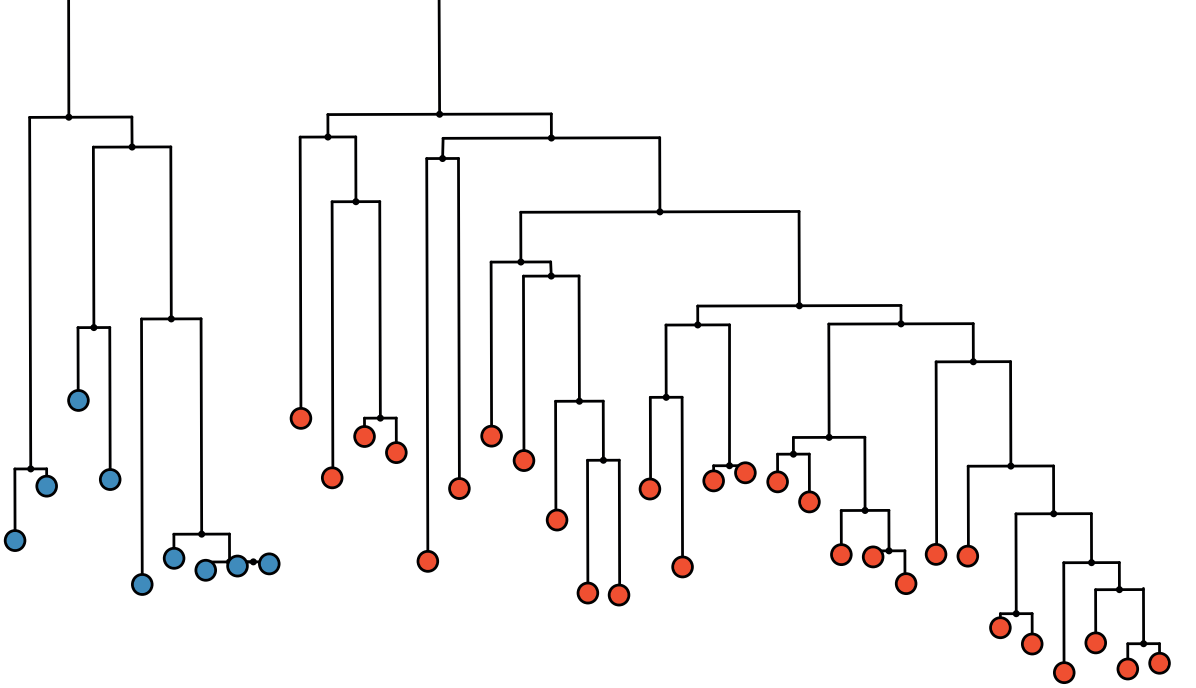
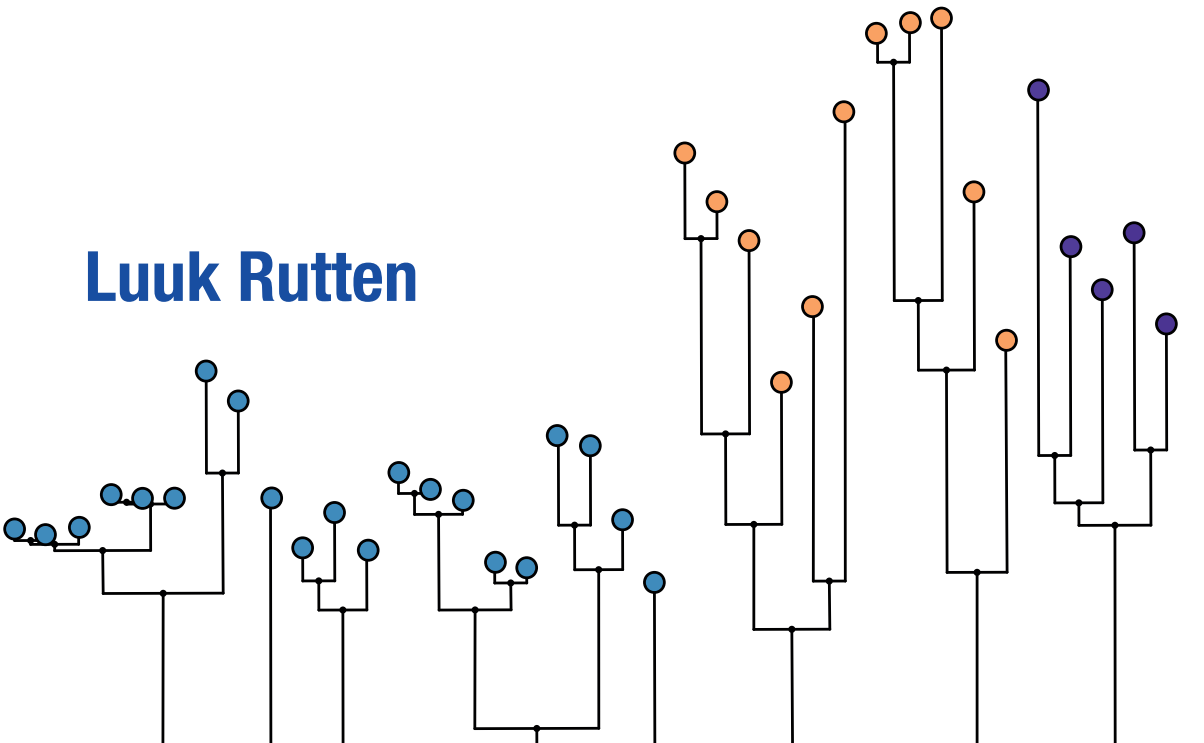


The Evolution of Nitrogen-fixing root Nodules

Analysis of conserved signalling modules in legumes and Parasponia

Luuk Rutten



Propositions

1. Multi-functionality, rather than high stringency, represents the ancestral function of the LYK-I clade LysM-type receptor kinases (This thesis).
2. The duplication of the NOD FACTOR PERCEPTION (NFP) gene was a driver for the evolution of nodule symbiosis (This thesis).
3. A priori equal weights for gains and losses in evolutionary rate models is a flawed assumption (Werner et al. 2014).
4. Engineered Honeybee symbionts, which combat the varroa destructor mite, cannot solve the global pollinator decline (Leonard et al. 2020).
5. Improvisation is to be preferred over detailed planning.
6. The ignorance for lab safety increases with the number of safety rules in place.

Propositions belonging to the thesis, entitled:

The evolution of Nitrogen-fixing root Nodules; analysis of conserved signalling modules in legumes and *Parasponia*

Luuk Rutten, 21st of October 2020

Leonard, Sean P., J. Elijah Powell, Jiri Perutka, Peng Geng, Luke C. Heckmann, Richard D. Horak, Bryan W. Davies, Andrew D. Ellington, Jeffrey E. Barrick, and Nancy A. Moran. 2020. "Engineered Symbionts Activate Honey Bee Immunity and Limit Pathogens." Science 367 (6477): 573–76.

Werner, Gijsbert D. A., William K. Cornwell, Janet I. Sprent, Jens Kattge, and E. Toby Kiers. 2014. "A Single Evolutionary Innovation Drives the Deep Evolution of Symbiotic N2-Fixation in Angiosperms." *Nature Communications* 5 (January): 4087.

Evolution of Nitrogen-fixing root Nodules

Analysis of conserved signalling modules in legumes and *Parasponia*

Lukas Johannes Joseph Rutten

Thesis committee

Promotor

Prof. Dr Ton A.H.J Bisseling

Professor of Molecular Biology

Wageningen University & Research

Co-promotor

Dr Rene Geurts

Associate professor, Laboratory of Molecular Biology

Wageningen University & Research

Committee members

Prof. Dr Toby Kiers, VU Amsterdam, Department of Ecological Science.

Prof. Dr Dolf. Weijers, Wageningen University & Research, Laboratory of Biochemistry

Prof. Dr ir. Bart. P.H.J Thomma, Wageningen University & Research, Laboratory of Phytopathology

Dr. Kasper Røjkjær Andersen, Aarhus University, Department of Molecular Biology and Genetics

This research was conducted under the auspices of the Graduate School

Experimental Plant Sciences.

DOI: https://doi.org/10.18174/517889

ISBN: 978-94-6395-350-4

Evolution of Nitrogen-fixing root Nodules

Analysis of conserved signalling modules in legumes and *Parasponia*

Lukas Johannes Joseph Rutten

Thesis

Submitted in the 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 21 October 2020. 16:00

Authored by: Lukas Johannes Joseph Rutten.

Phd Thesis with references, with summary in English.

Entitled: Evolution of Nitrogen-fixing root nodules; Analysis of conserved signalling modules in Legumes and Parasponia

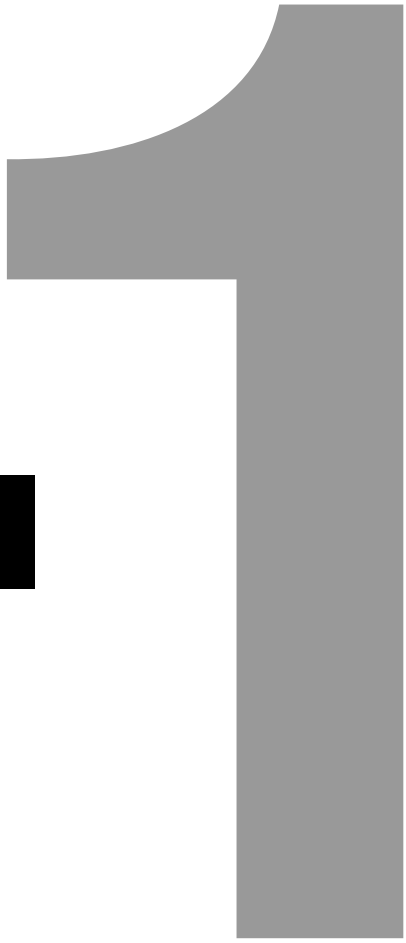
Wageningen University, Wageningen, The Netherlands (2020)

DOI: 10.18174/517889

ISBN: 978-94-6395-350-4

Table of contents

CHAPTER 1



General Introduction

Plant-microbe symbiosis as nitrogen acquisition strategy

One of the largest challenges of a plant in a non-aqueous environment is the acquisition of key nutrients to support growth. The first colonization of land by vascular plants is correlated with the origin of plant resource acquisition structures such as roots (Raven and Edwards 2001). A major strategy for resource acquisition strategies is the formation of symbiotic interactions with bacteria or fungi, which have originated earlier than the formation of plants roots. The earliest forms of plant life on land faced a continuous struggle to find symbionts for nutrient acquisition (Selosse and Le Tacon 1998; Brundrett 2002; Wang et al. 2010; Yue et al. 2012).

Symbiotic associations occur in different levels of intimacy, from loosely attached bacteria or fungi at the plant surface to endosymbiosis inside plant cells. One of the hallmarks of plant-microbe symbioses is the nitrogen-fixing nodule endosymbiosis, mostly known because of economically important legume species such as peas and beans. In this interaction bacteria are housed intracellularly in so-called nodules; specialized organs formed on the plant root or stems. Inside the nodules plants provide the optimal conditions for the bacteria to convert atmospheric nitrogen into ammonia, which they provide to the plants in exchange for photosynthates. This interaction is however limited to a relatively small number of plant species. Only about 2.5% of the angiosperm families is able to form a nitrogen-fixing nodule endosymbiosis.

The overapplication of chemical fertilizer in agriculture leads to major environmental problems in terrestrial and aquatic ecosystems. For example nitrogen deposition causes a loss of biodiversity in natural habitats by competitive exclusion of characteristic species by more nitrophilic plants (Bobbink, Hornung, and Roelofs 1998; Choudhury and Kennedy 2005). While leaching of nitrogen into aquatic ecosystems may cause algal blooms, which can reach toxic levels for existing plant and animal life (Camargo and Alonso 2006). Therefore scientist have considered it a major objective to engineer a nitrogen fixing endosymbiosis in major crop plants (Myriam Charpentier and Oldroyd 2010; Mus et al. 2016). However in order to engineer a crop plant one must first answer the question how this intricate nitrogen-fixing nodule endosymbiosis evolved, to know the adaptations required to engineer into crops such as rice or maize.

A nitrogen-fixing endosymbiosis can occur with three different types of bacteria: i. Filamentous Actinobacteria form the genus *Frankia*, nodulating a paraphyletic assembly of 25 genera distributed of 8 taxonomic families. ii. Rhizobia a paraphyletic group of - α, β and γ-Proteobacteria, nodulating Legumes (Fabaceae) and *Parasponia* (Cannabaceae) and iii. *Nostoc spp.*, a cyanobacterium infecting plants of the genus *Gunnera (Gunneraceae)*. Common to these three types of symbiosis is that bacterial

entry is preceded by host cell divisions and that once inside the cell the bacteria are enveloped by a host derived membrane. The *Gunnera-Nostoc* symbiosis is relatively unstudied and occurs only in the Gunnera genus the single member of the Gunnerales order. Cyanobacterial associations such as with Cyanobacteria *Anabaena* and *Nostoc* are usually known to occur extracellularly in like leaf cavities of the fern *Azolla* and coralloid roots of different Cycad species. The signalling cues on how the intracellular accommodation in Gunnera is achieved is unknown.

Rhizobium and/or *Frankia* nodulation occurs in several taxonomic lineages and has received more attention than cyanobacterial associations. These nodulating lineages are relatively closely related, with a distribution over four taxonomic orders; Fabales, Fagales, Cucurbitales and the Rosales. These orders are known as the Nitrogen-Fixing Clade (NFC) (Soltis et al. 1995). However, outside of the legume dominated Fabales order, nodulation is relatively rare. Thus even within the NFC, nodulation seems to be the exception rather than the rule.

The nitrogen fixing endosymbiosis in this clade of plants have fascinated researchers for decades. Due to their economic importance, most studies towards nodulation have been conducted on rhizobium-legume symbiosis. The development of two model legume systems *Medicago truncatula* (Medicago) and *Lotus japonicus* (Lotus), chosen for their small genome sizes and diploid genome, has greatly increased the speed of discoveries of genetic components required for symbiosis. Unfortunately, there is still a severe lack of knowledge on other nodulation systems, such as *Frankia*-based nodulation. In this introduction, I will summarize the essentials of these discoveries, with a focus on symbiotic signalling and signal transduction. In order to finally come to a strategy for providing insight in the evolutionary requirements for the evolution of nitrogen fixing symbiosis.

What are root Nodules and are they Novel?

The root nodules in the nitrogen fixing endosymbiosis serve four essential functions. (i) Nodules contain a population of cells that are permissive for intracellularly infection by the symbiont. (ii) Selective access to the nodule interior allows hosting of a clonal population of the symbiotic partner of choice. (iii) Nodules are optimized for nutrient exchange between host and microsymbiont. (iv) Nodules provide the optimal physiological conditions for nitrogen fixation to take place.

In all nitrogen-fixing symbioses, including *Gunnera-Nostoc*, cells that are permissive to infections undergo cell divisions initiated by host-microbial signal exchange. This means that host infection relies on the induction of the cell cycle for infection and nodule formation. This is a common feature of all endosymbiosis including that of the interaction with Arbuscular Mycorrhizal fungi (Bainard et al. 2011). In the upcoming sections I will focus on the signal exchange between host and symbiont.

A single signalling pathway for endosymbiosis.

Knowledge on host-symbiont exchange signals has mainly come from legume-rhizobium interactions. Recently actinorhizal plant - *Frankia* research has become more amendable. One of the major findings was that legumes, Parasponia and actinorhizal plants use a conserved set of genes to establish contact with their symbionts. The second major finding is that this conservation results from the recruitment of genes that function an older type of endosymbiosis; The Arbuscular Mycorrhizal symbiosis(AM-symbiosis). The AM-symbiosis is a symbiosis between fungi of the glomeromycota lineage and plant roots. This symbiosis is shared by ~72% of land plants and generally considered to be 450 million years old(Brundrett and Tedersoo 2018). The AM fungi facilitate the uptake of micronutrients, phosphate and fixed-nitrogen from the soil, which are exchanged for photosynthates in symbiotic structures called arbuscules in the root cortex. The overlap between the genetic networks controlling nodule and endomycorrhizal symbioses reveals a single pathway used by plants to establish endosymbiosis with either fungi or bacteria. This genetic network was therefore called the common symbiosis signalling pathway (CSSP) (Oldroyd and Downie 2006).

The CSSP is a major signalling cascade that spans from membrane localized receptor proteins to downstream transcription factor hubs. In the case of rhizobium-legume symbiosis, this CSSP is activated upon rhizospheric signalling. Rhizobia, in response to root exudates containing flavonoids, respond by the synthesis of lipo-chitooligosaccharides (LCOs, also known as Nodulaton (Nod) factors). These molecules consists of a β-1,4-n-acetylglucosamine (chitin) backbone, linked to a modified fatty acid chain. Several decorations on the chitin backbone, which may serve as host range determinants, can be present. These components are structurally similar to the more simple LCO molecules produced by AM fungi. Perception of rhizobium LCO signal molecules by the plants is provided by a heteromeric complex of multiple LysM-type receptor kinases and the LRR-type SYmbiotic Receptor Kinase (LjSYMRK/mtDMI2) (Stracke et al. 2002; Limpens et al. 2005). A downstream signalling cascade, which involves an enzyme in the mevalonate biosynthesis MtHMGR1 (Venkateshwaran et al. 2015), induces nuclear calcium oscillations. Calcium oscillations are common to all types of endosymbioses (Ehrhardt, Wais, and Long 1996; Navazio et al. 2007; Granqvist et al. 2015; Myriam Charpentier et al. 2016). To achieve calcium oscillations a network of proteins in the nuclear envelope are required, among which are LjCASTOR/MtDMI1, LjPOLLUX, MtCNGC15a-c that represent calcium channels (M. Charpentier, Sun, and Martins 2016; Kim et al. 2019). The calcium signal is decoded by a calcium calmodulin-dependent kinase (CCaMK), which interacts with and subsequently phosphorylates

the transcription factor LjCYCLOPS/MtIPD3 (Lévy et al. 2004; Yano et al. 2008; Horváth et al. 2011). The cascade up to this point is conserved between legume-rhizobium and arbuscular mycorrhizal signalling.

Symbiosis signalling by LysM receptor kinases

LysM-type receptor kinases are plasma membrane localized receptors, with an intracellular serine/threonine kinase and an extracellular domain with three Lysin Motif regions (Willmann and Nürnberger 2012). The Lysin motif is a ubiquitous molecular structure found in almost all living organisms except archaea. It is 42-48 amino acids long and has a symmetrical β-αα-β structure. LysM motifs were first described in bacteria where the proteins are used for peptidoglycan (PGN) binding, but they can be involved in the perception of various molecules with sugar chain backbones (Zhang, Cannon, and Stacey 2009). Legume LCO LysM-type receptor kinases have evolved upon a series of gene duplications giving rise to a large family. The Nod factor receptor complex is made up of at least two structurally different LysM-type receptors (Radutoiu et al. 2003; Moling et al. 2014). A LYR-type receptor, which kinases show no autophosphorylation activity, caused by a lack of several crucial amino acid in the kinase domain. And a LYK-type receptor, which has an active kinase. Both these types of LysM-type receptors are common to all plants and represent a large gene family (Buendia et al. 2018). Important LYK-type receptors, such as *Arabidopsis thaliana* AtCERK1, are involved in transducing chitin (CO) responses in defense. Heteromeric complex formation is also common for defense since AtCERK1 forms a complex with AtLYK4, which represents a LYR-type receptor (D.-X. Xue et al. 2019; Faulkner et al. 2013). The LYK-type receptor family is expanded in legumes, which has created a large number of LYK receptors which can function in recognition of LCO/CO or other yet unknown molecules. This legume expansion has allowed functional differentiation of receptor function (De Mita et al. 2014). The LYK-type receptor LjNFR1/MtLYK3 in legumes is required for signal transduction to activate the CSSP in rhizobium signalling (Radutoiu et al. 2003; Smit et al. 2007), while recent work indicates the chitin signalling is transduced by a separate LYK-type receptor LjCERK6/MtLYK9 (Bozsoki et al. 2017). It is generally believed that such gene duplications were instrumental in the evolution of LCO recognition. Since both CO and LCO signalling are involved in the establishment of mycorrhizal signalling, both these LYK-type receptors of legumes have roles in the establishment of AM-symbiosis (Leppyanen et al. 2017; Feng et al. 2019; Gibelin-Viala et al. 2019). The LYR-type LCO receptor in legumes is LjNFR5/MtNFP (Radutoiu et al. 2003; Limpens et al. 2003). LjNFR5/MtNFP is known to be important for LCO-recognition by direct binding on the second LysM domain (Broghammer et al. 2012; Gough and Jacquet 2013). These receptors are also major determinants for host specificity, which is in part encoded in the Nod factor structure (Dénarié,

Debellé, and Promé 1996; Bek et al. 2010). Yet also the NFP-clade is present in most non-nodulating plant species, resulting in speculation that another duplication, giving rise to MtNFP/MtLYR1 and LjNFR5/LjLYS11 was important for the evolution of stringent LCO perception (Arrighi et al. 2006; Young et al. 2011; Gough et al. 2018). Besides this receptor pair important for LCO binding numerous others LysM-type receptors have been implicated in a role for symbiosis. Such as Exopolysacharide receptor LjEPR/MtLYK10 and LCO binding LysM-type receptor MtLYR3 (Fliegmann et al. 2013, 2016; Kawaharada et al. 2017). Although given the major roles of LysM-type receptors as signalling receptors of the rhizosphere this may not be surprising (Zgadzaj et al. 2016).

Studies in the non-legume rhizobium nodulator *Parasponia andersonii* (Parasponia, prefix *Pan*) and the two actinorhizal plant species, *Casuarina glauca* (Casuarina, prefix *Cg*) and *Alnus glutinosa,* (Alnus, prefix *Ag*) revealed that symbiotic calcium oscillations also underlie these symbiosis. Further genetic studies in *Datisca glomerata* (Datisca, prefix *Dg*), Parasponia and Casuarina revealed symbiotic conservation of PanNFP1, CgSYMRK/DgSYMRK, CgCCaMK/PanCCaMK and CgNIN (Markmann, Giczey, and Parniske 2008; Gherbi et al. 2008; R. Op den Camp et al. 2011; Svistoonoff et al. 2013; Clavijo et al. 2015). The CSSP is known to be activated by other molecules than LCOs in the case of Casuarina and Alnus, although calcium oscillation remains a conserved feature (Clavijo et al. 2015; Chabaud et al. 2016). A common origin of LCO based nodulation is also highlighted by the discovery of LCO biosynthesis genes in *Frankia* strains of Cluster-II, which are expressed in symbiotic context (Van Nguyen et al. 2016; Ktari et al. 2017; Salgado et al. 2018; Persson et al. 2015). Cluster-II strains are hard to culture which suggests they represent obligate symbionts (Gtari et al. 2015). This makes research towards *Frankia* actinorhizal genes of strains in cluster-II difficult. **In Chapter 2,** I discuss the commonalities of these symbiotic pathways used by Arbuscular Mycorrhiza, rhizobium nodulation and frankia nodulation in more detail, with more focus on the establishment of symbiont recognition, specificity and the origin of their respective LCO/CO signal molecules. **In Chapter 4,** I focus on the role of the LysM-type receptor gene family of Parasponia. I identify PanLYK3 a LYK-type receptor and the homolog of legume LCO/CO receptors. I show that PanLYK3 has a dual function in symbiosis signalling and immunity. Indicating that specificity may not underlie the evolution of rhizobium symbiosis. Further I investigate the roles of PanNFP1 and the newly identified PanNFP2 **(Chapter 3.)** in LCO-signalling. Previous work on the LysM-type receptor kinases family of Parasponia uncovered a dual role for PanNFP1 (originally named PaNFP), in rhizobium nodulation and mycorrhization of Parasponia (R. Op den Camp et al. 2011). I show that PanNFP1 or PanNFP2 seem to have no, or only an additive role in mycorrhization of Parasponia.

The induction of transcription factors by the CSSP

Upon legume LCO perception the transcription factor CYCLOPS is activated and induces expression of the transcription factor encoding genes *NODULE INCEPTION* (*NIN*) and the *ETHYLENE RESPONSIVE FACTOR REQUIRED FOR NODULATION 1* (*ERN1*) and *ERN2* (Yano et al. 2008; L. Schauser et al. 1999; Laloum et al. 2014; Marion R. Cerri et al. 2016, 2017). These transcription factors are responsible for a large part of the transcriptional reprogramming required for rhizobium infection and nodule formation. Though these genes are not transcriptionally activated during AM signalling. Arbuscular mycorrhization triggers expression of different transcription factor encoding genes, such as GRAS-type transcription factor *REQUIRED FOR ARBUSCULAR MYCORRHIZA 1* (RAM1) (Gobbato et al. 2013; Pimprikar et al. 2016). Though several common transcription factor components exist as well; eg. the GRAS-type transcription factors *NODULATION SIGNALLING PATHWAY 1* (NSP1) and NSP2 that interact with RAM1(Hohnjec et al. 2015; L. Xue et al. 2015). NSP1 and NSP2 form a dimeric complex and are required for the induction of *NIN* and *ERN1*, during rhizobium symbiosis (M. R. Cerri et al. 2012; Hirsch et al. 2009). The role of *NSP1* and *NSP2* in nodulation is conserved in Parasponia (van Zeijl et al. 2018). These GRAS-type transcription factors are also required for the production of strigolactones (W. Liu et al. 2011). Strigolactones are important signalling hormones in arbuscular mycorrhization development and nodulation (Foo and Davies 2011; De Cuyper et al. 2014; Gutjahr 2014), since *MtDWARF27*, as strigolactone biosynthesis gene, was shown to be rapidly induced upon rhizobium LCO signalling (van Zeijl et al. 2015).

The essential role of the CSSP in nodulation is highlighted by the spontaneously formed pseudonodules upon artificial activation of this pathway in the absence of symbionts. This can be achieved by expression of the autoactive forms of *CCaMK* or *CYCLOPS*, both in legumes and non legumes (Gleason et al. 2006; Yano et al. 2008; R. Op den Camp et al. 2011; Svistoonoff et al. 2013). Or by the over-expression of *NIN* or *SYMRK* (Marsh et al. 2007; Ried, Antolín-Llovera, and Parniske 2014). At the same time, mutants in *ccamk*, *cyclops*, *symrk*or *nsp1* can be complemented by homologs of non nodulating species (Markmann 2008; Yokota et al. 2010; Banba et al. 2008; Saha et al. 2016). This indicates that the function of the CSSP between mycorrhizal plants and nodulating plants is conserved. It thus leaves a major question: Why nodules can not be triggered in all plant species that can establish an AM symbiosis?

Infection mechanisms and nitrogen-fixation strategies

Infection and nitrogen fixation in legumes

There are remarkable differences that take place during the infection and nitrogen-fixation processes of all these different nodule symbioses. The main differences appear to be governed by the two different types of microsymbiont gram- rhizobia bacteria and gram+ *Frankia* species. *Frankia* are filamentous Actinobacteria and generally slow growing. Rhizobia and *Frankia* both possess genes to fix nitrogen organised together with nodulation genes on symbiotic plasmids, or symbiotic islands in the genome. In the case of rhizobia, exchange of these symbiotic genes is common. Hence explaining why such a large diversity of bacteria has evolved the capacity to establish a nodule symbiosis with legumes and/or Parasponia (Bailly et al. 2007; Marchetti et al. 2014). After activation of the CSSP in legumes, infection proceeds with the entrapment of rhizobia by a root hair cell. The root hair forms a tight curl in which a rhizobia microcolony is able to penetrate the cell. The process requires activation of the cell cycle, which is shown to be activated by LCO signalling (W. C. Yang et al. 1994; Breakspear et al. 2014). A wall enclosed tubular structure called the infection thread is made. The plant continuously supplies vesicles to the tip of the growing infection thread, which depends on the formation of a host controlled infectome complex. Several important proteins have been identified which play a role in the formation of the infectosome, such as RPG, LIN and VAPYRIN (Arrighi et al. 2008; Kiss et al. 2009; Murray et al. 2011; Bapaume et al. 2019). During infection LCO signalling takes place continuously, for which membrane receptor stabilization of LYK3 in microdomains is important (Haney et al. 2011; Liang et al. 2017). Mutations in SYMBIOTIC REMORIN 1 (SYMREM1) and FLOTELLIN 4 (FLOT4), disrupts the formation of microdomain membrane rafts (Tóth et al. 2012). The infection thread continue to grow towards the root cortex. Yet uninfected root cortical cells guide the infection thread by forming cytoplasmic bridges called Pre- Infection- Threads (PITs) (Timmers, Auriac, and Truchet 1999). PIT formation is associated with slow calcium spiking and activation of the cell cycle (Sieberer et al. 2012). Since these cells do not divide, often endoreduplication occurs (Libbenga and Harkes 1973). Distal to the infection thread, a nodule primordium is initiated in the root cortex. Legume nodules can be classified in two forms; determinate and indeterminate nodules. Determinate nodules display a short lived meristem formed at the periphery of the nodule, giving rise to a spherical shape. Indeterminate nodules show a meristem at the apex, resulting in a nodule with continuous growth. Both these nodule types have a stem-like ontogeny with vascular bundles at the periphery, containing a central zone of infected cells (Katharina Pawlowski and Bisseling 1996). Inside the nodule rhizobium bacteria are released from the infection thread, however remain enveloped in host derived membrane. These droplets of bacteria are called symbiosomes. Inside the

symbiosomes rhizobium bacteria differentiate into the symbiotic nitrogen fixing state called bacteroids. The transformation of rhizobia into bacteroids is associated with an increase in size and the endoreduplication of the bacterial genome and the genome of the host cell (Penterman et al. 2014; Suzaki et al. 2014). Plant mutants that cannot undergo endoreduplication cannot be infected effectively (Yoon et al. 2014). Some legumes of the Inverted Repeat Lacking Clade (e.g. *M. truncatula*), show advanced control over bacterial differentiation. Here the bacteria undergo terminal bacteroid differentiation into enlarged rod- or even Y-shaped bacteroids governed by NCR-peptides (Haag et al. 2012; Guefrachi and Nagymihaly 2014). Rhizobia are unable to fix nitrogen outside of symbiosis, due to the oxygen dilemma of nitrogen fixation (Katharina Pawlowski 2008). The nitrogenase enzyme complex is highly sensitive to oxidation, yet nitrogen fixation requires high amounts of energy and therefore high respiration. Therefore, legumes provide oxygen protection in the form of high amounts of leghemoglobin (Ott et al. 2005). Hemoglobin allows the transport of high amounts of oxygen, yet keeps the freely diffusible oxygen concentration minimal (Ott et al. 2005).

Infection and fixation by Frankia in actinorhyzal species.

Unlike rhizobia, *Frankia* are able to fix nitrogen during saprophytic growth and in symbiosis. They have special capabilities of forming vesicles in which layers sterol lipids, called hopanoids, protect the Nitrogenase enzyme complex from harmful oxygen (A. M. Berry et al. 1993). *Frankia* occur in four taxonomic lineages called clusters, three of which represent symbiotic species (cluster-I to -III). These taxonomic clusters encode in part also a host range restriction. *Frankia c*luster-I strains mainly nodulate actinorhizal Fagales species such as *Casuarina*, *Alnus* and *Myrica* spp. *Frankia* cluster-II, of which some are known to posses LCO biosynthesis genes, mainly nodulate Cucurbitales and some Rosales species. Cucurbitales species of *Datisca* and *Coriaria* are well known to be nodulated by Cluster-II. Cluster-II species may also nodulate species of the *Rosaceae* and *Rhamnaceae* in Rosales like *Ceanothus*, *Discaria, Dryas and Purshia*. Cluster-III *Frankia* species nodulate Elaeagnaceae, Rhamnaceae in the Rosales and *Gymnostoma* and *Myrica* in the Fagales (Katharina Pawlowski and Demchenko 2012; Svistoonoff, Hocher, and Gherbi 2014). This does not mean that individual plants may not have more stringent host range determinants.

Besides these major differences in symbionts, between the actinorhizal plant species there are also large differences in infection methods. Actinorhizal plants in the Fagales allow *Frankia* to infect by a root hair curling-based mechanism similar to legumes. This means that upon signal exchange a root hair curl is made by which a transcellular infection thread grows (Svistoonoff et al. 2003). In the case of actinorhizal plants in the Fagales order, the *Frankia* hyphae grows towards a developing pre-nodule, in

which they infect dividing cortical cells. The nodule vascular bundle emerges from the pericycle giving rise to a nodule with a central vascular bundle. In nodulating Rosales such as *Elaeagnus* and *Discaria*, transcellular infection threads do not exist. Here, infection proceeds intercellularly between epidermal and cortex cells (Valverde and Wall 1999; Miller and Baker 1985). Detailed studies in *Discaria* suggest this is a plant dependent process (Valverde and Wall 1999). The plant secretes a dense matrix into the extracellular space, which is associated with repositioning of the host nucleus and endoplasmic reticulum (Imanishi et al. 2018). Similar mechanisms may exist in Cucurbitales although the early infection process is not well understood in these species. However, both in the nodulation of Rosales and Cucurbitales no pre-nodules are formed. Yet nodules emerge from the pericycle opposite protoxylem poles when the infection structures reach the inner cell layers. Intracellular infection then proceeds inside the nodule cortex (Katharina Pawlowski and Bisseling 1996). Inside the nodule, Frankia triggers the formation of short penetration structures in nodule cells, which are analogous to infection threads(R. H. Berg 1999). Yet due to the filamentous nature of *Frankia*, release into symbiosomes is not possible. Nevertheless, like rhizobia, *Frankia* filaments proliferate throughout the host cell during this stage of differentiation. Infecting *Frankia* remain enveloped in host cell membrane and cell wall, resembling fixation threads (Katharina Pawlowski and Demchenko 2012).

In general all actinorhizal plants produce a nodule with a central vasculature and multiple lobes of infection, with the exception of Datisca, which forms single lobed nodules (R. Howard Berg, Langenstein, and Silvester 1999; Katharina Pawlowski and Demchenko 2012). Another exception of Datisca is that upon infection the host cell becomes multinucleate, hinting at a role for cell cycle activation in *Frankia* intracellular infection as well, similar as found for legumes (R. Howard Berg, Langenstein, and Silvester 1999). An Actinorhizal nodule may resemble in some way a lateral root, yet represent a different structure. Since no root cap nor epidermis is formed. Inside the nodule cells *Frankia* proceeds to make vesicles, its nitrogen-fixing structure. The orientation of the vesicles may differ between the symbionts. In Alnus and most Fagales they tend to face the outside of the cell (Sasakura et al. 2006), while in Datisca and Coriaria vesicles are elongated and organised around a large central vacuole (R. Howard Berg, Langenstein, and Silvester 1999). At the base of the elongated vesicles numerous mitochondria are arrayed. This structural organization is believed to alleviate oxygen stress (Alison M. Berry et al. 2011). In most nodulating Rosales species, nitrogen-fixing vesicles are uniformly distributed across the cell. A notable exception in respect of vesicles formation occurs in the nodules of Casuarina (Fagales), in which no vesicles are formed. Here, *Frankia* relies on a host protection mechanism, which is the expression of a class 2 hemoglobin, similarly as occurs is

in legumes in the form of nodule-specific leghemoglobin expression (Jacobsen-Lyon et al. 1995). The expression of hemoglobin, or truncated hemoglobins in nodules of some *Alnus, Myrica* and *Datisca* seems to be related to the scavenging of nitric oxide (Sasakura et al. 2006; Anne B. Heckmann et al. 2006; K. Pawlowski et al. 2007).

Infection and fixation in Parasponia

The five tropical tree species of the genus *Parasponia* in the Cannabaceae family seem to represent the only lineage outside of the legumes to form nodules with rhizobium. In structure Parasponia nodules represent an actinorhizal nodule. Parasponia is entered by a form of crack entry, in which the bacteria enter the root system by the extracellular spaces. Parasponia induces epidermal and outer cortical cell divisions to provide space for the microsymbiont to enter the root (Lancelle and Torrey 1984). The crude organisation of these divisions suggest that little evolutionary time elapsed to improve the infection mechanism. Parasponia therefore also has remarkably little capacity to exclude inefficient rhizobial microsymbionts (M. J. Trinick and Hadobas 1989). The nodule structure itself looks like a Rosales actinorhizal nodule as seen in for example Ceanothus or Discaria (Valverde and Wall 1999; Q. Liu and Berry 1991). Parasponia nodules have a central vascular bundle, indeterminate meristem and the capacity to form branching nodules (Price, Mohapatra, and Gresshoff 1984; M. J. Trinick 1979). Inside the Parasponia nodule, uptake of bacteria proceeds directly by the formation of infection threads (M. J. Trinick 1979; Lancelle and Torrey 1984). Parasponia hosts its rhizobial microsymbionts in fixation threads, an ancestral character, in which bacteria remain in thread coated with host derived cell wall (M. J. Trinick and Hadobas 1988). Infection thread symbioses are common to some clades of legumes like in the genus *Chamaecrista (Naisbitt, James, and Sprent 1992)*. Inside the fixation thread rhizobia differentiate, indicated by an enlarged size and poly-β-hydroxybutyrate accumulation (M. J. Trinick 1979; Michael J. Trinick, Goodchild, and Miller 1989). Parasponia also uses hemoglobin to help protect the rhizobial nitrogenase. However, this gene originates for a different class when compared to legumes, namely class 1, suggesting convergent evolution of this trait (Appleby, Tjepkema, and Trinick 1983; Wittenberg et al. 1986; Sturms et al. 2010).

The recruitment of Nodule Inception

Although nodule structure and organization as observed in different taxonomic lineages seems rather diverse, there is at least a partially conserved genetic network essential for nodule formation. Besides the use of the CSSP a key transcription factor NIN, seems to underlie all these forms of nodule symbiosis.

NIN represents the first symbiosis gene that has been identified by forward genetics (L. Schauser et al. 1999). Yet it is also one of the most elusive modulators of the genetic pathway controlling nodulation. *NIN* represents a major player in rhizobium symbiotic gene expression. The gene encodes a transcription factor with a conserved RWP-RK domain and is most likely recruited specifically into nodulation. *NIN* belongs to a small gene family of RWP-RK transcription factors, which is known as NIN-LIKE PROTEINs (NLPs). Orthologs of *NIN* and *NLPs* can be found among all plant species. Studies in Arabidopsis and in Lotus revealed roles for NLPs in nitrate signalling (Konishi and Yanagisawa 2013; Leif Schauser, Wieloch, and Stougaard 2005).

NODULE INCEPTION in legumes

NIN is transcriptionally activated within hours of LCO recognition, first in the epidermis, where it is required for root-hair based infection. *NIN* expression in the epidermis is sufficient for root hair curling and subsequent infection thread formation. The process of curling and infection thread formation relies on activation of the cell cycle to provide membrane to the growing infection threads (W. C. Yang et al. 1994; Breakspear et al. 2014). NIN proteins bind specificity to modified NRE-like cis regulatory elements. These elements are present in symbiosis-responsive genes, such as the CCAAT-type transcriptional regulators *LjNF-YA1, LjNF-YA2* and *LjNF-YB1*, which encode subunits of a nuclear factor Y complex (Laloum et al. 2014; Soyano et al. 2013), and *NPL*, which encodes a nodulation pectate lyase required for infection thread formation (Xie et al. 2012). *NIN* expression in the pericycle and inner cortex is however required for subsequent nodule primordium formation. This indicates that *NIN* is a central regulator in nodulation; controlling infection thread formation in the epidermis and the induction of cell divisions in the inner cortex. Since LCO-induced pericycle and cortical cell divisions can be initiated hours after signalling and the fact that LCOs are immobile molecules this suggest that *NIN* induction in the pericycle requires a second messenger. Or that NIN protein is transported to the inner cell layers (Vernié et al. 2015).

Besides its responsiveness to LCO signalling, *NIN* induction depends on cytokinin. The gain of function mutation of cytokinin receptor *LjLHK1* (*snf2*) in Lotus induces spontaneous formation of nodules, which requires *NIN* expression (Anne Birgitte Heckmann et al. 2011; Murray et al. 2007). Similarly, the exogenous application of cytokinin can induce formation of pseudonodules, this response is however dependent on *NSP1*, *NSP2* and *NIN (Anne Birgitte Heckmann et al. 2011)*. Epidermal expression of *NIN* is sufficient to induce cytokinin accumulation in the pericycle. A large amount (~73%) of LCO induced gene expression depends on *MtCRE1* (the ortholog of *LjLHK1*) cytokinin receptor (Van Zeijl et al. 2015). Furthermore, NIN binds the promoters of CLE peptide-encoding genes *LjCLE-RS1* and *LjCLE-RS2*, key

players in the autoregulation of nodulation *(Kassaw et al. 2017; Hastwell, Gresshoff, and Ferguson 2015)*. The autoregulation of nodulation pathway relies on the production of the CLE peptides in active nodules, which are transported to the shoot. CLE peptides are perceived by a CLAVATA1-like Receptor named LjHAR1/MtSUNN (Okamoto and Kawaguchi 2015; Mortier et al. 2012; Okamoto et al. 2009) and mutations in this receptor cause excessive nodulation. Perception of CLE peptides results in a shoot signal that is transported back to the roots. This signal is believed to be cytokinin (Sasaki et al. 2014). Providing additional evidence for the wiring of NIN and cytokinin in the nodulation pathway, by negative regulation of infection by cytokinin (Mortier et al. 2012). The NIN transcription factor is also recruited into the nodulation program of non-legumes, such as the actinorhizal plant *Casuarina,* in which knockdown of the *NIN* transcript attenuated nodule development (Clavijo et al. 2015). *CgNIN* is also transcriptionally activated by *Frankia* diffusible factors (Chabaud et al. 2016; Clavijo et al. 2015). Additionally, NIN was shown to be highly expressed in nodules of *Datisca* (Cucurbitales) and *Ceanothus* (Rosales) (Demina et al. 2013; Salgado et al. 2018). Taken-together, it implies some degree of functional conservation in *NIN* regulation and functioning in nodulating species

The recruitment of the NIN transcription factor is common to all nitrogen fixing plants including Parasponia **(Chapter 3)**. The discovery of a novel *nin* mutant in Lotus called daphne showed that the functions of NIN in the epidermis and in the outer cortex can be uncoupled. The daphne mutant entails a chromosomal translocation 7 kb upstream of NIN (Yoro et al. 2014). In Lotus daphne cortical expression of *NIN* is lost, resulting in hyperinfection in the epidermis. In **Chapter 6,** we identify a similar nod- mutant -FN8113- in Medicago, by screening a fast neutron bombardment population. FN8113 represents a novel Medicago *nin* mutant in which infection and nodule organogenesis are uncoupled similar to the Lotus daphne mutant. Furthermore, we identify a remote upstream cis regulatory region required for the expression of *NIN* in the pericycle, and we show that this region is essential for nodule organogenesis. This region contains putative cytokinin response elements, and is conserved in eight more legume species.

Predisposition in the evolution of nodulation

So what are nodules? In the simplest definition nodules represent a novel lateral root derived structure, driven by signalling through the CSSP. An increasing amount of evidence points to the fact that many plants possess these signalling mechanisms and transcriptional modules, for their respective roles in arbuscular mycorrhizal infection, nitrate signalling and lateral root formation. It is therefore clear that nodulation does not seem to rely on completely novel components.

Were there multiple independent origins of Nodulation?

Nodule formation itself is clearly the invention of only a very limited group of plant species. As mentioned before this group of plant is evolutionary linked by a common ancestor and the the taxonomic clade encompassing all nodulating plants is referred to as the Nitrogen Fixation Clade (NFC). So, how many times has nodulation evolved? And what has driven its evolution? Researchers have tried to answer these questions by phylogenetic approaches, simply looking at the occurrence of nodulating and non nodulating clades over the phylogeny of the NFC. The most parsimonious solution, requiring the least evolutionary events, is that nodulation evolved multiple times independently (Jeff J. Doyle 1994; Soltis et al. 1995; Swensen 1996). This would also explain the different types of symbionts; three taxonomic clusters of *Frankia* Actinobacteria and the diverse assembly of rhizobia found in nodules. Additionally it would support the structural differences of nodules. While legume nodules look like stems with vasculature in the periphery, Frankia nodules all possess a lateral root like ontogeny. So, how many times has nodulation evolved? The hypothesis changes over time with the increase in understanding of the species phylogeny. Some studies predict five gains in the legume family (Fabales) for rhizobium nodulation (Jeff J. Doyle 2011). However, since most of these gains take place in the paraphyletic sister group of the Papilionoid legumes the Mimosoid, Caesalpinoid, Cassia (MCC-clade) assembly a single origin and multiple losses for the fabales family is also plausible (Janet I. Sprent, Ardley, and James 2017; Azani et al. 2017; J. J. Doyle 2016). Although this view conflicts with the many differences in fixation structures, nodule anatomy and host ranges (J. J. Doyle 2016; Janet I. Sprent 2007). In addition to the possibility of multiple origins in the legume clade, nine origins of *Frankia* nodulation are predicted (Li et al. 2015). These independent origins may be correlated to favorable climate conditions for nitrogen fixation millions of years ago. The ancestor of the nitrogen-fixing clade evolved in the early Cretaceous (ca. 92–110 Mya). With all of the four major orders predicted to have originated at around 90 Mya. The oldest known nodule fossils are from the late Santonian 83,6-86.3 Mya (Georgia et al. 1999). It is predicted that most of the nodulating actinorhizal lineages evolved in the late Cretaceous 101,8-71,4 Mya, with only two lineages originating in the late Eocene the genus *Ceanothus* and the Colletieae tribe that contains *Discaria (Li et al. 2015)*. This makes the actinorhizal symbiosis considerably older than the legume-rhizobium symbiosis. In the Fabales order nodulation is supposed to have evolved between 60-70 Mya followed by rapid speciation (J. J. Doyle and Luckow 2003). This rapid speciation may have been driven by ancient polyploidy events(Lavin, Herendeen, and Wojciechowski 2005; Cannon et al. 2015).

Was the evolution of Nodulation driven by a precursor state?

A major outstanding question remains why only this specific group, the NFC, evolved the capacity to nodulate? A hypothesis, as old as the recognition that nitrogen-fixing species belong to a single clade, is the existence of a precursor state or "predisposition" for nodulation (Soltis et al. 1995). In this way in the first common ancestor of the NFC, an innovation happened, and this innovation made it more likely for its descendants to evolve nodulation. This predisposition hypothesis has gained popularity over time. Although much debate exists over what it implies exactly. These vastly different timescales for the evolution of nodulation make the nature of the predisposition trait even more cryptic. If a single origin of the precursor state is implied it means retention of the precursor state for millions of years in some lineages. This means that the precursor state, must be a functional trait in non nodulators.

Modelling approaches, aimed to couple the likelihood of the evolutionary pattern to models with a single, multiple, or no occurrences of a precursor state, revealed that the evolution of a precursor state was a prerequisite to explain the recurrent patterns of the phylogeny. However, it also revealed that nodulation itself and the elusive precursor state can be lost (Werner et al. 2014). Nodulation represents a rather complex trait, which can be easily abolished by a single mutation in any of the genes essential for nodule formation and/or functioning. Conceptually it is therefore much easier to lose the nodulation trait rather than to gain it. The model also predicted that in order to keep nodulation, optimizing mutations accumulate giving rise to "stable fixers". This means that in these species, nodulation becomes embedded in their way of life, making loss of the trait more difficult. This feature may be reflected in the fact that legumes occupy habitats all over the world, also those with moderate or high amounts of nitrogen. Legumes seem to require high nitrogen concentrations, predominantly for their nitrogen-rich seeds and leafs. Legumes predominantly live a nitrogen rich lifestyle, meaning they boost photosynthesis using high leaf nitrogen content (McKey 1994). While the nitrogen-rich seeds kickstart seedling development (Adams et al. 2016). It was predicted that most papilionoid legumes belong to the "stable fixer" category and may be unlikely to lose nodulation due to their high nitrogen demand (Werner et al. 2014, 2015).

Alternative scenarios, The loss of mutualism

The possible loss of mutualism was already recognized in 1995 as an alternative hypothesis to explain the phylogenetic distribution of this trait (Soltis et al. 1995). The possibility of all nodules being homologous structures and therefore derived from a common ancestor was however largely dismissed as a possibility. This was mainly based on structural evidence (Janet I. Sprent 2007). Further, it would imply a large amount of independent losses, which would not be the most parsimonious solution

(Roy and Bousquet 1996; Werner et al. 2014; J. J. Doyle and Luckow 2003; Li et al. 2015; Soltis et al. 1995). Homologous nodulation involved, by definition, only a single recruitment event for each gene. In this case, genes recruited for nodulation would not only be homologous but also orthologous, which means the nodulation genes are derived by speciation, not duplication. Findings of non homologous recruitment, for example were distant paralogues are recruited, provides support for non-homology. It is generally not uncommon for evolutionary events to recurrently recruit similar modules of development. A prime example of this is the recurrent evolution of C4 photosynthesis, in which similar existing enzymatic pathways have been recruited multiple times. With an estimated origin of 22-24 times in Grasses, the existence of a predisposition for evolving C4 was postulated(Christin et al. 2013). It is clear that many genes involved in nodulation in the different species are direct orthologs, such as *NIN*, *SYMRK* and *CYCLOPS*. Nonhomologous nodulation would imply convergent recruitment of these genes in symbiosis. This would however not be uncommon in evolution. Since evolution often acts on existing pathways, recruitment of CSSP or *NIN* could simply be a prerequisite for evolving nitrogen-fixing nodules from the precursor state. This hypothesis implies that the precursor state may have something to do with the regulation of the CSSP or the regulation/targets of NIN (Soyano and Hayashi 2014; Markmann 2008). Evidence of orthology of symbiosis genes does therefore not directly infer homologous nodulation.Genomic comparison to uncover adaptive innovations

Studying *Parasponia* to discover genetic adaptations underlying rhizobium symbiosis.

A strategy to detect the nature of the precursor state and the adaptive innovations that are required to evolve nodulation would be to compare predicted recent gains of nodulation to non-symbiotic sister species (Delaux, Radhakrishnan, and Oldroyd 2015). The chances of success for such studies greatly depends on the genetic distance of the species that will be compared. One of the most recent predicted gains of nodulation would be *Parasponia* in the Cannabaceae lineage, as it is predicted to have evolved its ability to nodulate with rhizobium recently. The *Parasponia* genus only consists of five species and is native to Indonesia, Papua New Guinea and various volcanic island in the South Pacific. *Parasponia* spp. thus grows in a rather specific niche, volcanic soils, which are usually devoid of mineral nitrogen but high in minerals (Shipley 1919; Fujimura et al. 2016). *Parasponia* is closely related to *Trema*, which has a pantropical distribution (M. Q. Yang et al. 2013). *Trema,* according to the previously discussed model, is predicted to be still in precursor state(Werner et al. 2014).

The *Parasponia* lineage seems to represent a unique evolutionary replicate of rhizobium nodulation. Distinct features of the Parasponia - rhizobium interaction has prompted researchers to speculate this interaction was rather young (Behm, Geurts, and Kiers 2014). Parasponia is nodulated by a diverse set of rhizobia, and can form nodules with bacteria of the genus *Bradyrhizobium, Mesorhizobium, Sinorhizobium* and *Rhizobium*. Thus it deploys a remarkable lack of specificity for its symbiont (R. H. M. Op den Camp et al. 2012). Further, it was shown that Parasponia activates this symbiosis by triggering calcium spiking by treatment with LCOs, while this was not the case for *Trema* spp*. (Granqvist et al. 2015)*. As mentioned above, its lateral root-like nodule ontogeny suggests a less derived state. The most compelling indication was however that Parasponia seems to have co-opted its Hemoglobin independently and very recently. The fact that hemoglobin of *Trema spp.* is not adapted for advanced oxygen carrying, although the proteins have 96% identity, supports the hypothesis that Parasponia - rhizobium symbiosis represents a recent evolutionary origin (Sturms et al. 2010; Kakar et al. 2011).

In **Chapter 3**, we sequence and compare the genomes of *Parasponia andersonii* and *Trema orientalis* (Trema) to find gene gains which would correlate to the nitrogen fixation trait. Contrary to our initial expectation, no genes which correlate to the nitrogen fixation trait were identified in Parasponia. Rather we observed a pattern of gene loss in non symbiotic relative Trema. We identify seven genes consistently lost in *Trema* species, among which are *NIN*, *NFP2*, the ortholog of legume nod factor receptor NFP, and *RPG*, a protein exclusively expressed in legume infection threads. These losses occurred in parallel in different *Trema* lineages. This pattern of gene loss is shared by more distant Rosales species. These genes could represent the rewiring of LCO recognition, *NIN* for nodule formation and a common infection hub *RPG* for infection thread formation. The results presented in **Chapter 3** are therefore not in line with the long standing hypothesis of independent origins of nodulation and the occurrence of a predisposition event. On the contrary the results imply a role for gene loss to explain the evolutionary pattern of symbiosis. Besides gene losses in Trema and other Rosales species, we also identify a specific gene loss in *Parasponia*. *Parasponia* species have lost the direct ortholog the Lotus lysM receptor kinase *EXOPOLYSACHARIDE RECEPTOR* (*EPR*), responsible for host range recognition in Lotus. This raised questions on the evolution of the LysM-type receptor gene family of Parasponia. This is a focus of **Chapter 4,** where I determine which Parasponia LysM-type receptors are essential for LCO,CO signalling and its involvement in rhizobium nodulation and arbuscular mycorrhization.

During the genomic comparisons, we also identify two other large receptor kinase families with a putative role in nodulation. These will be the focus of the final experimental Chapter 5. The *CYSTEINE RICH RECEPTOR KINASE 11* (*CRK11*)

and the *LECTIN RECEPTOR KINASE 1* (*LEK1*). Both of these genes are lost in parallel in non nodulating *Trema* relatives, and both have a nodule specific expression pattern Parasponia. In addition, we noted 14 other members the *CRK* family have a nodule enhanced expression pattern. Therefore we decided to evaluate the role of this receptor kinase family in Parasponia - rhizobium nodulation. We reveal a variable cluster of *CRK* genes in the Parasponia genome with role in infection thread progression in the nodule. While PanLEK1 appears to be the only legume lectin involved in nodulation with a strong role in defense response suppression in the nodule.

References

Adams, Mark Andrew, Tarryn L. Turnbull, Janet I. Sprent, and Nina Buchmann. 2016. "Legumes Are Different: Leaf Nitrogen, Photosynthesis, and Water Use Efficiency." *Proceedings of the National Academy of Sciences* 113 (15): 4098–4103.

Appleby, C. A., J. D. Tjepkema, and M. J. Trinick. 1983. "Hemoglobin in a Nonleguminous Plant, Parasponia: Possible Genetic Origin and Function in Nitrogen Fixation." *Science* 220 (4600): 951–53.

Arrighi, Jean-François, Annick Barre, Besma Ben Amor, Anne Bersoult, Lidia Campos Soriano, Rossana Mirabella, Fernanda de Carvalho-Niebel, et al. 2006. "The *Medicago Truncatula* Lysin Motif-Receptor-like Kinase Gene Family Includes NFP and New Nodule-Expressed Genes." *Plant Physiology* 142 (September): 265–79.

Arrighi, Jean-François, Olivier Godfroy, Françoise de Billy, Olivier Saurat, Alain Jauneau, and Clare Gough. 2008. "The *RPG* Gene of *Medicago Truncatula* Controls Rhizobium-Directed Polar Growth during Infection." *Proceedings of the National Academy of Sciences of the United States of America* 105 (28): 9817–22.

Azani, Nasim, Marielle Babineau, C. Donovan Bailey, Hannah Banks, Ariane R. Barbosa, Rafael Barbosa Pinto, James S. Boatwright, et al. 2017. "A New Subfamily Classification of the Leguminosae Based on a Taxonomically Comprehensive Phylogeny." *Taxon* 66 (1): 44–77.

Bailly, Xavier, Isabelle Olivieri, Brigitte Brunel, Jean-Claude Cleyet-Marel, and Gilles Béna. 2007. "Horizontal Gene Transfer and Homologous Recombination Drive the Evolution of the Nitrogen-Fixing Symbionts of Medicago Species." *Journal of Bacteriology* 189 (14): 5223–36.

Bainard, L. D., J. D. Bainard, S. G. Newmaster, and J. N. Klironomos. 2011. "Mycorrhizal Symbiosis Stimulates Endoreduplication in Angiosperms." *Plant, Cell & Environment* 34 (9): 1577–85.

Banba, Mari, Caroline Gutjahr, Akio Miyao, Hirohiko Hirochika, Uta Paszkowski, Hiroshi Kouchi, and Haruko Imaizumi-Anraku. 2008. "Divergence of Evolutionary Ways among Common Sym Genes: CASTOR and CCaMK Show Functional Conservation between Two Symbiosis Systems and Constitute the Root of a Common Signaling Pathway." *Plant & Cell Physiology* 49 (11): 1659–71.

Bapaume, Laure, Sabine Laukamm, Geoffrey Darbon, Corinne Monney, Felix Meyenhofer, Nadja Feddermann, Min Chen, and Didier Reinhardt. 2019. "VAPYRIN Marks an Endosomal Trafficking Compartment Involved in Arbuscular Mycorrhizal Symbiosis." *Frontiers in Plant Science* 10 (June): 1–19.

Behm, Jocelyn E., Rene Geurts, and E. Toby Kiers. 2014. "*Parasponia*: A Novel System for Studying Mutualism Stability." *Trends in Plant Science* 19 (12): 757–63.

Bek, Anita S., Jørgen Sauer, Mikkel B. Thygesen, Jens Ø. Duus, Bent O. Petersen, Søren Thirup, Euan James, Knud J. Jensen, Jens Stougaard, and Simona Radutoiu. 2010. "Improved Characterization of Nod Factors and Genetically Based Variation in LysM Receptor Domains Identify Amino Acids Expendable for Nod Factor Recognition in Lotus Spp. E-Xtra *." *Molecular Plant-Microbe Interactions: MPMI* 58 (1): 58–66.

Berg, R. H. 1999. "Frankia Forms Infection Threads." *Canadian Journal of Botany. Journal Canadien de Botanique* 77 (9): 1327–33.

Berg, R. Howard, Birgit Langenstein, and Warwick B. Silvester. 1999. "Development in the Datisca-Coriaria Nodule Type." *Canadian Journal of Botany. Journal Canadien de Botanique* 77 (9): 1334–50.

Berry, Alison M., Alberto Mendoza-Herrera, Ying Yi Guo, Jennifer Hayashi, Tomas Persson, Ravi Barabote, Kirill Demchenko, Shuxiao Zhang, and Katharina Pawlowski. 2011. "New Perspectives on Nodule Nitrogen Assimilation in Actinorhizal Symbioses." *Functional Plant Biology: FPB* 38 (8-9): 645–52.

Berry, A. M., O. T. Harriott, R. A. Moreau, S. F. Osman, D. R. Benson, and A. D. Jones. 1993. "Hopanoid Lipids Compose the *Frankia* Vesicle Envelope, Presumptive Barrier of Oxygen Diffusion to Nitrogenase." *Proceedings of the National Academy of Sciences of the United States of America* 90 (13): 6091–94.

Bobbink, Roland, Michael Hornung, and Jan G. M. Roelofs. 1998. "The Effects of Air-Borne Nitrogen Pollutants on Species Diversity in Natural and Semi-Natural European Vegetation." *The Journal of Ecology* 86 (5): 717–38.

Bozsoki, Zoltan, Jeryl Cheng, Feng Feng, Kira Gysel, Maria Vinther, Kasper R. Andersen, Giles Oldroyd, Mickael Blaise, Simona Radutoiu, and Jens Stougaard. 2017. "Receptor-Mediated Chitin Perception in Legume Roots Is Functionally Separable from Nod Factor Perception." *Proceedings of the National Academy of Sciences* 114 (38): E8118–27.

Breakspear, Andrew, Chengwu Liu, Sonali Roy, Nicola Stacey, Christian Rogers, Martin Trick, Giulia Morieri, 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." *The Plant Cell* 26 (12): 4680–4701.

Broghammer, A., L. Krusell, M. Blaise, J. Sauer, J. T. Sullivan, N. Maolanon, M. Vinther, et al. 2012. "Legume Receptors Perceive the Rhizobial Lipochitin Oligosaccharide Signal Molecules by Direct Binding." *Proceedings of the National Academy of Sciences of the United States of America* 109 (34): 13859–64.

Brundrett, Mark C. 2002. “Coevolution of Roots and Mycorrhizas of Land Plants.” *The New Phytologist* 154 (2): 275–304.

Brundrett, Mark C., and Leho Tedersoo. 2018. “Evolutionary History of Mycorrhizal Symbioses and Global Host Plant Diversity.” *The New Phytologist* 220 (4): 1108–15.

Buendia, Luis, Ariane Girardin, Tongming Wang, Ludovic Cottret, and Benoit Lefebvre. 2018. “LysM Receptor-Like Kinase and LysM Receptor-Like Protein Families: An Update on Phylogeny and Functional Characterization.” *Frontiers in Plant Science* 9 (October): 1531.

Camargo, Julio A., and Alvaro Alonso. 2006. “Ecological and Toxicological Effects of Inorganic Nitrogen Pollution in Aquatic Ecosystems: A Global Assessment.” *Environment International* 32 (6): 831–49.

Cannon, Steven B., Michael R. McKain, Alex Harkess, Matthew N. Nelson, Sudhansu Dash, Michael K. Deyholos, Yanhui Peng, et al. 2015. “Multiple Polyploidy Events in the Early Radiation of Nodulating and Nonnodulating Legumes.” *Molecular Biology and Evolution* 32 (1): 193–210.

Cerri, Marion R., Lisa Frances, Audrey Kelner, Joelle Fournier, Patrick H. Middleton, Marie-Christine Auriac, Kirankumar S. Mysore, et al. 2016. “The Symbiosis-Related ERN Transcription Factors Act in Concert to Coordinate Rhizobial Host Root Infection.” *Plant Physiology* 171 (June): 00230.2016.

Cerri, Marion R., Quanhui Wang, Paul Stolz, Jessica Folgmann, Lisa Frances, Katja Katzer, Xiaolin Li, et al. 2017. “The *ERN1* Transcription Factor Gene Is a Target of the CCaMK/CYCLOPS Complex and Controls Rhizobial Infection in *Lotus Japonicus*.” *The New Phytologist*, 323–37.

Cerri, M. R., L. Frances, T. Laloum, M-C Auriac, A. Niebel, G. E. D. Oldroyd, D. G. Barker, J. Fournier, and F. de Carvalho-Niebel. 2012. “Medicago Truncatula ERN Transcription Factors: Regulatory Interplay with NSP1/NSP2 GRAS Factors and Expression Dynamics throughout Rhizobial Infection.” *Plant Physiology* 160 (December): 2155–72.

Chabaud, Mireille, Hassen Gherbi, Elodie Pirolles, Virginie Vaissayre, Daniel Moukouanga, Claudine Franche, Didier Bogusz, Louis S. Tisa, David G. Barker, and Sergio Svistoonoff. 2016. “Chitinase-Resistant Hydrophilic Symbiotic Factors Secreted by *Frankia* Activate Both Ca2+ Spiking and *NIN* Gene Expression in the Actinorhizal Plant *Casuarina Glauca*.” *The New Phytologist* 209 (1): 86–93.

Charpentier, M., J. Sun, and T. V. Martins. 2016. “Nuclear-Localized Cyclic Nucleotide–gated Channels Mediate Symbiotic Calcium Oscillations.” https://science.sciencemag.org/content/352/6289/1102.short?casa_token=2gZL8QTriRoAAAAA:C8aRsQXLv4uZ9z3U9DnYg9j8HEJh6GES6Dfwh7kO6sCzXocQTwryMLj0r_UcW-Q83Gcz300aBo7mh-NA.

Charpentier, Myriam, and Giles Oldroyd. 2010. “How Close Are We to Nitrogen-Fixing Cereals?” *Current Opinion in Plant Biology* 13 (5): 556–64.

Charpentier, Myriam, Jongho Sun, Teresa Vaz Martins, Guru V. Radhakrishnan, Kim Findlay, Eleni Soumpourou, Julien Thouin, et al. 2016. “Symbiotic Calcium Oscillations.” *Science* 352 (6289): 1102–5.

Choudhury, A. T. M. A., and I. R. Kennedy. 2005. “Nitrogen Fertilizer Losses from Rice Soils and Control of Environmental Pollution Problems.” *Communications in Soil Science and Plant Analysis* 36 (11-12): 1625–39.

Christin, Pascal Antoine, Colin P. Osborne, David S. Chatelet, J. Travis Columbus, Guillaume Besnard, Trevor R. Hodkinson, Laura M. Garrison, Maria S. Vorontsova, and Erika J. Edwards. 2013. “Anatomical Enablers and the Evolution of C4 Photosynthesis in Grasses.” *Proceedings of the National Academy of Sciences of the United States of America* 110 (4): 1381–86.

Clavijo, Fernando, Issa Diedhiou, Virginie Vaissayre, Laurent Brottier, Jennifer Acolatse, Daniel Moukouanga, Amandine Crabos, et al. 2015. “The *Casuarina NIN* Gene Is Transcriptionally Activated throughout *Frankia* Root Infection as Well as in Response to Bacterial Diffusible Signals.” *The New Phytologist* 208: 887–903.

De Cuyper, C., J. Fromentin, R. E. Yocgo, A. De Keyser, B. Guillotin, K. Kunert, F-D Boyer, and S. Goormachtig. 2014. “From Lateral Root Density to Nodule Number, the Strigolactone Analogue GR24 Shapes the Root Architecture of *Medicago Truncatula*.” *Journal of Experimental Botany* 66 (1): 137–46.

Delaux, Pierre Marc, Guru Radhakrishnan, and Giles Oldroyd. 2015. “Tracing the Evolutionary Path to Nitrogen-Fixing Crops.” *Current Opinion in Plant Biology* 26: 95–99.

Demina, Irina V., Tomas Persson, Patricia Santos, Marian Plaszczyca, and Katharina Pawlowski. 2013. “Comparison of the Nodule vs. Root Transcriptome of the Actinorhizal Plant Datisca Glomerata: Actinorhizal Nodules Contain a Specific Class of Defensins.” *PloS One* 8 (8). https://doi.org/10.1371/journal.pone.0072442.

De Mita, Stéphane, Arend Streng, Ton Bisseling, and René Geurts. 2014. “Evolution of a Symbiotic Receptor through Gene Duplications in the Legume-Rhizobium Mutualism.” *The New Phytologist* 201 (3): 961–72.

Dénarié, Jean, Frédéric Debellé, and Jean-Claude Promé. 1996. “Rhizobium Lipo-Chitooligosaccharide Nodulation Factors: Signaling Molecules Mediating Recognition and Morphogenesis.” *Annual Review of Biochemistry* 65 (1):

503–35.

Doyle, Jeff J. 1994. “Phylogeny of the Legume Family: An Approach to Understanding the Origins of Nodulation.” *Annual Review of Ecology and Systematics* 25: 325–49.

Doyle, Jeff J. 2011. “Phylogenetic Perspectives on the Origins of Nodulation.” *Molecular Plant-Microbe Interactions: MPMI* 24 (November): 1289–95.

Doyle, J. J. 2016. “Chasing Unicorns: Nodulation Origins and the Paradox of Novelty.” *American Journal of Botany* 103 (11): 1–4.

Doyle, J. J., and M. A. Luckow. 2003. “The Rest of the Iceberg. Legume Diversity and Evolution in a Phylogenetic Context.” *Plant Physiology* 131 (3): 900–910.

Ehrhardt, David W., Rebecca Wais, and Sharon R. Long. 1996. “Calcium Spiking in Plant Root Hairs Responding to Rhizobium Modulation Signals.” *Cell* 85 (5): 673–81.

Faulkner, Christine, Elena Petutschnig, Yoselin Benitez-Alfonso, Martina Beck, Silke Robatzek, Volker Lipka, and Andrew J. Maule. 2013. “LYM2-Dependent Chitin Perception Limits Molecular Flux via Plasmodesmata.” *Proceedings of the National Academy of Sciences of the United States of America*. https://doi.org/10.1073/pnas.1203458110.

Feng, Feng, Jongho Sun, Guru V. Radhakrishnan, Tak Lee, Zoltán Bozsóki, Sébastien Fort, Aleksander Gavrin, et al. 2019. “A Combination of Chitooligosaccharide and Lipochitooligosaccharide Recognition Promotes Arbuscular Mycorrhizal Associations in Medicago Truncatula.” *Nature Communications* 10 (1): 5047.

Fliegmann, Judith, Sophie Canova, Christophe Lachaud, Sandra Uhlenbroich, Virginie Gasciolli, Carole Pichereaux, Michel Rossignol, et al. 2013. “Lipo-Chitooligosaccharidic Symbiotic Signals Are Recognized by LysM Receptor-like Kinase LYR3 in the Legume Medicago Truncatula.” *ACS Chemical Biology* 8: 1900–1906.

Fliegmann, Judith, Alain Jauneau, Carole Pichereaux, Charles Rosenberg, Virginie Gasciolli, Antonius C. J. J. Timmers, Odile Burlet-Schiltz, Julie Cullimore, and Jean-Jacques Bono. 2016. “LYR3, a High-Affinity LCO-Binding Protein of *Medicago Truncatula* , Interacts with LYK3, a Key Symbiotic Receptor.” *FEBS Letters* 590: 1477–87.

Foo, Eloise, and Noel W. Davies. 2011. “Strigolactones Promote Nodulation in Pea.” *Planta* 234 (5): 1073–81.

Fujimura, Reiko, Seok-Won Kim, Yoshinori Sato, Kenshiro Oshima, Masahira Hattori, Takashi Kamijo, and Hiroyuki Ohta. 2016. “Unique Pioneer Microbial Communities Exposed to Volcanic Sulfur Dioxide.” *Scientific Reports* 6: 19687.

Georgia, Central, A. Preliminary Patrick S. Herendeen, Susana Magallon-, Conspectus Of, T. H. E. Allon, Richard Lupia, and R. Peter. 1999. “A Preliminary Conspectus of the Allon Flora from the Late Cretaceous (Late Santonian) of Author (S): Patrick S . Herendeen , Susana Magallon-Puebla , Richard Lupia , Peter R . Crane and Jolanta Kobylinska Source : Annals of the Missouri Botanical Gard.” *Annals of the Missouri Botanical Garden. Missouri Botanical Garden* 86 (2): 407–71.

Gherbi, Hassen, Katharina Markmann, Sergio Svistoonoff, Joan Estevan, Daphné Autran, Gabor Giczey, Florence Auguy, et al. 2008. “SymRK Defines a Common Genetic Basis for Plant Root Endosymbioses with Arbuscular Mycorrhiza Fungi, Rhizobia, and *Frankia* Bacteria.” *Proceedings of the National Academy of Sciences of the United States of America* 105 (12): 4928–32.

Gibelin-Viala, Chrystel, Emilie Amblard, Virginie Puech-Pages, Maxime Bonhomme, Magali Garcia, Adeline Bascaules-Bedin, Judith Fliegmann, et al. 2019. “The Medicago Truncatula LysM Receptor-like Kinase LYK9 Plays a Dual Role in Immunity and the Arbuscular Mycorrhizal Symbiosis.” *The New Phytologist* 223 (3): 1516–29.

Gleason, C., S. Chaudhuri, T. B. Yang, A. Munoz, B. W. Poovaiah, and G. E. D. Oldroyd. 2006. “Nodulation Independent of Rhizobia Induced by a Calcium-Activated Kinase Lacking Autoinhibition.” *Nature* 441 (7097): 1149–52.

Gobbato, Enrico, Ertao Wang, Gillian Higgins, Syeda Asma Bano, Christine Henry, Michael Schultze, and Giles Ed Oldroyd. 2013. “*RAM1* and *RAM2* Function and Expression during Arbuscular Mycorrhizal Symbiosis and *Aphanomyces Euteiches* Colonization.” *Plant Signaling & Behavior* 8 (10): e26049.

Gough, Clare, Ludovic Cottret, Benoit Lefebvre, and Jean-Jacques Bono. 2018. “Evolutionary History of Plant LysM Receptor Proteins Related to Root Endosymbiosis.” *Frontiers in Plant Science* 9 (July): 1–9.

Gough, Clare, and Christophe Jacquet. 2013. “Nod Factor Perception Protein Carries Weight in Biotic Interactions.” *Trends in Plant Science* 18 (10): 566–74.

Granqvist, Emma, Jongho Sun, Rik Op Den Camp, Petar Pujic, Lionel Hill, Philippe Normand, Richard J. Morris, et al. 2015. “Bacterial-Induced Calcium Oscillations Are Common to Nitrogen-Fixing Associations of Nodulating Legumes and Nonlegumes.” *The New Phytologist* 207 (3): 551–58.

Gtari, Maher, Faten Ghodhbane-Gtari, Imen Nouioui, Amir Ktari, Karima Hezbri, Wajdi Mimouni, Imed Sbissi, et al. 2015. “Cultivating the Uncultured: Growing the Recalcitrant Cluster-2 Frankia Strains.” *Scientific Reports* 5

(August): 13112.

Guefrachi, I., and M. Nagymihaly. 2014. “Extreme Specificity of NCR Gene Expression in Medicago Truncatula.” *BMC …*. http://www.biomedcentral.com/1471-2164/15/712.

Gutjahr, Caroline. 2014. “Phytohormone Signaling in Arbuscular Mycorhiza Development.” *Current Opinion in Plant Biology* 20 (August): 26–34.

Haag, Andreas F., Bernhard Kerscher, Sergio Dall’Angelo, Monica Sani, Renato Longhi, Mikhail Baloban, Heather M. Wilson, Peter Mergaert, Matteo Zanda, and Gail P. Ferguson. 2012. “Role of Cysteine Residues and Disulfide Bonds in the Activity of a Legume Root Nodule-Specific, Cysteine-Rich Peptide.” *The Journal of Biological Chemistry* 287 (14): 10791–98.

Haney, Cara H., Brendan K. Riely, David M. Tricoli, Doug R. Cook, David W. Ehrhardt, and Sharon R. Long. 2011. “Symbiotic Rhizobia Bacteria Trigger a Change in Localization and Dynamics of the Medicago Truncatula Receptor Kinase LYK3.” *The Plant Cell* 23 (7): 2774–87.

Hastwell, April H., Peter M. Gresshoff, and Brett J. Ferguson. 2015. “The Structure and Activity of Nodulation-Suppressing CLE Peptide Hormones of Legumes.” *Functional Plant Biology: FPB* 42 (3): 229–38.

Heckmann, Anne B., Kim H. Hebelstrup, Knud Larsen, Nuno M. Micaelo, and Erik Ø. Jensen. 2006. “A Single Hemoglobin Gene in Myrica Gale Retains Both Symbiotic and Non-Symbiotic Specificity.” *Plant Molecular Biology* 61 (4-5): 769–79.

Heckmann, Anne Birgitte, Niels Sandal, Anita Søndergaard Bek, Lene Heegaard Madsen, Anna Jurkiewicz, Mette Wibroe Nielsen, Leila Tirichine, and Jens Stougaard. 2011. “Cytokinin Induction of Root Nodule Primordia in *Lotus Japonicus* Is Regulated by a Mechanism Operating in the Root Cortex.” *Molecular Plant-Microbe Interactions: MPMI* 24 (11): 1385–95.

Hirsch, Sibylle, Jiyoung Kim, Alfonso Muñoz, Anne B. Heckmann, J. Allan Downie, and Giles E. D. Oldroyd. 2009. “GRAS Proteins Form a DNA Binding Complex to Induce Gene Expression during Nodulation Signaling in *Medicago Truncatula*.” *The Plant Cell* 21 (2): 545–57.

Hohnjec, Natalija, Lisa F. Czaja-Hasse, Claudia Hogekamp, and Helge Küster. 2015. “Pre-Announcement of Symbiotic Guests: Transcriptional Reprogramming by Mycorrhizal Lipochitooligosaccharides Shows a Strict Co-Dependency on the GRAS Transcription Factors NSP1 and RAM1.” *BMC Genomics* 16 (1): 994.

Horváth, Beatrix, Li Huey Yeun, Ágota Domonkos, Gábor Halász, Enrico Gobbato, Ferhan Ayaydin, Krisztina Miró, et al. 2011. “*Medicago Truncatula* IPD3 Is a Member of the Common Symbiotic Signaling Pathway Required for Rhizobial and Mycorrhizal Symbioses.” *Molecular Plant-Microbe Interactions: MPMI* 24 (11): 1345–58.

Imanishi, Leandro, Mireille Chabaud, Iltaf Abdou-pavy, Andrea Genre, Lukas Brichet, Alice Vayssi, Elodie Pirolles, et al. 2018. “Cell Remodeling and Subtilase Gene Expression in the Actinorhizal Plant Discaria Trinervis Highlight Host Orchestration of Intercellu- Lar Frankia Colonization,” 1018–30.

Jacobsen-Lyon, K., E. O. Jensen, J. E. Jørgensen, K. a. Marcker, W. J. Peacock, and E. S. Dennis. 1995. “Symbiotic and Nonsymbiotic Hemoglobin Genes of Casuarina Glauca.” *The Plant Cell* 7 (2): 213–23.

Kakar, Smita, Ryan Sturms, Andrea Tiffany, Jay C. Nix, Alan A. DiSpirito, and Mark S. Hargrove. 2011. “Crystal Structures of *Parasponia* and *Trema* Hemoglobins: Differential Heme Coordination Is Linked to Quaternary Structure.” *Biochemistry* 50 (20): 4273–80.

Kassaw, Tessema, Stephen Nowak, Elise Schnabel, and Julia A. Frugoli. 2017. “ROOT DETERMINED NODULATION1 Is Required for M. Truncatula CLE12, but Not CLE13 Peptide Signaling through the SUNN Receptor Kinase.” *Plant Physiology* 174 (August): 00278.2017.

Kawaharada, Yasuyuki, Mette W. Nielsen, Simon Kelly, Euan K. James, Kasper R. Andersen, Sheena R. Rasmussen, Winnie Füchtbauer, et al. 2017. “Differential Regulation of the Epr3 Receptor Coordinates Membrane-Restricted Rhizobial Colonization of Root Nodule Primordia.” *Nature Communications* 8 (February): 14534.

Kim, Sunghoon, Weizhong Zeng, Shane Bernard, Jun Liao, Muthusubramanian Venkateshwaran, Jean-Michel Ane, and Youxing Jiang. 2019. “Ca2+-Regulated Ca2+ Channels with an RCK Gating Ring Control Plant Symbiotic Associations.” *Nature Communications* 10 (1): 3703.

Kiss, Ernö, Boglárka Oláh, Péter Kaló, Monica Morales, Anne B. Heckmann, Andrea Borbola, Anita Lózsa, et al. 2009. “LIN, a Novel Type of U-box/WD40 Protein, Controls Early Infection by Rhizobia in Legumes.” *Plant Physiology* 151 (November): 1239–49.

Konishi, Mineko, and Shuichi Yanagisawa. 2013. “Arabidopsis NIN-like Transcription Factors Have a Central Role in Nitrate Signalling.” *Nature Communications* 4: 1617.

Ktari, Amir, Imen Nouioui, Teal Furnholm, Erik Swanson, Faten Ghodhbane-Gtari, Louis S. Tisa, and Maher Gtari. 2017. “Permanent Draft Genome Sequence of Frankia Sp. NRRL B-16219 Reveals the Presence of Canonical Nod Genes, Which Are Highly Homologous to Those Detected in Candidatus Frankia Dg1 Genome.” *Standards in*

Genomic Sciences 12 (1): 1–10.

Laloum, Tom, Ma€ El Baudin, Lisa Frances, Agnes Lepage, Benjamin Billault-Penneteau, Marion R. Cerri, Federico Ariel, et al. 2014. "Two CCAAT-Box-Binding Transcription Factors Redundantly Regulate Early Steps of the Legume-Rhizobia Endosymbiosis." *The Plant Journal: For Cell and Molecular Biology* 79: 757–68.

Lancelle, S. A., and J. G. Torrey. 1984. "Early Development of Rhizobium-Induced Root Nodules of *Parasponia Rigida*. I. Infection and Early Nodule Initiation." *Protoplasma* 123: 26–37.

Lavin, Matt, Patrick S. Herendeen, and Martin F. Wojciechowski. 2005. "Evolutionary Rates Analysis of Leguminosae Implicates a Rapid Diversification of Lineages during the Tertiary." *Systematic Biology* 54 (4): 575–94.

Leppyanen, Irina V., Vlada Y. Shakhnazarova, Oksana Y. Shtark, Nadezhda A. Vishnevskaya, Igor A. Tikhonovich, and Elena A. Dolgikh. 2017. "Receptor-Like Kinase LYK9 in Pisum Sativum L. Is the CERK1-Like Receptor That Controls Both Plant Immunity and AM Symbiosis Development." *International Journal of Molecular Sciences* 19 (1). https://doi.org/10.3390/ijms19010008.

Lévy, Julien, Cécile Bres, René Geurts, Boulos Chalhoub, Olga Kulikova, Gérard Duc, Etienne-Pascal Journet, et al. 2004. "A Putative Ca2+ and Calmodulin-Dependent Protein Kinase Required for Bacterial and Fungal Symbioses." *Science* 303 (5662): 1361–64.

Liang, Pengbo, Thomas F. Stratil, Claudia Popp, Macarena Marin, Jessica Folgmann, Kirankumar S. Mysore, Jiangqi Wen, and Thomas Ott. 2017. "Symbiotic Root Infections in Medicago Truncatula Require Remorin-Mediated Receptor Stabilization in Membrane Nanodomains." *bioRxiv*, 179036.

Libbenga, K. R., and P. A. A. Harkes. 1973. "Initial Proliferation of Cortical Cells in the Formation of Root Nodules in Pisum Sativum L." *Planta* 114 (1): 17–28.

Li, Hong-Lei, Wei Wang, Peter E. Mortimer, Rui-Qi Li, De-Zhu Li, Kevin D. Hyde, Jian-Chu Xu, Douglas E. Soltis, and Zhi-Duan Chen. 2015. "Large-Scale Phylogenetic Analyses Reveal Multiple Gains of Actinorhizal Nitrogen-Fixing Symbioses in Angiosperms Associated with Climate Change." *Scientific Reports* 5 (1): 14023.

Limpens, Erik, Carolien Franken, Patrick Smit, Joost Willemse, Ton Bisseling, and René Geurts. 2003. "LysM Domain Receptor Kinases Regulating Rhizobial Nod Factor-Induced Infection." *Science* 302 (5645): 630–33.

Limpens, Erik, Rossana Mirabella, Elena Fedorova, Carolien Franken, Henk Franssen, Ton Bisseling, and René Geurts. 2005. "Formation of Organelle-like N2-Fixing Symbiosomes in Legume Root Nodules Is Controlled by DMI2." *Proceedings of the National Academy of Sciences of the United States of America* 102 (29): 10375–80.

Liu, Qinqin, and A. M. Berry. 1991. "M IDI The Infection Process and Nodule Initiation in the Frankia-Ceanothus Root Nodule Symbiosis A Structural and Histochemical Study," 82–92.

Liu, Wei, Wouter Kohlen, Alessandra Lillo, Rik Op den Camp, Sergey Ivanov, Marijke Hartog, Erik Limpens, et al. 2011. "Strigolactone Biosynthesis in *Medicago Truncatula* and Rice Requires the Symbiotic GRAS-Type Transcription Factors NSP1 and NSP2." *The Plant Cell* 23 (10): 3853–65.

Marchetti, Marta, Alain Jauneau, Delphine Capela, Philippe Remigi, Carine Gris, Jacques Batut, and Catherine Masson-Boivin. 2014. "Shaping Bacterial Symbiosis With Legumes by Experimental Evolution." *Molecular Plant-Microbe Interactions: MPMI* 27 (9): 956–64.

Markmann, Katharina. 2008. "Functional Adaptation of the Plant Receptor-Kinase Gene SYMRK Paved the Way for the Evolution of Root Endosymbioses with Bacteria."

Markmann, Katharina, Gábor Giczey, and Martin Parniske. 2008. "Functional Adaptation of a Plant Receptor-Kinase Paved the Way for the Evolution of Intracellular Root Symbioses with Bacteria." *PLoS Biology* 6 (3): 0497–0506.

Marsh, John F., Alexandra Rakocevic, Raka M. Mitra, Lysiane Brocard, Jongho Sun, Alexis Eschstruth, Sharon R. Long, Michael Schultze, Pascal Ratet, and G. E. D. D. Oldroyd. 2007. "*Medicago Truncatula* NIN Is Essential for Rhizobial-Independent Nodule Organogenesis Induced by Autoactive Calcium/calmodulin-Dependent Protein Kinase." *Plant Physiology* 144 (1): 324–35.

McKey, D. 1994. "Legumes and Nitrogen: Th Evolutionary Ecology of a Nitrogen-Demanding Lifestyle." In *Advances in LEgume Systematics 5: The Nitrogen Factor.*, edited by J. I. Sprent and D. McKey, V, 211–28. Kew.

Miller, I. M., and D. D. Baker. 1985. "PROTOPL6SMA The Initiation , Development and Structure of Root Nodules in Elaeagnus Angustifolia L . (Elaeagnaceae)" 119: 107–19.

Moling, Sjef, Anna Pietraszewska-Bogiel, Marten Postma, Elena Fedorova, Mark a. Hink, Erik Limpens, Theodorus W. J. Gadella, and Ton Bisseling. 2014. "Nod Factor Receptors Form Heteromeric Complexes and Are Essential for Intracellular Infection in Medicago Nodules." *The Plant Cell* 26 (10): 4188–99.

Mortier, Virginie, Eva De Wever, Marnik Vuylsteke, Marcelle Holsters, and Sofie Goormachtig. 2012. "Nodule Numbers Are Governed by Interaction between CLE Peptides and Cytokinin Signaling." *The Plant Journal: For Cell and Molecular Biology* 70 (3): 367–76.

Murray, Jeremy D., Bogumil J. Karas, Shusei Sato, Satoshi Tabata, Lisa Amyot, and Krzysztof Szczyglowski. 2007.

"A Cytokinin Perception Mutant Colonized by Rhizobium in the Absence of Nodule Organogenesis." *Science* 315 (5808): 101–4.

Murray, Jeremy D., Rajasekhara Reddy Duvvuru Muni, Ivone Torres-Jerez, Yuhong Tang, Stacy Allen, Megan Andriankaja, Guangming Li, et al. 2011. "Vapyrin, a Gene Essential for Intracellular Progression of Arbuscular Mycorrhizal Symbiosis, Is Also Essential for Infection by Rhizobia in the Nodule Symbiosis of Medicago Truncatula." *The Plant Journal: For Cell and Molecular Biology* 65: 244–52.

Mus, Florence, Matthew B. Crook, Kevin Garcia, Amaya Garcia Costas, Barney A. Geddes, Evangelia D. Kouri, Ponraj Paramasivan, et al. 2016. "Symbiotic Nitrogen Fixation and the Challenges to Its Extension to Nonlegumes." *Applied and Environmental Microbiology* 82 (13): 3698–3710.

Naisbitt, T., Ek K. James, and Ji Janet I. Sprent. 1992. "The Evolutionary Significance of the Legume Genus Chamaecrista, as Determined by Nodule Structure." *The New Phytologist* 122 (3): 487–92.

Navazio, Lorella, Roberto Moscatiello, Andrea Genre, Mara Novero, Barbara Baldan, Paola Bonfante, and Paola Mariani. 2007. "A Diffusible Signal from Arbuscular Mycorrhizal Fungi Elicits a Transient Cytosolic Calcium Elevation in Host Plant Cells." *Plant Physiology* 144 (2): 673–81.

Okamoto, Satoru, and Masayoshi Kawaguchi. 2015. "Shoot HAR1 Mediates Nitrate Inhibition of Nodulation in Lotus Japonicus." *Plant Signaling & Behavior* 10 (5): e1000138.

Okamoto, Satoru, Erika Ohnishi, Shusei Sato, Hirokazu Takahashi, Mikio Nakazono, Satoshi Tabata, and Masayoshi Kawaguchi. 2009. "Nod Factor/nitrate-Induced *CLE* Genes That Drive HAR1-Mediated Systemic Regulation of Nodulation." *Plant & Cell Physiology* 50 (1): 67–77.

Oldroyd, Giles E. D., and J. Allan Downie. 2006. "Nuclear Calcium Changes at the Core of Symbiosis Signalling." *Current Opinion in Plant Biology* 9 (4): 351–57.

Op den Camp, Rik H. M., Elisa Polone, Elena Fedorova, Wim Roelofsen, Andrea Squartini, Huub J. M. Op den Camp, Ton Bisseling, and René Geurts. 2012. "Deploys a Broad Rhizobium Host Range Strategy Resulting in Largely Variable Symbiotic Effectiveness." *Molecular Plant-Microbe Interactions: MPMI* 25 (7): 954–63.

Op den Camp, Rik, Arend Streng, Stéphane De Mita, Qingqin Cao, Elisa Polone, Wei Liu, Jetty S. S. Ammiraju, et al. 2011. "LysM-Type Mycorrhizal Receptor Recruited for Rhizobium Symbiosis in Nonlegume *Parasponia*." *Science* 331: 909–12.

Ott, Thomas, Joost T. Van Dongen, Catrin Günther, Lene Krusell, Guilhem Desbrosses, Helene Vigeolas, Vivien Bock, Tomasz Czechowski, Peter Geigenberger, and Michael K. Udvardi. 2005. "Symbiotic Leghemoglobins Are Crucial for Nitrogen Fixation in Legume Root Nodules but Not for General Plant Growth and Development." *Current Biology: CB* 15 (6): 531–35.

Pawlowski, Katharina. 2008. "Nodules and Oxygen." *Plant Biotechnology* 25 (3): 291–98.

Pawlowski, Katharina, and Ton Bisseling. 1996. "Rhizobial and Actinorhizal Symbioses: What Are the Shared Features?" *The Plant Cell* 8 (10): 1899.

Pawlowski, Katharina, and Kirill N. Demchenko. 2012. "The Diversity of Actinorhizal Symbiosis." *Protoplasma* 249 (4): 967–79.

Pawlowski, K., K. R. Jacobsen, N. Alloisio, R. Ford Denison, M. Klein, J. D. Tjepkema, T. Winzer, A. Sirrenberg, C. Guan, and A. M. Berry. 2007. "Truncated Hemoglobins in Actinorhizal Nodules of Datisca Glomerata." *Plant Biology* 9 (6): 776–85.

Penterman, Jon, Ryan P. Abo, Nicole J. De Nisco, Markus F. F. Arnold, Renato Longhi, Matteo Zanda, and Graham C. Walker. 2014. "Host Plant Peptides Elicit a Transcriptional Response to Control the Sinorhizobium Meliloti Cell Cycle during Symbiosis." *Proceedings of the National Academy of Sciences of the United States of America* 111 (9): 3561–66.

Persson, Tomas, Kai Battenberg, Irina V. Demina, Theoden Vigil-Stenman, Brian Vanden Heuvel, Petar Pujic, Marc T. Facciotti, et al. 2015. "*Candidatus* Frankia Datiscae Dg1, the Actinobacterial Microsymbiont of *Datisca Glomerata*, Expresses the Canonical *nod* Genes *nodABC* in Symbiosis with Its Host Plant." *PloS One* 10 (5): e0127630.

Pimprikar, Priya, Samy Carbonnel, Michael Paries, Katja Katzer, Verena Klingl, Monica J. J. Bohmer, Leonhard Karl, et al. 2016. "A CCaMK-CYCLOPS-DELLA Complex Activates Transcription of *RAM1* to Regulate Arbuscule Branching." *Current Biology: CB* 26 (8): 987–98.

Price, G. D., S. S. Mohapatra, and P. M. Gresshoff. 1984. "Structure of Nodules Formed by Rhizobium Strain ANU289 in the Nonlegume *Parasponia* and the Legume Siratro (*Macroptilium Atropurpureum*)." *Botanical Gazette* 145 (4): 444–51.

Radutoiu, S., L. H. Madsen, E. B. Madsen, H. H. Felle, Y. Umehara, M. Gronlund, S. Sato, et al. 2003. "Plant Recognition of Symbiotic Bacteria Requires Two LysM Receptor-like Kinases." *Nature* 425 (6958): 585–92.

Raven, J. A., and D. Edwards. 2001. "Roots: Evolutionary Origins and Biogeochemical Significance." *Journal of*

Experimental Botany 52 (Spec Issue): 381–401.

Ried, Martina Katharina, Meritxell Antolín-Llovera, and Martin Parniske. 2014. "Spontaneous Symbiotic Reprogramming of Plant Roots Triggered by Receptor-like Kinases." *eLife* 3: 1–17.

Roy, Alice, and Jean Bousquet. 1996. "The Evolution of the Actinorhizal Symbiosis through Phylogenetic Analysis of Host Plants." *Acta Botanica Gallica: Bulletin de La Societe Botanique de France* 143 (7): 635–50.

Saha, Sudip, Anindita Paul, Laura Herring, Ayan Dutta, Avisek Bhattacharya, Sandip Samaddar, Michael B. Goshe, and Maitrayee DasGupta. 2016. "Gatekeeper Tyrosine Phosphorylation of SYMRK Is Essential for Synchronising the Epidermal and Cortical Responses in Root Nodule Symbiosis." *Plant Physiology* 171 (May): 01962.2015.

Salgado, Marco G., Robin van Velzen, Thanh Van Nguyen, Kai Battenberg, Alison M. Berry, Daniel Lundin, and Katharina Pawlowski. 2018. "Comparative Analysis of the Nodule Transcriptomes of Ceanothus Thyrsiflorus (Rhamnaceae, Rosales) and Datisca Glomerata (Datiscaceae, Cucurbitales)." *Frontiers in Plant Science* 9 (November): 1629.

Sasaki, Takema, Takuya Suzaki, Takashi Soyano, Mikiko Kojima, Hitoshi Sakakibara, and Masayoshi Kawaguchi. 2014. "Shoot-Derived Cytokinins Systemically Regulate Root Nodulation." *Nature Communications* 5 (September): 4983.

Sasakura, Fuyuko, Toshiki Uchiumi, Yoshikazu Shimoda, Akihiro Suzuki, Katsumi Takenouchi, Shiro Higashi, and Mikiko Abe. 2006. "A Class 1 Hemoglobin Gene from *Alnus Firma* Functions in Symbiotic and Nonsymbiotic Tissues to Detoxify Nitric Oxide." *Molecular Plant-Microbe Interactions: MPMI* 19 (4): 441–50.

Schauser, Leif, Wioletta Wieloch, and Jens Stougaard. 2005. "Evolution of NIN-like Proteins in Arabidopsis, Rice, and Lotus Japonicus." *Journal of Molecular Evolution* 60: 229–37.

Schauser, L., A. Roussis, J. Stiller, and J. Stougaard. 1999. "A Plant Regulator Controlling Development of Symbiotic Root Nodules." *Nature* 402 (6758): 191–95.

Selosse, M. A., and F. Le Tacon. 1998. "The Land Flora: A Phototroph-Fungus Partnership?" *Trends in Ecology & Evolution* 13 (1): 15–20.

Shipley, J. W. 1919. "The Nitrogen Content of Volcanic Ash in the Katmai Eruption of 1912 (Part V of Series)." *The Ohio Journal of Science* 19 (4): 213–23.

Sieberer, Björn J., Mireille Chabaud, Joëlle Fournier, Antonius C. J. Timmers, and David G. Barker. 2012. "A Switch in Ca 2+ Spiking Signature Is Concomitant with Endosymbiotic Microbe Entry into Cortical Root Cells of Medicago Truncatula." *The Plant Journal: For Cell and Molecular Biology* 69: 822–30.

Smit, Patrick, Erik Limpens, Rene Geurts, Elena Fedorova, Elena Dolgikh, Clare Gough, and Ton Bisseling. 2007. "Medicago LYK3, an Entry Receptor in Rhizobial Nodulation Factor Signaling." *Plant Physiology* 145 (September): 183–91.

Soltis, D. E., P. S. Soltis, D. R. Morgan, S. M. Swensen, B. C. Mullin, J. M. Dowd, and P. G. Martin. 1995. "Chloroplast Gene Sequence Data Suggest a Single Origin of the Predisposition for Symbiotic Nitrogen Fixation in Angiosperms." *Proceedings of the National Academy of Sciences of the United States of America* 92 (7): 2647–51.

Soyano, Takashi, and Makoto Hayashi. 2014. "Transcriptional Networks Leading to Symbiotic Nodule Organogenesis." *Current Opinion in Plant Biology* 20: 146–54.

Soyano, Takashi, Hiroshi Kouchi, Atsuko Hirota, and Makoto Hayashi. 2013. "NODULE INCEPTION Directly Targets NF-Y Subunit Genes to Regulate Essential Processes of Root Nodule Development in Lotus Japonicus." *PLoS Genetics* 9 (3): e1003352.

Sprent, Janet I. 2007. "Evolving Ideas of Legume Evolution and Diversity: A Taxonomic Perspective on the Occurrence of Nodulation." *The New Phytologist* 174 (1): 11–25.

Sprent, Janet I., Julie Ardley, and Euan K. James. 2017. "Biogeography of Nodulated Legumes and Their Nitrogen-Fixing Symbionts." *The New Phytologist* 215 (1): 40–56.

Stracke, Silke, Catherine Kistner, Satoko Yoshida, Lonneke Mulder, Shusei Sato, Takakazu Kaneko, Satoshi Tabata, et al. 2002. "A Plant Receptor-like Kinase Required for Both Bacterial and Fungal Symbiosis." *Nature* 417: 959–62.

Sturms, Ryan, Smita Kakar, James Trent, and Mark S. Hargrove. 2010. "Trema and Parasponia Hemoglobins Reveal Convergent Evolution of Oxygen Transport in Plants." *Biochemistry* 49 (19): 4085–93.

Suzaki, Takuya, Momoyo Ito, Emiko Yoro, Shusei Sato, Hideki Hirakawa, Naoya Takeda, and Masayoshi Kawaguchi. 2014. "Endoreduplication-Mediated Initiation of Symbiotic Organ Development in Lotus Japonicus." *Development* , no. May: 1–5.

Svistoonoff, Sergio, Faiza Meriem Benabdoun, Mathish Nambiar-veetil, Leandro Imanishi, Jocelyne Bonneau, Luis Wall, Nadia Ykhlef, Charles Rosenberg, Didier Bogusz, and Claudine Franche. 2013. "The Independent Acquisition of Plant Root Nitrogen-Fixing Symbiosis in Fabids Recruited the Same Genetic Pathway for Nodule Organogenesis." *PloS One* 8 (5): e64515.

Svistoonoff, Sergio, Valérie Hocher, and Hassen Gherbi. 2014. “Actinorhizal Root Nodule Symbioses: What Is Signalling Telling on the Origins of Nodulation?” *Current Opinion in Plant Biology* 20 (August): 11–18.

Svistoonoff, Sergio, Laurent Laplaze, Florence Auguy, John Runions, Robin Duponnois, Jim Haseloff, Claudine Franche, and Didier Bogusz. 2003. “cg12 Expression Is Specifically Linked to Infection of Root Hairs and Cortical Cells during Casuarina Glauca and Allocasuarina Verticillata Actinorhizal Nodule Development.” *Molecular Plant-Microbe Interactions: MPMI* 16 (7): 600–607.

Swensen, Susan M. 1996. “The Evolution of Actinorhizal Symbioses: Evidence for Multiple Origins of the Symbiotic Association.” *American Journal of Botany* 83 (11): 1503–12.

Timmers, A. C., M. C. Auriac, and Georges Truchet. 1999. “Refined Analysis of Early Symbiotic Steps of the Rhizobium-Medicago Interaction in Relationship with Microtubular Cytoskeleton Rearrangements.” *Development* 126 (16): 3617–28.

Tóth, Katalin, Thomas F. Stratil, Esben B. Madsen, Juanying Ye, Claudia Popp, Meritxell Antolín-Llovera, Christina Grossmann, et al. 2012. “Functional Domain Analysis of the Remorin Protein LjSYMREM1 in Lotus Japonicus.” *PloS One* 7 (1). https://doi.org/10.1371/journal.pone.0030817.

Trinick, Michael J., David J. Goodchild, and Celia Miller. 1989. “Localization of Bacteria and Hemoglobin in Root Nodules of Parasponia Andersonii Containing Both Bradyrhizobium Strains and Rhizobium Leguminosarum Biovar Trifolii.” *Applied and Environmental Microbiology* 55 (8): 2046–55.

Trinick, M. J. 1979. “Structure of Nitrogen-Fixing Nodules Formed by Rhizobium on Roots of Parasponia Andersonii Planch.” *Canadian Journal of Microbiology* 25 (5): 565–78.

Trinick, M. J., and P. A. Hadobas. 1988. “Biology of the Parasponia-Bradyrhizobium Symbiosis.” *Plant and Soil* 110 (2): 177–85.

———. 1989. “Competition by Bradyrhizobium Strains for Nodulation of the Nonlegume Parasponia Andersonii.” *Applied and Environmental Microbiology* 55 (5): 1242–48.

Valverde, Claudio, and Luis Gabriel Wall. 1999. “Time Course of Nodule Development in the Discaria Trinervis (Rhamnaceae) - Frankia Symbiosis.” *The New Phytologist* 141 (2): 345–54.

Van Nguyen, Thanh, Daniel Wibberg, Kai Battenberg, Jochen Blom, Brian Vanden Heuvel, Alison M. Berry, Jörn Kalinowski, and Katharina Pawlowski. 2016. “An Assemblage of *Frankia* Cluster II Strains from California Contains the Canonical *nod* Genes and Also the Sulfotransferase Gene *nodH*.” *BMC Genomics* 17 (1): 796.

Van Zeijl, Arjan, Rik H. M. Op Den Camp, Eva E. Deinum, Tatsiana Charnikhova, Henk Franssen, Huub J. M. Op Den Camp, Harro Bouwmeester, Wouter Kohlen, Ton Bisseling, and René Geurts. 2015. “Rhizobium Lipo-Chitooligosaccharide Signaling Triggers Accumulation of Cytokinins in Medicago Truncatula Roots.” *Molecular Plant* 8 (8): 1213–26.

Venkateshwaran, Muthusubramanian, Dhileepkumar Jayaraman, Mireille Chabaud, Andrea Genre, Allison J. Balloon, Junko Maeda, Kari Forshey, et al. 2015. “A Role for the Mevalonate Pathway in Early Plant Symbiotic Signaling.” *Proceedings of the National Academy of Sciences of the United States of America* 112 (31): 9781–86.

Vernié, Tatiana, Jiyoung Kim, Lisa Frances, Yiliang Ding, Jongho Sun, Dian Guan, Andreas Niebel, Miriam L. Gifford, Fernanda de Carvalho-Niebel, and Giles E. D. Oldroyd. 2015. “The NIN Transcription Factor Coordinates Diverse Nodulation Programs in Different Tissues of the Medicago Truncatula Root.” *The Plant Cell*, tpc.15.00461.

Wang, Bin, Li Huey Yeun, Jia-Yu Xue, Yang Liu, Jean-Michel Ané, and Yin-Long Qiu. 2010. “Presence of Three Mycorrhizal Genes in the Common Ancestor of Land Plants Suggests a Key Role of Mycorrhizas in the Colonization of Land by Plants.” *The New Phytologist* 186 (2): 514–25.

Werner, Gijsbert D. A., William K. Cornwell, Johannes H. C. Cornelissen, and E. Toby Kiers. 2015. “Evolutionary Signals of Symbiotic Persistence in the Legume–rhizobia Mutualism.” *Proceedings of the National Academy of Sciences* 112 (33): 10262–69.

Werner, Gijsbert D. A., William K. Cornwell, Janet I. Sprent, Jens Kattge, and E. Toby Kiers. 2014. “A Single Evolutionary Innovation Drives the Deep Evolution of Symbiotic N2-Fixation in Angiosperms.” *Nature Communications* 5 (January): 4087.

Willmann, Roland, and Thorsten Nürnberger. 2012. “How Plant Lysin Motif Receptors Get Activated: Lessons Learned from Structural Biology.” *Science Signaling*. https://doi.org/10.1126/scisignal.2003274.

Wittenberg, J. B., B. A. Wittenberg, Q. H. Gibson, M. J. Trinick, and C. A. Appleby. 1986. “The Kinetics of the Reactions of Parasponia Andersonii Hemoglobin with Oxygen, Carbon Monoxide, and Nitric Oxide.” *The Journal of Biological Chemistry* 261 (29): 13624–31.

Xie, Fang, Jeremy D. Murray, Jiyoung Kim, Anne B. Heckmann, Anne Edwards, Giles E. D. Oldroyd, and J. Allan Downie. 2012. “Legume Pectate Lyase Required for Root Infection by Rhizobia.” *Proceedings of the National Academy of Sciences* 109 (2): 633–38.

Xue, De-Xing, Chun-Lian Li, Zhi-Ping Xie, and Christian Staehelin. 2019. "LYK4 Is a Component of a Tripartite Chitin Receptor Complex in Arabidopsis Thaliana." *Journal of Experimental Botany*, July. https://doi.org/10.1093/jxb/erz313.

Xue, Li, Haitao Cui, Benjamin Buer, Vinod Vijayakumar, Pierre-Marc Delaux, Stefanie Junkermann, Marcel Bucher, Nina Gerlach, Stefanie Junkermann, and Marcel Bucher. 2015. "Network of GRAS Transcription Factors Involved in the Control of Arbuscule Development in *Lotus Japonicus*." *Plant Physiology* 167 (March): 854–71.

Yang, M. Q., Robin Van Velzen, F. T. Bakker, and Ali Sattarian. 2013. "Molecular Phylogenetics and Character Evolution of Cannabaceae." *Taxon* 62: 473–85.

Yang, W. C., C. de Blank, I. Meskiene, H. Hirt, J. Bakker, A. van Kammen, H. Franssen, and T. Bisseling. 1994. "Rhizobium Nod Factors Reactivate the Cell Cycle during Infection and Nodule Primordium Formation, but the Cycle Is Only Completed in Primordium Formation." *The Plant Cell* 6 (10): 1415–26.

Yano, Koji, Satoko Yoshida, Judith Müller, Sylvia Singh, Mari Banba, Kate Vickers, Katharina Markmann, et al. 2008. "CYCLOPS, a Mediator of Symbiotic Intracellular Accommodation." *Proceedings of the National Academy of Sciences of the United States of America* 105 (51): 20540–45.

Yokota, K., T. Soyano, H. Kouchi, and M. Hayashi. 2010. "Function of GRAS Proteins in Root Nodule Symbiosis Is Retained in Homologs of a Non-Legume, Rice." *Plant & Cell Physiology* 51 (9): 1436–42.

Yoon, Hwi Joong, Md Shakhawat Hossain, Mark Held, Hongwei Hou, Marilyn Kehl, Alexandre Tromas, Shusei Sato, et al. 2014. "Lotus Japonicus SUNERGOS1 Encodes a Predicted Subunit A of a DNA Topoisomerase VI That Is Required for Nodule Differentiation and Accommodation of Rhizobial Infection." *The Plant Journal: For Cell and Molecular Biology* 78 (5): 811–21.

Yoro, Emiko, Takuya Suzaki, Koichi Toyokura, Hikota Miyazawa, Hidehiro Fukaki, and Masayoshi Kawaguchi. 2014. "A Positive Regulator of Nodule Organogenesis, NODULE INCEPTION, Acts as a Negative Regulator of Rhizobial Infection in Lotus Japonicus." *Plant Physiology* 165 (2): 747–58.

Young, Nevin D., Frédéric Debellé, Giles E. D. Oldroyd, Rene Geurts, Steven B. Cannon, Michael K. Udvardi, Vagner a. Benedito, et al. 2011. "The *Medicago* Genome Provides Insight into the Evolution of Rhizobial Symbioses." *Nature* 480 (7378): 520–24.

Yue, Jipei, Xiangyang Hu, Hang Sun, Yongping Yang, and Jinling Huang. 2012. "Widespread Impact of Horizontal Gene Transfer on Plant Colonization of Land." *Nature Communications* 3: 1152.

Zeijl, Arjan van, Wei Liu, Ting Ting Xiao, Wouter Kohlen, Wei-Cai Yang, Ton Bisseling, and Rene Geurts. 2015. "The Strigolactone Biosynthesis Gene *DWARF27* Is Co-Opted in Rhizobium Symbiosis." *BMC Plant Biology* 15 (1): 260.

Zeijl, Arjan van, Titis A. K. Wardhani, Maryam Seifi Kalhor, Luuk Rutten, Fengjiao Bu, Marijke Hartog, Sidney Linders, et al. 2018. "CRISPR/Cas9-Mediated Mutagenesis of Four Putative Symbiosis Genes of the Tropical Tree Parasponia Andersonii Reveals Novel Phenotypes." *Frontiers in Plant Science* 9 (March): 1–14.

Zgadzaj, Rafal, Ruben Garrido-Oter, Dorthe Bodker Jensen, Anna Koprivova, Paul Schulze-Lefert, and Simona Radutoiu. 2016. *Root Nodule Symbiosis in Lotus Japonicus Drives the Establishment of Distinctive Rhizosphere, Root, and Nodule Bacterial Communities*. Vol. 113.

Zhang, Xue-Cheng, Steven B. Cannon, and Gary Stacey. 2009. "Evolutionary Genomics of LysM Genes in Land Plants." *BMC Evolutionary Biology* 9 (1): 183.

CHAPTER 2

2

Commonalities in symbiotic plant-microbe signalling

Rens Holmer[1], Luuk Rutten[1], Wouter Kohlen, Robin van Velzen & Rene Geurts

Wageningen University, Department of Plant Science, Laboratory of Molecular Biology, Droevendaalsesteeg 1, 6708PB Wageningen, The Netherlands
[1]authors contributed equally.

This chapter is published by Elsevier in: Advances in Botanical Research, Volume 82; How plants communicate with their biotic environment, Chapter 8, pages 187-221

Abstract

Plants face the problem that they have to discriminate symbionts from a diverse pool of soil microbes, including pathogens. Studies on different symbiotic systems revealed commonalities in plant-microbe signalling. In this chapter we focus on four intimate symbiotic interactions: two mycorrhizal ones, with arbuscular- and ectomycorrhizal fungi, and two nitrogen-fixing ones, with rhizobium and *Frankia* bacteria. Comparing these systems uncovered commonalities in the way plants attract their symbiotic partners. Especially flavonoids, and in a lesser extent strigolactones, are pivotal plant signals that are perceived by the microsymbiont. In response, signal molecules are exuded by the microbes to trigger symbiotic responses in their host plant. Strikingly, microbes that establish an endosymbiotic relation with their host plant, namely arbuscular mycorrhizal fungi, rhizobium and *Frankia* bacteria, make use of a symbiotic signalling network that is highly conserved in plants. The use of flavonoids as attractants for symbiotic microbes, in combination with the use of a common plant signalling network to establish endosymbioses, raises questions about how plants manage to discriminate their microbial partners.

1. Introduction

High throughput sequencing approaches have uncovered an overwhelming diversity of soil microbes. Plants affect this microbial community - directly or indirectly - with their root systems. For example, roots exude substantial amounts of organic and amino acids, polymerized sugars (e.g. mucilage) as well as release border cells and dead root cap cells, which all form a nutrient source for many microbes (Jones, Nguyen, & Finlay, 2009). On top of that more specific secondary metabolites are exuded that manipulate the microbial community by acting as antimicrobial agents or as attractants. Conversely, soil microbes can affect plant growth. For example, microbes can promote plant growth by improving nutrient availability, or inducing resistance against biotic and abiotic stresses (Coleman-Derr & Tringe, 2014; Mendes, Garbeva, & Raaijmakers, 2013). On the other hand, pathogenic microbes can induce resource loss and disease. In this complex plant root microbiome network the plant must therefore discriminate between bacteria and fungi that provide an advantage and those that act as commensals or even pathogens. In this chapter we will focus on the molecular communication in a symbiotic context which occurs in plant roots and the rhizosphere. Plants establish several intricate long-term mutualistic relationships with microbes that are hosted intercellularly (ecto) or intracellularly (endo). Here, we will discuss the commonalities of four intimate symbiotic interactions. Thereby we will focus on two key stages of the interaction: attraction of the microbial partner, and subsequent microbe-induced signalling to establish a symbiosis.

2. Intimate plant root-microbe symbioses

Plant root symbiosis occur at different levels of engagement, ranging from loosely attached microbes that provide a certain advantage to the plant to bacteria that are intracellularly accommodated as organelle-like structures (Mendes et al., 2013; Van Loon, 2007). The best studied plant root symbioses are those with arbuscular mycorrhizal and ectomycorrhizal fungi, and those with rhizobium and *Frankia* nitrogen fixing bacteria, together encompassing a diverse range of plant and microbial species.

2.1 Mycorrhizal symbiosis

Mycorrhizal symbioses - the symbiotic interactions between some soil fungi and plant roots - can occur in several forms. Of these the ancient arbuscular (endo-) mycorrhiza and the much younger forms of ectomycorrhiza are best studied. Based on fossil records arbuscular mycorrhizal symbiosis is estimated to be at least ~400-460 million years old, and evolved in a period that coincides with colonization of terrestrial habitats by plants (Redecker, Kodner, & Graham, 2000; Remy, Taylor, Hass, & Kerp, 1994; Simon, Bousquet, Levesque, & Lalonde, 1993). Still today the vast majority of land plant species establish an arbuscular mycorrhizal symbiosis, underlining the ecological importance of this interaction (Wang & Qiu, 2006). The fungi that establish an arbuscular mycorrhizal symbiosis belong to a distinct taxonomic phylum, the Glomeromycota. This phylum possibly represents more than 1000 species, though only less than 300 have been characterized to a certain level of detail (Redecker et al., 2013). Arbuscular mycorrhizal fungi are obligate biotrophs. Their hyphae penetrate plant roots intercellularly and form intracellular feeding structures - called arbuscules - in root cortical cells **(Figure 1A)**. Arbuscules are surrounded by a plant-derived membrane, but are largely deprived of plant cell wall material (Raffaella Balestrini & Bonfante, 2014). At this symbiotic interface nutrients are exchanged. Minerals - especially phosphates and nitrates - taken up by the fungal extraradical mycelium are delivered to the plant in return for carbohydrates. Arbuscules remain functional for several days, after which they collapse and disappear, leading to a reversion of the plant cell to its asymbiotic cortical fate.

Ectomycorrhizal symbiosis can occur between diverse groups of plant and fungal species, as a result of several independent evolutionary events (Martin et al., 2016). Overall this type of symbiosis can occur in about 2% of all land plants, including all dominant tree species in temperate forests, such as pines (*Pinus*), Douglas firs (*Pseudotsuga*), oaks (*Quercus*), willows (*Salix*), beeches (*Fagus*) and birches (*Betula*) (Smith & Read, 2008; Tedersoo et al., 2010). Ectomycorrhizal fungi belong to several taxonomic phyla including Basidiomycota, Ascomycota, and Zygomycota (Tedersoo & Smith, 2013) which are all closely related to species with

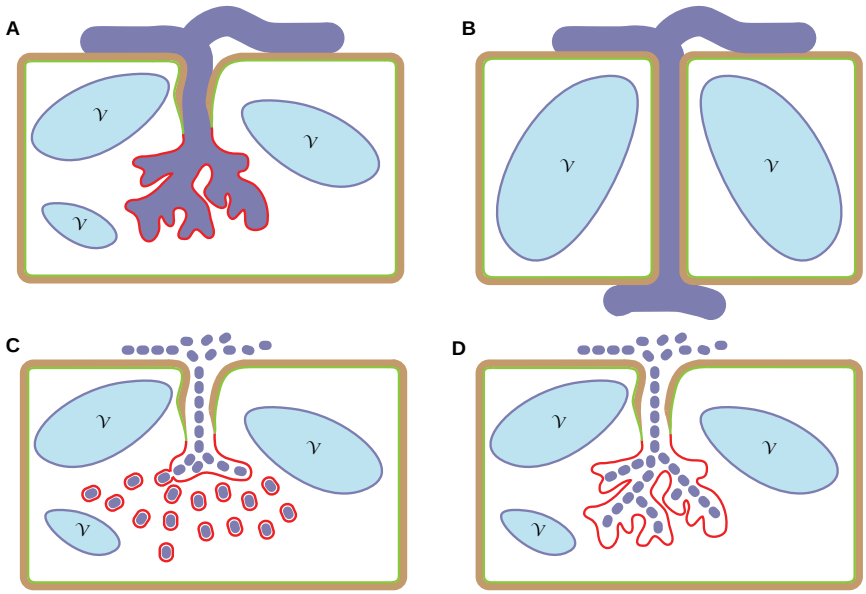


Figure 1. Schematic representation of the cellular mode of infection of the four symbioses discussed in this chapter. Green lines indicate plant cell membrane, red lines indicate the plant cell membrane-derived symbiotic interface in the form of the periarbuscular membrane for AM, symbiosomes for rhizobium/legumes, and fixation threads for *Frankia*/actinorhizal plants and for rhizobium/*Parasponia*. **(A)** Hyphae of an endomycorrhizal fungus penetrating the cell and forming a feeding structure named arbuscule. Arbuscules are not surrounded by a plant-derived cell wall. **(B)** Hyphae of an ectomycorrhizal fungus growing intercellularly. **(C)** Rhizobium bacteria released within transient organelle-like structures - named symbiosomes - in nodule cells of most legumes. **(D)** *Frankia* in actinorhizal plants and some rhizobia in parasponia and in some basal legumes infect cells of nodules through fixation threads. Fixation threads are largely deprived of plant cell wall. The bacteria in fixation threads remain in contact with the apoplast. blue: vacuole (v); purple: microbe; brown: plant cell wall; green: plant plasma membrane; red: plant-derived endosymbiotic membrane.

a saprotrophic lifestyle. Saprophytic fungi have an extensive repertoire of genes encoding degrading enzymes that can effectively mobilize resources, in particular nitrogen and phosphorus, from a variety of organic substrates (Plett & Martin, 2011). However, compared to their saprophytic sister clades, ectomycorrhizal fungi only have a limited set of genes encoding plant cell wall–degrading enzymes (e.g., pectin lyases and pectinases) (Kohler et al., 2015). Nevertheless, expression studies indicate that some of the plant cell wall-degrading enzymes that have been maintained may function during symbiosis (Balestrini & Bonfante, 2005).

Despite the diverse and paraphyletic groups of plant and fungal species that establish ectomycorrhizal symbioses, there is a remarkable resemblance in symbiotic phenotypes. The fungi preferentially colonise newly formed lateral roots. Upon hyphal attachment, they modulate root growth which allows them to colonise the

root apoplast, forming a so-called Hartig net. The penetration depth of the hyphae is variable, but it typically comprises several layers of cortical cells, excluding the endodermis. In contrast to the endomycorrhizal symbiosis root cells are not invaded intracellularly **(Figure 1B)**. Ultimately, many fungal hyphae cover the root surface forming a thick, multi-layered 'mantle', insulating the infected lateral root. The molecular mechanisms underlying Hartig net development remain elusive. Recent efforts, using a combination of genome sequencing and reverse genetic studies, have provided new insights in the early symbiosis signalling. For example in the fungus *Laccaria bicolor* it was found that an aquaporin (LbAQP1) is essential for Hartig net development and the expression of effector genes. *LbAQP* expression is triggered upon direct root contact and functions as a transport facilitator for plant signalling molecules, most likely H_20_2, NO or CO_2. (Navarro-Rodenas, Xu, Kemppainen, Pardo, & Zwiazek, 2015). Additionally *L. bicolor* produces auxin (IAA) in its mycelium that triggers auxin-related responses in the plant root (Vayssières et al., 2015). This finding is in line with pioneer work that showed that increased mycorrhizal activity is associated with increased auxin biosynthesis by the fungus (Gea, Normand, Vian, & Gay, 1994). Together, these studies make clear that plant-fungal signalling intertwines with plant auxin homeostasis and possibly used reactive oxygen species to establish a symbiotic interaction.

2.2 Nitrogen fixing endosymbioses

A selective, though diverse, group of plant species is able to establish an endosymbiosis with nitrogen-fixing (diazotrophic) bacteria. These bacteria belong either to the genus *Frankia* or to the paraphyletic group of bacteria known as rhizobia. *Frankia* and Rhizobia strains gained the symbiotic trait by horizontal gene transfer. The genus *Frankia* is a diverse assemblage of filamentous sporangia-forming actinobacterial species that can be saprophytic, facultative symbiotic, or obligate symbiotic. The *Frankia* genus can be separated in four separate bacterial clusters based on phylogenetic analysis, with only three of them that can establish symbiosis (Gtari, Tisa, & Normand, 2013). Plant species that can form a nitrogen-fixing endosymbiosis with *Frankia* bacteria (~230 species known as actinorhizal plants) are dispersed over 25 genera and 8 taxonomic families, suggesting multiple evolutionary origins of this symbiosis (Pawlowski & Demchenko, 2012).

Relative to *Frankia*, rhizobia are even more diverse, representing 15 genera in eight families of α-, β- and γ-Proteobacteria (Remigi, Zhu, Young, & Masson-Boivin, 2015). Nitrogen-fixing symbiosis with rhizobia is prominent in the legume family (*Fabaceae*), but can also occur in *Parasponia*, a genus in the Cannabis family (*Cannabaceae*). Based on the phylogenetic distance between *Fabaceae* and *Cannabaceae* it is most probable that - similarly to the actinorhizal symbiosis - there are multiple origins for rhizobial symbiosis (Behm, Geurts, & Kiers, 2014; Geurts, Lillo, & Bisseling,

2012). The formation of specific nodule-like structures (root nodules) by the host plant, in which the bacteria proliferate and fix nitrogen, is common for both types of endosymbiosis with diazotrophic bacteria.

The reason for this may be that rhizobia and *Frankia* bacteria are generally not able to infect differentiated cells of the plant root. Only cells of the future nodule that are mitotically activated by the microsymbiont can be infected, suggesting that these cells are developmentally reprogrammed (Geurts, Xiao, & Reinhold-Hurek, 2016). The nodules are optimized to facilitate growth of the microbial partner, which, once inside nodule cells, differentiates in its symbiotic form and fixes atmospheric nitrogen into ammonia in exchange for carbohydrates.

Variation exists in the way the nitrogen fixing bacterial partner is hosted. In most legume nodules, rhizobia are hosted in transient organelle-like structures, called symbiosomes. Symbiosomes are released from intracellular infection threads that have guided the rhizobium bacteria from the epidermis towards the newly formed nodule. Hundreds of symbiosomes surrounded by a plant-derived membrane, often containing only one bacterium, can be present in a single nodule cell. This membrane forms a symbiotic interface where nutrient exchanges take place between the bacteria and the cytoplasm of the host cell **(Figure 1C)**. In *Parasponia* and actinorhizal plants symbiosomes are not formed. Instead, the bacteria remain in thread-like structures, known as fixation threads **(Figure 1D)**. Fixation threads differ from the penetrating infection thread by a reduction of plant cell wall material. Fixation threads also occur in a few legume species and may represent a more ancestral form of bacterial endosymbiosis than symbiosomes (Behm et al., 2014).

Of the four intimate symbiotic interactions that are central here, three have evolved more than once: the symbioses with rhizobia, *Frankia* and ectomycorrhizal fungi. This suggests an evolutionary advantage of root symbiosis for both partners. Interestingly, several studies indicate that similar mechanisms have been co-opted in all four symbiotic interactions. Below we will discuss the commonalities in signalling mechanisms between the four symbioses central in this chapter.

3. Recognition and attraction of symbiotic partners

As outlined above, not all plants are able to form an intricate root microbe symbiosis; nor are all soil microbes symbiotic. Consequently, symbiotic partners need to recognize each other. Microbes recognize potential host plants by root exudates. Indeed, plants can exude signalling molecules to attract their symbiotic partner. Common signals in symbiotic partner recognition are exuded flavonoids, which play a role in all four symbioses. In addition, it was noted that exuded strigolactones can act as signal molecules, especially in arbuscular mycorrhizal symbiosis. Strikingly, both types of molecules function also as endogenous plant signals.

3.1 Flavonoids induce microbial responses

Flavonoids are a subclass of plant polyphenolic compounds and are a major class of secondary metabolites. As is typical for plant secondary metabolites, flavonoids are diverse: ~9000 chemical structures have so far been reported (Ferrer, Austin, Stewart, & Noel, 2008). Flavonoids are synthesized through the phenylpropanoid pathway. A chalcone synthase produces the chalcone scaffolds from which all other flavonoids are derived and is the first enzyme specific for flavonoid production (Falcone Ferreyra, Rius, & Casati, 2012). A series of enzymatic reactions can alter the chalcone scaffold into a huge diversity of compounds. Flavonoids are typically categorized in subclasses based on these enzymatic reactions. The major subclasses of flavonoids include phlobaphenes, flavones, flavanones, flavonols, aurones, isoflavonoids, anthocyanins and condensed tannins. Many flavonoids are known to have glycosylated forms, i.e. quercitrin is formed by the addition of the deoxy sugar rhamnose to the flavonol quercetin, whereas rutin is formed by the addition of the disaccharide rutinose. Such small modifications can have drastic consequences for the observed effects in symbioses.

The involvement of flavonoids in symbioses has been described for all four types of symbiosis discussed in this review (arbuscular mycorrhizal, ectomycorrhizal, rhizobial and actinorhizal symbioses). For two compounds positive effects in all four symbioses have been described **(Figure 2).** Naringenin positively influences

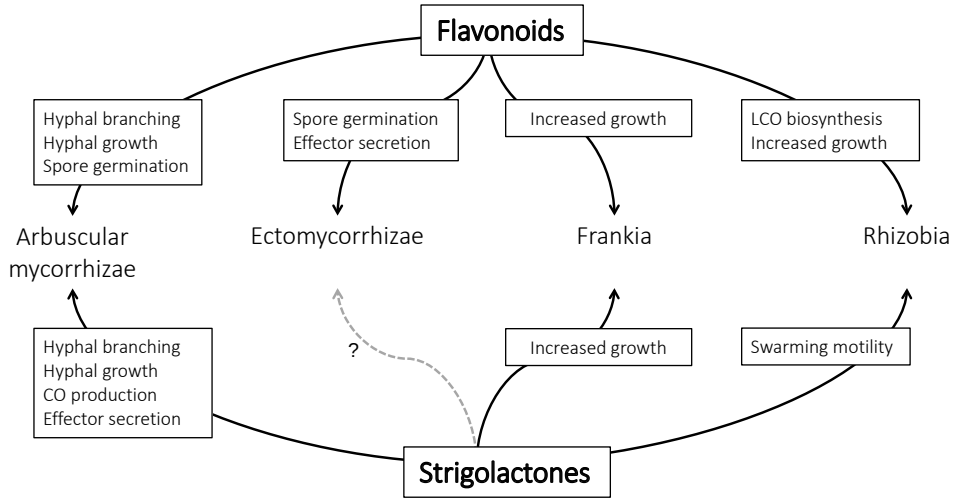


Figure 2. Flavonoids and Strigolactones are generic attractants for microsymbionts. Increased growth of *Frankia* can possibly be attributed to either flavonoids, strigolactones, or even other components as total root exudates were used to demonstrate this (Beauchemin et al., 2012). COs = Chito-Oligosaccharides, LCOs = Lipo-Chito-Oligosaccharides.

arbuscular mycorrhizal colonization (Garg & Singla, 2016) and rhizobium symbiosis (Weston & Mathesius, 2013), enhances spore germination of the ectomycorrhizal fungus *Suillis bovinus* (Kikuchi, Matsushita, Suzuki, & Hogetsu, 2007) and restores *Frankia* nodulation in a chalcone synthase mutant of the actinorhizal plant *Casuarina glauca* (Abdel-Lateif et al., 2013). Quercetin has been reported to stimulate spore germination, hyphal branching and growth of arbuscular mycorrhizal fungi (Bécard, Douds, & Pfeffer, 1992; Tsai & Phillips, 1991), the growth rate of rhizobium bacteria (Hartwig, Joseph, & Phillips, 1991), the actinorhizal nodulation (Sayed & Wheeler, 1999) and it also stimulates the production of the symbiotic effector protein MiSSP7 in the ectomycorrhizal fungus *L. bicolor* (Plett & Martin, 2012).

The molecular mode of action of naringenin and quercetin is not always known. Best studied is the effect of naringenin - and other flavonoids - in rhizobia, where flavonoids target NodD proteins. Rhizobial NodD proteins belong to the class of LysR-type transcriptional regulators that are activated upon the binding of external signals (Honma, Asomaning, & Ausubel, 1990; Mulligan & Long, 1989). Binding of a flavonoid molecule causes a conformational change which results in an increased binding affinity for specific cis regulatory elements. In case of NodD this element is known as the nod box (Chen et al., 2005). Rhizobia generally have several operons that contain a nod box in their promoter region. Most prominent are the genes encoding an ABC transporter (NodI and NodJ) and 3 genes encoding the enzymes N-acetylglucosaminyltransferase (NodC), a chitooligosaccharide deacetylase (NodB) and a N-acyltransferase (NodA). These proteins are essential for biosynthesis and secretion of lipo-chitooligosaccharide molecules (LCOs), which act as potent symbiotic signal molecules (see section 4) (Limpens, van Zeijl, & Geurts, 2015; G. E. D. Oldroyd, 2013).

In case of arbuscular mycorrhizae and *Frankia* it remains elusive whether flavonoids trigger biosynthesis of similar symbiotic signalling molecules, despite the fact that flavonoids have a positive effect on both symbioses (Auguy et al., 2011; Garg & Singla, 2016). LCOs and short chain chitin oligomers (tetra and pentameric COs) have been shown to be produced by the mycorrhizal fungus *Rhizophagus irregularis*, but their biosynthetic pathways have not yet been uncovered (Lin et al., 2014; Maillet et al., 2011; Genre et al 2013; Tisserant et al., 2013). In case of symbiotic *Frankia* species, LCO biosynthesis genes are not common, and only found in a representative of a relatively isolated taxonomic lineage (cluster 2): namely (candidatus) *Frankia datiscae* strain DG1 (Persson et al., 2015). For this strain it was found that *nodA, nodB, nodC, nodI* and *nodJ* are expressed when the bacteria occupy *Datisca glomerata* root nodules (Beauchemin et al., 2012). Therefore, it is most probable that *F. datiscae* LCO signals play a symbiotic role.

Other flavonoids have been described to be involved in one or a few of the discussed symbioses, but were never tested in the other types of symbiosis. Nevertheless, these observations can shed an interesting light on the symbiotic role of flavonoids. Especially interesting is the described host specificity in the legume rhizobia interaction (Reddy, María, & Soto, 2007), which in part is determined by recognition of specific flavonoids. Whereas a specific flavonoid can induce expression of the LCO biosynthetic *nodABC* operon in one bacterium, the same compound can have a negative effect in another bacterium. For example, the flavonoid coumestrol positively influences the symbiosis between *Glycine max* and *Sinorhizobium fredii* USDA191 (Kosslak, Bookland, Barkei, Paaren, & Appelbaum, 1987) but negatively influences the symbiosis between *Medicago sativa* and *Sinorhizobium meliloti* 1021 (Zuanazzi et al., 1998). In this context, it is also relevant to note that the composition of root exudates may vary depending on the developmental stage of the root. For example, studies in *Medicago sativa* indicate that flavonoids with a positive effect on the symbiosis are exuded in the elongation and differentiation zone that is susceptible to rhizobium infection. Whereas repelling flavonoids are exuded in the adjacent regions of the root (i.e. the root tip and the more mature part of the root) (Peters & Long, 1988; Zuanazzi et al., 1998). However, such studies have not been further extended to see whether similar mechanisms exist in other nodulating taxa.

In addition to direct application of flavonoids to microbial cultures, reverse genetic studies in plants have been conducted. In most studies chalcone synthase genes were targeted. Chalcone synthase knockdown in actinorhizal *Casuarina glauca*, or in the legume model *Medicago truncatula* results in impaired nodulation (Abdel-Lateif et al., 2013; Wasson, Pellerone, & Mathesius, 2006). In both plant systems this phenotype can be restored by the application of naringenin. In contrast, no effect was reported on the arbuscular mycorrhizal symbiosis when using a chalcone-synthase double-mutant in maize (Becard, Taylor, Douds, Pfeffer, & Doner, 1995). The fact that a chalcone synthase maize mutant can be normally mycorrhizal with different fungal species demonstrates that flavonoids are not essential signals for this symbiosis. Nevertheless, flavonoids may act as facultative signals, and may play a role in host selection, by activating certain fungi over others (Ellouze et al., 2012).

The importance of flavonoids in root nodule symbiosis may be the result of the fact that flavonoids also act as endogenous plant signals that control auxin transport (Brown et al., 2001; Mathesius et al., 1998; Wasson et al., 2006). Based on quantitative modeling and experimental studies it is hypothesized that a transient decrease in auxin efflux can lead to formation of a local auxin maximum, which is the onset of nodule development (Deinum, Geurts, Bisseling, & Mulder, 2012; Hirsch, Bhuvaneswari, Torrey, & Bisseling, 1989). Such a function of flavonoids in nodulation is supported by the finding that naringenin can restore nodulation in

the *Medicago truncatula* cytokinin signalling mutant *Mtcre1* (Ng, Perrine-Walker, Wasson, & Mathesius, 2015). This study demonstrates that naringenin not only acts as an attractant of symbiotic microbes, but also functions as an endogenous plant signal, which - in a symbiotic context - acts downstream, or in parallel, to rhizobium-induced cytokinin signalling.

3.2 Dual Role of Strigolactones

Strigolactones are carotenoid-derived terpenoid lactones, often composed of four rings. Three rings form a tricyclic lactone, which is connected to the fourth butenolide ring via an enol ether bridge (Pandey, Sharma, & Pandey, 2016). Strigolactones are known as endogenous plant hormones that control several steps in plant development (Brewer, Koltai, & Beveridge, 2013). Over the last decade major advances have been made on the identification of the strigolactone biosynthesis and perception pathway. A carotenoid isomerase (named DWARF27 (D27) in most species), two carotenoid cleaving dioxygenases (named CCD7 and CCD8), and a cytochrome P450 (possibly MAX1 in *Arabidopsis thaliana*) are sequentially required to produce the strigolactone backbones: either 4-deoxyorobanchol or 5-deoxystrigol (Alder et al., 2012; Y. Zhang et al., 2014). It is postulated that this backbone can be further decorated to produce the wealth of different strigolactone metabolites (Al-Babili & Bouwmeester, 2015). In plants the strigolactone receptor was identified as a α/β hydrolase (named OsDWARF14 (D14) in rice *(Oryza sativa*) and AtDAD2 in *Arabidopsis thaliana*). Together with a specific F-box protein named OsDWARF3/AtMAX2 it forms the SCF E3 ubiquitin ligase complex required for strigolactone signalling (Hamiaux et al., 2012; Zhao et al., 2015).

The discovery that strigolactones stimulate hyphal branching in the arbuscular mycorrhizal fungus *Gigaspora margarita* (Akiyama, Matsuzaki, & Hayashi, 2005) has launched an interest in the involvement of these compounds in symbiotic signalling. The observation that strigolactones induce hyphal branching in *G. margarita* - at very low concentrations - has led to the hypothesis that the induction of hyphal branching must be receptor-mediated (Akiyama, Ogasawara, Ito, & Hayashi, 2010). Furthermore, it was found that the synthetic strigolactone analog GR24 triggers mitochondrial activity in the arbuscular mycorrhizal fungi *Rhizophagus intraradices* and *Gigaspora rosea* (Besserer et al., 2006). However, it should be noted that in order to induce hyphal branching in G. rosea besides GR24, also the flavonoid quercetin is needed in the fungal growth medium (Besserer, Bécard, Jauneau, Roux, & Séjalon-Delmas, 2008). As quercetin is known to stimulate arbuscular mycorrhizal growth, hyphal branching and spore germination (Tsai & Phillips, 1991) this suggests that with this specific fungus strigolactones alone might not be sufficient to induce hyphal branching. In an independent experiment increased production of short chain COs upon application of GR24 was reported for *Rhizophagus irregularis* (Genre et

al., 2013) (Figure 2). In addition, a putative effector protein (RiSIS1) was identified in a screening of upregulated genes in GR24-treated *Rhizophagus irregularis* (Tsuzuki, Handa, Takeda, & Kawaguchi, 2016). Using host-induced gene silencing the RiSIS1 gene was knocked down during infection, which resulted in significant suppression of colonization and stunted arbuscules. This suggests that RiSIS1 is a strigolactone-induced effector protein.

Application of GR24 to four ectomycorrhizal species revealed no effect on hyphal branching (Steinkellner et al., 2007) (Figure 2). This suggests strigolactones play a less important or different role in this type of symbiosis. In contrast, a negative effect of GR24 was observed on growth and branching of a range of phytopathogenic fungi (Dor, Joel, Kapulnik, Koltai, & Hershenhorn, 2011), including species previously found not to be affected by GR24 (Steinkellner et al., 2007). It should be noted that these effects were only observed when relatively high concentrations of GR24 were used (Dor et al., 2011) and as such it remains unclear whether these concentrations were biologically relevant.

Apart from the beneficial effects in arbuscular mycorrhizal symbiosis, several studies revealed effects of strigolactones in the rhizobium/legume symbiosis **(Figure 2).** Exogenous application of GR24 increases *Medicago sativa* nodule number when inoculated with *Sinorhizobium meliloti* (Soto et al., 2010). Interestingly, the same study reports that the bacterial growth and *nodC* expression are not affected by GR24, leading the authors to hypothesize that the effect of GR24 is on the plant. However, more recently it was suggested that GR24 might affect *Sinorhizobium meliloti* by promoting bacterial swarming motility (Peláez-Vico, Bernabéu-Roda, Kohlen, Soto, & López-Ráez, 2016). In an independent experiment in Medicago truncatula low concentrations (0.1 µM) of GR24 also resulted in increased nodule numbers, but higher concentrations (2-5 µM) resulted in reduced nodule numbers and lateral root density (De Cuyper et al., 2014). Taken together this suggests that strigolactones act mainly as plant hormones involved in developmental programs during rhizobial symbiosis. In line with this, the strigolactone biosynthesis gene *MtD27* was shown to be inducible by rhizobium LCOs three hours after application and that this induction is regulated by the common symbiotic signalling pathway (Van Zeijl et al., 2015).

Mutants and knockdown experiments of strigolactone biosynthesis genes in several species shed light on the dual role of strigolactones in symbioses. Whereas often symbiotic phenotypes are observed, it is not trivial to decide whether these phenotypes are an effect of a change in direct signalling between host and symbiont, or whether a change in hormonal balance causes a difference in plant developmental program. Carotenoid cleavage dioxygenases (CCD) are among the most studied strigolactone biosynthetic enzymes in a symbiotic context. Mutation or knockdown of *ccd7/8* in

several plant species results in reduced mycorrhizal colonization (Gomez-Roldan et al., 2008; Kohlen et al., 2012; Kretzschmar et al., 2012; J. Liu et al., 2013; Vogel et al., 2010). In addition, nodulation was reported to be impaired in the *Lotus japonicus CCD7* knockdown and both *ccd7* and *ccd8* mutants of pea (Foo, Yoneyama, Hugill, Quittenden, & Reid, 2013; J. Liu et al., 2013). The importance of strigolactones in mycorrhizal colonization is further supported by the identification of a strigolactone transporter in *Petunia x hybrida*. The knockout of the ABC transporter PhPDR1 resulted in significantly reduced orobanchol levels in root exudates, which had effects on *Gigaspora margarita* and *Rhizophagus irregularis* mycorrhization efficiency. Plants show reduced colonization due to reduced mycorrhizal growth, branching and spore germination (Kretzschmar et al., 2012). The GRAS transcriptional regulators NSP1 and NSP2 were identified as regulators of strigolactone biosynthesis in rice and *Medicago truncatula* by regulating D27 expression (W. Liu et al., 2011). *Medicago truncatula nsp1* and *nsp2* mutants are not capable of forming nodules (Catoira et al., 2000; G. E. Oldroyd & Long, 2003). The *nsp1* mutant and the *nsp1/nsp2* double mutant produce no detectable amounts of strigolactones, whereas the *nsp2* mutant has a reduced and different strigolactone composition. Interestingly, mycorrhizal colonization of the *nsp1/nsp2* double mutant by *Rhizophagus irregularis* was only mildly reduced (W. Liu et al., 2011). In addition, the *Lotus japonicus nsp1* mutant is unable to form nodules, however infection by the arbuscular mycorrhizal fungus *Rhizophagus irregularis* was unaffected (Heckmann et al., 2006).

The rice and pea F-box mutants *Osd3/Psrms4* are markedly reduced in mycorrhizal colonization (Foo et al., 2013; Yoshida et al., 2012). This suggests that strigolactone perception *in planta* plays a role in AM colonization. Strikingly, in the pea *Psrms4* mutant nodule numbers are increased (Foo et al., 2013), indicating that the effect of strigolactones in nodulation is regulated differently compared to mycorrhization.

Interestingly, a severe mycorrhization phenotype in rice could be complemented by introducing a copy of *OsD14-LIKE* gene (Gutjahr et al., 2015). *OsD14-Like* is paralogous to *OsD14* and has strong similarities with the *Arabidopsis thaliana* karrikin receptor *AtKAI2*, which is responsible for detecting the smoke compound karrikin. OsD14 and OsD14-LIKE have been reported to have partially overlapping, but also distinct, functions for strigolactone and karrikin responses, as the *Atkai2* mutants are insensitive to karrikins but weakly responsive to strigolactones (Scaffidi et al., 2014). In addition, it was recently demonstrated that in *Arabidopsis thaliana* AtD14 and AtD14-like have different affinities for specific strigolactone stereoisomers (Scaffidi et al., 2014). This could indicate that the perception of specific strigolactones is regulated by multiple receptor complexes.

Taken together, the involvement of strigolactones in arbuscular mycorrhiza symbiosis is relatively well described, although several details remain unclear. A possible involvement in rhizobial symbiosis is just starting to be discovered. However, given the distinct nature of both symbioses, the mechanisms by which strigolactones function are likely different between the two. For ectomycorrhizal and actinorhizal symbioses no clear data on the involvement of strigolactones is available yet. As strigolactones are plant hormones involved in key developmental processes it is not surprising that ectomycorrhizal hosts were found to possess the strigolactone biosynthetic genes (Garcia, Delaux, Cope, & Ane, 2015).

4. A Conserved signalling pathway for Endosymbioses

As mentioned above, arbuscular mycorrhizal fungi, rhizobia and some basal *Frankia* species produce LCO signals in a symbiotic context, whereas no evidence has been found that LCO signals are playing a role in ectomycorrhizal symbiosis. This suggests that LCO signalling is a feature of microbes that establish an endo- rather than an ectosymbiosis.

LCOs are prominent signal molecules that are perceived by the host plant and set in motion symbiotic responses. Genetic studies in legumes, rice, *Parasponia* and the actinorhizal plant species *Datisca glomerata* (nodulated by *Frankia* sp. harboring LCO biosynthesis genes), but also in *Casuarina glauca*, a species that is nodulated by *Frankia* spp. that lack LCO biosynthesis genes uncovered a common symbiotic signalling network. This conserved symbiotic network stretches from transmembrane receptor kinases to a network of transcription factors that control the readout of symbiotic signalling (G. E. D. Oldroyd, 2013). A hallmark of endosymbiotic signalling is the induction of regular oscillations of the nuclear calcium concentration. To achieve this a complex of nuclear envelope-localized proteins are essential, including a potassium-permeable channel (encoded by *MtDMI1, LjCASTOR, LjPOLLUX*), a cyclic nucleotide–gated calcium channel, and a calcium ATPase (Capoen et al., 2011; Charpentier et al., 2016; Imaizumi-Anraku et al., 2005; Lévy et al., 2004). The induced calcium oscillations are decoded by a calcium-/calmodulin-dependent kinase (CCaMK), which is the onset of a transcriptional network (Soyano & Hayashi, 2014). Besides some common elements, like the CCaMK interacting transcription factor LjCYCLOPS, the activated network varies between arbuscular mycorrhizal and root nodule symbioses. For example, activation of the *NIN* transcription factor is essential for root nodule formation in legumes and *Casuarina glauca*, whereas it is not for arbuscular mycorrhizal symbiosis (Clavijo et al., 2015; Marsh et al., 2007; Schauser, Roussis, Stiller, & Stougaard, 1999). Conversely, arbuscular mycorrhizal symbiosis requires activation of GRAS transcription regulators such as *MtRAM1* in *Medicago*, which is not essential for root nodule formation (Gobbato et al.,

2012). Despite this divergence in transcriptional responses, the common symbiotic signalling genes are conserved in angiosperm and gymnosperm species that form an arbuscular mycorrhizal symbiosis. By contrast, plants that exclusively establish an ectomycorrhizal symbiosis - e.g. *Pinaceae* species - have lost several of these genes (Garcia et al., 2015). This supports the idea that ectomycorrhizal symbioses are founded on different signalling cues than arbuscular mycorrhizal and root nodule endosymbioses.

4.1 LCO Signalling

Most comprehensive studies on symbiotic signalling have been done in the legume model systems *Lotus japonicus* and *Medicago truncatula*. Both species have evolved to interact with a specific rhizobial species (*Mesorhizobium loti* for *Lotus japonicus* and *Sinorhizobium meliloti* for *Medicago truncatula*). By using these symbiotic models it was revealed that rhizobium LCOs are specifically recognized by a heteromeric complex of receptor-like kinases (LysM-RLKs) containing Lysine motif (LysM) domains: named LjNFR1 and LjNFR5 in *Lotus japonicus*, and MtLYK3 and MtNFP in *Medicago truncatula* (Arrighi et al., 2006; Limpens et al., 2003; Madsen et al., 2003; Radutoiu et al., 2003). The LysM domain is a ubiquitous molecular structure of 42-48 amino acids with a symmetrical βααβ folding. LysM domain-containing proteins were first described in bacteria to bind peptidoglycan (Buist, Steen, Kok, & Kuipers, 2008). In legumes LjNFR1/MtLYK3 and LjNFR5/MtNFP harbor 3 LysM domains in the receptor region which are essential to recognize specific rhizobium LCOs (Broghammer et al., 2012). In addition, it was found in *Lotus japonicus* that LjNFR5 interacts also with LjSYMRK, a LRR-type receptor that commits an essential function in symbiotic signalling (Antolín-Llovera, Ried, & Parniske, 2014). Interestingly, SYMRK is also essential for arbuscular mycorrhizal symbiosis, whereas both LjNFR1/MtLYK3 and LjNFR5/MtNFP only play an additive role in arbuscular mycorrhizal symbiosis (Oldroyd, 2013). Arbuscular mycorrhizal LCOs are known to trigger lateral root formation in *Medicago truncatula*, a response that is abolished in the *Mtnfp* knockout mutant (Maillet et al., 2011). *Mtlyk3* and *Ljnfr1* mutants display only a reduced level of infection when inoculated with a low dose of arbuscular mycorrhizal spores (X. Zhang et al., 2015). Nevertheless, transcriptome studies in *Medicago truncatula* revealed that MtNFP is playing a prominent role in arbuscular mycorrhizal LCO-induced transcriptional changes (Czaja et al., 2012). Two reasons may explain this discrepancy between knockout phenotype and function. Firstly, the weak arbuscular mycorrhizal symbiosis phenotype of the *Ljnr1/Mtlyk3* and *Ljnfr5/ Mtnfp* knockout mutants may be the result of gene redundancy in Lotus japonicus and Medicago truncatula. Both rhizobium LCO receptors evolved upon gene duplication events, giving rise to closely related homologs (De Mita, Streng, Bisseling, & Geurts, 2014; Op den Camp et al., 2011; Young et al., 2011). Expression studies of these

homologous genes show that they may also function in rhizobium and/or arbuscular mycorrhizal symbiosis (Rasmussen et al., 2016; Young et al., 2011). Secondly, it was found that arbuscular mycorrhizal fungi not only produce LCOs, but also short chain chitooligosaccharides (tetra and pentameric COs) as symbiotic signals (Genre et al., 2013). Such COs trigger in part similar symbiotic responses as reported for arbuscular mycorrhizal LCOs, though lack the capacity to promote lateral root formation (Maillet et al., 2011). LCO and CO signals may be perceived by different (symbiotic) receptor complexes.

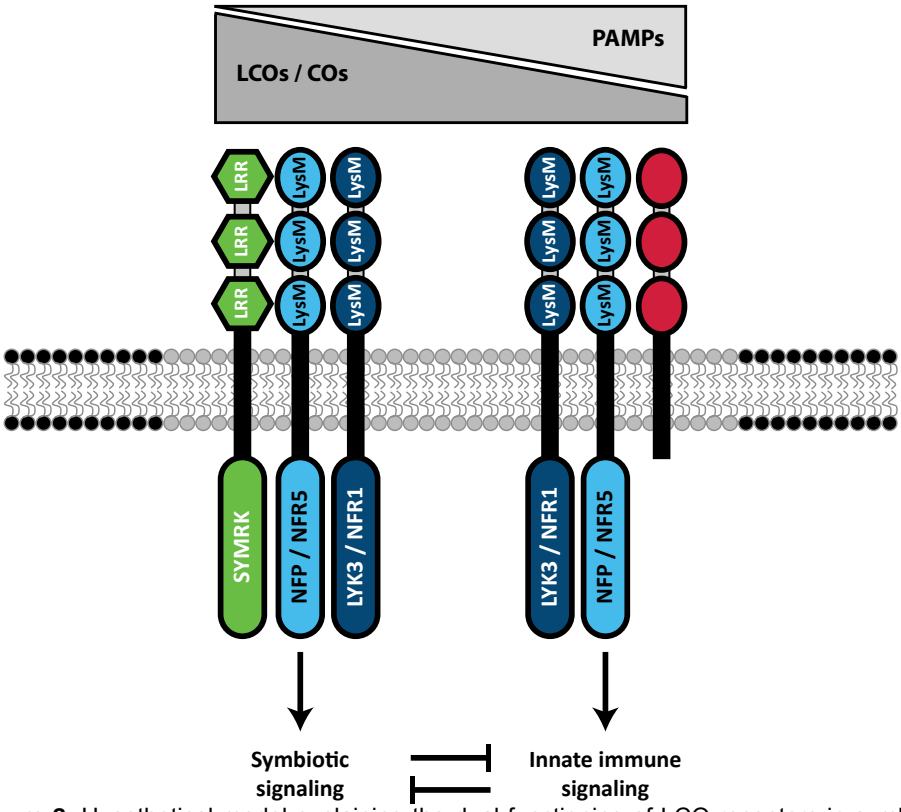


Figure 3. Hypothetical model explaining the dual functioning of LCO receptors in symbiotic and innate immune signalling as uncovered in *Medicago truncatula* and *Lotus japonicus*. Symbiotic signals and pathogen associated molecular patterns (PAMPs) are perceived by NFP/NFR5-NFR1/LYK3 receptor complexes. To commit either symbiotic or innate immune signalling a third receptor is essential. For symbiotic signalling this receptor may be SYMRK, as it interacts with NFP/NFR5. To induce innate immune responses this receptor has not been identified yet, but may have similarities to CeBIP in rice. OsCEBiP binds chitin oligomers and forms a heteromeric complex with the rice homolog of NFR1/LYK3 (OsCERK1) to activate innate immune signalling.

Non-legume systems provided additional support for a function of *NFR1/LYK3* and *NFR5/NFP* homologous genes in arbuscular mycorrhizal symbiosis. Reverse genetic studies in *Parasponia andersonii* and tomato (*Solanum lycopersicum*) revealed an essential role for putative NFR5/NFP orthologs in arbuscular mycorrhizal symbiosis (Buendia, Wang, Girardin, & Lefebvre, 2016; Op den Camp et al., 2011). In rice (*Oryza sativa*) it was demonstrated that the putative ortholog of NFR1/LYK3 - CHITIN-ELICITOR RECEPTOR KINASE 1 (OsCERK1)- plays such role (Miyata et al., 2014; X. Zhang et al., 2015). In Frankia no reverse genetic studies on LysM-RKs have been published yet. However, it is tempting to speculate that in actinorhizal plant species that can be nodulated by cluster 2 *Frankia* species, close homologs of *NFR1/LYK3* and/or *NFR5/NFP* play a symbiotic role in LCO perception.

The finding that COs and the chitin innate immune receptor OsCERK1 commit symbiotic functions uncovered a functional overlap between pathogenicity and symbiosis. Subsequent studies in *Lotus japonicus* and *Medicago truncatula* revealed four lines of supportive evidence for such dual function of LCO receptors. (I.) Rhizobium LCOs transiently trigger defence-related gene expression in an LjNFR1-dependent manner (Nakagawa et al., 2011). (II.) MtNFP has a function in defence against fungal and oomycete pathogens (Ben et al., 2013; Rey et al., 2013; Rey, Chatterjee, Buttay, Toulotte, & Schornack, 2015). (III.) Ectopic expression of both receptors - LjNFR5-LjNFR1 or MtNFP-MtLYK3 - in *Nicotiana benthamiana* leaves triggers a hypersensitive response (HR) (Broghammer et al., 2012; Pietraszewska-Bogiel et al., 2013). (IV.) Ectopic expression of MtNFP in *Medicago truncatula* triggers a premature cell death in nodules (Moling et al., 2014). This, and other studies, made also clear that LCO receptors are under tight post-translational control in legumes, probably to prevent pathogenic responses. For example, in *Medicago truncatula* nodules MtNFP and MtLYK3 accumulate only in nodule cells where infection takes place, but both receptors are rapidly removed from the membrane surrounding the rhizobium infection thread (Moling et al., 2014). Furthermore, it was found that LCO receptors are located in lipid-raft-like micro-domains in the plasma membrane, which play an important role in complex formation and receptor turnover (Haney & Long, 2010; Lefebvre et al., 2010). Taken together, these data suggest that dual functioning of LCO receptors in defence and symbiosis is a conserved feature in legumes and non-legume species.

The biological function of the overlap of LCO receptors in symbiotic and innate immune signalling remains unclear. However, a challenging model can be postulated (Limpens et al., 2015). In this model competition between receptors occurs to form multimeric complexes that differ in their functioning. Presence of LCOs (and/or short-chain COs) results in preferential formation of symbiotic receptor complexes at the expense of the formation of complexes that act in innate immunity **(Figure 3)**. In

legumes, such innate immune receptor complex has not yet been characterized. However, studies in rice revealed that perception of chitin oligomers requires an additional LysM-domain-containing receptor, which lacks an intracellular kinase domain (Kaku et al., 2006). This CHITIN ELICITOR BINDING PROTEIN (OsCEBiP) binds chitin oligomers and forms a heteromeric complex with OsCERK1 to activate chitin-triggered defence responses (Hayafune et al., 2014; Shimizu et al., 2010). Such innate immune receptor complex may also have a function in symbiosis. It is known that several typical innate immune responses, such as calcium influx, production of reactive oxygen (ROS) species, and focal exocytosis are associated with rhizobial and arbuscular mycorrhizal infection (Brewin, 2004). Rhizobium triggers formation of infection threads, which are tip-growing structures. ROS production is thought to facilitate the oxidative cross-linking of the infection thread matrix to allow the formation of a tube-like infection thread (Brewin, 2004). In a scenario that innate immune responses play a symbiotic role, the spatiotemporal regulation of receptor complexes becomes crucial to prevent HR.

4.2 Bypassing LCO signalling

Besides LCO-mediated signalling, alternative routes occur to mediate symbiotic responses. For example, many *Frankia* species (clusters 1 and 3) do not possess the machinery to produce LCOs (Tisa, Beauchemin, Gtari, Sen, & Wall, 2013). Furthermore, there are some legume lineages - e.g. several *Aeschynomene* species - that are nodulated by *Bradyrhizobium* strains that lack the highly conserved *nodABC* operon necessary for LCO synthesis (Fabre et al., 2015; Giraud et al., 2007). Nevertheless, studies in actinorhizal plant species *Casuarina glauca* and *Alnus glutinosa* using the non-LCO producing *Frankia* strain Cci3i, revealed that both SYMRK and CCaMK are essential to establish a symbiotic interaction, and activation of symbiotic signalling induces calcium oscillations (Chabaud et al., 2015; Franche et al., 2011; Gherbi et al., 2008; Granqvist et al., 2015). This strongly suggests that the underlying signalling pathway to establish an endosymbiosis is highly conserved, but can be activated by different signalling inputs.

The way non-LCO-producing rhizobia and Frankia achieve activation of the common symbiosis signalling pathway may vary. One way is by producing effector-like molecules that are secreted via the type III secretion system (T3SS). This mechanism is used by several rhizobium strains (Okazaki et al., 2016), and studies in soybean revealed that such effectors can bypass NFR1-NFR5 based signalling (Okazaki, Kaneko, Sato, & Saeki, 2013). However, additional mechanisms may also occur. For example, in case of *Aeschynomene* legumes the common symbiosis signalling pathway can also be activated in a T3SS-independent way (Fabre et al., 2015; Okazaki et al., 2016).

The current hypothesis is that *Frankia* strains of clusters 1 and 3 produce signalling molecules upon host recognition, of which the chemical nature is still poorly understood, but most probably different from LCOs. A first characterization of such signals came from studies on *Frankia* sp. strain CcI3i that nodulates *Casuarina glauca*. The signalling molecules produced by this strain are of low molecular weight, in the range of 500–5000 Dalton. Moreover, rhizobium and arbuscular mycorrhizal LCOs typically accumulate in the organic fraction upon a butanol extraction, whereas, in the case of *Frankia* CcI3i exudates, only water fractions could induce symbiotic responses (i.e. calcium oscillation). Furthermore, a chitinase treatment on the active water fractions did not affect their signalling capacity (Chabaud et al., 2015). This makes it highly unlikely that this strain produces LCO-type symbiotic signal molecules.

Studies with other *Frankia* strains revealed that, at least within a taxonomic cluster, the symbiotic signals are to a certain level conserved. For example, *Alnus glutinosa* and *Casuarina glauca* are nodulated by two different *Frankia* strains of the same cluster 1. Despite this strain specificity, *Frankia* sp. strain AC14a that nodulates *Alnus glutinosa* induces also calcium oscillation responses in *Casuarina glauca*. By contrast, the more distant BCU110501 strain of cluster 3 was unable to induce such responses (Chabaud et al., 2015). This suggests that the symbiotic signals produced by *Frankia* species are partially conserved within a taxonomic cluster, but may differ in a broader phylogenetic context.

5. Repressing Immunity

Although innate immune responses may be an integral part of the symbiotic infection process, it is essential that severe immune responses are avoided. Immune responses are controlled by two antagonistic hormones jasmonic acid and salicylic acid. The latter hormone is a major signal in resistance to biotrophic pathogens, whereas defence against necrotrophic mainly relies on jasmonic acid (Pieterse, Van der Does, Zamioudis, Leon-Reyes, & Van Wees, 2012). Both hormones act antagonistically, such that activation of jasmonic acid signalling compromises salicylic acid-dependent innate immune responses, and vice versa.

Studies in legumes suggest that repression of innate immunity is in part controlled by LCO signalling. In alfalfa (*Medicago sativa*) evidence was found that LCO signalling suppresses salicylic acid-dependent responses. LCO-deficient or incompatible rhizobia induce accumulation of salicylic acid, whereas compatible strains trigger a decrease of this defence hormone (Martinez-Abarca et al., 1998). Similarly, studies in pea (*Pisum sativum*) showed that endomycorrhizal fungi only trigger a transient increase in salicylic acid levels, which is repressed during prolonged colonization. In contrast, in a symbiosis deficient *ccamk* knockout mutant salicylic acid levels remain

high upon inoculation with endomycorrhizal fungi, suggesting that this suppression is based on activation of the symbiosis signalling network (Blilou, Ocampo, Garcıa-Garrido, & García-Garrido, 1999). Interestingly, defence responses in non-legumes (*Zea mays, Setaria viridis*), and even in non-AM plants (*Arabidopsis thaliana*) seem to be downregulated upon LCO perception, however it is currently unclear how this downregulation is linked to JA and SA signalling (Liang et al., 2013; Tanaka et al., 2015).

The Jasmonic acid - salicylic acid balance is in part controlled by DELLA GRAS-type transcriptional regulators (Navarro et al., 2008). DELLAs promote jasmonic acid signalling by binding JAZ (JASMONATE ZIM-DOMAIN) repressor proteins (Hou, Lee, Xia, Yan, & Yu, 2010). JAZ proteins repress jasmonic acid signalling upon binding with the MYC2 transcriptional activator (Boter, Ruı, & Abdeen, 2004; Hou et al., 2010). As MYC2 activity promotes DELLA accumulation, this results in a feedforward loop in jasmonic acid signalling (Wild et al., 2012; Yang et al., 2012). Several experiments indicate that endomycorrhizal fungi and rhizobium exploit this pathway, thereby indirectly reducing salicylic acid responses. *Della* knockout mutants in *Medicago truncatula* and rice are impaired in nodulation and/or arbuscule formation (Floss, Levy, Lévesque-Tremblay, Pumplin, & Harrison, 2013; Fonouni-Farde et al., 2016; Pimprikar et al., 2016; Yu et al., 2014). These phenotypes can be mimicked by application of gibberellins, whereas ectopic expression of a dominant active *DELLA* allele (*MtDELLA1Δ18*) promotes symbiotic responses (Floss et al., 2013; Jin et al., 2016; Pimprikar et al., 2016). Interestingly, the dominant active allele can also complement the cyclops symbiotic signalling mutant (Floss et al., 2013). This is likely due to the fact that in *Medicago truncatula* the DELLA1 protein was found to be able to form a complex with CYCLOPS and CCaMK, together activating the RAM1 GRAS-type transcriptional regulator (Pimprikar et al., 2016). Taken together this suggests that MtDELLA1 by interacting with JAZ proteins plays an important role in the LCO signalling network and the promotion of endomycorrhizal symbiosis through the modulation of jasmonic acid- salicylic acid balance.

Besides LCO triggered repression of immunity, plant immunity can also be manipulated by microbe secreted effector proteins. Studies in arbuscular mycorrhiza and ectomycorrhiza uncovered several small secreted effector proteins that are produced by the arbuscular mycorrhizal fungus *R. irregularis* and the ectomycorrhizal fungus *L. bicolor* (Lin et al., 2014; Martin et al., 2008; Tisserant et al., 2013). The mode of action of two such effector proteins has been characterized.

The *Rhizophagus irregularis* effector protein RiSP7 is secreted into *Medicago truncatula* root cells, where it localizes in the nucleus and interacts with a defence controlling ethylene-responsive transcription factor (MtERF19) (Kloppholz, Kuhn, &

Requena, 2011). In *Medicago truncatula* roots this gene is highly expressed upon pathogenic interaction, but only transiently during arbuscular mycorrhizal colonization. Ectopic expression of *RiSP7* in *Medicago truncatula* roots positively affects mycorrhizal colonization, while reducing defence responses. Intriguingly, RiSP7 has some similarity to the secreted NodO protein of *Rhizobium leguminosarum*, which enhances LCO signalling in the host plant. However, localization studies suggest that NodO localizes in the plant membrane, rather than acting as a nuclear effector (Economou, Hamilton, Johnston, & Downie, 1990; Sutton, Lea, & Downie, 1994).

The ectomycorrhizal fungus *Laccaria bicolor* expresses the *LbMiSSP7* gene encoding a secreted effector protein in response to plant exuded flavonoids (Plett & Martin, 2012). In black cottonwood poplar (*Populus trichocarpa*) it was shown that LbMiSSP7 is secreted in root cells where it localizes in the nucleus. There it stabilizes a JAZ protein (PtJAZ6) by direct interaction (Plett et al., 2014). As outlined above, JAZ proteins are repressors of jasmonic acid triggered immunity. Generally, JAZ proteins are degraded upon interaction with the F-box protein COI1 (CORONATINE-INSENSITIVE 1). This degradation is triggered by jasmonic acid signalling. LbMiSSP7 interaction to PtJAZ6 affects formation of the JAZ–COI1 complex. This prevents the jasmonic acid-dependent degradation of JAZ, resulting in reduced plant immune responses. Given that jasmonic acid is a negative regulator of ectomycorrhizal symbiosis, counteracting this plant innate immune response promotes the plant-fungus interaction.

6. Perspectives in symbiotic signalling

Central questions for future research will be on specificity of symbiotic signalling. How can a single symbiotic network that is conserved in most land plants trigger distinct root phenotypes? Since the symbiotic signalling network is basically conserved in most plant species the differences in the readout may be determined by yet unknown factors, such as the hormonal balance and/or the nutrient status of the root. For example, recently it was shown that Medicago truncatula lateral roots have an increased sensitivity to rhizobium LCOs compared with the main root. This indicates that susceptibility of a plant root varies, depending on the developmental and/or nutrient status (Sun et al., 2015).

Additional questions concerning specificity can also be addressed concerning the plant exuded flavonoids that act as attractants for symbiotic microbes. As shown for naringenin, these compounds are perceived by a diverging range of symbionts. Most probably this range extends to other soil borne microbes, most of which will not be symbiotic. Therefore, perhaps exuded flavonoids do not act as specific signals, but rather are more generic signals to which any root microbe can respond. For example, it was reported that exuded flavonoids may play a role also in phosphate and iron

acquisition (Cesco, Neumann, Tomasi, Pinton, & Weisskopf, 2010). In addition, the finding that flavonoids - similar to strigolactones - have a dual function, not only act as an attractant, but also function as endogenous plant signal interfering with auxin homeostasis, provides novel leads in symbiosis research.

Extending the range of model systems that are amenable for molecular genetic studies provided novel insights in symbiotic signalling. Establishment of new protocols for culturing Frankia and arbuscular mycorrhizal fungi, host induced gene silencing to trigger fungal gene expression, transformation of the ectomycorrhizal fungus Laccaria bicolor, the actinorhizal plants *Datisca, Casuarina*, and the non-legume rhizobia host *Parasponia*, in combination with microbial genome sequencing has opened new avenues. Although unravelling symbiotic signalling in these systems is still in its infancy, the recent findings that have been achieved are already groundbreaking. As mentioned above, it was demonstrated that especially in the endosymbioses (*Frankia*, rhizobium and arbuscular mycorrhiza) commonalities occur in symbiotic signalling (Gherbi et al., 2008; Op den Camp et al., 2011). One such commonality is that symbionts recognize plant secreted flavonoids and strigolactones. Another common theme is the use of LCO or CO signals of microbial origin of which biosynthesis is activated upon recognition of plant exuded molecules like flavonoids and/or strigolactones. LCOs/COs activate a conserved symbiotic network in plants that controls the diverse signalling output of the different symbiotic interactions (Parniske, 2008). Furthermore, it became apparent that LCO induced signalling can be bypassed. Especially in Frankia this appears to be a common strategy. Nevertheless, first studies indicate that LCO-independent signalling relies on the same symbiotic signalling network as identified in LCO dependent systems. Uncovering the nature of the non-LCO signal molecules in *Frankia* and rhizobia will add a new building brick in the symbiotic signalling network.

In the last decade new insights in the molecular aspects of root symbiosis were mainly generated by studying legume models *Medicago truncatula* and *Lotus japonicus*. With new model species in place in combination with next generation sequence technologies, this field will be revolutionized in the years to come.

Funding information

R.G. is supported by NWO-VICI 865.13.001 and W.K. by NWO-VENI 863.15.010

References

Abdel-Lateif, K., Vaissayre, V., Gherbi, H., Verries, C., Meudec, E., Perrine-Walker, F., … Hocher, V. (2013). Silencing of the chalcone synthase gene in *Casuarina glauca* highlights the important role of flavonoids during nodulation. *The New Phytologist*, *199*, 1012–1021. http://doi.org/10.1111/nph.12326

Akiyama, K., Matsuzaki, K., & Hayashi, H. (2005). Plant sesquiterpenes induce hyphal branching in arbuscular mycorrhizal fungi. *Nature*, *435*(June), 824–827. http://doi.org/10.1038/nature03608

Akiyama, K., Ogasawara, S., Ito, S., & Hayashi, H. (2010). Structural requirements of strigolactones for hyphal branching in AM fungi. *Plant & Cell Physiology*, *51*(7), 1104–17. http://doi.org/10.1093/pcp/pcq058

Al-Babili, S., & Bouwmeester, H. J. (2015). Strigolactones, a Novel Carotenoid-Derived Plant Hormone. *Annual Review of Plant Biology*, *66*(1), 161–186. http://doi.org/10.1146/annurev-arplant-043014-114759

Alder, a., Jamil, M., Marzorati, M., Bruno, M., Vermathen, M., Bigler, P., … Al-Babili, S. (2012). The Path from -Carotene to Carlactone, a Strigolactone-Like Plant Hormone. *Science*, *335*(6074), 1348–1351. http://doi.org/10.1126/science.1218094

Antolín-Llovera, M., Ried, M. K., & Parniske, M. (2014). Cleavage of the SYMBIOSIS RECEPTOR- LIKE KINASE ectodomain promotes complex formation with NOD FACTOR RECEPTOR 5. *Current Biology*, *24*, 422–427. Journal Article.

Arrighi, J. F., Barre, A., Ben Amor, B., Bersoult, A., Soriano, L. C., Mirabella, R., … Gough, C. (2006). The Medicago truncatula lysine motif-receptor-like kinase gene family includes NFP and new nodule-expressed genes. *Plant Physiol*, *142*(1), 265–279. Journal Article. http://doi.org/10.1104/pp.106.084657

Auguy, F., Abdel-Lateif, K., Doumas, P., Badin, P., Guerin, V., Bogusz, D., & Hocher, V. (2011). Activation of the isoflavonoid pathway in actinorhizal symbioses. *Functional Plant Biology*, *38*(8–9), 690–696. http://doi.org/10.1071/FP11014

Balestrini, R., & Bonfante, P. (2005). The interface compartment in arbuscular mycorrhizae: A special type of plant cell wall? *Plant Biosystems - An International Journal Dealing with All Aspects of Plant Biology*, *139*(1), 8–15. http://doi.org/10.1080/11263500500056799

Balestrini, R., & Bonfante, P. (2014). Cell wall remodeling in mycorrhizal symbiosis: a way towards biotrophism. *Frontiers in Plant Science*, *5*(June), 237. http://doi.org/10.3389/fpls.2014.00237

Beauchemin, N. J., Furnholm, T., Lavenus, J., Svistoonoff, S., Doumas, P., Bogusz, D., … Tisa, L. S. (2012). Casuarina root exudates alter the physiology, surface properties, and plant infectivity of Frankia sp. strain CcI3. *Applied and Environmental Microbiology*, *78*(2), 575–580. http://doi.org/10.1128/AEM.06183-11

Bécard, G., Douds, D. D., & Pfeffer, P. E. (1992). Extensive In Vitro Hyphal Growth of Vesicular-Arbuscular Mycorrhizal Fungi in the Presence of CO 2 and Flavonols Extensive In Vitro Hyphal Growth of Vesicular-Arbuscular Mycorrhizal Fungi in the Presence of CO2 and Flavonols. *Applied and Environmental Microbiology*, *58*(3), 821–825.

Bécard, G., Taylor, L. P., Douds, D. D., Pfeffer, P. E., & Doner, L. W. (1995). Flavonoids Are Not Necessary Plant Signal Compounds in Arbuscular Mycorrhizal Symbioses. *Molecular Plant-Microbe Interactions*. http://doi.org/10.1094/MPMI-8-0252

Behm, J. E., Geurts, R., & Kiers, E. T. (2014). Parasponia: a novel system for studying mutualism stability. *Trends in Plant Science*, *19*(12), 757–763. http://doi.org/10.1016/j.tplants.2014.08.007

Ben, C., Toueni, M., Montanari, S., Tardin, M.-C., Fervel, M., Negahi, A., … Gentzbittel, L. (2013). Natural diversity in the model legume Medicago truncatula allows identifying distinct genetic mechanisms conferring partial resistance to Verticillium wilt. *Journal of Experimental Botany*, *64*, 317–32. http://doi.org/10.1093/jxb/ers337

Besserer, A., Bécard, G., Jauneau, A., Roux, C., & Séjalon-Delmas, N. (2008). GR24, a synthetic analog of strigolactones, stimulates the mitosis and growth of the arbuscular mycorrhizal fungus Gigaspora rosea by boosting its energy metabolism. *Plant Physiology*, *148*(1), 402–413. http://doi.org/10.1104/pp.108.121400

Besserer, A., Puech-Pagès, V., Kiefer, P., Gomez-Roldan, V., Jauneau, A., Roy, S., … Séjalon-Delmas, N. (2006). Strigolactones stimulate arbuscular mycorrhizal fungi by activating mitochondria. *PLoS Biology*, *4*(7), 1239–1247. http://doi.org/10.1371/journal.pbio.0040226

Blilou, I., Ocampo, J. a., Garcıa-Garrido, M., & García-Garrido, J. M. . (1999). Resistance of pea roots to endomycorrhizal fungus or Rhizobium correlates with enhanced levels of endogenous salicylic acid. *Journal of Experimental Botany*, *50*(340), 1663–1668. http://doi.org/10.1093/jxb/50.340.1663

Boter, M., Ruı, O., & Abdeen, A. (2004). Conserved MYC transcription factors play a key role in jasmonate signaling both in tomato and Arabidopsis. *Genes & Development*, *2*, 1577–1591. http://doi.org/10.1101/gad.297704.tion

Brewer, P. B., Koltai, H., & Beveridge, C. A. (2013). Diverse roles of strigolactones in plant development. *Molecular Plant*, *6*(1), 18–28. http://doi.org/10.1093/mp/sss130

Brewin, N. J. (2004). Plant cell wall remodelling in the Rhizobium-legume symbiosis. *Critical Reviews in Plant Sciences*, *23*(4), 293–316. http://doi.org/10.1080/07352680490480734

Broghammer, a., Krusell, L., Blaise, M., Sauer, J., Sullivan, J. T., Maolanon, N., … Stougaard, J. (2012). Legume receptors perceive the rhizobial lipochitin oligosaccharide signal molecules by direct binding. *Proceedings of the National Academy of Sciences of the United States of America*, *109*(34), 13859–13864. http://doi.org/10.1073/pnas.1205171109

Brown, D. E., Rashotte, a M., Murphy, a S., Normanly, J., Tague, B. W., Peer, W. a, … Muday, G. K. (2001). Flavonoids act as negative regulators of auxin transport in vivo in arabidopsis. *Plant Physiology*, *126*(2), 524–535. http://doi.org/10.1104/pp.126.2.524

Buendia, L., Wang, T., Girardin, A., & Lefebvre, B. (2016). The LysM receptor-like kinase SlLYK10 regulates the arbuscular mycorrhizal symbiosis in tomato. *New Phytologist*, *210*(1), 184–195. http://doi.org/10.1111/nph.13753

Buist, G., Steen, A., Kok, J., & Kuipers, O. P. (2008). LysM, a widely distributed protein motif for binding to (peptido) glycans. *Molecular Microbiology*, *68*(April), 838–847. http://doi.org/10.1111/j.1365-2958.2008.06211.x

Capoen, W., Sun, J., Wysham, D., Otegui, M. S., Venkateshwaran, M., Hirsch, S., … Oldroyd, G. E. D. (2011). Nuclear membranes control symbiotic calcium signaling of legumes. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(34), 14348–14353. http://doi.org/10.1073/pnas.1107912108

Catoira, R., Galera, C., de Billy, F., Penmetsa, R. V, Journet, E. P., Maillet, F., … Denarie, J. (2000). Four genes of Medicago truncatula controlling components of a nod factor transduction pathway. *Plant Cell*, *12*(9), 1647–1666. Journal Article. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11006338

Cesco, S., Neumann, G., Tomasi, N., Pinton, R., & Weisskopf, L. (2010). Release of plant-borne flavonoids into the rhizosphere and their role in plant nutrition. *Plant Soil*, *329*, 1–25. http://doi.org/10.1007/s11104-009-0266-9

Chabaud, M., Gherbi, H., Pirolles, E., Vaissayre, V., Moukouanga, D., Franche, C., … Svistoonoff, S. (2015). Rapid report Chitinase-resistant hydrophilic symbiotic factors secreted by Frankia activate both Ca2+ spiking and NIN gene expression in the actinorhizal plant Casuarina glauca. *New Phytol.* http://doi.org/10.1111/nph.13732

Charpentier, M., Sun, J., Martins, T. V., Radhakrishnan, G. V, Findlay, K., Soumpourou, E., … Oldroyd, G. E. D. (2016). Nuclear-localized cyclic nucleotide–gated channels mediate symbiotic calcium oscillations. *Science*, *352*(6289), 1102–1105. http://doi.org/10.1126/science.aae0109

Chen, X.-C., Feng, J., Hou, B.-H., Li, F.-Q., Li, Q., & Hong, G.-F. (2005). Modulating DNA bending affects NodD-mediated transcriptional control in Rhizobium leguminosarum. *Nucleic Acids Research*, *33*(8), 2540–8. http://doi.org/10.1093/nar/gki537

Clavijo, F., Diedhiou, I., Vaissayre, V., Brottier, L., Acolatse, J., Moukouanga, D., … Svistoonoff, S. (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. http://doi.org/10.1111/nph.13506

Coleman-Derr, D., & Tringe, S. G. (2014). Building the crops of tomorrow: Advantages of symbiont-based approaches to improving abiotic stress tolerance. *Frontiers in Microbiology*, *5*(JUN), 1–6. http://doi.org/10.3389/fmicb.2014.00283

Czaja, L. F., Hogekamp, C., Lamm, P., Maillet, F., Martinez, E. a., Samain, E., … Hohnjec, N. (2012). Transcriptional Responses toward Diffusible Signals from Symbiotic Microbes Reveal MtNFP- and MtDMI3-Dependent Reprogramming of Host Gene Expression by Arbuscular Mycorrhizal Fungal Lipochitooligosaccharides. *Plant Physiology*, *159*(4), 1671–1685. http://doi.org/10.1104/pp.112.195990

De Cuyper, C., Fromentin, J., Yocgo, R. E., De Keyser, A., Guillotin, B., Kunert, K., … Goormachtig, S. (2014). From lateral root density to nodule number, the strigolactone analogue GR24 shapes the root architecture of Medicago truncatula. *Journal of Experimental Botany*, *66*(1), 137–146. http://doi.org/10.1093/jxb/eru404

De Mita, S., Streng, A., Bisseling, T., & Geurts, R. (2014). Evolution of a symbiotic receptor through gene duplications in the legume-rhizobium mutualism. *New Phytologist*, *201*(3), 961–972. http://doi.org/10.1111/nph.12549

Deinum, E. E., Geurts, R., Bisseling, T., & Mulder, B. M. (2012). Modeling a Cortical Auxin Maximum for Nodulation: Different Signatures of Potential Strategies. *Frontiers in Plant Science*, *3*(May), 1–19. http://doi.org/doi: 10.3389/fpls.2012.00096

Dor, E., Joel, D. M., Kapulnik, Y., Koltai, H., & Hershenhorn, J. (2011). The synthetic strigolactone GR24 influences the growth pattern of phytopathogenic fungi. *Planta*, *234*(2), 419–427. http://doi.org/10.1007/s00425-011-1452-6

Economou, A., Hamilton, W. D., Johnston, A. W., & Downie, J. A. (1990). The Rhizobium nodulation gene nodO encodes a Ca2(+)-binding protein that is exported without N-terminal cleavage and is homologous to haemolysin and related proteins. *The EMBO Journal*, *9*(2), 349–54. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/2303029

Ellouze, W., Hamel, C., Cruz, A. F., Ishii, T., Gan, Y., Bouzid, S., & St-Arnaud, M. (2012). Phytochemicals and spore

germination: At the root of AMF host preference? *Applied Soil Ecology*, *60*, 98–104. http://doi.org/10.1016/j.apsoil.2012.02.004

Fabre, S., Gully, D., Poitout, A., Patrel, D., Arrighi, J.-F., Giraud, E., … Cartieaux, F. (2015). The Nod factor-independent nodulation in Aeschynomene evenia required the common plant-microbe symbiotic "toolkit." *Plant Physiology*, *169*(December), pp.01134.2015. http://doi.org/10.1104/pp.15.01134

Falcone Ferreyra, M. L., Rius, S. P., & Casati, P. (2012). Flavonoids: biosynthesis, biological functions, and biotechnological applications. *Frontiers in Plant Science*, *3*(September), 1–15. http://doi.org/10.3389/fpls.2012.00222

Ferrer, J. L., Austin, M. B., Stewart, C., & Noel, J. P. (2008). Structure and function of enzymes involved in the biosynthesis of phenylpropanoids. *Plant Physiology and Biochemistry*, *46*(3), 356–370. http://doi.org/10.1016/j.plaphy.2007.12.009

Floss, D. S., Levy, J. G., Lévesque-Tremblay, V., Pumplin, N., & Harrison, M. J. (2013). DELLA proteins regulate arbuscule formation in arbuscular mycorrhizal symbiosis. *Proceedings of the National Academy of Sciences*, *110*(51), E5025-5034. http://doi.org/10.1073/pnas.1308973110

Fonouni-Farde, C., Tan, S., Baudin, M., Brault, M., Wen, J., Mysore, K. S., … Diet, A. (2016). DELLA-mediated gibberellin signalling regulates Nod factor signalling and rhizobial infection. *Nature Communications*, *7*, 12636. http://doi.org/10.1038/ncomms12636

Foo, E., Yoneyama, K., Hugill, C. J., Quittenden, L. J., & Reid, J. B. (2013). Strigolactones and the regulation of pea symbioses in response to nitrate and phosphate deficiency. *Molecular Plant*, *6*(1), 76–87. http://doi.org/10.1093/mp/sss115

Franche, C., Gherbi, H., Benabdoun, M., Svistoonoff, S., Hocher, V., & Bogusz, D. (2011). New insights in the molecular events underlying actinorhizal nodulation in the tropical tree Casuarina glauca. *BMC Proceedings*, *5*(Suppl 7), O33. http://doi.org/10.1186/1753-6561-5-S7-O33

Garcia, K., Delaux, P.-M., Cope, K. R., & Ane, J.-M. (2015). Molecular signals required for the establishment and maintenance of ectomycorrhizal symbioses. *New Phytologist*, *208*, 79–87. http://doi.org/10.1080/15411790701409895

Garg, N., & Singla, P. (2016). Stimulation of nitrogen fixation and trehalose biosynthesis by naringenin (Nar) and arbuscular mycorrhiza (AM) in chickpea under salinity stress. *Plant Growth Regulation*, 1–18. http://doi.org/10.1007/s10725-016-0146-2

Gea, L., Normand, L., Vian, B., & Gay, G. (1994). Structural Aspects Of Ectomycorrhiza Of Pinus Pinaster (Ait) Sol Formed By an Iaa Overproducer Mutant Of Hebeloma Cylindrosporum Romagnesi. *New Phytologist. Dec*, *128*(4), 659–670. http://doi.org/10.1111/j.1469-8137.1994.tb04030.x

Genre, A., Chabaud, M., Balzergue, C., Puech-Pagès, V., Novero, M., Rey, T., … Barker, D. G. (2013). Short-chain chitin oligomers from arbuscular mycorrhizal fungi trigger nuclear Ca2+ spiking in Medicago truncatula roots and their production is enhanced by strigolactone. *New Phytologist*, *198*, 190–202. http://doi.org/10.1111/nph.12146

Geurts, R., Lillo, A., & Bisseling, T. (2012). Exploiting an ancient signalling machinery to enjoy a nitrogen fixing symbiosis. *Current Opinion in Plant Biology*, *15*(4), 438–443. http://doi.org/10.1016/j.pbi.2012.04.004

Geurts, R., Xiao, T. T., & Reinhold-Hurek, B. (2016). What Does It Take to Evolve A Nitrogen-Fixing Endosymbiosis? *Trends in Plant Science*, *21*(3), 199–208. http://doi.org/10.1016/j.tplants.2016.01.012

Gherbi, H., Markmann, K., Svistoonoff, S., Estevan, J., Autran, D., Giczey, G., … Bogusz, D. (2008). SymRK defines a common genetic basis for planGherbi, H., Markmann, K., Svistoonoff, S., Estevan, J., Autran, D., Giczey, G., … Bogusz, D. (2008). SymRK defines a common genetic basis for plant root endosymbioses with arbuscular mycorrhiza fungi, rhizobia,. *Proceedings of the National Academy of Sciences of the United States of America*, *105*(12), 4928–4932. http://doi.org/10.1073/pnas.0710618105

Giraud, E., Moulin, L., Vallenet, D., Barbe, V., Cytryn, E., Avarre, J.-C., … Sadowsky, M. (2007). Legumes symbioses: absence of Nod genes in photosynthetic bradyrhizobia. *Science (New York, N.Y.)*, *316*(5829), 1307–1312. http://doi.org/10.1126/science.1139548

Gobbato, E., Marsh, J. F., Vernié, T., Wang, E., Maillet, F., Kim, J., … Oldroyd, G. E. D. (2012). A GRAS-Type Transcription Factor with a Specific Function in Mycorrhizal Signaling. *Current Biology*, *22*(23), 2236–2241. http://doi.org/10.1016/j.cub.2012.09.044

Gomez-Roldan, V., Fermas, S., Brewer, P. B., Puech-Pagès, V., Dun, E. A., Pillot, J.-P., … Rochange, S. F. (2008). Strigolactone inhibition of shoot branching. *Nature*, *455*(7210), 189–194. http://doi.org/10.1038/nature07271

Granqvist, E., Sun, J., Op Den Camp, R., Pujic, P., Hill, L., Normand, P., … Oldroyd, G. E. D. (2015). Bacterial-induced calcium oscillations are common to nitrogen-fixing associations of nodulating legumes and nonlegumes. *New Phytologist*, *207*(3), 551–558. http://doi.org/10.1111/nph.13464

Gtari, M., Tisa, L. S., & Normand, P. (2013). Diversity of Frankia Strains, Actinobacterial Symbionts of Actinorhizal Plants. In R. Aroca (Ed.), *Symbiotic Endophytes* (Vol. 37, pp. 123–148). Berlin, Heidelberg: Springer. http://doi.org/10.1007/978-3-642-39317-4

Gutjahr, C., Gobbato, E., Choi, J., Riemann, M., Johnston, M. G., Summers, W., … Paszkowski, U. (2015). Rice perception of symbiotic arbuscular mycorrhizal fungi requires the karrikin receptor complex. *Science*, *350*(6267), 1521–1524. http://doi.org/10.1126/science.aac9715

Hamiaux, C., Drummond, R. S. M. S. M., Janssen, B. J. J., Ledger, S. E. E., Cooney, J. M. M., Newcomb, R. D. D., & Snowden, K. C. C. (2012). DAD2 is an α/β hydrolase likely to be involved in the perception of the plant branching hormone, strigolactone. *Current Biology*, *22*(21), 2032–2036. http://doi.org/10.1016/j.cub.2012.08.007

Haney, C. H., & Long, S. R. (2010). Plant flotillins are required for infection by nitrogen-fixing bacteria. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(1), 478–483. http://doi.org/10.1073/pnas.0910081107

Hartwig, U. a, Joseph, C. M., & Phillips, D. a. (1991). Flavonoids Released Naturally from Alfalfa Seeds Enhance Growth Rate of Rhizobium meliloti. *Plant Physiology*, *95*(3), 797–803. http://doi.org/10.1104/pp.95.3.797

Hayafune, M., Berisio, R., Marchetti, R., Silipo, A., Kayama, M., Desaki, Y., … Shibuya, N. (2014). Chitin-induced activation of immune signaling by the rice receptor CEBiP relies on a unique sandwich-type dimerization. *Proceedings of the National Academy of Sciences of the United States of America*, *111*, E404–E413. http://doi.org/10.1073/pnas.1312099111

Heckmann, A. B., Lombardo, F., Miwa, H., Perry, J. A., Bunnewell, S., Parniske, M., … 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*(4), 1739–1750. Journal Article. http://doi.org/DOI 10.1104/pp.106.089508

Hirsch, A. M., Bhuvaneswari, T. V., Torrey, J. G., & Bisseling, T. (1989). Early nodulin genes are induced in alfalfa root outgrowths elicited by auxin transport inhibitors. *Proceedings of the National Academy of Sciences of the United States of America*, *86*(4), 1244–8. http://doi.org/10.1073/pnas.86.4.1244

Honma, M. A., Asomaning, M., & Ausubel, F. M. (1990). Rhizobium meliloti nodD genes mediate host-specific activation of nodABC. *Journal of Bacteriology*, *172*(2), 901–11. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/2298703

Hou, X., Lee, L. Y. C., Xia, K., Yan, Y., & Yu, H. (2010). DELLAs Modulate Jasmonate Signaling via Competitive Binding to JAZs. *Developmental Cell*, *19*(6), 884–894. http://doi.org/10.1016/j.devcel.2010.10.024

Imaizumi-Anraku, H., Takeda, N., Charpentier, M., Perry, J., Miwa, H., Umehara, Y., … Hayashi, M. (2005). Plastid proteins crucial for symbiotic fungal and bacterial entry into plant roots. *Nature*, *433*(7025), 527–531. http://doi.org/10.1038/nature03237

Jin, Y., Liu, H., Luo, D., Yu, N., Dong, W., Wang, C., … Wang, E. (2016). DELLA proteins are common components of symbiotic rhizobial and mycorrhizal signalling pathways. *Nature Communications*, *7*, 12433. http://doi.org/10.1038/ncomms12433

Jones, D. L., Nguyen, C., & Finlay, R. D. (2009). Carbon flow in the rhizosphere: Carbon trading at the soil-root interface. *Plant and Soil*, *321*(1–2), 5–33. http://doi.org/10.1007/s11104-009-9925-0

Kaku, H., Nishizawa, Y., Ishii-Minami, N., Akimoto-Tomiyama, C., Dohmae, N., Takio, K., … Shibuya, N. (2006). Plant cells recognize chitin fragments for defence signaling through a plasma membrane receptor. *Proceedings of the National Academy of Sciences of the United States of America*, *103*, 11086–11091. http://doi.org/10.1073/pnas.0508882103

Kikuchi, K., Matsushita, N., Suzuki, K., & Hogetsu, T. (2007). Flavonoids induce germination of basidiospores of the ectomycorrhizal fungus Suillus bovinus. *Mycorrhiza*, *17*(7), 563–570. http://doi.org/10.1007/s00572-007-0131-8

Kloppholz, S., Kuhn, H., & Requena, N. (2011). A secreted fungal effector of glomus intraradices promotes symbiotic biotrophy. *Current Biology*, *21*(14), 1204–1209. http://doi.org/10.1016/j.cub.2011.06.044

Kohlen, W., Charnikhova, T., Lammers, M., Pollina, T., Tóth, P., Haider, I., … López-Ráez, J. a. (2012). The tomato carotenoid cleavage dioxygenase8 (SlCCD8) regulates rhizosphere signaling, plant architecture and affects reproductive development through strigolactone biosynthesis. *New Phytologist*, *196*(2), 535–547. http://doi.org/10.1111/j.1469-8137.2012.04265.x

Kohler, A., Kuo, A., Nagy, L. G., Morin, E., Barry, K. W., Buscot, F., … Martin, F. (2015). Convergent losses of decay mechanisms and rapid turnover of symbiosis genes in mycorrhizal mutualists. *Nature Genetics*, *47*(4), 410–5. http://doi.org/10.1038/ng.3223

Kosslak, R. M., Bookland, R., Barkei, J., Paaren, H. E., & Appelbaum, E. R. (1987). Induction of Bradyrhizobium japonicum common nod genes by isoflavones isolated from Glycine max. *Proceedings of the National Academy of Sciences of the United States of America*, *84*(21), 7428–32. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/16593884

Kretzschmar, T., Kohlen, W., Sasse, J., Borghi, L., Schlegel, M., Bachelier, J. B., … Martinoia, E. (2012). A petunia ABC protein controls strigolactone-dependent symbiotic signalling and branching. *Nature*. Nature Publishing Group. http://doi.org/10.1038/nature10873

Lefebvre, B., Timmers, T., Mbengue, M., Moreau, S., Hervé, C., Tóth, K., … Ott, T. (2010). A remorin protein interacts with symbiotic receptors and regulates bacterial infection. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(5), 2343–2348. http://doi.org/10.1073/pnas.0913320107

Lévy, J., Bres, C., Geurts, R., Chalhoub, B., Kulikova, O., Duc, G., … Debellé, F. (2004). A putative Ca2+ and calmodulin-dependent protein kinase required for bacterial and fungal symbioses. *Science (New York, N.Y.)*, *303*(5662), 1361–1364. http://doi.org/10.1126/science.1093038

Liang, Y., Cao, Y., Tanaka, K., Thibivilliers, S., Wan, J., Choi, J., … Stacey, G. (2013). Nonlegumes respond to rhizobial Nod factors by suppressing the innate immune response. *Science (New York, N.Y.)*, *341*(6152), 1384–1387. http://doi.org/10.1126/science.1242736

Limpens, E., Franken, C., Smit, P., Willemse, J., Bisseling, T., & Geurts, R. (2003). LysM domain receptor kinases regulating rhizobial Nod factor-induced infection. *Science (New York, N.Y.)*, *302*(5645), 630–633. http://doi.org/10.1126/science.1090074

Limpens, E., van Zeijl, A., & Geurts, R. (2015). Lipochitooligosaccharides Modulate Plant Host Immunity to Enable Endosymbioses. *Annual Review of Phytopathology*, *53*(1), 311–334. http://doi.org/10.1146/annurev-phyto-080614-120149

Lin, K., Limpens, E., Zhang, Z., Ivanov, S., Saunders, D. G. O., Mu, D., … Huang, S. (2014). Single nucleus genome sequencing reveals high similarity among nuclei of an endomycorrhizal fungus. *PLoS Genetics*, *10*, e1004078. http://doi.org/10.1371/journal.pgen.1004078

Liu, J., Novero, M., Charnikhova, T., Ferrandino, A., Schubert, A., Ruyter-Spira, C., … Cardinale, F. (2013). Carotenoid cleavage dioxygenase 7 modulates plant growth, reproduction, senescence, and determinate nodulation in the model legume Lotus japonicus. *Journal of Experimental Botany*, *64*(7), 1967–1981. http://doi.org/10.1093/jxb/ert056

Liu, W., Kohlen, W., Lillo, A., Op den Camp, R., Ivanov, S., Hartog, M., … Geurts, R. (2011). Strigolactone biosynthesis in *Medicago truncatula* and rice requires the symbiotic GRAS-type transcription factors NSP1 and NSP2. *Plant Cell*, *23*(10), 3853–65. http://doi.org/10.1105/tpc.111.089771

Madsen, E. B., Madsen, L. H., Radutoiu, S., Sato, S., Kaneko, T., Tabata, S., & Sandal, N. (2003). A receptor kinase gene of the LysM type is involved in legume perception of rhizobial signals. *Nature*, *425*, 637–640.

Maillet, F., Poinsot, V., André, O., Puech-Pagès, V., Haouy, A., Gueunier, M., … Dénarié, J. (2011). Fungal lipochitooligosaccharide symbiotic signals in arbuscular mycorrhiza. *Nature*, *469*(7328), 58–63. http://doi.org/10.1038/nature09622

Marsh, J. F., Rakocevic, A., Mitra, R. M., Brocard, L., Sun, J., Eschstruth, A., … Oldroyd, G. E. D. D. (2007). Medicago truncatula NIN is essential for rhizobial-independent nodule organogenesis induced by autoactive Calcium/Calmodulin-Dependent Protein Kinase. *Plant Physiology*, *144*(1), 324–335. JOUR. http://doi.org/10.1104/pp.106.093021

Martin, F., Aerts, A., Ahrén, D., Brun, A., Danchin, E. G. J., Duchaussoy, F., … Grigoriev, I. V. (2008). The genome of Laccaria bicolor provides insights into mycorrhizal symbiosis. Nature, 452(7183), 88–92. http://doi.org/10.1038/nature06556

Martin, F., Kohler, A., Murat, C., Veneault-Fourrey C., & David S. Hibbett, D.S. (2016) Unearthing the roots of ectomycorrhizal symbioses. Nature reviews Microbiology. doi:10.1038/nrmicro.2016.149

Martinez-Abarca, F., Herrera-Cervera, J. A., Bueno, P., Sanjuan, J., Bisseling, T., & Olivares, J. (1998). Involvement of Salicylic Acid in the Establishment of the Rhizobium meliloti -Alfalfa Symbiosis. *Molecular Plant Microbe Interactions*, *11*, 153–155. http://doi.org/10.1094/MPMI.1998.11.2.153

Mathesius, U., Schlaman, H. R., Spaink, H. P., Of Sautter, C., Rolfe, B. G., & Djordjevic, M. A. (1998). Auxin transport inhibition precedes root nodule formation in white clover roots and is regulated by flavonoids and derivatives of chitin oligosaccharides. *Plant Journal*, *14*(1), 23–34. Journal Article. http://doi.org/10.1046/j.1365-313X.1998.00090.x

Mendes, R., Garbeva, P., & Raaijmakers, J. M. (2013). The rhizosphere microbiome: Significance of plant beneficial, plant pathogenic, and human pathogenic microorganisms. *FEMS Microbiology Reviews*, *37*(5), 634–663. http://doi.org/10.1111/1574-6976.12028

Miyata, K., Kozaki, T., Kouzai, Y., Ozawa, K., Ishii, K., Asamizu, E., … Nakagawa, T. (2014). Bifunctional plant receptor, OsCERK1, regulates both chitin-triggered immunity and arbuscular mycorrhizal symbiosis in rice. *Plant and Cell Physiology*, *55*, in press.

Moling, S., Pietraszewska-Bogiel, A., Postma, M., Fedorova, E., Hink, M. a., Limpens, E., … Bisseling, T. (2014). Nod

Factor Receptors Form Heteromeric Complexes and Are Essential for Intracellular Infection in Medicago Nodules. *The Plant Cell*, *26*(October), 4188–4199. http://doi.org/10.1105/tpc.114.129502

Mulligan, J. T., & Long, S. R. (1989). A family of activator genes regulates expression of Rhizobium meliloti nodulation genes. *Genetics*, *122*(1), 7–18. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/2731734

Nakagawa, T., Kaku, H., Shimoda, Y., Sugiyama, A., Shimamura, M., Takanashi, K., … Kouchi, H. (2011). From defence to symbiosis: limited alterations in the kinase domain of LysM receptor-like kinases are crucial for evolution of legume-Rhizobium symbiosis. *The Plant Journal*, *65*(2), 169–80. http://doi.org/10.1111/j.1365-313X.2010.04411.x

Navarro-Rodenas, A., Xu, H., Kemppainen, M., Pardo, A. G., & Zwiazek, J. J. (2015). Laccaria bicolor aquaporin LbAQP1 is required for Hartig net development in trembling aspen (Populus tremuloides). *Plant, Cell and Environment*, *38*(11), 2475–2486. http://doi.org/10.1111/pce.12552

Navarro, L., Bari, R., Achard, P., Lisón, P., Nemri, A., Harberd, N. P., & Jones, J. D. G. (2008). DELLAs control plant immune responses by modulating the balance of jasmonic acid and salicylic acid signaling. *Current Biology*, *18*, 650–655. http://doi.org/10.1016/j.cub.2008.03.060

Ng, J., Perrine-Walker, F., Wasson, A., & Mathesius, U. (2015). The Control of Auxin Transport in Parasitic and Symbiotic Root–Microbe Interactions. *Plants*, *4*(3), 606–643. http://doi.org/10.3390/plants4030606

Okazaki, S., Kaneko, T., Sato, S., & Saeki, K. (2013). Hijacking of leguminous nodulation signaling by the rhizobial type III secretion system. *Proceedings of the National Academy of Sciences*, *110*(42), 17131–17136. http://doi.org/10.1073/pnas.1302360110

Okazaki, S., Tittabutr, P., Teulet, A., Thouin, J., Fardoux, J., Chaintreuil, C., … Giraud, E. (2016). Rhizobium–legume symbiosis in the absence of Nod factors: two possible scenarios with or without the T3SS. *The ISME Journal*, *10*, 64–74. http://doi.org/10.1038/ismej.2015.103

Oldroyd, G. E. D. (2013). Speak, friend, and enter: signalling systems that promote beneficial symbiotic associations in plants. *Nature Reviews. Microbiology*, *11*(4), 252–63. http://doi.org/10.1038/nrmicro2990

Oldroyd, G. E., & Long, S. R. (2003). Identification and characterization of nodulation-signaling pathway 2, a gene of Medicago truncatula involved in Nod factor signaling. *Plant Physiol*, *131*(3), 1027–1032. Journal Article. http://doi.org/10.1104/pp.102.010710

Op den Camp, R. H. M., De Mita, S., Lillo, A., Cao, Q., Limpens, E., Bisseling, T., & Geurts, R. (2011). A Phylogenetic Strategy Based on a Legume-Specific Whole Genome Duplication Yields Symbiotic Cytokinin Type-A Response Regulators. *Plant Physiology*, *157*(December 2011), 2013–2022. http://doi.org/10.1104/pp.111.187526

Op den Camp, R., Streng, A., De Mita, S., Cao, Q., Polone, E., Liu, W., … Geurts, R. (2011). LysM-type mycorrhizal receptor recruited for rhizobium symbiosis in nonlegume Parasponia. *Science (New York, N.Y.)*, *331*(December), 909–912. http://doi.org/10.1126/science.1198181

Pandey, A., Sharma, M., & Pandey, G. K. (2016). Emerging Roles of Strigolactones in Plant Responses to Stress and Development. *Frontiers in Plant Science*, *7*, 434. http://doi.org/10.3389/fpls.2016.00434

Parniske, M. (2008). Arbuscular mycorrhiza: the mother of plant root endosymbioses. *Nature Reviews Microbiology*, *6*, 763–775. http://doi.org/10.1038/nrmicro1987

Pawlowski, K., & Demchenko, K. N. (2012). The diversity of actinorhizal symbiosis. *Protoplasma*, *249*(4), 967–979. http://doi.org/10.1007/s00709-012-0388-4

Peláez-Vico, M. A., Bernabéu-Roda, L., Kohlen, W., Soto, M. J., & López-Ráez, J. A. (2016). Strigolactones in the Rhizobium-legume symbiosis: Stimulatory effect on bacterial surface motility and down-regulation of their levels in nodulated plants. *Plant Science : An International Journal of Experimental Plant Biology*, *245*, 119–27. http://doi.org/10.1016/j.plantsci.2016.01.012

Persson, T., Battenberg, K., Demina, I. V., Vigil-Stenman, T., Vanden Heuvel, B., Pujic, P., … Berry, A. M. (2015). Candidatus Frankia Datiscae Dg1, the actinobacterial microsymbiont of Datisca glomerata, expresses the canonical nod genes nodABC in symbiosis with its host plant. *Plos One*, *10*(5), e0127630. http://doi.org/10.1371/journal.pone.0127630

Peters, N. K., & Long, S. R. (1988). Alfalfa Root Exudates and Compounds which Promote or Inhibit Induction of Rhizobium meliloti Nodulation Genes. *Plant Physiology*, *88*(2), 396–400. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/16666315

Pieterse, C. M. J., Van der Does, D., Zamioudis, C., Leon-Reyes, A., & Van Wees, S. C. M. (2012). Hormonal Modulation of Plant Immunity. *Annual Review of Cell and Developmental Biology*, *28*, 489–521. http://doi.org/10.1146/annurev-cellbio-092910-154055

Pietraszewska-Bogiel, A., Lefebvre, B., Koini, M. a., Klaus-Heisen, D., Takken, F. L. W., Geurts, R., … Gadella, T. W. J. (2013). Interaction of Medicago truncatula Lysin Motif Receptor-Like Kinases, NFP and LYK3, Produced in

Nicotiana benthamiana Induces Defence-Like Responses. *PLoS ONE*, *8*(6), 16–18. http://doi.org/10.1371/journal.pone.0065055

Pimprikar, P., Carbonnel, S., Paries, M., Katzer, K., Klingl, V., Bohmer, M. J. J., … Gutjahr, C. (2016). A CCaMK-CYCLOPS-DELLA Complex Activates Transcription of RAM1 to Regulate Arbuscule Branching. *Current Biology*, *26*(8), 987–998. http://doi.org/http://dx.doi.org/10.1016/j.cub.2016.01.069

Plett, J. M., Daguerre, Y., Wittulsky, S., Vayssières, A., Deveau, A., Melton, S. J., … Martin, F. (2014). Effector MiSSP7 of the mutualistic fungus Laccaria bicolor stabilizes the Populus JAZ6 protein and represses jasmonic acid (JA) responsive genes. *Proceedings of the National Academy of Sciences of the United States of America*, *111*(22), 8299–304. http://doi.org/10.1073/pnas.1322671111

Plett, J. M., & Martin, F. (2011). Blurred boundaries: Lifestyle lessons from ectomycorrhizal fungal genomes. *Trends in Genetics*, *27*(1), 14–22. http://doi.org/10.1016/j.tig.2010.10.005

Plett, J. M., & Martin, F. (2012). Poplar root exudates contain compounds that induce the expression of MiSSP7 in Laccaria bicolor. *Plant Signaling & Behavior*, *7*(1), 12–5. http://doi.org/10.4161/psb.7.1.18357

Radutoiu, S., Madsen, L. H., Madsen, E. B., Felle, H. H., Umehara, Y., Gronlund, M., … Stougaard, J. (2003). Plant recognition of symbiotic bacteria requires two LysM receptor-like kinases. *Nature*, *425*(6958), 585–592. Journal Article. http://doi.org/10.1038/nature02039

Rasmussen, S. R., Füchtbauer, W., Novero, M., Volpe, V., Malkov, N., Genre, A., … Radutoiu, S. (2016). Intraradical colonization by arbuscular mycorrhizal fungi triggers induction of a lipochitooligosaccharide receptor. *Scientific Reports*, *6*(May), 29733. http://doi.org/10.1038/srep29733

Reddy, P. M., María, M. R., & Soto, D. L. D. (2007). Flavonoids as Signaling Molecules and Regulators of Root Nodule Development. *Dynamic Soil, Dynamic Plant*, *Global Sci*.

Redecker, D., Kodner, R., & Graham, L. E. (2000). Glomalean fungi from the Ordovician. *Science (New York, N.Y.)*, *289*(5486), 1920–1921. http://doi.org/10.1126/science.289.5486.1920

Redecker, D., Schüßler, A., Stockinger, H., Stürmer, S. L., Morton, J. B., & Walker, C. (2013). An evidence-based consensus for the classification of arbuscular mycorrhizal fungi (Glomeromycota). *Mycorrhiza*, *23*, 515–531. http://doi.org/10.1007/s00572-013-0486-y

Remigi, P., Zhu, J., Young, J. P. W., & Masson-Boivin, C. (2015). Symbiosis within symbiosis: evolving nitrogen-fixing legume symbionts. *Trends in Microbiology*. http://doi.org/10.1016/j.tim.2015.10.007

Remy, W., Taylor, T. N., Hass, H., & Kerp, H. (1994). Four hundred-million-year-old vesicular arbuscular mycorrhizae. *Proceedings of the National Academy of Sciences of the United States of America*, *91*(December), 11841–11843. http://doi.org/10.1073/pnas.91.25.11841

Rey, T., Chatterjee, A., Buttay, M., Toulotte, J., & Schornack, S. (2015). Medicago truncatula symbiosis mutants affected in the interaction with a biotrophic root pathogen. *New Phytologist*, *206*, 497–500. http://doi.org/10.1111/nph.13233

Rey, T., Nars, A., Bonhomme, M., Bottin, A., Huguet, S., Balzergue, S., … Jacquet, C. (2013). NFP, a LysM protein controlling Nod factor perception, also intervenes in Medicago truncatula resistance to pathogens. *New Phytologist*, *198*, 875–886. http://doi.org/10.1111/nph.12198

Sayed, W. F., & Wheeler, C. T. (1999). Effect of the flavonoid quercetin on culture and isolation ofFrankia fromCasuarina root nodules. *Folia Microbiologica*, *44*(1), 59–62. http://doi.org/10.1007/BF02816223

Scaffidi, A., Waters, M. T., Sun, Y. K., Skelton, B. W., Dixon, K. W., Ghisalberti, E. L., … Smith, S. M. (2014). Strigolactone Hormones and Their Stereoisomers Signal through Two Related Receptor Proteins to Induce Different Physiological Responses in Arabidopsis. *Plant Physiology*, *165*(3), 1221–1232. http://doi.org/10.1104/pp.114.240036

Schauser, L., Roussis, A., Stiller, J., & Stougaard, J. (1999). A plant regulator controlling development of symbiotic root nodules. *Nature*, *402*(6758), 191–195. Journal Article. Retrieved from <Go to ISI>://000083716400053

Shimizu, T., Nakano, T., Takamizawa, D., Desaki, Y., Ishii-Minami, N., Nishizawa, Y., … Shibuya, N. (2010). Two LysM receptor molecules, CEBiP and OsCERK1, cooperatively regulate chitin elicitor signaling in rice. *Plant Journal*, *64*, 204–214. http://doi.org/10.1111/j.1365-313X.2010.04324.x

Simon, L., Bousquet, J., Levesque, R. C., & Lalonde, M. (1993). Origin and diversification of endomycorrhizal fungi and coincidence with vascular land plants. *Nature*, *365*, 67–69.

Smith, S. E., & Read, D. (2008). *Mycorrhizal Symbiosis*. *Mycorrhizal Symbiosis* (3rd ed.). New York: Academic Press. http://doi.org/10.1016/B978-012370526-6.50015-5

Soto, M. J., Fernández-Aparicio, M., Castellanos-Morales, V., García-Garrido, J. M., Ocampo, J. a., Delgado, M. J., & Vierheilig, H. (2010). First indications for the involvement of strigolactones on nodule formation in alfalfa (Medicago sativa). *Soil Biology and Biochemistry*, *42*(2), 383–385. http://doi.org/10.1016/j.soilbio.2009.11.007

Soyano, T., & Hayashi, M. (2014). Transcriptional networks leading to symbiotic nodule organogenesis. *Current Opinion in Plant Biology*, *20*, 146–154. http://doi.org/10.1016/j.pbi.2014.07.010

Steinkellner, S., Lendzemo, V., Langer, I., Schweiger, P., Khaosaad, T., Toussaint, J. P., & Vierheilig, H. (2007). Flavonoids and strigolactones in root exudates as signals in symbiotic and pathogenic plant-fungus interactions. *Molecules*, *12*(7), 1290–1306. http://doi.org/10.3390/12071290

Sun, J., Miller, J. B., Granqvist, E., Wiley-Kalil, A., Gobbato, E., Maillet, F., … Oldroyd, G. E. D. (2015). Activation of Symbiosis Signaling by Arbuscular Mycorrhizal Fungi in Legumes and Rice. *The Plant Cell*, *27*(March), 823–838. http://doi.org/10.1105/tpc.114.131326

Sutton, J. M., Lea, E. J., & Downie, J. A. (1994). The nodulation-signaling protein NodO from Rhizobium leguminosarum biovar viciae forms ion channels in membranes. *Proceedings of the National Academy of Sciences of the United States of America*, *91*(21), 9990–4. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/7524090

Tanaka, K., Cho, S. H., Lee, H., Pham, A. Q., Batek, J. M., Cui, S., … Stacey, G. (2015). Effect of lipo-chitooligosaccharide on early growth of C<inf>4</inf> grass seedlings. *Journal of Experimental Botany*, *66*(19), 5727–5738. http://doi.org/10.1093/jxb/erv260

Tedersoo, L., May, T. W., & Smith, M. E. (2010). Ectomycorrhizal lifestyle in fungi: Global diversity, distribution, and evolution of phylogenetic lineages. *Mycorrhiza*, *20*(4), 217–263. http://doi.org/10.1007/s00572-009-0274-x

Tedersoo, L., & Smith, M. E. (2013). Lineages of ectomycorrhizal fungi revisited: Foraging strategies and novel lineages revealed by sequences from belowground. *Fungal Biology Reviews*, *27*(3–4), 83–99. http://doi.org/10.1016/j.fbr.2013.09.001

Tisa, L. S., Beauchemin, N., Gtari, M., Sen, A., & Wall, L. G. (2013). What stories can the Frankia genomes start to tell us? *Journal of Biosciences*, *38*(4), 719–726. http://doi.org/10.1007/s12038-013-9364-1

Tisserant, E., Malbreil, M., Kuo, A., Kohler, A., Symeonidi, A., Balestrini, R., … Martin, F. (2013). Genome of an arbuscular mycorrhizal fungus provides insight into the oldest plant symbiosis. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(50), 20117–22. http://doi.org/10.1073/pnas.1313452110

Tsai, S. M., & Phillips, D. A. (1991). Flavonoids released naturally from alfalfa promote development of symbiotic glomus spores in vitro. *Applied and Environmental Microbiology*, *57*(5), 1485–1488.

Tsuzuki, S., Handa, Y., Takeda, N., & Kawaguchi, M. (2016). Strigolactone-induced putative secreted protein 1 is required for the establishment of symbiosis by the arbuscular mycorrhizal fungus Rhizophagus irregularis. *Molecular Plant-Microbe Interactions : MPMI*, *29*(4), 1–59. http://doi.org/10.1094/MPMI-10-15-0234-R

Van Loon, L. C. (2007). Plant responses to plant growth-promoting rhizobacteria. *European Journal of Plant Pathology*, *119*(3), 243–254. http://doi.org/10.1007/s10658-007-9165-1

Van Zeijl, A., Liu, W., Xiao, T. ting, Kohlen, W., Yang, W.-C., Bisseling, T., & Geurts, R. R. (2015). The strigolactone biosynthesis gene DWARF27 is co-opted in rhizobium symbiosis. *BMC Plant Biology*, *15*(1), 260. http://doi.org/DOI 10.1186

Vayssières, A., Pěnčík, A., Felten, J., Kohler, A., Ljung, K., Martin, F. M., & Legue, V. (2015). Development of the Populus-Laccaria bicolor ectomycorrhiza modifies root auxin metabolism, signalling and response. *Plant Physiology*, *169*(September), pp.114.255620. http://doi.org/10.1104/pp.114.255620

Vogel, J. T., Walter, M. H., Giavalisco, P., Lytovchenko, A., Kohlen, W., Charnikhova, T., … Klee, H. J. (2010). SlCCD7 controls strigolactone biosynthesis, shoot branching and mycorrhiza-induced apocarotenoid formation in tomato. *The Plant Journal : For Cell and Molecular Biology*, *61*(2), 300–11. http://doi.org/10.1111/j.1365-313X.2009.04056.x

Wang, B., & Qiu, Y. L. (2006). Phylogenetic distribution and evolution of mycorrhizas in land plants. *Mycorrhiza*, *16*(5), 299–363. http://doi.org/10.1007/s00572-005-0033-6

Wasson, A. P., Pellerone, F. I., & Mathesius, U. (2006). Silencing the flavonoid pathway in Medicago truncatula inhibits root nodule formation and prevents auxin transport regulation by rhizobia. *The Plant Cell*, *18*(7), 1617–1629. http://doi.org/10.1105/tpc.105.038232

Weston, L. A., & Mathesius, U. (2013). Flavonoids: Their Structure, Biosynthesis and Role in the Rhizosphere, Including Allelopathy. *Journal of Chemical Ecology*, *39*(2), 283–297. http://doi.org/10.1007/s10886-013-0248-5

Wild, M., Davière, J.-M., Cheminant, S., Regnault, T., Baumberger, N., Heintz, D., … Achard, P. (2012). The Arabidopsis DELLA RGA-LIKE3 is a direct target of MYC2 and modulates jasmonate signaling responses. *Plant Cell*, *24*, 3307–19. http://doi.org/10.1105/tpc.112.101428

Yang, D.-L., Yao, J., Mei, C.-S., Tong, X.-H., Zeng, L.-J., Li, Q., … He, S. Y. (2012). Plant hormone jasmonate prioritizes defence over growth by interfering with gibberellin signaling cascade. *Proceedings of the National Academy of Sciences*, *109*(19), E1192–E1200. http://doi.org/10.1073/pnas.1201616109

Yoshida, S., Kameoka, H., Tempo, M., Akiyama, K., Umehara, M., Yamaguchi, S., … Shirasu, K. (2012). The D3 F-box protein is a key component in host strigolactone responses essential for arbuscular mycorrhizal symbiosis. *New Phytologist*, *196*(4), 1208–1216. http://doi.org/10.1111/j.1469-8137.2012.04339.x

Young, N. D., Debellé, F., Oldroyd, G. E. D., Geurts, R., Cannon, S. B., Udvardi, M. K., … Roe, B. a. (2011). The Medicago genome provides insight into the evolution of rhizobial symbioses. *Nature*, *480*(7378), 520–524. JOUR. http://doi.org/10.1038/nature10625

Yu, N., Luo, D., Zhang, X., Liu, J. J., Wang, W., Jin, Y., … Wang, E. (2014). A DELLA protein complex controls the arbuscular mycorrhizal symbiosis in plants. *Cell Research*, *24*(1), 130–3. http://doi.org/10.1038/cr.2013.167

Zhang, X., Dong, W., Sun, J., Feng, F., Deng, Y., He, Z., … Wang, E. (2015). The receptor kinase CERK1 has dual functions in symbiosis and immunity signalling. *Plant Journal*, *81*(2), 258–267. http://doi.org/10.1111/tpj.12723

Zhang, Y., van Dijk, A. D. J., Scaffidi, A., Flematti, G. R., Hofmann, M., Charnikhova, T., … Bouwmeester, H. J. (2014). Rice cytochrome P450 MAX1 homologs catalyze distinct steps in strigolactone biosynthesis. *Nature Chemical Biology*, *10*(12), 1028–33. http://doi.org/10.1038/nchembio.1660

Zhao, L.-H., Zhou, X. E., Yi, W., Wu, Z., Liu, Y., Kang, Y., … Xu, H. E. (2015). Destabilization of strigolactone receptor DWARF14 by binding of ligand and E3-ligase signaling effector DWARF3. *Cell Research*, *25*(11), 1219–1236. http://doi.org/10.1038/cr.2015.122

Zuanazzi, J. A. S., Clergeot, P. H., Quirion, J. C., Husson, H. P., Kondorosi, A., & Ratet, P. (1998). Production of Sinorhizobium meliloti nod gene activator and repressor flavonoids from Medicago sativa roots. *Molecular Plant-Microbe Interactions*, *11*(8), 784–794. http://doi.org/10.1094/MPMI.1998.11.8.784

CHAPTER 3

3

Comparative genomics of the non-legume Parasponia reveals insights into evolution of nitrogen-fixing rhizobium symbioses

Robin van Velzen[1†], Rens Holmer[1,2†], Fengjiao Bu[1‡], Luuk Rutten[1‡], Arjan van Zeijl[1], Wei Liu[3], Luca Santuari[1], Qingqin Cao[1,4], Trupti Sharma[1,] Defeng Shen[1], Yuda Roswanjaya[1], Titis A.K. Wardhani[1], Maryam Seifi Kalhor[1], Joelle Jansen[1], D. Johan van den Hoogen[1], Berivan Güngör[1], Marijke Hartog[1], Jan Hontelez[1], Jan Verver[1], Wei-Cai Yang[3], Elio Schijlen[5], Rimi Repin[6], Menno Schilthuizen[7,8], M. Eric Schranz[9], Renze Heidstra[1], Kana Miyata[1], Elena Fedorova[1], Wouter Kohlen[1], Ton Bisseling[1], Sandra Smit[2], & Rene Geurts[1*]

[1]Wageningen University, Department of Plant Science, Laboratory of Molecular Biology, Wageningen, The Netherlands.[2]Wageningen University, Department of Plant Science, Bioinformatics Group, Wageningen, The Netherlands. [3]Institute of Genetics and Developmental Biology, Chinese Academy of Sciences, Beijing, China. [4]Beijing University of Agriculture, College of Biological Science and Engineering & Beijing Collaborative Innovation Center for Eco-Environmental Improvement with Forestry and Fruit trees, Beijing, China. [5]Wageningen UR, Plant Research, BU Bioscience, Wageningen, The Netherlands. [6]Sabah Parks, Kota Kinabalu, Sabah, Malaysia. [7]Naturalis Biodiversity Center, Leiden, The Netherlands. [8]Institute for Tropical Biology and Conservation, Universiti Malaysia Sabah, Malaysia. [9]Wageningen University, Department of Plant Science, Laboratory of Biosystematics, Wageningen, The Netherlands.
[†,‡]These authors contributed equally

This Chapter is published in PNAS, May 15, 2018 115 (20) E4700-E4709; https://doi. org/10.1073/pnas.1721395115

Abstract

Nodules harboring nitrogen-fixing rhizobia are a well-known trait of legumes, but nodules also occur in other plant lineages either with rhizobia or the actinomycete *Frankia* as microsymbiont. It is generally assumed that nodulation evolved independently multiple times. However, molecular genetic support for this hypothesis is lacking, as the genetic changes underlying nodule evolution remain elusive. We conducted genetic and comparative genomics studies using *Parasponia* species (Cannabaceae), the only non-legumes that can establish nitrogen-fixing nodules with rhizobium. Intergeneric crosses between *Parasponia andersonii* and its non-nodulating relative *Trema tomentosa* demonstrated that nodule organogenesis, but not intracellular infection, is a dominant genetic trait. Comparative transcriptomics of *P. andersonii* and the legume *Medicago truncatula* revealed utilization of at least 290 orthologous symbiosis genes in nodules. Among these are key genes that in legumes are essential for nodulation, including *NODULE INCEPTION* (*NIN*) and *RHIZOBIUM-DIRECTED POLAR GROWTH (RPG)*. Comparative analysis of genomes from three *Parasponia* species and related non-nodulating plant species show evidence of parallel loss in non-nodulating species of putative orthologs of *NIN*, *RPG*, and *NOD FACTOR PERCEPTION*. Parallel loss of these symbiosis genes indicates that these non-nodulating lineages lost the potential to nodulate. Taken together, our results challenge the view that nodulation evolved in parallel and raises the possibility that nodulation originated ~100 million years ago in a common ancestor of all nodulating plant species, but was subsequently lost in many descendant lineages. This will have profound implications for translational approaches aimed at engineering nitrogen-fixing nodules in crop plants.

Introduction

Nitrogen sources such as nitrate or ammonia are key nutrients for plant growth, but their availability is frequently limited. Some plant species in the related orders Fabales, Fagales, Rosales, and Cucurbitales -collectively known as the nitrogen-fixing clade- can overcome this limitation by establishing a nitrogen-fixing endosymbiosis with either *Frankia* or rhizobium bacteria (Soltis et al. 1995). These symbioses require specialized root organs, known as nodules, that provide optimal physiological conditions for nitrogen fixation (Udvardi and Poole 2013). For example, nodules of legumes (Fabaceae, order Fabales) contain a high concentration of hemoglobin that is essential to control oxygen homeostasis and protect the rhizobial nitrogenase enzyme complex from oxidation (Ott et al. 2005; Udvardi and Poole 2013). Legumes, such as soybean (*Glycine max*), common bean (*Phaseolus vulgaris*), and peanut (*Arachis hypogaea*) represent the only crops that possess nitrogen-fixing nodules, and engineering this trait in other crop plants is a long-term vision in sustainable agriculture (Burrill and Hansen 1917; Stokstad 2016).

Nodulating plants represent ~10 related clades that diverged >100 million years ago, supporting a shared evolutionary origin of the underlying capacity for this trait (Soltis et al. 1995). Nevertheless, these nodulating clades are interspersed with many non-nodulating lineages. This has led to two hypotheses explaining the evolution of nodulation (Soltis et al. 1995). (i) Nodulation has a single origin in the root of the nitrogen-fixation clade, followed by multiple independent losses. (ii) Nodulation originated independently multiple times, preceded by a single hypothetical predisposition event in a common ancestor of the nitrogen-fixing fixation clade. The latter of these hypotheses is most widely accepted (Swensen 1996; Doyle 1998, 2011; Werner et al. 2014; Li et al. 2015; Doyle 2016; Martin, Uroz, and Barker 2017).

Genetic dissection of rhizobium symbiosis in two legume models -*Medicago truncatula* (medicago) and *Lotus japonicus* (lotus)- has uncovered symbiosis genes that are essential for nodule organogenesis, bacterial infection, and nitrogen fixation (Dataset S1). These include genes encoding LysM-type receptors that perceive rhizobial lipo-chitooligosaccharides (LCOs, also known as Nod factors) and transcriptionally activate the *NODULE INCEPTION* (*NIN*) transcription factor (Limpens et al. 2003; Madsen et al. 2003; Radutoiu et al. 2003; J. F. Arrighi et al. 2006; Marsh et al. 2007; Broghammer et al. 2012). Expression of *NIN* is essential and sufficient to set in motion nodule organogenesis (Schauser et al. 1999; Marsh et al. 2007; Soyano et al. 2013; Vernié et al. 2015). Some symbiosis genes have been co-opted from the more ancient and widespread arbuscular mycorrhizal symbiosis (Parniske 2008; Oldroyd 2013). However, causal genetic differences between nodulating and non-nodulating species have not been identified (Rene Geurts, Xiao, and Reinhold-Hurek 2016).

To obtain insight in the molecular genetic changes underlying evolution of nitrogen-fixing root nodules we conducted comparative studies using *Parasponia* (Cannabaceae, order Rosales). The genus *Parasponia* is the only lineage outside the legume family establishing a nodule symbiosis with rhizobium (Clason 1936; Trinick 1973; Akkermans, Abdulkadir, and Trinick 1978; Becking 1992). Similar as shown for legumes, nodule formation in *Parasponia* is initiated by rhizobium-secreted LCOs (Marvel, Torrey, and Ausubel 1987; R. Op den Camp et al. 2011; Granqvist et al. 2015). This suggests that *Parasponia* and legumes utilize a similar set of genes to control nodulation, but the extent of common gene utilization between distantly related nodulating species remains unknown. The genus *Parasponia* represents a clade of five species that is phylogenetically embedded in the closely related *Trema* genus (Yang et al. 2013). Like *Parasponia* and most other land plants, *Trema* species can establish an arbuscular mycorrhizal symbiosis **(SI Appendix, Figure. S1)**. However, they are non-responsive to rhizobium LCOs and do not form nodules (Becking 1992; Granqvist et al. 2015). Taken together, *Parasponia* is an excellent

system for comparative studies with legumes and non-nodulating *Trema* species to provide insights into the molecular genetic changes underlying evolution of nitrogen-fixing root nodules.

Results

Nodule organogenesis is a genetically dominant trait

First, we took a genetics approach for understanding the rhizobium symbiosis trait of *Parasponia* by making intergeneric crosses **(SI Appendix, Table S1)**. Viable F_1 hybrid plants were obtained only from the cross *Parasponia andersonii* (2n=20) x *Trema tomentosa* (2n=4x=40) **(Figure. 1A, SI Appendix, Figure. S2)**. These triploid hybrids (2n=3x=30) were infertile, but could be propagated clonally. We noted that F_1 hybrid plants formed root nodules when grown in potting soil, similar as earlier observations for *P. andersonii* (R. H. M. Op den Camp et al. 2012). To further investigate the nodulation phenotype of these hybrid plants, clonally propagated plants were inoculated with two different strains; *Bradyrhizobium elkanii* strain WUR3 (R. H. M. Op den Camp et al. 2012) or *Mesorhizobium plurifarium* strain BOR2. The latter strain was isolated from the rhizosphere of *Trema orientalis* in Malaysian Borneo and showed to be an effective nodulator of *P. andersonii* **(SI Appendix, Figure. S3)**. Both strains induced nodules on F_1 hybrid plants **(Figure. 1B,D,E; SI Appendix, Figure. S4)** but, as expected, not on *T. tomentosa*, nor on any other *Trema* species investigated. Using an acetylene reduction assay we noted that, in contrast to *P. andersonii* nodules, in F_1 hybrid nodules of plant H9 infected with *M. plurifarium* BOR2 there is no nitrogenase activity **(Figure. 1C)**. To further examine this discrepancy, we studied the cytoarchitecture of these nodules. In *P. andersonii* nodules, apoplastic *M. plurifarium* BOR2 colonies infect cells to form so-called fixation threads (**Figure. 1F,H-J)**, whereas in F_1 hybrid nodules these colonies remain apoplastic, and fail to establish intracellular infections **(Figure. 1G,K)**. To exclude the possibility that the lack of intracellular infection is caused by heterozygosity of *P. andersonii* where only a nonfunctional allele was transmitted to the F_1 hybrid genotype, or by the particular rhizobium strain used for this experiment, we examined five independent F_1 hybrid plants either inoculated with *M. plurifarium* BOR2 or *B. elkanii* WUR3. This revealed a lack of intracellular infection structures in nodules of all F_1 hybrid plants tested, irrespective which of both rhizobium strains was used **(Figure. 1G,K, SI Appendix, Figure. S4)**, confirming that heterozygosity of *P. andersonii* does not play a role in the F_1 hybrid infection phenotype. These results suggest, at least partly, independent genetic control of nodule organogenesis and rhizobium infection. Because F_1 hybrids are nodulated with similar efficiency as *P. andersonii* **(Figure. 1B)**, we conclude that the network controlling nodule organogenesis is genetically dominant.

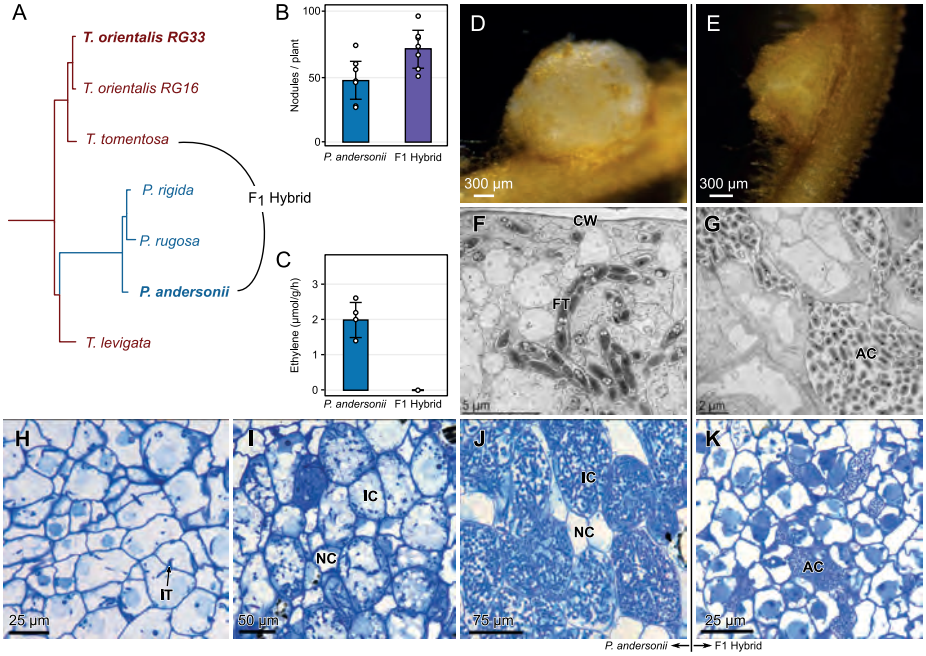


Figure 1. Nodulation phenotype of *Parasponia andersonii* and interspecific *P. andersonii* x *Trema tomentosa* F_1 hybrid plants. **A** Phylogenetic reconstruction based on whole chloroplast of *Parasponia* and *Trema*. The *Parasponia* lineage (blue) is embedded in the *Trema* genus (red). Species selected for interspecific crosses are indicated, species used for reference genome assembly are in bold. All nodes had a posterior probability of 1. **B** Mean number of nodules on roots of *P. andersonii* and F_1 hybrid plants (n=7). **C** Mean nitrogenase activity in acetylene reductase assay of *P. andersonii* and F_1 hybrid nodules (n=4). Barplot error bars indicate standard deviations; dots represent individual measurements **D** *P. andersonii* nodule. **E** F_1 hybrid nodule. **F,G** Ultrastructure of nodule tissue of *P. andersonii* **F** and F_1 hybrid **G**. Note the intracellular fixation thread (FT) in the cell of *P. andersonii* in comparison with the extracellular, apoplastic colonies of rhizobia (AC) in the F_1 hybrid nodule. **H-J** Light microscopy images of *P. andersonii* nodules in three subsequent developmental stages. **H** Stage 1: initial infection threads (IT) enter the host cells. **I** Stage 2: progression of rhizobium infection in nodule host cell, **J** Stage 3: nodule cells completely filled with fixation threads. Note difference in size between the infected (IC) and non-infected cells (NC). **K** Light microscopy image of F1 hybrid nodule cells. Note rhizobium colonies in apoplast, surrounding the host cells (AC) Nodules have been analysed 6 weeks post inoculation with *Mesorhizobium plurifarium* BOR2. Abbreviations: FT: fixation thread, CW: cell wall, AC: apoplastic colony of rhizobia, IT: infection threads, IC: infected cell, NC: non-infected cell.

Parasponia and *Trema* genomes are highly similar

Based on preliminary genome size estimates using FACS measurements, three *Parasponia* and five *Trema* species were selected for comparative genome analysis **(SI Appendix, Table S2)**. K-mer analysis of medium-coverage genome sequence data (~30x) revealed that all genomes had low levels of heterozygosity, except those of *Trema levigata* and *T. orientalis* accession RG16 **(SI Appendix, Figure. S5)**.

Based on these k-mer data we also generated more accurate estimates of genome sizes. Additionally, we used these data to assemble chloroplast genomes based on which we obtained additional phylogenetic evidence that *T. levigata* is sister to *Parasponia* **(Figure. 1A, SI Appendix, Figure. S6-8)**. Graph-based clustering of repetitive elements in the genomes (calibrated with the genome size estimates based on k-mers) revealed that all selected species contain roughly 300 Mb of non-repetitive sequence, and a variable repeat content that correlates with the estimated genome size that ranges from 375 to 625 Mb **(SI Appendix, Figure. S9, Table S3).** Notably, we found a *Parasponia*-specific expansion of ogre/tat LTR retrotransposons comprising 65 to 85 Mb **(SI Appendix, Figure. S9b)**. We then generated annotated reference genomes using high-coverage (~125X) sequencing of *P. andersonii* (accession WU1) (R. Op den Camp et al. 2011) and *T. orientalis* accession RG33 **(SI Appendix, Tables S4-5)**. These species were selected based on their low heterozygosity levels in combination with relatively small genomes. *T. tomentosa* was not used for a high-quality genome assembly because it is an allotetraploid **(SI Appendix, Figure. S5, Tables S2-3)**.

We generated orthogroups for *P. andersonii* and *T. orientalis* genes and six other Eurosid species, including arabidopsis (*Arabidopsis thaliana*) and the legumes medicago and soybean. From both *P. andersonii* and *T. orientalis* approximately 35,000 genes could be clustered into >20,000 orthogroups **(Dataset S2, SI Appendix, Table S6,** note that there can be multiple orthologous gene pairs per orthogroup). Within these orthogroups we identified 25,605 *P. andersonii* - *T. orientalis* orthologous gene pairs based on phylogenetic analysis as well as whole genome alignments (**SI Appendix, Table S6)**. These orthologous gene pairs had a median percentage nucleotide identity of 97% for coding regions **(SI Appendix, Figure. S10-11).** This further supports the recent divergence of the two species and facilitates their genomic comparison.

Common utilization of symbiosis genes in *Parasponia* and medicago

To assess commonalities in the utilization of symbiosis genes in *Parasponia* species and legumes we employed two strategies. First, we performed phylogenetic analysis of close homologs of genes that were characterized to function in legume-rhizobium symbiosis. This revealed that *P. andersonii* contains putative orthologs of the vast majority of these legume symbiosis genes (96 out of 126; **Dataset S1, S3)**. Second, we compared the sets of genes with enhanced expression in nodules of *P.andersonii* and medicago. RNA sequencing of *P. andersonii* nodules revealed 1,719 genes that are functionally annotated and have a significantly enhanced expression level (fold change > 2, $p < 0.05$, DESeq2 Wald test) in any of three nodule developmental stages compared with uninoculated roots (**SI Appendix, Figure. S12, Dataset S4)**. For medicago, we generated a comparable data set of

2,753 nodule-enhanced genes based on published RNA sequencing data (Roux et al. 2014a). We then determined the overlap of these two gene sets based on orthogroup membership and found that 382 orthogroups comprise both *P. andersonii* and medicago nodule-enhanced genes. This number is significantly larger than is to be expected by chance (permutation test, $p < 0.00001$) **(Dataset S5, SI Appendix, Figure. S13)**. Based on phylogenetic analysis of these orthogroups we found that in 290 cases putative orthologs have been utilized in both *P. andersonii* and medicago root nodules (Dataset S5, S6). Among these 290 commonly utilized genes are 26 putative orthologs of legume symbiosis genes; e.g. the LCO-responsive transcription factor *NIN* and its downstream target *NUCLEAR TRANSCRIPTION FACTOR-YA1* (*NFYA1*) that are essential for nodule organogenesis (Schauser et al. 1999; Combier

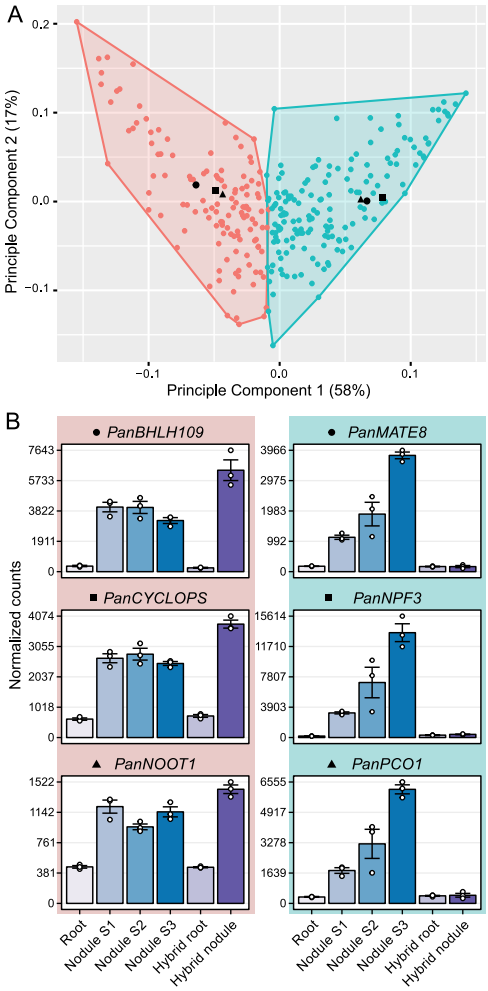


Figure 2. Clustering of commonly utilized symbiosis genes based on expression profile in *Parasponia andersonii*. **A** Principal component analysis plot of the expression profile of 290 commonly utilized symbiosis genes in 18 transcriptome samples: *P. andersonii* roots and nodules (stage 1-3), hybrid roots and nodules (line H9). All samples have three biological replicates. First two components are shown, representing 75% of the variation in all samples. Colors indicate clusters (K-means clustering using Pearson correlation as distance measure, k=2) of genes with similar expression patterns. The three genes with the highest pearson correlation to the cluster centroids are indicated as black dots, triangles, and squares, and their expression profiles are given in panel **B**. Cluster 1 (pink) represents genes related to nodule organogenesis: these genes are upregulated in both *P. andersonii* and hybrid nodules. Cluster 2 (green) represents genes related to infection and fixation: these genes are highly upregulated in *P. andersonii* nodules but do not respond in the hybrid nodule. *PanBHLH109: BASIC HELIX-LOOP-HELIX DOMAIN CONTAINING PROTEIN 109*; *PanNOOT1: NODULE ROOT 1*; *PanMATE8: MULTI ANTIMICROBIAL EXTRUSION PROTEIN 8*; *PanNPF3*: *NITRATE/PEPTIDE TRANSPORTER FAMILY 3*; *PanPCO1: PLANT CYSTEINE OXIDASE 1.*

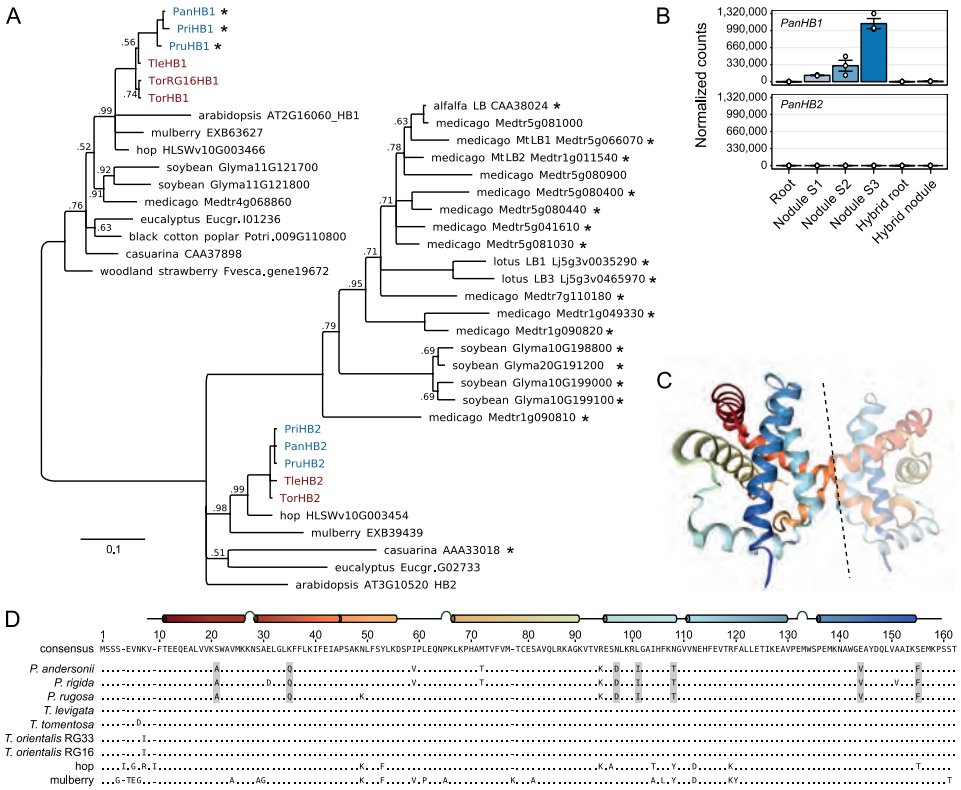


Figure 3. *Parasponia*-specific adaptations in class 1 hemoglobin protein HB1. **A** Phylogenetic reconstruction of class 1 (OG0010523) and class 2 hemoglobins (OG0002188). Symbiotic hemoglobins are marked with an asterisk; legumes and the actinorhizal plant casuarina have recruited class 2 hemoglobins for balancing oxygen levels in their nodules. Conversely, *Parasponia* has recruited a class 1 hemoglobin *PanHB1* confirming parallel evolution of symbiotic oxygen transport in this lineage. *Medicago truncatula* (Medtr); *Glycine max* (Glyma), *Populus trichocarpa* (Potri); *Fragaria vesca* (Fvesca); *Eucalyptus grandis* (Eugr); *Arabidopsis thaliana* (AT). Node values indicate posterior probabilities below 1; Scale bar represents substitutions per site. *Parasponia* marked in blue, *Trema* in red. **B** Expression profile of *PanHB1* and *PanHB2* in *P. andersonii* roots, stage 1-3 nodules, and in *P. andersonii* x *T. tomentosa* F_1 hybrid roots and nodules (line H9). Expression is given in DESeq2 normalized read counts, error bars represent standard error of three biological replicates, dots represent individual expression levels. **C** Crystal structure of the asymmetric dimer of PanHB1 as deduced by Kakar *et al.* 2011(Kakar et al. 2011). Dashed line separates the two units. **D** Protein sequence alignment of class 1 hemoglobins from *Parasponia* spp., *Trema* spp., hop (*Humulus lupulus*), and mulberry (*Morus notabilis*). Only amino acids that differ from the consensus are drawn. A linear model of the crystal structure showing alpha helices and turns is depicted above the consensus sequence. There are seven amino acids (marked grey) that consistently differ between all *Parasponia* and all other sampled species: Ala(21), Gln(35), Asp(97), Ile(101), Thr(108), Val(144), and Phe(155). These differences therefore correlate with the functional divergence between *P. andersonii* PanHB1 and *T. tomentosa* TtoHB1 (Kakar et al. 2011; Sturms et al. 2010).

et al. 2006; Soyano et al. 2013; Baudin et al. 2015), and *RHIZOBIUM DIRECTED POLAR GROWTH* (*RPG*) involved in intracellular infection (J.-F. Arrighi et al. 2008). Of these 26, five are known to function also in arbuscular mycorrhizal symbiosis (namely *VAPYRIN*, *SYMBIOTIC REMORIN,* the transcription factors *CYCLOPS* and *SAT1*, and a cysteine proteinase gene) (Kistner et al. 2005; Deguchi et al. 2007; Yano et al. 2008; Pumplin et al. 2010; Horváth et al. 2011; Murray et al. 2011; Tóth et al. 2012; Chiasson et al. 2014). To further assess whether commonly utilized genes may be co-opted from the ancient and widespread arbuscular mycorrhizal symbiosis we determined which fraction is also induced upon mycorrhization in medicago based on published RNA sequencing data (Afkhami and Stinchcombe 2016). This revealed that only 8% of the commonly utilized genes have such induction in both symbioses (Dataset S5).

By exploiting the insight that nodule organogenesis and rhizobial infection can be genetically dissected using hybrid plants we classified these commonly utilized genes into two categories based on their expression profiles in roots and nodules of both *P. andersonii* and F1 hybrids **(Figure. 2)**. The first category comprises 126 genes that are upregulated in both *P. andersonii* and hybrid nodules and that we associate with nodule organogenesis. The second category comprises 164 genes that are only upregulated in the *P. andersonii* nodule and that we therefore associate with infection and/or fixation **(Dataset S5)**. Based on these results we conclude that *Parasponia* and medicago utilize orthologous genes that commit various functions in at least two different developmental stages of the root nodule.

Lineage-specific adaptation in *Parasponia* HEMOGLOBIN 1

A notable exception to the pattern of common utilization in root nodules are the oxygen-binding hemoglobins. Earlier studies showed that *Parasponia* and legumes have recruited different hemoglobin genes (Sturms et al. 2010). Whereas legumes use class II LEGHEMOGLOBIN to control oxygen homeostasis, *Parasponia* recruited the paralogous class I HEMOGLOBIN 1 (HB1) for this function **(Figure. 3A,B)**. Biochemical studies have revealed that *P. andersonii* PanHB1 has oxygen affinities and kinetics that are adapted to their symbiotic function, whereas this is not the case for *T. tomentosa* TtoHB1 (Kakar et al. 2011; Sturms et al. 2010). We therefore examined HB1 from *Parasponia* species, *Trema* species, and other non-symbiotic Rosales species to see if these differences are due to a gain of function in *Parasponia* or a loss of function in the non-symbiotic species. Based on protein alignment we identified *Parasponia*-specific adaptations in 7 amino acids **(Figure. 3C,D)**. Among these is Ile(101) for which it is speculated to be causal for a functional change in *P. andersonii* HB1 (Kakar et al. 2011). Hemoglobin-controlled oxygen homeostasis is crucial to protect the rhizobial nitrogen-fixing enzyme complex Nitrogenase in legume rhizobium-infected nodule cells (Udvardi and Poole 2013; Ott

et al. 2005). Therefore, *Parasponia*-specific gain of function adaptations in *HB1* may have comprised an essential evolutionary step towards functional nitrogen-fixing root nodules with rhizobium endosymbionts.

Parallel loss of symbiosis genes in *Trema* and other relatives of *Parasponia*

Evolution of complex genetic traits is often associated with gene copy number variations (CNVs) (Żmieńko et al. 2014). To test if CNVs were associated with the generally assumed independent evolution of nodulation in *Parasponia*, we focused on two gene sets: (i) close homologs and putative orthologs of the genes that were characterized to function in legume-rhizobium symbiosis, and (ii) genes with a nodule-enhanced expression and functional annotation in *P. andersonii* (these sets partially overlap and add up to 1,825 genes; **SI Appendix**, **Figure. S14**). We discarded *Trema*-specific duplications as we considered them irrelevant for the nodulation

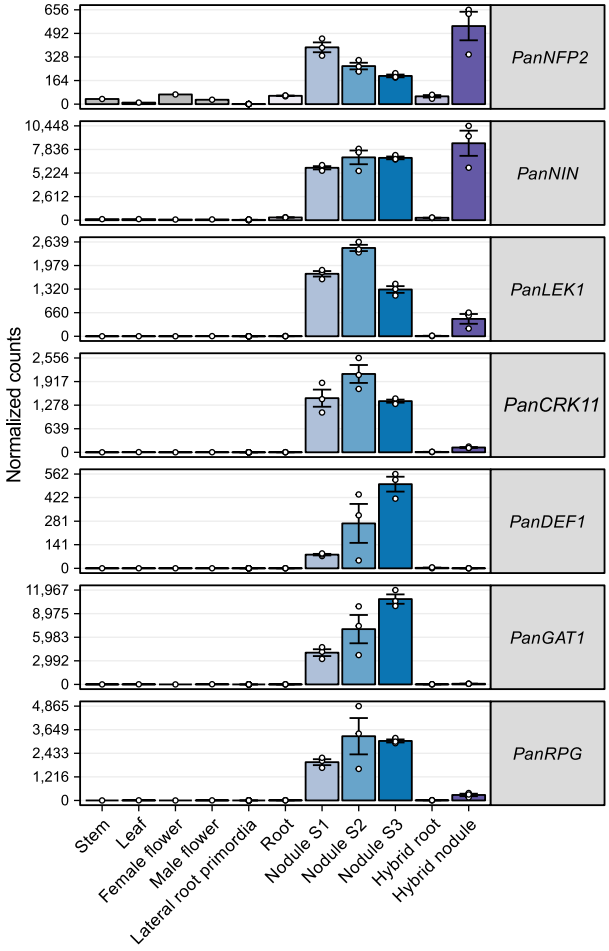


Figure 4. Expression profile of *Parasponia andersonii* symbiosis genes that are lost in *Trema* species. Expression of symbiosis genes in *P. andersonii* stem, leaf, female and male flowers, lateral root primordia, roots and 3 nodule stages (S1-3), and in F_1 hybrid roots and nodules (line H9). Expression is given in DESeq2 normalized read counts, error bars represent standard error of three biological replicates for lateral root primordia, root, and nodule samples. Dots represent individual expression levels. *PanNFP2: NOD FACTOR PERCEPTION 2*; *PanNIN: NODULE INCEPTION*; *PanLEK1: LECTIN RECEPTOR KINASE 1*; *PanCRK11: CYSTEINE-RICH RECEPTOR KINASE 11*; *PanDEF1: DEFENSIN 1*; *PanRPG: RHIZOBIUM DIRECTED POLAR GROWTH.*

phenotype. To ensure that our findings are consistent between the *Parasponia* and *Trema* genera and not due to species-specific events, we analyzed the additional draft genome assemblies of two *Parasponia* and two *Trema* species **(SI Appendix, Table S5)**. As these additional draft genomes were relatively fragmented, we sought additional support for presence absence of genes by mapping sequence reads to the *P. andersonii* and *T. orientalis* reference genomes and by genomic alignments. This procedure resulted in only 11 consistent CNVs in the 1,817 symbiosis genes examined, further supporting the recent divergence between *Parasponia* and *Trema* **(SI Appendix, Figure. S15**). Due to the dominant inheritance of nodule organogenesis in F_1 hybrid plants, we anticipated finding *Parasponia*-specific gene duplications that could be uniquely associated with nodulation. Surprisingly, we found only one consistent *Parasponia*-specific duplication in symbiosis genes; namely for a *HYDROXYCINNAMOYL-COA SHIKIMATE TRANSFERASE* (*HCT*) **(SI Appendix, Figure. S16-17).** This gene has been investigated in the legume forage crop alfalfa (*Medicago sativa)*, where it was shown that HCT expression correlates negatively with nodule organogenesis (Shadle et al. 2007; Gallego-Giraldo et al. 2014).

Therefore, we do not consider this duplication relevant for the nodulation capacity of *Parasponia*. Additionally, we identified three consistent gene losses in *Parasponia* among which is the ortholog of *EXOPOLYSACCHARIDE RECEPTOR 3* that in lotus inhibits infection of rhizobia with incompatible exopolysaccharides (Y. Kawaharada et al. 2015; Yasuyuki Kawaharada et al. 2017) **(SI Appendix, Figure. S18-20, Table S7)**. Such gene losses may have contributed to effective rhizobium infection in *Parasponia* and their presence in *T. tomentosa* could explain the lack of intracellular infection in the F1 hybrid nodules. However, they cannot explain the dominance of nodule organogenesis in the F1 hybrid.

Contrary to our initial expectations, we discovered consistent loss or pseudogenization of seven symbiosis genes in *Trema* **(SI Appendix, Figure. S21-23, Table S7)**. Based on our current sampling, these genes have a nodule-specific expression profile in *P. andersonii*, suggesting that they function exclusively in symbiosis **(Figure. 4)**. Three of these are orthologs of genes that are essential for establishment of nitrogen-fixing nodules in legumes: *NIN, RPG,* and the LysM-type LCO receptor *NFP*/*NFR5*. In the case of *NFP/NFR5*, we found two close homologs of this gene, *NFP1* and *NFP2*, a duplication that predates the divergence of legumes and *Parasponia* **(Figure. 5)**. In contrast to *NFP1*, *NFP2* is consistently pseudogenized in *Trema* species **(Figure. 5; SI Appendix, Figure. S22-23)**. In an earlier study we used RNA interference (RNAi) to target *PanNFP1* (previously named *PaNFP*), which led to reduced nodule numbers and a block of intracellular infection by rhizobia as well as arbuscular mycorrhiza (R. Op den Camp et al. 2011). However, we cannot rule out that the RNAi construct unintentionally also targeted *PanNFP2*, as both genes are ~70%

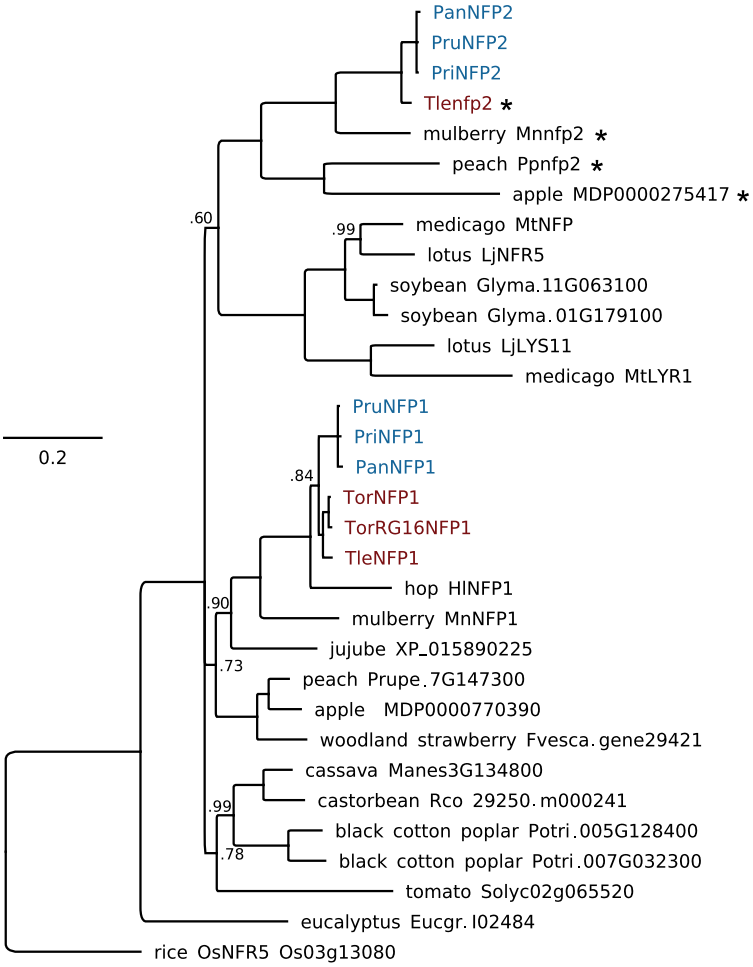


Figure 5. *Parasponia NFP2* are putative orthologs of legume LCO receptors MtNFP/LjNFR5. Phylogenetic reconstruction of the NFP/NFR5 orthogroup based on kinase domain. Protein sequences deduced from pseudogenes are marked with an asterisk. Included species: *Parasponia andersonii* (Pan); *P. rigida* (Pri); *P. rugosa* (Pru); *Trema orientalis* RG33 (Tor); *T. orientalis* RG16 (TorRG16); *T. levigata* (Tle); medicago (*Medicago truncatula*, Mt); lotus (*Lotus japonicus*, Lj); soybean (*Glycine max*, Glyma); peach (*Prunus persica*, Ppe); woodland strawberry (*Fragaria vesca*, Fvesca); black cotton poplar (*Populus trichocarpa*, Potri); eucalyptus (*Eucalyptus grandis*, Eugr); jujube (*Ziziphus jujube*), apple (*Malus* x *domestica*), mulberry (*Morus notabilis*), hops (*Humulus lupulus*), cassave (*Manihot esculenta*), rice (*Oryza sativa*), tomato (*Solanum lycopersicum*), and castor bean (*Ricinus communis*). Node numbers indicate posterior probabilities below 1, scale bar represents substitutions per site. *Parasponia* proteins are marked in blue, *Trema* in red.

identical in the 422 bp RNAi target region. Therefore, the precise functioning of both receptors in rhizobium and mycorrhizal symbiosis remains to be elucidated. Based on phylogenetic analysis the newly discovered *PanNFP2* is the ortholog of the legume *MtNFP/LjNFR5* genes encoding rhizobium LCO receptors required for nodulation, while *PanNFP1* is most likely a paralog **(Figure. 5)**. Also, *PanNFP2* is significantly higher expressed in nodules than *PanNFP1* **(SI Appendix, Figure. S25)**. Taken together, this indicates that *PanNFP2* may represent a key LCO receptor required for nodulation in *Parasponia*.

Based on expression profiles and phylogenetic relationships we postulate that also *Parasponia NIN* and *RPG* commit essential symbiotic functions similar as in other nodulating species **(Figure. 3; SI Appendix, Figure. S25-28)** (Schauser et al. 1999; Marsh et al. 2007; Borisov et al. 2003; Clavijo et al. 2015; J.-F. Arrighi et al. 2008). Compared with uninoculated roots, expression of *PanRPG* is >300 fold higher in *P. andersonii* nodules that become intracellularly infected (nodule stage 2), whereas in F_1 hybrid nodules -which are devoid of intracellular rhizobium infection- *PanRPG* this difference is less than 20-fold **(Figure. 3)**. This suggests that *PanRPG* commits a function in rhizobium infection, similar as found in medicago (J.-F. Arrighi et al. 2008). The transcription factor *NIN* has been studied in several legume species as well as in the actinorhizal plant casuarina (*Casuarina glauca*) and in all cases shown to be

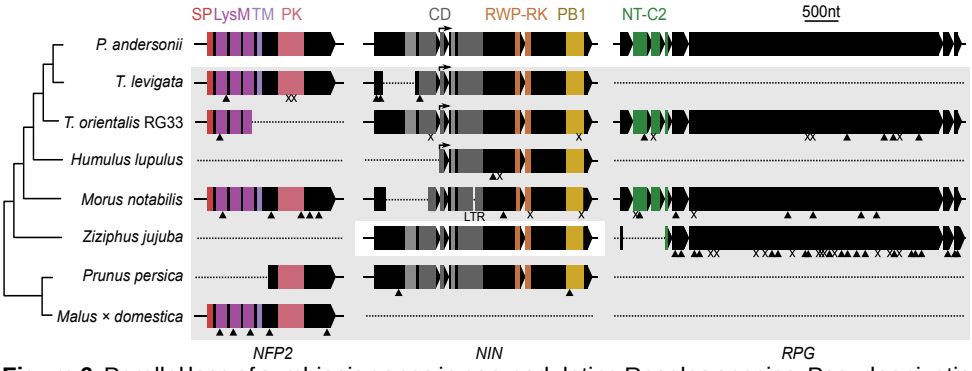


Figure 6. Parallel loss of symbiosis genes in non-nodulating Rosales species. Pseudoenization or loss of *NFP2*, *NIN,* and *RPG* in two phylogenetically independent *Trema* lineages, *Humulus lupulus* (hop), *Morus notabilis* (mulberry), *Ziziphus jujuba* (jujube), *Prunus persica* (peach), and *Malus x domestica* (apple). In *H. lupulus NIN* is pseudogenized, whereas *NFP2* and *RPG* were not found (this may due to the low N50 of the publicly available assembly). In *Z. jujuba NFP2* is lost and *RPG* is pseudogenized, but *NIN* is intact. In *Fragaria vesca* all three genes are lost (not shown). Introns are indicated but not scaled. Triangles indicate frame-shifts; X indicates premature stop codons; LTR indicates long terminal repeat retrotransposon insertion (not scaled); arrows indicate alternative transcriptional start site in *NIN*. SP = signal peptide (red); LysM: 3 Lysin Motif domains (magenta); TM = transmembrane domain (lilac); PK = protein kinase (pink); CD = 4 conserved domains (grey); RWP-RK: conserved amino acid domain (orange); PB1 = Phox and Bem1 domain (yellow); NT-C2 = N-terminal C2 domain (green).

essential for nodule organogenesis (Schauser et al. 1999; Marsh et al. 2007; Borisov et al. 2003; Clavijo et al. 2015). Loss of *NIN* and possibly *NFP2* in *Trema* species can explain the genetic dominance of nodule organogenesis in the *Parasponia* x *Trema* F1 hybrid plants.

Next, we assessed whether loss of these symbiosis genes also occurred in more distant relatives of *Parasponia*. We analysed non-nodulating species representing 6 additional lineages of the Rosales clade; namely hop (*Humulus lupulus,* Cannabaceae) (Natsume et al. 2015), mulberry (*Morus notabilis,* Moraceae) (He et al. 2013), jujube (*Ziziphus jujuba*, Rhamnaceae) (Huang et al. 2016), peach (*Prunus persica,* Rosaceae) (Verde et al. 2013), woodland strawberry (*Fragaria vesca*, Rosaceae) (Shulaev et al. 2011), and apple (*Malus x domestica,* Rosaceae) (Velasco et al. 2010). This revealed a consistent pattern of pseudogenization or loss of *NFP2, NIN* and *RPG* orthologs; the intact jujube *ZjNIN* being the only exception **(Figure. 6)**. We note that for peach *NIN* was previously annotated as protein-coding gene (Verde et al. 2013). However, based on comparative analysis of conserved exon structures we found two out-of-frame mutations **(SI Appendix, Figure. S28)**. We therefore conclude that also in peach the *NIN* gene is pseudogenized. Because the pseudogenized symbiosis genes are largely intact in most of these species and differ in their deleterious mutations, the loss of function of these essential symbiosis genes should have occurred relatively recently and in parallel in at least seven Rosales lineages.

Discussion

Here we present the nodulating non-legume *Parasponia* as a comparative system to obtain insights in molecular genetic changes underlying evolution of nitrogen-fixing root nodules. We show that nodulation is a genetically dominant trait and that *P. andersonii* and the legume medicago share a set of 290 genes that have a nodule enhanced expression profile. Among these are *NIN* and *RPG,* two genes that in legumes are essential for nitrogen-fixing root nodulation (Schauser et al. 1999; Borisov et al. 2003; Marsh et al. 2007; J.-F. Arrighi et al. 2008). Both these genes as well as a putative ortholog of the NFP/NFR5-type LysM receptor for rhizobium LCO signal molecules -named *NFP2* in *Parasponia*- are consistently pseudogenized or lost in *Trema* and other non-nodulating species of the Rosales order. This challenges the current view on the evolution of nitrogen-fixing plant-microbe symbioses.

Evolution of nodulation is generally viewed as a two-step process: first an unspecified predisposition event in the ancestor of all nodulating species, bringing species in the nitrogen-fixing clade to a precursor state for nodulation. Subsequently, nodulation originated in parallel; eight times with *Frankia* and twice with rhizobium (Soltis et al. 1995; Swensen 1996; Doyle 1998, 2011; Werner et al. 2014; Li et al. 2015; Doyle

2016; Martin, Uroz, and Barker 2017). This hypothesis is most parsimonious and suggests a minimum number of independent gains and losses of symbiosis. Based on this hypothesis it is currently assumed that non-host relatives of nodulating species are generally in a precursor state for nodulation (Werner et al. 2014).

Our results are difficult to explain under the hypothesis of parallel origins of nodulation. The functions of *NFP2, NIN,* and *RPG* currently cannot be linked to any non-symbiotic processes. Therefore it remains obscure why these symbiosis genes were maintained over an extended period of time in non-nodulating plant species, and were subsequently independently lost. Additionally, the hypothesis of parallel origins of nodulation would imply convergent recruitment of at least 290 genes to commit symbiotic functions in *Parasponia* and legumes. Because these 290 genes encode proteins with various predicted functions (e.g. from extracellular signaling receptors to sugar transporters; **Dataset S5**), as well as comprise at least two different developmental expression patterns (nodule organogenesis and intracellular infection and/or fixation; **Figure. 2, Dataset S5**), this would imply parallel evolution of a genetically complex trait.

Alternatively, the parallel loss of symbiosis genes in non-nodulating plants can be interpreted as parallel loss of nodulation (Soltis et al. 1995). Under this hypothesis nodulation possibly evolved only once in an ancestor of the nitrogen-fixing clade. Subsequently, nodulation was lost in most descendant lineages. This single gain-massive loss hypothesis fits our data better in two ways. First, a single gain explains the origin of the conserved set of at least 290 symbiosis genes utilized by both *Parasponia* and medicago, because they then result from the same ancestral recruitment event. Second, it more convincingly explains the parallel loss of symbiosis genes in non-nodulating plants, because then gene loss correlates directly with loss of nodulation. Additionally, the single gain-massive loss model eliminates the predisposition event, a theoretical concept that currently cannot be addressed experimentally. We therefore favor this alternative hypothesis over the currently most widely held assumption of parallel origins of nodulation.

Loss of nodulation is not controversial, as it is generally considered to have occurred at least 20 times in the legume family (Li et al. 2015; Werner et al. 2014). Nevertheless, the single gain-massive loss hypothesis implies many more evolutionary events than the current hypothesis of parallel gains. On the other hand, it is conceptually easier to lose a complex trait, such as nodulation, rather than to gain it (Doyle 2016). Genetic studies in legumes demonstrated that nitrogen-fixing symbioses can be abolished by a single knockout mutation in tens of different genes, among which are *NFP/NFR5, NIN*, and *RPG* **(Dataset S1)**. Because parsimony implies equal weights for gains and losses, it may therefore not be the best way to model the evolution of nodulation.

Preliminary support for the single gain-massive loss hypothesis can be found in fossil records. Putative root nodule fossils have been discovered from the late Cretaceous (approximately 84 million years ago), which corroborates our hypothesis that nodulation is much older than is generally assumed (Herendeen et al. 1999). Legumes are the oldest and most diverse nodulating lineage, however the earliest fossils that can be definitively assigned to the legume family appeared in the late Palaeocene (approximately 65 million years ago) (Bruneau et al. 2008). Notably, the age of the nodule fossils coincides with the early diversification of the nitrogen-fixing clade that has given rise to the 4 orders Fabales, Rosales, Cucurbitales, and Fagales (Li et al. 2015). As it is generally agreed that individual fossil ages provide minimum bounds for dates of origins it is therefore not unlikely that the last common ancestor of the nitrogen-fixing clade was a nodulator.

Clearly, the single gain-massive loss hypothesis that is supported by our comparative studies using *Parasponia* requires further substantiation. First, the hypothesis implies that many ancestral species in the nitrogen-fixing clade were able to nodulate. This should be further supported by fossil evidence. Second, the hypothesis implies that actinorhizal plant species maintained *NIN, RPG,* and possibly *NFP2* (the latter only in case LCOs are used as symbiotic signal) (Nguyen et al. 2016). Third, these genes should be essential for nodulation in these actinorhizal plants as well as in *Parasponia*. This can be shown experimentally, as was done for *NIN* in casuarina (Clavijo et al. 2015).

Loss of symbiosis genes in non-nodulating plant species is not absolute, as we observed a functional copy of *NIN* in jujube. This pattern is similar to the pattern of gene loss in species that lost endomycorrhizal symbiosis where occasionally endomycorrhizal symbiosis genes have been maintained in non-mycorrhizal plants (Delaux et al. 2015; Kamel et al. 2017). Conservation of *NIN* in *Ziziphus jujube* suggests that this gene has a non-symbiotic function. Contrary to *NFP2*, which is the result of a gene duplication near the origin of the nitrogen-fixing clade, functional copies of *NIN* are also present in species outside the nitrogen-fixing clade **(SI Appendix, Figure. S26)**. This suggests that these genes may have retained -at least in part- an unknown ancestral non-symbiotic function in some lineages within the nitrogen-fixing clade. Alternatively, *NIN* may have acquired a new non-symbiotic function within some lineages in the nitrogen-fixing clade.

As hemoglobin is crucial for rhizobium symbiosis in legumes (Ott et al. 2005), it is striking that *Parasponia* and legumes do not use orthologous copies of hemoglobin genes in their nodules (Sturms et al. 2010). Superficially, this seems inconsistent with a single gain of nodulation. However, hemoglobin is not crucial for all nitrogen-fixing nodule symbioses because several *Frankia* microsymbionts possess intrinsic

physical characteristics to protect the Nitrogenase enzyme for oxidation (Winship, Martin, and Sellstedt 1987; Silvester and Winship 1990; Silvester et al. 2007; Silvester, Harris, and Tjepkema 1990). In line with this, *Ceanothus* spp. (Rhamnaceae, Rosales) - which represent actinorhizal nodulating relatives of *Parasponia* - do not express a hemoglobin gene in their *Frankia*-infected nodules (Silvester and Winship 1990; Silvester et al. 2007; Silvester, Harris, and Tjepkema 1990). Consequently, hemoglobins may have been recruited in parallel after the initial gain of nodulation as parallel adaptations to rhizobium microsymbionts. Based on the fact that *Parasponia* acquired lineage-specific adaptations in HB1 that are considered to be essential for controlling oxygen homeostasis in rhizobium root nodules (Sturms et al. 2010; Kakar et al. 2011), a symbiont switch from *Frankia* to rhizobium may have occurred recently in an ancestor of the *Parasponia* lineage.

Our study provides novel leads for attempts to engineer nitrogen-fixing root nodules in agricultural crop plants. Such a translational approach is anticipated to be challenging (Rogers and Oldroyd 2014), and the only published attempt so far, describing transfer of 8 LCO signaling genes, was unsuccessful (Untergasser et al. 2012). Our results suggest that transfer of symbiosis genes may not be sufficient to obtain functional nodules. Even though F1 hybrid plants contain a full haploid genome complement of *P. andersonii* they lack intracellular infection. This may be due to haploinsufficiency of *P. andersoni*i genes in the F1 hybrid, or due to an inhibitory factor in *T. tomentosa*. For example, inhibition of intracellular infection may be the result of a dominant negative factor, or the result of heterozygosity negatively affecting the formation of e.g. LysM receptor complexes required for appropriate perception of microsymbionts. Such factors may also be present in other non-host species. Consequently, engineering nitrogen-fixing nodules may require gene knockouts in non-nodulating plants to overcome inhibition of intracellular infection. *Trema* may be the best candidate species for such a (re)engineering approach, due to its high genetic similarity with *Parasponi*a and the availability of transformation protocols (Cao et al. 2012). Therefore, the *Parasponia-Trema* comparative system may not only be suited for evolutionary studies, but also can form an experimental platform to obtain essential insights for engineering nitrogen-fixing root nodules.

Materials and Methods

Parasponia - Trema intergeneric crossing and hybrid genotyping

Parasponia and *Trema* are wind-pollinated species. A female-flowering *P. andersonii* individual WU1.14 was placed in a plastic shed together with a flowering *T. tomentosa* WU10 plant. Putative F_1 hybrid seeds were germinated (see SI Appendix, Supplementary Methods) and transferred to potting soil. To confirm the hybrid genotype a PCR marker was used that visualizes a length difference in the promoter region of *LIKE-AUXIN 1* (*LAX1*) (primers: LAX1-f: ACATGATAATTTGGGCATGCAACA, LAX1-r: TCCCGAATTTTCTACGAATTGAAA, amplicon size *P. andersonii*: 974 bp; *T. tomentosa*: 483 bp). Hybrid plant H9 was propagated *in vitro* (Davey et al. 1993; R. Op den Camp et al. 2011). The karyotype of the selected plants was determined according to Geurts and De Jong 2013 (René Geurts and de Jong 2013).

Assembly of reference genomes

Cleaned DNA sequencing reads were *de novo* assembled using ALLPATHS-LG (release 48961) (Gnerre et al. 2011) After filtering of any remaining adapters and contamination, contigs were scaffolded with two rounds of SSPACE-standard (v3.0) (Boetzer et al. 2011) with the mate-pair libraries using default settings. We used the output of the second run of SSPACE scaffolding as the final assembly (See SI Appendix, Supplementary Methods for full details and parameter choices). Validation of the final assemblies showed that 90-100% of the genomic reads mapped back to the assemblies **(SI Appendix, Table S4)**, and 94-98% of CEGMA (Parra, Bradnam, and Korf 2007) and BUSCO (Simão et al. 2015) genes were detected **(SI Appendix, Table S5)**.

Annotation of reference genomes

Repetitive elements were identified following the standard Maker-P recipe (http://weatherby.genetics.utah.edu/MAKER/wiki/index.php/Repeat_Library_Construction-Advanced accessed October 2015) as described on the GMOD site: (i) RepeatModeler with Repeatscout v1.0.5, Recon v1.08, RepeatMasker version open4.0.5, using RepBase version 20140131 (Bao, Kojima, and Kohany 2015) and TandemRepeatFinder; (ii) GenomeTools: LTRharvest and LTRdigest (Gremme, Steinbiss, and Kurtz 2013); (iii) MITEhunter with default parameters (Han and Wessler 2010). We generated species-specific repeat libraries for both *P. andersonii* and *T. orientalis* separately and combined these into a single repeat library, filtering out sequences that are >98% similar. We masked both genomes using RepeatMasker with this shared repeat library.

To aid the structural annotation we used 11 *P. andersonii* and 6 *T. orientalis* RNA sequencing datasets **(SI Appendix, Table S8)**. All RNA-seq samples were assembled *de novo* using genome-guided Trinity (Grabherr et al. 2011), resulting

in one combined transcriptome assembly per species. In addition all samples were mapped to their respective reference genomes using BWA-MEM and processed into putative transcripts using cufflinks (Trapnell et al. 2010) and transdecoder (Haas et al. 2013). As protein homology evidence, only Swiss-Prot (UniProt Consortium 2015) entries filtered for plant proteins were used. This way we only included manually verified protein sequences and prevented the incorporation of erroneous predictions. Finally, four gene-predictor tracks were used: 1) SNAP (Korf 2004), trained on *P. andersonii* transdecoder transcript annotations; 2) SNAP, trained on *T. orientalis* transdecoder transcript annotations; 3) Augustus (Stanke et al. 2008) as used in the BRAKER pipeline, trained on RNA-seq alignments (Hoff et al. 2016); 4) GeneMark-ET as used in the BRAKER pipeline, trained on RNA-seq alignments (Lomsadze, Burns, and Borodovsky 2014).

First, all evidence tracks were processed by Maker-P (Campbell et al. 2014). The results were refined with EVidenceModeler (EVM) (Haas et al. 2008), which was used with all the same tracks as Maker-P, except for the Maker-P blast tracks and with the addition of the Maker-P consensus track as additional evidence. Ultimately, EVM gene models were preferred over Maker-P gene models, except when there was no overlapping EVM gene model. Where possible, evidence of both species was used to annotate each genome (i.e. *de novo* RNA-seq assemblies of both species were aligned to both genomes).

To take maximum advantage of annotating two highly similar genomes simultaneously we developed a custom reconciliation procedure involving whole genome alignments. The consensus annotations from merging the EVM and Maker-P annotations were transferred to their respective partner genome using nucmer (Kurtz et al. 2004) and RATT revision 18 (Otto et al. 2011) (i.e. the *P. andersonii* annotation was transferred to *T. orientalis* and *vice versa*), based on nucmer whole genome alignments **(SI Appendix, Fig. S10)**. Through this reciprocal transfer, both genomes had two candidate annotation tracks. This allowed for validation of annotation differences between *P. andersonii* and *T. orientalis*, reduced technical variation, and consequently improved all downstream analyses. After automatic annotation and reconciliation 1,693 *P. andersonii* genes and 1,788 *T. orientalis* genes were manually curated. These were mainly homologs of legume symbiosis genes and genes that were selected based on initial data exploration.

To assign putative product names to the predicted genes we combined BLAST results against Swiss-Prot, TrEMBL, and nr with InterProScan results (custom script). To annotate GO terms and KEGG enzyme codes Blast2GO was used with the nr BLAST results and InterProScan results. Finally, we filtered all gene models with hits to InterPro domains that are specific to repetitive elements.

Orthogroup inference

To determine relationships between *P. andersonii* and *T. orientalis* genes, as well as with other plant species we inferred orthogroups with OrthoFinder version 0.4.0 (Emms and Kelly 2015). Since orthogroups are defined as the set of genes that are descended from a single gene in the last common ancestor of all the species being considered, they can comprise orthologous as well as paralogous genes. Our analysis included proteomes of selected species from the Eurosid clade: *Arabidopsis thaliana* TAIR10 (Brassicaceae, Brassicales) (Swarbreck et al. 2008) and *Eucalyptus grandis* v2.0 (Myrtaceae, Myrtales) from the Malvid clade (Myburg et al. 2014); *Populus trichocarpa* v3.0 (Salicaeae, Malpighiales) (Tuskan et al. 2006), legumes *Medicago truncatula* Mt4.0v1 (Young et al. 2011) and *Glycine max* Wm82.a2.v1 (Fabaceae, Fabales) (Schmutz et al. 2010), *Fragaria vesca* v1.1 (Rosaceae, Rosales) (Shulaev et al. 2011), *P. andersonii* and *T. orientalis* (Cannabaceae, Rosales) from the Fabid clade (Dataset S2). Sequences were retrieved from phytozome (www.phytozome.net).

Gene copy number variant detection

To assess orthologous and paralogous relationships between *Parasponia* and *Trema* genes, we inferred phylogenetic gene trees for all 21,959 orthogroups comprising *Parasponia* and/or *Trema* genes using the neighbor-joining clustering algorithm (Saitou and Nei 1987). Based on these gene trees, for each *Parasponia* gene its relationship to other *Parasponia* and *Trema* genes was defined as follows. 1) orthologous pair: the sister lineage is a single gene from the *Trema* genome suggesting that they are the result of a speciation event; 2) inparalog: the sister lineage is a gene from the *Parasponia* genome, suggesting that they are the result of a gene duplication event; 3) singleton: the sister lineage is a gene from a species other than *Trema*, suggesting that the *Trema* gene was lost; 4) multi-ortholog: the sister lineage comprises multiple genes from the *Trema* genome, suggesting that the latter are inparalogs. For each *Trema* gene, relationship was defined in the same way but with respect to the *Parasponia* genome **(SI Appendix, Table S6)**. Because phylogenetic analysis relies on homology we assessed the level of conservation in the multiple-sequence alignments by calculating the trident score using MstatX (https://github.com/gcollet/MstatX) (Valdar 2002). Orthogroups with a score below 0.1 were excluded from the analysis. Examination of orthogroups comprising >20 inparalogs revealed that some represented repetitive elements; these were also excluded. Finally, orthologous pairs were validated based on the whole-genome alignments used in the annotation reconciliation.

Nodule-enhanced genes

To assess gene expression in *Parasponia* nodules, RNA was sequenced from the three nodule stages described above as well as uninoculated roots (SI Appendix, Table S8). RNA-seq reads were mapped to the *Parasponia* reference genome with HISAT2 version 2.02 (Kim, Langmead, and Salzberg 2015) using an index that includes exon and splice site information in the RNA-seq alignments. Mapped reads were assigned to transcripts with featureCounts version 1.5.0 (Liao, Smyth, and Shi 2014). Normalization and differential gene expression were performed with DESeq2. Nodule enhanced genes were selected based on >2.0 fold-change and p<=0.05 in any nodule stage compared with uninoculated root controls. Genes without functional annotation or orthogroup membership or from orthogroups with low alignment scores (<0.1 trident score, see above) or representing repetitive elements were excluded from further analysis. To assess expression of *Parasponia* genes in the hybrid nodules, RNA was sequenced from nodules and uninoculated roots. Here, RNA-seq reads were mapped to a combined reference comprising two parent genomes from *P. andersonii* and *T. tomentosa*. To assess which genes are nodule-enhanced in medicago we re-analyzed published RNA-seq read data from Roux *et al.* (archived at NCBI under SRA study SRP028599) (Roux et al. 2014b). To assess which of these genes may be co-opted from the ancient and widespread arbuscular mycorrhizal symbiosis we generated a set of 575 medicago genes induced upon mycorrhization in medicago by re-analyzing published RNA-seq read data from Afkhami and Stinchcombe (archived at NCBI under SRA study SRP078249) (Afkhami and Stinchcombe 2016) Both medicago data sets were analysed as described above for *Parasponia* but using the medicago genome and annotation version 4.0v2 as reference (Young et al. 2011).

To assess common recruitment of genes in nodules from *Parasponia* and medicago we counted orthogroups comprising both *P. andersonii* and medicago nodule-enhanced genes. To assess whether this number is higher than expected by chance we performed the hypergeometric test as well as three different permutation tests where we randomized either the *Parasponia* gene set, the medicago gene set, or both sets with 10,000 permutations. We then determined putative orthology between the *Parasponia* and medicago genes within the common orthogroups based on phylogenetic analysis. *Parasponia* and medicago genes were considered putative orthogroups if they occur in the same subclade with more than 50% bootstrap support; otherwise they were considered close homologs.

Supplemental data

Supplemental data and data files belonging to this chapter are available at Proc Natl Acad Sci USA online (https://doi.org/10.1073/pnas.1721395115).

Availability of data and materials

The data reported in this study are tabulated in the additional information files; sequence data are archived at NCBI (https://www.ncbi.nlm.nih.gov) under BioProject numbers PRJNA272473 and PRJNA272482; draft genome assemblies, phylogenetic datasets, and orthogroup data are archived at the Dryad Digital Repository: https://doi.org/10.5061/dryad.fq7gv88. All analyzed data can also be browsed or downloaded through a web portal on www.parasponia.org. All custom scripts and code are available on https://github.com/holmrenser/parasponia_code.

Acknowledgments

We thank Shelley James, Thomas Marler, Giles Oldroyd, and Johan van Valkenburg for providing germplasm and Ries de Visser (IsoLife) for supporting ARA assays.

References

Afkhami, Michelle E., and John R. Stinchcombe. 2016. "Multiple Mutualist Effects on Genomewide Expression in the Tripartite Association between *Medicago Truncatula*, Nitrogen-Fixing Bacteria and Mycorrhizal Fungi." *Molecular Ecology* 25 (19): 4946–62.

Akkermans, A. D. L., S. Abdulkadir, and M. J. Trinick. 1978. "Nitrogen-Fixing Root Nodules in Ulmaceae." *Nature* 274 (5667): 190–190.

Arrighi, Jean-François, Olivier Godfroy, Françoise de Billy, Olivier Saurat, Alain Jauneau, and Clare Gough. 2008. "The *RPG* Gene of *Medicago Truncatula* Controls *Rhizobium*-Directed Polar Growth during Infection." *Proceedings of the National Academy of Sciences of the United States of America* 105 (28): 9817–22.

Arrighi, J. F., A. Barre, B. Ben Amor, A. Bersoult, L. C. Soriano, R. Mirabella, F. de Carvalho-Niebel, et al. 2006. "The *Medicago Truncatula* Lysine Motif-Receptor-like Kinase Gene Family Includes *NFP* and New Nodule-Expressed Genes." *Plant Physiology* 142 (1): 265–79.

Bao, Weidong, Kenji K. Kojima, and Oleksiy Kohany. 2015. "Repbase Update, a Database of Repetitive Elements in Eukaryotic Genomes." *Mobile DNA* 6 (June): 11.

Baudin, Maël, Tom Laloum, Agnès Lepage, Carolina Rípodas, Federico Ariel, Lisa Frances, Martin Crespi, et al. 2015. "A Phylogenetically Conserved Group of Nuclear Factor-Y Transcription Factors Interact to Control Nodulation in Legumes." *Plant Physiology* 169 (4): 2761–73.

Becking, J. H. 1992. "The Rhizobium Symbiosis of the Nonlegume *Parasponia*." In *Biological Nitrogen Fixation*, edited by G. Stacey, R. H. Burris, and H. J. Evans, 497–559. New York: Routledge, Chapman and Hall.

Boetzer, Marten, Christiaan V. Henkel, Hans J. Jansen, Derek Butler, and Walter Pirovano. 2011. "Scaffolding Pre-Assembled Contigs Using SSPACE." *Bioinformatics* 27 (4): 578–79.

Borisov, Alexey Y., Lene H. Madsen, Viktor E. Tsyganov, Yosuke Umehara, Vera A. Voroshilova, Arsen O. Batagov, Niels Sandal, et al. 2003. "The *Sym35* Gene Required for Root Nodule Development in Pea Is an Ortholog of *Nin* from *Lotus Japonicus*." *Plant Physiology* 131 (3): 1009–17.

Broghammer, A., L. Krusell, M. Blaise, J. Sauer, J. T. Sullivan, N. Maolanon, M. Vinther, et al. 2012. "Legume Receptors Perceive the Rhizobial Lipochitin Oligosaccharide Signal Molecules by Direct Binding." *Proceedings of the National Academy of Sciences of the United States of America* 109 (34): 13859–64.

Bruneau, Anne, Marjorie Mercure, Gwilym P. Lewis, and Patrick S. Herendeen. 2008. "Phylogenetic Patterns and Diversification in the Caesalpinioid Legumes." *Botany* 86 (7): 697–718.

Burrill, Thomas Jonathan, and Roy Hansen. 1917. "Is Symbiosis Possible between Legume Bacteria and Non-Legume Plants?" *Bulletin - University of Illinois Agricultural Experiment Station* 202: 115–81.

Campbell, Michael S., Meiyee Law, Carson Holt, Joshua C. Stein, Gaurav D. Moghe, David E. Hufnagel, Jikai Lei, et al. 2014. "MAKER-P: A Tool Kit for the Rapid Creation, Management, and Quality Control of Plant Genome Annotations." *Plant Physiology* 164 (2): 513–24.

Cao, Qingqin, Rik Op den Camp, Maryam Seifi Kalhor, Ton Bisseling, and Rene Geurts. 2012. "Efficiency of *Agrobacterium Rhizogenes*–mediated Root Transformation of *Parasponia* and *Trema* Is Temperature Dependent." *Plant Growth Regulation* 68 (3): 459–65.

Chiasson, David M., Patrick C. Loughlin, Danielle Mazurkiewicz, Manijeh Mohammadidehcheshmeh, Elena E. Fedorova, Mamoru Okamoto, Elizabeth McLean, et al. 2014. "Soybean SAT1 (Symbiotic Ammonium Transporter 1) Encodes a bHLH Transcription Factor Involved in Nodule Growth and NH4+ Transport." *Proceedings of the National Academy of Sciences of the United States of America* 111 (13): 4814–19.

Clason, E. W. 1936. "The Vegetation of the Upper-Badak Region of Mount Kelut (East Java)." *Bulletin Du Jardin Botanique de Buitenzorg, Sér. 3* 13: 509–18.

Clavijo, Fernando, Issa Diedhiou, Virginie Vaissayre, Laurent Brottier, Jennifer Acolatse, Daniel Moukouanga, Amandine Crabos, et al. 2015. "The *Casuarina NIN* Gene Is Transcriptionally Activated throughout *Frankia* Root Infection as Well as in Response to Bacterial Diffusible Signals." *The New Phytologist* 208: 887–903.

Combier, Jean-Philippe Philippe, Florian Frugier, Françoise De Billy, Adnane Boualem, Fikri El-yahyaoui, Sandra Moreau, Tatiana Vernié, et al. 2006. "MtHAP2-1 Is a Key Transcriptional Regulator of Symbiotic Nodule Development Regulated by microRNA169 in *Medicago Truncatula*." *Genes & Development* 20 (1990): 3084–88.

Davey, M. R., G. Webster, G. Manders, F. L. Ringrose, J. B. Power, and E. C. Cocking. 1993. "Effective Nodulation of Micro-Propagated Shoots of the Non-Legume *Parasponia Andersonii* by *Bradyrhizobium*." *Journal of Experimental Botany* 44 (5): 863–67.

Deguchi, Yuichi, Mari Banba, Yoshikazu Shimoda, Svetlana A. Chechetka, Ryota Suzuri, Yasuhiro Okusako, Yasuhiro Ooki, et al. 2007. "Transcriptome Profiling of *Lotus Japonicus* Roots during Arbuscular Mycorrhiza Development and Comparison with that of Nodulation." *DNA Research: An International Journal for Rapid Publication of Reports*

on Genes and Genomes 14 (3): 117–33.

Delaux, Pierre-Marc, Guru V. Radhakrishnan, Dhileepkumar Jayaraman, Jitender Cheema, Mathilde Malbreil, Jeremy D. Volkening, Hiroyuki Sekimoto, et al. 2015. "Algal Ancestor of Land Plants Was Preadapted for Symbiosis." *Proceedings of the National Academy of Sciences of the United States of America* 112 (43): 13390–95.

Doyle, Jeff J. 1998. "Phylogenetic Perspectives on Nodulation: Evolving Views of Plants and Symbiotic Bacteria." *Trends in Plant Science* 3 (12): 473–78.

———. 2011. "Phylogenetic Perspectives on the Origins of Nodulation." *Molecular Plant-Microbe Interactions: MPMI* 24 (November): 1289–95.

———. 2016. "Chasing Unicorns: Nodulation Origins and the Paradox of Novelty." *American Journal of Botany* 103 (11): 1865–68.

Emms, David M., and Steven Kelly. 2015. "OrthoFinder: Solving Fundamental Biases in Whole Genome Comparisons Dramatically Improves Orthogroup Inference Accuracy." *Genome Biology* 16 (August): 157.

Gallego-Giraldo, Lina, Kishor Bhattarai, Catalina I. Pislariu, Jin Nakashima, Yusuke Jikumaru, Yuji Kamiya, Michael K. Udvardi, Maria J. Monteros, and Richard A. Dixon. 2014. "Lignin Modification Leads to Increased Nodule Numbers in Alfalfa." *Plant Physiology* 164 (3): 1139–50.

Geurts, René, and Hans de Jong. 2013. "Fluorescent In Situ Hybridization (FISH) on Pachytene Chromosomes as a Tool for Genome Characterization." *Methods in Molecular Biology* 1069: 15–24.

Geurts, Rene, Ting Ting Xiao, and Barbara Reinhold-Hurek. 2016. "What Does It Take to Evolve a Nitrogen-Fixing Endosymbiosis?" *Trends in Plant Science* 21 (3): 199–208.

Gnerre, Sante, Iain Maccallum, Dariusz Przybylski, Filipe J. Ribeiro, Joshua N. Burton, Bruce J. Walker, Ted Sharpe, et al. 2011. "High-Quality Draft Assemblies of Mammalian Genomes from Massively Parallel Sequence Data." *Proceedings of the National Academy of Sciences of the United States of America* 108 (4): 1513–18.

Grabherr, Manfred G., Brian J. Haas, Moran Yassour, Joshua Z. Levin, Dawn A. Thompson, Ido Amit, Xian Adiconis, et al. 2011. "Full-Length Transcriptome Assembly from RNA-Seq Data without a Reference Genome." *Nature Biotechnology* 29 (7): 644–52.

Granqvist, Emma, Jongho Sun, Rik Op den Camp, Petar Pujic, Lionel Hill, Philippe Normand, Richard J. Morris, J. Allan Downie, Rene Geurts, and Giles E. D. Oldroyd. 2015. "Bacterial-Induced Calcium Oscillations Are Common to Nitrogen-Fixing Associations of Nodulating Legumes and Non-Legumes." *The New Phytologist* 207 (3): 551–58.

Gremme, Gordon, Sascha Steinbiss, and Stefan Kurtz. 2013. "GenomeTools: A Comprehensive Software Library for Efficient Processing of Structured Genome Annotations." *IEEE/ACM Transactions on Computational Biology and Bioinformatics / IEEE, ACM* 10 (3): 645–56.

Haas, Brian J., Alexie Papanicolaou, Moran Yassour, Manfred Grabherr, Philip D. Blood, Joshua Bowden, Matthew Brian Couger, et al. 2013. "De Novo Transcript Sequence Reconstruction from RNA-Seq Using the Trinity Platform for Reference Generation and Analysis." *Nature Protocols* 8 (8): 1494–1512.

Haas, Brian J., Steven L. Salzberg, Wei Zhu, Mihaela Pertea, Jonathan E. Allen, Joshua Orvis, Owen White, C. Robin Buell, and Jennifer R. Wortman. 2008. "Automated Eukaryotic Gene Structure Annotation Using EVidenceModeler and the Program to Assemble Spliced Alignments." *Genome Biology* 9 (1): R7.

Han, Yujun, and Susan R. Wessler. 2010. "MITE-Hunter: A Program for Discovering Miniature Inverted-Repeat Transposable Elements from Genomic Sequences." *Nucleic Acids Research* 38 (22): e199.

He, Ningjia, Chi Zhang, Xiwu Qi, Shancen Zhao, Yong Tao, Guojun Yang, Tae-Ho Lee, et al. 2013. "Draft Genome Sequence of the Mulberry Tree *Morus Notabilis*." *Nature Communications* 4 (September): ncomms3445.

Herendeen, Patrick S., Susana Magallon-Puebla, Richard Lupia, Peter R. Crane, and Jolanta Kobylinska. 1999. "A Preliminary Conspectus of the Allon Flora from the Late Cretaceous (late Santonian) of Central Georgia, USA." *Annals of the Missouri Botanical Garden. Missouri Botanical Garden*, 407–71.

Hoff, Katharina J., Simone Lange, Alexandre Lomsadze, Mark Borodovsky, and Mario Stanke. 2016. "BRAKER1: Unsupervised RNA-Seq-Based Genome Annotation with GeneMark-ET and AUGUSTUS." *Bioinformatics* 32 (5): 767–69.

Horváth, Beatrix, Li Huey Yeun, Ágota Domonkos, Gábor Halász, Enrico Gobbato, Ferhan Ayaydin, Krisztina Miró, et al. 2011. "Medicago Truncatula IPD3 Is a Member of the Common Symbiotic Signaling Pathway Required for Rhizobial and Mycorrhizal Symbioses." *Molecular Plant-Microbe Interactions: MPMI* 24 (11): 1345–58.

Huang, Jian, Chunmei Zhang, Xing Zhao, Zhangjun Fei, Kangkang Wan, Zhong Zhang, Xiaoming Pang, et al. 2016. "The Jujube Genome Provides Insights into Genome Evolution and the Domestication of Sweetness/acidity Taste in Fruit Trees." *PLoS Genetics* 12 (12): e1006433.

Kakar, Smita, Ryan Sturms, Andrea Tiffany, Jay C. Nix, Alan A. Dispirito, and Mark S. Hargrove. 2011. "Crystal Structures of *Parasponia* and *Trema* Hemoglobins: Differential Heme Coordination Is Linked to Quaternary

Structure." *Biochemistry* 50 (20): 4273–80.

Kamel, Laurent, Michelle Keller-Pearson, Christophe Roux, and Jean-Michel Ané. 2017. "Biology and Evolution of Arbuscular Mycorrhizal Symbiosis in the Light of Genomics." *The New Phytologist* 213 (2): 531–36.

Kawaharada, Yasuyuki, Mette W. Nielsen, Simon Kelly, Euan K. James, Kasper R. Andersen, Sheena R. Rasmussen, Winnie Füchtbauer, et al. 2017. "Differential Regulation of the Epr3 Receptor Coordinates Membrane-Restricted Rhizobial Colonization of Root Nodule Primordia." *Nature Communications* 8 (February): 14534.

Kawaharada, Y., S. Kelly, M. Wibroe Nielsen, C. T. Hjuler, K. Gysel, A. Muszyński, R. W. Carlson, et al. 2015. "Receptor-Mediated Exopolysaccharide Perception Controls Bacterial Infection." *Nature* 523 (7560): 308–12.

Kim, Daehwan, Ben Langmead, and Steven L. Salzberg. 2015. "HISAT: A Fast Spliced Aligner with Low Memory Requirements." *Nature Methods* 12 (4): 357–60.

Kistner, Catherine, Thilo Winzer, Andrea Pitzschke, Lonneke Mulder, Shusei Sato, Takakazu Kaneko, Satoshi Tabata, et al. 2005. "Seven *Lotus Japonicus* Genes Required for Transcriptional Reprogramming of the Root during Fungal and Bacterial Symbiosis." *The Plant Cell* 17 (8): 2217–29.

Korf, Ian. 2004. "Gene Finding in Novel Genomes." *BMC Bioinformatics* 5 (May): 59.

Kurtz, Stefan, Adam Phillippy, Arthur L. Delcher, Michael Smoot, Martin Shumway, Corina Antonescu, and Steven L. Salzberg. 2004. "Versatile and Open Software for Comparing Large Genomes." *Genome Biology* 5 (2): R12.

Liao, Yang, Gordon K. Smyth, and Wei Shi. 2014. "featureCounts: An Efficient General Purpose Program for Assigning Sequence Reads to Genomic Features." *Bioinformatics* 30 (7): 923–30.

Li, Hong-Lei, Wei Wang, Peter E. Mortimer, Rui-Qi Li, De-Zhu Li, Kevin D. Hyde, Jian-Chu Xu, Douglas E. Soltis, and Zhi-Duan Chen. 2015. "Large-Scale Phylogenetic Analyses Reveal Multiple Gains of Actinorhizal Nitrogen-Fixing Symbioses in Angiosperms Associated with Climate Change." *Scientific Reports* 5 (September): 14023.

Limpens, Erik, Carolien Franken, Patrick Smit, Joost Willemse, Ton Bisseling, and René Geurts. 2003. "LysM Domain Receptor Kinases Regulating Rhizobial Nod Factor-Induced Infection." *Science* 302 (5645): 630–33.

Lomsadze, Alexandre, Paul D. Burns, and Mark Borodovsky. 2014. "Integration of Mapped RNA-Seq Reads into Automatic Training of Eukaryotic Gene Finding Algorithm." *Nucleic Acids Research* 42 (15): e119.

Madsen, Esben Bjørn, Lene Heegaard Madsen, Simona Radutoiu, Shusei Sato, Takakazu Kaneko, Satoshi Tabata, and Niels Sandal. 2003. "A Receptor Kinase Gene of the LysM Type Is Involved in Legume Perception of Rhizobial Signals." *Nature* 425: 637–40.

Marsh, John F., Alexandra Rakocevic, Raka M. Mitra, Lysiane Brocard, Jongho Sun, Alexis Eschstruth, Sharon R. Long, Michael Schultze, Pascal Ratet, and G. E. D. D. Oldroyd. 2007. "*Medicago Truncatula* NIN Is Essential for Rhizobial-Independent Nodule Organogenesis Induced by Autoactive Calcium/calmodulin-Dependent Protein Kinase." *Plant Physiology* 144 (1): 324–35.

Martin, Francis M., Stéphane Uroz, and David G. Barker. 2017. "Ancestral Alliances: Plant Mutualistic Symbioses with Fungi and Bacteria." *Science* 356 (6340). https://doi.org/10.1126/science.aad4501.

Marvel, D. J., J. G. Torrey, and F. M. Ausubel. 1987. "*Rhizobium* Symbiotic Genes Required for Nodulation of Legume and Nonlegume Hosts." *Proceedings of the National Academy of Sciences of the United States of America* 84 (March): 1319–23.

Murray, Jeremy D., Rajasekhara Reddy Duvvuru Muni, Ivone Torres-Jerez, Yuhong Tang, Stacy Allen, Megan Andriankaja, Guangming Li, et al. 2011. "Vapyrin, a Gene Essential for Intracellular Progression of Arbuscular Mycorrhizal Symbiosis, Is Also Essential for Infection by Rhizobia in the Nodule Symbiosis of Medicago Truncatula." *The Plant Journal: For Cell and Molecular Biology* 65 (2): 244–52.

Myburg, Alexander A., Dario Grattapaglia, Gerald A. Tuskan, Uffe Hellsten, Richard D. Hayes, Jane Grimwood, Jerry Jenkins, et al. 2014. "The Genome of *Eucalyptus Grandis*." *Nature* 510 (7505): 356–62.

Natsume, Satoshi, Hiroki Takagi, Akira Shiraishi, Jun Murata, Hiromi Toyonaga, Josef Patzak, Motoshige Takagi, et al. 2015. "The Draft Genome of Hop (*Humulus Lupulus*), an Essence for Brewing." *Plant & Cell Physiology* 56 (3): 428–41.

Nguyen, Thanh Van, Daniel Wibberg, Kai Battenberg, Jochen Blom, Brian Vanden Heuvel, Alison M. Berry, Jörn Kalinowski, and Katharina Pawlowski. 2016. "An Assemblage of *Frankia* Cluster II Strains from California Contains the Canonical Nod Genes and Also the Sulfotransferase Gene nodH." *BMC Genomics* 17 (1): 796.

Oldroyd, Giles E. D. 2013. "Speak, Friend, and Enter: Signalling Systems That Promote Beneficial Symbiotic Associations in Plants." *Nature Reviews. Microbiology* 11 (4): 252–63.

Op den Camp, Rik H. M., Elisa Polone, Elena Fedorova, Wim Roelofsen, Andrea Squartini, Huub J. M. Op den Camp, Ton Bisseling, and René Geurts. 2012. "Nonlegume *Parasponia Andersonii* Deploys a Broad Rhizobium Host Range Strategy Resulting in Largely Variable Symbiotic Effectiveness." *Molecular Plant-Microbe Interactions: MPMI* 25 (7): 954–63.

Op den Camp, Rik, Arend Streng, Stéphane De Mita, Qingqin Cao, Elisa Polone, Wei Liu, Jetty S. S. Ammiraju, et al. 2011. “LysM-Type Mycorrhizal Receptor Recruited for Rhizobium Symbiosis in Nonlegume *Parasponia*.” *Science* 331 (December): 909–12.

Otto, Thomas D., Gary P. Dillon, Wim S. Degrave, and Matthew Berriman. 2011. “RATT: Rapid Annotation Transfer Tool.” *Nucleic Acids Research* 39 (9): e57.

Ott, Thomas, Joost T. van Dongen, Catrin Günther, Lene Krusell, Guilhem Desbrosses, Helene Vigeolas, Vivien Bock, Tomasz Czechowski, Peter Geigenberger, and Michael K. Udvardi. 2005. “Symbiotic Leghemoglobins Are Crucial for Nitrogen Fixation in Legume Root Nodules but Not for General Plant Growth and Development.” *Current Biology: CB* 15 (6): 531–35.

Parniske, Martin. 2008. “Arbuscular Mycorrhiza: The Mother of Plant Root Endosymbioses.” *Nature Reviews. Microbiology* 6: 763–75.

Parra, Genis, Keith Bradnam, and Ian Korf. 2007. “CEGMA: A Pipeline to Accurately Annotate Core Genes in Eukaryotic Genomes.” *Bioinformatics* 23 (9): 1061–67.

Pumplin, Nathan, Stephen J. Mondo, Stephanie Topp, Colby G. Starker, J. Stephen Gantt, and Maria J. Harrison. 2010. “Medicago Truncatula Vapyrin Is a Novel Protein Required for Arbuscular Mycorrhizal Symbiosis.” *The Plant Journal: For Cell and Molecular Biology* 61 (3): 482–94.

Radutoiu, S., L. H. Madsen, E. B. Madsen, H. H. Felle, Y. Umehara, M. Gronlund, S. Sato, et al. 2003. “Plant Recognition of Symbiotic Bacteria Requires Two LysM Receptor-like Kinases.” *Nature* 425 (6958): 585–92.

Rogers, Christian, and Giles E. D. Oldroyd. 2014. “Synthetic Biology Approaches to Engineering the Nitrogen Symbiosis in Cereals.” *Journal of Experimental Botany* 65 (8): 1939–46.

Roux, Brice, Nathalie Rodde, Marie-Françoise Jardinaud, Ton Timmers, Laurent Sauviac, Ludovic Cottret, Sébastien Carrère, et al. 2014a. “An Integrated Analysis of Plant and Bacterial Gene Expression in Symbiotic Root Nodules Using Laser-Capture Microdissection Coupled to RNA Sequencing.” *The Plant Journal: For Cell and Molecular Biology* 77 (6): 817–37.

———. 2014b. “An Integrated Analysis of Plant and Bacterial Gene Expression in Symbiotic Root Nodules Using Laser-Capture Microdissection Coupled to RNA Sequencing.” *The Plant Journal: For Cell and Molecular Biology* 77 (6): 817–37.

Saitou, N., and M. Nei. 1987. “The Neighbor-Joining Method: A New Method for Reconstructing Phylogenetic Trees.” *Molecular Biology and Evolution* 4 (4): 406–25.

Schauser, L., A. Roussis, J. Stiller, and J. Stougaard. 1999. “A Plant Regulator Controlling Development of Symbiotic Root Nodules.” *Nature* 402 (6758): 191–95.

Schmutz, Jeremy, Steven B. Cannon, Jessica Schlueter, Jianxin Ma, Therese Mitros, William Nelson, David L. Hyten, et al. 2010. “Genome Sequence of the Palaeopolyploid Soybean.” *Nature* 463 (7278): 178–83.

Shadle, Gail, Fang Chen, M. S. Srinivasa Reddy, Lisa Jackson, Jin Nakashima, and Richard A. Dixon. 2007. “Down-Regulation of Hydroxycinnamoyl CoA: Shikimate Hydroxycinnamoyl Transferase in Transgenic Alfalfa Affects Lignification, Development and Forage Quality.” *Phytochemistry* 68 (11): 1521–29.

Shulaev, Vladimir, Daniel J. Sargent, Ross N. Crowhurst, Todd C. Mockler, Otto Folkerts, Arthur L. Delcher, Pankaj Jaiswal, et al. 2011. “The Genome of Woodland Strawberry (*Fragaria Vesca*).” *Nature Genetics* 43 (2): 109–16.

Silvester, W. B., R. H. Berg, C. R. Schwintzer, and J. D. Tjepkema. 2007. “Oxygen Responses, Hemoglobin, and the Structure and Function of Vesicles.” In *Nitrogen-Fixing Actinorhizal Symbioses*, edited by Katharina Pawlowski and William E. Newton, 105–46. Nitrogen Fixation: Origins, Applications, and Research Progress 6. Springer Netherlands.

Silvester, W. B., S. L. Harris, and J. D. Tjepkema. 1990. “Oxygen Regulation and Hemoglobin.” In *The Biology of Frankia and Actinorhizal Plants*, edited by Christa R. Schwintzer and John D. Tjepkema, 157–76. New York: Academic Press.

Silvester, W. B., and L. J. Winship. 1990. “Transient Responses of Nitrogenase to Acetylene and Oxygen in Actinorhizal Nodules and Cultured *Frankia*.” *Plant Physiology* 92 (2): 480–86.

Simão, Felipe A., Robert M. Waterhouse, Panagiotis Ioannidis, Evgenia V. Kriventseva, and Evgeny M. Zdobnov. 2015. “BUSCO: Assessing Genome Assembly and Annotation Completeness with Single-Copy Orthologs.” *Bioinformatics* 31 (19): 3210–12.

Soltis, D. E., P. S. Soltis, D. R. Morgan, S. M. Swensen, B. C. Mullin, J. M. Dowd, and P. G. Martin. 1995. “Chloroplast Gene Sequence Data Suggest a Single Origin of the Predisposition for Symbiotic Nitrogen Fixation in Angiosperms.” *Proceedings of the National Academy of Sciences of the United States of America* 92 (7): 2647–51.

Soyano, Takashi, Hiroshi Kouchi, Atsuko Hirota, and Makoto Hayashi. 2013. “NODULE INCEPTION Directly Targets *NF-Y* Subunit Genes to Regulate Essential Processes of Root Nodule Development in *Lotus Japonicus*.” *PLoS*

Genetics 9 (3): e1003352.

Stanke, Mario, Mark Diekhans, Robert Baertsch, and David Haussler. 2008. “Using Native and Syntenically Mapped cDNA Alignments to Improve de Novo Gene Finding.” *Bioinformatics* 24 (5): 637–44.

Stokstad, E. 2016. “The Nitrogen Fix.” *Science* 353 (6305): 1225–27.

Sturms, Ryan, Smita Kakar, James Trent, and Mark S. Hargrove. 2010. “*Trema* and *Parasponia* Hemoglobins Reveal Convergent Evolution of Oxygen Transport in Plants.” *Biochemistry* 49 (19): 4085–93.

Swarbreck, David, Christopher Wilks, Philippe Lamesch, Tanya Z. Berardini, Margarita Garcia-Hernandez, Hartmut Foerster, Donghui Li, et al. 2008. “The *Arabidopsis* Information Resource (TAIR): Gene Structure and Function Annotation.” *Nucleic Acids Research* 36 (Database issue): D1009–14.

Swensen, Susan M. 1996. “The Evolution of Actinorhizal Symbioses: Evidence for Multiple Origins of the Symbiotic Association.” *American Journal of Botany* 83 (11): 1503–12.

Tóth, Katalin, Thomas F. Stratil, Esben B. Madsen, Juanying Ye, Claudia Popp, Meritxell Antolín-Llovera, Christina Grossmann, et al. 2012. “Functional Domain Analysis of the Remorin Protein LjSYMREM1 in Lotus Japonicus.” *PloS One* 7 (1). https://doi.org/10.1371/journal.pone.0030817.

Trapnell, Cole, Brian A. Williams, Geo Pertea, Ali Mortazavi, Gordon Kwan, Marijke J. van Baren, Steven L. Salzberg, Barbara J. Wold, and Lior Pachter. 2010. “Transcript Assembly and Quantification by RNA-Seq Reveals Unannotated Transcripts and Isoform Switching during Cell Differentiation.” *Nature Biotechnology* 28 (5): 511–15.

Trinick, M. J. 1973. “Symbiosis between *Rhizobium* and the Non-Legume, *Trema Aspera*.” *Nature* 244: 459–60.

Tuskan, G. A., S. DiFazio, S. Jansson, J. Bohlmann, I. Grigoriev, U. Hellsten, N. Putnam, et al. 2006. “The Genome of Black Cottonwood, *Populus Trichocarpa* (Torr. & Gray).” *Science* 313 (5793): 1596–1604.

Udvardi, Michael, and Philip S. Poole. 2013. “Transport and Metabolism in Legume-Rhizobia Symbioses.” *Annual Review of Plant Biology* 64 (March): 781–805.

UniProt Consortium. 2015. “UniProt: A Hub for Protein Information.” *Nucleic Acids Research* 43 (Database issue): D204–12.

Untergasser, Andreas, Gerben J. M. Bijl, Wei Liu, Ton Bisseling, Jan G. Schaart, and René Geurts. 2012. “One-Step *Agrobacterium* Mediated Transformation of Eight Genes Essential for Rhizobium Symbiotic Signaling Using the Novel Binary Vector System pHUGE.” *PloS One* 7 (10): e47885.

Valdar, William S. J. 2002. “Scoring Residue Conservation.” *Proteins* 48 (2): 227–41.

Velasco, Riccardo, Andrey Zharkikh, Jason Affourtit, Amit Dhingra, Alessandro Cestaro, Ananth Kalyanaraman, Paolo Fontana, et al. 2010. “The Genome of the Domesticated Apple (*Malus* × *Domestica* Borkh.).” *Nature Genetics* 42 (10): 833–39.

Verde, Ignazio, Albert G. Abbott, Simone Scalabrin, Sook Jung, Shengqiang Shu, Fabio Marroni, Tatyana Zhebentyayeva, et al. 2013. “The High-Quality Draft Genome of Peach (*Prunus Persica*) Identifies Unique Patterns of Genetic Diversity, Domestication and Genome Evolution.” *Nature Genetics* 45 (5): 487–94.

Vernié, Tatiana, Jiyoung Kim, Lisa Frances, Yiliang Ding, Jongho Sun, Dian Guan, Andreas Niebel, Miriam L. Gifford, Fernanda de Carvalho-Niebel, and Giles E. D. Oldroyd. 2015. “The NIN Transcription Factor Coordinates Diverse Nodulation Programs in Different Tissues of the *Medicago Truncatula* Root.” *The Plant Cell* 27 (12): 3410–24.

Werner, Gijsbert D. A., William K. Cornwell, Janet I. Sprent, Jens Kattge, and E. Toby Kiers. 2014. “A Single Evolutionary Innovation Drives the Deep Evolution of Symbiotic N_2-Fixation in Angiosperms.” *Nature Communications* 5 (June): 4087.

Winship, Lawrence J., Kendall J. Martin, and Anita Sellstedt. 1987. “The Acetylene Reduction Assay Inactivates Root Nodule Uptake Hydrogenase in Some Actinorhizal Plants.” *Physiologia Plantarum* 70 (2): 361–66.

Yang, Mei-Qing, Robin Van Velzen, Freek T. Bakker, Ali Sattarian, De-Zhu Li, and Ting-Shuang Yi. 2013. “Molecular Phylogenetics and Character Evolution of Cannabaceae.” *Taxon* 62 (3): 473–85.

Yano, Koji, Satoko Yoshida, Judith Müller, Sylvia Singh, Mari Banba, Kate Vickers, Katharina Markmann, et al. 2008. “CYCLOPS, a Mediator of Symbiotic Intracellular Accommodation.” *Proceedings of the National Academy of Sciences of the United States of America* 105 (51): 20540–45.

Young, Nevin D., Frédéric Debellé, Giles E. D. Oldroyd, Rene Geurts, Steven B. Cannon, Michael K. Udvardi, Vagner a. Benedito, et al. 2011. “The *Medicago* Genome Provides Insight into the Evolution of Rhizobial Symbioses.” *Nature* 480 (7378): 520–24.

Żmieńko, Agnieszka, Anna Samelak, Piotr Kozłowski, and Marek Figlerowicz. 2014. “Copy Number Polymorphism in Plant Genomes.” *TAG. Theoretical and Applied Genetics. Theoretische Und Angewandte Genetik* 127 (1): 1–18.

CHAPTER 4

4

Duplication of symbiotic Lysin Motif-receptors predates the evolution of nitrogen-fixing nodule symbiosis

Luuk Rutten[#1], Kana Miyata[#1], Yuda Purwana Roswanjaya[1,2], Rik Huisman[1], Fengjiao Bu[1], Marijke Hartog[1], Sidney Linders[1], Robin van Velzen[1,3], Arjan van Zeijl[1], Ton Bisseling[1], Wouter Kohlen[1] and Rene Geurts[1,*]

[1]Laboratory of Molecular Biology, Department of Plant Sciences, Wageningen University, Droevendaalsesteeg 1, 6708PB Wageningen, The Netherlands. [2]Centre of Technology for Agricultural Production, Agency for the Assessment and Application of Technology (BPPT), Jakarta, Indonesia. [3]Biosystematics Group, Department of Plant Sciences, Wageningen University, 6708 PB, Wageningen, The Netherlands.

[#]These authors contributed equally to this work

This chapter is published in Plant Physiology, July 21, 2020, DOI:10.1104/pp.19.01420

Abstract

Rhizobium nitrogen-fixing nodule symbiosis occurs in two taxonomic lineages: legumes (Fabaceae) and *Parasponia* (Cannabaceae). Both symbioses are initiated upon the perception of rhizobium-secreted lipo-chitooligosaccharides (LCOs), called Nod factors. Studies in the model legumes *Lotus japonicus* (lotus) and *Medicago truncatula* (medicago) showed that rhizobium LCOs are perceived by a heteromeric receptor complex of distinct Lysin Motif (LysM)- type transmembrane receptors named NOD FACTOR RECEPTOR1 (LjNFR1)-LjNFR5 (lotus) and LYSM DOMAIN CONTAINING RECEPTOR KINASE3 (MtLYK3)-NOD FACTOR PERCEPTION (MtNFP) (medicago). Recent phylogenomic comparative analyses indicated that the nodulation trait of legumes, *Parasponia,* as well as of so-called actinorhizal plants that establish a symbiosis with diazotrophic *Frankia* bacteria, share an evolutionary origin about 110 million years ago. However, the evolutionary trajectory of LysM-type LCO receptors remains elusive. By conducting phylogenetic analysis, trans-complementation studies, and CRISPR-Cas9 mutagenesis in *Parasponia andersonii* (parasponia)*,* we obtained insight to the origin of LCO receptors essential for nodulation. We identified four LysM-type receptors controlling nodulation in parasponia*P. andersonii*: *PanLYK1, PanLYK3, PanNFP1* and *PanNFP2*. These genes evolved upon ancient duplication events predating and coinciding with the origin of nodulation. Phylogenetic and functional analysis associated the occurrence of a functional NFP2-orthologous receptor to LCO-driven nodulation. Legumes and *Parasponia* use orthologous LysM-type receptors to perceive rhizobium LCOs, suggesting a shared evolutionary origin of LCO-driven nodulation. Furthermore, we found that both PanLYK1 and PanLYK3 are essential for intracellular arbuscule formation of mutualistic endomycorrhizal fungi. PanLYK3 also acts as a chitin oligomer receptor essential for innate immune signalling, demonstrating functional analogy to CHITIN ELECITOR RECEPTOR KINASE (CERK)-type receptors.

Introduction

Nitrogen availability is a critical factor for plant growth, but fixed nitrogen in the form of nitrate or ammonia in soils is limited. Plants have acquired different strategies to overcome this limitation. One such strategy is establishing a nodule endosymbiosis with nitrogen-fixing *Frankia* or rhizobium bacteria. Inside nodules, physiological conditions are created that allow the bacteria to convert atmospheric dinitrogen (N_2) into ammonia that can be used by the plant. Carbohydrates of plant origin fuel this energy demanding process. The unique character of nitrogen-fixing nodule symbiosis has raised the interest of plant researchers for more than a century, ultimately aiming to transfer this trait to non-leguminous crop species (Burrill and Hansen, 1917; Rogers and Oldroyd, 2014; Huisman and Geurts, 2019).

The *Frankia* and rhizobium nitrogen-fixing nodulation trait occurs in ten paraphyletic lineages within the orders Fabales, Fagales, Cucurbitales and Rosales, collectively known as the nitrogen-fixing clade (Soltis et al., 1995). Based on phylogenomic comparisons of nodulating and non-nodulating plant species, it is hypothesized that the nitrogen-fixing nodule symbiosis with rhizobium or *Frankia* bacteria has a shared evolutionary origin, dating to about 110 million years ago (Griesmann et al., 2018; van Velzen et al., 2018a; van Velzen et al., 2018b). Subsequently, the nodulation trait most probably was lost multiple times, which is associated with pseudogenization of two key genes essential for nodule organogenesis and bacterial infection; the transcription factor *NODULE INCEPTION* (*NIN*) and the coiled-coil protein-encoding gene *RHIZOBIUM POLAR GROWTH* (*RPG*) (Griesmann et al., 2018; van Velzen et al., 2018b). These two genes likely experienced genetic adaptations, allowing them to function exclusively in nodulation. However, insight into the evolutionary trajectory of signalling receptors involved in recognition of bacterial signals and subsequent activation of the pathways leading to nodule organogenesis and bacterial infection remains elusive.

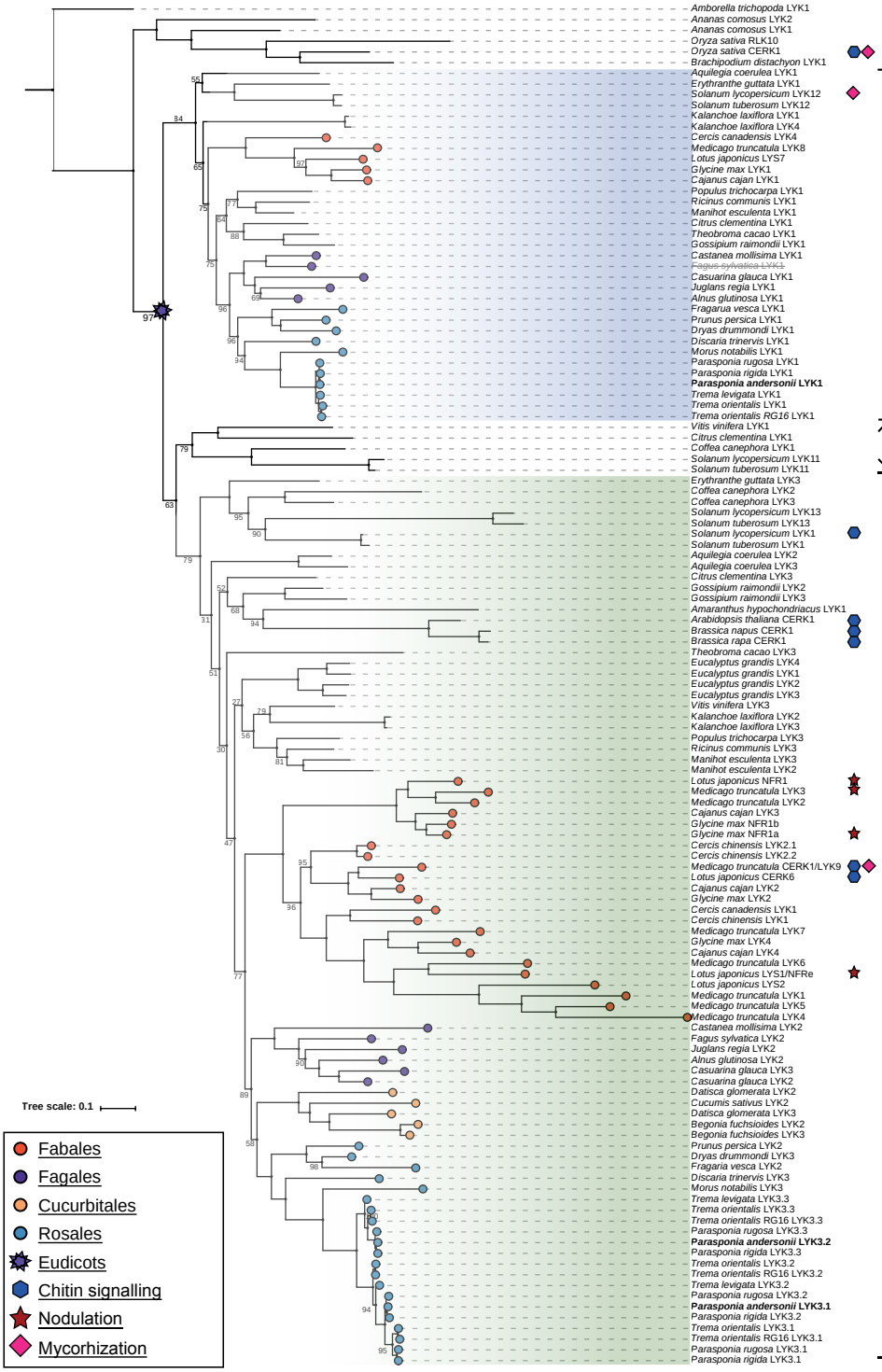
The nitrogen-fixing nodulation trait is best studied in the legume models *Lotus japonicus* (lotus) and *Medicago truncatula* (medicago) (Fabaceae, Fabales). Both these legumes recognize their rhizobium microsymbionts by the structural characteristics of secreted lipo-chitooligosaccharides (LCOs, also known as Nod factors). Perception of these molecules triggers nodule development (Wang et al., 2012). LCO signalling is also the basis of rhizobium-induced nodulation in the non-legume *Parasponia* (Cannabaceae, Rosales) (Marvel et al., 1987; Op den Camp et al., 2011; van Velzen et al., 2018b). Additionally, it was found that diazotrophic *Frankia* strains of a basal taxonomic lineage (so-called cluster-II strains) possess LCO biosynthesis genes, but the nodulating strains of two other taxonomic clusters do not (Pawlowski and Demchenko, 2012; Persson et al., 2015; Nguyen et al., 2016; Van Nguyen et al., 2019). LCOs, as well as chitin oligomers (COs), are also used

Figure 1. Phylogeny reconstruction of the LYK-I orthogroup, containing known CO and LCO receptors, based on 127 sequences from 47 species. Two main subgroups are recognized in Eudicots, LYK-Ia (blue) and LYK-Ib (green). Note the presence of both variants in *Aquilegia coeralia*, a basal Eudicot in the Ranunculales. A subset of proteins is unresolved. Parasponia proteins are in bold. *Parasponia* and *Trema* LYK3.1 and LYK3.2 represent protein variants of LYK3. Deduced pseudo-proteins are depicted in grey/strikethrough. Proteins with known functions in nodulation, mycorrhization, and/or chitin-innate immune signalling are indicated. Bootstrap values indicate IQ-tree UF-bootstrap support%, values >98 are not shown. Tree sScale bar represents substitutions per site. A complete list of species and accession numbers can be found in **Data set S1**.

by arbuscular mycorrhiza (AM) fungi to signal their hosts (Maillet et al., 2011; Genre et al., 2013). Perception of these AM signals requires a plant LysM-type receptor that also is essential for chitin innate immune signalling; e.g. CHITIN ELECITOR RECEPTOR KINASE1 (OsCERK1) in rice (*Oryza sativa*) (Miyata et al., 2014; Zhang et al., 2015; He et al., 2019). This suggests that nodulating bacteria co-opted LCO signalling from the widespread AM symbiosis and/or innate immune signalling (Parniske, 2008; Gough and Cullimore, 2011; Geurts et al., 2012).

Genetic and biochemical studies in lotus and medicago demonstrated that rhizobium LCOs are perceived specifically by a heteromeric complex containing two distinct LysM-type receptors, named NOD FACTOR RECEPTOR1 (LjNFR1) and LjNFR5 in lotus, and LYSM DOMAIN CONTAINING RECEPTOR KINASE3 (MtLYK3) and NOD FACTOR PERCEPTION (MtNFP) in medicago (Limpens et al., 2003; Madsen et al., 2003; Radutoiu et al., 2003; Arrighi et al., 2006; Radutoiu et al., 2007; Broghammer et al., 2012). Other receptors may modulate the LCO response, such as LjNFRe, a homolog of LjNFR1 in lotus (Murakami et al., 2018). The LysM-type receptor family can be divided into two subclasses; named LYK and LYR, characterized by having a functional or dead kinase domain (Arrighi et al., 2006). Together these make up 11 orthogroups, two of which include legume LCO receptors (Buendia et al., 2018). Within legumes, the orthogroup that includes LjNFR1/MtLYK3 (named LYK-I clade) expanded upon gene duplications, allowing functional separation of rhizobium-induced signalling, AM symbiosis and chitin-triggered innate immune responses (De Mita et al., 2014; Bozsoki et al., 2017; Buendia et al., 2018; Gibelin-Viala et al., 2019). Likewise, the *LjNFR5*/*MtNFP* (orthogroup LYR-IA) experienced a gene duplication early in the legume clade (Young et al., 2011; Buendia et al., 2018).

Data on symbiotic LysM-type receptors in nodulating non-legumes are scarce. Only in *Parasponia andersonii* (parasponia) has a receptor functioning in nodulation been identified; named *PanNFP1*, which is a close homolog of *LjNFR5/MtNFP* (Op den Camp et al., 2011). Besides *PanNFP1*, *Parasponia* species possess a homologous receptor, named *NFP2*, which is more closely related to *LjNFR5/MtNFP* and transcriptionally activated in root nodules. Interestingly, this receptor is


Amborella trichopoda LYK1
Ananas comosus LYK2
Ananas comosus LYK1
Oryza sativa RLK10
Oryza sativa CERK1
Brachipodium distachyon LYK1
Aquilegia coerulea LYK1
Erythranthe guttata LYK1
Solanum lycopersicum LYK12
Solanum tuberosum LYK12
Kalanchoe laxiflora LYK1
Kalanchoe laxiflora LYK4
Cercis canadensis LYK4
Medicago truncatula LYK8
Lotus japonicus LYS7
Glycine max LYK1
Cajanus cajan LYK1
Populus trichocarpa LYK1
Ricinus communis LYK1
Manihot esculenta LYK1
Citrus clementina LYK1
Theobroma cacao LYK1
Gossipium raimondii LYK1
Castanea mollisima LYK1
Fagus sylvatica LYK1
Casuarina glauca LYK1
Juglans regia LYK1
Alnus glutinosa LYK1
Fragarua vesca LYK1
Prunus persica LYK1
Dryas drummondi LYK1
Discaria trinervis LYK1
Morus notabilis LYK1
Parasponia rugosa LYK1
Parasponia rigida LYK1
Parasponia andersonii LYK1
Trema levigata LYK1
Trema orientalis LYK1
Trema orientalis RG16 LYK1
Vitis vinifera LYK1
Citrus clementina LYK1
Coffea canephora LYK1
Solanum lycopersicum LYK11
Solanum tuberosum LYK11
Erythranthe guttata LYK3
Coffea canephora LYK2
Coffea canephora LYK3
Solanum lycopersicum LYK13
Solanum tuberosum LYK13
Solanum lycopersicum LYK1
Solanum tuberosum LYK1
Aquilegia coerulea LYK2
Aquilegia coerulea LYK3
Citrus clementina LYK3
Gossipium raimondii LYK2
Gossipium raimondii LYK3
Amaranthus hypochondriacus LYK1
Arabidopsis thaliana CERK1
Brassica napus CERK1
Brassica rapa CERK1
Theobroma cacao LYK3
Eucalyptus grandis LYK4
Eucalyptus grandis LYK1
Eucalyptus grandis LYK2
Eucalyptus grandis LYK3
Vitis vinifera LYK3
Kalanchoe laxiflora LYK2
Kalanchoe laxiflora LYK3
Populus trichocarpa LYK3
Ricinus communis LYK3
Manihot esculenta LYK3
Manihot esculenta LYK2
Lotus japonicus NFR1
Medicago truncatula LYK3
Medicago truncatula LYK2
Cajanus cajan LYK3
Glycine max NFR1b
Glycine max NFR1a
Cercis chinensis LYK2.1
Cercis chinensis LYK2.2
Medicago truncatula CERK1/LYK9
Lotus japonicus CERK6
Cajanus cajan LYK2
Glycine max LYK2
Cercis canadensis LYK1
Cercis chinensis LYK1
Medicago truncatula LYK7
Glycine max LYK4
Cajanus cajan LYK4
Medicago truncatula LYK6
Lotus japonicus LYS1/NFRe
Lotus japonicus LYS2
Medicago truncatula LYK1
Medicago truncatula LYK5
Medicago truncatula LYK4
Castanea mollisima LYK2
Fagus sylvatica LYK2
Juglans regia LYK2
Alnus glutinosa LYK2
Casuarina glauca LYK3
Casuarina glauca LYK2
Datisca glomerata LYK2
Cucumis sativus LYK2
Datisca glomerata LYK3
Begonia fuchsioides LYK2
Begonia fuchsioides LYK3
Prunus persica LYK2
Dryas drummondi LYK3
Fragaria vesca LYK2
Discaria trinervis LYK3
Morus notabilis LYK3
Trema levigata LYK3.3
Trema orientalis LYK3.3
Trema orientalis RG16 LYK3.3
Parasponia rugosa LYK3.3
Parasponia andersonii LYK3.2
Parasponia rigida LYK3.3
Trema orientalis LYK3.2
Trema orientalis RG16 LYK3.2
Trema levigata LYK3.2
Parasponia rugosa LYK3.2
Parasponia andersonii LYK3.1
Parasponia rigida LYK3.2
Trema orientalis LYK3.1
Trema orientalis RG16 LYK3.1
Parasponia rugosa LYK3.1
Parasponia rigida LYK3.1
LYK-Ia clade
unresolved
LYK-Ib clade
Tree scale: 0.1
Fabales
Fagales
Cucurbitales
Rosales
Eudicots
Chitin signalling
Nodulation
Mycorhization

pseudogenized in non-nodulating Rosales species (van Velzen et al., 2018b). To obtain insight into the evolution of LysM-type LCO receptors that are essential for nodulation, we used parasponia as a comparative system to legumes. The genus *Parasponia* represents five tropical tree species, which form nitrogen-fixing nodules with LCO producing rhizobium species that also nodulate legumes (van Velzen et al., 2018b). *Parasponia* and legumes diverged at the root of the nitrogen-fixing clade >100 million years ago (Li et al., 2015; van Velzen et al., 2018a). The microbial symbionts of the ancestral nodulating plants remain elusive, and it is probable that *Parasponia* and legumes accepted rhizobium as a microbial partner in parallel (van Velzen et al., 2018a). In line with this, *Parasponia* provides a unique comparative system to obtain insight into evolutionary trajectories of different LCO receptors that are essential for nodulation.

Results

Phylogeny reconstruction of orthogroups representing LysM-type LCO receptors

To obtain insight into the LysM-type receptor family of parasponia, we analysed it phylogenetically. We identified 16 parasponia genes encoding putative LysM-type receptors that all grouped in known orthogroups except one **(Figure S1; Table S1)**. Genetic studies in legumes uncovered only two orthogroups that contain proteins with a known function in rhizobium LCO signalling; these are named LYK-I and LYR-IA (Buendia et al., 2018). Parasponia has two gene copies in both these orthogroups.

LYK-I is the largest orthogroup, containing the functional legume LCO receptors *MtLYK3/LjNFR1* and *LjNFR*e (Limpens et al., 2003; Radutoiu et al., 2003; Murakami et al., 2018). Besides these, the LYK-I orthogroup also includes chitin innate immune receptors of medicago *MtLYK9/MtCERK1,* lotus *LjCERK6*, arabidopsis (*Arabidopsis thaliana*) *AtCERK1*, tomato (*Solanum lycopersicum*) *SlLYK1* and rice *OsCERK1* (Limpens et al., 2003; Miya et al., 2007; Wan et al., 2008; Shimizu et al., 2010; Miyata et al., 2014; Zhang et al., 2015; Bozsoki et al., 2017; Carotenuto et al., 2017; Liao et al., 2018; Gibelin-Viala et al., 2019; He et al., 2019). OsCERK1 and MtLYK9/MtCERK1 have also been found to function in AM symbiosis (Miyata et al., 2014; Zhang et al., 2015; Feng et al., 2019; Gibelin-Viala et al., 2019). Two parasponia genes are part of this orthogroup; named *PanLYK1* and *PanLYK3*.

A more exhaustive phylogenetic reconstruction was conducted using gene orthologs of additional species to obtain insight into the evolutionary relationships of these genes when compared to LCO and CO receptors. Notably, LysM-type receptors of the recently sequenced nodulating actinorhizal plants and non-nodulating relatives were included (Griesmann et al., 2018). The resulting phylogeny largely resembled Rosid species trees as reconstructed on the basis of plastid-coding genes (Wang

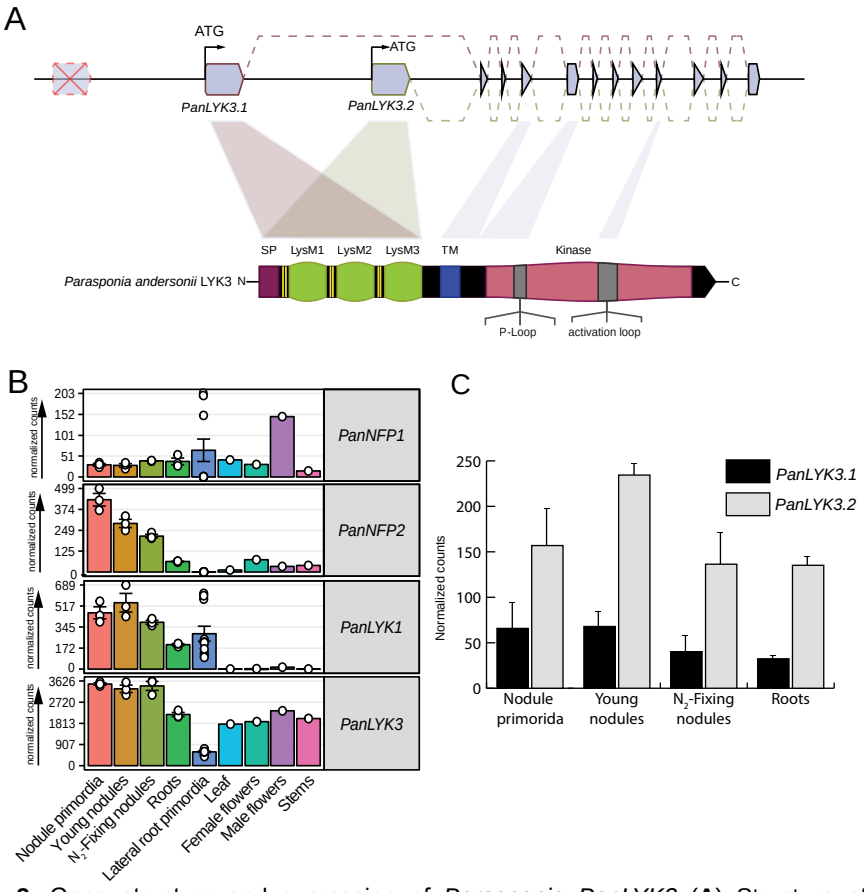


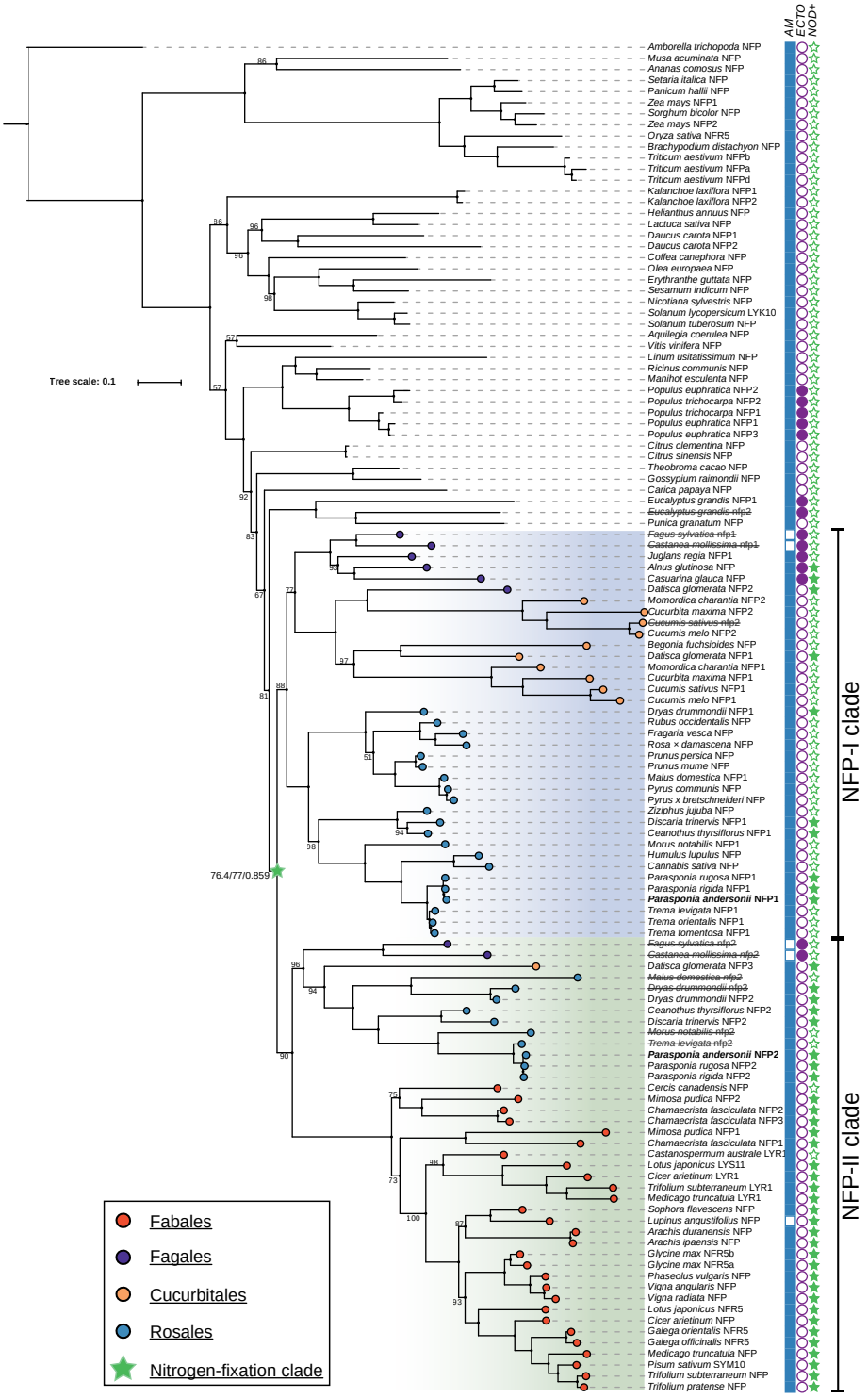
Figure 2. Gene structure and expression of *Parasponia PanLYK3* (**A**) Structure of the *PanLYK3* gene model and encoded proteins. *PanLYK3* possesses two transcriptional start sites resulting in two protein variants, which differ in the extracellular region containing the LysM-domains and are encoded by exon1. Red cross indicates a third upstream copy of exon1 lost in *Pparasponia andersonii,* maintained in other *Pparasponia* and/ *Ttrema* species. (**B**) Expression profile of *PanNFP1, PanNFP2*, *PanLYK1* and *PanLYK3* in different plant tissues. Expression is given in DESeq2-normalized read counts; error bars represent SE of biological replicates. Dots represent individual expression levels. The analysis is based on data presented in Van Velzen *et. al.* 2018 (van Velzen et al., 2018b). (**C**) Relative expression of the *PanLYK3.1* and *PanLYK3.2* transcriptional variants based on RNA-seq reads splicing into the second exon. Data are represented as mean ± SE (n=3). The analysis is based on data presented in Van Velzen *et. al.* 2018 (van Velzen et al., 2018b)

et al., 2009; Gonçalves et al., 2019). Our analysis revealed that *PanLYK1* and *PanLYK3* originated from an ancient duplication, dividing this orthogroup into two subgroups that we named LYK-Ia and LYK-Ib. This duplication does not coincide with the birth of the nitrogen-fixing clade, but rather has occurred in an ancestral eudicot **(Figure 1; Data set S1)**. The only studied member in the LYK-Ia orthogroup

Figure 3. Phylogeny reconstruction of LYRI-A orthogroup, containing known legume LCO receptors, based on 122 sequences from 87 species. A gene duplication in the root of the nitrogen-fixing clade is recognized; resulting in two subclades named NFP-I (blue) and NFP-II (green). The symbiotic capacities of the species are marked by filled (positive) and unfilled (negative) symbols: AM symbiosis (blue squares), ectomycorrhizal symbiosis (purple circles) and nodulation (green stars). Parasponia PanNFP1 and PanNFP2 are in bold. Deduced psuedo-proteins are depicted in grey/strikethrough. Values indicate IQ-tree UF-bootstrap support%. values >98 are not shown. Branch support for the nitrogen-fixing clade indicates aSH-aLRT / UF-Bootstrap / approximate Mr.Bayes support, respectively. Tree sScale bar represents substitutions per site. A list of species and accession numbers can be found in **Data set S2**.

is tomato *SlLYK12*, and knockdown of this gene by virus-induced gene silencing (VIGS) substantially reduces mycorrhizal colonization (Liao et al., 2018). The LYK-Ib clade represents several functionally characterized genes, including the chitin innate immune receptors and legume rhizobium LCO receptors. Legumes exhibit an increased number of genes in the LYK-Ib subclade, which are the result of tandem duplications (Limpens et al., 2003; Radutoiu et al., 2003; Zhu et al., 2006). These duplications may have driven neofunctionalization of LCO receptors in legumes (De Mita et al., 2014). In *Parasponia*, no gene duplications have occurred in the LYK-Ib clade (represented by *PanLYK3*) nor in the LYK-Ia clade (represented by *PanLYK1*). In contrast, parasponia *PanLYK3* experienced a duplication of exclusively the first exon. To determine whether this duplication is specific for the *Parasponia* genus we analysed the *LYK3* genomic region of two additional *Parasponia* and three non-nodulating species of the closely related genus *Trema*. This revealed that the duplication of *LYK3* exon1 is present in all species investigated and occurred twice, where the most distal exon 1 copy was lost in parasponia **(Figure 2A, Figure S2A)**. The encoded pre-mRNAs both splice into a shared second exon **(Figure 2)**. Each exon1 copy contains a putative transcription and translation start site, which allows for differential expression of the variants **(Figure 2B-C)**. Genes of the LYK-I clade have a highly conserved intron-exon structure (Zhang et al., 2009). In most cases, the first exon encodes the extracellular domain comprising the signal peptide and three LysM motifs. So, the parasponia *PanLYK3* gene encodes two protein variants, named PanLYK3.1 and PanLYK3.2, that differ in their extracellular domain **(Figure S2B).**

The LYR-IA orthogroup represents the legume LCO receptors MtNFP, LjNFR5 and pea (*Pisum sativum*) PsSYM10 (Madsen et al., 2003; Arrighi et al., 2006; Buendia et al., 2016; Miyata et al., 2016). Previously, we have shown that *Parasponia* species harbour two genes in this orthogroup, *PanNFP1* and *PanNFP2* in parasponia, of which the latter is more closely related to *MtNFP*/*LjNFR5* (van Velzen et al., 2018b). *PanNFP1* and *PanNFP2* originated from an ancient duplication. Phylogenetic


AM
ECTO
NOD+
Amborella trichopoda NFP
Musa acuminata NFP
Ananas comosus NFP
Setaria italica NFP
Panicum hallii NFP
Zea mays NFP1
Sorghum bicolor NFP
Zea mays NFP2
Oryza sativa NFR5
Brachypodium distachyon NFP
Triticum aestivum NFPb
Triticum aestivum NFPa
Triticum aestivum NFPd
Kalanchoe laxiflora NFP1
Kalanchoe laxiflora NFP2
Helianthus annuus NFP
Lactuca sativa NFP
Daucus carota NFP1
Daucus carota NFP2
Coffea canephora NFP
Olea europaea NFP
Erythranthe guttata NFP
Sesamum indicum NFP
Nicotiana sylvestris NFP
Solanum lycopersicum LYK10
Solanum tuberosum NFP
Aquilegia coerulea NFP
Vitis vinifera NFP
Linum usitatissimum NFP
Ricinus communis NFP
Manihot esculenta NFP
Populus euphratica NFP2
Populus trichocarpa NFP2
Populus trichocarpa NFP1
Populus euphratica NFP1
Populus euphratica NFP3
Citrus clementina NFP
Citrus sinensis NFP
Theobroma cacao NFP
Gossypium raimondii NFP
Carica papaya NFP
Eucalyptus grandis NFP1
Eucalyptus grandis nfp2
Punica granatum NFP
Fagus sylvatica nfp1
Castanea mollissima nfp1
Juglans regia NFP1
Alnus glutinosa NFP
Casuarina glauca NFP
Datisca glomerata NFP2
Momordica charantia NFP2
Cucurbita maxima NFP2
Cucumis sativus nfp2
Cucumis melo NFP2
Begonia fuchsioides NFP
Datisca glomerata NFP1
Momordica charantia NFP1
Cucurbita maxima NFP1
Cucumis sativus NFP1
Cucumis melo NFP1
Dryas drummondii NFP1
Rubus occidentalis NFP
Fragaria vesca NFP
Rosa × damascena NFP
Prunus persica NFP
Prunus mume NFP
Malus domestica NFP1
Pyrus communis NFP
Pyrus x bretschneideri NFP
Ziziphus jujuba NFP
Discaria trinervis NFP1
Ceanothus thyrsiflorus NFP1
Morus notabilis NFP1
Humulus lupulus NFP
Cannabis sativa NFP
Parasponia rugosa NFP1
Parasponia rigida NFP1
Parasponia andersonii NFP1
Trema levigata NFP1
Trema orientalis NFP1
Trema tomentosa NFP1
Fagus sylvatica nfp2
Castanea mollissima nfp2
Datisca glomerata NFP3
Malus domestica nfp2
Dryas drummondii nfp3
Dryas drummondii NFP2
Ceanothus thyrsiflorus NFP2
Discaria trinervis NFP2
Morus notabilis nfp2
Trema levigata nfp2
Parasponia andersonii NFP2
Parasponia rugosa NFP2
Parasponia rigida NFP2
Cercis canadensis NFP
Mimosa pudica NFP2
Chamaecrista fasciculata NFP2
Chamaecrista fasciculata NFP3
Mimosa pudica NFP1
Chamaecrista fasciculata NFP1
Castanospermum australe LYR1
Lotus japonicus LYS11
Cicer arietinum LYR1
Trifolium subterraneum LYR1
Medicago truncatula LYR1
Sophora flavescens NFP
Lupinus angustifolius NFP
Arachis duranensis NFP
Arachis ipaensis NFP
Glycine max NFR5b
Glycine max NFR5a
Phaseolus vulgaris NFP
Vigna angularis NFP
Vigna radiata NFP
Lotus japonicus NFR5
Cicer arietinum NFP
Galega orientalis NFR5
Galega officinalis NFR5
Medicago truncatula NFP
Pisum sativum SYM10
Trifolium subterraneum NFP
Trifolium pratense NFP
Tree scale: 0.1
76.4/77/0.859
NFP-I clade
NFP-II clade
Fabales
Fagales
Cucurbitales
Rosales
Nitrogen-fixation clade

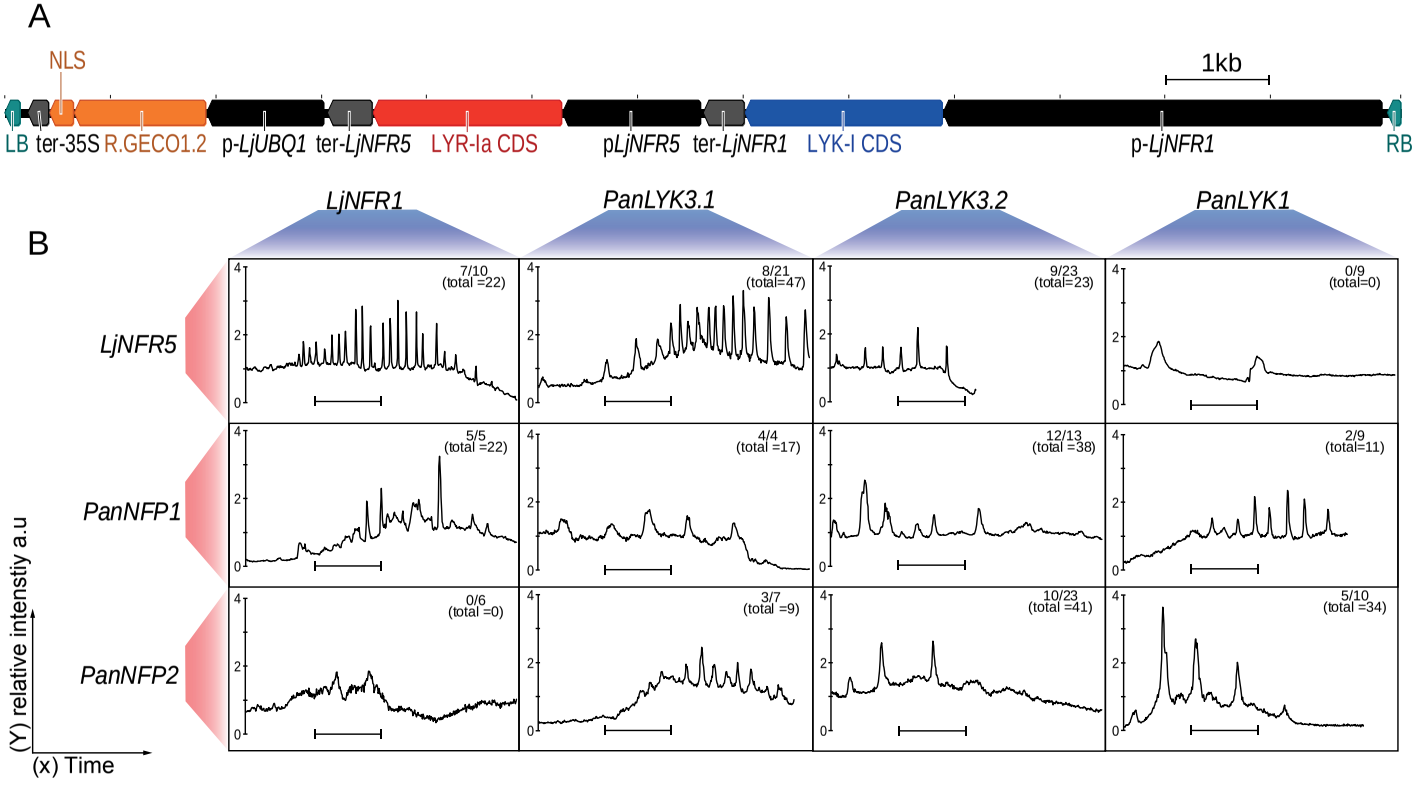

A
NLS
1kb
LB
ter-35S
R.GECO1.2
p-LjUBQ1
ter-LjNFR5
LYR-Ia CDS
pLjNFR5
ter-LjNFR1
LYK-I CDS
p-LjNFR1
RB
B
LjNFR1
PanLYK3.1
PanLYK3.2
PanLYK1
LjNFR5
PanNFP1
PanNFP2
7/10 (total =22)
8/21 (total=47)
9/23 (total=23)
0/9 (total=0)
5/5 (total =22)
4/4 (total =17)
12/13 (total =38)
2/9 (total=11)
0/6 (total =0)
3/7 (total =9)
10/23 (total =41)
5/10 (total =34)
(Y) relative intenstiy a.u
(x) Time

Figure 4. Parasponia *PanNFP1*, *PanNFP2*, *PanLYK1* and *PanLYK3* complement a lotus *Ljnfr1;Ljnfr5* mutant for rhizobium-induced Ca^{2+} oscillation. **(A)** Schematic representation of the T-DNA region of the binary construct used for *A. rhizogenes*-based root transformation of a lotus *Ljnfr1;Ljnfr5* double mutant. cDNA clones of LYK-I (marked blue) or LYR-Ia type genes (marked red) were cloned in identical fashion. cDNA clones were inserted between native promoter (marked black) pLjNFR1 (4,171bp) or pLjNFR5 (1,314bp), and native terminator (marked gray) sequences ter-LjNFR1 (394 bp) or ter-NFR5 (432 bp). pLjUBQ1::R.GECO1.2-nls:CaMV35S-ter (marked orange) was used to visualize nuclear calcium oscillation. The left border (LB) and right border (RB) (marked green) flank the T-DNA region. **(B)** Representative traces of nuclear Ca^{2+} oscillation, as observed in different combinations of LYK-I (red) and LYR-Ia (blue) type receptors introduced in a lotus *Ljnfr1;Ljnfr5* double mutant. Note that the receptor combinations *PanLYK1;LjNFR5* and *LjNFR1;PanNFP2* didn't complement for Ca^{2+} oscillation. Traces were recorded ~10 min post-application of LCOs extracted from *M. loti* R7A (~10^{-9} M). Numbers denote spiking roots vs the number of roots analyzed. The number in brackets denotes the total number of spiking nuclei observed. Scale bar = 10 minutes. Y-axis is the relative fluorescence intensity compared to defined baseline in arbitrary units.

reconstruction, including additional nodulating and non-nodulating species, supported the occurrence of NFP-I and NFP-II subclades in the LYR-IA orthogroup and showed that this duplication associates with the origin of the nitrogen-fixing clade **(Figure 3; Data set S2)**. Several Actinorhizal species possess gene copies in both NFP subclades; including *Datisca glomerata*, *Dryas drummondii*, and *Ceanothus thyrsiflorius.* All these species nodulate with diazotrophic *Frankia* species of taxonomic cluster-II, which possess LCO biosynthesis genes. An NFP-II-type orthologous gene is notably absent in Actinorhizal species that are exclusively nodulated by *Frankia* species of cluster-I or cluster-III that lack LCO biosynthesis genes; e.g. *Alnus glutinosa* and *Casuarina glauca* **(Figure 3)** (Pawlowski and Demchenko, 2012; Griesmann et al., 2018; Salgado et al., 2018; Van Nguyen et al., 2019). In line with what was reported for the non-nodulating Rosales species (van Velzen et al., 2018b), NFP-II-type pseudogenes can be found in the genomes of the non-nodulating Fagales species *Castanea mollissima* and *Quercus fagus*. This shows a strict association of the presence of a functional NFP-II-type gene and LCO-driven nodulation, suggesting that the NFP-II subclade represents LCO receptors that function exclusively in nodulation.

Parasponia PanNFP1, PanNFP2, PanLYK1 and PanLYK3 can perceive rhizobium LCOs

Based on the orthologous relation to legume LCO receptors, we considered *PanLYK3* (both variants) and *PanNFP2* as the most likely candidates to encode rhizobium LCO receptors in parasponia. We noted that, in contrast to *PanLYK3*, *PanLYK1* is exclusively expressed in roots and nodule tissue **(Figure 2B),** suggesting this gene may also function in a symbiotic context. Therefore, we decided to include this gene in further studies. Finally, we included also *PanNFP1*, since an earlier study

based on RNA interference (RNAi) in transformed parasponia roots showed that this gene functions in nodulation (Op den Camp et al., 2011). To test whether these four parasponia genes can function as rhizobium LCO receptors, we conducted two complementary experiments. First, we introduced parasponia receptor pairs into a lotus *Ljnfr1;Ljnfr5* double mutant aiming to determine whether these parasponia *P. andersonii* receptors can trans-complement for LCO-induced Ca^{2+} oscillation. Second, we generated CRISPR-Cas9 knockout mutants in parasponia to study their role in nodulation.

We selected lotus for trans-complementation studies as its microbial host *Mesorhizobium loti* strain R7A can also nodulate parasponia **(Figure S3A-C)**. By using *A. rhizogenes*-mediated root transformation, we tested six combinations of parasponia heterodimeric receptor pairs under control of the promoter and terminator of lotus LjNFR1 and LjNFR5 **(Figure 4A)**. These promoters showed to be functional in complementation of the lotus *Ljnfr1-1;Ljnfr5-2* double mutant **(Figure S3D-H**). For the trans-complementation constructs, we included the nuclear localized calcium sensor R-GECO1.2, allowing visualization of nuclear Ca^{2+} oscillations (Zhao et al., 2011). In wild-type lotus roots, Ca^{2+} oscillation was most strong in young root hair cells, whereas this response is not recorded in the *Ljnfr1-1;Ljnfr5-2* double mutant (Figure S3I,J; movie S1) (Miwa et al., 2006). Analysing the transgenic roots expressing parasponia receptor combinations revealed that nine out of eleven tested combinations elicit Ca^{2+} oscillation, although less regular in shape and frequency when compared to the positive control **(Figure 4B; movie S2)**. Interestingly, the receptor combinations *PanLYK1;LjNFR5* and *LjNFR1;PanNFP2* did not elicit any Ca^{2+} oscillation response, whereas both parasponia receptors are -at least partially- functional as an *M. loti* LCO receptor when combined with a parasponia counterpart **(Figure 4B)**. Upon inoculation with *M. loti* R7A, only nodule-like structures were observed on roots trans-complemented with different parasponia receptor combinations (4 weeks post-inoculation), but not with heterologous receptor pairs **(Table S2)**. We sectioned the largest nodule-like structures, which were present on *PanLYK3.2;PanNFP2* and *PanLYK1;PanNFP1* transformed plants. This showed the absence of intracellular rhizobium infections **(Figure S3K-P)**. Taken-together, the trans-complementation studies of a lotus *Ljnfr1;Ljnfr5* mutant indicated that all four parasponia receptors -*PanLYK1, PanLYK3, PanNFP1* and *PanNFP2*- have the potential to function as receptors for *M. loti* LCOs, but none could fully trans-complement a lotus *Ljnfr1-1;Ljnfr5-2* double mutant for nodulation.

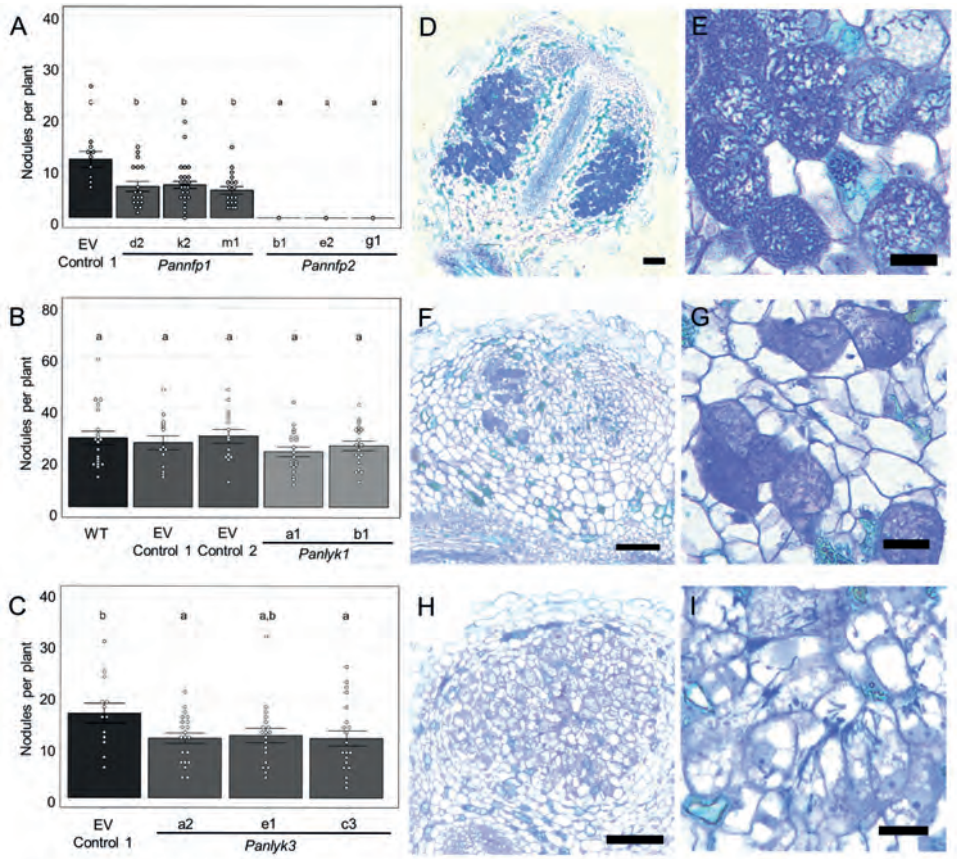


Figure 5. Parasponia *Pannfp1*, *Pannfp2* and *Panlyk3* mutants are affected in nodulation. Data are represented as mean ± SE, dots represent individual data points. Letters denote statistical significance based on one-way ANOVA and Tukey post-hoc contrasts P>0.05.
(**A**) Nodule numbers in parasponia CRISPR-Cas9 mutant lines *Pannfp1* d2 (n=18), k2 (n=31) and m1 (n=19) and *Pannfp2* b1 (n=19), e2 (n=10) and g1 (n=9), 5 wpi with *M. plurifarium* BOR2. EV Control 1 (n=12) represents a positive control line transformed with a binary vector not containing sgRNAs. (**B**) Nodule numbers in parasponia CRISPR-Cas9 mutant lines *Panlyk1* a1 (n=19) and b1 (n=20), 5 wpi with *M. plurifarium* BOR2. EV control 1 (n=14) and EV control 2 (n=14) represent two independent positive control lines transformed with a binary vector not containing sgRNAs. WT (n=20) represent untransformed plantlets. (**C**) Nodule numbers in parasponia CRISPR-Cas9 mutant lines *Panlyk3* a2 (n=21), c3 (n=21) and e1 (n=19), 5 wpi with *M. plurifarium* BOR2. EV Control 1 (n=14). (**D,E,F,G,H,I**) Toluidine blue-stained section of representative nodules grown with *M. plurifarium* BOR2 (**D**) Wild type parasponia transformed with an EV-1 construct expressing Cas9. Scale bar: 100 µm. (**E**) Infected nodule cells containing fixation threads formed on EV-1 plants. Scale bar: 20 µm. **(F)** Infected nodule of *Panlyk3* line e2. Note patches of infected cells. Scale bar: 100 µm. (**G**) infected nodule cells of the *Panlyk3* line e2 containing fixation threads. Scale bar: 20 µm. (**H**) Empty nodule of *Panlyk3* line e2. Note the absence of fully infected cells. Scale bar: 100 µm. (**I**) Nodule cells of the *Panlyk3* line e2 containing infection threads but no fixation threads. Scale bar: 20 µm.

***Parasponia PanNFP1, PanNFP2, PanLYK1* and *PanLYK3* function in nodulation**

We recently established an efficient *Agrobacterium tumefaciens*-mediated transformation protocol for parasponia, which allows the generation of CRISPR-Cas9 mutant plantlets in a ~3 month timeframe (van Zeijl et al., 2018; Wardhani et al., 2019). This enabled us to test by mutagenesis whether *PanLYK1, PanLYK3, PanNFP1* and *PanNFP2* are essential for rhizobium-induced nodule formation. We aimed to generate small deletions of 100-300 bp in the area covering the LysM domains by using two or three single guide RNAs (sgRNAs) that have no potential high identity off-targets. In the case of *PanLYK3* the transmembrane domain was targeted in order to mutate both alternative start variants. Additionally, we targeted specifically *PanLYK3.1* and *PanLYK3.2* by designing specific guides on the first exon. Selected single guides only had off-targets with at least three mismatches or two indels, based on alignments to the parasponia reference genome. Shoots regenerated after *A. tumefaciens*-mediated co-cultivation were genotyped using PCR and subsequent sequence analysis to detect potential mutations at the CRISPR target sites. Only T_0 shoots with a >75 bp deletion between the two target sites or edits generating a frameshift were considered for propagation and subsequent further evaluation. At least two independent mutant alleles were generated per gene, with the exception of *Panlyk3.1* for which only a single suitable allele could be identified **(Data set S3)**. Putative off-target sites that occur in coding sequence regions were amplified by PCR and subsequently sequenced by sanger sequencing. Subsequently, *PanNFP1* was sequenced in *PanNFP2* lines, and *PanNFP2* in *PanNFP1* lines **(Data set S3)**. No off-target mutations at these locations were identified. The selected tissue culture lines were *in vitro* propagated and rooted, so they could be used for experimentation.

We compared the nodulation phenotype of *Panlyk1, Panlyk3, Pannfp1* and *Pannfp2* knockout mutants in independent experiments, using empty vector (EV) transformed lines as control **(Figure 5; Figure S4)**. All three independent *Pannfp2* mutant lines showed to be unable to form nodules or nodule-like structures (5 weeks post inoculation, wpi) with strain *Mesorhizobium plurifarium* BOR2, demonstrating the requirement for this gene in the nodulation trait **(Figure 5A)**. Additionally, we noted a reduced nodulation efficiency of all three independent *Pannfp1* mutant lines. This is in line with earlier findings using RNAi to target *PanNFP1* in *A. rhizogenes*-transformed parasponia roots (Op den Camp et al., 2011), demonstrating that *Pannfp1* controls nodulation efficiency, but is not essential for rhizobium intracellular infection. Previously, we reported that *PanNFP1* RNAi-nodules have a strong infection phenotype when inoculated with the *Sinorhizobium fredii* strain NGR234 (Op den Camp et al., 2011). We did not observe such an infection phenotype in nodules induced by *M. plurifarium* BOR2 on *Pannfp1* knockout mutant plants **(Figure S4)**. In order to determine whether the *Pannfp1* infection phenotype is strain

dependent, we nodulated plants, also with *S. fredii* NGR234. This strain showed to be less optimal under the chosen conditions (agroperlite supplemented with EKM medium and *S. fredii* NGR234.pHC60 at OD 0.05). In an effort to optimize nodulation efficiency with this strain, we used river sand and scored nodulation 8 weeks post-inoculation. Under these conditions, no difference between *Pannfp1* and EV-control was observed. Nodules formed on *Pannfp1* mutant plants were infected normally **(Figure S4)**.

Similarly to *Pannfp1* mutant plants inoculated with *M. plurifarium* BOR2, we found a reduced nodulation efficiency in parasponia *Panlyk3* knockout mutants, but not in *Panlyk3.1* and *Panlyk3.2* variant specific mutant alleles, nor in *Panlyk1* mutants **(Figure 5; Figure S4)**. To determine whether nodules formed on *Panlyk1* and *Panlyk3* mutants have an infection phenotype, we analysed thin sections. In contrast to legumes, parasponia doesn't guide rhizobia in infection threads towards the nodule primordia. Instead, rhizobia enter via apoplastic cracks in epidermis and cortex, and only form infection threads to penetrate nodule cells. Once inside, infection threads develop into fixation threads, which are wider -having two phyla of bacteria aligned compared to one in infection threads- and possess a thinner cell wall (Lancelle and Torrey, 1984; Lancelle and Torrey, 1985). *Panlyk1* mutant nodules showed no defects in infection thread structure or the transition from infection threads to fixation

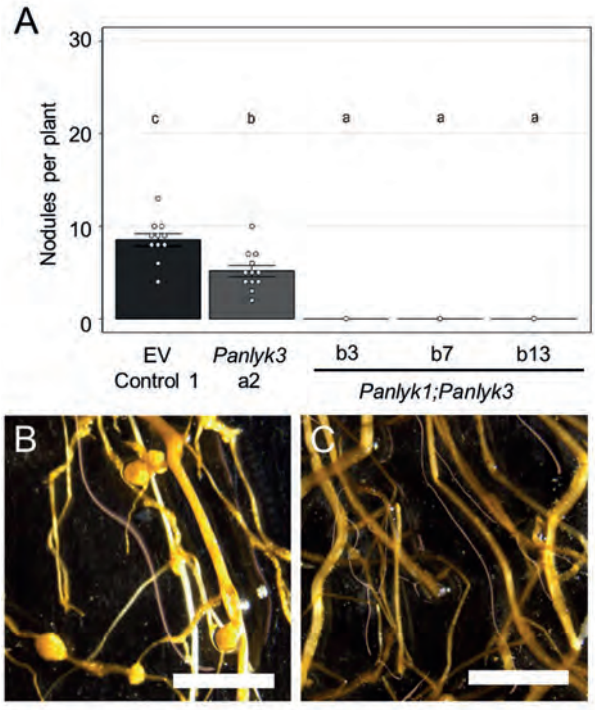


Figure 6. Parasponia *PanLYK1* and *PanLYK3* act redundantly in nodulation.
(**A**) Average nodule numbers per plant in EV control 1 (n=11) and re-transformed *Panlyk3* a2 line (n=12) and *Panlyk1;Panlyk3* double mutant lines b3 (n=10), b7 (n=5) and b13 (n=10), 5 wpi with *Mesorhizobium plurifarium* BOR2. Data are represented as mean ± SE, dots represent individual data points. Letters denote statistical significance based on one-way ANOVA and Tukey post-hoc contrasts P>0.05. (**B**) Roots with nodules of EV control 1, 5 wpi with *M. plurifarium* BOR2. Scale bar: 5 mm. (**C**) Roots without nodules of the *Panlyk1;Panlyk3* double mutant (line b3) 5 wpi with *M. plurifarium* BOR2 Scale bar: 5 mm

threads. In the case of *Panlyk3*, nodules were relatively small and had diverse phenotypes. Out of 45 sectioned nodules of the line *Panlyk3*-e2, 22 were infected like wild type, 15 contained only infection threads, but no fixation threads, and 8 showed an intermediate phenotype with few infected cells **(Figure 5 F-I. Figure S4)**. To confirm that the infection phenotype is a result of a full *Panlyk3* knockout mutation, we sectioned 28 nodules of the independent knockout line *Panlyk3*-c3. This revealed similar results; 11 nodules normally infected, 11 contained only infection threads, and 6 nodules with an intermediate phenotype. Next, we determined whether this infection phenotype is controlled specifically by or either *PanLYK3.1* or *PanLYK3.2,* which showed not to be the case **(Figure S4)**. As ~50% of the nodules formed on the parasponia *Panlyk3* mutant plants displayed a wild-type phenotype, it suggests redundancy in gene functioning. Interestingly, *S. fredii* NGR234 could not nodulate *Panlyk3* mutants, which suggest this strain is fully dependent on PanLYK3 controlled signal transduction **(Figure S4)**.

As parasponia did not experience any gene duplication events in the LYK-Ib clade, *PanLYK1* in the LYK-Ia clade is the closest homolog of *PanLYK3*. In order to investigate whether the *PanLYK1* gene is functionally redundant with *PanLYK3* in cases of *M. plurifarium* BOR2 inoculation, we generated a *Panlyk1;Panlyk3* double mutant. To do so, a binary construct with the two sgRNAs targeting *PanLYK1* was used for re-transformation of the *Panlyk3* mutant (line a2). We obtained three independent *Panlyk1;Panlyk3* mutants **(data set S3)**. *M. plurifarium* BOR2 inoculation experiments revealed that all *Panlyk1;Panlyk3* double mutant lines were unable to form any nodule or nodule-like structure **(Figure 6)**. To confirm that the nodulation minus phenotype in the *Panlyk1;Panlyk3* lines is not due to any off-target mutation, we conducted complementation studies using *A. rhizogenes*-mediated root transformation. As the putative promoter of *PanLYK3* is rather complex due to the occurrence of alternative transcriptional start sites (Figure 2), we used the *LjNFR1* promoter, as well as the constitutive *AtUBQ10* and *CaMV35S* promoters, to drive a CRISPR-resistant allele of *PanLYK3.1* (*PanLYK3cr*). Compound plants carrying transgenic roots expressing *PanLYK3cr* could be nodulated by *M. plurifarium* BOR2 **(Figure S5)**. Together, this showed that in parasponia, *PanLYK1* and *PanLYK3* act redundantly in root nodule formation. (For complementation studies of *Pannf2*, see below).

The results demonstrate that parasponia *PanLYK1, PanLYK3, PanNFP1* and *PanNFP2* function in rhizobium LCO-driven nodulation. *PanLYK3* and *PanNFP2* are orthologous to legume *LjNFR1/MtLYK3* and *LjNFR5/MtNFP*, indicating a shared evolutionary origin of LCO-driven nodulation in both taxonomic lineages. As

PanLYK1 and *PanLYK3* evolved from a duplication predating the emergence of the nitrogen-fixing clade, it suggests that LCO signalling is an ancestral function of these LYK-I receptors.

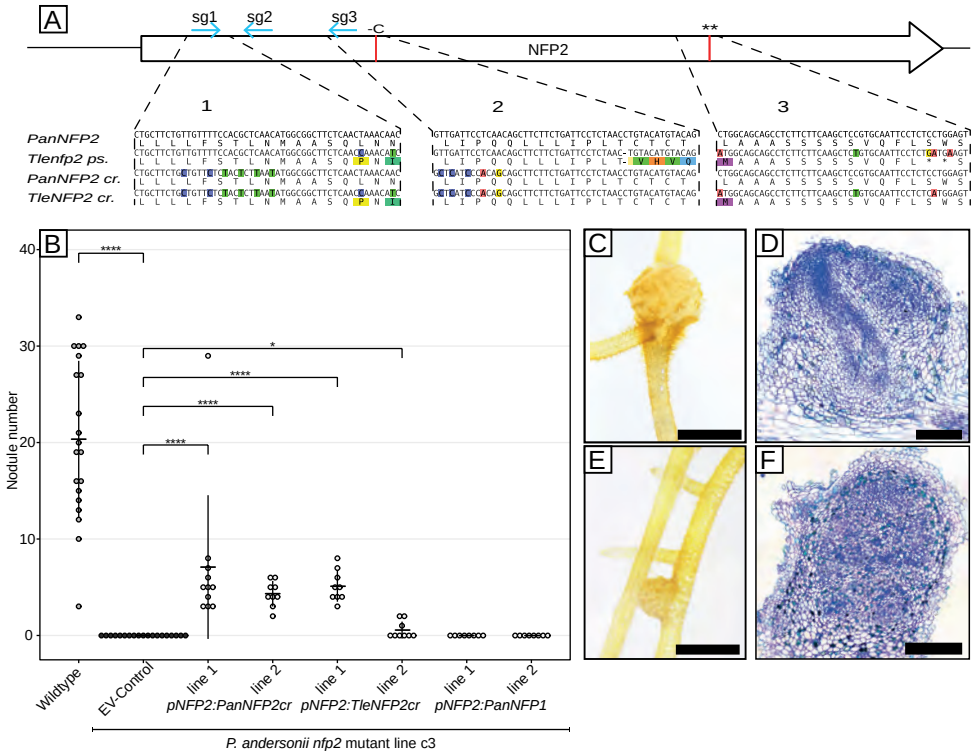


Figure 7. A repaired *Trema levigata nfp2* pseudogene can replace *PanNFP2* for nodule formation. (A) Schematic representation of NFP2 coding region with indicated replacements to avoid CRISPR targeting of inserted *NFP2* genes of *P. andersonnii* (*PanNFP2cr*) and a repaired *T. levigata* (*TleNFP2cr*). Blue arrows: Guide RNA target sites. Red lines: *TremaT. levigata* mutations..Region 1. Replacement of six codons at the sg1 site. Region 2. Replacement of five codons at the sg3 site plus repair of the *T. levigata* indel (red line). Region 3. Repair of the double stop codon in *Trema levigata*. (red line, black asterisks). The replacement of five codons at the sg2 site is not shown. **(B)** *PanNFP2cr* and repaired *TleNFP2cr* can restore nodulation in the *Pannfp2* mutant line C3 when driven by the *PanNFP2* promoter, whereas *PanNFP1* cannot. Nodulation scored 5 wpi with *M. plurifarium* BOR2. Error bars represent the SD of the mean, statistical significance by Mann–Whitney–Wilcoxon (MWW). P>0.05 not significant (ns), P<0.05 *, P<0.01 **, P<0.001, ***P <0.0001, ****P <0.00001. **(C,D)** Nodule and section of a *pNFP2:PanNFP2cr* line 1. **(E,F)** Nodule and section of *pNFP2:TleNFP2cr* line 1. **(C,E)** scale bar 2mm **(D,F)** scale bar 100µm.

A repaired *Trema levigata NFP2* pseudogene, but not *PanNFP1,* can functionally complement a parasponia *nfp2* mutant

PanNFP1 and *PanNFP2* differ in expression pattern. Whereas both genes are expressed in root tissue, only *PanNFP2* is upregulated in nodules **(Figure 2)** (Van Velzen et al., 2018b). We questioned whether the difference in symbiotic functioning between both genes is the result of regulatory evolution. To test this, we first identified a functional promoter region of *PanNFP2*. *A. tumefaciens* mediated transformation showed that a 2.75 kbps *PanNFP2* upstream region can be used to functionally complement the parasponia *Pannfp2* mutant when using a *PanNFP2* CRISPR-resistant allele (*PanNFP2cr*). Two independent lines formed 7±7 and 4±1 nodules 5 weeks post inoculation with *M. plurifarium* BOR2 **(Figure 7)**. However, when we

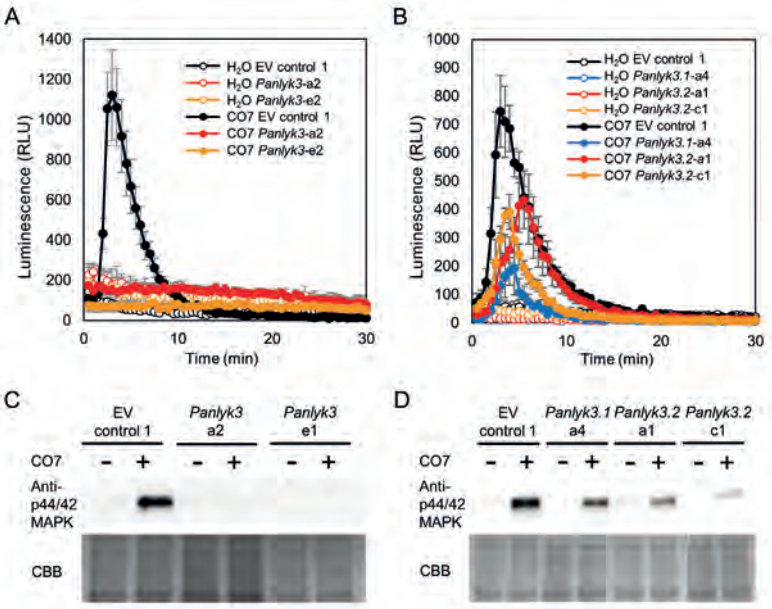


Figure 8. Parasponia *PanLYK3* is essential for chitin triggered immunity responses in roots. **(A,B)** Production of ROS measured upon treatment with 100 µM CO7 (filled circle) or H_2O (open circle) with (**A**) EV control 1 plants (black), *Panlyk3* line a2 (red) and *Panlyk3* line e1 (orange) (**B**) EV control 1 plants (black), *Panlyk3.1* line a4 (blue), *Panlyk3.2* line a1 (red) and *Panlyk3.2* line c1 (orange). For A and B, data are the average of at least three independent biological replicates ±SE. Luminescence is measured in relative light units RLU **(C,D)** Phosphorylation of mitogen activated protein kinase (MAPK) analysed by immunoblot using an anti-p44/42 MAPK antibody upon treatment with 100 µM CO7 (upper panel). Equal loading was confirmed by CBB staining (bottom panel). Results shown are a representative out of three independent experiments. (**C**) MAPK phosphorylation in root pieces of EV control 1, *Panlyk3* line a2 and *Panlyk3* line e1. (**D**) EV control 1, *Panlyk3.1* line a4, *Panlyk3.2* line a1 and *Panlyk3.2* line c1.

used *PanNFP1* driven by the *PanNFP2* promoter, no trans-complementation of the parasponia *nfp2* mutant phenotype was observed. This suggests that there is a functional difference in the encoded *PanNFP1* and *PanNFP2* receptors.

Next, we questioned whether the *nfp2* pseudogene as present in several non-nodulating Rosales species may have encoded a functional symbiosis receptor. To test this, we focussed on the the *nfp2* pseudogene of *Trema levigata*, as it has only three mutations that cause a disturbance of the open reading frame **(Figure 7).** We repaired these three mutations, using *PanNFP2* as a template, resulting in an engineered CRISPR resistant *TleNFP2cr* that encodes for a LysM-type receptor protein of 582 amino acids, similar to *PanNFP2* of parasponia. We tested whether *TleNFP2cr* driven by the *PanNFP2* promoter can trans-complement the parasponia *Pannfp2* mutant. *A. tumefaciens* transformation resulted in two lines that can form functional root nodules 5 weeks post-inoculation with *M. plurifarium* BOR2. This supports the hypothesis that *T. levigata nfp2* encoded a functional symbiosis receptor prior to the pseudogenization of this gene.

Parasponia *PanLYK3* is essential for chitin triggered immune responses and controls AM symbiosis in coherence with *PanLYK1*

Next, we aimed to determine whether the parasponia LysM-type receptors that control nodulation are also involved in other processes, as this may provide insights into ancestral functions of these genes. Some LysM-type receptors of the LYK-I clade are known to function in chitin-triggered immunity and/or the arbuscular mycorrhizal symbiosis; e.g. lotus LjCERK6, medicago MtLYK9/MtCERK1, arabidopsis AtCERK1, tomato SlLYK1 and rice OsCERK1 **(Figure 1)** (Miya et al., 2007; Wan et al., 2008; Shimizu et al., 2010; Bozsoki et al., 2017; Liao et al., 2018; Feng et al., 2019; Gibelin-Viala et al., 2019a; He et al., 2019). Similarly, some experimental evidence using transient silencing assays indicated that LysM-type receptors of the LYR-IA clade function in mycorrhization, including parasponia PanNFP1 (Op den Camp et al., 2011). In line with this, we aimed to confirm this phenotype in stable *Pannfp1* knockout mutants, and determine whether other parasponia symbiotic LysM-type receptors may function also in AM symbiosis and/or chitin-induced innate immunity signalling.

First, we investigated whether the parasponia LysM-type receptors mutants are affected in chitin-triggered immunity responses. To do so, two complementary assays were used; a chitin-induced ROS-burst production and MITOGEN-ACTIVATED PROTEIN KINASE3 (MAPK3) / MAPK6 phosphorylation assay. Chitin heptamers (CO7) effectively induced a ROS burst parasponia root segments at concentrations of <1 µM when incubated at 28°C, the regular growth temperature of *Parasponia* species. **(Figure 8A, Figure S6B)**. To test whether ROS bursts can also be triggered

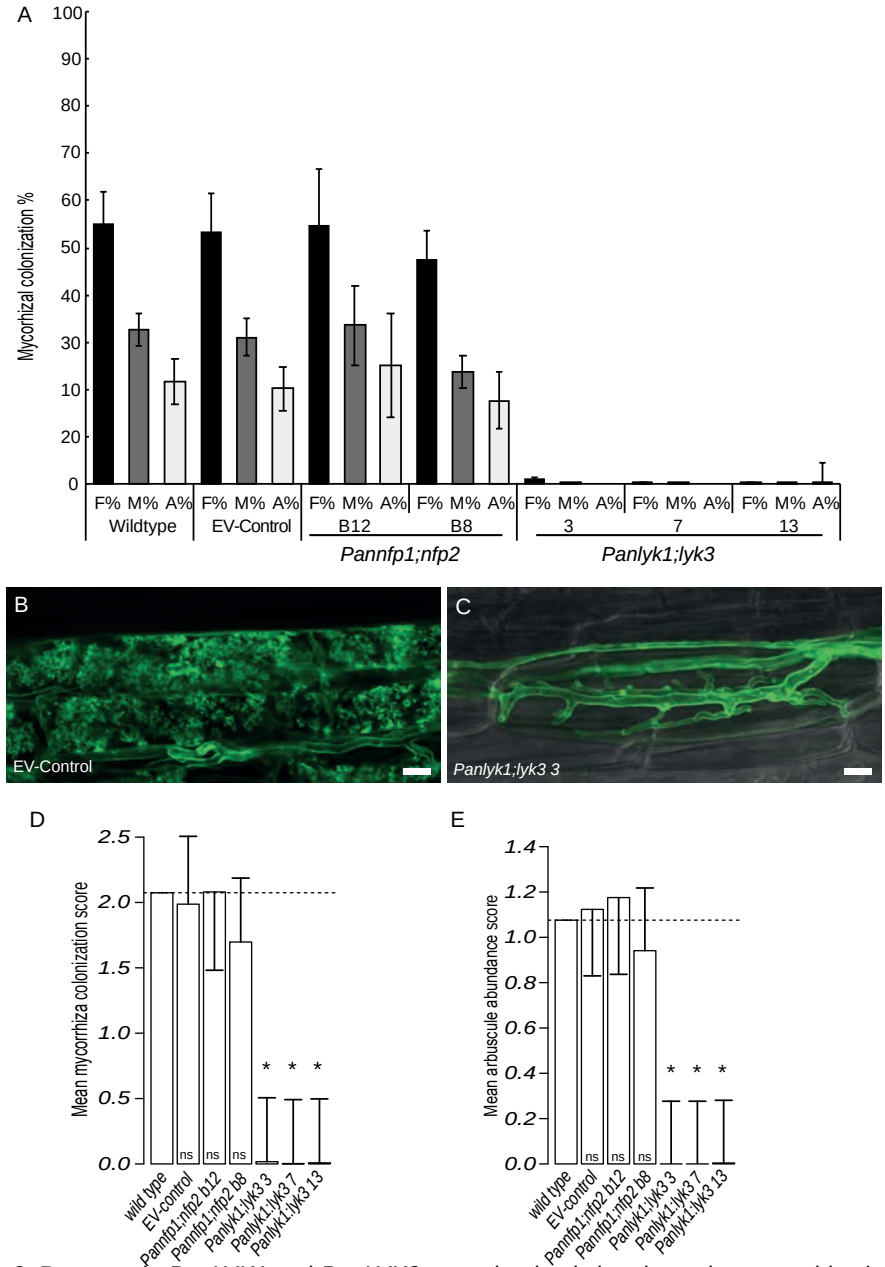


Figure 9. Parasponia *PanLYK1* and *PanLYK3* act redundantly in arbuscular mycorrhization. **(A)** The parasponia *Panlyk1;Panlyk3* double mutant shows a strongly reduced colonization compared to wild type and control parasponia roots. Parasponia *Pannfp1;Pannfp2* mutants are not significantly affected. Frequency and Arbuscule abundance classes according to Trouvelot et. al. 1986 (Trouvelot, A, Kough J L, Gianinazzi-Pearson V, 1986). F%: colonization frequency in the root system. M%: intensity of mycorrhizal colonization A%: Arbuscule abundance in the root system. Error bars represent the SE of 10 biological replicates Scored

at 6 weeks post inoculation using 250 spores of *Rhizophagus irregularis* strain DOAM197198 (Trouvelot, A, Kough J L, Gianinazzi-Pearson V, 1986). **(B)** Highly branched arbuscules formed in EV-Control plants 6 weeks PI stained with WGA-alexa488. Scale bar 10 µm. **(C)** Phenotype of stunted arbuscules formed in the *Panlyk1;Panlyk3* double mutant stained with WGA-alexa488. Scale bar 10 µm. **(D,E)** Statistical analysis of raw (observed) data. **(D)** Mean colonization frequency score (classes 0 to 5) and **(E)** Mean arbuscule score (classes 0 to 3). Classes presented in Trouvelot et. al. 1986 (Trouvelot, A, Kough J L, Gianinazzi-Pearson V, 1986). Reduced mycorrhizal colonization and arbuscule formation in *Panlyk1;Panlyk3* mutants is considered significant compared to Wwild type. Error bars represent the Bonferroni corrected Least Significant Difference. Error bars non overlapping with mean wild type value are considered significant. Dashed line indicates mean Wildtype score.

by rhizobium LCOs, we used the extracts of *M. loti* R7A and *Rhizobium tropici* CIAT899. These two strains can nodulate parasponia but produce structurally different LCOs (López-Lara et al., 1995; Folch-Mallol et al., 1996). However, neither triggered a ROS burst in parasponia roots **(Figure S6A)**. To determine whether CO7-induced ROS bursts were associated with phosphorylation of parasponia MAPK3 and MAPK6 homologs, we used an anti–phospho-p44/42 HsMAPK antibody, which detects phosphorylated MAPK3 and MAPK6 of different plant species (Yamaguchi et al., 2013; Bozsoki et al., 2017). Parasponia possesses a single *PanMAPK3* and a single *PanMAPK6* gene, which each encodes a protein with a conserved Thr202/Tyr204 phosphorylation site **(Figure S6C)**. Upon CO7 application (100 µM, 10 min.), a MAPK3/6 phosphorylation pattern can be detected, which is not observed upon application of *M. loti* or *R. tropici* LCO extracts **(Figure 8C; Figure S6D)**. Next, we determined whether parasponia LysM-type receptor mutants are affected in responses to chitin CO7 oligomers. *Pannfp1*, *Pannfp2* and also a newly created *Pannfp1;Pannfp2* double mutant showed a wild-type ROS-burst and MAPK3/6 phosphorylation profile **(Figure S6; Data set S3)**. Similarly, the *Panlyk1* mutant showed a ROS burst and MAPK3/6 phosphorylation profile, as did wild -type root segments **(Figure S6E,F).** In contrast, parasponia *Panlyk3* mutant lines lacked a chitin triggered ROS-burst and showed no p44/42 MAPK phosphorylation (**Figure 8)**. Individual exon knockout *Panlyk3.1* or *Panlyk3.2* mutants both showed ROS production and MAPK3/6 phosphorylation upon application of 100 µM CO7, however at reduced levels **(Figure 8B, D)**. Taken together, these data show that *PanLYK3* - which is the only parasponia gene in the LYK-Ib clade is essential for chitin innate immune signalling in roots.

Studies in parasponia, tomato, medicago and rice revealed that LYR-IA and LYK-I putative orthologous genes have functions in AM symbiosis (Miyata et al., 2014; Zhang et al., 2015; Buendia et al., 2016; Miyata et al., 2016; Carotenuto et al., 2017; Liao et al., 2018; Feng et al., 2019; Gibelin-Viala et al., 2019; He et al., 2019). Interestingly, we noted that the NFP-I-type gene is pseudogenized in European beech (*Fagus sylvatica*) and Chinese chestnut (*Castanea mollissima*).

Both species have lost AM symbiosis in favour of an ectomycorrhizal symbiosis **(Figure 3)** (Werner et al., 2018). We conducted an RNA-sequencing experiment on parasponia roots mycorrhized by *Rhizophagus irregularis* strain DOAM197198. Several marker genes for mycorrhization showed to be enhanced in expression in mycorrhized parasponia root samples; including *PanSTR1, PanSTR2, PanPT4, PanVPY, PanD27, PanRAD1* and *PanRAM1* **(Figure S7).** Also, this suggested that *PanNFP1* is expressed higher than *PanNFP2* under these conditions **(Figure S7)**. However, no significant differential regulation of any of the studied LysM-type receptor encoding genes was detected between phosphate starved control roots and mycorrhized root samples **(Figure S7)**. To determine whether parasponia symbiotic LysM-type receptors also function in AM symbiosis, we conducted three independent experiments using *in vitro* propagated mutant plantlets inoculated with 250 spores of *R. irregularis* DOAM197198. The average colonization and arbuscule formation frequency were scored 6 weeks post-inoculation. These experiments revealed substantial variation in mycorrhization efficiency between replicates, though no clear impaired AM symbiosis phenotype could be observed in any of the single mutants, including *Pannfp1.* Strikingly, *Panlyk1* showed a significant increase in colonization and arbuscule frequency **(Figure S8ABC)**. Analysing both double mutants -*Pannfp1;Pannfp2* and *Panlyk1;Panlyk3*- revealed a strong AM symbiosis phenotype only in the latter **(Figure 9, Figure S8)**. The fungal colonization of the *Panlyk1;Panlyk3* mutant was severely affected, with only a few infections observed. Confocal imaging of WGA-alexa488 stained roots showed that besides the level of colonization, also the morphology of the few arbuscules that were formed was affected in *Panlyk1;Panlyk3* plants. In wild type plants, many cortical cells were filled with arbuscules that were finely branched and occupied most of the cell. In contrast, the few hyphae that enter cortical cells in the *Panlyk1;Panlyk3* mutant were unable to form mature arbuscules, either because the fungus fails to switch to fine branching, or because a limited number of fine branches is made **(Figure 9)**. As both *Panlyk1* and *Panlyk3* single mutant plants do not show such impaired mycorrhizal phenotype, we conclude that both genes function in conjunction to control mycorrhizal infection.

Taken together, these experiments revealed that PanLYK1 and PanLYK3 can function in multiple processes, including rhizobium nodulation (PanLYK1 and PanLYK3), arbuscular mycorrhizal symbiosis (PanLYK1 and PanLYK3), and chitin innate immune signalling (PanLYK3). This suggests that no subfunctionalization of these receptors is required to allow functioning in the rhizobium nitrogen-fixing nodulation trait

Discussion

We used parasponia as a comparative system to legumes to obtain insight into the evolutionary trajectory of LysM-type rhizobium LCO receptors. By conducting phylogenetic analysis, trans-complementation studies in a lotus LCO receptor double mutant, and CRISPR-Cas9 mutagenesis in parasponia, we identified four LysM-type receptors that function in LCO-driven nodulation in a non-legume. Two of these, *PanLYK3* and *PanNFP2*, are putative orthologs to known legume rhizobium LCO receptors *LjNFR1/MtLYK3* and *LjNFR5/MtNFP,* respectively. As the *Parasponia* and legume lineages diverged early in the nitrogen-fixing clade (>100 MYA), the use of orthologous genes for rhizobium LCO perception supports the hypothesis of a shared evolutionary origin of LCO-driven nodulation. In contrast to legumes, symbiotic LysM-type receptors in *Parasponia* did not experience recent duplication events. Instead, the *Parasponia* symbiotic LysM-type LCO receptors evolved following two ancient duplications. We hypothesize that the *PanNFP1 - PanNFP2* duplication associates with the origin of the nitrogen-fixation clade, whereas in case of *PanLYK1* and *PanLYK3,* the duplication occurred prior to the birth of the nitrogen-fixing clade. This makes it most probable that the capability of these receptors to perceive LCOs predates the evolution of the nitrogen-fixing nodulation trait.

Currently, the *NFP1-NFP2* duplication cannot be precisely dated because legumes do not possess an NFP-I-type gene. This can be explained in two scenarios. (i) The *NFP1-NFP2* duplication occurred in the root of the nitrogen-fixing clade, and subsequently, the NFP-I-type gene got lost in the Fabales lineage. (ii) The *NFP1-NFP2* duplication occurred in an ancestor of the Fagales-Cucurbitales-Rosales lineages after the divergence of the Fabales order. The recent finding that ectopic expression of the *NFP*-type gene of two species outside of the nitrogen fixing clade (*Petunia hybrida PhLYK10* and tomato *SlLYK10*) can -at least partially- *trans* complement the medicago and lotus *Mtnfp* and *Ljnfr5* mutant phenotypes demonstrates that LCO receptor functionality is ancestral to the *NFP1-NFP2* duplication (Girardin et al., 2019). The putative promoters of *PhLYK10* and *SlLYK10* show a nodule-enhanced expression profile similar to that reported for *PanNFP2 (Girardin et al., 2019)*, which may support the second scenario, where the duplication of *NFP1-NFP2* has occurred only after the divergence of the Fabales clade. However, for such a scenario, it is essential that Fabales represents the most basal lineage in the nitrogen-fixing clade. To date, this remains unknown. For example, a recent phylogenetic study suggests, although with limited statistical support, that Fabales is sister to Fagales (Koenen et al., 2019). The phylogenetic analysis presented here (Figure 3) suggests that the first scenario is most probable (aSH-aLRT / UF-Bootstrap / approximation with Mr.Bayes support 76.4/77/0.859). Additionally, we searched for amino acid motifs in NFP-I and NFP-II type proteins and found an indel region in legume and non-legume

NFP-II type proteins that is distinct from NFP-I **(Figure S9)**. This also supports the hypothesis that NFP1-NFP2 duplicated at the root of the nitrogen-fixing clade. However, additional experiments are needed to definitively reject either scenario.

Trans complementation studies in a lotus *Ljnfr1;Ljnfr5* double mutant showed that parasponiaLCO receptors can only partially restore LCO signalling. This only partial complementation we did not anticipate, because of the shared microsymbiont *M. loti* that can nodulate parasponiaas well as lotus. One explanation for this limited functionality may be that such receptors function in larger multiprotein membrane domain complexes. In such a case, the parasponiaLCO receptors are not adapted to interact with associated lotus proteins. Additionally, legumes and *Paraspona* have diverged in the mode of rhizobium infection. Whereas rhizobium penetrates *Parasponia* roots apoplastically by crack entry, legumes are generally infected intracellularly via curled root hair cells. Phenotypic analysis of rhizobium infection in legumes suggests that a specific LCO receptor is involved in this process, the so-called entry receptor (Ardourel et al., 1994). Such entry receptors have not yet been fully characterized, but *MtLYK3* may carry out such functions, a s they control rhizobium infection (Limpens et al., 2003; Smit et al., 2007). It remains elusive whether such entry receptor functioning requires specific adaptations that did not occur in the *Parasponia LYK3* ortholog.

We showed that an engineered *T. levigata TleNFP2* receptor can functionally complement the parasponi*a Pannfp2* mutant, whereas *PanNFP1* cannot. This suggests that the NFP1 and NFP2 receptor proteins have functionally diverged. Based on the finding that NFP orthologous protein of *P. hybrida* (PhLYK10) and tomato (SlLYK10) can complement lotus *Ljnfr5* and medicago *Mtnfp* mutants, it can be hypothesized that in parasponia especially *PanNFP1* has experienced protein adaptations. However, it should be noted that the *trans* complementation studies presented here were conducted using the native *PanNFP2* promoter, whereas studies conducted with *PhLYK10* and *SlLYK10* were conducted with *CAMV35S* (Girardin et al., 2019). Such overexpression may mask differences in substrate affinity and/or specificity, under which native transcriptional regulation are biologically relevant. Our data demonstrate that the ancestor of *T. levigata* possessed a NFP2 receptor that can function in nodulation.

Mutant analysis in legumes demonstrated that rhizobium nodulation co-opted elements of an AM signalling pathway, including the LRR-type transmembrane receptor kinase lotus SYMBIOTIC RECEPTOR KINASE (LjSYMRK)/medicago DOES NOT MAKE INFECTIONS2 (MtDMI2), the nuclear envelope located cation ion channels LjCASTOR, LjPOLLUX/MtDMI1, the nuclear localized CALCIUM CALMODULIN KINASE LjCCaMK/MtDMI3, and the transcription factor LjCYCLOPS/

medicago INTERACTING PROTEIN OF DMI3 (MtIPD3) (Geurts et al., 2012). However, in legumes, rhizobium and AM fungi were shown to have independent perception mechanisms to activate this common symbiosis signalling pathway. In lotus and medicago, these consist of LjNFR1-LjNFR5 / MtLYK3-MtNFP for rhizobium LCOs and MtLYK9/MtCERK1 for AM signals (Geurts et al., 2012; Feng et al., 2019; Gibelin-Viala et al., 2019). MtLYK3 and MtLYK9/MtCERK1 both belong to the LYK-Ib clade and evolved upon legume specific duplication events **(Figure 1)** (De Mita et al., 2014). The strong phenotype in AM and nodule symbioses of the parasponia*Panlyk1;Panlyk3* knockout mutant demonstrates that such subfunctionalization is not causal for the evolution of rhizobium LCO receptors. In parasponia, both receptors function in conjunction to control AM and rhizobium nodulation. Additionally, PanLYK3 acts as a chitin innate immune receptor. Such multifunctionality has also been reported for MtLYK9/MtCERK1 in medicago and OsCERK1 in rice, which function both in AM symbiosis and chitin innate immune signalling (Miyata et al., 2014; Carotenuto et al., 2017; Feng et al., 2019; Gibelin-Viala et al., 2019). As monocots did not experience the LYK-Ia / LYK-Ib duplication, it demonstrates that committing multiple functions in symbioses and innate-immunity was ancestral to species of the nitrogen fixing clade but functionally diverted in the legume lineage.

The presence of NFP-type genes (LYR-IA orthogroup) in species outside of the nitrogen-fixing clade associates with the ability to establish an AM symbiosis **(Figure 3)** (Delaux et al., 2014; Gough et al., 2018). However, corresponding mutants have only a relatively weak phenotype in AM symbiosis (Buendia et al., 2016; Miyata et al., 2016; Girardin et al., 2019). Upon duplication of this gene, the NFP-I and NFP-II subclades may have inherited the ancestral function. As both the parasponia PanNFP1 and PanNFP2 receptors can partially complement LCO-induced Ca^{2+} oscillation in the lotus *Ljnfr1;Ljnfr5* double mutant, it supports that receptors of the NFP-I and NFP-II clades can act as an LCO receptor, which may reflect the ancestral function. Our observation that the presence of a functional gene in the NFP-II clade strictly associates with LCO-based nodulation suggests that this gene was co-opted to function in this trait. The importance of this LysM-type LCO receptor in the nitrogen-fixing nodulation trait is underlined by the complete block of nodulation in knockout mutants in legumes (e.g. lotus *Ljnfr5*, medicago *Mtnfp* and pea *Pssym10*) and *Parasponia* (parasponia *Pannfp2*) (Madsen et al., 2003; Arrighi et al., 2006). As *Parasponia* and legumes diverged at the root of the nitrogen-fixing clade, it suggests that the adaptations in the NFP-II clade are ancient and may have coincided with the birth of the nodulation trait.

The NFP-I type gene retained -at least in part- its ancestral function, indicated by its presence in non-nodulating species in the nitrogen-fixing clade that can establish an AM symbiosis. In cases where AM symbiosis is replaced by an ectomycorrhizal symbiosis such as in *Fagus sylvatica* or *Castanea mollissima,* the NFP-I type gene pseudogenized. However, phenotypic studies in stable parasponia mutants could not support the functioning of *PanNFP1* in AM symbiosis. These findings contradict our earlier observation that this gene functions in arbuscule formation (Op den Camp et al., 2011). The reason for this discrepancy may be due to the RNAi construct used, which may have off-target effects (van Velzen et al., 2018b). To determine whether this is the case, we have studied the expression of LysM-type RLK genes in two independent *PanNFP1* RNAi experiments. This revealed significant knockdown not only of *PanNFP1*, but also *PanNFP2*, which can explain the strong rhizobium nodulation and infection phenotype as reported by Op den Camp et al. (2011). We also found variable expression levels of other LysM-RLKs, including *PanLYK1* and *PanLYK3*, which may explain the reported mycorrhization phenotype on *PanNFP1* RNAi roots **(Figure S10)**. Studies presented here using CRISPR-Cas9 knockout mutant lines revealed substantial biological variation in mycorrhization efficiency of parasponia roots, which may have hindered the observation of minor quantitative AM symbiosis phenotypes. To rule out that PanNFP1 and PanNFP2 may function redundantly to control AM symbiosis, we analysed a *Pannfp1;Pannfp2* double mutant. Also, these lines showed to be effectively mycorrhized. Therefore, we conclude that our current mutant phenotype analysis does not find support for essential functioning of parasponia PanNFP1 and PanNFP2 in AM symbiosis by.

The study presented here provided insight into the evolutionary trajectory of symbiotic LCO LysM-type receptors. By using parasponia as a comparative system to legumes, we revealed two ancestral duplications of LysM-type LCO receptors that predate, and coincide with, the evolution of nitrogen-fixing nodules. The strict association of genes in the NFP-II clade with LCO-driven nodulation strongly suggests that this gene was co-opted to function specifically in this symbiosis, making NFP2s a target in approaches to engineer LCO-driven nodulation in non-leguminous plants.

MATERIAL & METHODS

LysM-type receptor phylogeny reconstructions

Orthogroups containing LysM-type receptor kinases of parasponia, generated in a previous study (van Velzen et al., 2018b), were combined and re-aligned into a single alignment using MafftV7.017. MrBayes3.2.6 was used to calculate phylogenetic relations under default parameters in Geneious R8.1.9 (Biomatters Ltd, UK) (Huelsenbeck and Bollback, 2001). Clades were named as published previously (Huelsenbeck and Bollback, 2001; Buendia et al., 2018). For clade LYK-I

and LYR-IA additional putative orthologs were collected from Phytozome and NCBI databases using BLAST with AtCERK and MtNFP protein sequences as query **(Table S1)**. Available Genomes from Fabales, Fagales, Cucurbitales and Rosales species were downloaded and local BLAST analysis was conducted using Geneious R8.1.9 (Biomatters Ltd, UK) to search for additional unannotated LYK-I and LYR-IA protein sequences. Pseudogenes were annotated manually based on the closest functional ortholog so that a protein sequence could be deduced. Correct protein sequences were aligned using MAFFTV7.017 and subsequently manually curated. The deduced amino acid sequence was subsequently added to the alignment if the alignment length was at least 70% of the parasponia protein. Phylogenetic analysis was performed using IQ-tree (Nguyen et al., 2015; Trifinopoulos et al., 2016), running the modelfinder extension to find the best substitution models (Kalyaanamoorthy et al., 2017). Branch support analysis was done using Sh.aLRT 1000 replicates, UF-BOOTSTRAP support 1000 iterations (Kalyaanamoorthy et al., 2017; Hoang et al., 2018) and approximate Bayes support. Branch supports shown are UF-Bootstrap support%. Best fit model for the LYK-I clade: JTT+I+G4. Best fit model for LYR-IA clade: JTT+I+G4. Resulting tree files were loaded into Interactive Tree Of Life (iTOL) v3 for editing (Letunic and Bork, 2016). The analysis was run at least three times. Trees were rooted to outgroup angiosperm species *Amborella trichopoda*. UF Bootstrap Branch supports >98 were omitted for visual clarity. Gene names, accession numbers and alignment file of identified homologs can be found in Data set S1 for LYK-I and Data set S2 for LYR-IA, and Table S1 for parasponia.

LYK3 alignment and variant detection

Genomic LYK3 regions of parasponia*, Parasponia rigida, Parasponia rugosa, Trema orientalis* RG16*, T. orientalis* RG33, and *Trema levigata* were extracted from the respective assemblies (van Velzen et al., 2018b) and Aligned using MAFFTV7.0.17 implemented in Geneious R8.1. Coding sequences of parasponia*, P.rigida, P.rugosa LYK3* protein variants were translated and aligned using MAFFTV7.0.17 implemented in Geneious R8.1 **(Data set S1)**.

Vector constructs

All vectors generated for this study were created using golden gate cloning (Engler et al., 2009). Backbones and binary vectors were derived from the golden gate molecular toolbox (Engler et al., 2014). Parasponia LysM-type receptor cDNA clones were sequence synthesized as level 0 modules, including silent mutations in golden gate BsaI or BpiI restriction sites. Golden gate compatible clones of LjNFR1 and LjNFR5 promoters, CDS and terminators were obtained from Arhus University, Denmark. The calcium signalling reporter pLjUBQ1:R-GECO1.2 was published previously (Kelner et al., 2018). The generation and assembly of parasponia CRISPR constructs were done as published previously (van Zeijl et al., 2018). For hairy root

transformation, a modified level 2 standard vector carrying spectinomycin instead of kanamycin resistance was created. All sgRNAs were expressed using the AtU6 promoter. All Golden Gate binary vectors were verified by restriction digestion and DNA sequencing before transformation. A list of primers and constructs can be found in **Table S3 and S4.**

Genotyping and off-target analysis

All sgRNA targets were designed using the Geneious R10 CRISPR design tool, which picks targets on the principles described in Doench et al. (2014). To be selected Guide RNAs must have no potential target sites in the genome with (i) Less than three mismatches or (ii) less than two indels. Known off-target locations in CDS regions were PCR amplified and sequenced. No off-target mutations at these sites were detected. Genotypes and known off-target locations of CRISPR mutants used in this study can be found in **Data set S3.** Primers used for the creation of sgRNAs and subsequent sequencing of mutants and off-targets are listed in **Table S4**.

Bacterial strains

We used *Mesorhizobium plurifarium* BOR2 (van Velzen et al., 2018b) and *Sinorhizobium fredii* NGR234.pHC60 expressing GFP (Trinick and Galbraith, 1980; Cheng and Walker, 1998; Op den Camp et al., 2011) for parasponia inoculation experiments. *M.loti* R7A.pHC60 (Cheng and Walker, 1998; Sullivan et al., 2002) was used for lotus inoculations. *M. loti* R7A and *Rhizobium tropici* CIAT899 (Martínez-Romero et al., 1991) containing plasmid pMP604 (Spaink et al., 1989) were used for LCO extraction. *A. rhizogenes* strain AR10 (Hansen et al., 1989b; Martínez-Romero et al., 1991) was used for lotus root transformation. *Agrobacterium tumefaciens* strain AGL-1 (Lazo et al., 1991) was used in parasponia transformation. *Agrobacterium* sp. MSU440 was used for parasponia hairy root transformations (Cao et al., 2012). The *Escherichia coli* strain DH5α was used to propagate plasmids and in all subsequent cloning steps.

Rhizobium LCO isolation

To isolate rhizobium LCOs the plasmid pMP604 containing an auto-active NodD protein was introduced in *M. loti* R7A and *R. tropici* CIAT899 (Spaink et al., 1989; López-Lara et al., 1995). LCOs were extracted from a 750 ml liquid culture, OD600=0.5, grown at 28°C in minimal medium (0.5.75mM% K_2HPO_4, 0.%7.35mM KH_2PO_4, 0%5.9mM KNO_3, 460 nM $CaCl_2$, 37.5μM $FeCl_3$, 2.07mM $MgSO_4$, 20.5nM biotin, 2.9nM Thiamine HCl , 8.1nM Nicotinic acid, 4.8nM Pyridoxine-HCl, 2.8nM Myo-inositol, 4.6nM Panthotenate and 1% w/v sucrose) by the addition of 150 mL 1-butanol and 1h shaking. The butanol phase was transferred and subsequently evaporated (water bath 40˚C). Pellet was dissolved in 75ml methanol, tested for

Nod-factor activity and stored at -20˚C for later use. The concentration of active LCOs was estimated by using *LjNIN* induction in lotus wild type Gifu roots, 3h post-application. The lowest active dilution was estimated to be ~10^{-10} M.

Lotus japonicus Agrobacterium rhizogenes root transformation

Lotus *Ljnfr1-1;Ljnfr5-2* double mutants (Madsen et al., 2003; Radutoiu et al., 2003) were used for LysM complementation assays and 'Gifu' wild-type as control. Seedlings for *A. rhizogenes* root transformation were moved to fresh half-strength B5 medium and co-cultivated for 1 weeks as described previously using *A. rhizogenes* strain AR10 (Stougaard et al., 1987; Hansen et al., 1989a; Stougaard, 1995). During root emergence plants were grown on 1% agar plates half-strength B5 media containing 0.03% w/v cefotaxime and 1% w/v sucrose. Plants were screened for transformed roots using nuclear-localized R.GECO1.2 fluorescence. Shoots with transformed roots were grown in Agroperlite (Maasmond-Westland, Netherlands) supplemented with modified ½ Hoagland's medium (Hoagland et al., 1950) containing 0.56 mM NH_4NO_3 and inoculated with *M. loti* R7A.pHC60 (expressing GFP) at OD600= 0.05. Plants were grown at 21°C under a 16h light/8h dark regime. For calcium oscillation analysis transformed plants were grown on ½ Hoagland's plates with 1% agar containing 0.56 mM NH_4NO_3 for 1 week. Plants were moved to N-free ½ hoaglands medium 1 week prior to imaging.

Calcium oscillation quantification

Calcium spiking experiments were performed on a Leica TCS SP8 HyD confocal microscope equipped with a water lens HC plan-Apochromat CS2 40x/1.0. Transformed root segments expressing R-GECO1.2 were selected and incubated with 500x diluted LCO extract (estimated to represent ~10^{-9} M) in nitrate-free ½ Hoagland's medium (Hoagland et al., 1950) on a glass slide with coverslip. Images were taken at 5s intervals for a minimum of 20 minutes per sample using an excitation wavelength of 552 nm and emission spectrum 585-620 nm. It is possible to monitor a large number of nuclei per root sample. However, only epidermal and especially root hairs showed to be responsive. Therefore, total nuclei numbers vary largely between samples. Video recordings of Imaged root samples were exported to ImageJ1.50i (Collins, 2007). The Geciquant ImageJ plugin was used for background subtraction and region of interest (ROI) selection (Srinivasan et al., 2015). Average pixel intensity of ROIs (individual nuclei) were measured. Average pixel values (0-255) per nucleus were plotted and a background R-GECO1.2 fluorescence baseline of 2x 1 minute (2 regions of 12 frames) was selected manually in a region of the trace where no spikes were occurring. Only nuclei with a minimum of three spikes with an amplitude of over 1.5 times background were considered as positive.

Parasponia growth conditions for propagation, transformation, mycorrhization and nodulation

Sequenced parasponia WU1 trees or their direct descendants, were used in all experiments (Op den Camp et al., 2011; van Velzen et al., 2018b). Prior to transformation or transfer to tissue culture parasponia trees are grown in a conditioned greenhouse at 28°C, 85% humidity and a 16/8 h day/night regime. Parasponia *in vitro* propagation, transformation, CRISPR-Cas9 mutagenesis and nodulation assays were done according to Van Zeijl et al 2018 (van Zeijl et al., 2018). Parasponia hairy root transformations were performed according to Cao et al 2012 (Cao et al., 2012).

Parasponia Nodulation assay and analysis

Rooted tissue culture plantlets for phenotyping assays were grown in crystal-clear polypropelene containers (1 L), with a gas exchange filter (OS140BOX, Duchefa Biochemie, Netherlands). Pots were half-filled with agraperlite (Maasmond-Westland, Netherlands) and watered with modified EKM medium [3 mM MES (C6H13NO4) pH 6.6, 2.08 mM MgSO4, 0.88 mM KH2PO4, 2.07 mM K2HPO4, 1.45 mM CaCl2, 0.70 mM Na2SO4, 0.375 mM NH4NO3, 15 µM Fe-citrate, 6.6 µM MnSO4, 1.5 µM ZnSO4, 1.6 µM CuSO4, 4 µM H3BO3, 4.1 µM Na2MoO4] (Becking, 1983)]. For nodulation assays, modified EKM medium (Becking, 1983) was inoculated with rhizobia (OD600 = 0.025) prior to planting the shoots. For inoculation with strain *S. fredii* NGR234.pHC60, containers were half-filled with sterilized river sand and watered with modified EKM-medium containing the bacteria at an OD600 = 0.05. All Nodules were fixed in buffer containing 4% w/v paraformaldehyde mixed with 3% v/v glutaraldehyde in 50 mM phosphate (pH = 7.4). A vacuum was applied for 2 hours during a total 48h incubation. Fixed nodules were embedded in Plastic, Technovit 7100 (Heraeus-Kulzer, Germany), according to manufacturer's recommendations. Sections (5 µm) were made using a RJ2035 microtome (Leica Microsystems). Sections were stained using 0.05% w/v Toluidine Blue O. Images were taken with a DM5500B microscope equipped with a DFC425c camera (Leica microsystems).

Parasponia mycorrhization assay

For mycorrhization experiments, pots were half-filled with sterilized river sand, watered with modified ½ strength Hoagland's medium containing 20 μM potassium phosphate. Pots were inoculated with 250 spores of *Rhizopagus irregularis* (Agronutrion-DAOM197198). In all experiments, plantlets in pots with closed lids were placed in a climate room at 28°C, 16/8 h day/night. Plants were watered with sterilized, demineralized water. Plants were harvested 6 weeks post inoculation with *Rhizopagus irregularis* (Agronutrion-DAOM197198). Root segments were treated with 10% (w/v) KOH and incubated at 90°C for 20 minutes. The root samples were then rinsed 6 times with water and stained with trypan blue at 90°C for 5 minutes. For

each mutant, ten plants were assessed and from each plant 30 root segments (each segment of approx. 1 cm long) were examined and mycorrhizal structures (hyphae, vesicles and arbuscules) were determined using the magnified line intersect method (Trouvelot, A, Kough J L, Gianinazzi-Pearson V, 1986) using a Leica CTR6000 microscope. For staining with WGA- Alexafluor 488 (Molecular Probes, Thermo Fisher Scientific, Waltham, MA, USA), roots were incubated in 10% (w/v) KOH at 60°C for 3 h. Then roots were washed three times in phosphate-buffered saline (PBS) (150 mM NaCl, 10 mM Na_2HPO_4 and 1.8 mM KH_2PO_4, pH 7.4), and incubated in 0.2 µg ml^{-1} WGA-Alexafluor 488 in PBS at room temperature for 16 h. For RNA-isolation, parasponia WT plants were grown according to conditions above. RNA was isolated according to protocols published in (Op den Camp et al., 2011; van Velzen et al., 2018b). Mock inoculated plants were harvested as control. Three independent biological replicates were taken per sample. Expression was determined using RNA-seq. Reads were mapped using kallisto (Bray et al., 2016). Expression values and differential expression were determined using sleuth (Pimentel et al., 2017). Differentially expressed genes (Benjamini-Hochberg multiple testing corrected q-value <= 0.05)

qPCR analysis of panNFPi cDNA samples.

PanNFPi cDNA samples were generated previously (Op den Camp et al., 2011). qPCR was performed in 10 µl reactions using 2x iQ SYBR Green Super-mix (Bio-Rad, United States). PCR reaction was executed on a CFX Connect optical cycler, according to the manufacturer's protocol (Bio-Rad, United States). Three technical replicates per cDNA sample were used. Data analysis and statistical analysis of biological replicates was performed using CFX Manager 3.0 software (Bio-Rad, United States). Gene expression was normalized against Reference genes *PanACTIN* and *PanEF1alpha.* Primers can be found in **Table S4.**

ROS assay

Parasponia plantlets were grown on rooting medium (van Zeijl et al., 2018) for 4 weeks at 28°C before the treatment. Roots, submerged in water, were cut into approximately 1cm pieces. Each well of a black 96 well flat bottom polystyrene plate (Nunc) was filled with 10 root pieces. 10 replicates per line were analysed. After filling the wells, the plate was kept 5 hours in 28 °C. After incubation, the water was replaced with 100 μl of assay solution containing 0.5 μM L-012 (FUJIFILM Wako Chemicals), 10 μg/ml Horseradish peroxidase (Sigma), and respective elicitors (; CO7 (ELICITYL) or LCOs extracted from *M. loti* or *R. tropici*) at described concentrations. As a mock treatment, 100 µl H_2O was added. The light emission was immediately measured at 30 second intervals for 30 minutes, using a Clariostar multi well-plate reader. All data are the average of at least three independent biological replicates.

Protein extraction from Parasponia and western blotting

Parasponia plantlets were grown on rooting medium (van Zeijl et al., 2018) for 4 weeks at 28°C before the treatment. About 200mg of roots were cut while submerged in water and collected in a PCR-tube. Root segments were incubated for 5 hours at 28 °C before treatment. Root pieces were treated with water containing 100 µM CO7 (ELICITYL) for 10 min. After incubation, roots were immediately frozen in liquid nitrogen. Samples were homogenized using metal beads. Total root protein was extracted in a buffer containing 50 mM Tris·HCl (pH 7.5), 150 mM KCl, 1 mM EDTA (pH 7.5), 0.1% w/v Triton X-100, 1 mM DTT, complete protease inhibitors (Roche), and phosstop (Roche). Amounts of extracted protein were measured with Qubit (Thermo fisher Scientific) and equal amounts of protein ~20 µg were electrophoresed by Mini-PROTEAN TGX stain free gels (BIORAD). A Trans-Blot Turbo Transfer system was used for blotting. To visualize phosphorylated MPK3/MPK6, the antibody for anti–phospho-p44/42 MAPK was used (no. 4370; Cell Signalling Technology). Anti-rabbit antibody (no. 7054; Cell Signalling Technology) were used as secondary antibody. Equal loading was confirmed by CBB staining.

Quantification and statistical analysis

Nodule number was quantified as Mean nodule number ± SE for all experiments. Replicate number is denoted in figure or figure legend. Additionally, all individual data points were plotted for graphical visualization of variation. Graphs and statistical analysis were performed using R studio 1.1.456 for nodulation experiments. Statistical tests on nodule numbers was done using One Way Analysis of variance (ANOVA) and a Tukey post-hoc test for multiple comparisons. Statistical significance was defined as a $p<0.05$. Levenes test for homogeneity of variance was used prior to running a one-way ANOVA. In cases where normality assumption was violated, alternative tests such as Mann–Whitney–Wilcoxon (MWW) were used as denoted in the figure legends. For the mycorrhization experiment a standard linear model was used to estimate the difference, and the corresponding least significant differences (LSD), of the knockout mutants with the wild type control. The LSD with respect to the control was Bonferroni adjusted to correct for multiple testing.

Accession numbers

Sequence data from this article can be found in the GenBank/EMBL data libraries under accession numbers as mentioned in Table S1.

Funding information

This work was supported by an NWO-VICI grant (865.13.001) to RG, the ENSA project funded by the Bill & Melinda Gates Foundation to the University of Cambridge to RG, NWO-VENI grant (863.15.010) to WK, a JSPS Overseas Research Fellowship to KM and The Ministry of Research, Technology and Higher Education of the Republic of Indonesia (RISET-PRO grant 8245-ID) to YPR.

Supplemental Data.

The supplemental data belonging to this chapter are available online at Plant Physiology: https://doi.org/10.1104/pp.19.0142

Supplemental Figure S1. Phylogeny reconstruction of orthogroups representing LysM-type receptors.

Supplemental Figure S2. Duplication of the *LYK3* first exon is conserved among *Parasponia* and *Trema* subspecies.

Supplemental Figure S3. Complementation of a lotus *Ljnfr1;Ljnfr5* mutant for LCO-induced calcium oscillation.

Supplemental Figure S4. Nodulation is affected in parasponia *Pannfp1*, *Pannfp2* and *Panlyk3* CRISPR-Cas9 mutants.

Supplemental Figure S5. Complementation of parasponia *Panlyk1;Panlyk3* double mutant.

Supplemental Figure S6. CO7 triggered ROS production and MPK phosphorylation in parasponia mutant lines.

Supplemental Figure S7. Expression of parasponia LYSM-type receptors during mycorrhization.

Supplemental Figure S8. Parasponia LysM-type receptor mutants can establish arbuscular mycorrhizal symbiosis.

Supplemental Figure S9. Conserved indel in NFP-II type receptor proteins.

Supplemental Figure S10. The PanNFPi RNAi construct has off target activity on *PanNFP2* and other LysM-type receptor kinases.

Supplemental Table S1. *Parasponia andersonii* LysM-type receptors.

Supplemental Table S2. Trans-complementation of lotus *Ljnfr1;Ljnfr5* for nodulation**.**

Supplemental Table S3. Constructs generated in this study.

Supplemental Table S4. Primers used in this study.

Supplemental Data Set S1. Sequence alignment of LYK-I type receptors in fasta format.

Supplemental Data Set S2. Sequence alignment of LYR-Ia type receptors in fasta format.

Supplemental Data Set S3. Genotyping and off target analysis of parasponia CRISPR-Cas9 mutants generated in this study.

Supplemental Movie S1. Calcium spiking in root hairs of lotus *Ljnfr1-1;Ljnfr5-2* double mutant complemented with *LjNFR1;LjNFR5*.

Supplemental Movie S2. Calcium spiking in root hairs of lotus *Ljnfr1-1;Ljnfr5-2* double mutant trans-complemented with *PanLYK3.1;LjNFR5.*

Acknowlegdements

We thank Jens Stougaard and Simona Radutoiu (Arhus University, Denmark) for sharing the *Lotus japonicus Ljnfr1-1;Ljnfr5-2* double mutant and the *LjNFR1* and *LjNFR5* constructs, Maryam Charpentier (John Innes Centre, United Kingdom) for sharing R-GECO1.2, Jinling Li and Bart Thomma (Wageningen University, The Netherlands) for the use of Clariostar multi-well reader and Joost van Heerwaarden (Wageningen University, The Netherlands) for statistical support. Furthermore, we would like to thank Arianne van Lierop, Jens van Kempen, Stan van der Wal and Vince Brouwer who contributed as undergraduate students to this research.

References

Ardourel M, Demont N, Debellé F, Maillet F, de Billy F, Promé JC, Dénarié J, Truchet G (1994) Rhizobium meliloti lipooligosaccharide nodulation factors: different structural requirements for bacterial entry into target root hair cells and induction of plant symbiotic developmental responses. Plant Cell 6: 1357–1374

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

Becking JH (1983) The Parasponia parviflora—Rhizobium symbiosis. Host specificity, growth and nitrogen fixation under various conditions. Plant Soil 75: 309–342

Bozsoki Z, Cheng J, Feng F, Gysel K, Vinther M, Andersen KR, Oldroyd G, Blaise M, Radutoiu S, Stougaard J (2017) Receptor-mediated chitin perception in legume roots is functionally separable from Nod factor perception. Proc Natl Acad Sci U S A 114: E8118–E8127

Bray NL, Pimentel H, Melsted P, Pachter L (2016) Erratum: Near-optimal probabilistic RNA-seq quantification. Nat Biotechnol 34: 888

Broghammer A, Krusell L, Blaise M, Sauer J, Sullivan JT, Maolanon N, Vinther M, Lorentzen A, Madsen EB, Jensen KJ, et al (2012) Legume receptors perceive the rhizobial lipochitin oligosaccharide signal molecules by direct binding. Proc Natl Acad Sci U S A 109: 13859–13864

Buendia L, Girardin A, Wang T, Cottret L, Lefebvre B (2018) LysM Receptor-Like Kinase and LysM Receptor-Like Protein Families: An Update on Phylogeny and Functional Characterization. Front Plant Sci 9: 1531

Buendia L, Wang T, Girardin A, Lefebvre B (2016) The LysM receptor-like kinase SlLYK10 regulates the arbuscular mycorrhizal symbiosis in tomato. New Phytol 210: 184–195

Burrill TJ, Hansen R (1917) Is symbiosis possible between legume bacteria and non-legume plants? Bulletin (University of Illinois (Urbana-Champaign campus). Agricultural Experiment Station); no. 202

Cao Q, den Camp RO, Kalhor MS, Bisseling T, Geurts R (2012) Efficiency of Agrobacterium rhizogenes–mediated root transformation of Parasponia and Trema is temperature dependent. Plant Growth Regul 68: 459–465

Carotenuto G, Chabaud M, Miyata K, Capozzi M, Takeda N, Kaku H, Shibuya N, Nakagawa T, Barker DG, Genre A (2017) The rice LysM receptor-like kinase OsCERK1 is required for the perception of short-chain chitin oligomers in arbuscular mycorrhizal signaling. New Phytol 214: 1440–1446

Cheng HP, Walker GC (1998) Succinoglycan is required for initiation and elongation of infection threads during nodulation of alfalfa by Rhizobium meliloti. J Bacteriol 180: 5183–5191

Collins TJ (2007) Introduction to ImageJ for Light Microscopy. Microsc Microanal 13: 1674–1675

Delaux P-M, Varala K, Edger PP, Coruzzi GM, Pires JC, Ané J-M (2014) Comparative phylogenomics uncovers the impact of symbiotic associations on host genome evolution. PLoS Genet 10: e1004487

De Mita S, Streng A, Bisseling T, Geurts R (2014) Evolution of a symbiotic receptor through gene duplications in the legume--rhizobium mutualism. New Phytol 201: 961–972

Doench JG, Hartenian E, Graham DB, Tothova Z, Hegde M, Smith I, Sullender M, Ebert BL, Xavier RJ, Root DE (2014) Rational design of highly active sgRNAs for CRISPR-Cas9-mediated gene inactivation. Nat Biotechnol 32: 1262–1267

Engler C, Gruetzner R, Kandzia R, 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 JDG, Patron NJ, Marillonnet S (2014) A golden gate modular cloning toolbox for plants. ACS Synth Biol 3: 839–843

Feng F, Sun J, Radhakrishnan GV, Lee T, Bozsóki Z, Fort S, Gavrin A, Gysel K, Thygesen MB, Andersen KR, et al (2019) A combination of chitooligosaccharide and lipochitooligosaccharide recognition promotes arbuscular mycorrhizal associations in Medicago truncatula. Nat Commun 10: 5047

Folch-Mallol JL, Marroqui S, Sousa C, Manyani H, López-Lara IM, van der Drift KM, Haverkamp J, Quinto C, Gil-Serrano A, Thomas-Oates J, et al (1996) Characterization of Rhizobium tropici CIAT899 nodulation factors: the role of nodH and nodPQ genes in their sulfation. Mol Plant Microbe Interact 9: 151–163

Genre A, Chabaud M, Balzergue C, Puech-Pagès V, Novero M, Rey T, Fournier J, Rochange S, Bécard G, Bonfante P, et al (2013) Short-chain chitin oligomers from arbuscular mycorrhizal fungi trigger nuclear Ca2+ spiking in Medicago truncatula roots and their production is enhanced by strigolactone. New Phytol 198: 190–202

Geurts R, Lillo A, Bisseling T (2012) Exploiting an ancient signalling machinery to enjoy a nitrogen fixing symbiosis. Curr Opin Plant Biol 15: 438–443

Gibelin-Viala C, Amblard E, Puech-Pages V, Bonhomme M, Garcia M, Bascaules-Bedin A, Fliegmann J, Wen J, Mysore KS, le Signor C, et al (2019) The Medicago truncatula LysM receptor-like kinase LYK9 plays a dual role in immunity and the arbuscular mycorrhizal symbiosis. New Phytol. doi: 10.1111/nph.15891

Girardin A, Wang T, Ding Y, Keller J, Buendia L, Gaston M, Ribeyre C, Gasciolli V, Auriac M-C, Vernié T, et al (2019) LCO Receptors Involved in Arbuscular Mycorrhiza Are Functional for Rhizobia Perception in Legumes. Curr Biol 29: 4249–4259.e5

Gonçalves DJP, Simpson BB, Ortiz EM, Shimizu GH, Jansen RK (2019) Incongruence between gene trees and species trees and phylogenetic signal variation in plastid genes. Mol Phylogenet Evol 138: 219–232

Gough C, Cottret L, Lefebvre B, Bono J-J (2018) Evolutionary History of Plant LysM Receptor Proteins Related to Root Endosymbiosis. Front Plant Sci 9: 923

Gough C, Cullimore J (2011) Lipo-chitooligosaccharide signaling in endosymbiotic plant-microbe interactions. Mol Plant Microbe Interact 24: 867–878

Griesmann M, Chang Y, Liu X, Song Y, Haberer G, Crook MB, Billault-Penneteau B, Lauressergues D, Keller J, Imanishi L, et al (2018) Phylogenomics reveals multiple losses of nitrogen-fixing root nodule symbiosis. Science 361: eaat1743

Hansen J, Jorgensen J-E, Stougaard J, Marcker KA (1989a) Hairy roots - a short cut to transgenic root nodules. Plant Cell Rep 8: 12–15

He J, Zhang C, Dai H, Liu H, Zhang X, Yang J, Chen X, Zhu Y, Wang D, Qi X, et al (2019) A LysM Receptor Heteromer Mediates Perception of Arbuscular Mycorrhizal Symbiotic Signal in Rice. Mol Plant 12: 1561–1576

Hoagland DR, Arnon DI, Others (1950) The water-culture method for growing plants without soil. Circular. California Agricultural Experiment Station 347:

Hoang DT, Chernomor O, von Haeseler A, Minh BQ, Vinh LS (2018) UFBoot2: Improving the Ultrafast Bootstrap Approximation. Mol Biol Evol 35: 518–522

Huelsenbeck JP, Bollback JP (2001) Empirical and hierarchical Bayesian estimation of ancestral states. Syst Biol 50: 351–366

Huisman R, Geurts R (2019) A roadmap towards engineered nitrogen-fixing nodule symbiosis. Plant Communications 100019

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

Kelner A, Leitão N, Chabaud M, Charpentier M, de Carvalho-Niebel F (2018) Dual Color Sensors for Simultaneous Analysis of Calcium Signal Dynamics in the Nuclear and Cytoplasmic Compartments of Plant Cells. Front Plant Sci 9: 245

Koenen EJM, Ojeda DI, Steeves R, Migliore J, Bakker FT, Wieringa JJ, Kidner C, Hardy OJ, Pennington RT, Bruneau A, et al (2019) Large-scale genomic sequence data resolve the deepest divergences in the legume phylogeny and support a near-simultaneous evolutionary origin of all six subfamilies. New Phytol.

Lancelle SA, Torrey JG (1984) Early development of Rhizobium-induced root nodules of Parasponia rigida. I. Infection and early nodule initiation. Protoplasma 123: 26–37

Lancelle SA, Torrey JG (1985) Early development of Rhizobium-induced root nodules of Parasponia rigida. II. Nodule morphogenesis and symbiotic development. Can J Bot 63: 25–35

Lazo GR, Stein PA, Ludwig RA (1991) A DNA transformation-competent Arabidopsis genomic library in Agrobacterium. Biotechnology 9: 963–967

Letunic I, Bork P (2016) Interactive tree of life (iTOL) v3: an online tool for the display and annotation of phylogenetic and other trees. Nucleic Acids Res 44: W242–5

Liao D, Sun X, Wang N, Song F, Liang Y (2018) Tomato LysM Receptor-Like Kinase SlLYK12 Is Involved in Arbuscular Mycorrhizal Symbiosis. Front Plant Sci 9: 1004

Li H-L, Wang W, Mortimer PE, Li R-Q, Li D-Z, Hyde KD, Xu J-C, Soltis DE, 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

Limpens E, Franken C, Smit P, Willemse J, Bisseling T, Geurts R (2003) LysM domain receptor kinases regulating rhizobial Nod factor-induced infection. Science 302: 630–633

López-Lara IM, Berg JDJ, Thomas-Oates JE, Glushka J, Lugtenberg BJJ, Spaink HP (1995) Structural identification of the iipo-chitin oligosaccharide nodulation signals of Rhizobium loti. Mol Microbiol 15: 627–638

Madsen EB, Madsen LH, Radutoiu S, Olbryt M, Rakwalska M, Szczyglowski K, Sato S, Kaneko T, Tabata S, Sandal N, et al (2003) A receptor kinase gene of the LysM type is involved in legume perception of rhizobial signals. Nature

425: 637–640

Maillet F, Poinsot 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

Martínez-Romero E, Segovia L, Mercante FM, Franco AA, Graham P, Pardo MA (1991) Rhizobium tropici, a novel species nodulating Phaseolus vulgaris L. beans and Leucaena sp. trees. Int J Syst Bacteriol 41: 417–426

Marvel DJ, Torrey JG, Ausubel FM (1987) Rhizobium symbiotic genes required for nodulation of legume and nonlegume hosts. Proc Natl Acad Sci U S A 84: 1319–1323

Miwa H, Sun J, Oldroyd GED, Downie JA (2006) Analysis of Nod-factor-induced calcium signaling in root hairs of symbiotically defective mutants of Lotus japonicus. Mol Plant Microbe Interact 19: 914–923

Miya A, Albert P, Shinya T, Desaki Y, Ichimura K, Shirasu K, Narusaka Y, Kawakami N, Kaku H, Shibuya N (2007) CERK1, a LysM receptor kinase, is essential for chitin elicitor signaling in Arabidopsis. Proc Natl Acad Sci U S A 104: 19613–19618

Miyata K, Hayafune M, Kobae Y, Kaku H, Nishizawa Y, Masuda Y, Shibuya N, Nakagawa T (2016) Evaluation of the Role of the LysM Receptor-Like Kinase, OsNFR5/OsRLK2 for AM Symbiosis in Rice. Plant Cell Physiol 57: 2283–2290

Miyata K, Kozaki T, Kouzai Y, Ozawa K, Ishii K, Asamizu E, Okabe Y, Umehara Y, Miyamoto A, Kobae Y, et al (2014) The bifunctional plant receptor, OsCERK1, regulates both chitin-triggered immunity and arbuscular mycorrhizal symbiosis in rice. Plant Cell Physiol 55: 1864–1872

Murakami E, Cheng J, Gysel K, Bozsoki Z, Kawaharada Y, Hjuler CT, Sørensen KK, Tao K, Kelly S, Venice F, et al (2018) Epidermal LysM receptor ensures robust symbiotic signalling in Lotus japonicus. Elife. doi: 10.7554/eLife.33506

Nguyen L-T, Schmidt HA, von Haeseler A, Minh BQ (2015) IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. Mol Biol Evol 32: 268–274

Nguyen TV, Wibberg D, Battenberg K, Blom J, Vanden Heuvel B, Berry AM, Kalinowski J, 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

Op den Camp R, Streng A, De Mita S, Cao Q, Polone E, Liu W, Ammiraju JSS, Kudrna D, Wing R, Untergasser A, et al (2011) LysM-type mycorrhizal receptor recruited for rhizobium symbiosis in nonlegume Parasponia. Science 331: 909–912

Parniske M (2008) Arbuscular mycorrhiza: the mother of plant root endosymbioses. Nat Rev Microbiol 6: 763–775

Pawlowski K, Demchenko KN (2012) The diversity of actinorhizal symbiosis. Protoplasma 249: 967–979

Persson T, Battenberg K, Demina IV, Vigil-Stenman T, Vanden Heuvel B, Pujic P, Facciotti MT, Wilbanks EG, 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

Pimentel H, Bray NL, Puente S, Melsted P, Pachter L (2017) Differential analysis of RNA-seq incorporating quantification uncertainty. Nat Methods 14: 687–690

Radutoiu S, Madsen LH, Madsen EB, Felle HH, Umehara Y, Grønlund 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

Radutoiu S, Madsen LH, Madsen EB, Jurkiewicz A, Fukai E, Quistgaard EMH, Albrektsen AS, James EK, Thirup S, Stougaard J (2007) LysM domains mediate lipochitin--oligosaccharide recognition and Nfr genes extend the symbiotic host range. EMBO J 26: 3923–3935

Rogers C, Oldroyd GED (2014) Synthetic biology approaches to engineering the nitrogen symbiosis in cereals. J Exp Bot 65: 1939–1946

Salgado MG, van Velzen R, Van Nguyen T, Battenberg K, Berry AM, Lundin D, Pawlowski K (2018) Comparative Analysis of the Nodule Transcriptomes of Ceanothus thyrsiflorus (Rhamnaceae, Rosales) and Datisca glomerata (Datiscaceae, Cucurbitales). Front Plant Sci 9: 1629

Shimizu T, Nakano T, Takamizawa D, Desaki Y, Ishii-Minami N, Nishizawa Y, Minami E, Okada K, Yamane H, Kaku H, et al (2010) Two LysM receptor molecules, CEBiP and OsCERK1, cooperatively regulate chitin elicitor signaling in rice. Plant J 64: 204–214

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

Soltis DE, Soltis PS, Morgan DR, Swensen SM, Mullin BC, Dowd JM, Martin PG (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

Spaink HP, Wijffelman CA, Okker RJ, Lugtenberg BE (1989) Localization of functional regions of the Rhizobium nodD product using hybrid nodD genes. Plant Mol Biol 12: 59–73

Srinivasan R, Huang BS, Venugopal S, Johnston AD, Chai H, Zeng H, Golshani P, Khakh BS (2015) Ca(2+) signaling in astrocytes from Ip3r2(-/-) mice in brain slices and during startle responses in vivo. Nat Neurosci 18: 708–717

Stougaard J (1995) Agobacterium rhizogenes as a vector for transforming higher plants: Application in Lotus corniculatus transformation. Plant gene transfer and expression protocols 49–61

Stougaard J, Abildsten D, Marcker K a. (1987) The Agrobacterium rhizogenes pRi TL-DNA segmen as a gene vector system for transformation of plants. Mol Gen Genet 207: 251–255

Sullivan JT, Trzebiatowski JR, Cruickshank RW, Gouzy J, Brown SD, Elliot RM, Fleetwood DJ, McCallum NG, Rossbach U, Stuart GS, et al (2002) Comparative sequence analysis of the symbiosis island of Mesorhizobium loti strain R7A. J Bacteriol 184: 3086–3095

Trifinopoulos J, Nguyen L-T, von Haeseler A, Minh BQ (2016) W-IQ-TREE: a fast online phylogenetic tool for maximum likelihood analysis. Nucleic Acids Research 44: W232–W235

Trinick MJ, Galbraith J (1980) The rhizobium requirements of the non-legume parasponia in relationship to the cross-inoculation group concept of legumes. New Phytol 86: 17–26

Trouvelot, A, Kough J L, Gianinazzi-Pearson V (1986) Mesure du taux de mycorhization VA d'un systeme radiculaire . Recherche de methodes d'estimation ayant une signification fonctionnelle. In GS Gianinazzi-Pearson V, ed, Physiological and Genetic Aspects of Mycorrhizae. INRA Press, Paris, pp 217–221

Van Nguyen T, Wibberg D, Vigil-Stenman T, Berckx F, Battenberg K, Demchenko KN, Blom J, Fernandez MP, Yamanaka T, Berry AM, et al (2019) Frankia-enriched metagenomes from the earliest diverging symbiotic Frankia cluster: They come in teams. Genome Biol Evol 11: 2273–2291

van Velzen R, Doyle JJ, Geurts R (2018a) A Resurrected Scenario: Single Gain and Massive Loss of Nitrogen-Fixing Nodulation. Trends Plant Sci. doi: 10.1016/j.tplants.2018.10.005

van Velzen R, Holmer R, Bu F, Rutten L, van Zeijl A, Liu W, Santuari L, Cao Q, Sharma T, Shen D, et al (2018b) 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

Wang D, Yang S, Tang F, Zhu H (2012) Symbiosis specificity in the legume--rhizobial mutualism. Cell Microbiol 14: 334–342

Wang H, Moore MJ, Soltis PS, Bell CD, Brockington SF, Alexandre R, Davis CC, Latvis M, Manchester SR, Soltis DE (2009) Rosid radiation and the rapid rise of angiosperm-dominated forests. Proc Natl Acad Sci U S A 106: 3853–3858

Wan J, Zhang X-C, Neece D, Ramonell KM, Clough S, Kim S-Y, Stacey MG, Stacey G (2008) A LysM receptor-like kinase plays a critical role in chitin signaling and fungal resistance in Arabidopsis. Plant Cell 20: 471–481

Wardhani TAK, Roswanjaya YP, 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. doi: 10.3791/59971

Werner GDA, Cornelissen JHC, Cornwell WK, Soudzilovskaia NA, Kattge J, West SA, Kiers ET (2018) Symbiont switching and alternative resource acquisition strategies drive mutualism breakdown. Proc Natl Acad Sci U S A. doi: 10.1073/pnas.1721629115

Yamaguchi K, Yamada K, Ishikawa K, Yoshimura S, Hayashi N, Uchihashi K, Ishihama N, Kishi-Kaboshi M, Takahashi A, Tsuge S, et al (2013) A receptor-like cytoplasmic kinase targeted by a plant pathogen effector is directly phosphorylated by the chitin receptor and mediates rice immunity. Cell Host Microbe 13: 347–357

Young ND, Debellé F, Oldroyd GED, Geurts R, Cannon SB, Udvardi MK, Benedito VA, Mayer KFX, Gouzy J, Schoof H, et al (2011) The Medicago genome provides insight into the evolution of rhizobial symbioses. Nature 480: 520–524

van Zeijl A, Wardhani TAK, Seifi Kalhor M, Rutten L, Bu F, Hartog M, Linders S, Fedorova EE, 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: 284

Zhang X-C, Cannon SB, Stacey G (2009) Evolutionary genomics of LysM genes in land plants. BMC Evol Biol 9: 183

Zhang X, Dong W, Sun J, Feng F, Deng Y, He Z, Oldroyd GED, Wang E (2015) The receptor kinase CERK1 has dual functions in symbiosis and immunity signalling. Plant J 81: 258–267

Zhao Y, Araki S, Wu J, Teramoto T, -F. Chang Y, Nakano M, Abdelfattah AS, Fujiwara M, Ishihara T, Nagai T, et al (2011) An Expanded Palette of Genetically Encoded Ca2 Indicators. Science 333: 1888–1891

Zhu H, Riely BK, Burns NJ, Ané J-M (2006) Tracing nonlegume orthologs of legume genes required for nodulation and arbuscular mycorrhizal symbioses. Genetics 172: 2491–2499

CHAPTER 5

5

Analysis of Nodulation correlated Receptor like kinases of *Parasponia* reveals novel phenotypes in the infection process

Luuk Rutten, Yuda Roswanjaya , Sidney Linders, Alexander van Hall, Arjan van Zeijl, Rene Geurts.

Wageningen University, Department of Plant Science, Laboratory of Molecular Biology, Droevendaalsesteeg 1, 6708PB Wageningen, The Netherlands

Manuscript in Preparation

Abstract

Parasponia in the Cannabaceae represents the only genus outside the Fabaceae which forms nitrogen-fixing root nodules with rhizobium. Recent phylogenomic analyses suggest that the nodulation trait has a single evolutionary origin, dated ~110 million years ago, followed by multiple losses of this trait. Phylogenetic comparisons between nodulating *Parasponia* and non-nodulating relative *Trema* species has identified seven genes that have nodule-specific expression and are pseudogenized in *Trema*. Two of these genes represent uncharacterized membrane-bound receptor-like kinases; Parasponia *LECTIN RECEPTOR-LIKE KINASE 1 (PanLEK1)* and *CYSTEINE RICH RECEPTOR-LIKE KINASE 11 (PanCRK11)*. Here we show that PanLEK1 is essential for defense response suppression in *Parasponia* nodules, while multiple PanCRKs play a vital role in infection thread progression in the nodule. Characterization of the respective L-lectin and Cysteine Rich Kinase gene families revealed a large *CRK* gene cluster which contained 12 nodule-enhanced *CRK* genes and specific adaptations in the PanLEK1 kinase domain. CRISPR-CAS9 mutagenesis of *PanLEK1* triggered defense response phenotypes in Parasponia nodules, while a ~1Mb deletion of the *CRK* gene cluster led to reduced infection levels. These results demonstrate novel roles for L-lectin receptor kinases and Cysteine rich kinases in rhizobium infection of the non-legume *Parasponia* and adds to the currently known roles of these gene families in the regulation of pattern triggered immunity and ROS perception.

Introduction

Plant growth is often limited by nutrient availability in the soil. Strategies to overcome a shortage of nutrients have been instrumental for plants to colonize land. Often these strategies involve the recruitment of symbiotic microbial partners, such as arbuscular mycorrhizal fungi and nitrogen-fixing bacteria. The level of symbiotic intimacy may vary from loose associations up to intracellular hosting of the microbe. The latter is a well-known strategy of legumes, *Parasponia* spp. and actinorhizal plants that establish a nitrogen-fixing nodule symbiosis with either rhizobia or *Frankia* bacteria (James 2000; Vessey, Pawlowski, and Bergman 2005). Nodules provide optimal physiological conditions for enzymatic conversion of dinitrogen into ammonia by the bacteria, allowing the transfer of fixed nitrogen to the host. In turn, the plant provides fixed carbon to the microsymbiont.

A commonality of nitrogen-fixing nodule endosymbioses is the ability of bacteria to bypass immune responses. The signalling pathway leading to successful intracellular infection is founded on the signalling pathway also used by arbuscular mycorrhizal fungi, and therefore is called the common symbiosis signalling pathway (CSSP) (Oldroyd 2013). The CSSP is activated by microbe secreted chitooligomers (COs),

lipochitooligomers (LCOs), and/or effectors leading to the activation of symbiosis-specific transcriptional networks. Genes encoding components of the CSSP are widespread, as over 70 percent of all land plants can establish arbuscular mycorrhizal symbiosis (Delaux et al. 2013; Hoeksema et al. 2018). In contrast, the nitrogen-fixing nodulation trait occurs only in 10 paraphyletic lineages of the related orders Fabales, Fagales, Cucurbitales and Rosales that are collectively known as the nitrogen-fixing clade (NFC) (Soltis et al. 1995). In two lineages, the nodulating symbiotic partner is rhizobium, namely in legumes (Fabaceae, Fabales) and *Parasponia* (Cannabaceae, Rosales), whereas plant species in the other eight lineages nodulate with *Frankia (Swensen and Mullin 1997)*. Recent phylogenomic analyses suggest that the nodulation trait in the NFC has a single evolutionary origin, dated ~110 million years ago, followed by multiple losses of this trait (R. van Velzen et al. 2018; Griesmann et al. 2018; Robin van Velzen, Doyle, and Geurts 2019). Yet, the genetic adaptations that were causal for the evolution of the nitrogen-fixing trait remain elusive.

Comparative analysis of three nodulating *Parasponia* and two non-nodulating *Trema* species identified seven genes that have nodule-specific expression and are pseudogenized in *Trema*. Among these are genes that are essential for nodule organogenesis and bacterial infection, like the LysM-type LCO receptor *NOD FACTOR PERCEPTION* (*NFP2*), the transcription factor *NODULE INCEPTION* (*NIN*) and the coiled-coil protein *RHIZOBIUM POLAR GROWTH* (*RPG*). However, the remaining genes have not been functionally characterized, and include two membrane-bound receptor-like kinases named *LECTIN-LIKE RECEPTOR 1* (*PanLEK1*) and *CYSTEINE RICH RECEPTOR KINASE 11* (*PanCRK11*) in *Parasponia andersonii (R. van Velzen et al. 2018)*. Both these receptor-like kinases belong to relatively large gene families; the L-type LecRLKs and CRKs. Both these families are greatly expanded in vascular plants (Vaattovaara et al. 2019; Bellande et al. 2017). Since these genes have a nodule-specific expression pattern in *P. andersonii* and got lost independently in two *Trema* species, we hypothesise that both receptor-like kinases commit a symbiotic function. Therefore, we explored their respective gene families in *P. andersonii* and studied their symbiotic functions.

Lectin receptor-like kinases are classified into three groups based on their extracellular lectin domain; G, C, and L-type. L-type lectin kinases, named after ‘Legume Lectins’, are predicted to bind sugar-like oligomers or small hydrophobic molecules in the extracellular space (André et al. 2005; Bouwmeester and Govers 2009; Wang and Bouwmeester 2017). For example, the *Arabidopsis thaliana* L-type lectin receptor AtDORN1 has been shown to bind extracellular ATP, which functions as an important ‘danger signal’ in pathogen infection (Choi et al. 2014). In recent years multiple other studies have implicated the involvement of Lectin receptor-like kinases in signalling upon recognition of Pathogen Associated Molecular Patterns

(PAMPs), although the ligands that bind to these receptors are often not evaluated (Singh et al. 2012; Wang et al. 2014; Yeh et al. 2015). The extracellular domains of the L-type Lectin receptor kinases are homologous to soluble lectins, initially discovered in legume seeds. Genes encoding soluble lectins are also expressed in legume roots and the encoded proteins are secreted in the extracellular space. Soluble legume lectins have sugar binding characteristics and can serve as host determinants in the legume-rhizobium interaction (Díaz et al. 1989). This has led to the lectin recognition hypothesis in which soluble legume lectins are important symbiont recognition proteins, interacting with sugar chains on the bacterial cell wall (Dazzo and Hubbell 1975; Hirsch 1999). Studies in the legume model *Medicago truncatula* showed that that ectopic expression of a truncated L-Lectin receptor-like kinase significantly increased nodulation efficiency (Navarro-Gochicoa et al. 2003). Therefore, the potential involvement of a L-lectin receptor-like kinase in *Parasponia* symbiotic signaling is intriguing.

The classical structure of CRKs is related to the Leucine Rich Repeat (LRR)- and S-locus lectin receptor kinases, and comprises of a signal peptide, generally two DUF26 motifs, a transmembrane domain and a kinase domain (Vaattovaara et al. 2019). The DUF26 motif consist of a core C-x8-C-x2-C structure and around this usually several more cysteine residues can be present, which may be essential for intramolecular disulfide bond formation. The DUF26 motif has homology to *Ginkgo biloba* GINKBILOBIN2 (GbGNK2), a small seed protein with antifungal activity (Miyakawa et al. 2007). The DUF26 motif occurs also in secreted proteins and receptor-like proteins that lack a kinase but have a transmembrane domain (Vaattovaara et al. 2019). The exact function of the DUF26 motif is not fully understood. Molecular modelling of proteins with a double DUF26 motif configuration reveals potential as carbohydrate recognition modules (Vaattovaara et al. 2019). This is in line with the finding that GbGNK2 as well as two DUF26 motif containing proteins of maize (*Zea mais* AFP1 and AFP2) bind mannose along with functioning in fungal defense (Miyakawa et al. 2009, 2014; Ma et al. 2018; Han et al. 2019). Therefore, DUF26 motif containing proteins might have the potential for binding sugar-like molecules (Vaattovaara et al. 2019).

CRKs are mainly known to regulate biotic and abiotic stress responses. Large scale phenotyping in *A. thaliana* has led to various abiotic and biotic stress phenotypes (Bourdais et al. 2015). Ectopic expression of *AtCRK4, AtCRK6, AtCRK28, AtCRK36 and AtCRK45* leads to increased resistance to *Pseudomonas syringae (Zhang et al. 2013; Yeh et al. 2015; Yadeta et al. 2017)*, while ectopic expression of *AtCRK5* leads to increase ABA sensitivity and increased drought tolerance (Lu et al. 2016). Several *CRK* genes are also implicated in regulating cell death in *A. thaliana* and rice (*Oryza sativa*) (Lee et al. 2017; Yadeta et al. 2017; Du et al. 2018). CRKs have

been hypothesized to act as sensors for apoplastic Reactive Oxygen Species (ROS). Mechanistically this is not well understood, tough the cysteine residues might be sensitive to ROS molecules (Qi et al. 2017). In symbiosis only a single CRK was found to be essential; namely *M. truncatula SymCRK* (also known as *M. truncatula DEFECTIVE IN NITROGEN FIXATION 5* (*MtDNF5*). This gene was found to be essential for symbiotic differentiation of rhizobium and suppression of defense responses in the nodule regulated by ethylene (Berrabah et al. 2014, 2018).

Given the involvement of both these receptor families in essential processes of microbe recognition a role of *PanCRK11* and *PanLEK1* in symbiotic establishment can be envisioned. Here we test whether both genes encode potential symbiotic signaling receptors. We characterize their respective gene families in *P. andersonii* and studied their roles in nodule formation by CRISPR-CAS9 mutagenesis.

Results

Characterization of Parasponia Cysteine and L-type Lectin receptor kinases

First, we carried out a phylogenetic reconstruction of the *CRK* gene family of the *Parasponia* and *Trema* lineages. In total 32 different *CRK* genes were identified among the two lineages. Of these, 29 genes showed to be present in both *P. andersonii* and *Trema orientalis. P. andersonii* carries 29 functional *CRK* genes, missing functional orthologs of *T. orientalis TorCRK8, TorCRK21* and *TorCRK32*. *T. orientalis* carries 31 *CRK* genes, missing a putative ortholog of *P. andersonii CRK11*, a gene that is specifically expressed in nodules **(Suppl. Table 3.)**. Loss of *CRK11* has been shown to be consistent among other *Trema* species whereas it is present in two additional *Parasponia* species, suggesting this gene may commit a nodule-specific function (R. van Velzen et al. 2018). We made use of previously assembled orthogroups and merged these with CRK protein sequences from *Solanum lycopersicum, Cucumis sativus, Datisca glomerata*, *Alnus glutinosa* and common bean (*Phaseolus vulgaris*) to create a larger dataset (Huang et al. 2009; Tomato Genome Consortium 2012; Griesmann et al. 2018; R. van Velzen et al. 2018; Quezada et al. 2019). The *CRK* gene family is diverse and fast evolving (Vaattovaara et al. 2019). We noted that among the selected proteins especially the extracellular domain region was highly variable. Although CRKs mostly occur with two DUF26 motifs in the extracellular region, it is not uncommon to find proteins with only a single or multiple DUF26 motifs **(Suppl. Table 3.)**. We found that no stable consistent results were obtained when aiming to infer the phylogeny from alignments of full length CRK protein sequences. Therefore we conducted phylogenetic reconstruction on the conserved kinase domains. Similar to previous phylogenies on CRK proteins a clear separation between a basal and a variable cluster can be found (Vaattovaara et al. 2019). The basal cluster consists out of clades I, II, III, IV and V, encompassing only four *P.*

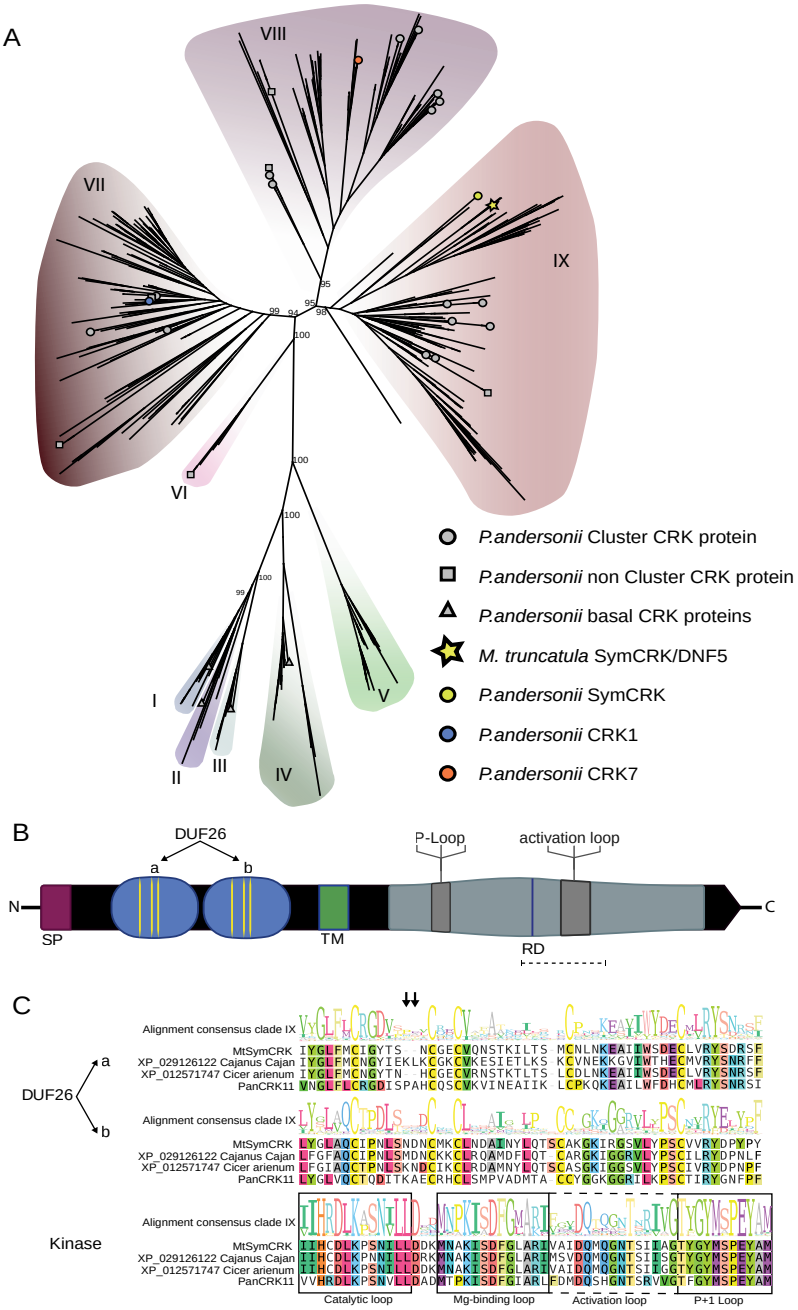


Figure 1. Phylogeny and gene structure of Cysteine-Rich Receptor inases (CRKs). **A,** Phylogenetic reconstruction based on kinase domain of CRK proteins of *P. andersonii, G. max, P. vulgaris, A. thaliana, T. orientalis*, *M. truncatula, F. vesca, C. sativus, D. glomerata, D. trinervis, A glutinosa.* The CRK family shows a separation between basal clades I-V and variable clades VI-IX. Node numbers indicate posterior probabilities leading to clades. **B,**

Domain organisation of CRK proteins. SP: signal peptide; DUF26: Domain of Unknown Function 26; TM: transmembrane spanning domain; P-loop: phosphate-binding loop; RD: conserved amino acids arginine and aspartic acid; activation loop: conserved region that upon phosphorylation triggers a conformational change. **C,** Alignments of DUF26a, DUF26b motifs and catalytic domains of the kinase region between legume MtSymCRK/MtDNF5 orthologs and PanCRK11. Note that *P. andersonii* CRK11 does not have arterations in the DUF26a or kinase domains. LOGO based on alignment of 85 complete protein sequences belonging to clade IX, supplemented with sequences from *Vigna angularis, Cajanus cajan, Cicer arietinum, Abrus precatorius, Vigna unguiculata.*

andersonii CRK genes **(Figure 1A)**. The variable cluster consists out of clades VI, VII, VIII and IX. The CRK proteins that cluster in the variable clade mostly separated into lineage dominated clades rather than intro true orthogroups, which indicates fast evolution of these genes. Since rearrangements are not uncommon between *CRK* genes, it is likely no true orthogroups can be inferred in the variable CRKs (Vaattovaara et al. 2019).

Nodule enhanced CRKs belong to the variable group and have a classical CRK structure with RD-kinase

In total *12 P. andersonii CRK* genes have a nodule enhanced expression pattern, of which *PanCRK1*, *PanCRK7* and *PanCRK11* are highest expressed in nodules **(Figure S1, Suppl. Table 3.)**. All of these 12 *PanCRK* genes belong to the variable cluster and are dispersed across all four clades. The encoded proteins of the nodule enhanced *P. andersonii CRK* genes have a classical configuration with a signal peptide, two DUF26 motifs in the extracellular domain (named DUF26a and DUF26b), a transmembrane region and serine/threonine-protein kinase domain **(Figure 1B)**. PanCRK11 seems closely related to MtSymCRK of *M. truncatula*, which groups in a small legume dominated lineage of clade IX. However, unlike MtSymCRK the PanCRK11 DUF26a and b motifs contain a conserved C-8x-C-2x-C structure, whereas MtSymCRK has a C-6x-C-2x-C structure in the DUF26a motif (Berrabah et al. 2014). Furthermore, PanCRK11 also has a classical RD-kinase, unlike the CD-kinase reported for MtSymCRK (Berrabah et al. 2014) **(Figure 1C)**. In line with this we hypothesize that PanCRK11 should be capable of auto-activation.

Many *Parasponia andersonii* variable clade CRKs occur in a genomic cluster

Previously it has been shown that *CRK* genes of the variable clades often occur in genomic clusters, for example in *A. thaliana* (Chen 2001; Bourdais et al. 2015), soybean (*Glycine max*) (Delgado-Cerrone et al. 2018) and common bean (Quezada et al. 2019). We found that in *P. andersonii* 20 out of 26 *CRK* genes of the variable clades were located in a single genomic cluster spanning over ~1 Mbps, whereas the six remaining *CRK* genes of the variable clades are in non-clustered locations **(Figure 2A)**. The cluster is largely colinear between *Parasponia andersonii* and *Trema orientalis***(Figure 2A, Figure S2.)**.

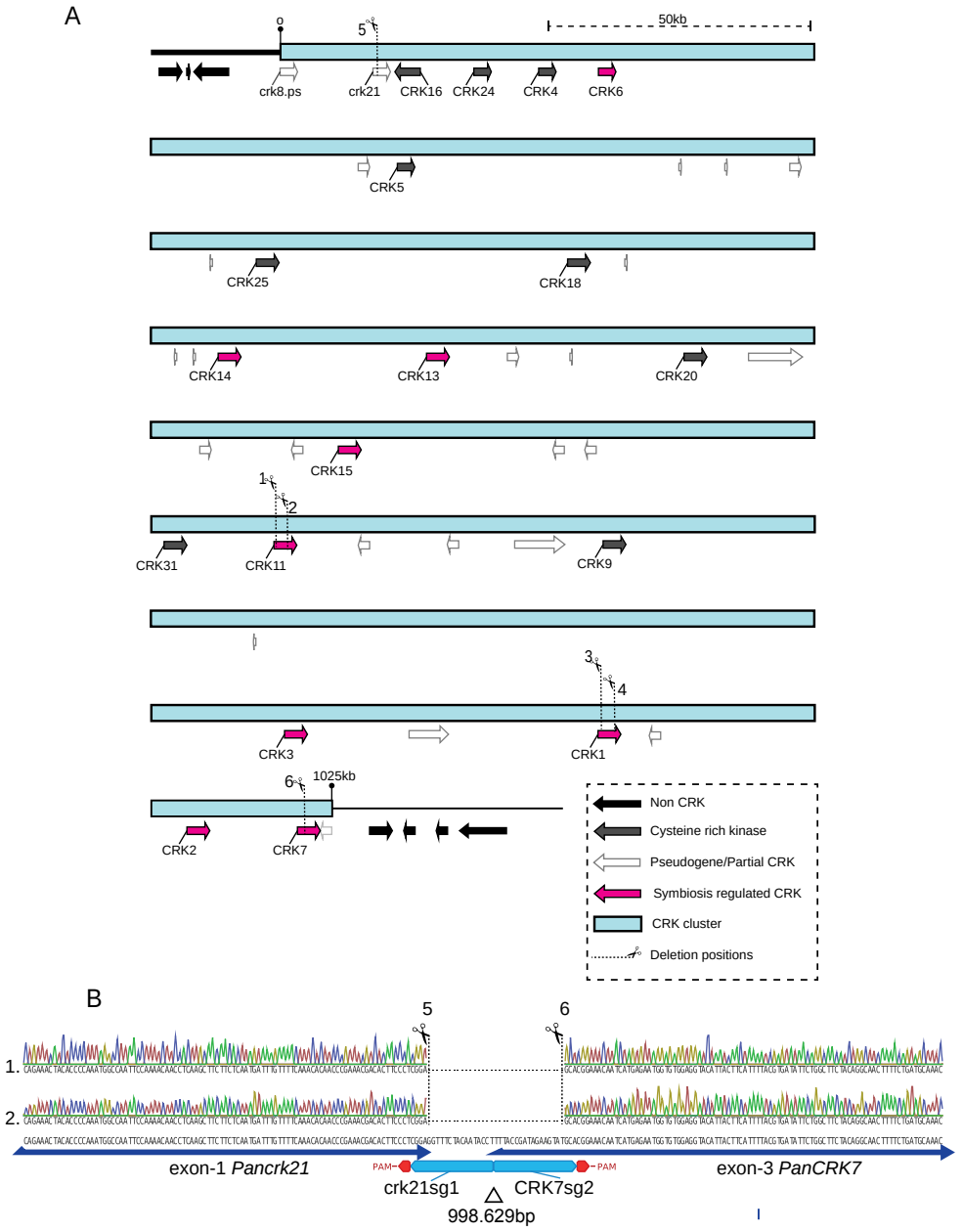


Figure 2. *Parasponia andersonii CRK* gene cluster. **A,** Schematic representation of *PanCRK* gene cluster organisation. CRISPR-Cas9 target sites indicated by scissors. *PanCRK11*, sites 1 & 2. *PanCRK1* sites 3&4, *PanCRK7* site 6, and *Panclfk* sites 5&6. Overall length of the *CRK* cluster is ~1Mbps. **B,** Genotype of the *PanCRK* cluster deletion in the *Panclfk* mutant, amplified and sequenced by primer set Ru_130 **(Suppl. table 1)**.

Phylogenetic reconstruction of L-type Lectin kinases.

A similar approach to the *CRK* phylogeny was conducted for the L-type lectin receptor kinases. Orthogroups were assembled into a large alignment and supplemented with extra sequences from *D. trinervis, A.glutinosa* and *C. glauca (Griesmann et al. 2018; R. van Velzen et al. 2018)*. The L-type lectin gene family in the *Parasponia* and *Trema* lineage consists of 26 members*. P. andersonii* contains 24 L-type Lectins kinases. *T. orientalis* is missing a *LEK1* ortholog, but in turn has two additional genes, *TorLEK9* and *TorLEK10*, which appear to be lost in *P. andersonii* **(Suppl. Table 4)**. The resulting phylogeny revealed 12 distinct clades, which except for clade II all contain *P. andersonii* and *T. orientalis* proteins **(Figure S3A)**. PanLEK1 groups in clade I. This clade can be separated into subclades Ia, Ib and Ic, with *PanLEK1* gene falling into clade Ia **(Figure 3A)**.

PanLEK1 is induced in Nodules and encodes a protein with a highly diverged kinase domain.

Among all *P. andersonii* L-type Lectin receptor kinases *PanLEK1* is the only gene with a significant nodule enhanced expression. Notably it appears that most other members of clade I are lower expressed in the nodules, with close homolog *PanLEK2* specifically expressed in above ground tissues **(Figure S3B)**. *PanLEK1* and *PanLEK2* are located in tandem, however the PanLEK1 protein appears to be highly dissimilar from PanLEK2, as well as other LEK proteins, indicated by its long branch length. The long branch length warranted a closer look at the protein structure. The PanLEK1 kinase domain contains several substitutions in amino acids that are generally conserved and essential for protein kinase activity. PanLEK1 contains amino acid substitutions in the P-LOOP, with two glycines substituted for charged amino acids, potentially influencing ATP binding. Further, the kinase domain lacks the RD-motif and shows no conservation in activation loop compared to the consensus sequence of kinase domains in clade Ia proteins **(Figure 3B,C)**. Based on these substantial substitutions we estimate the PanLEK1 kinase domain to be inactive(Dardick, Schwessinger, and Ronald 2012; Johnson, Noble, and Owen 1996). These kind of substitutions are not found in other L-type lectin kinases in *P.andersonii*.

PanLEK1 is essential for effective nodule formation

A protocol for the efficient stable transformation of Parasponia was developed recently (van Zeijl et al. 2018; Wardhani et al. 2019). We decided to target *PanLEK1,* using three guide RNAs. Three T_0 mutants lines, containing different mutations in *PanLEK1* were selected for experimentation **(Figure S4**). An EV-control line expressing *CAS9*, but without guide RNAs was generated as a control. These mutant lines were maintained in tissue culture and rooted for experimentation (Webster et al. 1995; Wardhani et al. 2019). Rooted plantlets were grown in Pots for five weeks with

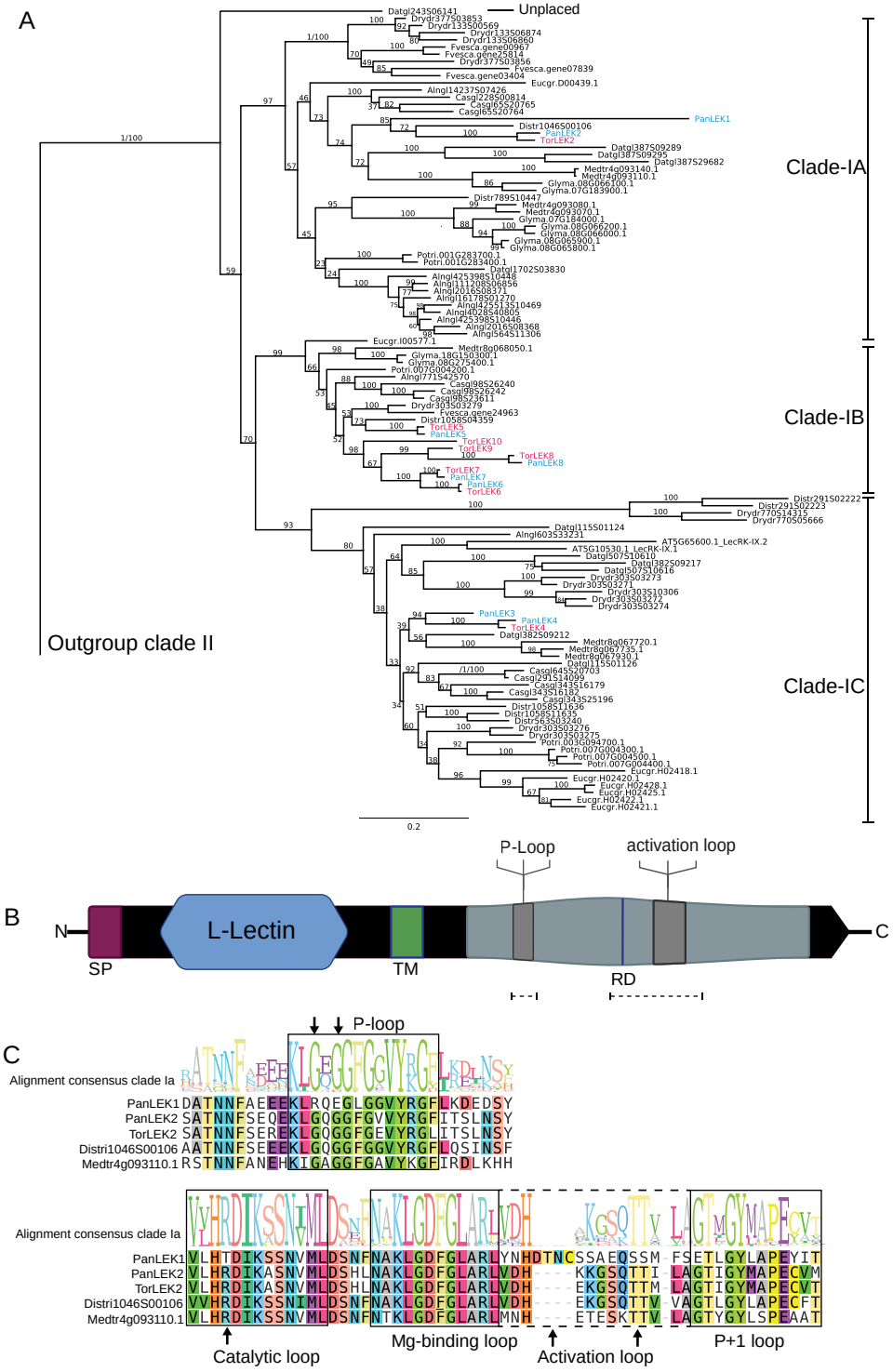

A
Unplaced
Clade-IA
Clade-IB
Clade-IC
Outgroup clade II
0.2
B
N
C
L-Lectin
SP
TM
RD
P-Loop
activation loop
C
P-loop
Alignment consensus clade Ia
PanLEK1 DATNNFAEEEKLRQEGLGGVYRGFLKDEDSY
PanLEK2 SATNNFSEQEKLGQGGFGVVYRGFITSLNSY
TorLEK2 SATNNFSERERKLGQGGFGEVYRGLITSLNSY
Distri1046S00106 AATNNFSEEEKLGQGGFGGVYRGFLQSINSF
Medtr4g093110.1 RSTNNFANEHKIGAGGFGAVYKGFIRDLKHH
Alignment consensus clade Ia
PanLEK1 VLHTDIKSSNVMLDSNFNAKLGDFGLARLYNHDTNCSSAEQSSM-FSETLGYLAPEYIT
PanLEK2 VLHRDIKASNVMLDSHLNAKLGDFGLARLVDH----KKGSQTTI-LAGTIGYMAPECVM
TorLEK2 VLHRDIKASNVMLDSHLNAKLGDFGLARLVDH----EKGSQTTM-LAGTIGYMAPECVT
Distri1046S00106 VVHRDIKSSNIMLDSNFNAKLGDFGLARLVDH----EKGSQTTV-VAGTLGYLAPECFT
Medtr4g093110.1 VLHRDIKSSNVMLDSNFNTKLGDFGLARLMNH----ETESKTTV-LAGTYGYLSPEAAT
Catalytic loop
Mg-binding loop
Activation loop
P+1 loop

Figure 3. Phylogeny and gene structure of L-type Lectin kinases. (**A),** Phylogenetic reconstruction of clade I L-type Lectin kinases of *P. andersonii, Trema orientalis, Discaria trinervis, Alnus glutinosa, Fragaria vesca, Eucalyptus grandis, Populus trichocarpa, Medicago truncatula, Glycine max.* (Figure S3A) (**B),** Protein structure of PanLEK1. SP: signal peptide, L-Lectin: lectin motif; TM: transmembrane spanning domain; P-loop: phosphate-binding loop; RD: conserved amino acids arginine and aspartic acid; activation loop: conserved region that upon phosphorylation triggers a conformational change. (**C),** Protein alignment of kinase P-loop and activation loop segment compared to clade I consensus and the selected proteins *P. andersonii* PanLEK2, *T. orientalis* TorLEK2, *Discaria trinervis* Distri104S00160 and *M. truncatula* Medtr4g093110. PanLEK1 contains substitutions in these domains, that are not present in paralog PanLEK2 or in homologous clade I proteins. LOGO based on alignment of 46 clade IA sequences (Suppl. file 4).

the *P. andersonii* compatible rhizobium strain *Mesorhizobium plurifarium* BOR2 (R. van Velzen et al. 2018). We noted a significant reduction in nodule number among all three *Panlek1* knockout lines. However, no consistent difference in biomass, nodule mass or nodule size was detected **(Figure 4A,B,C)**. Semi-thin sections of *Panlek1* nodules revealed a striking phenotype. Although *Panlek1* nodules showed no phenotype in infection thread progression, a striking accumulation of phenolic content in mature mutant nodules was observed. The phenotype was associated with patches of dead cell **(Figure 4D,E,F)**. This suggests a role for PanLEK1 in nodule maintenance.

A CRK cluster knockout shows infection phenotypes in Parasponia nodules

To find out the role of *P. andersonii CRK* genes in nodulation we decided to knock out *PanCRK1* and *PanCRK11* by using CRISPR-CAS9. We obtained three T_0-mutant lines for *PanCRK1* and two for *PanCRK11.* These mutations consist of small deletions/insertions at the target guide sites **(Figure S5A)**. Transgenic plantlets were grown in pots and nodulated with *M.plurifarium* BOR2 and harvested after four weeks of inoculation. However, no consistent phenotypes between CAS9 containing transgenic control and *Pancrk1* and *Pancrk11* mutant lines could be observed**(Figure S6A,B,C)**. Also, semi-thin sections revealed no obvious defects in infected cells or visible signs of activated defense responses**(Figure S6D-I)**.

Since 10 out of 12 CRK genes with a nodule enhanced expression occur in a cluster **(Figure 2A, Figure S1)**. we investigated whether these genes may have redundant functions. To do so, we aimed to remove the entire *CRK* cluster. Two guides located on either end of the cluster in *PanCRK7* and the pseudogene *Pancrk21* were used. Transgenic shoots were screened with cluster spanning primers. We observed a relatively low regeneration efficiency with this construct, and identified only a single transgenic shoot in which the ~1 Mbps cluster was effectively removed **(Figure 2B)**. This line is referred to as the *CRK* Cluster Full KO or *Panclfk*. Further we obtained two lines with only small deletions knocking out only *PanCRK7 Pancrk21.* Since *CRK21*

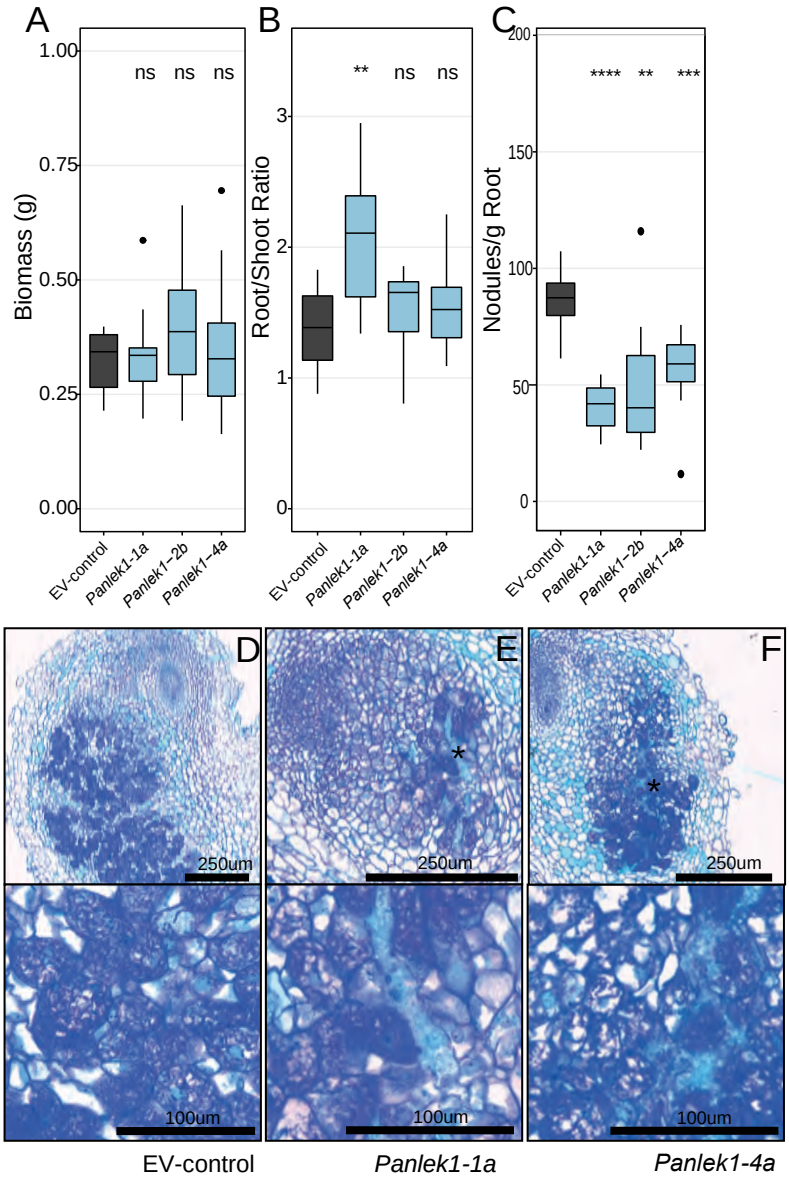


Figure 4. Phenotype of three independent *Panlek1* knockout lines. 4 weeks post inoculation with *M. plurifarium* BOR2 (**A-C),** Box plots of biomass (A)**,** root shoot ratio (B)**,** and nodule number / g. root weight (C). EV-control n=10, *Panlek1-1a* n=10, *Panlek1-2b* n=10, *Panlek1-4a* n=10 Student's T-tests relative to empty vector control (EV-control) samples. P>0.05 not significant (ns), P<0.05 *, P<0.01 ** P<0.001 ***P <0.0001 ****. **D,** Representative section of EV-control nodule, scale bar 250 µm, zoom in 100 µm. **F,** Representative section *Panlek1-1a* mutant nodule, scale bar 250 µm, zoom in 100 µm. **G,** Representative section *Panlek1-4a* mutant nodule, scale bar 250 µm, zoom in 100 µm. Note the accumulation of green/blue stained phenolic compounds in the infection zone (asterisk). Sections stained with toluidine blue.

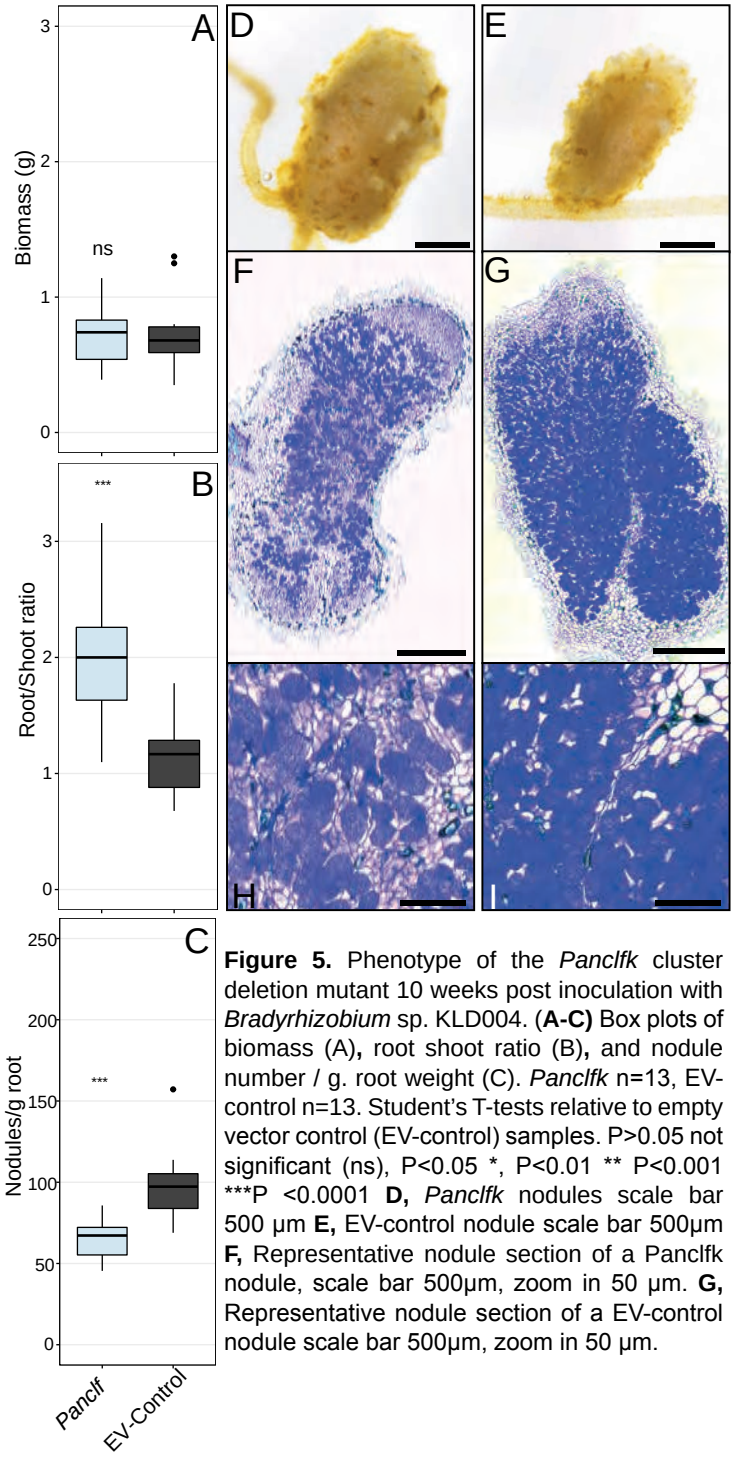
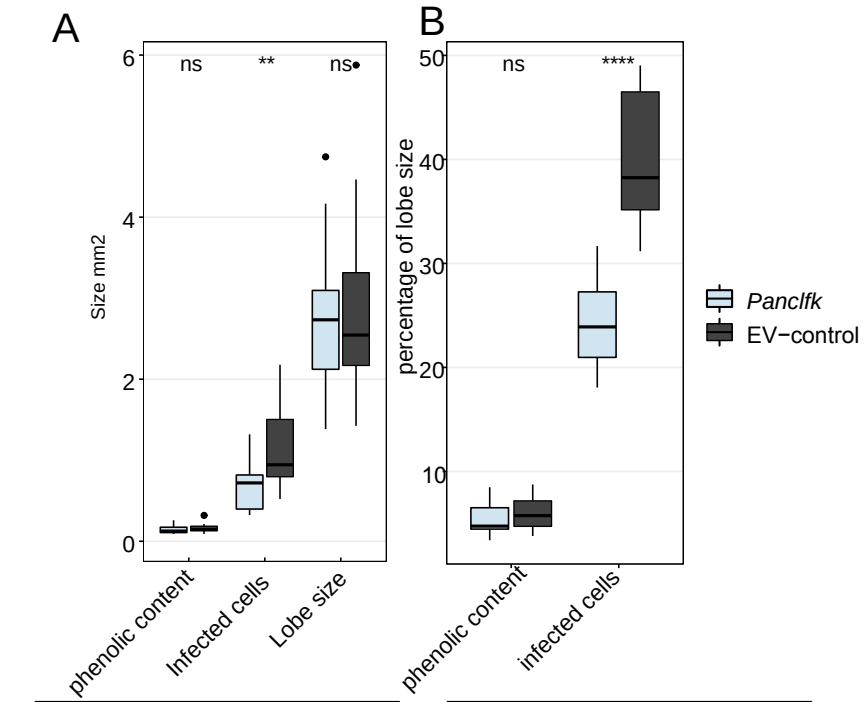


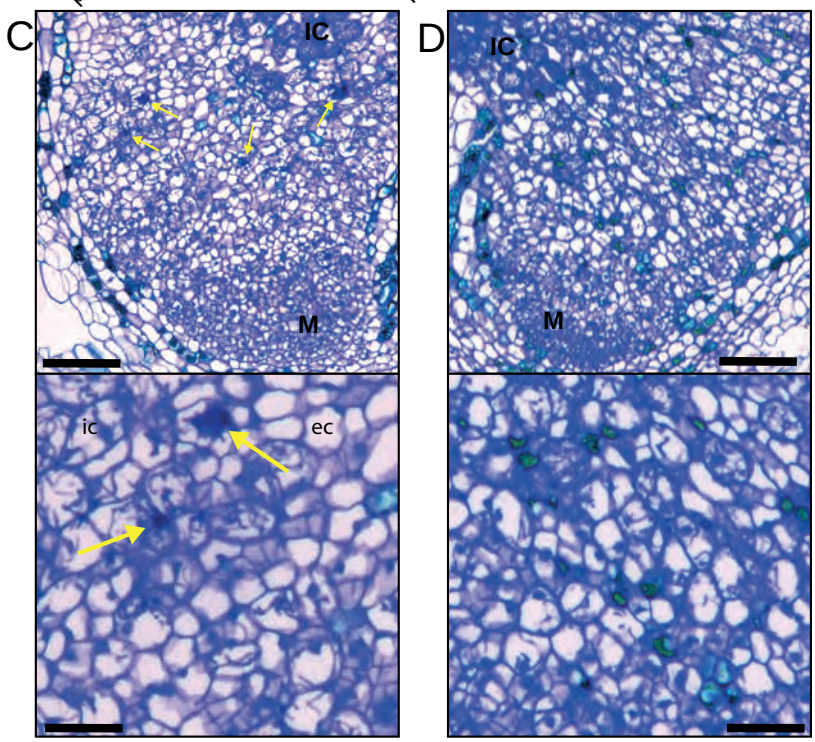
Figure 5. Phenotype of the *Panclfk* cluster deletion mutant 10 weeks post inoculation with *Bradyrhizobium* sp. KLD004. **(A-C)** Box plots of biomass (A)**,** root shoot ratio (B)**,** and nodule number / g. root weight (C). *Panclfk* n=13, EV-control n=13. Student's T-tests relative to empty vector control (EV-control) samples. P>0.05 not significant (ns), P<0.05 *, P<0.01 ** P<0.001 ***P <0.0001 **D,** *Panclfk* nodules scale bar 500 µm **E,** EV-control nodule scale bar 500µm **F,** Representative nodule section of a Panclfk nodule, scale bar 500µm, zoom in 50 µm. **G,** Representative nodule section of a EV-control nodule scale bar 500µm, zoom in 50 µm.

represents a pseudogene in *P. andersonii*, these mutants we kept as *Pancrk7-1* and *Pancrk7-3* **(Figure S5)**. To further characterize the Panclfk cluster deletion line we screened for presence/absence of *CRK* genes in the genome, confirming the absence of the 20 cluster located *CRK* genes and the presence of the remaining CRK genes **(Figure S5B)**.

The *Panclfk* line grew and rooted normally in tissue culture. In an initial experiment the *Panclfk* line was inoculated with *M. plurifarium* BOR2 for 4 weeks. Although a small reduction in number of nodules was observed, these differences were not significant, probably due to variation in plantlet size. Upon sectioning however, we noted a relatively longer infection zone in *Panclfk* mutant nodules, with the presence of thick bulging infection threads **(Figure S7)**. To better observe these phenotypes a second nodulation experiment was done using an additional rhizobium strain, one that is native to *Parasponia* and isolated from the volcanic mount Kelud. This *Bradyrhizobium sp.* strain KLD004 represents highly effective nodulator and nitrogen fixer of *Parasponia spp.*, although it showed to be somewhat slower in infecting pants when compared to the fast growing strain *M. plurifarium* BOR2. Therefore plants were examined 10 weeks post inoculation. At this time point the differences between *Panclfk* and EV-control plants were clearly visible. In comparison to control plants, Panclfk plants had a significantly higher root biomass and lower shoot biomass, a proxy for problems with nitrogen fixation or infection level **(Figure 5A,B,C)**. However successfully infected cells are indistinguishable from EV-control, with Infection threads progressing into thinner fixation threads. Sectioning revealed that the structure of *Panclfk* nodules was inherently more disordered, with the presence of large apoplastic colonies **(Figure 5F,G)**. These phenotypes cannot be observed in any of the *crk* single mutants inoculated with *Bradyrhizobium* sp*,* KLD004. These *Pancrk1*, *Pancrk7* and *Pancrk11* mutant plants formed nodules as effectively as EV-control, which resulted in similar biomass and root shoot ratio when compared to control plants **(Figure S8A,B,C)**. Nodules formed on single *crk* mutant plants had densely packed infection zones and no visible defects **(Figure S8D-K)**.

Figure 6. Nodules of the cluster deletion mutant *Panclfk* have a low infection level. with *Bradyrhizobium* sp. KLD004. **A,** Absolute area in mm^2 of phenolic content, infected area and lobe size for *Panclfk* mutant nodules compared to empty vector control (EV-control). **B,** Phenolic content and infected area relative to nodule lobe size. Number of nodules analyzed per line is 16. Student's T-tests relative to EV-control samples. P>0.05 not significant (ns), P<0.05 *, P<0.01 ** P<0.001 ***P <0.0001 **C,** Nodule meristem of *Panclfk* mutant nodule with thick apoplastic colonies (yellow arrows), cell containing infection threads indicated by (ic) infected cells. Empty cell indicated by (ec) **D,** Nodule meristem of an EV-control nodule. Scale bars 500um zoom ins 250um.





Panclfk meristem region

EV-control meristem region

Next, we quantified the infection level of *PanClfk* nodules using nodule sections through central lobe tissue. Significant reduction was found for both the absolute and the relative infected area of *PanClfk* nodules versus EV-control. Interestingly no difference in lobe size or the area containing phenolic cells could be observed **(Figure 6A,B; Figure S9)**. This implies that *P. andersonii* CRKs are involved in regulating the infection process in cells derived from the nodule meristem. The meristematic region of *PanClfk* nodules contains thicker and shorter infection threads, which is not observed EV-control or in single mutants**(Figure 6C,D Figure S8H-K)**. Larger apoplastic colonies also occur in the infection zone. This suggest that in P. andersonii CRKs act redundantly to control infection thread progression in the apex of the nodule.

Discussion

We studied two large classes of transmembrane receptor kinases and evaluated their role in the rhizobium nitrogen-fixing nodule symbiosis in the non-legume *P. andersonii*. The involvement of two members *PanLEK1* and *PanCRK11* was predicted by genomic comparisons of the *Parasponia* and *Trema* lineage (R. van Velzen et al. 2018). Both these genes belong to diverse groups of transmembrane protein kinases, the Cysteine rich receptor kinases and Legume Lectin receptor like kinases. Both genes were studied by CRISPR-Cas9 mediated reverse genetics, which revealed novel symbiotic phenotypes. *Panlek1* knockout mutants make less nodules, and within these nodules patches of death cells are observed, suggesting a role for this transmembrane receptor kinase in nodule formation and maintenance. In contrast, *CRKs* act redundantly, and only by deleting a large cluster of 20 *CRK* genes, a role of these transmembrane receptor kinases in infection thread progression was revealed.

PanLEK1 is a member of a large gene family of L-type lectin receptor-like kinases, and is the only member of this clade with a significant nodule enhanced expression. Interestingly, closely related paralogous genes are specifically downregulated in nodules. PanLEK1 showed to have non conserved substitutions in its kinase domain and is therefore predicted to be inactive. Given the loss of *LEK1* in *Trema* species, it is likely that this gene neo-functionalized to fulfill a role in defense response repression or microsymbiont recognition. Based on the nodulation specific expression pattern and the downregulation of other Lectin receptor kinases, we argue that PanLEK1 has a specific role the promotion rhizobium nitrogen fixing symbiosis, which to date makes it the first L-type lectin kinase proven known to act in nitrogen-fixing nodule symbiosis. Clade-I L-type lectin receptor kinases have been functionally analysed also in *A. thaliana.* It was shown that AtLecRK-IX.1 and ALecRK-IX.2 are involved in resistance to *Phytophthora* spp, however kinase activity was indispensable for this

function (Wang et al. 2015). PanLEK1 may function as a dominant inhibitor of such defense responses controlled by lectin receptor kinases, due to its alternative kinase domain.

CRKs have been studied in several plant species -e.g. *A. thaliana*, common bean, soybean, and apple (*Malus x domestica*)- and in all cases they are encoded by a large gene family (Chen 2001; Delgado-Cerrone et al. 2018; Quezada et al. 2019; Zuo et al. 2019). We found that in *P. andersonii* a large number of the *CRK* genes is functioning in symbiotic context. *P.andersonii* contains 12 *CRK* genes that are significantly upregulated in nodules, all of which belong to the fast evolving variable cluster (Vaattovaara et al. 2019). The specific function of the *Parasponia*-specific and nodule expressed gene *PanCRK11* is less clear. We generated mutants of *PanCRK11* as well as of two other nodule enhanced *CRK* genes -*PanCRK1* and *PanCRK7-;* genes belonging to the three largest expanding variable taxonomic clades **(Figure 1)**. None of the single mutants has a clearly distinguishable phenotype in nodulation. Therefore we hypothesized that nodule expressed *P. andersonii CRK* genes are -at least in part- functionally redundant. We managed to create a knockout line in which a large chromosomal segment of ~1 Mbps containing 20 *CRK* genes was effectively removed. One of the largest CRISPR deletions reported in plants. This *Panclfk* complete cluster deletion mutant showed a phenotype in nitrogen-fixing nodule symbiosis. *Panclfk* mutant plants showed a reduced nodulation efficiency and were affected in intracellular infection..

Parasponia CRKs most likely do not function as pattern recognition receptors, as was suggested for MtSymCRK. PanCRK11 is related to *M. truncatula* SymCRK, since both belong to the same variable clade IX. Within clade IX MtSymCRK groups in a legume specific subclade. Only Papilionoideae legumes seem to share the alterations found the DUF26a motif and the kinase domain of MtSymCRK **(Figure 1)**. In *MtSymcrk / Mtdnf5* mutants bacteria are released but do not differentiate into symbiosomes. This developmental defect triggers major defense like responses in the *Mtsymcrk* allele, but this affect is less pronounced in the *Mtdnf5* allele (Oa et al. 2006; Domonkos et al. 2013; Berrabah et al. 2014; Lang, Smith, and Long 2018). *P. andersonii* houses its bacteria in fixation threads, which are still formed in *Pancflk* mutant nodules, though in a lower frequency. The reduced infection level coincides with larger apoplastic colonies and more cell death. However, the mutant does not trigger generic defense responses, like accumulation of phenolic compounds as seen in *Panlek1* mutants. The reduction in the number of nodules correlated with a lower number of infection threads in the nodule meristem, resulting in a lower infection level in the nodule. The results point to a slow progression of infection threads in the Panclfk nodule meristem and early infection stages. Therefore we hypothesize a role of *P. andersonii* CRKs in perceiving signals promoting infection thread progression.

These signals could be related to more generic signaling roles previously predicted for CRK proteins, such as perception of ROS. The low oxygen environment in nodules, created by hemoglobin, coupled to high respiratory demands leads to increased production of ROS (Wittenberg et al. 1986; Fukudome et al. 2016; Günther et al. 2007). At early stages of rhizobium infection production of ROS is coupled to the detection of rhizobium secreted lipo-chitooligosaccharide signal molecules (known as Nod factors) (Ramu, Peng, and Cook 2002), and hydrogen peroxide accumulation can be detected in the infection threads of *M. truncatula(Jamet et al. 2007)*. Also it was shown that the production of hydrogen peroxide is essential for rhizobium infection, as well as the differentiation of rhizobia into their symbiotic form (Jamet et al. 2003; Lambert et al. 2011; Montiel et al. 2012; Andrio et al. 2013; Arthikala et al. 2014). Although *Parasponia* relies on crack entry for initial rhizobial infection into the root, infection threads are ultimately formed in the nodule primordia and persist in the region basal to the nodule meristem. In peanut (*Arachis hypogaea*) ROS signals are also important for the crack entry infection process (Muñoz et al. 2015). Similar results were obtained for *Sesbania rostrata*, which can be infected by roo hair-based infection threads as well as crack entry(D'Haeze et al. 2003). Clearly the production of ROS and infection thread progression are extensively linked. *A. thaliana* AtCRK2, belonging to the basal cluster of CRKs, has been shown to interact and directly phosphorylate RESPIRATORY BURST OXIDASE HOMOLOG D (AtRBOHD) at the C-terminus to regulate ROS bursts in planta (Kimura et al. 2020). Suppression of the RBOH complex is required for rhizobial colonization in Medicago truncatula nodules (Yu et al. 2018). Given the large amounts of links of CRKs to ROS sensing in abiotic and biotic stress interactions in non symbiotic conditions, the co-option of CRKs to function in ROS signalling in symbiosis seems likely (Bourdais et al. 2015; Du et al. 2018; Idänheimo et al. 2014; Lee et al. 2017). Therefore a role for *P. andersonii* CRKs in ROS sensing and possibly regulating ROS production in the infection stage is probable.

Taken together, we conclude that *Parasponia* recruited *PanCRK11* into symbiotic signalling. And although a highly similar MtsymCRK has been recruited to regulate symbiosome differentiation in legumes, the role of CRKs in the non-legume *Parasponia* is clearly not equivalent.

Materials and methods

Phylogeny constructions

Orthogroups generated in a previous study containing Cysteine rich kinases were combined in a single database supplemented with *Solanum Lycopersicum, Cucumis sativus, Datisca glomerata*, *Alnus glutinosa* and *Phaseolus vulgaris* to create a large CRK dataset(Huang et al. 2009; Tomato Genome Consortium 2012; Griesmann et al. 2018; Quezada et al. 2019; R. van Velzen et al. 2018). For LEK1 a similar approach was conducted, this time orthogroups were supplemented with sequences from *Discaria trinervis*, *Alnus glutinosa* and *Dastisca glomerata (Griesmann et al. 2018)*. Blasts were carried out using a local Blast setup implemented in GeneiousR8.1.9 using each Parasponia orthogroup sequence. Kinase domains including transmembrane regions were extracted. Sequences were aligned using MafftV7.017. Phylogeny analysis was performed using IQ-tree1.6.12 running model finder to find the best substitution model (JTT+F+G+R4 for Cysteine rich kinases, JTT+R4 for Lectin-kinase) (Nguyen et al. 2015; Kalyaanamoorthy et al. 2017). UF-bootstrap was set to 1000 iterations and abayes support was calculated for both phylogenies(Hoang et al. 2018). For visual clarity only branches leading to clades are shown. Alignment visualizations and LOGOs were done using Geneious R8.1.9. Genomic regions of Parasponia and Trema belonging to the CRK cluster were aligned using Mummer3 (Kurtz et al. 2004).

Vector constructs

Binary constructs generated for this study were created using Golden gate cloning and backbones and several inserts were derived from the golden gate molecular toolbox(Engler et al. 2009, 2014). Domestication of parts by removal of BsaI or BpiI sites was carried out as described in Engler et al 2014. The generation and assembly of CRISPR-CAS9 constructs for Parasponia were done as published previously(van Zeijl et al. 2018; Wardhani et al. 2019). Vectors were verified using restriction digestion and sequencing of the sg target containing region.

Guide RNAs were designed using the CRISPR design tool implemented in Geneious R10, which is based on variables described in Doench et al 2014(Doench et al. 2014). Selected guide target sites have no potential off targets in the Parasponia genome with less than two indels or three mismatches. A list of constructs generated for this study is listed in **(Table S2)**.

Genotyping

CRISPR deletions were detected using Phire plant direct PCRs (Thermo fisher, F130WH) using manufacturer's recommendations. Target sites were amplified with spanning primers listed in **(Table S1)**. Potential deletion candidate shoots were re-genotyped during shoot propagation to ensure line homogeneity. Genotypes of

obtained lines are visualized in **(Figure S4)** and **(Figure S5)**. To validate the deletion for *panclfk* markers for all Parasponia *CRK* genes were amplified for presence absence analysis.

Bacterial strains

All golden gate cloning steps and plasmid propagations were carried out in *Escheria coli* DH5a. *Agrobacterium tumefaciens* AGL-1 was used for all *Parasponia* stable transformations. *Mesorhizobium plurifarium* BOR2 was used for *Parasponia* nodulation assays and inoculations(R. van Velzen et al. 2018). *Bradyrhizobium sp.* strain KLD004, derived from nodules of *Parasponia* growing at Mt.Kelud Indonesia, was used as secondary nodulation strain as effective fixer and native symbiont.

Plant growth conditions

Parasponia andersonii WU1 or its direct descendants were used as starting material for transformations. Prior to co cultivation *Parasponia* trees are grown in a greenhouse at 28°C, 85% humidity and a 16/8 h day/night. *Parasponia* in vitro tissue culture, CRISPR mutagenesis and rooting were done according to van Zejil et al 2018(van Zeijl et al. 2018; Wardhani et al. 2019) Rooted tissue culture plants for phenotyping were transferred to polypropylene containers 1L, fitted with a gas exchange lid. (OS140BOX, Duchefa Biochemie, Netherlands). Pots were half filled with 2/3rd Agroperlite(Maasmond-Westland, Netherlands) 1/3rd Sterilized Steamed river sand and watered with modified EKM medium (3 mM MES (C6H13NO4) pH 6.6, 2.08 mM MgSO4, 0.88 mM KH2PO4, 2.07 mM K2HPO4, 1.45 mM CaCl2, 0.70 mM Na2SO4, 0.375 mM NH4NO3, 15 µM Fe-citrate, 6.6 µM MnSO4, 1.5 µM ZnSO4, 1.6 µM CuSO4, 4 µM H3BO3, 4.1 µM Na2MoO4)(Becking 1983). Bacterial inoculation density was set to OD600 0.025.

Histochemical analysis

Nodules were harvested and fixed in 5% glutaraldehyde in 50mM sodium phosphate buffer PH7.4 Vacuum was applied for 2 hours on a total incubation time of 48h. Fixed nodules were embedded in plastic Technovit 7100 (Heraeus-Kulzer, Germany) to suppliers specifications. RJ2035 microtome was used for thin sections (5 µm), subsequently stained with 0.05% Toluidine Blue O. Images were made using a DM5500B microscope equipped with a DFC425c camera (Leica microsystems). For sectioning of the larger older nodules a different fixative containing 4% paraformaldehyde mixed with 3% glutaraldehyde in 50mM sodium phosphate buffer PH7.4 was used. Incubation time was increased to 7 days to fully immerse the large nodule samples.

Quantification and Statistical analysis

Infection levels were quantified using a custom ImageJ macro running in Fiji (Schindelin et al. 2012). The macro is based on differential thresholding in the Red/Green/Blue channels of the image to effectively generate binary selections. The binary selections can be converted into Regions Of Interest (ROI). Pixel lengths were scaled to microns based on image scale bar. Non colonizable areas such as vascular bundles were manually excised from the pictures to exclude them from the analysis. Graphs and statistical analysis were performed using R studio 1.1.456 replicate number is denoted in the figure legend. Statistical analysis was done using student's T-tests relative to EV-control samples. Significance values were denoted as P>0.05 NS, P<0.05 *, P<0.01 ** P<0.001 ***P <0.0001 ****. Levene's test for homogeneity of variance was used prior to running t-tests.

Additional Supplemental files belonging to this chapter (not included in this thesis, available on request

Supplemental File 1. Nexus file of Cysteine Rich receptor Kinase consensus Tree

Supplemental File 2. Fasta file of CRK clade IX alignment

Supplemental File 3. Nexus file of Lectin kinase consensus Tree

Supplemental File 4. Fasta file of Lectin RK clade I alignment

Supplemental File 5. Genbank sequence Parasponia CRK-cluster

Supplemental File 6. Genbank sequence of Trema CRK-cluster

Supplemental data

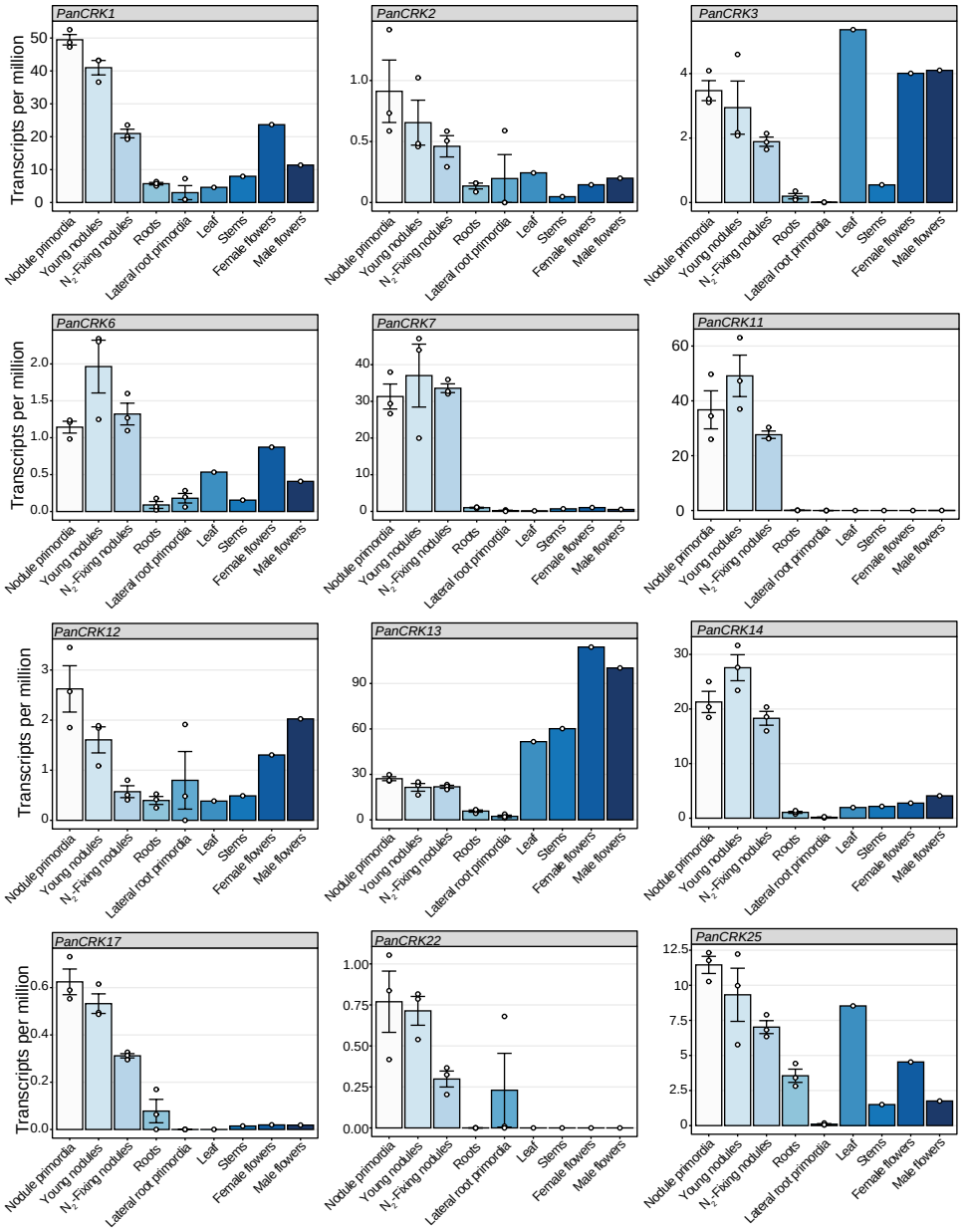


Figure S1. Expression profile of nodule enhanced *Parasponia andersonii CRK* genes. Analysis is based RNA-seq data presented in van Velzen et al 2018 (R. van Velzen et al. 2018). Expression in Transcrips per Million (TPM), replicates denoted by open dots.

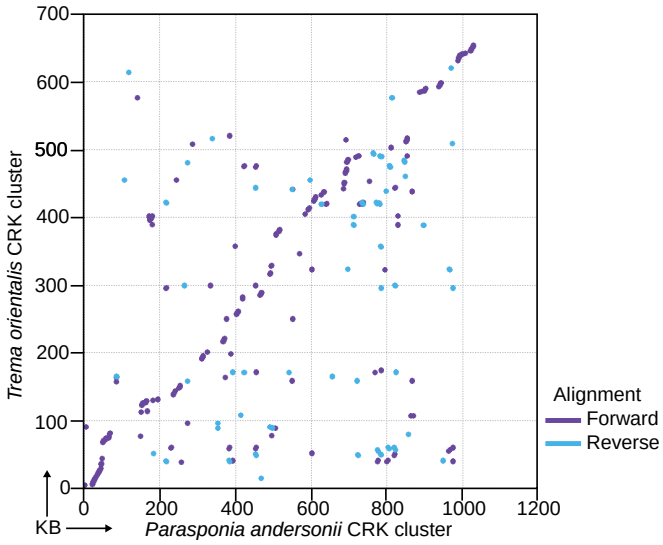


Figure S2. Dotplot of *CRK* cluster comparison of *Parasponia andersonii* and *Trema orientalis*. Alignment of CRK cluster regions made using Mummer3. The CRK cluster of both species is largely collinear. Increased size of *P. andersonii* is mainly due to an increased repeat content.

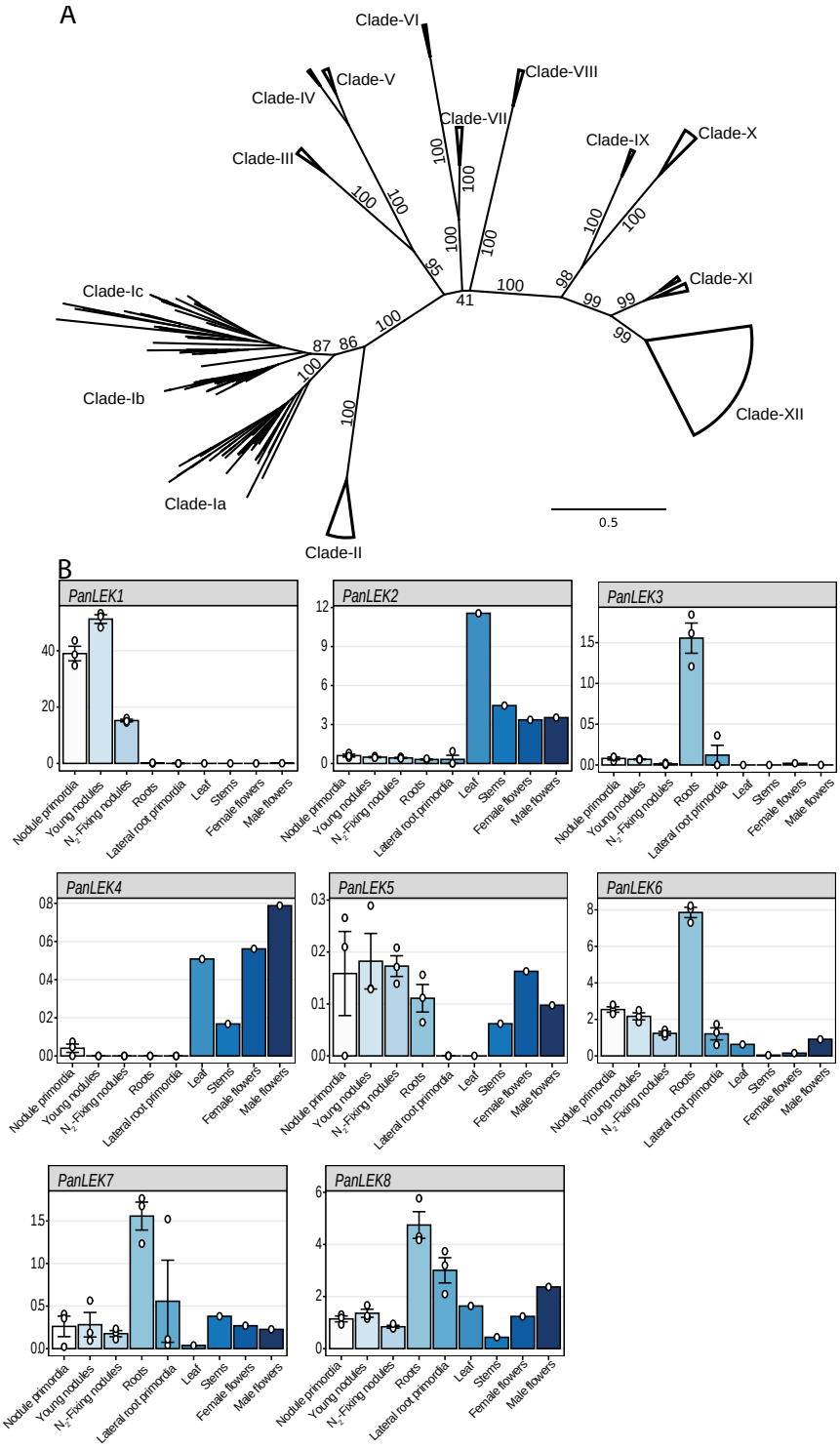

A
Clade-VI
Clade-V
Clade-IV
Clade-VII
Clade-VIII
Clade-III
Clade-IX
Clade-X
Clade-XI
Clade-Ic
Clade-Ib
Clade-Ia
Clade-II
Clade-XII
0.5
B
PanLEK1
PanLEK2
PanLEK3
PanLEK4
PanLEK5
PanLEK6
PanLEK7
PanLEK8
Nodule primordia
Young nodules
N2-Fixing nodules
Roots
Lateral root primordia
Leaf
Stems
Female flowers
Male flowers

Figure S3. Phylogeny L-type lectin receptor kinases and expression of *Parasponia andersonii* clade I *LEK* genes. **A,** Phylogenetic reconstruction, based on protein kinase alignments of L-type lectin receptor kinases, separates twelve distinct clades. PanLEK1 is part of in clade I (see also Supplemental file 3 and 4). **B,** Expression of *P. andersonii* clade I *LEK* genes. Only *PanLEK1* has a nodule specific expression profile.

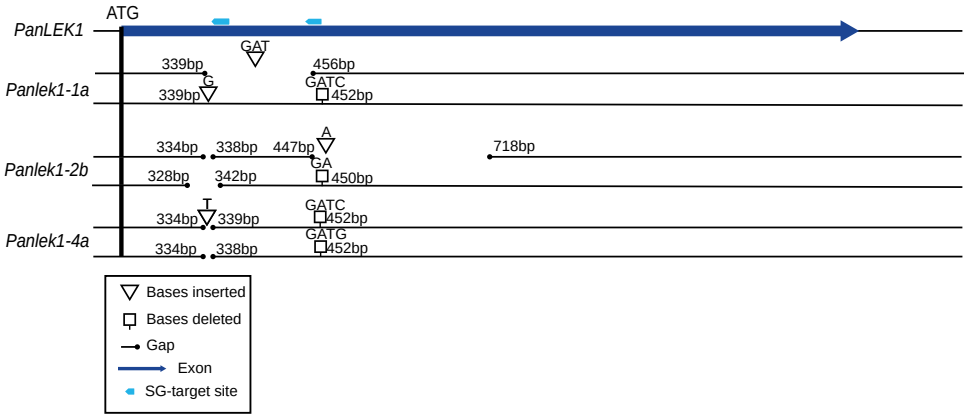


Figure S4. Gene model and genotype of *Panlek1* mutant lines. *PanLEK1* encodes a single exon gene. Position of CRISPR target sites is indicated by light blue blocks. *Panlek1* T_o -mutants represent biallelic mutations with small and larger deletions at the target regions. Mutant lines were genotyped and sequenced with primer pair Ru_296 **(Table S1)**.

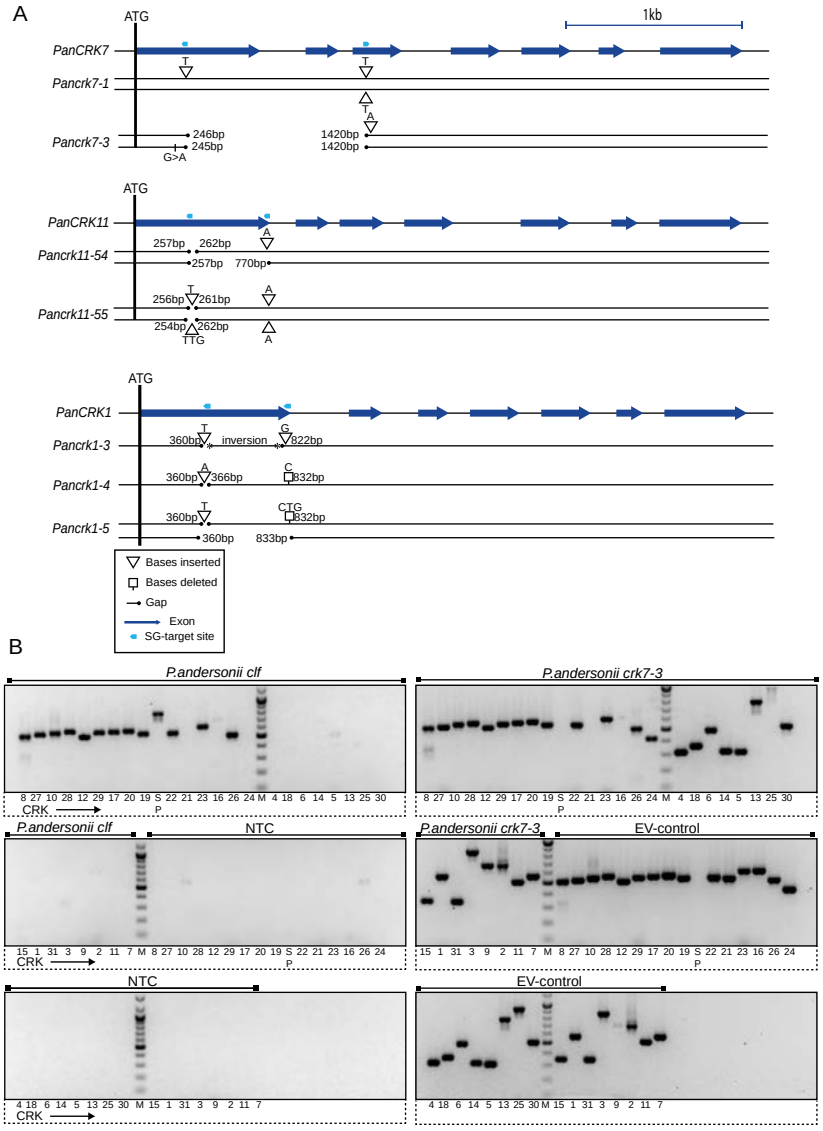


Figure S5. Gene model and genotype of *Pancrk* mutants. Position of CRISPR target sites is indicated by light blue blocks. **A,** *Pancrk11*, *Pancrk7-3* and *Pancrk1-5* Tmutants carry bi-allelic mutations, with different small insertions or deletions at or between the target sites. *Pancrk1-3* and *Pancrk1-4* carry homozygous mutations. **B,** Additional genotyping on the Panclfk deletion line that lacks *Pancrk21, PanCRK16, PanCRK24, PanCRK4, PanCRK6, PanCRK5, PanCRK25, PanCRK18, PanCRK14, PanCRK13, PanCRK20, PanCRK15, PanCRK31, PanCRK11, PanCRK9, PanCRK13, PanCRK1, PanCRK2* and *PanCRK7*. PCR markers for all *PanCRK* genes were tested on genomic DNA of *Panclfk*, *Pancrk7-3* and EV-control. Note the absence of products in the *Panclfk* mutant and the presence of the cluster spanning PCR amplicon SP (primer pair Ru_130, **Table S1**). The SP amplicon cannot be amplified in the *Pancrk7-3* mutant line and EV-control.

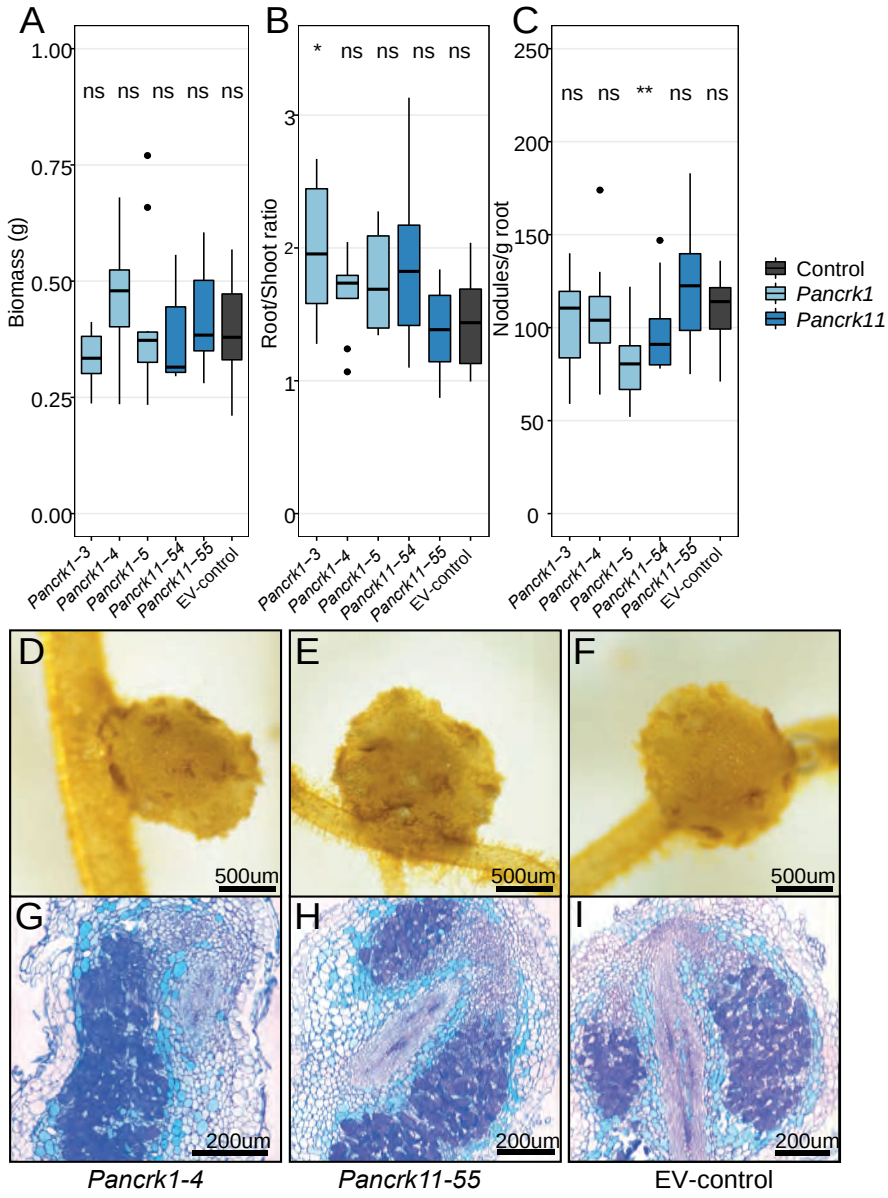


Figure S6. Phenotype of *Pancrk1* and *Pancrk11* single mutants 4 weeks post inoculation with *M.plurifarium* BOR2. (**A-C**) Box plots of biomass (A)**,** root shoot ratio (B), and nodule number / g. root weight (C). EV-control n=10, *Pancrk1-3* n=10, *Pancrk1-5* n=10, *Pancrk11-5*4 n=10, *Pancrk11=55* n=10. Student's T-tests relative to empty vector control (EV-control) samples. P>0.05 not significant (ns), P<0.05 *, P<0.01 ** P<0.001 ***P <0.0001 (**D),** *Pancrk1-4* mutant nodule scale bar 500 µm (**E),** *Pancrk11-55* mutant nodule scale bar 500 µm (**F),** EV-control nodule scale bar 500 µm (**G,H,I)** Representative nodule sections, scale bars 200 µm. **G,** *Pancrk1-4* (**H),** *Pancrk11-55* and (**I),** EV-control.

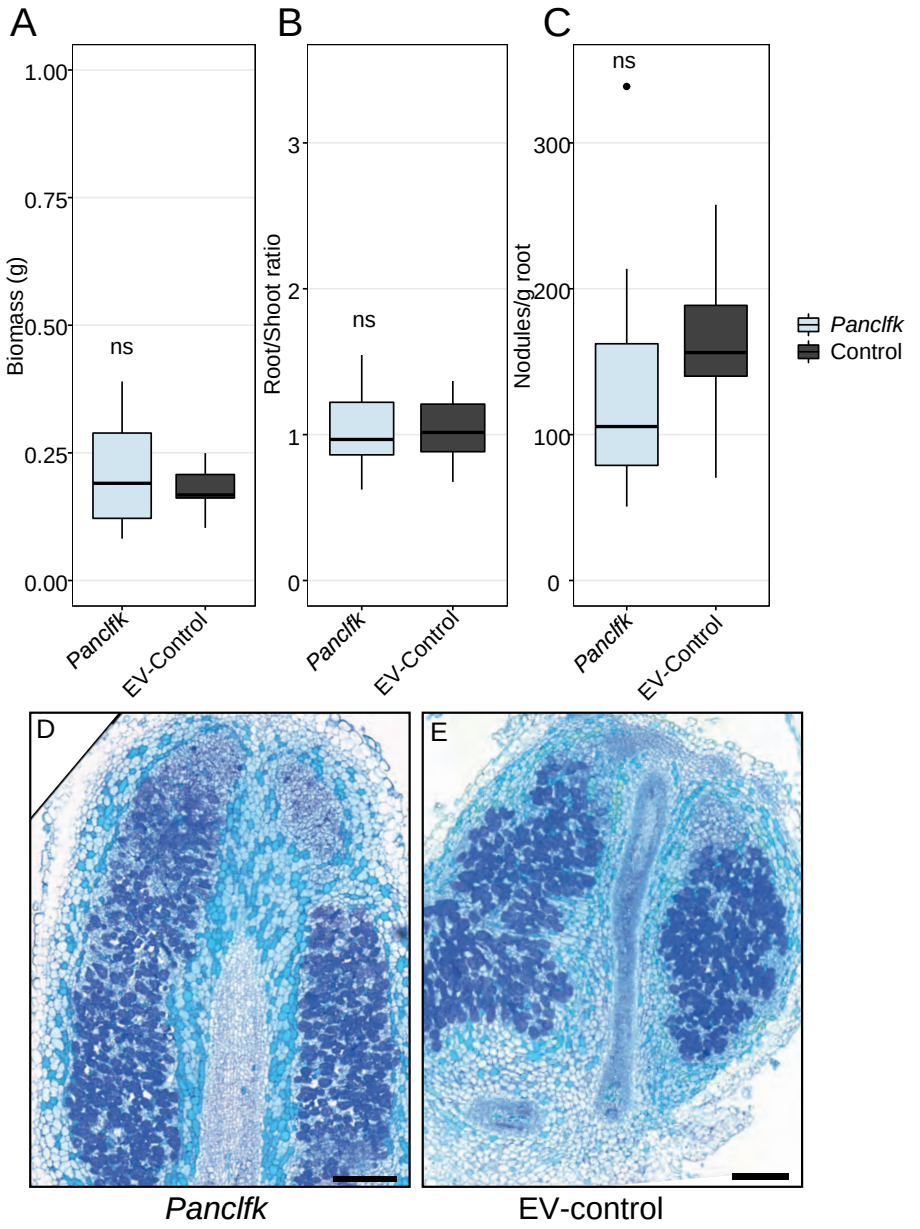


Figure S7. Phenotype of Panclfk CRK cluster deletion mutant 4 weeks post inoculation with *M. plurifarium* BOR2. (**A-C**) Box plots of biomass (A)**,** root shoot ratio (B), and nodule number / g. root weight (C)**.** EV-control n=10, *Panclfk* n=9 T-tests relative to empty vector control (EV-control) samples. P>0.05 not significant (ns), P<0.05 *, P<0.01 ** P<0.001 ***P <0.0001 **D,** representative section of *Panclfk* deletion mutant nodule, scale bar 500 µm **E,** EV-control nodule scale bar 500 µm.

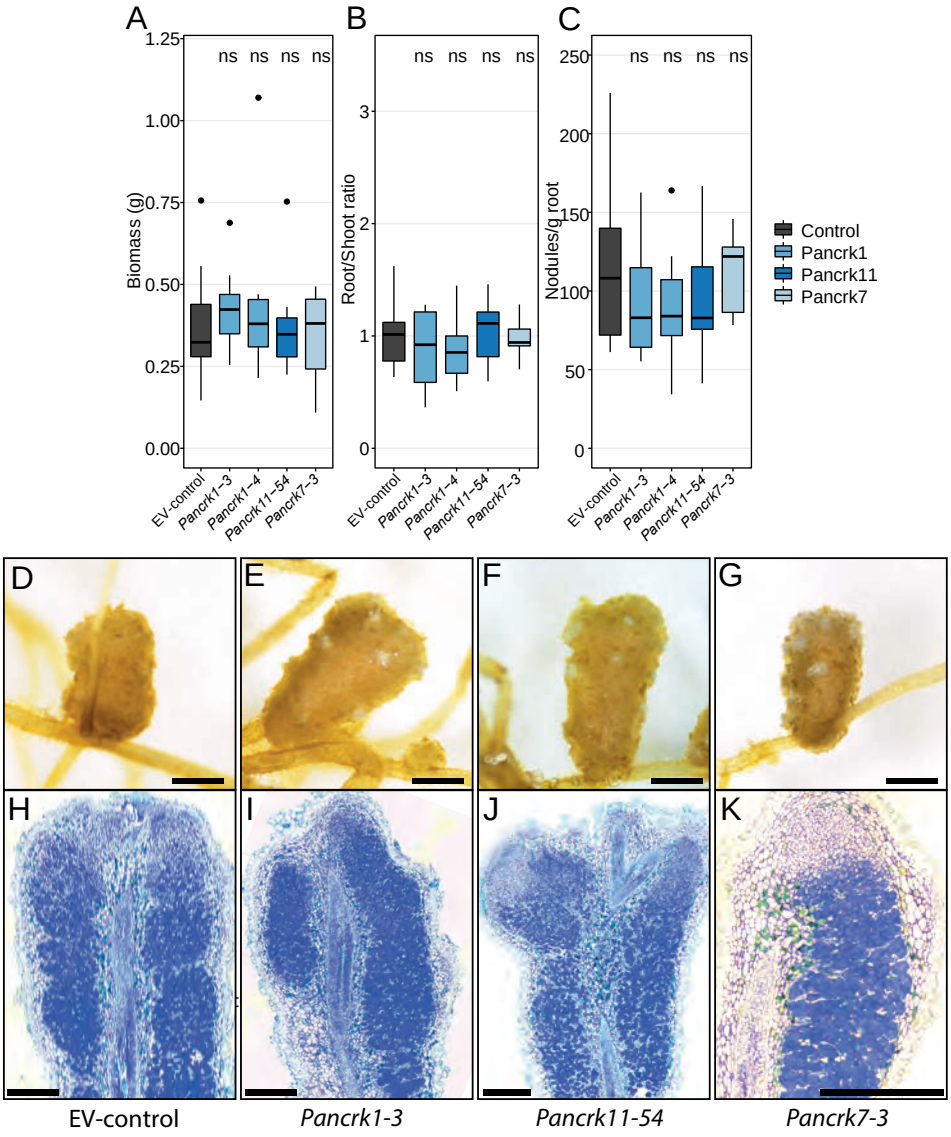


Figure S8. Phenotype of Pancrk1, *Pancrk7* and *Pancrk11* single mutants. 8 weeks post inoculation with *Bradyrhizobium* sp. KLD004. (**A-C**) Box plots of biomass (A)**,** root shoot ratio (B), and nodule number / g. root weight (C). EV-control n=12, *Pancrk1-3* n=12*, Pancrk1-4* n=11*, Pancrk11-54* n=13*, Pancrk7-3* n=9. Student's T-tests relative to EV control samples. P>0.05 not significant (ns), P<0.05 *, P<0.01 ** P<0.001 ***P <0.0001 **D,** EV-control nodule scale bar 500 µm **E,** *Pancrk1-3* mutant nodule scale bar 500 µm **F,** *Pancrk11-54 mutant* nodule scale bar 500 µm **G,** *Pancrk7-3* mutant nodule, scale bar 500 µm, (**H,I,J,K)** Representative nodule sections, scale bars 500 µm. **H,** EV-control **I,** *Pancrk1-3* **J,** *Pancrk11-54* **K,** *Pancrk7-3*

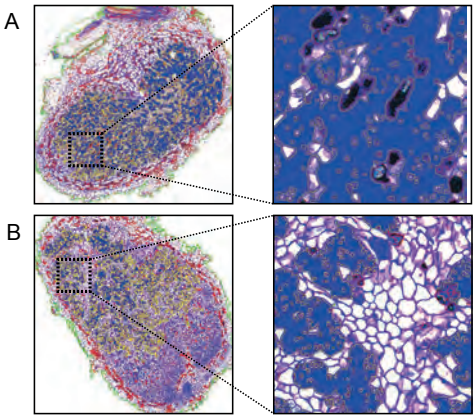


Figure S9. Selections made by custom IMAGEJ macro for quantification of infection level in nodule sections. Green selection: total nodule lobe, Yellow selection: infected cell, Red selection: phenolic cells. (**A),** EV-control nodule with zoom in. (**B),** *Panclfk* deletion mutant nodule with zoom in.

Supplemental table 1. Primers used for this study

Primer code	Primer name	Purpose	5'3'-sequence	Size
Ru_230	PanCRK26m_F	Genotyping clfk	CAAGGATTGCGTGGCATTCG	503
Ru_231	PanCRK26m_R	Genotyping clfk	CCTTGTTGATGGAGGCGGT	
Ru_232	PanCRK23m_F	Genotyping clfk	GTCCATAAAAGGGGTAAACTTCGT	600
Ru_233	PanCRK23m_R	Genotyping clfk	GAGTCCTCGCCAACCTCAC	
Ru_234	Pancrk8m_F	Genotyping clfk	AAGATGGAAGGGAAGTGGCC	511
Ru_235	Pancrk8m_R	Genotyping clfk	ACCCTTGCCATTCCGAAGTT	
Ru_236	Pancrk21m_F	Genotyping clfk	ACCGACTGTTCATCTGGCAG	527
Ru_237	Pancrk21m_R	Genotyping clfk	CAGTGATGTCCCAGGTGCAT	
Ru_238	PanCRK16m_F	Genotyping clfk	GCGTCAGCGAGCTCGTGA	599
Ru_239	PanCRK16m_R	Genotyping clfk	TCCAAAGTTTTTCTCTCTAATCTTTGC	
Ru_240	PanCRK24m_F	Genotyping clfk	GCTCCCACGACTGTCAAGAA	414
Ru_241	PanCRK24m_R	Genotyping clfk	GGCGAAAAGAACAGCAAGCA	
Ru_242	PanCRK4m_F	Genotyping clfk	TTCCACGGCCAGAGAAGTTC	307
Ru_243	PanCRK4m_R	Genotyping clfk	GAACCACCCGATTGACCCTT	
Ru_244	PanCRK6m_F	Genotyping clfk	TGCTTGCTGTTCTTTTCGCC	479
Ru_245	PanCRK6m_R	Genotyping clfk	AGCCACCACAACCTTCCAAA	
Ru_246	PanCRK5m_F	Genotyping clfk	TCCTCCAGCGATGACTCTGA	305
Ru_247	PanCRK5m_R	Genotyping clfk	AAAGCCGCCTTCTCAGTCAA	
Ru_248	PanCRK25m_F	Genotyping clfk	TGCAAGCCAAGAAGTCACCA	926
Ru_249	PanCRK25m_R	Genotyping clfk	CGCTAATGCTGAAGACGTGC	
Ru_250	PanCRK18m_F	Genotyping clfk	TCCAGGGAACGCTTCTGAAC	352
Ru_251	PanCRK18m_R	Genotyping clfk	TGTTTCTTGGGGTCTGCACA	
Ru_252	PanCRK14m_F	Genotyping clfk	TATAACGTCACTCGAGCCGC	313
Ru_253	PanCRK14m_R	Genotyping clfk	TATCACGGCCTCCTTCTCCA	
Ru_254	PanCRK13m_F	Genotyping clfk	TCGCCACCAGAGAAGTCAAC	768
Ru_255	PanCRK13m_R	Genotyping clfk	TTGAAAAGTCTACCGCCCCC	
Ru_256	PanCRK20m_F	Genotyping clfk	AACAGAGCGATTCAACGGGT	506
Ru_257	PanCRK20m_R	Genotyping clfk	CACCACACCAATTGGCACAG	
Ru_258	PanCRK15m_F	Genotyping clfk	TCTTGGAGAAGCCGTTGGTC	348
Ru_259	PanCRK15m_R	Genotyping clfk	CCAACAAGACCACGGCAATG	
Ru_260	PanCRK31m_F	Genotyping clfk	TCGAAAGTGCTCTTGCTCGT	346

Ru_261	PanCRK31m_R	Genotyping clfk	GCAACACCACGAAGGCAATT	
Ru_262	PanCRK11m_F	Genotyping clfk	TGCATGCTGTTATGGAGGCA	505
Ru_263	PanCRK11m_R	Genotyping clfk	TGCCGCTACATATCACCCAA	
Ru_264	PanCRK9m_F	Genotyping clfk	CTCGTCGGGACAAGAATCCG	683
Ru_265	PanCRK9m_R	Genotyping clfk	ACCCGTAAATGCTTCGGTTT	
Ru_266	PanCRK3m_F	Genotyping clfk	TGGCAGCTGCTTCTTCTTCT	850
Ru_267	PanCRK3m_R	Genotyping clfk	ACTACAGGCGGTAAAGCTGG	
Ru_268	PanCRK1m_F	Genotyping clfk	CGTCCTCTCCACCCTTGTTC	564
Ru_269	PanCRK1m_R	Genotyping clfk	CAACTGCGAGCGATGAGTTG	
Ru_270	PanCRK2m_F	Genotyping clfk	AGCCACTAATCATCATCAAGCCA	710
Ru_271	PanCRK2m_R	Genotyping clfk	CGTAACAACTGCGAGCGATGATTC	
Ru_272	PanCRK7m_F	Genotyping clfk	CTCGGACGCTCTATCCACAC	562
Ru_273	PanCRK7m_R	Genotyping clfk	TCCATCAACAGCCCCAAACA	
Ru_274	PanCRK17m_F	Genotyping clfk	TACGGCAGCTACCTTTGCAA	550
Ru_275	PanCRK17m_R	Genotyping clfk	GCGGAGGGGGTGGAAAAATA	
Ru_276	PanCRK19m_F	Genotyping clfk	AGCCTGAATTCGCTCGAACA	529
Ru_277	PanCRK19m_R	Genotyping clfk	CCACCTCCTCTCCACCCTTA	
Ru_278	PanCRK12m_F	Genotyping clfk	ATTCGCCTCTCTCGTTGCAA	503
Ru_279	PanCRK12m_R	Genotyping clfk	TTCTTCTCGTCGGCCTTGTC	
Ru_280	PanCRK2m_F	Genotyping clfk	TGGCCTTCTCCAATGCAGAG	534
Ru_281	PanCRK2m_R	Genotyping clfk	ATAGGTGCTGCTTGAGGTGG	
Ru_282	PanCRK10m_F	Genotyping clfk	ACTGTAACGAAGGCGCTTCA	538
Ru_283	PanCRK10m_R	Genotyping clfk	TGGGTCAAATCAGGTGTGCA	
Ru_284	PanCRK27m_F	Genotyping clfk	CATTCTTTGGCGCTTCCGAG	520
Ru_285	PanCRK27m_R	Genotyping clfk	TCGTTGACGCTCTTTCTGCT	
Ru_286	PanCRK28m_F	Genotyping clfk	CGTGCACAAACGGTCAACAT	551
Ru_287	PanCRK28m_R	Genotyping clfk	GTCTTCGAAGCGTCCTCCAA	
Ru_288	PanCRK29m_F	Genotyping clfk	AGAACATCACGGACAGCGTT	540
Ru_289	PanCRK29m_R	Genotyping clfk	AGCACCCTGTAAACATGGCT	
Ru_290	PanCRK30m_F	Genotyping clfk	TGTCTCTCCGACCCGAGAAT	562
Ru_291	PanCRK30m_R	Genotyping clfk	AACCCTCACCACGAACAGTC	
Ru_130	seq_CRKcluster-SP_F	Genotyping clfk	TCGTGTTCCCAGGTAGTTACC	
Ru_130	seq_CRKcluster-SP_R	Genotyping clfk	AGCTTTCCCTGCCATTTCGA	
Ru_296	Seq_LEK1_F	Genotyping LEK1	GGGACATCTCATCAGGAAGGC	

Ru_296	Seq_LEK1_R	Genotyping LEK1	CCGCGAGAAAAAATACAAAG-TAAATAC
Ru_292	seq_crk1_F	Genotyping CRK1	CGTCCTCTCCACCCTTGTTC
Ru_293	seq_crk1_R	Genotyping CRK1	TCTTCCCCTGCCTCTGCTAT
Ru_294	seq_crk11_F	Genotyping CRK11	GCGTGCCAATTCTCGTCTTC
Ru_295	seq_crk11_R	Genotyping CRK11	TGGAGTGCAGAAAGCTACCG
Ru_96	seq_crk7_F	Genotyping CRK7	CTCGGACGCTCTATCCACAC
Ru_96	seq_crk7_R	Genotyping CRK7	CCACCTTCTCCGAGCTTGTT
Ru_23	seq_lvl0_F	Contruct validation	TACCGCCTTTGAGTGAGCTG
Ru_23	seq_lvl0_R	Contruct validation	GTCTCATGAGCGGATACATATTT-GAATG
Ru_24	seq_lvl1_F	Contruct validation	GAACCCTGTGGTTGGCATGCACA-TAC
Ru_24	seq_lvl1_R	Contruct validation	CTGGTGGCAGGATATATTGTGGTG
Ru_25	seq_lvl2_F	Contruct validation	GTGGTGTAAACAAATTGACGC
Ru_25	seq_lvl2_R	Contruct validation	GGATAAACCTTTTCACGCCC
Ru_73	Kanamycin_F	transformation validation	AAAAGCGGCCATTTTCCACC
Ru_73	Kanamycin_R	transformation validation	GATGGATTGCACGCAGGTTC
Ru_74	AtCAS9_mF	transformation validation	TTCGATCTCGCTGAGGATGC
Ru_74	AtCAS9_mR	transformation validation	TAGCGAGAGGTCCCACGTAG
sgRNA_rev		CRISPR assembly	tgtggtctcaAGCGTAATGCCAACTTTGTAC
CRK7sgRNA1		CRISPR assembly	tgtggtctcaattGCGTTACTGGAAACATCACCTgttttagagctagaaatagcaag
CRK7sgRNA2		CRISPR assembly	tgtggtctcaattGTTTACCGATAGAAGTATGCAgttttagagctagaaatagcaag
CRK21sgRNA1		CRISPR assembly	tgtggtctcaattGAGGTATTGTAGAAACCTCCGgttttagagctagaaatagcaag
CRK21sgRNA2		CRISPR assembly	tgtggtctcaattGCCGGATACAGTTCCGGCAAAgtttttagagctagaaatagcaag
CRK1sgRNA1		CRISPR assembly	tgtggtctcaattGCGTATGTCTGCACCTGATTGgttttagagctagaaatagcaag
CRK1sgRNA2		CRISPR assembly	tgtggtctcaattGTTGTGGTACTAGCCAGTGGgttttagagctagaaatagcaag
CRK11sgRNA1		CRISPR assembly	tgtggtctcaattGCAACTCTGGCAATGCGCAGgttttagagctagaaatagcaag
CRK11sgRNA2		CRISPR assembly	tgtggtctcaattGTTACCTGAATAACTAGTTGGgttttagagctagaaatagcaag
Lek1sgRNA1		CRISPR assembly	gtggtctcaattGCTAGCGTTGTCCGTTACAAGgttttagagctagaaatagcaag
Lek1sgRNA2		CRISPR assembly	gtggtctcaattGACGTGTTCTCGAAGATCGGGgttttagagctagaaatagcaag

Supplemental table2 part 1: Level 1 modules generated for this study

ID	Backbone	Purpose	L2 position	Antibiotic resistance	promoter + utr	gene	tag/fusion (Nt-Ct)	terminator
MOB1001	pICH47802	plant selection	R1	Ampicillin	Pro Atu Nos + 5U TMVΩ	NPT-II		Ter Atu Ocs
MOB1093	pICH47751	CRISPR	F3	Ampicillin	pICSL01009:AtU6p,	CRK7sg1 PCR product		
MOB1094	pICH47761	CRISPR	F4	Ampicillin	pICSL01009:AtU6p,	CRK7sg2 PCR product		
MOB1095	pICH47772	CRISPR	F5	Ampicillin	pICSL01009:AtU6p,	CRK21sg1 PCR product		
MOB1096	pICH47791	CRISPR	F6	Ampicillin	pICSL01009:AtU6p,	CRK21sg2 PCR product		
MOB1191	pICH47822	CRISPR	R3	Ampicillin	pICSL01009:AtU6p,	CRK1sg1PCR product		
MOB1192	pICH47831	CRISPR	R4	Ampicillin	pICSL01009:AtU6p,	CRK1sg2 PCR product		
MOB1193	pICH47822	CRISPR	R3	Ampicillin	pICSL01009:AtU6p,	CRK11sg1 PCR product		
MOB1199	pICH47751	CRISPR	F3	Ampicillin	pICSL01009:AtU6p,	PanLEKsg1 PCR product		
MOB1200	pICH47761	CRISPR	F4	Ampicillin	pICSL01009:AtU6p,	PanLEK1sg2 PCR product		
MOB1194	pICH47831	CRISPR	R4	Ampicillin	pICSL01009:AtU6p,	CRK11sg2 PCR product		
MOB1190	pICH47742	CRISPR	F2	Ampicillin	pICH41388::p35S	AtCAS9	pAGM5331 NLS	pICH41414 t35S

Supplemental table 2 part 2. Level 2 Binary constructs used in this study

ID	Purpose	Backbone	Res	Pos1	Pos2	Pos3	Pos4	**Pos5**	**Pos6**
MOB2084	CRISPR EV-control	pICSL4723	kan	pNOS::NP-TII:tOCS MOB1001	35S::NLS-aCas9-MOB1190				
EC74839	CRISPR panCRK1	pICSL4723	kan	pNOS::NP-TII:tOCS MOB1001	35S::NLS-aCas9-MOB1190	MOB1191 CRK1sg1	MOB1192 CRK1sg2		
EC74840	CRISPR Pan-CRK11	pICSL4723	kan	pNOS::NP-TII:tOCS MOB1001	35S::NLS-aCas9-MOB1190	MOB1193 CRK11sg1	MOB1193 CRK11sg2		
MOB2101	CRISPR PanCRK cluster	pICSL4723	kan	pNOS::NP-TII:tOCS MOB1001	35S::NLS-aCas9-MOB1190	MOB1091 CRK7sg1	MOB1092 CRK7sg2	MOB1093 CRK21sg1	MOB1094 CRK21sg1
MOB2180	CRISPR PanLEK1	pICSL4723	kan	pNOS::NP-TII:tOCS MOB1001	35S::NLS-aCas9-MOB1190	MOB1199 LEK1sg1	MOB1200 LEK1sg2		

Supplemental Table 3. Gene IDs of P.andersonii and T.orientalis CRK genes

Gene	CRK-type	Pan I.D.	Tor I.D.	Clade
CRK1	Variable, clustered	PanWU01x14_284910	Tor-RG33x02_160020	VII
CRK2	Variable, clustered	PanWU01x14_367340	Tor-RG33x02_160040	VII
CRK3	Variable, clustered	PanWU01x14_284950	Tor-RG33x02_160000	VII
CRK4	Variable, clustered	PanWU01x14_292600	Tor-RG33x02_188500	VIII
CRK5	Variable, clustered	PanWU01x14_292640	Tor-RG33x02_188480	VIII
CRK6	Variable, clustered	PanWU01x14_292610	Tor-RG33x02_188490	VIII
CRK7	Variable, clustered	PanWU01x14_121600	Tor-RG33x02_160050	VIII
CRK8	Variable, clustered	PanWU01x14_292550_ps	Tor-RG33x02_188540	VII
CRK9	Variable, clustered	PanWU01x14_361810	Tor-RG33x02_159910	VIII
CRK10	Variable, non clustered	PanWU01x14_342760	Tor-RG33x02_328350	IX
CRK11	Variable, clustered	PanWU01x14_285030	ND	IX
CRK12	Variable, non clustered	PanWU01x14_171080	Tor-RG33x02_304800	IX
CRK13	Variable, clustered	PanWU01x14_356760	Tor-RG33x02_159720	IX
CRK14	Variable, clustered	PanWU01x14_356740	Tor-RG33x02_159700	IX
CRK15	Variable, clustered	PanWU01x14_285100	Tor-RG33x02_159800	IX
CRK16	Variable, clustered	PanWU01x14_292580	Tor-RG33x02_188520	VIII
CRK17	Variable, non clustered	PanWU01x14_350690	Tor-RG33x02_198740	VIII
CRK18	Variable, clustered	PanWU01x14_292700	Tor-RG33x02_159660	VII
CRK19	Variable, non clustered	PanWU01x14_252840	Tor-RG33x02_285020	IX
CRK20	Variable, clustered	PanWU01x14_364320	Tor-RG33x02_159760	VIII
CRK21	Variable, clustered	PanWU01x14_292570_ps	Tor-RG33x02_188530	VI
CRK22	Variable, non clustered	PanWU01x14_236300	Tor-RG33x02_310050	VII
CRK23	Variable, non clustered	PanWU01x14_301470	Tor-RG33x02_063470	VIII
CRK24	Variable, clustered	PanWU01x14_292590	Tor-RG33x02_188510	VIII
CRK25	Variable, clustered	PanWU01x14_292690	Tor-RG33x02_188470	VIII
CRK26	Variable, non clustered	PanWU01x14_240930	Tor-RG33x02_272860	III
CRK27	basal	PanWU01x14_263500.1	Tor-RG33x02_052430	IV
CRK28	basal	PanWU01x14_263470.1	Tor-RG33x02_052390	II
CRK29	basal	PanWU01x14_066090.1	Tor-RG33x02_287430	I
CRk30	basal	PanWU01x14_208650.1	Tor-RG33x02_199880	IX
CRK31	Variable, clustered	PanWU01x14_285040	Tor-RG33x02_159870	VII
CRK32	Variable, clustered	ND	Tor-RG33x02_159920	

Supplemental Table 4. Gene IDs *P.andersonii* and *T.orientalis* Lectin kinases

Gene	Pan I.D	Tor I.D.	Clade
LEK1	PanWU01x14_069780.1	-	IA
LEK2	PanWU01x14_069800.1	TorRG33x02_104060.1	IA
LEK3	PanWU01x14_358790.1	-	IC
LEK4	PanWU01x14_361850.1	TorRG33x02_071330.1	IC
LEK5	PanWU01x14_361860.1	TorRG33x02_057420.1	IB
LEK6	PanWU01x14_086330.1	TorRG33x02_057430.1	IB
LEK7	PanWU01x14_086340.1	TorRG33x02_071320.1	IB
LEK8	PanWU01x14_086360.1	TorRG33x02_057400.1	IB
LEK9	-	TorRG33x02_057390.1	IB
LEK10	-	TorRG33x02_057360.1	IB
LEK11	PanWU01x14_266520.1	TorRG33x02_329060.1	III
LEK12	PanWU01x14_266510.1	TorRG33x02_329040.1	III
LEK13	PanWU01x14_266530.1	TorRG33x02_329070.1	III
LEK14	PanWU01x14_045430.1	TorRG33x02_320830.1	IV
LEK15	PanWU01x14_045450.1	TorRG33x02_320810.1	IV
LEK16	PanWU01x14_074010.1	TorRG33x02_067090.1	V
LEK17	PanWU01x14_066640.1	TorRG33x02_074270.1	IX
LEK18	PanWU01x14_166220.1	TorRG33x02_041870.1	X
LEK19	PanWU01x14_228390.1	TorRG33x02_246520.1	X
LEK20	PanWU01x14_351380.1	TorRG33x02_134670.1	XI
LEK21	PanWU01x14_256070.1	TorRG33x02_269950.1	XI
LEK22	PanWU01x14_351370.1	TorRG33x02_134680.1	XII
LEK23	PanWU01x14_281460.1	TorRG33x02_098010.1	XII
LEK24	PanWU01x14_089790.1	TorRG33x02_023970.1	VII
LEK25	PanWU01x14_036990.1	TorRG33x02_044150.1	VI
LEK26	PanWU01x14366190.1	TorRG33x02_241560.1	VIII

References

André, Sabine, Hans-Christian Siebert, Mitsuru Nishiguchi, Kiyoshi Tazaki, and Hans-Joachim Gabius. 2005. “Evidence for Lectin Activity of a Plant Receptor-like Protein Kinase by Application of Neoglycoproteins and Bioinformatic Algorithms.” *Biochimica et Biophysica Acta (BBA) - General Subjects* 1725 (2): 222–32.

Andrio, Emilie, Daniel Marino, Anthony Marmeys, Marion Dunoyer de Segonzac, Isabelle Damiani, Andrea Genre, Stéphanie Huguet, Pierre Frendo, Alain Puppo, and Nicolas Pauly. 2013. “Hydrogen Peroxide-Regulated Genes in the Medicago Truncatula-Sinorhizobium Meliloti Symbiosis.” *The New Phytologist* 198 (1): 179–89.

Arthikala, Manoj Kumar, Rosana Sánchez-López, Noreide Nava, Olivia Santana, Luis Cárdenas, and Carmen Quinto. 2014. “RbohB, a Phaseolus Vulgaris NADPH Oxidase Gene, Enhances Symbiosome Number, Bacteroid Size, and Nitrogen Fixation in Nodules and Impairs Mycorrhizal Colonization.” *The New Phytologist* 202 (3): 886–900.

Becking, J. H. 1983. “The Parasponia Parviflora -Rhizobium Symbiosis. Host Specificity, Growth and Nitrogen Fixation under Various Conditions.” *Plant and Soil* 75 (3): 309–42.

Bellande, Kevin, Jean Jacques Bono, Bruno Savelli, Elisabeth Jamet, and Hervé Canut. 2017. “Plant Lectins and Lectin Receptor-like Kinases: How Do They Sense the Outside?” *International Journal of Molecular Sciences* 18 (6). https://doi.org/10.3390/ijms18061164.

Berrabah, Fathi, Thierry Balliau, El Hosseyn Aït-Salem, Jeoffrey George, Michel Zivy, Pascal Ratet, and Benjamin Gourion. 2018. “Control of the Ethylene Signaling Pathway Prevents Plant Defenses during Intracellular Accommodation of the Rhizobia.” *The New Phytologist* 219 (1): 310–23.

Berrabah, Fathi, Marie Bourcy, Alexis Eschstruth, Anne Cayrel, Ibtissem Guefrachi, Peter Mergaert, Jiangqi Wen, et al. 2014. “A nonRD Receptor-like Kinase Prevents Nodule Early Senescence and Defense-like Reactions during Symbiosis.” *The New Phytologist* 203 (4): 1305–14.

Bourdais, Gildas, Paweł Burdiak, Adrien Gauthier, Lisette Nitsch, Jarkko Salojärvi, Channabasavangowda Rayapuram, Niina Idänheimo, et al. 2015. “Large-Scale Phenomics Identifies Primary and Fine-Tuning Roles for CRKs in Responses Related to Oxidative Stress.” *PLoS Genetics* 11 (7): 1–36.

Bouwmeester, Klaas, and Francine Govers. 2009. “Arabidopsis L-Type Lectin Receptor Kinases: Phylogeny, Classification, and Expression Profiles.” *Journal of Experimental Botany* 60 (15): 4383–96.

Chen, Z. 2001. “A Superfamily of Proteins with Novel Cysteine-Rich Repeats.” *Plant Physiology* 126 (2): 473–76.

Choi, Jeongmin, Kiwamu Tanaka, Yan Liang, Yangrong Cao, Sang Yeol Lee, and Gary Stacey. 2014. “Extracellular ATP, a Danger Signal, Is Recognized by DORN1 in Arabidopsis.” *Biochemical Journal* 463 (3): 429–37.

Dardick, Chris, Benjamin Schwessinger, and Pamela Ronald. 2012. “Non-Arginine-Aspartate (non-RD) Kinases Are Associated with Innate Immune Receptors That Recognize Conserved Microbial Signatures.” *Current Opinion in Plant Biology*. https://doi.org/10.1016/j.pbi.2012.05.002.

Dazzo, F. B., and D. H. Hubbell. 1975. “Cross-Reactive Antigens and Lectin as Determinants of Symbiotic Specificity in the Rhizobium-Clover Association.” *Applied Microbiology* 30 (6): 1017–33.

Delaux, Pierre Marc, Nathalie Séjalon-Delmas, Guillaume Bécard, and Jean Michel Ané. 2013. “Evolution of the Plant-Microbe Symbiotic ‘Toolkit.’” *Trends in Plant Science* 18 (6): 298–304.

Delgado-Cerrone, Leonardo, Alfonso Alvarez, Eilyn I. D. Mena, Iné I. D. Ponce de Leó, and Marcos MontesanoID. 2018. “Genome-Wide Analysis of the Soybean CRK-Family and Transcriptional Regulation by Biotic Stress Signals Triggering Plant Immunity.” https://doi.org/10.1371/journal.pone.0207438.

D’Haeze, Wim, Riet De Rycke, René Mathis, Sofie Goormachtig, Sophie Pagnotta, Christa Verplancke, Ward Capoen, and Marcelle Holsters. 2003. “Reactive Oxygen Species and Ethylene Play a Positive Role in Lateral Root Base Nodulation of a Semiaquatic Legume.” *Proceedings of the National Academy of Sciences of the United States of America* 100 (20): 11789–94.

Díaz, Clara L., Leo S. Melchers, Paul J. J. Hooykaas, Ben J. J. Lugtenberg, and Jan W. Kijne. 1989. “Root Lectin as a Determinant of Host-Plant Specificity in the Rhizobium-Legume Symbiosis.” *Nature* 338 (6216): 579–81.

Doench, John G., Ella Hartenian, Daniel B. Graham, Zuzana Tothova, Mudra Hegde, Ian Smith, Meagan Sullender, Benjamin L. Ebert, Ramnik J. Xavier, and David E. Root. 2014. “Rational Design of Highly Active sgRNAs for CRISPR-Cas9-Mediated Gene Inactivation.” *Nature Biotechnology* 32 (12): 1262–67.

Domonkos, Agota, Beatrix Horvath, John F. Marsh, Gabor Halasz, Ferhan Ayaydin, Giles E. D. Oldroyd, and Peter Kalo. 2013. “The Identification of Novel Loci Required for Appropriate Nodule Development in Medicago Truncatula.” *BMC Plant Biology* 13 (October): 157.

Du, D., M. Liu, Y. Xing, X. Chen, Y. Zhang, M. Zhu, X. Lu, et al. 2018. “Semi-Dominant Mutation in the Cysteine-Rich Receptor-like Kinase Gene, ALS1, Conducts Constitutive Defence Response in Rice.” *Plant Biology*. https://doi.org/10.1111/plb.12896.

Engler, Carola, Ramona Gruetzner, Romy Kandzia, and Sylvestre Marillonnet. 2009. “Golden Gate Shuffling: A One-Pot DNA Shuffling Method Based on Type IIs Restriction Enzymes.” *PloS One* 4 (5): e5553–e5553.

Engler, Carola, Mark Youles, Ramona Gruetzner, Tim-Martin Ehnert, Stefan Werner, Jonathan D. G. Jones, Nicola J. Patron, and Sylvestre Marillonnet. 2014. “A Golden Gate Modular Cloning Toolbox for Plants.” *ACS Synthetic Biology* 3 (11): 839–43.

Fukudome, Mitsutaka, Laura Calvo-Begueria, Tomohiro Kado, Ken-Ichi Osuki, Maria Carmen Rubio, Ei-Ichi Murakami, Maki Nagata, et al. 2016. “Hemoglobin LjGlb1-1 Is Involved in Nodulation and Regulates the Level of Nitric Oxide in the Lotus Japonicus-Mesorhizobium Loti Symbiosis.” *Journal of Experimental Botany* 67 (17): 5275–83.

Griesmann, Maximilian, Yue Chang, Xin Liu, Yue Song, Georg Haberer, Matthew B. Crook, Benjamin Billault-Penneteau, et al. 2018. “Phylogenomics Reveals Multiple Losses of Nitrogen-Fixing Root Nodule Symbiosis.” *Science* 1743 (May): eaat1743.

Günther, Catrin, Armin Schlereth, Michael Udvardi, and Thomas Ott. 2007. “Metabolism of Reactive Oxygen Species Is Attenuated in Leghemoglobin- Deficient Nodules of Lotus Japonicus.” *Molecular Plant-Microbe Interactions: MPMI* 20 (12): 1596–1603.

Han, Li Bo, Yuan Bao Li, Fu Xin Wang, Wen Yan Wang, Jun Liu, Jia He Wu, Nai Qin Zhong, et al. 2019. “The Cotton Apoplastic Protein CRR1 Stabilizes Chitinase 28 to Facilitate Defense against the Fungal Pathogen Verticillium Dahliae.” *The Plant Cell* 31 (2): 520–36.

Hirsch, Ann M. 1999. “Role of Lectins (and Rhizobial Legume Nodulation Ann M Hirsch in.” *Current Opinion in Plant Biology* 2: 320–26.

Hoang, Diep Thi, Olga Chernomor, Arndt von Haeseler, Bui Quang Minh, and Le Sy Vinh. 2018. “UFBoot2: Improving the Ultrafast Bootstrap Approximation.” *Molecular Biology and Evolution* 35 (2): 518–22.

Hoeksema, Jason D., James D. Bever, Sounak Chakraborty, V. Bala Chaudhary, Monique Gardes, Catherine A. Gehring, Miranda M. Hart, et al. 2018. “Evolutionary History of Plant Hosts and Fungal Symbionts Predicts the Strength of Mycorrhizal Mutualism.” *Communications Biology* 1 (1): 116.

Huang, Sanwen, Ruiqiang Li, Zhonghua Zhang, Li Li, Xingfang Gu, Wei Fan, William J. Lucas, et al. 2009. “The Genome of the Cucumber, Cucumis Sativus L.” *Nature Genetics* 41 (12): 1275–81.

Idänheimo, Niina, Adrien Gauthier, Jarkko Salojärvi, Riccardo Siligato, Mikael Brosché, Hannes Kollist, Ari Pekka Mähönen, Jaakko Kangasjärvi, and Michael Wrzaczek. 2014. “The Arabidopsis Thaliana Cysteine-Rich Receptor-like Kinases CRK6 and CRK7 Protect against Apoplastic Oxidative Stress.” *Biochemical and Biophysical Research Communications*. https://doi.org/10.1016/j.bbrc.2014.02.013.

James, E. K. 2000. “Nitrogen Fixation in Endophytic and Associative Symbiosis.” *Field Crops Research* 65 (2-3): 197–209.

Jamet, Alexandre, Karine Mandon, Alain Puppo, and Didier Hérouart. 2007. “H 2 O 2 Is Required for Optimal Establishment of the Medicago sativa/Sinorhizobium Meliloti Symbiosis.” *Journal of Bacteriology* 187 (23): 8741–45.

Jamet, Alexandre, Samuel Sigaud, Ghislaine Van de Sype, Alain Puppo, and Didier Hérouart. 2003. “Expression of the Bacterial Catalase Genes During *Sinorhizobium Meliloti-Medicago Sativa* Symbiosis and Their Crucial Role During the Infection Process.” *Molecular Plant-Microbe Interactions: MPMI* 16 (3): 217–25.

Johnson, Louise N., Martin E. M. Noble, and David J. Owen. 1996. “Active and Inactive Protein Kinases: Structural Basis for Regulation.” *Cell* 85 (2): 149–58.

Kalyaanamoorthy, Subha, Bui Quang Minh, Thomas K. F. Wong, Arndt von Haeseler, and Lars S. Jermiin. 2017. “ModelFinder: Fast Model Selection for Accurate Phylogenetic Estimates.” *Nature Methods* 14 (6): 587–89.

Kimura, Sachie, Kerri Hunter, Lauri Vaahtera, Huy Cuong Tran, Matteo Citterico, Aleksia Vaattovaara, Anne Rokka, et al. 2020. “CRK2 and C-Terminal Phosphorylation of NADPH Oxidase RBOHD Regulate Reactive Oxygen Species Production in Arabidopsis.” *The Plant Cell*, February. https://doi.org/10.1105/tpc.19.00525.

Kurtz, Stefan, Adam Phillippy, Arthur L. Delcher, Michael Smoot, Martin Shumway, Corina Antonescu, and Steven L. Salzberg. 2004. “Versatile and Open Software for Comparing Large Genomes.” *Genome Biology* 5 (2): R12.

Lambert, Annie, Daniel Marino, Emilie Andrio, Etienne G. J. Danchin, Elodie Oger, Alain Puppo, and Nicolas Pauly. 2011. “A Medicago Truncatula NADPH Oxidase Is Involved in Symbiotic Nodule Functioning.” *The New Phytologist* 189: 580–92.

Lang, Claus, Lucinda S. Smith, and Sharon R. Long. 2018. “Characterization of Novel Plant Symbiosis Mutants Using a New Multiple Gene-Expression Reporter Sinorhizobium Meliloti Strain.” *Frontiers in Plant Science* 9 (February): 76.

Lee, Dong Sook, Young Cheon Kim, Sun Jae Kwon, Choong Min Ryu, and Ohkmae K. Park. 2017. “The Arabidopsis Cysteine-Rich Receptor-like Kinase crk36 Regulates Immunity through Interaction with the Cytoplasmic Kinase

BIK1." *Frontiers in Plant Science*. https://doi.org/10.3389/fpls.2017.01856.

Lu, Kai, Shan Liang, Zhen Wu, Chao Bi, Yong Tao Yu, Xiao Fang Wang, and Da Peng Zhang. 2016. "Overexpression of an Arabidopsis Cysteine-Rich Receptor-like Protein Kinase, CRK5, Enhances Abscisic Acid Sensitivity and Confers Drought Tolerance." *Journal of Experimental Botany* 67 (17): 5009–27.

Ma, Lay Sun, Lei Wang, Christine Trippel, Artemio Mendoza-Mendoza, Steffen Ullmann, Marino Moretti, Alexander Carsten, et al. 2018. "The Ustilago Maydis Repetitive Effector Rsp3 Blocks the Antifungal Activity of Mannose-Binding Maize Proteins." *Nature Communications* 9 (1). https://doi.org/10.1038/s41467-018-04149-0.

Miyakawa, Takuya, Ken Ichi Hatano, Yumiko Miyauchi, You Ichi Suwa, Yoriko Sawano, and Masaru Tanokura. 2014. "A Secreted Protein with Plant-Specific Cysteine-Rich Motif Functions as a Mannose-Binding Lectin That Exhibits Antifungal Activity." *Plant Physiology* 166 (2): 766–78.

Miyakawa, Takuya, Ken Ichi Miyazono, Yoriko Sawano, Ken Ichi Hatano, and Masaru Tanokura. 2009. "Crystal Structure of Ginkbilobin-2 with Homology to the Extracellular Domain of Plant Cysteine-Rich Receptor-like Kinases." *Proteins: Structure, Function, and Bioinformatics* 77 (1): 247–51.

Miyakawa, Takuya, Yoriko Sawano, Ken-Ichi Miyazono, Ken-Ichi Hatano, and Masaru Tanokura. 2007. "Crystallization and Preliminary X-Ray Analysis of Ginkbilobin-2 from *Ginkgo Biloba* Seeds: A Novel Antifungal Protein with Homology to the Extracellular Domain of Plant Cysteine-Rich Receptor-like Kinases." *Acta Crystallographica. Section F, Structural Biology and Crystallization Communications* 63 (9): 737–39.

Montiel, Jesús, Noreide Nava, Luis Cárdenas, Rosana Sánchez-López, Manoj Kumar Arthikala, Olivia Santana, Federico Sánchez, and Carmen Quinto. 2012. "A Phaseolus Vulgaris NADPH Oxidase Gene Is Required for Root Infection by Rhizobia." *Plant & Cell Physiology* 53 (10): 1751–67.

Muñoz, V., F. Ibáñez, M. Tordable, M. Megías, and A. Fabra. 2015. "Role of Reactive Oxygen Species Generation and Nod Factors during the Early Symbiotic Interaction between Bradyrhizobia and Peanut, a Legume Infected by Crack Entry." *Journal of Applied Microbiology* 118 (1): 182–92.

Navarro-Gochicoa, Maria Téresa, Sylvie Camut, Antonius C. J. Timmers, Andreas Niebel, Christine Hervé, Emmanuel Boutet, Jean Jacques Bono, Anne Imberty, and Julie V. Cullimore. 2003. "Characterization of Four Lectin-Like Receptor Kinases Expressed in Roots of Medicago Truncatula. Structure, Location, Regulation of Expression, and Potential Role in the Symbiosis with Sinorhizobium Meliloti." *Plant Physiology* 133 (4): 1893–1910.

Nguyen, Lam-Tung, Heiko A. Schmidt, Arndt von Haeseler, and Bui Quang Minh. 2015. "IQ-TREE: A Fast and Effective Stochastic Algorithm for Estimating Maximum-Likelihood Phylogenies." *Molecular Biology and Evolution* 32 (1): 268–74.

Oa, Gene Expression W., Colby G. Starker, Adriana L. Parra-colmenares, Lucinda Smith, Raka M. Mitra, and Sharon R. Long. 2006. "Nitrogen Fixation Mutants of Medicago Truncatula Fail to Support Plant and Bacterial Symbiotic" 140 (February): 671–80.

Oldroyd, Giles E. D. 2013. "Speak, Friend, and Enter: Signalling Systems That Promote Beneficial Symbiotic Associations in Plants." *Nature Reviews. Microbiology* 11 (4): 252–63.

Qi, Junsheng, Jinlong Wang, Zhizhong Gong, and Jian Min Zhou. 2017. "Apoplastic ROS Signaling in Plant Immunity." *Current Opinion in Plant Biology* 38: 92–100.

Quezada, Elsa Herminia, Gabriel Xicoténcatl García, Manoj Kumar Arthikala, Govindappa Melappa, Miguel Lara, and Kalpana Nanjareddy. 2019. "Cysteine-Rich Receptor-like Kinase Gene Family Identification in the Phaseolus Genome and Comparative Analysis of Their Expression Profiles Specific to Mycorrhizal and Rhizobial Symbiosis." *Genes* 10 (1). https://doi.org/10.3390/genes10010059.

Ramu, Senthil K., Hui-Mei Peng, and Douglas R. Cook. 2002. "Nod Factor Induction of Reactive Oxygen Species Production Is Correlated with Expression of the Early Nodulin Gene rip1 in Medicago Truncatula." *Molecular Plant-Microbe Interactions: MPMI* 15 (6): 522–28.

Schindelin, Johannes, Ignacio Arganda-Carreras, Erwin Frise, Verena Kaynig, Mark Longair, Tobias Pietzsch, Stephan Preibisch, et al. 2012. "Fiji: An Open-Source Platform for Biological-Image Analysis." *Nature Methods* 9 (7): 676–82.

Singh, Prashant, Yi Chun Kuo, Swati Mishra, Chia Hong Tsai, Chih Cheng Chien, Ching Wei Chen, Marie Desclos-Theveniau, et al. 2012. "The Lectin Receptor Kinase-VI.2 Is Required for Priming and Positively Regulates Arabidopsis Pattern-Triggered Immunity." *The Plant Cell* 24 (3): 1256–70.

Soltis, D. E., P. S. Soltis, D. R. Morgan, S. M. Swensen, B. C. Mullin, J. M. Dowd, and P. G. Martin. 1995. "Chloroplast Gene Sequence Data Suggest a Single Origin of the Predisposition for Symbiotic Nitrogen Fixation in Angiosperms." *Proceedings of the National Academy of Sciences of the United States of America* 92 (7): 2647–51.

Swensen, Susan M., and Beth C. Mullin. 1997. "Phylogenetic Relationships among Actinorhizal Plants. The Impact of Molecular Systematics and Implications for the Evolution of Actinorhizal Symbioses." *Physiologia Plantarum* 99 (4): 565–73.

Tomato Genome Consortium. 2012. “The Tomato Genome Sequence Provides Insights into Fleshy Fruit Evolution.” *Nature* 485 (7400): 635–41.

Vaattovaara, Aleksia, Benjamin Brandt, Sitaram Rajaraman, Omid Safronov, Andres Veidenberg, Markéta Luklová, Jaakko Kangasjärvi, et al. 2019. “Mechanistic Insights into the Evolution of DUF26-Containing Proteins in Land Plants.” *Communications Biology* 2 (1). https://doi.org/10.1038/s42003-019-0306-9.

Velzen, Robin van, Jeff J. Doyle, and Rene Geurts. 2019. “A Resurrected Scenario: Single Gain and Massive Loss of Nitrogen-Fixing Nodulation.” *Trends in Plant Science* 24 (1): 49–57.

Velzen, R. van, R. Holmer, F. Bu, L. Rutten, A. van Zeijl, W. Liu, L. Santuari, et al. 2018. “Comparative Genomics of the Nonlegume Parasponia Reveals Insights into Evolution of Nitrogen-Fixing Rhizobium Symbioses.” *Proceedings of the National Academy of Sciences of the United States of America* 115 (20). https://doi.org/10.1073/pnas.1721395115.

Vessey, J. Kevin, Katharina Pawlowski, and Birgitta Bergman. 2005. “Root-Based N2-Fixing Symbioses: Legumes, Actinorhizal Plants, Parasponia Sp. and Cycads.” *Plant and Soil*. https://doi.org/10.1007/s11104-005-5881-5.

Wang, Yan, and Klaas Bouwmeester. 2017. “L-Type Lectin Receptor Kinases: New Forces in Plant Immunity.” *PLoS Pathogens* 13 (8): 1–7.

Wang, Yan, Klaas Bouwmeester, Patrick Beseh, Weixing Shan, and Francine Govers. 2014. “Phenotypic Analyses of Arabidopsis T-DNA Insertion Lines and Expression Profiling Reveal That Multiple L-Type Lectin Receptor Kinases Are Involved in Plant Immunity.” *Molecular Plant-Microbe Interactions: MPMI* 27 (12): 1390–1402.

Wang, Yan, Jan H. G. Cordewener, Antoine H. P. America, Weixing Shan, Klaas Bouwmeester, and Francine Govers. 2015. “Arabidopsis Lectin Receptor Kinases LecRK-IX.1 and LecRK-IX.2 Are Functional Analogs in Regulating Phytophthora Resistance and Plant Cell Death.” *Molecular Plant-Microbe Interactions: MPMI* 28 (9): 1032–48.

Wardhani, Titis A. K., Yuda Purwana Roswanjaya, Simon Dupin, Huchen Li, Sidney Linders, Marijke Hartog, Rene Geurts, and Arjan van Zeijl. 2019. “Transforming, Genome Editing and Phenotyping the Nitrogen-Fixing Tropical Cannabaceae Tree Parasponia Andersonii.” *Journal of Visualized Experiments: JoVE* 2019 (150). https://doi.org/10.3791/59971.

Webster, G., P. R. Poulton, E. C. Cocking, and M. R. Davey. 1995. “The Nodulation of Micro-Propagated Plants of *Parasponia Andersonii* by Tropical Legume Rhizobia.” *Journal of Experimental Botany* 46 (290): 1131–37.

Wittenberg, J. B., B. A. Wittenberg, Q. H. Gibson, M. J. Trinick, and C. A. Appleby. 1986. “The Kinetics of the Reactions of Parasponia Andersonii Hemoglobin with Oxygen, Carbon Monoxide, and Nitric Oxide.” *The Journal of Biological Chemistry* 261 (29): 13624–31.

Yadeta, Koste A., James M. Elmore, Athena Y. Creer, Baomin Feng, Jessica Y. Franco, Jose Sebastian Rufian, Ping He, Brett Phinney, and Gitta Coaker. 2017. “A Cysteine-Rich Protein Kinase Associates with a Membrane Immune Complex and the Cysteine Residues Are Required for Cell Death.” *Plant Physiology* 173 (1): 771–87.

Yeh, Yu-Hung, Yu-Hsien Chang, Pin-Yao Huang, Jing-Bo Huang, and Laurent Zimmerli. 2015. “Enhanced Arabidopsis Pattern-Triggered Immunity by Overexpression of Cysteine-Rich Receptor-like Kinases.” *Frontiers in Plant Science* 6 (May): 322.

Yu, Haixiang, Aifang Xiao, Ru Dong, Yuqian Fan, Xianpeng Zhang, Chao Liu, Chao Wang, et al. 2018. “Suppression of Innate Immunity Mediated by the CDPK-Rboh Complex Is Required for Rhizobial Colonization in Medicago Truncatula Nodules.” *The New Phytologist* 220 (2): 425–34.

Zeijl, Arjan van, Titis A. K. Wardhani, Maryam Seifi Kalhor, Luuk Rutten, Fengjiao Bu, Marijke Hartog, Sidney Linders, et al. 2018. “CRISPR/Cas9-Mediated Mutagenesis of Four Putative Symbiosis Genes of the Tropical Tree Parasponia Andersonii Reveals Novel Phenotypes.” *Frontiers in Plant Science* 9 (March): 1–14.

Zhang, Xiujuan, Xiaomin Han, Rui Shi, Guanyu Yang, Liwang Qi, Ruigang Wang, and Guojing Li. 2013. “Arabidopsis Cysteine-Rich Receptor-like Kinase 45 Positively Regulates Disease Resistance to Pseudomonas Syringae.” *Plant Physiology and Biochemistry: PPB / Societe Francaise de Physiologie Vegetale* 73 (December): 383–91.

Zuo, Cunwu, He Liu, Qianqian Lv, Zhongjian Chen, Yuzhen Tian, Juan Mao, Mingyu Chu, Zonghuan Ma, Zeshan An, and Baihong Chen. 2019. “Genome-Wide Analysis of the Apple (Malus Domestica) Cysteine-Rich Receptor-Like Kinase (CRK) Family: Annotation, Genomic Organization, and Expression Profiles in Response to Fungal Infection.” *Plant Molecular Biology Reporter / ISPMB*, no. Fischer 2000. https://doi.org/10.1007/s11105-019-01179-w.

CHAPTER 6

6

A remote *cis*-Regulatory Region is required for *NIN* expression in the pericycle to initiate nodule primordium formation in *Medicago Truncatula*

Jieyu Liu[a,‡], Luuk Rutten[a,‡], Erik Limpens[a], Tjitse van der Molen[a], Robin van Velzen[a], Rujin Chen[b,c], Yuhui Chen[b,c], Rene Geurts[a], Wouter Kohlen[a], Olga Kulikova[a,e], and Ton Bisseling[a,e]

[a]Laboratory of Molecular Biology, Department of Plant Sciences, Graduate School Experimental Plant Sciences, Wageningen University, 6708 PB Wageningen, The Netherlands. [b]Plant Biology Division, The Samuel Roberts Noble Foundation, Ardmore, OK 73401, USA. [c]School of Life Sciences, Lanzhou University, 730000 Lanzhou, China. [e]Corresponding authors: ton. bisseling@wur.nl; olga.kulikova@wur.nl

[‡]These authors contributed equally to this work

This chapter is published in: The Plant Cell, First online: January 4, 2019, doi:10.1105/tpc.18.00478

Abstract

The legume-rhizobium symbiosis results in nitrogen fixing root nodules, and their formation involves both intracellular infection initiated in the epidermis and nodule organogenesis initiated in inner root cell layers. *NODULE INCEPTION* (*NIN*) is a nodule-specific transcription factor essential for both processes. These NIN-regulated processes occur at different times and locations in the root, demonstrating a complex pattern of spatiotemporal regulation. We show that regulatory sequences sufficient for the epidermal infection process are located within a 5 kb region directly upstream of the *NIN* start codon in *Medicago truncatula*. Furthermore, we identify a remote upstream *cis*-regulatory region required for the expression of *NIN* in the pericycle, and we show that this region is essential for nodule organogenesis. This region contains putative cytokinin response elements, and is conserved in eight more legume species. Both the cytokinin receptor *CRE1*, which is essential for nodule primordium formation, and the B-type response regulator *RR1* are expressed in the pericycle in the susceptible zone of the uninoculated root. This, together with the identification of the cytokinin responsive elements in the *NIN* promoter, strongly suggests that *NIN* expression is initially triggered by cytokinin signalling in the pericycle to initiate nodule primordium formation.

Introduction

The formation of nitrogen-fixing nodules is induced by rhizobium bacteria on the roots of legumes. It involves several processes: the induction of intracellular infection by rhizobia; nodule organogenesis; and a negative feedback loop that determines the number of nodules (Downie, 2014; Kosslak and Bohlool, 1984). Strikingly, the transcription factor NODULE INCEPTION (NIN), which is specifically expressed during nodulation, plays a key role in all of these processes (Schauser et al., 1999; Marsh et al., 2007; Soyano et al., 2014). These processes occur at different time points and locations, suggesting that *NIN* has a complex spatiotemporal regulation of expression that is regulated by distinct *cis*-regulatory sequences in its promoter. However, although *NIN* was identified almost two decades ago in *Lotus japonicus* (Lotus) (Schauser et al., 1999) and more than ten years ago in *Medicago truncatula* (Medicago) (Marsh et al., 2007), the promoter regions required for full complementation of *nin* knockout mutants have not been identified. Currently, it is unclear how NIN is involved in the multiple steps of the nodulation process.

In Medicago, nodule organogenesis starts with the local mitotic activation of pericycle cells, and subsequently cell division extends to the more outward located endodermis and cortex (Xiao et al., 2014). *Sinorhizobium meliloti* bacteria invade roots through tube-like structures called infection threads. Formation of infection threads in root hairs requires the prior induction of root hair curling. A tight curl is

formed when the tip of the curling root hair touches the shank of the hair and the root hair stops growing and forms an infection chamber. Microcolonies of rhizobia then develop within these chambers, in which the rhizobia can induce formation of the infection thread.

In *nin* null mutants, extensive root hair curling and deformation are induced by bacteria, a proper infection chamber fails to be established and only few bacteria are present within curled root hairs (Fournier et al., 2015; Schauser et al., 1999; Marsh et al., 2007). Wild-type NIN induces infection thread formation by triggering expression of genes required for infection thread formation such as *NF-YA1*, which encodes a subunit of a nuclear factor Y complex, and *NPL,* which encodes a nodulation pectate lyase (Soyano et al., 2013; Laporte et al., 2014; Xie et al., 2012). Subsequently, infection threads grow towards the nodule primordia. There, rhizobia are released into nodule primordium cells derived from the cortex. These cells become infected cells which host thousands of nitrogen-fixing bacteria. NIN is also required for autoregulation of nodulation, a negative feedback system involving root-to-shoot communication to determine the optimal number of nodules. This autoregulation mechanism includes the induction of CLAVATA3/ESR-RELATED (*CLE*) genes by NIN, and the *CLE*-encoded peptides induce systemic signalling between root and shoot, suppressing the formation of new nodule primordia (Soyano et al., 2014).

Expression of *NIN* is induced in the epidermis upon perception of nodulation (Nod) factors, which are lipochitooligosaccharides (LCOs) secreted by rhizobia (Vernié et al., 2015; Van Zeijl et al., 2015). Nod-factor signalling induces Ca2+ spiking, which activates the nuclear-localized calcium and calmodulin-dependent kinase (CCaMK) (Ehrhardt et al., 1996; Mitra et al., 2004). CCaMK phosphorylates *CYCLOPS,* a transcription factor that activates *NIN* expression (Yano et al., 2008; Singh et al., 2014). At about 24 h post inoculation (hpi), formation of both infection threads and nodule primordium, are initiated in Medicago roots (Xiao et al., 2014). At this developmental stage, Nod factor signalling occurs exclusively in the epidermis because rhizobia are present only there and Nod factors are immobile molecules (Goedhart et al., 2000). Therefore, NIN can induce infection thread formation in a cell-autonomous way in the epidermis, but it remains unclear how NIN can induce nodule primordium formation in inner root cell layers. It has been postulated that *NIN* expression can be induced in these root layers by a mobile signal that is generated upon Nod factor signalling in the epidermis (Hayashi et al., 2014). Alternatively, NIN proteins produced in the epidermis may be transported to the inner root layers (Vernié et al., 2015; Jardinaud et al., 2016).

In addition to regulation of *NIN* by CYCLOPS, *NIN* expression depends on cytokinin signalling. Exogenous application of cytokinin is sufficient to trigger *NIN* expression and also formation of structures resembling nodules (Heckmann et al., 2011; Gonzalez-Rizzo et al., 2006; Plet et al., 2011). Notably, Nod factor application results in the accumulation of cytokinin (Van Zeil et al., 2015). Furthermore, the induction of *NIN* expression by either Nod factors or cytokinin requires the cytokinin receptor CRE1, which plays a key role in nodule organogenesis (Van Zeijl et al., 2015; Gonzalez-Rizzo et al., 2006; Plet et al., 2011). Studies of the weak *nin* allele of the Lotus *daphne* mutant (Yoro et al., 2014) have provided valuable insight into the involvement of different *cis*-regulatory sequences in the *NIN* promoter that regulate infection and nodule organogenesis. In *daphne* roots, rhizobium infection and primordium formation are uncoupled. Formation of nodule primordia is completely absent, and increased numbers of infection threads are formed in the epidermis. The *daphne* mutation is caused by a large insertion that is ~7 kb upstream of the *NIN* start codon. This suggests that this 7-kb region includes essential *cis*-regulatory regions that are required for infection thread formation in the epidermis but are insufficient for the activation of cortical cell divisions. To determine how NIN induces these processes, we must identify the precise *cis*-regulatory regions in the *NIN* promoter that drive proper spatiotemporal *NIN* transcription.

Here, we have identified a conserved *NIN* promoter region that is essential for nodule organogenesis in Medicago. This region contains several putative cytokinin response elements and regulates *NIN* expression in the pericycle, where the cytokinin receptor (*CRE1*) and a B-type response regulator (*RR1*) are constitutively expressed. This reveals a key role for the pericycle in formation of nodule primordia, in which *NIN* expression is most likely activated by cytokinin signalling.

Results

Isolation of a Medicago *nin* mutant in which infection and nodule organogenesis are uncoupled.

By screening a plant population obtained from Medicago seeds that were mutagenized by fast neutron bombardment (Noble Research Institute, LLC., Ardmore USA), we identified a Nod- mutant which we have named FN8113. Three weeks post inoculation (wpi) with *Sinorhizobium meliloti*, the FN8113 mutant formed excessive numbers of infection threads, but nodulation was strongly impaired (compare **Figures 1A and 1B** with **Figures 1C and 1D**). We quantified the infection thread number in FN8113 and wild type roots at 2 wpi. The number of infection threads in FN8113 roots was more than tenfold the number in wild type roots **(Figure 1E).** Root hair curling in FN8113 resembled that of wild type, as entrapped bacteria formed colonies and infection threads were formed (compare **Figure 1F and 1G**). The majority of infection threads

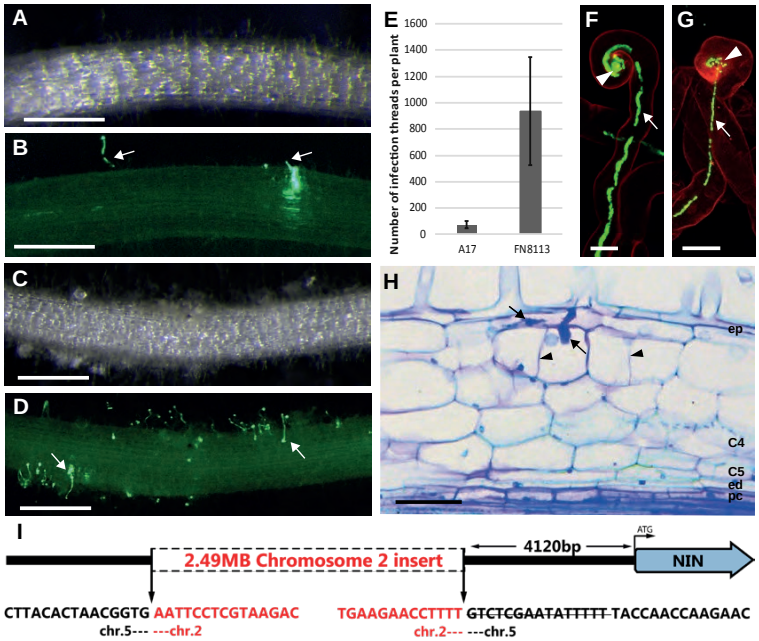


Figure 1. In the *Medicago nin* Mutant *daphne-like*, Infection and Nodule Organogenesis Are Uncoupled. **(A)** to **(D)** Images of wild-type and mutant roots. These transmitted light stereo macroscopy images **(A)** and **(C)** and corresponding green fluorescence stereo macroscopy images **(B)** and **(D)** were taken at 3 weeks post inoculation (wpi). Roots of *daphne-like* (FN8113) mutant plants **(C)** and **(D)** have an excessive number of infection threads in comparison to wild type A17 roots **(A)** and **(B)**. Bars = 2 mm. **(E)** Quantification of infection thread number. The number of infection threads per root was counted at 2 wpi in both A17 roots (n = 12) and FN8113 roots (n = 12). Data are mean ± SD. **(F)** and **(G)** Infection thread formation in mutant and wild type roots. These confocal images of roots stained with propidium iodide at 1 wpi show that a bacterial colony (arrowhead) is formed inside a daphne-like curled root hair and an infection thread (arrow) is initiated **(F)** like in a wild type root hair **(G)**. S. meliloti containing constitutively expressed GFP was used as inoculum. ep, epidermis; C4, C5, cortical cell layers 4 and 5; ed, endodermis; pc, pericycle. Bars = 10 µm. **(H)** Longitudinal plastic section of daphne-like root at3 wpi. The section stained with toluidine blue displays an infection thread (arrow). The infection threads in a mutant can occasionally reach cortical cell layers and induce some cell divisions (arrowhead). Bars = 50 µm. **(I)** Schematic representation of the chromosome translocation at the *NIN* locus in the *daphne-like* mutant. The strikethrough indicates a 15-bp deleted sequence.

were arrested in root hairs, but longitudinal sections of roots showed that a few infection threads could reach cortical cell layers **(Figure 1H)**. Occasionally, some cortical cells had divided locally around infection threads. However, cell divisions were not induced in the inner root cell layers, where nodule primordia are initiated in wild type Medicago plants.

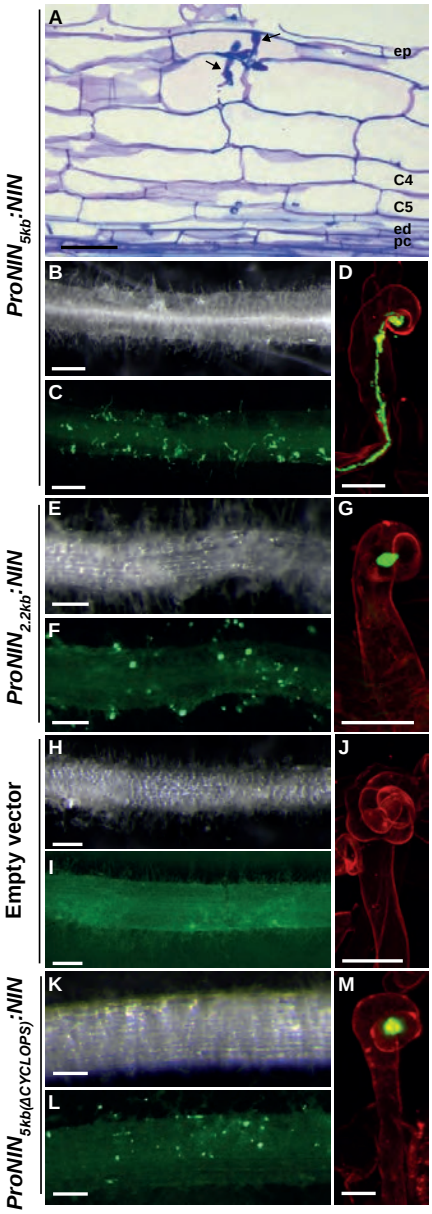


Figure 2. The Infection Process in *Medicago nin-1* roots is partially rescued by introducing ProNIN5kb:NIN or ProNIN2.2kb:NIN. **(A)** to **(D)** Phenotype of *nin-1* roots transformed with ProNIN5kb:NIN at 4 weeks post inoculation (wpi) with *S. meliloti* constitutively expressing GFP. The transmitted light microscopy image of a longitudinal plastic section of transgenic root stained with toluidine blue displays infection threads (arrow) that occasionally can reach cortical cell layers **(A)**. Bar = 50 µm. The transmitted light stereo macroscopy image (B) and corresponding green fluorescence stereo macroscopy image **(C)** show excessive infection thread formation in transgenic roots. Bars = 2 mm. This confocal image of a transgenic root stained with propidium iodide shows an infection thread initiated in the curled root hair **(D)**. Bar = 10 µm. **(E)** to **(G)** Phenotype of *nin-1* roots transformed with ProNIN2.2kb:NIN at 4 wpi with *S. meliloti* constitutively expressing GFP. The transmitted light stereo macroscopy image **(E)** and corresponding green fluorescence stereo macroscopy image **(F)** display numerous curled root hairs with bacterial colonies in transgenic roots. Bars = 2 mm. In this confocal image of a transgenic root stained with propidium iodide **(G)**, colonies are formed inside the chamber of the root hair curl, but an infection thread does not develop. Bar = 10 µm. **(H)** to **(J)** Phenotype of *nin-1* roots transformed with empty vector at 4 wpi with *S. meliloti* constitutively expressing GFP. The transmitted light stereo macroscopy image **(H)** and corresponding green fluorescence stereo macroscopy image **(I)** show that the transgenic root forms neither infection threads nor bacterial colonies. Bars = 2 mm. A transgenic root stained with propidium iodide shows excessive root hair curling **(J).** Bar = 10 µm. **(K)** to **(M)** Phenotype of *nin-1* roots transformed with ProNIN5kb(Δcyclops):NIN at 4 wpi with *S. meliloti* constitutively expressing GFP. The transmitted light stereo macroscopy image **(K)** and corresponding green fluorescence stereo macroscopy image **(L)** display many bacterial colonies in transgenic root hairs. Bars = 2 mm. This confocal image of a transgenic root stained with propidium iodide shows that colony is formed inside the chamber of the root hair curl but infection thread is not initiated **(M)**. Bar = 10 µm.

A segregating F2 population resulting from a cross between FN8113 (cv Jemalong A17) and Jemalong A20, showed an approximately 3:1 ratio of Nod+:Nod- plants (118 F2 plants; 84 Nod+: 34 Nod-). This indicates that FN8113 has a single recessive mutation that is responsible for its Nod- phenotype. Simple sequence repeat markers were used to determine the position of the mutation, which was localized to the end of chromosome 5, where *NIN* is located. Next, whole-genome sequencing was used to identify the mutation in this region, and this revealed a translocation of a ~2.49 Mbp region from chromosome 2 into chromosome 5. This was inserted 4120 bp upstream of the *NIN* start codon (–4120). In addition, a small deletion of 15 bp between –4121 and –4135 was detected **(Figure 1I)**. No mutations were found in the *NIN* coding sequence. FN8113 was shown to be a *nin* mutant because its Nod- phenotype could be complemented with a biologically functional *NIN* promoter driving *NIN* (described below). Because the phenotype of FN8113, as well as the nature of its mutation, are strikingly similar to Lotus *daphne*, we named the FN8113 mutant *daphne-like*.

The 5kb upstream region of Medicago *NIN* contains discrete regulatory sequences that affect root hair curling and infection.

The phenotype of *daphne-like* strongly suggests that *NIN* regulatory sequences required for primordium formation are located more than 4120 bp upstream of the *NIN* start codon. In addition, this phenotype indicates that the regulatory sequences located within this 4120-bp region are sufficient for proper root hair curling and infection thread formation. We tested this by using a construct containing the 5-kb region upstream of the start codon to drive expression of *NIN*. We introduced this construct, *ProNIN5kb:NIN*, into Medicago *nin-1* (null mutant, Marsh et al., 2007) roots by *Agrobacterium rhizogenes*-mediated root transformation. At 4 wpi, 41 of 44 analyzed transgenic roots showed excessive infection thread formation **(Figure 2B to 2D)**. Despite the numerous infections, these roots did not form nodules, except for one root on which four nodules were observed. As other transgenic roots of this composite plant have no nodules, we assume that this is caused by transgene insertion. Longitudinal sections of infected transgenic roots confirmed that cell divisions were not induced in the pericycle, endodermis and inner cortical cell layers **(Figure 2A)**. Infection threads were arrested in the epidermis, but occasionally some of these reached the cortex. Thus, the 5 kb promoter region is sufficient for infection thread formation, but it lacks regulatory sequences for primordium formation.

Interestingly, a single putative CYCLOPS/IPD3 binding site is located about –3 kb upstream of the start codon **(Figures 3 and 4A, Supplemental Figure 1, Supplemental Table 1)** (Singh et al., 2014). We therefore checked whether the function of NIN in the epidermis fully depends on this putative CYCLOPS binding site by using the –2.2 kb region **(Figure 4A)** to drive *NIN* expression. The *ProNIN2.2kb:NIN* construct was introduced into *nin-1* by *A. rhizogenes*-mediated

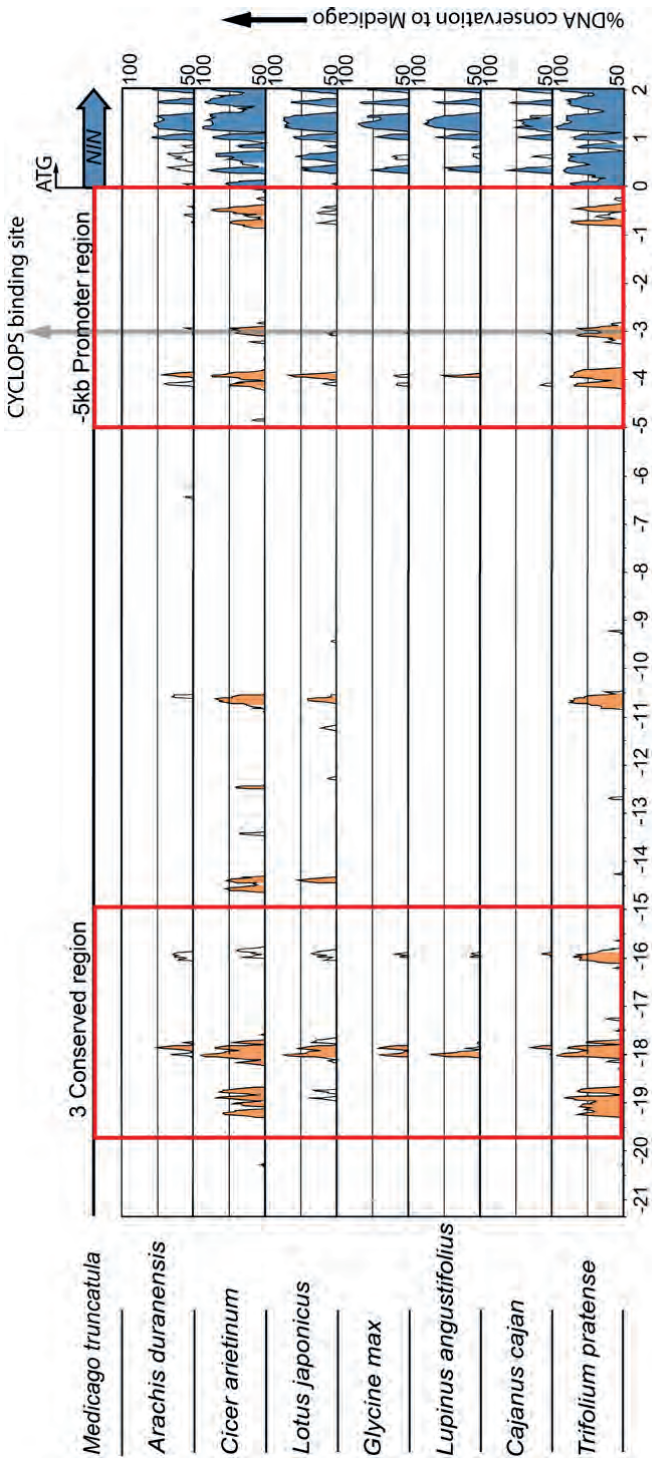


Figure 3. mVISTA alignment of genomic DNA sequences including 2 kb from the start codon of *NIN* and 5'-Upstream regions from eight legume species. Peaks indicate the level of identity with Medicago on a scale of 50% to 100%. Identities lower than 50% were not scored. The figure shows 2 kb of each NIN sequence downstream of the start codon (blue) and the 5'- upstream DNA sequence (orange). The red rectangle on the right next to the NIN coding sequence indicates the –5 kb promoter region. The red rectangle on the left indicates the three conserved regions (3C). The grey arrow shows the putative CYCLOPS binding site.

root transformation. The *nin-1* null mutant has excessive root hair curling but fails to form infection threads **(Figure 2H to 2J)**. Although all 37 analysed transgenic roots at 4 wpi showed tight root hair curls that enclosed bacterial colonies, infection threads were rare **(Figure 2E to 2G)**. Of 298 curled root hairs in *ProNIN2.2kb:NIN* transgenic roots containing a bacterial colony, only ~3% had an infection thread. This shows that root hair curling and establishment of infection chambers do not rely on the putative CYCLOPS binding site. By contrast, ~70% of these curled root hairs (n = 324) formed infection threads in *ProNIN5kb:NIN* transgenic roots.

These results indicate that the –5 kb to –2.2 kb region contains regulatory sequences that are critical for infection thread formation. The observed phenotype is reminiscent of that of Lotus and Medicago *cyclops-3/ipd3-2* mutants (Yano et al., 2008; Horváth et al., 2011), which do not form infection threads but show formation of bacterial colonies in tightly curled root hairs. Therefore, the –2.2 kb region can activate *NIN* expression in the epidermis, and the expression level is sufficient for tight root hair curling, allowing rhizobia to form a colony inside the curl. However, additional regulatory sequences located between –5 kb and –2.2 kb, probably involving the putative CYCLOPS binding site, are required for efficient infection thread formation. To test this, we analysed *nin-1* roots transformed with *NIN* driven by the –5 kb promoter in which the putative CYCLOPS binding site was deleted (*ProNIN5kb(Δcyclops):NIN*) **(Figure 2K to 2M)**. Due to this mutation, the number of curled root hairs with a colony (similar in size to the one formed in wild type roots) that initiated an infection thread dropped from 70% to 7% (n = 434). This shows that the putative CYCLOPS-binding site within the *NIN* promoter is essential for efficient infection thread formation.

A conserved region with putative Cytokinin Response Elements is located ~18 kb upstream of the *NIN coding region in Medicago truncatula.*

The *daphne-like* mutant, as well as *nin-1* transformed with *ProNIN5kb:NIN,* can induce formation of infection threads but not nodule primordia. Based on this, we hypothesized that the regulatory elements required for NIN-induced nodule primordium formation are located upstream of –5 kb. This resembles the Lotus *daphne* mutant, which contains a chromosomal insertion at ~7 kb upstream of the *NIN* start codon. Therefore, we expected that such remote regulatory regions would be conserved in Lotus and Medicago and probably in other legumes. To test this, we compared the genomic DNA sequences spanning from the *NIN* coding region to the first upstream gene in eight legume species (*Medicago truncatula*, *Lotus japonicus, Arachis duranensis*, *Cicer arietinum*, *Glycine max*, *Lupinus angustifolius*, *Cajanus cajan* and *Trifolium pratense*). Based on the high level of identity (50%- 100%) among all these species we identified DNA sequences with three conserved regions (3C) far upstream of the *NIN* start codon **(Figure 3 and Supplemental Table 1)**.

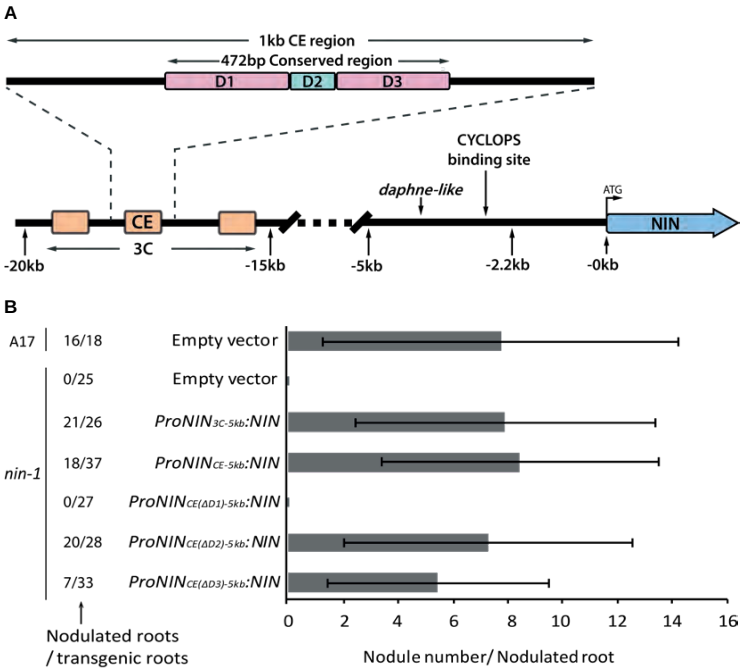


Figure 4. The CE Region in the *NIN* promoter is essential for nodule organogenesis. **(A)** Schematic illustration of the Medicago truncatula NIN promoter. Three conserved remote regions (3C, orange) were identified among the eight legume species studied here. The second 3C region is most conserved, and it includes about ten putative B-type cytokinin signalling RESPONSE REGULATOR (RR) binding sites and is named the CE region (for cytokinin response elements-containing). CE region in turn contains a highly conserved 472-bp sequence that was divided into three parts named domains D1, D2 and D3. Domains D1 and D3 (purple) contain six and three putative cytokinin response elements, respectively, whereas D2 (green) contains a putative AP2-binding site and a single putative cytokinin response element. **(B)** The number of nodules formed on wild type (A17) roots transformed with empty vector and *nin-1* roots transformed with the constructs carrying *NIN* driven by different parts of the NIN promoter as indicated. *S. meliloti* containing constitutively expressed GFP was used as inoculum. Nodule numbers per nodulated root were counted at 4 wpi. Data are means ± SD.

In Medicago, 3C is located 15 to 20 kb upstream of the *NIN* start codon, and in Lotus it is located between 42 to 49 kb upstream **(Supplemental Table 1)**. The levels of identity in conserved regions of 3C are similar to that of the *NIN* coding region (Figure 3). The second region in 3C is most conserved and includes about ten putative B-type cytokinin signalling RESPONSE REGULATOR (RR) binding sites **(Figure 4 and Supplemental Figure 2)** (Sheen, 2002; Heyl and Schmülling, 2003; Hosoda et al., 2002; Imamura et al., 2003). Therefore, we named this middle region the CE region (for cytokinin response elements-containing). Because cytokinin

signalling is essential for nodule organogenesis and induction of *NIN* expression, the CE region may be involved in regulation of *NIN* expression during initiation of nodule primordium formation.

The CE Region contains regulatory elements required for nodule organogenesis. To determine whether 3C (~4 kb) contains regulatory sequences necessary for nodule primordium formation, we fused it to the (upstream) –5 kb region (*ProNIN3C-5kb:NIN*), as the latter is sufficient for infection. *ProNIN3C-5kb:NIN* was introduced into *nin-1* by *A. rhizogenes*-mediated root transformation. 21 of 26 analysed transgenic roots (at 4 wpi) formed, on average, eight nodules per root **(Figure 4B)**. As the CE region (~1 kb) contains several putative cytokinin response elements, we tested whether this region is sufficient to trigger primordium formation. To this end, we transformed *nin-1* with the CE region fused to the –5 kb region driving *NIN* (*ProNINCE-5kb:NIN*). This resulted in 18 out of 37 transgenic roots forming on average eight nodules per root **(Figure 3B)**. This demonstrates that the CE region contains regulatory sequences that are required for primordium formation. Furthermore, the number of nodules formed on *ProNINCE-5kb:NIN* expressing roots was similar to the number on wild type roots transformed with an empty vector control (Figure 3B). This suggests that the autoregulation of nodulation mechanism is also activated (Soyano et al., 2014). In addition, 12 out of 22 *nin-1* roots transformed with *ProNINCE-5kb:NIN* displayed wild type like infection thread numbers, indicating that the excessive infection phenotype can be partially rescued by including the CE region in *NIN* promoter.

Normal nodules are pink due to the presence of leghemoglobin, which keeps the oxygen level low in the infected cells of the fixation zone so that the anaerobic process of nitrogen fixation can proceed. Pink nodules were formed on *nin-1* roots transformed with either *ProNIN3C-5kb:NIN* or *ProNINCE-5kb:NIN*. Longitudinal sections of these nodules showed a zonation similar to wild type nodules: a meristem at the apex; the infection zone, where rhizobia are released from the infection thread and subsequently divide and begin to enlarge; and the fixation zone, where rhizobia have reached their fully enlarged shape and are able to fix nitrogen **(Figure 5A to 5D)**. Nodules formed by inoculation with *S. meliloti* carrying the *PronifH:GFP* reporter, showed that nitrogenase *nifH* was expressed in the fixation zone, confirming that these nodules are functional **(Supplemental Figure 3)**. Thus, CE in combination with the –5 kb region is sufficient to induce wild type-like nodule organogenesis.

***NIN* expression cannot be induced by cytokinin in the *daphne-like* mutant.** Because the *ProNINCE-5kb:NIN* construct can fully restore nodulation ability in *nin-1* roots, we used it to verify that *daphne-like* is indeed a *nin* allele. Therefore, *daphne-like* was transformed with *ProNINCE-5kb:NIN* **(Figures 5E and 5F)**, and 15

of 17 transgenic roots analyzed at 4 wpi formed on average about seven nodules per root. The excessive infection phenotype in the *daphne-like* background was rescued by *ProNINCE-5kb:NIN* in 11 of these 17 transgenic roots. This result shows that *daphne-like* is a *nin* mutant. Its phenotype is most likely caused by the 2.49 Mbp insertion by which the CE region is positioned too far away from the transcription start to contribute to the correct expression of *NIN* for nodule primordium formation.

To test whether the CE region is sufficient to complement nodule organogenesis in *daphne-like*, we used a minimal –46 bp CaMV 35S promoter (Benfey and Chua, 1990) fused to the CE region *(ProNINCE-35Smin:NIN)* **(Figures 5G and 5H)**. We found that 37 out of 45 transgenic *daphne-like* roots had formed on average four nodules per root at 4 wpi. This indicates not only that the CE region is sufficient to induce nodule organogenesis, but also that, in combination with the –5 kb region, more nodules (about seven per root) can be formed. The ability to form nodules can be rescued in *daphne-like* by the CE region driving *NIN* expression. Therefore, it is likely that the CE region in *daphne-like* cannot regulate the expression of *NIN*. Because

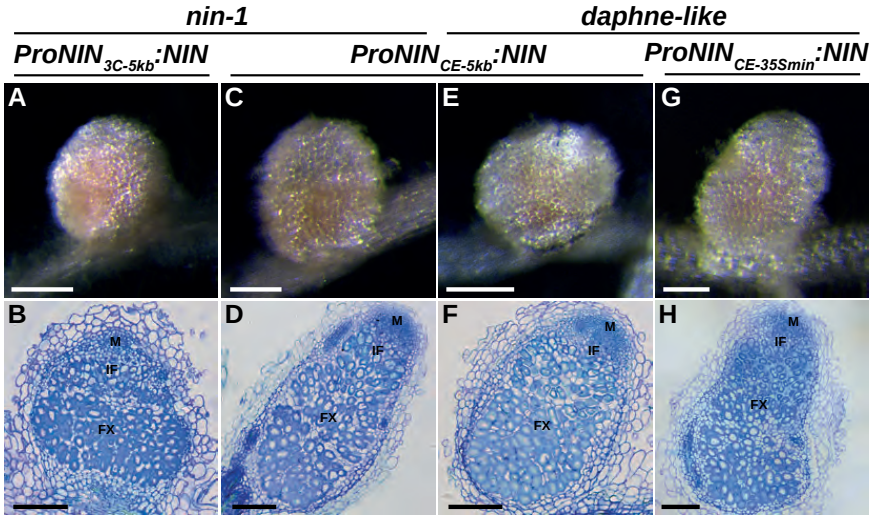


Figure 5. Non-nodulating phenotypes of *nin-1* and *daphne-like* are rescued by *A. rhizogenes*-mediated transformation with ProNIN3C-5kb:NIN, ProNINCE-5kb:NIN and ProNINCE-35Smin:NIN. **(A)** to **(D)** Nodules formed on *nin-1* roots transformed with the indicated constructs at 4 wpi with S. meliloti. Transmitted light macroscopy images of nodules **(A)** and **(C)**. Nodules are pink due to the presence of leghaemoglobin. Bars = 2 mm. Longitudinal plastic sections of these nodules stained with toluidine blue display normal zonation. **(B)** and **(D)**. M, meristem; IF, infection zone; FX, fixation zone. Bars = 200 µm. **(E)** to **(H)** Nodules formed on daphne-like roots transformed with the indicated constructs at 4 wpi with S. meliloti. Transmitted light macroscopy images of nodules **(E)** and **(G)**. Bars = 2 mm. Longitudinal plastic sections of these nodules stained with toluidine blue display normal zonation **(F)** and **(H)**. M, meristem; IF, infection zone; FX, fixation zone. Bars = 200 µm.

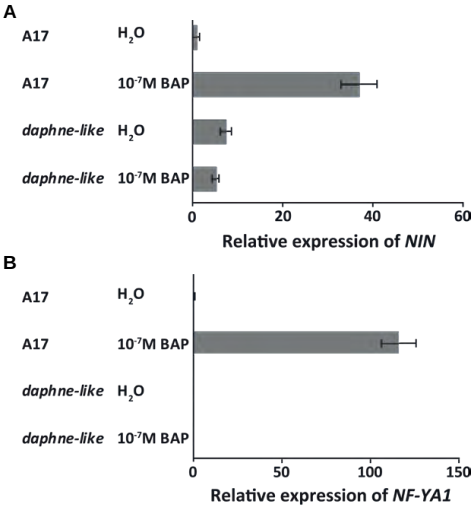


Figure 6. *NIN* and *NF-YA1* expression cannot be Induced by cytokinin application in *daphne-like* mutants.
Quantitative real-time RT-PCR analysis of NIN **(A)** and NF-YA1 **(B)** expression in wild type (A17) and *daphne-like* roots 16 h after application of 10-7M benzylaminopurine (BAP) or water. Data are means ± SEM of three biological replicates.

the CE region contains several putative response elements, we hypothesised that *NIN* expression would not be induced by cytokinin in *daphne-like*. To test this, we compared the induction of *NIN* expression by cytokinin vs. water (as control) in wild type (A17) and *daphne-like.* We found that 16 h after 10-7M benzylaminopurine (BAP) application, *NIN* expression level increased 37 fold compared with the control, and *NF-YA1* expression level increased over a hundred fold in wild-type, while both *NIN* and *NF-YA1* expression levels in *daphne-like* were not changed **(Figures 6A and 6B)**. This suggests that the CE region is required for the induction of *NIN* expression by cytokinin.

A Domain with six putative Cytokinin Response Elements is essential for nodule primordium formation.

Because cytokinin is known to be a positive regulator of nodule primordium formation (Suzaki et al., 2013), we tested whether the putative cytokinin response elements within the CE region are essential for primordium formation. To this end, we made several deletions in the CE region, which contains a 472-bp region that is highly conserved in all eight legume species studied here **(Figures 3 and 4A and Supplemental Figure 2)**. We divided this 472-bp region into three parts named domains one to three (D1 to D3). D1 and D3 contain six and three putative cytokinin response elements, respectively, whereas domain 2 (D2) contains a putative AP2-binding site as well as a single cytokinin response element **(Figure 4A and Supplemental Figure 2)**.

Several studies have shown that transcription factors of the AP2 family, including ERN (ethylene response factor required for nodulation) are involved in regulating nodulation (Andriankaja et al., 2007; Middleton et al., 2007; Wang et al., 2014). To

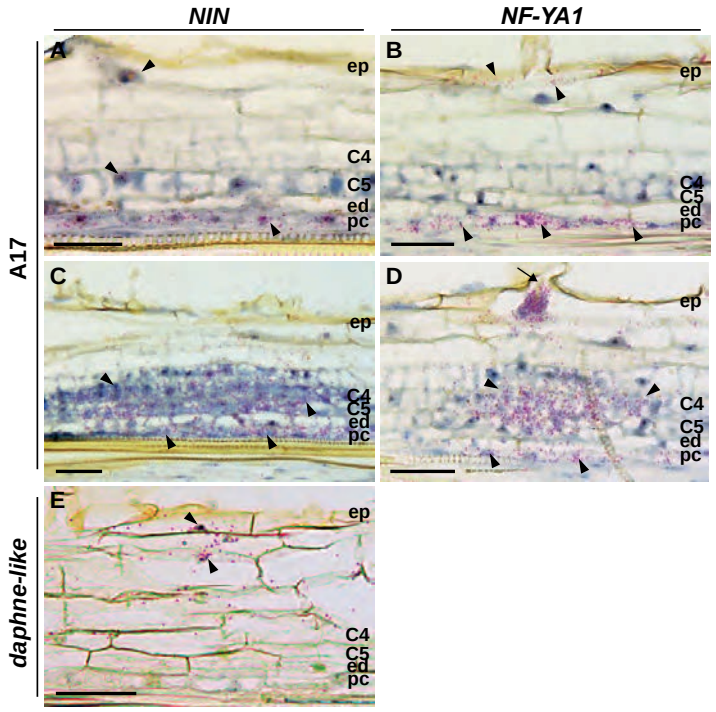


Figure 7. *NIN* and *NF-YA1* expression patterns in medicago Wild Type (A17) nodule primordia and *daphne-like* mutant.
(A) to **(D)** RNA in situ localization of NIN **(A)** and **(C)** and NF-YA1 **(B)** and **(D)** in nodule primordia at 2 days post inoculation (dpi) **(A)** and **(B)** and at 3 dpi **(C)** and **(D)**. The arrow indicates an infection thread.**(E)** RNA in situ localization of NIN in roots of the *daphne-like* mutant at 2 dpi. Hybridization signals are visible as red dots (arrowheads). ep, epidermis; C4, C5, cortical cell layers 4 and 5; ed, endodermis; pc, pericycle. Bars = 50 µm.

investigate their respective contribution to nodule primordium formation, the D1, D2 or D3 regions were separately deleted from the 1-kb CE region **(Figure 4A and Supplemental Figure 2)**, and the modified CE regions were fused to the –5 kb region to drive *NIN* expression. These three constructs were introduced into *nin-1* by *A. rhizogenes*-mediated root transformation. Our results show that deletion of D1 eliminated nodulation ability (**Figure 4B** and **Supplemental Figure 4A**), whereas deletion of D2 had no significant effect on nodulation (**Figure 4B** and **Supplemental Figure 4B**). Deletion of D3 caused a reduction of the relative number of roots with nodules from 49% to 21% and also reduced the average nodule number per root from eight to five (**Figure 4B** and Supplemental **Figure 4C**). These results show that regulatory sequences in D1 are essential for NIN-regulated nodule primordium

formation and suggest that the putative cytokinin response elements within D1 are likely responsible. In contrast, the putative AP2 binding site in D2 is not essential for nodule organogenesis.

Induction of *NIN* expression in inner root cell layers occurs in a non-cell-autonomous manner.

It was shown that the 2.2 kb upstream region of Medicago *NIN* is activated in the epidermis, 24h after Nod factor application (Vernié et al., 2015). Because this promoter region lacks the regulatory sequences required for nodule organogenesis, we wanted to determine whether expression of *NIN* is induced in inner root cell layers during primordium formation. We studied the localization pattern of *NIN* mRNA in nodule primordia via *in situ* hybridization. Plants were grown on plates and spot inoculated with *S. meliloti*. We analyzed a primordial stage at 2 dpi in which the pericycle cells have divided and some anticlinal divisions have occurred in the inner cortical cell layers C4 and C5 **(Figure 7A,B)**. We also analyzed cells at a slightly later stage at 3 dpi when cortical cells have divided more extensively **(Figure 7CD)**. At both stages, the infection thread had not yet reached the primordia. At the younger stage, *NIN* mRNA occurred in pericycle and epidermis, but it was hardly detectable in the divided cortical cells **(Figure 7A)**. The highest expression level occurred in the pericycle-derived cells. At the stage, when cortical cells have divided more extensively, the expression level of *NIN* in cortex derived cells was similar to that in the pericycle **(Figure 7C)**. This shows that expression of *NIN* was first strongly induced in the pericycle after which it extends to the other inner cell layers. *NF-YA1* is a known direct target of NIN (Soyano et al., 2013). Like *NIN,* it is expressed in the epidermis, where it regulates rhizobial infection (Laporte et al., 2014). To test whether NIN might also regulate *NF-YA1* expression in the primordia, we performed RNA *in situ* hybridization using *NF-YA1* as a probe. This analysis showed that *NF-YA1* expression is similar to *NIN* expression, as it is also first induced in pericycle and most likely cortical cell division precedes *NIN* and *NF-YA1* expression **(Figure 7B,D)**. This suggest that *NF-YA1* is regulated by *NIN* in both pericycle and other nodule primordium cells.

Therefore, rhizobia present in the epidermis induce *NIN* and *NF-YA1* expression in the pericycle-derived cells. Furthermore, because Nod factors are immobile molecules (Goedhart et al., 2000) that do not diffuse to the inner cell layers, *NIN* and *NF-YA* expression in the inner cell layer is most likely induced by a mobile signal generated in the epidermis where Nod factor signaling takes place.

The CE Region is required for induction of *NIN* expression in the pericycle.
We wanted to test whether the CE region is required for *NIN* expression in the inner cell layers, we compared expression patterns of *ProNINCE-5kb:GUS* and *ProNIN5kb:GUS* in roots. We first introduced these constructs into wild type Medicago (A17) roots *by A. rhizogenes* mediated transformation. We analyzed an early stage of primordium development when pericycle cells have divided and some anticlinal divisions have occurred in the inner cortical cell layers similar to the early stage tested for *in situ*. Both constructs were expressed in epidermis, pericycle, and endodermis and a lower signal was detected in some cortical cells **(Figure 8A,B)**. This result is surprising considering that *ProNIN5kb:NIN* is not sufficient for primordium formation in the *nin-1* background. Therefore, we hypothesized that expression of *ProNIN5kb:GUS* in inner cell layers is induced by endogenous NIN that is produced in the wild-type background. This implies that *NIN* expression in the inner layers is regulated by a positive feedback loop involving NIN itself, and that the essential *cis*-regulatory elements required for this are located in the –5 kb promoter region.

To test this hypothesis, we introduced *ProNINCE-5kb:GUS* and *ProNIN5kb:GUS* into *daphne-like* by *A. rhizogenes* mediated transformation. In *daphne-like*, infection threads can be formed indicating that *NIN* is induced in the epidermis and the production of the mobile signal might not be affected. However, nodule primordium formation is impaired, indicating there is no NIN production in the inner cell layers. Indeed, *ProNIN5kb:GUS* transgenic roots showed *GUS* expression only in epidermis and outer cortex **(Figure 8C)**, whereas no expression was observed in the pericycle cells. In contrast, *ProNINCE-5kb:GUS* transgenic roots showed *GUS* expression in epidermis, outer cortex and in the pericycle **(Figure 8D**). In this case, cell division was not induced in the pericycle, due to the absence of NIN. Taken together, these results demonstrate that the CE region regulates *NIN* expression in the pericycle prior to cell division in wild type roots. This means that the CE region is required for the initial induction of *NIN* expression in the pericycle. In addition, the expression of *ProNINCE-5kb:GUS* in the pericycle of *daphne-like* is weak which is consistent with the involvement of NIN in a feedback loop by (directly or indirectly) positively regulating its own expression.

To further demonstrate that the CE region is required for *NIN* expression in the pericycle, we studied *NIN* expression in *daphne-like* roots using RNA *in situ* hybridization at 2 dpi with rhizobia. In contrast to wild type **(Figure 7A)**, *NIN* is expressed in the epidermis and outer cortex but not in the pericycle **(Figure 7E).** This result supports the idea that that CE region is required for *NIN* expression in the pericycle.

Induction of *NIN* in the pericycle depends on *NIN* expression in the epidermis. It is likely that a mobile signal generated by Nod factor signaling in the epidermis induces *NIN* expression in the pericycle. If true, *NIN* expression in the pericycle would depend on *NIN* induction in the epidermis. To test this, we introduced *ProNINCE-5kb:GUS* and *ProNIN5kb:GUS* into *nin-1* by hairy root transformation. In both cases, GUS was present only in the epidermis and outer cortex, and not in the pericycle at 3 dpi **(Figure 8,F**). This suggests that NIN is required in the epidermis, probably for the generation of the mobile signal, in order to induce *NIN* expression in pericycle cells.

***CRE1* and *RR1* are expressed in the pericycle of uninoculated roots.** Rhizobium-induced *NIN* expression in the pericycle is dependent on the CE region and formation of precedes nodule primordia. The occurrence of multiple B-type RR response regulatory elements in the CE region suggests that the cytokinin signalling machinery is important for *NIN* transcriptional activation in the pericycle. To examine this, we determined the expression pattern of the cytokinin receptor *CRE1* and its

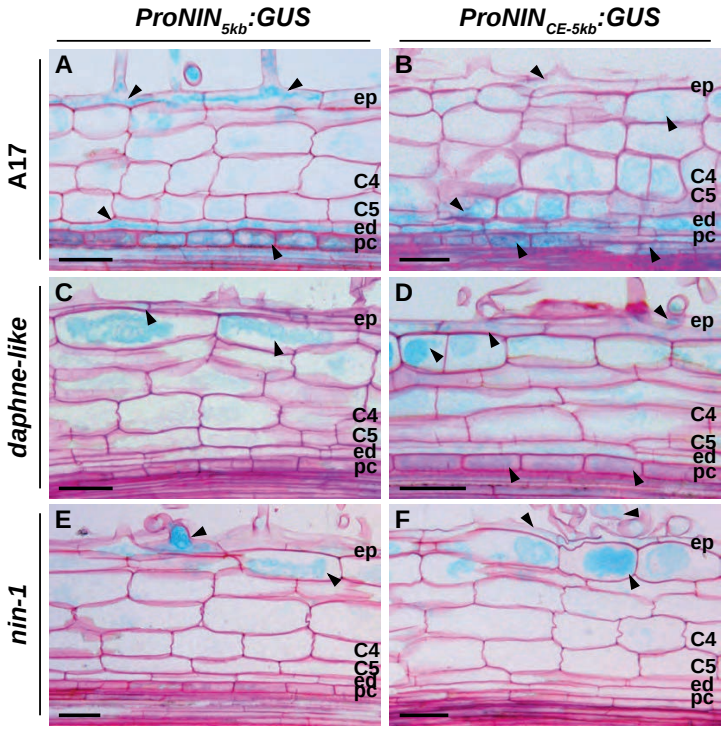


Figure 8. The CE Region is required for Rhizobium-induced *NIN* expression in the pericycle. **(A)** to **(F)** Tissue-specific ProNIN5kb:GUS and ProNINCE-5kb:GUS expression patterns in wild type and *nin* mutants at 2 dpi. Arrowheads indicate GUS expression (light blue) in wild type **(A)** and **(B)**, in *daphne-like* **(C)** and **(D)** and in *nin-1* **(E)** and **(F)** roots. ep, epidermis; C4, C5, cortical cell layers 4 and 5; ed, endodermis; pc, pericycle. Bars = 50 µm.

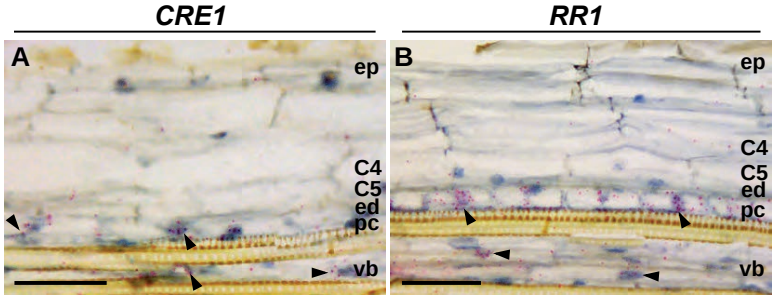


Figure 9. *CRE1* and *RR1* are expressed in the pericycle of uninoculated roots
RNA in situ localization of CRE1 **(A)** and RR1 **(B)** in the susceptible zone of uninoculated roots. For in situ hybridization, root tips of ~1 cm of 4-d-old seedlings were used. Hybridization signals are visible as red dots (arrowhead). ep, epidermis; C4, C5, cortical cell layers 4 and 5; ed, endodermis; pc, pericycle; vb, vascular bundle. Bar = 50 µm

putative target the B-type RESPONSE REGULATOR RR1, which is known to be expressed during nodule formation (Gonzalez-Rizzo et al., 2006). Using RNA *in situ* hybridization, we found that *CRE1* is actively transcribed in pericycle and vasculature cells of uninoculated roots, but not in endodermal or cortical cells **(Figure 9A)**. Also, mRNA of the B-type RR1 was present at the highest level in pericycle cells, and to a lower extent in root vasculature cells **(Figure 9B)**. Therefore, both *CRE1* and *RR1* have been already expressed in the pericycle by the time rhizobial signalling starts, suggesting that, initially, only this layer is responsive to cytokinin.

Discussion

In this study, we show that remote upstream regulatory sequences (the CE region) are required for proper regulation of *NIN* expression and *Medicago truncatula* nodule organogenesis. By contrast, regulatory sequences required for the infection process are located within a 5 kb region directly upstream of the start codon. The CE region contains several putative cytokinin response elements and domain 1 (D1), which contains six of these elements, is essential for nodule primordia formation. The CE region appears to be important for cytokinin induced expression of *NIN,* as *daphne-like* has lost this ability. Formation of nodule primordium initiates with *NIN* induction in the pericycle, and subsequently it extends to the cortical cells. The fact that *CRE1* and *RR1* are expressed in the pericycle supports the idea that cytokinin perception is necessary for the induction of *NIN* at the start of primordium formation.

In animals, many genes have been identified that are regulated by remote *cis*-regulatory elements that can be megabases away from the transcription start site. By contrast, in plants only a few remote *cis*-regulatory sequences are known (Shlyueva et al., 2014; Weber et al., 2016; Symmons and Spitz, 2013). One of the best characterized remote *cis*-regulatory sequence is the enhancer of *booster1* (*b1*)

in *Zea mays,* which is located 100 kb upstream of the gene (Stam et al., 2002). It has been shown that gene activation by remote enhancers can be associated with chromatin loop formation that brings the enhancer in close proximity to the promoter, a process that can facilitate assembly of transcription complexes (Cook, 2003; Nolis et al., 2009; Deng et al., 2012). The distance between the CE region and the transcription start site varies in the legume species studied here. In *L. angustifolius* it is about –7 kb, whereas in Lotus it is about –45 kb. We demonstrated that the CE region fused to the –5-kb promoter can rescue nodule organogenesis in Medicago. This shows that the sequences between the CE region and the –5 kb region are not essential for nodule organogenesis. However, we cannot exclude the possibility that in this region there are regulatory sequences required for fine tuning *NIN* expression.

During the infection process, NIN participates in a mechanism wherein root hair growth stops when a proper curl is formed. Regulatory sequences required for this process are located within the –2.2 kb promoter region. The fact that this region lacks the putative CYCLOPS binding site implies that in addition to CYCLOPS (IPD3 in Medicago), another transcription factor or factors is involved in regulating *NIN* expression in the epidermis. Because this –2.2 kb region is not sufficient for efficient infection thread formation, we assume that the expression level of *NIN* in the epidermis remains below the threshold level required for infection thread formation, whereas this level can be reached by the –5 kb promoter region which includes the putative CYCLOPS binding site **(Figure 10)**.

We present a model for the regulation of expression of *NIN* in **Figure 10**. After the rapid induction of *NIN* in the epidermis, *NIN* is subsequently induced in the pericycle. The latter most likely precedes the mitotic activation of pericycle cells. The induction of *NIN* in the pericycle requires the presence of the CE region and involves a positive feedback loop including NIN itself. The proposed feedback loop was based on our observation that expression of *ProNIN5kb:GUS* in the Medicago wild type background was induced in nodule primordia, despite the fact that this promoter region is not sufficient to trigger primordium formation. This result is similar to what was found in Lotus where a promoter region of *NIN* that does not trigger primordium formation was sufficient to drive expression of GUS in primordia (Yoro et al., 2014; Heckmann et al., 2011; Kosuta et al., 2011).

Our conclusion that nodule primordium formation requires the induction of *NIN* expression in inner root layers is consistent with the observation that nodule organogenesis is restored in the Lotus *daphne* mutant by *NIN* driven by a heterologous Arabidopsis enhancer that is active in endodermis and cortex (Yoro et al., 2014). When we transformed the Medicago *nin* null mutant with the *ProNIN2.2kb:NIN* construct, nodule organogenesis was not restored. By contrast, Vernie et al. 2015

reported the formation of nodules on a Medicago *nin* null mutant transformed with a similar construct. However, the nodule number was very low and nodules were observed a long time (50 d) after inoculation. To determine whether these structures are indeed genuine nodules and not simply modified lateral roots, analysis of sections is required.

Deletion of sequences within the CE region, which contains six putative cytokinin response elements, blocks primordium formation. We hypothesize that cytokinin signalling in the pericycle induces *NIN* expression. This hypothesis is supported by the fact that the expression of the cytokinin receptor (*CRE1*) and the B-type response regulator (*RR1*) is observed in the pericycle before rhizobial signalling is initiated. This agrees with a previous study showing that a *CRE1* promoter region driving *GUS* expression is specifically expressed in endodermis/pericycle cells opposite

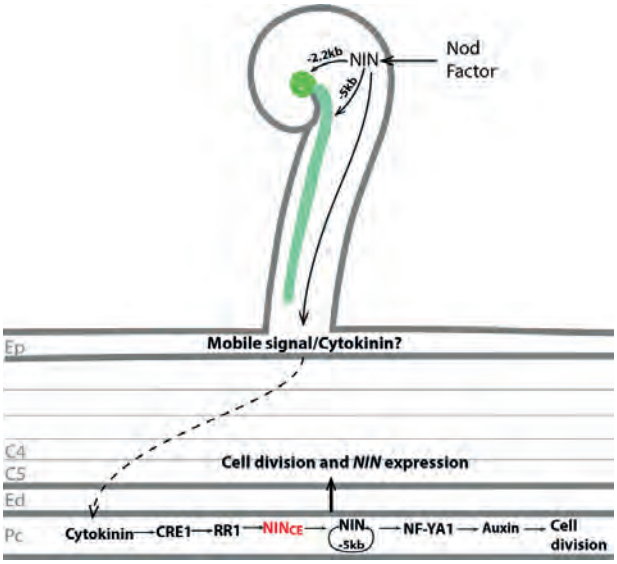


Figure 10. Proposed Model of *NIN* Function during Initiation of Nodule Primordia. After perception of the Nod factor, *NIN* expression is induced in the epidermis. The –5-kb regulatory region of the NIN promoter is sufficient for both tight root hair curling and infection thread formation. By contrast, expression driven by the –2.2 kb region is sufficient only for the tight root hair curling and formation of bacterial colonies inside the curl. A mobile signal is generated in the epidermis in a NIN-dependent manner, and this signal translocates to the pericycle. Whether or not this mobile signal is cytokinin or an unknown signal, it causes cytokinin accumulation in the inner root cell layers. The CRE1 receptor in the pericycle perceives cytokinin and activates the B-type RR1, which further activates NIN expression. NIN directly or indirectly regulates its own expression via a positive feedback loop, and the –5 kb promoter region is sufficient for this feedback regulation. NIN directly activates *NF-YA1* expression and stimulates further cell divisions. Later, the NIN-induced response in pericycle contributes to cell division and *NIN* expression in the endodermis and cortical cells. ep, epidermis; C4, C5, cortical cell layers 4 and 5; ed, endodermis; pc, pericycle

the protoxylem poles (Boivin et al., 2016), the sites where nodule primordia are formed (Heidstra et al., 1997). The involvement of the CE region in cytokinin induced *NIN* expression is indicated by the *daphne-like* mutant, which has lost this ability. However, it remains to be proven that the CE region is sufficient for the cytokinin induced *NIN* expression. Our attempts to show this in *Agrobacterium*-mediated hairy root transformation were inconclusive due to high basal levels of *NIN* expression in the absence of applied cytokinin in hairy roots.

The CE region is conserved in the eight legume species that we studied. They belong to different clades of the legume Papilionoideae subfamily, representing the Genistoids, IRLC, Robinioids, Milletioids and Dalbergioids clades. This suggests that the regulation of *NIN* expression by cytokinin is conserved in this subfamily. After the induction of *NIN* in the pericycle, *NIN* expression extends to the endodermis and inner cortex. In young nodule primordia in which cortical cells have divided anticlinally **(Figure 7A,B)**, expression of both *NIN* and *NF-YA1* are highest in pericycle, and it is hardly detectable in the divided cortical and endodermal cells. This suggests that NIN-induced responses in the pericycle contribute to cell division in endodermis and cortical cells **(Figure 10)**. At a later stage of development, *NIN* is expressed in the dividing cortical cells **(Figure 7C,D)**. How *NIN* expression is regulated in these cells remains to be studied. Expression of cytokinin biosynthesis genes as well as bioactive cytokinin accumulation is induced by Nod factor signalling (Van Zeijl et al., 2015). As it is a mobile molecule, it is likely that this results in the accumulation of cytokinin in (at least) the pericycle. Whether cytokinin biosynthesis in the pericycle is triggered by an as yet unknown mobile signal generated in epidermis or whether cytokinin itself is this mobile signal that accumulates in the pericycle is not known.

Cell division in nodule primordia correlates with auxin accumulation, and this occurs before the first cell division (Mathesius et al., 1998; Suzaki et al., 2012). Auxin accumulation (*DR5* expression) depends on NIN, as it does not occur in a *nin* null mutant (Suzaki et al., 2012). Furthermore, ectopic expression of both *NIN* and *NF-YA1* is sufficient to induce abnormal cell division during lateral root development (Soyano et al., 2013), suggesting that their expression causes the local accumulation of auxin. Therefore, we hypothesize that cytokinin signalling in the pericycle triggers *NIN* expression and that this results in the local accumulation of auxin, which subsequently triggers mitotic activity (Figure 10). This is supported by a previous study showing that *STY* genes are targets of NF-YA1 (Hossain et al., 2016). *STY* genes encode transcription factors that have been shown to regulate YUCCA auxin biosynthesis genes in *Arabidopsis* (Eklund et al., 2010a, 2010b; Sohlberg et al., 2006). If this is the case, then during nodule primordium formation, NIN induced *NF-YA* expression in the pericycle might induce the local production of auxin which subsequently induces cell division in pericycle, endodermis and cortex.

Materials and Methods

Plant material and growth, hairy root transformation and inoculation with Rhizobia

Medicago (*Medicago truncatula*) ecotype Jemalong A17 was used as the wild type. *Agrobacterium* msu 440-mediated hairy root transformation was performed according to Limpens et al., 2004. Medicago plants were grown in perlite saturated with low nitrate [0.25 mM Ca(NO3)2] Färhaeus (Fa) medium (Catoira et al., 2000) at 21°C under a 16h light/8h dark regime. After one week of growth, plants were inoculated with *S. meliloti 2011* constitutively expressing GFP or carrying the *PronifH:GFP* reporter (OD600 = 0.1, 1 mL per plant). Plants growing on Fä plates were spot inoculated with 0.5 µL of rhizobium suspension per root.

DNA constructs

DNA fragments of *NIN* including the 3'UTR and promoter regions were generated by PCR using Medicago genomic DNA as a template and Phusion high-fidelity DNA polymerase (Finnzymes) with the specific primers listed in Supplemental Table 2. The DNA fragments used for pENTR-D-TOPO cloning (Invitrogen) were amplified with forward primers containing an extra 5'-CACC sequence. Forward primers containing an attB4 site (GGGGACAACTTTGTATAGAAAAGTTGNN) and reverse primers with an attB1 site (GGGGACTGCTTTTTTGTACAAACTTGN) were used to generate DNA fragments for cloning into pDONOR P4-P1 by BP recombination (Invitrogen). The forward primers with attB2 (GGGGACAGCTTTCTTGTACAAAGTGGAA) and reverse primers with attB3 (GGGGACAACTTTGTATAATAAAGTTGC) were used to amplify DNA fragments for cloning into pDONOR P2-P3. To generate deletions (D1/D2/D3) in the CE region and deletion of the putative CYCLOPS binding site in the –5 kb region, two rounds of PCR were performed. In the first round, two DNA fragments that are separated by the deletion were amplified with specific primers to introduce a 15 bp overhang **(Supplemental Table 2)**. Subsequently, the PCR products were purified and mixed and 5 µL of this mixture was used as a template in a second round of PCR with ProNINCE-F and ProNINCE-R or ProNIN5kb-F and ProNIN5kb-R primers (Supplemental Table 2). This allowed creation of a single amplicon with a deletion in either the CE or the –5 kb regions. The Entry vectors were recombined into the modified Gateway binary vector pKGW-RR-MGW (Ovchinnikova et al., 2011) using Multisite LR recombination (Invitrogen).

Histological analysis and microscopy

Transgenic roots carrying the *ProNIN:GUS* constructs were incubated in GUS buffer [3% sucrose, 10 mM EDTA, 2 mM k-ferrocyanide, 2 mM k-ferricyanide, 0.5mg/mL X-Gluc in 0.1M phosphate buffer (pH = 7)] at 37°C for 1 to 2 h. Embedding of plant tissue in plastic, sectioning and tissue staining were performed as described in Xiao

et al., 2014. Sections were analysed using a DM5500B microscope equipped with a DFC425C camera (Leica). Bright-field and fluorescence images of transgenic roots and nodules were taken using a stereo macroscope (M165 FC, Leica). Confocal images were taken with an SP8 (Leica) microscope, using excitation wavelengths of 488 nm and 543 nm for GFP and propidium iodide respectively.

RNA isolation and qRT-PCR

RNA was isolated from one-week-old A17 and *daphne-like* roots using the EZNA Plant RNA mini kit (Omega Bio-tek, Norcross, GA, USA). For cDNA synthesis, 1 µg of this RNA was used with the iScript cDNA synthesis kit (Bio-Rad). Real-time qPCR was performed in 10 µL reactions using SYBR Green Supermix (Bio-Rad) and a CFX real-time system (Bio-Rad). Gene expression levels were determined using the primers listed in Supplemental Table 2. The gene expression was normalized using *ACTIN2* as a reference gene.

Quantification of colonies, infection threads and nodules

To quantify the number of curled root hairs containing colonies or infection threads, more than 20 transgenic roots (5 to 10 cm long) were cut into fragments of ~1 cm and randomly selected for counting. To quantify the nodule number per root, 5 to 10 cm long transgenic roots were selected.

RNA *in situ* hybridization

Medicago roots were fixed with 4% paraformaldehyde mixed with 3% glutaraldehyde in 50 mM phosphate buffer (pH = 7.4) and were then embedded in paraffin (Paraplast X-tra, McCormick Scientific). Root sections (7 µm) were prepared using a RJ2035 microtome (Leica). RNA *in situ* hybridization was conducted using Invitrogen™ ViewRNA™ ISH Tissue 1- Plex Assay kits (ThermoFisher Scientific) and was performed according to the user manual, which can be accecced at https://cdn.panomics.com/. RNA ISH probe sets were designed and synthesized by request at ThermoFisher Scientific. Catalogue numbers of probes for Medicago genes: VF1-20312 for *NIN*, VF1-6000865 for *CRE1*, VF1-6000866 for *RR1* and VF-20311 for *NF-YA1*. A typical probe set consisted of ~20 pairs of oligonucleotide probes (20-nt long) that hybridize to specific regions across the target mRNA. Each probe was composed of a region of ~20 nucleotides, a short linker region, and a tail sequence. The two tail sequences (double Z) together form a site for signal amplification. This design controls increased background by reducing the chance of a nonspecific hybridization event being amplified. For the nodulation specific genes, we used uninoculated roots as a negative control. For ISH with *CRE1* and *RR1* performed on non-inoculated roots of 4-d-old seedlings, we used *ENOD2* (nodule-specific gene) probe set as a negative control. Images were taken with an AU5500B microscope equipped with a DFC425c camera (Leica).

Map-based cloning of *daphne-like*

A segregating F2 populations resulting from a cross between FN8113 (cv Jemalong A17) and Jemalong A20 (118 plants) was made. DNA was extracted using a standard CTAB DNA miniprep method (Taylor and Powell, 1982). Initially, simple sequence repeat markers based on Mun et al. 2006 (Mun et al., 2006) were used to determine the global chromosomal location of the FN8113 locus. Subsequently, additional SSR markers were developed for the FN8113 locus on chromosome 5, and were used for chromosome walking. PCR was performed using 100 ng of genomic DNA and was analysed on 2.5 % agarose gels. The SSR marker JH5.17 (Supplemental Table 2) on BAC clone CU424494 showed the closest linkage to the FN8113 locus. No crossovers were found at the distal end of chromosome 5. Next, whole genome sequencing (Illumina Hiseq2000, paired-end) was used to identify mutations in the genomic region identified from the genetic mapping. The genomic sequence of the mutated region is provided in **Supplemental Data Set 1**. Cleaned DNA sequence reads were mapped against the Medicago genome (Young et al., 2011) using the bwa_mem algorithm (Li and Durbin, 2010). Clipped reads and mismapped mate pairs revealed an interchromosomal translocation, and this was further confirmed by aligning reads spanning the mutation to the genome using BLASTN (https://blast.ncbi.nlm.nih.gov/Blast.cgi).

Alignment of upstream regions of *NIN*

Most of the alignment work used Geneious v8.1.9 (https://www.geneious.com) (Kearse et al., 2012). The Medicago NIN protein sequence was analyzed using custom BLAST databases and Geneious v8.1.9 (Altschul et al., 1990; Kearse et al., 2012). A diverse selection of legume species with a good quality of publicly available genomic sequences were used: *Medicago truncatula* (Young et al., 2011; Tang et al., 2014), *Lotus japonicus* (Sato et al., 2008), *Arachis duranensis* (Bertioli et al., 2016), *Cicer arietinum* (Varshney et al., 2013), *Glycine max* (Schmutz et al., 2010), *Lupinus angustifolius* (Hane et al., 2017), *Cajanus cajan* (Varshney et al., 2012) and *Trifolium pratense* (De Vega et al., 2015). Selected *NIN* scaffolds (Supplemental Table 1) and up to 80 kb of upstream sequence and 10 kb of downstream sequence of *NIN* were extracted. Selected sequences were custom aligned using the mVISTAs web-based alignment tool (http://genome.lbl.gov/vista/mvista) (Frazer et al., 2004). The alignment program selected was the shuffle-lagan global alignment program, which detects rearrangements (Brudno et al., 2003). In addition to this larger scale alignment, individual alignments were made using MAUVE as a Geneious plugin (Darling et al., 2004). This better allowed for more precise determination of conserved sequences relative to the *NIN* start codon in all species. A complete overview of detected conserved regions can be found in **Supplemental Table 1**.

Alignment of CE regions and prediction of binding sites

Detected conserved sequences of CE regions for selected scaffolds (Supplemental Table 1)were aligned using MAFFTv7.017 as Geneious plugin (Katoh, 2002). Conservedbinding sites were predicted by using PlantPAN2.0 (Chow et al., 2016). Some sites were manually added based on homology with known putative B-type RR binding sequences (Heyl and Schmülling, 2003; Hosoda et al., 2002b; Imamura et al., 2003).

Accession Numbers

Sequence data from this article can be found in the GenBank/EMBL or Mt4.0v1 databases under the following accession numbers: *NIN* (Medtr5g099060), *CRE1* (Medtr8g106150), *NF-YA1* (Medtr1g056530), and *RR1* (Medtr3g102600).

Supplemental Data belonging to this chapter

Supplemental data is available online at The Plant Cell: doi:10.1105/tpc.18.0047

Supplemental Figure 1. MAFFT Alignment of the Putative CYCLOPS Binding Site of Eight Legume Species.

Supplemental Figure 2. MAFFT Alignment of the 472-bp Conserved Region of Eight Legume Species.23

Supplemental Figure 3. *nifH* Expression Is Induced in *ProNINCE-5kb:NIN* Transgenic *nin-1* Root Nodules.

Supplemental Figure 4. Phenotype of *nin-1* Transformed with *ProNINCE(ΔD1/D2/D3)-5kb:NIN* Constructs.

Supplemental Table 1. Sequence Information of Aligned Species.

Supplemental Table 2. Primers Used in This Study.

Supplemental Data Set 1. Genomic Sequence of the Mutated Region in FN8113.

Acknowledgements

This research was supported by the European Commission (ERC-2011-AdG-294790), NWO (Netherlands Organization for Scientific Research)-VENI (863.15.010), NWO-VICI (865.13.001), and by the China Scholarship Council (JL: 201506300062).

References

Altschul, S.F., Gish, W., Miller, W., Myers, E.W., and Lipman, D.J. (1990). Basic local alignment search tool. J. Mol. Biol. 215: 403–410.

Andriankaja, A., Boisson-Dernier, A., Frances, L., Sauviac, L., Jauneau, A., Barker, D.G., and de Carvalho-Niebel, F. (2007). AP2-ERF transcription factors mediated Nod factor dependent Mt ENOD11 activation in root hairs via a novel cis-regulatory motif. Plant Cell 19: 2866–2885.

Benfey, P.N. and Chua, N.-H. (1990). The Cauliflower Mosaic Virus 35S Promoter: Combinatorial Regulation of Transcription in Plants. Science (80). 250: 959–966.

Bertioli, D.J. et al. (2016). The genome sequences of Arachis duranensis and Arachis ipaensis, the diploid ancestors of cultivated peanut. Nat. Genet. 48: 438–446.

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.

Brudno, M., Malde, S., Poliakov, A., Do, C.B., Couronne, O., Dubchak, I., and Batzoglou, S. (2003). Glocal alignment: Finding rearrangements during alignment. Bioinformatics 19.

Catoira, R., Galera, C., Billy, F. De, Penmetsa, R.V., Journet, E., Maillet, F., Rosenberg, C., Cook, D., Gough, C., and Dénarié, J. (2000). Four Genes of Medicago truncatula Controlling Components of a Nod Factor Transduction Pathway. 12: 1647–1665.

Chow, C.N., Zheng, H.Q., Wu, N.Y., Chien, C.H., Huang, H. Da, Lee, T.Y., Chiang-Hsieh, Y.F., Hou, P.F., Yang, T.Y., and Chang, W.C. (2016). PlantPAN 2.0: An update of Plant Promoter Analysis Navigator for reconstructing transcriptional regulatory networks in plants. Nucleic Acids Res. 44: D1154–D1164.

Cook, P.R. (2003). Nongenic transcription, gene regulation and action at a distance. J. Cell Sci. 116: 4483–4491.

Darling, A.C.E., Mau, B., Blattner, F.R., and Perna, N.T. (2004). Mauve : Multiple Alignment of Conserved Genomic Sequence With Rearrangements Mauve : Multiple Alignment of Conserved Genomic Sequence With Rearrangements. Genome Res. 14: 1394–1403.

Deng, W., Lee, J., Wang, H., Miller, J., Reik, A., Gregory, P.D., Dean, A., and Blobel, G.A. (2012). Controlling long-range genomic interactions at a native locus by targeted tethering of a looping factor. Cell 149: 1233–1244.

De Vega, J.J. et al. (2015). Red clover (Trifolium pratense L.) draft genome provides a platform for trait improvement. Sci. Rep. 5.

Downie, J.A. (2014). Legume nodulation. Curr. Biol. 24: R184–R190.

Ehrhardt, D.W., Wais, R., and Long, S.R. (1996). Calcium spiking in plant root hairs responding to rhizobium modulation signals. Cell 85: 673–681.

Eklund, D.M., Staldal, V., Valsecchi, I., Cierlik, I., Eriksson, C., Hiratsu, K., Ohme-Takagi, M., Sundstrom, J.F., Thelander, M., Ezcurra, I., and Sundberg, E. (2010a). The Arabidopsis thaliana STYLISH1 Protein Acts as a Transcriptional Activator Regulating Auxin Biosynthesis. Plant Cell 22: 349–363.

Eklund, D.M., Thelander, M., Landberg, K., Staldal, V., Nilsson, A., Johansson, M., Valsecchi, I., Pederson, E.R.A., Kowalczyk, M., Ljung, K., Ronne, H., and Sundberg, E. (2010b). Homologues of the Arabidopsis thaliana SHI/STY/LRP1 genes control auxin biosynthesis and affect growth and development in the moss Physcomitrella patens. Development 137: 1275–1284.

Fournier, J., Teillet, A., Chabaud, M., Ivanov, S., Genre, A., Limpens, E., Carvalho-niebel, F. De, and Barker, D.G. (2015). Remodeling of the Infection Chamber before Infection Thread Formation Reveals a Two-Step Mechanism for Rhizobial Entry into the Host Legume Root Hair. Plant Physiol. 167: 1233–1242.

Frazer, K.A., Pachter, L., Poliakov, A., Rubin, E.M., and Dubchak, I. (2004). VISTA: Computational tools for comparative genomics. Nucleic Acids Res. 32: 273–279.

Goedhart, J., Hink, M.A., Visser, A.J.W.G., Bisseling, T., and Gadella, T.W.J. (2000). In vivo fluorescence correlation microscopy (FCM) reveals accumulation and immobilization of Nod factors in root hair cell walls. Plant J. 21: 109–119.

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 Online 18: 2680–2693.

Hane, J.K. et al. (2017). A comprehensive draft genome sequence for lupin (Lupinus angustifolius), an emerging health food: insights into plant–microbe interactions and legume evolution. Plant Biotechnol. J. 15: 318–330.

Hayashi, T., Shimoda, Y., Sato, S., Tabata, S., Imaizumi-Anraku, H., and Hayashi, M. (2014). Rhizobial infection does not require cortical expression of upstream common symbiosis genes responsible for the induction of

Ca2+spiking. Plant J. 77: 146–159.

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., van Kammen, A., and Bisseling, T. (1997). Ethylene provides positional information on cortical cell division but is notinvolved in Nod factor-induced root hair tip growth in Rhizobium-legumeinteraction. Development 124: 1781–1787.

Heyl, A. and Schmülling, T. (2003). Cytokinin signal perception and transduction. Curr. Opin. Plant Biol. 6: 480–488.

Horváth, B. 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.

Hosoda, K., Imamura, A., Katoh, E., Hatta, T., Tachiki, M., Yamada, H., Mizuno, T., and Yamazaki, T. (2002). Molecular Structure of the GARP Family of Plant Myb-Related DNA Binding Motifs of the Arabidopsis Response Regulators. Plant Cell 14: 2015–2029.

Hossain, M.S. et al. (2016). Lotus japonicus NF-YA1 Plays an Essential Role During Nodule Differentiation and Targets Members of the SHI/STY Gene Family. Mpmi 29: 950–964.

Imamura, A., Kiba, T., Tajima, Y., Yamashino, T., and Mizuno, T. (2003). In vivo and in vitro characterization of the ARR11 response regulator implicated in the His-to-Asp phosphorelay signal transduction in Arabidopsis thaliana. Plant Cell Physiol. 44: 122–131.

Jardinaud, M.-F. et al. (2016). A Laser Dissection-RNAseq Analysis Highlights the Activation of Cytokinin Pathways by Nod Factors in the *Medicago truncatula* Root Epidermis. Plant Physiol. 171: 2256–2276.

Katoh, K. (2002). MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. Nucleic Acids Res. 30: 3059–3066.

Kearse, M. et al. (2012). Geneious Basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data. Bioinformatics 28: 1647–1649.

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

Kosuta, S., Held, M., Hossain, M.S., Morieri, G., MacGillivary, A., Johansen, C., Antolín-Llovera, M., Parniske, M., Oldroyd, G.E.D., Downie, A.J., Karas, B., and Szczyglowski, K. (2011). Lotus japonicus symRK-14 uncouples the cortical and epidermal symbiotic program. Plant J. 67: 929–940.

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

Li, H. and Durbin, R. (2010). Fast and accurate long-read alignment with Burrows-Wheeler transform. Bioinformatics 26: 589–595.

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.

Marsh, J.F., Rakocevic, A., Mitra, R.M., Brocard, L., Sun, J., Eschstruth, A., Long, S.R., Schultze, 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.

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

Middleton, P.H. et al. (2007). An ERF Transcription Factor in Medicago truncatula That Is Essential for Nod Factor Signal Transduction. Plant Cell Online 19: 1221–1234.

Mitra, R.M., Gleason, C. a, Edwards, A., Hadfield, J., Downie, J.A., Oldroyd, G.E.D., and Long, S.R. (2004). A Ca2+/calmodulin-dependent protein kinase required for symbiotic nodule development: Gene identification by transcript-based cloning. Proc. Natl. Acad. Sci. U. S. A. 101: 4701–4705.

Mun, J.H. et al. (2006). Distribution of microsatellites in the genome of Medicago truncatula: A resource of genetic markers that integrate genetic and physical maps. Genetics 172: 2541–2555.

Nolis, I.K., McKay, D.J., Mantouvalou, E., Lomvardas, S., Merika, M., and Thanos, D. (2009). Transcription factors mediate long-range enhancer-promoter interactions. Proc. Natl. Acad. Sci. 106: 20222–20227.

Ovchinnikova, E. et al. (2011). IPD3 controls the formation of nitrogen-fixing symbiosomes in pea and Medicago Spp. Mol. Plant. Microbe. Interact. 24: 1333–44. 29

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.

Sato, S. et al. (2008). Genome structure of the legume, Lotus japonicus. DNA Res. 15: 227–239.

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

Schmutz, J. et al. (2010). Genome sequence of the palaeopolyploid soybean. Nature 463: 178–183.

Sheen, J. (2002). Phosphorelay and transcription control in cytokinin signal transduction. Science (80-.). 296: 1650–1652.

Shlyueva, D., Stampfel, G., and Stark, A. (2014). Transcriptional enhancers: From properties to genome-wide predictions. Nat. Rev. Genet. 15: 272–286.

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.

Sohlberg, J.J., Myrenås, M., Kuusk, S., Lagercrantz, U., Kowalczyk, M., Sandberg, G., and Sundberg, E. (2006). STY1 regulates auxin homeostasis and affects apical-basal patterning of the Arabidopsis gynoecium. Plant J. 47: 112–123.

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: 14607–12.

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.

Stam, M., Belele, C., Dorweiler, J.E., and Chandler, V.L. (2002). Differential chromatin structure within a tandem array 100 kb upstream of the maize b1 locus is associated with paramutation. Genes Dev. 16: 1906–1918.

Suzaki, T., Ito, M., and Kawaguchi, M. (2013). Genetic basis of cytokinin and auxin functions during root nodule development. Front. Plant Sci. 4: 1–6.

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

Symmons, O. and Spitz, F. (2013). From remote enhancers to gene regulation: Charting the genome's regulatory landscapes. Philos. Trans. R. Soc. B Biol. Sci. 368.

Tang, H. et al. (2014). An improved genome release (version Mt4.0) for the model legume Medicago truncatula. BMC Genomics 15: 1–14.

Taylor, B,, Powell, A. (1982). Isolation of plant DNA and RNA. Focus 4: 4–6.

Van Zeijl, A., Op Den Camp, R.H.M., Deinum, E.E., Charnikhova, T., Franssen, H., Op Den Camp, H.J.M., Bouwmeester, H., Kohlen, W., Bisseling, T., and Geurts, R. (2015). Rhizobium Lipo-chitooligosaccharide Signaling Triggers Accumulation of Cytokinins in Medicago truncatula Roots. Mol. Plant 8: 1213–1226.

Varshney, R.K. et al. (2013). Draft genome sequence of chickpea (Cicer arietinum) provides a resource for trait improvement. Nat Biotechnol 31: 240–246.

Varshney, R.K. et al. (2012). Draft genome sequence of pigeonpea (Cajanus cajan), an orphan legume crop of resource-poor farmers. Nat. Biotechnol. 30: 83–89.

Vernié, 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.

Wang, Y., Wang, L., Zou, Y., Chen, L., Cai, Z., Zhang, S., Zhao, F., Tian, Y., Jiang, Q., Ferguson, B.J., Gresshoff, P.M., and Li, X. (2014). Soybean miR172c targets the repressive AP2 transcription factor NNC1 to activate ENOD40 expression and regulate nodule initiation. Plant Cell 26: 4782–4801.

Weber, B., Zicola, J., Oka, R., and Stam, M. (2016). Plant Enhancers: A Call for Discovery. Trends Plant Sci. 21: 974–987.

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. 109: 633–638.

Yano, K. et al. (2008). CYCLOPS, a mediator of symbiotic intracellular accommodation. Proc. Natl. Acad. Sci. U. S. A. 105: 20540–20545.

Yoro, E., Suzaki, T., Toyokura, K., Miyazawa, H., Fukaki, H., and Kawaguchi, M. (2014). A positive regulator of nodule organogenesis, NODULE INCEPTION, acts as a negative regulator of rhizobial infection in Lotus japonicus. Plant Physiol. 165: 747–58.

Young, N.D. et al. (2011). The Medicago genome provides insight into the evolution of rhizobial symbioses. Nature 480: 520–524.

CHAPTER 7

7

General Discussion

Introduction

In the recent decades, the favoured hypothesis to explain the distribution of nitrogen-fixing nodule symbiosis was the occurrence of independent origins, shaped by a single predisposition event the last common ancestor of all nodulating plant species (Soltis et al. 1995; Werner et al. 2014). The commonalities in gene regulation, such as the use of transcription factor *NODULE INCEPTION* (*NIN)*, or the conservation of the common symbiosis signalling pathway, were regarded as the result of convergent evolution. In order to find gene gains that would correlate with a gain of function event, phylogenetic comparisons between nodulating and Non-nodulating lineages were suggested (Doyle 2016). *Parasponia* represents an excellent example for such a comparative study since it is nested within the non-nodulating *Trema* lineage. However, no high likelihood gene gains were identified that correlated to the nitrogen-fixation trait in *Parasponia*. In contrast, we identified multiple parallel losses of symbiosis genes in non-nodulating *Trema*, and other non-nodulating relatives in the Rosales order. These genes include orthologs of important symbiotic regulators like *NOD FACTOR NOD FACTOR PERCEPTION 2* (*NFP2*), symbiotic program regulator *NIN* and *RHIZOBIUM POLAR GROWTH* (*RPG*). All of these symbiotic genes have been proven to be indispensable for symbiotic infection (**Chapter 3.**) (L. Schauser et al. 1999; Arrighi et al. 2008). A complementary set of genome comparisons throughout the nitrogen-fixation clade has identified similar gene loss events of *NIN* and *RPG*. In this case, many more non-nodulating lineages throughout the nitrogen fixation clade were sampled (Griesmann et al. 2018). Strikingly no gene gains were identified in nodulating species that would explain the nitrogen-fixing trait. The most likely explanations for symbiotic gene loss events is that they happened independently and in parallel. These results are difficult to explain under the hypothesis of independent origins of nodulation. These findings have some implications for our understandings on the origins of root Nodulation. The hypothesis of independent origins of nodulation represents the most parsimonious scenario, requiring fewer evolutionary events, which can be supported by ancestral state phylogenetic reconstructions (Soltis et al. 1995; Swensen 1996; Werner et al. 2014; Li et al. 2015). Yet, in order to explain its confinement to the nitrogen fixation clade, it also presents the scientific community with a hard to conceptualize predisposition event. Since no tangible evidence for the existence of a symbiotic precursor state was identified to date, the data favours an alternative hypothesis; a single gain of nodulation in a common ancestor of the nitrogen-fixation clade, followed by loss of this trait in many lineages. This scenario was proposed as an alternative hypothesis, yet largely disregarded, since the earliest recognition that nitrogen-fixing root nodules occur in a single taxonomic clade (Soltis et al. 1995). Assuming that all nodulating plants are derived from a single nodulating common ancestor, it

can be predicted that many genes recruited for nodulation are orthologous. In this case, we find that *Parasponia* and *Medicago* have commonly recruited over 290 genes in nodulation **(Chapter 3.)**. Following their recruitment in symbiosis, some of these genes likely lost (in part) their ancestral functions. Outside of the nitrogen fixation clade, orthologs of *NIN, RPG* and *NFP* are clearly present. In these species, they likely perform a function independent of symbiotic interaction with rhizobia or *Frankia*. In the case of *RPG* and *NFP* this function possibly lies in the interaction with arbuscular mycorrhizal fungi. During the evolution of nitrogen fixation, these three genes have been recruited into symbiosis for important symbiotic function. *NFP2* for the perception of symbiotic LCO signals, *RPG* for infection thread progression and *NIN* as a transcription factor to control infection, nodule organogenesis and the integration with nitrate homeostasis. A mutation of any of such genes abolishes an effective nitrogen-fixing interaction in both legumes and in non-legumes. Inside the nitrogen fixation clade, the presence of these genes exclusively correlates with the ability to form nitrogen-fixing nodules.

Given that a single mutation can abolish an effective symbiotic interaction, losing nodulation is conceptually much easier than gaining it. Therefore an a priori equal weight for gains and losses in complex traits represents a fundamentally flawed assumption. Replacing the origin of a symbiotic precursor state with nodulation itself thus represent a simpler solution to explain the current distribution of nodulation. When a single origin of nodulation is inferred on the species three the number of losses required to explain the distribution of nodulation is at least 7 losses in Fagales, 5 in Cucurbitales, 17 in Rosales and 36 in Fabales (Robin van Velzen, Doyle, and Geurts 2019). Similar examples of parallel loss events and subsequent gene erosion have been shown in the loss of arbuscular mycorrhizal interactions in Brassicales and parasitic plant species (Delaux et al. 2014). Another example would be the loss of limbs in certain groups of vertebrates such as snakes and whales. Limb loss is associated with the loss of transcriptional enhancers and regulating genes (McGowen, Gatesy, and Wildman 2014; Infante et al. 2015; Roscito et al. 2018). In other traits, where parallel origins are predicted, such as the evolution of multicellularity in fungi, the role of losses has been recognized as at least a contributing driver for the observed evolutionary pattern (Nagy, Kovács, and Krizsán 2018). However, the co-option of existing genes and pathways was probably the dominant mechanism, for example recruiting genes for hyphal morphogenesis, while gene duplication was apparently less prevalent (Kiss et al. 2019). Therefore it seems that rather than the gain of completely new developmental modules, the distribution of nodulation can be best explained by the single co-option of existing developmental and signalling modules and subsequent loss of the trait. This results in two major still

outstanding questions. (i) What were the potential drivers for the parallel loss of the nitrogen-fixing symbiosis trait? (ii) How were existing pathways co-opted for rhizobial signalling and symbiotic organ development?

(I) Potential drivers for the parallel loss of the nitrogen-fixing symbiosis trait

The energetic cost of nitrogen fixation

A nitrogen-fixation nodule symbiosis in itself seems like the perfect nitrogen acquisition strategy. However, biological nitrogen-fixation is a costly process. Sixteen Moles of ATP are required to fix a mole of dinitrogen. This means nodulation has to be tightly controlled in regard to the available nitrate by autoregulation both in legumes, actinorhizal plants and *Parasponia* (Arnone, Kohls, and Baker 1994; Osipova et al. 2012; Krusell et al. 2011; Nishida et al. 2018; Dupin, Geurts, and Kiers 2020). Increased availability of nitrogen thus limits the amounts of effective nodulation and therefore limits the effective advantage of nodulation. Loss of nitrogen-fixing ability in high nitrate environments may be an adaptive strategy in specific niches. Since nitrate availability is likely only a local driver, it would not necessarily be able to drive the fixation of deleterious fix^- alleles in a large population. The presence of microbial "cheaters" or increased pathogen susceptibility has also been postulated as a potential fitness cost in mutualism (Friesen 2012; Jones et al. 2015). In general, all nodulating plants are vulnerable to the possibility of microbial "cheaters", bacteria that can enter the nodule but do not provide fixed nitrogen. A reason for this is that symbiotic signalling, needed for nodule formation and bacterial infection, mainly relies on cues which do not directly infer nitrogen-fixing capacity (Herder and Parniske 2009). Although host-symbiont co-evolution has driven many legumes to highly stringent symbiont recognition mechanisms, any symbiont can become a cheater by the mutation of only a few genes (Denison and Kiers 2004; Sachs, Ehinger, and Simms 2010). Generally, in legumes, ineffective nodules are combated by host sanctions, but this process may not have been as efficient in some lineages (Oono, Anderson, and Denison 2011; Friesen 2012). Similarly, pathogenic soil bacteria can be turned into colonizers by transfer of a symbiotic plasmid (Marchetti et al. 2014). Similar to nitrogen availability, the presence of microbial cheaters is likely a local selection driver (Denison and Kiers 2004).

Climate change as a global driver of evolution

Since nodulation is lost in parallel in different symbiotic lineages, selection pressure on a global scale is likely causal for the loss of the symbiosis trait. Such factors may represent global changes in temperature or atmospheric CO_2 concentration. Changes in atmospheric CO_2 levels have been a driving force behind the evolution of plant anatomy and physiology, such as C4-photosynthesis (P.-A. Christin et al. 2008;

Gerhart and Ward 2010; P. A. Christin et al. 2013; Heyduk et al. 2019). Atmospheric CO_2 concentrations have gradually decreased over the last 100 million years. During the Eocene–Oligocene transition ~34 Mya, the Late Oligocene ~29–23 Mya, and the Middle Miocene ~14 Mya the decline in CO_2 concentrations intensified (Zhang et al. 2013; Robin van Velzen, Doyle, and Geurts 2019). Elevated CO_2 levels showed to increase the nitrogen fixation capacity of plants, while low CO_2 attenuates fixation ability (Rogers et al. 2006; Vogel and Curtis 1995). Therefore during the drop in atmospheric CO_2 levels over the last 100-million years fixed carbon could have become a larger growth inhibitor than nitrogen availability. This would not only explain the global loss pattern of the nitrogen-fixing trait, but also the different timing of these losses (Robin van Velzen, Doyle, and Geurts 2019).

Alternative types of symbiosis

An alternative scenario to climatic factors as the sole driver of loss of the nitrogen-fixing nodulation trait is the gain of different, possibly more cost-effective, strategies for nutrient acquisition. Such strategies represent the formation of ectomycorrhizal associations (Werner et al. 2018). Ectomycorrhizal associations prove to be an effective phosphate and nitrogen acquisition strategy in a forest environment(Muller et al. 2007; Hoeksema et al. 2018). The presence of new mutualisms for nitrogen acquisition may in part contribute to the stabilization of potential loss of function alleles in a population. It is very hard, if not impossible, to infer what drove symbiotic gene losses to fixation in any given population tens of millions of years ago. Since mutations happened at different time points and in different species. The mechanism can very well be different for each lineage and most likely it is a combination of all of the above factors in some cases.

Evidence of a host-switch in *Parasponia*

In **Chapter 3.** we established that nitrogen-fixing symbiosis in the *Parasponia* lineage is an ancestral character, rather than a recent gain. However, as discussed in the introduction several lines of evidence suggest that *Parasponia* gained several symbiotic characteristics relatively recently **(Chapter 1.)**. One of these is the expression of cluster-I hemoglobin (HB1) in *Parasponia* nodules. *Parasponia* HB1 contains several adaptive mutations making it an efficient oxygen scavenger, while HB1 of non-nodulating *Trema* species does not have these mutations (Sturms et al. 2010; Kakar et al. 2011; R. van Velzen et al. 2018). Without the expression of hemoglobin an effective nitrogen-fixing symbiosis with rhizobia would not be possible. *Frankia* Actinobacteria, unlike rhizobia, are able to protect their Nitrogenase for oxidation by the formation of hopanoid containing vesicles (Berry et al. 1993). Therefore additional oxygen protection mechanisms are generally not required in actinorhizal symbiosis. Some Actinorhizal plants, like *Casuarina*, offer also oxygen

protection by hemoglobin expression, which most probably evolved more recently than hopanoid-based protection (Jacobsen-Lyon et al. 1995). The oxygen protection system of *Frankia* allows for a simpler evolutionary trajectory, since the evolution of elaborate oxygen protection structures is not an a priori requirement for effective nitrogen fixation with *Frankia*. Therefore the recent gain of HB1-based oxygen protection system in *Parasponia*, may, in fact, be a recent gain of rhizobium as microsymbiont over a pre-existing *Frankia* symbiosis. This implicates that a host switch occurred in *Parasponia*, from *Frankia* to rhizobium.

Actinorhizal symbiosis is generally considered to be older than rhizobium symbiosis. It can be hypothesized that nodulation first evolved with an LCO-producing *Frankia* strain. Different proteobacteria may have obtained LCO biosynthesis and nitrogen fixation genes by horizontal gene transfer (Bailly et al. 2007; Persson et al. 2015). Possibly first to γ-proteobacteria of the *Burkholderiales* order (Bontemps et al. 2010). This subsequently gave rise to the massive diversity of nodulating bacteria in the α-proteobacteria(Bontemps et al. 2010; Remigi et al. 2016). A host switch is also predicted to have occurred in the Fabales order giving rise to rhizobium nodulation (Robin van Velzen, Doyle, and Geurts 2019).

Several other lines of evidence suggest a recent gain of rhizobium symbiosis in the *Parasponia* lineage. First, it appears that all *Parasponia* species have lost the ortholog of thr *Lotus* LysM-receptor EPR3 **(Chapter 3.).** In *Lotus,* the LjEPR3 receptor is responsible for recognising exopolysaccharide decorations on the symbiont surface (Kawaharada et al. 2017). The loss of this receptor in *Parasponia* is peculiar since it is conserved in most nodulating and non-nodulating plants. The loss of *PanEPR* could have alleviated host restrictions on symbiont decorations, allowing the co-option of rhizobium as a symbiont. Further, *Parasponia* is promiscuous, interacting with a wide range of rhizobium symbionts belonging to different clades of α-Proteobacteria (Op den Camp et al. 2012). Another peculiarity is that *Parasponia* is highly susceptible to natural *Agrobacterium* transformation events. With *Parasponia* containing nine natural T-DNA insertions in its genome (Matveeva and Otten 2019). While *Agrobacterium* transformants occur in nature, the exceptionally large number of near to complete T-DNA sequences in *Parasponia* could be an indication of a recent gain of rhizobium symbiosis. The close relationship of rhizobia and agrobacteria together with the recent adoption of rhizobia as a symbiont may have increased the susceptibility to *Agrobacterium* infection. In line with this *Parasponia* nodules can be readily infected by *Agrobacterium* carrying SYM-plasmids, leading to ineffective, though infected, nodules (Bender et al. 1987). Similarly *Parasponia* commonly hosts ineffective strains in its nodules such as *Rhizobium leguminosarum* or *Rhizobium tropici (Trinick, Goodchild, and Miller 1989)* **(Chapter 3.)**. Therefore although nodules may be formed by many strains on *Parasponia,* many of these interactions

may prove ineffective. This supports the hypothesis that rhizobium was co-opted relatively recently as a host and only a few mechanisms for symbiont control are in place. Host-symbiont switches during the course of evolution of nitrogen fixation may be common. Symbiont switches, whether it is within or outside a taxonomic lineage, almost certainly require Host genome evolution. Therefore the plethora of different symbionts and different requirements for each symbiont may have driven the large variety of nodule types we see today.

(ii) The co-option of pre-existing pathways to evolve nodulation

Recruitment of LysM-type receptor kinases in the infection process

In **Chapter 3.** we established that among the symbiosis genes lost in *Trema* was putative LCO-receptor NFP2. This warranted a closer look at the evolution of LysM-type receptor kinases in the nitrogen fixation clade. LysM-type receptors belonging to the clade of NFP, also called LYR-Ia, are present in most plant species in or outside of the nitrogen fixation clade. Exceptions usually involve species that do not form an interaction with arbuscular mycorrhiza. In **Chapter 3.** and **4.**, phylogenetic comparisons have shown that *NFP* has duplicated at or near the root of the nitrogen fixation clade, giving rise to two subclades named NFP-I and NFP-II. NFP-I and NFP-II orthologous genes can be found in species of all four orders of the nitrogen fixation clade; Fabales, Fagales, Cucurbitales and Rosales. The clade of NFP-II contains known Nod Factor receptors of *Lotus, Medicago,* pea, soybean and *Parasponia,* while an NFP-I orthologous gene is present in most lineages except Fabales. NFP-I orthologs are not present in the currently available genomes for Fabales species. No pseudogenes could be detected which clearly belong to the NFP-I-clade in these genomes. The order Fables, Fagales, Cucurbitales and Rosales have arisen near-simultaneously during the evolution, while the branch of the legumes is often considered as the earliest split (Janssens et al. 2020; Robin van Velzen, Doyle, and Geurts 2019). Therefore the possibility exists that the duplication in the NFP (LYR-Ia) clade happened in the ancestor of Rosales, Fagales and Cucurbitales and thus was not present at the origin of all four orders. However, the close relationship of all NFP-II sequences suggests this is not the case **(Chapter 4).** Therefore it is plausible that the duplication happened at the root of the nitrogen fixation clade, although it remains best to approach the phylogeny with some caution.

A strict correlation between the presence of an NFP-II-type ortholog and nodulation exists in Cucurbitales and Rosales. NFP-II-type pseudogenes can be found in the nodulating Fagales *Castanea mollissima* and *Fagus sylvatica*, indicating a recent loss **(Chapter 4.)**. In Fagales, Rosales and Cucurbitales the NFP-I and NFP-II type gene copy were initially retained, while the NFP-II-type was favoured for LCO-signalling during nodulation, the NFP-I-type may be to be involved in arbuscular

mycorrhizal signalling or defence responses. In early Fabales ancestors, the NFP-I-ortholog was lost, leaving again a single NFP-II-type gene for both functions. Later, independent duplications of the NFP-II ortholog may have again allowed this functional divergence between a specific receptor for rhizobium symbiosis and one for mycorrhizal symbiosis. In species that can establish and ectomycorrhizal symbiosis rather than endomycorrhiza and that do not nodulate, such as *Castanea mollissima* and *Fagus sylvatica* (Fagales), both NFP-type were lost relatively recently, indicated by remnants of both types in their genome **(Chapter 4.)**. This may be due to the complete switch to ectomycorrhizal fungi as a nutrient acquisition strategy. Interestingly, also the Fagales nodulators *Casuarina* and *Alnus* lost the NFP-II-type gene. Both plants have *Frankia* cluster-I species as microsymbiont. Since cluster-I *Frankia* strains do not produce LCOs for symbiotic signalling, it is plausible that the evolution of a different signalling molecular allowed for the loss of the LCO receptor NFP-II in *Casuarina* and *Alnus*. In any case, it is clear that the presence of an NFP-II type NFP gene correlates with LCO-based nodulation. The correlation between NFP-II and nodulation is less strict in the Fabales, since *Cercis*, a known non-nodulating legume, contains an apparently functional NFP-II-type gene, possibly for arbuscular mycorrhizal interactions. In many legumes, independent duplications of this NFP-II-type gene has occurred, for example in the Papilionidae where a duplication gave rise to MtNFP/LjNFR5 and MtLYR1/LjLYS11 receptors in *Medicago* and *Lotus* (Gough et al. 2018). It appears that this duplication allowed a functional divergence of NFP-II functions since MtLYR1 and its direct ortholog LjLYS11 is considered a mycorrhizal responsive LysM-receptor (Gomez et al. 2009; Rasmussen et al. 2016). For example, *Castanospermum australe*, a known non-nodulator in the Papilionidae, specifically lost its *MtNFP/LjNFR5* orthologous gene but kept its LYR1-type copy **(Figure 3, Chapter 4.)**.

The co-option of LYR-Ia LysM-type receptors for the nodulation specific LCO-signal may have happened only once in evolutionary history and likely was depending on the duplication event at the root of the nitrogen fixation clade. Since in *Parasponia* NFP1 (as part of the NFP-I clade) does not support symbiosis on its own, the divergence of the ancestral function over the two copies is likely (**Chapter 4.**). Strikingly an LYR-Ia LysM-type protein of *Petunia hybrida* can functionally complement a legume for nodule organogenesis, indicating that a priori functional divergence of LYR-Ia is not required for co-option (Girardin et al. 2019). While the role of the LYR-Ia clade appears to lie in arbuscular mycorrhizal signalling, the phenotypes of LYR-Ia type mutants are often relatively mild. In *Parasponia* even a *Pannfp1;Pannfp2* double mutant did not show a phenotype in arbuscular mycorrhizal infection level. Also in *Medicago,* the *Mtnfp* mutant has no phenotype. This could be explained that during arbuscular mycorrhizal infection a combination of chitin oligosaccharide (CO) and

LCO signalling is used for infection, by which knocking out the myc-LCO signalling pathway only has minor phenotypic effects (Feng et al. 2019). Perhaps due to the duality of LCO-signalling in two different symbiotic interactions the importance of myc-LCOs in arbuscular mycorrhizal symbiosis may have become less in the nodulating lineage. Gene silencing and mutant analysis indicates that LYR-Ia type genes such as *Petunia hybrida PtlYK10* and tomato *SlLYK10* are important for mycorrhizal infection, with both studies showing reduced colonization levels (Buendia et al. 2016; Girardin et al. 2019)

The second receptor that acts in LCO perception is part of the LYK-I clade. These LYK-I clade LysM-type receptors are of outstanding importance for LCO perception in legumes. Legumes have developed a highly specific LysM-type receptor for rhizobium Nod factor signalling during Nodulation; named MtLYK3/LjNFR1 in *Medicago* and *Lotus*. This receptor evolved on a series of gene duplications in the Fabales lineage (De Mita et al. 2014). These legume duplications in the LYK-I-clade are not present in any other nodulating lineages. While a duplication happened early in the eudicots, giving rise to the LYK-Ia and LYK-Ib subclades, no duplications correlate with the origin of the nitrogen fixation clade **(Figure 1, Chapter 4).** The MtLYK3/LjNFR1 LysM-type receptor is specific for nodulating genera in the Fabales and is known to be lost in *Cercis (De Mita et al. 2014)*. These duplications also gave rise to several other LysM-type receptors such as MtLYK9/LjCERK6, which are important for immunity and arbuscular mycorrhizal signalling (Bozsoki et al. 2017; Leppyanen et al. 2017; Feng et al. 2019; Gibelin-Viala et al. 2019). While the functions of most other copies are not yet known, LjLYS1 now renamed LjNFRe plays additive roles in epidermal nod factor signalling in Lotus (Murakami et al. 2018). The overlap of LCO and CO signalling in arbuscular mycorrhizal symbiosis makes it difficult to separate their respective roles during mycorrhizal infection. In Medicago, a *Mtlyk9/Mtnfp* double mutant significantly reduces colonization, since in this mutant both CO by LYK9 and LCO signalling by NFP are affected (Feng et al. 2019). In the *Parasponia/Trema* lineage, no duplications have occurred in the LYK-I clade. However, the *LYK3* gene in the LYK-Ib clade underwent triplication of the first exon, giving rise to different protein variants. Though, single exon knockouts do not have deleterious effects on chitin signalling or rhizobium infection **(Chapter 4.)**. Interestingly *Parasponia LYK3* encodes a trifunctional receptor. *PanLYK3* is essential for chitin signalling in defence responses. Although the *Panlyk3* mutants are still able to nodulate, they have severe problems with infection in most nodules. Interestingly *PanLYK3* is partially functionally redundant with *PanLYK1* from the LYK-Ia clade. Only a double LYK-I clade mutant could fully block nodule formation and arbuscular mycorrhizal interactions **(Figure 6 and 9, Chapter 4.)**. In general, chitin signalling appears to be the ancestral function of the LYK-I-clade. Outside of

the nitrogen fixation clade, chitin signalling appears to be the major role of LYK-Ib clade members, such as arabidopsis *AtCERK1* or tomato *SlLYK1(Miya et al. 2007; Liao et al. 2018)*. While little experimental evidence exists for LYK-Ia clade members, tomato *SlLYK12* is involved in arbuscular mycorrhizal infections (Liao et al. 2018). This is in line with the role of *PanLYK1* and *PanLYK3* in arbuscular mycorrhization. Rice OsCERK, which falls well outside of the eudicot duplication event, also has a dual function in arbuscular mycorrhizal symbiosis and immunity signalling (Miyata et al. 2014; Carotenuto et al. 2017; He et al. 2019). Since most non-nodulating species retained both LYK-Ia and LYK-Ib copies it appears that these receptors may not have specialized for a specific role in nodule initiation and symbiont recognition outside of the Fabales lineage.

The co-option and regulation of the transcription factor NODULE INCEPTION in nodulation

Outside of the legume family homologs of all symbiosis genes are ubiquitously present. Therefore nodulation most likely evolved by recruitment of existing pathways rather than the evolution of completely new genes and interaction. Due to the large overlap of symbiosis regulated genes in *Medicago* and *Parasponia*, some of these genes may have been recruited only once in evolutionary history. Others may be recruited in parallel at later stages, because of restrictions in the physiological requirements of nodules (e.g. hemoglobins). A gene that is essential for nodulation in all plants tested is the transcription factor *NIN*. *NIN* is nearly ubiquitously lost in non-nodulating genera in the nitrogen fixation clade. Although functional gene models of *NIN* may be found in *Ziziphus jujuba*, and truncated *NIN* versions are expressed in *Trema, Cannabis* and *Humulus* **(Chapter 3.)**. Previous work has suggested that NIN has functionally separable roles in the epidermis and nodule organogenesis (Yoro et al. 2014). While epidermal infection relies on the presence of elements located relatively close to the *NIN* transcription start site, such as a binding element for the symbiotic transcription factor CYCLOPS (Singh et al. 2014), complementation of a legume *nin* mutant in for nodule organogenesis was as of yet difficult. In rare cases, successful nodules were reported with the expression of short nin promoters or under ectopic expression of *NIN (Clavijo et al. 2015)*.

The first 5 kbp of upstream sequence of the *NIN* transcriptional start site contains all cis-regulatory elements necessary for LCO-induced epidermal expression. A promoter of this length is not enough for effective organogenesis **(Chapter 6.)**. This is further indicated by the similarities in the mutant phenotypes between the *Lotus daphne* mutant and the *Medicago FN8113* mutant that both have a chromosomal translocation between *NIN* and the first upstream located gene (Yoro et al. 2014) (**Chapter 6.)**. By genome comparisons of a diverse set of legumes, we detected an upstream region in the *NIN* promoter of legumes that is required for the induction

of the gene in the pericycle during nodule organogenesis. This enhancer element contains cytokinin response elements and is conserved in legumes. The element was named the CE region, for Cytokinin response Elements containing region. Cytokinin biosynthesis genes and the accumulation of bioactive cytokinin are induced upon Nod factor signalling (Van Zeijl et al. 2015). It has long been known that nodule organogenesis can be initiated by exogenous cytokinin application in legumes. Both the cytokinin receptor MtCRE1 and B-type response regulator MtRR1 are expressed in the pericycle of uninoculated roots. Therefore it is probable that this enhancer is at least in part regulated by cytokinin. Besides Cytokinin and CYCLOPS, an important regulator of *NIN* transcription appears to be the AP2 transcription factor ERN1 (M. Liu et al. 2019). Since the CE element contains a GCC-box motif, a potential AP2 recognition site, ERN1 may target the *NIN* promoter in the CE enhancer element. This would potentially explain why *NIN* induction by cytokinin is much lower in the *ern1* mutant (M. Liu et al. 2019). In legumes cytokinin may feature primarily as a feed-forward enhancer, to activate primordium formation more quickly below the developing nodule primordium, possibly in concert with other transcription factors such as ERN1. Cytokinin application does however not induce pseudo-nodules in non-legumes, suggesting that the cytokinin-NIN feed-forward loop is legume specific (Gauthier-Coles, White, and Mathesius 2018). *NIN* activation in the pericycle also reduces colonization of rhizobia in the epidermis, which ensures the entrapment of mostly a single colony in a nodule primordium. The inhibition of epidermal infection could not always be completely restored by using the CE element plus the 5 kbp upstream promoter region complementation **(Chapter 6.)**. The mechanism of how this inhibition of epidermis infection is achieved is largely unknown. Therefore it is probable that other important regulating sites, primarily negative regulators, bind to other regions of the *NIN* promoter.

NIN as a master regulator of the cell cycle and lateral root programme

Nodules in some way resemble lateral roots, however, represent a derived and distinct structure. Increasing amounts of evidence continue to show the overlap over lateral roots and nodules, such as the involvement of PLETHORA transcription factors in nodule meristem maintenance and the expression of the transcription factor *MtWOX5* that acts as root quiescent centre marker in the nodule vascular meristem (Franssen et al. 2015; Blilou et al. 2005). In order to generate cell divisions in the pericycle and inner cortex, an auxin maximum needs to be established. This process is somewhat similar to the formation of lateral roots, only the auxin maxima of nodules are wider and less precise (Eva E. Deinum 2015). Also in actinorhizal plants, a similar broad pattern of an auxin maximum may underlie nodule formation (Imanishi et al. 2014). Modelling approaches have generated hypothesis on how these maxima are generated, with roles for the differential organization of PIN auxin

efflux carriers, AUX1-LAX auxin influx carriers and localized auxin production (Roy et al. 2017; Takanashi, Sugiyama, and Yazaki 2011). Slightly different locations of the auxin maxima may also drive the difference between determinate and indeterminate legume nodules (Eva Elisabeth Deinum et al. 2012). The generation of this auxin maximum may depend on *NIN* expression in the inner cell layers induced via a mobile signal from the epidermis (Eva E. Deinum, Kohlen, and Geurts 2016). Recently increasing evidence has strengthened this view. Recent work has coupled *NIN* expression in the inner cell layers to the induction of *ASYMMETRIC LEAVES 2-LIKE18/LATERAL ORGAN BOUNDARIES DOMAIN 16*a (*ASL18/LBD16a*) (Schiessl et al. 2019; Soyano et al. 2019). Orthologs of these genes are required for lateral root development in non-legumes (J. Liu et al. 2014). The expression of *LBD16a* induces auxin biosynthesis via transcriptional induction of *STYLISH* (*STY*) and *YUCCAs* (*YUC*), promoting the formation of auxin maximum (Schiessl et al. 2019). The use of a key transcription factor in lateral root development for nodule development, results in a large overlap in transcriptional regulation in both developmental processes. This raises questions as to how specificity is retained. The discovery of a *Medicago* mutant with a nodule to root conversion phenotype (*MtNOOT1*), reveals that additional transcriptional regulators may have been recruited to regulate this aspect (J.-M. Couzigou et al. 2012; J. M. Couzigou et al. 2016).

Why NIN from the *NIN-LIKE PROTEIN (NLP)* family has been specifically recruited for nodulation, remains a central question. While LCO-signalling transduction was clearly co-opted from the common symbiotic signalling pathway, NIN itself is not required for arbuscular mycorrhization (Kumar et al. 2020). Therefore NIN represents a unique component in the regulation of nodule symbiosis signalling. It was shown that members of the *NLP* gene family are required for nitrate regulation of root development and growth responses. Arabidopsis NLPs are central in nitrate signalling responses, and it was shown that all NLP proteins can bind to the Nitrate Responsive Element, which is highly similar to the NIN-binding element (Soyano et al. 2013; Konishi and Yanagisawa 2014). In Arabidopsis AtNLP6 and AtNLP7 interact with TEOSINTE BRANCHED CELL FACTOR 20 (TCP20) through the Phox and Bem1-domain (PB1) (Guan et al. 2017). During nitrate starvation, TCP20-NLP6/7 heterodimers accumulate in the nucleus to control the expression of nitrate assimilation and signalling genes. Curiously this correlates with the down-regulation of the G2/M cell-cycle marker gene, CYCB1 (Guan et al. 2017). In order to target NLPs to the nucleus phosphorylation of NLPs by subgroup III Ca^{2+} SENSOR PROTEIN KINASEs CPKs) is required (Marchive et al. 2013; K.-H. Liu et al. 2017). Targets of NLP transcription factors include *AtLBD37, AtLBD38* and *ATLBD39 (Yanagisawa 2014; Rubin et al. 2009; Alvarez et al. 2020)*. This close link

of NLPs with LBD transcription factors may explain how NIN proteins target *LBD16* to regulate the lateral root developmental program (Schiessl et al. 2019; Soyano et al. 2019).

While NLPs are mostly post-transcriptionally regulated by modifying their subcellular localization, in nodulating plants *NIN* is under stringent transcriptional control. This major difference in regulation is not readily explained. Legume NIN proteins contain degraded domains in the N-terminal region, where in case of NLPs phosphorylation may take place, this makes it likely that phosphorylation is no longer important (Leif Schauser, Wieloch, and Stougaard 2005; Suzuki, Konishi, and Yanagisawa 2013). In non-legumes, like *Casuarina* NIN and *Parasponia* NIN, these domains show conservation with the NLP domains (Clavijo et al. 2015; Bu et al. 2019), indicating that there may be NIN-interacting proteins, like the CPKs, targeting these domains. Recently the nitrate unresponsive mutant *nrsym1* was found to encode LjNLP4, to regulate nodule number according to exogenous nitrate concentrations (Nishida et al. 2018). A system which was found to be functioning independently of the autoregulation of nodulation pathway. It highlights that multiple NLP transcription factors were recruited into a symbiotic role. The recruitment of NIN in the nodulation gene expression programme seems to have resulted in a loss of the original NIN function, indicated by the widespread loss in non-nodulating species in the nitrogen fixation clade. Research toward the role of NIN transcription factors in non-nodulating plants closely related to the nitrogen fixation clade could help our understanding of this important transcription factor.

Future for the engineering of Nodulation.

In this thesis we have found evidence that the nitrogen-fixing endosymbiosis with *Frankia* or rhizobia may have evolved only once. While the evolution of nodulation did not require the emergence of new genes it represented a rewiring of transcriptional modules previously unconnected. While the CE-element is a small part of the *NIN*-regulation puzzle, there will be more work to find transcriptional enhancers in other symbiosis genes. Meanwhile, another strategy might represent the replacement of lost genes in non-nodulating species which experienced a recent loss of the trait. In *Trema* species only a limited number of symbiosis genes have eroded from their genomes. This means that while *Trema* does not nodulate, restoration of nodulation may be feasible. When we take a closer look at the genes lost in *Trema NFP2, NIN* and *RPG* are obviously essential engineering candidates for nodule organogenesis and infection. Similar to *PanNFP2, PanNIN* is essential for nodulation in *Parasponia*. *Pannin* knockouts do not form nodules (Bu et al. 2019). However besides the seven genes identified, which were consistently lost in *Trema* species, several more genes showed inconsistent losses in single *Trema* lineages **(Chapter 3., 4. and**

5.). Therefore in any given non-nodulating species in the number of gene losses might be far higher than the genes we know today. In order to build a functional gene expression system capable of supporting symbiosis, our understanding of plant transformation and the integration of large DNA constructs needs to increase. Meanwhile, another problem in repairing the lost symbiosis genes is not the loss of the genes themselves, but the loss of their regulatory sequences. Identifying gene regulatory networks and conserved transcription factor binding sites may prove a major Bio-informatic challenge. *Trema* may serve as the first proof of principle that restoration of the symbiosis trait is feasible in the near future.

References

Alvarez, José M., Anna-Lena Schinke, Matthew D. Brooks, Angelo Pasquino, Lauriebeth Leonelli, Kranthi Varala, Alaeddine Safi, Gabriel Krouk, Anne Krapp, and Gloria M. Coruzzi. 2020. "Transient Genome-Wide Interactions of the Master Transcription Factor NLP7 Initiate a Rapid Nitrogen-Response Cascade." *Nature Communications* 11 (1): 1157.

Arnone, John A., Steven J. Kohls, and Dwight D. Baker. 1994. "NITRATE EFFECTS ON NODULATION AND NITROGENASE ACTIVITY OF ACTINORHIZAL CASUARINA STUDIED IN SPLIT-ROOT SYSTEMS." *Soil* 26 (5): 599–606.

Arrighi, Jean-François, Olivier Godfroy, Françoise de Billy, Olivier Saurat, Alain Jauneau, and Clare Gough. 2008. "The *RPG* Gene of *Medicago Truncatula* Controls Rhizobium-Directed Polar Growth during Infection." *Proceedings of the National Academy of Sciences of the United States of America* 105 (28): 9817–22.

Bailly, Xavier, Isabelle Olivieri, Brigitte Brunel, Jean-Claude Cleyet-Marel, and Gilles Béna. 2007. "Horizontal Gene Transfer and Homologous Recombination Drive the Evolution of the Nitrogen-Fixing Symbionts of Medicago Species." *Journal of Bacteriology* 189 (14): 5223–36.

Bender, Gregory L., Walter Goydych, Barry G. Rolfe, and Murali Nayudu. 1987. "The Role of Rhizobium Conserved and Host Specific Nodulation Genes in the Infection of the Non-Legume Parasponia Andersonii." *Molecular & General Genetics: MGG* 210 (2): 299–306.

Berry, A. M., O. T. Harriott, R. A. Moreau, S. F. Osman, D. R. Benson, and A. D. Jones. 1993. "Hopanoid Lipids Compose the *Frankia* Vesicle Envelope, Presumptive Barrier of Oxygen Diffusion to Nitrogenase." *Proceedings of the National Academy of Sciences of the United States of America* 90 (13): 6091–94.

Blilou, I., J. Xu, M. Wildwater, V. Willemsen, I. Paponov, J. Friml, R. Heidstra, M. Aida, K. Palme, and B. Scheres. 2005. "The PIN Auxin Efflux Facilitator Network Controls Growth and Patterning in Arabidopsis Roots." *Nature* 433 (7021): 39–44.

Bontemps, Cyril, Geoffrey N. Elliott, Marcelo F. Simon, Fábio B. Dos Reis Júnior, Eduardo Gross, Rebecca C. Lawton, Nicolau Elias Neto, et al. 2010. "Burkholderia Species Are Ancient Symbionts of Legumes." *Molecular Ecology* 19 (1): 44–52.

Bozsoki, Zoltan, Jeryl Cheng, Feng Feng, Kira Gysel, Maria Vinther, Kasper R. Andersen, Giles Oldroyd, Mickael Blaise, Simona Radutoiu, and Jens Stougaard. 2017. "Receptor-Mediated Chitin Perception in Legume Roots Is Functionally Separable from Nod Factor Perception." *Proceedings of the National Academy of Sciences* 114 (38): E8118–27.

Buendia, Luis, Tongming Wang, Ariane Girardin, and Benoit Lefebvre. 2016. "The LysM Receptor-like Kinase SlLYK10 Regulates the Arbuscular Mycorrhizal Symbiosis in Tomato." *The New Phytologist* 210 (1): 184–95.

Bu, Fengjiao, Luuk Rutten, Yuda Purwana Roswanjaya, Olga Kulikova, Marta Rodriguez-Franco, Thomas Ott, Ton Bisseling, Arjan van Zeijl, and Rene Geurts. 2019. "Mutant Analysis in the Nonlegume Parasponia Andersonii Identifies NIN and NF-YA1 Transcription Factors as a Core Genetic Network in Nitrogen-Fixing Nodule Symbioses." *The New Phytologist* n/a (n/a). https://doi.org/10.1111/nph.16386.

Carotenuto, Gennaro, Mireille Chabaud, Kana Miyata, Martina Capozzi, Naoya Takeda, Hanae Kaku, Naoto Shibuya, Tomomi Nakagawa, David G. Barker, and Andrea Genre. 2017. "The Rice LysM Receptor-like Kinase Os CERK 1 Is Required for the Perception of Short-Chain Chitin Oligomers in Arbuscular Mycorrhizal Signaling." *The New Phytologist* 214 (4): 1440–46.

Christin, Pascal-Antoine, Guillaume Besnard, Emanuela Samaritani, Melvin R. Duvall, Trevor R. Hodkinson, Vincent Savolainen, and Nicolas Salamin. 2008. "Oligocene CO2 Decline Promoted C4 Photosynthesis in Grasses." *Current Biology: CB* 18 (1): 37–43.

Christin, Pascal Antoine, Colin P. Osborne, David S. Chatelet, J. Travis Columbus, Guillaume Besnard, Trevor R. Hodkinson, Laura M. Garrison, Maria S. Vorontsova, and Erika J. Edwards. 2013. "Anatomical Enablers and the Evolution of C4 Photosynthesis in Grasses." *Proceedings of the National Academy of Sciences of the United States of America* 110 (4): 1381–86.

Clavijo, Fernando, Issa Diedhiou, Virginie Vaissayre, Laurent Brottier, Jennifer Acolatse, Daniel Moukouanga, Amandine Crabos, et al. 2015. "The *Casuarina NIN* Gene Is Transcriptionally Activated throughout *Frankia* Root Infection as Well as in Response to Bacterial Diffusible Signals." *The New Phytologist* 208: 887–903.

Couzigou, Jean Malo, Kevin Magne, Samuel Mondy, Viviane Cosson, Jonathan Clements, and Pascal Ratet. 2016. "The Legume NOOT-BOP-COCH-LIKE Genes Are Conserved Regulators of Abscission, a Major Agronomical Trait in Cultivated Crops." *The New Phytologist* 209 (1): 228–40.

Couzigou, Jean-Malo, Vladimir Zhukov, Samuel Mondy, Ghada Abu el Heba, Viviane Cosson, T. H. Noel Ellis, Mike Ambrose, et al. 2012. "NODULE ROOT and COCHLEATA Maintain Nodule Development and Are Legume Orthologs of Arabidopsis BLADE-ON-PETIOLE Genes." *The Plant Cell* 24 (11): 4498–4510.

Deinum, Eva E. 2015. “Computational and Experimental Evidence That Auxin Accumulation in Nodule and Lateral Root Primordia Occurs by Different Mechanisms ?.” In *Biological Nitrogen Fixation. Volume 2*, edited by Frans de Bruijn. John Wiley & Sons.

Deinum, Eva E., Wouter Kohlen, and René Geurts. 2016. “Quantitative Modelling of Legume Root Nodule Primordium Induction by a Diffusive Signal of Epidermal Origin That Inhibits Auxin Efflux.” *BMC Plant Biology* 16 (1): 254.

Deinum, Eva Elisabeth, René Geurts, Ton Bisseling, and Bela M. Mulder. 2012. “Modeling a Cortical Auxin Maximum for Nodulation: Different Signatures of Potential Strategies.” *Frontiers in Plant Science* 3: 1–19.

Delaux, Pierre Marc, Kranthi Varala, Patrick P. Edger, Gloria M. Coruzzi, J. Chris Pires, and Jean Michel Ané. 2014. “Comparative Phylogenomics Uncovers the Impact of Symbiotic Associations on Host Genome Evolution.” *PLoS Genetics* 10 (7). https://doi.org/10.1371/journal.pgen.1004487.

De Mita, Stéphane, Arend Streng, Ton Bisseling, and René Geurts. 2014. “Evolution of a Symbiotic Receptor through Gene Duplications in the Legume-Rhizobium Mutualism.” *The New Phytologist* 201 (3): 961–72.

Denison, R. Ford, and E. Toby Kiers. 2004. “Lifestyle Alternatives for Rhizobia: Mutualism, Parasitism, and Forgoing Symbiosis.” *FEMS Microbiology Letters* 237 (2): 187–93.

Doyle, J. J. 2016. “Chasing Unicorns: Nodulation Origins and the Paradox of Novelty.” *American Journal of Botany* 103 (11): 1–4.

Dupin, Simon E., René Geurts, E. Toby Kiers. 2020. “The Non-Legume Parasponia Andersonii Mediates the Fitness of Nitrogen-Fixing Rhizobial Symbionts Under High Nitrogen Conditions.” *Frontiers in Plant Science* 10: 1779.

Feng, Feng, Jongho Sun, Guru V. Radhakrishnan, Tak Lee, Zoltán Bozsóki, Sébastien Fort, Aleksander Gavrin, et al. 2019. “A Combination of Chitooligosaccharide and Lipochitooligosaccharide Recognition Promotes Arbuscular Mycorrhizal Associations in Medicago Truncatula.” *Nature Communications* 10 (1): 5047.

Franssen, Henk J., Ting Ting Xiao, Olga Kulikova, Xi Wan, Ton Bisseling, Ben Scheres, and Renze Heidstra. 2015. “Root Developmental Programs Shape the Medicago Truncatula Nodule Meristem.” *Development* 142 (17): 2941–50.

Friesen, Maren L. 2012. “Widespread Fitness Alignment in the Legume – Rhizobium Symbiosis.” *The New Phytologist*. https://doi.org/10.1111/j.1469-8137.2012.04099.x.

Gauthier-Coles, Christopher, Rosemary G. White, and Ulrike Mathesius. 2018. “Nodulating Legumes Are Distinguished by a Sensitivity to Cytokinin in the Root Cortex Leading to Pseudonodule Development.” *Frontiers in Plant Science* 9: 1901.

Gerhart, Laci M., and Joy K. Ward. 2010. “Plant Responses to Low [CO2] of the Past.” *The New Phytologist* 188 (3): 674–95.

Gibelin-Viala, Chrystel, Emilie Amblard, Virginie Puech-Pages, Maxime Bonhomme, Magali Garcia, Adeline Bascaules-Bedin, Judith Fliegmann, et al. 2019. “The Medicago Truncatula LysM Receptor-like Kinase LYK9 Plays a Dual Role in Immunity and the Arbuscular Mycorrhizal Symbiosis.” *The New Phytologist* 223 (3): 1516–29.

Girardin, Ariane, Tongming Wang, Yi Ding, Jean Keller, Luis Buendia, Mégane Gaston, Camille Ribeyre, et al. 2019. “LCO Receptors Involved in Arbuscular Mycorrhiza Are Functional for Rhizobia Perception in Legumes.” *Current Biology: CB* 29 (24): 4249–59.e5.

Gomez, S. Karen, Hélène Javot, Prasit Deewatthanawong, Ivone Torres-Jerez, Yuhong Tang, Elison B. Blancaflor, Michael K. Udvardi, and Maria J. Harrison. 2009. “Medicago Truncatula and Glomus Intraradices Gene Expression in Cortical Cells Harboring Arbuscules in the Arbuscular Mycorrhizal Symbiosis.” *BMC Plant Biology* 9 (1): 10.

Gough, Clare, Ludovic Cottret, Benoit Lefebvre, and Jean-Jacques Bono. 2018. “Evolutionary History of Plant LysM Receptor Proteins Related to Root Endosymbiosis.” *Frontiers in Plant Science* 9 (July): 1–9.

Griesmann, Maximilian, Yue Chang, Xin Liu, Yue Song, Georg Haberer, Matthew B. Crook, Benjamin Billault-Penneteau, et al. 2018. “Phylogenomics Reveals Multiple Losses of Nitrogen-Fixing Root Nodule Symbiosis.” *Science* 1743 (May): eaat1743.

Guan, Peizhu, Juan-José Ripoll, Renhou Wang, Lam Vuong, Lindsay J. Bailey-Steinitz, Dening Ye, and Nigel M. Crawford. 2017. “Interacting TCP and NLP Transcription Factors Control Plant Responses to Nitrate Availability.” *Proceedings of the National Academy of Sciences* 114 (9): 2419–24.

He, Jiangman, Chi Zhang, Huiling Dai, Huan Liu, Xiaowei Zhang, Jun Yang, Xi Chen, et al. 2019. “A LysM Receptor Heteromer Mediates Perception of Arbuscular Mycorrhizal Symbiotic Signal in Rice.” *Molecular Plant* 12 (12): 1561–76.

Herder, Griet Den, and Martin Parniske. 2009. “The Unbearable Naivety of Legumes in Symbiosis,” no. Box 1: 491–99.

Heyduk, Karolina, Jose J. Moreno-Villena, Ian S. Gilman, Pascal Antoine Christin, and Erika J. Edwards. 2019. “The Genetics of Convergent Evolution: Insights from Plant Photosynthesis.” *Nature Reviews. Genetics* 20 (8): 485–93.

Hoeksema, Jason D., James D. Bever, Sounak Chakraborty, V. Bala Chaudhary, Monique Gardes, Catherine A. Gehring, Miranda M. Hart, et al. 2018. “Evolutionary History of Plant Hosts and Fungal Symbionts Predicts the Strength of Mycorrhizal Mutualism.” *Communications Biology* 1 (1): 116.

Imanishi, Leandro, Francine M. Perrine-Walker, Adama Ndour, Alice Vayssières, Genevieve Conejero, Mikaël Lucas, Antony Champion, Laurent Laplaze, Luis Wall, and Sergio Svistoonoff. 2014. “Role of Auxin during Intercellular Infection of *Discaria Trinervis* by *Frankia*.” *Frontiers in Plant Science* 5: 399.

Infante, Carlos R., Alexandra G. Mihala, Sungdae Park, Jialiang S. Wang, Kenji K. Johnson, James D. Lauderdale, and Douglas B. Menke. 2015. “Shared Enhancer Activity in the Limbs and Phallus and Functional Divergence of a Limb-Genital Cis-Regulatory Element in Snakes.” *Developmental Cell* 35 (1): 107–19.

Jacobsen-Lyon, K., E. O. Jensen, J. E. Jørgensen, K. a. Marcker, W. J. Peacock, and E. S. Dennis. 1995. “Symbiotic and Nonsymbiotic Hemoglobin Genes of Casuarina Glauca.” *The Plant Cell* 7 (2): 213–23.

Janssens, Steven B., Thomas L. P. Couvreur, Arne Mertens, Gilles Dauby, Leo-Paul M. J. Dagallier, Samuel Vanden Abeele, Filip Vandelook, et al. 2020. “A Large-Scale Species Level Dated Angiosperm Phylogeny for Evolutionary and Ecological Analyses.” *Biodiversity Data Journal* 8 (January): e39677.

Jones, Emily I., Michelle E. Afkhami, Erol Akçay, Judith L. Bronstein, Redouan Bshary, Megan E. Frederickson, Katy D. Heath, et al. 2015. “Cheaters Must Prosper: Reconciling Theoretical and Empirical Perspectives on Cheating in Mutualism.” *Ecology Letters* 18 (11): 1270–84.

Kakar, Smita, Ryan Sturms, Andrea Tiffany, Jay C. Nix, Alan A. DiSpirito, and Mark S. Hargrove. 2011. “Crystal Structures of *Parasponia* and *Trema* Hemoglobins: Differential Heme Coordination Is Linked to Quaternary Structure.” *Biochemistry* 50 (20): 4273–80.

Kawaharada, Yasuyuki, Mette W. Nielsen, Simon Kelly, Euan K. James, Kasper R. Andersen, Sheena R. Rasmussen, Winnie Füchtbauer, et al. 2017. “Differential Regulation of the Epr3 Receptor Coordinates Membrane-Restricted Rhizobial Colonization of Root Nodule Primordia.” *Nature Communications* 8 (February): 14534.

Kiss, Enikő, Botond Hegedüs, Máté Virágh, Torda Varga, Zsolt Merényi, Tamás Kószó, Balázs Bálint, et al. 2019. “Comparative Genomics Reveals the Origin of Fungal Hyphae and Multicellularity.” *Nature Communications* 10 (1): 4080.

Konishi, Mineko, and Shuichi Yanagisawa. 2014. “Emergence of a New Step towards Understanding the Molecular Mechanisms Underlying Nitrate-Regulated Gene Expression.” *Journal of Experimental Botany* 65 (19): 1–12.

Krusell, Lene, Naoto Sato, Izumi Fukuhara, Bjã E. V. Koch, Christina Grossmann, Satoru Okamoto, Erika Oka-Kira, et al. 2011. “The Clavata2 Genes of Pea and Lotus Japonicus Affect Autoregulation of Nodulation.” *The Plant Journal: For Cell and Molecular Biology* 65 (6): 861–71.

Kumar, Anil, Donna R. Cousins, Cheng-Wu Liu, Ping Xu, and Jeremy D. Murray. 2020. “Nodule Inception Is Not Required for Arbuscular Mycorrhizal Colonization of Medicago Truncatula.” *Plants* 9 (1). https://doi.org/10.3390/plants9010071.

Leppyanen, Irina V., Vlada Y. Shakhnazarova, Oksana Y. Shtark, Nadezhda A. Vishnevskaya, Igor A. Tikhonovich, and Elena A. Dolgikh. 2017. “Receptor-Like Kinase LYK9 in Pisum Sativum L. Is the CERK1-Like Receptor That Controls Both Plant Immunity and AM Symbiosis Development.” *International Journal of Molecular Sciences* 19 (1). https://doi.org/10.3390/ijms19010008.

Liao, Dehua, Xun Sun, Ning Wang, Fengming Song, and Yan Liang. 2018. “Tomato LysM Receptor-Like Kinase SlLYK12 Is Involved in Arbuscular Mycorrhizal Symbiosis.” *Frontiers in Plant Science* 9 (July): 1004.

Li, Hong-Lei, Wei Wang, Peter E. Mortimer, Rui-Qi Li, De-Zhu Li, Kevin D. Hyde, Jian-Chu Xu, Douglas E. Soltis, and Zhi-Duan Chen. 2015. “Large-Scale Phylogenetic Analyses Reveal Multiple Gains of Actinorhizal Nitrogen-Fixing Symbioses in Angiosperms Associated with Climate Change.” *Scientific Reports* 5 (September): 14023.

Liu, J., L. Sheng, Y. Xu, J. Li, Z. Yang, H. Huang, and L. Xu. 2014. “WOX11 and 12 Are Involved in the First-Step Cell Fate Transition during de Novo Root Organogenesis in Arabidopsis.” *The Plant Cell* 26 (3): 1081–93.

Liu, Kun-Hsiang, Yajie Niu, Mineko Konishi, Yue Wu, Hao Du, Hoo Sun Chung, Lei Li, et al. 2017. “Discovery of nitrate–CPK–NLP Signalling in Central Nutrient–growth Networks.” *Nature* 545 (7654): 311–16.

Liu, Meng, Takashi Soyano, Koji Yano, Makoto Hayashi, and Masayoshi Kawaguchi. 2019. “ERN1 and CYCLOPS Coordinately Activate NIN Signaling to Promote Infection Thread Formation in Lotus Japonicus.” *Journal of Plant Research* 132 (5): 641–53.

Marchetti, Marta, Alain Jauneau, Delphine Capela, Philippe Remigi, Carine Gris, Jacques Batut, and Catherine Masson-Boivin. 2014. “Shaping Bacterial Symbiosis With Legumes by Experimental Evolution.” *Molecular Plant-Microbe Interactions: MPMI* 27 (9): 956–64.

Marchive, Chloé, François Roudier, Loren Castaings, Virginie Bréhaut, Eddy Blondet, Vincent Colot, Christian Meyer, and Anne Krapp. 2013. “Nuclear Retention of the Transcription Factor NLP7 Orchestrates the Early Response to

Nitrate in Plants." *Nature Communications* 4: 1713.

Matveeva, Tatiana V., and Léon Otten. 2019. "Widespread Occurrence of Natural Genetic Transformation of Plants by Agrobacterium." *Plant Molecular Biology* 101 (4-5): 415–37.

McGowen, Michael R., John Gatesy, and Derek E. Wildman. 2014. "Molecular Evolution Tracks Macroevolutionary Transitions in Cetacea." *Trends in Ecology & Evolution* 29 (6): 336–46.

Miya, Ayako, Premkumar Albert, Tomonori Shinya, Yoshitake Desaki, Kazuya Ichimura, Ken Shirasu, Yoshihiro Narusaka, Naoto Kawakami, Hanae Kaku, and Naoto Shibuya. 2007. "CERK1, a LysM Receptor Kinase, Is Essential for Chitin Elicitor Signaling in Arabidopsis."

Miyata, Kana, Toshinori Kozaki, Yusuke Kouzai, Kenjirou Ozawa, Kazuo Ishii, Erika Asamizu, Yoshihiro Okabe, et al. 2014. "Bifunctional Plant Receptor, OsCERK1, Regulates Both Chitin-Triggered Immunity and Arbuscular Mycorrhizal Symbiosis in Rice." *Plant & Cell Physiology* 55: in press.

Muller, Tobias, Meghan Avolio, Martin Olivi, Mariam Benjdia, Enno Rikirsch, Alexis Kasaras, Michael Fitz, Michel Chalot, and Daniel Wipf. 2007. "Nitrogen Transport in the Ectomycorrhiza Association: The Hebeloma Cylindrosporum-Pinus Pinaster Model." *Phytochemistry* 68 (1): 41–51.

Murakami, Eiichi, Jeryl Cheng, Kira Gysel, Zoltan Bozsoki, Yasuyuki Kawaharada, Christian Toftegaard Hjuler, Kasper Kildegaard Sørensen, et al. 2018. "Epidermal LysM Receptor Ensures Robust Symbiotic Signalling in Lotus Japonicus." *eLife* 7 (June). https://doi.org/10.7554/eLife.33506.

Nagy, László G., Gábor M. Kovács, Krisztina Krizsán. 2018. "Complex Multicellularity in Fungi: Evolutionary Convergence, Single Origin, or Both?" *Biological Reviews of the Cambridge Philosophical Society* 93(4):1778–94.

Nishida, Hanna, Sachiko Tanaka, Yoshihiro Handa, Momoyo Ito, Yuki Sakamoto, Sachihiro Matsunaga, Shigeyuki Betsuyaku, et al. 2018. "A NIN-LIKE PROTEIN Mediates Nitrate-Induced Control of Root Nodule Symbiosis in Lotus Japonicus." *Nature Communications* 9 (1). https://doi.org/10.1038/s41467-018-02831-x.

Oono, Ryoko, Carolyn G. Anderson, and R. Ford Denison. 2011. "Failure to Fix Nitrogen by Non-Reproductive Symbiotic Rhizobia Triggers Host Sanctions That Reduce Fitness of Their Reproductive Clonemates." *Proceedings. Biological Sciences / The Royal Society* 278 (1718): 2698–2703.

Op den Camp, Rik H. M., Elisa Polone, Elena Fedorova, Wim Roelofsen, Andrea Squartini, Huub J. M. Op den Camp, Ton Bisseling, and René Geurts. 2012. "Deploys a Broad Rhizobium Host Range Strategy Resulting in Largely Variable Symbiotic Effectiveness." *Molecular Plant-Microbe Interactions: MPMI* 25 (7): 954–63.

Osipova, M. a., V. Mortier, K. N. Demchenko, V. E. Tsyganov, I. a. Tikhonovich, L. a. Lutova, E. a. Dolgikh, and Sofie Goormachtig. 2012. "WUSCHEL-RELATED HOMEOBOX5 Gene Expression and Interaction of CLE Peptides with Components of the Systemic Control Add Two Pieces to the Puzzle of Autoregulation of Nodulation." *Plant Physiology* 158 (March): 1329–41.

Persson, Tomas, Kai Battenberg, Irina V. Demina, Theoden Vigil-Stenman, Brian Vanden Heuvel, Petar Pujic, Marc T. Facciotti, et al. 2015. "*Candidatus* Frankia Datiscae Dg1, the Actinobacterial Microsymbiont of *Datisca Glomerata*, Expresses the Canonical *nod* Genes *nodABC* in Symbiosis with Its Host Plant." *PloS One* 10 (5): e0127630.

Rasmussen, S. R., W. Füchtbauer, M. Novero, V. Volpe, N. Malkov, A. Genre, P. Bonfante, J. Stougaard, and S. Radutoiu. 2016. "Intraradical Colonization by Arbuscular Mycorrhizal Fungi Triggers Induction of a Lipochitooligosaccharide Receptor." *Scientific Reports* 6 (May): 29733.

Remigi, Philippe, Jun Zhu, J. Peter W. Young, and Catherine Masson-Boivin. 2016. "Symbiosis within Symbiosis: Evolving Nitrogen-Fixing Legume Symbionts." *Trends in Microbiology* 24 (1): 63–75.

Rogers, Alistair, Yves Gibon, Mark Stitt, Patrick B. Morgan, Carl J. Bernacchi, Donald R. Ort, and Stephen P. Long. 2006. "Increased C Availability at Elevated Carbon Dioxide Concentration Improves N Assimilation in a Legume." *Plant, Cell & Environment* 29 (8): 1651–58.

Roscito, Juliana G., Katrin Sameith, Genis Parra, Bjoern E. Langer, Andreas Petzold, Claudia Moebius, Marc Bickle, Miguel Trefaut Rodrigues, and Michael Hiller. 2018. "Phenotype Loss Is Associated with Widespread Divergence of the Gene Regulatory Landscape in Evolution." *Nature Communications* 9 (1): 4737.

Roy, Sonali, Fran C. Robson, Jodi Lorraine Stewart Lilley, Chengwu Liu, Xiaofei Cheng, Jiangqi Wen, Caitlin Bone, et al. 2017. "MtLAX2, a Functional Homologue of the Auxin Importer AtAUX1, Is Required for Nodule Organogenesis." *Plant Physiology* 174 (1): 326–38.

Rubin, Grit, Takayuki Tohge, Fumio Matsuda, Kazuki Saito, and Wolf-Rüdiger Scheible. 2009. "Members of the LBD Family of Transcription Factors Repress Anthocyanin Synthesis and Affect Additional Nitrogen Responses in Arabidopsis." *The Plant Cell* 21 (11): 3567–84.

Sachs, J. L., M. O. Ehinger, and E. L. Simms. 2010. "Origins of Cheating and Loss of Symbiosis in Wild Bradyrhizobium." *Journal of Evolutionary Biology* 23 (5): 1075–89.

Schauser, Leif, Wioletta Wieloch, and Jens Stougaard. 2005. "Evolution of NIN-like Proteins in Arabidopsis, Rice, and

Lotus Japonicus." *Journal of Molecular Evolution* 60: 229–37.

Schauser, L., A. Roussis, J. Stiller, and J. Stougaard. 1999. "A Plant Regulator Controlling Development of Symbiotic Root Nodules." *Nature* 402 (6758): 191–95.

Schiessl, Katharina, Jodi L. S. Lilley, Tak Lee, Ioannis Tamvakis, Wouter Kohlen, Paul C. Bailey, Aaron Thomas, et al. 2019. "NODULE INCEPTION Recruits the Lateral Root Developmental Program for Symbiotic Nodule Organogenesis in Medicago Truncatula." *Current Biology: CB* 29 (21): 3657–68.e5.

Singh, Sylvia, Katja Katzer, Jayne Lambert, Marion Cerri, and Martin Parniske. 2014. "CYCLOPS, a DNA-Binding Transcriptional Activator, Orchestrates Symbiotic Root Nodule Development." *Cell Host & Microbe* 15 (2): 139–52.

Soltis, D. E., P. S. Soltis, D. R. Morgan, S. M. Swensen, B. C. Mullin, J. M. Dowd, and P. G. Martin. 1995. "Chloroplast Gene Sequence Data Suggest a Single Origin of the Predisposition for Symbiotic Nitrogen Fixation in Angiosperms." *Proceedings of the National Academy of Sciences of the United States of America* 92 (7): 2647–51.

Soyano, Takashi, Hiroshi Kouchi, Atsuko Hirota, and Makoto Hayashi. 2013. "NODULE INCEPTION Directly Targets NF-Y Subunit Genes to Regulate Essential Processes of Root Nodule Development in Lotus Japonicus." *PLoS Genetics* 9 (3): e1003352.

Soyano, Takashi, Yoshikazu Shimoda, Masayoshi Kawaguchi, and Makoto Hayashi. 2019. "A Shared Gene Drives Lateral Root Development and Root Nodule Symbiosis Pathways in Lotus." *Science* 366 (6468): 1021–23.

Sturms, Ryan, Smita Kakar, James Trent, and Mark S. Hargrove. 2010. "Trema and Parasponia Hemoglobins Reveal Convergent Evolution of Oxygen Transport in Plants." *Biochemistry* 49 (19): 4085–93.

Suzuki, Wataru, Mineko Konishi, and Shuichi Yanagisawa. 2013. "The Evolutionary Events Necessary for the Emergence of Symbiotic Nitrogen Fixation in Legumes May Involve a Loss of Nitrate Responsiveness of the NIN Transcription Factor." *Plant Signaling & Behavior* 8 (February 2015): 6–10.

Swensen, Susan M. 1996. "The Evolution of Actinorhizal Symbioses: Evidence for Multiple Origins of the Symbiotic Association." *American Journal of Botany* 83 (11): 1503–12.

Takanashi, K., A. Sugiyama, and K. Yazaki. 2011. "Involvement of Auxin Distribution in Root Nodule Development of *Lotus Japonicus*." *Planta* 234 (1): 73–81.

Trinick, M. J., D. J. Goodchild, and C. Miller. 1989. "Localization of Bacteria and Hemoglobin in Root-Nodules of Parasponia-Andersonii Containing Both Bradyrhizobium Strains and Rhizobium-Leguminosarum Biovar Trifolii." *Applied and Environmental Microbiology* 55 (8): 2046–55.

Van Zeijl, Arjan, Rik H. M. Op Den Camp, Eva E. Deinum, Tatsiana Charnikhova, Henk Franssen, Huub J. M. Op Den Camp, Harro Bouwmeester, Wouter Kohlen, Ton Bisseling, and René Geurts. 2015. "Rhizobium Lipo-Chitooligosaccharide Signaling Triggers Accumulation of Cytokinins in Medicago Truncatula Roots." *Molecular Plant* 8 (8): 1213–26.

Velzen, Robin van, Jeff J. Doyle, and Rene Geurts. 2019. "A Resurrected Scenario: Single Gain and Massive Loss of Nitrogen-Fixing Nodulation." *Trends in Plant Science* 24 (1): 49–57.

Velzen, R. van, R. Holmer, F. Bu, L. Rutten, A. van Zeijl, W. Liu, L. Santuari, et al. 2018. "Comparative Genomics of the Nonlegume Parasponia Reveals Insights into Evolution of Nitrogen-Fixing Rhizobium Symbioses." *Proceedings of the National Academy of Sciences of the United States of America* 115 (20). https://doi.org/10.1073/pnas.1721395115.

Vogel, C. S., and P. S. Curtis. 1995. "Leaf Gas-Exchange and Nitrogen Dynamics of N2-Fixing Field-Grown Alnus Glutinosa under Elevated Atmospheric CO2." *Global Change Biology* 1: 55–61.

Werner, Gijsbert D. A., Johannes H. C. Cornelissen, William K. Cornwell, Nadejda A. Soudzilovskaia, Jens Kattge, Stuart A. West, and E. Toby Kiers. 2018. "Symbiont Switching and Alternative Resource Acquisition Strategies Drive Mutualism Breakdown." *Proceedings of the National Academy of Sciences of the United States of America* 115 (20): 5229–34.

Werner, Gijsbert D. A., William K. Cornwell, Janet I. Sprent, Jens Kattge, and E. Toby Kiers. 2014. "A Single Evolutionary Innovation Drives the Deep Evolution of Symbiotic N2-Fixation in Angiosperms." *Nature Communications* 5 (January): 4087.

Yanagisawa, Shuichi. 2014. "Transcription Factors Involved in Controlling the Expression of Nitrate Reductase Genes in Higher Plants." *Plant Science: An International Journal of Experimental Plant Biology* 229 (December): 167–71.

Yoro, Emiko, Takuya Suzaki, Koichi Toyokura, Hikota Miyazawa, Hidehiro Fukaki, and Masayoshi Kawaguchi. 2014. "A Positive Regulator of Nodule Organogenesis, NODULE INCEPTION, Acts as a Negative Regulator of Rhizobial Infection in Lotus Japonicus." *Plant Physiology* 165 (2): 747–58.

Zhang, Yi Ge, Mark Pagani, Zhonghui Liu, Steven M. Bohaty, and Robert DeConto. 2013. "A 40-Million-Year History of Atmospheric CO2." *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 371 (2001): 20130096.

Thesis summary

Nitrogen represents one of the most important elements for plant growth. Therefore various plant lineages have established a symbiotic relationship with nitrogen-fixing bacteria. One of the hallmarks of plant-microbe symbioses is the nitrogen-fixing nodule endosymbiosis. In this interaction bacteria are housed intracellularly in so-called nodules; specialized organs formed on the plant root or stems. Inside the nodules plants provide the optimal conditions for the bacteria to convert atmospheric nitrogen into ammonia, which they provide to the plants in exchange for photosynthates. Only about 2.5% of the angiosperm families is able to form a nitrogen-fixing nodule endosymbiosis. The over application of chemical fertilizer in agriculture leads to major environmental problems in terrestrial and aquatic ecosystems. Excessive nitrogen deposition causes a loss of biodiversity in natural habitats and nitrogen leaching into the surface waters causes excessive algal blooms. However, without the use of chemical fertilizers the food demands in the world as we know it would not be attainable. Therefore a major aim of the scientific community considers it a major aim to engineer a form of Nitrogen-fixing endosymbiosis in major crop plants, such as rice, wheat or maize.

The nitrogen-fixing endosymbiosis with rhizobia or filamentous actinobacteria of the genus *Frankia* only occurs in four related taxonomic orders. The Fabales, Fagales, Rosales and the Cucurbitales. Collectively called the Nitrogen Fixation Clade (NFC) Within this clade nodulation is restricted to 10 families scattered among mostly non nodulating families. A nitrogen-fixing endosymbiosis can occur with three different types of bacteria: i. Filamentous Actinobacteria form the genus *Frankia*, nodulating a paraphyletic assembly of 25 genera distributed of 8 taxonomic families. ii. Rhizobia a paraphyletic group of -α, β and γ-Proteobacteria, nodulating only two families, the Legumes (Fabaceae) and *Parasponia* (Cannabaceae). In order to explain this distribution a hypothesis, is the existence of a precursor state or “predisposition” for nodulation. In this way in the first common ancestor of the NFC, an innovation happened, which made it more likely for its descendants to evolve nodulation. Most of the knowledge on rhizobium symbiosis comes from model legumes *Medicago truncatula* and *Lotus japonicus*. In these lineages it was discovered that Nodulation shares is initiated by the perception of Lipo-chitooligosaccharides (LCOs), which are structurally similar to the signals produced by Arbuscular Mycorrhizal fungi(AM). These obligate biotrophic fungi colonize the roots of most land plants, where they exchange nutrients for photosynthates. Besides the similarity of the signalling molecule, many of the downstream signalling components between AM-symbiosis

and rhizobium symbiosis are shared. Therefore the signalling pathway is called the Common Symbiosis Signalling Pathway (CSSP). Many common aspects of symbiosis signalling were reviewed in **Chapter 2** in more detail.

The genus *Parasponia*, represents the only lineage to nodulate with rhizobium outside of the legume family. It was previously estimated to have gained its symbiosis relatively recently, given its close relationship to the non-nodulating *Trema* lineage. In order to find this evolutionary precursor state and the innovations that would have initiated a symbiosis with rhizobia. We set out to compare nodulator *Parasponia* with non nodulating relative *Trema* in **Chapter 3**. In contrast to our initial expectations with did not find any gene gains that would correlate to the nodulation trait. Rather we found a pattern of gene loss in close relatives of the nitrogen-fixing *Parasponia* in the Rosales lineage. In addition we discovered a large overlap in the nodule enhanced gene set of model legume *Medicago* and *Parasponia*. Three genes were found to be consistently lost in close relatives of *Parasponia*, transcription factor Nodule Inception(NIN), LCO-receptor Nod Factor Perception (NFP2) and a protein related to infection Rhizobial Polar Growth (RPG). In legumes these three genes are essential for rhizobial infection and nodule formation.

In **Chapter 4,** I continued work on the gene family of Lysin-Motif receptor like kinases (Lysm-RKs) in Parasponia, which includes the putative LCO-receptor NFP2. In legumes it was discovered that the LYR-type Nod factor receptor MtNFP/LjNFR5, functions as a heterodimer with a LYK type receptor MtLYK1/LjNFR1. Here I discovered that Parasponia uses at least four LysM-RKs for rhizobial Nod-factor recognition. Of these four receptors, two are of the LYK-type and have intact kinases with phosphorylating ability. These are named PanLYK1 and PanLYK3. These receptors evolved upon an ancient duplication in the eudicots. In addition to a role in rhizobium symbiosis, PanLYK3 is also involved in chitin triggered immunity, indicating a dual functionality for this receptor. The second receptor PanLYK1 has no major phenotype as a single mutant. However only a double *panlyk1-panlyk3* mutant can complete block nodule formation and Arbuscular Mycorrhizal infection in Parasponia. This indicates that there seems to be functional overlap but also a distinction between the two LYK-I type receptors.

The other two receptors, PanNFP1 and PanNFP2, represent two LYR-type LysM-RKs, with inactive kinase conformations. The duplication which gave rise to these two copies of LYR-I type receptors happened in an ancestor of the NFC. The presence of an NFP-II-type ortholog strictly correlates to the presence of nodulation. The loss of the NFP2 copy in non-nodulating lineages indicates that this receptor is committed to functioning as a stringent LCO-perception protein in symbiotic context. While *pannfp2* mutants cannot be infected by Rhizobia or form nodules, they have

no apparent phenotype in AM-symbiosis, which further supports its specialized role. PanNFP1 most likely has a role in AM-symbioses as was suggested previously, which would be in line with its presence in most non-nodulating lineages in the NFC-clade. However, besides a minor phenotype in Nodulation a functional role for this receptor in AM-colonization could not be supported.

Besides the loss of LCO-receptor NFP2. Orthologues of two other receptors, PanLEK1 and PanCRK11, were consistently lost in the *Trema* lineage. In **Chapter 5,** I characterized their respective gene families and show that PanLEK1 plays an essential role in regulating defense responses in the Nodule. Mutants in *panlek1* show a reduction in nodule number and a accumulation of phenolic compounds in the nodule. PanCRK11 belongs to a large gene family of Cysteine-Rich Receptor like kinases (CRKs), many of which are regulated in symbiotic context. A genomic cluster of 20 CRKs was targeted by CRISPR-CAS9, which resulted in a reduction of nodule number and infection level in the mutant. These results hint towards a role for cysteine richt kinase receptors in regulating infection thread progression.

Most of the genes used the context of Nodulation exist outside of the NFC, where they have a different function. The recruitment of these genes in nodulation required novel Cis-regulatory elements in their respective promoters. Finding these Cis-regulatory elements may prove essential for future engineering efforts in crop species, since they allow the correct spatio-temporal gene expression. NIN-represents one of the most central transcription factors in Nodulation, however its regulation in both Symbiotic infection and Nodule formation was to date not well understood. In **Chapter 6**, we discovered a novel a Cis-regulatory element required to initiate NIN-expression in the pericycle. This element, proved to be essential for functional complementation of a *nin* mutant. The expression of this element is in part regulated by cytokinin, which by itself is capable of stimulating nodule organogenesis.

The findings in **Chapter 3,4 and 5** are not in line with a independent origin of Nodulation, rather they suggest a Single origin of Nodulation in the NFC. This scenario would imply a widespread loss of the Nodulation trait. In the general discussion **Chapter 7,** I discuss the potential drivers which could have led to the widespread loss of the nitrogen fixing endosymbiosis. I discuss the implications of these findings for the potential of engineering the nitrogen fixation trait in crop species.

List of Publications

L. Rutten*, K. Miyata*, Y. P. Roswanjaya, R. Huisman, F. Bu, M. Hartog, S. Linders, R. van Velzen, T. Bisseling, W. Kohlen and Rene Geurts. 2020. The duplication of two Symbiotic LysM-receptors predates the evolution of Nitrogen fixing symbiosis. *Plant Physiology Jul 2020, pp.01420.2019; DOI: 10.1104/pp.19.01420*

Holmer, R.*, L. Rutten*, W. Kohlen, R. van Velzen, and R. Geurts. 2017. "Commonalities in Symbiotic Plant-Microbe Signalling." *Advances in Botanical Research*. https://doi.org/10.1016/bs.abr.2016.11.003.

Liu, Jieyu*, Luuk Rutten*, Erik Limpens, Tjitse van der Molen, Robin van Velzen, Rujin Chen, Yuhui Chen, et al. 2019. "A Remote Cis-Regulatory Region Is Required for NIN Expression in the Pericycle to Initiate Nodule Primordium Formation in Medicago Truncatula." *The Plant Cell*, January, tpc.00478.2018.

Velzen, R. van, R. Holmer, F. Bu, L. Rutten, A. van Zeijl, W. Liu, L. Santuari, et al. 2018. "Comparative Genomics of the Nonlegume Parasponia Reveals Insights into Evolution of Nitrogen-Fixing Rhizobium Symbioses." *Proceedings of the National Academy of Sciences of the United States of America* 115 (20). https://doi.org/10.1073/pnas.1721395115.

Bu, Fengjiao, Luuk Rutten, Yuda Purwana Roswanjaya, Olga Kulikova, Marta Rodriguez-Franco, Thomas Ott, Ton Bisseling, Arjan van Zeijl, and Rene Geurts. 2019. "Mutant Analysis in the Non legume Parasponia Andersonii Identifies NIN and NF-YA1 Transcription Factors as a Core Genetic Network in Nitrogen-Fixing Nodule Symbioses." *The New Phytologist* https://doi.org/10.1111/nph.16386.

Zeijl, Arjan van, Titis A. K. Wardhani, Maryam Seifi Kalhor, Luuk Rutten, Fengjiao Bu, Marijke Hartog, Sidney Linders, et al. 2018. "CRISPR/Cas9-Mediated Mutagenesis of Four Putative Symbiosis Genes of the Tropical Tree Parasponia Andersonii Reveals Novel Phenotypes." *Frontiers in Plant Science* 9 (March): 1–14.

*shared first authorship.

To keep track of future updates on publications, scan the qr-code or visit the link:

https://orcid.org/0000-0002-6239-2404

Acknowledgements

Hallo lezers van deze thesis. Het is natuurlijk vanzelfsprekend dat je een thesis niet in je eentje kan schrijven. Hier zal ik proberen iedereen die een bijdrage heeft geleverd aan deze thesis te bedanken. En ik hoop dat ik uiteindelijk niemand ben vergeten.

Ten eerste was deze thesis er natuurlijk nooit geweest zonder mijn dagelijkse begeleider en copromotor Rene. Jij hebt alweer een tijd geleden het lef gehad om te beginnen aan het onderzoek aan stikstoffixatie in Parasponia. Een onbekend terrein voor velen. Je hebt me de kans gegeven om na mijn master thesis door te gaan met promotieonderzoek. Ik moet je daarom bedanken voor het vertrouwen dat je op dat moment hebt getoond. Het onderzoek mag dan af en toe andere resultaten hebben opgeleverd dan vooraf verwacht, toch was je nooit uit het veld te slaan.

Het MolBi-lab had uiteraard niet bestaan zonder mijn Promotor Ton Bisseling. Die in de beginjaren van mijn Phd traject, ook als hoofd van EPS aan mijn opleiding verbonden was. Ton, bedankt voor je scherpzinnige opmerkingen tijdens werk discussies en op mijn vragen. Officieel op weg naar je pensioen, maar in de praktijk zijn kunnen we nog lang niet zonder onze Ton. Mijn excuses dat ik als stand-in in enkele MolBi-cabaret uitvoeringen ben ingezet. De volgende die ik graag wil bedanken is Arjan. Arjan jij was mijn begeleider tijdens mijn Msc thesis, toen begeleidde je Mij, Fleur, Titis en Max tegelijk. Later werden we collega's. Zonder jou hadden we Parasponia onderzoek lang niet zo ver gekregen als dat we nu zijn. En of je het nu wil of niet je had een behoorlijke voorbeeldfunctie in het lab. Tegelijk was je zo ongeveer het middelpunt in het Parasponia Team. Ook de lunchpauzes konden niet zonder jou discussies. Jouw tegenpool in die lunch discussies was natuurlijk Sidney. Want op het moment dat Sidney zich bij het parasponia team voegde, werden jullie een vast team. Met gezonde samenwerking hebben jullie beide hard gewerkt aan Parasponia transformaties. Sidney bedankt voor alle hulp en assistentie die je hebt geboden gedurende je Periode bij MolBi. Het is nu weer een stuk stiller op het lab. Rens, ik heb je regelmatig lastig gevallen met vragen naar informatie, genomen, expressie data en modelling. Je was een onmisbare schakel binnen het Parasponia team, veel succes met het afronden van jouw thesis binnenkort. Robin, jij hebt me met allerlei aspecten geholpen, phylogenetische bomen, alignments en wat zo niet meer. Je hebt hard gewerkt aan het genoom paper van parasponia, wat uiteindelijk het startpunt is geworden voor mijn verder thesis. Wouter, ik wil je bedanken voor alle extra discussies die regelmatig voerden over de resultaten, maar ook over de zwaardere tijden die je tegenkomt. Ook je hulp en kennis is van pas gekomen bij het schrijven van hoofdstuk 2 en 6. Erik, ik wil je bedanken vanwege je kennis die

je in kon brengen tijdens mijn werk discussies. Rik ook jou wil ik bedanken, vooral voor je praktische tips in het lab, je discussies tijdens de koffie en ook voor je hulp bij hoofdstuk 4. Hopelijk hebben we ook nog een vruchtbare samenwerking in de toekomst. Verder wil ik graag Tijs bedanken, jij hebt me in mijn Bsc thesis al het vertrouwen gegeven dat ik nodig had. Ook tijdens mijn Phd kon ik bij jou altijd wat luchtige opbeurende woorden vinden.

Ook alle andere leden binnen de MolBi and PCB staff zou ik graag bedanken bij hun hulp: Henk, voor het simpel oplossen van wereldproblemen tijdens de koffie. Jan Verver, voor het getrommel en al zijn hulp bij het kloneren, DNA isoleren en alle andere vragen. Jan Hontelez, voor de hulp tijdens het geven van mijn eerste toolbox. Jac, voor de praatjes tijdens de koffie die altijd wat uitliepen. Marijke, voor al het werk dat ze gestoken heeft in het onderhouden van mijn Parasponia plantjes en voor de praktische vragen bij de tissue culture. Joan, voor de VriMiBo en de literatuur discussies. Renze voor het lenen van zijn golden gate plasmids. Maria en Marie-José, die altijd snel te hulp konden schieten bij vragen, betalingen, formulieren, bestellingen maar ook gewoon voor de gezelligheid en de sfeer die jullie brengen tijdens de koffie.

I would also like to thank all of my fellow Parasponia Team members: Titis, for the extensive work on further developing the Parasponia protocols and for increasing the social bonding in the lab. Yuda, for your enormous working ethos, helping me characterize the LysM-mutants for mycorrhization. But also for your great cooking during dinners. Kana, for your extreme professionalism. You did a load of work with me on the characterizing the LysM-receptors in Chapter 4. Without you the work wouldn't be what it is now. I cannot thank you enough for that. Simon, from whom you can always receive help with figuring out Plots in R. Fengjiao, who always sat right in front of me. I enjoyed working with you on the Parasponia NIN, NF-YA manuscript, which you have recently defended in your thesis. Sultan, you have only just started, but already you bring a smile to the Molbi Lab. I wish you the best of luck in the future. Finally i would like to thank Elena for her help with sectioning and microscopy.

Olga and Jieyu, I would like to thank you for trusting my random piece of alignment which turned out to be the CE-element. Together our work sprouted into a great paper which is now published in Plant Cell and makes up Chapter 6. of my thesis.

Further i would like to thank all other members, former members, phd students and staff of the MolBi, Plant Cell or Cell Biology Lab. Ting Ting, Huchen, Defeng, Jianyong, Xu, Jelle, Yinshan, Fang, Lucas, Martinus, Asma, Ashenafi, Comfort, Amina, Peng, Tian, Ole, Sabine, Adam, Mengmeng, Bandan, Jeroen, Aniek, Steven, Elysa, Han, Wenkun, Guiling, Zhichun, Fuxi, Siqi, Jundi, Maryam, Trupti, Gorge, Kavya, Bandan, Eva, Wilma, Joel, Anneke, Menno, Ikram, Vera, Merijn, Alejandra,

Peter, Kiki, Carolien, Kerstin, Marcel, Norbert, Hannie, Henk k, Viola en niet te vergeten Ben. Without any one of you my phd in the Molbi lab would have been different. Thank you.

Tot slot blijven natuurlijk over mijn vrienden en familie. Thijs, als het weer eens even tegen zat ging ik met jou een rondje fietsen, aan fietsen klussen of gewoon een biertje dingen. Verder wil ik natuurlijk bedanken mijn fietsmaten bij W.S.W.V Hellingproof en voormalig Teamgenoten bij Computron cyclingteam, waarbij fietsen mijn uitlaatklep mocht zijn. Verder wil ik nog alle leden van C.V. Hakuna Matata bedanken voor de leuke uitstapjes.

Pap, Mam, Eveline, Stef, Anouk, Pieter, Dyantha en Leon bedankt voor alle gezellige weekenden. Su, lieve meid. Toen ik jou zag staan in t MolBi-lab was ik verkocht. Bedankt voor alle lieve steun en aansporing tijdens het schrijfwerk. En ook voor het nog grondig herzien van de tekst. En als aller-allerlaatste Kedi en Abbey, voor het doorstaan van alle gevraagde en ongevraagde knuffels.

Ten slotte wens ik jullie allen veel wijsheid en geluk!

Groeten Luuk.

Curriculum Vitae.

Lukas Johannes Joseph Rutten, was born on the 26th of January 1991 in Deurne, Noord-Brabant, The netherlands. After obtaining his high school degree from the ‘Elzendaal College’ in Boxmeer, he moved to wageningen to start his Bachelor in Biology. He specialized with a major on Organismal and Developmental biology and did a minor on Ecology. He completed his bachelor thesis on Actin filament dynamics in July 2012. Which was performed at the laboratory of Cell Biology, under the supervision of Dr ir. Tijs Ketelaar. Due to his growing interest in plant biology he decided to continue with a Master in Plant adaptation. For his Master Thesis he Joined a project led by Dr ir. Rene Geurts at the laboratory of Molecular Biology. He completed his Msc thesis entitled: “A Comparative genomic study between the genera Parasponia and Trema, the effects of auxin on Nodule formation”, in January 2014. In March 2014, he Joined the group of Prof dr Maarten Korneef at the Max Planck Institute for plant breeding research in Cologne, Germany. Here he worked under the supervision of Dr Ralph Bours to work on the Genetic Mapping of shoot and flowering time traits in arabidopsis. This internship was completed in Oktober 2014. Soon after the completion, In December 2014, he received his Msc Certificate. In January 2015, he received an invitation to re-join the Laboratory of Molecular Biology at Wageningen University to perform his PhD research. Under supervision of Dr ir. Rene Geurts and promoter Prof. Dr Ton Bisseling, he studied how Parasponia engages in symbiosis with nitrogen-fixing rhizobium bacteria. The results of this work are presented in this thesis.

Colofon

This work was conducted in the laboratory of Molecular Biology, Wageningen University, Gelderland, The Netherlands. with suppport by an NWO-VICI grant (865.13.001) to RG, the ENSA project funded by the Bill & Melinda Gates Foundation to the University of Cambridge to RG, NWO-VENI grant (863.15.010) to WK, a JSPS Overseas Research Fellowship to KM and The Ministry of Research, Technology and Higher Education of the Republic of Indonesia (RISET-PRO grant 8245-ID) to YPR

Design and Print

Cover Image Luuk Rutten, edited by proefschriftmaken

Printed by: Proefschriftmaken. www.proefschriftmaken.nl