was observed (**Supplemental Figure 11. A**). However, as this induction is absent in the *Mtdmi3* mutant it suggests that, if it is a biological relevant response, it is dependent on LCO perception and not on direct cytokinin signalling. As suspected both *MtARR4* and *MtNIN* were significantly induced when treated with *S. meliloti* 2011, demonstrating these root samples to be responsive (**Supplemental Figure 11. B**). No induction was observed in *Mtdmi3* mutant background treated with either WUR077 or *S. meliloti* 2011, suggesting that neither strain secretes cytokinin (like) molecules (**Supplemental Figure 11. A-B**).



**Figure 3. The role of the histidine kinase (HK) cytokinin receptors during nodulation in *C. mimosoides*.** (A) Normalised gene expression of *ARR4*, *ARR9*, *ARR4* and *ARR8* after 24 and 48 hours with (+) or without (-, mock treatment) *Bradyrhizobium* sp WUR077 inoculation ( $n=3$ , student *t*-test). (B-D) Normalised gene expression of *HK2*, *HK3* and *HK4* in compound plants bearing transgenic roots expressing EV, (B) *HK2i* (C) *HK3i* or (D) *HK4i* ( $n>10$ , student *t*-test). (E) Number of nodules on compound plants bearing transgenic roots expressing EV, *HK2i*, *HK3i* and *HK4i* analysed at 4 weeks post inoculation with *Bradyrhizobium* sp. WUR077 ( $n>10$ , One-way ANOVA followed by post-hoc Tukey HSD for multiple comparisons). (A-E) Bars represent averages  $\pm$  SEM, (A-D) Significance:  $p<0.001$  \*\*\*\*,  $p<0.01$  \*\*\*,  $p<0.05$  \*. (E) bars with different letters

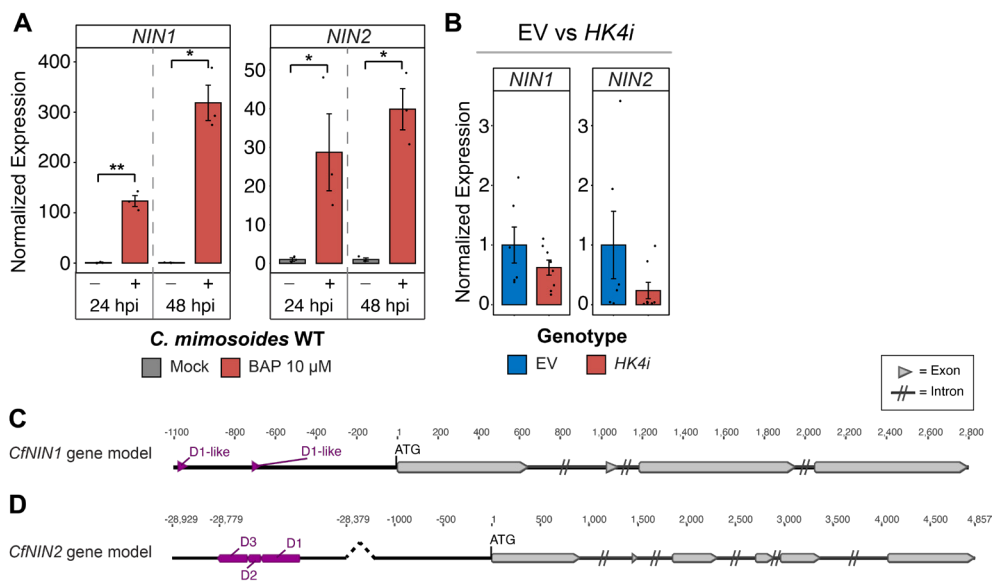
In model Papilionoideae, the cytokinin receptor MtCRE1/LjLHK1 is required for nodule initiation (Murray et al., 2007; Plet et al., 2011; Held et al., 2014; Boivin et al., 2016a). To test whether this receptor is also required for nodule initiation in legumes outside of the papilionoid subclade, we designed three RNA-interference constructs to specifically target unique genomic regions encoding CHASE domain of *HK2*, *HK3*, and *HK4*, respectively. The designed constructs successfully knocked down each HK they targeted, even though with some variation (**Figure 3. B-D**). *HK2i* had some effect on the expression levels of *HK4* (**Figure 3. B**), and *vice versa* (**Figure 3. D**), however *HK3* was only affected by *HK3i* (**Figure 3. C**). Knockdown of *HK3* had no effect on nodulation. Whereas, silencing of *HK2* and/or *HK4* resulted in a clear reduction in nodule numbers (**Figure 3. E**). The effect on nodule numbers in both lines could be the result of redundancy between *HK2* and *HK4*. Functional redundancy between multiple HKs is also described in *Lotus* and *Medicago* nodulation (Held et al., 2014; Boivin et al., 2016a). However, it is possible that the observed reduction in nodule numbers in both lines is due to the partial knockdown of either *HK2* or *HK4*. In addition, roughly 10% of nodules formed on *HK4i* are impaired in infection thread development, compared to the EV control (**Figure 3. F, I, J**), suggesting a possible role of CK signalling during intracellular infection. This phenotype was not observed in any of the nodules formed on either *HK2i* or *HK3i* (**Figure 3. G, H, J**), indicating a possible specificity for *HK4* in this phenotype.

During early nodulation events in Papilionoideae, NIN and *HK4* are part of a feedforward loop (Vernié et al., 2015; Liu et al., 2019c). As we demonstrated that NIN and HKs are also required during *C. mimosoides* nodulation, the question arises if a similar interconnection exists.

To test if either *NIN1* and/or *NIN2* is cytokinin responsive we treated plate grown *C. mimosoides* with 1  $\mu$ M BAP for 24 and 48 hours. Indeed, both *NIN1* and *NIN2* were highly induced after 24 hours (**Figure 4. A**), which the expression of *NIN1* increased further at 48 hours and the expression of *NIN2* stabilised (**Figure 4. A**). Next, we questioned if the symbiotic induction of *NIN* is dependent on any of the cytokinin receptors. To this end, we analysed the expression of *NIN1* and *NIN2* in *HK2i*, *HK3i* and *HK4i*. Although not significant both *NIN1* and *NIN2* were slightly lower in *HK4i* (**Figure 4. B**). In contrast, although again not significant, *NIN1* and *NIN2* expression was higher in *HK2i* and *HK3i* (**Supplemental Figure 12. A, B**). In *Medicago*, the expression of *MtCRE1* is also induced

#### Figure 3 (continued).

differ significantly ( $p < 0.05$ ). (**F-I**) Representative micrographs Toluidine Blue stained longitudinal sections of mature EV, *HK2i*, *HK3i* and *HK4i* nodules ( $n=10-23$ ). (**J**) Percentage nodules lacking symbiosomes in EV, *HK2i*, *HK3i* and *HK4i* harvested 4 weeks post inoculation with *Bradyrhizobium* sp WUR077 (Normal, fully infected nodules. Inf-, nodules display only infection threads and apoplastic colonisation).



**Figure 4. The putative NIN-Cytokinin Signalling crosstalk in *C. mimosoides*.** (A) Normalised gene expression of *NIN1* and *NIN2* after 24 and 48 hours with (+) or without (-, mock treatment) exogenous application of benzylaminopurine (BAP) ( $n=3$ , student *t*-test). (B) Normalised gene expression of *NIN1* and *NIN2* in compound plants bearing transgenic roots expressing EV or *HK4i*, ( $n>10$ , student *t*-test). (C-D) Schematic gene model of the *C. fasciculata* (C) *NIN1* and (D) *NIN2*. Exons highlighted in grey. Position of the predicted cytokinin response CE element (D1-D2-D3) or D1-like in the corresponding promoter region highlighted in purple.

during early nodulation. To test the similar hypothesis, we analysed the expression level of *HK4* in EV, *NIN1i*, *NIN2i* and *NIN1NIN2i* in *C. mimosoides*. Again, no significant changes were observed (Supplemental Figure 12. C).

It was recently reported that in the Papilionoideae, a highly conserved cis-regulatory element (the so-called CE motive) is required for the mitotic activation of pericycle cells during nodule initiation in *M. truncatula* (Liu et al., 2019c). This CE motive contains several putative type B-RR binding motifs and is located ~17 kb upstream of the *NIN* start codon. We utilised the *C. fasciculata* genome database once more and detected a similar motif located ~28kb upstream of the *NIN2* start codon (Figure 4. C). No CE motive was detected upstream *NIN1*. However, a few putative type B-RR binding sites were predicted closer (~1 kb) to the start site of both *NIN1*.

## Discussion

The symbiotic association between legumes and rhizobia is the most extensively studied interaction within the nitrogen-fixing clade, with a particular focus on the Papilionoideae. In this sub-clade, nitrogen-fixing symbiosis is a predominant trait, whereas in the Caesalpinioideae sub-clade it is less common. As a result, only a few studies have been performed in the Caesalpinioideae. The genus *Chamaecrista* is particularly interesting, as species belonging to this genus display a diversity in infection modes. Some *Chamaecrista* species develop persistent fixation threads, whereas others produce the more 'advanced' symbiosomes (Naisbitt et al., 1992; Sprent et al., 2013; Ibáñez et al., 2017). In *C. mimosoides* we observed infection threads that are released as symbiosomes, which agrees with previous descriptions (Naisbitt et al., 1992).

Santos and colleagues (2017) analyzed nodules from six different *Chamaecrista* species (i.e. *C. desvauxii*, *C. rotundifolia*, *C. supplex*, *C. pascuorum*, *C. serpens* and *C. flexuosa*) and concluded that all species developed indeterminate nodules (Santos et al., 2017). However, Ren (2018) concluded that in *C. fasciculata* a determinate nodule is formed, but that a secondary cluster of dividing cells at the nodule apex is activated later during nodule development. In *C. mimosoides*, we observed that morphology during nodule initiation closely resembles that of *C. fasciculata* nodules, suggesting that in this species a similar secondary cluster of cells is activated at a later stage. This indicates that, at least in some *Chamaecrista* species, a determinate nodule is formed that displays an indeterminate growth phenotype (Ren, 2018). It is possible that this secondary cluster of dividing cells was previously mistaken for a nodule meristem (Santos et al., 2017), making the reactivation of growth in a by definition determinate nodule the default *Chamaecrista* nodule type.

Recent phylogenetic studies comparing nodulating to non-nodulating species revealed that genes essential for nodulation, i.e. *NIN* (*NODULE INCEPTION*), *NFP* (*NOD FACTOR PERCEPTION*), and *RPG* (*RHIZOBIUM-DIRECTED POLAR GROWTH*) are lost, or pseudogenized, in the non-nodulating species (Griesmann et al., 2018; van Velzen et al., 2018). This led to the revival of the single gain, massive loss hypothesis (Soltis et al., 1995). Genetic analysis in *Medicago* and *Lotus*, demonstrated the importance of NIN during rhizobia nodulation in Papilionoideae. NIN was also shown to be required for nodule initiation in the only non-legume nodulating with rhizobia, *Parasponia andersonii* (Rosales) and the actinorhizal species *Casuarina glauca* (Fagales) (Bu et al., 2020, Clavijo et al., 2015). In addition, NIN is required for *Frankia*-induced root hair deformation, suggesting a role for NIN during infection as well (Clavijo et al., 2015).

Combined, this suggests a conserved function for *NIN* during nodulation exists within the whole nitrogen-fixation clade. The fact that we found a similar importance for *NIN* in *C. mimosoides* is in line with this hypothesis.

We discovered two copies of *NIN* in both *C. fasciculata* and *C. mimosoides*. *Chamaecrista* species did not share the whole genome duplication that occurred in Papilionoideae (Cannon et al., 2010), therefore it is unknown what gave rise to this second copy of *NIN*. Nevertheless, the fact that both copies are transcriptionally activated by *Bradyrhizobium* sp WUR077 application at an early stage, suggests that both are able to function during nodule initiation. In model species like *M. truncatula* and *L. japonicus*, it was demonstrated that during nodule initiation *NIN* expression is activated twice, once in the epidermis and later a second time in the pericycle and inner cortex (Breakspear et al., 2014; Yoro et al., 2014; Vernié et al., 2015; Liu et al., 2019c). In line with this, *NIN* has been shown to function in both nodule initiation and rhizobia infection (Breakspear et al., 2014; Yoro et al., 2014; Vernié et al., 2015; Liu et al., 2019c). The two copies of *NIN* present in *Chamaecrista* could potentially share these functions between them, a hypothesis possibly supported by the fact that the promoter regions of *NIN1* and *NIN2* differ considerably. Although our *NIN1* or *NIN2* knockdown experiments experienced co-silencing of the other *NIN* copy to some degree, they suggested both copies to be involved in nodule initiation. Nevertheless, our results leave the possibility that *NIN2* might be the only copy of *NIN* in *C. mimosoides* involved in bacteria release. However, as the observed effect on nodule initiation and infection could also be the result of the cumulative reduction of both genes, it remains to be seen if more specific knockdown or knockout lines targeting *NIN1* or *NIN2* separately can be created in *C. mimosoides* to answer this question. In addition, *in situ* hybridization or promotor GUS analysis could elucidate the timing and localization of the expression of both genes to shed light on possible symbiotic specificity.

Similarly, we were able to demonstrate that, like in *M. truncatula* and *L. japonicus*, nodule initiation in *C. mimosoides* is dependent on cytokinin perception. Although our methods are not able to properly discriminate between *HK2* and *HK4* silencing, the nodulation phenotypes (i.e. number of nodules and intracellular infection) clearly correlate with *HK4* knockdown. The role of cytokinin perception, possibly through *HK4*, has never been shown outside the Papilionoidea and therefore is an important discovery in light of the single gain, massive loss hypothesis. In that sense it would be interesting to see if any of the HKs fulfil a similar function in *Parasponia andersonii* or any of the actinorhizal plant species. The observed infection phenotypes in *HK4i* are in line with phenotypes described in *M. truncatula* and *L. japonicus*, although here additional silencing of *MtHK2* and *MtHK3* or the creation of *lhk1-1 lhk1a-1 lhk3* triple mutants, in *M. truncatula* or *L. japonicus* respectively, are required to observe these phenotypes (Held et al., 2014,

Boivin et al., 2016b). This suggests that in *C. mimosoides* infection is regulated by HK4, although with the off-target effect on *HK2* a similar redundancy as in Papilionoideae cannot be excluded.

Many cytokinin responsive and biosynthesis genes are transcriptionally activated during early LCO signalling in *M. truncatula* and *L. japonicus* (van Zeijl et al., 2015a; Vernié et al., 2015; Jardinaud et al., 2016; Reid et al., 2017; Liu et al., 2019c). Although two copies of Type A cytokinin Response Regulators (*ARR4* and *ARR9*) are induced 24 hours post inoculation with WUR077, cytokinin biosynthesis genes orthologous to those activated in *M. truncatula* (*IPT1*, *CYP735a*, *LOG3*) were not. This suggests that either cytokinin perception is elevated independent of cytokinin levels, or that cytokinin biosynthesis is activated using different genes or post translational.

The current model of nodule initiation in model Papilionoideae centers around a NIN-cytokinin feedforward loop. Shortly after the perception of rhizobia LCOs, *NIN* is activated in a *MtCRE1/LjLHK1* dependent manner (van Zeijl et al., 2015a) and exogenous cytokinin application is able to induce the expression of *NIN* (Liu et al., 2019c). In addition, a cis-regulatory element containing ten Type B-RR binding motifs which could be linked to the cytokinin related induction of *NIN* was discovered (Liu et al., 2019c). The fact that this cis-regulatory element is likely present in the upstream region of *NIN2* in *C. fasciculata*, but not *NIN1* suggest that if this feed forward loop exists in *Chamaecrista* it might be restricted to *NIN2*. Although the additional type B-RR binding motifs identified close to the *NIN1* start codon could be involved in the cytokinin responsiveness of *NIN1*, the fact that both copies of *NIN* are responsive to exogenous cytokinin application suggest a more complex regulation. Even though not statistically significant, *NIN1* and *NIN2* expression showed to be induced less in the *HK4i* background compared to EV control, whereas in *HK2i* and *HK3i* both were induced more. This further suggests a it is possible that a NIN-cytokinin feedforward loop similar to that found in Papilionoideae could also exist in *Chamaecrista* species. However, whether this holds true for other Caesalpinioideae remains to be seen.

Altogether, we demonstrate *C. mimosoides* is a good model to study legume nodulation outside the Papilionoideae sub-clade. Our data indicates that NIN and cytokinin signaling (likely through HK4) fulfil conserved roles during nodulation in Caesalpinioideae. In addition, our data suggest that a similar NIN-cytokinin feed forward loop as described in Papilionoideae could also exist in Caesalpinioideae.

## Materials and Methods

### Plant materials and growth conditions

*C. mimosoides* seeds were scarified using concentrated  $H_2SO_4$  for 10 minutes, and subsequently washed 6-7 times with MQ water. Next, seeds were sterilized using commercial bleach (~4% hypochlorite) for 6 minutes and washed 6-7 times with MQ water. Afterwards, seeds were incubated in sterile MQ water at room temperature for ~3 hours, and placed on filter paper topped agar plates containing Fåhrens medium. All plates were sealed for 2/3 by parafilm and covered by aluminum foil. To synchronize germination, seeds were placed at 4°C in darkness for 1 day and subsequently transferred to room temperature to induce germination for 2 days. Plates were positioned upside down to enable the emerging roots to grow vertically. To perform nodulation assays and exogenous cytokinin applications on plates, germinated seedlings were grown on square 12cm Fåhrens medium (Fahraeus, 1957) containing plates using a low-nitrate concentration [0.25 mM  $Ca(NO_3)_2$ ] for 10 days (21°C/18°C 16h/8h light/dark regime). The seedling roots were either treated with *Bradyrhizobium* sp WUR077 ( $OD_{600} = 0.015$ ) or benzylaminopurine (10  $\mu$ M, BAP) and tissues (root susceptible zone) were harvested at 24h and 48h, snap frozen and stored at -80°C until used further (**Supplemental File 6. A**).

### Vectors and constructs

For RNAi-mediated knockdown of *CmHK2*, *CmHK3*, *CmHK4*, *CmNIN1*, and *CmNIN2* (**Supplemental Table 1**), the corresponding target sequences were cloned into pENTR-D-TOPO (Invitrogen, Carlsbad, USA) and recombined into the DsRed-modified gateway vector pK7GWIWG2(II)-RR driven by the CaMV35S promoter (Limpens et al., 2004). This resulted in the creation of binary constructs pK7GWIWG2(II)-RR-p35S-*CmHK2*-RNAi, pK7GWIWG2(II)-RR-p35S-*CmHK3*-RNAi, pK7GWIWG2(II)-RR-p35S-*CmHK4*-RNAi, pK7GWIWG2(II)-RR-p35S-*CmNIN1*-RNAi, pK7GWIWG2(II)-RR-p35S-*CmNIN2*-RNAi and pK7GWIWG2(II)-RR-p35S-*CmNIN1NIN2*-RNAi. As empty vector control, the binary plasmid pK7GWIWG2(II)-RR-p35S-RNAi-control was used as previously described (van Zeijl et al., 2015b).

### Plant transformation

Transformation was performed as previously described by Limpens (2004) with minor modifications. Co-cultivation of seedlings with *A. rhizogenes* MSU440 was performed for 7 days and roots developed during this period were removed before plants were transferred to root emergence medium. In addition, 25 mg/L kanamycin was added to the emergence medium to select for transformation events. At 15-20 days post transformation, the DsRed fluoroform was used to identify transformed roots.

Untransformed roots were removed and plants with sufficient roots left were transferred to vermiculite:river sand mixture (1:9) supplemented with Fåhres medium [0.25 mM  $\text{Ca}(\text{NO}_3)_2$ ], inoculated with *Bradyrhizobium* sp WUR077, and grown for an additional four weeks (21°C/18°C 16h/8h light/dark regime). After this, nodule numbers were determined, root material was collected to validate gene knockdown efficiency. In addition, nodules were harvested at 4 wpi for further analysis.

### Histochemical analysis and microscopy

For plastic sections, nodules of all transformed plants were fixed in 4% paraformaldehyde (w/v), 5% glutaraldehyde (v/v) in 50 mM phosphate buffer (pH 7.2) at 4°C for 24 h. Subsequently, the samples were dehydrated using an ethanol series and embedded in Technovit 7100 (Heraeus Kulzer, Hanau, Germany) according to the manufacturer's instructions. Semithin sections were cut using a Leica Ultra-cut microtome (Leica Microsystems, Wetzlar, Germany) to 5 µm thickness for nodules and were stained with 0.05% Toluidine Blue. Images were analysed using the DM5500B microscope equipped with a DFC425C camera (Leica). Bright-field and fluorescence images of transgenic roots and nodules were taken using the M165FC stereomicroscope (Leica).

### RNA isolation and qRT-PCR analysis

RNA was isolated from root samples using the EZNA Plant RNA mini kit (Omega Bio-tek). For cDNA synthesis, either 250 ng or 100 ng RNA was used with the iScript cDNA synthesis kit (Bio-Rad) final volume 20 µL and subsequently diluted either 10 or 4-fold, respectively. Real-time qPCR was performed using SYBR Green Supermix (Bio-Rad) and the CFX real-time system (Bio-Rad) at 10 µL volume per reaction and three technical replicates. Level of gene expression was determined with gene specific primer combinations and normalized using UBC and *EF1a* for *C. mimosoides* and *M. truncatula* samples, respectively (**Supplemental Table 2**). Primer efficiency was determined based on a dilution series of total cDNA.

### Phylogenetic constructions

TBLASTX using default settings was used to identify putative homologues of *NIN*, *HK2*, *HK3* and *HK4/CRE1* in *C. fasciculata* and *C. mimosoides* using *M. truncatula* CDSs of the aforementioned genes as the query. To verify the predicted gene model in *C. fasciculata* genome database and to search for open reading frames (ORFs) in the mRNA sequence of *C. mimosoides* transcriptomic database, we used the ORFfinder online tool (<https://www.ncbi.nlm.nih.gov/orffinder/>) and Fgenesh+ Protein prediction method (<http://www.softberry.com/berry.phtml>). All hits were manually curated to assure that correct protein coding transcripts were found. Protein and nucleotide alignments were performed based on MAFFT method (Katoh et al., 2002). The phylogenetic trees were

reconstructed by maximum likelihood-based FastTree (Price et al., 2009) method. All steps in this procedure were done in Geneious Software v8.1.

### **Statistical analysis**

Graph visualization and statistical analysis were performed using R studio Version 1.1.463. Bars represent mean  $\pm$  SE for all experiments. Elements in the boxplot illustrate the lowest and highest data points, the sample median, and the first and third quartiles. All individual data points are plotted to visualize actual sample variation. When appropriate, the R package Agricolae (Mendiburu, 2020) was used to perform One Way Analysis of variance (ANOVA) followed by post hoc Tukey-HSD test was used to determine statistical significance ( $p < 0.05$ ). In other cases, students t-tests were used.

### **Author Contributions**

Conceptualization, T.A.K.W., A.v.Z., and W.K.; Methodology, T.A.K.W.; Investigation, T.A.K.W., L.R., J.K.; Writing – Original Draft, T.A.K.W.; Writing – Review & Editing, T.A.K.W., and W.K.; Funding Acquisition, R.G., and W.K.; Supervision, W.K.

### **Acknowledgments**

The authors would like to thank Lotte Azink, Sverre Beijk, Kirsten Bronsvort, Bas Janssen, Mathangi Lakshmipathi, Thijs de Vroet, and Çağlar Yıldız for their contribution in this project, as a part of 'Molecular Biology Toolbox' project. This work was supported by NWO-VENI (863.15.010) to W.K and NWO-VICI (865.13.001) to R.G.

## Supplemental Information

### Supplemental Table 1. Sequences of RNAi-targeting *HK2* in *C. mimosoides* designed from the transcriptomic database of *C. mimosoides*.

AACTGTTTGAGTCATGGGTAATGGAGTTATGTGCAACTCAGCATTTGTCTGAATTGGTCTAACTTGGTTAAAGGAT  
GCTGGTGGATTTTGGTTGGAATTACAATGTGCTGCAAATTGTCTGGCTTCTTTTATGTTGGGTGGAGTCCAAATCC  
GAACTGGTTGCTCAGCAGAAGCAGATGCAACATTCTGTCTCAGGAGAAATCCAAAGGGTGGTGGCAGTGGAGGAAA

### Supplemental Table 2. Sequences of RNAi-targeting *HK3* in *C. mimosoides* designed from the transcriptomic database of *C. mimosoides*.

CTTTTGGCTTTATGTTGCTGGATTGTCTCTGTAATTTACATCAACTGGTTTAGCAACGGTGTGATTATGGACACCA  
AGGTGGTTTACTTGGTGGTGGTGGTGGTAGTAGCAAGATTTGGCTCAATTGGTGGGAGAAGATCTCCGG  
CAAGAGCGGTAAAATCTACCAACTACAACGAGTATATTGGGTACAAGAAAGCGAAGCGAACTCGGCGGAGATGG  
ATTTTCATG

### Supplemental Table 3. Sequences of RNAi-targeting *HK4* in *C. mimosoides* designed from the transcriptomic database of *C. mimosoides*.

GCTGAAAATAGAAATGCTGCAATGAATGGAAAATCTGAGACTTATCTAAATGGGGGATCAGACGAAGTTATGCATA  
TGTCTGGTGGTTATCAGTTCAAAACCCTAAGTGGATGTGAAGCAGCTGATGAACGGAACAGCTGGGATAATTTTAA  
GCATCTAATTGCAGATGAAGAATTTGCTTTGATGCTTCAAGTAAGAAGGTGACCACTGTTCAAGCTTCAGAGCTA  
GTCAC

### Supplemental Table 4. Sequences of RNAi-targeting *NIN1* in *C. mimosoides* designed from the genomic database of *C. fasciculata*.

AACACAAACAGTCTTGAACAGCATAAAGATTGGGAGATGCAGCATCAGCAGGAAGAATATATGTATGGTGGTGGGG  
AAGTAGTTGAGAATTGTGGATATGGGAGGTTGTGTGAGGCAATCCAGAAGCAGCATTGATTGATGAGTTGCTTGT  
ACAAGGATGCTGGGTAGAAGCCAGTGGTGT

### Supplemental Table 5. Sequences of RNAi-targeting *NIN2* in *C. mimosoides* designed from the genomic database of *C. fasciculata*.

TGGAGGAGTATGATGGGGCATTGAGAGCTTGTGAGAAAGTGCAGCATCAGTATCATCAGAGGTATCATTAGTTAT  
GGATTTCATTGATGACTTGTAGTACACGTTGCTGGCTAGAAACATCATCATACATCAACAACAACCTCAATAAT  
AGCAACAACAACCACCTTCTTCC

**Supplemental Table 5. Primers used in this study.**

Name	Sequence	Purpose
CmiHK2-RNAi_Fw	Ggggacaagtttgtacaaaaaagcaggcttc AACTGTTTGCAGTCATGGGT	Gateway
CmiHK2-RNAi_Rv	Ggggaccactttgtacaagaaagctgggtc CCCTTTGGATTCTCTGACA	Gateway
CmiHK3-RNAi_Fw	Ggggacaagtttgtacaaaaaagcaggcttc CTTTTGGCTTTATGTTGCTGG	Gateway
CmiHK3-RNAi_Rv	Ggggaccactttgtacaagaaagctgggtc CATGAAAATCCATCTCCGCC	Gateway
CmiHK4-RNAi_Fw	Ggggacaagtttgtacaaaaaagcaggcttc GCTGAAAATAGAAATGCTGCA	Gateway
CmiHK4-RNAi_Rv	Ggggaccactttgtacaagaaagctgggtc GAACAGTGGTCACCTTCTTA	Gateway
CmiHK2_Fw	GTATCCAATCAGGTGCCACTAG	qRT-PCR
CmiHK2_Rv	TGGGCTCTTAACTTCATTCGT	qRT-PCR
CmiHK3_Fw	GTTGCCCATATCATTTCTGTCTG	qRT-PCR
CmiHK3_Rv	AATGTAAGTATAACTCCGAGCCG	qRT-PCR
CmiHK4_Fw	CCCTAGCCATTCTCGTTTCTAC	qRT-PCR
CmiHK4_Rv	TCAGTGTCAACTACCCTTTGTG	qRT-PCR
CmiNIN1_Fw	CAAACCCACAACCACAACCTAC	qRT-PCR
CmiNIN1_Rv	CACCTCGCCAATTAGTTACCT	qRT-PCR
CmiNIN2_Fw	TGATTGAGATTGGGTGCCTCC	qRT-PCR
CmiNIN2_Rv	CACCACCACCACCACCATAA	qRT-PCR
CmiARR4_Fw	GCAATTTCTGGGACTTCACG	qRT-PCR
CmiARR4_Rv	TCCAGGCATGCAGTAATCAG	qRT-PCR
CmiARR5_Fw	CTGTGGATTCTGGTAGTAAGGC	qRT-PCR
CmiARR5_Rv	GAAGATGTGTTTGCCTCGATC	qRT-PCR
CmiARR8_Fw	ATCCGGCGAACTTCATGTC	qRT-PCR
CmiARR8_Rv	AAAGCTCTGGTTCCACTATCG	qRT-PCR
CmiARR9_Fw	GATGAGCAGAGGAACGCAGA	qRT-PCR
CmiARR9_Rv	TCCTGGACGGGACATTCTCT	qRT-PCR
CmiIPT1_Fw	AGATTCATCGCCTCCATTCC	qRT-PCR
CmiIPT1_Rv	AGGAATTGTGGAGAATCCGG	qRT-PCR
CmiCYP735_Fw	TCTCACGACATTGTTTCCCG	qRT-PCR
CmiCYP735_Rv	TGGCGTATTGAGAGGAAG	qRT-PCR
CmiLOG3_Fw	AGGCTTTATCTCACCACCGG	qRT-PCR
CmiLOG3_Rv	TCTATCTTCCCACCAACTTG	qRT-PCR
CmiUBC_Fw	ATGAGGGTGGTGTCTTCCAG	qRT-PCR

CmiUBC_Rv	GGCTCCAAGCATTTTCAA	qRT-PCR
MtUBQ10_Fw	CCCTTCATCTTGTCTTCGTCTG	qRT-PCR
MtUBQ10_Rv	CACCTCCAATGTAATGGTCTTTCC	qRT-PCR
MtNIN_Fw	GGGAGAAAGTCCGGGGACAA	qRT-PCR
MtNIN_Rv	GACACACACCGATGCTCTTTGC	qRT-PCR
MtRR4_Fw	TTGTTCCTCGGGTTAAAGGTG	qRT-PCR
MtRR4_Rv	GCGTGGCAAGATATTTTCAGAAG	qRT-PCR
MtRR9_Fw	TCCTCAGAGAATGTCCATCAAGG	qRT-PCR
MtRR9_Rv	TGTGGTTTCAGCTTGTTACATC	qRT-PCR

Supplemental Table 6. *C. fasciculata* nucleotide sequences**>*C. fasciculata* NIN1-CDS\_scaffold2145\_cov214**

ATGCAGCATCAGCAGGAAGAATATATGTATGGTGGTGGGGAAGTAGTTGAGAATTTGGGATATGGGAGGTTGTGTG  
 AGGCAATCCCAGAAGCAGCATGTATGATGAGTTGCTTGTACAAGGATGCTGGGTAGAAGCCAGTGGTGTATATC  
 TCATCATGTGAATCAGCAGCAGTACATGTCTCATCAGAATAATTCACAAAAAATGGTGGATCGGACCAATTCCA  
 TCTTCTTCAGTGAAAGAAAGATTACTAGTTGCCGTAGCCTACTTGAGACACTACTACTACACCAACATTAGTA  
 ATATTCAGATATGGGTCCCTCCAAGTCCCACAAATCTTCACCCCTCACTACGCTTATGATGATCATGAAGCGAATAC  
 AGAAGATGTCACTAGTGTTCCTACCAAACTACTTCAACAGTACTACTGGATTACACCTATTAATAATGTTTCGC  
 TTCTTCAGAACTCATCACTGTTTATCCTCACAGCATTAACACAGGATGATGGTGGTGAGCCCTCTTTCGCTGTCT  
 TTCAAAGAGGCAGTGGTGCATGCCTTGGTGTTCCTCAGATTCTCATGGATCACCAATCTAACATATCAACTATCC  
 CCCACAACCACAACACTACTTCTCCATGCAATCATGCAGTTGAGTTGAAGAGTGGTCATCAGAATTGATCCCTCCT  
 GCTGCTACTAATAATAAGTAACTAATTGGGAGGTGGAGCGAGGTGCAATAATAGAAGAGATAATAGAGATATGA  
 GGAGTGTGTGGAAGACTCATAATTTGGCTCTGGTCTGACGTGGGGTCCCTGCATCCAGCAAGGGAAGAGTGGATG  
 CAGTAGTGAAGAATGTGTATCCACGGTGGATTGTGCTTGTAGTGGGAGATGTAGACATAGTGGGATTCAGGAAC  
 GCGTCAGTGTAGTGGCATTGTGGGAAGGCAAGGATAGTTGGAACAGCTTTAACAAACAACAAAACCGTGTTTTG  
 CAACAGACATAAGTGGCTTTAGTAGAGCTGAATATCCCTTGGCTCTCCATGCCACCATTGTTGGATTGCATGCTGC  
 TCTTGCTATTCCTCTCTACTTCAGGACTAATACAGCAACAGCTGATTTTGTGTTGGAATTTTCTTACCTAAGGAT  
 TGCCGAGATAGTGAACAACAGAGGAGATTGTGAAGTCTTGTGCATGGTGGTACACCAAGCTGCTTGCCGGAGTA  
 ATTTACATGTGGAAGTAGTTGGCTTTGAACAAGATCATGATGATCATGAATTAATGAAAGAAAACAAACAGAAATC  
 AGCTTAGTGTCTGCTGATTGGCAGCTTGAAGAGGCAAAAGCAGGAGTTCAAATGACAACAAGCATGTGTCTTACA  
 ACTCTGAAGAGAATATGCAGACAACATGGAATAACAAGGTGGCCTTCAAGGAAAATCAAGAAGGTAGGCCACTCTC  
 TCAAGAAATCCAACTAGTGTGATTGAGTTGAGGGAGTGGAGGGTGCCTTACATTGGCTCCTTCTATTCAAAA  
 CTTTCCACAATTGACTACCACTCTTCTGCCTCTTGAATCATCATAATAATTTCTAATCCTCCAACCTCCAACAAT  
 ATAATTACAACACTAACCACTTAATTAACAATATGGTAATGAAGGGTCCACAAGTTTAACCAATCCAAATCTT  
 CACCTCCAGTTGAGTGGGAGTGAACCTATAATAATGCAGAAGATGCAGGGTCAAATAGGTTCTCATCAGATGA  
 CACTGAGAGTTTGTAGTTTTCAGAGTGAAGCTAGTTTGGAGATGAAAACATCCGTTTTCAGCTTCCCACCAATTTGG  
 GGTCTTAGAGATTTGCAAAATGGAGATTGTAGAGCTTCAATCTAAACCTAAAGCAGCTCTACCAATGTTGATC  
 TCAAGTATCTGGATGATGATGGTGAATGGTGCTTTTAAACATGCGATGCTGATTTTGAAGAATGTAAGACTTGCA  
 TTTGCATACTACACCATCCTCCTCTCATCTCATACTACCCTTACTATCAGACTCTCCTTGTGTTCAACCACTGCT  
 TTGCCTTCGAATCTACCAATTCATTACATCAATTATATCAATATCATTTTTAA

**>*C. fasciculata* NIN2-CDS\_scaffold2145\_cov214**

ATGGAGGAGTATGATGGGGCATTGTGAGAGCTTGTGAGAAAGTGCAGCATCAGTATCATCAGAGGTATCATTAGTTA  
 TGGATTTTCATTGATGACTTGTTAGTACACGGTTGCTGGCTAGAAACATCATCATACATCAACAACAACCTCAATAA  
 TAGCAACAACAACCACTTCTTCCATTTACAACTGTATAATACCTTCTTCTTCTTCTTCTTCTTCTTCTTCACTACT  
 ACTTCACACTACTACCCTAATAATAATTACTTACAACACTACTGTTTCTGGACATCCTCAGACAAATCTTTCATCAT  
 CATCAACAACATCATCAGAAATCTATCATCATCAACATCATCAACAACAACAACATAGAAAGCTTTGTTCTTGA  
 CATCACCACCAGTAATACTAATAATAATGATGAATTATTAGTGGGAAAAAGATGGTGGATTGGACCAAGGCA  
 AATCCAGGTCTTCTTCATCAGTGAAGGAAAGATTAGTGCTTGCTGTTGGGTGCTTGAAGGAATTATATTACACCA  
 CCACCAACAACAACAACAACAACAACAACACTGCTAATAAAAAATAATTTTCATGATTGAGATTGGGTGCCTCCTCA  
 TCATCATGAAGATTATTACAGAAAGTTGCATCATCATAATGATTATTATGGTGGTGGTGGTGTGTTAGGTTTTTC  
 AGGAGCCATGAGTATCCATATGCTAATTATTATTTGGTGAATGGATCACAACAATACTATGATAATGTCCGAGGAGC  
 CATGTTTGGCTCTTCTCATGTTTGTAGAGAGGAAGTGGGACATGTCTTGGTGTGTTGAGATTTTAACTCATCA  
 TCATGCTAATCATCACAACAACAACAACATCAACAACATCAATAGTAATCTCTCCCATATCTCCAGGCTGTTGAT  
 TTTAGGAGCTGCCAGAACCTCATCCCTCCTCCCATGAAGGTATTTGATGAGTGGTACCAAGCTGCACATAAATGAGA  
 TAGTGGATGTGTTATCAACAGTATGCAAGACTCATAATTTGCCTTTAGTCTCTCACTTGGGACCCCTGTGTTCAACA  
 AAACAAGACTGGATGTCTCCCATTAATAATAATAACAACAACAACAACATCTTCAAGCATGATCAGATTTCAT  
 CAAAACAACAACAATTATATGTGGTGTGTGTCCACTGTGGATTCTGCTTGCTACATAGGTGACTTAGACTTATTAG  
 GCTTCCAAGAGCCTGCTCTGAGTACCATCTTTTCAGAGGTCAAGGCATCGTTGGAACCGCCTTCTCAACCGCCAA  
 ACCGTGTTTTGCAACCGACATAACCGCCTTCACCAAGCCGGAATACCCCTCTTCTCACCACGCTACTATGTTTGGC  
 TTGCTATGCTGCTGTGCTATTCTCTAAGGAGTTCTGTGTATACAGGATCATCAGATTTTGTGCTGGAGTTTTCT  
 TGCCAAAAGATTGTCATGACATTGAGCAGCAGAGAAGCAGATGCTGAATTCGCTTGGTGGTTCACCAAGCTTG  
 TTTCAGAAGCTTACATGTTGTTTCTTCAAGTTGTTGATAATCAGGAAGCAGAATTCATGTTTAAACAATAATAAC  
 CATCCATCTTCAAAGGAATCATCATGGATTGCTCATATGATGGAGGTTCAAATCAGCAAAAGGAAAAAGGTGTTT  
 CTGTTTCATTGGAATACCTTGAAGAGCCAAAGCAGAGTTCAAAGTGACTACTAATTTGGGATATTAGTGGTAGTCA  
 TCATGGAGGGTTATGCCATAATGGACAAGTTTTTCTTTCAGATTTTGGACAAGTTTCATCAGCAGCAGCAAGCTCA  
 GCTTCAAGAGCCACCGTGGAGGGCGGAGGAGAGTCCATAGTTACGGGGGACGCCGGTCTCCGGGGGAGAAAAAT

CCGGCGAGAAGAGACGGACTAAGCGGGAGAAGACTATTAGCTTGCCTGTTCTCAGACAATACTTTGCAGGAAGTCT  
TAAAGATGCTGCCAAGAGCATTGGTGATGTCTTACCACACTCAAAAGGATATGCAGACAACATGGGATCACAAAGG  
TGGCCCTCAAGGAAAAATAAGAAAGTAGGACACTCATTAAAGAACTTCAGCTTGTGATTGATTCTGTTCAAGGTG  
CTGAGGGTGCCATTCAAATAGGGTCTTCTATTCAAGTTTTCCAGAATTAAGCACCACAGATTTGTCTTCTGGTGT  
CCCTCACAGTTCTCAATCTCTAGTAACAACATTAGTGACCATTCCAAGAAAACCCATCATCATCATCATCAA  
CATCATCATAGATTACTAATGGCTTACTGAAATCTCCACCCTCTTGTTCAGCCAGACTTCAACCATCATCA  
ACAATAAGCCATCAATAACAATGGAGATATTATCATGACAGAGAACCAGAACGCTTCAATAGCAGAGTTCA  
TACTCAAAATCATGCTCTCATTGAAGCAGAATTGCTTGAGGATGCAAAGCATAATCATAATCATCATCATCTTGAG  
GCTCTTCTCCATTGCCAGCTGAAATACTACTCATCATGGAGGGGCTTTTAGAGTGAAAGCAACTTTTGGGGATG  
AAAAATCAGATTAGCTTGCACCCCAATTGGGGATTGAGAGAATTGCAATGGAGATCACAAGAAGATTCAATCT  
CAATAATAATATTAATCATGATGTGATGATGAACAACCTTGATCTCAAGTATCTTGATGATGATGGTGAATGGGT  
CTTTTGACATGTGATGATCTTGGAGGTGTAAGGACATATACAGATCATCTTCAGAGCCGTACTATGAGGTCT  
CTCTTTTTCAAGCTTCTCTATGAATCTAGCATCAGATACTTTTGGGAATACAAGCAGCAGCCCATCTTGA

### >*C. fasciculata* HK2-CDS\_scaffold3128\_cov210

ATGTCTGTGAATCGAAAGCTTTCTTCGCAAAGTTCCAAGTTAGGGAAAGCAAATAATGAGGCTCTGCATGAGTCTA  
ATTCTCTCAAGAAATGCAAGAGGAAGAAACCCTTGTTCTTCTTCTGCGTTTTTCTTGCTGTTGTTGTTATTGCTGT  
GGAACTGTGTGTTTTTTCATGAGTTTTTACAGTGAATCATCATTGTTGATGAAGAAGGATAAAAGGTGGGATGAT  
TATGAAGAGAGAGCTCGAATCTTGGCTTCAGCGTTACAATGTACAGCAAGAGCAGCTTCATGCTCTGATTCTCTTT  
TCTCTGTATCAGATCAGATTTTATTGCAATGTATTGACGAAAGAGGAGTGAAAAAGCTTTTAAAGCAGTGACTTTGC  
CAGTACTTCACAGTCGATATGTCCAGAGAATCGGGAGTGCCAAAAGGAGCATACATGTGTCCACAGAAATGTAGAA  
CTCATAGAACAATGCCAATTCTCGATGATTATGTCCAGAGAAATCTGGAATCATCACCGATAGTGGAaaaaactg  
TATCATTTGGCCTCACACTGTGAAATTTGTCAAATTTGGTTAAATGGTTTGTCTCCATGGAAAGAACTGTTGCGCT  
ATGGTAATGGAGTTATGTGCAACTCAGCATTTGCTGAATTTGGTATAACTTGGTTAAAGGAGGCTGGTGGGTTTTG  
GTTGGAATTACAATGTGCTGCAAAATTTGCTGGCTTCTTTTTTGTGGGTGGAGTCCAAATCCGAAACTGGTGTCTC  
AGCAGAAGCAGATGCAACATTCTGTCCAGGAGAAATCCAAGGGTGGTGGCGGTGGAGGAAAAGGCTTTTAGTTGC  
ATTCGTGTTAGTTGGAATAGTTGCATCCATTTGGTTATTTTGGCGCATGAATAAGGATAATACGATGGCAAGAGAA  
GAAATGCTTACTAATATGTGTGATGAACGAGCCCGTATGCTGCAGGATCAATTTAATGTGAGCATGAACCATGTTCT  
ATGCCCTTGGCTAATCTTGTCTCCACTTTTCCCATAGTAAACATCCTTCTGCTATTGACCAAAAACTTTTCGAGA  
ATATACAGAGAGGACGGCATTGAGAGACCCTTACTAGTGGTGTGCTTATGCTTTGAAAGTTCTCCACTCTGAC  
AGGGAGCCTTTTGAAGCTGCATGGATGGACAATTAAGAAAATGGAACAGAGAATGAGACACTTGTCCAAGATT  
GTATTCAGAAAATCTGGATCCTGCACCCATTCAAGATGAATATGCACCAAGTATATTTGCTCAAGAGACTGTTTC  
CCATATTGTATCTATTGCATGATGTGAGGAAGGAAGACCGTGAGAATATCTTGCAGCAAGGGCACTGGAAG  
GGGTCTTAATCTCTCTTTTAAAGTACTGAAGTCCAATTACCTGGGCGTTGACTCACATTTGCTGTCTATAATG  
CTAATTTTCCACCAATGTACACCAGAGCAGCGCATTGAAGCTACTGTTGGGTATTGGGTGCATCCTATGATGT  
TCCATCACTTGTGGACAAGCTCCTGCATCAACTTGCCAGCAAGCAAACCATTTGTTGTAATGTTTATGATACAA  
AATGCATCTGCACCATCACTATGTATGGCACTGATGTTGTTGATACTGGCCTACTACATATAAGCAACTTAGACT  
TTGGGGATCCTCTGCGAAAACATGAGATGCATTCAGGTTCAAGCATAGGCGCTCACTTACCTTGGACAGCAATCAG  
TGCATCAGGTGCAGTACTTGTATTACATTTGCTTCTGGTTATATCTTTTACGAGCAATTAATCGAATAGCAAGA  
GTGGAGGATGACTTTCGTGAAATGAGGGAGCTCAAGGTCGAGCTGAAGCTGCAGATGTGGCAAAGTCTCAGTTTC  
TTGCAACGGTTTCACATGAGATCAGGACTCCGATGAATGGTGTTTTAGGTATGCTGCAAAATGCTGATGGACACTGA  
ATTAGATGAAGGTGAGTGGATTATGCTCAGACTGCTCACAAGAGTGGTAAAGATCTTATATCAGTGATAAATGAG  
GTTCTTGATCAGGCCAAGATCGAGGCAGGAAAACCTTGAAGTGAAGCTGTAGAGTTTGATCCACGTTCTGTTCTGG  
ATGAAGTTCTATCACTGTTCTCTGACCAATCTAGTGAGAAAGGAATTGAGTTGGCTGTTTATGATATCCAATCAGGT  
GCCACTAGTTGTCTATTGGTGATCCTAAACGCTTCCGGCACATAATTACTAATCTTGTGGGAAATTCAGTGAAGTTC  
ACTCATAATAAAGGACACGTGTTTGTATCGGTTCTATCTTACAAATGAAGTTAAGAACCCACTACATATTATGGATG  
CAGTACTGAGACAAGGGATCCAGGATATGTCAGACAGAACATGTAATACATTAAGTGGGTTTCCAGTTGTTAATAG  
ATGGAAAAGCTGGGCAAAATTTCAAAAAATTAAGCTGTCTGAATTCATGGACGACCCCTGAAATGATTCAATTGCTA  
GTAACAGTTGAGGATCAGGTATTGGAAATTCCTTTGGATGCACAAATGCCCATATTACACCTTTTATGCAAGCTG  
ATAGTTCCACCTCAGCAACATATGGTGGGACTGGAATAGGATTGAGCATTAGTAAATGCTAGTTGATCTCATGGG  
AGGCGAAATTTGATTGTAAGTGAACCTGGAATTTGGCAGTACTTTTCTTTTACTGGAAGTTTATAGGATAGGAAAA  
ACAACATCTGATGCAAGTGGCAGCATTGCAATCCATCTATTTCTGAATTTCAAGGATTGAGAGCTTTAGTAGTGG  
ATAGAAGAAAAATCCGTGCTGAAGTTACAAGATATTATCTGCAGAGGTTGGGAATGTCTGTGGATGTTACTTTTAG  
CTTGAATCAGCATCTTGTCTATCTGTTGACTTGCATATGAGTATTTCTAGACAGTTGGCAATGATCTTTATA  
GACAAAAGATGTTCTCGATAAGGATGGCATCAAGAAACAGAGGCTGAAAAATGTCAAAGGAGACCCAAAAAATTTCC  
CAAAGATTTTTCTCTTGGTGAACCACTAAGCCCTGAAGAGCATGATGAACCTCAAACTCAAGGCGTTATTGATGA  
TGTACTAATGAAACCTCTTTGGCTTAGTGCTTTGATCCACTGTTACAGAGAAGCCCTTGAAGTGAGAGAAGCAA  
ATAGACAGAAAGAAAGAAATCGAATCTAGGAACTTGCTTAGACAGAAATCGATTTTAGTGGTTGATGATAATGCAG  
TAAACAGAAAGTAGCGGAAGGTGTTCTGCAAGATACGGGCGAAAGTTACCTGTGTGGATGGTGGCAGGGCTGC

TTTGAAAGTTCTCAAGCCACCACATGTTTTTGTATGCTTGTTCATGGATCTCCAGATGCCAGAAATGGATGGATTT  
GAAGTAAC TAGGCAAAATCCGCAGTATGAAAAAGAGGTAATGAGAAAATTGCATGTGGGGAAGCATCAGCCAAGA  
TGTTTGGTGATGTGGCTCATTGGCACATGCCGATATTGGCTATGACGGCTGATGTAAC T CAGGCCTCAAACGAGGA  
GTGCAGTCTGTGCGGGATGGATGACTATGTGTCCAAGCCATTGCAAGAAGAGCAGCTCTATATGGCTTTGTACAGT  
TGTTTCGGGTCCGGTTCTGGTTCTTACAATGCAAGTTAA

### >*C. fasciculata* HK3-CDS\_scaffold1568\_cov222

ATGAGTTTGTCTATGTAGTTGGATTTGGTCTAAAGGTAGGACATCTACTTTTGGCTTTATGTTGCTGGATTGTCT  
CTGTAAATTTACATCAACTGGTTTATAGCAACGGTGTGATTATGGACACCAAGGTTGGTTTACTTGGTGGTGGTGG  
TAGCACTAAGATGTGGCTCAATTGGTGGGAGAAGATCTCCGGCAAGAGCTGTAAATCTACCAACACTATAACGAG  
TATATTTGGGTACAAGAAAGCCAAGCGAATCGGGCAAGATGGGTTTTCATCGCATGGCTTGTGTTTTGGTTCATAG  
TATGTTTGTGCACCTTCTTCTATATGCATTTAATGCTACTGAGAGGAGGAAGGAGACTCTAGCTAGTATGTGTGA  
TGAGAGGGCTAGGATGCTACAAGATCAATTCAATGTTAGTATGAACCATATTCAAGCCATGTCAATTTCTATTGCT  
ACCTTCCACCATGCCAAATCCCCCTCTGCCATCGATCAGAGAACGTTTGTAGGTACACAGAAAGAACTCTTTTG  
AGAGGCCTCTTACAAGTGGTGTGCTTATGCTGTGAGGGTGTCTCCATTTCTGAAAGGGAGCAATTTGAGAAGCATCA  
AGGGTGGACTATTAAAGAGGATGGATAACGTTGAGCAGAACCCGGTTCATAAGGATGACTATGCTACTGAGGCCTTA  
GAGCCTTACCCGGTTCAGGAAGAATATGCTCCTGTCATTTTGGCCAGGAACTGTTGCCCATATCATTCTGTCTG  
ATGTCGTTTTCTGAAAGGAAGATCGTAAAAATGTGTTGCGAGCAAGAGAATTAGGGAAGGGGTTTTAACTGCTCC  
TTTCAGGCTTCTCAAATCAAATCGGCTCGGAGTTATACTTACATTTGCCGTCTACAAGACAGAACTTCCACCAAAT  
GCAACCCCAATGAAAGGATTCAAGCAACAGATGGGTATCTTGGAGGTGTCTTTGATGTTGAGTCATTGGTGGAAA  
AATTACTTCAACAACCTTGCCAGCAAGCAAACTGTAATTGTTAATGTGTATGATACTACTAATCACACCCGCCCAAT  
TGCCATGTATGGCTCGAATGTGACAGATGATGGGTTTCATCATATCAGCTCCCTGAATTTCCGGAGATCCTTTCAGG  
AAGCATGAGATGCAGCTGCAGGTTCAAACATAAACCTCCATGGCCATGGACGGCAATTTTACTTCAATTTGGCATCT  
TTATTTATGTCACCTTCTTCTTGGGCATATTGCCCATGCTACTTTGAATCATATTCTAAAGTAGAGGATGATTGCCA  
TGAGATGACGGAACCTAAAAAGCGAGCTGAGGCAGCTGATGTTGCAAAATCCCAGTTCTTGGCTACTGTTTCCCAT  
GAGATCAGAACACCAATGAATGGTGTCTAGGAATGTTGCATATGCTTTATGGACACAGATTTAGATGTACACAGC  
AAGAATATGTCAGAACAGCACAGGCTAGTGGAAAAGCTCTAGTGTCATTTATAAATGAGGTTTTTGGACCAGGCGAA  
GATTGAATCTGGTAAGCTAGAGCTTGAAGCAGTCCTGTTTGATTACGAGCAGTTCTAGATGATGTATTGTCTACTA  
TTCTCCGAGAAATCTCAAGAAAAAGGGGTAGAGTTGGCAGTTTATATATACAGATCAGGTTCTGAACTGCTAGTAG  
GTGATCCAGGAAGGTTTCGGCAAAATAATTACCAATCTCATGGGTAACCTCAATTAAGTTCACGGACAAAGGACACAT  
CTTTGTGACCGTCCATCTTATTGAGGAGGTTGTGAATTCATTGAAGTTGCAACGTCAAAGAATACATTGAGTGGC  
TTTCTGTGCTGTATAGACGCCAAAGTTGGAAGGGATTGAGTACTGTTAATCAAGACGAACTCTTGGTTCTTTCT  
CATCCTCTGCTCTGATCGCATTAACACTTATTGTATCTGTTGAGGATACTGGAGAAGGTATTCCTCTAGAAGCTCA  
GTCCCGTATATTCACTCCATTCATGCAGGTTGGTCCATCCATCTCCCGAAACATGGGGGAACAGGTATCGGGCTA  
AGCATTAGCAAGTGTGTTGGTTGGCCTCATGAATGGGGAATTTGATTTGTGAGCACACCCAAAAACAGGATCTACTT  
TCACATTTACTGCTGTCTTACCAATGGATGTCCCACTTCAAATGAGTGTAAAAGTCAGCAATCAACAACCAACC  
CAATCTTGATCCTCAGAAATTTAGGGCATGACTGCATTAGTTGTAGACCTTAGACCTGTCCGAGCTAAAGTGTC  
AAATATCACATTCAACGGCTTGGCATTACAGTAGAAATAGCTTCAGATTTAAATCAAGGTTTGTCTACCTTGAGCA  
ATGGGAATATAGTTATTAATATGTTGTCTGTTGAACAGAGGTTGGAAGAGGATTTATCATCTCGTCCGAGTT  
TGTCAATAACACAAGGAAAATTTGATCGTGGTGTCTCTCCTAAGTTATTCAATTTCTGCTAATTTCTAGTAGTTCCGA  
ATGAGTTCTGTAAAGTTTCGATGTTTCAATCCAAGTGTATCACAAAACCTCTGAGAGCAAGTATGCTTGTCTGCTT  
CACTGCAGCGAGCCATGGGTGTTGGAACAAGGGGAATCCTCGAAATGGGGAGCTCCCTAGTCTGTCTCTCCGCCA  
TCTTCTGCTTGGGAGGAAAATTTCTATTGTAGATGACAATAGTGTGAACCGCACAGTAGCAGCTGGTGTCTTAAAGA  
AAATATGGAGCTGATGTGTTTGTGTTAATAGTGGAAAAGAGCCATCTCGTTGGAAGCCACCCCTAGAGAAATTTG  
ACGCTGTCTTTCATGGATATTCAAATGCCAGAAATGGATGGTTTTTGAAGCTACTAGGAGGATACGTGAGATGGAACG  
CAATGTGAACAGAGTTGAATTGCTGTGGAAGCCTTAGAGAATATTACAACTGGCATGTGCCATTTTAGCCATG  
ACTGCAGACGTTATCCAGGCAACACATGAGGAATGCTTGAAGTGTGGAATGGATGGATATGTTTCAAAACCGTTTG  
AAGCTGAACAACTTTATCGGGAAGTTTACGATTTTTTCCAATCTTCT

### >*C. fasciculata* HK4-CDS\_scaffold13101\_cov108

ATGGGTATGAAGATTCGGAGCCAGCAACAGCACCACCCTGTGGCTTTGAGGTTCCATGAGCAATGGGAACTAAAA  
GAAAATACACTTTCAATCAGGCTCAAAGAAATTTGGCTTCCAAAGTATCTGCTGCTATGGATCATGGTGATGCTATT  
AATCAGTTGGTGCATCTACAAGAACATGGATGCCGGTAATAAGTCAAAGGAGAGAGACTCTGGGCAGCCTCTGT  
GAGCAAAAGGGCCAGAATGTTACAAGATCAATTCAAGTGTAGTGTAAATCATGTCCATGCCTTAGCCATTCTCGTTT  
CTACCTTCCATTACTACATAAAACCATCTGCAATTGACCAGGAAACCTTTGCCGAGTATACAGCCAGAACAGCAT  
TGAGCGGCTATTGAATGGGTTGGCCTATGCACAAAGGGTAGTTGACTCTGAAAGAGAAAACCTTTGAGAAAAA  
CATGGATGGGTAATAAAGACAATGAAAGAGAGCCTTCTCCAGTTTCAGATGAGTACGCACCAAGTATATTGCC  
AAGAACTATCTCTTATCTTGATCTATTGATATGATGTCAGGAGAGGAGGACCGAGAGAACATTTGAGGGCTAG

AGCCACTGGAAAAGCTGTTCTGACTAGCCCCCTTCAGGTTGTTGGGTTCTCATCATCTTGGTGTAAGTCTTGCATTC  
 CCTGTTTACAAGTCTAAGCTCCCTCCAAAGCCTACTGTAGAAGAACGCATTGAGGCAACTGCAGGATATGTTGGAG  
 GATCCTTTGATGTTGAGTCCCTTGTGGAGAATTTACTTGGGCAGCTTGTGGTAACCAAGCAATTTTGGTGATGT  
 GTATGATATCACAACTCTACTAACCCCTCTAATCATGTATGGTAACCAATATGAAGAAGGTGATAAGTCTCTTGT  
 CATGTAAGCAAGCTTGATTTTGGAGATCCTTACAGGAAACATCAGATGATTTGTAGGTATCATCAGAAGGCACCAA  
 CACCTTGGACAGGACTTATCCTACCATTCCTGTTCTTTGTGATTGGTTTATTAGTGGGATACATTTTATATGGTGC  
 TGCAATTCATATTGTCAAAGTTGAAGATGATTTCCATGAAATGGAGAGACTAAAAGTTTCGAGCAGAAGCTGCTGAT  
 GTTGCCAAATCCCAGTTTTTGGCAACTGTCTCTCATGAAATTAGAACACCAATGAATGGAATCCTAGGAATGCTCG  
 CTCTGCTTCTAGATACAGAACTGAGCTCAACTCAAAGAGATTATGGTCAGACTGCCCAAGCATGTGGAAGGCACT  
 GATAGCTTTAATAAATGAGGTGCTTGACCGAGCTAAAATTGAAGCTGGCAAGTTAGAGCTTGAAGCAGTTCCATTT  
 GACATTCGTTCCATACTTGATGATGTCTTTTATTGTTTTCTGAGAAGTCAAGACACAAAGGTTTAGAGTTGGCAG  
 TCTTTGTTTTCTGATAAAGTTCCAGATATTGTTATGGGGATCCTGGGAGATTCAGACAAATAATAACAAATCTTGT  
 TGGCAACTCTGTTAAATTCAGTGAACGAGGACATATACTTGTAAAGGTCATTTAGCTGAAAAATAGAAATGCTGCA  
 ATGAATGGAAAATCTGAGACTTATCTAAATGGGGGATCAGACGAAGTTATGCATATGTCTGGTGGTTATCAGTTCA  
 AAACCCTAAGTGGATGTGAAGCAGCTGATGAACGGAACAGCTGGGATAATTTTAAAGCATTTAATTCAGATGAAGA  
 ATTTTGCTTTGATGCTTCAAGTAAGAAGGTGACCACTGTTCAAGCTTCAGAGCTAGTCACTTTGATGGTCTGTGTG  
 GAGGATACTGGAATTGGAATTCCTTTTTCTGCCAGGATAGGATTTTCATGCCCTTTGTGCAAGCAGACAGCTCAA  
 CCTCTCGAAACTATGGTGGAACTGGTATTGGCTTGAGCATCAGTAAGTGTCTGGTTGAACTAATGGGTGGTCAAAT  
 AAATTTTATAAGCCGCCCCAGGTTGGGAGCACATTTTCCATCAGAGCAGTTTTTGGGACATATAAGAAAAGTTCA  
 AACAGCGATGTGAAGAACTTAATGTGGAAGATCTACCTTGTAGTTTTTAGAGGACTGAAAGCCATAGTGGTTGATG  
 GAAAACCTGTTAGAGCTGCTGTGACTAGATACCATCTGAAGAGACTAGGAATACTAGTTAAAGTTGCAAATAACAT  
 CAACAAGGCTGTTGCTTTTATGTGGAAAAATGGTTCCTTGACATCGGGATTATCCAGCCTGATATAATTTTGGTT  
 GAGAAGGACTCGTGGATTGTTGGAGAGGATGGGAGCTTCAATGTATGGCAACTGGACTGGAACAGAATGGGCATA  
 CATTGAAGCTGCCTAAAATGATCCTTCTTGCCACCAATATTAGTAATGTTGTGCTTGATAAAGCAAAAGCCGAGG  
 TTTTGCTGATACTGTGATCATGAAACCTCTGAGAGCTAGTATGGTGGCTGCGTGTCTTCAGCAAGTTCTTGGGGTG  
 GGGAAGAAGAGGCAACACGGAAAAAGACATGCCCCAACGGTTCCAGTTTTCTACAGAGCCTTCTTTGTGGAAAGAAAA  
 TTCTCGTGGTCGATGACAATGTGGTAACCGGAGGGTTGCAGCAGGTGCATTAATAAATTTGGAGCTGATGTTAA  
 ATGTGTTGAAAGTGGAAGCTGCGCTTGAAATGCTTCAATTGCCCCACAGTTTTGATGCTTGCTTCATGGACATT  
 CAAATGCCTGAAATGGACGGGTTTGAAGCAACTCGTCGAATTCGGGTGATGGAAGCAAGGCAATGAGCAGACAA  
 ATGGTGGATCTGTGGAAGGAAATGAATCGAAAAGTGAGTGGCATTGTTGCAATACTAGCCATGACAGCTGATGTGAT  
 TCATGCTACCTATGAGGAGTGCCTGAAATGTGGGATGGATGGCTATGTCTCGAAACCCCTTGAGGAAGAGAATCTC  
 TATCAAGCAGTTGCAAAGTTTTTCAAATCCAAGCCTACCTCAGACTCATGA

**>C. mimosoides NIN1-CDS transcript 1109620.1**

**>C. mimosoides NIN2-CDS transcript 105950.1**

58

TTTAGTTACGGGGGACGCCGGTCTCTCCGGGGGAGAAAAATCCGGCGAGAAGAGACGGACTAAGCGGGAGAAGACTA  
TTAGCTTGCTGTCTCAGACAGTACTTTGCAGGAAGTCTTAAAGATGCTGCCAAGAGCATTGGTGTATGCTCTAC  
CACACTTAAAGGATATGCAGACAACATGGGATCACAAGGTGGCCCTCAAGGAAAAATAAGAAAGTAGGACACTCA  
TTAAGGAACTTCAGCTTGTGATTGATTCTGTACAAGGTGCTGAGGGTGCCATTCAAATAGGGTCACTTATTCAA  
GTTTTCCAGAAATTAAGCACCACAGATTTGTCCTCTGGTGTCCCCACAGTTCTCAATCTCTAGTAACAACATTAG  
TGACCATTCCAAGAAAACCATCATCATCATCAACAACATCAACATCAACCAACATCATCACTGAGATT  
ACTAATG

### >*C. mimosoides* HK2-CDS transcript\_0037839

ATGTCTGTGAATCGAAAGCTTTCGCAAAGTTCCAAGTTAGGGAAGCAAATGAGGCTCTGCATGAGTCTAATTCTC  
TCAAGAAATGCAAGAGGAAGAAGCCCTTGTGTTCTGCGTTTTTCTTGCTGTTGTTGTTGCTGCTGTTGGAACTGT  
TTGTTTTTTCATGAGTTTTTCACAGTGAATCATTGTTGATGAGGAAGGATAAAAGGTGGGATGATTATGAAGAGAGA  
TTCGAATCTTGCTTCAGCGTTACAATGTCAGCAAGAAGCAGCTTCATGCTCTGATTTCCTTTTTCTCTGTATCAG  
ATCAGATTTTATTGCAATGTATTGACGAAAGAGGAGTGAAAAAGCTTTTAAAGCAGTGAATGTTGCCAGTACTTCACA  
GTCGATATGTCCAGAGAATCGGGAGTGCCAAAAGGAGCATAACATGTGTACGGGAGAACGTAGAAGTTATCGAACAA  
TGCCCAATTTCTCGATGATTATCTCCAGAGAAATCTGGAATCATCACTGATAGTGAAAAAACTGTATCATTGGCCT  
CACACACTGTGAAATTTGTCAAATTTGGTTAATGGTTTGCTCCATGGAAGAAGCTGTTTGCGATCATGGGTAATGGA  
GTTATGTGCAACTCAGCATTGTCTGAATTTGGTCTAAGTTGGTTAAAGGATGCTGGTGGATTGTTGGTTGAATTACA  
ATGTGCTGCAAAATTTGCTGGCTTCTTTTATGTTGGGTGGAGTCCAAATCCGAAACTGGTTGCTCAGCAGAAGCAGA  
TGCAACATTTCTGTCCAGGAGAAATCCAAAGGGTGGTGGGAGTGGAGGAAAGGTTTTAGTTACATTCTGTGTTAGT  
TGGAATAGTTGCATCCATTGTTTATTTTGGCGCATGAATAAGGATAATAGGATGGCAAGAGAGAAATGCTTACT  
AATATGTGTGATGAACGAGCCCGTATGCTGCAGGATCAATTTAATGTGAGCATGAACCATGTTTCATGCCTTGGCTA  
ATCTTGTCTCCACTTTTACCACAGTAACAATCCTTCTGCTATTGACCAGAAAATTTCCGAGAAATATACAGAGAG  
ACGGCATTTGAGAGACCCTTACTAGTGGTGTGCTTATGCTTTGAAAGTTCTCCACTCTGACAGGGAGCATTTT  
GAAAAGCTGCATGGGTGGACAATTAAGAAAAATGGAAACAGAGAACGAGACACTTGTCCAAGATTGTATTCCAGAAA  
ATCTGGATCCTGCACCCATTCAAGATGAATATGCACCAGTTATATTTGCTCAAGAGACTGTTTCCCATATTGTATC  
TATTGACATGATGTCCAGGAAGGAAGACCGTGAGAATATCTTGCAGAGCAAGGGCAACTGGAAGGGGGTCTTAAGT  
TCTCCTTTTAACTACTGAAGTCCAATCACCTGGGCGTTGTACTCACATTTGCTGCTATAATGCTGATTTTCCAC  
CAAATGCTACACCAGAGCAGCGCATTTGAAGCTACTGTTGGGTATTTGGGTGATGTTGCTGATGTTCTGCATCATTGT  
GGACAAGCTCCTGCATCAACTTGCCAGCAAGCAAACCATTTGTTGAAATGTTTATGATACAACATATGCATCTGCA  
CCCATCACTATGTATGGCACTGATGTTGTTGATACTGGCCTACTACATATAAGCAACTTAGACTTTGGGGATCCTC  
TGCGAAAACATGAATGCATTGCAGGTTCAAGCATAGGCGTCCATTACCTTGGACAGCAATCAGTGCATCAGGTGC  
AGTACTTATTATTACATTGCTTCTAGGTTATATCTTTTACGCAGCAATTAATCGAATAGCAAGAGTGGAGGATGAC  
TATCGTGAAATGAGGGAGCTCAAGGTCCGAGCTGAAGCTGCAGATGTGGCAAGTCTCAGTTTCTTGCACCGGTTT  
CACATGAGATCAGGACTCCGATGAATGGTGTGTTAGGTATGCTGCAATGCTGATGGACACTGAATTAGATGAAGT  
CCAGATGGATTATGCTCAGACTGCTCACAAGAGTGGTAAAGATCTTATATCAGTGATAAATGAGGTTCTTGATCAG  
GCCAAGATCGAGGAGGAAACTTGAAGCTCGAAGCTGTAGAGTTGATCCACGTTCTGTTCTGGATGAAGTTCTAT  
CACAGTTCTCTGACCAACTAGTGAGAAAGGAATTGAGTTGGCTGTTTATGTATCCAATCAGGTGCCACTAGTTGT  
CATTGGTGATCCTTAACGCTTCCGGCACATAAATTAATACTTGTGGGAAATTCAGTGAAGTTCACTCATAATAAG  
GGACACGTGTTGTATCGGTTTCATCTTACGAATGAAGTTAAGAGCCCACTTCATAATATGGATGCAGTACTGAGAC  
AAGGGATCCAGGATATGTCCAGACAGAATATGAATACATTAAGTGGGTTTCCGTTGTTAATAGATGGAAGCTG  
GGCAAAATTTCAAAAATTAAGCTGCTGAATTCATGGACGACCTGAAATGATTCAATTGCTAGTAACAGTTGAG  
GATACAGGTATTGGAATTCCTTTAGATGCACAATGCCGCATATTTACACCTTTTATGCAAGCTGATAGTTCCACCT  
CACGACATATGGCGGACTGGAATAGGATTGAGCATTAGTAATGCCAGTTGATCTCATGGAGGCGAAATTTGG  
ATTTGTAAGTGAACCTGGAATTTGGCAGTACTTTTTCTTTTACTGGAACCTTTAGGATAGGAAAAACAACATCTGAT  
GCAAAGTGGCAGCATTGCAATCCATCTATTTCTGAATTTCAAGGATTGAGAGCTTTAGTAGTGGATAGAAGAAAA  
TCCGTGCTGAAGTTACAAGATATTATCTGCAGAGGTAGGAATGCTGTGGATGTTACTTTTAGCCTGGAATCAGC  
ATACTCTTGTCTATCTGGTACTTGCAATATGAGACAGTTGGCAATGATTCTTATAGACAAGGATGTTCTGGATAAG  
GATGCCATCAAGAAAGAGGCCGAAAAATGTCAAGAGGAGACCGAAAAATTTCCAAAGATTTTCTCTGGTGA  
ACCACCTAAGCCCTGAAGAGCATGATGAATCAACTCAAACTCAAGGCGTTATTGATAATGTACTAATGAAACCTCTTTG  
GCTTAGTGCTTTGATCCACTGTTATAGAGAAGCCCTTGGAAAGTGAAGAAGCAATAGACAGAAAGAAAGAAATCG  
AATCTAGGAAACTTGTCTAGACAGAAATCGATTTTAGTGGTTGATGATAATGCAGTAAACAGAAAGTAGCGGAAG  
GTGTTCTGCAAGATACGGGGCGAAAGTTACCTGTGTGGATGGTGGCAGGGTGCTTTAAAAGTTCTCAAGCCACC  
ACATGTTTTTGTGCTTTTTCATGGATCTCCAGATGCCAGAAATGGATGGATTGAGTTAGGTAACATAGGCAAAATCCGC  
TGATGTGAAAAAGAGTAAATGAGAAAAATGCGATGTGGGAAGCATCGCGAAGATTGTTGGTATGTTGTTCTATT  
GGCAGATGCCGATATTGGCTATGACGGCTGATGTAACCTCAGGCTTCAACGAGGAGTGCAGTCTGTGCGGGATGGA  
TGACTATGTGTCCAGCCATTTGAAGAAGAGCAGCTCTATATGGCTTTGTCACGTTGTTTCCGGTCCGGTCTGGT  
TCTTACAATTCAAGTTAA

ATGAGTTTGGCTCTATGTAGTTGGATTGGTCTAAAGGTAGGCATCTACTTTTGGCTTTATGTTGCTGGATTGTCT  
CTGTAATTTACATCAACTGGTTTAGCAACGGTGTGATTATGGACACCAAGGTTGGTTTACTTGGTGGTGGTGGTGG  
TGGTAGTAGTAGCAAGATTTTGGCTCAATTTGGTGGGAGAAGATCTCCGGCAAGAGCGGTAAAACTCTACCAACATAC  
AACGAGTATATTTGGGTACAAGAAAGCGAAGCGAACTCGCGGAGATGGATTTTCAATGGCATGGGTTGGGTTTGGT  
TCATAATAGCTTTGATTATCTTCTCTTCTTATTTTACTCTTATGCTGCTGAGAGGAGGAAGACTTTAGCTAGTAT  
GTGTGATGAGAGGGCTAGGATGCTTCAAGATCAGTTCAATGTTAGTATGAACCATATTTCAAGCCATGTCAATTCTC  
ATTGCAACCTTCCACCATGCCAAGTCCCCCTCTGCCATCGATCAGAGAACGTTTGCTAGGTACACAGAAAGAACTT  
CTTTTGAGAGGCCTCTTACAAGTGGTGTGTTTCTTATGCTGAGGGTGCTCCATTCTGAAAGGGAGCAATTTTGAGAA  
CCAACAGGGTGGACTATTAAGAGGATGGATAAGTCTGAGCAGAACCCTGGTTCATAAGAATGACTATGCTACCGAG  
GCCTTAGACCTTACCCTGTCAGGAAGAATATGCTCCTGTCAATTTTGCCACGGAAGTGTGCCCATATCATTTT  
CTGTGATGTGCTTTCTGGAAGGAAGATCGTGAAAATGTGTTGCGCGCAGAGAATTAGGGAAAGGGGTTTTAAC  
TGCTCCTTTTCAAGCTTCTCAAAATCAAAATCGGCTCGGAGTTTACTTTACATTTGCCGCTCTACAAGACAGAACTTCCA  
CCAAATGCAACCCCAAATGAAAGGATTCAAGCAACAGATGGGTATCTTGGAGGCGTCTTTGATGTTGAGTCAATTGG  
TGGAAAAATTTACTTCAACAACCTTCCGACAGCAAACTGTAATTTGTAAATGTGATGCTACTACTAATCACACCCG  
CCGATTTGCCATGTATGGCTCGAATGTGCAGATGATGGGTTTCATCATATGACTGCTGAACTTCCGAGATCCTT  
TTCAGGAAGCATGAGATGCACTGAGGTTCAAACATAAACCTCCATGGCCATGGACGGCAATTTTTACTTCAATTG  
GCATCTTTATTATTGCACTCCTTCTTGGGCATATTGCTCATGCTACTTTAAATCATATTGCTAAAGTAGAGGATGA  
TTGCCATGAGATGAAGGAACTTAAAAAGCGAGCAGAGGCGAGTGTGTTGCAAGTCCCAGTTCTTGCTACTGTT  
TCCCATGAGATCAGAACCAATGAATGGTGTCTAGGAATGTTGCATATGCTTATGGACAGAGATTTAGATGTCA  
CACAGCAAGAATATGTGAGAAGCGCAGGGTAGTGAGAAAGCTCTTGTTGCTACTTATAAATGAGGTTTGGAGCA  
GGCGAAGATTGAATCGGGTAAGCTAGAGCTTGAAGCAGTCTGTTTGATTATCAGAGCAGTTCTAGATGATGATTG  
TCACTATTCTCCGAGAAATCTCAAGAAAAAGGGGTAGAGTTGGCAGTTTATATATCAGATCAGGTTCTCTGAATTGC  
TAGTAGGTGATCCAGGAAGGTTTTCGCAAAATTAATTCCAATCTCATGGGTAAGTCCATTAAAGTTACAGACAAAGG  
ACACATTTTTGTGACCGTCCATCTTATTGAGGAGGTTGTCAATTTCAATTTGAAGTTGCAACGTCAAAGAATACCTTG  
ATGGCTTTTCTGCTGTATAGCGCAAAGTTTGAAGGATTCAAGTGTAAATCAGAGCAACCTCTTGTT  
CTTTCTCATCTCTGCTGATGATCGCATCACTTATTTGATCTGTGAGGATACTGGAGAAGGATTTCTCTAGA  
AGCCAGTCCCGTATATTACTCCATTTCATGCAGGTTGGTCCATCCATCTCCCGAAACATGGGGGAACAGGTATC  
GGGCTAAGCATAGCAAGTGTGTTGGTTGGCTCATGATGGGGAATTTGGATTTGTGAGCACACCCAAAATAGGA  
CTACTTTTCACTTACTGTCTTCACTTCAATGGATGTCCCCTTCAATGAGTGTAAAAGTCAGCAAAATCAACA  
CCAACCCCAATCTTGATCTCTCAGAAATTTAGGGCATGACTGTCATTGTTGTAGACCTAGACAGTCCGAGCTAAA  
GTGTCAGATATCACATTCACAGCGTGGTATTCACGTAGAATAGTCTCAGATTTAAATCAGGTTTGTCTACCA  
TGAGCAATGGGAATATAGTTATTAATATGGTTCTGGTTGAACAGGAGGTCTGGGATAGAGATTCTATCATCTCGT  
CCAGTTTGTCAATAACACAAAGGAAAATGATGGTGGTGTCTCTCCTAAGTTATTATCTTGTCAATTCTAGTAGT  
TTCCGGAATGAGTTCTGTAACGTTTCAATGTTTCAACATCCAAGTGTATCATCAAAAACCTCTGAGAGCTAGTATGCTT  
CTGCTTCACTGCAGCGACCATGGGTGTGGAACAAGGGGAATCTCGAATGGGAGCTCCCTAGTCTGTCTCT  
CCGCATCTCTCTGCTTGGGAGGAAAAATCTCATCTGATAGACAAATGATGTGAACCGCAGTAGCAGCTGGTGCT  
TTAAGAAAAATATGGAGCTGATGTGGTTTGTGTTAATAGTGGAAAAGAAGCCATCTCGTTGCTGAAGCCACCCCATG  
AGTTTGACGCTGTTTCATGGATATTCAAATGCCAGAAATGATGGTTTGAAGCTACTAGGAGGATTTCGTGAGAT  
GGAACGCAACGTCAGACAGATTGAATTGCTGTGGAGGCTTAGAGAATTTACAAACTGGCATGTGCCATTTTA  
GCCATGACTGCAGATGTTATCCAGGCAACCCATGAGGAATGCTTGAAGTGTGAATGGATGGATATGTTTCAAAAC  
CGTTTGAAGCTGAACAACTTTATCCGGAAGTTTCAAGGTTTTTCCAATCTCTTAG

ATGATGGGTATGAAGATTAAAGGCCAGCAACAGCACCACCGTGGCTTTGAGGTTCCATGACCAAATGGGAATA  
AAAGAAAATACACTTTTCATTAGGCTCAAAGAAATTGGCTTCCAAGTATCTGCTGCTATGGATCATGGTGATGCT  
ATTAATCAGTTGGTGTCATCTACAAGAACATGGATGCCGTAATAAAGTCAAAAGGAGAGACTCTGGGCAGCCCT  
TGTGACGAGAGGCCAGAATGTTACAAGATCAATTCAGTGTAGTGTAAATCATGCTTCCATGCCCTAGCCATTCTCG  
TTTCTACCTTCCATTATACATAAACCCATCTGCAATTGACAGGAAACCTTTGCCGAGTATACAGCCAGAACAGC  
ATTTGAGCGGCCATTACTGAATGGGGTGGCCTATGCACAAGGGTAGTTGACACTGAAAGAGAAAACTTTGAGAAA  
CAGCATGGATGGGTAATAAAGACAATTGAAAGAGAGCGCTTCTCCAGTTCGAGATGAGTACGCACACAGTGATATTTG  
CCCAAGAAACTATCTCTTATCTTGGATCTATTGATATGATGTCTCAGGAGAGGAGCCAGAGAACACTTTGAGGCC  
TAGAGCACTGGAAAGCTGTTCTGACTAGCCCTTCAGGTTGTTGGGTTCTCATCATCTTGGTGATGTTGTGACA  
TTCCCTGTTTACAAGTCTAAGCTCCCTCCAAGCCTACTGTAGAAGAACGCATTGAGGCAACTCGAGGATATGTTG  
GAGGATCCTTTGATGTTGAGTCCCTTGTGGAGAATTTACTTGGGCAGCTTGTGGTAACCAAGCAATTTTGGTGCA  
TGTGTATGATATCACAACTCTACTAACCCCTCTAATCATGTATGGTAACCAATATGAAGAAGTGATAGTCTCTT  
GTACATGTAAGCAAGCTGATTTGGAGATCCTTACAGGAAACATCAGATGATTTGAGGTATCATCAGAAGGCC  
CAACACTTTGAGCAGGACTTATCTCATCATTCTGTCTTTGTGATTTGGTTTATTAGTGGGATACATTTTATATG

TGCTGCAATTCATATTGTCAAAGTTGAAGATGATTTCCATGAAATGGAGAGACTAAAAGTTCGAGCAGAAGCTGCT  
GATGTTGCCAAATCCCAGTTTTTGGCAACTGTCTCTCATGAAATTAGAACCACCAATGAATGGAATCCTAGGAATGC  
TCGCTCTGCTTCTAGATACAGAAGTGAAGTCAACTCAAAGAGATTATGGTCAGACTGCCCAAGCATGTGGAAGGC  
ACTGATAGCTTTAATAAATGAGGTGCTTGACCGAGCTAAAATTGAAGCTGGCAAGTTAGAGCTTGAAGCAGTTCCA  
TTTGACATTCGTTCCATACTTGATGATGTCTTTTCATTGTTTTCTGAGAAGTCAAGACACAAAGGTTTAGAGTTGG  
CAGTCTACGTTTCTGATAAAGTCCAGATATTGTTATGGGGGATCCTGGAAGATTGAGACAAATAATAACAAATCT  
TGTTGGCAACTCTGTAAATTCACCTGAACGAGGACATATACTTGTAAAGTCCATTTAGCTGAAAATAGAAATGCT  
GCAATGAATGGAAAATCTGAGACTTATCTAAATGGGGGATCAGACGAAGTTATGCATATGTCTGGTGGTTATCAGT  
TCAAACCCCTAAGTGGATGTGAAGCAGCTGATGAACGGAACAGCTGGGATAATTTTAAAGCATCTAATTGCAGATGA  
AGAAATTTGCTTTGATGCTTCAAGTAAGAAAGGTGACCACTGTTCAAGCTTCAGAGCTAGTCACTTTGATGGTCTGT  
GTGGAGGATACTGGAATTGGAATTCCTTTTTCTGCCAGGATAGGATTTTCATGCCCTTTGTGCAAGCAGACAGCT  
CAACTTCTCGAAACTATGGTGGAACTGGTATTGGCTTGAGCATCAGTAAGTGTCTGGTTGAACCTAATGGGTGGTCA  
AATAAACTTTATAAGCCGGCCCCAGGTTGGAAGCACATTTTCCCTTCACAGCAATTTTGGGACATATAAGAAAAGT  
TCCAACAGCGATGTGAAGAACTTAATGTGGAAGATCTACCTGGTAGTTTTAGAGGACTGAAAGCCATAGTGGTTG  
ATGGAACACCTGTTAGAGCTGCTGTGACTAGATACCATCTGAAGAGACTAGGAATACTAGTTCAAGTTCAGAAATAA  
CATCAACAAGGCTGTTGCTTTATGTGGAAAAATGGTTCCTTGACATCGGGATTATCCAGCCTGATATAATTTTG  
GTTGAGAAGGACTCGTGGATTGTTGGAGAGGATGGGAGCTTCAATGTATGGCAACTGGACTGGAAACAGAATGGCC  
ATACATTGAAGCTGCCTAAATGATCCTTCTTGCCACCAATATTAGTAATGTTGTGCTTGATAAAGCAAAGGCCG  
AGGTTTTGCTGATACTGTGATCATGAAACCTCTGAGAGCTAGTATGGTGGCTGCATGTCTTCAGCAAGTCTTTGGG  
GTGGGGAAGAGAGGCAACACGGAAGACATGCCCAACGGTTCAGTTTCTACAGAGCCTTCTTTGTGGAAGA  
AAATCTAGTGGTCGATGACAATGTGGTAAACCGGAGGGTTCGAGCAGGTGCATTAAAAAGTTTGGAGCTGATGT  
GAAATGTGTTGAAAGTGGCAAGCTGCGCTTGAATGCTTCAATTGCCCCACAGTTTGTATGCTTGTCTTCAGAC  
ATTCAAATGCCTGAAATGGATGGGTTTGAAGCAACACGTCGAATTCGGGTGATGGAAGCAAGGCAATGAACAGA  
CAAATGGTGGATCTGTGGAAGAAATGAATCGAAAAGTGAAGTGGCATTGCTTATACTAGCCATGACAGCTGATGT  
GATTCATGCTACCTATGAGGAGTGCTGAAATGTGGGATGGATGGCTATGTCTCGAAACCCCTTTGAGGAAGAGAAT  
CTCTATCAAGCAGTTGCAAAGTTTTTCAAATCCAAGCCTACCTCAGACTCATGA

### >*C. mimosoides* *ARR4-CDS\_transcripts\_0028539*

ATGGACGGCTATGATCTCTCGCCGCCACACTCCGACGAGCAAGTCCATGTCTTAGCCGTCGATGATAGCCTCGTCG  
ATCGGAAGGTCATCGAGCGTCTTCTCAAGATTTTCAAGCTTGTAAGTTACTGCTGTTGATAGTGGGATAAGAGCCCT  
GCAATTTCTGGGACTTCACGAAGAGAAAAAACCCTCTGAATCCGATGAAGCCTTCGAGGGCTTAAACATAAAG  
GTGGATCTTATAATCACTGATTACTGCATGCCTGGAATGACCGGCTATGAATTGCTCAAGAAAAATCAAGGAATCGT  
CCAGTTCCGAGAAATCCCGTCTGATTATGTCTCTGAGAACGTTTTACCGGCATAGACAGATGCTTAGAAGA  
AGGAGCAGAGGAATTCATAGTGAAGCCAGTGAAGTTATCTGACGTGAAACGGTTAAGACATTACATGACGACAAGA  
GAGATTATTAAGCGGGAAGTGGTCAAGAAGACAGATCCGACGGCTGTGAGAGGATCAACAAGCGAAAGCTGCAAT  
CTCCACCGTCCGTTTCTCTCATCGTCAACCGTGCAGTTCGCCGTCGACTTCACCTCCTTGCTTGATTCCCCAAT  
CCGACGGCTTAAATGACCACCAACGATTGA

### >*C. mimosoides* *ARR5-CDS\_transcripts\_0075599*

ATGGGATTGGCTGCAGATTCCCAATTTTCATGTTCTTGCTGTGGATGATAGTCTGATTGATAGGAAGCTCATAGAGA  
AGCTGCTCAGAACTCTTCTTATCAAGTTACTACTGTGATTCTGGTAGTAAGGCTTTAGAGTTTCTGGCTTTCG  
TGAGAGTGACCAGATCGACGCAACACATCTTCTGTGTCTCCTAGCAATCATCAGGAAGTGGAGGTGAATCTTATT  
ATTACAGACTACTGTATGCCTGGCATGACAGGCTATGATTTACTCAAGAACTCAAGGTATGCTTTGTTTGGTGTCT  
CTGGTTTCCATTCTTTGTCTTCTTTCTTTTTTCATTATATCCCCACTTTTCTCGGTCTGTACCCCAAGAATA  
A

### >*C. mimosoides* *ARR8-CDS\_transcripts\_0015157*

ATGAGTGCCCGCGCAGCTTATTCGGCTCAATTTGCCGCACGTTCTCCAGGAATCCCCCTCGGGATCCGGCGAAC  
TTCATGTCTCGCCGTCGATGACAGCCACGTTGATCGGAAGGTCATCGAACGATTGCTCAAAATCTCTTCGTGCAA  
AGTGACGGTTGTGATAGTGAACAGAGCTTTGCAATATTTGGGATTGGATGGAGAAAAGAAATCCATTGGGTTT  
GATAGTGTGAAGTTAATCTGATCATGACGGATTATTCATGCCCGGGATGACAGGATATGAACCTTCTCAAGAGGA  
TCAAGGATCATCTATATTAGAGAGATCCAGTGGTATCATGTCTTCTGAGAACATCTTGACTCGAATTGATAG  
GTGTTTGGAGGAAGGAGCAGAGGAGTTTCTATTGAAGCCAGTAAAGTTATCAGATGTGAACCGCTCAGAGATTTC  
ATCATGAGCGGTGAGAAGAAGGAAAGAGAGAAGAAATCAAACAAAAGAGGCGATCCGATGATCAGTTTGTCTC  
CATCTCGCCACCTGCCTTGTCTCTAAGAGATTAAATGTGA

**>C. mimosoides ARR9-CDS\_transcripts\_0028537**

TTGAAAATTCTTGGTTTGAATGAAGATGAGCAGAGGAACGCAGATACTAATATCTCTGTTGCTTCTGAAAGCCATC  
AGGATGTAGAAGTAAACCTTATCATAACGGATTACTGCATGCCAGGAATGACAGGCTATGATCTTCTGAGAAAAAT  
TAAGGAATCTAAATCTCTCAAAGACATACCAGTAGTGATTATGTCTCAGAGAATGTCCCGTCCAGGATCAACAGA  
TGCTTAGAAGAAGGAGCAGAAGAATTTTTCTGAAACAGTTTCCAGCAGTCAGATGTAATAAGCTAAAGCCACATT  
TGATGAAATCAAAAGCTAAGGAAGAGGAGGAGGACCAATTCAATTAATAGCAAAAGAAAGGGCATGGAAGAAACCTA  
TTCACCAGATAAAACCAGAGCAAAAAATATGCAGTTAG

**>C. mimosoides IPT1-CDS\_transcripts\_0041582**

ATGAATATTATGTCTCTCTGTCTCCAAACAAGTACAGCCCCCTAATGAATTTCCAAAGTGCCCTGAACATGGAGA  
ACTTGTTCATCACCAGCGCAAAGACAAAGTGGTGGTGATCATCGGCGCCACCGGCACCGGCAATCGAAGCTGGC  
CATCGACCTGGCCACCCGGTTCCTCCCGCCGAGATCGTCAACTCCGACAAAATGCAAGTGACAAAGGCCCTCAAC  
ATCACCACCAACAAGTACCGGAAGACGAATGCCACGGCGTCCCTCACCACCTCTCGGCATCATCGAGCATCCCA  
ACTCCAATTTACCGCCGACGACTTCTCCCTTCACGCCCTCTCCGCCATTCACTCCATCGTCCACCGCGATTCCCT  
CCCCATCGTCGCGCGGCTCCAAATTCGTACATCGAAGCCCTCTCCACCACCGCTTCCACTTCGATTGCTGCTTC  
CTCTGGGTGACGTTGCTCTCCCGTCTCAATTCCCTTCTCTCCGATCGGGTCGACCGCATGATCGAATCGGGTC  
AACTCGACGAGGTCCGACCCTTCTTCGACCCGGAAGGAGATTACTCCCGAGGAATTCGACGAGCAATCGGAGTCCC  
CGAATTCATGACTATTTTCAAGCAGAGGCCTCCGGAGCCGATGAAACGACGAAGAAGAGCTTCTAGAAGGCCGCC  
ATTGCGAAGATCAAAGCCAACAATTGCTCGCTGGCGTATCGTCAGCTTCAGAAGATTTCATCGCCTCCATTCCATGT  
GGAAGCGAAACATGTACCGTCTCGACGCCACCCAAGCCTTTATCAAGTCCGGCCACGAAGCAGAGGAGGCCCTGGCA  
CAATGACGTGCTTGCCTGCAAGAGCCACCGGATTCTCCACAATTCCTCTACGACGATGCTCCCATGCTTCCGCCACCG  
CCACTCATGACGGCGCCGCAAAAACAGGTCGTAGCCGTCGACGAGCATCTCACTAG

**>C. mimosoides IPT9-CDS\_transcripts\_0073486**

ATGATTACGGTGGATTCTGCAGTGTTCTGTACATGCTTCTCGCGGTTCTTTCTTCTCGGAGAAACCTCTACTCC  
AAGCAGCTCAGCCCTCATTGCTTCCGCGCGCCGCGCGCGCTTTTCGGCCACCGTATCGTCGGTCTCCGAAAC  
TCATAAGAAGGACAAAGAGAAGGTCATTGTGATTTCTGGTCCCACTGGGTCTGGTAAAAGTCGGCTTGCATTG  
GAGCTCGCCAAGCGCTCAATGGCGAGATTGTTAGTGCGGATTCTGTCCAGGTATACCAGGGTCTTGATGTTGGAT  
CAGCCAAGCCCTCCTTAAATGATAGAAAGGAGGTACCACATCATCTGGTTGACATATTGCACCCCTTCTGAAGATTA  
TTCTGTTGGACAATTTTTGAGGATGCAAGGCAAGCTACCAGGTGTATCTTGACAATGATCGTGTTCCTATAGTT  
GTTGGGGAACTGGTTTGATTTAAGATGGTTCATTTATGGAAGCCTGATGTCCCAAGGCCCTCGCCAGAGGCTG  
CATCTGAAGCATATCTGAGTTAGCTCAATTACAGAGAATGGACAGTGGGATGAAGCATGGAGTTGGTTGTAAA  
GGCGGGTGACCCGAAGGCTTTGTTTATACTCCCAATGACTGGTATCGTTTACGACGTAGCCTTGAAATATCAAG  
TCTAGTGGATCACCTCCATCTGCTTTTCGGGTACCATACGATTCTTTCAGGGAACAGGGAGACTCTAGTGTTCAG  
AAAAATCTGACCCCTCTTATGTCAATGCCATTGATGATGGAATGAAACAACATCTCTCAGATTTGGACTATGAATT  
CATCTGTTTTTTCTTTCCAATCAAAGACTTCACCTCTATAGATCAATTGACTATCGCTGTGAAGATATGCTACTA  
GGAAGAGATGGGATATTGCTGAGGCTCAATGTCTTCTTAATATGGGTCTTCATCCAAATTCAAATTCGCAACAC  
GAGCAATTGGTTATAGACAGGCCATGGAGTACTTGCTAAAAATGCAGACAGCAAGGGGGTCAGAGTTCAGCTCGAGA  
CTTCTTCAGCTTTTATATGAATTTCAAAAAGCATCTAGGAATTTTGCCAAAAGGCAAAATGACATGGTTTCGGAAAT  
GAACATATATACCATTGGCTTGATGCTTCTAAACCTCTGGAAACTATTCTAGACTTTATCCTTAATTTCTTACCATG  
ATAACAGTGAAGTATCATTGTTCTGGAGACCTAAAAATGCCAGAGATATTCACAACAGACGAGAAGTTAATGA  
ACTAAAGTCTTACCAGCCGAAAAATAGGCATTTCTTGAGTGGAAGAAGATTGTTCCCGGTTCTGGACTGGATATGG  
AGAACTCAAAGGTGA

**>C. mimosoides IPT-CDS\_transcripts\_0074205**

ATGACCATCTCAATGGTAATGTGACAGACCACCAACGCAGCCACTAATAAACGGTGGCGCCAAAGTAAAGAACGTGT  
CGACCAGTCTTCCACGGCTGAAGGAGAAGGTGGTGGTGGTGGTGGGAGCCACCGGCACCGGAAAGTCAGGGCTATC  
GACGACCTCGCCGGCCATTTTCGGCGGAGATAATCAATTCGACAAAAATGCAAGCCTATGAAGGACTCGACATA  
GTCACGAACAAGATTCACAGAGGAAGAACAGAGTGCGGTCGCTCACCATTGCTAGGAACGCAGAAATCCGAATCCG  
AATTCACGGCGGCCAATTTCTGCGACATGGCTTCACTCGCCATTGAAACAATCACGGGTGCGGGAAGGCTCCGAT  
CATCGTCGAGGCTCGAATTCGTACATGGAGGCTTAAATCGACGACGACGACTACAAATTCGATCGAGGTACGAT  
TTCTGCTGTGTGGATCGACGTGGCGATGCCGGTGCTTCATTGGTATGTGGGAGAGAGGGTGGATCAAATGTTAG  
AGAAAGGAATGGTGAGCGAATTGAGAGAAATTTCAATCCAAAAGGGGATTACTCGCTAGGGATTAGGAAAGCGAT  
TGGATTGCCGGAATTCGATAAATACTTTAGGAGCGAAGCGAGTGTGGATGAGAAAAACAGGGGAGAGGCTGATGGAG  
GAAGCGGTGAGGGAGATGAAGGAGAATACCAGGAAATTTGGCGGGAAGCAGTTGGGGAAGATCCAGAGGCTGAGGA  
ATGTGAAGAGGTGGAAGATGCAGCGCGTGGATGCCACGCGGTTTGTAGTGCAGGTCGGGGAAATGGAGGATGATGA  
GGTGTGGAGGAAGTGGTCACGGAGCCAGTGTTATGATGGTACATCGCTTTCTTCAAGAATCACTGCTGCCAAT  
CTCAGATTGCCAGCTCGTCGTTTCAGACTACTGCCATTGCTTCTTCTTCTTCTTCTTCT

**>C. mimosoides IPT-CDS\_transcripts\_0002444**

ATGACCATCTCCACCTCTGTGGTAAATGAACAGAATAACACAACCACTCATTGATATTCCTTCGAGTGGTAGAGAAT  
TGAATCTGAGCAATCTTCCACGGCAGAAAGAGAAGGTAGTGATTGTGATGGGAGCCACCGGCACCGGAAAATCCAG  
GCTGTCCATCGACCTCGCCACCTTATTTCTGCAGAAATCATCAACTCCGACAAAATGCAAGTCTACAAAGGCCTT  
GACATTGTACCAACAAAATCACCAGGGAAGAGCAACGCGGTGTGCCTCACCATTGTCTGGCAACACAAAATCCCG  
ATTCCGAGTTTACAGCCACCAATTTCGCGGAATCGGCGTCGCTTGCAATTGACTCAATCACGAGGCGTGGAAATCT  
TCCCATCGTCGTGGGAGGATCGAACTCGTACATCGAGGCGCTAATCGACGATGACGACTATAAGTTCGGATCGAGG  
TATGATTTCTGTTGTGTGTGGGTAGACGTTTCAATGCCCGTTCTTCGCGAGTACGTGGCGGAGCGGGTAGACCAA  
TGATCGAGAATGGAATGGTAAATGAAGTGAGGCGGTTTTCAAATCCAAAGGGAGATTATTCGCGAGGGATTAGGAA  
AGCGATAGGGGTTCAGAAATTCGATGAGTATTTCCGTGCGGATGAAGAAACCAGAGGGAGGCTACTGGAAGAGGCG  
GTTAGGGAATGAAGGAGAACACGTGGAAATTGGCGCGGAAACAGTTGGGGAATTCGAGAGGCTGAGGAATGTGA  
AGAGATGGAAGATTGAGCGAGTGGATGCGACGCCTGTGTTTCGAAAGCGTGGGAAAGAGCAGAGGAGGCATGGAA  
GAAGATGGTTGCAGAGCCAAGTGCGATGATCGTCGCAAGTTTCTTTACAATTCAGCACTGTCAATGTCGACTAC  
CATCATCAACCATCGAAAGTGCCGTGCTACGGCTATGTGCTGAATTATGGTGA

**>C. mimosoides CYP735A-CDS\_transcripts\_0047612**

ATGGTGGCCGTCATCGTTCTCCTCTTCTCCATGTGTGTTGATCCTCTTAAGAGTTGCGTATGAAACGGTGTGGT  
GTTACTTTTTAAACCCAAGACGCATCCATCTGATAATGAAAAACAAGCGTGCCTGGTCCAAAACCTCGGTTTCAT  
CACCGGCAACCTCTCGACATGGCGTCCCTCTTCTCCGCCGCCACTTCCGGCGACATGTCCTCCATCTCTCACGAC  
ATTGTTTCCCGCGTCTCCTCCTCATTTCCGCCCGCTGGTCCGCCGCTACGCGCAAACGCTTCTCTTGGAAACGGCA  
CCGAGCCTCGTCTCTGTCTCACCAGAAACCCACCTCATCAAAGACTTCTCTCCAAATACGCCACCGTCTCCGGCCG  
CTCTTGGCAACACTCCAGGGTTCCAAGCATTTCATCGGCGGCGCCTTCTCATGGCCAACGGTAGCCCTCTGGTAT  
CACCAGAGACACATCGTCCAACCGGCATTATGGGAGATAAACTCAAAGGGTATGCTGGGCACATGGTGAATGCA  
CGAAGGAGATGCTAGAATCGTGGCGAATGTGGTGGAGAATGGACAGAGTGAGGTGGAGATCGGTGAATACGTCAC  
CAGACTCACTGCTGACATCATTTTCGAGAACAGAATTCCGCACCTAGTTATGAGAAGGGGAAACAGATTTCCATCTT  
CTTACCATGCTTCAGTCTCTGTGCGCTCAGGCCAGCCGTCATCTCTGTTTCTCGGTAGCCGGTATTTTCCGAGCA  
AGTACAACAGGGAGATAAAGGCGATGAAGAAGGAGGTGGAGAGATTATTAAGAGAAATAATAGAGAGCAGGAAGGA  
CTGTGTGGAATAGGGAGGAGCAATTCGTACGGGAATGATTTGCTGGGGATGCTGTTAGATGAGATGAAGAAGAAT  
AAGAATGAAAAGAGAGGGGATGGGTTTAGCTTGAATCTTCAGATGATTATGGACGAATGCAAGACGTTCTTCTTCG  
CCGGACATGAAACCACCGCCTGTTGCTGACGTGGACTGTCATGTTGCTCGCCACCAATCTCTGAATGGCAGAAATA  
GGTCAGAGAAGAAGTCACACAGGTCTGCAATGGTGAACTCCTGCACTCGATCACCTCCCCAACCTCTCTCTCTTG  
CATATGGTGATAAACGAGGCATTGAGGCTATACCCACCGGCACAGTGCTGCCAAGAATGGTATTTGAGGACATAG  
TGCTTGGTGACCTTCATATCCAAAGGGAGTCTCTGTTTGGATTCCGGTTTTAGCTATCCATCATAGTAAGTCACT  
GTGGGGTGAGGATGCCAATGACTTCAATCCAAACAGGTTTCGCTTCTAAGTCCTTACACCAGGTCGCTTTATTTCCC  
TTCCGTATTGGTCTGAGAACTGTGGGACAAAATTTTGCTATTATGGAAGCTAAGATCATTTTGGCTATGTTAA  
TCTCCAAATTCCTTACCATTTCGATAATTACCGTCATGCCCTGTCTTCTCACTATCAAGCCCAATA  
TGGAGTTCAAGTTTGTCTCAAGCCCATCATCCATACAAAGTAA

**>C. mimosoides LOG3-CDS\_transcripts\_0141097**

ATGGGGCTCGTTTTCTCAGGCAGTTCATGATGGTGGGCGCCATGTTCTAGGGGTATCCCCAGAACTCTTATGCCAA  
GAGAGATCACAGGTGATCCTATTGGCGAAGTGAGGGAAGTATCAAATATGCACCAAAGAAAGGCTGAAATGGCTCG  
CCAAGCTGATGCTTTATTGCTCTCCCTGGAGGGTATGGCACTTGGAAGAGCTGCTTGAAGTAATTACATGGGCT  
CAGCTTGGGATACATCGCAAACCCGTGGGGCTGTGTAATGTGGATGGATACTACAATTTCTTGTGTCGTTTCATTG  
ATCAGGCCGTTGATGAAGGCTTTATCTCACAACGGCAGTCGTATTATGTGCTGCACCCAGTGCACCAACAATT  
GGTCAAAGAGCTAGAGGAGCATGTACCTGAGCAAGATGAGGTGGTATCCAAGTTGGTGTGGGAAGATAGACTAAGT  
TTCGTGCTGAATCGTCAGAGGTTGCCATGTGA

**>C. mimosoides LOG-CDS\_transcripts\_0091732**

ATGGAGATGGAGAAGATTTATGAAGAGAGTGGGATGAAGAAATCGAGGTTTCAGACGCATCTGTGTGTTCTGTGGTA  
GCAGTCTTGGCAAGAATCTAGCTACCAGCTCGTGTCTATTCAACTTGCTAAACAATTCGTAGAGAGGAAGATAGA  
CTTGTGTATGGAGGAGGAAGCATTGGGTTGATGGGACTCATCTCCCAAGCAGTGTTCGATGGGGTTCGTCATGTC  
CTTGGAGTAATCCCAAAGTCCCTCATGCCAGAGAGATAACAGGGGAAACAGTAGGAGAAGTACGAGCAGTATCGG  
GTATGCACCAACGCAAAGCCGAATGGCTCGACAAGCCGATGCTTTTATCGCTTTCGACGTTGGTATATGGCACCT  
GGAGAACTATTGGAAGTCATCACTGGGCTCAACTTGAATTACAGATAAACCCGTGGGGTTCGTGACAGTATGAT  
GGGTATTACAACCTACTGCTGTCAATTCATAGATAAAGCTGTGGACGAAGGCTTCGTTACTCTGCTGCCCGTCACA  
TTATTGTCTCTGCCAAACTGCCACGACCTCATGTCCAACCTTGAGGTAATCTTAAATTCATTCTTTATGGC  
TGTATACCTGA

**>C. mimosoides LOG-CDS\_transcripts\_0005938**

ATGGAGATGGAGAAGATTTATGAAGAGAGTGGGATGAAGAAATCGAGGTTTCAGACGCATCTGTGTGTTCTGTGGTA  
GCAGTCCTGGCAAGAATCCTAGCTACCAGCTCGCTGCTATTCAACTTGCTAAACAATTCGTAGAGAGGAAGATAGA  
CTTTGGTGATGGAGGAGGAAGCATTTGGGTTGATGGGACTCATCTCCCAAGCAGTGTTTTGATGGGGTTCGTATGTC  
CTTGGAGTAATCCCAAAGTCCCTCATGCCAGAGAGATAACAGGGGAAACAGTAGGAGAAGTACGAGCAGTATCGG  
GTATGCACCAACGCAAAGCCGAAATGGCTCGACAAGCCGATGCTTTTATCGCCTTGCCAGGTGGTTATGGCACCCCT  
GGAGAACTATTGGAAGTCATCACCTGGGCTCAACTTGAATTACGATAAAACCCGTGGGGTTGCTGAACGTAGAT  
GGGTATTACAACCTCACTGCTGTCTATTATAGATAAAGCTGTGGACGAAGGCTTCGTACTCTGCTGCCCGTCACA  
TTATTGTCCTGCCCCAAACTGCCACGACCTCATGTCCAACCTTGAGGAATACGTGCCGGAGCATTGTGGTGTGGC  
TCCAAAGCTAAGTTGGGAGATGGAACAACAATTGGTTTCAGTCACTACTGCCAAGTCTGATATTTACCGTTGA

**>C. mimosoides LOG-CDS\_transcripts\_0021288**

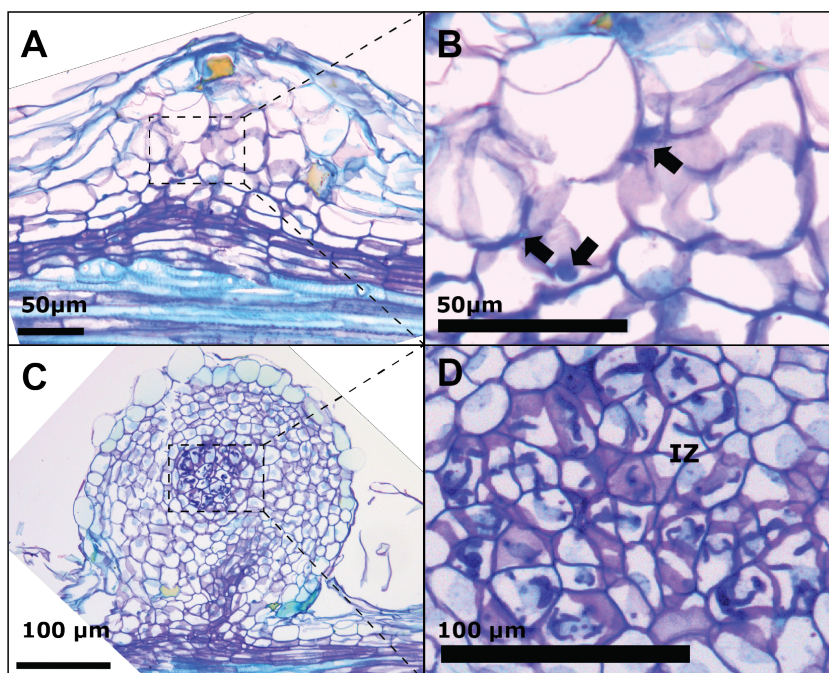
ATGGAGAGAGAAAATCATGAGATGAAACAACAATCAAAATTCAGAGGATTTGCGTCTTCTGTGGGAGTAGTCCTG  
GCAACAAAAGTAGCTATAAGGATGCTGCTATTGAGCTTGGAAGAGAATTGGTGTCTAGAAAATATTGATCTGGTTTA  
TGGAGAGGAGCAGCATTTGGATTAATGGGGTTGATTCTCAAGCTGTGTACAATGGTGGTCGCCATGTGATAGGAGTT  
ATTCCCAAGACACTCATGCCAGAGAGATTACTGGAGAACTGTGGGAGAAGTGAAGGCAGTAGCAGACATGCATC  
AAAGAAAAGCCGAAATGGCCAGGCACTCTGACGCTTTCATTGCCTTACCCGGTGGCTATGGTACACTTGAGGAGCT  
TCTTGAAGTCATAACTTGGGCTCAACTTGGCATCCATGATAAACCTGTGGGATTGTTAAATGTGGATGGATACTAC  
AATTCTTGTCTATCCTTCATTGATAAAGCAGTGAAGAAGGCTTCATTAGCCCTAATGCTCGTCACATTATTGTCT  
CTGCCCCAAGTGCAAGGAGCTTGTCAAAAAGATGGAGGAATATTTCCACGACATGAAATAGTTGCTTCCAAGCT  
AAGCTGGGAAAGCGAGCAGATTAAGAACTCCTCAACATATGAGATCTTAAGGTGA

**>C. mimosoides LOG-CDS\_transcripts\_038413**

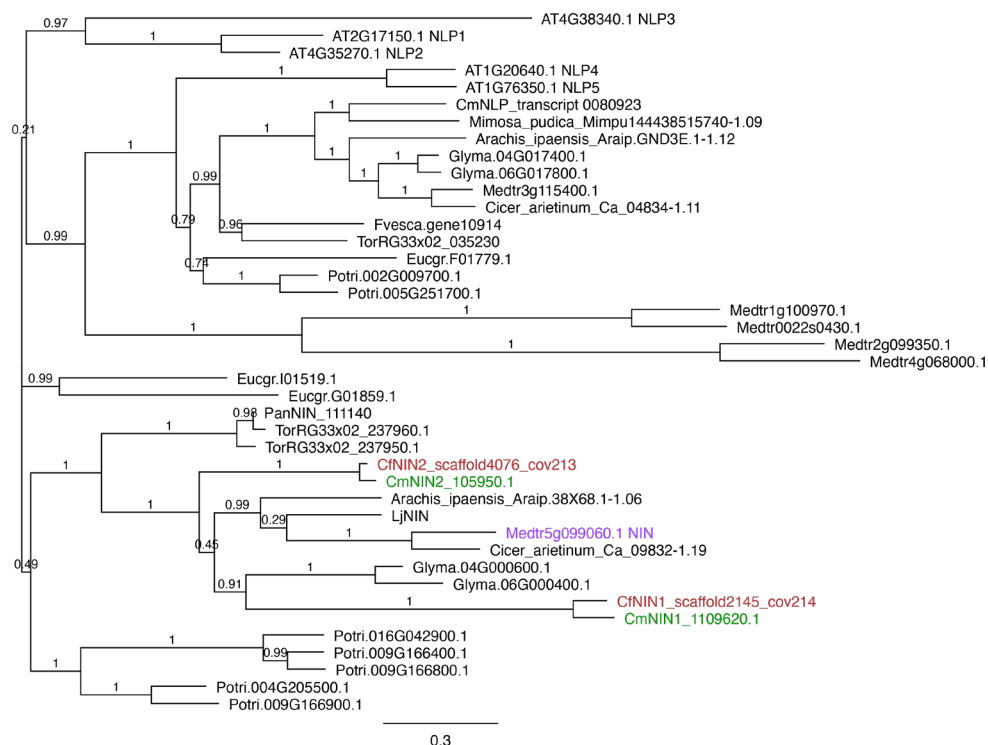
ATGGAGATAGAGAGAGATAGAGATGAAACAATCTAGATTCAAGAGGATTTGTGTGTTTTGTGGAAGTAGCCCTG  
GGAAGAAGACCGACTATAAGGATGCTGCTATTGAGCTTGGAAGAGAATTGGTGTCAAGGAATATCAATCTGGTGTA  
TGGAGGAGGTAGCATTTGGTTTGATGGGCTTGATTCTCAAGCTGTTTATAATGGTGGTCGTATGTCTATAGGGGTT  
ATTCCCAAGACGCTCATGCCAGAGAGATAACAGGAGAAACAGTAGGGGAAGTGAAGCAGTAGCAGATATGCATC  
AAAGGAAAGCTGAGATGGCTAAGCATTCTGATGCCTTTATTGCCTTACCCGGTGGGTATGGGACACTTGAGGAGCT  
TCTTGAAGTCATAACTTGGGCTCAACTAGGCATCCATGATAAACAGTGGGATTATTAACGTGGATGGCTACTAC  
AATTCTTTGTGTCTTTCATTGACAAAGCAGTAGAAGAAGGCTTCATTAGCCCCAACGCTCGCCATATAATTGTCT  
CTGCCCCAACGCCAAAAGAGCTCATAAAAAAGATGGAGGAATATACTCCAAAACATGAGAAAGTTGCGTCCAAGCT  
GAGCTGGGAAATTGAGCAGTTAATGTGA

**>C. mimosoides LOG-CDS\_transcripts\_085230**

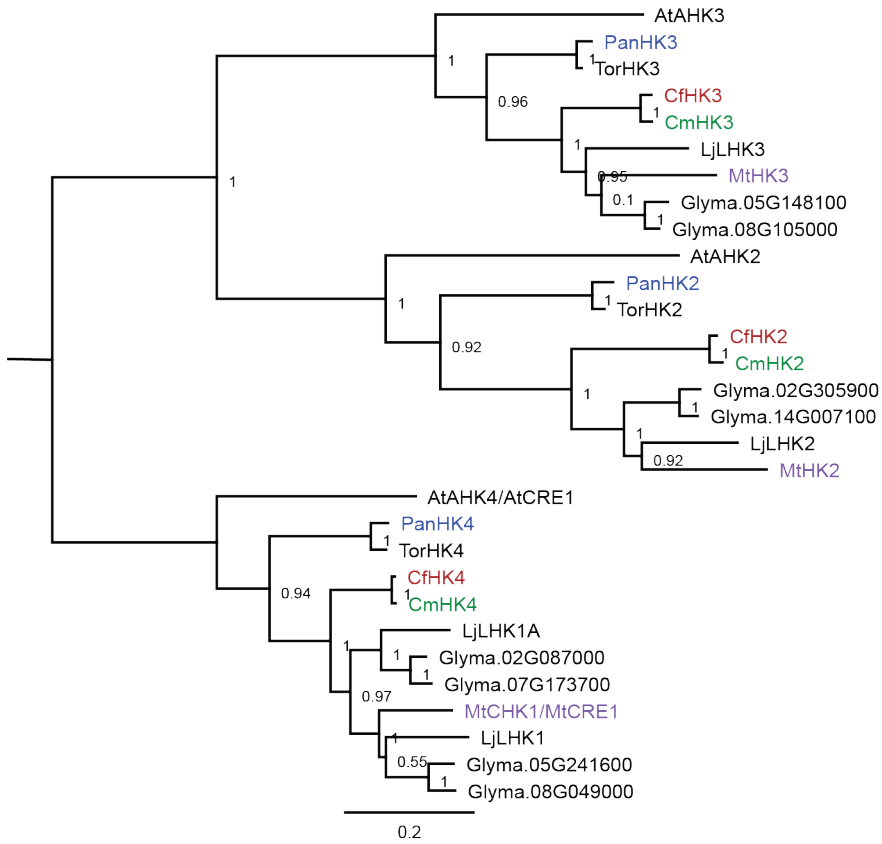
ATGGAGATAGAGAGTGAGATAAAGCAGTCAAGGTTTAAGAGGATTTGTGTGTTTTGTGGGAGTAGTCAAGGGAAGA  
AGAGTAGCTATCAAGATGCTGCCATTGAGCTTGGAAGGAATTGGTGGCAAGGAACATTGATCTAGTCTACGGAGG  
GGGAAGCATTGGTCTAATGGGTTTGGTCTCACAAGCAGTTCATGATGGCGGTTCGGCATGTCATTGGAGTTATTCCC  
AAGACGCTCATGCCTCGAGAGCTAACTGGCGAAACAGTGGGAGAGGTGAAGGCAGTGGCAGATATGCACCAACGGA  
AGGCTGAGATGGCTAAGCATTGATGCCTTTATTGCCTTACCAGGAGGATATGGGACACTAGAAGAGCTCCTCGA  
AGTCATAACCTGGGCTCAACTTGGGATTATGACAAACCGGTGGGATTATTAATGTTGATGGATACTACAATTCA  
TTGCTGTCGTTTGTGACAAAGCCGTGGAAGAGGATTTATTAATCCAATGCTCGCCACATCATTGTATCAGCAC  
CAACACCTAAGGAAGTGGTCAAGAAATTGGAGGATTACGTGCCATGTGATGAAAGTGTTGCTCCTAAGTTGAGCTG  
GCAGATTGAACAGCTGGGTTACCTCAAGAGTATGATATCTCAAGGTAG



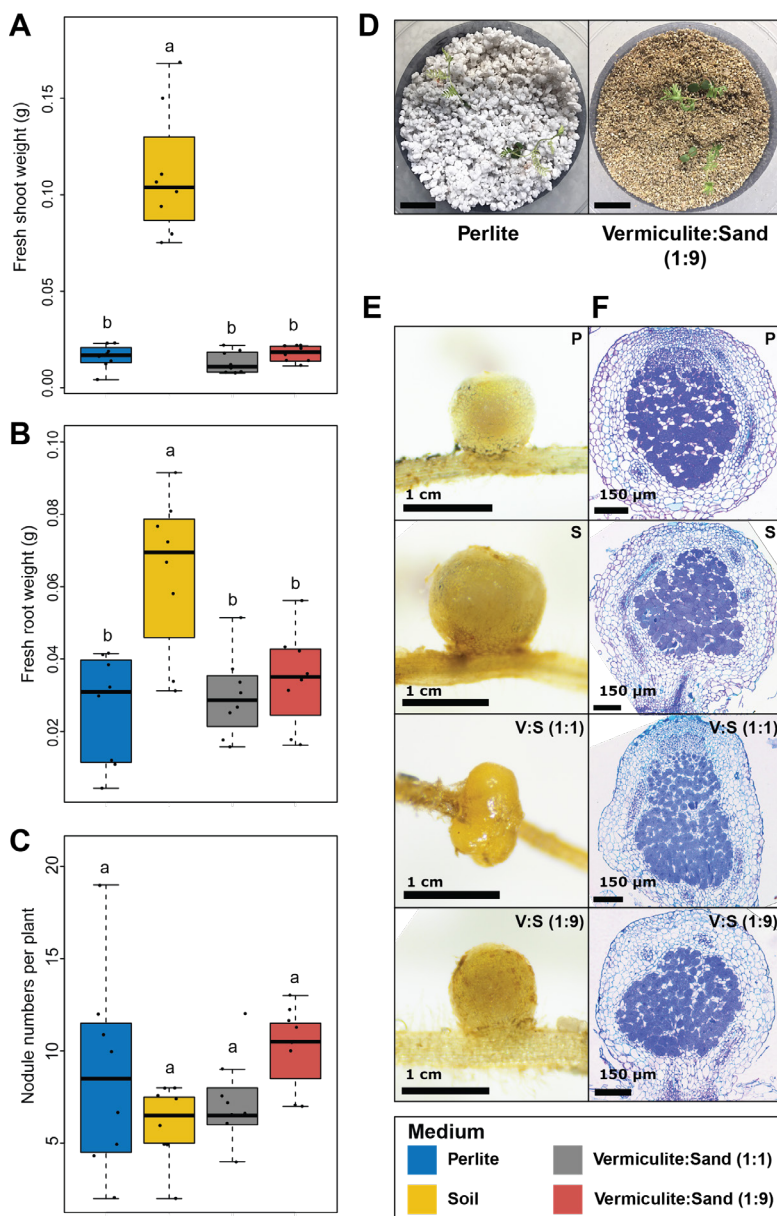
**Supplemental Figure 1. Representative micrographs of longitudinal sections of *C. mimosoides* nodules highlighting infection. (A) Nodule primordia with intercellular infection. (B) Zoom in highlighting intercellular infections and apoplastic colonies (black arrows). (C) Young nodule with intracellular infection. (D) Zoom in highlighting infection threads (IZ).**



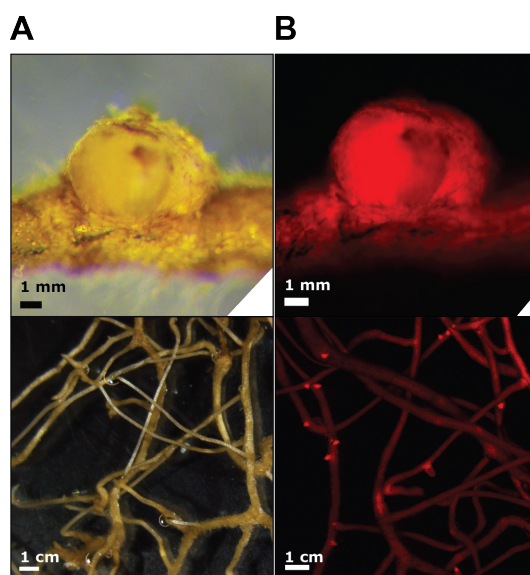
**Supplemental Figure 2. Phylogenetic analysis of NIN.** The phylogenetic tree of NIN was reconstructed using Geneious Software v 8.1 (FastTree) based on maximum-likelihood method (bootstrap analysis 1000x). Included species are; *Chamaecrista mimosoides* (Cm), *Chamaecrista fasciculata* (Cf), *Parasponia andersonii* (Pan), *Arabidopsis thaliana* (At), *Medicago truncatula* (Medtr), *Glycine max* (Glyma), *Lotus japonicus* (Lj), *Eucalyptus grandis* (Eucgr), *Populus tricarpha* (Potri), *Mimosa pudica*, *Arachis ipaensis*, *Cicer arietinum*, and *Fragaria vesca* (Fvesca). *C. mimosoides* and *C. fasciculata* NIN proteins are highlighted by green and red colour, respectively. Branch label shows percentage of consensus support.



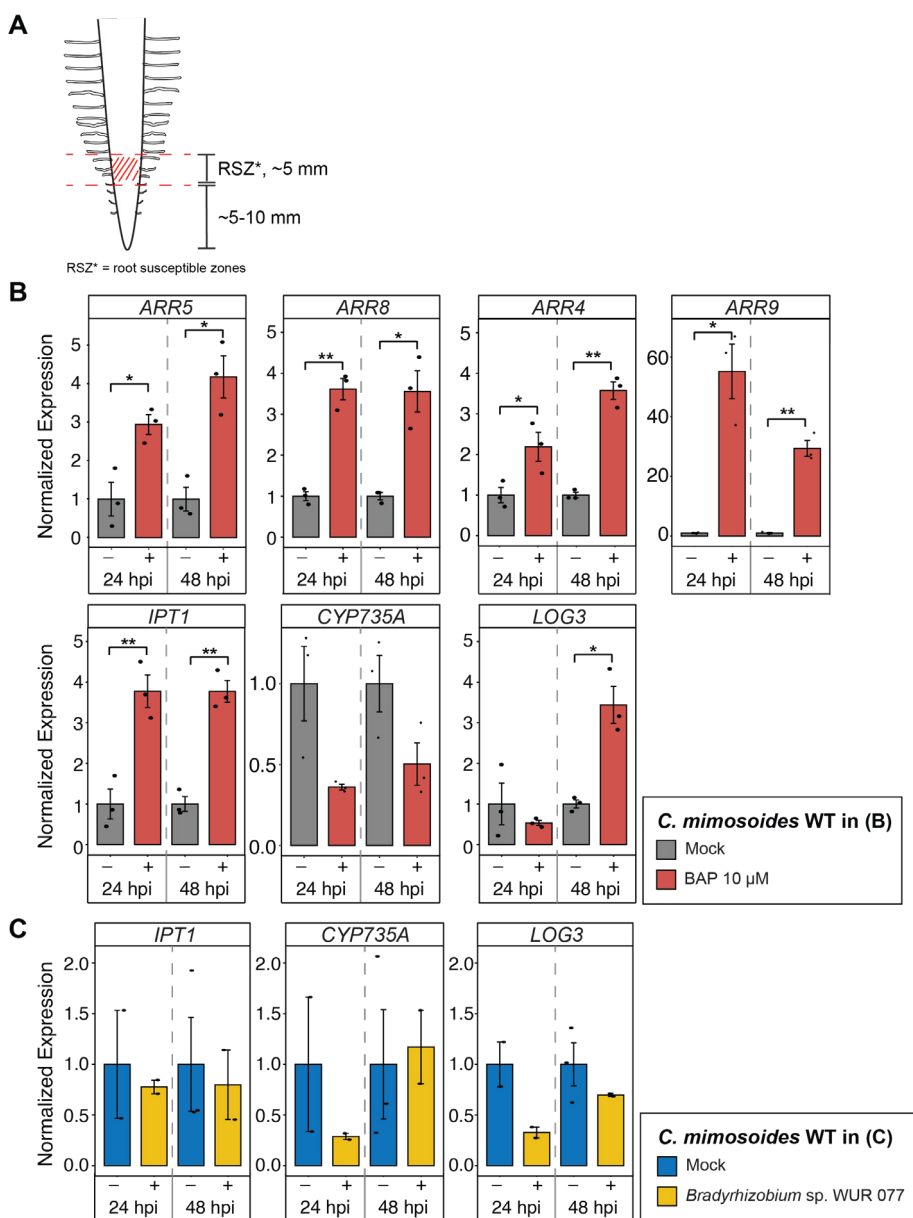
**Supplemental Figure 3. Phylogenetic analysis of the histidine kinase cytokinin receptors.** The phylogenetic tree of HKs was reconstructed using Geneious Software v 8.1 (FastTree) based on maximum-likelihood method (bootstrap analysis 1000x). Include species are; *Chamaecrista mimosoides* (Cm), *Chamaecrista fasciculata* (Cf), *Parasponia andersonii* (Pan), *Arabidopsis thaliana* (At), *Medicago truncatula* (Medtr), *Glycine max* (Glyma), *Lotus japonicus* (Lj), *Eucalyptus grandis* (Eucgr), and *Populus tricocarpa* (Potri). *C. mimosoides* and *C. fasciculata* cytokinin receptor proteins are highlighted by green and red colour, respectively. Branch label shows percentage of consensus support.



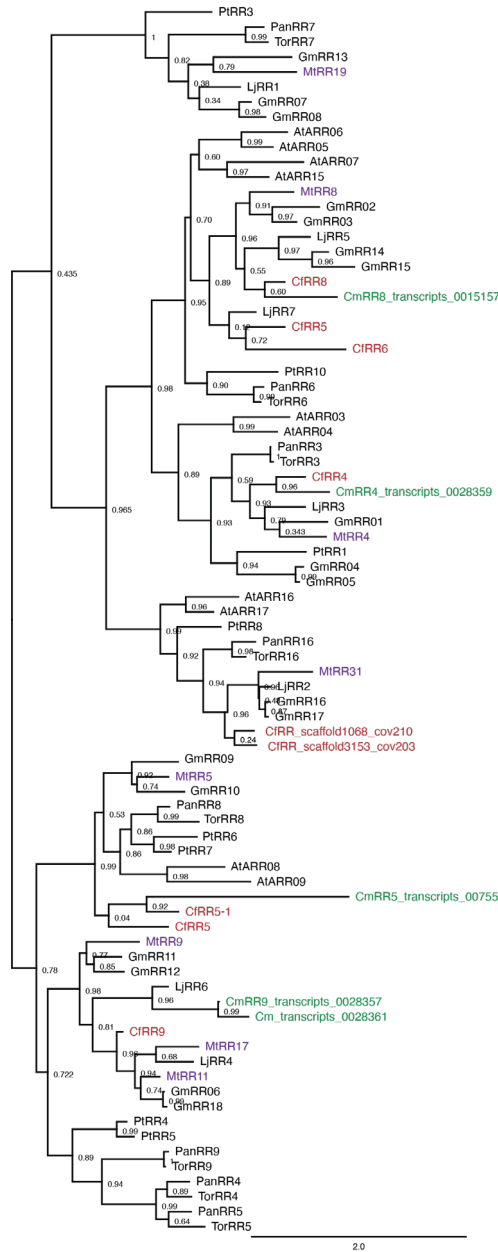
**Supplemental Figure 4. Nodulation assay of *C. mimosoides* on different growth substrates.** 14-day old seedlings grown on perlite, soil, vermiculite:sand, v:v, 1:1, or vermiculite:sand, v:v, 1:9 and inoculated with *Bradyrhizobium* sp WUR077. **(A)** fresh shoot weight, **(B)** fresh root weight and **(C)** number of nodules of plants grown 4 weeks post inoculation ( $n=5$ ,  $\pm$  SEM. Bars with different letters differ significantly ( $p < 0.05$ , One-way ANOVA and followed by post-hoc Tukey HSD). **(D)** Representative photograph of shoot quality of plants grown in perlite or vermiculite:sand (1:9) ( $n=5$ ). Representative bright-field images **(E)** and sections **(F)** of nodules grown on (P) perlite, (S) soil, (V:S (1:1)) vermiculite:sand, v:v, 1:1 or (V:S (1:1)) vermiculite:sand v:v, 1:9 ( $n=10$ ).



**Supplemental Figure 5. Representative images of *Agrobacterium rhizogenes* transformed roots and nodules in *C. mimosoides* containing the DsRED fluorescence cassette. (A) Representative bright-field and (B) the corresponding red fluorescence images. Nodules were harvested and analysed at 4 weeks post inoculation with *Bradyrhizobium* sp WUR077. (C) Transformation efficiency scored 4 weeks post inoculation.**

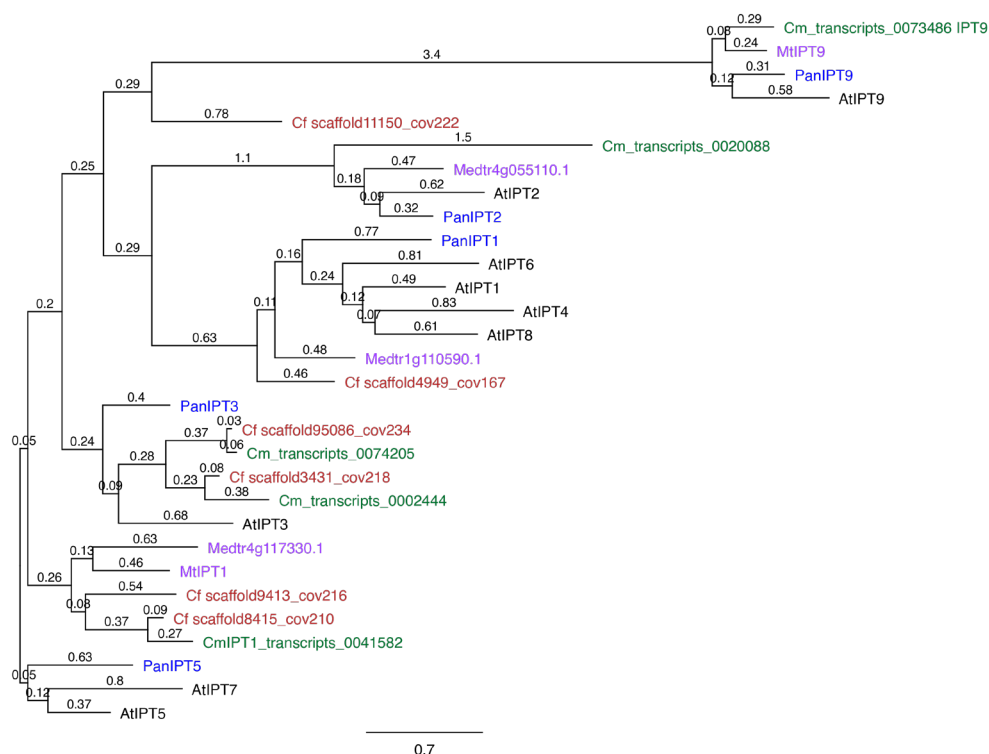


**Supplemental Figure 6. Relative expression of putative cytokinin biosynthesis and signalling genes upon treatment with Benzylaminopurine (BAP) or *Bradyrhizobium* sp WUR077. (A)** schematic representation of the root susceptible zone which was treated and harvested for gene expression analysis. Normalized gene expression of samples treated for 24 or 48 hours with (+) or without (-, mock treatment) **(B)** Benzylaminopurine (BAP) or **(C)** *Bradyrhizobium* sp WUR077. Genes analysed; *ARR4,5,8,9*, (Putative Type A Cytokinin Response Regulators), *IPT1*, *CYP735a*, *LOG3* (Putative Cytokinin biosynthesis).  $n=2-3$ , bars represent averages  $\pm$  SEM, Significance:  $p<0.01$  '\*\*\*',  $p<0.05$  '\*\*', student *t*-test.

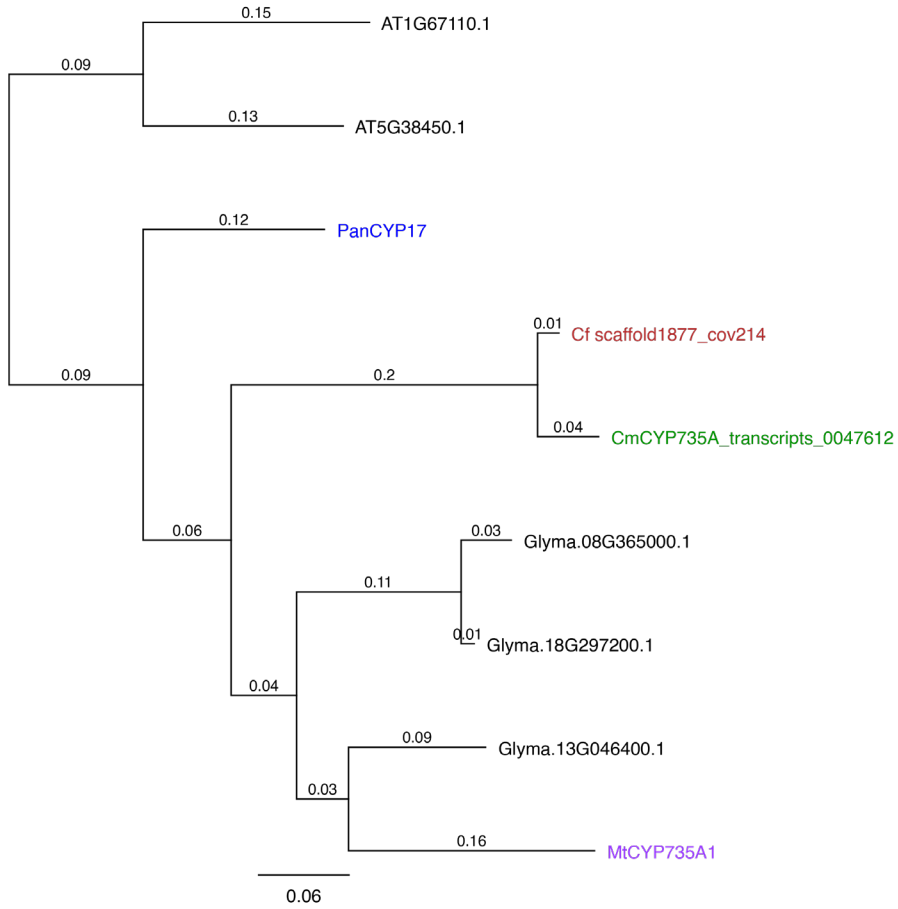


**Supplemental Figure 7. Phylogenetic analysis of Type A Cytokinin Response Regulators (ARR).**

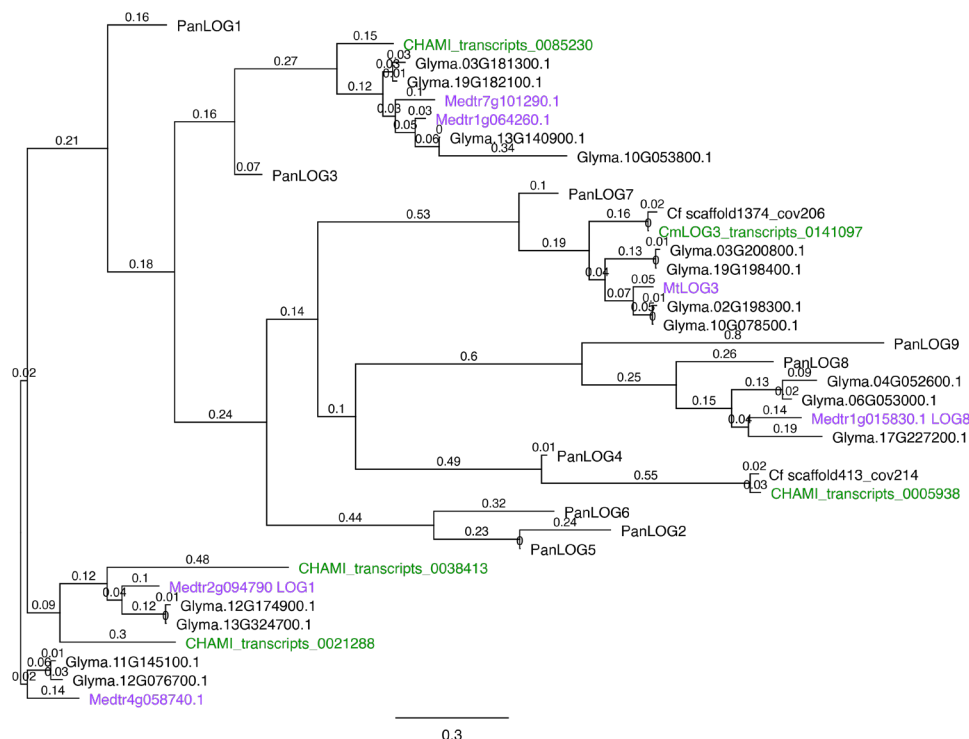
The phylogenetic tree of ARRs was reconstructed using Geneious Software v 8.1 (FastTree) based on maximum-likelihood method (bootstrap analysis 1000x). Include species are; *Chamaecrista mimosoides* (Cm), *Chamaecrista fasciculata* (Cf), *Parasponia andersonii* (Pan), *Arabidopsis thaliana* (At), *Medicago truncatula* (Mt), *Glycine max* (Gly), *Lotus japonicus* (Lj), and *Populus trichocarpa* (Pt). *C. mimosoides* and *C. fasciculata* Type A Response Regulators (ARRs) are highlighted by green and red colour, respectively. Branch label shows percentage of consensus support.



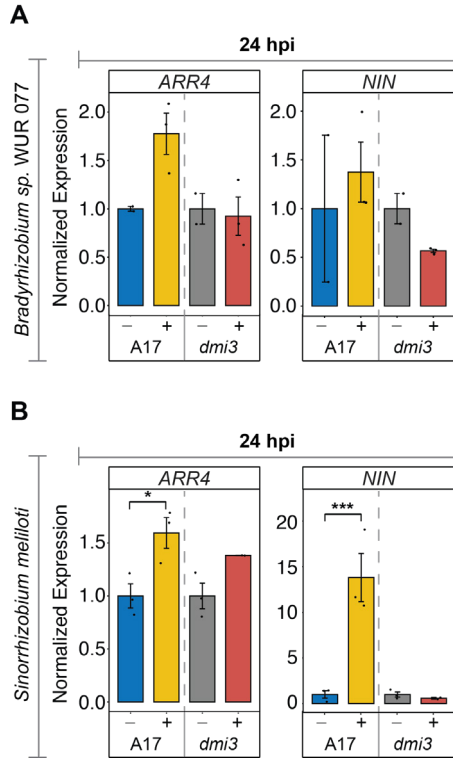
**Supplemental Figure 8. Phylogenetic analysis of isopentenyl transferase (IPT).** The phylogenetic tree of IPTs was reconstructed using Geneious Software v 8.1 (FastTree) based on maximum-likelihood method (bootstrap analysis 1000x). Include species are; *Chamaecrista mimosoides* (Cm), *Chamaecrista fasciculata* (Cf), *Parasponia andersonii* (Pan), *Arabidopsis thaliana* (At), *Medicago truncatula* (Medtr), *Glycine max* (Glyma), *Lotus japonicus* (Lj), *Eucalyptus grandis* (Eucgr), and *Populus tricarpha* (Potri). *C. mimosoides* and *C. fasciculata* IPT enzymes are highlighted by green and red colour, respectively. Branch label shows percentage of consensus support.



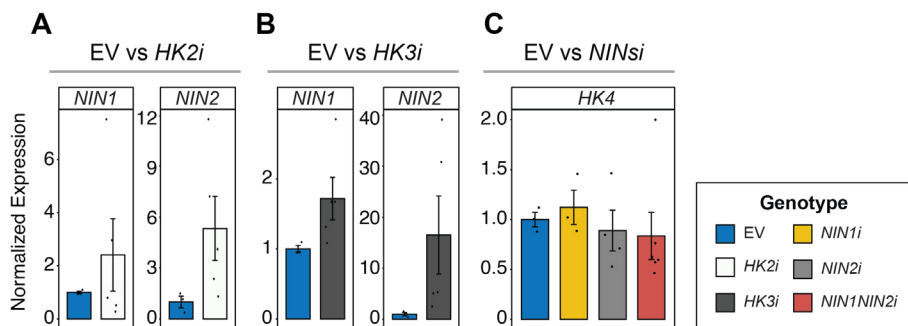
**Supplemental Figure 9. Phylogenetic analysis of CYP735A.** The phylogenetic tree CYP735a of was reconstructed using Geneious Software v 8.1 (FastTree) based on maximum-likelihood method (bootstrap analysis 1000x). Include species are; *Chamaecrista mimosoides* (Cm), *Chamaecrista fasciculata* (Cf), *Parasponia andersonii* (Pan), *Arabidopsis thaliana* (At), *Medicago truncatula* (Medtr), *Glycine max* (Glyma), *Lotus japonicus* (Lj), *Eucalyptus grandis* (Eucgr), and *Populus tricarpha* (Potri). *C. mimosoides* and *C. fasciculata* CYP735a enzymes are highlighted by green and red colour, respectively. Branch label shows percentage of consensus support.



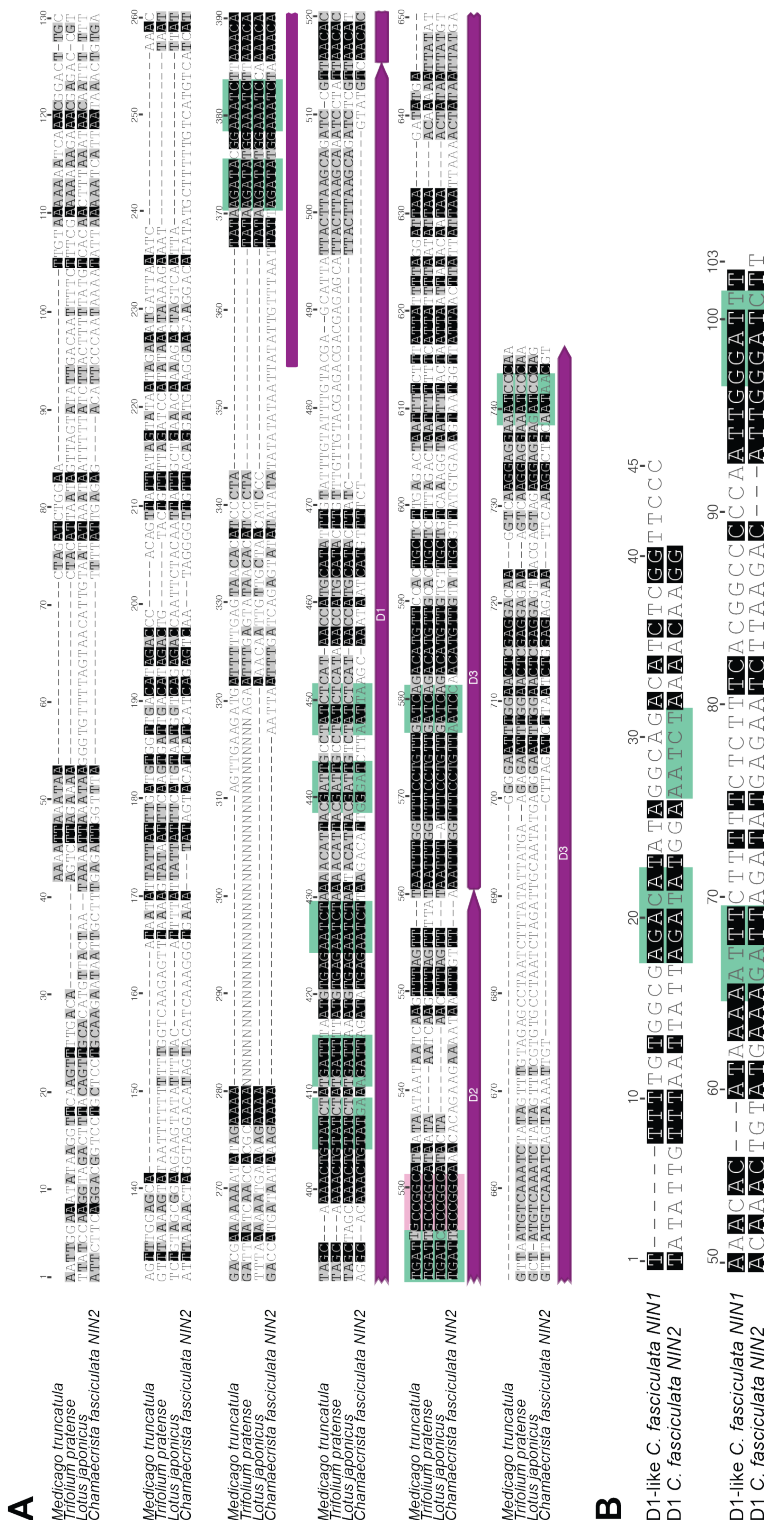
**Supplemental Figure 10. Phylogenetic tree analysis of LONELY GUY (LOG).** The phylogenetic tree of LOG was reconstructed using Geneious Software v 8.1 (FastTree) based on maximum-likelihood method (bootstrap analysis 1000x). Include species are; *Chamaecrista mimosoides* (Cm), *Chamaecrista fasciculata* (Cf), *Parasponia andersonii* (Pan), *Arabidopsis thaliana* (At), *Medicago truncatula* (Medtr), *Glycine max* (Glyma), *Lotus japonicus* (Lj), *Eucalyptus grandis* (Eucgr), and *Populus tricarpha* (Potri). *C. mimosoides* and *C. fasciculata* LOG enzymes are highlighted by green and red colour, respectively. Branch label shows percentage of consensus support.



**Supplemental Figure 11. Relative gene expression of the early symbiotic responsive genes *ARR4* and *NIN* in *M. truncatula* A17 WT and *dmi3* mutant.** Normalized gene expression of samples treated for 24 hours with (A) *Bradyrhizobium* sp WUR077 or (B) *Sinorhizobium meliloti*. Bacteria treatment (+), mock treatment (-), n=2-3, bars represent averages  $\pm$  SEM, Significance:  $p < 0.001$  \*\*\*\*,  $p < 0.05$  \*\*, student *t*-test.



**Supplemental Figure 12. Relative gene expression of *NIN1*, *NIN2* and *HK4* in transgenic *C. mimosoides* roots expressing different RNAi knockdown constructs.** Normalized expression level of *NIN1* and *NIN2* in **(A)** EV vs *HK2i* or **(B)** EV vs *HK3i*. **(C)** Normalized expression level of *HK4* in *NIN1i*, *NIN2i* or *NIN1NIN2i*.  $n=3-8$ , bars represent averages  $\pm$  SEM,  $p>0.05$  ' ', student *t*-test.



**Supplemental Figure 13. MAFFT alignment of the upstream promoter regions of *NIN1* and *NIN2* in *C. fasciculata*.** Level of conservation is indicated by shading; black = 100%, grey = 80–100%, and white = less than 60% identity. **(A)** The ~20 kb upstream region of *NIN2* aligned with the corresponding region of *Medicago truncatula*, *Lotus japonicus* and *Trifolium pratense*. Highlighted in purple are the three domains (D1, D2 and D3) containing the putative AP2 (highlighted in pink) and Type B Response Regulator (highlighted in green) binding sites as previously described by Liu and colleagues (2019c). **(B)** a ~1 kb upstream region of *NIN1* (D1-like) aligned with the putative D1 located ~28 kb upstream of *NIN2*, containing three putative Type B Response Regulator (highlighted in green) binding sites.



# Chapter 3

## **Transforming, Genome Editing and Phenotyping the Nitrogen-Fixing Tropical Cannabaceae Tree *Parasponia andersonii***

**Titis A. K. Wardhani<sup>1,2</sup>, Yuda Purwana Roswanjaya<sup>1,2</sup>, Simon Dupin<sup>1,3</sup>, Huchen Li<sup>1,4</sup>, Sidney Linders<sup>1</sup>, Marijke Hartog<sup>1</sup>, Rene Geurts<sup>1</sup>, Arjan van Zeijl<sup>1</sup>**

<sup>1</sup>Laboratory of Molecular Biology, Department of Plant Sciences, Wageningen University & Research, Wageningen, The Netherlands

<sup>2</sup>Center of Technology for Agricultural Production, Agency for the Assessment and Application of Technology (BPPT), Jakarta, Indonesia

<sup>3</sup>Department of Ecological Science, Faculty of Earth and Life Sciences, Vrije Universiteit Amsterdam, Amsterdam, The Netherlands

<sup>4</sup>Beijing Advanced Innovation Center for Tree Breeding by Molecular Design, Beijing University of Agriculture, Beijing, China

**Published in *J. Vis. Exp.* (150), e59971, doi:10.3791/59971 (2019)**

Not all those who wander are lost.

—J.R.R. Tolkien, *The Lord of the Rings, The Fellowship of the Ring*





## Summary

*Parasponia andersonii* is a fast-growing tropical tree that belongs to the Cannabis family (Cannabaceae) and can form nitrogen-fixing root nodules in association with the rhizobium. Here, we describe a detailed protocol for reverse genetic analyses in *P. andersonii* based on *Agrobacterium tumefaciens*-mediated stable transformation and CRISPR/Cas9-based genome editing.

## Abstract

*Parasponia andersonii* is a fast-growing tropical tree that belongs to the Cannabis family (Cannabaceae). Together with 4 additional species, it forms the only known non-legume lineage able to establish a nitrogen-fixing nodule symbiosis with rhizobium. Comparative studies between legumes and *P. andersonii* could provide valuable insight into the genetic networks underlying root nodule formation. To facilitate comparative studies, we recently sequenced the *P. andersonii* genome and established *Agrobacterium tumefaciens*-mediated stable transformation and CRISPR/Cas9-based genome editing. Here, we provide a detailed description of the transformation and genome editing procedures developed for *P. andersonii*. In addition, we describe procedures for the seed germination and characterization of symbiotic phenotypes. Using this protocol, stable transgenic mutant lines can be generated in a period of 2-3 months. Vegetative in vitro propagation of T0 transgenic lines allows phenotyping experiments to be initiated at 4 months after *A. tumefaciens* co-cultivation. Therefore, this protocol takes only marginally longer than the transient *Agrobacterium rhizogenes*-based root transformation method available for *P. andersonii*, though offers several clear advantages. Together, the procedures described here permit *P. andersonii* to be used as a research model for studies aimed at understanding symbiotic associations as well as potentially other aspects of the biology of this tropical tree.

## Keywords

*Parasponia andersonii*, Cannabaceae, tree, wood, *Agrobacterium tumefaciens*, stable transformation, CRISPR-Cas9, genome editing, nodulation, rhizobium, mycorrhization

## Introduction

*Parasponia andersonii* is a tropical tree belonging to the Cannabaceae family (Cannabaceae) and is native to Papua New Guinea and several Pacific Islands (Clason, 1936; Soepadmo, 1974; Becking, 1992). Together with 4 additional *Parasponia* species, it represents the only non-legume lineage that can establish a nitrogen-fixing nodule symbiosis with rhizobia. This symbiosis is well studied in the legume (Fabaceae) models *Medicago truncatula* and *Lotus japonicus*, which has resulted in acquiring detailed knowledge of the molecular genetic nature of nodule formation and functioning (Oldroyd, 2013). Additionally, it was demonstrated that the root nodule symbiosis in legumes is founded on the much older, and widespread arbuscular mycorrhizal symbiosis (van Velzen et al., 2018). Phylogenomic comparisons suggest that the nitrogen-fixing nodule symbioses of legumes, *Parasponia*, as well as, the so-called actinorhizal plant species that host diazotrophic Frankia bacteria, have a shared evolutionary origin (Griesmann et al., 2018; van Velzen et al., 2018; van Velzen et al., 2019). To determine whether the genes identified to be involved in the legume nodule formation are the part of a conserved genetic basis, studies on non-legume species are essential. To this end, we propose to use *P. andersonii* as a comparative research model, alongside legumes, to identify the core genetic networks underlying root nodule formation and functioning.

*P. andersonii* is a pioneer that can be found on the slopes of volcanic hills. It can meet growth speeds of 45 cm per month and reach lengths of up to 10 meters (Becking, 1979). *P. andersonii* trees are wind-pollinated, which is facilitated by the formation of separate male and female flowers (Becking, 1992; van Zeijl et al., 2018). We recently sequenced and annotated the diploid genome ( $2n = 20$ ; 560 Mb/1C) of *P. andersonii*, and assembled draft genome sequences of 2 additional *Parasponia* species; *P. rigida* and *P. rugosa* (van Velzen et al., 2018). This revealed ~35,000 *P. andersonii* gene models that can be clustered in >20,000 orthogroups together with genes from *M. truncatula*, soybean (*Glycine max*), *Arabidopsis thaliana*, woodland strawberry (*Fragaria vesca*), *Trema orientalis*, black cotton poplar (*Populus trichocarpa*) and eucalypt (*Eucalyptus grandis*) (van Velzen et al., 2018). Additionally, transcriptomic comparisons between *M. truncatula* and *P. andersonii* identified a set of 290 putative orthologues that display a nodule-enhanced expression pattern in both species (van Velzen et al., 2018). This provides an excellent resource for comparative studies.

To study the gene function in *P. andersonii* roots and nodules, a protocol for *Agrobacterium rhizogenes*-mediated root transformation has been established (Cao et al., 2012). Using this protocol, compound plants bearing transgenic roots can be generated in a relatively short time frame. This method is, also, widely applied in the legume-symbiosis

research (Boisson-Dernier et al., 2001; Kumagai and Kouchi, 2003; Limpens et al., 2004). However, the disadvantage of this method is that only roots are transformed and that each transgenic root represents an independent transformation event, resulting in substantial variation. Also, the transformation is transient and transgenic lines cannot be maintained. This makes *A. rhizogenes*-based root transformation less suited for CRISPR/Cas9-mediated genome editing. Additionally, *A. rhizogenes* transfers its root inducing locus (rol) genes to the plant genome, which once expressed interfere with hormone homeostasis (Nilsson and Olsson, 1997). This makes studying the role of plant hormones in *A. rhizogenes*-transformed roots challenging. To overcome these limitations, we recently developed a protocol for *Agrobacterium tumefaciens*-based transformation and CRISPR/Cas9-mediated mutagenesis of *P. andersonii* (van Zeijl et al., 2018).

Here, we provide a detailed description of the *A. tumefaciens*-based transformation procedure and reverse genetics pipeline developed for *P. andersonii*. Additionally, we provide protocols for the downstream handling of transgenic plantlets, including assays to study symbiotic interactions. Using the protocol described here, multiple transgenic lines can be generated in a 2-3 months period. In combination with CRISPR/Cas9-mediated mutagenesis, this allows efficient generation of knockout mutant lines. These mutant lines can be vegetatively propagated in vitro (Davey et al., 1993; Webster et al., 1995; van Zeijl et al., 2018), which allows sufficient material to be generated to start phenotypic characterization at 4 months after the transformation procedure has been initiated (van Zeijl et al., 2018). Together, this set of procedures should allow any lab to adopt *P. andersonii* as a research model for studies aimed at understanding rhizobial and mycorrhizal associations, as well as potentially other aspects of the biology of this tropical tree.

## Protocol

### 1. Grow *P. andersonii* trees in the greenhouse

1.1. Germinate *P. andersonii* WU1 seeds (Op den Camp et al., 2011a).

1.1.1. Use fresh *Parasponia* berries or soak dried berries in water for 2 h to rehydrate. Squash berries on a piece of tissue paper or rub against the inside of a tea sieve to remove the seeds.

1.1.2. Disinfect seeds using commercial bleach (~4% hypochlorite) for 15-20 min and subsequently wash the seeds 6 times using sterilized water.

1.1.3. Transfer the seeds to sterile 200  $\mu$ L PCR tubes. Fill the tubes with sterilized water, such that the seeds are completely submerged. Incubate the tubes for 10 days in a thermocycler running the following program: 30 cycles (7 °C for 4 h, 28 °C for 4 h). Do not use a heated lid, as this might kill the seeds.

1.1.4. Prepare SH-0 plates (see **Supplemental Table 1**). Transfer the seeds to SH-0 plates and incubate at 28 °C, 16 h:8 h day:night. Close plates with 2 layers of elastic sealing foil to prevent drying during incubation at 28 °C.

1.2. After seedlings have developed their first set of true leaves (~3-4 weeks after incubation at 28 °C), transfer seedlings to pots filled with commercial potting soil and cover the seedlings with a translucent plastic cup to prevent desiccation. Place pots in a 28 °C climate room or greenhouse, ~85% relative humidity, under a 16 h:8 h day:night regime.

1.3 After 1 week, remove the translucent plastic cup. Water the pots regularly and when trees grow bigger supplement with fertilizer to sustain growth.

## 2. Cloning of constructs for CRISPR/Cas9-mediated mutagenesis of *P. andersonii*

NOTE: Standard binary transformation vectors can be used for the stable transformation of *P. andersonii*. Here, as an example, is a procedure to generate constructs for CRISPR/Cas9-mediated mutagenesis using modular cloning (e.g., Golden Gate ) (Weber et al., 2011).

2.1. Identify guide RNA target sequences for the gene(s) of interest, using bioinformatics software featuring a built-in CRISPR design tool. Choose guide RNA sequences located at the 5'-end of the coding sequence of the target gene to increase the chance of obtaining full knockouts. Make sure to check for off-target effects by searching against the *P. andersonii* genome (van Velzen et al., 2018).

NOTE: Use 2 sgRNAs per target gene, preferably 200-300 bp apart. This may generate deletions that can be identified by PCR and subsequently by agarose gel electrophoresis.

2.2. Generate level 1 Golden Gate constructs containing the sgRNA sequences.

2.2.1. Design primers to amplify each individual sgRNA by inserting the 20 bp guide sequence at the position of N(20) in the following primer sequence: 5'-TGTGGTCTCAATTGN(20) GTTTTAGAGCTAGAAATAGCAAG-3'.

NOTE: If the guide sequence equals GN(19), remove the G at the 5' end of the guide sequence before inserting in the primer sequence.

2.2.2. PCR amplify sgRNAs from pICH86966::AtU6p::sgRNA\_PDS20 using the forward primers designed at step 2.2.1 and the universal reverse primer: 5'-TGTGGTCTCAAGCGTAATGCCAACTTTGTAC-3'. Use a high-fidelity heat-stable DNA polymerase and the following PCR conditions: 98 °C for 30 s; 30 cycles (98 °C for 10 s; 53 °C for 20 s; 72 °C for 10 s); 72 °C for 7 min. Successful PCR reactions yield a 165 bp amplicon.

2.2.3. Column-purify the PCR amplicon using a commercial PCR purification kit. Subsequently, set up Golden Gate reactions to clone sgRNAs behind the Arabidopsis thaliana AtU6p small RNA promoter: 10 ng of the sgRNA PCR amplicon, 150 ng of pICSL01009::AtU6p20, 60 ng of the appropriate level 1 acceptor vector, 2 µL of T4 ligase buffer, 2 µL 0.1% of bovine serum albumin (BSA), 0.5 µL of BsaI, 0.5 µL of T4 ligase, fill to 20 µL with ultra-pure water. Ensure that all sgRNAs are cloned in the same orientation to prevent hairpin formation.

2.2.4 Incubate reactions in a thermocycler running the following program: 37 °C for 20 s; 26 cycles (37 °C for 3 min; 16 °C for 4 min); 50 °C for 5 min; 80 °C for 5 min. Transform Golden Gate reactions to *Escherichia coli* and plate on LB medium (Bertani G, 1951) containing ampicillin (50 mg/L), X-Gal (200 mg/L) and IPTG (1 mM).

NOTE: Prepare stock solutions of IPTG and X-Gal in ultra-pure water and dimethylformamide, respectively. Filter sterilize the ampicillin and IPTG stock solutions and store all stocks at -20 °C. Wear gloves when handling dimethylformamide.

2.2.5. Select white colonies and isolate plasmids using a commercial plasmid isolation kit. Sequence verify isolated plasmids before continuing with Golden Gate level 2 assembly.

2.3. Assemble level 2 Golden Gate constructs for the stable transformation.

2.3.1. Perform a Golden Gate reaction using the level 1 AtU6p::sgRNA constructs (generated under section 2.2) as well as pICH47802::NPTII, pICH47742::35Spro::ΩNLS-aCas9::35Ster, the level 2 acceptor pICSL4723 and the appropriate end-linker (see (Engler et al., 2014)). Perform reactions as following: use ~100 fmol of each donor vector and ~20 fmol of the acceptor vector and add 2 µL of T4 ligase buffer, 2 µL of 0.1% BSA, 0.5 µL of Bpil, 0.5 µL of T4 ligase, fill to 20 µL with ultra-pure water.

NOTE: The level 1 plasmids pICH47802::NPTII, pICH47742::35Spro::ΩNLS-aCas9::35Ster need to be cloned first (see **Supplemental Table 2**), as is described for sgRNAs under section 2.2 (Nekrasov et al., 2013; Engler et al., 2014; Fauser et al., 2014).

2.3.2. Incubate reactions as under step 2.2.4 and transform into *E. coli*. Plate on LB medium containing kanamycin. Next day, select white colonies and isolate plasmids. Determine the correct plasmid assembly by restriction enzyme-digestion analysis.

2.4. Transform level 2 constructs to *Agrobacterium tumefaciens* strain AGL1 (Lazo et al., 1991).

### 3. Stable transformation of *P. andersonii*

3.1. Inoculate 2 LB plates containing the appropriate antibiotics with *A. tumefaciens* strain AGL1 transformed with the construct of interest. Incubate plates at 28 °C for 2 days.

3.2. Harvest young branches from greenhouse-grown trees. Use about 5 branches of 5-8 cm in length for each transformation. Ensure to only use healthy non-infected branches. Remove the leaves by cutting them as such that ~1 cm<sup>2</sup> of leaf tissue is left at the end of each petiole. Discard the leaves.

3.3. Disinfect tissue for 15 min using 1:1-diluted commercial bleach (~2% hypochlorite after dilution) containing a few drops of polysorbate 20. Then, rinse the tissue 6 times with autoclaved water.

NOTE: This step, as well as, the following steps need to be conducted inside a laminar down-flow cabinet to keep tissue sterile.

3.4. Re-suspend the *A. tumefaciens* cells from 1-2 plates in 25 mL of infiltration medium (see **Supplemental Table 1**) containing acetosyringone (20 mg/L) and a non-ionic surfactant (0.001% v/v) to reach an optical density (OD600) of ~5.

NOTE: Prepare the acetosyringone stock solution in 70% ethanol and store at -20 °C. The non-ionic surfactant needs to be filter-sterilized before adding to the infiltration medium.

3.5. Cut both the stem and petiole tissue in pieces of ~1 cm in length inside the *A. tumefaciens* suspension, thereby creating fresh wounds at both sides. Leave tissue pieces in the *A. tumefaciens* suspension for 10-30 min.

3.6. Prepare the rooting medium (see **Supplemental Table 1**) and add acetosyringone (20 mg/L) after autoclaving. Dry tissue pieces on a sterile piece of filter paper and place them on the medium (~10 explants/plate). Incubate plates in dark at 21 °C for 2 days. NOTE: Allow the medium to cool down to ~60 °C prior to adding acetosyringone.

3.7. After 2 days, inspect plates for fungal or obvious bacterial contamination (bacteria other than *A. tumefaciens*). Contaminated plates need to be discarded.

3.8. Prepare liquid SH-10 medium (see **Supplemental Table 1**). After autoclaving, add polysorbate 20 (0.01%, v/v). Transfer tissue pieces to 10 mL of SH-10 containing polysorbate 20. During a period of at least 10 min, gently agitate every 2-3 min to wash the tissue.

3.9. Wash two additional times with fresh SH-10 containing polysorbate 20. These times, a 2-3 min incubation time per washing step is enough.

3.10. Prepare the rooting medium (see **Supplemental Table 1**). After autoclaving, add cefotaxime (300 mg/L) and kanamycin (50 mg/L) and pour plates. For the secondary transformations (transformations of transgenic kanamycin-resistant lines), apply hygromycin (15 mg/L) selection.

3.11. Dry tissue pieces on sterile pieces of filter paper. Afterwards, transfer tissue pieces to the plates prepared in step 3.10.

3.12. Incubate plates for 7 days at 28 °C, 16 h:8 h day:night. Every 2 days check plates for fungal or bacterial contamination and excessive growth of *A. tumefaciens*. In case of contamination, transfer non-infected pieces to a fresh plate.

3.13. After 7 days, transfer tissue pieces to propagation medium (see **Supplemental Table 1**) containing cefotaxime (300 mg/L) and kanamycin (50 mg/L). Incubate plates at 28 °C, 16 h:8 h day:night. Refresh plates once a week until transgenic shoots develop. Ensure to only transfer non-infected tissue pieces to fresh plates. Discard the pieces that are overgrown by *A. tumefaciens*.

3.14. Once putatively-transgenic shoots are  $\geq 1$  cm in length, cut shoots and culture them independently in the propagation medium containing cefotaxime (300 mg/L) and kanamycin (50 mg/L). To ensure that shoots represent independent transformants, take only a single shoot from each side of an explant.

3.15. Vegetatively propagate putatively-transgenic shoots as described under step 5.2.

#### 4. Genotyping of putatively-transgenic shoots

4.1. Design primers spanning the sgRNA recognition site(s). To allow PCR amplicon sequencing, choose primers 150-250 bp away from the sgRNA recognition site(s).

4.2. Cut a leaf tip (~5 mm) from each transgenic shoot to be genotyped. Also, harvest a wild-type control sample.

4.3. Perform 50  $\mu$ L PCR reactions using the primers designed at step 4.1 and a commercial kit to directly amplify DNA from plant samples. Alternatively, PCR reactions can be performed on purified DNA using a high-fidelity polymerase.

4.4. Separate PCR amplicons on a 1.5-2% agarose gel.

4.5. Analyze the results from gel electrophoresis. Check for samples producing multiple bands (more than 1 allele) and PCR amplicons with sizes different from wild type, which indicates the presence of medium-sized indels.

4.6. Sequence PCR amplicons to identify the exact mutations. For samples producing a single PCR amplicon, PCR products can be sequenced directly. Samples that produce more than 1 band after gel electrophoresis or that appear to be heterozygous after direct sequencing of the PCR amplicon, need to be cloned into a blunt-end cloning vector first. Subsequently, sequence multiple clones for each sample to identify all possible alleles present in the sample.

4.7. Align sequencing results to the gene of interest and inspect the alignment to check for mutations near the sgRNA target site(s). Subsequently, check whether these mutations create frameshifts. Discard lines with > 2 alleles, and lines containing in-frame mutations.

4.8. Select several lines for further analysis.

4.9. Propagate selected lines as described under step 5.2.

4.10. When lines have developed several new shoots, take new samples from  $\geq 3$  leaf tips and repeat steps 4.3-4.7. Determine whether the mutations present in each of the samples originating from the same line as well as the original PCR sample are identical.

Lines that yield the same mutations in all samples are homogeneously mutated and can be used for further experimentation. Discard lines that do not yield the same results as these lines are chimeric.

## 5. Preparation of rooted *P. andersonii* plantlets for experimentation

5.1. Initiate a new tissue culture line of *P. andersonii*.

5.1.1. Harvest axillary buds, young adventitious shoots or leaf tissue from healthy trees. Alternatively, seedlings can be used as a starting material.

5.1.2. Disinfect tissue using 1:1-diluted commercial bleach (~2% hypochlorite after dilution) containing a few drops of polysorbate 20 for 15 min. Afterwards, rinse tissue 6 times using autoclaved water.

NOTE: This step, as well as, the following steps need to be conducted inside a laminar downflow or laminar crossflow cabinet to keep tissue sterile.

5.1.3. Transfer tissue to propagation medium (see **Supplemental Table 1**). Close plates with 2 layers of elastic sealing foil and incubate plates at 28 °C, 16 h:8 h day:night.

5.1.4. Inspect plates every few days during the first 2 weeks to ensure that tissue is free from fungal or bacterial contamination.

5.2. Propagate tissue by placing ~10 shoots on a fresh plate of propagation medium and close the plate with 2 layers of elastic sealing foil. Incubate plates at 28 °C, 16 h:8 h day:night. Repeat this step every 4 weeks.

5.3. When shoots are >1 cm in length, cut shoots at their base and place them on the rooting medium (see **Supplemental Table 1**). About 10 shoots can be placed on a single rooting plate. Position shoots upright by inserting the basal tip of the shoot into the medium. Roots appear at 10-14 days after incubation of the plates at 28 °C, 16 h:8 h day:night.

NOTE: Do not root all shoots but keep part for tissue culture propagation (see step 5.2).

## 6. Nodulation of *P. andersonii* plantlets in pots

6.1. Prepare rhizobium inoculum.

6.1.1. Inoculate 10 mL of liquid YEM medium (see **Supplemental Table 3**) from a single colony of *Mesorhizobium plurifarum* BOR26 and incubate at 28 °C for 2 days.

NOTE: *M. plurifarum* BOR2 is preferred as it efficiently nodulates *P. andersonii*. However, other rhizobium strains can also be used for nodulation of *P. andersonii* (e.g. *Bradyrhizobium elkanii* WUR3 (Op den Camp et al., 2012), *Rhizobium tropici* CIAT899 (Graham et al., 1982; Martinez-Romero et al., 1991) or *Bradyrhizobium* sp. Kelud2A4).

6.1.2. Use the 10 mL culture to inoculate a larger volume of liquid YEM medium. The volume of this culture is dependent on the number of pots that need to be inoculated.

6.1.3. Prepare liquid EKM medium (see **Supplemental Tables 4, Supplemental Table 5**). Centrifuge the bacterial culture for 10 min at 3,500 x g to harvest the cells. Subsequently, re-suspend the bacterial pellet in liquid EKM (use approximately the same volume as the original YEM culture) and determine the optical density (OD600).

6.2. For ~20 pots, prepare 3 L of liquid EKM medium and inoculate with the rhizobial suspension prepared at step 6.1.3. to reach OD600 = 0.025.

6.3. Mix 3 L of EKM containing rhizobia with 1,250 g of perlite. Subsequently, add 210 g of this mixture to sterile translucent polypropylene pots. Alternatively, instead of perlite, use sand as a substrate for nodulation assays.

6.4. Plant 1-3 *P. andersonii* plantlets in each pot. Also, prepare several pots containing *P. andersonii* plantlets transformed with the CRISPR-control construct (see **Supplemental Table 2**). Weigh several pots to be able to determine water loss during the experiment. Cover the bottom of each pot to shield the roots from light exposure.

6.5. Incubate pots in a climatized growth room (28 °C, 16 h:8 h day:night) for 4-6 weeks. Once a week, weigh several pots to determine water loss. If water loss exceeds 10 mL, supplement with ultra-pure water to compensate for the loss.

6.6. After 4-6 weeks, clean the roots from perlite and determine nodule numbers using a binocular to examine the nodulation efficiency.

## 7. Nodulation of *P. andersonii* plantlets on plates

7.1. Prepare cellophane membranes (Felten et al., 2009).

7.1.1. Cut the cellophane membrane to fit into a square 12 cm x 12 cm Petri dish. Cut the membranes a bit shorter at the top to allow space for the shoots to grow.

7.1.2. To increase the permeability of cellophane membranes, boil the membranes in EDTA solution (1 g/L) for 20 min. Afterwards, rinse at least 6x with demineralized water to remove the EDTA.

NOTE: As the dry membrane tend to wrinkle when in contact with water, submerge the dry membranes one by one into the solution.

7.1.3. Arrange the membranes horizontally in a thin layer of water in a round glass plate. Sterilize the membranes by autoclaving twice.

7.2. Place 1 autoclaved cellophane membrane on a square 12 x 12 cm Petri dish containing agar-solidified EKM medium (see **Supplemental Table 4, Supplemental Table 5**). Place two 3-week old rooted *P. andersonii* plantlets (see section 5) or 4-week old seedlings (see section 1.1) on the top of the membrane. Ensure to only pick plantlets or seedlings with roots that have white root tips, indicating that these roots are still growing.

7.3. Gently cover the roots with a second cellophane membrane, creating a sandwich layer. Seal the plate with 3 layers of elastic sealing foil. Wrap the bottom half of the plates with aluminium foil, to cover the roots from light exposure.

7.4. Incubate the plates in a climatized growth room (28 °C, 16 h:8 h day:night) for 3-4 weeks. Mark the position of the root tips to follow the root growth over time.

7.5. If the EKM plates start to dry out due to prolonged incubation, transfer the plants to fresh EKM plates a few days ahead of bacterial inoculation.

7.6. Prepare the bacterial inoculum as described at step 6.1.

7.7. Remove the top cellophane membrane and apply 1 mL of rhizobium culture (OD<sub>600</sub> = 0.025) to the roots. Subsequently, place a new cellophane membrane on the inoculated roots. Wrap the outside of the plate using aluminum foil to cover the roots from light exposure.

7.8. After 4 weeks, examine nodule numbers using a binocular to determine the nodulation efficiency.

## 8. Nodulation of *P. andersonii* seedlings in pouches

8.1. Germinate *P. andersonii* seeds as described in section 1.1. After the cotyledons have fully emerged (~12 days on SH-0 plates at 28 °C), transfer the seedlings to pouches.

8.2. To prepare the pouches, tear the folded section of the paper wick and add 7 mL of modified EKM medium (see **Supplemental Table 4**, **Supplemental Table 5**).

8.3. Insert 1 or 2 seedlings by placing the roots in between both sheets of paper that form the paper wick and the front plastic sheet of the pouch.

8.4. Shield the roots from light exposure, by folding aluminium foil around the pouch. Suspend the pouches in a plastic box covered with a translucent lid to maintain high humidity. Place the box in a climatized growth room (28 °C, 16 h:8 h day:night).

8.5. Compensate for water evaporation by adding sterile ultra-pure water, as such that the paper wick remains humid (avoid standing water at the bottom of the pouch). After the first week, this generally requires adding 2-3 ml every 4 days.

8.6. Prepare the bacterial inoculum as described at step 6.1.

8.7. After seedlings have been grown for 10-12 days in pouches, inoculate the root system with 500 µL of rhizobium culture (OD600 = 0.025).

8.8. Follow the nodule formation through time. Four weeks post inoculation, nodules can be counted and harvested to determine nodulation efficiency.

## 9. Nodule cytoarchitecture analysis

9.1. Collect 10-15 nodules in a 2 mL tube containing fixative (5% glutaraldehyde in 0.1 M phosphate buffer, pH 7.2). Apply vacuum for ½-1 h and incubate overnight at 4 °C. During this period, the samples sink to the bottom of the tube.

NOTE: The fixative solution can be stored at 4 °C for ~2-4 weeks prior usage. Make sure to wear gloves when working with tissue fixative.

9.2. Wash the nodules 2x with 0.1 M phosphate buffer, pH 7.2. Apply 10 min intervals between each washing step.

9.3. Dehydrate the samples by subsequently incubating in 30%, 50%, 70%, and 100% ethanol. To ensure that all water is removed from the samples, repeat the 100% ethanol step 3x. Apply 10 min intervals between each dehydration step.

9.4. Prepare polymerization mixture I (PM-I) by adding 1 pack of Hardener I to 2.5 mL of PEG400 mixed with 100 mL of HEMA (2-hydroxyethyl methacrylate)-based resin solution. Stir the solution for ~15 min to completely dissolve the Hardener I. Subsequently, store PM-I at -20 °C.

9.5. Remove the ethanol from step 9.3. and infiltrate the samples in the following order: PM-I:100% ethanol (1:3, v/v), PM-I:100% ethanol (1:1, v/v), and PM-I:100% ethanol (3:1, v/v). Incubate the samples in each solution at RT for ½-1 h or until the samples sink to the bottom.

9.6. Incubate samples overnight at 4 °C in 100% PM-I solution.

9.7. Prepare polymerization mixture II by mixing PM-I and Hardener II in a 15:1 (v/v) ratio. Fill the plastic mold with the polymerization solution, orient the samples horizontally at the bottom of the mold, and cover with a piece of elastic sealing foil. Avoid the formation of air bubbles.

NOTE: As the solution starts to polymerize upon exposure to RT, try to orient the samples as quickly as possible in the plastic holder. Polymerization is completed after overnight incubation at RT, or 1 h at 37 °C.

9.8. Remove the elastic sealing foil cover from step 9.7 and place a holder to the polymerized samples. To mount the holder to the samples, dissolve 10 mL of methyl methacrylate-based resin powder in 5 mL of methyl methacrylate-based resin solution. Quickly add the solution to the hole in the top of the holder.

NOTE: Perform the polymerization step in the fume hood (~30 min at RT).

9.9. Microtome section samples to a thickness of 4-5 µm. Place a microscope slide on a 58 °C hot plate and add a large drop of water to each slide. Place the sections on the top of the water. Once the water has evaporated, the sections will adhere to the slide.

9.10. Stain slides by immersing in 0.05 % (w/v) toluidine blue for 2 min. Subsequently, rinse slides 3x with ultra-pure water. Slides can be observed using a bright-field microscope.

## 10. Mycorrhization of *P. andersonii* plantlets

### 10.1. Prepare *Rhizophagus irregularis* spores' inoculum

10.1.1. Prepare a stack of polyester woven filters with the following sizes (top to bottom): 210  $\mu\text{m}$ , 120  $\mu\text{m}$ , and 36  $\mu\text{m}$  mesh size.

10.1.2. Pipette the required amount of a commercial spore suspension onto the stack of polyester filters. Rinse the filters 3x with 100 mL of autoclaved demineralized water. The spores are retained on the surface of the 36  $\mu\text{m}$  filter.

NOTE: Prepare the spore suspension in the laminar crossflow cabinet to prevent contamination.

10.1.3. Disassemble the polyester stack and keep the 36  $\mu\text{m}$  filter only. Repeat the washing step with autoclaved demineralized water for at least 6x.

10.1.4. Place the filter on a Petri dish and re-suspend the spores in autoclaved demineralized water. Use a volume of water equal to the volume of the spore suspension used in step 10.1.2. Transfer the spore suspension to a sterile tube by pipetting.

10.1.5. Place 5 drops of 20  $\mu\text{L}$  of the spore suspension on a glass slide and count the number of spores using a bright-field microscope. Convert spore counts into a ratio of spores/mL and dilute the spore suspension until it reaches 250 spores/mL. Store the spore suspension at 4 °C.

10.2. Perform mycorrhization assay. To this end, add 800 g of autoclaved sand supplemented with 70 mL of ½-Hoagland medium to sterile translucent polypropylene pots (see **Supplemental Tables 6-7**). Mix sand and medium directly in the pot by shaking vigorously.

10.3. Place one *P. andersonii* plantlet in each pot, and pipette 1 mL of the spore suspension directly onto the root of the *P. andersonii* plantlet. Ensure to include several pots containing *P. andersonii* plantlets transformed with a CRISPR-control construct (see Supplementary **Supplemental Table 1**).

10.4. Incubate pots in a climatized growth room (28 °C, 16 h:8 h day:night) for 6 weeks.

10.5. Take out the plants from the pots and wash the roots with running water to remove as much sand as possible.

10.6. Cut roots in 1 cm long pieces and boil the root pieces in 10% KOH (w/v) for 20 min at 90 °C. Subsequently, place the boiled roots on a cell strainer with a 100 µm mesh size and rinse 3x with 50 mL of water.

10.7. Stain roots with 0.05% (w/v) trypan blue in lactoglycerol (300 mL of lactic acid; 300 mL of glycerol; and 400 mL of demineralized water) for 5 min at 90 °C in a water bath or heating block. Subsequently, transfer roots to 30% glycerol. The root samples can be stored at RT.

10.8. Place 15-25 root fragments on a single microscope slide. Add 30% glycerol and cover with a cover glass and press until root pieces become flat. Observe the root fragments using a bright-field microscope and score the mycorrhizal colonization.

NOTE: A method to score mycorrhization is described according to Trouvelot et al. (Trouvelot et al., 1986). This method uses several classes (%F, %M, and %A), which allows rapid estimation of the level of mycorrhizal colonization of each root fragment and abundance of arbuscules.

## Results

*P. andersonii* trees can be grown in a conditioned greenhouse at 28 °C and ~85% relative humidity (**Figure 1. A**). Under these conditions, trees start flowering at 6-9 months after planting. Female *P. andersonii* flowers produce berries that each contains a single seed. During maturation, the berries change color; first from green to white and subsequently from white to brown (**Figure 1. B**). Seeds extracted from the ripened brown berries, germinate well after a 10-day temperature cycle and a 7-day incubation on SH-0 plates (**Figure 1. C**). Germinated seeds continue to develop into young seedlings that can be used for experimentation after ~4 weeks (**Figure 1. D**).

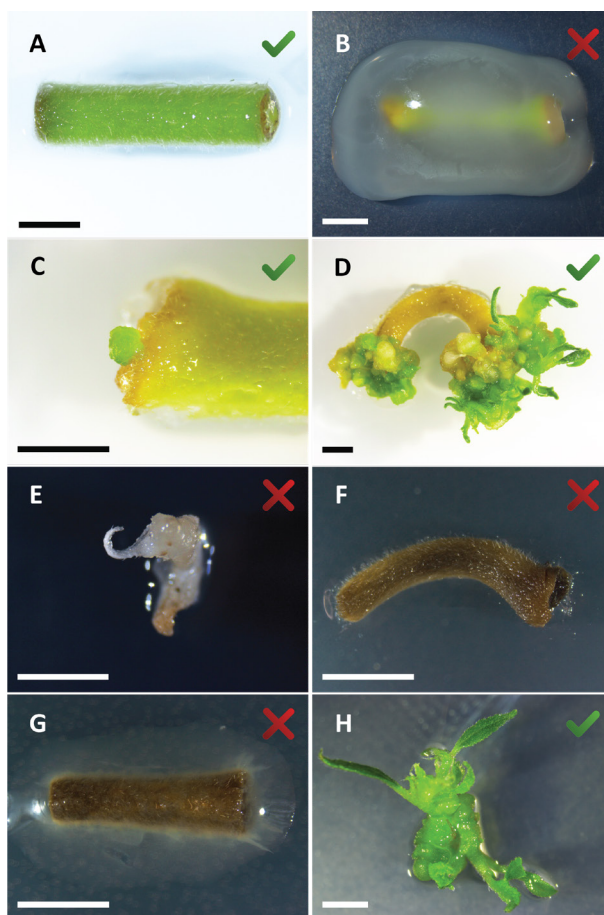
We have previously shown that petioles and segments of young *P. andersonii* stems can be efficiently transformed using *A. tumefaciens* strain AGL1 (van Zeijl et al., 2018). At the start of the transformation procedure, the tissue explants are co-cultivated with *A. tumefaciens* for 2 days at 21 °C (**Figure 2. A**). Prolonged co-cultivation results in the over-colonization of the tissue explants by *A. tumefaciens* and should, therefore, be prevented (**Figure 2. B**). After the co-cultivation period, tissue explants are transferred to selective media, which promotes outgrowth of transformed tissue. Two to three weeks later, small green micro-calli are generally observed along the original wound surface (**Figure 2. C**). These calli should continue to grow and develop 1 or more putatively-



**Figure 1: Representative images of a *P. andersonii* tree, seeds and seedlings.** (A) Six-month old *P. andersonii* tree grown in potting soil in a greenhouse conditioned at 28 °C. (B) Representative image depicting *P. andersonii* berries at various stages of maturation. Young *P. andersonii* berries (unripe) will change color from green to white and finally to brown (ripe) upon ripening. (C) *P. andersonii* seeds incubated on SH-0 medium for 1 week. A black circle indicates a germinated seedling. (D) Four-week old *P. andersonii* seedlings grown in SH-0 medium. Scale bars are equal to 25 cm in (A) and 1 cm in (B-D).

transformed shoots at 6-8 weeks after the transformation procedure has been initiated (Figure 2. D). At this stage, transformation efficiencies typically range from ~10-30% for transformations initiated with tissue explants taken from mature and partly woody branches (Supplemental Table 8).

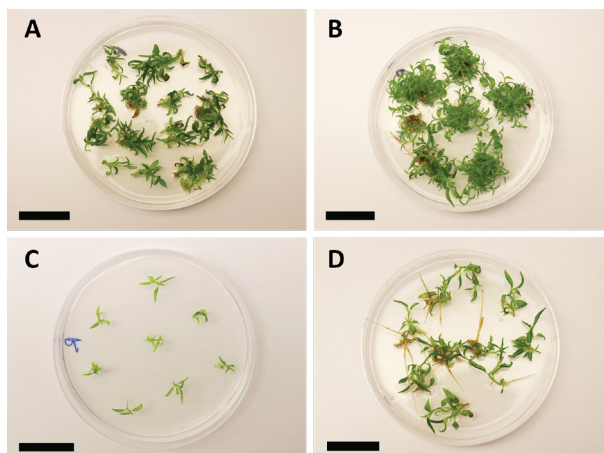
If transformations are initiated with explants taken from the young and rapidly-growing tips of branches that are not yet bearing flowers, transformation efficiencies of ~65-75% can be achieved (Supplemental Table 8). Occasionally, whitish calli are formed on the side of an explant that is not in contact with the medium and, therefore, do not experience kanamycin selection. These calli are often not transgenic and any shoots formed from these calli will generally bleach and die after direct contact with kanamycin-containing medium (Figure 2. E). In case the transformation rate is low and/or the starting material was suboptimal, tissue pieces might turn brown (Figure 2. F) and suffer from over-



**Figure 2. Representative images of explants at different stages of the stable transformation procedure.**

**(A)** Explant co-cultivated with *A. tumefaciens*. **(B)** Explant overgrown by *A. tumefaciens* during the first 2 weeks post transformation. **(C)** Transgenic micro-callus formed near the wound site of an explant at 2.5 weeks post co-cultivation. **(D)** Representative image of an explant at 6 weeks post co-cultivation showing the emergence of shoots from (transgenic) calli. **(E)** Representative image of a shoot that becomes whitish and eventually dies when in direct contact with kanamycin-containing medium. This shoot is most likely non-transgenic and escaped kanamycin selection when attached to the explant. **(F)** Representative image of an unsuccessfully transformed explant. **(G)** Representative image of an unsuccessfully transformed explant overgrown by *A. tumefaciens*. **(H)** Single transgenic shoot grown on propagation medium at 8 weeks post co-cultivation with *A. tumefaciens*. Scale bars equal 2.5 mm. Boxes containing green check marks or red crosses indicate successful or unsuccessful transformation of explants, respectively.

proliferation by *A. tumefaciens* (**Figure 2. G**). To prevent *A. tumefaciens* from spreading and overgrowing nearby explants, regular refreshment of the medium is required, and severely infected explants need to be removed. Once individual transgenic shoots are placed in the propagation medium, over-proliferation by *A. tumefaciens* is generally not occurring anymore (**Figure 2. H**).

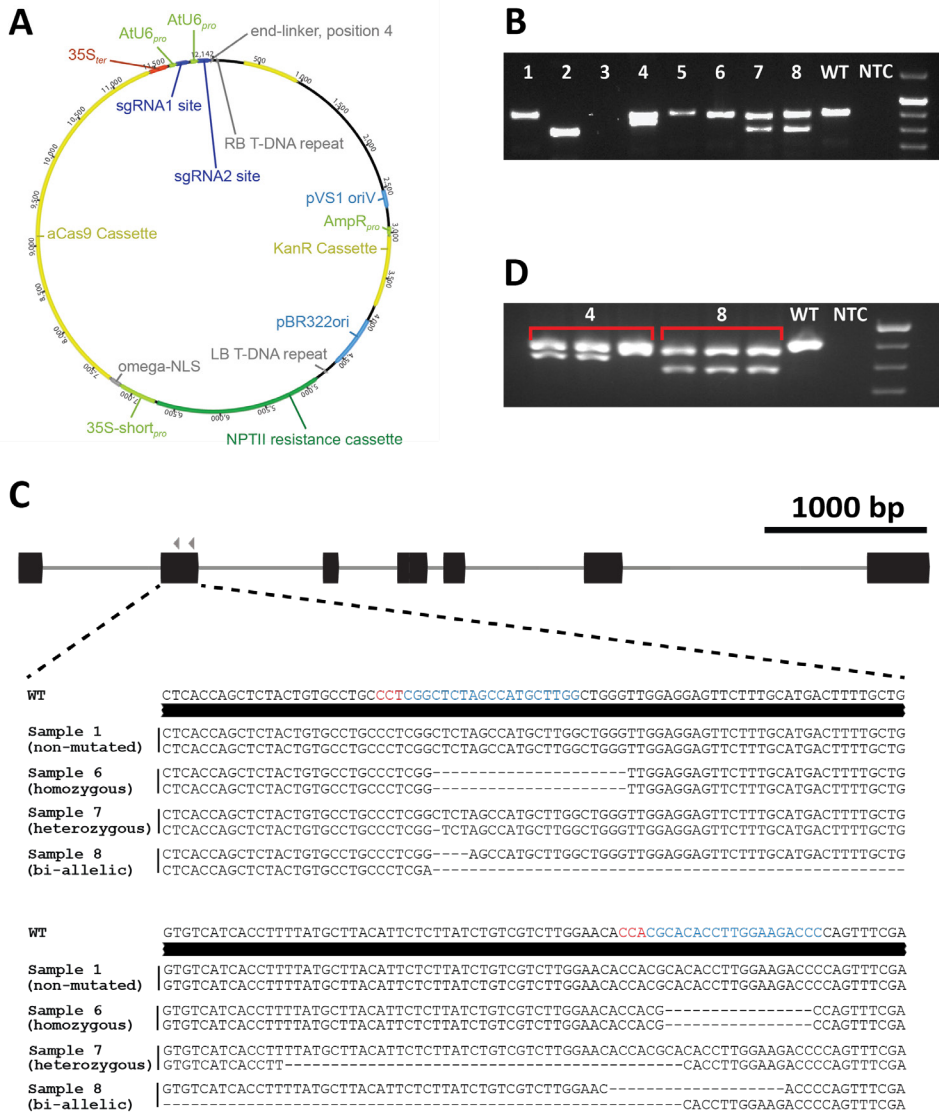


**Figure 3. Representative images of *in vitro* propagation.** (A) Shoots grown on propagation medium. The image was taken at 1 week after plates were refreshed. (B) Shoots grown on propagation medium. The image was taken at 4 weeks after plates were refreshed. (C) Freshly cut shoots placed on rooting medium. (D) Shoots incubated on rooting medium for 2 weeks. Note the presence of roots. Scale bars are equal to 2.5 cm.

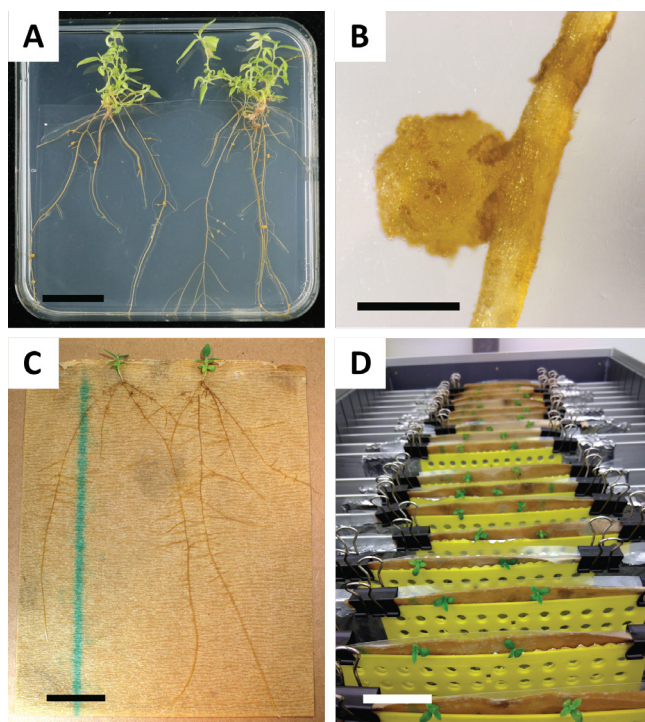
Transgenic shoots can be multiplied through *in vitro* propagation, which will give rise to tens of shoots in a period of one month (**Figure 3. A-B**). These shoots can be placed on rooting medium, which should induce root formation after ~2 weeks (**Figure 3. C-D**). Rooted plantlets can be subsequently used for experimentation.

To create knockout mutant lines, we make use of CRISPR/Cas9-mediated mutagenesis. To this end, we make use of a binary vector containing the kanamycin resistance gene NPTII, a Cas9-encoding sequence driven by the CaMV35S promoter and 2 sgRNAs per target gene that are expressed from the AtU6p small RNA promoter (Nekrasov et al, 2013). A graphical representation of the construct used for CRISPR/Cas9-mediated mutagenesis of *P. andersonii* is provided in **Figure 4. A**. Using this method, genome editing is observed in ~40% of putatively-transformed shoots (van Zeijl et al, 2018). To identify mutant lines, putatively-transformed shoots are genotyped for mutations at the sgRNA target site(s) using primers spanning the targeted region.

An example of the expected results is given in **Figure 4**. As can be seen from the photo taken after gel electrophoresis, several samples produce a PCR amplicon with similar size to the wild type (e.g., lines 1, 5, and 6 in **Figure 4. B**). These plants may contain small indels that cannot be visualized by agarose gel electrophoresis or remain unedited by the Cas9 enzyme. Additionally, several samples yield bands that are different in size from the wild type (e.g., lines 2, 4, 7 and 8 in **Figure 4. B**). In these lines, 1 (lines 4, 7 and 8) or both (line 2) alleles contain larger indels that can be easily visualized. The exact nature of the mutations at the target site(s) is revealed after PCR amplicon sequencing.

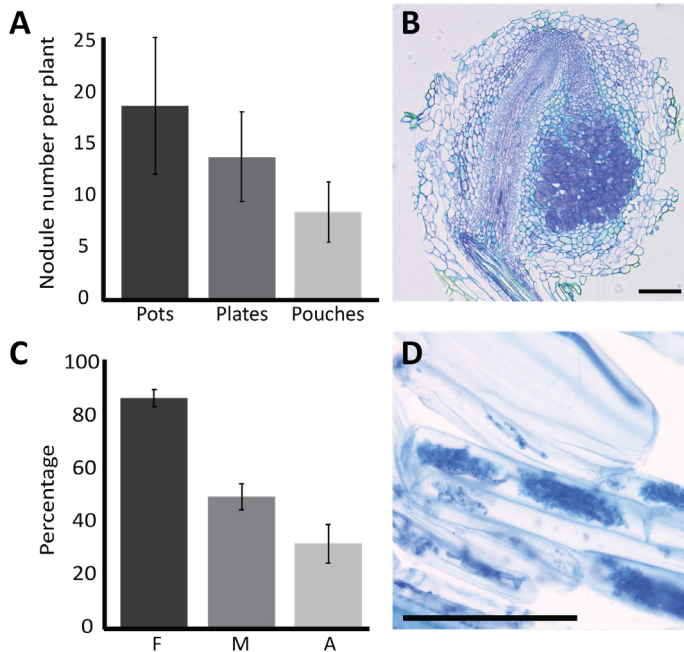


**Figure 4. Representative results after genotyping of *P. andersonii* T0 transgenic CRISPR/Cas9 mutant lines.** (A) Representative map of a binary vector used for CRISPR/Cas9-mediated mutagenesis of *P. andersonii*. (B) Representative result after PCR-based genotyping of potential CRISPR/Cas9 mutant lines using primers spanning the sgRNA target site(s). Shown is an image after agarose gel electrophoresis of amplicons. Samples taken from individual transgenic lines are indicated by numbers. Wild type (WT) and no template control (NTC) indicate lanes containing positive and negative controls, respectively. (C) Schematic representation of mutant alleles obtained after CRISPR/Cas9-mediated gene editing. Highlighted in blue and red colors are the sgRNA target sites and PAM sequences, respectively. (D) Representative result after PCR-based screening for potential chimeric mutant lines. Shown is an image after agarose gel electrophoresis of 3 individual samples taken from mutant lines 4 and 8. Note that transgenic mutant line 4 is chimeric.



**Figure 5. Representative images of nodulation assays in plates and pouches.** (A) Nodulation on plates containing agar-solidified EKM medium and inoculated with *M. plurifarium* BOR2 for 4 weeks. (B) Representative image of a *P. andersonii* root nodule. The image was taken at 4 weeks post inoculation with *M. plurifarium* BOR2. (C) Nodulation in pouches containing liquid EKM medium. Seedlings were inoculated with *Bradyrhizobium* sp. Kelud2A4 for 5 weeks. (D) Representative image of a complete setup used for the nodulation in pouches. Scale bars are equal to 2.5 cm in (A,C), 1 mm in (B), and 5 cm in (D).

As can be seen from **Figure 4. C**, both small indels of 1-4 bp, as well as, larger deletions can be obtained after CRISPR/Cas9 mutagenesis. In **Figure 4. C**, the sequence of line 1 is identical to that of the wild type, indicating that this line escaped editing and, therefore, should be discarded. Among the lines that contain mutations, heterozygous, homozygous and bi-allelic mutants can be identified (**Figure 4. C**). However, heterozygous mutants are generally rare (van Zeijl et al, 2018). Homozygous or bi-allelic knockout mutants can be propagated vegetatively to obtain sufficient material for phenotypic analysis. As phenotypic analysis is performed in the  $T_0$  generation, it is important to check whether mutant lines might be chimeric. To this end, genotyping needs to be repeated on at least 3 different samples taken from each mutant line. If the genotyping results are identical to each other and the original genotyping sample (e.g., line 8 in **Figure 4. D**), the line is homogeneously mutated and can be used for further analysis. However, if the genotyping results differ between independent samples (e.g., line 4 in **Figure 4. D**), the mutant line is chimeric and needs to be discarded.



**Figure 6. Representative results of the nodulation and mycorrhization assays. (A)** Representative bar graph showing the number of nodules formed per plant at 4 weeks post inoculation with *M. plurifarium* BOR2 in pots or on plates and at 5 weeks post inoculation with *Bradyrhizobium* sp. Kelud2A4 in pouches. Data represent mean  $\pm$  SD ( $n = 10$ ). **(B)** Representative image of a longitudinal section through a nodule formed at 4 weeks post inoculation with *M. plurifarium* BOR2. The section is stained with toluidine blue. **(C)** Representative bar graph showing quantification of mycorrhization. Variables quantified according to Trouvelot et al (1986) are F, the frequency of analyzed root fragments that are mycorrhized; M, the intensity of infection; A, the abundance of mature arbuscules in the total root system. Mycorrhization was quantified at 6 weeks post inoculation with *R. irregularis* (strain DAOM197198). Data represent mean  $\pm$  SD ( $n = 10$ ). **(D)** Representative image of mature arbuscules present in *P. andersonii* root cortical cells at 6 weeks post inoculation with *R. irregularis*. Scale bars equal 75  $\mu$ m.

Inoculation of *P. andersonii* with *M. plurifarium* BOR2 results in the formation of root nodules (**Figure 5**). As can be seen in **Figure 5. A**, these nodules are distributed along the root system. Nodules of *P. andersonii* are light brown in color but can be easily discriminated from the root tissue based on their shape (**Figure 5. B**).

Inoculation experiments in pots and subsequent growth for 4-6 weeks typically result in the formation of ~10-30 nodules (**Figure 6. A**). A similar number of nodules is formed after inoculation of EKM plate-grown *P. andersonii* plantlets at 4 weeks after inoculation (**Figure 6. A**). In pouches, *P. andersonii* seedlings typically form ~5-15 nodules at 5 weeks post inoculation (**Figure 5. C-D, 6. A**). To analyze the nodule cytoarchitecture, nodules can be sectioned and observed using bright-field microscopy. **Figure 6. B** shows an example of a longitudinal section through the middle of a *P. andersonii* nodule. This

section shows the central vascular bundle of a *P. andersonii* nodule, which is flanked by nodule lobes containing infected cells (**Figure 6. B**). *P. andersonii* plantlets can also be mycorrhized. After 6 weeks of inoculation with *R. irregularis*, mycorrhizal colonization frequency typically reaches > 80% (**Figure 6. C**). At this time point, generally ~30% of the cells contain arbuscules (**Figure 6. C**). A representative image of a *P. andersonii* root segment containing arbuscules is shown in **Figure 6. D**.

## Discussion

Legumes and the distantly-related Cannabaceae genus *Parasponia* represent the only two clades of plant species able to establish an endosymbiotic relationship with nitrogen-fixing rhizobia and form root nodules. Comparative studies between species of both clades are highly relevant to provide insights into the core genetic networks allowing this symbiosis. Currently, genetic studies are mainly done in legumes; especially the two model species *M. truncatula* and *L. japonicus*. To provide an additional experimental platform and facilitate comparative studies with a nodulating non-legume, we describe here a detailed protocol for stable transformation and reverse genetic analyses in *P. andersonii*. The presented protocol uses *in vitro* propagation of T<sub>0</sub> transgenic *P. andersonii* lines, allowing phenotypic analysis to be initiated within 4 months after *A. tumefaciens* co-cultivation. This is substantially faster than current protocols that have been established for stable transformation of legumes (Wang, 2015). This makes *P. andersonii* an attractive research model.

The protocol described here contains several critical steps. The first of which concerns seed germination. To prepare *P. andersonii* seeds for germination, seeds need to be isolated from the berries. This is done by rubbing the berries on a piece of tissue paper or against the inside of a tea sieve. This procedure needs to be performed gently in order to prevent damage to the seed coat. If the seed coat gets damaged, bleach could enter the seed during sterilization, which reduces seed viability. To break seed dormancy, seeds are subjected to a 10-day temperature cycle. However, despite this treatment, germination is not entirely synchronized. Generally, the first seeds show radicle emergence after 7 days, but others might take several days longer to germinate.

Critical points in the transformation procedure concern the choice of the starting material and the duration of the co-cultivation step. To reach efficient transformation, it is best to use healthy and young stems or petioles of non-sterile greenhouse-grown plants as the starting material. In order to induce the growth of young branches, it is advisable to trim *Parasponia* trees every 2-3 months and refresh trees once a year. Additionally, the co-cultivation step needs to be performed for 2 days only. Prolonged co-cultivation

promotes over-colonization of tissue explants by *A. tumefaciens* and generally reduces transformation efficiency. To prevent over-colonization by *A. tumefaciens* it is also important to regularly refresh the plates on which the explants are cultivated. In case over-colonization does occur, tissue explants could be washed (see Section 3.8) to remove *A. tumefaciens* cells. We advise adding bleach to the SH-10 solution used for washing (final concentration: ~2% hypochlorite). It is important to note that this additional washing step might not work on heavily-infected explants (**Figure 2. B**). In case a transformation with a CRISPR/Cas9 construct yields only a limited number of putatively-transformed shoots or if mutagenesis of a particular gene is expected to cause problems in regeneration, it is advisable to include an empty vector control construct as the positive control. Lastly, it is important to ensure that all transgenic lines that are selected are resulting from independent T-DNA integration events. Therefore, we instruct to take only a single putatively-transgenic shoot from each side of an explant. However, we realize that this reduces the potential number of independent lines. If many lines are required, researchers could decide to separate putatively-transformed calli from the original explants when these calli are  $\geq 2$  mm in size and culture these calli independently. In this way, multiple lines could be isolated from each explant, which raises the number of potential transgenic lines.

In the current protocol, transgenic lines of *P. andersonii* are propagated vegetatively through in vitro propagation. The advantage of this is that many transgenic plantlets can be generated in a relatively short time period. However, this method also has several limitations. Firstly, the maintenance of  $T_0$  transgenic lines through in vitro propagation is labor intensive and could result in unwanted genetic or epigenetic alterations (Larkin and Scowcroft, 1981; Smulders and de Klerk, 2011). Secondly,  $T_0$  lines still contain a copy of the T-DNA, including the antibiotic resistance cassette. This limits the number of possible re-transformations, as different selection markers are required for each re-transformation. Currently, we have only tested transformation using kanamycin or hygromycin selection (data not shown). Furthermore, the presence of the Cas9-encoding sequence and sgRNAs in the  $T_0$  transgenic lines complicates complementation studies. Complementation assays are possible but require the sgRNA target site(s) to be mutated as such that gene-editing of the complementation construct is prevented. Thirdly, a disadvantage of working with  $T_0$  lines is that CRISPR/Cas9 mutants might be chimeric. To prevent phenotypic analysis of chimeric mutant lines, we recommend repeating the genotyping analysis after in vitro propagation on at least 3 different shoots. Although, the number of chimeric mutants obtained using the protocol described here is limited, they are occasionally observed (van Zeijl et al, 2018). To overcome the limitations of working with  $T_0$  lines, *P. andersonii* mutant lines could be propagated generatively. *P. andersonii* trees are dioecious and wind-pollinated (Soepadmo, 1974). This means that

each transgenic line needs to be manipulated as such that male and female flowers are produced on a single individual, and subsequently grown as such that cross pollination does not occur. As *P. andersonii* is a fast-growing tree it requires a substantial amount of space in a tropical greenhouse (28 °C, ~85% relative humidity). Therefore, although technically possible, generative propagation of *P. andersonii* transgenic lines is logistically challenging.

In the protocol section, we described 3 methods for nodulation of *P. andersonii*. The advantage of the plate and pouch systems is that the roots are easily accessible, which may allow spot-inoculation of bacteria and following nodule formation over time. However, the plate system is quite labor intensive, which makes it less suited for large-scale nodulation experiments. A disadvantage of the pouch system is that it is difficult to prevent fungal contamination. Pouches are not sterile, and therefore fungal growth is often observed on the top half of the pouch. However, this does not affect *P. andersonii* growth, and therefore does not interfere with nodulation assays. Additionally, the pouch system is only suitable for seedlings. Despite several attempts, we have been unable to grow plantlets obtained through in vitro propagation in pouches.

The *P. andersonii* reverse genetics pipeline described here offers a substantial improvement compared to the existing *A. rhizogenes*-based root transformation method (Cao et al, 2012). Using the described procedures, stable transgenic lines can be generated efficiently and can be maintained via in vitro propagation. In contrast, *A. rhizogenes* transformation is transient and only results in the formation of transgenic roots. Because each transgenic root results from an independent transformation, *A. rhizogenes* transformation-based assays suffer from substantial phenotypic variation. This variation is much less in case of stable lines, although in vitro propagation also creates some level of variation. Because of this reduced variation and the fact that multiple plantlets could be phenotyped for each stable line, stable lines are more suited for quantitative assays compared to *A. rhizogenes*-transformed roots. Additionally, the stable transformation does not depend on the introduction of the *A. rhizogenes* root inducing locus (*rol*) that affects the endogenous hormone balance (Nilsson et al, 1997). Therefore, stable lines are better suited for reverse genetic analysis of genes involved in hormone homeostasis compared to *A. rhizogenes*-transformed roots. A more general advantage of *P. andersonii* as research model is that it did not experience a recent whole genome duplication (WGD). The legume Papilionoideae subfamily, which includes the model legumes *M. truncatula* and *L. japonicus*, as well as the Salicaceae (order Malpighiales) that includes the model tree *Populus trichocarpa* experienced WGDs ~65 million years ago (Tuskan et al., 2006; Cannon et al., 2006). Many paralogous gene copies resulting from these WGDs are retained in the genomes of *M. truncatula*, *L. japonicus*

and *P. trichocarpa* (Tuskan et al., 2006; Sato et al., 2008; Young et al., 2011), which creates redundancy that might complicate reverse genetic analyses. As *P. andersonii* did not experience a recent WGD, reverse genetic analyses on *P. andersonii* might be less affected by redundant functioning of paralogous gene copies.

Taken together, we provide a detailed protocol for reverse genetic analysis in *P. andersonii*. Using this protocol, single mutant lines can be efficiently generated in a timeframe of 2-3 months<sup>10</sup>. This protocol can be extended to create higher order mutants through multiplexing of sgRNAs targeting different genes simultaneously, as shown for other plant species (Xing et al., 2014; Lowder et al., 2015; van Zeijl, 2017). Additionally, the stable transformation procedure described here is not limited to CRISPR/Cas9 gene-targeting but could also be used to introduce other types of constructs (e.g., for promoter-reporter assays, ectopic expression or trans-complementation). We established *P. andersonii* as a comparative research model to study mutualistic symbioses with nitrogen-fixing rhizobia or endomycorrhizal fungi. However, the protocols described here also provide tools to study other aspects of the biology of this tropical tree, such as wood formation, the development of bi-sexual flowers or the biosynthesis of Cannabaceae-specific secondary metabolites.

## Protocol Visualisation (Video) and List of Materials

The visualized protocol and information on material and chemical compounds used in this article is available online at [JoVE Journal](#).

## Disclosures

The authors have nothing to disclose.

## Acknowledgments

The authors like to acknowledge Mark Youles, Sophien Kamoun and Sylvestre Marillonnet for making Golden Gate cloning parts available through the Addgene database. Additionally, we would like to thank E. James, P. Hadobas, and T. J. Higgins for *P. andersonii* seeds. This work was supported by The Netherlands Organization for Scientific Research (NWO-VICI grant 865.13.001; NWO-Open Competition grant 819.01.007) and The Ministry of Research, Technology and Higher Education of the Republic of Indonesia (RISET-PRO grant 8245-ID).

## Supplemental Information

**Supplemental Table 1. Composition of Schenk-Hildebrandt-based media used for growing *P. andersonii* seedlings, stable transformation, and in vitro propagation (Schenk and Hildebrandt, 1972)\*.**

Compound	SH-0	SH-10	Propagation medium	Rooting medium	Infiltration medium
SH-basal salt medium	3.2 g	3.2 g	3.2 g	3.2 g	3.2 g
SH-vitamin mixture	1 g	1 g	1 g	1 g	1 g
Sucrose	-	10 g	20 g	10 g	10 g
BAP (1 mg/mL)	-	-	1 mL (4.44 $\mu$ M)	-	-
IBA (1 mg/mL)	-	-	100 $\mu$ L (0.49 $\mu$ M)	1 mL (4.92 $\mu$ M)	-
NAA (1 mg/mL)	-	-	-	100 $\mu$ L (0.54 $\mu$ M)	-
1 M MES pH=5.8	3 mL	3 mL	3 mL	3 mL	3 mL
1 M KOH	Adjust pH to 5.8	Adjust pH to 5.8	Adjust pH to 5.8	Adjust pH to 5.8	Adjust pH to 5.8
Daishin agar	8 g	-	8 g	8 g	-

\*Dissolve solid compounds into 750 mL of ultra-pure water before adding liquid stocks. Afterwards, fill the complete medium to 1 L. Prepare BAP, IBA, NAA stocks in 0.1 M KOH and store at -20 °C.

**Supplemental Table 2. Overview of level 1 and level 2 constructs used for CRISPR/Cas9 mutagenesis.**

Construct	Description	Level	Backbone	Contains	Reference
1	npIIL resistance cassette	1	pICH47802	pICSL70004:npIIL	-
2	35Spro:QNLS-Cas9:35Ster	1	pICH47742	pICH41388:35Spro pAGM5331:QNLS pICH41308:aCas9 pICH41414:35Ster	Engler et al. (2014) Engler et al. (2014) Fauser et al. (2014) Engler et al. (2014)
3	sgRNA1	1	pICH47751	pICSL01009:AtU6p, corresponding PCR amplicon	Nekrasov et al. (2013)
4	sgRNA2	1	pICH47761	pICSL01009:AtU6p, corresponding PCR amplicon	Nekrasov et al. (2013)
5	CRISPR/Cas9_ctrl	2	pICSL4723	1R: construct 1, 2F: construct 2, end-link pICH41744	Engler et al. (2014)
6	CRISPR/Cas9_module	2	pICSL4723	1R: construct 1, 2F: construct 3, 3F: construct 4, end-link pICH41766	Engler et al. (2014)

**Supplemental Table 3. Composition of Yeast-Mannitol (YEM) medium used for growing rhizobium\*.**

Before autoclaving:		
Compound	Amount per liter	Final concentration
Mannitol	5 g	27.45 mM
Na-Gluconate	5 g	22.92 mM
Yeast extract	0.5 g	-
MgSO <sub>4</sub> ·7H <sub>2</sub> O	0.2 g	0.81 mM
NaCl	0.1 g	1.71 mM
K <sub>2</sub> HPO <sub>4</sub>	0.5 g	2.87 mM
After autoclaving:		
Compound	Amount per liter	Final concentration
1.5 M CaCl <sub>2</sub>	1 mL	1.5 mM

\*Adjust the pH to 7.0 and fill with ultra-pure water to 1 L. To prepare the agar-solidified YEM medium, add 15 g of microagar before autoclaving.

**Supplemental Table 4. Composition of 1 L modified EKM medium used for *P. andersonii* nodulation assay (Becking, 1983)\*.**

Before autoclaving:			
Compound	Stock concentration	Amount per liter medium	Final concentration
KH <sub>2</sub> PO <sub>4</sub>	0.44 M	Add 2 mL	0.88 mM
K <sub>2</sub> HPO <sub>4</sub>	1.03 M	Add 2 mL	2.07 mM
500x micro-elements stock solution	-	Add 2 mL	-
MES pH=6.6	1 M	Add 3 mL	3 mM
HCl	1 M	Adjust pH to 6.6	-
Ultra-pure water	-	Fill to 990 mL	-
After autoclaving:			
Compound	Stock concentration	Amount per liter medium	Final concentration
MgSO <sub>4</sub> ·7H <sub>2</sub> O	1.04 M	2 mL	2.08 mM
Na <sub>2</sub> SO <sub>4</sub>	0.35 M	2 mL	0.70 mM
NH <sub>4</sub> NO <sub>3</sub>	0.18 M	2 mL	0.36 mM
CaCl <sub>2</sub> ·2H <sub>2</sub> O	0.75 M	2 mL	1.5 mM
Fe(III)-citrate	27 mM	2 mL	54 μM

\*The composition of the 500x micro-elements stock solution is listed in Supplemental Table 4. To prepare 2% agar-solidified EKM medium, add 20 g of Daishin agar before autoclaving. Autoclave the MgSO<sub>4</sub>·7H<sub>2</sub>O, Na<sub>2</sub>SO<sub>4</sub>, CaCl<sub>2</sub>·2H<sub>2</sub>O, and Fe(III)-citrate stocks to sterilize. Filter sterilize NH<sub>4</sub>NO<sub>3</sub> stock solution to sterilize.

**Supplemental Table 5. Composition of the 500x micro-elements stock solution used for preparing modified EKM medium\*.**

Compound	Amount per liter	Stock concentration
MnSO <sub>4</sub>	500 mg	3.31 mM
ZnSO <sub>4</sub> ·7H <sub>2</sub> O	125 mg	0.43 mM
CuSO <sub>4</sub> ·5H <sub>2</sub> O	125 mg	0.83 mM
H <sub>3</sub> BO <sub>3</sub>	125 mg	2.02 mM
Na <sub>2</sub> MoO <sub>4</sub> ·2H <sub>2</sub> O	50 mg	0.21 mM

\*Store the micro-elements stock solution at 4 °C.

**Supplemental Table 6. Composition of ½-Hoagland medium used for mycorrhization assays (Hoagland et al., 1950)\*.**

Compounds	Stock concentration	Amount per liter medium	Final concentration
K <sub>2</sub> HPO <sub>4</sub>	20 mM	1 mL	0.2 mM
NH <sub>4</sub> NO <sub>3</sub>	0.28 M	10 mL	2.8 mM
MgSO <sub>4</sub>	40 mM	10 mL	0.4 mM
K <sub>2</sub> SO <sub>4</sub>	40 mM	10 mL	0.4 mM
Fe(II)-EDTA	9 mM	10 mL	0.9 mM
CaCl <sub>2</sub>	80 mM	10 mL	0.8 mM
50x micro-elements stock solution	-	10 mL	-

\*The composition of the 50x micro-elements stock solution is listed in Supplemental Table 7. Prepare the Fe(II)-EDTA solution by combining FeSO<sub>4</sub>·7H<sub>2</sub>O (9 mM) and Na<sub>2</sub>-EDTA (9 mM) into 1 stock solution, and store at 4 °C. Adjust the pH of the medium to 6.1 using 1 M KOH and fill with ultra-pure water to 1 L.

**Supplemental Table 7. Composition of the 50x micro-elements stock solution used for preparing ½-Hoagland medium.**

Compounds	Amount per liter	Stock concentration
H <sub>3</sub> BO <sub>3</sub>	71.1 mg	1.15 mM
MnCl <sub>2</sub> ·4H <sub>2</sub> O	44.5 mg	0.22 mM
CuSO <sub>4</sub> ·5H <sub>2</sub> O	3.7 mg	23.18 μM
ZnCl <sub>2</sub>	10.2 mg	74.84 μM
Na <sub>2</sub> MoO <sub>4</sub> ·2H <sub>2</sub> O	1.2 mg	4.96 μM

**Supplemental Table 8. Transformation efficiency of *P. andersonii*. Here, transformation efficiency is defined as the percentage of explants that form at least 1 transgenic callus or shoot.**

Age of explants	Transformation efficiency*
Young	69.4 ± 6.2% (n = 2)
Mature	18.3 ± 10.2% (n = 15)

\*Transformation efficiency was scored at 6 weeks post transformation and is depicted as mean ± SD. n indicates the number of transformation experiments from which the transformation efficiency was determined.





# Chapter 4

## **CRISPR/Cas9-Mediated Mutagenesis of Four Putative Symbiosis Genes of the Tropical Tree *Parasponia andersonii* Reveals Novel Phenotypes**

**Arjan van Zeijl<sup>1</sup>, Titis A. K. Wardhani<sup>1</sup>, Maryam Seifi-Kalhor<sup>1</sup>, Luuk Rutten<sup>1</sup>, Fengjiao Bu<sup>1</sup>, Marijke Hartog<sup>1</sup>, Sidney Linders<sup>1</sup>, Elena Fedorova<sup>1,2</sup>, Ton Bisseling<sup>1</sup>, Wouter Kohlen<sup>1</sup> and Rene Geurts<sup>1</sup>**

<sup>1</sup>Laboratory of Molecular Biology, Department of Plant Sciences, Wageningen University & Research, Wageningen, The Netherlands

<sup>2</sup>Current address: K.A.Timiryazev Institute of Plant Physiology, Russian Academy of Sciences, Moscow, Russia

**Published in *Front. Plant Sci.* 9:284. doi: 10.3389/fpls.2018.00284 (2018)**

Getting over a painful experience is much like crossing monkey bars.  
You have to let go at some point in order to move forward.

—C. S. Lewis





## Abstract

*Parasponia* represents five fast-growing tropical tree species in the Cannabaceae and is the only plant lineage besides legumes that can establish nitrogen-fixing nodules with rhizobium. Comparative analyses between legumes and *Parasponia* allows identification of conserved genetic networks controlling this symbiosis. However, such studies are hampered due to the absence of powerful reverse genetic tools for *Parasponia*. Here, we present a fast and efficient protocol for *Agrobacterium tumefaciens*-mediated transformation and CRISPR/Cas9 mutagenesis of *Parasponia andersonii*. Using this protocol, knock-out mutants are obtained within three months. Due to efficient micro-propagation, bi-allelic mutants can be studied in the T<sub>0</sub> generation, allowing phenotypic evaluation within six months after transformation. We mutated four genes - *PanHK4*, *PanEIN2*, *PanNSP1*, and *PanNSP2* - that control cytokinin, ethylene or strigolactone hormonal networks and that in legumes commit essential symbiotic functions. Knockout mutants in *Panhk4* and *Panein2* displayed developmental phenotypes, namely reduced procambium activity in *Panhk4* and disturbed sex differentiation in *Panein2* mutants. The symbiotic phenotypes of *Panhk4* and *Panein2* mutant lines differ from those in legumes. In contrast, *PanNSP1* and *PanNSP2* are essential for nodule formation, a phenotype similar as reported for legumes. This indicates a conserved role for these GRAS-type transcriptional regulators in rhizobium symbiosis, illustrating the value of *Parasponia* trees as a research model for reverse genetic studies.

## Keywords

*Parasponia andersonii*, rhizobium, nodule, symbiosis, CRISPR/Cas9, stable transformation

## Introduction

*Parasponia* are tropical tree species belonging to the Cannabis family (Cannabaceae) and are known as the only non-legume plants that can establish a nitrogen-fixing endosymbiosis with rhizobium (Akkermans et al., 1978; Clason, 1936; Trinick, 1973). The *Parasponia* genus consists of five species indigenous to the Malay Archipelago and Papua New Guinea, where they grow on the slopes of volcanic mountains (Becking, 1992; Clason, 1936; Soepadmo, 1974). *Parasponia* spp. are typical fast-growing pioneer plants, capable of covering nitrogen-poor eroded soils in a relatively short time span (Becking, 1992). Under suitable greenhouse conditions, young *Parasponia* trees can grow at speeds exceeding 45 centimeters per month, and fix up to 850 kg N ha<sup>-1</sup> year<sup>-1</sup> in association with rhizobium (Trinick, 1980; Trinick, 1981; Trinick and Hadobas, 1989). As *Parasponia* is the only non-legume that can establish rhizobium symbiosis, it may represent a valuable model to study the core genetic networks underlying this symbiosis (Behm et al., 2014; Geurts et al., 2012; Geurts et al., 2016).

Like legumes, *Parasponia* develops specialized root nodular organs to host the rhizobium partner. Nodules provide the rhizobium bacteria with suitable environmental conditions to convert atmospheric nitrogen into ammonium. The Cannabaceae and legume family (Fabaceae) diverged about a hundred million years ago (Wang et al., 2009), underlining that the rhizobium symbiosis in legumes and *Parasponia* evolved largely independent (Li et al., 2015). This is reflected in the distinct nodule-types found in both lineages (Behm et al., 2014). Legume nodules possess a large central zone of infected cells, which is surrounded by peripheral vascular bundles. In contrast, *Parasponia* nodules have a central vascular bundle and infected cells in the peripheral zone, giving these nodules a lateral root-like appearance. Nevertheless, initial comparative studies revealed that both symbioses are founded on conserved signaling networks. In legumes as well as *Parasponia*, root nodule formation is induced upon recognition of rhizobial secreted lipochitooligosaccharide (LCO) signals (Granqvist et al., 2015; Marvel et al., 1987; Op den Camp et al., 2011a). Research on model legumes, like *Medicago truncatula* and *Lotus japonicus*, showed that the perception of these symbiotic signals requires a signaling cascade that has been co-opted from the much older endomycorrhizal symbiosis (Geurts et al., 2012; Oldroyd, 2013). In legumes, activation of the LCO signaling network results in a massive transcriptional reprogramming, requiring among others the GRAS-type transcriptional regulators NODULATION SIGNALLING PATHWAY 1 (NSP1) and NSP2 and the cytokinin receptor MtCRE1/LjLHK1 (Gonzalez-Rizzo et al., 2006; Heckmann et al., 2006; Kaló et al., 2005; Murray et al., 2007; Plet et al., 2011; Smit et al., 2005; Tirichine et al., 2007). Subsequent nodule formation is tightly controlled by regulatory feedback loops, including negative regulation by ethylene signaling (Miyata et al., 2013; Penmetsa et al., 2008; van Zeijl et al., 2015a).

A reference quality genome sequence for *Parasponia andersonii* and draft genome sequences of two additional *Parasponia* species have been generated (van Velzen et al., 2017). Mining these genomes uncovered ~1,800 putative symbiosis genes, of which hundreds are close homologs of legume symbiosis genes (van Velzen et al., 2017). Initial reverse genetic studies in *P. andersonii*, using a transient *Agrobacterium rhizogenes*-based root transformation system, revealed that at least two genes - *NOD FACTOR PERCEPTION 1* (*PanNFP1*) and *CALCIUM AND CALMODULIN-DEPENDENT PROTEIN KINASE* (*PanCCaMK*) - commit conserved functions in the *Parasponia* and legume LCO signaling pathways (Op den Camp et al., 2011a). We argue that a more comprehensive comparative analysis between legumes and *Parasponia* will allow identification of conserved genetic networks that are essential to establish symbiosis with rhizobium. However, to use *Parasponia* as an effective research model - alongside the legume models *M. truncatula* and *L. japonicus* - efficient transformation and genome editing tools are required.

Here, we exploit an efficient *in vitro* micro-propagation system available for *P. andersonii* to establish stable transformation and CRISPR/Cas9-mediated mutagenesis for this species (Cao et al., 2012; Davey et al., 1993; Webster et al., 1995). We show that using *Agrobacterium tumefaciens*-mediated transformation, stable transgenic lines of *P. andersonii* can be obtained in ~3-4 months. Additionally, we show that *P. andersonii* is amenable to targeted mutagenesis using the CRISPR/Cas9 system. As ~40% of the resulting T<sub>0</sub> lines harbor bi-allelic mutations, these can be phenotyped upon *in vitro* propagation. As proof of concept, we mutated four genes in *P. andersonii* that in legumes control hormonal pathways as well as commit symbiotic functions. These include: the GRAS-type transcriptional regulators *NSP1* and *NSP2* that are essential for nodule organogenesis (Heckmann et al., 2006; Kaló et al., 2005; Smit et al., 2005) and control strigolactone biosynthesis by mediating *DWARF27* (*D27*) expression (Liu et al., 2011; van Zeijl et al., 2015b); the cytokinin receptor *HISTIDINE KINASE 4* (*HK4*) that in legumes is essential for nodule organogenesis (Gonzalez-Rizzo et al., 2006; Murray et al., 2007; Plet et al., 2011); and the ethylene signaling hub *ETHYLENE INSENSITIVE 2* (*EIN2*) that is a negative regulator of nodulation in legumes (Miyata et al., 2013; Penmetsa and Cook, 1997; Penmetsa et al., 2008).

## Results

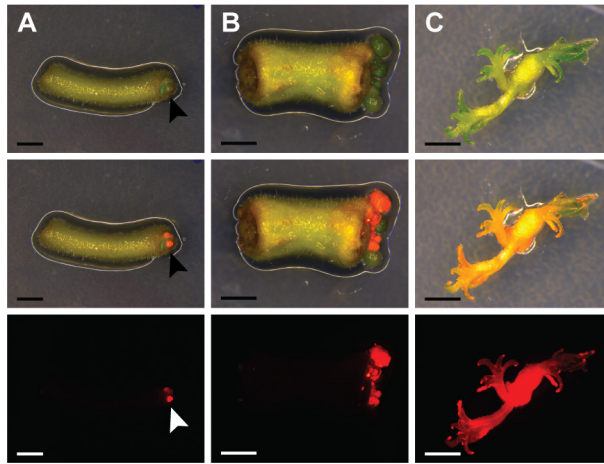
### ***Agrobacterium tumefaciens*-mediated transformation of *Parasponia***

To establish a protocol for stable transformation of *P. andersonii*, we first determined the most optimal conditions for regeneration of non-transgenic tissue. We compared regeneration efficiencies of nine tissue explant types in combination with eleven

different media, including the propagation and root-inducing media previously used for *P. andersonii* (**Supplemental Tables 4-5**) (Cao et al., 2012; Op den Camp et al., 2011a). This revealed that young stem pieces and petioles placed on original propagation medium regenerate plantlets most efficiently (**Supplemental Table 4**). Next, we questioned whether stem pieces and petioles could be transformed efficiently using *A. tumefaciens*. To this end, we used *A. tumefaciens* AGL1 carrying a binary transformation vector containing in its T-DNA the kanamycin resistance gene *NPTII* and the red fluorescent protein *DsRED1*. Co-cultivation of *A. tumefaciens* and *P. andersonii* stem or petiole explants was conducted in darkness for two days at 21°C to promote T-DNA transfer (Cao et al., 2012). Afterwards, tissue explants were placed on selective medium and incubated at 28°C in the light. These latter conditions are most favorable for *P. andersonii* regeneration (Cao et al., 2012). From day eight onwards, DsRED1-fluorescent cells could be observed near the wound surface indicating a successful transfer of the T-DNA.

Recent research on *Arabidopsis thaliana* showed that acquisition of pluripotency requires activation of a root developmental program (Kareem et al., 2015). We tested whether an initial culturing period on root-inducing medium further improves the transformation efficiency of *P. andersonii*. This showed to be the case (**Supplemental Table 6**). About half of the explants formed regenerative calli at four weeks after co-cultivation (**Figure 1. A**). When 2 mm in size, transgenic calli were separated from tissue explants, which stimulated shoot formation (**Figure 1. B-C**). Two to three months after the start of transformation, a single shoot was selected from each explant to ensure that the transgenic lines represent independent transformation events. These shoots can be genotyped and vegetatively propagated (**Supplemental Figure 1**). The latter allows clonal multiplication of individual transgenic lines in a period of ~4-6 weeks, which means that phenotyping assays could be initiated at ~4 months after the start of transformation.

To characterize the resulting transgenic *P. andersonii* lines at the molecular level, we selected - based on red fluorescence - twenty independent transformants for further analyses. PCR reactions using primers amplifying a sequence near the right T-DNA border indicated complete T-DNA integration in 19 out of 20 lines (**Supplemental Table 7**). To determine whether the transformation procedure might affect ploidy level of the regenerated transgenic lines, we estimated genome size based on flow cytometry. This showed no effect of the transformation procedure on the genome size of transgenic lines (**Supplemental Table 7**). To estimate the number of T-DNA integrations, we used quantitative RT-PCR (qRT-PCR) as well as Southern blotting. This showed an overall low T-DNA copy number, varying between one and three integrations per line (**Supplemental Table 7**). We selected three transgenic lines with a single T-DNA integration to examine



**Figure 1. *P. andersonii* stem and petiole explants can be efficiently transformed using *Agrobacterium tumefaciens*.**

(A) Petiole explant at four weeks after transformation using *A. tumefaciens*. Arrowheads indicate transgenic micro-calli. (B) Stem explant at five weeks after transformation using *A. tumefaciens*. (C) Small transgenic shoots at ten weeks after transformation. Explants were incubated on root-inducing medium for nine days, prior to transfer to propagation medium. DsRED fluorescence indicates transgenic tissue. Scale bars are equal to 2.5 mm. Shown from top to bottom are bright-field images, overlays of bright-field and DsRED fluorescence and DsRED fluorescence images.

T-DNA stability. In greenhouse-grown trees as well as *in vitro* propagated material, DsRED1 fluorescence could still be observed at 6–12 months after transgenic lines were selected (**Supplemental Figures 1–2**). This indicates that *trans*-genes remain stably integrated into the *P. andersonii* genome and actively transcribed, even after multiple rounds of vegetative propagation. Taken together, the protocol described above allows generating *A. tumefaciens*-transformed *P. andersonii* plantlets within three months, which can be phenotyped upon vegetative propagation.

### ***Parasponia* is amenable to CRISPR/Cas9-mediated mutagenesis**

To test whether CRISPR/Cas9 could be used for targeted mutagenesis in *P. andersonii*, we aimed at mutating the *P. andersonii* putative orthologues of  *EIN2*,  *MtCRE1/LjLHK1*,  *NSP1* and  *NSP2*. These genes were selected, because they control legume root nodule formation as well as commit essential non-symbiotic functions in hormone homeostasis. Putative orthologues of all four genes were previously identified from the *P. andersonii* genome and named  *PanEIN2*,  *PanHK4*,  *PanNSP1* and  *PanNSP2*, respectively (van Velzen et al., 2017). Phylogenetic reconstruction based on protein sequences confirmed that these represent the most likely orthologues of legume symbiotic genes (**Supplemental Figures 3–6**). To mutate  *PanEIN2*,  *PanHK4*,  *PanNSP1* and  *PanNSP2*, three single guide RNAs (sgRNAs) targeting  *PanHK4* and  *PanNSP2* and single sgRNAs targeting  *PanEIN2* and  *PanNSP1* were placed under an  *A. thaliana AtU6* small RNA promoter (Nekrasov et

al., 2013). These were cloned into a binary transformation vector containing the *NPTII* kanamycin resistance gene as well as a Cas9-encoding sequence fused to an N-terminal nuclear-localization signal and driven by the CaMV 35S promoter (Engler et al., 2014; Fauser et al., 2014). The resulting constructs as well a control construct containing only the *NPTII*- and Cas9-encoding sequences were transformed to *P. andersonii* using the method described above. For all constructs, transgenic shoots were obtained, although in case of the construct targeting *PanHK4* regeneration took considerably longer (up to six months). Genotyping of regenerated shoots showed that > 85% contained the Cas9 gene, indicating successful transformation. Potential mutations at any of the target sites were identified through PCR amplification and subsequent sequencing of the PCR product. This revealed mutations at the target site in about half of the transgenic shoots examined, of which the majority were bi-allelic (**Table 1**).

**Table 1.** Mutation frequency in CRISPR/Cas9 transgenic lines.

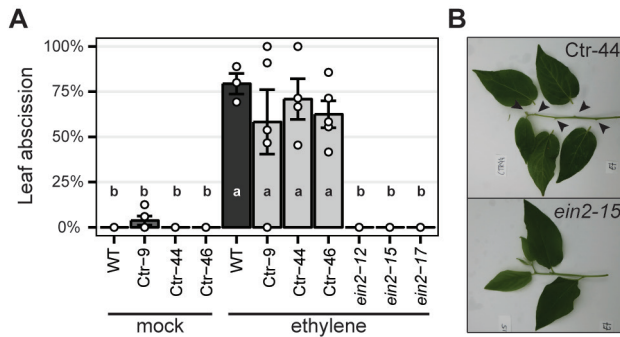
Target gene	No. of sgRNA's	No. of lines	Non-mutated <sup>a</sup>	Mutated		
				Bi-allelic	Heterozygous	Unknown <sup>b</sup>
NSP1	1	29	15 (51.7%)	11 (37.9%)	0	3 (10.3%)
NSP2	3	29	13 (44.8%)	10 (34.5%)	3 (10.3%)	3 (10.3%)
EIN2	1	9	1 (11.1%)	6 (66.7%)	2 (22.2%)	0
HK4	3	26	13 (50.0%)	12 (46.2%)	1 (3.8%)	0
Total		93	42 (45.2%)	39 (41.9%)	6 (6.5%)	6 (%)

<sup>a</sup> This includes an unknown number of individuals that are not transgenic.

<sup>b</sup> Sequencing of the PCR product indicates that plants are mutated, but the exact mutation and zygosity were not determined.

Most mutations represent small insertion and deletions but also larger deletions and inversions were identified, some of which occur in between two target sites (**Supplemental Figures 7-10**). In case of *PanHK4*, most mutants contained small in-frame deletions of three or six base pairs that most likely do not disrupt protein function. In fact, only two bi-allelic knockout mutants could be identified (**Supplemental Figure 8**). For the remaining three constructs, multiple bi-allelic knockout mutants were identified of which three individuals were selected for further studies (for an overview of mutant alleles see **Supplemental Figures 7-10**).

For phenotypic evaluation, *P. andersonii* T<sub>0</sub> transgenic lines are propagated vegetatively. Therefore, we first evaluated whether any of the mutant lines might be chimeric. To this end, tissue samples were taken from at least three different positions and genotyped for the corresponding target mutation. For each of the mutant lines, except *Pannsp2-9*, the same mutations were retrieved, suggesting that genome editing occurred soon after



**Figure 2. *Panein2* mutants are insensitive to ethylene treatment.**

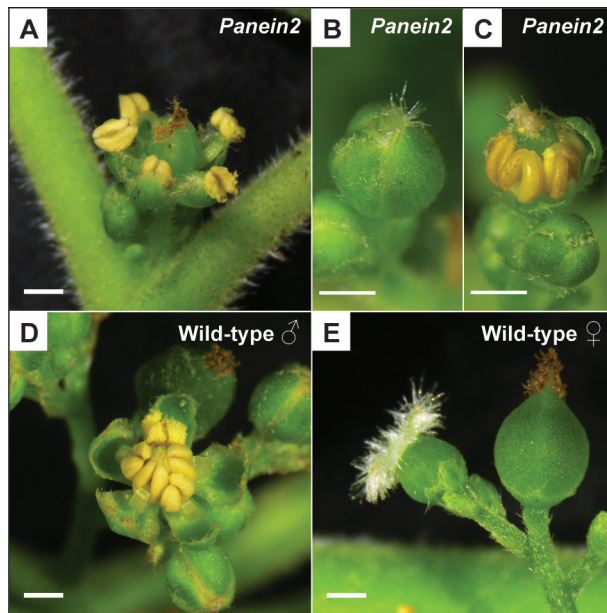
**(A)** Percentage of abscised leaves after mock or ethylene treatment. Data represent means of 3-5 biological replicates  $\pm$  SEM. Dots represent measurement values of biological repeats. Different letters indicate statistical significance ( $p < 0.05$ ) as determined by ANOVA in combination with Tukey post-hoc test. **(B)** Representative images showing abscission of leaves on a transgenic control line (Ctr-44), but not on a *Panein2* mutant. Abscission points are indicated by arrowheads.

T-DNA integration. In case of *Pannsp2-9*, chimeric mutations were detected at the first of three target sites (**Supplemental Figure 10. C**). However, the nature of the mutations at the second and third target site prevent that gene function could be restored in this line. Therefore, all eleven mutants are suitable for phenotypic evaluation. This proves that CRISPR/Cas9 can be used to efficiently mutagenize *P. andersonii* in the  $T_0$  generation.

### Non-symbiotic phenotypes in *Parasponia ein2*, *hk4*, *nsp1* and *nsp2* mutant lines

To characterize the resulting *Panein2*, *Panhk4*, *Pannsp1* and *Pannsp2* mutant lines, we studied their non-symbiotic phenotypes. *PanEIN2* putatively encodes a central component of the ethylene signaling pathway and therefore *Panein2* mutants are expected to be ethylene insensitive. One phenotype triggered in response to ethylene treatment is abscission of leaves and flowers, as shown in amongst others common bean (*Phaseolus vulgaris*), cotton (*Gossypium hirsutum*) and citrus (*Citrus clementina*) (Agustí et al., 2009; Brown, 1997; Jackson and Osborne, 1970). We exploited this phenotype to assess ethylene sensitivity of *Panein2* mutants. To this end, the tips of young shoot branches of greenhouse grown trees were exposed to ethylene gas. Within three days, ethylene triggered abscission of ~65% of treated leaves on wild-type *P. andersonii* as well as control transgenic lines (**Figure 2**). In contrast, leaf abscission was not observed on *Panein2* mutant trees (**Figure 2**). This demonstrates that *Panein2* mutants are indeed ethylene insensitive.

Inspection of *Panein2* mutant trees revealed an additional non-symbiotic phenotype. These trees form bisexual flowers containing both male and female reproductive organs (**Figure 3. A-C**). In contrast, wild-type *P. andersonii* trees form unisexual flowers that



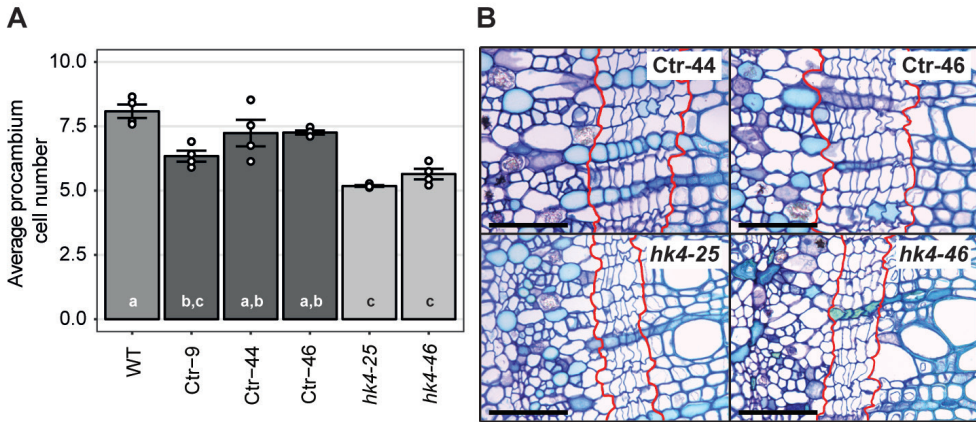
**Figure 3. *Panein2* mutants form bisexual flowers.**

(A) Mature *Panein2* flower. Note the presence of both stamen and a carpel inside *Panein2* flowers. (B) Immature *Panein2* flower. Note the presence of stigmata, indicating presence of a carpel inside the flower. (C) Immature *Panein2* flower of which sepals have been removed to show the presence of stamen. (D) Mature wild-type male flower. (E) Mature wild-type female flowers. Left, young female flower. Right, older female flower. Scale bars are equal to 1 mm.

contain either stamens or carpels (Becking, 1992) (**Figure 3. D-E**). This suggests that ethylene is involved in the regulation of *Parasponia* sex type.

Cytokinins are important regulators of cambial activity, as shown in *A. thaliana* and poplar (*Populus tremula* x *tremuloides*) (Bhalerao and Fischer, 2017; Matsumoto-Kitano et al., 2008; Nieminen et al., 2008). To determine whether reduced cytokinin sensitivity in *Panhk4* mutant lines affects the activity of the procambium, we sectioned young primary stems, 5 cm below the apical meristem. This showed that procambium activity is reduced in *Panhk4* mutant lines compared to transgenic controls (**Figure 4**). Therefore, we conclude that PanHK4-mediated cytokinin signaling is required for regulation of *P. andersonii* secondary growth.

Expression studies in *M. truncatula* previously identified a set of genes downregulated in roots of *Mtnsp1* and *Mtnsp2* mutants (Liu et al., 2011). Among these are *DWARF27* (*MtD27*; Medtr1g471050) and *MORE AXILLARY BRANCHING 1* (*MtMAX1*; Medtr3g104560) that are putatively involved in strigolactone biosynthesis (Cardoso et al., 2014; Liu et al., 2011; van Zeijl et al., 2015b; Zhang et al., 2014). We identified

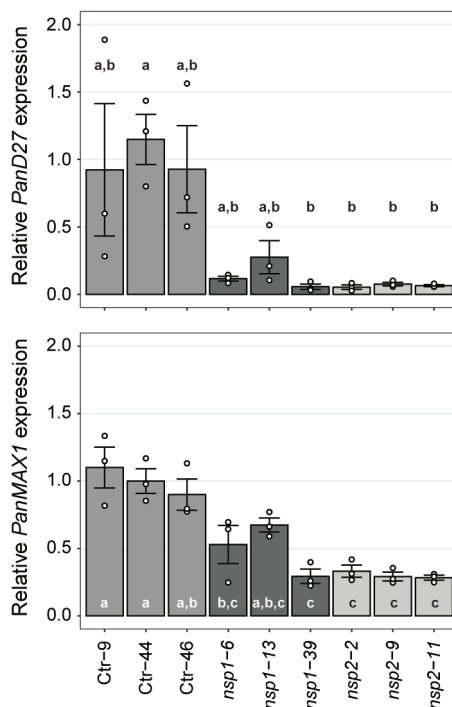


**Figure 4. *Panhk4* mutants display reduced procambium activity.**

**(A)** Average number of procambium cells in wild type, transgenic control (Ctr) and *Panhk4* mutant stems. Data represent means of four biological replicates  $\pm$  SEM. Dots represent measurement values of biological repeats. Different letters indicate statistical significance ( $p < 0.05$ ) as determined by ANOVA in combination with Tukey post-hoc test. **(B)** Representative images of stem cross-sections of control transgenic (Ctr) and *Panhk4* mutant lines. The location of the procambium is indicated by red lines. Scale bars represent 50  $\mu$ m.

putative *P. andersonii* orthologues of these genes (**Supplemental Figures 11-12**) and compared their expression levels in young root segments of three *Pannsp1*, *Pannsp2* and control plants by qRT-PCR. This showed that expression of *PanD27* and *PanMAX1* is reduced in roots of *Pannsp1* and *Pannsp2* mutant lines (**Figure 5**). We noted that *Pannsp1* mutant lines differ in the level of *PanD27* and *PanMAX1* expression. Both genes have an intermediate expression level in *Pannsp1-6* and *Pannsp1-13*, compared to *Pannsp1-39* and *Pannsp2* mutants (**Figure 5**). The three *Pannsp1* mutant lines differ from each other in the type of mutations that were created. *Pannsp1-6* and *Pannsp1-13* contain a one base pair insertion and five base pair deletion close to the 5'-end of the coding region, respectively. These mutations are immediately followed by a second in-frame ATG that in wild-type PanNSP1 encodes a methionine at position 16. In contrast, *Pannsp1-39* contains a large 232 base pair deletion that removes this in-frame ATG (see **Supplemental Figure 9**). Together, this suggests that *Pannsp1-6* and *Pannsp1-13* might represent weak alleles. Overall, these data suggest that regulation of *D27* and *MAX1* expression by NSP1 and NSP2 is conserved between *M. truncatula* and *P. andersonii*.

Taken together, we showed that EIN2, HK4, NSP1, and NSP2 in *P. andersonii* commit non-symbiotic functions in hormonal homeostasis. These functions are in line with what is described for other plant species, suggesting that the generated *P. andersonii* lines represent true mutants.



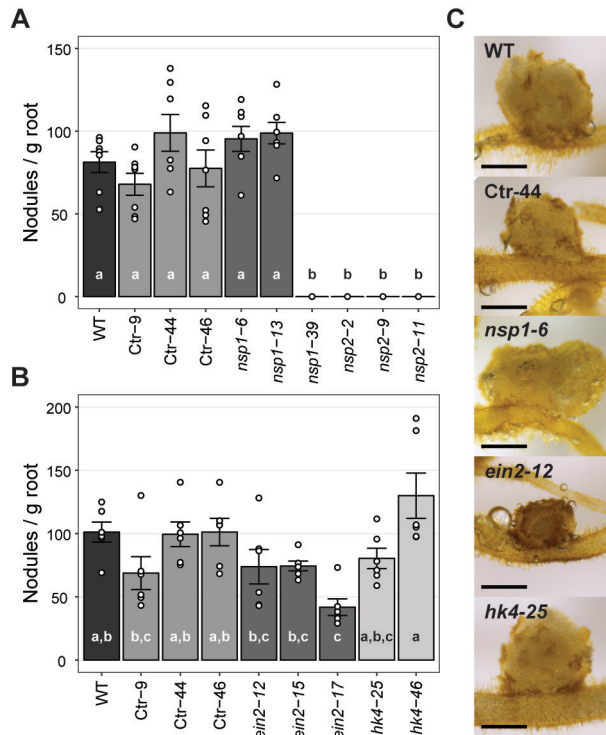
**Figure 5. Expression of *PanD27* and *PanMAX1* is reduced in *Pannsp1* and *Pannsp2* mutant roots.**

Relative expression of *PanD27* (A) and *PanMAX1* (B) in roots of transgenic control (Ctr) and *Pannsp1* and *Pannsp2* mutant lines. Data represent means of three biological replicates  $\pm$  SEM. Dots represent measurement values of biological repeats. Different letters indicate statistical significance ( $p < 0.05$ ) as determined by ANOVA in combination with Tukey post-hoc test.

### Nodulation phenotypes of *Parasponia Panein2* and *Panhk4* mutants differ from their legume counterparts

To determine whether *PanEIN2*, *PanHK4*, *PanNSP1* and *PanNSP2* perform similar functions during nodule formation as their legume orthologous, *P. andersonii* mutant plantlets were inoculated with *Mesorhizobium plurifarium* BOR2 (van Velzen et al., 2017). Nodulation phenotypes were examined one month after inoculation.

The strong *Pannsp1-39* mutant allele and all three *Pannsp2* mutant lines are unable to form root nodules (Figure 6. A and Supplemental Figure 13). This is similar as described for *M. truncatula*, *L. japonicus* and *Pisum sativum* *nsp1* and *nsp2* mutants (Heckmann et al., 2006; Kaló et al., 2005; Shtark et al., 2016; Smit et al., 2005). In contrast, the weak *Pannsp1* alleles *Pannsp1-6* and *Pannsp1-13* could be nodulated similar as wild-type or control transgenic plants (Figure 6. A and Supplemental Figure 13), suggesting that

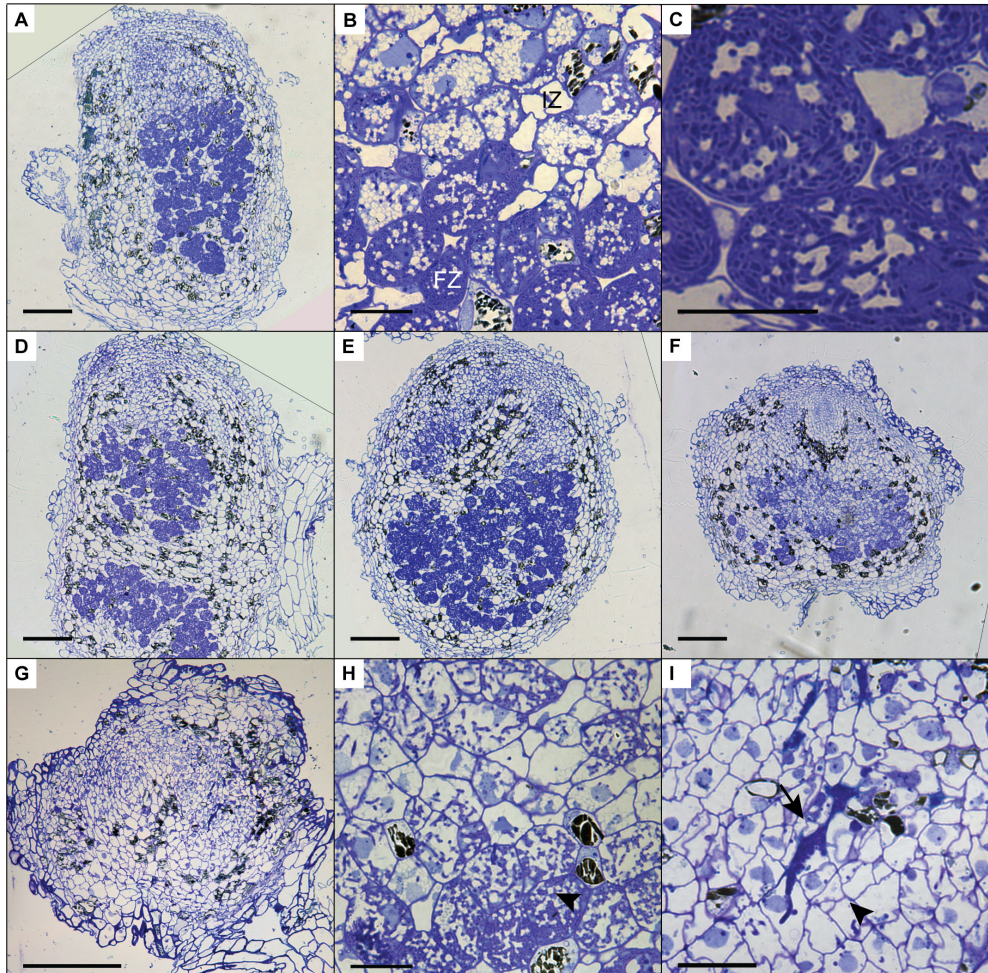


**Figure 6. Nodule formation on *P. andersonii* CRISPR/Cas9 mutant lines.**

**(A)** Nodule number per gram root fresh weight on wild type (WT), transgenic control (Ctr) and *Pannsp1* and *Pannsp2* mutant lines. Nodule number was determined at one month after inoculation with *Mesorhizobium plurifarium* BOR2. **(B)** Nodule number per gram root fresh weight on wild type (WT), transgenic control (Ctr) and *Panein2* and *Panhk4* mutant lines. Nodule number was determined at one month after inoculation with *Mesorhizobium plurifarium* BOR2. **(C)** Representative images of one month-old nodules. Note the dark color of *Panein2* nodules. Scale bars are equal to 0.5 mm. Data in **(A-B)** represent means of 5-7 biological replicates  $\pm$  SEM. Dots represent measurement values of biological repeats. Different letters indicate statistical significance ( $p < 0.05$ ) as determined by ANOVA in combination with Tukey post-hoc test. Data on nodule number and root weight are shown in **Supplemental Figure 13**.

residual PanNSP1 activity is sufficient to support root nodule formation. Overall, these data show that NSP1 and NSP2 functioning is essential for root nodule formation in *Parasponia*.

Analysis of the nodulation phenotype of *P. andersonii* *Panhk4* mutants showed that PanHK4 is not required for root nodule formation. Both *Panhk4* mutant lines formed a similar amount of nodules as wild type and transgenic controls (**Figure 6. B** and **Supplemental Figure 13**). This is different from the corresponding legume mutants - *M. truncatula* *Mtcre1* and *L. japonicus* *Ljlhk1* - that are generally not forming root nodules (Murray et al., 2007; Plet et al., 2011).



**Figure 7. Cytoarchitecture of CRISPR/Cas9 mutant nodules.**

(A) Longitudinal nodule sections of one a month-old nodule formed on transgenic control line Ctr-9. (B) Zoom in on cells in the infection (IZ) and fixation zone (FZ) of the transgenic control nodule shown in (A). Note the presence of small fragmented vacuoles in infected cells in the infection zone. (C) Zoom in on a cell in the fixation zone of a transgenic control nodule showing the presence of fixation threads. (D) Longitudinal nodule sections of one a month-old nodule formed on *Pannsp1-13*. (E) Longitudinal nodule sections of one a month-old nodule formed on *Panhk4-25*. (F) Longitudinal nodule sections of one a month-old nodule formed on *Panein2-15*. (G) Longitudinal nodule sections of one a month-old nodule formed on *Panein2-17*. (H) Zoom in on cells in the basal part of the *Panein2-15* nodule shown in (F). Indicated by an arrowhead are cells containing fixation threads. (I) Zoom in of the *Panein2-17* nodule shown in (G). Indicated by an arrowhead are infection threads. Indicated by an arrow are large apoplastic colonies. Scale bars in (A, D-G) are equal to 150  $\mu\text{m}$  and 25  $\mu\text{m}$  in (B-C, H-I).

The phenotype of *P. andersonii* *Panein2* mutants also differs from that of legume mutants. *M. truncatula ein2* mutants - as well as *L. japonicus* plants in which both EIN2-encoding genes have been silenced - form more nodules than wild type, which are clustered in distinct zones along the root (Miyata et al., 2013; Penmetsa and Cook, 1997). *Panein2* mutants do not form such nodule clusters and nodule number is not higher than wild type (**Figure 6. B** and **Supplemental Figure 13**). However, nodules formed on *Panein2* mutant plants are smaller and dark colored when compared to nodules of control plants (**Figure 6. C**). This suggests impaired nodule development in *P. andersonii ein2* mutants.

To determine the cytoarchitecture of *Panein2*, *Panhk4* and *Pannsp1-6/Pannsp1-13* mutant nodules, we sectioned ~10 nodules for each mutant line and studied these by light microscopy. Wild-type *P. andersonii* nodules harbor an apical meristem, followed by several cell layers that contain infection threads (**Figure 7. A**) (Op den Camp et al., 2012). Below this infection zone, 2-3 cell layers are present that display vacuolar fragmentation and increase in size compared to non-infected cells (**Figure 7. B**). These cells are followed by cells that are filled with fixation threads (**Figure 7. B-C**). The general cytoarchitecture of *Panhk4* and *Pannsp1-6/Pannsp1-13* mutant nodules does not differ from that of wild-type or transgenic control nodules (**Figure 7. A, D-E**), suggesting that these are functional. In contrast, in *Panein2* mutant nodules intracellular infection is hampered (**Figure 7. F-I**). Most (> 75%) *Panein2* mutant nodules harbor only infection threads as well as large apoplastic colonies (**Figure 7. I**). Some mutant nodules, harbor cells that contain fixation threads. However, even in the best nodules, fixation thread formation is severely delayed and many cells in the fixation zone still show vacuolar fragmentation (**Figure 7. I**). This shows that ethylene signaling is required for efficient fixation thread formation in *P. andersonii* nodules.

Taken together, these data reveal symbiotic mutant phenotypes for *nsp1*, *nsp2* and *ein2*, whereas no effect on nodule formation was found by knocking out *hk4* in *P. andersonii*. Interestingly, we uncovered a novel role for the ethylene signaling component EIN2 in intracellular infection of *P. andersonii* nodules.

## Discussion

Comparative studies between legumes and the Cannabaceae tree *Parasponia* can provide insights into ‘core’ genetic networks underlying rhizobium symbiosis (van Velzen et al., 2017). To facilitate such studies, we aimed to establish a reverse genetics platform for *P. andersonii* based on CRISPR/Cas9 genome editing. We show that using *A. tumefaciens* transformation, *P. andersonii* stable transgenic lines can be obtained in 3-4 months. In combination with CRISPR/Cas9 mutagenesis, this allows efficient generation of bi-allelic knockout mutants. As a proof-of-concept, we mutated four

genes that commit essential symbiotic functions in legumes as well as control different hormonal networks. Characterization of the resulting lines revealed both symbiotic as well as non-symbiotic mutant phenotypes. Therefore, we conclude that stable *A. tumefaciens*-mediated transformation in combination with CRISPR/Cas9 genome editing can be efficiently used for reverse genetic analysis in *P. andersonii*.

Plant transformation efficiency is the main bottleneck in plant genome editing (Altpeter et al., 2016; Ledford, 2016). Especially regeneration of an entire transgenic plant out of a single transformed cell remains difficult for most plant species. We took advantage of an efficient micro-propagation system available for *Parasponia* spp. to establish a protocol for stable transformation (Cao et al., 2012; Davey et al., 1993; Webster et al., 1995). About 8-12 weeks after cocultivation with *A. tumefaciens*, ~50% of explants develop transgenic shoots. This relatively high efficiency is, in part, obtained through an initial nine-day culturing period on root-inducing medium, before incubation on standard propagation medium. This adaptation in the protocol was inspired by a recent study that showed that regeneration of plant cells consists of two distinctive steps (Kareem et al., 2015). Regenerative competence is established through activation of a root developmental program, followed by activation of shoot promoting factors that are required to complete shoot regeneration (Kareem et al., 2015). The latter explains why transfer to propagation medium is required to regenerate *P. andersonii* transgenic shoots. However, this promoting effect of rooting medium on regeneration of transgenic shoots might differ between different explant types, as noted for *P. andersonii* stems and petioles (**Supplemental Table 3**).

An advantage of the *Parasponia* system is that  $T_0$  transgenic knockout mutants can be clonally propagated through *in vitro* micro-propagation (Cao et al., 2012; Davey et al., 1993; Webster et al., 1995). This allows a large number of rooted plantlets to be generated in a relatively short time span. As a result, phenotypic characterization can be initiated already at four months after the start of the transformation. However, a disadvantage of clonal propagation in combination with CRISPR/Cas9 mutagenesis is the possibility of obtaining chimeric mutants. Among the mutant lines we created, we identified one line (out of twelve) that was chimeric for one out of three CRISPR target sites (**Supplemental Figure 10C**). Most mutant lines were genetically homogeneous, suggesting that mutations are induced soon after T-DNA integration. This is consistent with results in poplar, which also revealed a low percentage of chimeric mutants (Fan et al., 2015). Since chimeras are observed occasionally, thorough genotypic analysis will be required when phenotyping is performed in the  $T_0$  generation. Besides vegetative propagation, *Parasponia* trees can also be propagated generatively. Under suitable greenhouse conditions, *Parasponia* trees flower within ~6-9 months and are self-compatible (Becking, 1992). However, *Parasponia* trees can be monoecious or

diecious and female flowers are wind pollinated (Soepadmo, 1974). This complicates selfing of trees and the production of pure seed badges. Additionally, *Parasponia* trees are fast growing and occupy a substantial amount of space in a tropical greenhouse (28°C, ~100% relative humidity), making generative propagation of multiple mutant lines logistically somewhat challenging. An alternative to generative propagation is *in vitro* maintenance of transgenic lines. Additionally, the fast and efficient transformation procedure presented here will allow recreation of a particular mutant in less than six months.

Among the mutants we created, *Panhk4* and *Panein2* showed symbiotic phenotypes that differ from corresponding legume mutants. *P. andersonii* *Panhk4* mutants form nodules with a wild-type cytoarchitecture, indicating that these nodules are most likely functional. Analysis of stem cross-sections showed that *Panhk4* mutants possess a reduced procambial activity. Similar phenotypes are observed in homologous mutants in *A. thaliana* (Mahonen et al., 2006a; Mahonen et al., 2006b). Procambium activity is slightly reduced in the orthologous receptor mutant *arabidopsis histidine kinase 4* (*ahk4*), whereas it is completely abolished in the *ahk2 ahk3 ahk4* triple mutant (Mahonen et al., 2006a; Mahonen et al., 2006b). The comparable phenotypes in cambium activity upon mutating histidine kinases suggests that *PanHK4* encodes a functional cytokinin receptor. *M. truncatula* and *L. japonicus* mutants in the cytokinin receptors orthologous to *PanHK4* are characterized as nodulation deficient (Murray et al., 2007; Plet et al., 2011). However, these mutants occasionally form nodules (Boivin et al., 2016; Held et al., 2014; Plet et al., 2011). This suggests redundant functioning of additional cytokinin receptors in both legume species. The *P. andersonii* genome also encodes two additional cytokinin receptors: *PanHK2* and *PanHK3* (van Velzen et al., 2017) (**Supplemental Figure 4**). Therefore, redundant functioning of one of these receptors cannot be excluded. In legumes, cell divisions associated with nodule development are initiated in the root cortex in response to epidermal perception of rhizobial signals (Xiao et al., 2014; Timmers et al., 1999). Cytokinin appears important for activation of this cortical organogenesis program (Vernie et al., 2015; Gamas et al., 2017). In *Parasponia*, cell divisions associated with nodule development are first observed in the epidermis, the cell layer that is in direct contact with the rhizobium bacteria (Lancelle and Torrey, 1984; Geurts et al., 2016). This difference in mitotically-responding tissues could create different dependencies on cytokinin signaling between legumes and *Parasponia*. However, whether this explains the absence of a symbiotic phenotype of *Panhk4* mutants requires further experimentation.

*Panein2* mutants are ethylene insensitive, as indicated by the absence of leaf abscission following ethylene treatment. Additionally, we noticed a disturbed sex differentiation in *Panein2* flowers. Functioning of ethylene in flower sex differentiation is known in

cucurbit species, like cucumber (*Cucumis sativus*) and melon (*Cucumis melo*) (Rudich et al., 1972; Tanurdzic and Banks, 2004; Yin and Quinn, 1995). Molecular genetic studies revealed that flower bud-specific expression of ACC synthase (ACS) genes, which are essential for biosynthesis of the ethylene precursor ACC, inhibits stamen development (Boualem et al., 2008; Boualem et al., 2015). In line with these findings in cucurbits, we hypothesize that EIN2-mediated ethylene signaling commits a similar function in sex differentiation in *Parasponia* species.

In symbiotic context, *EIN2* knockout mutations result in different phenotypes between *Parasponia* and legumes. In legumes, ethylene negatively regulates rhizobial infection and root nodule formation (Miyata et al., 2013; Penmetsa and Cook, 1997; Penmetsa et al., 2008). This is illustrated by the phenotype of the *M. truncatula ein2* mutant (named *sickle*) that forms extensive epidermal infection threads and clusters of small nodules (Penmetsa and Cook, 1997; Xiao et al., 2014). *P. andersonii ein2* mutants also form smaller nodules than wild type. However, in contrast to the *Mtein2* mutant, these nodules are regularly spaced on the root system. This suggests that in *Parasponia* ethylene signaling is not involved in regulating nodule number. Additionally, also the infection phenotype of *Panrein2* mutants differs from that in legumes. In *M. truncatula* and *L. japonicus*, interference with ethylene signaling increases the number of epidermal infection threads but does not affect intracellular colonization of nodule cells (Lohar et al., 2009; Nukui et al., 2004; Penmetsa and Cook, 1997). In contrast, in *P. andersonii Panrein2* mutants, intracellular colonization is hampered. Inside nodules, large apoplastic colonies are observed and fixation thread formation is severely reduced or even absent. This suggests that in *Parasponia* a functional ethylene signaling pathway is required for efficient intracellular infection of nodule cells.

Mutagenesis of the *NSP2* orthologue of *P. andersonii* indicated a conserved symbiotic role for this GRAS-type transcriptional regulator. In legumes, *NSP2* works in concert with *NSP1* to control root nodule formation (Hirsch et al., 2009). Mutagenesis of the *NSP1* orthologue of *P. andersonii* resulted in contrasting nodulation phenotypes. Two mutant lines, *Pannsp1-6* and *Pannsp1-13*, form nodules with a wild-type cytoarchitecture, whereas mutant line *Pannsp1-39* is unable to form nodules (**Figure 6-7**). However, all three mutants are affected in transcriptional regulation of strigolactone biosynthesis genes *PanD27* and *PanMAX1* (**Figure 5**). The three *Pannsp1* mutant lines differ from each other in the type of mutations that were created. *Pannsp1-6* and *Pannsp1-13* contain small deletions that are immediately followed by a second in-frame ATG that in wild-type PanNSP1 encodes a methionine at position 16. In contrast, *Pannsp1-39* contains a larger deletion that removes this in-frame ATG (see **Supplemental Figure 9**). Several reports have shown that alternative start codons are occasionally used to initiate transcription (Bazykin and Kochetov, 2011; Chabregas et al., 2003; Thatcher

et al., 2007). Therefore, *Pannsp1-6* and *Pannsp1-13* most probably represent weak alleles that still possess residual PanNSP1 function. Such residual levels of PanNSP1 are affecting the expression of strigolactone biosynthesis genes, but are still sufficient to allow nodule formation. Therefore, we argue that the *P. andersonii* *Pannsp1-39* line carries a knockout mutation, indicating that in *P. andersonii* both NSP1 and NSP2 are essential for rhizobium root nodule formation.

Taken together, we showed that *P. andersonii* can be efficiently transformed using *A. tumefaciens* and is amenable to targeted mutagenesis using CRISPR/Cas9. This protocol takes only marginally more time than the transient *A. rhizogenes* transformation system that is generally used to study root nodule formation (e.g. Boisson-Dernier et al., 2001; Cao et al., 2012; Kumagai and Kouchi, 2003; Limpens et al., 2004; Op den Camp et al., 2011a) but has several advantages. One of these is the absence of the *A. rhizogenes* root inducing locus (*rol*) that interferes with hormone homeostasis (Nilsson and Olsson, 1997). The protocol we developed will allow studies on *P. andersonii* symbiosis genes to determine to what extent legumes and *Parasponia* use a similar mechanism to establish a nitrogen-fixing symbiosis with rhizobium.

## Methods

### Plant materials and growth conditions

All experiments were conducted using *Parasponia andersonii* WU1 or offspring thereof (Op den Camp et al., 2011; van Velzen et al., 2017). *P. andersonii* trees were grown in a conditioned greenhouse at 28°C, 85% humidity and a 16/8h day/night regime. For *in vitro* culturing, *P. andersonii* was grown in an Elbanton growth cabinet at 28°C, 16/8h day/night. Growth of young *P. andersonii* plantlets for nodulation assays or qRT-PCR analysis was performed in 1 L crystal-clear polypropylene containers equipped with an gas exchange filter (OS140BOX, Duchefa Biochemie, The Netherlands). Pots were half-filled with agraperlite (Maasmond-Westland, The Netherlands) and watered with modified EKM medium (3 mM MES (C<sub>6</sub>H<sub>13</sub>NO<sub>4</sub>) pH 6.6, 2.08 mM MgSO<sub>4</sub>, 0.88 mM KH<sub>2</sub>PO<sub>4</sub>, 2.07 mM K<sub>2</sub>HPO<sub>4</sub>, 1.45 mM CaCl<sub>2</sub>, 0.70 mM Na<sub>2</sub>SO<sub>4</sub>, 0.375 mM NH<sub>4</sub>NO<sub>3</sub>, 15 µM Fe-citrate, 6.6 µM MnSO<sub>4</sub>, 1.5 µM ZnSO<sub>4</sub>, 1.6 µM CuSO<sub>4</sub>, 4 µM H<sub>3</sub>BO<sub>3</sub>, 4.1 µM Na<sub>2</sub>MoO<sub>4</sub>) (Becking, 1983) and placed in a climate room set at 28°C, 16/8h day/night. For nodulation assays, EKM medium was inoculated with *Mesorhizobium plurifarum* BOR2 (OD<sub>600</sub> = 0.025) (van Velzen et al., 2017).

### Vectors and constructs

For CRISPR/Cas9-mediated mutagenesis, binary transformation constructs were created using Golden Gate assembly (Engler et al., 2009). For an overview of all Golden Gate clones used in this study, see **Supplemental Table 1**. sgRNAs were designed

based on the principles described in Doench et al. (2014) and PCR amplified using specific forward primers and a universal reverse primer (**Supplemental Table 2**), using Addgene plasmid # 46966 as template (Nekrasov et al., 2013). These were cloned behind the AtU6p small RNA promoter and inserted behind the neomycin phosphotransferase II gene (*NPTII*) and an *A. thaliana* codon-optimized variant of Cas9 (Fauser et al., 2014) fused to an N-terminal nuclear localization signal and driven by the 35S promoter (**Supplemental Table 1**). As negative control, a binary vector was created containing only the NPTII- and NLS-Cas9-encoding sequences (**Supplemental Table 1**). To setup *P. andersonii* stable transformation, vector pKGWFS7-RR was used (Karimi et al., 2002).

### Phylogenetic reconstruction

Protein sequences of *Glycine max* (Wm82.a2.v1) (Schmutz et al., 2010), *Medicago truncatula* (Mt4.0v1) (Tang et al., 2014; Young et al., 2011) and *Populus trichocarpa* (v3.0) (Tuskan et al., 2006) were obtained through Phytozome 10 (<http://phytozome.jgi.doe.gov/>). Protein sequences of *Parasponia andersonii* (PanWU01x14\_asm01\_ann01) and *Trema orientalis* (TorRG33x02\_asm01\_ann01) were obtained from [www.parasponia.org](http://www.parasponia.org) (van Velzen et al., 2017). These sequences were mined using sequences from *Arabidopsis thaliana* (TAIR10, [www.arabidopsis.org](http://www.arabidopsis.org)) (Lamesch et al., 2012) and *M. truncatula*. Protein sequences were aligned using MAFFT v7.017 (Kato et al., 2002) implemented in Geneious 8.1.9 (Biomatters, New Zealand), using default parameter settings. Approximately-maximum-likelihood phylogenetic trees were constructed using FastTree (Price et al., 2009) implemented in Geneious 8.1.9. Mid-point rooting was applied for better tree visualization using FigTree v1.4.2 (<http://tree.bio.ed.ac.uk/software/figtree/>).

### Plant transformation

Stable transformation of *P. andersonii* was performed using *Agrobacterium tumefaciens* strain AGL1 (Lazo et al., 1991). *A. tumefaciens* was grown for two days on agar-solidified LB medium containing appropriate antibiotics. For each *P. andersonii* transformation, two Petri dishes (Ø 9 cm) of *A. tumefaciens* were used. Bacteria were scraped from plate and resuspended in 25 ml of infiltration medium (SH10 (**Supplemental Table 3**), 20 mg/l acetosyringone (Sigma, USA), 0.001% (v/v) Silwet L-77 ([www.arabidopsis.com](http://www.arabidopsis.com))). *P. andersonii* tissue explants used for transformation were harvested from mature trees grown under greenhouse conditions and sterilized in 2% commercial bleach for 15 minutes. Tissue explants were cut at both ends inside the *A. tumefaciens* suspension, creating fresh wound surfaces, and kept inside the suspension for about 20 minutes. Subsequently, excess liquid was removed from tissue explants using sterilized filter paper and explants were placed on co-cultivation medium (Root-inducing medium (**Supplemental Table 3**), 20 mg/l acetosyringone (Sigma, USA)). Plates were incubated for two days at 21°C in darkness. After two days, tissue explants were washed three

times using SH10 (**Supplemental Table 3**) and subsequently dried using filter paper. Tissue explants were placed on root-inducing medium containing 50 mg/l kanamycin and 300 mg/l cefotaxime and incubated at 28°C, 16h/8h (day/night). Nine days after transformation, tissue explants were transferred to propagation medium (**Supplemental Table 3**) containing 50 mg/l kanamycin and 300 mg/l cefotaxime. Plates were refreshed every other week. When regenerative calli reached ~2 mm in size they were separated from tissue explants to stimulate shoot formation. A single shoot was selected per tissue explant. These shoots were propagated on propagation medium (**Supplemental Table 3**), as previously described (Cao et al., 2012). Rooted plantlets were generated by placing individual shoots on root-inducing medium (**Supplemental Table 3**) (Cao et al., 2012).

### Characterization of transgenic lines

For T-DNA copy number estimates based on qPCR analysis, genomic DNA was isolated using the DNeasy Plant Mini Kit (Qiagen, Germany). qPCR was set up in a 10 µl reaction system with 2x iQ SYBR Green Super-mix (Bio-Rad, Hercules, USA) and 5 ng template DNA. The experimental setup and procedure were executed on a CFX Connect optical cycler, according to the manufacturer's protocol (Bio-Rad, USA). T-DNA copy number was estimated using two primer pairs amplifying part of the T-DNA and two primer pairs amplifying single copy *P. andersonii* genes (*PanAGT1* and *PanWU01x14\_asm01\_ann01\_338920*) that were selected based on a study by Duarte et al. (2010). Primer sequences are listed in **Supplemental Table 2**. Data analysis was performed using CFX Manager 3.0 software (Bio-Rad, USA). For T-DNA copy number estimates based on Southern blotting, genomic DNA was separately digested with *XbaI*, *HindIII* and *EcoRI*. Blots were hybridized with a 516 bp α-<sup>32</sup>P-labelled probe corresponding to part of the *nptII* gene that was amplified using primers *nptII\_Fw* and *nptII\_Rv* listed in **Supplemental Table 2**.

Genotyping of transgenic lines was performed using the Phire Plant Direct PCR Kit (Thermo Scientific, USA) and gene specific primers listed in **Supplemental Table 2**. Ploidy estimates of transgenic lines were determined by FACS as described by van Velzen et al. (2017).

To determine ethylene sensitivity of *Panein2* mutants, tips of young branches of four months-old trees were covered with 1 L plastic bags and injected with 1 ml of pure ethylene gas. After three days, bags were removed and leaf abscission examined. Total number of leaves on treated branches varied from 6-18.

### Microtome sectioning

Stem cross-sections were made from the primary stem, 5 cm below the apical meristem, of two month-old trees. Shoot tissue was fixed in 5 % glutaraldehyde and embedded in Technovit 7100 (Heraeus-Kulzer, Germany), according to the manufacturer's protocol. Semi-thin (7  $\mu$ m) sections were cut using a microtome (Reichert-Jung, Leica Microsystems, The Netherlands) and stained with 0.05% Toluidine Blue O. Images were taken using a Leica DM5500B microscope equipped with a DFC425C camera (Leica Microsystems, Germany). Average procambium cell number was quantified by averaging the number of cells within 25-40 cell files for each of the biological replicates. Nodule tissue fixation and embedding was performed as previously described (Fedorova et al., 1999). Semi-thin (0.6  $\mu$ m) sections were cut using a Leica Ultracut microtome (Leica Microsystems, Germany) and photographed as described above.

### RNA isolation and qRT-PCR analysis

RNA was isolated from snap-frozen root tips (~2-3 cm) as described by van Velzen et al. (2017). cDNA was prepared from 1  $\mu$ g of total RNA using the i-script cDNA synthesis kit (Bio-Rad, USA), following the manufacturer's instructions. RT-qPCR was set up as described above. Normalization was performed based on two stably expressed reference genes (*UNKNOWN 2* (*PanUNK2*) and *ELONGATION FACTOR 1 $\alpha$*  (*PanEF1 $\alpha$* )), chosen based on previous study (Bansal et al., 2015; Czechowski et al., 2005). All primer sequences are listed in **Supplemental Table 2**.

### Statistical analysis

Statistical differences were determined based on One-way ANOVA and Tukey post-hoc tests. Statistical analyses were performed using IBM SPSS Statistics 23.0 (IBM, USA).

### Data Availability Statement

All datasets analyzed for this study are included in the manuscript and the supplementary files. Gene identifiers for all *P. andersonii* genes used in this study can be found in **Supplemental Table 8**. Sequences can be downloaded from [www.parasponia.org](http://www.parasponia.org)

### Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## Author Contributions

Conceptualization, A.v.Z and R.G.; Methodology, A.v.Z, M.H., S.L. and W.K.; Investigation, A.v.Z., T.A.K.W, M.S.K., L.R., F.B., M.H., S.L., E.F. and W.K.; Formal Analysis, A.v.Z, T.A.K.W and E.F.; Visualization, A.v.Z.; Writing – Original Draft, A.v.Z.; Writing – Review & Editing, A.v.Z. and R.G.; Funding Acquisition, T.B. and R.G.; Supervision, R.G.

## Funding

This work was supported by NWO-VICI (865.13.001) to RG, NWO-VENI (863.15.010) to WK, European Research Council (ERC-2011-AdG294790) to TB and China Scholarship Councils (201303250067) to FB.

## Acknowledgments

The authors like to thank Renze Heidstra for help with FACS analysis and Michiel Lammers and Renze Heidstra for useful tips regarding CRISPR/Cas9 strategy. Ethylene gas was kindly provided by Arjen van de Peppel and Julian Verdonk. Golden Gate parts and cloning vectors were kindly provided by Mark Youles, Sophien Kamoun and Sylvestre Marillonnet through the Addgene database. The work described here has not been previously published, except in the form of a chapter in the publicly defended PhD thesis of Arjan van Zeijl (van Zeijl, 2017). The publication of this thesis has occurred in accordance with the policy of Wageningen University and Research, Wageningen, the Netherlands.

## Supplemental Information

**Supplemental Table 1.** List of Golden Gate<sup>a</sup> constructs used in this study.

Construct	Description	Level	Backbone	Contains <sup>b</sup>	Literature
1	<i>nptII</i> resistance cassette	1	pICH47802	pICSL70004: <i>nptII</i>	-
2	35S <sub>pro</sub> :ΩNLS-Cas9:35S <sub>ter</sub>	1	pICH47742	pICH41388:35S <sub>pro</sub> , pAGM5331:ΩNLS, pICH41308::aCas9, pICH41414:35S <sub>ter</sub>	Fauser et al. (2014)
3	PanEIN2sgRNA1	1	pICH47751	pICSL01009:AtU6p, corresponding PCR amplicon	Nekrasov et al. (2013)
4	PanNSP1sgRNA1	1	pICH47751	pICSL01009:AtU6p, corresponding PCR amplicon	Nekrasov et al. (2013)
5	PanNSP2sgRNA1	1	pICH47751	pICSL01009:AtU6p, corresponding PCR amplicon	Nekrasov et al. (2013)
6	PanNSP2sgRNA2	1	pICH47761	pICSL01009:AtU6p, corresponding PCR amplicon	Nekrasov et al. (2013)
7	PanNSP2sgRNA3	1	pICH47841	pICSL01009:AtU6p, corresponding PCR amplicon	Nekrasov et al. (2013)
8	CRISPR_ctrl	2	pICSL4723	1R: construct 1, 2F: construct 2, end-link pICH41744	-
9	CRISPR_PanEIN2	2	pICSL4723	1R: construct 1, 2F: construct 2, 3F: construct 3, end-link pICH41766	-
10	CRISPR_PanNSP1	2	pICSL4723	1R: construct 1, 2F: construct 2, 3F: construct 4, end-link pICH41766	-
11	CRISPR_PanNSP2	2	pICSL4723	1R: construct 1, 2F: construct 2, 3F: construct 5, 4F: construct 6, 5R: construct 7, end-link pICH41800	-

<sup>a</sup>Backbones and standard parts used for Golden Gate assembly are from Engler et al. (2014).

<sup>b</sup>1F, 2R, etc. indicate position and orientation in level 2 constructs. Construct followed by a number refers to constructs build during this study and described in this table.

**Supplemental Table 2.** Primers used in this study.

Name	Purpose	Sequence
PanEIN2_sgRNA1	Clone sgRNA	tgtggtctcaattGTGTTAATCCTGGA AAATGGGgttttagagctagaaatagcaag
PanNSP1_sgRNA1	Clone sgRNA	tgtggtctcaattGCCAGTCCATGATG TGATCCGgttttagagctagaaatagcaag
PanNSP2_sgRNA1	Clone sgRNA	tgtggtctcaattGAAGACGACCATGG CTGCGCGgttttagagctagaaatagcaag
PanNSP2_sgRNA2	Clone sgRNA	tgtggtctcaattGACTGGAACGTCC TTACCGGgttttagagctagaaatagcaag
PanNSP2_sgRNA3	Clone sgRNA	tgtggtctcaattGTAGTTGCACACCG CCTACCGgttttagagctagaaatagcaag
sgRNA_Rv	Clone sgRNA	tgtggtctca AGCGTAATGCCAACTTTGTAC
geno_Cas9_Fw	Amplify aCas9	TTCGATCTCGCTGAGGATGC
geno_Cas9_Rv	Amplify aCas9	TAGCGAGAGGTCCCACGTAG
geno_PanEIN2-crispr-Fw	Genotyping CRISPR mutants	CATTGCAACAGCCTGTGGAC
geno_PanEIN2-crispr-Rv	Genotyping CRISPR mutants	CTGAGCAAGACCCCTTCCAG
geno_PanHK4-crispr-Fw	Genotyping CRISPR mutants	ATGCAGAACCACCATCCGT
geno_PanHK4-crispr-Rv	Genotyping CRISPR mutants	GCATACTCATCCCGACAGG
geno_PanNSP1-crispr-Fw	Genotyping CRISPR mutants	TGGTTTCTCGTGGCCTTTGT
geno_PanNSP1-crispr-Rv	Genotyping CRISPR mutants	CTGTGGCCTTAGCTGAGCTT
geno_PanNSP2-crispr-Fw	Genotyping CRISPR mutants	ACTTCCACTCTGTCCCCGAA
geno_PanNSP2-crispr-Rv	Genotyping CRISPR mutants	GACCGGTGACTGAAGTGAGG
qPanD27_Fw	qRT-PCR	TCGGATCGCCATTCAACATC
qPanD27_Rv	qRT-PCR	GACAAGTTCCCCTGTTTTG
qPanMAX1_Fw	qRT-PCR	TGATGAGATCGTGGCCAAGAG
qPanMAX1_Rv	qRT-PCR	AAAAACGTTCTCGCAACCG
qPanUNK2_Fw	qRT-PCR	TGCCATTGGTGTTAGCTGTG
qPanUNK2_Rv	qRT-PCR	GTGTCTATCACTGCCTCTTTGC
qPanEF1α_Fw	qRT-PCR	AGACAAGGTTAAGCGTGCAG
qPanEF1α_Rv	qRT-PCR	TGCAACTGGGCAACAACTC
qTCN-1_Fw	T-DNA copy number	AAGCGCGTTACAAGAAAGCC
qTCN-1_Rv	T-DNA copy number	ACGTTGCCCCGCATAATTACG
qTCN-3_Fw	T-DNA copy number	TGCACGACCACGCATTAATG
qTCN-3_Rv	T-DNA copy number	ACGATGCCATGTTTCATCTGC
PanAGT1_Fw	T-DNA copy number	AGCGCTTGATCTCCTTTTCG
PanAGT1_Rv	T-DNA copy number	TGCTTTACCTTGTGGCAGTC
Pan338920_Fw	T-DNA copy number	TGAAGCACAAACACGATGG

Pan338920_Rv	T-DNA copy number	TGCTTTCAGCCAGTTAACCG
npII_Fw	Southern blotting	CACAACAGACAATCGGCTGC
npII_Rv	Southern blotting	TGATATTCGGCAAGCAGGCA

**Supplemental Table 3.** Composition of tissue culture media.

Component	Medium		
	SH10	Propagation	Root-inducing
Schenk & Hildebrandt basal salt medium <sup>a,b</sup>	1x	1x	1x
Schenk & Hildebrandt vitamin mixture <sup>a,b</sup>	1x	1x	1x
Sucrose <sup>a</sup>	1% (w/v)	1% (w/v)	2% (w/v)
BAP (6-Benzylaminopurine) <sup>c</sup>	-	1 mg/l	-
IBA (Indole-3-butyric acid) <sup>c</sup>	-	0.1 mg/l	1 mg/l
NAA (1-Naphthalene-acetic acid) <sup>c</sup>	-	-	0.1 mg/l
MES (2-ethanesulfonic acid) <sup>c</sup>	3 mM, pH = 5.8	3 mM, pH = 5.8	3 mM, pH = 5.8
Daishin agar <sup>a</sup>	-	0.8% (w/v)	0.8% (w/v)

<sup>a</sup> Duchefa Biochemie, Haarlem, The Netherlands

<sup>b</sup> Schenk and Hildebrandt, 1972

<sup>c</sup> Sigma, St. Louis, USA

**Supplemental Table 4.** Regeneration efficiency of different explant types. Regeneration was determined for non-transgenic material incubated on propagation medium.

Explant type	Regeneration efficiency
Young leaves	-
Mature leaves	+/-
Midveins	+/-
Stem pieces (5 mm)	++
Stem pieces (1 mm)	+/-
Stem pieces, cut length-wise	+
Petioles	+++
Shoot apical meristems	-
Tissue culture shoots	+

+++ , means that > 90% of explants develop multiple regenerative calli.

- , indicates that no regeneration was observed, tissue turned dark brown in color.

**Supplemental Table 5.** Hormone composition of tissue culture media used in this study.

Hormone	Propagation medium	Root-inducing medium	Alternative media								
			1	2	3	4	5	6	7	8	9
BAP (mg/L)	1	-	0.33	1	3	0.33	1	3	0.33	1	3
IBA (mg/L)	0.1	1	-	-	-	-	-	-	-	-	-
NAA(mg/L)	-	0.1	0.03	0.03	0.03	0.1	0.1	0.1	0.3	0.3	0.3

**Supplemental Table 6.** Regeneration of transgenic material after co-cultivation and continuous culturing on propagation medium (Prop) or co-cultivation and culturing for seven additional days on root-inducing medium and subsequent transfer to propagation medium (Root -> prop). Regeneration was scored 6 weeks after start of transformation.

	Prop	Root -> prop
Petioles	+	++
Stem pieces (5 mm)	+/-	++

++, means that > 50% of explants develop regenerative calli.

-, means that < 20% of explants develop regenerative calli.

**Supplemental Table 7.** Molecular characterization of transgenic lines.

Line <sup>a</sup>	T-DNA copy number estimate		Ploidy	Presence of right border sequences <sup>d</sup>
	qRT-PCR <sup>b</sup>	Southern blot <sup>c</sup>		
1	2.9	na	Diploid	+
3*	1.0	1	Diploid	+
4	1.2	na	Diploid	+
12	0.9	na	Diploid	+
13	1.0	2	Diploid	+
18	1.9	na	Diploid	+
19	2.9	3	Diploid	+
21	1.1	na	Diploid	+
31	1.9	2	Diploid	+
33*	1.0	1	Diploid	+
37	2.1	na	Diploid	+
43*	0.9	1	Diploid	+
47	2.8	na	Diploid	+
49	1.1	na	Diploid	+
50	nd	na	Diploid	-
52	1.0	1	Diploid	+
54	0.5	2	Diploid	+
55	0.8	na	Diploid	+
59	0.4	na	Diploid	+
61	0.9	na	Diploid	+

<sup>a</sup> Asterisks indicate lines selected for further analyses.

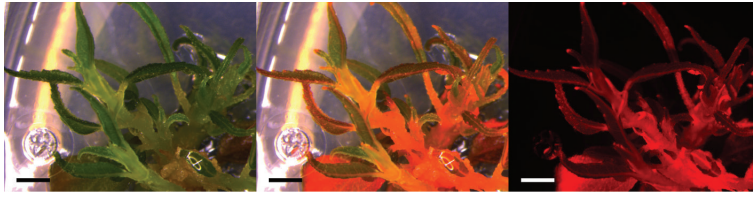
<sup>b</sup> These numbers indicate measurement values from qRT-PCR on genomic DNA. nd indicates that T-DNA was not detected.

<sup>c</sup> na indicates samples not analyzed by Southern blot.

<sup>d</sup> PCR was performed using primers amplifying a sequence close to the T-DNA right border. +/- indicates presence/absence of a PCR amplicon.

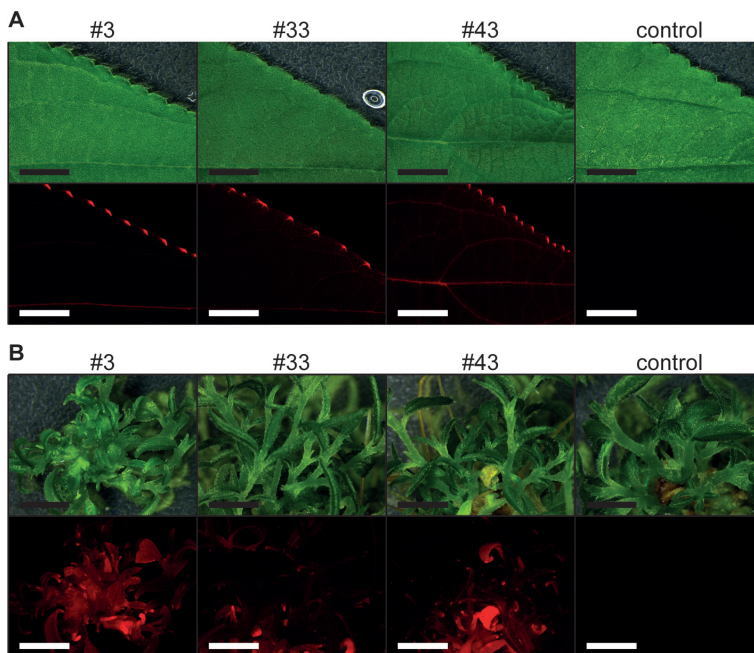
**Supplemental Table 8.** GeneIDs for all *P. andersonii* genes used in this study. GeneIDs refer to *P. andersonii* gene models, which can be searched for on [www.parasponia.org](http://www.parasponia.org).

Name	GeneID
PanEIN2	PanWU01x14_asm01_ann01_090380
PanHK2	PanWU01x14_asm01_ann01_002510
PanHK3	PanWU01x14_asm01_ann01_222240
PanHK4	PanWU01x14_asm01_ann01_103390
PanNSP1	PanWU01x14_asm01_ann01_334190
PanSCL26	PanWU01x14_asm01_ann01_236490
PanSCL34	PanWU01x14_asm01_ann01_107700
PanSCL16	PanWU01x14_asm01_ann01_130270
PanSHR	PanWU01x14_asm01_ann01_108080
PanNSP2	PanWU01x14_asm01_ann01_157260
PanSCL28	PanWU01x14_asm01_ann01_101430
PanSCL18	PanWU01x14_asm01_ann01_295370
PanSCL17	PanWU01x14_asm01_ann01_245480
PanSCL27	PanWU01x14_asm01_ann01_245490
PanRAM1	PanWU01x14_asm01_ann01_146390
PanD27	PanWU01x14_asm01_ann01_194930
PanD27L1	PanWU01x14_asm01_ann01_057040
PanD27L2	PanWU01x14_asm01_ann01_143920
PanMAX1	PanWU01x14_asm01_ann01_182160
PanUNK2	PanWU01x14_asm01_ann01_211960
PanEF1 $\alpha$	PanWU01x14_asm01_ann01_070680
PanAGT1	PanWU01x14_asm01_ann01_040540
-	PanWU01x14_asm01_ann01_338920



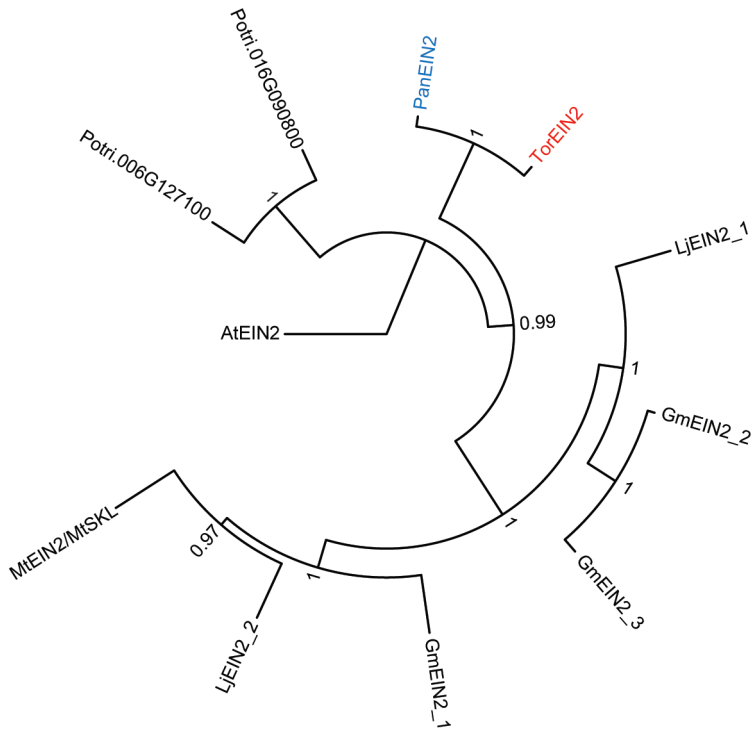
**Supplemental Figure 1. *P. andersonii* transgenic shoots can be effectively propagated.**

Transgenic *P. andersonii* shoots propagated *in vitro*. Images were taken six months after transformation. Scale bars are equal to 2.5 mm. Shown from left to right are bright-field images, overlays of bright-field and DsRED fluorescence and DsRED fluorescence images.



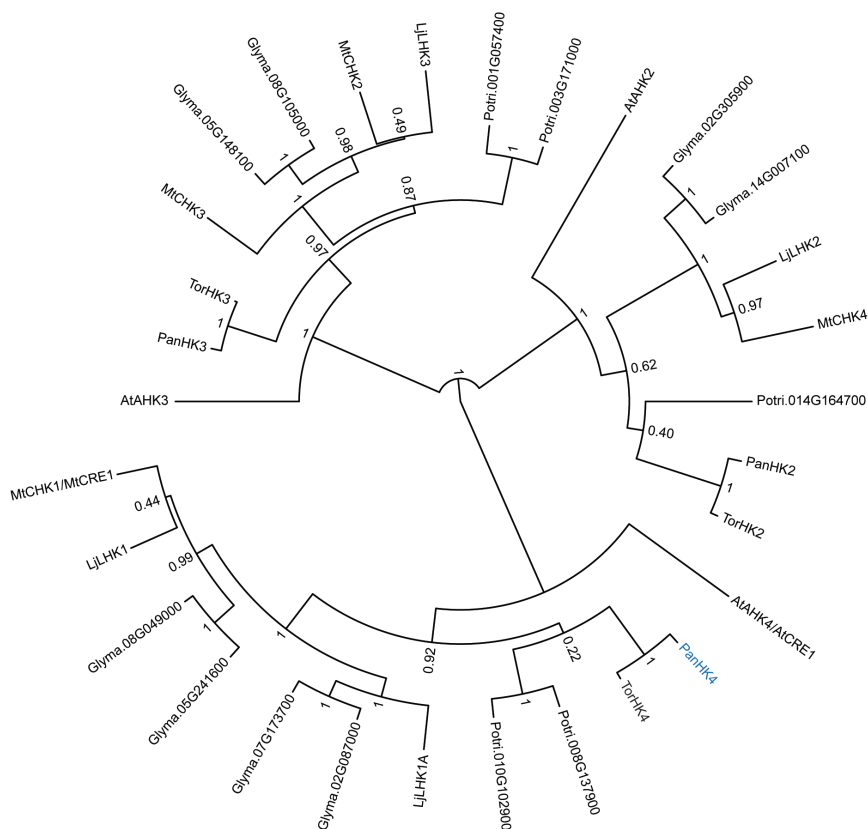
**Supplemental Figure 2. *Trans*-genes remain stably integrated in the *P. andersonii* genome.**

Bright field (top) and DsRED1 fluorescence (bottom) images of leaves harvested from mature trees grown under greenhouse conditions (**A**) or *in vitro* propagated material (**B**). Images are shown for transgenic lines 3, 33 and 43 and control plants not expressing the DsRED1 marker. Images were taken one year after transgenic lines were selected. Scale bars are equal to 5 mm.



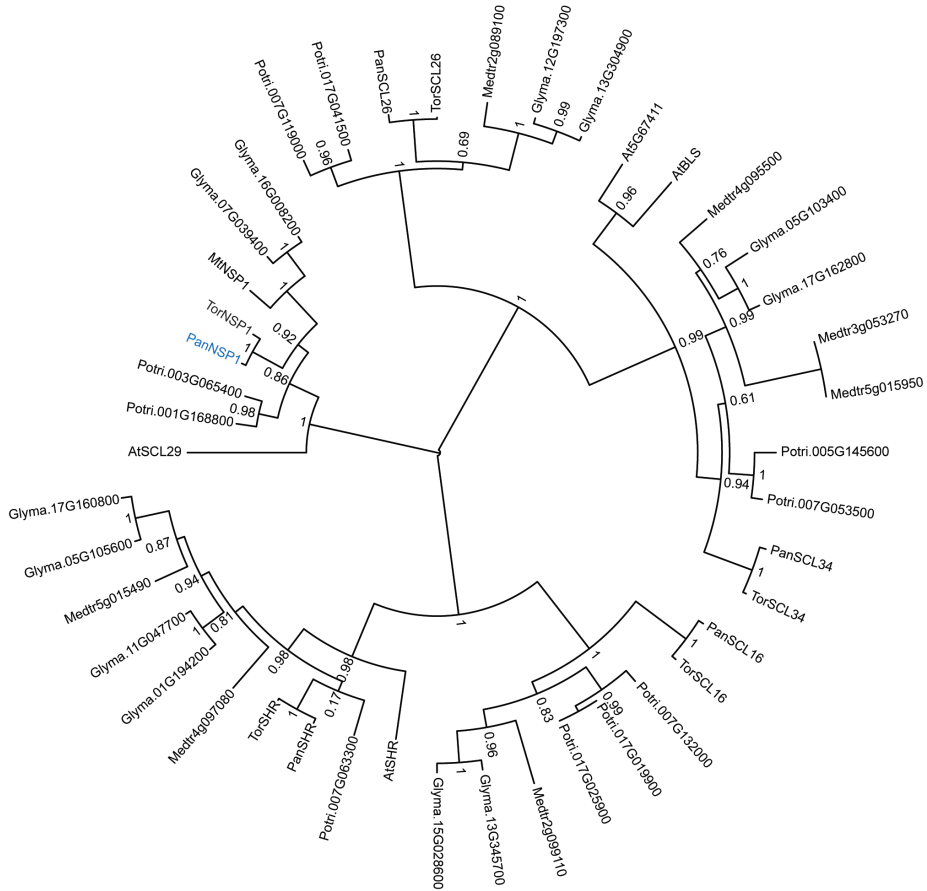
### Supplemental Figure 3. Phylogeny of EIN2 proteins.

Phylogeny was reconstructed based on an alignment of EIN2 proteins from *Arabidopsis thaliana* (At), soybean (*Glycine max*, Gm), *Lotus japonicus* (Lj), *Medicago truncatula* (Mt), poplar (*Populus trichocarpa*, Potri), *Parasponia andersonii* (Pan) and *Trema orientalis* (Tor). Branch support is indicated by FastTree support values (Price et al., 2009). Terminals are labeled by their gene name or gene identifier. The EIN2 protein of *P. andersonii* is highlighted in blue. Mid-point rooting was applied for better tree visualization.



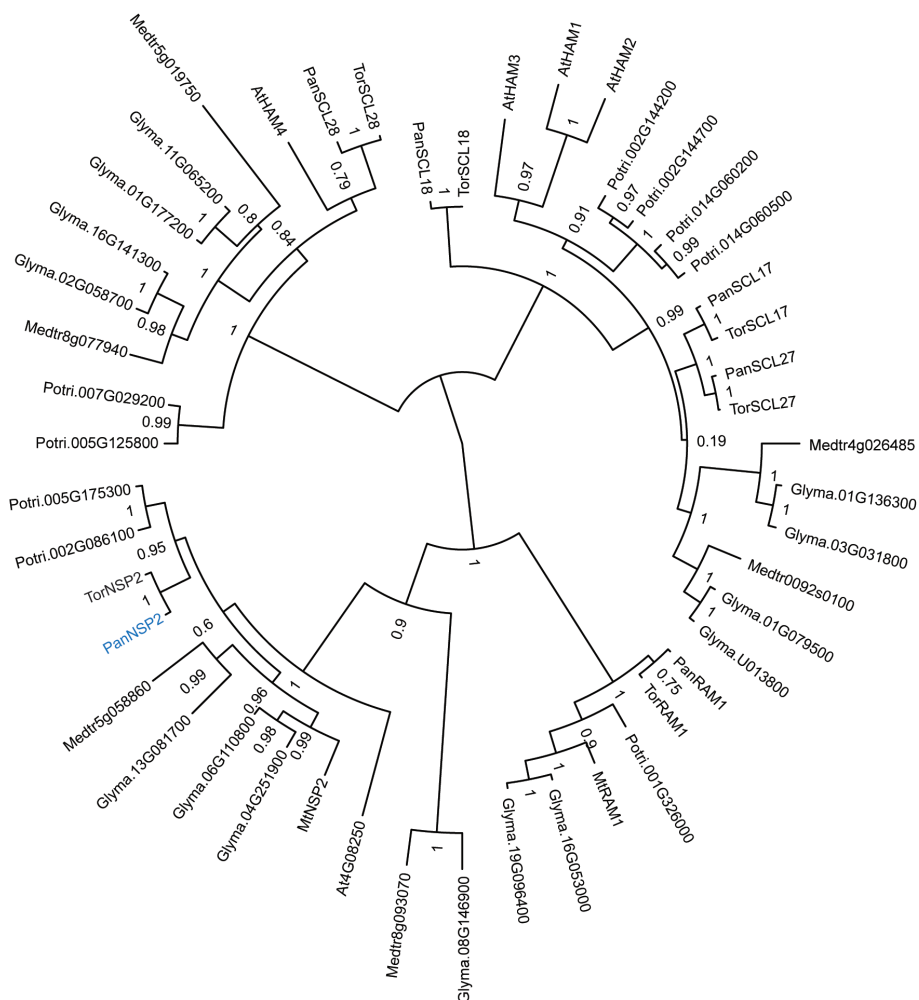
#### Supplemental Figure 4. Phylogeny of the histidine kinase cytokinin receptor family.

Phylogeny was reconstructed based on an alignment of cytokinin receptor proteins from *Arabidopsis thaliana* (At), soybean (*Glycine max*, Gm), *Lotus japonicus* (Lj), *Medicago truncatula* (Mt), poplar (*Populus trichocarpa*, Potri), *Parasponia andersonii* (Pan) and *Trema orientalis* (Tor). Branch support is indicated by FastTree support values (Price et al., 2009). Terminals are labeled by their gene name or gene identifier. The MtCRE1/LjLHK1 putative orthologue of *P. andersonii* is highlighted in blue. Mid-point rooting was applied for better tree visualization.

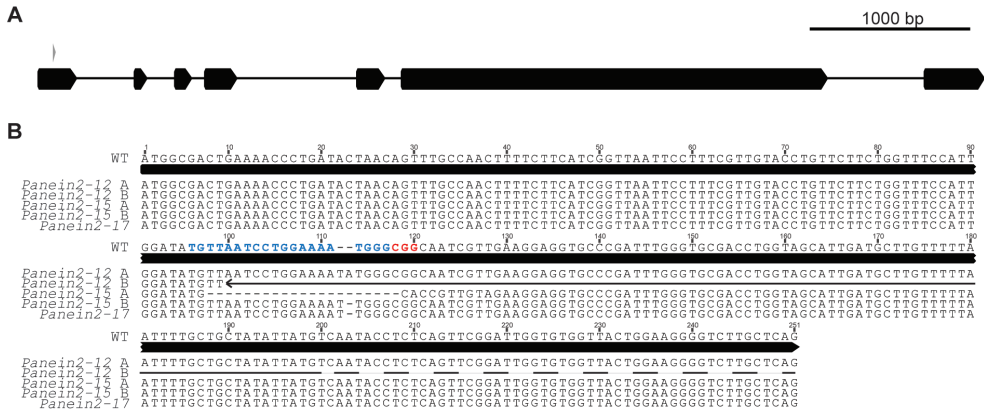


### Supplemental Figure 5. Phylogeny of NSP1 and related GRAS proteins.

Phylogeny was reconstructed based on an alignment of GRAS proteins from *Arabidopsis thaliana* (At), soybean (*Glycine max*, Gm), *Medicago truncatula* (Mt), poplar (*Populus trichocarpa*, Potri), *Parasponia andersonii* (Pan) and *Trema orientalis* (Tor). Branch support is indicated by FastTree support values (Price et al., 2009). Terminals are labeled by their gene name or gene identifier. The MtNSP1 putative orthologue of *P. andersonii* is highlighted in blue. Mid-point rooting was applied for better tree visualization.



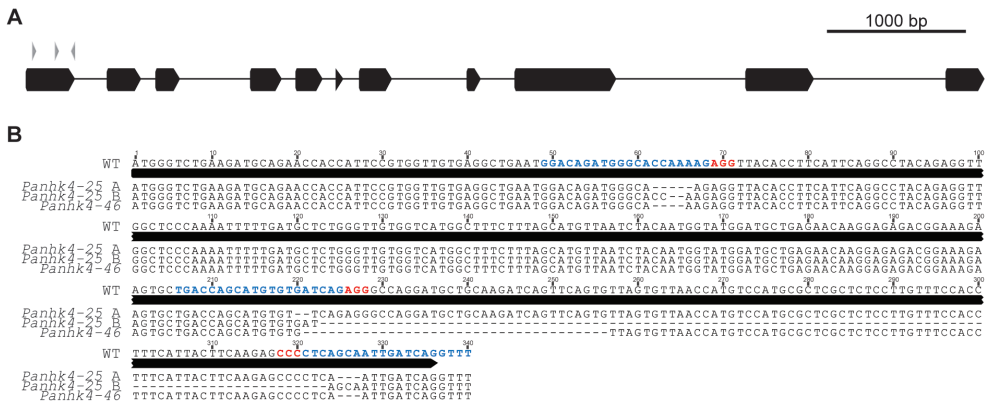
Phylogeny was reconstructed based on an alignment of GRAS proteins from *Arabidopsis thaliana* (At), soybean (*Glycine max*, Gm), *Medicago truncatula* (Mt), poplar (*Populus trichocarpa*, Potri), *Parasponia andersonii* (Pan) and *Trema orientalis* (Tor). Branch support is indicated by FastTree support values (Price et al., 2009). Terminals are labeled by their gene name or gene identifier. The MtNSP2 putative orthologue of *P. andersonii* is highlighted in blue. Mid-point rooting was applied for better tree visualization.



### Supplemental Figures 7. CRISPR mutant alleles of *Panein2* mutant lines.

(A) Schematic representation of *PanEIN2* gene model. Indicated by a grey arrowhead is the location of the sgRNA target site.

(B) Sequence alignment of the first exon of *PanEIN2* in wild type (WT) and *Panein2* mutant lines. For bi-allelic mutant lines, both alleles are shown (A and B alleles). Highlighted in blue and red are the sgRNA target site and PAM sequence, respectively. Arrow in sequence indicates a large (> 100 bp) insertion in *Panein2-12* allele B.



### Supplemental Figures 8. CRISPR mutant alleles of *Panhk4* mutant lines.

(A) Schematic representation of *PanHK4* gene model. Indicated by grey arrowheads are the locations of three sgRNA target sites.

(B) Sequence alignment of the first exon of *PanHK4* in wild type (WT) and *Panhk4* mutant lines. For bi-allelic mutant lines, both alleles are shown (A and B alleles). Highlighted in blue and red are the sgRNA target sites and PAM sequences, respectively.

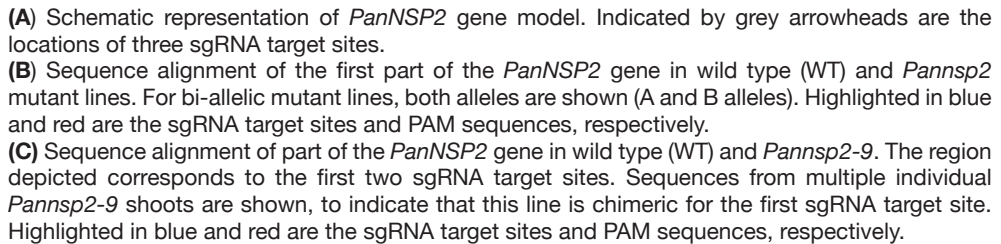


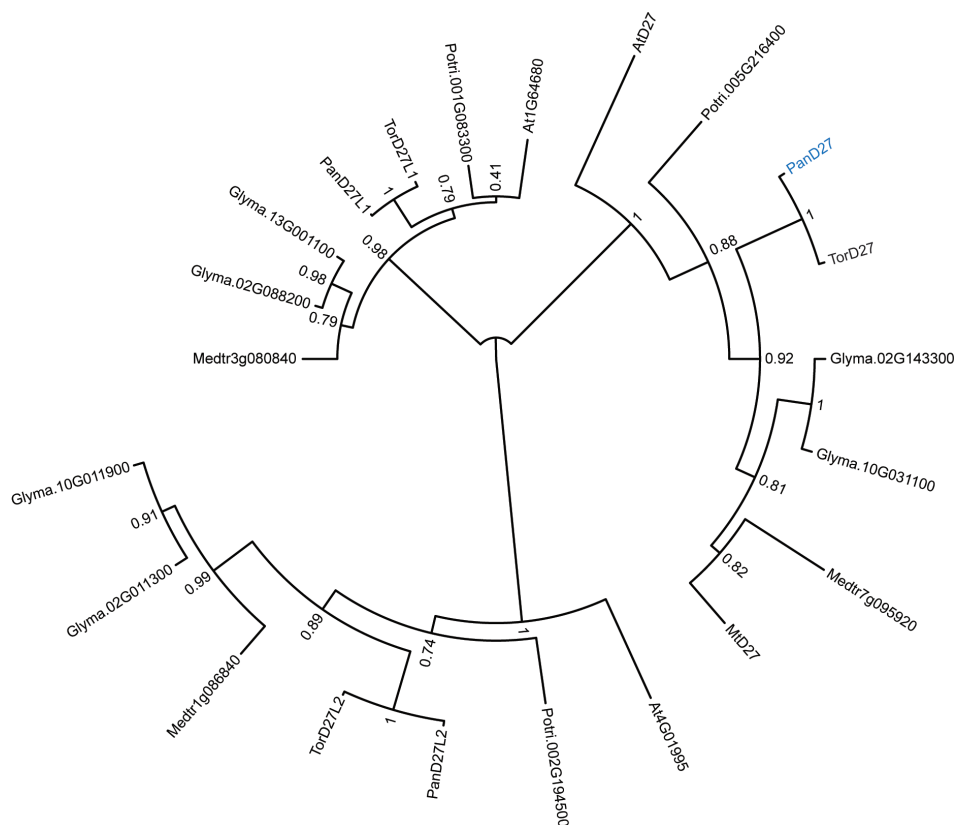
**Supplemental Figures 9. CRISPR mutant alleles of *Pannsp1* mutant lines.**

**(A)** Schematic representation of *PanNSP1* gene model. Indicated by a grey arrowhead is the locations of the sgRNA target site.

**(B)** Sequence alignment of the first part of *PanNSP1* in wild type (WT) and *Pannsp1* mutant lines. Highlighted in blue and red are the sgRNA target site and PAM sequence, respectively. Highlighted in green are in-frame ATGs present in *Pannsp1-6* and *Pannsp1-13*.

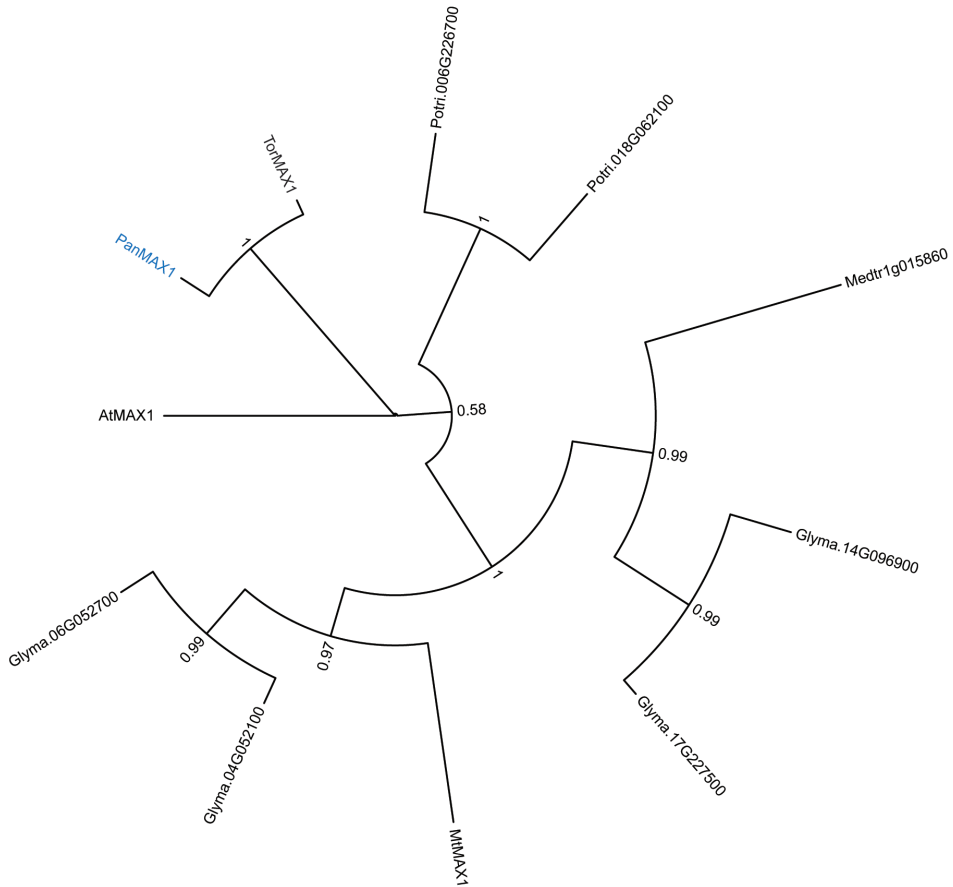
**Supplemental Figures 10. CRISPR mutant alleles of *Pannsp2* mutant lines.**





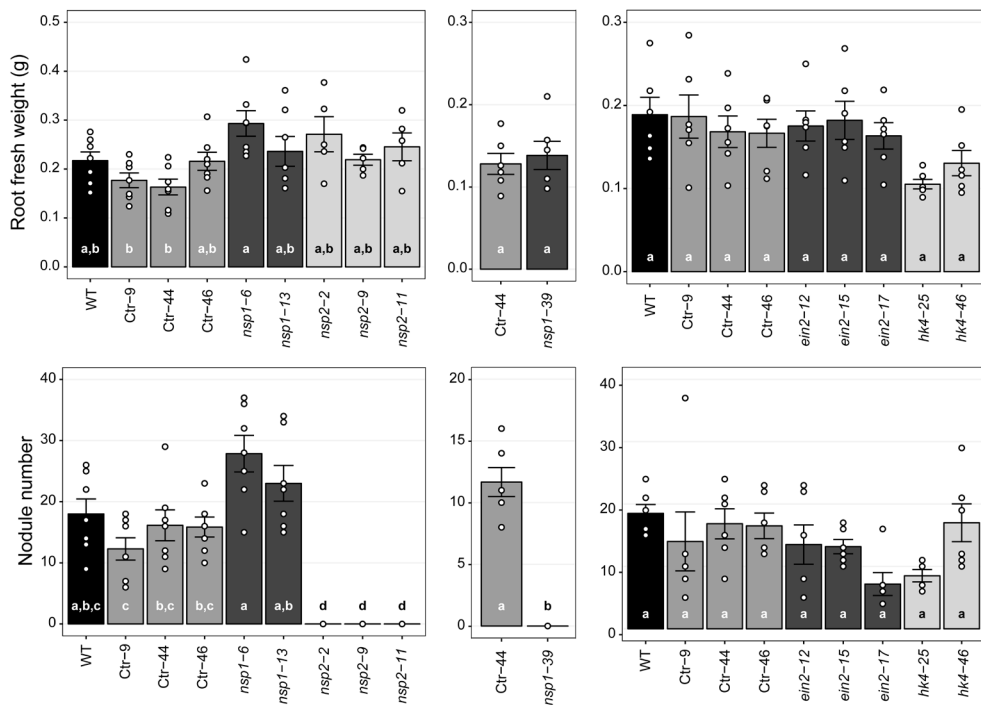
### Supplemental Figures 11. Phylogeny of D27 and D27-like proteins.

Phylogeny was reconstructed based on an alignment of D27 and D27-like proteins from *Arabidopsis thaliana* (At), soybean (*Glycine max*, Gm), *Medicago truncatula* (Mt), poplar (*Populus trichocarpa*, Potri), *Parasponia andersonii* (Pan) and *Trema orientalis* (Tor). Branch support is indicated by FastTree support values (Price et al., 2009). Terminals are labeled by their gene name or gene identifier. The *P. andersonii* putative orthologue of MtD27 is highlighted in blue. Mid-point rooting was applied for better tree visualization.



### Supplemental Figures 12. Phylogeny of MAX1 proteins.

Phylogeny was reconstructed based on an alignment of MAX1 proteins from *Arabidopsis thaliana* (At), soybean (*Glycine max*, Gm), *Medicago truncatula* (Mt), poplar (*Populus trichocarpa*, Potri), *Parasponia andersonii* (Pan) and *Trema orientalis* (Tor). Branch support is indicated by FastTree support values (Price et al., 2009). Terminals are labeled by their gene name or gene identifier. The *P. andersonii* putative orthologue of MtMAX1 is highlighted in blue. Mid-point rooting was applied for better tree visualization.



**Supplemental Figure 13. Nodule formation on *P. andersonii* CRISPR/Cas9 mutant lines.**

Nodule number and fresh root weight data belonging to **Figure 6**. Nodule number and root fresh weight was determined at one-month post inoculation with *Mesorhizobium plurifarium* BOR2. Data represent means of 5-7 biological replicates  $\pm$  SEM. Dots represent measurement values of biological repeats. Different letters indicate statistical significance ( $p < 0.05$ ) as determined by ANOVA in combination with Tukey post-hoc test. WT, wild type.





# Chapter 5

## **The *Parasponia andersonii* Cytokinin Receptor *HISTIDINE KINASE3* is Required for Intracellular *Rhizobium* Infection of Nodule Cells**

**Titis A. K. Wardhani<sup>1</sup>, Arjan van Zeijl<sup>1</sup>, Sidney Linders<sup>1</sup>, Kerstin Gühl<sup>1</sup>,  
Ton Bisseling<sup>1</sup>, Wouter Kohlen<sup>1</sup> and René Geurts<sup>1</sup>**

<sup>1</sup>Department of Plant Sciences, Cluster of Plant Developmental Biology, Wageningen University, Droevendaalsesteeg 1, 6708 PB, Wageningen, The Netherlands.

***Manuscript in preparation***

*“For me, becoming isn’t about arriving somewhere or achieving a certain aim. I see it instead as forward motion, a means of evolving, a way to reach continuously toward a better self. The journey doesn’t end”.*  
—Michelle Obama, in her book *“Becoming”*





## Abstract

The nitrogen-fixing nodulation trait occurs in ten lineages of the orders Fabales, Fagales, Cucurbitales and Rosales. A unique feature of this mutualistic interaction between plant and diazotrophic rhizobium or *Frankia* bacteria is the uptake of bacteria into nodule cells. During nodule development, plant cells experience a short developmental window in which they are promiscuous for such intracellular infection. Based on studies in legumes, it is speculated that cytokinin signalling may play a pivotal role in the developmental fate of these susceptible cells. However, due to the multiple roles of cytokinin in nodulation, clear evidence of this hypothesis is lacking. We used the Cannabaceae species *Parasponia andersonii* to investigate the role of cytokinin signalling in a nodulating non-legume and found that the *HISTIDINE KINASE3* (*PanHK3*) cytokinin receptor is essential for infectability of nodule cells. Like legumes, *Parasponia* species can be nodulated by rhizobium. By exploiting the two-component signalling reporter TCSn, we confirm that like in legumes, rhizobium induces cytokinin signalling in epidermal and outer cortical cell layers of *P. andersonii*. This spatiotemporal expression pattern is mirrored by *PanHK3*, but not by the two other homologous genes *PanHK2* and *PanHK4*. Subsequent CRISPR-Cas9 mutant analysis revealed that *Panhk3* knockout plants form nodules devoid of intracellular rhizobium infection. This demonstrates that a cytokinin signalling receptor commits an essential role in the infectability of *P. andersonii* nodule cells. As the nodulation trait in Fabales, Fagales, Cucurbitales and Rosales has a shared evolutionary origin, we hypothesize a conserved role of cytokinin signalling in intracellular nodule infection.

## Introduction

Some plants of the taxonomic orders Fabales, Fagales, Cucurbitales and Rosales, can establish a nitrogen-fixing nodule symbiosis with either rhizobium or filamentous *Frankia* bacteria. In this mutualistic interaction, nodules serve as niche organs, accommodating the diazotrophic bacteria intracellularly. Here, bacteria find the appropriate physiological conditions to convert atmospheric dinitrogen ( $N_2$ ) into ammonia, a reaction that is fuelled by carbohydrates obtained from the plant (Udvardi and Poole, 2013). In return, the plant receives ammonia. Comparative genomics revealed a shared evolutionary origin of nodulation in these four taxonomic orders, which are collectively known as the nitrogen-fixing clade (Griesmann et al., 2018; van Velzen et al., 2018; van Velzen et al., 2019). A major objective in symbiosis research is to identify the key evolutionary events that led to nodulation, and allowed diazotrophic bacteria to infect plant cells intracellularly.

The uptake of bacteria into living plant cells is considered as a distinct feature of the nitrogen-fixing nodulation trait (Parniske, 2018). It is argued that bacterial-induced cell divisions are a prerequisite for such intracellular hosting, which may explain why nitrogen-fixing endosymbioses with rhizobium or *Frankia* require the formation of nodules (Geurts et al., 2016). Cells that emerge from an active nodule meristem experience a short developmental window in which they are promiscuous for intracellular infection. This suggests that the fate of these nodule cells is distinct when compared to for example, cells that emerge from a lateral root meristem. However, the underlying signals that control the specific fate of 'infectable' cells remain elusive.

Physiological, molecular and genetic studies in the legume models *Lotus japonicus* and *Medicago truncatula* (Fabaceae, Fabales) uncovered an important role of cytokinin in rhizobium-induced nodule organogenesis. Rhizobium secreted signal molecules namely lipo-chitooligosaccharide (LCO), which is essential for nodule formation and infection, can also induce the expression of cytokinin biosynthesis genes (van Zeijl et al., 2015a; Reid et al., 2016; Reid et al., 2017). Significant transcriptional changes downstream of LCO perception are dependent on the presence of a functional cytokinin receptor *CYTOKININ RESPONSE1* (*MtCRE1*) in *M. truncatula* roots (van Zeijl et al., 2015a; Jardinaud et al., 2016; Schiessl et al., 2019). The importance of this *MtCRE1* in nodulation, as well as its ortholog *LOTUS HISTIDINE KINASE1* (*LjLHK1*) in *L. japonicus*, is shown in three complementary experiments. (i) *L. japonicus* *Ljlhk1* and *M. truncatula* *Mtcre1* knockout mutants are affected in nodule organogenesis (Murray et al., 2007; Plet et al., 2011). (ii) A gain-of-function mutation in the region encodes extracellular ligand binding CHASE domain of *LjLHK1*/*MtCRE1* triggers spontaneous nodule formation in absence of bacteria (Tirichine et al., 2006; Tirichine et al., 2007; Hayashi et al., 2010;

Ovchinnikova et al., 2011). (iii) Exogenously applied cytokinin, or cytokinin analogs secreted by microbes, showed to induce nodule organogenesis in a series of nodulating legume species, a response that is dependent on *LjLHK1/MtCRE1* in *L. japonicus* and *M. truncatula* (Relic et al., 1993; Cooper and Long, 1994; Bauer et al., 1996; Mathesius et al., 2000; Heckmann et al., 2011; Gauthier-Coles et al., 2019). Functioning of *LjLHK1/MtCRE1* in nodule organogenesis requires several symbiotic transcription factors; among them is NODULE INCEPTION (NIN) (Gonzalez-Rizzo et al., 2006; Tirichine et al., 2007; Hayashi et al., 2010; Plet et al., 2011). In legumes, *NIN* is LCO and cytokinin responsive, and expression is observed in the epidermis as well as in cortical cells that give rise to the nodule primordium (Gonzalez-Rizzo et al., 2006; Heckmann et al., 2011; Plet et al., 2011; Vernié et al., 2015; Liu et al., 2019c). Legume *nin* knockout mutants are blocked in rhizobium intracellular infection and nodule organogenesis (Schauser et al., 1999; Borisov et al., 2003; Marsh et al., 2007). The latter response requires a ~500 bp conserved *cis*-regulatory promoter element, which contains several putative cytokinin B-type RESPONSE REGULATOR (RR) binding sites (Liu et al., 2019c). This suggests that in legumes pericycle *NIN* expression is cytokinin dependent.

Studies in *Arabidopsis thaliana* showed that cytokinins promote differentiation of cells derived from root meristems. Cytokinin signalling activates B-type RRs post-transcriptionally, which negatively regulate auxin signalling, as well as promote the onset of endoreduplication (Dello Ioio et al., 2008; Takahashi and Umeda, 2014). Such a direct effect on cell fate determination by cytokinins in root cells, together with the observation of active cytokinin signalling in nodule primordia, and in the infection zone of legume nodules (Lohar et al., 2004; Held et al., 2014; Reid et al., 2017), make cytokinin a potential signal to play a role in achieving the cell identity required for infectability of nodule cells (Miri et al., 2016). However, data obtained in legumes that support this hypothesis were limited and, to some extent, contradicting. An *M. truncatula* knockout mutant in the B-type RR *MtRRB3* shows a reduced nodulation efficiency, as well as a decrease in the number of intracellular infection threads (Tan et al., 2020). In contrast, an *L. japonicus* knockout mutant in *LjLHK1* showed to be hyper-infected in epidermal root hair cells, suggesting a negative role of cytokinin in epidermal infections (Murray et al., 2007; Held et al., 2014). Several factors may hamper studies on the role of cytokinin signalling in rhizobium intracellular infection. Among others these include; gene redundancy in cytokinin signalling genes (Op den Camp et al., 2011b; Held et al., 2014; Tan et al., 2019; Tan et al., 2020), the entanglement of nodule organogenesis and intracellular cortical infection, two processes that are dependent on the cell cycle machinery (Yang et al., 1994), and the complex regulation of root hair-based infection, which may require different signalling pathways when compared to cortical infection (Miri et al., 2016).

Several lines of evidence indicate that cytokinin signalling plays a less prominent role in nodulation of non-legumes. Studies in the actinorhizal plants *Alnus glutinosa* (Betulaceae, Fagales), *Casuarina glauca* (Casuarinaceae, Fagales) and *Datisca glomerata* (Datisceae, Cucurbitales) showed that exogenously applied cytokinin does not trigger nodule organogenesis (Gauthier-Coles et al., 2019). Like in legumes, *NIN* was found to be essential for nodule organogenesis in *C. glauca* and *Parasponia andersonii* (Cannabaceae, Rosales) (Clavijo et al., 2015; Bu et al., 2020). However, there were no reports that *NIN* is cytokinin-responsive in these nodulating non-legumes, nor has it been reported that nodulating non-legumes possess the putative cytokinin responsive *cis*-regulatory element that is conserved in the legume *NIN* promoter region (Clavijo et al., 2015; Liu et al., 2019c; Bu et al., 2020). Lastly, a *P. andersonii* knockout mutant in the *LjLHK1/MtCRE1* orthologous cytokinin receptor *PanHK4* has no obvious symbiotic phenotype (van Zeijl et al., 2018). Nevertheless, studies in *D. glomerata* showed higher cytokinin levels in root nodules when compared to roots, suggesting a role of cytokinin in non-legume nodules (Demina et al., 2019).

To obtain evidence on the role of cytokinin signalling in a nodulating non-legume, we used *P. andersonii*, since a fast and efficient transformation protocol has been established for this species (van Zeijl et al., 2018; Wardhani et al., 2019). *Parasponia* represents five tropical tree species, and are the only non-legumes known to nodulate with rhizobia. Like in legumes, nodulation in *Parasponia* is driven by LCO signalling, though its nodule ontogeny and mode of rhizobium infection is different (Lancelle and Torrey, 1984; Lancelle and Torrey, 1985; Op den Camp et al., 2011a; Rutten et al., 2020; Shen et al., 2020). By spatiotemporal analysis of the cytokinin two-component signalling reporter *TCSn* and *PanHK* promoter reporter constructs, we showed that rhizobium-induced activation of cytokinin signalling in the root epidermis and nodule primordia mirrors the transcriptional regulation of *PanHK3*. Subsequent CRISPR-Cas9 mutagenesis revealed a role of this cytokinin receptor in intracellular infection by rhizobium.

## Results

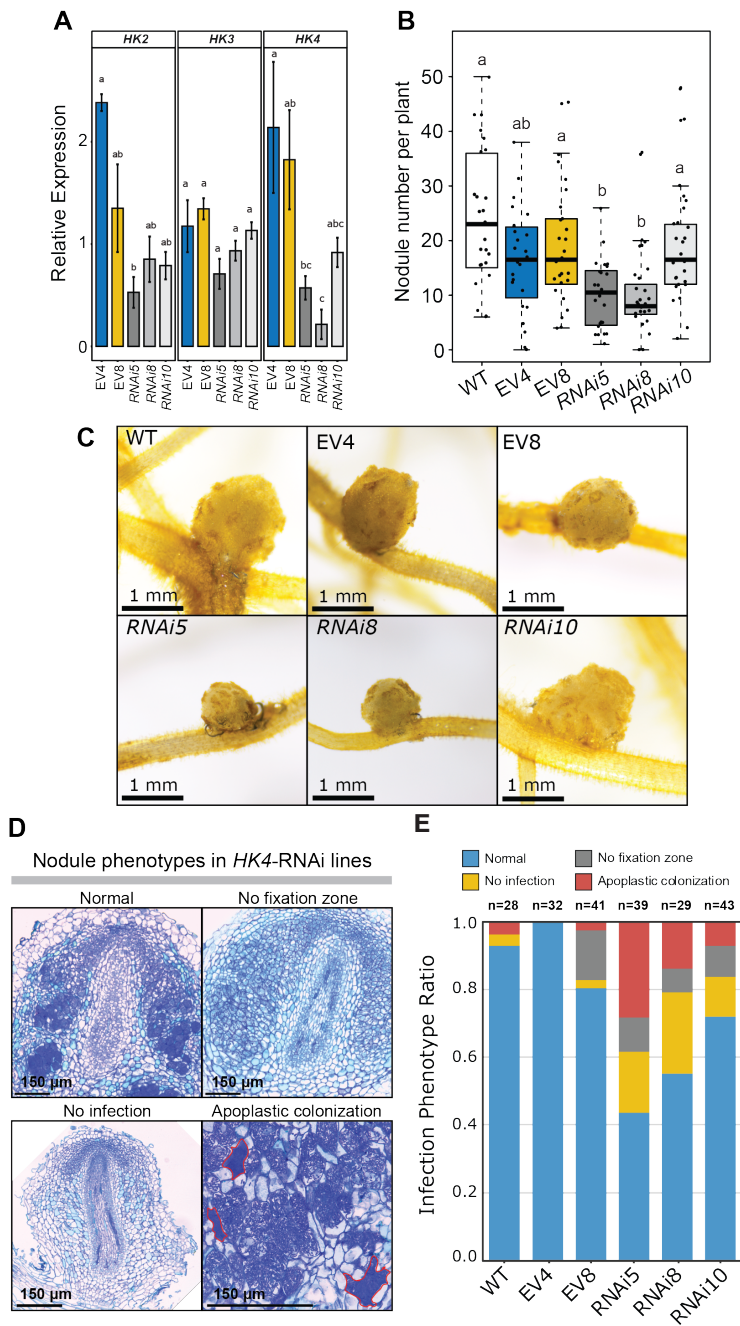
### A symbiotic role of cytokinin in *Parasponia* nodules

A central question that remains is whether cytokinin commits an essential function in nodulation of non-legumes. To this end, we studied the expression of genes involved in cytokinin homeostasis in *P. andersonii* rhizobium-induced nodules and nodule primordia. We noted a higher expression of an *ISOPENTENYL TRANSFERASE* (*PanIPT5*) in nodule primordia and young nodules, and a higher expression of two putative *CYTOKININ OXIDASES* (*PanCKX2* and *PanCKX7*) in mature nodules (**Supplemental Figure 1**). This suggests that biosynthesis and degradation of cytokinin is occurring in *P. andersonii* nodules.

Earlier studies revealed that *Pandersonii Panhk4* knockout mutants do not have a phenotype in nodulation (van Zeijl et al. 2018). We conducted two additional studies on *PanHK4*; RNA interference (RNAi), and constitutive expression of a dominant active allele. By mimicking a missense mutation in the CHASE domain (L263F) of *PanHK4*, similar to that reported for dominant active *LjLHK1* and *MtCRE1* alleles (Tirichine et al., 2006; Tirichine et al., 2007; Ovchinnikova et al., 2011), we created a dominant active allele (*PanHK4\**) (**Supplemental Figure 2**). Ectopic expression of *PanHK4\** using the *CaMV 35S* promoter in *P. andersonii* strongly affected plant development. Only a single line could be obtained, and *in vitro* propagated (**Supplemental Figure 3. A**). Transgenic plantlets ectopically expressing *PanHK4\** were dwarf and highly branched (**Supplemental Figure 3. C**). Only occasionally roots were formed that remained short and relatively thick (**Supplemental Figure 3. C**). Therefore, this *CaMV35S:PanHK4\** line could not be used for further studies on nodulation. Expression studies on leaf tissue revealed increased expression of A-type *RR* genes, which were known to be transcriptionally activated upon cytokinin signalling (**Supplemental Figure 3. B and 4**). This suggests that *PanHK4\** was active. To our surprise we also noted a >20-fold increase of *PanHK2* expression in *CaMV35S:PanHK4\** leaves when compared to control plants (**Supplemental Figure 3. A**). This led us to speculate that in *P. andersonii* a positive feedback loop may occur between *PanHK4* and *PanHK2*.

In a second experiment, we conducted the RNAi approach based on a hairpin construct targeting the second exon encoding the *PanHK4* CHASE domain (**Supplemental Figure 5**). This region shows only 78% and 70% homology to the two other *P. andersonii* cytokinin receptors *PanHK2* and *PanHK3*, respectively. Therefore, we anticipated that specific silencing of *PanHK4* in the transgenic RNAi-lines. We successfully generated three independent RNAi-lines (RNAi5, RNAi8 and RNAi10), and two empty vector control lines (EV4 and EV8). To determine the knockdown levels, qRT-PCR was conducted on RNA isolated from non-inoculated roots. This revealed that *PanHK4* knockdown levels varied between 50% (line RNAi10) and 90% (line RNAi 8) (**Figure 1. A**). We also noted a lower expression of *PanHK2*, and to a lesser extent *PanHK3*, most prominently in the RNAi5 line (**Figure 1. A**). This reduced expression of all three *PanHK* genes can be the result of co-silencing, or alternatively, reduced activity of a *PanHK4*-controlled positive feedback loop.

We noticed that *in vitro* propagated RNAi plantlets were slightly smaller than the EVi controls and wild-type plantlets. Therefore, we quantified shoot and root fresh weight of plantlets grown for 4 weeks in a closed pot system. Results showed a slightly reduced shoot biomass of RNAi5 and RNAi8, despite no difference in root biomass (**Supplemental Figure 6**).



**Figure 1. Phenotypic analysis on *PanHK4*-RNAi stable lines.**  
(A) Relative expression of *HK2*, *HK3* and *HK4* in non-inoculated roots of 2 independent empty vector control lines (EV) and 3 independent RNAi lines. Relative expression was determined by qRT-PCR.  
(B) Nodulation efficiency of wild type, empty vector control and RNAi lines at 4 wpi with *M. plurifarium* BOR2. Data are presented as the average number of nodules per plant.

Next, we scored for potential nodulation phenotypes on the RNAi lines. We noticed that the nodule numbers were reduced on lines RNAi5 and RNAi8, 4 weeks post inoculation with *M. plurifarium* BOR2 (**Figure 1. B**). Also, nodules on the RNAi lines were somewhat smaller in size when compared to wild-type and EVi-control nodules (**Figure 1. C**). Sectioning of nodules revealed that RNAi5 and RNAi8 possessed a significant number of nodules with infection phenotypes that included the absence of intracellular infection, intracellular infection but no fixation thread formation, or occurrence of large apoplastic colonies (**Figure 1. D-E**). The latter is probably the result of the death of nodule cells, which creates large apoplastic spaces. Taken together, these results suggest that one, or several *PanHK* genes, may function in rhizobium infection.

### Rhizobium induces cytokinin signalling in the epidermis and cells of the nodule primordium

To determine in which cells cytokinin signalling is activated upon rhizobium inoculation, we generated transgenic *P. andersonii* lines harbouring the cytokinin-responsive reporter element *TCSn* driving  $\beta$ -*GLUCURONIDASE* (*TCSn:GUS*). This reporter contains B-type *RR* cis-regulatory elements, and has been demonstrated to be cytokinin-responsive in several plant species (Zürcher et al. 2013; Wybouw and De Rybel 2019). Five independent *P. andersonii* *TCSn:GUS* lines were selected and vegetatively propagated. These lines showed a similar *TCSn:GUS* activity under non-symbiotic conditions (**Supplemental Figure 7**). In roots, GUS staining was observed in the columella, the root meristem and the vasculature (**Figure 2. A-B**). This expression pattern in roots is similar to that previously reported for *A. thaliana* and *M. truncatula* (Zürcher et al., 2013; Jardinaud et al., 2016). Roots of *in vitro* grown *P. andersonii* plantlets of all five lines were inoculated with the flavonoid-independent LCO producing strain *Mesorhizobium loti* R7A.pMP604 (Rutten et al., 2020). This strain possesses a plasmid encoding a dominant active NodD protein, which activates LCO biosynthesis genes independently of plant root-secreted

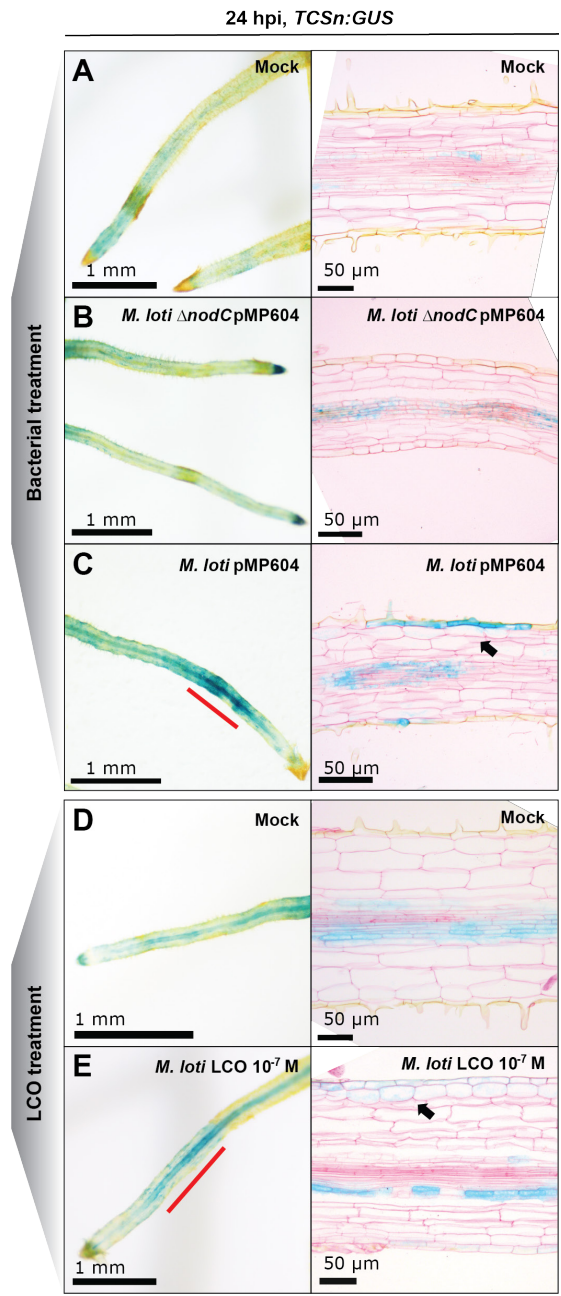
#### Figure 1 (continued).

**(C)** Representative images of nodules collected from RNAi and control lines at 4 wpi with *M. plurifarium* BOR2.

**(D)** Representative images of sections made from RNAi nodules. The various phenotypes observed among sections of these nodules are shown; nodules with a wild-type cytology, nodules in which an infection zone is absent, nodules that contain apoplastic colonization (highlighted with a red line), and non-infected nodules.

**(E)** Quantification of infection phenotypes observed among nodules from the wild type, empty vector and RNAi lines. Representative images of the quantified phenotypes are shown in **(D)**.

Data in **(A)** represents means of 4-6 biological replicates  $\pm$  SEM. Data in **(B)** represents ~25-30 plants collected from 3 different experiments. WT, wild type; EV, stable transgenic empty vector control lines; RNAi, stable transgenic RNAi lines. Statistical analysis was performed using One-way ANOVA and post hoc Tukey HSD test ( $p < 0.05$ ) for multiple group comparisons. Bars with different letters differ significantly.



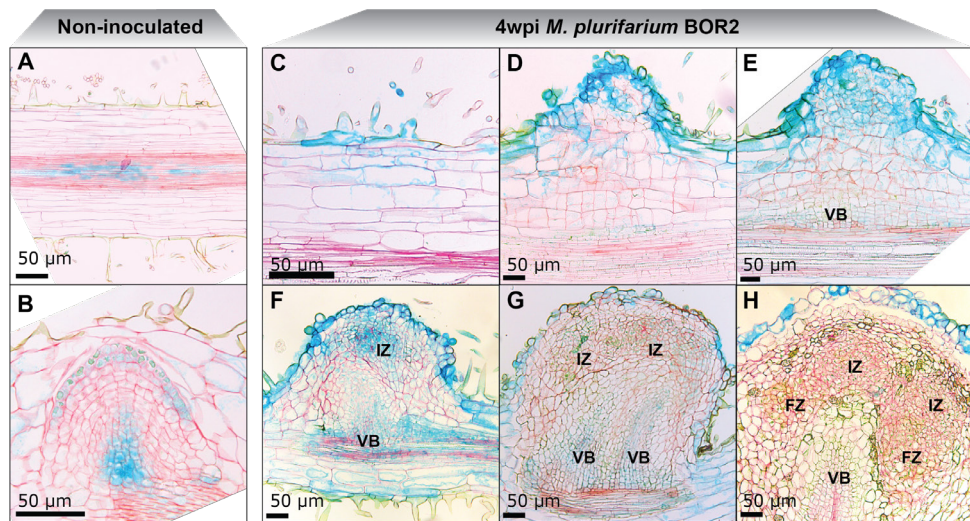
**Figure 2. Rhizobial-induced of cytokinin signalling reporter *P. andersonii* *TCSn:GUS* activity on lateral roots.**

Spatiotemporal expression of *TCSn:GUS* activity in mock-inoculated roots (**A**) or roots inoculated with *M. loti* R7A.pMP604 (**B**) or *M. loti* R7A  $\Delta$ nodC pMP604 (**C**) for 24 h. Spatiotemporal expression of *TCSn:GUS* activity in mock-inoculated roots (**D**) or roots inoculated *M. loti* LCOs  $10^{-7}$  M (**E**) for 24 h. Sections were counterstained with ruthenium red.

flavonoids (Spaink et al., 1987). As a negative control, *M. loti* R7A  $\Delta nodC$ .pMP604 was used, which is unable to synthesize LCOs due to a mutation in *nodC* (Rodpothong et al., 2009; Rutten et al., 2020). Root systems were stained for GUS activity at 24h post inoculation. *TCSn* showed to be highly responsive towards inoculation with *M. loti* R7A.pMP604. Several lateral roots (30.23% in line A2 (n=13/47)) showed dark blue regions about 1-2 mm above the root apical meristem within 30 minutes after incubation in GUS-staining solution (**Figure 2. A**). This region is known to be susceptible for nodulation (**Supplemental Figure 8**). In contrast to the dark blue regions observed on lateral roots, the primary roots displayed more faint staining that was indistinguishable from the negative control.

We selected the lateral roots of *TCSn:GUS* line A2 for sectioning and subsequent microscopic analysis. In longitudinal root sections of mock-treated and negative control plants, GUS activity is observed in the inner region of the root (**Figure 2. A**). In contrast, upon inoculation of *M. loti* R7A.pMP604 intense GUS-activity was observed also in the epidermis (**Figure 2. A**). This suggests that LCO-producing rhizobia trigger cytokinin signalling specifically in the epidermis. To confirm these findings, and demonstrate that epidermal activation of *TCSn* is an LCO-specific response, we applied purified *M. loti* LCOs ( $10^{-7}$  M) on roots of *in vitro* grown *TCSn:GUS* line A2, (36.84%, n=14/38) and stained for GUS activity at 24h post application. This revealed a similar induction of *GUS* expression in the epidermis, though substantially weaker when compared to *M. loti*.pMP604-inoculated roots (**Figure 2. B**). Taken together, these data show that LCO-secreting rhizobia can induce cytokinin signalling in lateral roots of the nodulating non-legume *P. andersonii*.

To investigate whether cytokinin signalling remains activated during nodule organogenesis, the *TCSn:GUS* spatial expression pattern was analysed in nodulating plants. To this end, *TCSn:GUS* reporter line A2 was inoculated with the highly-compatible strain *Mesorhizobium plurifarum* BOR2 and nodules at different developmental stages were isolated at 4 weeks post inoculation, and subsequently stained in GUS buffer for 30 min. This revealed intense blue staining of nodule primordia and more mature nodules (**Supplemental Figure 9**). Compared to legumes, rhizobium infection in *Parasponia* spp. is facilitated through crack-entry, which can occur via the apoplastic space between dividing epidermal and outer cortical cells (Lancelle and Torrey, 1984; Bender et al., 1987). Sections of young nodule primordia showed an intense GUS activity, especially in epidermal cells (**Figure 3. C-D**). When nodules emerged, GUS activity increased in dividing cortical cells that will subsequently become infected by rhizobium (**Figure 3. E-F**). In more mature nodules, intense GUS staining was restricted to the outermost cell layer and vasculature (**Figure 3. G-H**).



**Figure 3. Spatiotemporal expression of the *TCSn:GUS* reporter at different stages of nodule development**

**(A)** Longitudinal section through a non-inoculated root showing activity of the *TCSn:GUS* reporter in the vasculature.

**(B)** Activity of the *TCSn:GUS* reporter in an emerging lateral root primordium.

**(C-E)** Activity of the *TCSn:GUS* reporter during subsequent stages of nodule primordium development. Samples were harvested at 4 wpi with *M. plurifarum* BOR2.

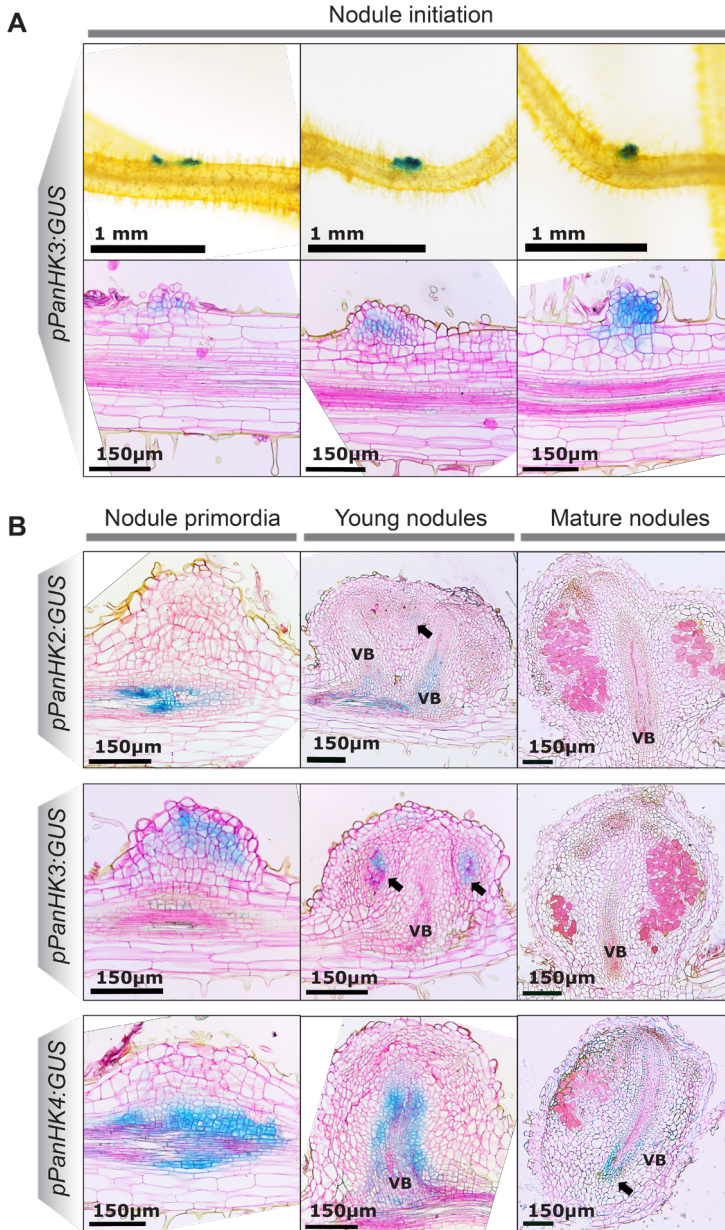
**(F-H)** Activity of the *TCSn:GUS* reporter in young **(F)** and developing nodules **(G-H)**. Samples were harvested at 4 wpi with *M. plurifarum* BOR2.

IZ, infection zone; VB, nodule vascular bundle; FZ, fixation zone. Sections were counterstained using ruthenium red.

Taken together, these data demonstrate that cytokinin signalling is activated in the *P. andersonii* root epidermis by LCO-secreting rhizobia, and is maintained in nodule cortical cells that ultimately become infected.

### ***PanHK3* is transcriptionally induced by rhizobium and expressed in infected zone**

As RNAi suggests that cytokinin receptors may function in rhizobium intracellular infection, we questioned whether this is controlled in a redundant fashion, or whether a specific *PanHK* gene has been recruited to commit such function in nodulation. To obtain insights in this, we first conducted promoter reporter studies, using the ~3.5-3.7 kb regions upstream of the translational start sites of *PanHK2*, *PanHK3*, and *PanHK4* as putative promoters. These reporter constructs were introduced into *P. andersonii* by *A. tumefaciens*-mediated transformation. For each of the three constructs, 2-3 independent lines were selected. For each of the three constructs, independent transgenic lines showed very similar non-symbiotic expression profiles. But, unlike the *TCSn:GUS* lines, which required only ~30 minutes of incubation in GUS-staining solution, the *PanHK* promoter-GUS lines were incubated for ~7 hours. *pPanHK2:GUS* and *pPanHK4:GUS*



**Figure 4. Differential activity of the putative promoters of cytokinin receptors *PanHK2*, *PanHK3* and *PanHK4* during *P. andersonii* nodule development**

**(A)** Representative bright-field microscopic pictures of (top), and longitudinal sections through (bottom), nodule initiation sites on roots of *PanHK3:GUS* line B3. Samples were taken at 4 wpi with *M. plurifarium* BOR2.

**(B)** Sections through nodule primordia, young nodules and mature nodules formed on roots harbouring the *pPanHK2:GUS* (line B1), *pPanHK3:GUS* (line B3), and *pPanHK4:GUS* (line B2) reporter constructs. Nodules were collected 4 wpi with *M. plurifarium* BOR2.

Data represent representative images of two to three independent lines with 8-10 biological replicates each. Arrowheads point to sites of weaker GUS activity. VB, nodule vascular bundle.

were active in the root meristematic zone, emerging lateral root primordia, and in the vasculature of almost the entire root system. In contrast, *PHK3::GUS* activity was detected in only a 1–2 mm region just above the root tip (**Supplemental Figure 10**).

Subsequently, the *pPanHK* reporter-*GUS* lines were inoculated with *M. plurifarium* BOR2, and investigated for *GUS* staining at 4 weeks post inoculation. Again, *pPanHK2::GUS* and *pPanHK4::GUS* showed a similar spatiotemporal activity. We did not detect any activity of *pPanHK2* and *pPanHK4* at the stage of nodule initiation. However, both putative promoters were active in the vasculature of the developing nodule (**Figure 4. A**). In case of *pPanHK2::GUS*, we also observed some faint expression in the nodule infection zone. In contrast, *pPanHK3::GUS* was induced in dividing epidermal and cortical cells, and at later stages in the nodule cells that become infected by rhizobium (**Figure 4. A–B**). In mature nodules, activity of *pPanHK2* and *pPanHK3* was no longer detected, neither in the nodule vascular bundle, the fixation zone, nor in the nodule outer cortex. Only *pPanHK4* remained active in the nodule vascular bundle (**Figure 4. A**). Taken-together, this shows that the expression pattern of the *pPanHK3-GUS* reporter mirrors the activity of the *TCSn* cytokinin reporter in response to rhizobium-induced nodulation.

### ***P. andersonii* PanHK3 controls rhizobium intracellular infection**

To determine whether *PanHK3* commits an essential function in nodulation, we used CRISPR-Cas9 mutagenesis to create *Panhk3* knockout lines. Likewise, we aimed to mutate *PanHK2*, as well as to generate different combinations of double mutants (*Panhk2;Panhk3*, *Panhk2;Panhk4*, and *Panhk3;Panhk4*). We used single guide (sg) RNAs targeting exon 2 encoding the CHASE-domain. This strategy only resulted in *Panhk3* mutant lines, whereas *Panhk2*, or any combination of double mutant could not be obtained, regardless of the concentration of synthetic hormones that we used in the culture medium or transformation strategy. This suggests that a mutation in *PanHK2*, or simultaneous mutagenesis of both *PanHK3* and *PanHK4*, blocks the ability of *P. andersonii* cells to regenerate and form shoots.

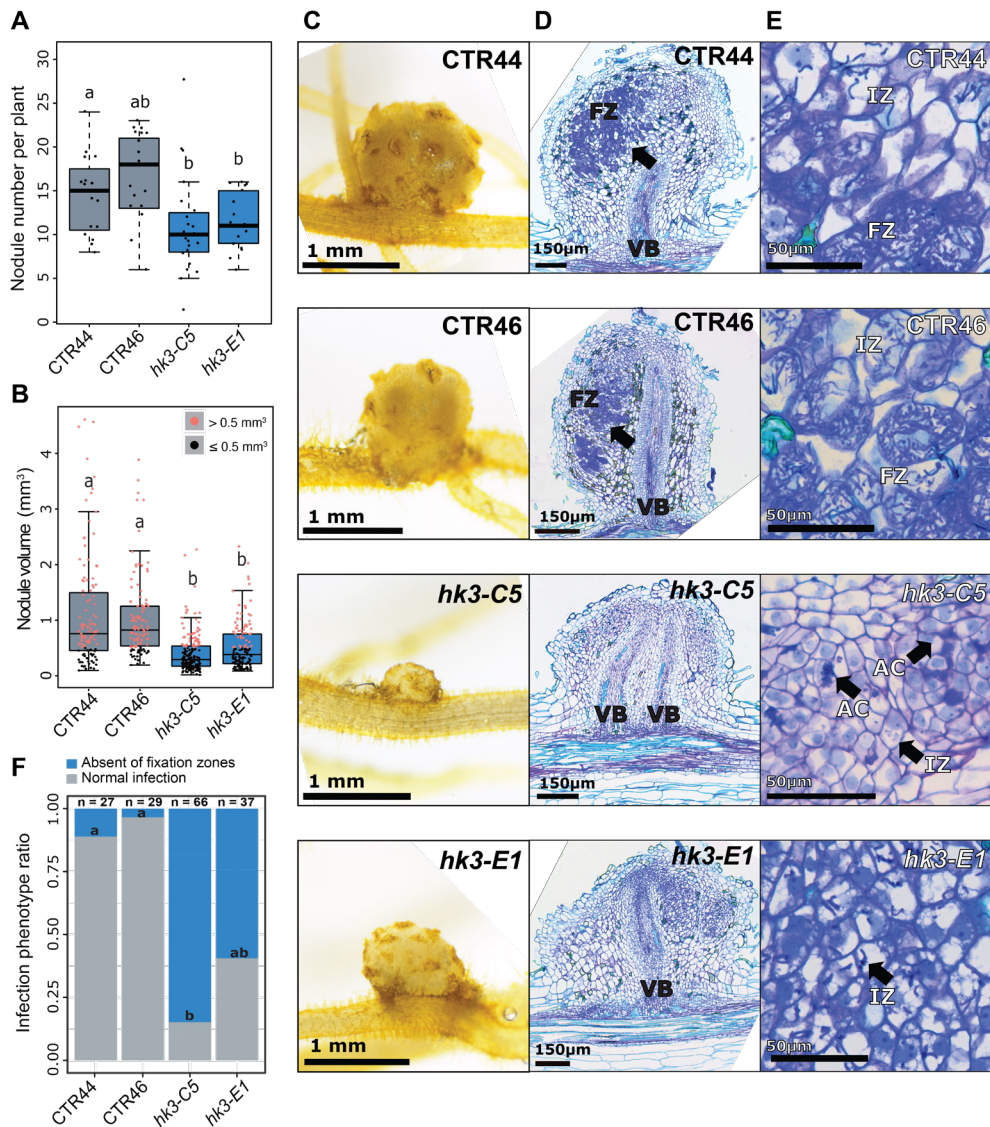
Two independent *Panhk3* mutant lines were obtained. Line *Panhk3* C5 is homozygous carrying a ~151 bp deletion, while the *Panhk3* line E1 is bi-allelic with an in-frame mutation in one of the alleles (**Supplemental Figure 11**). *Panhk3* mutant plantlets did not have any obvious non-symbiotic phenotypes. We performed a nodulation assay with *M. plurifarium* BOR2 on both *Panhk3* mutant lines, and nodules were scored at 4 weeks post inoculation. As control, we included two empty vector lines, CTR44 and CTR46, which express *Cas9* and the *NPTII* kanamycin resistance marker gene. Both *Panhk3* mutant lines developed nodules, though nodule numbers were somewhat reduced when compared to the empty vector controls (**Figure 5. A**). More strikingly,

*Panhk3* mutant nodules were significantly smaller when compared to control nodules (**Figure 5. B**). About 73.5% of *Panhk3*-C5 nodules had a volume of  $\leq 0.5 \text{ mm}^3$ , which is half the volume of control nodules (**Figure 5. B**). This suggests that *Panhk3* C5 nodules experienced an arrest or delay in their development. In the case of the *Panhk3*-E1 line, about 63.8% of the nodules were normal in size, suggesting that this line has a weaker symbiotic phenotype (**Figure 5. B**). Subsequently, we sectioned nodules of *Panhk3*-C5 (n=66), *Panhk3*-E1 (n=37), and both control lines (n=27 for CTR44 and n=29 for CTR46). This showed that 85% of *Panhk3*-C5 and 59.4% of *Panhk3*-E1 mutant nodules were devoid of intracellular rhizobium infection (**Figure 5. D-G**). In contrast, in the control lines (CTR44 and CTR46) absence of a fixation zone was only observed in nodule primordia and young nodules (**Figure 5. F**). In *Panhk3*-C5 nodules, we also observed an increase of apoplastic colonization (**Figure 5. E**). Apoplastic colonization is often recognized as a 'pathogenic-type' of infection, where the bacteria are proliferated and surrounded by an electron-dense cell wall material (Lancelle and Torrey, 1985). In the *Panhk3*-E1 mutant line, apoplastic colonization is much less observed. Sections of the few large nodules found on both *Panhk3*-C5 and *Panhk3*-E1 mutant lines, showed a wild-type cytoarchitecture. Altogether, based on the spatiotemporal expression and functional analysis of *PanHK3*, we conclude that *PanHK3* is important for rhizobium intracellular infection.

## Discussion

Current hypothesis on evolutionary perspective of root nodulation is that the symbiotic signalling pathway likely utilizes a common set of genes shared within the nodulating lineages (Griesmann et al., 2018; van Velzen et al., 2018; van Velzen et al., 2019). In legumes, several lines of evidence demonstrated that the activation of cytokinin signalling is an essential step leading to nodule organogenesis. Furthermore, it is speculated that cytokinin signalling also plays a pivotal role in infection, however, clear evidence is lacking. Here, we performed a series of experiments to gain insights in the role of cytokinin signalling in nodulation of the nonlegume *Parasponia andersonii*. By exploiting the two-component signalling reporter *TCSn*, we demonstrated that rhizobium induces cytokinin signalling in *P. andersonii* root epidermal cells and young nodule primordia. This spatiotemporal expression pattern is mirrored by the *PanHK3* histidine kinase receptor. As *Panhk3* knockout plants form many nodules devoid of intracellular rhizobium infection, we conclude that this cytokinin signalling receptor commits an essential role in the infectability of *P. andersonii* nodule cells.

In *L. japonicus*, *TCSn* activity is first activated in cortical cells in response to rhizobia, leading to initiation of nodule primordia and only later epidermal activity is observed. This symbiotic *TCSn* activation requires a functional *LjLHK1* gene (Held et al., 2014;



**Figure 5. *Panhk3* mutants develop smaller nodules that display defects in intracellular infection** (A) Number of nodules formed per plant at 4 wpi with *M. plurifarium* BOR2.

(B) Volume of control and *Panhk3* mutant nodules. Nodules were harvested from at 4 wpi with *M. plurifarium* BOR2.

(C) Representative bright-field microscopic pictures of nodules formed on control and *Panhk3* mutant lines at 4 wpi with *M. plurifarium* BOR2.

(D) Mid-plane sections through nodules formed on control and *Panhk3* mutant lines at 4 wpi with *M. plurifarium* BOR2.

(E) Zoomed-in pictures of the nodule sections shown in (D).

(F) Ratiometric quantification of the infection phenotype in control and *Panhk3* mutant nodules at 4 wpi with *M. plurifarium* BOR2. Quantification was performed by scoring the ratio of infected nodules vs nodules that were devoid of a fixation zone.

Reid et al., 2017). In contrast, in *M. truncatula*, *TCSn* activity is observed first in the epidermis and outer cortex in response to *S. meliloti* inoculation (4 to 72 hpi) and only later become strongly induced in the inner cortex (Jardinaud et al., 2016). This suggests that in legumes variations occurs in cytokinin signalling. Similar as in *M. truncatula*, in *P. andersonii*, *TCSn* activation in response to rhizobium-inoculation is firstly observed in epidermal cells. These cells subsequently divide and become associated with nodule development (Lancelle and Torrey, 1984; Geurts et al., 2016). The responsive region upon LCO perception in *P. andersonii* *TCSn::GUS* roots was relatively small (~1-2 mm) and 30-37% of lateral roots was responsive. This hindered us to conduct additional analysis like quantification of cytokinins and gene expression in response to rhizobia. Nevertheless, the observed *TCSn::GUS* activity in response to rhizobium inoculation points to a strong symbiotic cytokinin responses in *P. andersonii*. As nodulation in legumes and *Parasponia* has a shared evolutionary origin, we postulate that symbiotic cytokinin signalling is a conserved module of the nodulation trait.

Reverse genetic studies revealed a function of *PanHK3* in rhizobium intracellular infection. Basically, three phenotypes were observed; (i) nodules devoid from intracellular rhizobium infection, (ii) nodules that contained infected cells, but were hampered in formation of fixation threads, and (iii) young nodules that were intracellularly infected normally, but possessed in increased apoplastic colonization. Large apoplastic colonization is the result of the death of a nodule cell and is occasionally observed in wild type nodules of *Parasponia* species (Lancelle and Torrey, 1985). An increased number of large apoplastic colonies is only observed in *Panhk3* mutant nodules that were infected rather normally, but not in nodules devoid of intracellular infection (but with rhizobia in the apoplast), nor in nodules with arrested infection threads. This suggests that a rhizobium infected nodule cell becomes more vulnerable during fixation thread formation, and that cytokinin may repress a cell death response. Interestingly, studies in *A. thaliana* revealed a function of the orthologous receptor *AtHK3* in repressing cell death triggered by a circadian stress regime (Nitschke et al., 2016), as well as stomatal immunity against the bacterial pathogen *Pseudomonas syringae* pv. Tomato (Arnaud et al., 2017). The latter response requires cytokinin-mediated regulation of reactive oxygen species (ROS) homeostasis. As ROS and ROS-producing RBOH enzymes have shown to be critical in rhizobium infection in legumes (Montiel et al., 2016), it is tempting to speculate a conserved functioning of *HK3* orthologous genes in different plant species and biological processes.

### Figure 5 (continued).

Data in (A) and (B) were generated from 17-23 biological replicates of 2 independent experiments. Statistical analysis was performed using One-way ANOVA and followed by post hoc Tukey HSD test ( $p < 0.05$ ). FZ, fixation zone; VB, nodule vascular bundle; IZ, infection zone; AC, apoplastic colonization. Bars with different letters differ significantly.

A fraction of the nodules formed on *Panhk3* mutant plantlets have a wild type appearance. This suggests that intracellular infection can occur independent of cytokinin signalling, or alternatively, the occurrence of some level of gene redundancy. To our opinion, the latter is more likely, because weak activity of the *PanHK2* putative promoter was observed in the infection zone of wild type nodules (**Figure 4.B**). As *PanHK4* RNAi and expression of a dominant active allele suggest that feedback loops occur in transcriptional regulation of *PanHK* genes, *Panhk3* knockout may also have caused transcriptional compensation by other *HK* genes. Similar in legumes, gene redundancy may have masked rhizobium infection phenotypes in *hk* mutants. *L. japonicus* and *M. truncatula* possess four *HK* genes that all are transcriptionally induced in nodule primordia (Held et al., 2014; Boivin et al., 2016). In *L. japonicus* *LjLHK1* commits an important role in nodule organogenesis, its symbiotic functioning is partially redundant with *LjLHK1A* and *LjLHK3* (Held et al., 2014). Similarly, redundancy is found in the symbiotic functioning of *MtCRE1* in *M. truncatula* (Boivin et al., 2016). This redundancy may explain why in legumes only limited evidence has been found that indicates the importance of cytokinin signalling in the infectability of nodule cells. The finding that *PanHK3* plays an essential role in rhizobium of *P. andersonii* nodules, underlines the importance of a complementary research model to identify key mechanisms in nitrogen-fixing nodulation.

## Methods

### Plant Materials and growth conditions

All experiments were performed using *P. andersonii* WU1 or offspring thereof (van Velzen et al., 2018a; Wardhani et al., 2019). Plants were grown and maintained as described previously (van Zeijl et al., 2018; Wardhani et al., 2019). Plantlets for nodulation assays were multiplied via *in vitro* propagation, rooted, and inoculated with *Mesorhizobium plurifarium* BOR2 ( $OD_{600} = 0.05$ ) (van Velzen et al., 2018a; van Zeijl et al., 2018; Wardhani et al., 2019). For early induction of cytokinin signalling, we made use of *Mesorhizobium loti* R7A transformed with pMP604 ( $OD_{600} = 0.05$ ) (Spaink et al., 1987; Díaz et al., 2000), *M. loti* R7A  $\Delta nodC$  pMP604 ( $OD_{600} = 0.05$ ), and extracted LCOs from *M. loti* R7A ( $10^{-7}$  M) (Rutten et al., 2020). Nodulation efficiencies were calculated by plotting the nodule number per plant. Nodule volume size estimates were quantified using ellipsoid volume form ( $V = \frac{4}{3}\pi abc$ ), of which the length of *a*, *b*, *c* axes were analysed using IMAGEJ software. Comparisons were made based on the average nodule volume size per plant using at least nine replicate plants.

### Vectors and constructs

For RNAi-mediated knockdown of *PanHK4*, a fragment of 82 bp (**Supplemental Figure 5**) was used to target the second exon of *PanHK4* (**Supplemental Table 1-2**). The amplicon was cloned into pENTR-D-TOPO (Invitrogen, Carlsbad, USA) and recombined

into the DsRed-modified gateway vector pK7GWIWG2(II)-RR driven by the CaMV 35S promoter (Limpens et al., 2004), which resulted in the binary constructs pK7GWIWG2(II)-RR-p35S-PanHK4-RNAi. For the empty vector control of RNAi, we used the binary plasmid pK7GWIWG2(II)-RR-p35S-RNAi-control as previously described (van Zeijl et al., 2015b). For the constitutively active form of PanHK4, we introduced a point mutation which led to a substitution of leucine (L) to phenylalanine (F) at amino acid position 263. The mutated *PanHK4*-CDS was cloned behind the CaMV 35s promoter, and assembled in a binary transformation vector as described previously (van Zeijl et al., 2018; Wardhani et al., 2019). For CRISPR/Cas9-genome editing, single-guide RNAs (sgRNAs) were designed using the 'Find CRISPR Targets' function implemented in Geneious Software v 9.1.5 (Biomatters, Auckland, New Zealand), and subsequently checked against the *P. andersonii* genome database for high-identity off-targets. Two sgRNAs were used to target the second coding exon of *PanHK3* (**Supplemental Figure 8**). Selected sgRNAs were amplified using sequence-specific forward primers and a universal reverse primer (**Supplemental Table 1-2**), using Addgene plasmid no. 46966 as template (Nekrasov et al., 2013). Constructs for CRISPR/Cas9-mediated mutagenesis were assembled as described previously (van Zeijl et al., 2018; Wardhani et al., 2019). For spatiotemporal expression studies, the putative promoter sequences of *PanHK2*, *PanHK3* and *PanHK4*, encompassing ~3.5-3.7 kb upstream of the translational start site, were obtained by DNA synthesis (**Supplemental Table S3-5**). To allow for Golden Gate cloning, *BpiI* and *BsaI* restriction sites in the putative promoter sequences were mutated by introducing single nucleotide substitutions (Engler et al., 2014). Subsequently, the putative promoter sequences were cloned in front of a  $\beta$ -glucuronidase-encoding sequence and 35S terminator, and assembled in binary transformation vectors as previously described (Engler et al., 2014). The Gene Identifiers and Gene Identifiers accession nos. of the used *P. andersonii* genes ([www.parasponia.org](http://www.parasponia.org)) are: PanWU01x14\_002510 (*HK2*), PanWU01x14\_222240 (*HK3*), PanWU01x14\_205610 (*HK4*).

### Plant transformation

*Agrobacterium tumefaciens*-mediated transformation and genotyping were performed based on previously published protocols (van Zeijl et al., 2018; Wardhani et al., 2019). Primers used for genotyping are listed in **Supplemental Table S2**. For promoter-GUS reporter studies, three to five independent lines were examined for each construct.

### Histochemical analysis and microscopy

Root and nodule samples of the *P. andersonii* *TCSn::GUS*, *PanHK2::GUS*, *PanHK3::GUS*, and *PanHK4::GUS* lines were incubated in GUS buffer (3% (w/v) sucrose, 10 mM EDTA, 2 mM k-ferrocyanide, 2 mM k-ferricyanide, and 0.5 mg/ml 5-bromo-4-chloro-3-indolyl- $\beta$ -D-glucuronic acid cyclohexylammonium salt (X-Gluc), in 0.1 M phosphate buffer (pH 7.2) at 37°C for 30 min in case of *TCSn::GUS* lines, and 7 h for all putative

cytokinin promoter reporter lines. For plastic sections, root segments and/or nodules of all transgenic lines were fixed in 4% paraformaldehyde (w/v), 5% glutaraldehyde (v/v) in 50 mM phosphate buffer (pH 7.2) at 4°C for 24 h. Subsequently, the samples were dehydrated using an ethanol series and embedded in Technovit 7100 (Heraeus Kulzer, Hanau, Germany) according to the manufacturer's instructions. Semithin sections were cut using a Leica Ultra-cut microtome (Leica Microsystems, Wetzlar, Germany) to 5 µm thickness for nodules formed on *PanHK4-RNAi* and *Panhk3* mutant lines and 8 µm thickness for GUS-stained samples. Sections were stained with 0.05% Toluidine Blue or 0.1% Ruthenium Red. Images were photographed using a Leica DM5500B microscope equipped with a DFC425C camera (Leica Microsystems).

### RNA isolation and qRT-PCR analysis

RNA was isolated from snap-frozen root segments, and cDNA was prepared from 1 µg of total RNA using the iScript cDNA synthesis kit (Bio-Rad), following the manufacturer's instructions. Ten microliter quantitative reverse transcription polymerase chain reaction (qRT-PCR) reactions were set up using iQ SYBR Green Supermix (Bio-Rad) and 5 ng template cDNA. Quantification was performed using a CFX Connect optical cycler, according to the manufacturer's protocol (Bio-Rad). Normalization was performed based on the stably expressed reference gene ELONGATION FACTOR 1a (*PanEF1a*; van Zeijl et al., 2018). Primers used for qRT-PCR analysis are listed in **Supplemental Table S2**.

### Statistical analysis

Bars represent mean  $\pm$  SE for all experiments. Elements in the boxplot illustrate the lowest and highest data points, the sample median, and the first and third quartiles. Graphs visualisation and statistical analysis were generated by R studio Version 1.1.463. Statistical analysis was performed using One-Way ANOVA, followed by post-hoc Tukey HSD ( $p < 0.05$ ) for multiple group comparison (Mendiburu, 2020). Bars labelled with identical letters are deemed not significantly different.

### Phylogenetic reconstruction

Phylogenetic tree reconstruction was performed using FastTree implemented in Geneious Software v 8.1 (Price et al., 2009). As input for phylogenetic reconstruction a protein alignment generated using MAFFT alignment (Katoh et al., 2002) was used. Protein sequences of *P. andersonii* (*Pan*) were retrieved from [www.parasponia.org](http://www.parasponia.org), *Chamaecrista fasciculata* (*Cfa*) were retrieved from <http://gigadb.org/dataset/101045>, and for *Arabidopsis thaliana* (*At*), *Medicago truncatula* (*Medtr*), *Glycine max* (*Glyma*), *Populus tricarpha* (*Pt*), *Fragaria vesca* (*Fvesca*), and *Eucalyptus grandis* (*Eucgr*) were retrieved from <https://phytozome.jgi.doe.gov/pz/portal.html>.

## Author Contributions

Conceptualization, T.A.K.W., A.v.Z. and R.G.; Methodology, T.A.K.W. and A.v.Z.; Investigation, T.A.K.W., A.v.Z., S.L., and K.G.; Writing – Original Draft, T.A.K.W.; Writing – Review & Editing, T.A.K.W., A.v.Z., W.K., and R.G. Funding Acquisition, R.G., and T.B.; Supervision, R.G.

## Acknowledgments

The authors would like to thank Dugald Reid (Aarhus University, Denmark) for providing the TCSn level 0 construct and Simona Radutoiu (Aarhus University, Denmark) for providing *M. loti* R7A  $\Delta nodC$ . Golden Gate parts and cloning vectors were kindly provided by Mark Youles, Sophien Kamoun and Sylvestre Marillonnet through the Addgene database. This work was supported by NWO-FSC Joint Research project (846.11.005) to T.B. and R.G., NWO-VICI (865.13.001) and ENSA project funded by the Bill & Melinda Gates Foundation to the University of Cambridge to R.G.

## Supplemental Information

**Supplemental Table 1. List of Golden Gate constructs used in this study.**

Construct	Description	Level	Backbone	Contains <sup>2</sup>	Literature
1	<i>nptII</i> resistance cassette	1	pICH47802	pICSL70004: <i>nptII</i>	(van Zeijl et al., 2018)
2	35S <sub>pro</sub> ::ΩNLS-Cas9:35S <sub>ter</sub>	1	pICH47742	pICH41388:35S <sub>pro</sub> , pAGM5311:ΩNLS, pICH41308::aCas9, pICH41414:35S <sub>ter</sub>	(Fauser et al., 2014; van Zeijl et al., 2018)
3	<i>PanHK4*</i> CDS	1	pICH47751	pICH51277:35SShort-pro, L0 <i>PanHK4*</i> -CDS, pICH41414:35S	-
4	35S: <i>GUS</i>	1	pICH47751	pICH51277:35SShort-pro, pICH7511: <i>GUS</i> -CDS, pICH41414:35S	-
5	PanHK3sgRNA1	1	pICH47751	pICSL01009:AtU6p, corresponding PCR amplicon	(Nekrasov et al., 2013)
6	PanHK3sgRNA2	1	pICH47761	pICSL01009:AtU6p, corresponding PCR amplicon	-
7	PanHK2pro	1	pICH47751	Pro+5U'TR <i>PanHK2</i> , pICH7511: <i>GUS</i> -CDS, pICH41414:35S	-
8	PanHK3pro	1	pICH47751	Pro+5U'TR <i>PanHK3</i> , pICH7511: <i>GUS</i> -CDS, pICH41414:35S	-
9	PanHK4pro	1	pICH47751	Pro+5U'TR <i>PanHK4</i> , pICH7511: <i>GUS</i> -CDS, pICH41414:35S	-
10	<i>TCSn</i> : <i>GUS</i>	1	pICH47751	pro <i>TCSn</i> , pICH7511: <i>GUS</i> -CDS, pICH41414:35S	-
11	<i>PanHK4*</i>	2	pICSL4723	1R: construct 1, 2F: construct 3, end-link pICH41744	-
12	<i>Pan35S</i> : <i>GUS</i>	2	pICSL4723	1R: construct 1, 2F: construct 4, end-link pICH41744	-
13	CRISPR_ <i>PanHK3</i>	2	pICSL4723	1R: construct 1, 2F: construct 5, 3F: construct 6, end-link pICH41766	-
14	p <i>PanHK2</i> : <i>GUS</i>	2	pICSL4723	1R: construct 1, 2F: construct 7, end-link pICH41744	-
15	p <i>PanHK3</i> : <i>GUS</i>	2	pICSL4723	1R: construct 1, 2F: construct 8, end-link pICH41744	-

16	pPanHK4:GUS	2	pICSL4723	1R: construct 1, 2F: construct 9, end-link pICH41744	-
17	<i>TCSn</i> :GUS	2	pICSL4723	1R: construct 1, 2F: construct 10, end-link pICH41744	-

Position and orientation in level 2 Golden Gate modules is depicted by a number followed by either F or R for forward or reverse orientation, respectively.

**Supplemental Table 2. Primers used in this study.**

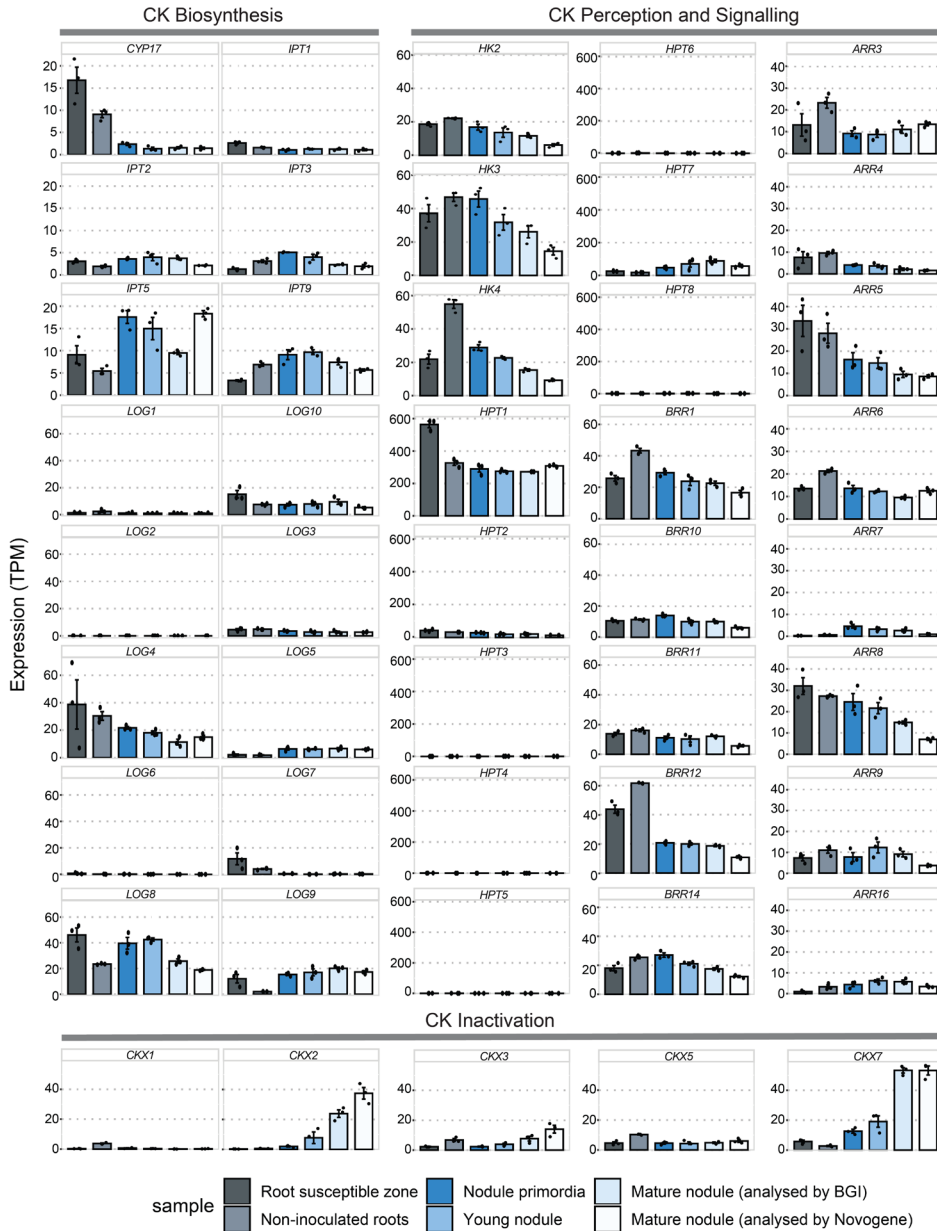
Name	Purpose	Sequence
PanHK3_sgRNA1	Clone sgRNA	TGTGGTCTCAATTGTGAAAGACCTCTCAC GAGTGGTTTTAGAGCTAGAAATAGCAAG
PanHK3_sgRNA2	Clone sgRNA	TGTGGTCTCAATTGATATGTGTGACGAG CGAGCCGTTTTAGAGCTAGAAATAGCAAG
sgRNA_Rv	Clone sgRNA	tgtggtctcaAGCGTAATGCCAACCTTTGTAC
PanHK3-cripsr_geno_Fw	Genotyping CRISPR mutants	AGAGTCGGGAAAAACATGGTG
PanHK3-cripsr_geno_Rv	Genotyping CRISPR mutants	GCAATGGAAAAGGAGCAGTT
PanHK2 <sub>pro</sub> _geno_Fw	Genotyping PanHK2 <sub>pro</sub> :GUS line	CATAGGTGGGTTTTGGTTACT
PanHK2 <sub>pro</sub> _geno_Rv	Genotyping PanHK2 <sub>pro</sub> :GUS line	TCATTGTTGCCTCCCTG
PanHK3 <sub>pro</sub> _geno_Fw	Genotyping PanHK3 <sub>pro</sub> :GUS line	ACTTGTGTTGGTTACTGCT
PanHK3 <sub>pro</sub> _geno_Rv	Genotyping PanHK3 <sub>pro</sub> :GUS line	TCATTGTTGCCTCCCTG
PanHK4 <sub>pro</sub> _geno_Fw	Genotyping PanHK4 <sub>pro</sub> :GUS line	CGTCTGTTGACTGGCAGGTA
PanHK4 <sub>pro</sub> _geno_Rv	Genotyping PanHK4 <sub>pro</sub> :GUS line	GGCACAGCACATCAAAGAGA
qPanARR3_Fw	qRT-PCR	ATGCCTGGAATGACTGGATATGAAC
qPanARR3_Rv	qRT-PCR	TGAAATCCTCTGCTCCTTCCTCC
qPanARR6_Fw	qRT-PCR	ATGTTGATAITTCGATGGGTCTGG
qPanARR6_Rv	qRT-PCR	CTCGTTCCACTTTCTACAGCAGTG
qPanARR7_Fw	qRT-PCR	ACCCCTTGACCATTCTTGTC
qPanARR7_Rv	qRT-PCR	TGTTGCCATTGCTTTTACCAG
qPanARR8_Fw	qRT-PCR	TGGGAATGGCTACAACAGAGT
qPanARR8_Rv	qRT-PCR	TCTGGGAAACAGAAGGTGT
qPanHK2_Fw	qRT-PCR	AGCCACCCCATAGTTTGTATG
qPanHK2_Rv	qRT-PCR	TTTTCGTACCCTTCCACAGTG
qPanHK3_Fw	qRT-PCR	CTGGCAAATCTAATGGAAAGGG
qPanHK3_Rv	qRT-PCR	CTGGCAAATCTAATGGAAAGGG
qPanHK4_Fw	qRT-PCR	TCGCTCTCCTTGTTCCAC
qPanHK4_Rv	qRT-PCR	CCCTCTCCATTGCTTTATAGTCC

[illegible]

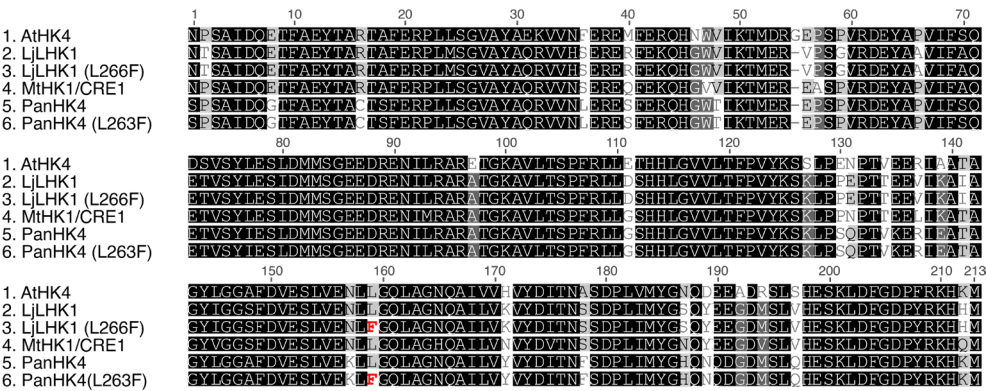
**Supplemental Table 4. Sequences of putative promoter regions of *P. andersonii* HK3.** Yellow colour highlights 5'UTR. Nucleotides highlighted in red indicate nucleotide replacements to remove *BsaI* or *BpII* sites.

CTCAGAAAAATCCCAGAGCAATACAGTATATATCTCATACAATCAATCCCATCACAGATATTCGAGAAAACTAAAA  
ATCACATTTGTTCTCATATATTCGAGAAAAATAGTTTATTCAAACACGATATAAAACAGAAGTAATTGATTCCCT  
ACTTCGCTCGAATAACAATACCACGCTATCGAATCCTCACAGTGTCTGTTGACGATCGCCCGGAGGGTCGGGAACCA  
ATTCCAAGCCCTCGATCGTCTCTTACGAAAACTAACTTCTCATTGGAGCTTCAATCCATAAACGGACATTTTC  
ATTGAATTCCTAACCCCAAAATAGATAGGTTAGACATATAACTTGTCCAATTCGGGTCGGGTTCCACACCCGGCT  
GGTGGCATGCAAAACGAGGTAAATCTCGACAGTCCGATCTCGATTTCTGTGCCGATCACCACCATAGGCAATATA  
TATATCAAACCTTCTCCATTTTCTGAGTATATTCAAAAATGACAACAAAGCACAATCCAACGGCCGGATCAAAAA  
GTCAAACCTTAAACTCAAAATATCCAAAAATGGCAACTTTCAATCCAGGAAAAATGGGGTTGAGATATATATACCTT  
AATGATCCTTGAACGAAACGAAATTCACGGGACCGGTTAGAGGTGCAATGGTGGCCGGAATCGTCGCCGTGACG  
AGAAACCGATTTCGGCCTCTATAGAGATTCTGGCAATGCAGACGTTGGCGGCTCGGAGTTTACGCGCGCCTGG  
CGAGTCGACGCCGTCGGTGGCATGGCGGACGGTGGCTGGCGCGGCTCGGCATGTCTGCGGTGGTGGCGCG  
AGAGAGCTGAGAGAGAAGAGAGAAAAATCGGGGAAGAAGAGAGAGATGAGAAAAAGAAAGGACAGGCCACATGGCC  
ATATCTTATATAGCTGTGCAATTTTCAAAAAATGCACACAAGTCCCAAAATTTCAAATTTCTATCAATCAAGTCT  
CCTCAAAATTTACAATTTAATCCCTCTGTATAAAATTTTACAATTAATCCTCACAATATTATATCCAGAAAAAT  
AGTAAATTGTCTAAATTACCTTCGAAAATTCGGGTTTTTACACATCCAACAAGATTGACATACTCTATGTGTAT  
TTTTTGGCTTTACATATTTGACCCCAACACCCTTAGCCAAAACAGTAAGACAACTTGAATATGAGATTTTTTTG  
CGGAAGAGTGAGATTATCAATTTGATTCTAGCTAATCATTTTAGCTCGTACTTGTTCGTAACCTATTTAACCAA  
ATATCATTTTAAATGTGCGTTAAGATTAGAATACATCGATTCTATTTTAGCCAGCCACTTAAGACAATAGTTATTA  
TATATATTCATGAAGGTATGTATTATCTCTTACAAAGAGGATCATTGTTTTTATTTATAAAATATTTATTCACAT  
TAGGTAAAATTTGGGTAAGTATCTTGATTAATATCTAAAATAAATCAAGAAAAATGACATTGTGACTCGATTACT  
CAAAAATAATTATAATTAGTTATTTAATTAATAGCGTTTCTTAACATAATAGTTATTATATATATATATATATA  
TATATATATATATATTTACGAGAGGTATGCATTATCTCTATAAAGAGGATAATTATTTTTTATTTATAAAATATTT  
GTTCACTATAGTTAAGTTAGGTAAAGTATCTTAATTAATCCAAACATAAATCAAGAAAAATGCTATATGTGACT  
GGATTACTCAAAGTAATTATAATTAGTTATTTAATTAATAGAGTTTCTTAACATAATATTAACACAATAGTTATTA  
TATATATATATATATATATATATATATATTTACGAGAGGCATGTATGTATTATCTCTTACAAAGAGGATAATTA  
TTTTTATTTATAAAATATTTGTTAGAACCCTCTCATTAGTTATTTGCTCATAAACACGCCGATTTCTTCTTCATTC  
TAAATATTTCTAGGTTTCTCTACCTATTTACTCTATATAGTTTTATTGACTTTTTATTCTTTCTTATTTCT  
TTTGACGGCCACAACCACTCTACTACGAATTCGAATTTCAATGGTGACTTTTCCAGCATGATCAAGTGAGAAATTT  
AATTGATCTAAACACGTGATGCATTGATTAAGAATTAAGACTCTTATTTCAAAAAGAAATGAAAGAGGATTAATACT  
CGTTTTTCTATAACAAAACATGTGGATTAGATTCCTAAAGAGAGAAAGCAGCATAATAACACCTCGAGAGGTTCTT  
ATCACTGGGGGCTTTGGCTCTTTACCTACTCTTTTGACTATGGGTCTTGATGGCAATTCTGAAATGCATTGG  
GTATCTAATGCCATTTTCACAGAAGGGTGGCTGTTTGCTATTGAACTTCCGTTTCCATGCAAGGGCACTTTTGTC  
TAAAAAAAAGTACAGCCATACAAACACAGAGTTAACGACTTAGCGGTAGAAATTTGCTATTGGGTGCGGATGGA  
TAATAAGATAGGGATGTGAACACAGAGGAAAGTTGGAGAAAAAGCAGTCATTCTGCTGAGTGGAAGTATACCTT  
TACTGCCCTTAATCTTGCTAGCTTTATTTATAGTTTATTGTTTAAATCATTAAATACTGTAATATATTTTTGTATT  
TTATTTTATTTATTTTTAGAAAAAGGTTCTGTAAAGGATAGACTATGGACAAGATAATTGGCCTGTGCTAGCAG  
ACATCTCCACCATATTATTATCCCTCTCATTTTTTTTTTTGTTTGGGGTCTTTCTTTAATTTCTTCTTCTTCT  
TTAAATGTCATGAATCCACTCTTTTAGTTTCTTTCTCTCTTAAACCTGCTTCATCCACTACTGTTCTGATGC  
ATATTTTTTCAACTCCTGTGCAAGATCACAACAGTTTAGTTTACTAAACAGTCAAACCTCATACACCCACATGCT  
TAAATTTCTTTTGTGTTTGAATTCAGGGATATAAGGCTCTTGAATACTAGATTATTGTGAGATTCTTTGTTTAT  
GATTGATAAGTGAAGAAAAAAGAACTGGGAATCTTTTTGTTGTTACTTTGGCCATTCTTGGTTCTGGG  
CTTTGGTTGGTAGCGAAACAGGATAAAATCGAGTTCTGCTTGGAAATAGAAACCGAATCTTGGGATTCTTTGAT  
TCTAACTTGTGGGTATGTGAATTTATTGTCTTGGTAATATTGGGTTTTCTCTTGCAATTTCCCAAGTTGAAGTTTTT  
TGGGTGGATTGGATAAGATAGCATCTTTGTTATTGTGAATTTGGGGTTGTGGTGGTGGCTTTGTTATTATGCTCAT  
TTCTTCTGCTCTTCTGTTTACAAAGTTTGAACTTTTGTCCAAATTGTGTTGAAGCAGACATTTTCCCTTCTT  
TCTCTAGATTGATAGTGAAGTCTCTCTAATCTCTGACCCTGACTCTAACGCACACGCACGCGCGGGCGCACACA  
TCGTTTTACCCACCTGGAATTGGCAAGGAGGTGACACCCAGCTCTAAATTTGGTACTGTGACGAAGTTTAACTTTGG  
TTGGTTACTGCTGTGG

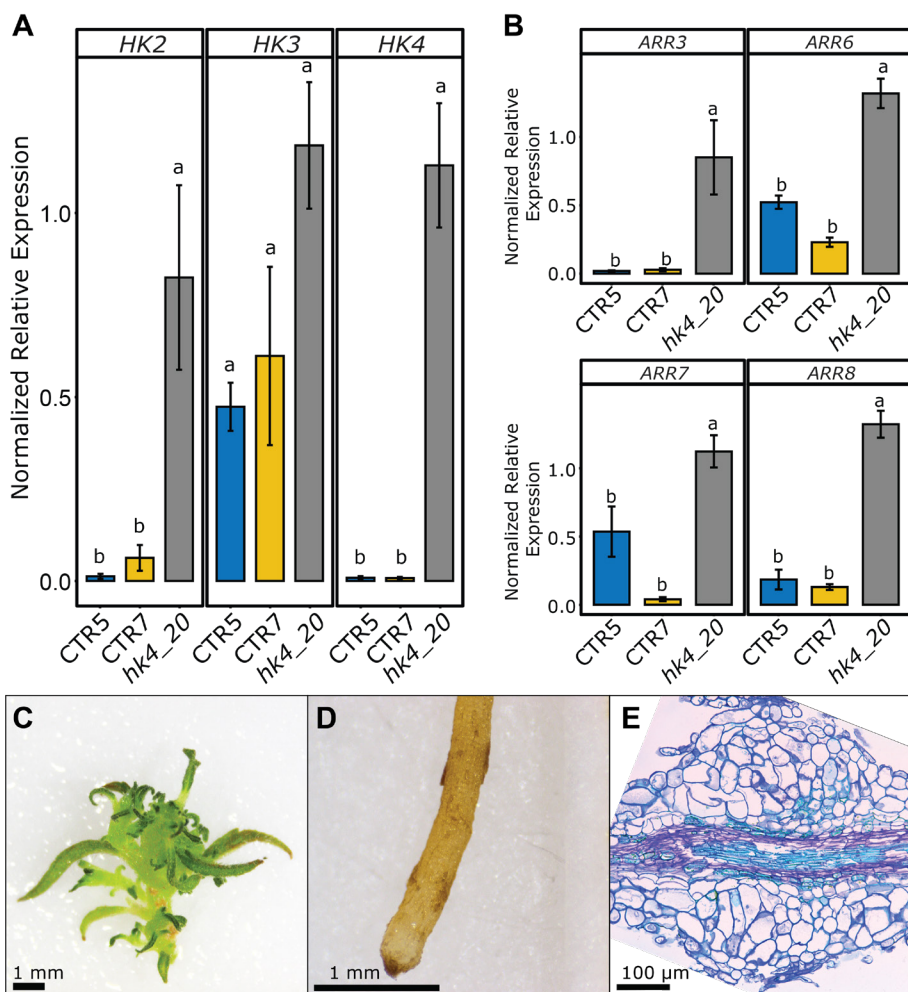
178



**Supplemental Figure 1. Expression of genes related to cytokinin biosynthesis, perception, signalling, and inactivation at different stages of nodule development in *P. andersonii*.** In general, the expression patterns showed that cytokinin homeostasis genes are more present at the beginning of nodule development, while it is decreasing as nodules are developed.



**Supplemental Figure 2. Amino acid alignment of the CHASE domains of the HK4 orthologues cytokinin receptors.** The single amino acid replacement (L263F) in the constitutively active *PanHK4\** is highlighted in red, which reflects the same amino acid replacement previously described in LjLHK1 autoactive/*Ljtnf2* (Tirichine, et al., 2007). *At*, *Arabidopsis thaliana*; *Lj*, *Lotus japonicus*; *Mt*, *Medicago truncatula*; *Pan*, *Parasponia andersonii*.



**Supplemental Figure 3. Constitutively active *P. andersonii* HK4\* line induces cytokinin perception and signalling, while affects plant growth and development.**

**(A)** A significant induction of *HK4* expression in the constitutively active *Panhk4\** line. The expression level of *HK2* is also increased, while the expression level of *HK3* is not statistically different compared to controls.

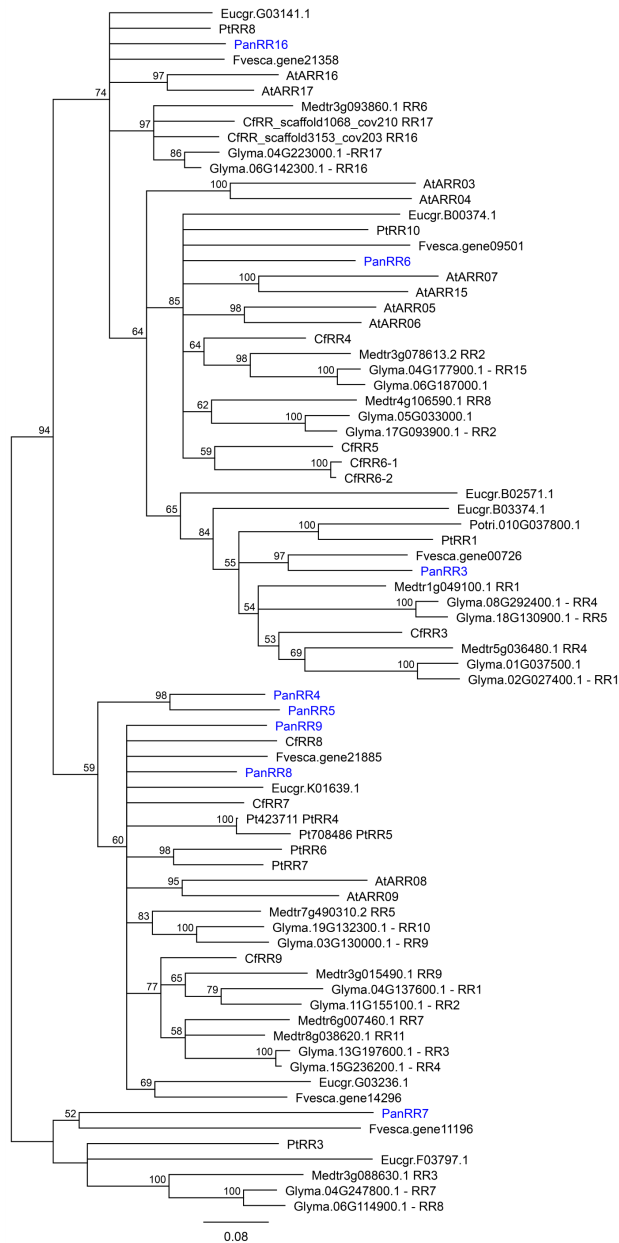
**(B)** A significant induction of cytokinin responsive genes (type A-RRs) in the constitutively active *Panhk4\** line.

**(C)** General size of shoots of the constitutively active *Panhk4\** line, which is maintained via *in vitro* propagation. Plantlets that are generated from this line are mostly dwarfed.

**(D)** General size of roots of the constitutively active *Panhk4\** line, which is maintained via *in vitro* propagation. Lateral root development in this line is heavily impaired.

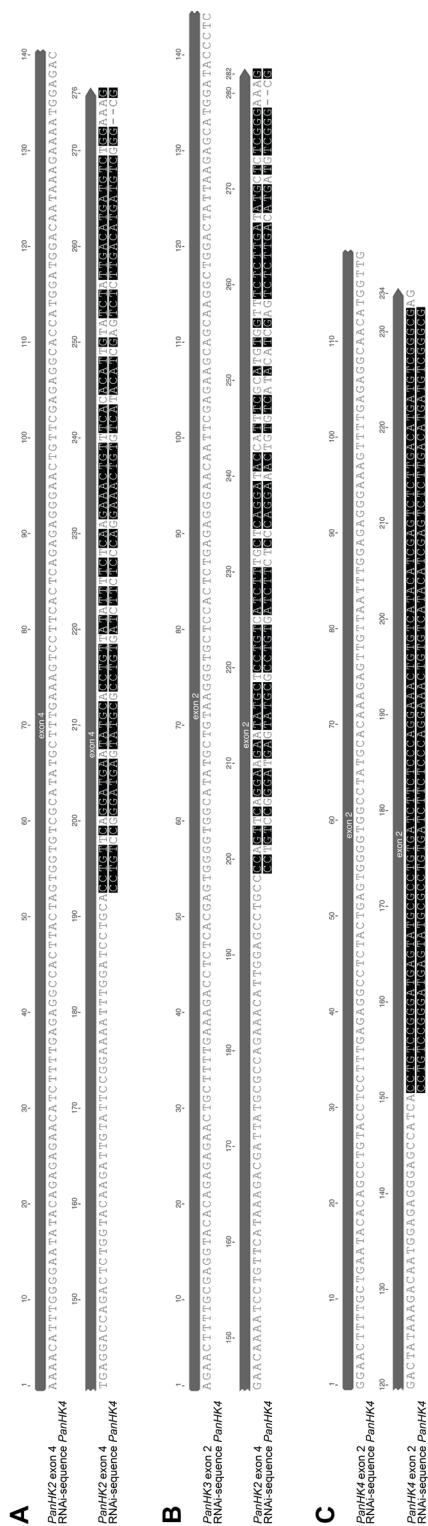
**(E)** Sectioned image of a root segment of the constitutively active *Panhk4\** line, which shows cell size enlargement.

In **(A-B)**, data represents means of 3 biological replicates  $\pm$  SEM. Statistical analysis was performed by One-way ANOVA and post-hoc Tukey HSD method ( $p < 0.05$ ), and samples that were deemed statistically different were labeled with different letters. In **(E)**, the plastic embedded sample was sectioned to 5  $\mu$ m thickness and stained with 0.05% Toluidine Blue.

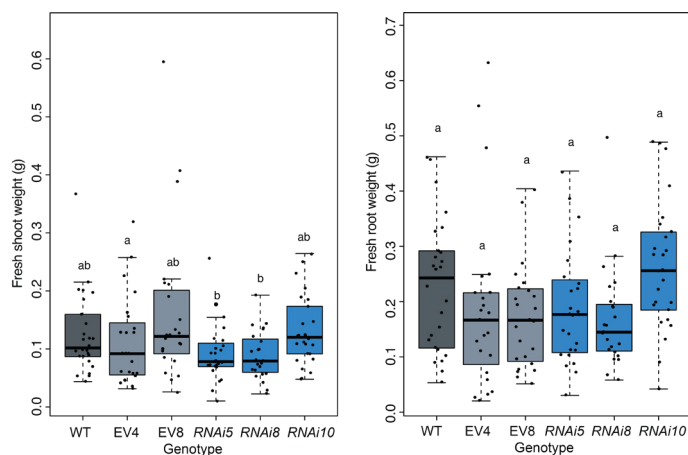


#### Supplemental Figure 4. Phylogenetic analysis of Type A-RRs.

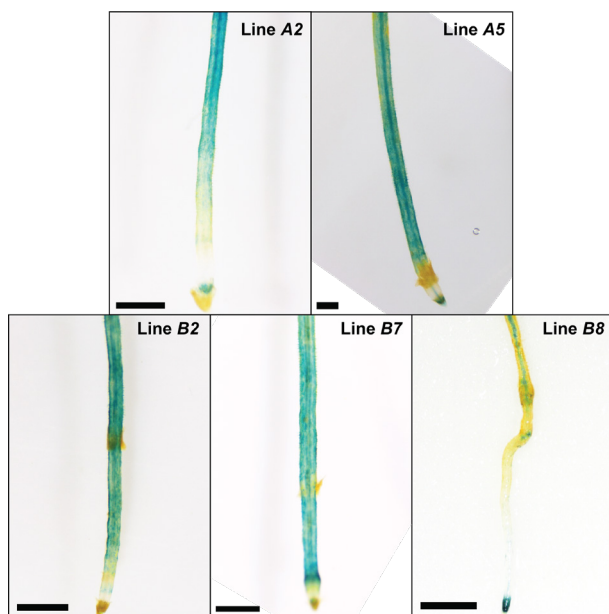
Protein alignment was performed using MAFFT alignment method and phylogenetic tree was reconstructed using Geneious Software v 8.1 (FastTree). The protein alignments used in this analysis were from following species: *Parasponia andersonii* (Pan), *Arabidopsis thaliana* (At), *Medicago truncatula* (Medtr), *Glycine max* (Glyma), *Chamaecrista fasciculata* (Cf), *Populus trichocarpa* (Pt), *Fragaria vesca* (Fvesca), *Eucalyptus grandis* (Eucgr). *P. andersonii* Type A-RR proteins are highlighted in blue colour. Branch label shows percentage of consensus support.



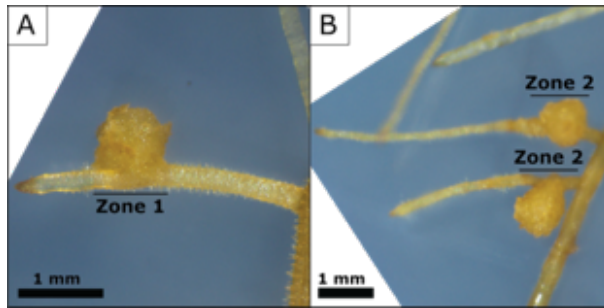
**Supplemental Figure 5. Nucleotide alignment of the RNAi fragment against the CHASE domains of the three histidine kinase cytokinin receptors of *P. andersonii*.** (A) Alignment with the exon 4 of *PanHK2* resulted in 78% of homology identity. (B) Alignment with the exon 2 of *PanHK3* resulted in 70% of homology identity. (C) Alignment with the exon 2 of *PanHK4* resulted in 100% of homology identity.



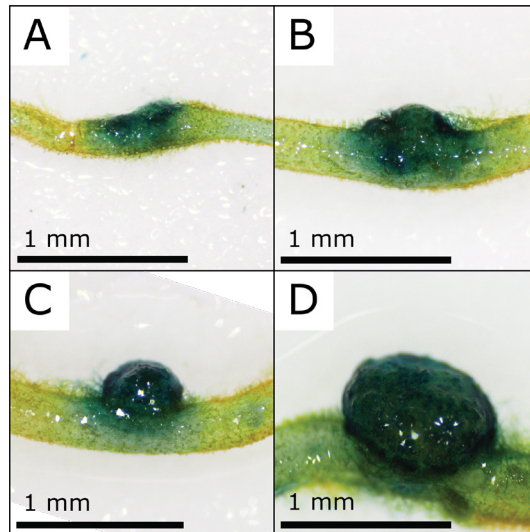
**Supplemental Figure 6. Fresh root and shoot weight of *P. andersonii* *hk4*-RNAi lines compared to controls (EV, empty vector control for RNAi) and wild-type (WT).** Rooted plantlets were grown in a closed pot system for 4 weeks, supplemented with modified-EKM medium containing 0.375 mM  $\text{NH}_4\text{NO}_3$ . Fresh shoots and roots were harvested and weighed accordingly. Statistical analysis was carried out using One-way ANOVA ( $p < 0.05$ ) and followed by Post-hoc Tukey HSD method. Data represented means of 20-30 samples from 3 independent experiments  $\pm$  SEM.



**Supplemental Figure 7. GUS-activity of five independent *TCSn:GUS* lines in non-inoculated roots.** Line A2, A5, B2, and B7 showed more similar GUS-activities, of which was observed on the root tip where the cell division program is active, continued to the elongation and maturation zone. Line B8 showed a strong GUS activity on the root tip and the inner region of maturation zone, while a very weak expression on the elongation zone, which may suggest less active of *TCSn:GUS* expression. All lines were grown in a square plate containing EKM-modified agar for 4 weeks.

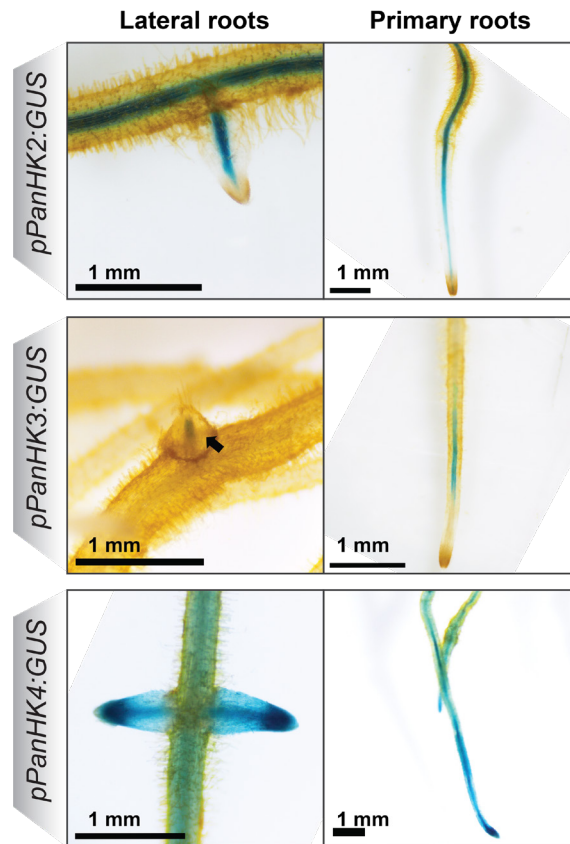


**Supplemental Figure 8. Definition of the 'root susceptible zone' in *P. andersonii* lateral roots.** Zone 1 (A) and Zone 2 (B) are part of lateral roots where the nodules are usually formed. Nodules are scored 4wpi with *M. plurifarium* BOR2.

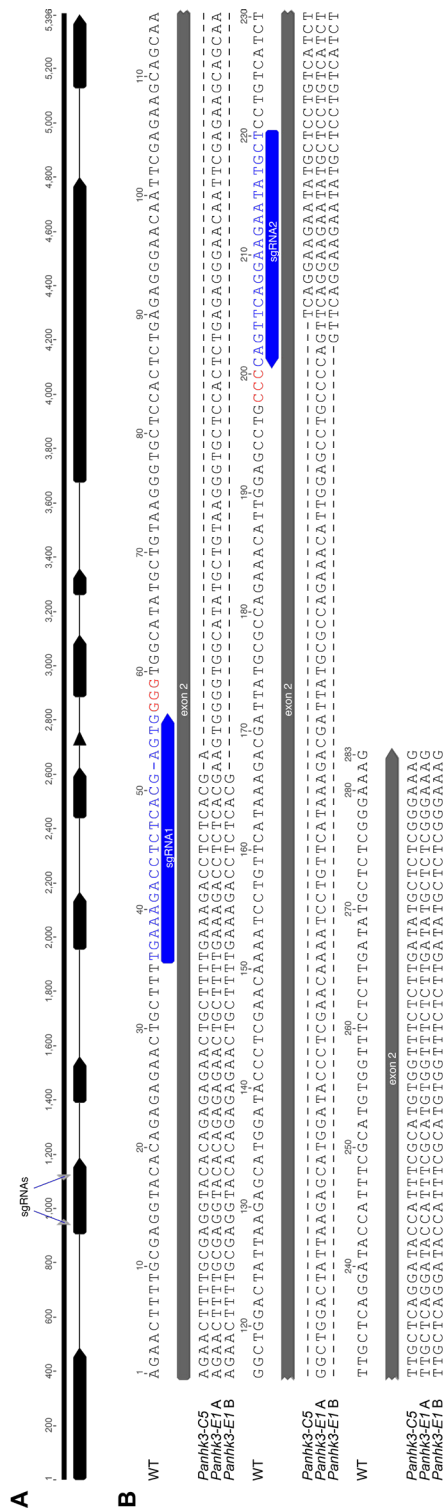


**Supplemental Figure 9. Activity of the *TCSn:GUS* reporter following rhizobial inoculation of line A2.**

Inoculation with *M. plurifarium* BOR2 for 4 weeks induces root nodules in *TCSn:GUS* line A2, in which GUS-activity was observed in nodule primordia (A-B), young nodules (C), and mature nodules (D).



**Supplemental Figure 10. Spatial expression of *Parasponia* cytokinin receptors in emerging lateral roots and primary roots.** In emerging lateral roots (left panels), GUS-activity of *pPanHK2:GUS* and *pPanHK4:GUS* are strongly detected in vascular bundle, while the expression of *pPanHK3:GUS* is lowly detected on the lateral root tip (indicated by an arrowhead). In primary roots (right panels), the *pPanHK2:GUS* and *pPanHK4:GUS* reporters are active in the vasculature and meristematic zone. The *pPanHK3:GUS* reporter is only expressed in the vasculature of the maturation zone of the primary root.



**Supplemental Figure 11. CRISPR mutant alleles of two independent *P. andersonii* *hk3* mutant lines.**  
(A) Schematic representation of the *PanHK3* gene model. Indicated by a grey arrowhead are the position of 2 sgRNA sites.  
(B) Nucleotide alignment of the third exon of *PanHK3* in wild-type (WT) and *Panhk3* mutant lines C5 and E1.



# Chapter 6

## **General Discussion:**

### **Synergy of Cytokinin Signalling and Transcriptional Signalling Networks in Symbiotic Rhizobia Association**

**Titis A. K. Wardhani<sup>1</sup>**

<sup>1</sup>Laboratory of Molecular Biology, Department of Plant Sciences, Wageningen University & Research, Wageningen, The Netherlands

*“Ing ngarsa sung tulada, ing madya mangun karsa, tut wuri handayani”*

(An educator should set a good example for the students, an educator should stimulate the creative ideas among students, an educator should be able to provide encouragements to the students).

—Ki Hajar Dewantara, Indonesian Educator.





## Current evolutionary perspectives on the nitrogen-fixing symbioses

Access to biologically active nitrogen is often the limiting factor for plant growth and development. To overcome nitrogen shortage, plant species belonging to ten taxonomic lineages can establish a nodule symbiosis with nitrogen-fixing bacteria. These diazotrophic bacteria can be divided into two groups; gram-positive filamentous *Frankia* sp. (*Actinobacteria*), and gram-negative rhizobia (alpha and beta-proteobacteria) (Young and Haukka, 1996; Svistoonoff et al., 2014; Sprent et al., 2017). The nitrogen-fixing symbioses occur exclusively in four orders of angiosperms (Fabales, Fagales, Cucurbitales, and Rosales), which define the so-called ‘nitrogen-fixing clade’ (Soltis et al., 1995). The symbiotic associations with rhizobia appear abundantly in the legume subfamilies Caesalpinioideae and Papilionoideae (Fabaceae, order Fabales), and in the genus *Parasponia* sp (Becking, 1992), which belongs to Cannabaceae (order Rosales). Plants in the remaining eight nodulating clades establish a mutualism with *Frankia* bacteria, and are named actinorhizal plants.

Earlier phylogenetic studies suggested that the nodulation capacity evolved independently multiple times (Soltis et al., 1995; Doyle, 2011; Werner et al., 2014). The restriction of nodulation to the nitrogen-fixing clade and interspersed with non-nodulating species, was explained by hypothesizing a predisposition event in the root of the nitrogen-fixing clade that allowed plants to gain nodulation (Soltis et al., 1995; Werner et al., 2014). However, recent comparative phylogenomic studies questioned this hypothesis. The studies uncovered that genes dedicated to nodulation, such as *NFP*, *NIN*, and *RPG* are lost or pseudogenized in the non-nodulating species within the nitrogen-fixing clade (Griesmann et al., 2018; van Velzen et al., 2018). This suggests that the loss of these genes resulted from the loss of nodulation trait. Given our current understanding, it is assumed that the evolutionary trajectory of root nodule symbiosis is that of a ‘single gain-parallel loss’ scenario, meaning that the nodulation capacity originates from a single common nodulating ancestor, while it encountered massive parallel loss in many descending lineages (Griesmann et al., 2018; van Velzen et al., 2018).

The occurrence of massive parallel loss in nodulation trait can be best explained by the environmental pressure, which may have occurred at a global scale. Such selective pressure could be a decreasing level of CO<sub>2</sub> atmospheric in the last ~110 MYA, in which the mentioned period coincides with the origin of the nitrogen-fixing clade. The nitrogen-fixing symbiosis is an energetically demanding process, which is a significant trade-off for the plant. Therefore, when atmospheric CO<sub>2</sub> levels were the limiting factor for photosynthesis, it became less favourable to maintain the nitrogen-fixing trait (van Velzen et al., 2019).

Assuming a single gain of the nodulation trait, at least two switches of nitrogen-fixing microsymbiont must have occurred (van Velzen et al., 2019). Since some basal *Frankia* strains demonstrate an intrinsic feature that protects the nitrogenase from oxygen exposure, as well as was found to produced similar lipo-chitooligosaccharide (LCO)-based signal molecules as rhizobia, it is argued that it was the original microsymbiont of the common ancestor of all nitrogen-fixing nodulation (Silvester et al., 2008; van Velzen et al., 2019). Subsequently, horizontal gene transfer may have resulted in the recruitment of LCO biosynthesis and nitrogen fixation genes by different proteobacteria (Persson et al., 2015; Remigi et al., 2015; Garrido-Oter et al., 2018), leading to the competition with *Frankia*. It is hypothesized that an ancestral legume and *Parasponia* have acquired independently rhizobium as microsymbiont (Remigi et al., 2015; van Velzen et al., 2019). Such symbiont switch between *Frankia* and rhizobia, may have resulted in specific adaptations in the host plant. One such scenario is the recruitment of plant hemoglobin genes to control oxygen homeostasis in the nodule. This adaptation is required for rhizobia to protect the nitrogenase from oxidation, since rhizobia lack the oxygen-regulating features that are present in *Frankia* (Sturms et al., 2010; van Velzen et al., 2018).

## Comparative studies as a tool to unravel key genetics in symbioses

The nitrogen-fixing symbioses require *de novo* organogenesis of root nodules. These nodules are the result of synchronised processes of perception of bacterial signals, infection events and nodule organogenesis (Oldroyd, 2013). Despite variation in rhizobial entry, nodule development and ontogeny across nodulating species, the root nodules serve a consensus function, which is to accommodate the energy source to bacteria in exchange of fixed nitrogen.

The potential of nitrogen-fixing plants to contribute to a sustainable agriculture is one of the driving forces in the research community to engineer the nodulation capacity into non-nodulating crops. To achieve this objective, it is crucial to decipher which genes are essential for the perception of bacterial signals, infection processes, and nodule organogenesis. Additionally, it is equally important to identify the mechanisms by which the genes are regulated. Genetic studies conducted in model legumes *Medicago truncatula* and *Lotus japonicus* (both belonging to the Papilionoideae subfamily) have revealed a great complexity and numerous molecular components underlying the nodulation trait (See review in (Roy et al., 2020)). However, in the light of engineering this trait, the exploration of only these models will result in a legume-biased knowledge of the symbiosis. Therefore, expanding the observations beyond the legume Papilionoideae subfamily is highly relevant.

Molecular insights in the nodulation trait of non-legumes can also be used in comparative studies. Such studies can be used e.g. to identify ortholog genes that are a conserved expression profile in nodules. By comparing the transcriptome of *M. truncatula* and *Parasponia andersonii*, ~290 shared nodule-enhanced genes could be identified (van Velzen et al., 2018). This suggests that the root nodulation in both species is orchestrated by the same symbiotic genes. Some of these genes have been identified also in forward genetic approaches in model legumes; e.g. like the transcription factors CYCLOPS and NODULE INCEPTION (NIN). However, for many these conserved nodule enhanced genes functional analysis has not been done. Therefore, incorporation of the omics studies with the genetics in models and non-models is essential to obtain its mechanistic insights and to identify the symbiosis genes that belong to the core genetic network controlling nodulation. Further, comparative genetic analyses can highlight the information on the specific adaptations that occurred later in a particular clade. These specific-lineage adaptations provide insights on the degree of flexibility of how the symbiotic signalling is integrated, and therefore opens the possibilities to incorporate features from multiple symbioses (Huisman and Geurts, 2020).

During nodule initiation and development, rewiring of plant hormonal signalling is one of the major targets of the symbiotic signalling pathway (Ferguson and Mathesius, 2014). Amongst many determinants, cytokinin signalling has been demonstrated to be essential and sufficient for nodule organogenesis (Cooper and Long, 1994; Fang and Hirsch, 1998a; Mathesius et al., 2000). Studies that dissect cytokinin signalling in relation to nodulation are mainly conducted in Papilionoideae legumes, especially using *M. truncatula* and *L. japonicus*. In this thesis, I questioned whether the recruitment of cytokinin signalling in root nodulation is a common feature within the nodulating species.

To perform comparative genetic analyses, it is important to consider several criteria in choosing new experimental systems. This includes the genetic diversity of the nitrogen-fixing clade. Also, the selected species should be suitable for laboratory experimentation, which includes reproducibility for seed production, synchronized seed germination, *in vitro* propagation, a homozygote and relatively small diploid genome, and protocols for regeneration and transformation (Huisman and Geurts, 2020). The latter is very important, as lack of efficient transformation protocols often serves as a bottle-neck in reverse genetic approaches. In this thesis, I presented an example of using two species that are distantly related from each other and represent the genetic diversity in the nodulation trait; *Chamaecrista mimosoides* (Caesalpinoideae, Fabales) (**Chapter 2**) and *P. andersonii* (Cannabaceae, Rosales) (**Chapter 3-5**).

*Chamaecrista* represents the basal legume subfamily Caesalpinoideae. This means that experimental studies on *C. mimosoides* may provide insights on the adaptations that led

to legume-type nodules. *C. fasciculata* has been proposed previously as a comparative system (Singer et al., 2009). The genome sequence of *C. fasciculata* is recently determined (Griesmann et al., 2018). However, *C. fasciculata* turned out to be difficult to reproduce under laboratory conditions (Lee and Bazzaz, 1982) and the transformation efficiency by *Agrobacterium rhizogenes* is relatively low (Wouter Kohlen, data not shown). This resulted in difficulties to perform functional genetic studies in this species. Therefore, we selected another species; namely *C. mimosoides*. *C. mimosoides* can be crossed under the laboratory conditions for seed production. In **Chapter 2**, I developed a hairy-root transformation protocol mediated by *A. rhizogenes*. Using this protocol, >90% of the roots are co-transformed. This shows that reverse genetic analysis is possible in *C. mimosoides*.

*Parasponia*, which is the only non-legume that can establish interactions with rhizobium, may serve as an alternative species to study the nodulation trait. Although *Parasponia* is a tree and exclusively distributed in eroded soils of volcanic areas (Becking, 1992), the plants can be grown under the laboratory conditions for seed production and protocols have been developed for efficient *in vitro* propagation. In **Chapter 3**, we successfully established a stable transformation protocol mediated by *Agrobacterium tumefaciens*. This protocol will allow the generation of transgenic lines in a 6 weeks' time frame. Subsequent *in vitro* propagation of selected T<sub>0</sub> lines will allow experimentation 3-4 months after start of the transformation process. In **Chapter 4**, we demonstrate that *Parasponia andersonii* is suited for efficient CRISPR/Cas9-targeted mutagenesis. This indicates that this plant species can be a robust experimental model for reverse genetic studies. In contrast to legumes, *P. andersonii* did not experienced recent whole genome duplication events. Therefore *P. andersonii* is a good model to identify the core gene set for rhizobial-induced nodulation.

Taken together, the development of transformation protocols for *C. mimosoides* and *P. andersonii* described in this thesis, provide important tools to study the functional conservation of nodulation genes. Using these tools, I performed investigations on the symbiotic role of cytokinin signalling in nodulation. Here, I will present the knowledge gained from this PhD project and further examine whether specific lineage adaptations also determine the nodulation capacity.

## Interplay of NIN and cytokinin signalling at early symbiotic interactions

Nodulation in legumes is initiated upon perception of rhizobium secreted lipochitooligosaccharide (LCO) signals, which are also known as Nod-Factors. Nod-Factors are recognized by a complex of LysM-type receptor kinases (RLKs) that are named MtNFP/MtLYK3 in *M. truncatula* and LjNFR1/LjNFR5 in *L. japonicus* (Limpen,

2003; Arrighi, 2006; Smit et al., 2007). Initially Nod-Factors signalling is still restricted to the root epidermis. Nod-Factors perception triggers activation of Common Symbiotic Signalling Pathway (CSSP, **Chapter 1**) that includes the initiation of a secondary signal; nuclear calcium spiking (Granqvist et al., 2015). This result in the transcriptional activation of several transcription factors (*NIN*, *ERN1*, *NSP1*, *NSP2*), which are essential for rhizobial infection and nodule organogenesis (Hirsch et al., 2009; Cerri et al., 2012; Vernié et al., 2015; Kawaharada et al., 2017a; Liu et al., 2019d). Shortly after the perception of Nod-Factors, elevated levels of the cytokinin isopentenyladenine (iP) and trans-zeatin (tZ) are detected in the root region where the root nodules are generally formed (van Zeijl et al., 2015a; Jardinaud et al., 2016). This cytokinin accumulation is in line with a strong induction of several cytokinin biosynthesis-related genes, such as: *IPT3*, *CYP735A*, *LOG2*, and some primary cytokinin type-A *RESPONSE REGULATOR (ARR)* genes (van Zeijl et al., 2015a; Jardinaud et al., 2016; Reid et al., 2017). Several studies on cytokinin biosynthesis genes have been performed in legumes. *L. japonicus* mutants in *LjIPT4* or knockdown of *LjIPT3* showed a reduction in nodule numbers (Chen et al., 2014; Reid et al., 2017). Also knockdown of *MtLOG1* in *M. truncatula* reduced nodule numbers (Mortier et al., 2014). In contrast, no such phenotypes are found in the case of a knockout mutation in the cytokinin biosynthesis gene *LjCYP735* (Reid et al., 2017). The importance of cytokinin in early Nod-Factors signalling became apparent in transcriptome studies using the *M. truncatula* cytokinin receptor mutant *Mtcre1*. This mutant shows a severe reduction in the number of Nod-Factors responsive genes (van Zeijl et al., 2015a; Jardinaud et al., 2016; Schiessl et al., 2019).

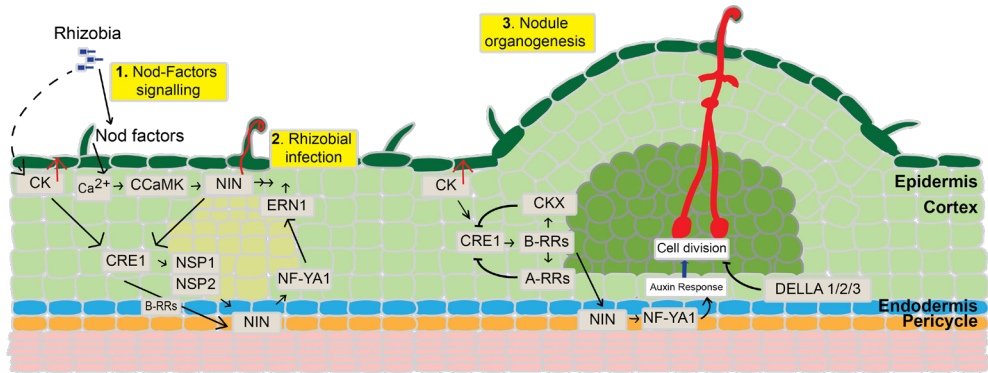
To determine whether Nod Factors-induced cytokinin signalling is a conserved feature in legumes, I studied early rhizobial induced responses in *C. mimosoides* roots. Plants inoculated with *Bradyrhizobium* sp. WUR077 (24 hpi and 48 hpi) showed that the cytokinin *ARR* genes; *ARR4* and *ARR9*, are transcriptionally induced. This induction coincided with the first cell divisions (**Chapter 2**). Unlike *M. truncatula* and *L. japonicus*, the first rhizobium-induced cell divisions in *C. mimosoides* are occurring in the outer most cortical cells and endodermis. (**Chapter 2**). This result shows that the cytokinin signalling is Nod-Factors responsive in a nodulating basal Caesalpinoid legume.

I aimed to conduct similar studies, also in the nodulating non-legume *P. andersonii*. Unfortunately, data Nod-Factors induced gene expression could not be generated. This was due to the difficulties in defining the Nod Factors-responsive regions in *P. andersonii* roots. Nevertheless, evidence that demonstrates the activation of cytokinin signalling in *P. andersonii* upon the Nod-Factors perception was obtained using the synthetic cytokinin-responsive promoter *TCSn*. *TCSn* contains several binding motifs for the cytokinin activated Type B-RR transcription factors (Zürcher et al., 2013). The *TCSn* promoter has been previously used in *M. truncatula* and *L. japonicus* to study

Nod-Factors-induced responses. This experiment showed that *TCSn* is mainly induced in dividing cortical cells (Jardinaud et al., 2016; Reid et al., 2017). In contrast, in *P. andersonii* *TCSn:GUS* was first detected in the epidermis of young lateral roots upon inoculation with *M. loti*.pMP604 or *M. loti* Nod factors at 24 hpi. At a later stage, *TCSn* activity is also observed in dividing epidermal and outer cortical cells that will form a nodule primordium. Such an *TCSn* expression pattern is unique to *P. andersonii*, since the first cell divisions upon rhizobium-induced signalling occur in the epidermis and most outer cortex (Lancelle and Torrey, 1984).

Studies in *M. truncatula* and *L. japonicus* showed that symbiotic functioning of cytokinin and NODULE INCEPTION (*NIN*) is intertwined. First, it was shown that *NIN* is induced by cytokinin, and exogenously applied cytokinin induced formation of nodule-like structures in a *NIN*-dependent manner (Vernié et al., 2015). Also, it was shown that *NIN* expression is cytokinin responsive and that the *NIN* gene of Papilionoid legumes, as well as the Caesalpionoid *C. fasciculata* (**Chapter 2**), possess a 500 bp conserved *cis* regulatory element that includes several type-B RR binding sites (Liu et al., 2019c). This putative Cytokinin-responsive Element (CE) is essential for nodule organogenesis (Liu et al., 2019c) and their formation involves both intracellular infection initiated in the epidermis and nodule organogenesis initiated in inner root cell layers. Interestingly, *NIN* induction in the epidermis is independent of the CE, but requires a *cis*-regulatory element containing a CYCLOPS binding site. Although *NIN* activation in the epidermis can occur in a cytokinin-independent manner, its activation is important for activation of cytokinin signalling in the cortical cells (Vernié et al., 2015). Also, induction of *NIN* in the cortex is dependent on the MtCRE1/LjLHK1 cytokinin receptor, whereas expression of *MtCRE1* is down-regulated in a *Mtnin* mutant background (Vernié et al., 2015). Taken-together these data suggest the occurrence of an feedforward activation loop between cytokinin signalling and cortical *NIN* expression (**Figure 1**).

Studies in *M. truncatula* and *L. japonicus* demonstrated the importance of the cytokinin receptor MtCRE1/LjLHK1 nodule formation. Loss of function mutants in these genes results in a dramatic reduction of nodule numbers (Murray et al., 2007; Plet et al., 2011). Similarly, RNAi of the *C. mimosoides* orthologous gene *CmHK4* results in decreased nodulation efficiencies (**Chapter 2**). In contrast, a gain-of-function mutation in *MtCRE1* and *LjLHK1* (*snf2*) results in nodule formation independent of the presence of rhizobium or Nod factors (Tirichine et al., 2007; Ovchinnikova et al., 2011). This demonstrates that activation of cytokinin signalling is sufficient to activate the nodule developmental program. Spontaneous nodule formation induced by expression of an auto-active the cytokinin-receptor MtCRE1<sup>(L267F)</sup>/Lj*snf2* requires the presences of GRAS-type transcription factors DELLA, NSP1 and NSP2 proteins, placing symbiotic cytokinin signalling upstream of these genes (Kaló et al., 2005; Tirichine et al., 2007; Madsen et



**Figure 1. A model for cytokinin signalling during rhizobial infection and nodule organogenesis.** (i) Nod-Factors signalling in epidermis induces NIN in a CCaMK-CYCLOPS dependent manner. (ii) Epidermal *NIN* activation is required for the induction of cytokinin signalling in the cortical cells. In the cortex, *NIN* activation is cytokinin dependent. The functionality of the cytokinin receptor MtCRE1/LjLHK1/CmHK4 in the cortical cells dictates the expression of *NIN*. (iii) Three GRAS-type transcription factors DELLA, NSP1 and NSP2 are essential for nodule development. (Adapted from Suzaki et al, 2013, and Roy et al, 2020)

al., 2010; Fonouni-Farde et al., 2016; Jin et al., 2016). Finally, analyses of a *L. japonicus* double mutant line of *Ljccamk*;*Ljsnf2* showed that spontaneous nodules are formed independent of CCaMK, indicating that the cytokinin signalling is positioned downstream of CCaMK (Madsen et al., 2010) (**Figure 1**).

A study in the *L. japonicus* *Ljlhk1* mutant showed that introduction of the *Arabidopsis thaliana* *AtHK4* gene under control of the *LjLHK1* promoter can functionally restore nodule organogenesis. This indicates that the legume cytokinin receptor protein did not experience adaptations allowing it to function in nodulation (Held et al., 2014). Also, in the case *AtHK4* is driven by its own promoter it can *trans*-complement nodule organogenesis, as was shown in *trans*-complementation experiments of the *M. truncatula* *Mtcre1* mutant. However, these nodules are affected in the infection process (Boivin et al., 2016). This data may suggest that cytokinin signalling is required also in the case of rhizobial infection, and that this requires specific adaptations in *cis* regulatory promoter elements of the *LjHK4*/*MtCRE1*/*AtHK4* cytokinin receptor gene.

We also studied the role of NIN and cytokinin signalling in nodulation of the non-legume *P. andersonii*. Transcriptomic data from different stages of nodules showed that several cytokinin biosynthesis-related genes, such as *PanIPT5*, *PanIPT9*, and *PanLOG8* are induced in nodule primordia. Cytokinin biodegradation-related genes, such as *PanCKX5* and *PanCKX7*, are transcriptionally activated as the nodules become more mature (**Chapter 5**). This data is further substantiated by the spatial expression of *P. andersonii* *TCSn:GUS* lines, in which the GUS signal was strongly observed on the dividing cells of epidermal and cortical cell layers, and was maintained in the nodule cortical cells

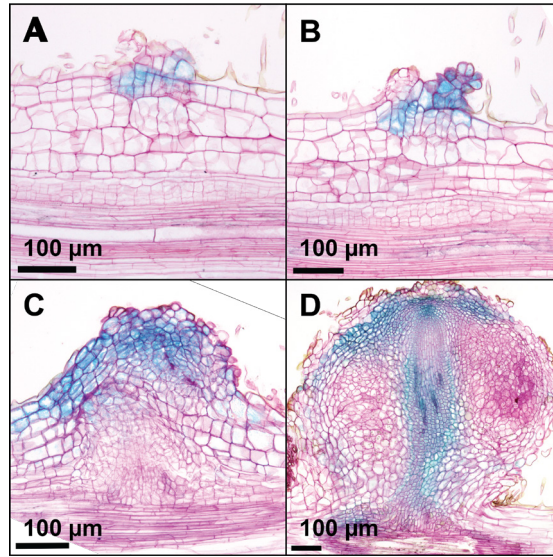
that are eventually infected. Like in legumes, *NIN* is essential for nodule initiation in *P. andersonii* (Bu et al., 2020). However, we could not observe a conserved CE *cis*-regulatory element in the putative promoter region of *PanNIN*. Also, when knocking out the *P. andersonii PanHK4* gene, which is orthologous to *MtCRE1/LjLHK1*, no nodule phenotype was observed (Van Zeijl et al., 2018; **Chapter 4**). This may suggest the occurrence of functional redundancy in three *P. andersonii HK* genes. Some level of functional redundancy was also observed in *L. japonicus*, a species that possesses four *HK* genes. Only knocking out three of these genes, *Ljlhk1;Ljlhk1a;Ljlhk3*, abolished the formation of cortical infections and nodule formation (Held et al., 2014). To obtain first insights whether also in *P. andersonii HK* genes commit symbiotic functions in redundant fashion, I conducted RNAi experiments. Only when partially knocking down *PanHK4* together with *PanHK2* and *PanHK3*, a reduction in nodule numbers was observed. This suggests that cytokinin signalling may play a role in *P. andersonii* nodule formation. However, some RNAi nodules were disturbed also in rhizobium infection. Therefore, we cannot rule out that the observed reduction in nodule numbers is a pleiotropic phenotype, which is the result from hampered rhizobium infection (**Chapter 5**).

## Other regulators in nodule organogenesis

Since root nodulation acquires reprogramming of cortical cells, therefore it is plausible that additional generic hormonal pathways have been recruited to function in nodule organogenesis. *In planta*, cytokinin and auxin work in concert in regulating shoot and root development (Fukaki and Tasaka, 2009; Su et al., 2011; Immanen et al., 2016). Experiments suggests that the onset of cortical cell divisions leading to nodule organogenesis is also the result of a fine tuning between cytokinin signalling and local auxin accumulation (Suzaki et al., 2013; Takahashi and Umeda, 2014; Xiao et al., 2014; Vernié et al., 2015; Reid et al., 2017).

The significance of auxin for nodule organogenesis is best supported by the finding that nodule initiation can be induced by the auxin transport inhibitors (Scheres et al., 1992; Fang and Hirsch, 1998b; Hirsch et al., 2009; Rightmyer and Long, 2011). The ability of auxin transport inhibitors to induce nodule initiation is also demonstrated by a partial rescue of nodulation of the *M. truncatula Mtre1* mutant (Ng et al., 2015). An activation of auxin responses is also observed in the spontaneous nodulating *Ljsnf2* mutant background, which implies that the induction of auxin is positioned downstream of cytokinin signalling and is activated in an LjLHK1-dependent manner (Suzaki et al., 2012).

Activation of auxin signalling during nodule organogenesis was also demonstrated by the using auxin responsive reporters, like DR5, GH3, or DR5/DII-VENUS (Pacios-



**Figure 2. Spatial expression analysis of the auxin responsive reporter DR5 at different stages of *P. andersonii* nodule development. (A-B)** At early stages of symbiotic initiation, the GUS activity was detected in a cluster of dividing epidermal and outer cortical cells. **(C)** GUS activity was observed in dividing cells of the nodule primordia, the emerging nodule vasculature and in the infection zone of the nodule (depicted by arrow). **(D)** In a mature nodule, the GUS activity was observed in the nodule vasculature and meristematic zone.

Bras et al., 2003; Suzaki et al., 2012; Nadzieja et al., 2019). Studies in *L. japonicus* and *M. truncatula*, showed that symbiotic auxin signalling is first observed in cell layers underneath rhizobium infected root hairs, and in cortical and pericycle cells that give rise to the nodule primordium (Pacios-Bras et al., 2003; Suzaki et al., 2012; Breakspear et al., 2014). In mature nodules, auxin signalling is maintained in the nodule vasculature and in the cluster of dividing cells at the nodule apex (Suzaki et al., 2012; Breakspear et al., 2014; Franssen et al., 2015; Nadzieja et al., 2019). Similar studies have been conducted also in *P. andersonii*. In *P. andersonii* DR5:GUS lines, a strong auxin response was observed in the dividing epidermal and outer cortical cells forming the nodule primordium (**Figure 2**). As the nodule becomes more mature, auxin signalling is restricted to the nodule meristem and vasculature. No signal was detected on the fixation zone of the nodule. This DR5 activity in *P. andersonii* is highly similar with the spatial patterns observed for the cytokinin reporter *TCSn* (**Chapter 5**), suggesting a spatial interconnection between cytokinin signalling and auxin response in nodule organogenesis.

Two recent studies in *M. truncatula* and *L. japonicus* exemplified the interconnection of cytokinin signalling to facilitate local auxin accumulation. By cytokinin activation of *NIN* expression, this transcription factor can activate expression of *LOB-DOMAIN PROTEIN 16* (*LBD16*) expression. LBD16 subsequently activates auxin biosynthesis via transcriptional induction of *STYLISH* (*STY*) and *YUCCA* (*YUC*) genes (Schiessl et al., 2019; Soyano et

al., 2019). These studies provided evidence that the nodule developmental program has recruited genetic components of the lateral root program, where LBD16 and STY/YUC are also essential for root primordium initiation (**Figure 1**).

Symbiotic cytokinin signalling was found also to interact with ethylene. A *M. truncatula* double mutant of *Mtcre1* and the ethylene-insensitive mutant allele *Mtsckl/ein2* resulted in suppression of hyper-nodule formation (Plet et al., 2011). Also, it was found that the burst of cytokinin molecules upon Nod-Factors induced signalling is enhanced in the *Mtsckl/ein2* mutant, when compared to wild-type roots (van Zeijl, et al., 2015a). Altogether, the data suggests that MtEIN2-dependent ethylene signalling acts as a negative regulator of symbiotic cytokinin signalling.

Combining the *M. truncatula della2* and *della3* mutants into *Mtcre1* gain-of-function mutant suppresses the spontaneous nodule formation (Jin et al., 2016). This phenotype can be explained since MtDELLA proteins regulate *MtERN1* and *MtNIN* expression during nodule organogenesis (Jin et al., 2016). In *M. truncatula*, DELLA1-mediated gibberellic acid signalling interplays with the CRE1-dependent pathway to regulate early nodulation in response to both Nod-factors and cytokinin signals, with dominant active DELLA induces the spontaneous nodule organogenesis (Fonouni-Farde et al., 2016). Other GRAS proteins that show positive regulation in nodulation are NSP1 and NSP2 (Kaló et al., 2005; Smit et al., 2005). Both NSP proteins are also required for strigolactone biosynthesis via regulating the *D27* and *MAX1* expression (Liu et al., 2011; van Zeijl, et al., 2015). In *P. andersonii*, mutations on *PanNSP1* and *PanNSP2* resulted in the abolishment of nodule formation (**Chapter 4**) (van Zeijl et al., 2018). This demonstrates that *NSP1* and *NSP2* functioning is conserved and essential for nodule formation, although their interconnection with the CK signalling still needs to be established in *Parasponia*.

## Emerging role of cytokinin signalling in rhizobial infections

In accordance with the involvement of cytokinin signalling in nodule initiation, several transcriptomic studies focusing on the epidermis- and root hairs-derived samples were performed to identify whether the cytokinin biosynthesis and signalling are also integrated in regulating early rhizobial infection (Liu et al., 2015; Damiani et al., 2016; Jardinaud et al., 2016). Propositions on the role of cytokinin in regulating the rhizobial infection came from several observations. Functional studies on *L. japonicus* cytokinin receptors, showed that the single *Ljlhk1* mutant or triple mutants *Ljlhk1;Ljlhk1a;Ljlhk3* suffered from hyper-infection in the epidermal cell. This suggests that cytokinin signalling negatively regulates epidermal infection (Murray et al., 2007; Held et al., 2014). Introduction of an epidermal-specific promoter driving *LjLHK1* genomic sequence did not complement the hyper infection phenotype in the *Ljlhk1* mutant. This suggests that this hyper-infection

phenotype in the *Ljlhk1* and the *Ljlhk1;Ljlhk1a;Ljlhk3* mutant background is an indirect effect, probably due to absence of cortical cell divisions (Miri et al., 2019). The *Ljlhk1* and *Mtcre1* mutants experienced a defect that abolishes progression of rhizobium infection into the cortical cells (Murray et al., 2007; Plet et al., 2011). This implies a positive regulation of cortical infection by cytokinin signalling. Such positive role of cytokinin signalling is also supported by *CmHK4 RNAi* studies in *C. mimosoides*, which showed ~10% of nodules having an impaired rhizobium infection (**Chapter 2**).

In *Parasponia*, the early infection event of rhizobia is accommodated by a crack-entry, which leads to small apoplastic colonies in between cells (Lancelle and Torrey, 1984; Bender et al., 1987). From there, cell wall-bound intracellular infection threads are formed in susceptible nodule cells. These infection threads ultimately branch into cell wall free fixation threads.

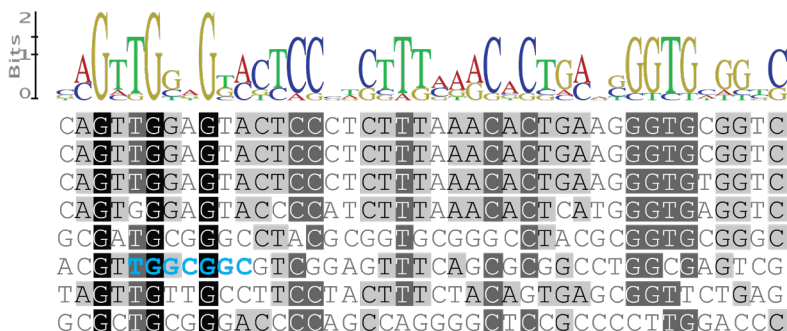
To dissect the spatial expression of the three *P. andersonii* *PanHK* genes (putative) promoter GUS reporter lines were created. Analysing these lines showed that only *pPanHK3:GUS* activity was associated with nodule initiation and in the nodule infection zone. This spatiotemporal expression is similar to the expression of *TCSn* in *P. andersonii*. In contrast, *pPanHK2:GUS* and *pPanHK4:GUS* lines showed GUS activity only in the nodule vasculature, and in case of *pPanHK4:GUS* in the outer cortical cells of mature nodules (**Chapter 5**). The overlapping expression of *pPanHK3* and *TCSn* in nodule primordia and the infection zone of the nodule indicates a putative role of cytokinin in *P. andersonii* nodulation. Subsequent CRISPR-Cas9 mutagenesis of *PanHK3* showed that this receptor functions in rhizobium intracellular fixation thread formation. (**Chapter 5**). About ~60-80% of *Panhk3* mutant nodules contain only infection threads as well as enlarged apoplastic colonies, but no fixation threads. A very similar nodulation phenotype is also observed in *Panin2* mutant nodules (**Chapter 4**). This suggests that both, cytokinin and ethylene signalling, act positively on the fixation thread formation.

In legumes, intracellular infection involves cell enlargement, which result from endoreduplication, a prerequisite for intracellular infection. Cytokinin signalling is proposed to be indirectly involved in controlling endoreduplication, by promoting the expression of the *CELL CYCLE SWITCH 52A* (*MtCCS52A*) gene (Takahashi et al., 2013; Tan et al., 2020). In *A. thaliana*, *AtCCS52A* is shown to be involved in the transition of mitotic cycles to endoreduplication cycles by acting as a substrate-specific activator of the anaphase-promoting complex (APC) ubiquitin ligase, which subsequently leads to the degradation of mitotic cyclins (Vinardell et al., 2003). Additionally, study in *A. thaliana* also demonstrated that ethylene inhibits cell proliferation in the root apical meristem through induction of endoreduplication (Street et al., 2015).

**A**

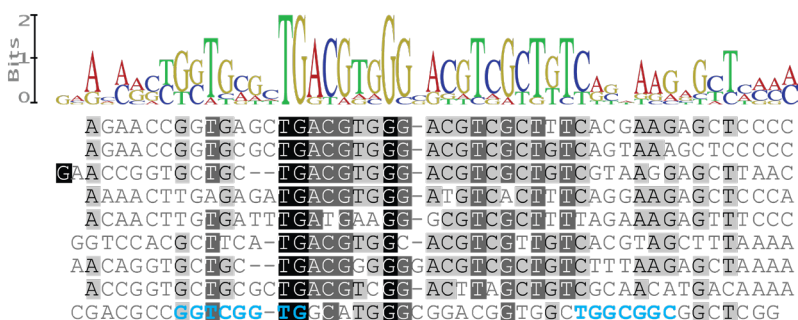
Sequence Logo

1. GmHK2-2
2. GmHK2-1
3. PvHK2
4. LjHK2
5. LjLHK3
6. **PanHK3**
7. PvHK3
8. LjLHK1A

**B**

Sequence Logo

1. CfHK3
2. TorHK3
3. PtHK3-2
4. LjLHK3
5. GmHK3-2
6. VvHK3
7. PtHK3-1
8. EgHK3-1
9. **PanHK3**



**Figure 3. Motif search analysis on promoter regions of cytokinin receptors from nodulating and non-nodulating species.** (A) and (B) show the predicted motifs from MEME analysis. Blue colour highlights the putative AP2/ERF binding domain (TOMTOM analysis), which is only conserved in the *PanHK3* promoter region. Species included in the analyses: Gm, *Glycine max*. Pv, *Phaseolus vulgaris*. Lj, *Lotus japonicus*. Pan, *Parasponia andersonii*. Cf, *Chamaecrista fasciculata*. Tor, *Trema tomentosa*. Pt, *Populus trichocarpa*. Vv, *Vitis vinifera*. Eg, *Eucalyptus grandis*.

The associated nodule phenotype of *Panhk3* mutant lines is a novel phenotype that is not previously observed in the legume-rhizobium interactions, as the main cytokinin perception in legumes depends on the functionality of *MtCRE1/LjLHK1/CmHK4* which controls the nodule organogenesis. In line with the emerging role of cytokinin signalling in rhizobial infection, the integration of *PanHK3* in nodule infection of *P. andersonii* offers a new perspective, which may also hint to lineage-specific adaptations; e.g. in the promoter of *PanHK3*. To this end, I analysed the promoter regions of *HK2*, *HK3*, and *HK4* genes of nodulating and non-nodulating species. This led me to find two conserved *cis* regulatory elements in the promoter of *HK3* and some *HK2* genes (Figure 3. A-B). The putative transcription factors that may bind to these domains remain unknown. However, when analysing *PanHK3*, I noted the occurrence of three putative AP2/ERF domains. Such domains may function as an ERN1 binding site (Cerri et al., 2017). The AP2/ERF domains are only found in the promoter region of *P. andersonii PanHK3* and are located ~2.7kb upstream of coding sequence. Similar AP2/ERF domains, together

with putative type B-RRs binding domains, were recently shown to be essential in the *M. truncatula* NIN promoter (Liu et al., 2019c). *ERN1* functions as a positive regulator of rhizobial infection (Cerri et al., 2017; Kawaharada et al., 2017). *L. japonicus* knockout mutations in *ERN1* results in a severe reduction of the number of infection threads formed, as well as a down-regulation of symbiotic gene expression, such as *EPR3* and *NF-YA1* (Kawaharada et al., 2017). Additionally, it was found that *ERN1* is involved in cytokinin-induced cortical cell division (Kawaharada et al., 2017). In line with this, it is relevant to investigate the importance of the putative *ERN1* binding sites in the *PanHK3* promoter of *P. andersonii*.

## Concluding remarks

Based on the studies in Papilionoid legumes, it is evident that hormonal networks are integrated in the rhizobial-induced root nodulation. Cytokinin signalling has been demonstrated as a major player, which commits functions that are spatial- and symbiotic context-dependent. Elaboration on the roles of cytokinin signalling in nodulation outside Papilionoid legumes as presented in this thesis, suggests that this hormone is important in nodulating of the basal legume sub-family Caesalpinoideae (*C. mimosoides*) and the distantly related non-legume species *P. andersonii*. Data included in this PhD thesis suggest that the integration of cytokinin signalling in nodulation is a general feature, which may have been recruited in early in the nitrogen-fixing clade.

For the future perspectives, identifying the components in the symbiotic pathway that regulate, and are regulated by, plant hormones will be essential. This information will provide insights on how the perception of Nod-Factors is amplified and translated into hormonal changes controlling the infection events and the nodule developmental program. Comparative genetic analyses by exploration of multiple nodulating species will help to distil the key genetic elements of the nodulation trait and the degree of variation in the regulatory networks that may occur in specific nodulating lineages. Altogether, to gain understanding in nodule organogenesis and rhizobium infection is a core strategy towards engineering the competent organ for nitrogen-fixation in the more expanding host range.

## Acknowledgments

The author would like to thank Defeng Shen for providing the data showing **Figure 2**.



# References





## References

- Agustí, J., Merelo, P., Cercós, M., Tadeo, F.R., and Talón, M.** (2009). Comparative transcriptional survey between laser-microdissected cells from laminar abscission zone and petiolar cortical tissue during ethylene-promoted abscission in citrus leaves. *BMC Plant Biol.* **9**:127.
- Akkermans, A.D.L., Abdulkadir, S., and Trinick, M.J.** (1978). Nitrogen-fixing root nodules in Ulmaceae. *Nature* **274**:190.
- Altpeter, F., Springer, N.M., Bartley, L.E., Blechl, A.E., Brutnell, T.P., Citovsky, V., Conrad, L.J., Gelvin, S.B., Jackson, D.P., Kausch, A.P., et al.** (2016). Advancing crop transformation in the era of genome editing. *Plant Cell* **28**:1510-1520.
- Ané, J.-M., Kiss, G. B., Riely, B. K., Penmetsa, R. V., Oldroyd, G. E. D., Ayax, C., Lévy, J., Debellé, F., Baek, J.-M., Kalo, P., et al.** (2004). *Medicago truncatula* DMI1 required for bacterial and fungal symbioses in legumes. *Science* (80-. ). **303**:1364-1367.
- APG III** (2009). An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG III. *Bot. J. Linn. Soc.* **161**:105-121.
- Ariel, F., Brault-Hernandez, M., Laffont, C., Huault, E., Brault, M., Plet, J., Moison, M., Blanchet, S., Ichante, J. L., Chabaud, M., et al.** (2012). Two Direct Targets of Cytokinin Signaling Regulate Symbiotic Nodulation in *Medicago truncatula*. *Plant Cell* **24**:3838-3852.
- Arnaud, D., Lee, S., Takebayashi, Y., Choi, D., Choi, J., Sakakibara, H., and Hwang, I.** (2017). Cytokinin-mediated regulation of reactive oxygen species homeostasis modulates stomatal immunity in arabidopsis. *Plant Cell* **29**:543-559.
- Arrighi, J.-F., Barre, A., Ben Amor, B., Bersoult, A., Soriano, L. C., Mirabella, R., de Carvalho-Niebel, F., Journet, E.-P., Gherardi, M., Huguet, T., et al.** (2006). The *Medicago truncatula* lysin motif-receptor-like kinase gene family includes NFP and new nodule-expressed genes. *Plant Physiol.* **142**:265-279.
- Arrighi, J. F., Godfroy, O., De Billy, F., Saurat, O., Jauneau, A., and Gough, C.** (2008). The RPG gene of *Medicago truncatula* controls Rhizobium-directed polar growth during infection. *Proc. Natl. Acad. Sci. U. S. A.* **105**:9817-9822.
- Bansal, R., Mittapelly, P., Cassone, B.J., Mamidala, P., Redinbaugh, M.G., and Michel, A.** (2015). Recommended reference genes for quantitative PCR analysis in soybean have variable stabilities during diverse biotic stresses. *PLoS One* **10**:e0134890.
- Bauer, P., Ratet, P., Crespi, M. D., Schultze, M., and Kondorosi, A.** (1996). Nod factors and cytokinins induce similar cortical cell division, amyloplast deposition and MsEnoD12A expression patterns in alfalfa roots. *Plant J.* **10**:91-105.
- Bazykin, G.A., and Kochetov, A.V.** (2011). Alternative translation start sites are conserved in eukaryotic genomes. *Nucleic Acids Res.* **39**:567-577.
- Becking, J. H.** (1979). Root-nodule symbiosis between Rhizobium and Parasponia (Ulmaceae). *Plant Soil* **51**:289-296.
- Becking, J. H.** (1983). The Parasponia parviflora -Rhizobium symbiosis. Host specificity, growth and nitrogen fixation under various conditions. *Plant Soil* **75**:309-342.
- Becking, J. H.** (1992). The Rhizobium symbiosis of the nonlegume Parasponia. *Biol. nitrogen Fixat.* Advance Access published 1992.
- Behm, J.E., Geurts, R., and Kiers, E.T.** (2014). *Parasponia*: a novel system for studying mutualism stability. *Trends Plant Sci.* **19**:757-763.
- Bender, G. L., Nayudu, M., Goydych, W., and Rolfe, B. G.** (1987). Early Infection Events in the Nodulation of the Non-Legume Parasponia andersonii by Bradyrhizobium **51**:285-293.
- Bertani G** (1951). Studies on lysogenesis. I. The mode of phage liberation by lysogenic *Escherichia coli*. *J. Bacteriol.* Advance Access published 1951.
- Bhalerao, R.P., and Fischer, U.** (2017). Environmental and hormonal control of cambial stem cell dynamics. *J. Exp. Bot.* **68**:79-87.

- Boisson-Dernier, A., Chabaud, M., Garcia, F., Becard, G., Rosenberg, C., and Barker, D. G. (2001). Agrobacterium rhizogenes-transformed roots of *Medicago truncatula* for the study of nitrogen-fixing and endomycorrhizal symbiotic associations. *Mol. Plant-Microbe Interact.* **14**:695–700.
- Boivin, S., Kazmierczak, T., Brault, M., Wen, J., Gamas, P., Mysore, K. S., and Frugier, F. (2016). Different cytokinin histidine kinase receptors regulate nodule initiation as well as later nodule developmental stages in *Medicago truncatula*. *Plant Cell Environ.* **39**:2198–2209.
- Boogerd, F. C., and Van Rossum, D. (1997). Nodulation of groundnut by *Bradyrhizobium*: A simple infection process by crack entry. *FEMS Microbiol. Rev.* **21**:5–27.
- Borisov, A. Y., Madsen, L. H., Tsyganov, V. E., Umehara, Y., Voroshilova, V. A., Batagov, A. O., Sandal, N., Mortensen, A., Schauser, L., Ellis, N., et al. (2003). The Sym35 gene required for root nodule development in pea is an ortholog of Nin from *Lotus japonicus*. *Plant Physiol.* **131**:1009–17.
- Boualem, A., Fergany, M., Fernandez, R., Troadec, C., Martin, A., Morin, H., Sari, M.A., Collin, F., Flowers, J.M., Pitrat, M., et al. (2008). A conserved mutation in an ethylene biosynthesis enzyme leads to andromonoecy in melons. *Science* **321**:836–838.
- Boualem, A., Troadec, C., Camps, C., Lemhemdi, A., Morin, H., Sari, M.A., Fraenkel-Zagouri, R., Kovalski, I., Dogimont, C., Perl-Treves, R., et al. (2015). A cucurbit androecy gene reveals how unisexual flowers develop and dioecy emerges. *Science* **350**:688–691.
- Breakspear, A., Liu, C., Roy, S., Stacey, N., Rogers, C., Trick, M., Morieri, G., Mysore, K. S., Wen, J., Oldroyd, G. E. D., et al. (2014). The root hair “infectome” of *Medicago truncatula* uncovers changes in cell cycle genes and reveals a requirement for auxin signaling in rhizobial infection. *Plant Cell* **26**:4680–4701.
- Brown, K.M. (1997). Ethylene and abscission. *Physiol. Plant.* **100**:567–576.
- Bu, F., Rutten, L., Roswanjaya, Y. P., Kulikova, O., Rodriguez-Franco, M., Ott, T., Bisseling, T., van Zeijl, A., and Geurts, R. (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. *New Phytol.* **226**:541–554.
- Cannon, S. B., Sterck, L., Rombauts, S., Sato, S., Cheung, F., Gouzy, J., Wang, X., Mudge, J., Vasdewani, J., Schiex, T., et al. (2006). Legume genome evolution viewed through the *Medicago truncatula* and *Lotus japonicus* genomes. *Proc. Natl. Acad. Sci.* **103**:14959–14964.
- Cannon, S. B., Ilut, D., Farmer, A. D., Maki, S. L., May, G. D., Singer, S. R., and Doyle, J. J. (2010). Polyploidy did not predate the evolution of nodulation in all legumes. *PLoS One* **5**:e11630.
- Cao, Q., Op den Camp, R., Seifi Kalhor, M., Bisseling, T., and Geurts, R. (2012). Efficiency of *Agrobacterium rhizogenes*-mediated root transformation of *Parasponia* and *Trema* is temperature dependent. *Plant Growth Regul.* **68**:459–465.
- Cardoso, C., Zhang, Y.X., Jamil, M., Hepworth, J., Charnikhova, T., Dimkpa, S.O.N., Meharg, C., Wright, M.H., Liu, J.W., Meng, X.B., et al. (2014). Natural variation of rice strigolactone biosynthesis is associated with the deletion of two *MAX1* orthologs. *Proc. Natl. Acad. Sci. U. S. A.* **111**:2379–2384.
- Cerri, M. R., Frances, L., Laloum, T., Auriac, M.-C., Niebel, A., Oldroyd, G. E. D., Barker, D. G., Fournier, J., and de Carvalho-Niebel, F. (2012). *Medicago truncatula* ERN transcription factors: regulatory interplay with NSP1/NSP2 GRAS factors and expression dynamics throughout rhizobial infection. *Plant Physiol.* **160**:2155–2172.
- Cerri, M. R., Wang, Q., Stolz, P., Folgmann, J., Frances, L., Katzer, K., Li, X., Heckmann, A. B., Wang, T. L., Downie, J. A., et al. (2017). The ERN1 transcription factor gene is a target of the CCaMK/CYCLOPS complex and controls rhizobial infection in *Lotus japonicus*. *New Phytol.* **215**:323–337.

- Chabregas, S.M., Lucbe, D.D., Van Sluys, M.A., Menck, C.F.M., and Silva-Filho, M.C.** (2003). Differential usage of two in-frame translational start codons regulates subcellular localization of *Arabidopsis thaliana* TH11. *J. Cell Sci.* **116**:285–291.
- Charpentier, M., Sun, J., Martins, T. V., Radhakrishnan, G. V., Findlay, K., Soumpourou, E., Thouin, J., Véry, A., Sanders, D., Morris, R. J., et al.** (2016). Nuclear-localized cyclic nucleotide-gated channels mediate symbiotic calcium oscillations. *Science* (80- ). **352**:1102–1105.
- Chen, Y., Chen, W., Li, X., Jiang, H., Wu, P., Xia, K., Yang, Y., and Wu, G.** (2014). Knockdown of LjIPT3 influences nodule development in lotus japonicus. *Plant Cell Physiol.* **55**:183–193.
- Cissoko, M., Hoher, V., Gherbi, H., Gully, D., Carré-Mlouka, A., Sane, S., Pignoly, S., Champion, A., Ngom, M., Pujic, P., et al.** (2018). Actinorhizal signaling molecules: Frankia root hair deforming factor shares properties with NIN inducing factor. *Front. Plant Sci.* **871**:1–12.
- Clason, E. W. W.** (1936). The vegetation of the upper-Badak region of mount Kelut (east java). *Bull. Jard. Bot. Buitenzorg Serie III*,:509–518.
- Clavijo, F., Diedhiou, I., Vaissayre, V., Brottier, L., Acolatse, J., Moukouanga, D., Crabos, A., Auguy, F., Franche, C., Gherbi, H., et al.** (2015). The Casuarina NIN gene is transcriptionally activated throughout Frankia root infection as well as in response to bacterial diffusible signals. *New Phytol.* **208**:887–903.
- Cooper, J. B., and Long, S. R.** (1994). Morphogenetic rescue of Rhizobium meliloti nodulation mutants by trans-zeatin secretion. *Plant Cell* **6**:215–225.
- Czechowski, T., Stitt, M., Altmann, T., Udvardi, M.K., and Scheible, W.R.** (2005). Genome-wide identification and testing of superior reference genes for transcript normalization in Arabidopsis. *Plant Physiol.* **139**:5–17.
- Damiani, I., Drain, A., Guichard, M., Balzergue, S., Boscari, A., Boyer, J. C., Brunaud, V., Cottaz, S., Rancurel, C., da Rocha, M., et al.** (2016). Nod factor effects on root hair-specific transcriptome of medicago truncatula: Focus on plasma membrane transport systems and reactive oxygen species networks. *Front. Plant Sci.* **7**:1–22.
- Davey, M. R., Webster, G., Manders, G., Ringrose, F. L., Power, J. B., and Cocking, E. C.** (1993). Effective nodulation of micro-propagated shoots of the non-legume *Parasponia andersonii* by *Bradyrhizobium*. *J. Exp. Bot.* **44**:863–867.
- Delaux, P. M., Radhakrishnan, G., and Oldroyd, G.** (2015). Tracing the evolutionary path to nitrogen-fixing crops. *Curr. Opin. Plant Biol.* **26**:95–99.
- Dello Ioio, R., Nakamura, K., Moubayidin, L., Perilli, S., Taniguchi, M., Morita, M. T., Aoyama, T., Costantino, P., and Sabatini, S.** (2008). A genetic framework for the control of cell division and differentiation in the root meristem. *Science* (80- ). **322**:1380–1384.
- Demina, I. V., Maity, P. J., Nagchowdhury, A., Ng, J. L. P., van der Graaff, E., Demchenko, K. N., Roitsch, T., Mathesius, U., and Pawlowski, K.** (2019). Accumulation of and Response to Auxins in Roots and Nodules of the Actinorhizal Plant *Datisca glomerata* Compared to the Model Legume *Medicago truncatula*. *Front. Plant Sci.* **10**:1–19.
- Doench, J.G., Hartenian, E., Graham, D.B., Tothova, Z., Hegde, M., Smith, I., Sullender, M., Ebert, B.L., Xavier, R.J., and Root, D.E.** (2014). Rational design of highly active sgRNAs for CRISPR-Cas9-mediated gene inactivation. *Nat. Biotechnol.* **32**:1262–1267.
- Domonkos, Á., Kovács, S., Gombár, A., Kiss, E., Horváth, B., Kováts, G. Z., Farkas, A., Tóth, M. T., Ayaydin, F., Bóka, K., et al.** (2017). NAD1 controls defense-like responses in *Medicago truncatula* symbiotic nitrogen fixing nodules following rhizobial colonization in a BacA-independent manner. *Genes (Basel)*. **8**:1–21.
- Doyle, J. J.** (1998). Phylogenetic perspectives on nodulation: evolving views of plants and symbiotic bacteria. *Trends Plant Sci.* **3**:473–478.

- Doyle, J. J. (2011). Phylogenetic perspectives on the origins of nodulation. *Mol. plant-microbe Interact.* **24**:1289–95.
- Duarte, J.M., Wall, P.K., Edger, P.P., Landherr, L.L., Ma, H., Pires, J.C., Leebens-Mack, J., and dePamphilis, C.W. (2010). Identification of shared single copy nuclear genes in *Arabidopsis*, *Populus*, *Vitis* and *Oryza* and their phylogenetic utility across various taxonomic levels. *BMC Evol. Biol.* **10**:61.
- Endre, G., Kereszt, A., Mihaceae, S., Kalo, P., and Kiss, G. B. (2002). A receptor kinase gene regulating symbiotic nodule development. *Nature* **417**:962–966.
- Engler, C., Gruetzner, R., Kandzia, R., and Marillonnet, S. (2009). Golden Gate shuffling: a one-pot DNA shuffling method based on Type IIs restriction enzymes. *PLoS One* **4**:e5553.
- Engler, C., Youles, M., Gruetzner, R., Ehnert, T.-M., Werner, S., Jones, J. D. G., Patron, N. J., and Marillonnet, S. (2014). A Golden Gate Modular Cloning Toolbox for Plants. *ACS Synth. Biol.* Advance Access published 2014, doi:10.1021/sb4001504.
- Fahraeus, G. (1957). The infection of clover root hairs by nodule bacteria studied by a simple glass slide technique. *Microbiology* **16**, 374–381. doi: 10.1099/00221287-16-2-374
- Fan, D., Liu, T.T., Li, C.F., Jiao, B., Li, S., Hou, Y.S., and Luo, K.M. (2015). Efficient CRISPR/Cas9-mediated targeted mutagenesis in *Populus* in the first generation. *Sci. Rep.* **5**:12217.
- Fang, Y., and Hirsch, A. M. (1998). Studying early nodulin gene *ENOD40* expression and induction by nodulation factor and cytokinin in transgenic alfalfa. *Plant Physiol.* **116**:53–68.
- Fausser, F., Schiml, S., and Puchta, H. (2014). Both CRISPR/Cas-based nucleases and nickases can be used efficiently for genome engineering in *Arabidopsis thaliana*. *Plant J.* **79**:348–359.
- Fedorova, E., Thomson, R., Whitehead, L.F., Maudoux, O., Udvardi, M.K., and Day, D.A. (1999). Localization of H<sup>+</sup>-ATPase in soybean root nodules. *Planta* **209**:25–32.
- Felten, J., Kohler, A., Morin, E., Bhalerao, R. P., Palme, K., Martin, F., Ditengou, F. A., and Legue, V. (2009). The Ectomycorrhizal Fungus *Laccaria bicolor* Stimulates Lateral Root Formation in Poplar and *Arabidopsis* through Auxin Transport and Signaling. *Plant Physiol.* **151**:1991–2005.
- Ferguson, B. J., and Mathesius, U. (2003). Signaling Interactions during Nodule Development. *J. Plant Growth Regul.* **22**:47–72.
- Ferguson, B. J., and Mathesius, U. (2014). Phytohormone regulation of legume-rhizobia interactions. *J. Chem. Ecol.* **40**:770–790.
- Fonouni-Farde, C., Tan, S., Baudin, M., Brault, M., Wen, J., Mysore, K. S., Niebel, A., Frugier, F., and Diet, A. (2016). DELLA-mediated gibberellin signalling regulates Nod factor signalling and rhizobial infection. *Nat. Commun.* **7**.
- Franssen, H. J., Thompson, D. V., Idler, K., Kormelink, R., van Kammen, A., and Bisseling, T. (1990). Nucleotide sequence of two soybean ENOD2 early nodulin genes encoding Ngm-75. *Plant Mol. Biol.* **14**:103–106.
- Franssen, H. J., Xiao, T. T., Kulikova, O., Wan, X., Bisseling, T., Scheres, B., and Heidstra, R. (2015). Root developmental programs shape the Medicago truncatula nodule meristem. *Dev.* **142**:2941–2950.
- Frebort, I., Kowalska, M., Hluska, T., Frebortova, J., and Galuszka, P. (2011). Evolution of cytokinin biosynthesis and degradation. *J Exp Bot* **62**:2431–2452.
- Fukaki, H., and Tasaka, M. (2009). Hormone interactions during lateral root formation. *Plant Mol. Biol.* **69**:437–449.
- Gage, D. J. (2004). Infection and invasion of roots by symbiotic, nitrogen-fixing rhizobia during nodulation of temperate legumes. *Microbiol. Mol. Biol. Rev.* **68**:280–300.
- Gamas, P., Brault, M., Jardinaud, M.-F., Frugier, F. (2017). Cytokinins in symbiotic nodulation: when, where, what for? *Trends Plant Sci.* **22**: 792–802.

- Garrido-Oter, R., Nakano, R. T., Dombrowski, N., Ma, K. W., McHardy, A. C., and Schulze-Lefert, P. (2018). Modular Traits of the Rhizobiales Root Microbiota and Their Evolutionary Relationship with Symbiotic Rhizobia. *Cell Host Microbe* **24**:155–167.e5.
- Gauthier-Coles, C., White, R. G., and Mathesius, U. (2019). Nodulating legumes are distinguished by a sensitivity to cytokinin in the root cortex leading to pseudonodule development. *Front. Plant Sci.* **9**:1–14.
- Geurts, R., Lillo, A., and Bisseling, T. (2012). Exploiting an ancient signalling machinery to enjoy a nitrogen fixing symbiosis. *Curr. Opin. Plant Biol.* **15**:438–443.
- Geurts, R., Xiao, T. T., and Reinhold-Hurek, B. (2016). What does it take to evolve a nitrogen-fixing endosymbiosis? *Trends Plant Sci.* **21**:199–208.
- Gonzalez-Rizzo, S., Crespi, M., and Frugier, F. (2006). The Medicago truncatula CRE1 cytokinin receptor regulates lateral root development and early symbiotic interaction with Sinorhizobium meliloti. *Plant Cell* **18**:2680–2693.
- Graham, P. H. H., Viteri, S. E. E., Mackie, F., Vargas, A. T. T., and Palacios, A. (1982). Variation in acid soil tolerance among strains of Rhizobium phaseoli. *F. Crop. Res.* **5**:121–128.
- Granqvist, E., Sun, J., Op den Camp, R., Pujic, P., Hill, L., Normand, P., Morris, R. J., Downie, J. A., Geurts, R., and Oldroyd, G. E. D. (2015). Bacterial-induced calcium oscillations are common to nitrogen-fixing associations of nodulating legumes and non-legumes. *New Phytol.* **207**:551–558.
- Griesmann, M., Chang, Y., Liu, X., Song, Y., Haberer, G., Crook, M. B., Billault-Penneteau, B., Laressergues, D., Keller, J., Imanishi, L., et al. (2018). Phylogenomics reveals multiple losses of nitrogen-fixing root nodule symbiosis. *Science* (80-. ). **361**:1–85.
- Groth, M., Takeda, N., Perry, J., Uchid, H., Dräxl, S., Brachmann, A., Sato, S., Tabata, S., Kawaguchi, M., Wang, T. L., et al. (2010b). NENA, a Lotus japonicus homolog of Sec13, is required for rhizodermal infection by arbuscular mycorrhiza fungi and rhizobia but dispensable for cortical endosymbiotic development. *Plant Cell* **22**:2509–2526.
- Haney, C. H., and Long, S. R. (2010). Plant flotillins are required for infection by nitrogen-fixing bacteria. *Proc. Natl. Acad. Sci. U. S. A.* **107**:478–483.
- Harberd, N. P., Belfield, E., and Yasumura, Y. (2009). The angiosperm gibberellin-GID1-DELLA growth regulatory mechanism: How an “inhibitor of an inhibitor” enables flexible response to fluctuating environments. *Plant Cell* **21**:1328–1339.
- Hayashi, T., Banba, M., Shimoda, Y., Kouchi, H., Hayashi, M., Imaizumi-Anraku, H., and Imaizumi-Anraku, H. (2010). A dominant function of CCaMK in intracellular accommodation of bacterial and fungal endosymbionts. *Plant J.* **63**:141–154.
- Heckmann, A. B., Lombardo, F., Miwa, H., Perry, J. A., Bunnewell, S., Parniske, M., Wang, T. L., and Downie, J. A. (2006). Lotus japonicus nodulation requires two GRAS domain regulators, one of which is functionally conserved in a non-legume. *Plant Physiol.* **142**:1739–1750.
- Heckmann, A. B., Sandal, N., Bek, A. S., Madsen, L. H., Jurkiewicz, A., Nielsen, M. W., Tirichine, L., and Stougaard, J. (2011). Cytokinin induction of root nodule primordia in Lotus japonicus is regulated by a mechanism operating in the root cortex. *Mol. Plant-Microbe Interact.* **24**:1385–1395.
- Heidstra, R., Yang, W. C., Yalcin, Y., Peck, S., Emons, A. M., Van Kammen, A., and Bisseling, T. (1997). Ethylene provides positional information on cortical cell division but is not involved in Nod factor-induced root hair tip growth in Rhizobium-legume interaction. *Development* **124**:1781–1787.
- Held, M., Hou, H., Miri, M., Huynh, C., Ross, L., Hossain, M. S., Sato, S., Tabata, S., Perry, J., Wang, T. L., et al. (2014). Lotus japonicus cytokinin receptors work partially redundantly to mediate nodule formation. *Plant Cell* **26**:678–694.

- Hiltbrand, R., Thomas, J., McCarthy, H., Dykema, K. J., Spurr, A., Newhart, H., Winn, M. E., and Mukherjee, A. (2016). A developmental and molecular view of formation of auxin-induced nodule-like structures in land plants. *Front. Plant Sci.* **7**:1–18.
- Hirose, N., Takei, K., Kuroha, T., Kamada-Nobusada, T., Hayashi, H., and Sakakibara, H. (2008). Regulation of cytokinin biosynthesis, compartmentalization and translocation. *J. Exp. Bot.* **59**:75–83.
- Hirsch, A. M., Bhuvaneswari, T. V., Torrey, J. G., and Bisseling, T. (1989). Early nodulin genes are induced in alfalfa root outgrowths elicited by auxin transport inhibitors. *Proc. Natl. Acad. Sci. U. S. A.* **86**:1244–8.
- Hirsch, A. M. (1992). Developmental biology of legume nodulation. *Dev. Biol.* **122**:211–237.
- Hirsch, S., Kim, J., Muñoz, A., Heckmann, A.B., Downie, J.A., and Oldroyd, G.E.D. (2009). GRAS proteins form a DNA binding complex to induce gene expression during nodulation signaling in *Medicago truncatula*. *Plant Cell* **21**:545–557.
- Hoagland, D. R., Arnon Revised, D. I., and Arnon, D. I. (1950). The Water-Culture Method for Growing Plants without Soil. *Circ Calif Agric Exp Stn.* **347**:1–32.
- Horváth, B., Yeun, L. H., Domonkos, Á., Halász, G., Gobbato, E., Ayaydin, F., Miró, K., Hirsch, S., Sun, J., Tadege, M., et al. (2011). *Medicago truncatula* IPD3 is a member of the common symbiotic signaling pathway required for rhizobial and mycorrhizal symbioses. *Mol. Plant-Microbe Interact.* **24**:1345–1358.
- Huisman, R. (2018). Rik HUISMAN Advance Access published 2018.
- Huisman, R., and Geurts, R. (2020). A Roadmap toward Engineered Nitrogen-Fixing Nodule Symbiosis. *Plant Commun.* **1**:100019.
- Huo, X., Schnabel, E., Hughes, K., and Frugoli, J. (2006). RNAi phenotypes and the localization of a protein::GUS fusion imply a role for *Medicago truncatula* PIN genes in nodulation. *J Plant Growth Regul* **25**:156–165.
- Ibáñez, F., Wall, L., and Fabra, A. (2017). Starting points in plant-bacteria nitrogen-fixing symbioses: Intercellular invasion of the roots. *J. Exp. Bot.* **68**:1905–1918.
- Imaizumi-Anraku, H., Takeda, N., Charpentier, M., Perry, J., Miwa, H., Umehara, Y., Kouchi, H., Murakami, Y., Mulder, L., Vickers, K., et al. (2005). Plastid proteins crucial for symbiotic fungal and bacterial entry into plant roots. *Nature* **433**:527–531.
- Immanen, J., Nieminen, K., Smolander, O. P., Kojima, M., Alonso Serra, J., Koskinen, P., Zhang, J., Elo, A., Mähönen, A. P., Street, N., et al. (2016). Cytokinin and Auxin Display Distinct but Interconnected Distribution and Signaling Profiles to Stimulate Cambial Activity. *Curr. Biol.* **26**:1990–1997.
- Itoh, H., Ueguchi-Tanaka, M., and Matsuoka, M. (2008). Chapter 6 Molecular Biology of Gibberellins Signaling in Higher Plants. *Int. Rev. Cell Mol. Biol.* **268**:191–221.
- Jackson, M.B., and Osborne, D.J. (1970). Ethylene, the natural regulator of leaf abscission. *Nature* **225**:1019–1022.
- Jardinaud, M.-F., Boivin, S., Rodde, N., Catrice, O., Kisiala, A., Lepage, A., Moreau, S., Roux, B., Cottret, L., Sallet, E., et al. (2016). A laser dissection-RNAseq analysis highlights the activation of cytokinin pathways by Nod factors in the *Medicago truncatula* root epidermis. *Plant Physiol.* **33**:pp.00711.2016.
- Jin, Y., Liu, H., Luo, D., Yu, N., Dong, W., Wang, C., Zhang, X., Dai, H., Yang, J., and Wang, E. (2016). DELLA proteins are common components of symbiotic rhizobial and mycorrhizal signalling pathways. *Nat. Commun.* **7**:12433.
- Kaló, P., Gleason, C., Edwards, A., Marsh, J., Mitra, R. M., Hirsch, S., Jakab, J., Sims, S., Long, S. R., Rogers, J., et al. (2005). Nodulation signaling in legumes requires NSP2, a member of the GRAS family of transcriptional regulators. *Science*. **308**:1786–1789.

- Kanamori, N., Madsen, L. H., Radutoiu, S., Frantescu, M., Quistgaard, E. M. H., Miwa, H., Downie, J. A., James, E. K., Felle, H. H., Haaning, L. L., et al. (2006). A nucleoporin is required for induction of Ca<sup>2+</sup> spiking in legume nodule development and essential for rhizobial and fungal symbiosis. *Proc. Natl. Acad. Sci. U. S. A.* **545**:359–364.
- Kareem, A., Durgaprasad, K., Sugimoto, K., Du, Y., Pulianmackal, A.J., Trivedi, Z.B., Abhayadev, P.V., Pinon, V., Meyerowitz, E.M., Scheres, B., et al. (2015). PLETHORA genes control regeneration by a two-step mechanism. *Curr. Biol.* **25**:1017–1030.
- Karimi, M., Inze, D., and Depicker, A. (2002). GATEWAY™ vectors for *Agrobacterium*-mediated plant transformation. *Trends Plant Sci.* **7**:193–195.
- Katoh, K., Misawa, K., Kuma, K. I., and Miyata, T. (2002). MAFFT: A novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Res.* **30**:3059–3066.
- Kawaharada, Y., James, E. K., Kelly, S., Sandal, N., and Stougaard, J. (2017). The ethylene responsive factor required for nodulation 1 (ERN1) transcription factor is required for infection-thread formation in lotus japonicus. *Mol. Plant-Microbe Interact.* **30**:194–204.
- Kieber, J. J., and Schaller, G. E. (2010). The perception of cytokinin: a story 50 years in the making. *Plant Physiol.* **154**:487–92.
- Kim, G. B., Son, S. U., Yu, H. J., and Mun, J. H. (2019). MtGA2ox10 encoding C20-GA2-oxidase regulates rhizobial infection and nodule development in *Medicago truncatula*. *Sci. Rep.* **9**:1–13.
- Kucho, K. I., Tamari, D., Matsuyama, S., Nabekura, T., and Tisa, L. S. (2017). Nitrogen fixation mutants of the actinobacterium *Frankia casuarinae* Ccl3. *Microbes Environ.* **32**:344–351.
- Kumagai, H., and Kouchi, H. (2003). Gene Silencing by Expression of Hairpin RNA in Lotus japonicus Roots and Root Nodules. *Mol. Plant-Microbe Interact.* **16**:663–668.
- Lamesch, P., Berardini, T.Z., Li, D.H., Swarbreck, D., Wilks, C., Sasidharan, R., Muller, R., Dreher, K., Alexander, D.L., Garcia-Hernandez, M., et al. (2012). The Arabidopsis Information Resource (TAIR): improved gene annotation and new tools. *Nucleic Acids Res.* **40**:D1202–D1210.
- 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**:26–37.
- Lancelle, S. A., and Torrey, J. G. (1985). Early Development of Rhizobium-Induced Root-Nodules of *Parasponia-Rigida*. 2. Nodule Morphogenesis and Symbiotic Development. *Can. J. Bot.* **63**:25–35.
- Larkin, P. J., and Scowcroft, W. R. (1981). Somaclonal variation — a novel source of variability from cell cultures for plant improvement. *Theor. Appl. Genet.* **60**:197–214.
- Lazo, G. R., Stein, P. A., and Ludwig, R. A. (1991). A DNA transformation-competent Arabidopsis genomic library in *Agrobacterium*. *Biotechnology* **9**:963–967.
- Ledford, H. (2016). A better way to hack plant DNA. *Nature* **539**:16–17.
- Lee, T. D., and Bazzaz, F. A. (1982). Regulation of Fruit Maturation Pattern in an Annual Legume. *Ecology* **63**:1374–1388.
- Lefebvre, B., Timmers, T., Mbengue, M., Moreau, S., Hervé, C., Tóth, K., Bittencourt-Silvestre, J., Klaus, D., Deslandes, L., Godiard, L., et al. (2010). A remorin protein interacts with symbiotic receptors and regulates bacterial infection. *Proc. Natl. Acad. Sci. U. S. A.* **107**:2343–2348.
- Lerouge, P., Roche, P., Faucher, C., Maillet, F., Truchet, G., Promé, J. C., and Dénarié, J. (1990). Symbiotic host-specificity of *Rhizobium meliloti* is determined by a sulphated and acylated glucosamine oligosaccharide signal. *Nature* **344**:781–784.

- Lévy, J., Bres, C., Geurts, R., Chalhoub, B., Kulikova, O., Duc, G., Journet, E.-P., Ané, J.-M., Lauber, E., Bisseling, T., et al. (2004). A putative Ca<sup>2+</sup> and calmodulin-dependent protein kinase required for bacterial and fungal symbioses. *Science* (80- ). **303**:1361–1364.
- Li, H.L., Wang, W., Mortimer, P.E., Li, R.Q., Li, D.Z., Hyde, K.D., Xu, J.C., Soltis, D.E., and Chen, Z.D. (2015). Large-scale phylogenetic analyses reveal multiple gains of actinorhizal nitrogen-fixing symbioses in angiosperms associated with climate change. *Sci. Rep.* **5**:14023.
- Liang, P., Stratil, T. F., Popp, C., Marín, M., Folgmann, J., Mysore, K. S., Wen, J., and Ott, T. (2018). Symbiotic root infections in *Medicago truncatula* require remorin-mediated receptor stabilization in membrane nanodomains. *Proc. Natl. Acad. Sci. U. S. A.* **115**:5289–5294.
- Limpens, E., Franken, C., Smit, P., Willemse, J., Bisseling, T., and Geurts, R. (2003). LysM domain receptor kinases regulating rhizobial Nod factor-induced infection. *Science* (80). **302**:630–633.
- Limpens, E., Ramos, J., Franken, C., Raz, V., Compaan, B., Franssen, H., Bisseling, T., and Geurts, R. (2004). RNA interference in *Agrobacterium rhizogenes*-transformed roots of *Arabidopsis* and *Medicago truncatula*. *J. Exp. Bot.* **55**:983–992.
- Limpens, E., and Bisseling, T. (2014). CYCLOPS: a new vision on rhizobium-induced nodule organogenesis. *Cell Host Microbe* **15**:127–129.
- Limpens, E., van Zeijl, A., and Geurts, R. (2015). Lipochitooligosaccharides Modulate Plant Host Immunity to Enable Endosymbioses. *Annu. Rev. Phytopathol.* **53**:311–334.
- Liu, W., Kohlen, W., Lillo, A., den Camp, R. O., Ivanov, S., Hartog, M., Limpens, E., Jamil, M., Smaczniak, C., Kaufmann, K., et al. (2011). Strigolactone biosynthesis in *Medicago truncatula* and rice requires the symbiotic GRAS-type transcription factors NSP1 and NSP2. *Plant Cell* **23**:3853–3865.
- Liu, C. W., Breakspear, A., Stacey, N., Findlay, K., Nakashima, J., Ramakrishnan, K., Liu, M., Xie, F., Endre, G., de Carvalho-Niebel, F., et al. (2019a). A protein complex required for polar growth of rhizobial infection threads. *Nat. Commun.* **10**:1–17.
- Liu, C. W., Breakspear, A., Guan, D., Cerri, M. R., Jackson, K., Jiang, S., Robson, F., Radhakrishnan, G. V., Roy, S., Bone, C., et al. (2019b). NIN acts as a network hub controlling a growth module required for rhizobial infection. *Plant Physiol.* **179**:1704–1722.
- Liu, J., Rutten, L., Limpens, E., Van Der Molen, T., Van Velzen, R., Chen, R., Chen, Y., Geurts, R., Kohlen, W., Kulikova, O., et al. (2019c). A remote cis-regulatory region is required for nin expression in the pericycle to initiate nodule primordium formation in *medicago truncatula*. *Plant Cell* **31**:68–83.
- Liu, M., Soyano, T., Yano, K., Hayashi, M., and Kawaguchi, M. (2019d). ERN1 and CYCLOPS coordinately activate NIN signaling to promote infection thread formation in *Lotus japonicus*. *J. Plant Res.* **132**:641–653.
- Liu, J., and Bisseling, T. (2020). Evolution of nin and NIN-like genes in relation to nodule symbiosis. *Genes (Basel)*. **11**:1–15.
- Lohar, D. P., Schaff, J. E., Laskey, J. G., Kieber, J. J., Bilyeu, K. D., and Bird, D. M. K. (2004). Cytokininins play opposite roles in lateral root formation, and nematode and Rhizobial symbioses. *Plant J.* **38**:203–214.
- Lohar, D., Stiller, J., Kam, J., Stacey, G., and Gresshoff, P. M. (2009). Ethylene insensitivity conferred by a mutated *Arabidopsis* ethylene receptor gene alters nodulation in transgenic *Lotus japonicus*. *Ann. Bot.* **104**:277–285.
- Lowder, L. G., Zhang, D., Baltes, N. J., Paul, J. W., Tang, X., Zheng, X., Voytas, D. F., Hsieh, T.-F., Zhang, Y., and Qi, Y. (2015). A CRISPR/Cas9 Toolbox for Multiplexed Plant Genome Editing and Transcriptional Regulation. *Plant Physiol.* **169**:971 LP – 985.

- Ma, W., Penrose, D. M., and Glick, B. R. (2002). Strategies used by rhizobia to lower plant ethylene levels and increase nodulation. *Can. J. Microbiol.* **48**:947–954.
- Madsen, E. B., Madsen, L. H., Radutoiu, S., Sato, S., Kaneko, T., Tabata, S., and Sandal, N. (2003). A receptor kinase gene of the LysM type is involved in legume perception of rhizobial signals. *Nature* **425**:637–640.
- Madsen, L. H., Tirichine, L., Jurkiewicz, A., Sullivan, J. T., Heckmann, A. B., Bek, A. S., Ronson, C. W., James, E. K., and Stougaard, J. (2010). The molecular network governing nodule organogenesis and infection in the model legume *Lotus japonicus*. *Nat. Commun.* **1**.
- Mahonen, A.P., Bishopp, A., Higuchi, M., Nieminen, K.M., Kinoshita, K., Tormakangas, K., Ikeda, Y., Oka, A., Kakimoto, T., and Helariutta, Y. (2006a). Cytokinin signaling and its inhibitor AHP6 regulate cell fate during vascular development. *Science* **311**:94–98.
- Mahonen, A.P., Higuchi, M., Tormakangas, K., Miyawaki, K., Pischke, M.S., Sussman, M.R., Helariutta, Y., and Kakimoto, T. (2006b). Cytokinins regulate a bidirectional phosphorelay network in *Arabidopsis*. *Curr. Biol.* **16**:1116–1122.
- Maillet, F., Poinso, V., André, O., Puech-Pagès, V., Haouy, A., Gueunier, M., Cromer, L., Giraudet, D., Formey, D., Niebel, A., et al. (2011). Fungal lipochitooligosaccharide symbiotic signals in arbuscular mycorrhiza. *Nature* **469**:58–63.
- Marsh, J. F., Rakocevic, A., Mitra, R. M., Brocard, L., Sun, J., Eschstruth, A., Long, S. R., Schultz, M., Ratet, P., and Oldroyd, G. E. D. (2007). *Medicago truncatula* NIN is essential for rhizobial-independent nodule organogenesis induced by autoactive calcium/calmodulin-dependent protein kinase. *Plant Physiol* **144**:324–335.
- Martinez-Romero, E., Segovia, L., Mercante, F. M., Franco, A. A., Graham, P., and Pardo, M. A. (1991). *Rhizobium tropici*, a novel species nodulating *Phaseolus vulgaris* L. beans and *Leucaena* sp. trees. *Int. J. Syst. Bacteriol.* **41**:417–426.
- Marvel, D.J., Torrey, J.G., and Ausubel, F.M. (1987). *Rhizobium* symbiotic genes required for nodulation of legume and nonlegume hosts. *Proc. Natl. Acad. Sci. U. S. A.* **84**:1319–1323.
- Masson-Boivin, C., Giraud, E., Perret, X., and Batut, J. (2009). Establishing nitrogen-fixing symbiosis with legumes: how many rhizobium recipes? *Trends Microbiol.* **17**:458–66.
- Masson-Boivin, C., and Sachs, J. L. (2018). Symbiotic nitrogen fixation by rhizobia — the roots of a success story. *Curr. Opin. Plant Biol.* **44**:7–15.
- Mathesius, U., Charon, C., Rolfe, B. G., Kondorosi, A., and Crespi, M. (2000). Temporal and spatial order of events during the induction of cortical cell divisions in white clover by *Rhizobium leguminosarum* bv. *trifolii* Inoculation or localized cytokinin addition. *Mol. Plant-Microbe Interact.* **13**:617–628.
- Matsumoto-Kitano, M., Kusumoto, T., Tarkowski, P., Kinoshita-Tsujimura, K., Václavíková, K., Miyawaki, K., and Kakimoto, T. (2008). Cytokinins are central regulators of cambial activity. *Proc. Natl. Acad. Sci. U. S. A.* **105**:20027–20031.
- Mendiburu, F. De (2020). *Agricolae: Statistical Procedures for Agricultural Research* Advance Access published 2020.
- Mergaert, P., Vanmontagu, M., Prome, J. C., and Holsters, M. (1993). Three unusual modifications, a D-arabinosyl, an N-methyl, and a carbamoyl group, are present on the Nod factors of *Azorhizobium caulinodans* strain ORS571. *Proc. Natl. Acad. Sci. U. S. A.* **90**:1551–1555.
- Miri, M., Janakirama, P., Held, M., Ross, L., and Szczygłowski, K. (2016). Into the Root: How Cytokinin Controls Rhizobial Infection. *Trends Plant Sci.* **21**:178–186.
- Miri, M., Janakirama, P., Huebert, T., Ross, L., McDowell, T., Orosz, K., Markmann, K., and Szczygłowski, K. (2019). Inside out: root cortex-localized LHK1 cytokinin receptor limits epidermal infection of *Lotus japonicus* roots by *Mesorhizobium loti*. *New Phytol.* **222**:1523–1537.

- Miyata, K., Kawaguchi, M., and Nakagawa, T. (2013). Two distinct  *EIN2*  genes cooperatively regulate ethylene signaling in  *Lotus japonicus* .  *Plant Cell Physiol.*  **54**:1469–1477.
- Moling, S., Pietraszewska-Bogiel, A., Postma, M., Fedorova, E., Hink, M. a., Limpens, E., Gadella, T. W. J., and Bisseling, T. (2014). Nod Factor receptors form heteromeric complexes and are essential for intracellular infection in  *Medicago*  nodules.  *Plant Cell*  **26**:4188–4199.
- Montiel, J., Arthikala, M. K., Cárdenas, L., and Quinto, C. (2016). Legume NADPH oxidases have crucial roles at different stages of nodulation.  *Int. J. Mol. Sci.*  **17**:1–12.
- Mortier, V., Wasson, A., Jaworek, P., De Keyser, A., Decroos, M., Holsters, M., Tarkowski, P., Mathesius, U., and Goormachtig, S. (2014). Role of LONELY GUY genes in indeterminate nodulation on  *Medicago truncatula* .  *New Phytol.*  **202**:582–593.
- Müller, B., and Sheen, J. (2008). Cytokinin and auxin interaction in root stem-cell specification during early embryogenesis.  *Nature*  **453**:1094–1097.
- Murray, J. D., Karas, B. J., Sato, S., Tabata, S., Amyot, L., and Szczyglowski, K. (2007). A Cytokinin Perception Mutant Colonized by  *Rhizobium*  in the Absence of Nodule Organogenesis.  *Science.*  **315**:101–104.
- Murray, J. D., Muni, R. R. D., Torres-Jerez, I., Tang, Y., Allen, S., Andriankaja, M., Li, G., Laxmi, A., Cheng, X., Wen, J., et al. (2011). Vapyrin, a gene essential for intracellular progression of arbuscular mycorrhizal symbiosis, is also essential for infection by rhizobia in the nodule symbiosis of  *Medicago truncatula* .  *Plant J.*  **65**:244–252.
- Nadzieja, M., Stougaard, J., and Reid, D. (2019). A Toolkit for High Resolution Imaging of Cell Division and Phytohormone Signaling in Legume Roots and Root Nodules.  *Front. Plant Sci.*  **10**:1–12.
- Naisbitt, T., James, E. K., and Sprent, J. I. (1992). The evolutionary significance of the legume genus  *Chamaecrista* , as determined by nodule structure.  *New Phytol.*  **122**:487–492.
- Nekrasov, V., Staskawicz, B. J., Weigel, D., Jones, J. D. G., and Kamoun, S. (2013). Targeted mutagenesis in the model plant  *Nicotiana benthamiana*  using Cas9 RNA-guided endonuclease.  *Nat. Biotechnol.*  **31**:691–693.
- Ng, J. L. P., Hassan, S., Truong, T. T., Hocart, C. H., Laffont, C., Frugier, F., and Mathesius, U. (2015).  *Flavonoids and Auxin Transport Inhibitors Rescue Symbiotic Nodulation in the Medicago truncatula Cytokinin Perception Mutant cre1* .
- Nguyen, T. Van, Wibberg, D., Battenberg, K., Blom, J., Vanden Heuvel, B., Berry, A. M., Kalinowski, J., and Pawlowski, K. (2016). An assemblage of  *Frankia*  Cluster II strains from California contains the canonical  *nod*  genes and also the sulfotransferase gene  *nodH* .  *BMC Genomics*  **17**:796.
- Nguyen, T. Van, and Pawlowski, K. (2017).  *Frankia*  and Actinorhizal Plants: Symbiotic Nitrogen Fixation. In  *Rhizotrophs: Plant Growth Promotion to Bioremediation* , pp. 237–261.
- Nieminen, K., Immanen, J., Laxell, M., Kauppinen, L., Tarkowski, P., Dolezal, K., Tahtiharju, S., Elo, A., Decourteix, M., Ljung, K., et al. (2008). Cytokinin signaling regulates cambial development in poplar.  *Proc. Natl. Acad. Sci. U. S. A.*  **105**:20032–20037.
- Nilsson, O., and Olsson, O. (1997). Getting to the root: The role of the  *Agrobacterium rhizogenes*   *rol*  genes in the formation of hairy roots.  *Physiol. Plant.*  **100**:463–473.
- Nitschke, S., Cortleven, A., Iven, T., Feussner, I., Havaux, M., Riefler, M., and Schmölling, T. (2016). Circadian stress regimes affect the circadian clock and cause jasmonic acid-dependent cell death in cytokinin-deficient arabidopsis plants.  *Plant Cell*  **28**:1616–1639.
- Nukui, N., Ezura, H., and Minamisawa, K. (2004). Transgenic  *Lotus japonicus*  with an ethylene receptor gene  *Cm-ERS1/H70A*  enhances formation of infection threads and nodule primordia.  *Plant Cell Physiol.*  **45**:427–435.

- Oldroyd, G. E., Engstrom, E. M., and Long, S. R. (2001). Ethylene inhibits the Nod factor signal transduction pathway of *Medicago truncatula*. *Plant Cell* **13**:1835–49.
- Oldroyd, G. E. D., Murray, J. D., Poole, P. S., and Downie, J. A. (2011). The Rules of Engagement in the Legume-Rhizobial Symbiosis. *Annu. Rev. Genet.* **45**:119–144.
- Oldroyd, G. E. D. (2013). Speak, friend, and enter: signalling systems that promote beneficial symbiotic associations in plants. *Nat. Rev. Microbiol.* **11**:252–263.
- Op den Camp, R., Streng, A., De Mita, S., Cao, Q., Polone, E., Liu, W., Ammiraju, J.S.S., Kudrna, D., Wing, R., Untergasser, A., et al. (2011a). LysM-type mycorrhizal receptor recruited for rhizobium symbiosis in nonlegume *Parasponia*. *Science* **331**:909–912.
- Op den Camp, R. H. M., De Mita, S., Lillo, a., Cao, Q., Limpens, E., Bisseling, T., and Geurts, R. (2011b). A Phylogenetic Strategy Based on a Legume-Specific Whole Genome Duplication Yields Symbiotic Cytokinin Type-A Response Regulators. *Plant Physiol.* **157**:2013–2022.
- Op den Camp, R.H.M., Polone, E., Fedorova, E., Roelofsen, W., Squartini, A., Op den Camp, H.J.M., Bisseling, T., and Geurts, R. (2012). Nonlegume *Parasponia andersonii* deploys a broad Rhizobium host range strategy resulting in largely variable symbiotic effectiveness. *Mol. Plant. Microbe Interact.* **25**:954–963.
- Ovchinnikova, E., Journet, E. P., Chabaud, M., Cosson, V., Ratet, P., Duc, G., Fedorova, E., Liu, W., Op Den Camp, R., Zhukov, V., et al. (2011). IPD3 controls the formation of nitrogen-fixing symbiosomes in pea and *Medicago* spp. *Mol. Plant-Microbe Interact.* **24**:1333–1344.
- Pacios-Bras, C., Schlaman, H. R. M., Boot, K., Admiraal, P., Langerak, J. M., Stougaard, J., and Spaink, H. P. (2003). Auxin distribution in *Lotus japonicus* during root nodule development. *Plant Mol. Biol.* **52**:1169–1180.
- Parniske, M. (2008). Arbuscular mycorrhiza: The mother of plant root endosymbioses. *Nat. Rev. Microbiol.* **6**:763–775.
- Parniske, M. (2018). Uptake of bacteria into living plant cells, the unifying and distinct feature of the nitrogen-fixing root nodule symbiosis. *Curr. Opin. Plant Biol.* **44**:164–174.
- Pawlowski, K., and Demchenko, K. N. (2012). The diversity of actinorhizal symbiosis. *Protoplasma* **249**:967–979.
- Penmetsa, R.V., and Cook, D.R. (1997). A legume ethylene-insensitive mutant hyperinfected by its rhizobial symbiont. *Science* **275**:527–530.
- Penmetsa, R. V., Uribe, P., Anderson, J., Lichtenzveig, J., Gish, J. C., Nam, Y. W., Engstrom, E., Xu, K., Schisel, G., Pereira, M., et al. (2008). The *Medicago truncatula* ortholog of Arabidopsis EIN2, *sickle*, is a negative regulator of symbiotic and pathogenic microbial associations. *Plant J.* **55**:580–595.
- Peoples, M. B., Brockwell, J., Herridge, D. F., Rochester, I. J., Alves, B. J. R., Urquiaga, S., Boddey, R. M., Dakora, F. D., Bhattarai, S., Maskey, S. L., et al. (2009). The contributions of nitrogen-fixing crop legumes to the productivity of agricultural systems. *Symbiosis* **48**:1–17.
- Persson, T., Battenberg, K., Demina, I. V., Vigil-Stenman, T., Vanden Heuvel, B., Pujic, P., Facciotti, M. T., Wilbanks, E. G., O'Brien, A., Fournier, P., 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**:e0127630.
- Pimprakar, P., Carbonnel, S., Paries, M., Katzer, K., Klingl, V., Bohmer, M. J., Karl, L., Floss, D. S., Harrison, M. J., Parniske, M., et al. (2016). A CCaMK-CYCLOPS-DELLA Complex Activates Transcription of RAM1 to Regulate Arbuscule Branching. *Curr. Biol.* Advance Access published 2016, doi:10.1016/j.cub.2016.01.069.

- Plet, J., Wasson, A., Ariel, F., Le Signor, C., Baker, D., Mathesius, U., Crespi, M., and Frugier, F. (2011). MtCRE1-dependent cytokinin signaling integrates bacterial and plant cues to coordinate symbiotic nodule organogenesis in *Medicago truncatula*. *Plant J.* **65**:622–633.
- Price, M. N., Dehal, P. S., and Arkin, A. P. (2009). Fasttree: Computing large minimum evolution trees with profiles instead of a distance matrix. *Mol. Biol. Evol.* **26**:1641–1650.
- Pumplin, N., Mondo, S. J., Topp, S., Starker, C. G., Gantt, J. S., and Harrison, M. J. (2010). *Medicago truncatula* Vapyrin is a novel protein required for arbuscular mycorrhizal symbiosis. *Plant J.* **61**:482–494.
- Radutoiu, S., Madsen, L. H., Madsen, E. B., Felle, H. H., Umehara, Y., Gronlund, M., Sato, S., Nakamura, Y., Tabata, S., Sandal, N., et al. (2003). Plant recognition of symbiotic bacteria requires two LysM receptor-like kinases. *Nature* **425**:585–592.
- Reid, D. E., Heckmann, A. B., Novak, O., Kelly, S., and Stougaard, J. (2015). CYTOKININ OXIDASE/DEHYDROGENASE3 maintains cytokinin homeostasis during root and nodule development in *Lotus japonicus*. *Plant Physiol.* **170**:pp.15.00650.
- Reid, D., Nadzieja, M., Novák, O., Heckmann, A. B., Sandal, N., and Stougaard, J. (2017). Cytokinin biosynthesis promotes cortical cell responses during nodule development. *Plant Physiol.* **175**:361–375.
- Relic, B., Talmont, F., Kopcinska, J., Golinowski, W., Prome, J. C., and Broughton, W. J. (1993). Biological activity of *Rhizobium* sp. NGR234 Nod-factors on *Macroptilium atropurpureum*. *Mol. Plant-Microbe Interact.* **6**:764–774.
- Relic, B., Perret, X., Estradagarcia, M. T., Kopcinska, J., Golinowski, W., Krishnan, H. B., Pueppke, S. G., and Broughton, W. J. (1994). Nod factors of *Rhizobium* are a key to the legume door. *Mol. Microbiol.* **13**:171–178.
- Remigi, P., Zhu, J., Young, J. P. W., and Masson-Boivin, C. (2015). Symbiosis within Symbiosis: Evolving Nitrogen-Fixing Legume Symbionts. *Trends Microbiol.* **xx**:1–13.
- Ren, G. (2018). The evolution of determinate and indeterminate nodules within the Papilionoideae subfamily Advance Access published 2018.
- Rightmyer, A. P., and Long, S. R. (2011). Pseudonodule formation by wild-type and symbiotic mutant *Medicago truncatula* in response to auxin transport inhibitors. *Mol. Plant-Microbe Interact.* **24**:1372–1384.
- Rodpohong, P., Sullivan, J. T., Songsrirote, K., Sumpton, D., Cheung, K. W. J.-T., Thomas-Oates, J., Radutoiu, S., Stougaard, J., and Ronson, C. W. (2009). Nodulation gene mutants of *Mesorhizobium loti* R7A-nodZ and nolL mutants have host-specific phenotypes on *Lotus* spp. *Mol. Plant. Microbe. Interact.* **22**:1546–1554.
- Rodriguez-Barrueco, C., and Bermudez De Castro, F. (1973). Cytokinin-induced pseudonodules on *Alnus glutinosa*. *Physiol Plant* **29**:277–280.
- Roy, S., Robson, F., Lilley, J., Liu, C. W., Cheng, X., Wen, J., Walker, S., Sun, J., Cousins, D., Bone, C., et al. (2017). MtLAX2, a functional homologue of the arabidopsis auxin influx transporter AUX1, is required for nodule organogenesis. *Plant Physiol.* **174**:326–338.
- Roy, S., Liu, W., Nandety, R. S., Crook, A., Mysore, K. S., Pislariu, C. I., Frugoli, J., Dickstein, R., and Udvardi, M. K. (2020). Celebrating 20 Years of Genetic Discoveries in Legume Nodulation and Symbiotic Nitrogen Fixation. *Plant Cell* **32**:15–41.
- Rudich, J., Kedar, N., and Halevy, A.H. (1972). Ethylene evolution from cucumber plants as related to sex expression. *Plant Physiol.* **49**:998–&.
- Rutten, L., Miyata, K., Roswanjaya, Y. P., Huisman, R., Bu, F., Hartog, M., Linders, S., Van Velzen, R., van Zeijl, A., Bisseling, T., et al. (2020). Duplication of symbiotic Lysine Motif-receptors predates the evolution of nitrogen-fixing nodule symbiosis. *Plant Physiol.* doi:10.1104/pp.19.01420.

- Saito, K., Yoshikawa, M., Yano, K., Miwa, H., Uchida, H., Asamizu, E., Sato, S., Tabata, S., Imaizumi-Anraku, H., Umehara, Y., et al. (2007). NUCLEOPORIN85 is required for calcium spiking, fungal and bacterial symbioses, and seed production in *Lotus japonicus*. *Plant Cell* **19**:610–624.
- Santos, J. M. F. dos, Casaes Alves, P. A., Silva, V. C., Kruschewsky Rhem, M. F., James, E. K., and Gross, E. (2017). Diverse genotypes of Bradyrhizobium nodulate herbaceous Chamaecrista (Moench) (Fabaceae, Caesalpinioideae) species in Brazil. *Syst. Appl. Microbiol.* **40**:69–79.
- Sato, S., Nakamura, Y., Kaneko, T., Asamizu, E., Kato, T., Nakao, M., Sasamoto, S., Watanabe, A., Ono, A., Kawashima, K., et al. (2008). Genome Structure of the Legume, *Lotus japonicus*. *DNA Res.* **15**:227–239.
- Schaller, G. E., Bishopp, A., and Kieber, J. J. (2015). The Yin-Yang of Hormones: Cytokinin and Auxin Interactions in Plant Development. *Plant Cell* **27**:44–63.
- Schauser, L., Roussis, A., Stiller, J., and Stougaard, J. (1999). A plant regulator controlling development of symbiotic root nodules. *Nature* **402**:191–195.
- Schenk, R. U., and Hildebrandt, A. C. (1972). Medium and techniques for induction and growth of monocotyledonous and dicotyledonous plant cell cultures. *Can. J. Bot.* **50**:199–204.
- Scheres, B., McKhann, H. I., Zalensky, A., Löbler, M., Bisseling, T., and Hirsch, A. M. (1992). The PsENOD12 gene is expressed at two different sites in afghanistan pea pseudonodules induced by auxin transport inhibitors. *Plant Physiol.* **100**:1649–1655.
- Schiessl, K., Lilley, J. L. S., Lee, T., Tamvakis, I., Kohlen, W., Bailey, P. C., Thomas, A., Luptak, J., Ramakrishnan, K., Carpenter, M. D., et al. (2019). NODULE INCEPTION Recruits the Lateral Root Developmental Program for Symbiotic Nodule Organogenesis in *Medicago truncatula*. *Curr. Biol.* **29**:3657–3668.e5.
- Schmutz, J., Cannon, S.B., Schlueter, J., Ma, J.X., Mitros, T., Nelson, W., Hyten, D.L., Song, Q.J., Thelen, J.J., Cheng, J.L., et al. (2010). Genome sequence of the palaeopolyploid soybean. *Nature* **463**:178–183.
- Schneiderberg, M., Schmitz, L., Cheng, X., Polman, S., Franken, C., Geurts, R., and Bisseling, T. (2018). A genetically and functionally diverse group of non-diazotrophic Bradyrhizobium spp. colonizes the root endophytic compartment of *Arabidopsis thaliana*. *BMC Plant Biol.* **18**:1–9.
- Shen, D., Xiao, T. T., van Velzen, R., Kulikova, O., Gong, X., Geurts, R., Pawlowski, K., and Bisseling, T. (2020). A Homeotic Mutation Changes Legume Nodule Ontogeny into Actinorhizal-Type Ontogeny. *Plant Cell* **32**:1868–1885.
- Shtark, O.Y., Sulima, A.S., Zhernakov, A.I., Kliukova, M.S., Fedorina, J.V., Pinaev, A.G., Kryukov, A.A., Akhtemova, G.A., Tikhonovich, I.A., and Zhukov, V.A. (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**:129–144.
- Sieberer, B. J., Chabaud, M., Timmers, A. C., Monin, A., Fournier, J., and Barker, D. G. (2009). A nuclear-targetedameleon demonstrates intranuclear Ca<sup>2+</sup> spiking in *Medicago truncatula* root hairs in response to rhizobial nodulation factors. *Plant Physiol.* **151**:1197–1206.
- Silvester, W. B., Berg, R. H., Schwintzer, C. R., and Tjepkema, J. D. (2008). *Nitrogen-fixing Actinorhizal Symbioses*. (ed. Pawlowski, K.) and Newton, W. E.).
- Singer, S. R., Maki, S. L., Farmer, A. D., Ilut, D., May, G. D., Cannon, S. B., and Doyle, J. J. (2009). Venturing beyond beans and peas: What can we learn from *Chamaecrista*? *Plant Physiol.* **151**:1041–1047.
- Singh, S., Katzer, K., Lambert, J., Cerri, M., and Parniske, M. (2014). CYCLOPS, a DNA-binding transcriptional activator, orchestrates symbiotic root nodule development. *Cell Host Microbe* **15**:139–152.

- Smit, P., Raedts, J., Portyanko, V., Debellé, F., Gough, C., Bisseling, T., and Geurts, R.** (2005). NSP1 of the GRAS protein family is essential for rhizobial nod factor-induced transcription. *Science*. **308**:1789–1791.
- Smit, P., Limpens, E., Geurts, R., Fedorova, E., Dolgikh, E., Gough, C., and Bisseling, T.** (2007). Medicago LYK3, an entry receptor in rhizobial nodulation factor signaling. *Plant Physiol.* **145**:183–191.
- Smulders, M. J. M., and de Klerk, G. J.** (2011). Epigenetics in plant tissue culture. *Plant Growth Regul.* **63**:137–146.
- Society, B., and Journal, A.** (2009). Environmental Contribution to Floral Trait Variation in *Chamaecrista fasciculata* ( Fabaceae : Caesalpinoideae ) Author ( s ): Joan E . Frazee and Robert J . Marquis Published by : *Botanical Society of America Stable URL* : <http://www.jstor.org/stable/24456>. *America (NY)*. **81**:206–215.
- Soepadmo, E.** (1974). Ulmaceae. *Flora Malesiana-Series 1, Spermatophyta* **8**:31–76.
- Soltis, D. E., Soltis, P. S., Morgan, D. R., Swensen, S. M., Mullin, B. C., Dowd, J. M., and Martin, P. G.** (1995). Chloroplast gene sequence data suggest a single origin of the predisposition for symbiotic nitrogen fixation in angiosperms. *Proc. Natl. Acad. Sci. U. S. A.* **92**:2647–2651.
- Soyano, T., Kouchi, H., Hirota, A., and Hayashi, M.** (2013). NODULE INCEPTION Directly Targets NF-Y Subunit Genes to Regulate Essential Processes of Root Nodule Development in *Lotus japonicus*. *PLoS Genet.* **9**:e1003352.
- Soyano, T., Hirakawa, H., Sato, S., Hayashi, M., and Kawaguchi, M.** (2014). NODULE INCEPTION creates a long-distance negative feedback loop involved in homeostatic regulation of nodule organ production. *Proc. Natl. Acad. Sci. U. S. A.* **111**:14607–14612.
- Soyano, T., Shimoda, Y., and Hayashi, M.** (2015). NODULE INCEPTION antagonistically regulates gene expression with nitrate in *Lotus japonicus*. *Plant Cell Physiol.* **56**:368–376.
- Soyano, T., Shimoda, Y., Kawaguchi, M., and Hayashi, M.** (2019). A shared gene drives lateral root development and root nodule symbiosis pathways in *Lotus*. *Science (80-. )*. **366**:1021–1023.
- Spaink, H. P., Okker, R. J. H., Wijffelman, C. A., Pees, E., and Lugtenberg, B. J. J.** (1987). Promoters in the nodulation region of the *Rhizobium leguminosarum* Sym plasmid pRL1J1. *Plant Mol. Biol.* **9**:27–39.
- Spaink, H. P., Sheeley, D. M., van Brussel, A. A. N., Glushka, J., York, W. S., Tak, T., Geiger, O., Kennedy, E. P., Reinhold, V. N., and Lugtenberg, B. J. J.** (1991). A novel highly unsaturated fatty acid moiety of lipo-oligosaccharide signals determines host specificity of *Rhizobium*. *Nature* **354**:125–130.
- Sprent, J. I., Ardley, J. K., and James, E. K.** (2013). From North to South: A latitudinal look at legume nodulation processes. *South African J. Bot.* **89**:31–41.
- Sprent, J. I., Ardley, J., and James, E. K.** (2017). Biogeography of nodulated legumes and their nitrogen-fixing symbionts. *New Phytol.* **215**:40–56.
- Stokkermans, T. J. W., and Peters, N. K.** (1994). Bradyrhizobium elkanii lipo-oligosaccharide signals induce complete nodule structures on *Glycine soja* Siebold et Zucc. *Planta* **193**:413–420.
- Stracke, S., Kistner, C., Yoshida, S., Mulder, L., Sato, S., Kaneko, T., Tabata, S., Sandal, N., Stougaard, J., Szczyglowski, K., et al.** (2002). A plant receptor-like kinase required for both bacterial and fungal symbiosis. *Nature* **417**:959–962.
- Street, I. H., Aman, S., Zubo, Y., Ramzan, A., Wang, X., Shakeel, S. N., Kieber, J. J., and Eric Schaller, G.** (2015). Ethylene inhibits cell proliferation of the arabidopsis root meristem. *Plant Physiol.* **169**:338–350.
- Sturms, R., Kakar, S., Trent, J., and Hargrove, M. S.** (2010). Trema and parasponia hemoglobins reveal convergent evolution of oxygen transport in plants. *Biochemistry* **49**:4085–4093.
- Su, Y. H., Liu, Y. B., and Zhang, X. S.** (2011). Auxin-cytokinin interaction regulates meristem development. *Mol. Plant* **4**:616–625.

- Suzaki, T., Yano, K., Ito, M., Umehara, Y., Sukanuma, N., and Kawaguchi, M.** (2012). Positive and negative regulation of cortical cell division during root nodule development in *Lotus japonicus* is accompanied by auxin response. *Development* **139**:3997–4006.
- Suzaki, T., Ito, M., and Kawaguchi, M.** (2013). Genetic basis of cytokinin and auxin functions during root nodule development. *Plant Cell* **4**:42.
- Svistoonoff, S., Hocher, V., and Gherbi, H.** (2014). Actinorhizal root nodule symbioses: What is signalling telling on the origins of nodulation? *Curr. Opin. Plant Biol.* **20**:11–18.
- Swensen, S. M.** (2010). The Evolution of Actinorhizal Symbioses : Evidence for Multiple Origins of the Symbiotic Association Author ( s ): Susan M . Swensen Source : American Journal of Botany , Vol . 83 , No . 11 ( Nov . , 1996 ) , pp . 1503-1512 Published by : Botanical Society **83**:1503–1512.
- Takahashi, N., Kajihara, T., Okamura, C., Kim, Y., Katagiri, Y., Okushima, Y., Matsunaga, S., Hwang, I., and Umeda, M.** (2013). Cytokinins control endocycle onset by promoting the expression of an APC/C activator in arabidopsis roots. *Curr. Biol.* **23**:1812–1817.
- Takahashi, N., and Umeda, M.** (2014). Cytokinins promote onset of endoreplication by controlling cell cycle machinery. *Plant Signal. Behav.* **9**:11–13.
- Takei, K., Sakakibara, H., and Sugiyama, T.** (2001). Identification of Genes Encoding Adenylate Isopentenyltransferase, a Cytokinin Biosynthesis Enzyme, in *Arabidopsis thaliana*. *J. Biol. Chem.* **276**:26405–26410.
- Takei, K., Yamaya, T., and Sakakibara, H.** (2004). *Arabidopsis* CYP735A1 and CYP735A2 encode cytokinin hydroxylases that catalyse the biosynthesis of trans-Zeatin. *J. Biol. Chem.* **279**:41866–41872.
- Tan, S., Debellé, F., Gamas, P., Frugier, F., and Brault, M.** (2019). Diversification of cytokinin phosphotransfer signaling genes in *Medicago truncatula* and other legume genomes. *BMC Genomics* **20**:1–19.
- Tan, S., Sanchez, M., Laffont, C., Boivin, S., Le Signor, C., Thompson, R., Frugier, F., and Brault, M.** (2020). A Cytokinin Signaling Type-B Response Regulator Transcription Factor Acting in Early Nodulation. *Plant Physiol.* **183**:1319–1330.
- Tang, H.B., Krishnakumar, V., Bidwell, S., Rosen, B., Chan, A.N., Zhou, S.G., Gentzbittel, L., Childs, K.L., Yandell, M., Gundlach, H., et al.** (2014). An improved genome release (version Mt4.0) for the model legume *Medicago truncatula*. *BMC Genomics* **15**:312.
- Tanurdzic, M., and Banks, J.A.** (2004). Sex-determining mechanisms in land plants. *Plant Cell* **16**:S61–S71.
- Thatcher, L.F., Carrie, C., Andersson, C.R., Sivasithamparam, K., Whelan, J., and Singh, K.B.** (2007). Differential gene expression and subcellular targeting of *Arabidopsis* glutathione S-transferase F8 is achieved through alternative transcription start sites. *J. Biol. Chem.* **282**:28915–28928.
- Timmers, A.C.J., Auriac, M-C., and Truchet, G.** (1999). Refined analysis of early symbiotic steps of the *Rhizobium-Medicago* interaction in relationship with microtubular cytoskeleton rearrangements. *Development* **126**:3617–3628.
- Tirichine, L., James, E. K., Sandal, N., and Stougaard, J.** (2006). Spontaneous root-nodule formation in the model legume *Lotus japonicus*: a novel class of mutants nodulates in the absence of rhizobia. *Mol. Plant-Microbe Interact* **19**:373–382.
- Tirichine, L., Sandal, N., Madsen, L. H., Radutoiu, S., Albrechtsen, A. S., Sato, S., Asamizu, E., Tabata, S., and Stougaard, J.** (2007). A gain-of-function mutation in a cytokinin receptor triggers spontaneous root nodule organogenesis. *Science* **315**:104–7.
- Trinick, M.J.** (1973). Symbiosis between *Rhizobium* and the non-legume, *Trema aspera*. *Nature* **244**:459–460.
- Trinick, M.J.** (1980). Growth of *Parasponia* in agar tube culture and symbiotic effectiveness of isolates from *Parasponia* spp. *New Phytol.* **85**:37–45.

- Trinick, M.J.** (1981). The effective rhizobium symbiosis with the non-legume *Parasponia andersonii*. In: Current Perspectives in Nitrogen Fixation--Gibson, A.H., and Newton, W.E., eds.: Australian Academy of Sciences. 480.
- Trinick, M.J., and Hadobas, P.A.** (1989). Biology of the *Parasponia-Bradyrhizobium* symbiosis. In: Nitrogen Fixation with Non-Legumes: Springer Netherlands. 25-33.
- Trouvelot, A., Kough, J. L., and Gianinazzi-Pearson, V.** (1986). Mesure du taux de mycorhization VA d'un système racinaire. Recherche de méthodes d'estimation ayant une signification fonctionnelle. *Physiol. Genet. Asp. Mycorrhizae* Advance Access published 1986.
- Truchet, G., Roche, P., Lerouge, P., Vasse, J., Camut, S., Debilly, F., Prome, J. C., and Denarie, J.** (1991). Sulphated lipo-oligosaccharide signals of *Rhizobium meliloti* elicit root nodule organogenesis in alfalfa. *Nature* **351**:670-673.
- Tuskan, G.A., and DiFazio, S., and Jansson, S., and Bohlmann, J., and Grigoriev, I., and Hellsten, U., and Putnam, N., and Ralph, S., and Rombauts, S., and Salamov, A., et al.** (2006). The genome of black cottonwood, *Populus trichocarpa* (Torr. & Gray). *Science*. **313**:1596-1604.
- Udvardi, M., and Poole, P. S.** (2013). Transport and metabolism in legume-rhizobia symbioses. *Annu. Rev. Plant Biol.* **64**:781-805.
- Ueguchi-Tanaka, M., Ashikari, M., Nakajima, M., Itoh, H., Katoh, E., Kobayashi, M., Chow, T., Hsing, Y. C., Kitano, H., Yamaguchi, I., et al.** (2005). GIBBERELLIN INSENSITIVE DWARF1 encodes a soluble receptor for gibberellin. *Nature* **437**:693-698.
- Van de Velde, W., Guerra, J. C. P., De Keyser, A., De Rycke, R., Rombauts, S., Maunoury, N., Mergaert, P., Kondorosi, E., Holsters, M., and Goormachtig, S.** (2006). Aging in legume symbiosis. A molecular view on nodule senescence in *Medicago truncatula*. *Plant Physiol.* **141**:711-720.
- van Velzen, R., Holmer, R., Bu, F., Rutten, L., van Zeijl, A., Liu, W., Santuari, L., Cao, Q., Sharma, T., Shen, D., et al.** (2017). Parallel loss of symbiosis genes in relatives of nitrogen-fixing non-legume *Parasponia*. bioRxiv, doi: 10.1101/169706.
- van Velzen, R., Holmer, R., Bu, F., Rutten, L., Van Zeijl, A., Liu, W., Santuari, L., Cao, Q., Sharma, T., Shen, D., et al.** (2018). Comparative genomics of the nonlegume *Parasponia* reveals insights into evolution of nitrogen-fixing rhizobium symbioses. *Proc Natl Acad Sci U S A* Advance Access published 2018, doi:10.1073/pnas.1721395115.
- van Velzen, R., Doyle, J. J., and Geurts, R.** (2019). A Resurrected Scenario: Single Gain and Massive Loss of Nitrogen-Fixing Nodulation. *Trends Plant Sci.* **24**:49-57.
- van Zeijl, A., Op den Camp, R.H.M., Deinum, E.E., Charnikhova, T., Franssen, H., Op den Camp, H.J.M., Bouwmeester, H., Kohlen, W., Bisseling, T., and Geurts, R.** (2015a). Rhizobium lipo-chitoooligosaccharide signaling triggers accumulation of cytokinins in *Medicago truncatula* roots. *Mol. Plant* **8**:1213-1226.
- van Zeijl, A., Liu, W., Xiao, T.T., Kohlen, W., Yang, W.-C., Bisseling, T., and Geurts, R.R.** (2015b). The strigolactone biosynthesis gene *DWARF27* is co-opted in rhizobium symbiosis. *BMC Plant Biol.* **15**:260.
- van Zeijl, A.** (2017). Dissecting hormonal pathways in nitrogen-fixing rhizobium symbioses. Wageningen University PhD thesis, Wageningen, The Netherlands, doi: 10.18174/419674.
- van Zeijl, A., Wardhani, T. A. K., Seifi Kalhor, M., Rutten, L., Bu, F., Hartog, M., Linders, S., Fedorova, E. E., Bisseling, T., Kohlen, W., et al.** (2018). CRISPR/Cas9-Mediated Mutagenesis of Four Putative Symbiosis Genes of the Tropical Tree *Parasponia andersonii* Reveals Novel Phenotypes. *Front. Plant Sci.* **9**:1-14.
- Venkateshwaran, M., Jayaraman, D., Chabaud, M., Genre, A., Balloon, A. J., Maeda, J., Forshey, K., den Os, D., Kwiecien, N. W., Coon, J. J., et al.** (2015). A role for the mevalonate pathway in early plant symbiotic signaling. *Proc. Natl. Acad. Sci. U. S. A.* **112**:9781-9786.

- Vernie, T., Kim, J., Frances, L., Ding, Y., Sun, J., Guan, D., Niebel, A., Gifford, M.L., de Carvalho-Niebel, F., and Oldroyd, G.E.D. (2015) The NIN transcription factor coordinates diverse nodulation programs in different tissues of the *Medicago truncatula* root. *Plant Cell* **27**:3410–3424.
- Vinardell, J. M., Fedorova, E., Cebolla, A., Kevei, Z., Horvath, G., Kelemen, Z., Tarayre, S., Roudier, F., Mergaert, P., Kondorosi, A., et al. (2003). Endoreduplication mediated by the anaphase-promoting complex activator CCS52A is required for symbiotic cell differentiation in *Medicago truncatula* nodules. *Plant Cell* **15**:2093–2105.
- Wang, H., Moore, M.J., Soltis, P.S., Bell, C.D., Brockington, S.F., Alexandre, R., Davis, C.C., Latvis, M., Manchester, S.R., and Soltis, D.E. (2009). Rosid radiation and the rapid rise of angiosperm-dominated forests. *Proc. Natl. Acad. Sci. U. S. A.* **106**:3853–3858.
- Wang, K. (2015). *Methods in Molecular Biology: Agrobacterium Protocols*. 3rd ed. (ed. Wang, K.) New York, Heidelberg, Dordrecht, London: Humana Press.
- Wang, Y., Yang, W., Zuo, Y., Zhu, L., Hastwell, A. H., Chen, L., Tian, Y., Su, C., Ferguson, B. J., and Li, X. (2019). GmYUC2a mediates auxin biosynthesis during root development and nodulation in soybean. *J. Exp. Bot.* **70**:3165–3176.
- Wardhani, T. A. K., Roswanjaya, Y. P., Dupin, S., Li, H., Linders, S., Hartog, M., Geurts, R., and van Zeijl, A. (2019). Transforming, genome editing and phenotyping the nitrogen-fixing tropical cannabaceae tree *Parasponia andersonii*. *J. Vis. Exp.* **2019**:1–17.
- Weber, E., Engler, C., Gruetzner, R., Werner, S., and Marillonnet, S. (2011). A modular cloning system for standardized assembly of multigene constructs. *PLoS One* **6**.
- Webster, G., Poulton, P. R., Cocking, E. C., and Davey, M. R. (1995). The nodulation of micro-propagated plants of *Parasponia andersonii* by tropical legume rhizobia. *J. Exp. Bot.* **46**:1131–1137.
- Werner, G. D. A., Cornwell, W. K., Cornelissen, J. H. C., and Kiers, E. T. (2015). Evolutionary signals of symbiotic persistence in the legume–rhizobia mutualism. *Proc. Natl. Acad. Sci.* **112**:201424030.
- Werner, G. D. A., Cornwell, W. K., Sprent, J. I., Kattge, J., Kiers, E. T., and Werner (2014). A single evolutionary innovation drives the deep evolution of symbiotic N<sub>2</sub>-fixation in angiosperms. *Nat. Commun.* **5**:4087.
- Xiao, T. T., Schilderink, S., Moling, S., Deinum, E. E., Kondorosi, E., Franssen, H., Kulikova, O., Niebel, A., and Bisseling, T. (2014). Fate map of *Medicago truncatula* root nodules. *Development* **141**:3517–3528.
- Xie, F., Murray, J. D., Kim, J., Heckmann, A. B., Edwards, A., Oldroyd, G. E. D., and Downie, J. A. (2012). Legume pectate lyase required for root infection by rhizobia. *Proc. Natl. Acad. Sci. U. S. A.* **109**:633–638.
- Xing, H.-L., Dong, L., Wang, Z.-P., Zhang, H.-Y., Han, C.-Y., Liu, B., Wang, X.-C., and Chen, Q.-J. (2014). A CRISPR/Cas9 toolkit for multiplex genome editing in plants. *BMC Plant Biol.* **14**:327.
- Yang, W. C., de Blank, C., Meskiene, I., Hirt, H., Bakker, J., van Kammen, A., Franssen, H., and Bisseling, T. (1994). Rhizobium nod factors reactivate the cell cycle during infection and nodule primordium formation, but the cycle is only completed in primordium formation. *Plant Cell* **6**:1415–26.
- Yano, K., Yoshida, S., Müller, J., Singh, S., Banba, M., Vickers, K., Markmann, K., White, C., Schuller, B., Sato, S., et al. (2008). CYCLOPS, a mediator of symbiotic intracellular accommodation. *Proc. Natl. Acad. Sci. U. S. A.* **105**:20540–20545.
- Yin, T.J., and Quinn, J.A. (1995). Tests of a mechanistic model of one hormone regulating both sexes in *Cucumis sativus* (Cucurbitaceae). *Am. J. Bot.* **82**:1537–1546.
- Yoro, E., Suzuki, T., Toyokura, K., Miyazawa, H., Fukaki, H., and Kawaguchi, M. (2014). A Positive regulator of nodule organogenesis, NODULE INCEPTION, acts as a negative regulator of rhizobial infection in *Lotus japonicus*. *Plant Physiol.* **165**:747–758.

- Young, J. P. W., and Haukka, K. E.** (1996). Diversity and phylogeny of rhizobia. *New Phytol.* **133**:87–94.
- Young, N. D. N., Debellé, F., Oldroyd, G. E. D. G., Geurts, R., Cannon, S. B., Udvardi, M. K., Benedito, V. A., Mayer, K. F. X., Gouzy, J., Schoof, H., et al.** (2011). The Medicago Genome Provides Insight into the Evolution of Rhizobial Symbioses. *Nature* **480**:520–524.
- Zhang, Y.X., van Dijk, A.D.J., Scaffidi, A., Flematti, G.R., Hofmann, M., Charnikhova, T., Verstappen, F., Hepworth, J., van der Krol, S., Leyser, O., et al.** (2014). Rice cytochrome P450 MAX1 homologs catalyze distinct steps in strigolactone biosynthesis. *Nat. Chem. Biol.* **10**:1028–1033.
- Zürcher, E., Tavor-Deslex, D., Lituiev, D., Enkerli, K., Tarr, P. T., and Müller, B.** (2013). A robust and sensitive synthetic sensor to monitor the transcriptional output of the cytokinin signaling network in planta. *Plant Physiol.* **161**:1066–75.





# Summary

**Titis A. K. Wardhani<sup>1</sup>**

<sup>1</sup>Laboratory of Molecular Biology, Department of Plant Sciences, Wageningen University & Research, Wageningen, The Netherlands





## Summary

Nitrogen-fixing symbioses provide a beneficial platform for plants to sustain the growth and development under the nitrogen-limiting environment. The nitrogen-fixing symbioses occur exclusively in a subset of genera of 10 taxonomic families. These nodulating plant genera form a paraphyletic clade in the orders Fabales, Fagales, Cucurbitales and Fagales, which collectively are known as 'Nitrogen Fixing Clade'. Nodulating species of eight families establish the symbiotic interaction with the filamentous bacteria *Frankia*, and therefore are classified as actinorhizal plants. Nodulating species of two other families interact with the nitrogen-fixing rhizobium; namely legumes (Fabaceae, Fabales) and *Parasponia* (Cannabaceae, Rosales). Both nitrogen-fixing microbes induce the formation of a niche organ, so-called root nodules. Despite varieties in nodule morphology across nodulating species, the root nodules serve a general purpose to accommodate the bacteria intracellularly and to provide a suitable physiological environment for the bacteria to convert the atmospheric nitrogen ( $N_2$ ) into ammonium.

A major aim in nitrogen-fixing nodulation research is to unravel the signalling pathways required for a successful mutualism. Recently, phylogenomic comparisons of nodulating and non-nodulating species postulated that the nodulation trait likely originated from a common ancestor of the Nitrogen-Fixing Clade. This implies that the nodulation trait encountered massive parallel loss during the divergence of this clade. Considering the contribution of nitrogen-fixing symbioses to sustainable plant growth under low nitrogen input farming, it leads to a question whether it is possible to extend the host range of nitrogen-fixing symbionts to non-nodulating crops (such as wheat, maize, or rice). To this end, it is important to identify the key genetic adaptations underlying these successful symbioses. To date, a massive body of knowledge on nodulation has been generated in molecular genetic studies on two model species; *Medicago truncatula* and *Lotus japonicus* which belong to the legume subfamily of Papilionoideae. In accordance with the single gain hypothesis, it implies that the core genetic adaptations underlying the nodulation trait are conserved in all nodulating plant species and orchestrated via a shared genetic network. Therefore, expanding the nodulation research outside the Papilionoideae clade is highly relevant to pinpoint the core set of nodulation genes, and to reveal the lineage-specific genetic adaptations that have shaped the symbiotic interactions.

In this PhD thesis, I utilise two alternative species to study the nodulation trait, namely *Chamaecrista mimosoides* and *Parasponia andersonii*. *C. mimosoides* belongs to the legume subfamily of Caesalpinoideae, in which the occurrence of nodulating species is far less frequent than in the subfamily of Papilionoideae. Caesalpinoideae also represents a basal node in the divergence of the legume family, and therefore functional studies

*Chamaecrista* species may offer insights in basal adaptations in the legume-rhizobium interactions. The second species, *P. andersonii* is a most distantly related nodulating species to legumes. As both lineages diverged at the root of the nitrogen-fixing clade, and therefore studying *Parasponia* in the context of rhizobial symbiosis may provide a platform to identify the core genetic adaptations underlying the nodulation trait.

Genetic screens in *M. truncatula* and *L. japonicus* uncovered synchronised processes leading to the formation of functional nodules. Rhizobial chitin oligomer-based signal molecules (namely, LCOs or Nod-Factors) are perceived by the plant by specific LysM-type receptors, which initiate signalling cascade that sets in motion nodule organogenesis and rhizobium intracellular infection. It was found that cytokinin signalling is an integral part of Nod-Factors induced signalling, and essential for nodule organogenesis. Cytokinin signalling is tightly linked to the functional activity of transcription factor NIN. The perception of cytokinin is mainly regulated by histidine kinase cytokinin receptor MtCRE1 in *M. truncatula* and LjLHK1 in *L. japonicus*. Knockout mutations in these receptors result in a dramatic reduction of nodule numbers, whereas a dominant-active mutant MtCRE1/LjLHK1 form can induce spontaneous nodule formation. This underlines the importance of cytokinin signalling in nodulation of *M. truncatula* and *L. japonicus*. However, it remains elusive whether symbiotic cytokinin signalling is conserved in nodulating species outside the legume Papilionoideae subfamily.

In **Chapter 2**, I show that the cytokinin-NIN regulation is conserved in *C. mimosoides*. Rhizobium inoculation triggers the expression of *Type A RR* cytokinin responsive genes that are orthologous to the LCO responsive MtARR4 and MtARR9 genes in *M. truncatula*. Also, two *NIN* genes are transcriptionally activated. This suggests that the cytokinin signalling and *NIN* expression are Nod-Factors responsive in *C. mimosoides*. This finding is consistent with the observations in the model legumes. Functional studies in *C. mimosoides* mediated by RNA interference (RNAi) targeting either single *CmNIN1* or *CmNIN2*, and double *CmNIN1;CmNIN2* resulted in a significant reduction in nodule numbers. In particular silencing of *CmNIN2*, resulted also in nodules devoid of a fixation zone. This suggests that *CmNIN2* may have an important function not only in nodule organogenesis, but also for successful release of bacteria from infection threads. A similar RNAi strategy was applied to target *C. mimosoides* cytokinin histidine kinase receptor *CmHK4*. Also, *CmHK4 RNAi* lines possess less functional nodules when compared to wild type control plants. This suggests that nodule organogenesis is HK4-dependent. Therefore, I hypothesize that the MtCRE1/LjLHK1/CmHK4-mediated cytokinin and *NIN* signalling is a conserved mechanism in nodule formation of Papilionoideae and Caesalpinoideae legumes. However, further studies in *C. mimosoides* are still required to substantiate the interconnection between the cytokinin signalling and NIN.

In **Chapter 3**, I present ‘a mini handbook’ for the use of *P. andersonii* as an experimental system. This chapter describes the elaborative protocols on performing reverse genetic analyses in *P. andersonii*. This protocol is based on the CRISPR/Cas9-mutagenesis in *Agrobacterium tumefaciens*-transformed plants. Using this protocol, stable transgenic lines can be generated within three months. Since the *in vitro* propagation is highly efficient in *P. andersonii*, phenotypic analyses of symbiotic responses can be studied in T<sub>0</sub> generation lines within six months after transformation. Therefore, I argue that *P. andersonii* is an excellent alternative experimental system to study mechanisms underlying the nitrogen-fixing nodulation.

In **Chapter 4**, I present a proof of concept of the reverse genetic analyses in *P. andersonii*, by targeting four putative symbiosis genes; namely *PanHK4*, *PanEIN2*, *PanNSP1*, or *PanNSP2*. These genes are known to be involved in the cytokinin and ethylene signalling, and the regulation of strigolactone biosynthesis, respectively. Additionally, these genes also exhibit essential functions in root nodulation of legumes. Loss-of-functions of *Pannsp1* and *Pannsp2* revealed a conserved role for NSP1 and NSP2 to regulate nodule organogenesis. Also, both mutants are affected in the expression of the strigolactone biosynthesis genes *D27* and *MAX1*. The *Panein2* mutant also displayed a symbiotic phenotype, though clearly distinct from what is found in legumes. Instead of displaying a hypernodulation phenotype as seen in legume *ein2* mutants, *Panein2* mutants developed nodules without fixation zones. Analyses of *Panhk4* mutants revealed no symbiotic phenotypes. However, these mutants, as well as *Panein2* showed non-symbiotic phenotypes that correlate with the reduced cytokinin and ethylene signalling. Altogether, this study provides a valuable example of utilising *P. andersonii* as an experimental system to study the rhizobial-induced symbiotic interaction, alongside the model legumes.

In **Chapter 5**, I present an in-depth analysis on the role of cytokinin signalling in the *P. andersonii*-rhizobium interaction. Utilising the cytokinin synthetic reporter TCSn, it is shown that the cytokinin signalling induced upon rhizobium and Nod factor application. TCSn remains active during nodule organogenesis in cells that become infected by rhizobium. By analysing the spatiotemporal expression of the three *P. andersonii* cytokinin HK genes, we found an expression pattern of *PanHK3* that largely overlaps with TCSn during nodulation. RNAi knockdown studies targeting different cytokinin receptors simultaneously (*PanHK2*, *PanHK3*) suggests some functional redundancy among three of the cytokinin receptors in regulating nodule organogenesis. Additionally, mutagenesis of *PanHK3* resulted in nodules that are devoid of fixation threads. Taken together, I argue that the cytokinin signalling is integrated in the rhizobial-induced nodulation of *P. andersonii*.

In summary, this thesis provides an example of comparative genetic analyses of the nodulation trait by utilising two species outside the legume Papilionoideae subfamily. Results that are presented in this thesis illustrate the integration of cytokinin signalling in the rhizobium-induced nodulation. This suggests that the Nod factors-induced cytokinin signalling represents a conserved mechanism in nodulation, which has been recruited in the common ancestor of the Nitrogen-Fixing Clade.

## Samenvatting

Een stikstof-bindende symbiose helpt planten om te blijven groeien en ontwikkelen onder stikstof-limiterende condities. Stikstof-bindende symbioses komen echter enkel voor in een subset van geslachten die behoren tot 10 taxonomische families die behoren tot de Deze stikstofbindende geslachten vormen samen een parafylitische clade in de orden Fabales, Fagales, Cucurbitales en Rosales, die gezamenlijk bekend staat als de 'stikstofbindende clade'. Stikstofbindende planten uit acht geslachten gaan een symbiotische relatie aan met bacteriën van het geslacht *Frankia*. Deze planten worden ook wel actinorhiza-planten genoemd. De stikstofbindende-planten uit twee andere families interacteren met stikstofbindende rhizobium; namelijk de vlinderbloemigen (Fabaceae, Fabales) en *Parasponia* (Cannabaceae, Rosales). Zowel rhizobium en *Frankia* induceren de aanleg van een niche orgaan, het wortelknolletje. Ondanks een verschil in morfologie van het wortelknolletje tussen verschillende nodulerende soorten, hebben de wortelknolletjes eenzelfde functie namelijk het intracellulair huisvesten van de bacteriën en het creëren van een geschikte fysiologische omgeving voor omzetting van atmosferische stikstof ( $N_2$ ) in ammonium.

Een belangrijk doel van het onderzoek aan stikstofbindende wortelknolletjes is het ontrafelen van de signaleringsroutes die essentieel zijn voor een succesvol mutualisme. Onlangs hebben fylogenomische vergelijkingen tussen wortelknolvormende en niet-wortelknolvormende soorten gepostuleerd dat het kenmerk wortelknolvorming ontstaan is in een gemeenschappelijke voorouder van de stikstofbindende clade. Aangezien stikstofbindende symbioses een bijdrage kunnen leveren aan duurzame plantengroei bij landbouw met een lage stikstofinbreng, een belangrijke vraag is of het mogelijk is om een stikstofbindende wortelknolsymbiose over te brengen naar niet-wortelknolvormende gewassen (zoals tarwe, mais of rijst). Om deze reden is het noodzakelijk de belangrijkste genetische veranderingen die ten grondslag liggen aan een succesvolle symbiose te identificeren. Tot op heden is er enorm veel kennis over wortelknolvorming opgedaan op basis van moleculair genetische studies aan twee modelsoorten; *Medicago truncatula* en *Lotus japonicus* die beiden behoren tot de vlinderbloemigen onderfamilie Papilionoideae. De hypothese dat wortelknolvorming eenmaal is ontstaan impliceert dat de genetische kernaanpassingen die ten grondslag liggen aan de eigenschap wortelknolvorming aanwezig moeten zijn in alle wortelknolvormende plantensoorten. Daarom is het uitbreiden van het onderzoek aan wortelknolvorming naar soorten buiten de Papilionoideae zeer relevant voor het identificeren van de meest belangrijke genen voor wortelknolvorming. Daarnaast kunnen op deze manier mogelijk genetische aanpassingen worden gevonden die hebben bijgedragen aan het verfijnen van de symbiose in specifieke soorten.

In dit proefschrift, maak ik gebruik van twee alternatieve soorten om knolvorming te bestuderen; *Chamaecrista mimosoides* en *Parasponia andersonii*. *C. Mimosoides* behoort tot de Caesalpinoideae, een onderfamilie van de vlinderbloemigen. In deze onderfamilie komt knolvorming veel minder vaak voor dan in de onderfamilie Papilionoideae. Daarnaast hebben de Caesalpinoideae zich vroeg afgesplitst van de rest van de vlinderbloemigen familie, en daarom kunnen functionele studies aan deze onderfamilie inzicht bieden in de basale veranderingen die plaats hebben gevonden bij het ontstaan van de interactie tussen vlinderbloemigen en rhizobium. De tweede knolvormende soort waarmee ik heb gewerkt, *P. andersonii*, is een verre verwant van de vlinderbloemigen. Hun verwantschap gaat terug tot aan het begin van de stikstof-bindende clade. Daarom kan het bestuderen van *Parasponia* in de context van rhizobium symbiose een platform bieden om de belangrijkste genetische adaptaties te identificeren die ten grondslag liggen aan de eigenschap knolvorming.

Genetische analyses in *M. truncatula* en *L. japonicus* brachten gesynchroniseerde processen aan het licht dit leiden tot de vorming van functionele wortelknolletjes. Signaal moleculen geproduceerd door rhizobium en bestaande uit chitine oligomeren worden herkend door specifieke LysM-type receptoren van de plant. Dit leidt tot het aanzetten van een signaaltransductie cascade die zowel de vorming van wortelknolletjes alsmede de infectie door rhizobium in gang zet. Er is aangetoond dat cytokinine signalering een integraal onderdeel vormt van de door Nod factor-geïnduceerde signalering, en daarnaast essentieel is voor wortelknolvorming. Cytokinine signalering is nauw verbonden met de functionele activiteit van de transcriptie factor NIN. De perceptie van cytokinine wordt voornamelijk gereguleerd door de histidine-kinase receptor *MtCRE1* in *M. truncatula* en *LjLHK1* in *L. japonicus*. Het aantal wortelknolletjes dat gevormd wordt op mutanten waarin deze receptoren zijn uitgeschakeld is sterk gereduceerd, terwijl een dominant-actieve vorm van *MtCRE1/LjLHK1* de inductie van spontane wortelknolletjes induceert. Dit illustreert de belangrijke rol die cytokinine signalering speelt in het proces van wortelknolvorming in *M. truncatula* en *L. japonicus*. Echter, het is nog onduidelijk of cytokinine signalering een even belangrijke rol speelt in wortelknolvorming in soorten buiten de vlinderbloemigen onderfamilie Papilionoideae.

In **hoofdstuk 2**, toon ik aan dat het cytokinine-NIN regulon geconserveerd is in *C. mimosoides*. Inoculatie met rhizobium leidt tot de expressie van de *Type A RR* cytokinine-responsieve genen die ortholoog zijn aan de LCO-responsieve *MtARR4* en *MtARR9* genen uit *M. truncatula*. Daarnaast worden ook twee *NIN* genen transcriptioneel geactiveerd. Dit suggereert dat cytokinine signalering en *NIN*-expressie Nod factor-responsief zijn in *C. mimosoides*. Deze vinding is consistent met observaties in model vlinderbloemigen. Functionele studies in *C. mimosoides* met behulp van RNA interference (RNAi) laten een significante afname in knolaantallen zien wanneer zowel afzonderlijk de

expressie van *CmNIN1* en *CmNIN2* ofwel tegelijkertijd wordt verlaagd. Met name na onderdrukking van de expressie van *CmNIN2* is in knolletjes de afwezigheid van een fixatie zone zichtbaar. Dit suggereert dat *CmNIN2* niet alleen een belangrijke rol speelt bij wortelknolvorming, maar ook bij de succesvolle intracellulaire infectie door rhizobium. Een vergelijkbare strategie is gebruikt om met behulp van RNAi de expressie van de *C. mimosoides* cytokinine receptor *CmHK4* te verlagen. Planten waarin de expressie van *CmHK4* is verlaagd maken significant minder wortelknolletjes dan controle planten. Dit suggereert dat wortelknolvorming afhankelijk is van HK4. Daarom is mijn hypothese dat MtCRE1/LjLHK1/CmHK4-afhankelijke cytokinine en *NIN*-signalering een geconserveerd mechanisme vormt in wortelknolvorming in de Papilionoideae en Caesalpinoideae vlinderbloemigen. Verdere studies aan *C. Mimosoides* zijn nodig om de onderlinge verbinding tussen cytokinine signaling en *NIN* bij deze soort te onderbouwen.

In **hoofdstuk 3**, presenteer ik een ‘mini handboek’ voor het gebruik van *P. andersonii* als experimenteel systeem. Dit hoofdstuk beschrijft uitgebreide protocollen voor analyses aan *P. andersonii* op basis van reverse genetica. Dit protocol is gebaseerd op mutagenese door middel van CRISPR/Cas9 in planten getransformeerd met *Agrobacterium tumefaciens*. Met behulp van dit protocol kunnen binnen drie maanden een transgene lijnen worden gegenereerd. Aangezien het vermeerderen van *P. andersonii* *in vitro* zeer efficiënt is, kan fenotypering al plaatvinden in de T<sub>0</sub> generatie. Op deze manier is het mogelijk om al binnen zes maanden na transformatie symbiotische fenotypes te analyseren. Daarom stel ik dat *P. andersonii* een uitstekend alternatief experimenteel systeem is om de mechanismen te bestuderen die ten grondslag liggen aan wortelknolvorming.

In **hoofdstuk 4**, laat ik de kracht van een reverse genetische aanpak in *P. andersonii* zien door vier genen met een veronderstelde rol in de symbiose te muteren; *PanHK4*, *PanEIN2*, *PanNSP1* en *PanNSP2*. Deze genen zijn achtereenvolgens betrokken bij cytokinine en ethyleen signaling, en de regulatie van strigolacton biosynthese. Daarnaast spelen deze genen een essentiële rol in wortelknolvorming in vlinderbloemigen. Mutanten waarin *Pannsp1* en *Pannsp2* niet langer functioneel zijn, laten zien dat NSP1 en NSP2 een geconserveerde rol spelen in wortelknolvorming. Ook laten een beide mutanten een verlaging zien in de expressie van de strigolacton biosynthese genen *D27* en *MAX1*. De *Panein2* mutant heeft ook een symbiotisch fenotype, al is dit fenotype beduidend anders dan beschreven in vlinderbloemigen. In plaats van een veelvoud aan wortelknolletjes zoals bekend is van vlinderbloemigen mutanten, ontwikkeld de *Panein2* mutant wortelknolletjes waarin geen fixatie zone aanwezig is. Analyse van *Panhk4* mutanten heeft geen symbiotisch fenotype laten zien. Echter, zowel deze alsmede de *Panein2* mutanten laten wel niet-symbiotische fenotypes zien die correleren aan respectievelijk een verlaging van cytokinine en ethyleen signaling. Alles bij elkaar toont

dit hoofdstuk hoe *P. andersonii* als experimenteel systeem gebruikt kan worden naast de model vlinderbloemigen voor het bestuderen van de door rhizobium-geïnduceerde symbiotische interactie.

In **hoofdstuk 5**, presenteer ik een grondige analyse van de rol van cytokinine signalering in de interactie tussen rhizobium en *P. andersonii*. Door gebruik te maken van de synthetische reporter *TCSn*, wordt aangetoond dat cytokinine signalering geïnduceerd wordt nadat rhizobium of Nod factoren zijn toegediend. De *TCSn* reporter blijft gedurende de wortelknolontwikkeling actief in cellen die geïnfecteerd worden door rhizobium. Door het analyseren van het expressie patroon van de drie cytokinine receptoren van *P. andersonii*, hebben wij gevonden dat het expressie patroon van *PanHK3* tijdens knolontwikkeling grotendeels overlapt met dat van de *TCSn* reporter. Het tegelijkertijd verlagen van de expressie van twee cytokinine receptoren (*PanHK2*, *PanHK3*) suggereert dat er een zekere mate van functionele overlap tussen de verschillende cytokinine receptoren tijdens wortelknolvorming. Daarnaast leidde het muteren van *PanHK3* tot de vorming van wortelknolletjes waarin geen fixatie draden aanwezig waren. Op basis van dit alles, is mijn bewering dat cytokinine signalering een geïntegreerd onderdeel uitmaakt van rhizobium-geïnduceerde wortelknolvorming bij *P. andersonii*.

Samengevat, dit proefschrift biedt een voorbeeld van een vergelijkende genetische analyse van het kenmerk wortelknolvorming door gebruik te maken van twee soorten buiten de vlinderbloemigen onderfamilie Papilionoideae. De resultaten in dit proefschrift illustreren dat cytokinine signalering een geïntegreerde rol speelt in de door rhizobium geïnduceerde wortelknolvorming. Dit suggereert dat de door Nod factoren geïnduceerde cytokinine signalering een geconserveerd onderdeel uitmaakt van wortelknolvorming, en dat dit is gerekruteerd in de gemeenschappelijke voorouder van de stikstofbindende clade.

## Ringkasan

Interaksi mutualisme antara tanaman dan bakteri penambat nitrogen memberikan manfaat bagi pertumbuhan dan perkembangan tanaman, khususnya pada kondisi lahan kekurangan unsur nitrogen. Simbiosis ini terjadi secara eksklusif dalam 10 famili yang tersebar dalam Ordo Fabales, Fagales, Cucurbitales dan Fagales yang bersifat parafiletik dan secara kolektif dikenal sebagai 'Klade Penambat Nitrogen'. Delapan famili membentuk interaksi simbiosis dengan bakteri berfilamen *Frankia*, yang selanjutnya diklasifikasikan sebagai tumbuhan aktinoriza. Dua famili lain; yaitu legum (Fabaceae, Fabales) dan *Parasponia* (Cannabaceae, Rosales), berinteraksi dengan bakteri penambat nitrogen rhizobium. Simbiosis antara tumbuhan dengan *Frankia* atau rhizobium terjadi dalam tingkat intraseluler yang menghasilkan pembentukan organ baru pada akar tanaman yang disebut sebagai 'bintil akar atau nodul'. Meskipun terdapat perbedaan morfologi nodul pada spesies inang, secara umum nodul berfungsi untuk menyediakan lingkungan fisiologis yang sesuai bagi bakteri untuk mengubah nitrogen atmosfer ( $N_2$ ) menjadi amonium.

Penelitian tentang simbiosis penambat nitrogen telah berlangsung selama puluhan tahun dengan tujuan utama adalah untuk menginvestigasi regulasi genetik yang diperlukan dalam simbiosis ini. Dua penelitian terbaru yang menyajikan analisis filogenomik antara spesies nodulasi dan non-nodulasi menunjukkan bahwa, kemampuan tumbuhan untuk bernodulasi diadaptasi satu kali oleh nenek moyang pada Klade Penambat Nitrogen, namun mengalami kehilangan materi genetik yang terjadi secara masif dan paralel selama divergensi spesies pada klade ini.

Terinspirasi oleh manfaat yang besar dari simbiosis pengikat nitrogen dalam memproduksi amonium yang dibutuhkan oleh tumbuhan, hal ini menimbulkan pertanyaan apakah mungkin untuk merekayasa kemampuan bersimbiosis ini ke tanaman non-nodulasi di luar Klade Penambat Nitrogen, seperti gandum, jagung, atau padi, sehingga menghasilkan varietas yang berkelanjutan. Untuk menjawab tantangan tersebut, sangat penting untuk mengidentifikasi kunci adaptasi genetik yang mendasari keberhasilan simbiosis ini. Sampai saat ini, pengetahuan tentang mekanisme nodulasi telah dihasilkan dari berbagai studi genetika molekuler pada dua spesies model; *Medicago truncatula* dan *Lotus japonicus* yang keduanya termasuk dalam sub-famili legum Papilionoideae. Sesuai dengan hipotesis 'nenek moyang tunggal' untuk asal kemampuan bersimbiosis, hal ini mengimplikasikan bahwa materi genetik inti pada tumbuhan nodulasi diatur melalui jaringan genetik yang sama. Oleh karena itu, penelitian tentang simbiosis penambat nitrogen di luar subfamili Papilionoideae sangat relevan dilakukan, sehingga

dapat mengungkap materi genetik inti dari simbiosis ini dan mengidentifikasi adaptasi genetik yang terjadi secara spesifik di setiap garis keturunan dalam Klade Penambat Nitrogen. Tujuan inilah yang merupakan landasan pemikiran dari disertasi PhD ini.

Untuk mempelajari mekanisme genetik pada nodulasi, saya menggunakan dua model spesies yaitu *Chamaecrista mimosoides* dan *Parasponia andersonii*. *C. mimosoides* termasuk dalam subfamili legum Caesalpinoideae, dimana peristiwa nodulasi lebih jarang terjadi dibandingkan dalam subfamili Papilionoideae. Caesalpinoideae juga mewakili simpul keturunan basal dalam divergensi famili legum, sehingga studi genetik pada spesies *Chamaecrista* dapat merepresentasikan adaptasi basal dalam interaksi legum-rhizobium. Spesies kedua, *P. andersonii* adalah spesies nodulasi yang memiliki hubungan kekerabatan yang jauh dengan legum. Mempelajari *Parasponia* dalam konteks simbiosis dengan rhizobium membuka peluang untuk mengidentifikasi materi genetik inti yang penting untuk mengatur nodulasi yang telah dikonfirmasi fungsinya di famili legum, apakah memberikan fungsi genetik yang sama untuk *Parasponia*.

Studi genetik pada *M. truncatula* dan *L. japonicus* menunjukkan bahwa proses infeksi rhizobium dan pembelahan sel pada organogenesis nodul harus terjadi secara sinergi untuk menghasilkan nodul yang fungsional. Simbiosis ini ditandai dengan penerimaan sinyal molekul bakteri (*Nod-factors*) oleh reseptor spesifik pada tanaman (LysM) yang mengaktifkan kaskade persinyalan untuk organogenesis nodul dan infeksi intraseluler rhizobium. Studi pada model legum juga menunjukkan bahwa persinyalan sitokinin merupakan bagian integral dari regulasi nodulasi yang diinduksi oleh *Nod-factors* ini. Persinyalan sitokinin juga terkait erat dengan aktivitas fungsional dari faktor transkripsi NIN, yang merupakan pengatur utama dalam nodulasi. Persepsi sitokinin diatur oleh reseptor sitokinin histidin kinase MtCRE1 di *M. truncatula* dan LjLHK1 di *L. japonicus*. Mutasi genetik pada reseptor ini menghasilkan pengurangan jumlah nodul secara dramatis, sedangkan mutan dominan-aktif dari MtCRE1/LjLHK1 menghasilkan pembentukan struktur nodul meskipun tanpa diinokulasi dengan rhizobium. Hal ini menggarisbawahi pentingnya pensinyalan sitokinin pada nodulasi *M. truncatula* dan *L. japonicus*. Namun demikian, apakah pensinyalan sitokinin dalam konteks symbiosis juga dipertahankan pada spesies nodulasi di luar subfamili legum Papilionoidea belum diketahui.

**Bab 2**, bertujuan untuk mengkaji apakah regulasi sitokinin-NIN juga ditunjukkan di dalam spesies *C. mimosoides*. Dalam studi ini, inokulasi *C. mimosoides* dengan rhizobium berhasil menginduksi ekspresi gen tipe *A-RR* yang merupakan ortologis terhadap gen *MtARR4* dan *MtARR9* pada *M. truncatula*, yang telah diketahui responsif terhadap *Nod-factors*. Selain itu, inokulasi rhizobium juga menginduksi ekspresi gen

*NIN1* dan *NIN2* di *C. mimosoides*. Observasi ini selain menunjukkan bahwa persinyalan sitokinin dan ekspresi *NIN* pada *C. mimosoides* responsif terhadap *Nod-factors*, juga menunjukkan konsistensi pengamatan seperti yang diketahui pada model legum. Studi fungsi genetik yang dilakukan menggunakan interferensi RNA (*RNAi*) dengan target *C. mimosoides CmNIN1*, *CmNIN2*, atau *CmNIN1-NIN2* ganda menghasilkan penurunan jumlah nodul yang signifikan. Khususnya pada galur tanaman yang mengalami peredaman gen *CmNIN2* menghasilkan nodul tanpa zona fiksasi. Hal ini menunjukkan bahwa *CmNIN2* memiliki fungsi penting tidak hanya dalam organogenesis nodul, tetapi juga untuk keberhasilan diferensiasi bakteri untuk menjadi pemfiksasi. Strategi *RNAi* serupa diterapkan untuk meredam ekspresi gen *C. Mimosoides* sitokinin reseptor histidin-kinase *CmHK4*. Galur *CmHK4 RNAi* menghasilkan jumlah nodul yang berkurang secara signifikan jika dibandingkan dengan tanaman kontrol. Hal ini menunjukkan bahwa organogenesis nodul bergantung pada HK4. Oleh karena itu, saya berhipotesis bahwa sitokinin dan pensinyalan NIN yang dimediasi MtCRE1/LjLHK1/CmHK4 adalah mekanisme yang penting dalam pembentukan bintil akar di legum Papilionoideae dan Caesalpinoideae. Meskipun demikian, studi lebih lanjut pada *C. mimosoides* diperlukan untuk mengidentifikasi secara spesifik interkoneksi antara pensinyalan sitokinin dan NIN.

**Bab 3**, menyajikan buku panduan dalam menggunakan *P. andersonii* sebagai model eksperimental. Protokol yang dipaparkan pada bab ini diperuntukkan untuk menganalisis genetik pada *P. andersonii*. Protokol tersebut didasarkan pada CRISPR/Cas9-mutagenesis untuk tanaman yang ditransformasi dengan *Agrobacterium tumefaciens*. Dengan mengaplikasikan protokol ini, jalur transgenik yang stabil dapat dihasilkan dalam kurun waktu tiga bulan. Selanjutnya, melalui perbanyakan *in vitro* yang sangat efisien pada *P. andersonii*, analisis fenotipik dari galur generasi T0 dapat dilakukan dalam waktu enam bulan setelah transformasi. Oleh karena itu, berdasarkan tingkat efisiensi protokol ini, saya berpendapat bahwa *P. andersonii* adalah sistem eksperimental yang sangat baik dan dapat digunakan sebagai alternatif spesies untuk mempelajari mekanisme nodulasi.

**Bab 4**, menyajikan bukti konsep analisis genetik pada *P. andersonii*, dengan melakukan mutasi tunggal pada gen yang berfungsi pada simbiosis; yaitu *PanHK4*, *PanEIN2*, *PanNSP1*, atau *PanNSP2*. Gen ini diketahui terlibat dalam persinyalan sitokinin dan etilen, dan regulasi biosintesis strigolakton. Selain itu, gen ini juga menunjukkan fungsi esensial dalam regulasi nodul. Kehilangan fungsi *Pannsp1* dan *Pannsp2* menghasilkan data fenotipik yang konsisten dengan observasi yang telah diketahui di model legum, yaitu mempengaruhi pembentukan nodul. Selain itu, galur mutan *Pannsp1* dan *Pannsp2* juga mempengaruhi ekspresi gen biosintesis strigolakton *D27* dan *MAX1*. Mutan *Panein2* menunjukkan fenotipe simbiosis, meskipun jelas berbeda dari apa yang

ditemukan di legum. Alih-alih menampilkan fenotipe hipernodulasi seperti yang terlihat pada mutan legum *ein2*, mutan *Pan<sub>ein2</sub>* memproduksi nodul tanpa zona fiksasi. Analisis mutan *Pan<sub>hk4</sub>* tidak menunjukkan fenotipe simbiosis. Namun, mutan ini, serta *Pan<sub>ein2</sub>* menunjukkan fenotipe non-simbiosis yang berkorelasi dengan reduksi persinyalan sitokinin dan etilen. Secara keseluruhan, penelitian ini memberikan contoh yang relevan dalam memanfaatkan *P. andersonii* sebagai sistem eksperimental untuk mempelajari interaksi simbiosis yang diinduksi oleh rhizobial, di samping legum model.

**Bab 5**, menyajikan analisis mendalam tentang peranan persinyalan sitokinin dalam interaksi *P. andersonii*-rhizobium. Dengan menggunakan reporter sintetik sitokinin TCSn, terlihat bahwa persinyalan sitokinin di *P. andersonii* teraktivasi setelah diinokulasi dengan rhizobium atau *Nod*-factors. Reporter TCSn tetap terdeteksi selama organogenesis nodul, khususnya untuk sel yang terinfeksi oleh rhizobium. Dengan menganalisis ekspresi temporal dari tiga gen *P. andersonii* reseptor HK, pola ekspresi PanHK3 sebagian besar serupa dengan pola ekspresi TCSn selama organogenesis nodul. Studi RNAi yang meredam ekspresi gen reseptor sitokinin secara bersamaan (*Pan<sub>HK2</sub>*, *Pan<sub>HK3</sub>*, *Pan<sub>HK4</sub>*) menunjukkan redundansi fungsional di antara tiga reseptor sitokinin dalam mengatur organogenesis nodul. Temuan utama pada Bab ini adalah analisis fenotipik dari galur mutan *Pan<sub>hk3</sub>* yang menghasilkan nodul tanpa zona fiksasi atau berkurangnya zona fiksasi secara signifikan. Mempertimbangkan berbagai data dan observasi yang dipaparkan dalam Bab ini, saya berpendapat bahwa pensinyalan sitokinin terintegrasi dalam nodulasi *P. andersonii* yang diinduksi oleh rhizobium.

Secara ringkas, disertasi PhD ini memberikan contoh analisis genetik komparatif terhadap kemampuan nodulasi yang memanfaatkan dua spesies di luar subfamili legum Papilionoideae. Hasil dalam tesis ini menggambarkan integrasi pensinyalan sitokinin pada nodulasi yang diinduksi oleh rhizobium. Hal ini menunjukkan persinyalan sitokinin yang diinduksi oleh *Nod*-factors merupakan mekanisme basal yang telah direkrut dalam nenek moyang dari Klade Penambat Nitrogen.





# Appendices





# Acknowledgments

## Disclaimer - Long Session

A PhD journey is never solitude. Sometimes it is messy, yet it can be rewarding at the same time. For all the ups and downs during this period of seeking knowledge, I am grateful for having an opportunity to be challenged, to overcome the obstacles, and finally to get through one of the lifetime experiences. This journey was indeed very personal, although I must say that it would not happen without any contributions and supports from many people, both for professional and personal development. Here, I would like to express my gratitude to whom I was lucky enough to have an encounter with.

**René**, our first meeting was back in 2013, where we discussed about the possibility of research project for my MSc thesis. I instantly noticed that you were very enthusiastic explaining about *Parasponia* along with its potentials for studying the rhizobium-induced symbiosis. From the meeting in 2013, it was also my very first time knowing about *Parasponia*, which is apparently grown in a place that is dear to my heart; Indonesia. Since then, I grew this unnecessary emotional bonding with *Parasponia* that led me to a wonderful time in MOLBI. In 2015, you kindly offered me a PhD position regardless of hurdles that we had to encounter at the beginning. Although this is not a thesis that can get us a Nobel-Prize (as we usually joke about), yet this thesis is something that I am proud of. To this end, I am really thankful for your patience, understanding and guidance as my co-promotor, as well as your critical comments during our countless discussions. All of those have helped me to shape my PhD research direction, to be an independent researcher, and to always think critically, yet to be optimistic in this highly competitive scientific climate. I also would like to extend my appreciation for another opportunity to be in MOLBI a 'little' longer by offering me a project in cassava, which is (again) as another species that is related very well with Indonesia.

**Ton**, I would like to express my gratitude for being my promotor, who I admire your seniority in science. You usually ask questions during the MOLBI Friday afternoon meetings, which sometimes could be very difficult to answer. However, I am really thankful for being in this working environment, as it indeed helped me to think critically about my research project.

**Arjan**, you've been part of this PhD journey since the very beginning. We argued a lot during the process, but it was rewarded when we managed to make the work published. Thank you for your all scientific contributions in my PhD projects that your name always appears in every of my experimental chapters.

**Wouter**, I would like to thank for your tremendous help in the Chapter 2 - *Chamaecrista* project and your expertise as a plant physiologist. Your critical comments in my *Parasponia* project were also very valuable that further helped me to formulate the Chapter 5 even better. I am thankful for our current collaborations, while looking forward to our future collaborations, which hopefully can translate our hard-work into some published articles.

**Henk**, I would like to appreciate all of your questions during the MOLBI Friday afternoon meetings, as well as your great suggestions during the writing time of my PhD thesis. I admire your critical thinking, although it could make my research data (or, basically any other research) become less significant at all. I really appreciate your suggestion to add another legume illustration in my PhD thesis cover, unfortunately it did not work. Above of all, your perspectives in society, knowledge and attitude in science are inspiring.

**Sidney**, not a single day in MOLBI went by without us listening to your jokes or stories. This particular social skill of yours was indeed a great treat when we had to be working in the lab for all day long doing the boring job of transferring tissue cultures or harvesting a lot of plants. Thank you for being a fun colleague. Most importantly, thank you for all of your technical help during my PhD project, especially during the process of establishing the stable transformation protocol in *Parasponia* that we published it as Chapter 3 and 4. Even though I have less people to talk about Marvel's movie in the office now, but I really wish you a very successful career wherever you are.

**Yuda**, what would I do without you around as my colleague (both in Indonesia and the Netherlands), but most importantly, as my friend? People would always think of you as a nice and sweet person, but they don't know about you at all. You actually can be very strong with your opinions. Thank you for your scientific contribution in the Chapter 3, all the countless talks, the 'curhat' sessions, your reliable cooking skills, and your patience of having a stubborn friend like me. It is sad that you were not around when I am defending my thesis, but I hope that your supports will always be around regardless of the place and time. I am looking forward to our more collaborations in the future.

**Luuk**, at the beginning, especially during our MSc thesis in MOLBI, I thought that you were very introvert person, but this impression changes completely over time. You can actually talk about a lot of things and you are very helpful. I appreciated all of your help, especially during the early time of my PhD, when we wanted to set up the protocol to detect or quantify the early LCO-responses in *Parasponia*. Although it did not end up perfectly, yet we now have something that is quite workable. I also would like to appreciate your contribution in Chapter 2, which hopefully we can prepare for the submission soon.

**Kana-san**, I enjoyed having you around as a colleague, although it was a quite short period of time. I would like to thank for the shared information during the time that you were setting up the calcium-spiking quantification, which further inspired the selection of ‘root susceptible zone’ in *Parasponia*. I wish you a great success in your scientific career as well as your personal life. I am looking forward to any opportunities for future collaborations.

The current and former MOLBI lab technicians; **Marijke, JanV, and Kerstin**, the go-to persons when we encounter troubles in the lab. **Marijke**, thank you for your hard-work in establishing the stable transformation in *Parasponia* together with Sidney and Arjan, which we managed to create many mutant lines, as well as to publish the related work. **JanV**, thank you for your knowledge and skills for such an elaborative RNA isolation protocol, which works very well in all of my samples. Additionally, thank you for sharing your musical skills with us (your drumming session of the centrifuge), which made the works in the lab become more fun! **Kerstin**, thank you for your help in the cytokinin extraction and UPLC/MS-MS analyses of my *Parasponia* root samples.

The current and former *Parasponia* team members; **Simon, Joel, Sultan, Robin, Rens, and Fengjiao**. **Simon**, I would like to thank for your contribution in Chapter 3 and for all Marvel’s related talks that we had. **Joel**, thank you for your contribution in the Chapter 2, which hopefully we can opt for submission soon. **Robin**, I would like to thank for giving me a nice impression that Dutch people can actually be generous. On the first day that I came back again to Wageningen for starting my PhD, you bought me a lunch as I did not have the Dutch bank account, yet. However, most importantly, I would like to thank for the shared data that I can use to recreate a figure in my Chapter 1, as well as your critical questions that you asked during my regular presentations. **Rens**, thank you for being a nice colleague and creating a nice working environment. **Sultan**, I wish you a great success with your experiments and the rest of your PhD. **Fengjiao**, thank you for the friendship that we shared and for being my housemate during a short period of pandemic time. I really wish you a great success in your professional and personal life.

**Rik**, I am sorry that you don’t receive the 3-pages long of acknowledgments, as some people say that less is also better ;) I enjoyed our times as the Youtube musicians although we never make money out of it, hopefully we can make more videos/songs in the future. Most importantly, thank you for your suggestions during my work discussions and thesis writing. I enjoy having you back in MOLBI/*Parasponia* team and finally witness you to work on two things that you ‘reaaaaaally like the most’; transcription factor and hormone. I am looking forward to our collaborations (i.e. our luck?) of ‘forcing the capacity of Nature’ to establish the root nodule symbiosis.

**Defeng**, thank you for trusting me with your data of DR5 lines, which I incorporated it in my General Discussion chapter. I wish you a good life in Germany and the professional career afterwards.

**Maria and Marie-José**, I would like to thank for all of your help related to the administrative works. Otherwise, my stay in the foreign country will not be very smooth. Marie-José, thank you for being attentive and giving me moral support, especially when I heard about the big flood that happened in Indonesia, which could potentially affect my family.

I am also grateful for the BSc/MSc students that I supervised or am currently supervising during my PhD: **Ilona, Lonne, Victor, Pleun, Thomas, and Kim**. Thank you for your help in one way and another. I hope that you enjoyed and learned something from the project, as much as I learned a lot from supervising each and every one of you. I wish you a great success with your future endeavours!

All staffs, former and current members of MOLBI/PDB/Cell Biology; **Erik, Joan, Renze, Viola, Tijs, Carolien, Huchen, Asma, Xu, Jing, Zhichun, Jundi, Yueyang, Lucas, Martinus, Ashenafi, Comfort, Amina, Yinshan, Tian, Mengmeng, Renan, Adam, Bandan, Jeroen, Aniek, Han, Wenkun, Guiling, Trupti, Menno, Anneke, Ikram, Vera, Merijn, Norbert, Alejandra, Henk K, Kiki, Hannie, and Ben**. The presence of each and every one of you have created both a professional and an enjoyable working environment in MOLBI. **Olga**, thank you for all your suggestions when I faced troubles with fixing samples or sectioning. I always enjoy our talks during the coffee break. **Ting Ting**, thank you for your shared experience, especially during the early time of my PhD in regard to handle cross section samples. I wish you all the best in life. **Peng**, thank you for being a nice colleague as well as the photographer for my defence (if by that time I can have my defence in Aula). I always thought that you are a quiet and shy person, but you certainly can make some jokes! I wish you all the best for the last stretch of your PhD. **Jelle**, thank you for being a nice colleague and a friend to talk about Marvel's movies (or basically, any other movies). Together with **Fang, Asma**, Simon, Yuda, and Jieyu, I enjoyed our trip to Paris for the Adam Kondorosi symposium. I wish you all the best with the rest of your PhD journey/professional career. **Renan**, I enjoyed our talks when we supervised the Gene Technology course, which made the supervising activity became more fun. I wish you a great success with your PhD!

My paranymph; **Teteh Dewi**, who would have thought that our first encounter in 2013 could turn into a friendship that I truly cherish? It feels like yesterday that we enjoyed the MSc courses, and then we were finally back in the Netherlands to pursue our PhD. Thank you for being a friend to talk to, a travel-partner, a family when we are both away

from our own families. Regardless of the defence situation, be it online or on-location, I am grateful for your moral support and your willingness to keep me accompanied in this important day. I wish you all the best for the rest of your PhD journey, as well as your personal life with Kak Zul.

My paronymph; **Jieyu**, thank you for the friendship that we share during our PhD times, for all the casual scientific discussions that we had, for being the shopping advisor and a very reliable guide when we were in Wuhan. I really enjoyed our museum visit together with the delicious noodles from a small restaurant in Wuhan. I wish you a great success with the last stretch of your PhD, as well as the future career that makes you feel content and happy.

Being thousands of kilometres away from home and family would not be bearable without the presence of the old and new friends who bring a new definition of 'the family'. **Mbak Nila** and **Mas Anto**, thank you for being my family representatives when my parents cannot visit NL due to the pandemic. Thank you for all these years of our friendship, your help and supports. I wish you all the best for your family! **Ammare**, **Tsara** and **Farah**, who are growing up so fast! Thank you for all the fun times that we shared, and I wish you all good times at school! **Mbak Tika**, thank you for being my housemate, a sister, a friend, a travel-partner, and thank you for all your help and supports. I wish you a wonderful life! **Mbak Linda**, **Mbak Atik**, **Mbak Aviv**, and **Mbak Nurmi**, thank you for the friendship, for being sisters to me, for all the countless talks that we had, for all the times of highs and lows that we get through together, I am really grateful for your presences. I wish you all the best in life. **Nuning**, my K-pop partner, the one who understands the fangirling of chasing Super Junior or BTS concerts all the way to Paris. Thank you for the friendship that we shared, all the talks and moral supports, I always admire your spirit of determination in life. I wish you all the best for your professional and personal life. **Pak Dikky-Mbak Aulia**, **Pak Fajar-Mbak Nurul**, **Ika**, **Mbak Vivi**, and **Mbak Diah**, thank you for the friendship, for your shared life experiences that I can learn from, for the supports and delicious meals that we shared, and thank you for being my families in the Netherlands. I am truly grateful for meeting each and every one of you during my PhD journey in Wageningen. **Zulhaj**, thank you for the friendship, the help and all the fun times of playing Pokémon Go! Together with **Gumi-Gendis**, I enjoyed our board games and karaoke nights, the delicious meals, and thank you for being very nice neighbours. All the best for all of you! **Mas Taufik**, thank you for the friendship, the support, and the spirit of your *mangayubago*! I wish you a success with your PhD and a joyful life with **Teh Pini**. **Bang Emil**, **Mbak Lina** and **Nathans**, thank you for the friendship and all the fun bike rides around Wageningen. I wish you all the best for your family! **Bli Indra**, the one who always teases me, yet supports me at the same time. Thank you for the friendship, all the laughs and jokes that we shared. I wish you a good life with Titi!

**Hanna**, thank you for the friendship, all the fun videos from ‘Wageningen Challenge’ that you created together with **Bening**. Thank you for your visit in Wageningen, I enjoyed our short trip to Texel, and I am very much looking forward to our future trips when the pandemic is over, either to Vaasa for Bening, or probably for a Green Day concert? ;) Regardless, I really wish you all the best. **Belinda**, I still remember your ‘cucumber signature dish’ that you brought for our potluck dinner, although what we wanted was simply the raw cucumbers ;) I thank you for the friendship, our fun trip to Lisse, the picnic and barbeque times. I wish you all the best in life. **Teh Novi-Mas Indra**, thank you for the friendship that we shared, the supports, the food, the *Sundanese* conversations, which most of the times made me lost in translation, and all the fun times of board game nights. I wish you both a life filled with happiness, joy and love. **Cristina**, thank you for being one of my nice housemates, a friend to talk to, a partner to bake the best carrot cake ever. I wish you all the best with your PhD journey! **Nadia**, even though your stay in Wageningen was only a year, it feels like that we have known each other for a very long time. Thank you for the friendship, your moral supports, our talks in almost about anything and our similar taste in music. I wish you a life that brings you to happiness, inspirations, and content. **Ce Ayu**, it is always nice to have a chance to meet you every time I was in Jakarta. I admire your strong characteristic of going for something that matters to you, regardless of challenges ahead. Thank you for the friendship and the supports. I wish you all the best in life. **WUR 2011-2014 (Ce Wardah, Aris, Ibnu, Bang Zulham, Mbak Depi, Mas Dwi, Mbak Anti, Annisa, Pakde Mahfut, Ferdi, Natasha, Bang Zein, Wempy, Mbak Aty, Mbak Berlianti)**, who made my first two years of living in the Netherlands filled with good memories! Thank you to each and every one of you for the friendship that we shared.

I am also thankful for the presence of the Indonesian PhD community in Wageningen, which reflects multidisciplinary PhD projects from many different professional backgrounds. This platform does not only serve as a place to share tips and trick of doing the PhD, but also inspires us to strengthen our networking through series of symposiums (Wageningen Indonesia Scientific Exposure), which opens the possibilities to any scientific collaborations. Here, I would like to mention: **Mbak Shinta, Mbak Uma, Mas Fahriz, Mbak Ami, Mas Sahri, Koh Sony, Alim, Mbak Eka, Pak Eko, Pak Iman, Pak Ery, Satria, Margi, Riahna, Windi, Vina, Mbak Nadya, Martha, Silvia, Lenny, Calvin, Mega, Mbak Dian, Mbak Nurul, Pak Woles, Pak Yohanes, Pak Sakti, Mbak Uun, Pak Gede**. Thank you for all the good times, inspirations, and the friendship that we shared.

Thinking about Indonesia, it would be less significant without the presence of my good friends whom I have known for years: **Mira, Prio, Husni**. I really thank you for your kindness, inspirations, warm-hearted friendships, and your endless supports that you

always want to see me happy. It was also beautiful to witness that all of us went for our passions, either to study in Germany or The Netherlands. I really wish you and your family a beautiful life. **Mamet, John, Ita, Om Chicco, Mol and Yudhi**; we have our own adventures in life, be it in different cities, or even in the different continents. Yet, you are always there. Thank you for everything! **Emmy, Fatima**, your two girls are truly inspirational! It is very easy for us to get dragged into philosophical talks that hours of it are just gone like a blink of eye. Our first meeting was rather random, yet we are able to establish the friendship that is also dear to my heart. I look forward to meeting you again and resuming our talks, or simply just having delicious meals with you. My BPPT colleagues; **Mbak Dwi, Mbak Winda, Mbak Lianna, Teh Henti**, thank you for being fun colleagues and sisters to me. I truly appreciate for all the supports and the beauty of friendships that you share for the past years. Former HLPO-UGM members: **Mbak Wulan, Mbak Wida, Bung Unggul, Bung Atmo, Linda, Bung Edwin, Uni Hesty**, I would like to thank you for the friendship, inspirations, creativities, and your spirits to always want to do something beneficial for the communities.

**Dek Elok**, my sister, the one who her beauty is going beyond the meaning of the name itself, thank you for your love and supports. Thank you for the jokes and laughter during our countless video calls, shared stories, our childhood memories, and your kind understanding that I might not always be in a close distant. You know that you are always being a part in every journey that I take.

**Papi-Mami**, thank you for your unconditional love and supports. Your faith in me have led me to be a strong and independent person. This surely has become a strong foundation when I have to get through so many difficulties in life. Your patience and understanding have allowed me to achieve so many dreams in life. Thank you for teaching Dek Elok and me the importance of education, even though it means that we have to be very far away from each other. Too bad that we are currently experiencing the global pandemic, which makes the overseas trips become very much restricted. Although it is almost impossible to could have possibly explained how thankful I am to have both of you as my parents, I hope that this thesis could speak a little part of it. Terima kasih!

Lastly, I would like to dedicate a special gratitude to **Dr. F** and **Dr. K**, who brought me a fancy rice cooker all the way from Roermond to Wageningen. Yes, as an Asian person, I indeed think about a rice cooker very seriously.

Wageningen, The Netherlands  
During the second wave of COVID-19 pandemic  
**Titis A.K. Wardhani**

## About the Author

*From the perspective of Ika Nurziah*



Titis Anggraeny Kusuma Wardhani was born in Pacitan, East Java, Indonesia, on 13<sup>th</sup> of April 1986. Both of her parents are of Javanese descent, but she grew up and lived in Bima, West Nusa Tenggara until she completed her high school. She then moved to Yogyakarta to study Chemistry at Gadjah Mada University, Indonesia. She graduated *cum laude* with BSc in Chemistry in 2008. She later worked as a researcher at Agency of the Assessment and Application of Technology (BPPT). She was assigned with agriculture related projects during her work at BPPT. Although the projects that were assigned to her, to some extent, not related with her degree in chemistry, she performed and delivered the projects well. This was also when she began to train herself as a plant scientist and grew an interest in molecular biology.

In 2012, Titis was awarded with NFP (Netherlands Fellowship Programme) Scholarship to study Master of Science in Molecular Life Science at Wageningen University & Research (WUR). During her master study, she did two master theses at Laboratory of Biochemistry and Molecular Biology. Her master thesis in Molecular Biology would later be a foundation of her PhD research. Aside from academic activities, Titis had been an active member of Indonesian Student Association in Wageningen (PPI Wageningen). There, she was assigned as a public relation staff. Along with other Indonesian students at WUR, Titis helped to promote cultural exchange between Indonesia and Netherlands, as well as shared their insight for Indonesian students who would like to pursue their study in the Netherlands. Titis was completed her master study in 2014 and came back to Indonesia to resume her work at BPPT until 2015.

A chance to come back to her 'second home' of The Netherlands came when she was offered a PhD position at Laboratory of Molecular Biology-WUR, in 2015. There, she follows up her research on *Parasponia* species, to investigate the influence of cytokinin signalling in the *Parasponia*-rhizobium interactions, which the results are presented in this thesis. As she describes it to her friends, she considers her PhD as a resilient journey, where she grows academically, professionally, and also where she finds a space to self-develop. Throughout her PhD, she also has a chance to lead a

bilateral symposium (Wageningen-Indonesia Scientific Exposure 2017), a platform to bring the research expertise from WUR and Ministries/Research Institutes/Universities from Indonesia. Additionally, her PhD project allows her to collaborate and work closely with leading experts of nitrogen-fixing symbioses in the world.

Titis maintains work-life balance throughout her PhD year. She loves traveling, baking and cooking. On the weekend, she can be found with her guitar jamming or covering some music, even writing her own poetry. Titis also does painting and sketching, which she proudly hangs on her house for the guests to enjoy. Titis is due to complete her PhD at Laboratory of Molecular Biology, WUR, later this year. She is a thorough and hard-working person, but many of her friends would agree that it is her relentless nature toward accomplishing her dreams that makes her who she is right now.

Written by:

**Ika Nurziah**

Former “*Javastraat*” Housemate

MSc in Plant Biotechnology, WUR

PhD Candidate, International Agriculture DTP,

Rothamsted Research-University of Nottingham, UK

## List of Publications

- Wardhani, T. A. K.**, Roswanjaya, Y. P., Dupin, S., Li, H., Linders, S., Hartog, M., Geurts, R., van Zeijl, A. (2019). Transforming, Genome Editing and Phenotyping the Nitrogen-fixing Tropical Cannabaceae Tree *Parasponia andersonii*. *J. Vis. Exp.* (150), e59971, doi:10.3791/59971.
- van Zeijl, A., **Wardhani, T. A. K.**, Seifi Kalhor, M., Rutten, L., Bu, F., Hartog, M., Linders, S., Fedorova, E. E., Bisseling, T., Kohlen, W., et al. (2018). CRISPR/Cas9-Mediated Mutagenesis of Four Putative Symbiosis Genes of the Tropical Tree *Parasponia andersonii* Reveals Novel Phenotypes. *Front. Plant Sci.* 9:1–14.
- van Velzen, R., Holmer, R., Bu, F., Rutten, L., van Zeijl, A., Liu, W., Santuari, L., Cao, Q., Sharma, T., Shen, D., Roswanjaya, Y.P., **Wardhani, T.A.K.**, Kalhor, M.S., Jansen, J., van den Hoogen, D.J., Güngör, B., Hartog, M., Hontelez, J., Verver, J., Wei-Cai Yang, Schijlen, E.G.W.M., Repin, R., Schilthuizen, M., Schranz, M.E., Heidstra, R., Miyata, K., Fedorova, E., Kohlen, W., Bisseling, T., Smit, S., Geurts, R. (2018). Comparative genomics of the nonlegume *Parasponia* reveals insights into evolution of nitrogen-fixing rhizobium symbioses. *Proc. Natl. Acad. Sci. U. S. A.* 115:E4700–E4709.
- Wardhani, T.A.K.**, Rutten, L., van Zeijl, A., Klein, J., Geurts, R., Kohlen, W. (2020). Dissecting the Role of Cytokinin Signalling and NIN during nodulation in *Chamaecrista mimosoides*. *To be submitted*.
- Wardhani, T.A.K.**, van Zeijl, A., Linders, S., Gühl, K., Bisseling, T., Kohlen, W., Geurts, R. (2020). The *Parasponia andersonii* Cytokinin Receptor HISTIDINE KINASE3 is Required for Intracellular Rhizobium Infection of Nodule Cells. *To be submitted*.

## About the Thesis Cover

Thesis cover shows representative illustrations of two plant species used in this PhD thesis. For aesthetic purpose, the illustrations were arranged in a such way and did not necessarily reflect the real dimension size.

On the top part of the thesis cover; stretched from the front, the spine, and to the back is *Parasponia andersonii* berries and leaves. At the bottom part, stretched from the front, the spine, and to the back is the illustration of the nodulated roots of *P. andersonii*.

On the right bottom corner of the thesis cover, it depicts the illustration of *Chamaecrista mimosoides* flowers and leaves.

Distributed randomly as background molecules with low opacity are cytokinin molecules (isopentenyladenine, *trans*-zeatin, and *cis*-zeatin). These cytokinin molecules are a class of plant hormones, which are frequently detected in plant tissues.

# Colophon

The research described in this thesis was financially supported by the Dutch Science Organization (NWO) – VICI (865.13.001), the European Research Council (ERC-2011-AdG294790), and ENSA Project funded by Bill & Melinda Gates Foundation to the University of Cambridge to the Wageningen University (René Geurts).

Financial support from Wageningen University for printing this thesis is gratefully acknowledged.

**Cover design:** Titis Anggraeny Kusuma Wardhani

**Layout:** Titis Anggraeny Kusuma Wardhani

**Printed by:** Digiforce - Proefschriftmaken, Ede, NL

